Proceedings
of the
Steklov Institute of Mathematics

1979, Issue 2

Geometric Problems in the Theory of Infinite-Dimensional Probability Distributions

by

V. N. Sudakov

Translation of

ТРУДЫ
ордена Ленина
МАТЕМАТИЧЕСКОГО ИНСТИТУТА
имени В. А. СТЕКЛОВА

Том 141 (1976)

АКАДЕМИЯ НАУК

СОЮЗА СОВЕТСКИХ СОЦИАЛИСТИЧЕСКИХ РЕСПУБЛИК

ТРУДЫ

ордена Ленина

МАТЕМАТИЧЕСКОГО ИНСТИТУТА

имени В. А. СТЕКЛОВА

CXLI

ГЕОМЕТРИЧЕСКИЕ ПРОБЛЕМЫ ТЕОРИИ БЕСКОНЕЧНОМЕРНЫХ ВЕРОЯТНОСТНЫХ РАСПРЕДЕЛЕНИЙ

В. Н. СУДАКОВ

Издательство "Наука"
Москва 1976

Ответственный редактор академик С. М. НИКОЛЬСКИЙ

Translated by H. H. McFaden

AMS (MOS) subject classifications (1970). Primary 28–02, 46C05, 60–02, 60B99, 60C05, 60G99, 62B05, 62B20, 62E15; Secondary 05B20, 47B99, 52A45, 60E05, 60F20, 62P20, 15A51.

ABSTRACT. Questions in measure theory related to the geometric character of the methods of investigation used make up the subject of this monograph. The first chapter is of an auxiliary nature. The problem of extending a generalized random process to a measure in a Banach space is studied in the second chapter. The case of Gaussian random processes is considered in particular detail, and new results in the geometry of Hilbert space are used. The third chapter deals with the problem of the existence of an independent complement to a pair of given measurable decompositions of a measure space; the basic result here can be regarded as an improved continuous analogue of the Birkhoff-von Neumann theorem on the decomposition of doubly stochastic matrices. Here the investigation is concerned with the existence of measures on subsets of product spaces with given marginal distributions and other additional properties. The results of the third chapter are applied in determining conditions for the existence of an optimal one-to-one plan for transfer of mass in the Monge problem on a finite-dimensional space with a Minkowski metric.

Library of Congress Cataloging in Publication Data

Sudakov, V. N.

Geometric problems in the theory of infinite-dimensional probability distributions.

(Proceedings of the Steklov Institute of Mathematics; 1979, issue 2 ISSN 0081-5438)

Translation of Geometricheskie problemy teorii beskonechnomernykh veroîatnostnykh raspredeleniĭ.

Number in Russian series statement: t. 141 (1976)

Bibliography: p.

1. Distribution (Probability theory) 2. Measure theory. I. Title. II. Series: Akademiîa nauk SSSR. Matematicheskiĭ institut. Proceedings; 1979, issue 2.

QA1.A413 1979, issue 2 [QA273.6] 510'.8s [519.2] 79-11640

ISBN 0-8218-3041-4

PROCEEDINGS OF THE STEKLOV INSTITUTE OF MATHEMATICS
IN THE ACADEMY OF SCIENCES OF THE USSR

TABLE OF CONTENTS

Introduction 1

CHAPTER I. Auxiliary information 7

CHAPTER II. Sample functions of random processes. Correlation theory 11

§1. Statement of the problem 12

§2. Problem of the size of a convex solid angle in Hilbert space and the infinite-dimensional Cauchy distribution 23

§3. Cauchy measure and Gaussian measure 30

§4. The *GB* property and mixed volumes 33

§5. Monotonicity of the functional h_1 41

§6. Mixed volumes and the continuity of paths 49

§7. ϵ-entropy conditions 54

§8. The non-Gaussian case 59

§9. Remark on Borel realizations 67

CHAPTER III. Independence and combinations of measurable decompositions of a probability space 69

§10. The Birkhoff-von Neumann problem 70

§11. Probability measures on subsets of direct products 129

§12. Marginal sufficiency of statistics 153

§13. Conditions for the existence of a one-to-one optimal plan in the problem of transport of mass in Minkowski spaces 160

Bibliography 173

INTRODUCTION

This monograph is a collection of some results, published previously but mostly without detailed proofs, of investigations in the theory of probability distributions. Of the three chapters making up this book, the first contains only auxiliary information. The division of the whole content into two chapters (II and III) corresponds to the two main themes of the investigations. The second and third chapters have in common the geometric character of the problems studied and the related definite unity of methods, although in content these chapters are formally independent of each other. A study of the measure of a solid angle in a Hilbert space occupies one of the central places in the second chapter, which deals with the properties of sample functions of random processes; the problem of the properties of the sample functions is studied in terms of the geometry of a certain subset of the Hilbert space determined by the random process and the relevant property of the realization. On the other hand, the basic theorem of the third chapter, whose proof makes up the content of §10, can be naturally stated as a result on the extreme points of the infinite-dimensional analogue of the so-called "Hungarian polyhedron" determined by specifying the marginal distributions of two statistics (two measurable decompositions) and consisting of all the measures having the given marginal distributions. The remaining sections of the third chapter are closely related in substance to this result or are directly based on it.

The first area of probability theory in which the methods of functional analysis received a wide application, beginning with the work of Kolmogorov [61]–[64] and Cramér [20], was the theory of stationary random processes, where the spectral decomposition of the one-parameter group of linear operators associated with the process becomes a fundamental tool of investigation. Later, Karhunen [54] introduced into use by the specialists in probability theory the spectral decomposition of the integral operator whose kernel is the correlation function of the random process.

It should be mentioned, however, that, although the possibility of regarding an arbitrary random process as a probability measure in a space of functions on the parameter set appeared from the time of the proof of Kolmogorov's theorem on the extendibility of a compatible system of finite-dimensional distributions to a measure, the advantage of subsequently carrying through this point of view became sufficiently clear only in the 50's, when a further widening of the area of application of the methods of functional analysis in the theory of random processes took place. Here it is impossible not to mention the appearance in 1953 of Doob's monograph [24], especially the origin of the concept of a "generalized random process" (I. M. Gel′fand [34],

[35]; K. Itô [43]) and the systematic investigation of functional-analytic methods in the question of convergence of probability measures in separable metric and, in particular, in normed linear spaces, with application of the general results obtained to the study of convergence of random processes (Ju. V. Prohorov [86]). One of the important problems that presented itself to both Gel'fand and Prohorov, and to their students, was the problem of a criterion for the existence in a linear space (in particular, in a normed linear space) of a measure generating a given system of compatible finite-dimensional distributions (i.e., having a given positive definite functional as its characteristic functional).

R. A. Minlos [76], [77] and V. Sazonov [103], independently of each other and in different terms, established criteria for a Hilbert space (Sazonov, using the results of [86]) and for the space dual to a nuclear countably normed space (Minlos, after proving a conjecture of Gel'fand). The connection of the theorems of Minlos and Sazonov and their equivalence was observed by Kolmogorov [67] (see also [15]). It later bacame clear (Sudakov [115]) that for certain separable Banach spaces a criterion for the existence of a measure with a given characteristic functional cannot in principle be formulated in terms of the continuity (as in Sazonov's theorem) of this characteristic functional in some topology determined by the space under consideration. However, for Gaussian measures it was shown in [15] that such a critical topology always exists, though it has not been described in the non-Hilbert case (except in certain particular cases, for example, for the l^p spaces). It was not clear in terms of which mutual characteristics of the Banach space and the correlation operator determining the Gaussian weak distribution one should solve the problem of extension of a weak distribution to a measure; at the same time, the complete solution of the problem for the Hilbert case, given by the Minlos-Sazonov theorem, did not yield sufficiently nice necessary or sufficient conditions for such concrete Banach spaces as spaces of bounded or continuous functions.

The idea of using ϵ-entropy characteristics here is due independently to several authors. At the 1966 International Congress of Mathematicians in Moscow the present author presented conditions for extendibility of a Gaussian weak distribution to a measure in a space E dual to some separable Banach space F, expressed in terms of the ϵ-entropy of the unit ball $V_F \subset F$ with respect to the Hilbert norm generated by the correlation operator. It turned out that the condition $\rho(V_F) \leqslant 2$ is necessary, and the condition $\rho(V_F) < 2$ is sufficient for the extendibility of a Gaussian weak distribution to a measure in E, i.e. for the boundedness of the supremum of the set of Gaussian variables V_F ($\rho(A)$ is the entropy index). However, as was then remarked, necessary and sufficient conditions for the extendibility of a Gaussian weak distribution to a measure in an arbitrary (or even an arbitrary separable or the dual of a separable) Banach space cannot be formulated in terms of ϵ-entropy.

In 1967, independently of this, R. M. Dudley published an article [25] in which he proved that the condition $\int_0 (H_{V_F}(\epsilon))^{1/2}\, d\epsilon < \infty$, where $H_A(\epsilon)$ is the ϵ-entropy of the set A (weaker than the condition $\rho(V_F) < 2$, but, of course, more restrictive than the condition $\rho(V_F) \leqslant 2$), ensures the boundedness of the supremum of the set K of

Gaussian variables and even the continuity with probability 1 of the realizations $x(\omega)$, $x \in K$, of the Gaussian process in the relative (Hilbert) metric of the "parameter" set K, regarded as a subset of the Hilbert space of all random variables with variance. At the same time Dudley stated as a conjecture the assertion (already proved at that time) that the condition $\rho(K) > 2$ implies the unboundedness of the paths with probability 1. What was fundamentally new in Dudley's work was the study of the continuity of the sample functions with respect to the topology on K (with regard to the idea of using ϵ-entropy Dudley refers to V. Strassen).

As was subsequently observed ([121], [125]), our method enables us in a completely uniform way to get, besides necessary conditions (which were not the aim of Dudley's method), also sufficient conditions. Later, Dudley [26] improved the necessary (for the continuity of realizations) condition $\rho(K) < 2$, proving that the condition $\limsup_{\epsilon \to 0} \epsilon^2 \ln N(K, \epsilon) < \infty$ is necessary for the boundedness of realizations ($N(K, \epsilon)$ is the cardinality of a smallest ϵ-net). Our method allowed us [131] to sharpen somewhat the result and to show that for continuity it is necessary that the condition $\limsup \epsilon^2 \ln N(K, \epsilon) = 0$ holds (Dudley had remarked earlier that this condition was satisfied in all the examples considered by him). By the same token, necessary conditions for boundedness and continuity of the paths of Gaussian processes were separated for the first time in the language of ϵ-entropy.

Fernique proved recently [32] that Dudley's ϵ-entropy sufficient condition is also necessary in the class of stationary Gaussian processes; the ϵ-entropy methods had thereby led to a solution of an old problem in the theory of Gaussian processes. Jain and Marcus [44] then succeeded in completely eliminating mention of ϵ-nets in the stationary case, expressing the ϵ-entropy of the corresponding spiral in terms of the monotone rearrangement of the covariance function of the process and reformulating the Dudley-Fernique condition in this language. The continuity criteria based on the use of ϵ-entropy enable one, in particular, to re-prove all the continuity conditions for Gaussian processes that have appeared up to this point. The same can be said also about the ϵ-entropy conditions for discontinuity of paths. The ϵ-entropy conditions for boundedness and continuity of realizations of a Gaussian process x_t, $t \in T$, are formulated in terms of the metric on the parameter set T given by the formula $d(t, s) = (2(1 - B(s, t)))^{1/2}$, where $s, t \in T$, and $B(s, t)$ is the correlation function of the process (assumed to be standardized). However [121], it can be shown that any Banach space E with a countable total set of linear functionals on it is contained in the set of all linear forms on F that are bounded on V_F, and contains the set of all linear forms on F that are bounded and continuous on V_F in the Hilbert topology, where F is a given total set of linear functionals on E, and V_F is the polar of the unit ball $V_E \subset E$. Thus, any Banach space is "sandwiched" in the set-theoretic sense between the space of bounded and the space of continuous forms on V_F (which coincide when E is reflexive). Therefore, sufficient conditions for continuity, applied to the set V_F (which thus is regarded as parametric), are simultaneously sufficient conditions for the extendibility of a compatible system of finite-dimensional Gaussian distributions to a measure in E; and necessary conditions for boundedness can be regarded as necessary conditions for realizations of the process to have the "property E" with probability 1. We mention also that the zero-one law is proved for Gaussian measures of linear spaces in [115] and [121].

The essence of the method presented in [120], [121], [125] and [131] lies in the following. Exact criteria for continuity and boundedness of realizations of a Gaussian process K, regarded as a subset of the Hilbert space $L^2(\Omega, P)$, can be formulated in terms of the Minkowski mixed volume of the first degree of homogeneity $h_1(K)$ (Theorems 1 and 3), which is proportional to the integral of the supremum of the set of functions $K \subset L^2$ (Proposition 14). For this, it is proved that the problem of verifying the boundedness of realizations of the process K is equivalent to the problem of the positivity of the measure of the solid angle in the Hilbert space that is polar to the cone generated by a certain translate of the set K. The problem of the measure of a solid angle leads naturally to the consideration of an infinite-dimensional Cauchy measure, which is simply related to Gaussian measure (Propositions 6, 7, and 8, 10), since the conditional measures for the Gaussian measure under the decomposition of a linear Gaussian measure space into rays are δ-measures (Proposition 5).

The connection between the asymptotic behavior of the Cauchy measure of the homotheties of the set K° and the value of the functional $h_1(K)$ is described in Proposition 21. An estimate of the value of $h_1(K)$ in terms of the probability of exit from the unit level is given by Proposition 18 (a less sharp estimate can be obtained by using results of Skorohod [109] or Fernique [30]; to get our estimate we use the solution of the isoperimetric problem on the n-dimensional sphere). Using the "geometric" origin of the Cauchy measure and its connection with the functional $h_1(K)$, we can prove a monotonicity property for $h_1(K)$ (Theorem 2 and its Corollary) that is very essential for us (and is new even in the finite-dimensional version) and that permits us to estimate $h_1(K)$ from below and from above in terms of the ϵ-entropy of the set K by comparing the ϵ-nets and the ϵ-lattices for simple sets such that $h_1(K)$ can be calculated or estimated directly (Propositions 31, 33, 34, and their corollaries). In investigating the continuity of the paths of a process K we study the structure of the space of continuous forms on K (Propositions 26 and 27), we consider the "oscillation" of the process (the functional $\delta(K) = \sup\{d\colon \gamma(dK^\circ) = 0\}$), and we find its connection with the functional $h_1(K)$ (Proposition 30). Determining the connection between these quantities allows us to get a lower estimate of the oscillation $\delta(K)$ in terms of the ϵ-entropy characteristics of the process (Proposition 35), and an upper estimate is essentially contained in the estimates of Propositions 30 and 33. A definitive formulation of the ϵ-entropy conditions is given in Theorem 4.

The second chapter is concluded with a study of arbitrary (non-Gaussian, in general) processes $K \subset L^2(\Omega, P)$ ("with second moments"); we clarify what information about the boundedness or the continuity of the paths of the process is carried by the geometry of the set K (i.e. the correlation function of the process). Theorem 5 gives an exhaustive answer to this question.

The third chapter is the longest. Its main result is contained in §10. The complete proof presented in this section involves many technical details and hence may seem tedious (the proof here, as also in other places, is preceded by a brief sketch of the basic scheme of the arguments). The result can be stated very simply in terms of the theory of measurable decompositions: If two measurable decompositions ξ and η of a Lebesgue space (M, m) with purely continuous marginal measures are "quasi-independent" (the definitions are collected in Chapter I), then there exists a third measurable decomposition ζ that is simultaneously an independent complement to ξ and η (Theorem 8; [122], [124], [127]); an improvement that is important for subsequent applications is given in Theorem 8*.

Theorem 8 provides the solution of a problem posed by Birkhoff in [11]; it can be regarded as a substantially improved "continuous" version of the well-known Birkhoff-von Neumann theorem on the extreme points of the set of all doubly stochastic square matrices of given dimension (the extreme points turn out to be the doubly stochastic (0, 1)-matrices). While for infinite matrices, as the investigations of many authors have shown (a survey is given at the beginning of the third chapter), the situation is on the whole analogous to the finite-dimensional case, the picture in the continuous version turns out to be qualitatively more complicated, and for a long time it was not clear when it is possible to guarantee that a measure with given marginal distributions with respect to a pair of given measurable decompositions has a decomposition analogous to the Birkhoff-von Neumann decomposition of a doubly stochastic square matrix. The set of extreme points of the compact set of all probability measures with given marginal distributions has a very complicated structure and, anyway, does not consist only of measures analogous to the doubly stochastic (0, 1)-matrices, i.e. those corresponding to isomorphisms of the marginal measure spaces $(M/\xi, m/\xi)$ and $(M/\eta, m/\eta)$. This circumstance barred the use of the Choquet-Kreĭn-Mil′man theorem on representation of the points of a compact set as barycenters of measures on the set of extreme points, and made it necessary to find a direct proof.

Other points of view about the result contained in Theorem 8 (operator-theoretic, statistical, connection with the theory of Latin squares) are discussed in detail at the beginning of the third chapter.

The questions in §10 arose from statistical problems, especially from the investigations carried out by Ju. V. Linnik and his students concerning the Behrens-Fisher problem. The concluding papers [45], [71] in this direction, in which the existence of nonrandomized tests in the Behrens-Fisher problem for samples of different parity is proved, make essential use of a lemma of I. V. Romanovskiĭ and the author [93], which is a vector version of a result preceding the proof of Theorem 8 (a description of the set of extreme points in the compact set of bounded doubly stochastic densities with zero sets containing a given subset; see Propositions 43 and 43a).

In spite of all its unwieldiness, the proof of Theorem 8 has a central point. This is the proof of the approximation theorem (Theorem 7). Moreover, essential use is made of a condition for the existence of a subprobability measure $\overline{m}$ on M having given marginals and majorized by a given measure m (Theorem 6): to answer the question of the existence of such a measure $\overline{m}$ it suffices to answer the analogous question for 2×2 matrices. The work of authors who obtained related results (usually involving a less general situation) are carefully noted each time, as far as possible, in the footnotes.

The area of research to which the results in §10 belong can be called continuous combinatorics.

In §11 we study criteria for the existence of measures with given marginal distributions and concentrated on a fixed subset of the product of two spaces [129], [130]. Unlike all the other authors who have studied related problems (Kellerer [55]–[57] and Strassen [113]), we do not assume any additional structures (for example, topologies) on our measure spaces, so that the class of subsets of the product $X \times Y$ for which the solution of the problem is given (Theorem 9) is described in terms of pure measure theory, and in the topological case, for example, it turns out to be a strictly wider class than the closed subsets, for which the results of Strassen can be used. However, the main result of §11, of which essential use is made in the sequel, is contained in Theorem 10 and, apparently, cannot be obtained as simply as Theorem 9. Theorem 10 gives necessary and sufficient conditions for the existence of a doubly stochastic density on a particular subset of the product of two spaces, i.e., a measure that is absolutely continuous with respect to the product of its marginal distributions. The difficulty that must be overcome here has to do with the noncompactness of the set of doubly stochastic densities, which rules out the possibility of a simple passage to the limit.

§12 occupies an essentially independent position in the third chapter and is connected with the other sections only informally. In it we prove that in the case of an independent sample the marginal sufficiency of a statistic implies its sufficiency [126]. The apparatus used and the situation when we consider a number of measurable decompositions on one space relate §12 with the other parts of the third chapter. The author included this section as a memorial to Ju. V. Linnik, who suggested the problem to him and, with unexpected enthusiasm, pointed to the method of its solution.

Finally, §13 can serve as an example of the application of the basic results in §§10 and 11. Here we solve the problem of determining sufficient conditions for the existence of a one-to-one optimal plan in the Monge problem on transport of mass in Minkowski spaces. The method presented is essentially based on Kantorovič's theorem on the existence in the Monge problem of a potential function, by means of which the optimality of this or that plan of transport is tested.

In conclusion the author wishes to express his gratitude to L. N. Dovbyš, without whose constant help and support this work could not have appeared.

Chapter I

AUXILIARY INFORMATION

1. The structure of a measure space, measurable decompositions, Lebesgue space (7). 2. Other structures associated with a measure space (9). 3. Linear measure spaces (9).

In this chapter we recall some definitions of concepts to be used later, and we make more precise the meaning of terms for whose use there is no universally recognized tradition.

1. The basic object (more precisely, the basic structure) with which we shall have to deal is the structure of a measure space. In the following the word "measure" usually means a measure on a Lebesgue space, i.e. a probability (nonnegative normalized) measure on a σ-algebra of subsets of some set such that there exists a one-to-one reciprocally measurable mapping, defined almost everywhere, of this space onto a subset of full measure in a measure space that is the union of a (possibly degenerate) segment with Lebesgue measure and a (possibly zero) number of point masses. Every separable metric space with a Borel probability measure is a Lebesgue space.

We frequently consider other structures simultaneously on a measure space (topological, or algebraic, or both these and others). The compatibility of the structures usually lies in the fact that some natural σ-algebra of subsets determined by the topological (or algebraic) structure is a generating σ-algebra for the σ-algebra on which the measure is defined (thus, each Lebesgue measurable set is Borel mod 0, i.e., its symmetric difference with some Borel set has measure 0). In other words, the σ-algebra on which the measure is defined is the completion of such a natural σ-algebra (is obtained by adjoining all subsets whose outer and inner measures coincide). Of course, the completion of a σ-algebra depends on the specific form of the measure on it.

In the case of a Lebesgue space the σ-algebra on which the measure is defined is countably generated, i.e., is the completion of the smallest σ-algebra containing some countable collection of subsets. Such a system of subsets is called a basis.

A decomposition of a measure space is said to be measurable if it is a decomposition into the sets of constant value ("level sets") of some measurable function. The set-theoretic limit of a refining sequence of measurable decompositions is a measurable decomposition. A function is said to be measurable with respect to a decomposition ξ if it is measurable with respect to the σ-algebra of all measurable subsets made up of elements of ξ.

If ξ is a measurable decomposition of the space $(\Omega, \mathfrak{A}, \mu)$, then the measure $\mu/\xi = \mu\pi^{-1}$ is defined in the natural way on the set of elements $C^\xi \subset \Omega$, $C^\xi \in \xi$, which make up the quotient space Ω/ξ, by means of the canonical mapping $\pi\colon \Omega \longrightarrow \Omega/\xi$;

this measure will frequently be called the marginal measure, or marginal distribution, corresponding to the measurable decomposition ξ. The decomposition into points is usually denoted by ϵ, and the trivial decomposition by ν; thus $\mu/\epsilon = \mu$, and $\mu/\nu = \delta$ (the δ-measure).

A system of measures $\{\mu_C\}$ (C runs through the set of elements of a measurable decomposition ξ) on the space $(\Omega, \mathfrak{A})$ is said to be a system of conditional measures corresponding to ξ if $\mu_C C = 1$ for any $C \in \xi$, and for any $A \in \mathfrak{A}$ Fubini's formula holds:

$$\mu A = \int_{\Omega/\xi} \mu_C A d(\mu/\xi).$$

If $(\Omega, \mathfrak{A}, \mu)$ is a Lebesgue space, then there is a system of conditional measures for each measurable decomposition of this space.

Lattice operations $\vee$ and $\wedge$ are introduced in a natural way in the set of measurable decompositions of a Lebesgue space $(\Omega, \mathfrak{A}, \mu)$; if ξ and η are measurable decompositions, then $\xi < \eta$ means that $\xi \neq \eta$ and η is a subdecomposition of ξ. With this order relation the set of all decompositions is a complete lattice set (in the terminology of Bourbaki). A decomposition η is said to be a complement of the decomposition ξ if $\xi \vee \eta = \epsilon$ and $\xi \wedge \eta = \nu$. We usually write $\xi\eta$ for $\xi \vee \eta$.

Let ξ and η be measurable decompositions of a Lebesgue space $(\Omega, \mathfrak{A}, \mu)$. We consider the canonical mapping $\pi_\xi \times \pi_\eta \colon \Omega \longrightarrow \Omega/\xi \times \Omega/\eta$. If the image $\mu(\pi_\xi \times \pi_\eta)^{-1}$ of μ under this mapping coincides with the measure $\mu/\xi \times \mu/\eta$, then we say that ξ and η are independent. And if $\mu(\pi_\xi \times \pi_\eta)^{-1}$ is absolutely continuous with respect to $\mu/\xi \times \mu/\eta$, then ξ and η are said to be quasi-independent.

Let $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ be Lebesgue spaces. A measure m on the product $(X \times Y, \mathfrak{A} \mathbin{\overline{\otimes}} \mathfrak{B})$ is said to be (μ, ν)-doubly stochastic (or simply doubly stochastic) if $\mu = m\pi_X^{-1}$ and $\nu = m\pi_Y^{-1}$.

The characteristic property of Lebesgue spaces, which distinguishes them among all spaces with measures defined on countably generated σ-algebras, is their completeness with respect to any (or some kind of) basis. This means that the mapping determined by a basis $\{B_k, k = 1, \dots\}$ from the set Ω into the countable product $K = \{0, 1\}^\omega$ assigning to each point $x \in \Omega$ the element of K with kth coordinate equal to 1 or 0, depending on whether $x \notin B_k$, carries Ω into a set that is measurable (and of full measure) with respect to the Borel σ-algebra on the compact set K, completed with respect to the measure on K having the same marginal distributions with respect to each coordinate decomposition as the marginal distributions of the measure μ itself with respace to the decomposition $\xi_k = \{B_k, \Omega \backslash B_k\}$; and analogously for any finite sets of coordinates. (On a compact set any compatible system of marginal distributions is generated by a measure.)

Therefore, the countably generated measure spaces that are not Lebesgue spaces are isomorphic to nonmeasurable subsets of outer measure 1 in the compact metric space K with a Borel measure (or in a segment with Borel measure) and are thus nonreal objects. This remark justifies restriction to the class of Lebesgue spaces, although many (but not all) of the proofs of the propositions to follow are true for arbitrary or arbitrary countably generated measure spaces.

A measure of the form $\alpha\mu$, where $0 \leqslant \alpha \leqslant 1$ and μ is an ordinary measure, is called a subprobability measure.

2. Each countably generated measure space can be completed and thereby turned into a Lebesgue space. Many important structures associated with a measure space are not changed under completion, but at the same time regenerate the measure space in a canonical way in the class of complete spaces. The metric structure of the equivalence classes of measurable subsets is such a structure (the distance is defined as the measure of the symmetric difference). If $\{B_k\}$ and $\{B'_k\}$ are bases in two measure spaces, and the aforementioned measures on the compact set $K = \{0, 1\}^\omega$ corresponding to these bases coincide, a canonical isomorphism is thereby established between the completions of these spaces. Let $S_\mu = S(\Omega, \mathfrak{A}, \mu)$ be the space of all (classes of μ-equivalent) measurable functions on $(\Omega, \mathfrak{A}, \mu)$. Then S_μ is a ring and, simultaneously, a partially ordered space (a Riesz space). A ring isomorphism of two spaces S_{μ_1} and S_{μ_2}, as well as an isomorphism of their Riesz lattices, generates in a canonical way an isomorphism of the metric structures of the measurable sets and an isomorphism mod 0 of the corresponding Lebesgue spaces $(\Omega_1, \mathfrak{A}_1, \mu_1)$ and $(\Omega_2, \mathfrak{A}_2, \mu_2)$, so that a study of Lebesgue spaces can be carried out in terms of the corresponding rings of measurable functions or the Riesz lattices. Analogously, an isomorphism of the rings of bounded measurable functions $L^\infty(\Omega_1, \mathfrak{A}_1, \mu_1)$ and $L^\infty(\Omega_2, \mathfrak{A}_2, \mu_2)$ (and also of the natural Riesz lattices on these spaces) implies in a similar canonical way an isomorphism of the corresponding Lebesgue spaces. We remark that an isomorphism of the same spaces in the sense of an isomorphism of linear topological spaces does not lead to an isomorphism of the measure spaces $(\Omega_1, \mathfrak{A}_1, \mu_1)$ and $(\Omega_2, \mathfrak{A}_2, \mu_2)$, as is shown by the example of the spaces l^∞ and $L^\infty[0, 1]$.

3. We dwell in somewhat more detail on the case when the measure space under consideration is equipped with the additional structure of a linear space: on the structure of a linear measure space. A linear measure space is a Lebesgue space that is simultaneously a linear space, where the σ-algebra is generated by some collection of linear forms. If, as is natural, we consider linear measure spaces to within an isomorphism (i.e. we identify any two measure spaces for which there exist linear subsets of full measure that are isomorphic with respect to both structures simultaneously), then it turns out, for example, that there are as many Gaussian measure spaces as there are dimensions; in particular, there exists only one infinite-dimensional countably generated linear Gaussian measure space. In the following we understand a linear measure space to be such a space, considered to within the isomorphism described. However, when necessary, we speak of a particular linear measure space.

A linear measure space is specified if we are given a weak distribution, i.e. a compatible system of finite-dimensional distributions. A weak distribution is a linear mapping of a linear set F of measurable linear functionals generating the σ-algebra into some space $S_\mu = S(\Omega, \mathfrak{A}, \mu)$; that is, to each finite set of measurable linear functionals we assign a distribution in $\mathbf{R}^n$, with the observance of the obvious compatibility conditions. A weak distribution can be described by specifying on the space F the characteristic functional defined as the restriction to the image of F of the fundamental functional χ on S_μ defined by

$$\chi(f) = \int_\Omega e^{if(\omega)} d\mu.$$

It is convenient to state various characteristics of a weak distribution in terms of the functional χ.

If (E, μ) is a linear measure space, then we can consider the subset L of the space $S(E, \mu) = S_\mu$ consisting of all the measurable linear functionals. When a Gaussian random process is considered, it is sometimes convenient to work with a Gaussian measure in a linear space, and sometimes convenient to work with a so-called Gaussian subspace $H \subset L^2(\Omega, \mu) \subset S(\Omega, \mu)$ consisting of Gaussian measurable functions. Any two Gaussian subspaces of the Hilbert space $L^2(\Omega, \mu)$ that separate points are carried one into the other by an orthogonal transformation that is the adjoint of some automorphism of the Lebesgue space (Ω, μ). In any case a closed (in $L^2(\Omega, \mu)$) Gaussian subspace is a maximal Gaussian subspace consisting of the functions that are measurable with respect to a certain decomposition (into the maximal subsets on which all the functions of the Gaussian subspace are constant).

Let (E, γ) be a linear space (of countable dimension) with Gaussian measure, and L the space of all measurable linear functionals. Obviously, $L \subset L^2(E, \gamma)$. If $H^* \subset (E, \gamma)$ consists of all the continuous (with respect to the Hilbert norm) linear forms on L (the so-called kernel), then $H^* \subset E_0$ for any realization of (E, γ) in the form of a concrete measure space (E_0, γ). Moreover, H^* is a maximal subset having this property, and $\gamma H^* = 0$. The unit sphere of the Hilbert space H^* (the polar of the unit sphere of the Hilbert space $L \subset L^2$) is called the variance ellipsoid of the Gaussian measure γ. The kernel H^* coincides with the set of quasi-invariant translations of the space (E, γ). If E_0 is a Hilbert space, then the imbedding $H^* \subset E_0$ is Hilbert-Schmidt. Of course, the space H^* can be defined in a similar way for any measure μ for which $L \subset L^2(E, \mu)$.

Finally, we define the concept of barycenter of the measure μ in a linear measure space (E, μ). The point $x_\mu \in E$ is said to be the barycenter of μ if the set L of measurable linear functionals on (E, μ) is a subset of $L^1(E, \mu)$, and for any $f \in L$

$$\int_E f d\mu = (x_\mu, f).$$

The barycenter of the normalized restriction of a measure for which $L \subset L^2(E, \mu)$ to any subset of positive measure belongs to the kernel H^*. We mention here the Choquet-Kreĭn-Mil'man theorem: each point of a convex compact subset of a linear topological space is the barycenter of a measure on the set of extreme points of this compact set.

CHAPTER II

SAMPLE FUNCTIONS OF RANDOM PROCESSES. CORRELATION THEORY

§1. Statement of the problem. 0. Brief survey of the results of Chapter II (12). 1. Linear measure spaces, in particular, Gaussian (12). 2. The Gaussian subspace H and its kernel H^*. The zero-one law for Gaussian measures (13). 3. Problem of the Banach properties of realizations. Description of the features encountered in studying the problem of extension of a weak distribution to a measure in a Banach space (σ-algebras, distribution of the polar of the unit ball, etc.) (15). 4. Features of the reflexive case (20). 5. Every Banach space can be "sandwiched" between the space of bounded and the space of continuous linear forms on some subset of the dual space (21).– §2. Problem of the size of a convex solid angle in Hilbert space and the infinite-dimensional Cauchy distribution. 1. Brief survey of the contents of the section. The *GB* property depends on the intrinsic geometry of the set K. Problem of positivity of a solid angle in Hilbert space (23). 2. Problem of the measure of a solid angle polar to a given one (24). 3. Spherically invariant measures. Decomposition into rays; the conditional measures of a Gaussian space under a decomposition into rays are δ-measures. Schoenberg's theorem (26). 4. Connection between the problem of the measure of a solid angle and the problem of Cauchy measure (28).– §3. Cauchy measure and Gaussian measure. 1. Cauchy measure as a weighted mean of Gaussian measures (30). 2. Mutual estimates of Cauchy and Gaussian measures (31). 3. Problem of extension of Cauchy and Gaussian measures. The zero-one law for Cauchy measure. Cauchy measure of star-shaped subsets (32). 4. Monotonicty of Cauchy measure (32).– §4. The *GB*-property and mixed volumes. 1. Mixed volumes (33). 2. The functional $h_1(K)$, $K \subset H$, and its properties (34). 3. The relation between $h_1(K)$ and $\sup_{\omega \in K} x(\omega)$ (35). 4. Estimates of γK° in terms of $h_1(K)$. If $0 \in \operatorname{conv} K$ and $\gamma K^\circ > 1/2$, then $h_1(K) < \infty$. Derivation of an upper estimate for the quantity $h_1(K)$ in terms of the probability γK° of exit from the unit level (36). 5. The estimate $\kappa K^\circ \leqslant \kappa(h_1(K) I_0)^\circ$, where I_0 is a unit segment with endpoint at zero. Mutual estimates of the Gaussian and Cauchy measures and the value of the functional h_1. (38). 6. Theorem 1: equivalent restatements of the *GB*-property in terms of the value of $h_1(K)$, the Cauchy measure, and the magnitude of the solid angle (40). –§5. Monotonicity of the functional h_1. 1. Schlaefli's formula (41). 2. $(d\kappa/d\epsilon)(\epsilon K)^\circ|_{\epsilon=0} = -h_1(K)/\pi$ (42). 3. Local monotonicity of the functional h_1: $\partial h_1(K)/\partial l_{ij} \geqslant 0$ (l_{ij} is the length of an edge of the polyhedron K) (43). 4. Global monotonicity of the functional $h_1(K)$. Theorem 2: monotonicity of h_1. Corollary on comparison of the values of the functional h_1 (45).– §6. Mixed volumes and continuity of paths. 1. Structure of the space of continuous forms. Absence of continuity implies certain exit from some level (49). 2. Definition of the oscillation $\delta(K)$. Gaussian measure of the polar of the intersection of a set with a subspace and of the polar of a projection of it. Corollary: oscillation of an intersection and a projection. Monotonicity of Gaussian measure of convex sets under dilations. Proof of the relations $h_1(\mathrm{Pr}_j K) \searrow 2\pi\delta(K)$ and $h_1(K \cap L_j) \searrow 2\pi\delta(K)$. Theorem 3: equivalent restatements of the *GC*-property of a set K in terms of the values of the functional h_1, its projections, and sections (50).– §7. ϵ-entropy conditions. 1. Estimate of the values of h_1 for a simplex, $h_1(S_n) \leqslant (2\pi \ln n)^{1/2}$ ($n \geqslant 3$) (54). 2. Definition of the entropy index ρ. The relations

$$\rho(K) < 2 \Rightarrow \sum 2^{-k} (\log_2 N(K, 2^{-k}))^{1/2} < \infty,$$

and $h_1(K) \leqslant 22 \, \Sigma \, 2^{-k} (\log_2 N(K, 2^{-k}))^{1/2}$. Corollaries: $\rho(K) < 2 \Rightarrow K \in GB$, $K \in GC$, and γK° is estimated (54). 3. ϵ-lattices. Lower estimate of $h_1(K)$ in terms of the cardinality of an ϵ-lattice.

Corollary: if $\limsup \epsilon^2 \ln M = \infty$ (in particular, if $\rho(K) > 2$), then $h_1(K) = \infty$ and $K \notin GB$. Lower estimate of the oscillation with respect to the ϵ-entropy (56). 4. Impossibility of a complete solution of the problem in terms of ϵ-entropy (58). 5. Theorem 4: summary of ϵ-entropy conditions for boundedness, continuity, and estimates of the values of the functional h_1 and the magnitude of the oscillation δ (59).—§8. The non-Gaussian case. 1. Statement of the problem (59). 2. The s-characteristic of a set $K \subset H$. Theorem 5: complete description of the wide-sense processes that always have continuous realizations. The alternative: continuity—certain unboundedness (60).—§9. Remark on Borel realizations.

§1. Statement of the problem

0. In this chapter we study questions connected with the problem of extension of a weak distribution to a measure in a Banach space. Basically, we consider random processes with second moments and certain other processes that are closely related to them (for example, Cauchy processes; see below). The fundamental aim of the investigations is to determine conditions permitting us to judge from the correlation function of a process whether the sample functions of the random process belong with probability 1 to some Banach space of functions on a parameter set (sometimes understood in a somewhat extended sense).

In the case of Gaussian processes with zero mean the correlation function is known to carry complete information about the process, and our approach permits us in many cases to formulate necessary and sufficient conditions for the possibility of extending a weak Gaussian distribution to a measure in a Banach space. Study of the geometric characteristics introduced for a process leads to the establishment of several new inequalities for the Gaussian measures of convex sets that, in turn, allow us to obtain convenient (for checking) conditions in terms of the ϵ-entropy of the parameter set. A special role is played by conditions for the boundedness of realizations of a Gaussian process and for their continuity with respect to the metric that arises naturally on the parameter set.

For an arbitrary process with second moments and with given correlation function we prove the alternative: either its sample distribution functions are bounded and continuous (in the natural metric) with probability 1, or the finite-dimensional distributions of the process, which are compatible with the given correlation function, can be chosen in such a way that the realizations of the process are certainly unbounded. A criterion is given for checking which case actually holds.

1. As usual, let $(\Omega, \mathfrak{A}, P)$ be a probability space. We assume that the σ-algebra $\mathfrak{A}$ is countably generated (as is practically always the case) and that this measure space is complete ([92]; the terminology is discussed in the first chapter). We recall that an incomplete measure space is an object that is just as imaginary as a subset of a segment that is not Lebesgue measurable. Let $\{x_t(\omega), t \in T\}$ be a random process in the usual sense. We consider the space S_P of all measurable functions on the space $(\Omega, \mathfrak{A}, P)$, which is metrizable with respect to convergence in probability, and the closed linear subspace $L = L_P \subset S_P$ of it generated by the elements $x_t \in S_P$, $t \in T$. If we now assign to each finite set of elements $x_1(\omega), \ldots, x_n(\omega) \in L$ their joint distribution (a probability measure in $\mathbf{R}^n$), we get a compatible system of finite-dimensional distributions, or, in other words, a weak distribution or generalized random process [35], which determines a linear measure space (E, P) to within an isomorphism (see [141], and also Chapter I of the present book). It is always possible (by Kolmogorov's theorem [66] on the extension of a compatible system of finite-dimensional distributions to a measure, or by the Minlos–Sazonov theorem [77], [103]) to select a concrete form for the linear space E with measure P such that the elements of the space L are measurable linear functionals (for more detail see [141] and [139]). It frequently happens that it is not exactly the concrete form of the linear space E that is important, but only the structure of the measure space (E, P), which is completely determined by the original ordinary random process $x_t(\omega)$, $t \in T$. It is convenient to represent the passage from the measurable functions $x_t(\omega)$ to measurable linear functionals on a linear measure space as a process of linearization of the set Ω; the linear space E is the very same set Ω, and comparison of the measurable functions on (Ω, P) and the measurable linear functionals on (E, P) with the same distributions establishes (see Chapter I) an isomorphism mod 0 of the measure spaces (Ω, P) and (E, P).

It is easy to see that, to within an isomorphism of linear measure spaces, there is a unique linear space with an infinite-dimensional (of countable dimension) Gaussian measure, and from our point of view (which differs from the usual) it is frequently appropriate to speak, for example, not of "Gaussian measures in Hilbert space" or in some concrete separable Banach space, but simply of Gaussian measure (i.e., of linear Gaussian measure space). This remark is relevant, in particular, to the problem of the absolute continuity of the measures corresponding to two generalized random processes, the solution of which is, of course, determined only by the weak distributions and does not depend on the concretization of the space E. (The concept of isomorphism of measure spaces is discussed in [115]. The axiomatic approach is studied in Veršik's paper [139].)

A systematic feature of the following presentation is the consistent adherence to the point of view that in investigating the question of extending a weak Gaussian distribution to a measure it is more natural to consider, not the various Gaussian weak distributions on a fixed Banach space with the purpose of distinguishing those of them that can be extended to measures in this space, but the various Banach subspaces of linear Gaussian measure space with the purpose of determining their Gaussian measure.

2. We shall now study Gaussian measures in more detail. Let $\{x_t(\omega), t \in T\}$ be a Gaussian random process; this means that $L \subset S_P$ is a Gaussian subspace, i.e. a closed linear subspace H of the Hilbert space of random variables $L^2(\Omega, \mathfrak{A}, P)$ that consists of measurable functions with Gaussian distributions and separates the points of Ω. (If the latter condition were not satisfied, it would be sufficient to pass to the corresponding quotient space.) As shown by Veršik [140], a Gaussian subspace of $L^2(\Omega, \mathfrak{A}, P)$ with purely continuous measure P is unique to within an orthogonal transformation that is the adjoint of some automorphism of the measure space $(\Omega, \mathfrak{A}, P)$. The space H of Gaussian variables will always be equipped with the Hilbert norm induced by the inclusion $H \subset L^2(\Omega, \mathfrak{A}, P)$, so that the value of the covariance function $r(s, t)$, $s, t \in T$, coincides with the value of the usual scalar product of the elements x_t and x_s of H.

Everywhere in the following when we speak of Gaussian random processes, we shall for simplicity have in mind Gaussian processes with zero mean value, so that the norm of an element x of H coincides with its standard deviation σ_x.

Let (E, γ) be a linear Gaussian measure space. The linear space E (considered to within the complement of a measurable linear subspace of full measure) necessarily contains the subspace $H^* \subset E$ consisting of all continuous linear functionals on the Hilbert space H of Gaussian variables. (The subspace H^* consists of all the vectors such that translations by them preserve these measures, and therefore it does not depend on a concrete realization of the linear measure space.) In the infinite-dimensional case we always have $\gamma H^* = 0$. Thus, the set H of all measurable linear functionals on a linear Gaussian measure space coincides with the set of continuous linear functionals on the Hilbert space H^*.

PROPOSITION 1 (the zero-one law for Gaussian measures; cf. [115] and [121]).

1) *Let (E, γ) be a linear space with Gaussian measure γ, and $E_1 \subset E$ a measurable linear subspace. Then either $E_1 = E$ (mod γ), or $E_1 = \varnothing$ (mod γ). In other words, either $\gamma E_1 = 1$, or $\gamma E_1 = 0$.*

2) *Let $\{L_n\}$ be a sequence of closed subspaces of finite defect in the Gaussian space H. If the set $M \subset (E, \gamma)$ is such that for almost all $\omega \in E$ the inclusion $\omega \in M$ is determined by the values on the element ω of the functions $x(\omega) \in H$ that are orthogonal to all the subspaces L_n, beginning with an arbitrarily large number, then either $\gamma M = 0$, or $\gamma M = 1$.*

PROOF. 1) We choose in H an arbitrary orthonormal basis $\{e_k(\omega), k = 1, 2, \ldots\}$. We construct the mapping $e: (E, \gamma) \longrightarrow \mathbf{R}^\infty$ that acts according to the formula $e(\omega) = (e_1(\omega), e_2(\omega), \ldots) \in \mathbf{R}^\infty$. Under this mapping the Gaussian measure γ on E goes over into the standard countable-dimensional Gaussian measure γ_0: $\gamma_0 = \gamma e^{-1}$, where γ_0 is the product of a countable number of standard one-dimensional Gaussian measures. The space H in this concrete realization of the linear Gaussian measure space (E, γ) coincides with the space l^2, and the characteristic functional, which on the original space H equals

$$\chi_\gamma(x) = \int_E \exp(ix(\omega))\, d\gamma = \exp\left(-\frac{1}{2}\|x\|_H^2\right),$$

can now be written in the form

$$\chi_{\gamma_0}(\{x_k\}) = \exp\left(-\frac{1}{2}\sum x_k^2\right),$$

where $(x_k, k = 1, \ldots) \in l^2$.

If there existed a measurable linear subspace $E_1 \subset (E, \gamma)$ of intermediate γ-measure, we could associate with it a γ_0-measurable subspace $L_1 \subset \mathbf{R}^\infty$ of the same intermediate measure. Let $l^2 = \{(y_k, k = 1, \ldots)\} \subset \mathbf{R}^\infty$ be the space of sequences with convergent series of squares. (The space $l^2 \subset (\mathbf{R}^\infty, \gamma_0)$ is the space H^* mentioned above.)

We consider two cases.

I. Suppose that there is a vector $y \in \mathbf{R}^\infty$ such that $y \in l^2$, but $y \notin L_1$. It is well known ([30], Chapter IV, §5.2, Theorem 3) that the Gaussian measure γ is quasi-invariant with respect to translations by elements of the set $H^* = l^2 \subset \mathbf{R}^\infty$, from which it follows that $\gamma_0 L_1 = 0$, since in the opposite case we would have, by quasi-invariance, an uncountable number of disjoint sets $L_1 + \lambda y$, $\lambda \in \mathbf{R}$, with positive measure.

II. Suppose that $H^* \subset L_1$; then all the finite-dimensional subspaces spanned by elementary basis vectors of the space $\mathbf{R}^\infty$ (i.e., the vectors of the form $(0, \ldots, 0, 1, 0, \ldots) \in \mathbf{R}^\infty$) are also contained in L_1. This means that for an element $y \in \mathbf{R}^\infty$ the property of belonging to L_1 is determined only by the members of the sequence $y = (y_1, y_2, \ldots)$ that are arbitrarily far out. In other words, we apply the Kolmogorov zero-one law [66] to the event $L_1 \subset (\mathbf{R}^\infty, \gamma_0)$.

The proof of 1) is concluded. By passing to the space $(\mathbf{R}^\infty, \gamma_0)$ the assertion 2) is reduced directly to the Kolmogorov zero-one law. ●

Further, we mention that the topology defined by the norm and the topology defined by the inclusion $H \subset S$ (the topology of convergence in probability) coincide on the Gaussian space H. See, for example, [94].

3. Let B be a Banach space of functions on the parameter set T. We are interested in the question of whether it can be assumed that the sample functions of a particular process (for definiteness we speak first of Gaussian processes) belong to the space B with probability 1.

As mentioned, the parameter set T can be assumed to lie in the Gaussian space H (we assign to a point $t \in T$ the point $x_t \in H$). The imbedding $T \subset H$ permits us to carry a metric, and, together with it, a Borel structure, from the space H to the set T. To connect the Banach space structure on B and the structure of the weak distribution we assume that some subspace L of the space B^* of continuous linear functionals on B that separates the points of B consists of measurable linear functionals, i.e., is identified with some (generally nonclosed) subspace of H. As a rule (but not always) the functionals "value at the point $t \in T$" turn out to be continuous on B. The assumption that L does not coincide with B^* is interesting in the nonreflexive case, as will be clear from the examples given below.

First of all, we show that without loss of generality we can assume that the subspace $L \subset B^*$ is closed, i.e. is a Banach space, and that the image of its unit ball under the inclusion $L \subset H$ is precompact in H. We prove a more general assertion that is relevant not only to Gaussian weak distributions, and then we give some examples.

PROPOSITION 2. *Let B be a Banach space, $L \subset B^*$ a linear subspace that separates the points of B (i.e., is weakly dense in B^*) and contains a countable subset that is norm dense in it, and μ a probability measure on the σ-algebra $\mathfrak{A}_L$ of subsets of B generated by the functionals in L. Let $\overline{L}^B \subset B^*$ be the smallest sequentially weakly (in the topology $\sigma(B^*, B)$) closed subspace of B^* that contains L. Then the following assertions are true:*

1) *$\mathfrak{A}_L$ is countably generated.*

2) *The elements of $\overline{L}^B$ are measurable with respect to $\mathfrak{A}_L$ and can thus be considered as elements of the space L_μ of all μ-measurable linear functionals.*

3) *The subspace* $\overline{L}^B$ *is closed in* B^* *and is thus a Banach space.*

4) *The image of the unit ball* $V_{\overline{L}^B}$ *in* $\overline{L}^B$ *under the mapping* $\overline{L}^B \longrightarrow L_\mu \subset S_\mu$ *established in* 2) *is a bounded subset of the linear metric space* S_μ.

PROOF. The statement 1) follows from the assumption about the separability of the set $L \subset B^*$, since the pointwise limit of a sequence of functions is measurable with respect to a σ-algebra with respect to which all the members of the sequence are measurable. To prove 2) we observe that $\overline{L}^B$ can be obtained as the union of the nondecreasing transfinite sequence of successive sequential weak closures of L over all countable transfinite ordinal numbers (the Borel superstructure over L, i.e. the union of the Baire classes constructed beginning from the elements of L). Since there is no countable sequence of ordinal numbers that is cofinal in the set of all countable transfinite ordinal numbers, each weakly convergent sequence of elements of the space $\overline{L}^B$ thus constructed actually belongs to one of the spaces in the union (whose transfinite number can be determined as the supremum of the countable set of numbers of the spaces to which the members of the convergent sequence belong), and the limit function then belongs to the space that is next in order. We have proved the sequential weak closedness of the space $\overline{L}^B$ thus constructed; moreover, it is obvious that this space is the smallest sequentially weakly closed superspace of L, and its elements are measurable with respect to the σ-algebra $\mathfrak{A}_L$, since they appear in the Baire hierarchy over L. The statement 3) is now almost obvious; in fact, it is sufficient to verify the norm closedness for sequences, and norm convergence implies weak convergence. Before proceeding to the proof of the assertion 4), we observe that the mapping $\overline{L}^B \longrightarrow L_\mu$ can be assumed to be an imbedding (passing in the opposite case to a quotient space), i.e., we can assume that the measure μ is such that no $\sigma(B, L)$-closed subspace has full measure (taking in the opposite case instead of B the smallest such subspace, which exists and coincides with the polar of the subspace $L_0 \subset L$ of linear functionals having a δ_0-distribution). The space L_0 is norm closed in L (since it is sequentially weakly closed) and is clearly separable, so that its polar can be represented as the intersection of a countable number of hyperplanes, each of which is a subset of full measure.

We prove the boundedness of the set $V_{\overline{L}^B} \subset S_\mu$. Boundedness of some set $A \subset S_\mu$ means that for any neighborhood V of zero in S_μ there is a number $\lambda > 0$ such that $\lambda A \subset V$. In our case a fundamental system of neighborhoods in S_μ is formed by the sets of the form

$$V_\varepsilon = \{x(\omega) : \mu\{\omega \mid x(\omega) \mid \leqslant \varepsilon\} > 1 - \varepsilon\}.$$

For fixed $\epsilon > 0$ there is a number $r > 0$ such that $\mu(rV_B) > 1 - \epsilon$. Since for $x(\omega) \in V_{\overline{L}^B}$ we have that $|x(\omega)| \leqslant 1$ for $\omega \in V_B$ and $x \in V_{\overline{L}^B}$, we get from the bilinearity with respect to x and ω that $|x(\omega)| \leqslant \epsilon$ for $\omega \in rV_B$ and $x \in \epsilon r^{-1} V_{\overline{L}^B}$, which concludes the proof of Proposition 2. ●

REMARKS. 1) If B is a separable reflexive space (in particular, Hilbert space), then any linear set of functionals that separates points (i.e., is weakly dense in B^*) is also norm dense (for convex subsets of the dual space closedness and weak closedness are equivalent [12]), so that we can always assume that the weak distribution is given on the whole dual space.

2) In this case the image of the unit ball $V_{\overline{L}^B}$ of $\overline{L}^B$ in the space S_μ is compact in the topology of convergence in measure. For, as noted above, we can assume that the mapping $\overline{L}^B \longrightarrow S_\mu$ is monomorphic. Weak convergence implies convergence in measure, but in the weak topology the ball $V_{\overline{L}^B}$ is compact; hence it is compact in the topology of convergence in measure. In particular, in the case of a Gaussian measure in a separable reflexive space we can assume that the unit ball of B^* is a compact subset of H.

3) In the case of an arbitrary Banach space B the norm closure of the space $L \subset B^*$ does not necessarily coincide with the whole space B^*, even if B^* is separable (and L separates the points of B).

EXAMPLE. $B = c$ is the space of sequences converging to a limit. In this example $B^* = l^1 \times \mathbf{R}$, and we can take $L = \overline{L} = l^1 \times \{0\} \subsetneq B^*$. We note that the space l^1 is not sequentially weakly $\sigma(l^1, c)$-complete, and that the norm closed subspace $L \subset B^*$ is weakly dense in B^*.

4) The last observation can be given a general character: if the space B is separable, then always $\overline{L}^B = B^*$. Indeed, in this case $\overline{L}^B$ is weakly closed in B^*, because its intersection with any closed ball in B^* is sequentially weakly closed (as the intersection of two sequentially weakly closed sets) and hence weakly closed, since on the bounded sets in the dual space of a separable space the weak topology is metrizable, and it is sufficient to verify weak closure for weakly convergent sequences. On the other hand, by a theorem of Banach ([6], Chapter VIII, §5), the weak closedness of the intersection $\overline{L}^B \cap V_{B^*}$ implies the weak closedness of $\overline{L}^B$ and, consequently, by the weak density of it in B^* (totality), the condition $\overline{L}^B = B^*$.

5) The separability of B in the preceding item cannot be omitted.

EXAMPLE: $B = l^\infty, L = l^1$. Since l^1 is sequentially weakly $\sigma(l^1, l^\infty)$-complete [6], we have $\overline{L}^B = l^1$, although l^1 is weakly dense in $(l^\infty)^*$. We remark, further, that, although $\overline{L}^B$ does not necessarily coincide with B^* ($l^1 \subsetneq (l^\infty)^*$), it can happen that each functional in $(l^\infty)^*$ coincides almost everywhere with some functional in l^1 (the σ-algebra is assumed to be complete with respect to the given measure). In particular, this will be the case if in l^∞ we consider a Gaussian measure γ corresponding to a sequence of independent trials; moreover, $\gamma c_0 = 0$ (the last observation is due to B. S. Cirel'son).

6) For measures in a nonreflexive space the image in S_μ of the unit ball of $\overline{L}^B$ need not be totally bounded.

EXAMPLE. Let $B = l^\infty, L = l^1$, and let the measure μ be determined by the characteristic functional on E: $\chi(x) = \chi((x_1, \ldots)) = \Pi_1^\infty \cos x_k, x = (x_1, \ldots) \in l^1$ (i.e., corresponding to the distribution of a sequence of independent random variables taking the values $+1$ and -1 with probabilities $1/2$). In this case the space of all measurable linear functionals on the space (l^∞, μ) coincides with the space l^2, in which the set of unit vectors (i.e., the independent random variables mentioned) is contained in the unit ball of l^1.

7) If the measure μ is Gaussian, then, as will be clear from the following, the set $V_{\overline{L}^B} \subset H \subset S_\mu$ is relatively compact (totally bounded). In the following the order of its ϵ-entropy in the metric of H is estimated.

8) Even in the case of a Gaussian measure the set $V_{\overline{L}^B} \subset H$ need not be compact. Its closure in L_μ (or, what is the same, in H) consists of all μ-measurable linear forms $x(\omega)$ on the space (B, μ) for which ess $\sup_{\omega \in V_B} |x(\omega)| \leqslant 1$.

However, contrary to the case of a separable reflexive space B, in the general situation it can happen that some of these forms are not μ-equivalent to any "Baire" linear form over L, i.e. to any element of $\overline{L}^B$.

EXAMPLE. We consider the sequence of Gaussian variables $e_k(\omega) = c_k e'_k(\omega) + e_0(\omega)$, $k = 1, \ldots$, where $e_0(\omega)$ and $e'_k(\omega)$ are ortho-Gaussian.

If $c_k = O((\ln k)^{-1/2})$, then the realization of such a random process belongs to l^∞. The space L of measurable continuous linear forms on l^∞ in this example can be represented as the space $\mathbf{R}^\infty$: the linear span of the "coordinate functionals"; that is, we consider on Ω the smallest σ-algebra with respect to which all the random variables e_k, $k = 1, \ldots$, are measurable. The norm closure in $(l^\infty)^*$ of the space $\mathbf{R}^\infty = L$ is the space l^1, which is well known to be sequentially weakly complete, and consequently it coincides with $\overline{L}^B$. It is easy to verify that the unit ball of the space $H^*_\gamma \subset l^\infty$ (the "variance ellipsoid") for this Gaussian measure is the set

$$E^* = \left\{\omega = (y_1, \ldots) \in l^\infty : \omega = \lambda_1^2 z + \lambda_2 (1, 1, 1, \ldots),\right.$$

where

$$z \in E' = \left\{\omega : \sum_{k=1}^{\infty} \frac{y_k^2}{c_k^2} \leqslant 1\right\}, \quad \left.\lambda_1^2 + \lambda_2^2 \leqslant 1\right\},$$

and the pre-Hilbert norm induced on l^1 by its imbedding in H is given by

$$\| x \|^2 = \sum_{k=1}^{\infty} c_k^2 x_k^2 + \left(\sum_{k=1}^{\infty} x_k\right)^2$$

(we can also say that this functional defines the imbedding of $l^1 = \overline{L}^B$ into H). Obviously, $e_k \longrightarrow e_0$ (in probability, and hence in the mean square), i.e., $e_0 \in H$. On the other hand, the functional $e_0(\cdot)$ cannot be identified with any element of l^1. Indeed, when $c_k = o((\ln k)^{-1/2})$, we have for almost all ω that $c_k e'_k(\omega) \longrightarrow 0$ and $e_k(\omega) \longrightarrow e_0(\omega)$, i.e., $\gamma c = 1$, where, as usual, $c \subset l^\infty$ denotes the space of sequences converging to a limit. The functional $e_0(\cdot)$ is defined on the space c by the equation $e_0(\omega) = e_0((y_1, \ldots)) = \lim y_k$ and is equal to zero on the space c_0 of sequences converging to zero, i.e., on a total set of functionals in l^∞, and therefore it is not generated by an element of l^1. A continuous linear functional with unit norm defined on the set of full measure $c \subset l^\infty$ can, by the Hahn-Banach theorem, be extended to the whole space l^∞ with preservation of linearity and norm (a "Banach limit" of the bounded sequences), but each such Banach limit is nonmeasurable with respect to the σ-algebra $\mathfrak{A}_{l^1}$, and the construction of particular examples of such functionals is just as problematical as the construction of particular examples of subsets of a segment that are not Lebesgue measurable. In the case when $c_k = (\ln (k + 1))^{-1/2}$ and $\gamma c = 0$, the value of the functional e_0 can be obtained for almost all $\omega \in l^\infty$ as the limit of the sequence $f_n(\omega) = (\Sigma_1^n e_k(\omega))/n \in l^1$, which a fortiori does not converge in l^1.

9) We mention that a norm closed subspace B_1 of (B, γ) that contains the variance ellipsoid E^*_γ can have zero measure.

EXAMPLE. $B = l^\infty, B_1 = c_0, E^*_\gamma = \{\omega: \omega = (y_1, \ldots), \Sigma_1^\infty \ln(k+1)y_k^2 \leqslant 1\}, \mathfrak{A}_L = \mathfrak{A}_{l^1}$.

10) If the original random process x_t, $t \in T$, is such that in the Banach space B of functions on T all the functionals of the type "evaluation at a point" are continuous, i.e. the norm $\|\cdot\|_B$ majorizes the topology of pointwise convergence, then the parameter set T is a subset of the space L. In the following it is convenient in many cases to regard the unit ball $V_{\overline{L}^B} \subset L_\mu$ as the parameter set. Of course, if the functionals $x_t \in L_\mu$ are only μ-measurable, and not continuous functions on B, then the true parameter set is not contained in $\overline{L}^B$. Such is the case, for example, in clarifying (with the help of the Minlos–Sazonov theorem, or directly [94]) the question of when the realizations of a Gaussian process on some segment $[a, b]$ belong to $L^2[a, b]$, or, more precisely, when the classes of functions that are Lebesgue-equivalent on $[a, b]$ to realizations of the process x_t belong to L^2.

11) Finally, we touch on the question of measurability of realizations. Let $K \subset S_\mu$ be a convex bounded balanced set whose linear span $L = L(K)$ is a Banach space in the norm $\|\cdot\|_K$. On the set K we consider the σ-algebra $\mathfrak{B}$ induced from the space S_μ. We now consider the space $L^{(*)}$ of all linear forms on L that are bounded on K (linear functionals) and measurable with respect to $\mathfrak{B}$. The space $L^{(*)}$ is a subspace of the dual space L^* of $(L, \|\cdot\|_K)$, but, generally speaking, does not coincide with it.

EXAMPLE. Let $H \subset S_\mu$ be a Gaussian subspace, $e_k \in H$ an orthonormal basis in H, and $K = \{x: |\langle x, e_k\rangle| \leqslant c_k\}$, where $\langle\cdot,\cdot\rangle$ is the intrinsic scalar product in H. Then the space $(L, \|\cdot\|_K)$, where $L = L(K)$, is isometric to l^∞, and $L^{(*)}$ is isometric to l^1 (but not to the whole space $(l^\infty)^*$).

Given a set K of the indicated type, it is natural to consider the question of whether realizations of a random process with parameter set K are almost surely bounded. In other words, we can consider the question of extending the weak distribution that arises to a measure in L^*. The example given shows that L^* can contain functionals that are nonmeasurable with respect to the σ-algebra $\mathfrak{B}$ on the parameter set. However, it turns out that we can limit ourselves to the problem of extending the weak distribution to a measure in the space of Borel linear functionals $L^{(*)}$. Namely, it can be shown that if a weak distribution can be extended to a measure in L^*, then it can be extended to a measure in $L^{(*)}$, i.e. that it is always possible to assume that the realizations in our sense are Borel. (Compare with Doob's theorem on the existence of a measurable modification of a process in [24], p. 61. Doob's theorem coincides with the above assertion neither in the hypothesis nor in the conclusion.)

We consider that $(L^{(*)}, \mathfrak{A}_L, \mu)$ is a Lebesgue space; consequently, it suffices to look for a solution of the problem of finding conditions for the boundedness of realizations of processes with separable (in the metric of S_μ) parameter sets in terms of the linear measure space (E, μ), which is, by definition, a Lebesgue space. Although the subspace $L^{(*)} \subset L^*$ is norm closed, it is $\sigma(L^*, L)$-dense in L^*, and, as shown by 9), the assertion that, always, $\mu^*L^{(*)} = 1$ is not trivial (μ^* is the outer measure) (see §9).

4. In the following we assume that $\overline{L}^{B} = L$, in the opposite case extending the weak distribution from L to $\overline{L}^{B}$. Now it is true that we cannot assume that the Banach space L is separable; but it contains a norm closed weakly dense (in the topology $\sigma(L, B)$) separable subspace, i.e., it is in any case weakly separable. It is clear from the above presentation that the solution of the problem of extending a Gaussian weak distribution to a measure in some Banach space B lies in describing all the variance ellipsoids $E^* \subset B$ corresponding to Gaussian weak distributions that are extendible to a measure. In [15] the author gave necessary and sufficient conditions for the extension of a weak distribution to a measure in the space l^p ($p \leqslant 2$) precisely in terms characterizing the distribution of the variance ellipsoid of a Gaussian measure in this space (see also the more general result of Vahanija in [135] and [136]). The conditions amount to the structural boundedness of the ellipsoid E^* in the corresponding l space, i.e. to the condition $\Sigma_1^\infty \sigma_k^p < \infty$, where σ_k^2 is the variance of the kth coordinate functional. It is possible to describe a certain class of boundedly complete vector lattices [53] for which the structural boundedness of the variance ellipsoid E^* is a necessary and sufficient condition for the extendibility of a Gaussian weak distribution to a measure.

If B is separable and reflexive, then the unit ball V_L of the space L is compact in H [Remark 2) after Proposition 2] and $L = B^*$ and $B = L^*$; that is, B consists of all the linear forms that are bounded on V_L. Therefore, necessary or sufficient conditions for the boundedness (or continuity in the natural metric of the parameter set) of the sample functions are automatically necessary or sufficient conditions for the extendibility of a Gaussian weak distribution to a measure in the separable reflexive Banach space (the unit ball $V_L = V_{B^*} \subset H$ of the dual space serves as the parameter set).

Not so is the situation in the nonreflexive case. If (B, γ) is an arbitrary Banach space with a Gaussian measure, then, generally speaking, the space H of Gaussian variables does not contain a subset $K \subset H$ for which B is the space of all linear forms that are bounded on K or the space of all linear forms that are continuous on K in the topology induced from H.

EXAMPLE. We consider a sequence of independent Gaussian variables ξ_k with variances $\sigma_k^2 \longrightarrow 0$ such that the sample distribution sequence converges to zero with probability 1 (we recall that we always assume, unless a statement to the contrary is made, that the mean values of our Gaussian variables are equal to 0). Let (l^∞, γ) and (c_0, γ) be the spaces of sample distribution sequences l^∞ and c_0 with the Gaussian measure corresponding to this sequence of Gaussian variables and with the σ-algebras generated by the coordinate functionals.

1) *The space H of all measurable linear functionals on the space (l^∞, γ) does not contain a set K such that (l^∞, γ) is precisely the set of all linear forms on K (i.e., on the linear span $L(K)$) that are continuous on K in the topology induced by H.* Indeed, suppose that the spaces $L(K) \subset H$ and l^∞ are in duality (in the opposite case the measure γ is not reproduced by K). Necessarily, $L(K) \subset l^1$, where $l^1 \subset H$ is the space of all measurable linear functionals that are bounded on the unit ball V_{l^∞}. Since $l^\infty = (l^1)^*$, each strongly closed (with respect to the norm $\|\cdot\|_{l^1}$) subspace of l^1 is also

weakly closed and, therefore, cannot be in duality with l^∞. Consequently, $L(K)$ is dense in l^1. On the other hand, since, by assumption, the elements of l^∞ are continuous functions on $L(K) \subset l^1 \subset H$ (in the metric $\|\cdot\|_H$), they are continuous functions also on the closure of K in l^1 in the metric $\|\cdot\|_{l^1}$ (if $x_n \in L(K)$ and $x_n \to x$ in the norm of l^1, then the value on the element $x \in l^1$ of any functional $y \in l^\infty$ is equal to the limit $\lim(x_n, y)$). By the same token, the set $L(K)$ can be assumed to be closed in l^1, and consequently it coincides with l^1. But on l^1 with the metric $\|\cdot\|_H$ (for which $\|e_k(\omega)\|_H = \sigma_k$, where e_k is the kth coordinate functional, i.e., $H \ni e_k \leftrightarrow (0, 0, \dots, 0, 1, 0, \dots) \in l^1$), not all elements in l^∞ are continuous functions; for example, $e_k \xrightarrow{\|\cdot\|_H} 0$, but, if $y = (1, -1, 1, -1, \dots) \in l^\infty$, then (e_k, y) does not converge to a limit. With this, assertion 1) is proved.

2) *The space H of all measurable linear functionals on (c_0, γ) does not contain a set K such that (c_0, γ) is precisely the set of all linear forms that are bounded on K.* Indeed, in the opposite case c_0 would be the space dual to $L(K)$, the latter equipped with the norm dual to the norm of c_0, while c_0 is not the dual of any space (its unit ball does not contain extreme points, whereas the unit ball of any dual space is compact in the weak topology and therefore contains extreme points).

We now consider the space $(l^\infty, \gamma) \times (c_0, \gamma)$. From 1) and 2) it follows at once that the space $H \oplus H$ of all measurable linear functionals does not contain a subset K for which $(l^\infty, \gamma) \times (c_0, \gamma)$ coincides with the set of all continuous or with the set of all bounded (measurable) linear forms on K.

5. Thus, if we do not assume that the space B is reflexive, then the unit ball $V_L \subset H$ of L does not permit the description of the linear forms in B in some standard way, even under the assumption of completeness (the Banach property) for B. The set of all linear forms that are bounded (or bounded measurable) on V_L can be (though not necessarily, even in the nonreflexive case) considerably broader than the set of all continuous (in the metric $\|\cdot\|_H$) linear forms (in the norm dual to the norm of B each bounded form is clearly continuous). However, it turns out that continuity (in the metric $\|\cdot\|_H$) and boundedness are "extremal" properties. We prove that the space B is enclosed (with respect to inclusion) between the spaces of all continuous and all bounded forms on V_L.

PROPOSITION 3. *Suppose that a Gaussian weak distribution is given on the Banach space B by the inclusion $L \subset H$. Then B consists of linear forms that are bounded on V_L, and it contains the set of all linear forms that are continuous on $V_L \subset H$ with respect to the metric $\|\cdot\|_H$.*

PROOF. The first assertion is trivial (it means that $B \subset L^*$). To prove the second one, we use a theorem ([91], Chapter VI, §1, Proposition 2) saying that if two linear spaces (L, B) in duality are given, and A is a weakly closed bounded subset of L, then a linear form y on the linear span $L(A)$ belongs to the completion of B with respect to the A-norm if and only if $y^{-1}(0) \cap A$ is closed in the topology $\sigma(L, B)$. In our case the role of A is played by the unit ball V_L, and if y is a linear form that is continuous on V_L in the metric $\|\cdot\|_H$, then $y^{-1}(0) \cap V_L$ is closed in L in the Hilbert norm, and, along with this, in the topology $\sigma(L, B)$. Indeed, $\sigma(L, B)$ majorizes the topology $\sigma(L, H^*)$ (since $H^* \subset B$), and the latter is the weak topology of the Hilbert space H, which coincides on totally bounded sets [see Remark 7) after Proposition 2] with the topology generated by the Hilbert norm, so that $y^{-1}(0) \cap V_L$, which is closed in V_L in the Hilbert norm, is also $\sigma(L, B)$-closed, and it follows from the norm completeness of B that $y \in B$. ●

REMARK. Actually, in solving the problem of extendibility of a weak distribution to a measure in the Banach space B it suffices to assume that B consists only of linear forms that are measurable on V_L (Borel linear functionals) [see Remark 11) after Proposition 2]. The set of linear functionals in B that are Borel functions on $V_L \subset H$ is always a norm closed (and even sequentially weakly closed) subspace of B.

We summarize what has been said. Suppose that we are given a Banach space B and a Gaussian weak distribution on it determined by a continuous linear mapping (we can assume it is an imbedding) of some set L of linear forms on B (which can be assumed to be sequentially weakly complete) into a separable Hilbert space H of Gaussian random variables. We regard the unit ball $V_L \subset H$ of L as the parameter set of a Gaussian random process. If the realizations of this process are unbounded with probability 1 (or at least with positive probability), then by the same token they do not belong to B with probability 1, i.e., the Gaussian weak distribution does not extend to a countably additive measure in B. Conversely, if the realizations of a process are continuous in the natural metric on V_L with probability 1 (or at least with positive probability), then the weak distribution extends to a Gaussian measure in B. Moreover, if B is the dual space of L (with the norm induced by the duality (L, B)), or the subspace of the dual space of L consisting of all linear forms that are bounded on V_L and measurable with respect to the σ-algebra $\mathfrak{A}_H$ on V_L induced by the imbedding $V_L \subset H$, then the question of extending a Gaussian weak distribution to a measure in B is equivalent to the question of boundedness of realizations of a Gaussian process with parameter set V_L. Therefore, the basic problem considered below is the problem of boundedness of realizations of an arbitrary Gaussian process with zero mean. In passing we also obtain some additional conditions for the continuity of the sample functions (besides the fact that each condition that is necessary for boundedness is also necessary for continuity, and each condition that is sufficient for continuity is also sufficient for boundedness: every linear form defined and measurable on the whole Banach unit ball V_L is a norm-continuous linear functional and, in particular, is bounded on V_L).

Thus, we are interested in the class of subsets K of the Hilbert space H having the property that a Gaussian random process with natural parameter set K has realizations that are bounded in modulus. This property, and with it also the whole class of subsets having it, is called GB, following Dudley [25]. In other words, $K \in GB$ means that any subset of a Gaussian subspace H of the Hilbert space $L^2(\Omega, \mathfrak{A}, P)$ that is isometric to K is structurally bounded from above and below in the space $S(\Omega, \mathfrak{A}, P)$ of all measurable functions (in such a case K is also bounded in L^2). In fact, the structural boundedness of K means that there exists a modification of the process K for which the set $\{x(\omega): x \in K\}$ is bounded for almost all $\omega \in \Omega$.

A subset $K \subset H$ corresponding to a Gaussian process with realizations that are bounded and continuous on K in the natural (induced by H) topology is said to have the GC property. By definition, each GC-set is a GB-set. However, the converse is false: if $e_k \in H$ is an orthonormal sequence in a Gaussian space, then

$$K = \{(2 \cdot \ln k)^{-1/2} e_k\} \in GB,$$

but $K \notin GC$ (the paths of the process K are bounded with probability 1, but, with probability 1,

$$\liminf \left\{ \frac{1}{\sqrt{2 \ln k}} | e_k(\omega) | \right\} \geqslant 1,$$

while $(2 \cdot \ln k)^{-1/2} e_k \longrightarrow 0$ in the natural topology on K).

The basic aim of the immediate investigations is to find conditions permitting us to judge from the intrinsic geometry of the set K whether the property $K \in GB$ is satisfied, or, what is the same, to judge from the geometric properties of subsets of a Gaussian subspace $H \subset L^2$ whether they are structurally bounded.

§2. Problem of the size of a convex solid angle in Hilbert space and the infinite-dimensional Cauchy distribution

1. We reduce the problem of verifying the property $K \in GB$ to the problem of the size of a certain convex solid angle in Hilbert space, which enables us to obtain an equivalent geometric reformulation of the problem of bounded realizations and several important inequalities for the Gaussian measures of convex sets that reduce to convenient conditions in terms of the ϵ-entropy of K.

A Gaussian process is determined by the intrinsic geometry of the set K and its distribution in the Gaussian space $H \subset L^2(\Omega, P)$. However, the GB property is easily seen to depend neither on the isometric transformations of K (since they are the adjoints of measure-preserving transformations of the space (Ω, P)) nor on translations (since a translation is an addition to each of the random variables appearing in K of some single random variable). There arises the natural desire to give a characteristic of the GB property in terms determined only by the intrinsic geometry of K. This characteristic will be shown to be simultaneously a characteristic of every cone in Hilbert space whose polar cone has positive measure (precise definitions are given below). It turns out that the class of GB-sets coincides with the class of sets having a finite Minkowski mixed volume $h_1(K)$ of the first degree of homogeneity.

The more detailed investigation of the properties of the functional h_1 to be carried out later permits us to obtain an important and nontrivial characteristic of it that relates to the finite-dimensional case: the monotonicity of this functional on the class of polyhedrons with respect to comparison of the lengths of corresponding edges. This monotonicity property allows us in many cases to get estimates of the magnitude of $h_1(K)$ by comparison with those simple sets for which $h_1(K)$ can be calculated (rectangular parallelepipeds) or well estimated (finite sets of pairwise orthogonal vectors of equal length). The basic tool of the investigation is infinite-dimensional Cauchy measure, which, in turn, arises from the problem of the positivity of the measure of a solid angle polar to a given one.

Finally, we formulate a criterion for the continuity of realizations of a Gaussian process K, i.e. for the validity of the condition $K \in GC$, in terms of the functional h_1 (or in terms of the Cauchy measure).([1]) From this general criterion, which is difficult to apply directly, as for the GB property, we derive a convenient (for checking) ϵ-entropy condition that is necessary for GC-sets, but not necessary for GB-sets (the ϵ-entropy sufficient conditions available for the GB property are at the same time sufficient also for the GC property). This ϵ-entropy language for checking the GB (or GC) property turns out to be very successful, as was mentioned in the Introduction. Nevertheless, as examples show, an exact criterion for the verification of the GB property, as well as for the verification of the positivity of the measure of a solid angle, cannot be formulated in terms of ϵ-entropy.

2. We proceed to the problem of the measure of a solid angle in a Hilbert space. We consider closed convex cones with vertices at zero in a Hilbert space, including "wedges", i.e. cones containing linear subspaces not coinciding with the whole space. We first consider the finite-dimensional case. We can assign to each cone in the finite-dimensional Euclidean space $\mathbf{R}^n$ its measure: the measure of the solid angle, equal to the measure of the intersection of this cone with the surface of the unit ball with respect to the normalized spherically invariant measure on the surface of this ball. In infinite-dimensional Hilbert space there does not exist a spherically invariant measure on the surface of the unit ball (see, for example, [30]), but, for example, it is clearly natural to assign to each polyhedral cone (wedge) formed by the intersection of a finite number of closed half-spaces the "measure" possessed by the orthogonal projection of this wedge onto the orthogonal complement of the largest linear subspace contained in this wedge. For example, if we consider the cone formed by the two half-spaces $\{y: \langle x_1, y\rangle \leq 0\}$ and $\{y: \langle x_2, y\rangle \leq 0\}$, then it is natural to take the number $1/2 - (2\pi)^{-1} \cdot \arccos ((x_1, x_2)/(\|x_1\| \cdot \|x_2\|))$ to be the "measure" of the solid angle obtained by intersecting these half-spaces. If we now define the size of an arbitrary closed convex solid angle as the infimum of the "measures" of the wedges of finite defect containing it, then we obtain an additive (but not countably additive) function defined on the collection of closed convex cones and their finite unions. This "measure" is invariant with respect to orthogonal transformations but, of course, not with respect to arbitrary isometric transformations of the Hilbert space into itself; and, therefore, the magnitude of the "measure" of a convex cone (or its intersection with the unit sphere) depends not only on its intrinsic geometry, but also on its position in the containing Hilbert space. It is possible, however, to state the problem of the "measure" of a solid angle in such a way that its magnitude remains dependent only on the intrinsic geometry of the particular set, but not on the character of its imbedding in the containing space.

([1]) The original version of the theorem on ϵ-entropy criteria [121] was obtained without the help of mixed volumes. However, the method used in [121] is, in essence, based on the use of the properties of the functional $h_1(K)$; a clear separation of it is useful (even if we do not consider the monotonicity theorem obtained for it, which is a curiosity in itself) (see [125], and also [131], where, in particular, the author notes the relation $h_1(K) = (2\pi)^{1/2} \cdot E \sup \{x(\omega): x \in K\}$, so that the monotonicity property of h_1 is precisely equivalent to the monotonicity of $E \sup \{x(\omega): x \in K\}$, which was proved independently by Chevet [19]: see also [33] and Proposition 14 below).

We return to the finite-dimensional case. On the surface Σ_E of the unit ball V_E in n-dimensional Euclidean space we consider the normalized measure m that is invariant with respect to all orthogonal transformations. Let $F = \mathbf{R}^n$ be another copy of n-dimensional Euclidean space, whose elements serve as (continuous) linear functionals on E; the corresponding bilinear form is denoted by $\langle x, y\rangle$, $x \in F$, $y \in E$. In the finite-dimensional case it would be possible to identify the elements of E and F, but we do not do this, thinking ahead of the infinite-dimensional generalizations. However, it is in any case convenient to consider in the space F its intrinsic scalar product (x', x''), by means of which it is possible, for example, to measure the angles between rays in F.

To each subset Q of the surface Σ_F of the unit ball in F, or, what is the same, to each cone con Q in F, there corresponds a closed spherically convex subset $Q^\perp \subset \Sigma_E$, or, what is the same, a closed convex cone con $Q^\perp = (\text{con } Q)^\perp$, the polar with respect to the form $\langle x, y\rangle$:

$$(\text{con } Q)^\perp = \{y : y \in E,\ \forall x\ x \in \text{con } Q \Rightarrow \langle x, y\rangle \leqslant 0\}.$$

With what geometric characteristics of Q is the size of the solid angle of con $Q^\perp$ (i.e. the quantity $mQ^\perp$) connected? From spherical geometry it is well known that, for example, in the case $n = 3$ the quantity $mQ^\perp$ is equal to $1/2 - p/2\pi$, where p is the length of the perimeter of the spherical convex hull of the set $Q \subset \Sigma_F$ (i.e. the intersection of the convex hull of con Q and the sphere Σ_F).

However, in the case $n = 4$ even for the simplest set Q consisting of four points of Σ_F in general position, the measure of the spherical simplex polar to Q is not expressible in an elementary way (and, in general, at all simply) in terms of the pairwise angular distances between the points of Q. Even the calculation of the asymptotic behavior of the measure of a regular spherical simplex is not at all simple.

An important property of $mQ^\perp$ as a function of Q is the fact that it is completely determined by the intrinsic metric of Q induced by the imbedding $Q \subset \Sigma_F \subset F$ and does not depend on the dimension of the space F (while $Q^\perp$ itself is defined only simultaneously with the fixing of the whole space $\mathbf{R}^n$). We mention also that if $Q_1 \subset Q$, then $mQ_1^\perp \geqslant mQ^\perp$.

In the finite-dimensional case (i.e., if Q is isometric to a subset of the unit ball of a finite-dimensional space), we have $mQ^\perp = 0$ if and only if the spherically convex hull of Q contains pairs of points between which the angular distance is arbitrarily close to π, i.e., if and only if the closure of the spherically convex hull contains, together with some vector x, also the vector $-x$.

Much more complicated is the question of estimating the measure of a polar angle independently of the dimension n and thereby suitably for the infinite-dimensional case when Q is a subset of the unit ball of the Hilbert space $H = F$. Our basic problem consists of determining such estimates, and, especially, of determining conditions on the set Q for the positivity of the quantity $mQ^\perp$. Properly speaking, in the infinite-dimensional case we can state the problem of the positivity of $mQ^\perp$ without mention of what a spherically invariant measure is in this case, clarifying for which sets $Q \subset \Sigma_F$ the condition

$$\inf\{mQ_1^\perp : Q_1 \subset Q,\ \dim Q_1 < \infty\} > 0$$

is satisfied (as was explained, it is not necessary to indicate each time in which space we are considering the measure m). The value of this infimum is the exact analogue of $mQ^\perp$ for finite-dimensional Q. However, for our purposes it is useful also in the infinite-dimensional case to have a real (countably additive) measure m that is invariant with respect to rotations. Moreover, only an analysis of the specific properties of the infinite-dimensional distributions that do not exhibit themselves (exhibit themselves only asymptotically) in the finite-dimensional case enables us to find the approach to the problem that is presented below.

3. As already mentioned, no measure in the Hilbert space H^* of continuous linear functionals on $F = H$ can serve as the "spherically invariant" measure m; however, we can regard as spherically invariant any weak distribution on H^* whose characteristic functional, which is defined on H, depends only on the norm $\|\cdot\|_H$. Let (E, m) be a linear measure space given by such a spherically invariant characteristic functional. Each element of H is a measurable linear functional on (E, m). We show that $(\operatorname{con} Q)^\perp$, as well as $mQ^\perp$, where $Q \subset \Sigma_H$, can be correctly defined also in this case. For this, it suffices to set

$$(\operatorname{con} Q)^\perp = \{\omega : (\sup_{x \in Q} \{x\})(\omega) \leqslant 0\},$$

where the sup is taken in the sense of the usual partial order structure on the set S_m of all measurable functions. [By the Minlos–Sazonov theorem, we can also choose some concrete extension B of H^* in which the weak distribution extends to a measure m, and manage with the help of those elements of H (forming a dense subset of H) that are continuous functionals on B.] The spherical invariance of m means, in particular, that each orthogonal operator in H is the adjoint of some automorphism of the linear measure space (E, m). It is also natural to call this automorphism an m-orthogonal transformation of (E, m). In the case of a concrete realization (B, m) an m-orthogonal operator is defined only on a linear subspace of the Banach space B having full m-measure.

In the finite-dimensional case it is natural to call the measure m^Σ concentrated on the surface Σ of the unit ball an elementary spherically invariant measure. Any other spherically invariant measure m can be obtained by means of m^Σ in the following way:

$$m(A) = \int_0^\infty m^\Sigma(\sigma A)\, \alpha(d\sigma),$$

where $\alpha(d\sigma)$ is some Borel probability measure on $[0, \infty)$. If we consider the (measurable) decomposition λ of a finite-dimensional measure space (E, m^Σ) into the rays going out from the origin (it is assumed that the zero point does not have positive measure), then the conditional measures on the rays are obviously δ-measures concentrated at unit distance from the origin, i.e., the decomposition λ is equivalent to the decomposition into points, and for λ the conditional measures on the rays of (E, m) with an arbitrary spherically invariant measure are the measures $\alpha(d\sigma)$.

In the infinite-dimensional case the standard Gaussian measure plays the same role among the spherically invariant measures as the measure m^{Σ} in the finite-dimensional case.

PROPOSITION 4. *Let (E, m) be a linear measure space, and $m\{0\} = 0$. The decomposition λ of (E, m) into rays is measurable.*

PROOF. The decomposition λ is equal to the product of a countable number of measurable decompositions into pairs of half-spaces. ●

PROPOSITION 5. *Let (E, γ) be a linear Gaussian measure space. Then the decomposition λ into rays is equivalent (in the sense of equivalence of measurable decompositions of Lebesgue spaces) to the decomposition ϵ into points, i.e., the conditional measures on the rays are δ-measures (at some point of each ray).*

PROOF. We realize the space (E, γ) as the space $(\mathbf{R}^\infty, \gamma_0)$, where $\mathbf{R}^\infty = \{(y_1, y_2, \dots)\}$ is the space of all numerical sequences, and γ_0 is the standard Gaussian product measure. For almost all elements of $\mathbf{R}^\infty$ we have

$$\lim \frac{1}{n} \sum_{1}^{n} y_k^2 = 1, \tag{1}$$

but on each ray this equality cannot be satisfied at more than one point. The point at which it holds is the support of the conditional δ-measure. ●

COROLLARY 1 (Schoenberg's theorem [106]). *For the function $F(r)$, $r \geqslant 0$, to be a positive definite function in the n-dimensional space $\mathbf{R}^n = \{(x_1, \dots, x_n)\}$ for any n, where $r = (x_1^2 + \cdots + x_n^2)^{1/2}$, it is necessary and sufficient that $F(r)$ admits the representation*

$$F(r) = \int_0^\infty e^{-ur^2} \alpha(du).$$

Indeed, $\alpha(du)$ must be taken to be the conditional measure on a ray for an (infinite-dimensional) spherically invariant measure with characteristic functional $\chi(h) = F(\|h\|_H)$.

A direct proof of Schoenberg's theorem takes several pages of calculations (even in the simplified presentation in [1]).

COROLLARY 2. *Any spherically invariant measure m can be represented in the following form:*

$$m(A) = \int_0^\infty \gamma^\sigma(A)\, \alpha(d\sigma), \tag{2}$$

where $\gamma^\sigma(A) = \gamma(\sigma A)$ is the Gaussian measure determined by the characteristic functional $\chi(x) = \exp(-(\sigma\|x\|_H)^2/2)$.

(This is essentially a reformulation of Corollary 1.)

4. We return to the problem of the measure of a solid angle. It is much more common to work, not with subsets Q of the unit sphere and the polars $Q^{\perp}$ and $(\text{con } Q)^{\perp}$ of them and the cones they span, but with arbitrary subsets of the Hilbert space and the usual (for functional analysis) polars of them. We transform the problem of the positivity of the measure of a polar solid angle, formulating it in new terms. Let F be a Hilbert space (possibly finite-dimensional) and Q a set on its unit sphere. Let f_0 be an arbitrary unit vector in F for which

$$(f_0, x) > 0 \text{ for any } x \in Q \subset \Sigma_F \tag{3}$$

(if Q has a center of symmetry, it is natural to take it to be f_0).

In F we consider the hyperplane $H_0 = \{x: x \in F, (x, f_0) = 1\}$ tangent to Σ_F at f_0. We assign to the set Q the set $K = K(Q)$ of central projections of the points of Q onto the hyperplane H_0:

$$K(Q) = \{x : x \in H_0, \ x\|x\|^{-1} \in Q\}.$$

The set $Q = Q(K)$ is identically reproduced by K. We remark that con $K(Q)$ = con Q.

In the Gaussian measure space (E, γ) we now consider the hyperplane $E_0 = \{\omega: \omega \in E, \langle f_0, \omega\rangle = -1\}$. (This is a proper definition in the infinite-dimensional case, because the decomposition of (E, γ) into hyperplanes parallel to E_0 is measurable, and it is easy to write out explicitly the characteristics of the conditional Gaussian measure on each one-parameter family of such hyperspaces.) Next, let

$$H_0^* = \{\omega : \omega \in H^*, \ \langle f_0, \omega\rangle = -1\}.$$

It follows from (3) that each ray lying in $(\text{con } Q)^{\perp}$ intersects E_0.

We now construct the central projection π along the rays of the whole half-space $E' = \{\omega: \langle f_0, \omega\rangle < 0\}$ into the hyperplane E_0. Each ray lying in this half-space intersects E_0, and therefore we can consider the image $\gamma\pi^{-1}$ of γ (more precisely, the part of it that is concentrated in E') under the mapping π. Obviously,

$$\gamma(\text{con } Q)^{\perp} = \gamma\pi^{-1}((\text{con } Q)^{\perp} \cap E_0).$$

On the affine space H_0 we now consider a new linear structure compatible with the affine structure already there, taking f_0 as the new zero point. In other words, if $x_1, x_2 \in H_0$ and $+_0$ and $\cdot_0$ denote addition and mutiplication by scalars in the new sense, then

$$\alpha_1 \cdot_0 x_1 +_0 \alpha_2 \cdot_0 x_2 = \alpha_1(x_1 - f_0) + \alpha_2(x_2 - f_0) + f_0.$$

Similarly, we introduce a new linear structure in E, taking as the new zero point a vector $\omega_0 \in H_0^*$ for which $\|\omega_0\|_{H^*} = 1$. The linear spaces H_0 and H_0^* become Hilbert spaces if they are equipped with the norms

$$\|x\|_{H_0} = \|x - f_0\|, \quad \|\omega\|_{H_0^*} = \|\omega - \omega_0\|_{H^*}.$$

The Hilbert spaces H_0 and H_0^* are in a natural duality defined by the bilinear form $\langle x, \omega\rangle_0 = \langle x - f_0, \omega - \omega_0\rangle$, and we can regard the elements of H_0 as measurable linear functionals on $(E_0, \gamma\pi^{-1})$. We use the notation A° for the polar of the set A in the sense of the duality (H_0, H_0^*). If $A \subset H_0$, then, as shown above, the following set is well defined:

$$A^{\circ} = \{\omega : \langle x, \omega\rangle_0 \leqslant 1 \text{ for } x \in A\}.$$

Now let $Q \subset \Sigma_F$, $K = K(Q)$. Then

$$\begin{aligned} K^{\circ} &= \{\omega : \omega \in E_0, \langle x, \omega\rangle_0 \leqslant 1 \quad \text{for} \quad x \in K\} \\ &= \{\omega : \omega \in E_0, \langle x - f_0, \omega - \omega_0\rangle \leqslant 1 \text{ for } x \in K\} \\ &= \{\omega : \omega \in E_0, \langle x, \omega\rangle - \langle f_0, \omega\rangle - \langle x, \omega_0\rangle + \langle f_0, \omega_0\rangle \leqslant 1, \ x \in K\} \\ &= \{\omega : \omega \in E_0, \langle x, \omega\rangle \leqslant 0 \mid \text{for } x \in K\} = E_0 \cap (\operatorname{con} Q)^{\perp}. \end{aligned}$$

Thus, the polar K° of an arbitrary set $K \subset H_0$ coincides with the intersection of $(\operatorname{con} K)^{\perp}$ with the space (hyperplane) E_0; therefore, to find the γ-measure of $(\operatorname{con} Q)^{\perp}$ it is necessary to find the $(\gamma\pi^{-1})$-measure of the polar of $K = K(Q) \subset H_0$.

It remains to describe the measure $\kappa/2 = \gamma\pi^{-1}$ on E_0. If $\omega \in E'$, then $\pi(\omega) = -\omega/\langle f_0, \omega\rangle$. We represent E' in the form of a direct product,

$$E' = E_1 \times \mathbf{R}^{+}, \quad E_1 = \{\omega : \langle f_0, \omega\rangle = 0\}, \ \mathbf{R}^{+} = \{\lambda, \lambda \geqslant 0\},$$

assigning to the element $\omega \in E'$ the pair (ω_1, λ), where $\omega_1 = \omega + \omega_0\langle f_0, \omega\rangle$ and $\lambda = -\langle f_0, \omega\rangle$.

If $\omega = (\omega_1, \lambda) \in E_1 \times \mathbf{R}^{+}$, then $\pi(\omega) = (\omega_1/\lambda, 1) \in E_0$; moreover, ω_1 has a normal distribution that is orthogonal to the projection of the standard distribution γ, and λ is independent of ω_1 and has the distribution of the positive part of a one-dimensional standard Gaussian variable. It is well known (see, for example, [37], §24, Example 4) that the ratio of a standard multidimensional normal variable to a one-dimensional standard normal variable (or its modulus) that is independent of it has a multidimensional standard Cauchy distribution; consequently, the distribution $\kappa/2$ on E_1 in any finite-dimensional case, and hence in any case, is described by the characteristic functional

$$\chi_{1/2\kappa}(x) = \frac{1}{2}\exp(-\|x\|_{H_0}), \quad x \in H_0. \tag{4}$$

The coefficient 1/2 in front of the exponent appears because, in the sense of our definitions, the measure $\gamma\pi^{-1} = \kappa/2$ is not a probability measure, but a "semi-probability" one, i.e., $\kappa E_1/2 = 1/2$ (the measure of a convex nondegenerate solid angle is not greater than one-half). By exact analogy with the finite-dimensional case, we call the measure κ having characteristic functional

$$\chi_{\kappa}(x) = \exp(-\|x\|_H), \quad x \in H, \tag{5}$$

a Cauchy measure also in the infinite-dimensional case.

We summarize what has been said in the following statement.

PROPOSITION 6. *Let (E, γ) be a linear Gaussian measure space, and H the Hilbert space of all measurable linear functionals on (E, γ). Let Q be a subset of the unit sphere Σ_H of H, $f_0 \in Q$ a vector such that $(x, f_0)_H > 0$ for all $x \in Q$, and $K = K(Q) = H_0 \cap \operatorname{con} Q$, where $H_0 = \{x : (x, f_0)_H = 1\}$. Consider a Cauchy process with natural parameter set K_1 that is isometric to K and located in the Gaussian space H_1 such* *that the zero point of H_1 corresponds to the point $f_0 \in K$, i.e., a random process whose characteristic function is given on H_1 by (5), with H replaced by H_1. Then the measure of the solid angle* (con $Q)^{\perp}$ *is equal to one-half of the probability that realizations of such a Cauchy process do not exceed one*:

$$\gamma(\mathrm{con}\ Q)^{\perp} = \frac{1}{2}\varkappa K_1^{\circ}.$$

The proof of Proposition 6 was given above.

In the following we use the notation H both for the space of measurable linear functionals on a Gaussian measure space, and for the space of measurable linear functions on the space with the corresponding Cauchy measure; in this second case the Hilbert norm in H is, of course, not the trace of the norm in L^2, since $H \cap L^2(E, \kappa) = \{0\}$, but it is identically reproduced by the characteristic functional of the Cauchy measure according to (5).

§3. Cauchy measure and Gaussian measure

1. As does any spherically invariant infinite-dimensional measure, the Cauchy measure admits the representation (2), where $\alpha(d\sigma)$ is the conditional measure on a typical element l of the decomposition of the space (E, κ) into rays, and the scale on the half-line $\{\sigma\} = l = \mathbf{R}^+$ is chosen so that the conditional measure of the countable-dimensional normal spherically invariant distribution γ is the δ-measure at the point 1 (see Proposition 5). This choice of scale on the ray l is called standard (with respect to the measure γ). We determine the concrete form of the measure $\alpha(d\sigma)$.

PROPOSITION 7. *For any measurable set $A \subset (E, \kappa)$*

$$\varkappa(A) = \int_0^\infty \gamma^\sigma(A)\,\rho(\sigma)\,d\sigma, \tag{6}$$

where

$$\rho(\sigma) = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma^2}\exp\left(-\frac{1}{2\sigma^2}\right), \tag{7}$$

and the Gaussian measure γ^σ has the characteristic functional

$$\chi_{\gamma^\sigma}(x) = \exp\left(-\frac{1}{2}(\sigma\|x\|)^2\right).$$

PROOF. The following identity can be verified directly:

$$\exp(-r) = \int_0^\infty \exp\left(-\frac{1}{2}(r\sigma)^2\right)\left(\sqrt{\frac{2}{\pi}}\frac{1}{\sigma^2}\exp\left(-\frac{1}{2\sigma^2}\right)\right)d\sigma. \tag{8}$$

Indeed, denoting the right-hand side by y, we find that

$$y' = -r\sqrt{\frac{2}{\pi}}\int_0^\infty \exp\left(-\frac{1}{2}(r\sigma)^2\right)\exp\left(-\frac{1}{2\sigma^2}\right)d\sigma$$

and, after the substitution $\sigma = 1/r\tau$, we get that

$$y' = -\sqrt{\frac{2}{\pi}} \int_0^\infty \exp\left(-\frac{1}{2}\frac{1}{\tau^2}\right) \exp\left(-\frac{1}{2}\tau^2 r^2\right) \frac{1}{\tau^2}\, d\tau = -y. \tag{9}$$

Obviously, $y = 1$ for $r = 0$, and this, together with (9), proves (8). But, since $\exp(-(\sigma\|x\|)^2/2)$ is the characteristic functional of the Gaussian measure γ^σ, and $\exp(-\|x\|)$ is the characteristic functional of κ, this identity (8) is precisely equivalent to the equality (6) of interest to us. ●

The density $\rho(\sigma)$ is thus the density of the conditional measure on a ray for the Cauchy measure κ. However, we might expect this result, if we recall that the Cauchy measure characterizes the distribution of the ratio of a standard infinite-dimensional Gaussian variable to the modulus of an independent standard one-dimensional Gaussian variable, and on almost every ray the infinite-dimensional Gaussian variable is constant (Proposition 5), its value being used for choosing a scale of the ray. But the density $\rho(\sigma)$ characterizes exactly the distribution of the variable $|\xi|^{-1}$, where ξ is a standard Gaussian random variable. Henceforth, we say that the Cauchy measure κ is standard with respect to the Gaussian measure γ if the characteristic functional of κ has the form (5), where $\|x\|_H$ is the usual norm of the function $x(\cdot) \in H$ in $L^2(E, \gamma)$.

2. The expression for the conditional Cauchy measures permits us to get useful and simple estimates connecting the Cauchy and Gaussian measures of convex sets.

Proposition 8. *Let $Q \subset (E, \kappa)$ be a measurable subset that is star-shaped with respect to zero, κ a Cauchy measure on E, and γ a Gaussian measure on E with respect to which κ is standard. Then*:

1) $\gamma Q \geqslant \dfrac{\kappa Q - q}{1 - q}$, *where* $q = \sqrt{\dfrac{2}{\pi}} \displaystyle\int_0^1 \exp\left(-\frac{1}{2}\tau^2\right) d\tau \approx 0.3174$;

2) $\gamma Q \leqslant \dfrac{1}{q}\kappa Q.$

Proof. Let λ be the measurable decomposition of (E, κ) into rays, and let κ/λ and γ/λ denote the measures on the quotient spaces $(E, \kappa)/\lambda$ and $(E, \gamma)/\lambda$. Obviously, $\kappa/\lambda = \gamma/\lambda$. Let $l \in \lambda$ be an arbitrary ray, and $u(Q, l)$ the length of the segment $l \cap Q$ in the scale that is standard with respect to γ. By the definition of conditional measures,

$$\kappa Q = \int_{E/\lambda} d(\kappa/\lambda) \int_0^{u(Q,\, l)} \sqrt{\frac{2}{\pi}} \frac{1}{\sigma^2} \exp\left(\frac{1}{2\sigma^2}\right) d\sigma,$$

$$\gamma Q = (\gamma/l)(\{l : u(Q,\, l) \geqslant 1\}),$$

from which we find that

$$\kappa Q = \int_{\{l:\,u\leqslant 1\}} + \int_{\{l:\,u>1\}} \leqslant \int_{\{u\leqslant 1\}} d(\kappa/\lambda) \int_0^1 + \int_{\{u>1\}} d(\kappa/\lambda) \int_0^\infty \leqslant$$

$$\leqslant \int_0^1 \sqrt{\frac{2}{\pi}} \frac{1}{\sigma^2} \exp\left(\frac{1}{2\sigma^2}\right) d\sigma \cdot (\gamma/\lambda)(\{u \leqslant 1\}) + (\gamma/\lambda)(\{u > 1\})$$

$$= (1 - \gamma Q)\, q + \gamma Q \Rightarrow \gamma Q \geqslant \frac{Q - q}{1 - q}.$$

Similarly,

$$\varkappa Q \geqslant \int_{\{u \geqslant 1\}} d(\gamma/\lambda) \int_0^1 \sqrt{\frac{2}{\pi}} \frac{1}{\sigma^2} \exp\left(\frac{1}{2\sigma^2}\right) d\sigma = q\gamma Q. \bullet$$

3. The "geometric" origin of the Cauchy measure κ leads to the circumstance that is is frequently more convenient to get estimates for it than for Gaussian measure. At the same time, the problem of extending a Gaussian weak distribution in a linear space is equivalent to that of extending a Cauchy weak distribution.

PROPOSITION 9. *The zero-one law for linear subspaces is valid for any spherically invariant measure m (i.e., Proposition 1 holds with the measure m substituted for γ).*

PROOF. Since a linear subspace is invariant with respect to homotheties, any linear subspace $E_1 \subset (E, m)$ either has the property $\gamma^\sigma E_1 = 1$ for any $\sigma > 0$, or has the property $\gamma^\sigma E_1 = 0$ for $\sigma > 0$, and the required assertion then follows from the representation (Corollary 2 to Proposition 5). $\bullet$

PROPOSITION 10. *Let $V \subset (E, \kappa)$ be a measurable set that is star-shaped with respect to zero and such that $\kappa \bigcup_{n=1}^{\infty} nV = 1$. Then $\kappa V > 0$.*

PROOF. If $\kappa V = 0$, then $u(V, l) = 0$ for (κ/λ)-almost all l, and, since $u(nV, l) = nu(V, l)$, we would have $\kappa \bigcup_1^\infty nV = 0$ (see the formula for the conditional Cauchy measures). $\bullet$

COROLLARY. *If $V \subset E$ is a convex subset containing zero, then the condition $\gamma L(V) = 1$ is equivalent to the condition $\kappa V > 0$.*

4. Finally, we mention another useful property of a Cauchy measure: its special kind of monotonicity with respect to the measures of the polars of sets subject to a certain class of transformations. Originally this property permitted one to establish ϵ-entropy conditions without using the concept of the mixed volume h_1 [121] (in contrast to the method of presentation used below). In the sequel we prove the monotonicity of $h_1(K)$ with respect to the same class of transformations, which makes it possible not to use the monotonicity of the Cauchy measure. Therefore, we give the corresponding assertion here without proof. It is proved with the help of a theorem of Schlaefli more simply than the monotonicity property of the functional h_1 (Theorem 2).

PROPOSITION 11. *Let $A, B \subset H$, and suppose that there exists a mapping ψ of A onto B such that for $f, g \in A$*

$$\frac{(f,\ g)+1}{(\|f\|^2+1)^{1/2}(\|g\|^2+1)^{1/2}} \leqslant \frac{(\psi f,\ \psi g)+1}{(\|\psi f\|^2+1)^{1/2}(\|\psi g\|^2+1)^{1/2}}.$$

Then $\kappa A^\circ \leqslant \kappa B^\circ$.

§4. The GB property and mixed volumes

1. We recall that the Minkowski mixed volume of the first degree of homogeneity for a convex bounded subset $K \subset \mathbf{R}^n$ of n-dimensional Euclidean space is defined (see, for example, [14]) to be the coefficient $w_{n-1}^{(n)} = w_{n-1}^{(n)}(K)$ of ϵ^{n-1} in the expansion in powers of ϵ of the quantity

$$W(K,\ \varepsilon) = \mathrm{vol}_n(K+\varepsilon V) = w_0^{(n)}(K) + \varepsilon w_1^{(n)}(K) + \ldots + \varepsilon^n w_n^{(n)}(K) \qquad (10)$$

(which is a polynomial of degree n), where V is the unit ball in $\mathbf{R}^n$, and $\mathrm{vol}_n\, Q$ denotes the (n-dimensional) volume of Q. It is natural to consider the functionals $w_i^{(n)}(K)$ to within a multiplicative constant, and then the concrete form of the set V ceases to play a role. Another definition of the functional $w_i^{(n)}(K)$, $0 \leqslant i \leqslant n$, is the following:

$$w_i^{(n)}(K) = \int_{G_n^{n-i}} \mathrm{vol}_{n-i}\, \mathrm{Pr}_L\, K\, d\tau, \qquad (11)$$

where $\mathrm{Pr}_L\, K$ is the orthogonal projection of K onto the subspace $L \subset \mathbf{R}^n$, L runs through the set of all subspaces of dimension $n-i$, and τ is a measure on the corresponding Grassmann manifold G_n^{n-i} that is invariant with respect to the transformations adjoint to the orthogonal ones.

We are interested in the functional $a_n w_{n-1}^{(n)}(K)$, $a_n = \mathrm{const}$, of the first degree of homogeneity, henceforth denoted by $h_1(K)$, where we choose the normalizing factor a_n so that, for a rectilinear segment I, the value of $h_1(I)$ is equal to its length. With this normalization the value $h_1(K)$ for any finite-dimensional convex compact set K does not depend on the dimension of the Euclidean space $\mathbf{R}^n$ in which we consider the set K, but is determined only by its intrinsic metric. We define h_1 also on arbitrary (not necessarily convex) finite-dimensional affine sets, setting $h_1(K) = h_1(\mathrm{conv}\ K)$. For planar convex figures $K \subset \mathbf{R}^2$ the value $h_1(K)$ is the semiperimeter; for a convex body $K \subset \mathbf{R}^3$ the value $h_1(K)$ is proportional to the integral of the mean curvature. In particular, if K is a convex polyhedron, then

$$h_1(K) = \frac{1}{2\pi}\sum_k (\pi - \varphi_k)\, l_k,$$

where l_k is the length of an edge of K, and φ_k is the magnitude of the dihedral angle at this edge; the summation runs over all edges. More generally, let K be an arbitrary convex polyhedron, and P_i one of its vertices. We locate it at the origin of coordinates and let $Q_1 = \mathrm{con}\,(K; P_i) = \bigcup_{\lambda\to\infty} \lambda K$, a convex cone. We note that $\Sigma_i\, mQ_i^\perp = 1$. We now consider an edge P_iP_j and let K_{ij} be the orthogonal projection of K with kernel of dimension 1 and such that the line P_iP_j passes into the point P_{ij}, which is, thus, a vertex of the polyhedron K_{ij}. Then, analogously to the three-dimensional case,

$$h_1(K) = \sum_{i,\ j} l_{ij} m\,(\mathrm{con}\,(K_{ij};\ P_{ij}))^\perp, \qquad (12)$$

where m is the spherically invariant measure. (The formula (12) is easily derived directly from (10), applied to the case when K is a polyhedron.) For a ball $V_n \subset \mathbf{R}^n$ of unit radius the normalized value of the "generalized length" $h_1(V_n)$ can be calculated as

$$h_1(V_n) = \sqrt{\pi}\,(n-1)\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \sim \sqrt{\pi n} \quad (n > 1).$$

For a parallelepiped $K \subset \mathbf{R}^n$ the value of $h_1(K)$ is the sum of the lengths of all the edges, multiplied by 2^{1-n} (i.e., the sum of the lengths of the edges forming a maximal set of pairwise perpendicular edges).

2. We can now define the value of the functional $h_1(K)$ for subsets $K \subset H$ of an infinite-dimensional Hilbert space H, setting

$$h_1(K) = \sup_{K' \subset K,\ \dim L(K') < \infty} h_1(K').$$

PROPOSITION 12. *$h_1(K) = h_1(\overline{K})$ for any $K \subset H$.*

The proof follows from the fact that any finite-dimensional set $K_1 \subset \overline{K}$ can be approximated by a finite-dimensional set $K_1' \subset K$. ●

PROPOSITION 13. *Let $K_1 \subset K_2 \subset \cdots$, and $K = \bigcup_1^\infty K_n$. Then*

$$h_1(\overline{K}) = \lim h_1(K_n).$$

PROOF. By Proposition 12, it suffices to prove that $h_1(\bigcup_1^\infty K_n) = \lim h_1(K_n)$. We fix an arbitrary $\epsilon > 0$ and choose a finite-dimensional set $K_\epsilon \subset K$ such that $h_1(K_\epsilon) > h_1(K) - \epsilon$. Without loss of generality we can assume that all the sets K_n, together with K, are convex. We consider the expanding sequence of finite-dimensional affine spaces $L_1 \subset L_2 \subset \cdots$ spanned by the corresponding sets K_n. The set K_ϵ is contained in some space L_{n_0}; consequently

$$\lim_i h_1(K_i \cap L_{n_0}) \geq h_1(K_\epsilon) > h_1(K) - \epsilon.$$

By the obvious monotonicity of h_1 with respect to inclusion and the arbitrariness of ϵ, we get the desired conclusion. ●

We mention several of the simplest properties of the functional h_1.

1) Homogeneity of degree 1 (obvious).

2) Additivity with respect to algebraic addition: $h_1(\Sigma_i K_i) = \Sigma_i h_1(K_i)$ (the additivity is proved in the same way as for the finite-dimensional case).

3) Lower semicontinuity (Proposition 13).

4) Upper semicontinuity in the finite-dimensional case: if $\mathbf{R}^n \supset K_1 \supset K_2 \supset \cdots$ and $K = \bigcap K_n$, then $h_1(K) = \lim h_1(K_n)$.

In the infinite-dimensional case we have only the inequality and it is possible to give an example in which $<$ holds.

EXAMPLE. Let $K_1 \subset H$, $K_1 = \{\pm(2 \cdot \ln n)^{-1/2} e_n,\ n = 1, \ldots\}$, where $\{e_n\}$ is an orthonormal sequence in the Hilbert space H. It will be shown that $K \in GB$, and, therefore, $h_1(K) < \infty$; however, $K \notin GC$, and from this, as will be shown, it follows that if

$$K_m = \left\{ \pm \frac{1}{\sqrt{2 \ln n}} e_n, \quad n = m+1, \quad m+2, \ldots \right\},$$

then $h_1(K_m) \geqslant \alpha > 0$, although $\bigcap_m K_m = \varnothing$ and $\bigcap_m \overline{\operatorname{conv} K_m} = \{0\}$.

3. For what follows it is important to obtain a direct expression, analogous to (11) (for $i = n - 1$), for the quantity $h_1(K)$ with the help of Gaussian measure.

PROPOSITION 14. *Let $K \subset H$, where H is the Hilbert space of measurable linear functionals on a linear Gaussian measure space (E, γ). Then*

$$h_1(K) = \sqrt{2\pi} \int_E \sup_{x \in K} x(\omega)\, d\gamma.$$

PROOF. We first prove the assertion for a finite-dimensional set $K \subset H$. Let $K \subset \mathbf{R}^n$, and let E in this case denote the dual space of $\mathbf{R}^n$ (another copy of $\mathbf{R}^n$), equipped with the standard Gaussian measure. For such sets we use (11). Let λ be the decomposition of (E, γ) into rays l, so that $\omega = (l, \sigma)$, where $\omega \in E$, $l \in \lambda$, and σ is the distance from ω to the origin. If e is an arbitrary unit vector in the Euclidean space $\mathbf{R}^n \supset K$, then the length of the projection $\operatorname{vol}_1 \operatorname{Pr}_L K$ of the set K onto the line L running through e is

$$\sup_{x \in K} (x, e) - \inf_{x \in K} (x, e) = \sup_{x \in K} (x, e) + \sup_{x \in K} (x, -e).$$

Therefore,

$$w_{n-1}^{(n)}(K) = \int_{G_n^1} \operatorname{vol}_1 \operatorname{Pr}_L K\, d\tau = \int_\Sigma \sup_{x \in K} x(\omega)\, m^\Sigma(d\omega)$$

$$= \int_E \sup_{x \in K} m^\Sigma(d\omega) = \int_{E/\lambda} \sup x(1, l) \cdot (m^\Sigma/\lambda)(dl),$$

where m^Σ is the spherically invariant measure on the unit sphere $\Sigma \subset E = (\mathbf{R}^n)^*$. If now m is an arbitrary spherically invariant measure, and $q(d\sigma)$ is the conditional measure on (almost) every ray l, then, assuming that $\int_0^\infty \sigma q(d\sigma) = c < \infty$, we have

$$\int_E \sup_{x \in K} x(\omega)\, m(d\omega) = \int_{E/\lambda} (m/\lambda)(d\lambda) \int_0^\infty \sup_K x(\sigma, l)\, q(d\sigma)$$

$$= \int_{E/\lambda} (m/\lambda)(dl) \sup x(1, l) \int_0^\infty \sigma q(d\sigma) = c w_{n-1}^{(n)}(K).$$

In particular, for the Gaussian measure γ we get also that

$$h_1(K) = a_n w_{n-1}^{(n)}(K) = c_n \int_E \sup_K x(\omega)\, d\gamma. \tag{13}$$

To find the value of the constant c_n it suffices to take K to be a segment of unit length, placed symmetrically with respect to zero:

$$1=\frac{c_n}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{1}{2}|x|e^{-x^2/2}\,dx,\quad \text{i.e.}\quad c_n=\sqrt{2\pi}.$$

The fact that the normalizing constant does not depend on n permits us to assume that the measure γ in (13) is standard countable-dimensional Gaussian for any finite-dimensional K. If now K is an arbitrary set and $K_1 \subset K_2 \subset \cdots \subset K$ is a sequence of finite-dimensional sets such that $K \subset \bigcup_1^\infty K_n$, then $h_1(K_n) \longrightarrow h_1(K)$ (Proposition 13), and the sequence of functions $\{\sup_{x \in K_n} x(\omega)\}$ converges monotonically upwards to the function $\sup_{x\in K} x(\omega)$; therefore,

$$h_1(K)=\lim_n h_1(K_n)=\lim_n\sqrt{2\pi}\int_E \sup_{K_n} x(\omega)\,d\gamma=\sqrt{2\pi}\int_E \sup_K x(\omega)\,d\gamma. \bullet$$

COROLLARY. *If (B, γ) is a Banach space with a Gaussian finite- or countable-dimensional measure, then $h_1(V_L) = (2\pi)^{1/2} \int_B \|y\|_B\, d\gamma$, where $V_L \subset H$ is the unit ball of the space L, in duality with B, on whose elements the weak distribution is given.*

4. In particular, Proposition 14 allows us to get a lower estimate (by Čebyšev's inequality) of the quantity γK° in terms of $h_1(K)$, from which, by Proposition 1, it follows that the finiteness of h_1 is sufficient for the condition $K \in GB$. For our purposes it is especially important to get a sharp upper estimate of γK° in terms of $h_1(K)$ (from which, in particular, the necessity of the finiteness of $h_1(K)$ for the condition $K \in GB$ will follow).

We first study the problem of a "one-sided bound" for a Gaussian process (a detailed summary of the known results is found in [109]), and we consider the class GB^+ of processes whose realizations are bounded above with positive probability. The finiteness of $h_1(K)$ is trivially not necessary for the condition $K \subset GB^+$ (as shown by the example of the process $x_t(\omega) = tx(\omega)$, $t \in [0, +\infty)$, $x(\omega)$ a Gaussian variable). However, we show that if the convex hull of K contains the origin of coordinates and $\gamma K^\circ > 1/2$, then $h_1(K) < \infty$, and we get a sharp estimate. With this aim, we determine for what kind of convex set $K \subset \mathbf{R}^n$, $0 \in K$, $\gamma_n K^\circ = p > 0$ (γ_n is the standard Gaussian measure in $\mathbf{R}^n$), the maximum of $h_1(K)$ is attained.

PROPOSITION 15. *Let $V \subset (\mathbf{R}^n, \gamma_n)$ be a convex subset containing the origin. Then*

$$\frac{d}{d\lambda}\Big|_{\lambda=1}\gamma_n(\lambda V)=n\gamma_n V-\int_V \|x\|^2\,d\gamma_n,$$

where $\|\cdot\|$ is the Euclidean norm in $\mathbf{R}^n$.

PROOF. If $T_{k,\lambda}$ is the operator of dilation by a factor of λ in the direction of the axis x_k, then

$$\frac{d}{d\lambda}\int\limits_{T_{k,\lambda}}\cdots\int\limits_{V} c_n e^{-\frac{1}{2}\sum\limits_{i=1}^{n} x_i^2}\, dx_1 \ldots dx_n = \frac{d}{d\lambda}\int\limits_{V}\cdots\int \lambda e^{-\frac{1}{2}\sum\limits_{i\neq k} x_i^2+\lambda x_k^2}\, dx_1 \ldots dx_n$$
$$= \int\limits_{V}\cdots\int (1 - x_k^2)\, d\gamma_n,$$

from which, since

$$\left.\frac{d}{d\lambda}\right|_{\lambda=1} \gamma_n(\lambda V) = \sum_{i=1}^{n} \left.\frac{d}{d\lambda}\right|_{\lambda=1} \gamma_n(T_{k,\lambda} V),$$

we arrive at the desired conclusion. ●

COROLLARY. *In the class $\mathfrak{Y}_{n,p}$ of convex sets $V \subset \mathbf{R}^n$ containing the origin and such that $\gamma_n V = p$, the minimum of $(d/d\lambda)|_{\lambda=1}\, \gamma_n(\lambda V)$ is attained simultaneously with the maximum of $\int_V \|x\|^2\, d\gamma_n$.*

The next proposition is related to a lemma of Landau and Shepp [69]. The original preprint of Shepp [107] contained an error: the corresponding proposition was formulated for the class of centrally symmetric convex bodies; in this case the arguments connected with the solution of the isoperimetric problem on the sphere are not applicable, and for the correction the path presented below was suggested. This path turned out to be close to the course of the arguments in [69].

PROPOSITION 16. *Let $m_{n,r}$ be the spherically invariant measure in $\mathbf{R}^n$ concentrated on the surface of the ball of radius r, and $\mathfrak{W}_{n,r,p}$ the class of convex subsets $W \subset \mathbf{R}^n$ containing $\mathbf{0}$ and such that $m_{n,r} W = p$. Then for $p \leqslant 1/2$ the maximum of the (nonpositive) quantity $dm_{n,r}W/dr$ on the class $\mathfrak{W}_{n,r,p}$ is equal to zero, and is attained at an arbitrary convex cone with solid angle of magnitude p, while for $p > 1/2$ this maximum is attained when W is a half-space (containing $\mathbf{0}$).*

PROOF. The first part is obvious. Let $\Delta = \Delta(W, r)$ denote the boundary of the set $W \cap \{x\colon \|x\| = r\}$ on the surface Σ_r of the ball. We define on Δ a function $b(x)$ whose value at $x \in \Delta$ is the minimum of the distances to the origin from the hyperplanes passing through x and supporting W. Obviously,

$$\frac{d}{dr} m_{n,r} W = -c \int\limits_{\Delta} \frac{b(x)}{\sqrt{r^2 - b^2(x)}}\, dm'_\Delta, \tag{14}$$

where dm'_Δ is the "surface area" differential of the boundary Δ of the set $W \cap \{x\colon \|x\| = r\}$, and c is a positive constant independent of the choice of $W \in \mathfrak{W}_{n,r,p}$. Obviously, for any $W \in \mathfrak{W}_{n,r,p}$ and for $p > 1/2$ the value of $b(x)$ at any point $x \in \Delta \subset W$ is not less than the constant value of the function $b(x)$ constructed for a set $W_0 \in \mathfrak{W}_{n,r,p}$ that is a half-space. Since a "cap" ([105], [134]) is the solution of the isoperimetric problem on a sphere of arbitrary dimension, i.e., among all subsets $W \in \mathfrak{W}_{n,r,p}$ the minimum of the integral $\int_{\Delta(W,r)} dm'_\Delta$ is attained at a set W_0 that is a half-space $\{x\colon (x, y_0) \leqslant a\}$, we get that for a layer both the integrand and the total magnitude of the measure $m'_\Delta \Delta$ in (14) are minimized, from which the desired statement follows. ●

PROPOSITION 17. *On the class $\mathfrak{Y}_{n,p}$ the maximum of $\int_V \|x\|^2 \, d\gamma_n$ for $p \leqslant 1/2$ is attained for an arbitrary convex cone with center at 0, and for $p > 1/2$ this maximum is attained for a half-space $V_0 \in \mathfrak{Y}_{n,p}$.*

PROOF. The first part is obvious. Suppose that $p > 1/2$ and $V \in \mathfrak{Y}_{n,p}$. It follows from Proposition 16 that if $m_r(V) - m_r(V_0) \not\equiv 0$, then this difference changes sign for increasing r only one time, and from plus to minus (this is true for an arbitrary convex set $V \subset \mathbf{R}^n$ containing zero, independently of the condition $V \in \mathfrak{Y}_{n,p}$). Moreover, $\int_V \|x\|^2 \, d\gamma_n = \int_0^\infty r^2 m_r(V) \, d\chi_n$, where χ_n is the conditional measure on a ray (the square of the variable having the distribution χ_n has a χ^2 distribution with n degrees of freedom); therefore

$$\int_V \| x \|^2 \, d\gamma_n - \int_{V_0} \| x \|^2 \, d\gamma_n = \int_0^\infty r^2 \left(m_r(V) - m_r(V_0)\right) d\chi_n \leqslant 0,$$

since $\int_0^\infty m_r(V) \, d\chi_n = \int_0^\infty m_r(V_0) \, d\chi_n = p$, and r^2 is monotonically increasing. ●

We remark that instead of $\|x\|^2$ we could have used an arbitrary nondecreasing function of the length of the radius; see also [133].

PROPOSITION 18. *For a subset $V \subset \mathbf{R}^n$ in the class $\mathfrak{Y}_{n,p}$, $p > 1/2$,*

$$h_1(V^\circ) \leqslant \frac{1}{\Phi^{-1}(p)} e^{-\frac{1}{2}(\Phi^{-1}(p))^2} + p\sqrt{2\pi}.$$

Here $V^\circ \subset \mathbf{R}^n$ is the polar of V, and Φ^{-1} is the inverse of the function

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{1}{2} u^2\right) du.$$

PROOF. By Propositions 14 and 17, and the corollary to Proposition 15,

$$h_1(V^\circ) = \sqrt{2\pi} \int_{\mathbf{R}^n} \sup_{y \in V^\circ} (x, y) \, d_x \gamma_n = \sqrt{2\pi} \int_0^\infty \lambda d[\gamma_n(\lambda V)]$$

$$= \sqrt{2\pi} \int_0^1 \lambda d[\gamma_n(\lambda V)] + \sqrt{2\pi} \int_1^\infty \lambda d[\gamma_n(\lambda V)] \leqslant \sqrt{2\pi} \int_0^1 d[\gamma_n(\lambda V)]$$

$$+ \sqrt{2\pi} \int_1^\infty \lambda d[\gamma_n(\lambda V_0)] = p\sqrt{2\pi} + \int_1^\infty \lambda \Phi^{-1}(p) \exp\left(-\frac{1}{2}[\Phi^{-1}(p)\lambda]^2\right) d\lambda$$

$$= \frac{1}{\Phi^{-1}(p)} e^{-\frac{1}{2}[\Phi^{-1}(p)]^2} + p\sqrt{2\pi}. \quad ●$$

REMARK. *There are subsets V for which $h_1(V) = \infty$ in the class $\mathfrak{Y}_{n,p}$ for $p < 1/2$ (for example, cones with vertex at the origin).*

5. The following lemma is used later.

LEMMA (Shepp [107], via Proposition 16, or Landau and Shepp [69], directly).

Suppose given a spherically invariant measure m in $\mathbf{R}^n$ such that the Radon-Nikodým derivative dm^λ/dm, where $m^\lambda A = m((1/\lambda)A)$, depends monotonically on the norm of a point. Let V_0 be a half-space and V an arbitrary convex body containing the origin and such that $mV = mV_0$. Then $m(tV) \geqslant m(tV_0)$ for any $t \geqslant 1$.

The estimates in Propositions 8 and 18 permit us to write out two-sided estimates connecting $h_1(K)$, γK°, and κK°. The use of Proposition 21, to be proved below, is actually all right.

PROPOSITION 19. *Let K be a convex subset of a Hilbert space that contains zero, and I_0 a segment of length one with an endpoint at the origin. Then*

$$\varkappa K^\circ \leqslant \varkappa\,(h_1(K)\,I_0)^\circ.$$

PROOF. It is easy to see that Shepp's lemma can be carried over to the infinite-dimensional case by passage to the limit. From Proposition 7 it follows that the Cauchy measure satisfies the conditions of this lemma. We assume that

$$\varkappa K^\circ > \varkappa\,(h_1(K)\,I_0)^\circ.$$

Let $\lambda_0 < 1$ be such that

$$\varkappa K^\circ = \varkappa\,(\lambda_0 h_1(K)\,I_0)^\circ.$$

By Shepp's lemma,

$$\varkappa\,(\varepsilon K)^\circ \geqslant \varkappa\,(\varepsilon\lambda_0 h_1(K)\,I_0)^\circ$$

for any $\epsilon < 1$; consequently, since $\kappa\{\mathbf{0}\}^\circ = 1$,

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \varkappa\,(\varepsilon K)^\circ \geqslant \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \varkappa\,(\varepsilon\lambda_0 h_1(K)\,I_0)^\circ,$$

i.e., $h_1(K) \leqslant h_1(\lambda_0 h_1(K) I_0) = \lambda_0 h_1(K)$ (Proposition 21), which contradicts the assumption, since $\lambda_0 < 1$. ●

We now write out some estimates of the Gaussian and Cauchy measures of the polar of a convex set $K \subset H$ in terms of the value $h_1(K)$. Proposition 21 (proved later) is used for the proof of the inequality 2), but is not based on these inequalities.

PROPOSITION 20. *Let $K \subset H$ be a convex set containing the origin, and $\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^x \exp(-t^2/2)\,dt$. Then*:

$$1)\ \gamma K^\circ \geqslant 1 - \frac{1}{\sqrt{2\pi}} h_1(K);$$

$$2)\ \varkappa K^\circ \geqslant 1 - \frac{1}{\pi} h_1(K);$$

$$3)\ h_1(K) \leqslant \frac{1}{\Phi^{-1}(\gamma K^\circ)} e^{-\frac{1}{2}(\Phi^{-1}(\gamma K^\circ))^2} + \sqrt{2\pi}\,\gamma K^\circ, \quad \text{if} \quad \gamma K^\circ > \frac{1}{2};$$

$$4)\ \varkappa K^\circ \leqslant \frac{1}{2} + \frac{1}{\pi} \operatorname{arctg} \frac{1}{h_1(K)}.$$

PROOF. 1) By Proposition 14,

$$h_1(K) = \sqrt{2\pi}\int_E \sup_K x(\omega)\,d\gamma \geq \sqrt{2\pi}\int_{\{\sup_K x>1\}} d\gamma = \sqrt{2\pi}\,(1-\gamma K^\circ).$$

2) We prove that $\kappa(\epsilon K)^\circ$ is a convex function of ϵ. Preserving the notation used for the proof of Proposition 8, we can write

$$\varkappa(\varepsilon K)^\circ = \int_{E/\lambda} d(\varkappa/\lambda)\int_0^{u((\varepsilon K)^\circ,\, l)} \sqrt{\frac{2}{\pi}}\frac{1}{\sigma^2}\exp\left(-\frac{1}{2\sigma^2}\right)d\sigma = 1$$

$$-\int_{E/\lambda} d(\varkappa/\lambda)\int_0^{\frac{\varepsilon}{u(K^\circ,\, l)}} \sqrt{\frac{2}{\pi}}\exp\left(-\frac{t^2}{2}\right)dt.$$

The convexity of $\kappa(\epsilon K)^\circ$ is now clear from the fact that the Gaussian density appearing as the inside integrand is a decreasing function. The inequality 2) holds because, by Proposition 21, the slope of the graph of $\kappa(\epsilon K)^\circ$ at zero is $-h_1(K)/\pi$.

3) See Proposition 18. Another(2) (sharper) estimate could be obtained by using an idea of Fernique [30] (taking Proposition 14 into account):

$$h_1(K) \leqslant \sqrt{2\pi}\,\gamma K^\circ\left(1+\exp\left(-\frac{1}{24}\ln\frac{\gamma K^\circ}{1-\gamma K^\circ}\right)\right), \quad \gamma K^\circ > \frac{1}{2}.$$

4) See Proposition 19, computing $\kappa(h_1(K)I_0)^\circ$. From the proof of Shepp's lemma in his formulation the inequality $\kappa K^\circ \leqslant (2/\pi)\operatorname{arctg}(2/h_1(K))$ would follow for centrally symmetric sets K. ●

COROLLARY 1. *If* $\gamma K^\circ = 0$, *then* $h_1(K) \geqslant (2\pi)^{1/2}$ *(from* 1)). *If* $\gamma K_n^\circ \to 1$, *then* lim sup $h_1(K_n) \leqslant (2\pi)^{1/2}$ *(from* 3)).

COROLLARY 2. *If* $h_1(K) < \infty$, *then* $h_1(\lambda K) < (2\pi)^{1/2}$ *for some* $\lambda > 0$, *and then* $\gamma(\lambda K)^\circ > 0$ *by the inequality* 1), *i.e.*, $K \in GB$. *Conversely, if* $K \in GB$, *then, by Proposition* 1, $\gamma[(\lambda K)^\circ \cap (-\lambda K)^\circ] > 1/2$ *for some* $\lambda > 0$, *from which a fortiori* $\gamma(\lambda K)^\circ > 1/2$, *and* $h_1(K) < \infty$ *by* 3).

6. Summarizing what has been said, we formulate once more the statements proved above about the various equivalent forms of the property that a Gaussian process has bounded realizations.

THEOREM 1. *Let* $K \subset H$ *be a subset of the Hilbert space* H. *The following stetements are equivalent*:

1) *A random Gaussian process with natural parameter set* K *has with probability* 1 *sample functions that are bounded in modulus* ($K \in GB$).

2) $\kappa K^\circ > 0$, *where* κ *is the standard Cauchy measure.*

3) *If* f *is an arbitrary nonzero vector perpendicular to the affine hull of* K *and originating at some point of the convex hull of* K, *then the measure of the solid angle of the cone polar to the cone with vertex at the endpoint of* f *and generated by* K *is positive.*

4) $h_1(K) < \infty$.

(2) See also the paper of Skorohod [109].

Here h_1 is the normalized Minkowski mixed volume of first degree of homogeneity. The values γK°, κK°, and $h_1(K)$ are connected by the inequalities in Proposition 20.

§5. Monotonicity of the functional h_1

1. Although each of the properties 1)–4) in Theorem 1 characterizes the GB-sets and, consequently, can formally be used for solving the problem of extension of a Gaussian weak distribution to a measure in a Banach space, the practical verification of these properties for a process given by its correlation function is difficult. Nevertheless, a study of the geometrical characteristics of GB-sets leads to convenient estimates that allow us to give nice conditions that are easily checked for the GB and GC properties.

We recall Schlaefli's formula from spherical geometry; it has an application also in statistics.(3)

THEOREM (L. Schlaefli [104]). *Let Q be a spherical polyhedron on the unit sphere $\Sigma_n \subset \mathbf{R}^n$ with center at 0 having vertices at the points $\Pi_k \in \Sigma_n$ and given by the Gram matrix Γ of the system of vectors $\mathbf{0}\Pi_k$, $\Gamma = \Gamma_Q = (r_{jk})$, $r_{jk} = (\mathbf{0}\Pi_j, \mathbf{0}\Pi_k) = \cos\varphi_{jk}$, where φ_{jk} is the angle between the vertices Π_j and Π_k. Let m be a spherically invariant measure in $\mathbf{R}^n$. Then*

$$\frac{\partial}{\partial r_{jk}} m\,(\operatorname{con} Q)^\perp = -\frac{1}{2\pi\sqrt{1-r_{jk}^2}}\, m\,(\operatorname{con} Q_{\cdot jk})^\perp,$$

where $Q_{\cdot jk}$ is the spherical polyhedron on the sphere Σ_{n-2} determined by the Gram matrix $\Gamma_{\cdot jk} = \Gamma_{Q_{\cdot jk}} = (r_{\mu\nu\cdot jk})$ of its vertices, and

$$r_{\mu\nu\cdot jk} = \frac{\begin{vmatrix} r_{\mu\nu} & r_{\mu j} & r_{\mu k} \\ r_{j\nu} & 1 & r_{jk} \\ r_{k\nu} & r_{kj} & 1 \end{vmatrix}}{\begin{vmatrix} 1 & r_{\mu j} & r_{\mu k} \\ r_{j\mu} & 1 & r_{jk} \\ r_{k\mu} & r_{kj} & 1 \end{vmatrix}^{1/2} \cdot \begin{vmatrix} 1 & r_{\nu j} & r_{\nu k} \\ r_{j\nu} & 1 & r_{jk} \\ r_{k\nu} & r_{kj} & 1 \end{vmatrix}^{1/2}} \quad (\mu \neq j,\ k;\ \nu \neq j,\ k).$$

The matrix $\Gamma_{\cdot jk}$ is known in mathematical statistics as the matrix of partial (conditional) correlation coefficients [3]. The geometrical meaning of the polyhedron $Q_{\cdot jk}$ is the following. The vectors $\mathbf{0}\Pi_{i\cdot jk}$ that are directed from the center to the vertices of this polyhedron are the normalized projections of the vectors $\mathbf{0}\Pi_i$ ($i \neq j$, $i \neq k$) onto the orthogonal complement of the linear space spanned by the vectors $\mathbf{0}\Pi_j$ and $\mathbf{0}\Pi_k$. If all the elements of the matrix Γ are close to 1, then the spherical polyhedron Q is similar to a "planar" polyhedron (of dimension $n-1$), and then con $Q_{\cdot jk}$ is the cone obtained if the orthogonal projection of Q onto the subspace forming the orthogonal complement of the edge $\Pi_j\Pi_k$ is subjected to an infinite positive homothety with center at the projection of $\Pi_j\Pi_k$.

We make a remark with regard to the use of the words "polyhedron" and "edge". The dimension of the space in which we consider the polyhedra is unrestricted; therefore, except for "degenerate" cases, all polyhedrons under consideration are simplexes

(3) Slepian [110] first applied Schlaefli's theorem to the circle of problems under consideration.

(in a space of appropriate dimension). Similarly, by "edge" we always understand any diagonal (which in the "nondegenerate" case turns out to be a real edge).

Schlaefli's formula shows the monotonicity (in the small), with respect to the lengths of the edges, of the measure of the solid angle of the cone polar to the cone generated by a spherical polyhedron. Indeed, replacing differentiation with respect to r_{jk} by differentiation with respect to the angle φ_{jk}, we rewrite Schlaefli's formula in the form

$$\frac{\partial}{\partial\varphi_{jk}} m\,(\mathrm{con}\; Q)^{\perp} = -\frac{1}{2\pi} m\,(\mathrm{con}\; Q_{\cdot jk})^{\perp}, \tag{15}$$

from which it is obvious that the right-hand side is nonpositive.

Since the measure of the solid angle of the cone standing after the derivative sign on the left-hand side of (15) can be written in terms of the Cauchy measure, and for small sizes of the spherical polyhedron Q the lengths of the edges of its central projection $K = K(Q)$ onto the affine hyperplane tangent to the sphere $\Sigma \supset Q$ at some point in Q are close to the lengths of its own edges, it is natural to use Schlaefli's formula in trying to get a theorem on the monotonicity of the Cauchy measure of the polar of the set K with respect to contractive mappings (such an approach was used in [121]). However, the Cauchy measure of the polar of K depends not only on the intrinsic metric of K, but also on its position with respect to the zero point of the space H, while the GB property is invariant with respect to any translations. Moreover, it is good to use Schlaefli's formula for investigating small sets K that differ only slightly from their projections $Q(K)$ onto the unit sphere. It seems preferable to us, therefore, to work, not with Cauchy measure, but with the mixed volume h_1, for which Schlaefli's formula can be used to prove an interesting property of monotonicity with respect to a certain class of "contractive" mappings, and this monotonicity can be used to obtain convenient conditions for the finiteness of $h_1(K)$ and estimates of the Gaussian and Cauchy measures of the polar of K.

2. We first prove some auxiliary propositions.

PROPOSITION 21. *Let $K \subset H$ be a set such that the zero point belongs to its closed convex hull. Then*

$$\left.\frac{d\varkappa\,(\varepsilon K)^{\circ}}{d\varepsilon}\right|_{\varepsilon=0} = -\frac{1}{\pi} h_1(K).$$

PROOF. As before, let $u(l, K^{\circ})$ denote the length of the intersection $l \cap K^{\circ}$ (in the scale determined by the position of the δ-measure that is the conditional measure of Gaussian measure on the ray l). Let $\omega = (l, \sigma)$; then

$$(\sup_{x\in K} x)\,(l,\ 1) = \frac{1}{u\,(l,\ K^{\circ})}$$

and, consequently,

$$h_1(K) = \sqrt{2\pi} \int_E \frac{d\gamma}{u\,(l,\ K^{\circ})} = \sqrt{2\pi} \int_{E/\lambda} \frac{(\gamma/\lambda)\,(dl)}{u\,(l,\ K^{\circ})}. \tag{16}$$

Furthermore,

$$1-\varkappa(\varepsilon K)^\circ=\int_{E/\lambda}(\varkappa/\lambda)\,(dl,\int_{\frac{1}{\varepsilon}u(l,\,K^\circ)}^{\infty}\sqrt{\frac{2}{\pi}}\frac{1}{\sigma^2}\exp\left(-\frac{1}{2\sigma^2}\right)d\sigma$$

$$=\int_{E/\lambda}(\gamma/\lambda)\,(dl)\int_0^{\frac{\varepsilon}{u(l,\,K^\circ)}}\sqrt{\frac{2}{\pi}}\exp\left(-\frac{t^2}{2}\right)dt\quad\left(t=\frac{1}{\sigma}\right).$$

Finally, we get, by (16),

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\varkappa(\varepsilon K)^\circ=-\sqrt{\frac{2}{\pi}}\int_{E/\lambda}(\gamma/\lambda)\,(dl)\frac{e^{-\frac{1}{2}\left(\frac{\varepsilon}{u(l,\,K^\circ)}\right)^2}}{u\,(l,\;K^\circ)}\Bigg|_{\varepsilon=0}=-\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{2\pi}}h_1\,(K).\;\bullet$$

PROPOSITION 22. *In a Euclidean space suppose that the affine hyperplane L is tangent to the unit sphere Σ at the point x_0, and the points $x_1, x_2 \in L$ are each at a distance not greater than $\epsilon > 0$ from the point of tangency. Moreover, suppose that the segment x_1x_2 is rotated through an angle ψ around x_1. Then the following assertions are true:*

1) *The angle φ at which the segment x_1x_2 is seen from the center of the sphere Σ satisfies the inequalities*

$$\frac{|\,x_1x_2\,|}{1+\varepsilon^2}\leqslant\varphi\leqslant|\,x_1x_2\,|,$$

where $|x_1x_2|$ is the length of the segment joining x_1 and x_2.

2) *For a rotation of the segment through an angle ψ around x_1,*

$$\left|\frac{d\varphi}{d\psi}\right|\leqslant 4\varepsilon.$$

The proof is elementary.

3. We proceed to the proof of the monotonicity of h_1. From Proposition 14 it follows that

$$\varkappa\,(\varepsilon K)^\circ=1-\frac{\varepsilon}{\pi}h_1\,(K)+\varepsilon^2O\,(1).$$

If we differentiate this equation formally with respect to the length of some edge l_{ik} (without specifying, for the time being, what this means), and then divide by ϵ and set $\epsilon = 0$, then we get, successively,

$$\frac{\partial}{\partial l_{ik}}\varkappa\,(\varepsilon K)^\circ=-\frac{\varepsilon}{\pi}\frac{\partial}{\partial l_{ik}}h_1\,(K)+\varepsilon^2\frac{\partial}{\partial l_{ik}}O,$$

$$\frac{1}{\varepsilon}\frac{\partial}{\partial l_{ik}}\varkappa\,(\varepsilon K)^\circ=-\frac{1}{\pi}\frac{\partial}{\partial l_{ik}}h_1\,(K)+\varepsilon\frac{\partial}{\partial l_{ik}}O.$$

When ϵ is made small, the left-hand side, by Schlaefli's formula (15) (ϵK approximates $Q(\epsilon K)$, and l_{ik} approximates φ_{ik}), passes into

$$-\frac{1}{\pi}m\,(\mathrm{con}\,Q_{\cdot ik})^\perp=-\frac{1}{\pi}m\,(\mathrm{con}\,(K_{ik},\;P_{ik}))^\perp,$$

while the second term of the right-hand side vanishes, and we arrive at the equation

$$\frac{\partial}{\partial l_{ik}} h_1(K) = m(\mathrm{con}(K_{ik}, P_{ik}))^{\perp}.$$

In this argument we have not considered that the quantity $\kappa(\epsilon K)^{\circ}$ depends not only on the lengths of the edges l_{ik}, but also on the position of the zero point. Also, why we could write $\lim \epsilon(\partial/\partial l_{ij})O = 0$ was not justified. We now present this argument in a rigorous sense.

PROPOSITION 23. *Let K be a nondegenerate convex polyhedron specified by the lengths of all its edges l_{ij}, i.e., suppose that for some $\delta > 0$ there exist polyhedra with arbitrary lengths of sides l'_{ij} such that $|l_{ij} - l'_{ij}| < \delta$. Then*

$$\frac{\partial}{\partial l_{ij}} h_1(K) = m(\mathrm{con}(K_{ij}, P_{ij}))^{\perp} \geqslant 0. \tag{17}$$

PROOF. We consider the family of polyhedra K_λ whose members differ from one another by the length of only one fixed edge, so that

$$l_{i_0 j_0}^{(\lambda)} = l_{i_0 j_0} + \lambda, \quad l_{ij}^{(\lambda)} = l_{ij}, \quad \text{if} \quad (i, j) \neq (i_0, j_0).$$

By the nondegeneracy assumption, the family $\{K_\lambda\}$ contains polyhedra for all λ in a δ-neighborhood of zero. Let the total number of vertices of the polyhedron K be n. We now consider the unit sphere $\Sigma \subset \mathbf{R}^n$ tangent at the vertex P_{i_0} to the polyhedron ϵK_λ, which lies in some hyperplane of $\mathbf{R}^n$. Together with the polyhedron we consider its central projection $Q_{\lambda\epsilon} = Q(\epsilon K_\lambda)$ onto Σ. Let $\varphi_{ik}^{\epsilon}(\lambda)$ be the lengths of the projections of the edges ϵl_{ik} onto Σ. From Proposition 22 and the nondegeneracy of K it follows that

$$\left| \frac{d\varphi_{ij}^{\varepsilon}}{d\lambda} \right| \leqslant C_{ij}(K)\,\varepsilon^2, \tag{18}$$

$$\left| \frac{d\varphi_{i_0 j_0}^{\varepsilon}}{d\lambda} - \varepsilon \right| \leqslant C_{i_0 j_0}(K)\,\varepsilon^2. \tag{19}$$

Therefore

$$\varkappa(\varepsilon K_\lambda)^{\circ} = \varkappa(\varepsilon K_0)^{\circ} + \lambda \frac{d}{d\lambda} \varkappa(\varepsilon l_{ij}(\lambda)) \Big|_{\lambda=\tilde\lambda} = \varkappa(\varepsilon K_0)^{\circ} + \lambda \frac{d}{d\lambda} \varkappa(\{\varphi_{ij}^{\varepsilon}(\lambda)\}) \Big|_{\lambda=\tilde\lambda}$$

$$= \varkappa(\varepsilon K_0)^{\circ} + \lambda \left[\frac{\partial \varkappa}{\partial \varphi_{i_0 j_0}} \frac{d\varphi_{i_0 j_0}^{\varepsilon}}{d\lambda} + \sum_{i,j} \frac{\partial \varkappa}{\partial \varphi_{ij}} \frac{d\varphi_{ij}^{\varepsilon}}{d\lambda} \right] \Bigg|_{\lambda=\tilde\lambda}$$

$$= \varkappa(\varepsilon K_0)^{\circ} + \lambda\varepsilon \left[\frac{\partial \varkappa}{\partial \varphi_{i_0 j_0}} \frac{1}{\varepsilon} \left(\varepsilon + \left(\frac{d\varphi_{i_0 j_0}^{\varepsilon}}{d\lambda} - \varepsilon \right) \right) + \sum \frac{\partial \varkappa}{\partial \varphi_{ij}} \frac{1}{\varepsilon} \frac{d\varphi_{ij}^{\varepsilon}}{d\lambda} \right] \qquad (\tilde\lambda \in [0, 1]).$$

Consequently

$$\frac{1}{\varepsilon}(1 - \varkappa(\varepsilon K_\lambda)^{\circ})$$

$$= \frac{1}{\varepsilon}(1 - \varkappa(\varepsilon K)^{0}) - \lambda \left[\frac{\partial \varkappa}{\partial \varphi_{i_0 j_0}} \left(1 + \frac{1}{\varepsilon} \left(\frac{d\varphi_{i_0 j_0}^{\varepsilon}}{d\lambda} - \varepsilon \right) \right) + \sum \frac{\partial \varkappa}{\partial \varphi_{ij}} \frac{1}{\varepsilon} \frac{d\varphi_{ij}^{\varepsilon}}{d\lambda} \right] \Bigg|_{\lambda=\tilde\lambda}.$$

If for fixed λ, $|\lambda| < \delta$, we now let ϵ go to zero, then, by (18), (19), and Proposition 21, taking into account that, by Schlaefli's theorem, $|\partial\kappa/\partial\varphi_{ij}| \leqslant 1$, we find that

$$\frac{\frac{1}{\varepsilon}(1-\varkappa(\varepsilon K_\lambda)^\circ) - \frac{1}{\varepsilon}(1-\varkappa(\varepsilon K)^\circ)}{\lambda} \sim -\frac{\partial \varkappa(\varepsilon K)^\circ}{\partial \varphi_{i_0 j_0}};$$

consequently,

$$\varlimsup_{\varepsilon}\left(\sup_{\tilde\lambda\in(0,\delta)}\left(-\frac{\partial\varkappa(\varepsilon K_{\tilde\lambda})^\circ}{\partial\varphi_{i_0j_0}}\right)\right) \leqslant \frac{1}{\lambda}\left(\frac{1}{\pi}h_1(K_\lambda) - \frac{1}{\pi}h_1(K)\right) \leqslant \varlimsup_{\varepsilon}\left(\inf_{\tilde\lambda\in(0,\delta)}\left(-\frac{\partial\varkappa(\varepsilon K_{\tilde\lambda})^\circ}{\partial\varphi_{i_0j_0}}\right)\right)$$

and as $\delta \to 0$, since, by (15),

$$\frac{\partial\varkappa(\varepsilon K)^\circ}{\partial\varphi_{i_0j_0}} \to -\frac{1}{\pi} m \operatorname{con}(K_{i_0j_0}, P_{i_0j_0})^\perp,$$

we find at last that

$$\frac{d}{d\lambda} h_1(K_\lambda) = m \operatorname{con}(K_{i_0j_0}, P_{i_0j_0})^\perp \geqslant 0. \bullet$$

In the following we shall only need the qualitative part of the last assertion, which is expressed by the inequality.

REMARK. The inequality (17) means the local monotonicity of $h_1(K)$ as a function of the lengths of the edges. The condition of nondegeneracy used in the proof is, in fact, not essential, and (17) is carried over by continuity to all polyhedra for which it makes sense. (Every polyhedron can be approximated by nondegenerate ones.) The differentiation with respect to the length l_{ij} of an edge can be replaced by differentiation with respect to the direction of any vector with positive coefficients consisting of some of the l_{ij}.

4. We now prove the global monotonicity of $h_1(K)$ as a function of the numbers l_{ij}, $i = 1, \dots, n-1$, $i < j \leqslant n$.

We consider the space $\mathbf{R}^{n(n-1)/2}$ and in it the subset

$$D = \{(l_{ij} : i = 1, \dots, n-1,\ i < j \leqslant n)\} \subset \mathbf{R}^{\frac{n(n-1)}{2}}$$

of all points for which there exist polyhedra with lengths l_{ij} of edges (different points can correspond to congruent polyhedra). Let $\mathring{D} \subset D$ be the set of points corresponding to nondegenerate polyhedra. We first describe the set $D \subset \mathbf{R}^{n(n-1)/2}$ in other terms.

PROPOSITION 24. *For the numbers l_{ij}, $i = 1, \dots, n-1$, $i < j \leqslant n$, to be the lengths of the edges of some nondegenerate polyhedron, i.e., for the condition $l = (l_{ij}) \in \mathring{D}$ to hold, it is necessary and sufficient that the matrix*

$$\Lambda(R) = \begin{pmatrix} 1 & & & 1-\frac{l_{ik}^2}{2R^2} \\ & 1 & & \\ & & \ddots & \\ 1-\frac{l_{ik}^2}{2R^2} & & & 1 \end{pmatrix}_{n\times n}$$

is positive definite for all $R > R_0$ (for some R_0).

PROOF. *Necessity*. Let $l = (l_{ij}) \in \mathring{\mathcal{D}}$. We show that for some $R_0 = R_0(l)$ the matrix $\Delta(R)$ is positive definite. Let the polyhedron $K(l)$ with lengths of edges (l_{ij}) be located on some hyperplane in $\mathbf{R}^n$, and let $\Sigma_R \subset \mathbf{R}^n$ be a sphere of sufficiently large radius R that is tangent to this hyperplane at a point in $K(l)$. The unit vectors $e_{k,R}$ drawn from the center of Σ_R in the direction of the vertices P_k, $k = 1, \dots, n$, of $K(l)$ form on the unit sphere Σ_1 with the same center as Σ_R the vertices of a nondegenerate spherical simplex. The Gram matrix $\Gamma_{k,R} = ((e_{i,R}, e_{j,R}))_{n\times n}$ of these vectors is therefore positive definite. The angle $\alpha_{i,j,R}$ between the vectors $e_{i,R}$ and $e_{j,R}$ differs from l_{ij}/R by an infinitesimally small quantity having order $O(1/R^3)$, and the value $(e_{i,R}, e_{j,R})$ of the cosine of this angle (the element of the Gram matrix) differs from $1 - (\alpha_{i,j,R})^2/2$ by an infinitesimally small quantity having order $O(1/R^4)$; hence

$$(e_{i,R}, e_{j,R}) = 1 - \frac{1}{2}\left(\frac{l_{ij}}{R} + O\left(\frac{1}{R^3}\right)\right)^2 + O\left(\frac{1}{R^4}\right)$$

$$= 1 - \frac{1}{2}\frac{l_{ij}^2}{R^2} + \frac{\beta_{ij}}{R^4}, \quad \text{where } |\beta_{ij}| \leqslant C < \infty.$$

From the geometric meaning of the size of the determinant of the Gram matrix it follows immediately that for increasing radius R the quantity $\det \Gamma_R$ has order of decrease

$$\det \Gamma_R = O\left(\frac{1}{R^{2(n-1)}}\right),$$

and from the nondegeneracy of the system of vectors $\{e_{i,1}\}$ it follows that this estimate of the order of $\det \Gamma_R$ is sharp, i.e., for some $c > 0$

$$\det \Gamma_R > \frac{c}{R^{2(n-1)}} \quad \text{for} \quad R > 1. \tag{20}$$

(We recall that the determinant of the Gram matrix of a system of vectors is equal to the square of the volume of the parallelepiped determined by these vectors.)

The verification of the positive definiteness of the matrix $\Lambda(R)$ consists of verifying the positivity of its principal minors. We prove that for sufficiently large R they really are positive. The following equality is obvious:

$$\det\begin{pmatrix} 1 & (e_{1,R}, e_{2,R}) & \cdots\cdots & (e_{1,R}, e_{n,R}) \\ (e_{2,R}, e_{1,R}) & 1 & & (e_{2,R}, e_{n,R}) \\ \cdots & \cdots & \cdots\cdots & \cdots \\ (e_{n,R}, e_{1,R}) & (e_{n,R}, e_{2,R}) & \cdots\cdots & 1 \end{pmatrix}$$

$$= \det\begin{pmatrix} 1 & -\frac{1}{2}\frac{(l_{12}^{(j)})^2}{R^2} + \frac{\beta_{12}}{R^4} & -\frac{1}{2}\frac{(l_{13}^{(j)})^2}{R^2} + \frac{\beta_{13}}{R^4} & \cdots & -\frac{1}{2}\frac{(l_{1n}^{(j)})^2}{R^2} + \frac{\beta_{1n}}{R^4} \\ 1 & 0 & -\frac{1}{2}\frac{(l_{23}^{(j)})^2}{R^2} + \frac{\beta_{23}}{R^4} & \cdots & -\frac{1}{2}\frac{(l_{2n}^{(j)})^2}{R^2} + \frac{\beta_{2n}}{R^4} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & -\frac{1}{2}\frac{(l_{n2}^{(j)})^2}{R^2} + \frac{\beta_{n2}}{R^4} & -\frac{1}{2}\frac{(l_{n3}^{(j)})^2}{R^2} + \frac{\beta_{n3}}{R^4} & \cdots & 0 \end{pmatrix} +$$

$$
+\det\begin{pmatrix}
0 & 1 & -\frac{1}{2}\frac{(l_{13}^{(j)})^2}{R^2}+\frac{\beta_{13}}{R^4} & \cdots & -\frac{1}{2}\frac{(l_{1n}^{(j)})^2}{R^2}+\frac{\beta_{1n}}{R^4} \\
-\frac{1}{2}\frac{(l_{21}^{(j)})^2}{R^2}+\frac{\beta_{21}}{R^4} & 1 & -\frac{1}{2}\frac{(l_{23}^{(j)})^2}{R^2}+\frac{\beta_{23}}{R^4} & \cdots & -\frac{1}{2}\frac{(l_{2n}^{(j)})^2}{R^4}+\frac{\beta_{2n}}{R^4} \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
-\frac{1}{2}\frac{(l_{n1}^{(j)})^2}{R^2}+\frac{\beta_{n1}}{R^4} & 1 & -\frac{1}{2}\frac{(l_{n3}^{(j)})^2}{R^2}+\frac{\beta_{n3}}{R^4} & \cdots & 0
\end{pmatrix}+\cdots
$$

$$
+\det\begin{pmatrix}
0 & -\frac{1}{2}\frac{(l_{12}^{(j)})^2}{R^2}+\frac{\beta_{12}}{R^4} & \cdots\cdots & 1 \\
-\frac{1}{2}\frac{(l_{21}^{(j)})^2}{R^2}+\frac{\beta_{21}}{R^4} & 0 & \cdots\cdots & 1 \\
\cdots & \cdots & \cdots\cdots & \cdots \\
-\frac{1}{2}\frac{(l_{n1}^{(j)})^2}{R^2}+\frac{\beta_{n1}}{R^4} & -\frac{1}{2}\frac{(l_{n2}^{(j)})^2}{R^2}+\frac{\beta_{n2}}{R^4} & \cdots\cdots & 1
\end{pmatrix}
$$

$$
+\det\begin{pmatrix}
0 & -\frac{1}{2}\frac{(l_{12}^{(j)})^2}{R^2}+\frac{\beta_{12}}{R^4} & \cdots\cdots & -\frac{1}{2}\frac{(l_{1n}^{(j)})^2}{R^2}+\frac{\beta_{1n}}{R^4} \\
-\frac{1}{2}\frac{(l_{21}^{(j)})^2}{R^2}+\frac{\beta_{n1}}{R^4} & 0 & \cdots\cdots & -\frac{1}{2}\frac{(l_{2n}^{(j)})^2}{R^2}+\frac{\beta_{2n}}{R^4} \\
\cdots & \cdots & \cdots\cdots & \cdots \\
-\frac{1}{2}\frac{(l_{n1}^{(j)})^2}{R^2}+\frac{\beta_{n1}}{R^4} & -\frac{1}{2}\frac{(l_{n2}^{(j)})^2}{R^2}+\frac{\beta_{n2}}{R^4} & \cdots\cdots & 0
\end{pmatrix}.
$$

From this representation of $\det \Gamma_R$ it is clear that the principal term, which has order (as could have been foreseen) $O(1/R^{2(n-1)})$, does not depend on the values β_{ik}, which contribute only to the terms of order not less than $1/R^{2n}$, and, as mentioned, the coefficient of this principal term is trivially nonzero (see (20)). From this and from the boundedness of the quantities β_{ij} it follows that for sufficiently large R ($R > R_1$) the sign of the whole expression coincides with the sign of the principal term. Therefore, since $\det \Gamma_R > 0$ always, the determinant of the matrix $\Lambda(R)$, which is a principal part of $\det \Gamma_R$, is positive for $R > R_1$.

The positivity of the other principal minors for $R > R_k$, $k = 2, \dots, n$, is established similarly, and this concludes the proof of the necessity: $R_0 = \max_{1 \le k \le n} R_k$.

Sufficiency. Suppose that the matrix $\Lambda(R)$ is positive definite for some sequence $\{R_m\}$ of values R, $R_m \to \infty$. To prove the existence of a polyhedron with lengths of edges equal to l_{ik} we use the obvious closedness of the set $D \subset \mathbf{R}^n$ and construct a sequence of polyhedra K_m, each having n vertices, such that $l_{ij}^{(m)} \to l_{ij}$ (the $l_{ij}^{(m)}$ are the lengths of the edges of K_m). With this aim, we consider for each m a spherical polyhedron on the unit sphere Σ_1 for which $\Delta(R_m)$ is the Gram matrix. Then we subject Σ_1, together with the polyhedron, to a homothety that increases its radius to R_m, we consider some hyperplane tangent to the sphere Σ_{R_m} at a point of the spherical polyhedron on its surface, and we project this polyhedron from the center of the sphere onto this tangent hyperplane; let K_m be this projection. As $m \to \infty$ all the elements of the matrix $\Delta(R_m)$ converge to 1; hence the spherical polyhedron for which $\Lambda(R_m)$ is the

Gram matrix contracts, and its diameter converges to zero as $1/R_m$. Therefore, by Proposition 22, the lengths of the edges of the spherical polyhedra on the spheres Σ_{R_m} differ from the lengths of the edges of their central projections onto the tangent hyperplanes by quantities of order $O(1/R_m)$. On the other hand, the lengths of the edges of the spherical polyhedra on Σ_1 are equal, respectively, to $\arccos(1 - l_{ij}^2/R_m^2)$, and on Σ_{R_m} the lengths of the edges of the corresponding polyhedron are equal to

$$R_m \arccos\left(1 - \frac{l_{ij}^2}{R_m^2}\right) \to l_{ij}.$$

With this, it is proved that the polyhedra K_m approximate some polyhedron with lengths of edges l_{ij}. The proof of Proposition 24 is concluded. ●

For the proof of the global monotonicity of h_1 we prove that the set D is "quadratically convex".

PROPOSITION 25. *Let $K^{(0)}$ and $K^{(1)}$ be two polyhedra with the same number of vertices, and let $l_{ij}^{(k)}$, $k = 0, 1$, be the lengths of the corresponding edges. Then for any number λ, $0 \leqslant \lambda \leqslant 1$, there is a polyhedron K_λ with the same number of vertices for which the lengths of the corresponding edges are*

$$l_{ij}^{(\lambda)} = \sqrt{\lambda l_{ij}^{(1)^2} + (1-\lambda)\, l_{ij}^{(0)^2}}. \tag{21}$$

PROOF. We first assume that $K^{(0)}$ and $K^{(1)}$ are nondegenerate, i.e., that the points $l^{(0)}, l^{(1)} \in D$ corresponding to these polyhedra belong to the interior $\mathring{D}$ of D. Let $R_0 > 0$ be such that the matrices $\Delta^{(k)}(R)$, $k = 0, 1$, $R > R_0$, formed with respect to the numbers $l_{ij}^{(k)}$ are positive definite (Proposition 24). By the convexity of the cone of positive definite matrices, for each λ, $0 \leqslant \lambda \leqslant 1$, the matrices

$$\Lambda^{(\lambda)}(R) = \begin{pmatrix} 1 & & 1 - \dfrac{l_{ij}^{(\lambda)^2}}{2R^2} \\ & 1 & \\ & \ddots & \\ 1 - \dfrac{l_{ij}^{(\lambda)^2}}{2R^2} & & 1 \end{pmatrix}_{n\times n}, \qquad R > R_0,$$

are positive definite, and then there exists a polyhedron with lengths of edges $l_{ij}^{(\lambda)}$.

Now if one or both of the points $l^{(k)}$, $k = 0, 1$, lies on the boundary of D, we consider two sequences of points $l_m^{(k)}$ $(k = 0, 1;\ m = 1, 2, \ldots)$, $l_m^{(k)} \to l^{(k)}$, $l_m^{(k)} \in \mathring{D}$, and for each m we consider a curve $\lambda \to l_m^{(\lambda)}$, $\lambda \in [0, 1]$, lying in the set D as described above and joining the points $l_m^{(0)}$ and $l_m^{(1)}$. Obviously, the limit curve does not go outside of D and is described by the same equation (21) as the sequence curves. ●

The global monotonicity of h_1 is now simple to prove.

THEOREM 2. *Let $K^{(0)}$ and $K^{(1)}$ be two polyhedra with the same number of vertices such that $l_{ij}^{(0)} \leqslant l_{ij}^{(1)}$ for any pair of corresponding edges $l_{ij}^{(0)}$ and $l_{ij}^{(1)}$. Then $h_1(K^{(0)}) \leqslant h_1(K^{(1)})$.*

PROOF. Let $l^{(0)}$ and $l^{(1)}$ be the points of the subset D corresponding to $K^{(0)}$ and

$K^{(1)}$. By Proposition 25, these points can be joined by a curve $\lambda \to l^{(\lambda)} \in D$, $\lambda \in [0, 1]$, described by (21). By Proposition 23 and the accompanying remark, if b^λ denotes the direction of the tangent to this curve at the point $l^{(\lambda)}$, then

$$\frac{\partial}{\partial b^{(\lambda)}} h_1(K^{(\lambda)}) \geqslant 0$$

($K^{(\lambda)}$ corresponds to the point $l^{(\lambda)}$). Therefore, we get

$$h_1(K^{(1)}) - h_1(K^{(0)}) = \int_0^1 \frac{\partial}{\partial b^{(\lambda)}} h_1(K^{(\lambda)})\, dl^{(\lambda)} \geqslant 0. \bullet$$

COROLLARY. *Let K_1 and K_2 be subsets of the space H, and $Z_1 \subset K_1$ and $Z_2 \subset K_2$ subsets of them such that $K_1 \subset \overline{\operatorname{conv} Z_1}$ and $K_2 \subset \overline{\operatorname{conv} Z_2}$. If there exists an epimorphism $\varphi\colon Z_1 \to Z_2$ such that*

$$\|x' - x''\| \leqslant \|\varphi(x') - \varphi(x'')\|, \quad x',\ x'' \in Z_1,$$

then $h_1(K_1) \leqslant h_1(K_2)$.

Indeed, let $x_k \in Z_1$ be a dense subset of Z_1 such that $\{\varphi(x_k)\}$ is dense in Z_2. Then (Proposition 13)

$$h_1(K_1) = \lim_n h_1(\{x_k,\ k = 1, \dots, n\}),$$
$$h_1(K_2) = \lim_n h_1(\{\varphi(x_k),\ k = 1, \dots, n\})$$

and the required relation follows from Theorem 2.

§6. Mixed volumes and the continuity of paths

1. Here we prove a certain criterion for the continuity of realizations of a Gaussian process, formulated in terms of the functional h_1. As in the case of the *GB* property, this criterion appears at first glance to be difficult to apply, but it enables us later to obtain nice conditions in the more convenient ϵ-entropy language. We first prove an auxiliary statement. Let $K \subset H$ be a bounded, convex, balanced subset of Hilbert space, $F = L(K)$, and $F_{(*)}$ the space of all linear forms on F that are bounded on K and continuous on K in the Hilbert weak topology.

PROPOSITION 26. *Let $\{L_n, n = 1, \dots\}$ be a decreasing sequence of closed linear subspaces of H with finite defects and intersection zero. Then*

$$F_{(*)} = \bigcap_{\lambda > 0} \lambda \bigcup (K \cap L_n)^\circ, \tag{22}$$

where the polar is understood in the sense of the duality (F, F^*), F^* being the strong dual of the space $(F, \|\cdot\|_K)$ (the norm $\|\cdot\|_K$ is generated by the unit ball determined by the set K).

PROOF. We first observe that the right-hand side of (22) consists of those and only those elements $y \in F^*$ for which

$$m_n(y) = \sup_{x \in K \cap L_n} |y(x)| \underset{n \to \infty}{\to} 0. \tag{23}$$

Indeed, if $m_n(y) \geqslant \alpha > 0$, then for some sequence $x_n \in K \cap L_n$ we have $|y(x_n)| > \alpha/2 > 0$; consequently $y \notin (\alpha/2)(K \cap L_n)^\circ$ for all $n = 1, \dots$, i.e.,

$$y \notin \frac{\alpha}{2} \bigcup_n (K \cap L_n)^\circ$$

and a fortiori $y \notin \bigcap_\lambda \lambda \bigcup_n (K \cap L_n)^\circ$. Conversely, if $m_n(y) \longrightarrow 0$, then for any λ and any $n > N(\lambda)$ we have that $y \in \lambda (K \cap L_n)^\circ$, i.e., $y \in \lambda \bigcup_n (K \cap L_n)^\circ$, and $y \in \bigcap_\lambda \lambda \bigcup_n (K \cap L_n)^\circ$. Now suppose that $y \in F_{(*)}$. Then for any sequence $x_n \in K \cap L_n$ we have $y(x_n) \longrightarrow 0$, from which it follows that $m_n(y) \longrightarrow 0$, and, by the above, $y \in \bigcap_\lambda \lambda \bigcup_n (K \cap L_n)^\circ$.

Conversely, let $y \in \bigcap_\lambda \lambda \bigcup_n (K \cap L_n)^\circ$. We prove the continuity of y on K. It suffices to verify continuity with respect to sequences converging to zero. Let $x_k \in K$, $x_k \longrightarrow 0$ (weakly in H). From the weak convergence it follows that x_k is approximable in any closed subspace of finite defect with respect to the norm $\|\cdot\|_K$; in particular, any of the subspaces L_n (the quotient space by L_n is finite-dimensional, and on it all forms of convergence coincide). Since $y \in F^*$, i.e., is bounded, we have that

$$|y(x_k)| \leqslant m_n(y) + \|y\|_K^* \sup_{x \in K \cap L_n} \|x_K - x\|_K$$

and as $k \longrightarrow \infty$ we find

$$\lim_{k \to \infty} |y(x_k)| \leqslant m_n(y),$$

i.e., by (23), $y(x_k) \longrightarrow 0$, whence $y \in F_{(*)}$. ●

As a corollary of this characteristic of the set $F_{(*)}$ of linear forms that are bounded and continuous on K we get a statement of Dudley.

PROPOSITION 27 (Dudley [33]). *Let $K \subset H$ be a Gaussian process. If the measure of the space $F_{(*)}(K)$ of linear forms that are continuous and bounded on K is equal to zero, then for some $\epsilon > 0$ the realizations are not bounded in modulus by the number ϵ with probability* 1, *and conversely.*

PROOF. For any $\lambda > 0$ the set $M_\lambda = \lambda \bigcup_n (\mathrm{conv}\, K \cap L_n)^\circ$ satisfies the conditions of applicability of Kolmogorov's zero-one law (the assertion 2) of Proposition 1), since the inclusion $y \in M_\lambda$ depends only on the behavior of the values $y(x)$ for $x \in L_n$ with arbitrarily large number n. Consequently, if $\gamma F_{(*)} = \gamma \bigcap_\lambda M_\lambda = 0$, then $\gamma M_{\lambda_0} = 0$ for some $\lambda_0 > 0$, i.e., we can take $\epsilon = \lambda_0$. Conversely, if $\gamma M_{\lambda_0} = 0$, then also $\gamma F_{(*)} = \gamma \bigcap_\lambda M_\lambda = 0$. ●

2. We now proceed to the derivation of a criterion for the continuity of realizations of Gaussian processes in terms of the functional h_1.

DEFINITION. Let $K \subset H$ be a Gaussian random process. Then

$$\delta(K) = \sup\{d : \gamma(dK^\circ) = 0\}.$$

Thus, $K \in GB \Leftrightarrow \delta(K) < \infty$, and $K \in GC \Leftrightarrow \delta(K) = 0$ (Proposition 27).

PROPOSITION 28. *Let $K \subset H$ be a convex balanced set, and $L \subset H$ a closed subspace of finite defect. Then the following assertions are true:*

1) *If $\gamma((1 + \epsilon)K^\circ) = 0$ for some $\epsilon > 0$, then $\gamma(K \cap L)^\circ = 0$.*

2) *If* $\gamma K^\circ = 0$, *then also* $\gamma(\mathrm{Pr}_L\, K)^\circ = 0$ (Pr_L *is the operator of orthogonal projection onto* L).

PROOF. The second assertion follows at once form the fact that the conditional Gaussian measures of the set K° under the decomposition into the cosets in the factorization by L° are equal to zero. We prove 1). We can assume that the defect of the subspace L is 1. If $y_L \in H^* \subset E$, $\|y_L\|_{H^*} = 1$, is a functional on H having L as kernel, then $(K \cap L)^\circ = K^\circ + \{\lambda y_L : \lambda \in \mathbf{R}\}$. We consider the measurable decomposition ξ_L of (E, γ) into lines parallel to the line $\{\lambda y_L\}$. The space (E, λ) can be regarded as a product of measure spaces: the product of the line $\{\lambda y_L\}$ with standard one-dimensional Gaussian measure and the Gaussian measure space (E_1, γ_1), where E_1 can be identified with the subspace of E on which a function $x_L \in H$ orthogonal to L vanishes; the variance ellipsoid of the measure γ_1 coincides with the intersection of the unit ball V_{H^*} with E_1. By the boundedness of $K \subset H$, the line $\{\lambda y_L\}$ intersects K° in a nondegenerate segment, i.e., in a set of positive conditional (one-dimensional) Gaussian measure, and, because of the convexity of K° and the boundedness of $K \subset H$, each line parallel to the line $\{\lambda y_L\}$ and intersecting K° intersects $(1 + \epsilon)K^\circ$ in a nondegenerate segment. But from the condition $\gamma[(1 + \epsilon)K^\circ] = 0$ it then follows that the set of lines intersecting K° has (γ/ξ_L)-measure zero, with which the proof is concluded. A more exacting argument would show that it is even possible to assume that $\epsilon = 0$. ●

COROLLARY. *If* $K \subset H$ *is a convex balanced set and* $L \subset H$ *a closed subspace of finite defect, then*

$$\delta(K) = \delta(K \cap L) = \delta(\mathrm{Pr}_L K).$$

PROPOSITION 29. *Let* A_λ *be an operator in the space* (E, γ) *effecting a dilation by a factor of* λ $(\lambda > 1)$ *in some fixed direction. Then* $\gamma(A_\lambda V) \geqslant \gamma V$ *for any convex balanced subset* $V \subset E$.

PROOF. It suffices to limit ourselves to the case when $\gamma = \gamma_n$ is standard Gaussian measure in $\mathbf{R}^n$; then A_λ is an arbitrary selfadjoint operator with spectrum $(1, \ldots, 1, \lambda)$, $\lambda > 1$. Let $L \subset \mathbf{R}^n$ be the characteristic subspace corresponding to the eigenvalue 1 of A_λ, and let $L_t = L + te_n$, where $\|e_n\| = 1$ and $A_\lambda e_n = \lambda e_n$. Let Pr_L be the orthogonal projection onto L, and γ_{n-1} the standard (conditional) Gaussian measure on L $(\gamma_{n-1} = \gamma_n \mathrm{Pr}_L^{-1})$. Then

$$\mathrm{Pr}_L(A_\lambda V \cap L_t) = \mathrm{Pr}_L(V \cap L_{t/\lambda}). \tag{24}$$

From the convexity of the set V it follows that

$$\frac{\lambda+1}{2\lambda}(V \cap L_t) + \frac{\lambda-1}{2\lambda}(V \cap L_{-t}) \subset V \cap L_{t/\lambda}$$

or, what is equivalent,

$$\frac{\lambda+1}{2\lambda}\mathrm{Pr}_L(V \cap L_t) + \frac{\lambda-1}{2\lambda}\mathrm{Pr}_L(V \cap L_{-t}) \subset \mathrm{Pr}_L(V \cap L_{t/\lambda}). \tag{25}$$

The sets $\mathrm{Pr}_L(V \cap L_t) \subset L$ and $\mathrm{Pr}_L(V \cap L_{-t}) \subset L$ are centrally symmetric to each other. By a theorem of Zalgaller [146] (based on application of the Brunn–Minkowski inequality for mixed volumes to the left-hand side of (25); see, for example, [2], Chapter VIII, §3), if A and A' are two convex bodies that are centrally symmetric to each other and $0 \leqslant \alpha \leqslant 1$, then $\gamma(\alpha A + (1-\alpha)A') \geqslant \gamma A$. In our case $A = \Pr_L(V \cap L_t)$, and Zalgaller's theorem gives

$$\gamma_{n-1}\left(\frac{\lambda+1}{2\lambda}\Pr_L(V \cap L_t) + \frac{\lambda-1}{2\lambda}\Pr_L(V \cap L_{-t})\right) \geqslant \gamma_{n-1}\Pr_L(V \cap L_t).$$

By (25) and (24), we can write

$$\gamma_{n-1}\Pr_L(A_\lambda V \cap L_t) \geqslant \gamma_{n-1}\Pr_L(V \cap L_t), \tag{26}$$

but

$$\gamma_n V = \int_{-\infty}^{\infty} \gamma_{n-1}\Pr_L(V \cap L_t)\,\gamma_1(dt), \quad \gamma_n(A_\lambda V) = \int_{-\infty}^{\infty} \gamma_{n-1}\Pr_L(A_\lambda V \cap L_t)\,\gamma_1(dt)$$

and (26), which holds for each t, leads to the desired conclusion. ●

PROPOSITION 30. *Let $K \subset H \subset L^2(E, \gamma)$, $K \in GB$, be a convex balanced subset of the Gaussian space H, $\{L_j, j = 0, 1, \dots\}$ a sequence of subspaces such that $H = L_0 \supset L_1 \supset \cdots$, $\bigcap L_j = \{0\}$, $\dim L_j/L_{j+1} < \infty$, $j = 0, 1, \dots$, and $\Pr_j = \Pr_{L_j}$. Then:*

1) $\sup \Pr_j K \longrightarrow \delta(K)$ *in* L^2;
2) $\sup (K \cap L_j) \longrightarrow \delta(K)$ *in* L^2;
3) *if* $\lambda > \delta(K)$, *then* $\gamma(\lambda(\Pr_j K)^\circ) \nearrow 1$, *and if* $\lambda < \delta(K)$, *then* $\gamma(\lambda(\Pr_j K)^\circ) = 0$;
4) *if* $\lambda > \delta(K)$, *then* $\gamma(\lambda(K \cap L_j)^\circ) \nearrow 1$, *and if* $\lambda < \delta(K)$, *then* $\gamma(\lambda(K \cap L_j)^\circ) = 0$;
5) $h_1(\Pr_j K) \searrow (2\pi)^{1/2}\delta(K)$;
6) $h_1(K \cap L_j) \searrow (2\pi)^{1/2}\delta(K)$.

PROOF. Each of the Gaussian spaces L_j can be regarded as the space of (Gaussian) linear functionals on (E, γ) that are measurable with respect to the measurable decomposition η_j of (E, γ) into the subspaces parallel to the finite-dimensional subspace $L_j^\circ \subset H^* \subset E$. Let $\widetilde{L}_j \subset L^2(E, \gamma)$ denote the subspace consisting of all (and not only linear) functions that are measurable with respect to η_j, so that $L_j = H \cap \widetilde{L}_j$. $\Pr_j$ is the operator of passage to the conditional (in the narrow sense) mathematical expectation.

We first prove the assertions 2), 4), and 6). The sequence of sets $K \cap L_j$ decreases in the set-theoretical sense; therefore the sequence of nonnegative measurable functions $\sup (K \cap L_j)$ decreases monotonically; consequently it converges to a limit that is measurable with respect to each of the decompositions η_j, i.e. to a constant. By Proposition 28, this constant coincides with $\delta(K)$. The assertion 6) now follows from Proposition 14, and 4) follows from Proposition 28.

We now prove 1). If $\Pr$ is the operator of conditional (in the narrow sense) mathematical expectation, then $\sup \Pr K \leqslant \Pr \sup K$ (a generalization of the obvious inequality $\sup_k (x_k + y_k) \leqslant \sup_k x_k + \sup_k y_k$). For each $k \geqslant 0$ the sequence $\Pr_j \sup \Pr_k K$ ($j = k, k+1, \dots$) forms a martingale, and, by the well-known theorem of Doob [24], it converges in the mean square to some constant $c_k \geqslant 0$; moreover, it is easy to see that $c_0 \geqslant c_1 \geqslant \cdots$. Indeed, $c_k = E \Pr_j \sup \Pr_k K$ for any $j = k, k+1, \dots$; in particular, $c_k = E \sup \Pr_k K$, and, by the corollary to Theorem 2, since a projection operator does not increase distances, we find that $h_1(K) \geqslant h_1(\Pr_1 K) \geqslant h_1(\Pr_2 K) \geqslant \cdots$, i.e. (taking Proposition 14 into account) $c_0 \geqslant c_1 \geqslant c_2 \geqslant \cdots$; let $c = \lim c_k$. Now we can write ($j \geqslant k$)

$$E|\sup \Pr_j K - c| \leqslant E|\sup \Pr_j K - \Pr_j \sup \Pr_k K| + E|\Pr_j \sup \Pr_k K - c_k| + |c_k - c|,$$

and if k and j are sufficiently large, then the right-hand side is made arbitrarily small (by the monotonicity mentioned above,

$$E|\Pr_j \sup \Pr_k K - \sup \Pr_j K| = E(\sup \Pr_j K - \Pr \sup \Pr_k K) = c_k - c_j$$

and for fixed k, in view of the convergence $\Pr_j \sup \Pr_k K \longrightarrow c_k$, the second term is arbitrarily small if j is sufficiently large), whence the convergence of $\sup \Pr_j K$ to a constant c in the mean is proved. Actually, there is even mean-square convergence: by the monotonicity in k and by the well-known properties of martingales, the set of second moments of the variables $\Pr_j \sup \Pr_k K$ is bounded, so convergence in the mean implies convergence on a dense set, and (considering the boundedness of the norms) we have Hilbert weak convergence; moreover, by a property of martingales,

$$\|\Pr_j \sup \Pr_k K\|_{L^2} \underset{j}{\rightarrow} c_k,$$

and by monotonicity,

$$\|\Pr_j \sup \Pr_j K\|_{L^2} \equiv \|\sup \Pr_j K\|_{L^2} \rightarrow c;$$

weak convergence and convergence of the second moments imply convergence in the mean square (see, for example, [49], Chapter VIII, §1.3, Theorem 4). From Propositions 28 and 29 it follows that $c = \delta(K)$. To prove 3) it now suffices to use Proposition 29 (the limiting case $\lambda = +\infty$) and the second assertion of Proposition 28. The assertion 5) is proved just as 6). ●

COROLLARY. *If $\gamma K^\circ = 0$, then the constant $(2\pi)^{1/2}$ in the inequality $h_1(K) \geqslant (2\pi)^{1/2}$ is best possible (see Proposition 20, inequality 1)).*

Indeed, beginning with any set K for which $\gamma K^\circ = 0$ and $\delta(K) = 1$ (or $1 + \epsilon$), it suffices to consider the set $K \cap L_j$ for sufficiently large j.

We describe the qualitative picture by the following theorem, which has, in essence, already been proved.

THEOREM 3. *The validity of any of the following conditions is necessary and sufficient for a convex balanced subset K of the Hilbert space H to belong to the class GC:*

1) *There is a sequence $\{L_j, j = 0, 1, \ldots\}$ of subspaces of H having the properties $H = L_0 \supset L_1 \supset \cdots$, $\bigcap_j L_j = \{0\}$, and $\dim L_j/L_{j+1} < \infty$ such that $h_1(\Pr_j K) \longrightarrow 0$ ($\Pr_j$ is the projection onto L_j).*

2) *There is a sequence $\{L_j\}$ of subspaces having the indicated properties such that $h_1(K \cap L_j) \longrightarrow 0$.*

3) *For any sequence $\{L_j\}$ of subspaces having the indicated properties, $h_1(\Pr_j K) \rightarrow 0$.*

4) *For any sequence* $\{L_j\}$ *of subspaces having the indicated properties,* $h_1(K \cap L_j) \to 0$.

The theorem follows from 5) and 6) in Proposition 30. ●

Theorem 3 enables us to give convenient criteria for the continuity of the sample functions in terms of the ϵ-entropy of K.

§7. ϵ-entropy conditions

1. The criteria proved above for checking the conditions $K \in GB$ and $K \in GC$, formulated in terms of the functional h_1, and the properties established for this functional enable us to derive conditions for the continuity and boundedness of realizations of a Gaussian process directly in terms of the correlation function, i.e., the intrinsic metric of the set $K \subset H$: in terms of the ϵ-entropy of K in the metric induced from H.

PROPOSITION 31. *For a regular* $(n-1)$*-dimensional simplex* S_n *with edges of unit length the following inequality holds*:

$$h_1(S_n) \leqslant \sqrt{2\pi \ln n}.$$

PROOF. Let $e_1(\omega), \ldots$ be a sequence of independent standard Gaussian variables. As shown in [21] (p. 376), for $\hat{e}_n = \max\{e_1, \ldots, e_n\}$ we have the estimate

$$E\hat{e}_n = \sqrt{2 \ln n} - \frac{\ln \ln n + \ln 4\pi - 2\gamma}{2\sqrt{2 \ln n}} + O\left(\frac{1}{\ln n}\right) \quad (\gamma \text{ is Euler's constant}),$$

where, as follows from the proof, the remainder term is nonpositive. By Proposition 14, $h_1(S_n) = (2\pi)^{1/2} E(2^{-1/2}\hat{e}_n)$, from which the proof follows.(4) ●

2. DEFINITION (see, for example, [28] or [79]). The *entropy index*, or *entropy type*, of a precompact subset K of a metric space is defined to be the number

$$\rho(K) = \limsup_{\epsilon \to 0} (\log \log N(K, \epsilon))(\log(1/\epsilon))^{-1},$$

where $N(K, \epsilon)$ is the cardinality of a smallest ϵ-net of the set K.

PROPOSITION 32. *If* $\rho(K) < 2$, *then* $\Sigma\, 2^{-k} (\log_2 N(K, 2^{-k}))^{1/2} < \infty$.

PROOF. Suppose that $\rho(K) = p < 2$, and let $p < q < 2$. For sufficiently large k we have $k^{-1} \log_2 \log_2 N(K, 2^{-k}) < q < 2$, i.e., $\log_2 N(K, 2^{-k}) < 2^{qk}$, and

$$\Sigma 2^{-k}(\log_2 N(K, 2^{-k}))^{1/2} < \Sigma 2^{(q/2-1)k} < \infty. \quad ●$$

PROPOSITION 33. *For an arbitrary precompact subset* $K \subset H$ *the following inequality holds*:

$$h_1(K) \leqslant 22 \sum_{k=-\infty}^{\infty} 2^{-k} \sqrt{\log_2 N(K, 2^{-k})}.$$

PROOF. To prove this inequality we use the monotonicity of h_1 (a corollary of

(4) For an estimate of the values of h_1 on regular octahedra the author consulted Ibragimov [41]. With regard to the distribution of the largest order statistic see also [145] (§9.6(b)), where references are given.

Theorem 2) and construct a set $S \subset H$ for which there is a contractive mapping of a subset of it onto a dense subset of K, while $h_1(S)$ can be estimated by the right-hand side of the desired inequality. Let E_k be a minimal ϵ-net for K constructed for $\epsilon_k = 2^{-k}$. Let $N_{k_0} = N(K, \epsilon_{k_0}) = 1$, and $N_{k_0+1} > 1$. We can assume that $\cdots = E_{k_0-1} = E_{k_0}$. Let $E = \bigcup_{-\infty}^{\infty} E_k$. Let S_k be a regular simplex with length of edges equal to 1 and dimension N_k $(k = k_0 + 1, k_0 + 2, \ldots)$; moreover, the S_k are located in pairwise orthogonal subspaces of the Hilbert space H. We set $S = \Sigma_{k_0+1}^{\infty} 8\epsilon_k S_k$.

By the additivity property of h_1 we find (Proposition 31)

$$h_1(S) = 8\sum_{k=k_0+1}^{\infty} 2^{-k} h_1(S_k) \leqslant 8\sum_{k=k_0+1}^{\infty} 2^{-k}\sqrt{2\pi\ln(N_k+1)} \leqslant 8\sqrt{2\pi\ln 3}\sum_{k=-\infty}^{\infty}\frac{\sqrt{\log_2 N_k}}{2^k}.$$

It remains to describe a certain subset of S whose image under some contractive mapping is the set E. We describe an injective mapping of E into S whose inverse will be the desired contractive mapping. In the set E we introduce an order relation in the following way. Let $x \in E_k$ and $y \in E_{k+1}$. We say that y follows x if y lies in the ϵ_k-neighborhood of x. To each element $y \in E$, except for the element forming the ϵ_{k_0}-net, we can assign at least one predecessor. For each such element $y \in E$ we mark and fix one of its predecessors; in this way the elements of E are arranged in the form of a tree. The order relation corresponding to this fixed tree structure will be denoted by $\succ$. Thus, if $y \succ x$, then there exists a chain $x \prec z_1 \prec z_2 \prec \cdots \prec z_{p-1} \prec y$ for which $z_i \in E_{k+i}$, $x \in E_k$ and $y \in E_p$, and each member of this chain, except y, is the designated predecessor of the next member. Now if $x_i \in E_{k_i}$ $(i = 1, 2, 3)$, $x_3 \prec x_1$, and $x_3 \prec x_2$ $(k_3 < k_1, k_3 < k_2)$, then

$$\|x_3 - x_1\| \leqslant \sum_{k=k_3}^{k_1-1}\frac{1}{2^k} < \frac{1}{2^{k_3-1}}$$

and $\|x_3 - x_2\| < 2^{-k_3+1}$; consequently, $\|x_1 - x_2\| < 2^{-k_3+2}$.

Now let $a_k^{(0)}, a_k^{(1)}, \ldots, a_k^{(N_k)}$ be the vertices of the simplex $8\epsilon_k S_k$ (we can assume that $a_k^{(0)} = 0$ for all k), and let $x_k^{(1)}, \ldots, x_k^{(N_k)} \in E_k$ be the points of the ϵ_k-net E_k. We now assign to the point $x_{k_0}^{(1)} \in E_{k_0}$ the point

$$b_0 = \sum_{k=k_0+1}^{\infty} a_k^{(0)} \in S,$$

to the points $x_{k_0+1}^{(l)}$ the respective points

$$b_1^{(l)} = a_{k_0+1}^{(l)} + \sum_{k=2} a_k^{(0)} \in S,$$

and, in general, if the point $x_s^{(i_s)} \in E_s$, $s > k_0$, is such that $x_s^{(i_s)} \succ x_{s-1}^{(i_{s-1})} \succ \cdots \succ x_{k_0}^{(1)}$, then we place it into correspondence with the point

$$b_s^{(i_s)} = a_{k_0+1}^{(i_{k_0+1})} + \ldots + a_{s-1}^{(i_{s-1})} + a_s^{(i_s)} + \sum_{k=s+1}^{\infty} a_k^{(0)}.$$

Moreover, it is easily seen that if for the points $x_{k_1}^{(i_{k_1})} \in E_{k_1}$ and $x_{k_2}^{(i_{k_2})} \in E_{k_2}$ the infimum in the sense of the order on the tree is attained at $x_{k_3}^{(i_{k_3})} \in E_{k_3}$, i.e. if

$$b_{k_1}^{(i_{k_1})} = a_{k_0+1}^{(i_{k_0+1})} + \dots + a_{k_3}^{(i_{k_3})} + a_{k_3+1}^{(i'_{k_3+1})} + \dots + \sum_{k=k_1+1}^{\infty} a_k^{(0)},$$

$$b_{k_2}^{(i_{k_2})} = a_{k_0+1}^{(i_{k_0+1})} + \dots + a_{k_3}^{(i_{k_3})} + a_{k_3+1}^{(i''_{k_3+1})} + \dots + \sum_{k=k_2+1}^{\infty} a_k^{(0)},$$

where $a_{k_3+1}^{(i'_{k_3+1})} \neq a_{k_3+1}^{(i''_{k_3+1})}$, then

$$\left\| b_{k_1}^{(i_{k_1})} - b_{k_2}^{(i_{k_2})} \right\| \geqslant 8 \frac{1}{2^{k_3+1}} = \frac{1}{2^{k_3-2}} > \left\| x_{k_1}^{(i_{k_1})} - x_{k_2}^{(i_{k_2})} \right\|,$$

which proves the contractive character of the inverse of the mapping just described. ●

COROLLARY 1. *If $\rho(K) < 2$, then $K \in GB$, and γK° (and κK°) can be estimated from below by means of the inequalities in Propositions 32, 33, and 20.*

COROLLARY 2 (Dudley's theorem [33]). *If $\sum 2^{-k} (\log_2 N(K, 2^{-k}))^{1/2} < \infty$, then $K \in GC$; in particular, $\rho(K) < 2 \Rightarrow K \in GC$.*

Indeed, we can assume that the set K is convex and balanced. If the sequence $\{L_j\}$ of spaces satisfies the conditions of Proposition 30, then for any $\epsilon > 0$ only a finite number of members of the sequence $N(K \cap L_j, \epsilon)$, $j = 0, 1, \dots,$ are different from 1, from which it follows that the right-hand side of the inequality in Proposition 33, with $K \cap L_j$ substituted for K, converges to zero, and, by 6) in Proposition 30, it follows from this that $\delta(K) = 0$, i.e., $K \in GC$ (Proposition 27). ●

3. Let $M = M(K, \epsilon)$ be the cardinality of a largest subset $F \subset K$ consisting of elements such that the pairwise distances between them are greater than ϵ (an ϵ-lattice).

PROPOSITION 34. *If $M = M(K, \epsilon) \geqslant 10$, then for any $\epsilon > 0$*

$$h_1(K) \geqslant \left(1 - \frac{1}{e}\right) \sqrt{\frac{\pi}{2}}\, \varepsilon \sqrt{\ln M}.$$

PROOF. From the inequality 1) in Proposition 20 it follows that $h_1(K) \geqslant \sigma^{-1}(2\pi)^{1/2}(1 - \gamma(\sigma K)^\circ)$. If B_n is the simplex whose vertices are the vectors $2^{-1/2}e_1, \dots, 2^{-1/2}e_n$, where $e_1, \dots, e_n$ are unit basis vectors, then, by Theorem 2, we find that

$$h_1(K) \geqslant h_1(\varepsilon B_{M(K,\varepsilon)}).$$

Consequently (see, for example, [27], Chapter 5),

$$h_1(K) \geqslant \frac{\sqrt{2\pi}}{\sigma}(1 - \gamma(\sigma\varepsilon B_M)^\circ) = \frac{\sqrt{2\pi}}{\sigma}\left(1 - \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{2}/\varepsilon\sigma} \exp\left(-\frac{1}{2}u^2\right) du\right]^M\right)$$

$$= \frac{\sqrt{2\pi}}{\sigma}\left(1 - \left[1 - \frac{\sigma\varepsilon}{2\sqrt{\pi}} \exp\left(-\frac{1}{\sigma^2\varepsilon^2}\right)\left(1 - \alpha\left(\frac{\sqrt{2}}{\sigma\varepsilon}\right)\right)\right]^M\right), \quad \text{where } 0 < \alpha(z) < \frac{1}{z^2}.$$

For $\sigma\epsilon < 1$ we continue the chain of inequalities:

$$\geqslant \frac{\sqrt{2\pi}}{\sigma}\left(1 - \exp\left\{M\left[-\frac{\sigma\varepsilon}{4\sqrt{\pi}}\exp\left(-\frac{1}{\sigma^2\varepsilon^2}\right)\right]\right\}\right)$$

$$= \frac{\sqrt{2\pi}}{\sigma}\left(1 - \exp\left\{-\exp\left[-\frac{1}{\sigma^2\varepsilon^2} + \ln M + \ln\frac{\sigma\varepsilon}{4\sqrt{\pi}}\right]\right\}\right).$$

Since the function $z^2 \ln (zM/4\pi^{1/2})$ is monotonically increasing with respect to z, and since, as is easily checked, $z^2 \ln (zM/4\pi^{1/2}) > 1$ for $M \geqslant 10$ and $z = 2(\ln M)^{-1/2}$, we find that if σ and ϵ satisfy the condition $-(\sigma\epsilon)^2 + \ln M + \ln (\sigma\epsilon/4\pi^{1/2}) = 0$, then $\sigma\epsilon < 2(\ln M)^{-1/2} < 1$, and $\sigma < 2/\epsilon(\ln M)^{1/2}$, from which the desired inequality follows immediately. ●

COROLLARY. *If* $\limsup \epsilon^2 \ln M(K, \epsilon) = \infty$ *(in particular, if* $\rho(K) > 2$*), then* $h_1(K) = \infty$*, and* $K \notin GB$.

Indeed, it is well known [68] that $M(K, \epsilon) \geqslant N(K, \epsilon)$; hence, if $\rho(K) = p > 2$, then for some sequence $\epsilon_k \searrow 0$ we find that $\ln M(K, \epsilon_k) > q \ln (1/\epsilon_k)$ for some $q > 2$, i.e., $\epsilon_k^q \ln M(K, \epsilon_k) > 1$, and $\epsilon_k(\ln M(K, \epsilon_k))^{1/2} \longrightarrow \infty$; with regard to the condition $K \notin GB$ see Theorem 1. ●

If $\limsup_{\epsilon\to 0} \epsilon^2 \ln M(K, \epsilon) < \infty$, then the condition $K \in GB$ may or may not hold.

For example, let $a_k = (k \ln k)^{-1}$, $k = 2, \ldots$, and let π be the set of vertices of a rectangular parallelepiped whose edges have lengths a_k. By the additivity of the functional h_1 we have $h_1(\pi) = \Sigma\, a_k = \infty$, i.e., $\pi \notin GB$. We prove that

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log M(\pi, \varepsilon) = 0.$$

It is known [68] that $M(K, \epsilon) \leqslant N(K, \epsilon/2)$, and so it suffices to prove that $\epsilon^2 \log N(\pi, \epsilon) \to 0$. If for fixed $\epsilon > 0$ we choose a number $m = m(\epsilon)$ such that the subset $\pi_m \subset \pi$ composed of the vertices of some rectangular parallelepiped with lengths of edges $a_2, \ldots, a_{m+1}$ forms an ϵ-net in the set π, while the set π_{m-1} does not, then we get that $N(\pi, \epsilon) \leqslant 2^{m(\epsilon)}$, and $\log_2 N \leqslant m(\epsilon)$. If π_m is an ϵ-net in π, then $\Sigma_{k=m+2}^{\infty} (k \ln k)^{-2} \leqslant \epsilon^2$, and, by assumption, $\Sigma_{k=m+1}^{\infty} (k \ln k)^{-2} > \epsilon^2$. Since

$$\int_{m+2}^{\infty} \frac{dt}{t^2 \ln^2 t} = \frac{1}{(m+2)\ln^2 (m+2)} - 2\int_{m+2}^{\infty} \frac{dt}{t^2 \ln^3 t},$$

it follows from this that

$$\frac{1}{m \ln^2 m} > \frac{1}{(m+2)\ln^2 (m+2)} > \int_{m+2}^{\infty} \frac{dt}{t^2 \ln^2 t} > \sum_{k=m+3}^{\infty} \frac{1}{k^2 \ln^2 k}$$

$$= \sum_{k=m+1}^{\infty} \frac{1}{k^2 \ln^2 k} - \frac{1}{(m+2)^2 \ln^2 (m+2)} - \frac{1}{(m+3)^2 \ln^2 (m+3)} > \varepsilon^2 - \frac{2}{m^2 \ln^2 m}$$

and hence

$$\varepsilon^2 \log_2 N \leqslant \varepsilon^2 m(\varepsilon) < \frac{1}{\ln^2 m(\varepsilon)} + \frac{2}{m \ln^2 m} \xrightarrow[\varepsilon\to 0]{} 0.$$

A more careful calculation would show that $\int_0^1 \epsilon^2 d(\ln N(\pi, \epsilon)) > -\infty$. It is curious to note that with the help of the methods of estimating ϵ-entropy developed by Mitjagin in [79], it can be shown, using the Minlos–Sazonov theorem (or our estimates) that in the class of ellipsoids E (sets of the form $\{x: \Sigma (x, e_k)^2/\lambda_k^2 \leqslant 1\} \subset H$, where e_k is an orthonormal basis) the condition $\int_0^1 \epsilon^2 d(\ln N(E, \epsilon)) > -\infty$ is necessary and sufficient for the condition $E \in GB$. An example in which $K \in GB$ and

$$\limsup \epsilon^2 \ln M(K, \epsilon) > 0$$

was given by Dudley in [26]. In the same place Dudley remarked that in all known examples of GC-sets K the condition $\lim_{\epsilon \to 0} \epsilon^2 \log N(K, \epsilon) = 0$ holds. Our methods permit us to prove a more precise statement.

PROPOSITION 35. *The following inequality holds*:

$$\delta(K) \geqslant \frac{1}{2}\left(1 - \frac{1}{e}\right) \limsup_{\varepsilon \to 0} \varepsilon \sqrt{\ln M(K, \varepsilon)}.$$

PROOF. For any linear subspace $L \subset H$ of finite defect $d = \dim(H/L)$ the orthogonal projection $\mathrm{Pr}_L\, K$ of a precompact set K onto this subspace has the same order of growth of ϵ-entropy:

$$\limsup \varepsilon^2 \ln M(\mathrm{Pr}_L K, \varepsilon) = \limsup \epsilon^2 \ln M(K, \epsilon).$$

It suffices for us to observe that, since K is contained in the orthogonal sum of the set $\mathrm{Pr}_L\, K$ and some d-dimensional bounded set, we have the estimate $M(\mathrm{Pr}_L\, K, \epsilon) \geqslant M(K, \epsilon) c(K)/\epsilon^d$, where $c(K)$ is a constant. From this it follows that

$$\limsup \varepsilon \sqrt{\ln M(K, \varepsilon)} \leqslant \limsup \varepsilon \sqrt{\ln M(\mathrm{Pr}_L K, \varepsilon)}.$$

From Proposition 34 we get, as $\epsilon \to 0$,

$$h_1(K) \geqslant \limsup \left(1 - \frac{1}{e}\right) \sqrt{\frac{\pi}{2}}\, \varepsilon \sqrt{\ln M(K, \varepsilon)}$$

and consequently, for any subspace L, $\dim(H/L) < \infty$,

$$h_1(\mathrm{Pr}_L K) \geqslant \limsup \left(1 - \frac{1}{e}\right) \sqrt{\frac{\pi}{2}}\, \varepsilon \sqrt{\ln M(K, \varepsilon)}.$$

Now if $\{L_j\}$ is a sequence of such subspaces that decreases to zero, then by the assertion 5) in Proposition 30 we get, finally,

$$\delta(K) = \frac{1}{\sqrt{2\pi}} \lim_j h_1(\mathrm{Pr}_{L_j} K) \geqslant \frac{1}{2}\left(1 - \frac{1}{e}\right) \limsup \varepsilon \sqrt{\ln M(K, \varepsilon)}.$$

As was mentioned earlier, in the case of stationary Gaussian processes the ϵ-entropy language permits us to give necessary and sufficient conditions for the continuity of sample functions.

4. Unfortunately, we cannot hope to get a necessary and sufficient condition for the condition $K \in GB$ or $K \in GC$ in terms of ϵ-entropy. It was shown in [121] that for any increasing function $H(\epsilon)$ there is an ellipsoid E for which $H(a\epsilon) \leqslant \log N(E, \epsilon) \leqslant H(b\epsilon)$. Now let $H(\epsilon) = \log N(\pi, \epsilon)$, where π is the rectangular parallelepiped with lengths of edges $a_k = (k \ln k)^{-1}$ described above. As noted, $\int_0^1 \epsilon^2 dH(\epsilon) > -\infty$, from

which it follows that an ellipsoid E with the same growth of entropy belongs to the class GB, while $h_1(\pi) = \Sigma a_k = +\infty$, and $\pi \notin GB$. It can also be shown that it is not possible to characterize the class GB in terms of the sequence of finite-dimensional diameters. Regarding the GC property, it follows that, as will be clear from the next section, each ellipsoid having the GB property also has the GC property (just as any convex set that is the unit ball of a reflexive space); therefore, the example given with the parallelepiped π and the ellipsoid E, both having the same growth of entropy, shows that the class of GC-sets cannot be exactly characterized by entropy criteria.

5. The basic results of this section can be summarized in the following form (Pr_L is the orthogonal projection onto the subspace $L \subset H$).

THEOREM 4. *For the convex balanced precompact subset K of the Hilbert space H to belong to the class GB it is necessary that*

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log_2 N(K, \varepsilon) < \infty.$$

Moreover, if $M(\mathrm{conv}\, K, \epsilon) \geqslant 10$, then

$$h_1(K) \geqslant 0.65\varepsilon \sqrt{\ln M(K, \varepsilon)}, \quad \delta(K) \geqslant 0.31 \limsup_{\varepsilon\to 0} \varepsilon \sqrt{\ln M(K, \varepsilon)}$$

and, in particular, if $K \in GC$, then

$$\limsup \varepsilon^2 \log_2 N(K, \varepsilon) = 0.$$

Moreover,

$$h_1(K) < 22 \sum_{k=-\infty}^{\infty} 2^{-k} \sqrt{\log_2 N(K, 2^{-k})}$$

and, in particular, if $\Sigma_{-\infty}^{\infty} 2^{-k}(\log_2 N(K, 2^{-k}))^{1/2} < \infty$, then $K \in GC$.

§8. The non-Gaussian case

1. We now try to determine what the geometric characteristics can yield in investigating the questions of boundedness and continuity of realizations of arbitrary (not necessarily Gaussian) random processes. Let $\widetilde{K} \subset L^2(\Omega, \mu)$ be an arbitrary subset of the space of square-integrable functions, which, as before, we regard as a random process. When we are interested in the question of boundedness (with probability 1) of the sample functions $x(\omega)$, $x \in \widetilde{K}$, of this process, none of the structures on $\widetilde{K}$ connected with some property of the parametrization are essential; but if we are concerned with the continuity of the realization $x(\omega)$ of the process $\widetilde{K}$, then we always assume that $\widetilde{K}$ has the topology induced by the imbedding $\widetilde{K} \subset L^2$ (we shall be interested mainly in precompact sets $\widetilde{K}$, so we need not specify which topology on the Hilbert space H is used). It is not hard to see that the specification only of the geometry of $\widetilde{K}$ does not now provide, in all cases, an answer about boundedness or continuity of a realization. For example, from a geometric point of view the standard Wiener and Poisson processes, regarded on the same intervals of time, are equivalent. It is therefore advisable to try to find geometric characteristics of the process $\widetilde{K}$ that would guarantee its boundedness or continuity independently of the location of $\widetilde{K}$ in L^2, i.e., independently of the more subtle properties that are detectable only with the help of the

finite-dimensional distributions, the higher moments, the properties connected with the ring structure in L^2 or with the structure of the partially ordered set, etc. In other words, we want to isolate the whole class of those subsets of a Hilbert space that represent random processes with bounded (or continuous) realizations for any concrete realization of the Hilbert space with distinguished subset in the form of a Hilbert space $L^2(\Omega, \mu)$ of square-integrable functions with respect to a probability measure μ.

2. We first study the property of boundedness of realizations. To say that the process $\tilde{K} \subset L^2(\Omega, \mu)$ has bounded realizations with probability 1 is to say that the measurable function $\sup \tilde{K}$ with values in $\{\mathbf{R}, +\infty\}$ is almost everywhere finite (it is more precise to speak of the class of equivalence mod 0 of such functions), i.e., that the subset $\tilde{K} \subset L^2(\Omega, \mu) \subset S(\Omega, \mu)$ is structurally bounded in $S(\Omega, \mu)$.

We consider the following characteristic of a compact balanced subset K of the Hilbert space H.

DEFINITION (see [116]). Let $K \subset H$ be an arbitrary subset of a Hilbert space. Then the *s-characteristic* of K is defined to be the number

$$s(K) = \sup \{ \| \sup \tilde{K} \|^2_{L^2(\Omega, \mu)} \},$$

where $\tilde{K}$ is the image of K under a (linear) isometry of H onto some space $L^2(\Omega, \mu)$, the inside sup is taken in the sense of the natural partial order structure in $L^2(\Omega, \mu)$, and the outside sup runs over all possible such isometries. Sets K for which $s(K) < \infty$ are called *Schmidt sets*, or *s-sets*.

We remark that the value of $s(K)$ would not change if the outside sup were taken only over the discrete measures, i.e., with respect to all possible isometries of H onto l^2. As is shown below, the class of Schmidt sets coincides with the class of subsets of Schmidt ellipsoids; for ellipsoids E the value of $\|\sup \tilde{E}\|_{L^2(\Omega, \mu)}$, where $\tilde{E}$ is the image of E under an isometry of H onto $L^2(\Omega, \mu)$, does not depend on the choice of this isometry.

THEOREM 5. *Let $K \subset H$. If $s(K) < \infty$, then K is the image of some bounded subset of the unit ball of H under a Hilbert-Schmidt transformation whose square has trace equal to $s(K)$; moreover, the sample functions $x(\omega)$, $x \in \tilde{K}$, of any random process $\tilde{K} \subset L^2(\Omega, \mu)$ that is isometric to the set K are bounded and continuous (with respect to $x \in \tilde{K}$ in the relative topology). But if $s(K) = \infty$, then for some subset $\tilde{K} \subset L^2(\Omega, \mu)$ isometric to the set $K \subset H$ the sample functions of the random process $\tilde{K}$ are unbounded with probability 1, i.e., $\sup \tilde{K} \equiv +\infty$ (mod 0). The condition $s(K) < \infty$ follows from the finiteness of $\|\sup \tilde{K}\|_{L^2(\Omega, \mu)}$ for any $\tilde{K} \subset L^2$ that is isometric to $K \subset H$.*

We precede the proof by a number of auxiliary statements. From the following arguments it will be clear that we can limit ourselves to the consideration of convex symmetric (with respect to zero) compact sets $K \subset H$. The half-lengths of K with respect to the basis $\{e_k\}$ are defined to be the numbers $g_k(K) = \sup_{x \in K} |(x, e_k)|$. We prove that for each s-set K in an infinite-dimensional separable Hilbert space there is an ellipsoid containing it with the same s-characteristic. It is curious that in finite-dimensional spaces the last statement ceases to be true (although the fact of the existence of an s-ellipsoid containing K is then trivial).

The proof given below is close to the arguments in [116]. Other proofs of the statement characterizing s-sets as subsets of s-ellipsoids are now known (see [111] or [73]).

PROPOSITION 36. *Let $\mathcal{E} \subset H$ be an ellipsoid, $\{b_n\}$ the lengths of its semiaxes, and $\{e_k\}$ an arbitrary orthonormal basis; then*

$$s(\mathcal{E}) = \sum g_k^2(\mathcal{E}) = \sum b_n^2.$$

PROOF. It suffices to prove the second equality. Suppose that in the system $\{e_n^{(0)}\}$ the ellipsoid $\mathcal{E}$ is given by the inequality

$$\sum \frac{(x, e_n^{(0)})^2}{b_n^2} \leqslant 1, \quad x = \sum \xi_n e_n^{(0)}, \quad e_k = \sum_n a_{kn} e_n^{(0)}.$$

Obviously,

$$g_k^2(\mathcal{E}) = \sup (x, e_k)^2 = \sup (\sum \xi_n a_{kn})^2 = \sum b_n^2 a_{kn}^2, \quad \sum\sum b_n^2 a_{kn}^2 = \sum b_n^2. \quad \bullet$$

PROPOSITION 37. *Given the n points $x_1, \dots, x_n$ in the Euclidean space $\mathbf{R}^n$, let $K = \Gamma(x_1, \dots, x_n)$ be their convex balanced hull. There exists a unique ellipsoid $\mathcal{E}$ containing K and such that $s(\mathcal{E}) = s(K)$.*

PROOF. We first suppose that the vectors $x_1, \dots, x_n$ are linearly independent. The proof is carried out by induction on n. For $n = 2$ we locate K in a rectangle P so that the points x_1 and x_2 lie on adjacent sides of P, and the diagonals of P are parallel to $x_1 + x_2$ and $x_1 - x_2$. It is easy to see that there is an ellipse contained in P and passing through x_1 and x_2. Let e_1 and e_2 be unit vectors orthogonal to the sides of P. $\mathcal{E} \supset K$; consequently $s(\mathcal{E}) \geqslant s(K)$; but $s(\mathcal{E}) = g_1^2(K) + g_2^2(K) \leqslant s(K)$, i.e., $s(\mathcal{E}) = s(K)$. Since there is only one ellipse in P that passes through x_1 and x_2, every other ellipse has half-lengths in the system of coordinates e_1, e_2 that are not less than g_1 and g_2, respectively.

Now suppose that the proposition has been proved for $n - 1$. We consider $K = \Gamma(x_1, \dots, x_n) \subset \mathbf{R}^n$ (the vectors $x_1, \dots, x_n$ are assumed to be linearly independent). We consider an orthogonal basis $e_1, \dots, e_n$ in which $s(K)$ is attained and let P be the rectangular parallelepiped whose $(n-1)$-dimensional boundaries are orthogonal to the vectors e_k and are supporting hyperplanes for K (we call P the extremal parallelepiped for K). We remark that any boundary of P is the extremal parallelepiped for the projection onto it of K. In particular, the $(n-1)$-dimensional boundary G_n orthogonal to e_n is the extremal $(n-1)$-dimensional parallelepiped for the projections onto it of the points $x_1, \dots, x_{n-1}$. From this it follows that there exists an $(n-1)$-dimensional ellipsoid $\mathcal{E}_0$ lying in the hyperplane passing through the points $0, x_1, \dots, x_{n-1}$, containing these points, and contained in the cylinder with generatrix parallel to e_n and directrix G_n. There obviously exists a (one-parameter) family of balanced ellipsoids $\{\mathcal{E}_t\}$ contained in our cylinder and containing the points $x_1, \dots, x_{n-1}$ ($\mathcal{E}_0$ is a degenerate ellipsoid contained in our family). Let us now assume that the boundary G_n does not intersect $\mathcal{E}_0$. In this case there is an ellipsoid $\mathcal{E}_{t_0}$ for which G_n is a supporting hyperplane. Let x_n' be a point of tangency of G_n and $\mathcal{E}_{t_0}$. Considering the projections of

P and E_{t_0} onto the plane $L(e_k, e_n)$, $k = 1, \ldots, n-1$, and using the theorem for $n = 2$ and the remark about the extremality of the projections of P, we see that the coordinates of x'_n coincide with those of x_n.

It remains to show that the boundary G_n never intersects E_0. We translate G_n parallel to itself, and at each of its positions we mark a point x' on it from the following condition: for each $k < n$ the projection of the whole parallelepiped cut out from our cylinder by the hyperplanes of G_n and $-G_n$ onto the plane $L(e_k, e_n)$ is an extremal rectangle for the parallelogram with vertices at the projections of x_k and x'. The set T thus obtained contains all possible points of tangency of the ellipsoids in $\{E_t\}$ and the hyperplanes of G_n; however, in addition to them it can contain also other points, for example, lying on the other side of the hyperplane containing E_0 from the point e_n, but for which $(x', e_n) > 0$. We note first of all that $x_n \in T$. Next, considering that $(x_n, e_n) > 0$, we show that x_n and e_n lie on the same side of E_0. Indeed, if this is not so, then we consider $K' = \Gamma(x_1, \ldots, x_{n-1}, x''_n)$, where x''_n is symmetric to x_n with respect to the hyperplane of E_0. Since $s(K) = s(K')$, and x''_n is not contained in P (since $(x''_n, e_n) > (x_n, e_n)$), P cannot be the extremal parallelepiped for K. A direct calculation shows that the projection of T onto each coordinate plane $L(e_k, e_n)$ is a hyperbola $x_n = C_k/x_k$, from which it is clear that the curve T itself is a branch of a hyperbola. Each point $x \in T$ lying on the same side of E_0 as e_n lies on some hyperplane of G_n disjoint from the ellipsoid E_0 ($(x, e_n) > 0$). This concludes the proof for the case of points lying in general position. For linearly dependent vectors the proposition is now proved by passage to the limit.

The uniqueness of a minimal ellipsoid is proved just as in the case $n = 2$. •

REMARK. *Suppose that the linear span of the compact set $K = \Gamma(x_1, \ldots, x_n)$ has dimension m and that on the $(m-1)$-dimensional boundary of the minimal ellipsoid E there are k of the points (of course, $k \geqslant m$). Then $s(K) = s(E)$ already in k-dimensional space.* Indeed, we can regard the vectors $x_1, \ldots, x_n$ as limiting cases of linearly independent vectors; here the limiting extremal parallelepiped is not more than k-dimensional, since, if x_s lies inside the m-dimensional ellipsoid E, then the corresponding boundary of P, and hence also the opposite boundary, contains E, i.e., they merge. •

PROPOSITION 38. *Let the convex balanced compact set K have dimension n and lie in $\mathbf{R}^k$, $k \geqslant n(n+1)/2$. Then there is a unique ellipsoid $E \supset K$ such that $s(E) = s(K)$.*

It suffices to prove Proposition 38 for arbitrary convex polyhedra.

LEMMA 1. *Consider the compact (in its natural topology) space E of balanced ellipsoids in $\mathbf{R}^n$ that contain some compact set $M = \Gamma(x_1, \ldots, x_l)$ and are contained in a given (sufficiently large) ball. The functional s has a unique minimum on E.*

PROOF. Suppose that s takes a (local) minimum on some ellipsoid E. By moving a sufficient number of points $x_{l+1}, \ldots, x_N$ into E, it is possible to make E the minimal (smallest in the sense of the value of the functional s) ellipsoid for $K' = \Gamma(x_1, \ldots, x_N)$. But not more than l of the points lie on the boundary of E, so, by the remark at the end of the proof of Proposition 37, $s(E) = s(K') = s(K)$ in l-dimensional space, i.e., by the uniqueness proved, the ellipsoid E is minimal also for K. ●

LEMMA 2. *Suppose that the linear span of the compact set $M = \Gamma(x_1, \ldots, x_n)$ has dimension m, and E is the minimal ellipsoid for K. There are not more than $m(m + 1)/2$ points $x_{i_1}, \ldots, x_{i_N}$ such that E is already the minimal ellipsoid for $\Gamma(x_{i_1}, \ldots, x_{i_N})$.*

PROOF. By Lemma 1, E is the minimal ellipsoid of the set of points lying on its surface. Suppose that the number of such points is greater than $m(m + 1)/2$. From them we choose $m(m + 1)/2$ points for which the functional s takes a largest value on the minimal ellipsoid E' for these points. Suppose that E' is different from E. Since an m-dimensional balanced ellipsoid can be specified by means of $m(m + 1)/2$ points of its surface, in this sense there are "dependent" points among our points, i.e., points such that each ellipsoid passing through the remaining points passes also through them. We remark that outside the ellipsoid E' there is at least one of the points lying on the surface of E. We discard from the number of our $m(m + 1)/2$ points a "dependent" one and add a point x_s lying outside. The minimal ellipsoid E'' constructed for these new $m(m + 1)/2$ points is such that $s(E'') > s(E')$, and this contradicts the choice of E'. ●

The proof of Proposition 38 for an arbitrary convex polyhedron follows now from Lemma 2 and Proposition 37. ●

PROPOSITION 39. *Let $K \subset l^2$, $\|\sup K\| = 1$, and let $\Pr$ be a projection operator. Then $\|\sup \Pr(K)\| \leqslant 1$.*

For finite-dimensional coordinate spaces l^2_n the assertion can be verified directly. In the case of infinite-dimensional l^2 one should use the fact that $\bigcup_n l^2_n$ is dense in l^2. We mention that the analogous assertion for arbitrary selfadjoint operators with norm not exceeding 1 can be proved in the same way. ●

PROPOSITION 40. *Let H be infinite-dimensional, $K \subset H$ and $s(K) < \infty$. For any $\epsilon > 0$ there is an ellipsoid E such that $K \subset E$ and $s(E) \leqslant s(K) + \epsilon$.*

PROOF. If $H = H_1 \oplus H_2$, then $s(K) \leqslant s_{H_1}(\Pr_{H_1}(K)) + s_{H_2}(\Pr_{H_2}(K))$ (the notation is clear). Let H be decomposed into an infinite orthogonal sum of finite-dimensional subspaces $H = H_1 \oplus H_2 \oplus \cdots$, in each of which a balanced ellipsoid E_k is given, and let E be the sum of the ellipsoids E_k. Let the positive numbers $\lambda_1, \lambda_2, \ldots$ be such that $\Sigma\, \lambda_k^2 = 1$. Then E is contained in an ellipsoid E such that $s(E) = \Sigma\, s(E_k)\lambda_k^2$. In fact, the point $x = \{\xi_k\}$ belongs to E if

$$\sum_{k=1}^{n_1} \frac{\xi_k^2}{b_k^2} \leqslant 1, \quad \sum_{k=n_1+1}^{n_2} \frac{\xi_k^2}{b_k^2} \leqslant 1, \quad \ldots,$$

but in this case the ellipsoid E,

$$\sum_{k=1}^{\infty} \frac{\xi_k^2}{\lambda_k'^2 b_k^2} \leqslant 1, \quad \lambda_k' = \lambda_s, \qquad \text{if} \quad n_s \geqslant k > n_{s-1},$$

contains x, and hence all of E.

Now we approach the proof of the proposition in the following way. For simplicity we take $s(K) = 1$. We choose an arbitrary sequence $1, a_2, a_3, \ldots \searrow 0$. Select a finite-dimensional subspace H_1, $H = H_1 \oplus H^1$, such that $s(\Pr_{H_1}(K)) > 1 - a_2$; then $s(\Pr_{H^1}(K)) \leqslant a_2$. It is clear that $\Pr_{H_1}(K)$ can be moved into the finite-dimensional ellipsoid E_1, $s(E_1) \leqslant 1$, since $\Pr_{H_1}(K)$ can be regarded as $\Pr_{H_1}(\Pr_{H'_1}(K))$, where the dimension of H'_1 is $n(n+1)/2$ (n is the dimension of H_1); but

$$1 \geqslant s_{n(n+1)/2}(\Pr_{H'_1}(K)) \geqslant s_{n(n+1)/2}(\Pr_{H'_1}(K)).$$

Continuing this construction, we arrive at a decomposition of H into an infinite orthogonal sum $H = H_0 \oplus H_1 \oplus \cdots$, where the projection of K onto H_0 is zero, and the projections onto $H_1, H_2, \ldots$ can be included in ellipsoids $E_1, E_2, \ldots$ such that $s(E) \leqslant 1$, $s(E_2) \leqslant a_2$, etc. Since

$$\inf\left(\lambda_1^2 + \sum_{k=2}^{\infty} a_k \lambda_k^2\right) = 1$$

(the infimum runs over all sets of λ_k, $\Sigma\lambda_k^{-2} = 1$, and over all sets of a_k, $a_k \searrow 0$), the proposition is proved. ●

PROOF OF THEOREM 5. To prove that each s-set K is the image of some subset of the unit ball of the Hilbert space under a Hilbert-Schmidt transformation whose square has trace $s(K)$ it is sufficient to prove that K is contained in some ellipsoid E for which $s(E) = s(K)$.

We equip the space of ellipsoids in $\mathbf{R}^n$ with the natural topology. Fix a basis $\{e_k\}$. By means of a diagonal process we construct a sequence $\{E_k\}$ of ellipsoids in H containing K and such that their projections onto each finite-dimensional coordinate subspace converge to certain ellipsoids, and $s(E_k) \longrightarrow s(K)$. It is not hard to show that the projections onto any finite-dimensional subspace will then converge. The collection of limit ellipsoids of the projections forms a "spectrum", i.e., is such that if $L_1 \subset L_2$ are finite-dimensional spaces in H, then the limit ellipsoid in L_1 is the projection of the limit ellipsoid from L_2 onto L_1. This spectrum is generated by the projections of some closed convex balanced set E. Since all finite-dimensional projections of E are ellipsoids, E is also an ellipsoid. The inequality $s(E) \leqslant \lim s(E_k)$ and the inclusion $K \subset E$ are obvious.

REMARK. *The condition that H is infinite-dimensional is essential.* For example, for a regular hexagon Z in $\mathbf{R}^2$ there is no ellipse E containing Z for which $s(E) = s_{\mathbf{R}^2}(Z)$.

We now show that $s(K) < \infty$ if $\Sigma g_k^2(K) < \infty$ for any basis $\{e_k\}$, i.e., if

$$\|\sup \tilde{K}\|_{L^2(\Omega,\mu)} < \infty,$$

for all possible isometries of (H, K) onto $(L^2(\Omega, \mu), \tilde{K})$, at least for discrete measures μ. Let $s(K) = \infty$. There are othogonal unit vectors $e_1, \ldots, e_{n_1}$ such that $\Sigma_1^{n_1} g_k^2(K) \geqslant 1$. Moreover, by the preceding, $s(K) = \infty$ again holds for the projection of K onto the orthogonal complement H_1 of these vectors. In H_1 we make the same construction with $\Pr_{H_1}(K)$. Continuing the process, we construct a basis in which, obviously, $\Sigma g_k^2(K) = \infty$. The last assertion can be restated as follows: each subset of l^2 that, together with any of its unitary images, is order bounded is an s-set. Further, we note that the order in the sense of l^2 can be replaced everywhere by the order in the sense of L^2 (which is not isomorphic to the first). Indeed, the supremum of an s-ellipsoid with semiaxes $\{f_n(t)\}$ is $\Sigma f_n^2(t)$, and its norm is invariant with respect to unitary transformations. The role of the dense set in Proposition 39 is played by a certain set of step functions.

We now prove that the realizations $x(\omega)$ of the process $\widetilde{K} \subset L^2(\Omega, \mu)$ are continuous with respect to $x \in \widetilde{K}$ (in the relative topology on $\widetilde{K}$) with probability 1 if $\widetilde{K}$ is an s-set. Without loss of generality we can assume that $\widetilde{K}$ is convex. Let $E_{\widetilde{K}}$ be a normed space made up of the elements of the linear span of $\widetilde{K}$ and with the norm $\|\cdot\|_{\widetilde{K}}$ generated by this set. Moreover, let $\mathsf{E} \supset \widetilde{K}$ be an ellipsoid such that $s(\mathsf{E}) = s(\widetilde{K})$, and $E_{\mathsf{E}} \supset E_{\widetilde{K}}$ the Hilbert space with unit sphere E and norm $\|\cdot\|_{\mathsf{E}}$. It follows from the condition $s(\mathsf{E}) < \infty$ that the topology on E_{E} defined by the original norm $\|\cdot\|_{L^2(\Omega,\mu)}$ is majorized by the J-topology (see [102]) with respect to the topology of $\|\cdot\|_{\mathsf{E}}$. We now consider the positive definite functional $\varphi(x) = \int_\Omega e^{ix(\omega)}\, d\mu$ on E_{E}. The functional $\varphi(x)$ is obviously continuous in the norm $\|\cdot\|_{L^2(\Omega,\mu)}$, hence also in the J-topology with respect to the norm $\|\cdot\|_{\mathsf{E}}$; therefore, by the Minlos–Sazonov theorem (in Sazonov's form) the weak distribution determined by it can be extended to a completely additive measure μ^* in the Hilbert space of continuous linear functionals on E_{E}. Thus, the functions $x(\omega)$ forming the space E_{E} can be regarded as linear functionals on some Hilbert measure space (H^*, μ^*) that have the same finite-dimensional joint distributions. By the same token, a homomorphism g is induced [140] from the measure space (Ω, μ) onto (H^*, μ^*), acting in such a way that for an arbitrary (mod 0) element $\omega \in \Omega$ and an arbitrary $x \in E_{\mathsf{E}}$ we have that $x(\omega) = \langle x, \omega^*\rangle$, where $\omega^* = g(\omega) \in H^*$. But the bilinear form $\langle\cdot, \omega^*\rangle$ is continuous on the set $\mathsf{E} \subset E_{\mathsf{E}}$ in the relative topology for any fixed $\omega^* \in H^*$; therefore, for almost every $\omega \in \Omega$ the function $x(\omega) = \langle x, g(\omega)\rangle$ is continuous in x on the set E and a fortiori on the set $K \subset \mathsf{E}$. (We remark that we have actually obtained the Gel'fand–Kostjučenko theorem on the density of the system of generalized eigenvectors of a selfadjoint operator from the Minlos–Sazonov theorem; see [118].)

We now prove the second part of the assertion of the theorem. Let $K \subset H$, $s(K) = +\infty$. Then, as already shown, there is an orthonormal basis $\{e_k, k = 1, \ldots\}$ in H such that $\Sigma_{k=1}^\infty g_k^2(K) = \infty$, where $g_k(K) = \sup |(x, e_k)_H|$. Let the numbers $0 = m_0 < m_1 < m_2 < \cdots$ be such that

$$\sum_{m_j+1}^{m_{j+1}} g_k^2(K) \geqslant \sum_{k=1}^{m_j} g_k^2(K) + j + 1.$$

Now we describe a realization of H in the form of a subspace of the Hilbert space $L^2[0, 1]$ by constructing a certain orthogonal sequence $\{f_k(t), k = 1, \ldots\}$ and then associating the vectors $e_k \in H$ and $f_k \in L^2[0, 1]$. With this goal, let each of the functions $f_1(t), \ldots, f_{m_1}(t)$ take the three values h_k, 0, and $-h_k$, where the supports of the different functions are disjoint and have the whole segment [0, 1] as their union. We choose the values $h_1, \ldots, h_{m_1}$ so that $g_1(K)h_1 = g_2(K)h_2 = \cdots = g_{m_1}(K)h_{m_1}$. It is easy to see that if $\mathrm{Pr}_1^{(H)}$ denotes the orthogonal projection onto the subspace $L_1^{(H)} \subset H$ spanned by the vectors $e_1, \ldots, e_{m_1}$, and if Ψ denotes the linear isomorphism of $L_1^{(H)}$ onto the subspace L_1 of $L^2[0, 1]$ spanned by the vectors $f_1, \ldots, f_{m_1}$, then

$$\sup \Psi \mathrm{Pr}_1^{(H)} K = \Big(\sum_{k=1}^{m_1} g_k^2(K)\Big)^{1/2} = \text{const.}$$

The next sequence of functions $f_{m_1+1}, \ldots, f_{m_2}$ is defined so that, as before, each of them takes the three values h_k, 0, and $-h_k$, where $g_{m_1+1}(K)h_{m_1+1} = \cdots = g_{m_2}(K)h_{m_2}$, and, moreover, each function $f_i^+(t)$, $1 \leqslant i \leqslant m_1$ (which coincides with $f_i(t)$ where $f_i(t) = h_1$ and equals zero for the remaining $t \in [0, 1]$), is orthogonal to each function f_k, $m_1 + 1 \leqslant k \leqslant m_2$, and the same with respect to the functions $f_i^-(t)$, $1 \leqslant i \leqslant m_1$. It is easy to see that if $\mathrm{Pr}_2^{(H)}$ denotes the orthogonal projection onto the subspace $L_2^{(H)} \subset H$ spanned by the vectors $e_{m_1+1}, \ldots, e_{m_2}$, and if Ψ: $L_2^{(H)} \to L_2 \subset L^2[0, 1]$ is defined analogously to the preceding, then

$$\sup \Psi \mathrm{Pr}_2^{(H)} K = \left(\sum_{k=m_1+1}^{m_2} g_k^2(K)\right)^{1/2} = \text{const.}$$

Continuing this construction, making sure each time that the newly introduced functions f_k, which take three values, are orthogonal to the characteristic functions of all the sets on which the functions f_k defined earlier in the sequence are constant, we obtain an orthogonal sequence and linear isomorphism Ψ carrying $\{e_k\}$ into $\{f_k\}$; consequently, we get a set ΨK in $L^2[0, 1]$ isometric (and isometrically situated) with respect to $K \subset H$.

We show that $\sup \Psi K = +\infty$. With this goal, we decompose each element $y(t) \in \Psi K$ into an orthogonal series: $y = \Sigma_1^\infty y_j$, where $y_j \in L_j$. For any $m > 0$ we have

$$\sup \left\{\sum_{j=1}^{m} y_j\right\} \geqslant \sum_{j=1}^{m} \tilde{y}_j$$

(where $\tilde{y} = \Sigma_1^m y_j$ is the expansion of any fixed element $y \in \Psi K$);

$$\sum_{j=1}^{m} \tilde{y}_j \geqslant \tilde{y}_m - \sum_{j=1}^{m-1} |\tilde{y}_j| \geqslant \sup_K y_m - \sum_{j=1}^{m-1} \sup_K |y_j| \geqslant m$$

(by the choice of the numbers m_j); the sign of the modulus of y_j can be omitted, because K is balanced. Thus, $\sup \{\Sigma_1^m y_j\} \to \infty$. But it is clear that if $\{A_n\}$ is a sequence of ϵ_n-nets of the set $A \subset L^2$ and $\epsilon_n \to 0$, then

$$\limsup_{n\to\infty} \{\sup A_n\} \geqslant \sup A.$$

In our case the role of A is played by the set ΨK, and the role of A_n is played by its projection onto the subspace $L_1 \oplus \cdots \oplus L_n$. The theorem is completely proved. ●

In conclusion we mention that s-sets cannot be characterized in terms of the finite-dimensional diameters d_n. To prove this we construct two compact sets M_1 and $M_2 \supset M_1$ having the same sequence of finite-dimensional diameters, but of which only M_1 is an s-set. Let M_1 be an ellipsoid with semiaxes

$$b_k = \sqrt{\sum_{n=k+1}^{\infty} \frac{1}{n^2 \ln^2 n}},$$

and let $M_2 = \bigcap B_k$, where

$$B_k = \left\{x : x = \sum_{n=1}^{\infty} \xi_n e_n,\ \sum_{n=k}^{\infty} \xi_n^2 < \sum_{n=k}^{\infty} \frac{1}{n^2 \ln^2 n}\right\}.$$

It is not hard to check that for both M_1 and M_2

$$d_{k-1}^2 = \sum_{n=k+1}^{\infty} \frac{1}{n^2 \ln^2 n} = b_k^2, \quad \sum b_k^2 < \infty.$$

The compact set M_2 contains the parallelepiped

$$Q = \left\{x : x = \sum \xi_n e_n,\ |\xi_n| \leqslant \frac{1}{n \ln n}\right\}.$$

We show that Q is not an s-set.

LEMMA. *For the parallelepiped $P = \{x: x = \Sigma\, \xi_n e_n,\ |\xi_n| \leqslant \xi_n^{(0)}\}$ to be an s-set it is necessary and sufficient that $\Sigma\, \xi_n^{(0)} < \infty$ $(\xi_n^{(0)} \geqslant 0)$.*

Only the necessity needs to be proved. As is clear from the uniqueness in the finite-dimensional case, the axes of the minimal ellipsoid for a finite-dimensional parallelepiped are parallel to its edges. Let $\{b_k\}$ $(k = 1, \ldots, n)$ be the lengths of the axes of the minimal ellipsoid for the projection of P onto the subspace spanned by $e_1, \ldots, e_n$. By Hölder's inequality for $x = \Sigma_1^n\, \xi_k^{(0)} e_k$ we have

$$\sum b_k^2 \geqslant \left(\sum \xi_k^{(0)}\right)^2 \left(\sum \frac{\xi_k^{(0)2}}{b_k^2}\right)^{-1} = \left(\sum \xi_k^{(0)}\right)^2,$$

from which the lemma follows. Thus Q, and a fortiori M_2, is actually not an s-set. ●

As a simple example we mention now that segments of the Wiener spiral are not s-sets, since they are isometric to the corresponding segments of a Poisson process.

§9. Remark on Borel realizations

Let $K \subset L^2(\Omega, \mu)$ be a bounded convex balanced set. Suppose that the random process K has bounded realizations with probability 1. This means that the corresponding weak distribution can be extended to a measure in the strong dual of the space $(L(K), \|\cdot\|_K)$. If $(L(K), \|\cdot\|_K)$ is nonreflexive, then its dual can contain linear forms over $L(K)$ that are not Borel with respect to the Borel structure induced by the imbedding $K \subset L^2$. We can consider the strongly closed subspace $L(K)^{(*)} \subset (L(K), \|\cdot\|_K)^*$ consisting of all the linear forms on $L(K)$ that are bounded and Borel on K. It was hitherto not obvious that each weak distribution of the type considered that is extendible to a measure in $L(K)^*$ can be extended to a measure in $L(K)^{(*)}$; in fact, $L(K)^{(*)}$ is weakly dense in $L(K)^*$, and the situation could turn out to be analogous to the case with the spaces C and $L^\infty \supset C$, where there exist "nice" processes with a measure concentrated in L^∞, but not in C. However, it can be proved that, in the case described, if a weak distribution is extendible to a measure in $L(K)^*$, then $\mu^* L(K)^{(*)} = 1$ (μ^* is the outer measure), from which it follows that we can regard the realizations of the process K as Borel with probability 1 (compare with Doob's theorem in [24], p. 61, on the existence of a measurable modification).

CHAPTER III

INDEPENDENCE AND COMBINATIONS OF MEASURABLE DECOMPOSITIONS OF A PROBABILITY SPACE

0. Brief description of the problems discussed in the chapter. —§ 10. The Birkhoff-von Neumann problem. 1. The Birkhoff-von Neumann problem. A doubly stochastic integral operator and its kernel. Requirement of completeness of the measure space. Existence of a kernel in the case of Lebesgue spaces. Importance of completeness. Isomorphism operators and their kernels. Formulation in terms of operators (70). 2. The extreme points of the compact set M of all doubly stochastic measures. Impossibility of using the Choquet-Kreĭn-Mil'man theorem (75). 3. Other reformulations of the problem. Equivalence of the formulations. Condition of quasi-independence of decompositions. Terminology (76). 4. Existence of an independent nontrivial decomposition. Specifics of the continuous case in comparison with the matrix case. An example of Veršik (77). 5. Sketch of the general plan of proof of the basic theorem. Definitions and notation (79). 6. Description of the extreme points of the set of measures with given marginal distributions and majorized by a given measure. Generalization to the vector case. Application to the quasi-independent case. Corollary: existence of subsets of constant width. The case when $\xi \wedge \eta \neq \nu$ (81). 7. Description of pairs of marginal distributions of subprobability measures majorized by a given one. Theorem 6: reduction of the existence of a subprobability measure with given marginal distributions to the analogous problem for 2×2 matrices. Another formulation of Theorem 6. Remark on σ-finite measures (87). 8. Criterion for an isomorphism. Sets of the form $A \times B$, $\mu A + \nu B \geqslant c$, of arbitrarily small measure (94). 9. Definition of the functional Π_m. The case when $\Pi_m = 1$. Estimate of the closeness of marginal distributions (97). 10. Theorem 7: the approximation theorem. Version for the case when the decomposition $\xi \wedge \eta$ is purely continuous (99). 11. Definition of $\Pi_n C$, $\Pi'_n C$ and ΠC. Value of $\Pi' D$ when $D \subset \bigcup_k X_k \times Y_k$. Version of the Lebesgue theorem on points of density of a measurable set. Derivation of the relation $m(X_1 \times Y_1) \geqslant \mu X_1 + \nu Y_1 - 1$. For $X_1 \subset X$ there is a $Y^1 \subset Y$ such that $\nu Y^1 = 1 - \mu X_1$ and $\Pi'_m(M \setminus (X_1 \times Y^1)) < 1 + \epsilon$ (106). 12. Determination of how close measures subordinate to a given measure n can be to doubly stochastic ones (112). 13. Proof of the existence of an independent complement in the case of a bounded Radon-Nikodým derivative. Definition of the condition* У $(\epsilon; X_1, \ldots, X_k; Y_1, \ldots, Y_l)$. Existence of countable decompositions satisfying the condition У(117). 14. Elimination of the condition of a bounded derivative. Description of the construction. Theorem 8: Existence of an independent complement in the quasi-independent case. Generalization of Theorem 8 to the case when the decomposition $\xi \wedge \eta$ is purely continuous (122). 15. Application of the generalized Theorem 8 to measures on affine subspaces. The case when $\xi \vee \eta \neq \epsilon$. Discussion of the case of three factors. Examples (127). —§ 11 Probability measures on subsets of direct products. 0. Brief survey of the contents of the section. 1. Statement of the problem. Definitions (σ-algebras $\mathfrak{A} \overline{\otimes} \mathfrak{B}$ and $\mathfrak{A} \overline{\overline{\otimes}} \mathfrak{B}$, the Boolean algebra $\mathfrak{A}/\mu$, the σ-ring $\mathfrak{N}$). Discussion of the problem of σ-algebras of subsets of a product. The role of completeness of the measure space. The Boolean algebra $\Omega = \mathfrak{A}/\mu_X \overline{\otimes} \mathfrak{B}/\mu_Y$ and its properties (130). 2. Reformulation of the problem. The class $\mathfrak{F}$ of subsets of a product for which a solution is given. Definition of the classes P, αP and αP_k. Definition of the functional $\Pi(C)$ $(C \in \alpha P)$. The metric $\rho(C, D) = 2 - \Pi(C \Delta D)$ and the space $(\alpha P, \rho)$. Characteristics of the class $\mathfrak{N}$. Existence of a supremum for families in P. Definition of the class $\mathfrak{F}$ and verification of the condition $F \in \mathfrak{F}$ (133). 3. The space $M(F)$ of doubly stochastic measures on the set F, the topology τ on it, its compactness and

*Editor's note. У is the first letter of the Russian word условие which means "condition".

metrizability, and its structure. Theorem 9: condition for the nonemptiness of $M(F)$. Discussion (137). 4. Minimization of a set $F \in \mathfrak{F}$. Reformulation of Theorem 9 (140). 5. Problem of the existence of doubly stochastic densities on an arbitrary (and not only in the class $\mathfrak{F}$) measurable subset of a product. Strong minimization. The 2×2 matrices. Derivation of the relation $\int a(x)b(y)dm \geqslant \int a d\mu + \int b d\nu - 1$. Generalization of Theorem 6 to the σ-finite case. Theorem 10: criteria for the existence of a doubly stochastic density (141). —§12. Marginal sufficiency of statistics. Definitions and initial discussion. Theorem 11: sufficiency of a marginally sufficient statistic in the case of an independent sample. Counterexample. The case when the Radon-Nikodým derivatives of the one-dimensional distributions are bounded. General case. —§13. Conditions for the existence of a one-to-one optimal plan in the problem of transport of mass in Minkowski spaces. 1. Preliminary discussion and history of the problem. The Kantorovič-Rubinšteĭn metric. Counterexample. (160). 2. Statement of the result. Notation and definitions. Locally affine decompositions (161). 3. Sketch of the basic idea of the proof. Sets on which there always exists a measure that is the kernel of an isomorphism of the marginal distributions of an arbitrary probability measure on this set (163). 4. Proof of the basic result of the section (167).

0. This chapter deals with the study of a circle of problems connected with combinations of a number of measurable decompositions on a measure space. The language of the theory of decompositions is one of several equivalent languages suitable for the description of the problems relevant here. It would be possible to conduct the presentation in terms of, for example, rings of functions or operator theory, but we have chosen the language of "pure" measure theory for the exposition.

The main result (in the opinion of the author) in this chapter is the solution in §10 of a problem first posed, apparently, by Garrett Birkhoff (see [10], p. 266) that has attracted the attention of a whole series of mathematicians [42], [88], [83], [58], [89], [16], [59]. Certain results that bear an auxiliary character, as far as the solution of this problem goes, also have independent interest. Thus, Proposition 38 (in the multidimensional version) was used for a proof of the existence of a nonrandomized test in the Behrens-Fisher problem (see [45], [70], [71]). In §§11 and 12 the methods developed are used for the solution of problems that are not directly related to the Birkhoff problem (a brief presentation of the results is contained in the Introduction). Finally, in the last section of this chapter its basic results are applied to the solution of a problem in econometrics: the determination of conditions for the existence of a one-to-one plan in the Monge-Kantorovič problem on optimal transport of mass in a finite-dimensional Banach space (a Minkowski space).

The presentation of the results concerned with the Birkhoff problem is carried out in detail, so that not even the purely technical parts, sometimes involving quite laborious estimates, are omitted. In the applications the main attention is directed to the theoretical side of the problem, and some standard arguments used earlier (for example, the use of Zorn's lemma) are presented briefly. Nevertheless, all the auxiliary propositions are clearly separated everywhere.

§10. The Birkhoff-von Neumann problem

1. The well-known Birkhoff-von Neumann problem (see [11], or, for example, [10], p. 266) asserts that any square doubly stochastic matrix B (i.e., a matrix with nonnegative elements such that the sum of the elements in each row and in each

column is equal to 1) of order n can be represented in the form

$$B = \sum_{g \in G_n} p_g P_g,$$

where the p_g are nonnegative numerical coefficients, and P_g is the matrix of the permutation g of an n-element set, i.e., the P_g are square matrices with elements 0 and 1 ((0, 1)-matrices) such that in each column and in each row exactly one element is different from 0 (and equal to 1); G_n denotes the group of all such permutations g and contains, of course, $n!$ elements.

Number 111 in the "list" of unsolved problems in the second edition of Birkhoff's monograph *Lattice Theory* is to extend this theorem to the infinite-dimensional case, "under suitable hypotheses." The latter stipulation definitely shows that Birkhoff himself had in mind under the infinite-dimensional case not only infinite matrices, but also the "continuous" case (which is in principle different, as will be seen, from the discrete case). Birkhoff's problem for doubly stochastic infinite matrices was considered from various points of view by Isbell [42], Rattray and Peck [88], Kendall [58], and Révész [89]. The last-named proved, in essence, that each doubly stochastic infinite matrix can be represented as an integral with respect to a certain measure on the set of permutation matrices. In our view there is most ground for calling Révész' result the analogue of the finite-dimensional Birkhoff-von Neumann theorem in the case of infinite-dimensional matrices.

An infinite-dimensional analogue of a doubly stochastic matrix is the "kernel" of a doubly stochastic operator, i.e., a positive linear operator acting from some space of equivalence classes of measurable functions on a measure space $(X, \mathfrak{A}, \mu)$ into another such space of functions and carrying the function identically equal to 1 into the same unit (in this approach the "operator" is a more primary concept than its domain and image spaces, which can, in principle, be selected in various ways for one and the same operator; we call attention to the fact that no continuity of the operator is assumed—this is replaced by positivity). For definiteness, we can say that such an operator B acts from $L^\infty(X, \mathfrak{A}, \mu)$ to $L^\infty(Y, \mathfrak{B}, \nu)$.

A doubly stochastic integral operator (with which we shall have to deal in the following) is understood to be an operator B acting on a function f by the formula

$$(Bf)(y) = \int_X k(x, y) f(x) \mu(dx),$$

where $k(x, y)$ is a nonnegative function defined on the product $M = X \times Y$ of the measure spaces $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ that is integrable with respect to the product measure $\mu \times \nu$ and such that

$$\int_X k(x, y) \mu(dx) = 1, \quad \text{for } \nu\text{-almost all } y \in Y,$$

and

$$\int_Y k(x, y) \nu(dy) = 1, \quad \text{for } \mu\text{-almost all } x \in X.$$

For the spaces $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ it is always assumed that μ and ν are nonatomic probability measures and that the σ-algebras $\mathfrak{A}$ and $\mathfrak{B}$ of measurable subsets are countably generated, i.e., are completions (with respect to the measures μ and ν) of the smallest σ-algebras containing certain countable collections of subsets (countable bases [92]). It is well known that in this case the Banach algebras $L^\infty(X, \mathfrak{A}, \mu)$ and $L^\infty(Y, \mathfrak{B}, \nu)$ are isomorphic to each other in the sense of an isomorphism of the Banach algebra structures and of the partially ordered Banach space structures. Such an isomorphism does not imply an isomorphism of the measure spaces $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$, but it means an isomorphism of the Boolean σ-algebras of equivalence classes of measurable subsets of these spaces (see, for example, [38]) (the isomorphism of these Boolean σ-algebras follows already from the isomorphism of the Banach space structures of $L^\infty(X, \mathfrak{A}, \mu)$ and $L^\infty(Y, \mathfrak{B}, \nu)$. It is always possible, without changing these σ-algebras and hence without changing the spaces $L^\infty(X, \mathfrak{A}, \mu)$ and $L^\infty(Y, \mathfrak{B}, \nu)$ (as, in general, the spaces $S(X, \mathfrak{A}, \mu)$ and $S(Y, \mathfrak{B}, \nu)$ of all equivalence classes of measurable functions), to enlarge [92] $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ by adding new elements to X and Y in such a way that the enlarged measure spaces (do not confuse this with the completion of σ-algebras mentioned at the beginning of this paragraph!) $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ become isomorphic to each other as measure spaces, complete (see [92], and also Chapter I), and isomorphic to some standard "nice" (i.e., complete with a countable basis) measure space, for example, the segment $[0, 1]$ with Lebesgue measure, or the countable product of two points with one-half masses.

We recall that a complete measure space with a countably generated σ-algebra is called a Lebesgue space. An incomplete measure space is, in the exact sense of the word, just as exotic an object as a subset of a segment that is not Lebesgue measurable. In the following we assume that $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ are complete; however, this requirement is not necessary and is introduced only for convenience, since we could carry out the presentation only in terms of the various structures on $S(X, \mathfrak{A}, \mu)$ and $S(Y, \mathfrak{B}, \nu)$ that do not concern the measure spaces themselves.

The measure on $(X \times Y, \mathfrak{A} \mathbin{\overline{\otimes}} \mathfrak{B})$ for which the function $k(x, y)$ appearing in the definition of the doubly stochastic integral operator B is the density with respect to the product measure $\mu \times \nu$ is called the kernel of this operator. If $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ are Lebesgue spaces, then any positive measure m defined on the σ-algebra $\mathfrak{A} \mathbin{\overline{\otimes}} \mathfrak{B}$ (whose exact definition is given below) and whose marginal distributions, i.e. the measures $m_X = m\pi_X^{-1}$ and $m_Y = m\pi_Y^{-1}$ (where π_X and π_Y are the canonical projections $X \times Y \to X$ and $X \times Y \to Y$), are equal to μ and ν (doubly stochastic) generates an operator B (said to be doubly stochastic in this case) acting by the formula

$$(Bf)(y) = \int_{C_y} f(x)\, m_y(d\dot{x}) = \int_{X \times Y} f(x)\, d(m \mid y), \tag{1}$$

where $C_{y'} = \{(x, y)\colon\ y = y'\}$ is the corresponding element of the decomposition ξ_Y of $(X \times Y, \mathfrak{A} \mathbin{\overline{\otimes}} \mathfrak{B}, m)$ generated by the projection π_Y, and $m_{y'}$ is the conditional measure on this element, which can also be regarded as a measure $(m|y')$ on the whole space $(X \times Y, \mathfrak{A} \mathbin{\overline{\overline{\otimes}}} \mathfrak{B})$. A detailed study of the collection of measurable sets $\mathfrak{A} \mathbin{\overline{\overline{\otimes}}} \mathfrak{B}$

for the class of such doubly stochastic measures is contained in §11, and we shall not consider this question here. The measure m is also called the kernel of the operator B.

As is shown below, without the assumption of the completeness of $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ it cannot be asserted that any doubly stochastic operator is generated by some doubly stochastic measure on $X \times Y$, though the last term in (1) shows that even without the assumption of completeness, i.e. without the assumption of the existence of conditional distributions on the elements of the coordinate decomposition ξ_Y, the specification of a doubly stochastic measure on $X \times Y$ defines an operator (the "conditional mathematical expectation of the random variable f"). Nevertheless, the following statement holds.

PROPOSITION 4.1. *Let $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ be Lebesgue spaces and B a doubly stochastic operator acting from $L(X, \mathfrak{A}, \mu)$ to $L(Y, \mathfrak{B}, \nu)$. Then B is generated by some kernel m.*

PROOF. We find the measure m with the help of the system of conditional measures m_x on the elements of the coordinate decomposition $\xi_X = \xi$. This system of conditional measures can be obtained in the following way. Let $\xi^{(k)}$, $k = 1, \dots,$ be a refining sequence of coarsenings of the decomposition ξ_X generated by the projection π_X, where each $\xi^{(k)}$ is a decomposition of the whole space $X \times Y$ into a finite number of sets (for example, into 2^k sets that are projected onto X into 2^k disjoint subsets of equal measure). For each k and each element $C_{k,n}$, $n = 1, \dots, 2^k$, we consider the function $f_{k,n}(x, y)$ equal to 2^k on the set $C_{k,n}$ and zero outside it. On each element $C_{x'}$ of the decomposition ξ_X, i.e., on the set $\{(x, y) \in X \times Y,\ x = x'\}$, we now consider the measure $m_{x',k}$ that is absolutely continuous with respect to the measure $\nu\pi_Y^{x'}$ (where $\pi_Y^{x'}$ is the restriction of the operator π_Y to $C_{x'}$) and has density $(Bf_{k,n})\cdot(y)$ with respect to this measure. Let m_k be the measure on $(X \times Y,\ \mathfrak{A} \bar{\otimes} \mathfrak{B})$ having the family of measures $m_{x,k}$ as conditional measures on the elements of ξ_X, which makes sense, since

$$\int (Bf_{k,n})(y)\, \nu(dy) = 1$$

for any x. In §11 it is proved that the set M of all doubly stochastic measures is compact in the topology of uniform convergence on all subsets of the form $A \times B$, where $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$ (Proposition 70). Therefore, since on each subset of the form $A \times B$, where A is measurable with respect to the decomposition $\xi^{(l)}$ and $B \in \mathfrak{B}$, the values of the measures m_k stabilize as $k \longrightarrow \infty$ beginning with $k = l$, it follows that the sequence of measures m_k converges as $k \longrightarrow \infty$ to some limit m, which is the required kernel of the doubly stochastic operator B, because the operator with kernel m coincides with B on all the functions that are measurable with respect to one of the decompositions $\xi^{(k)}$, and hence coincides in general with B. •

As is clear from the proof, the system of conditional measures m_x on the elements C_x of ξ can be obtained in the following way. For (almost) all $x \in X$ the measure m_x is obtained as the limit of the measures $m_{x,k}$ in the sense described above, i.e., in particular, for each $B \in \mathfrak{B}$ and μ-almost all $x \in X$

$$m_{x,k}\left((\pi_Y^{\mp})^{-1}B\right)\to m_x\left((\pi_Y^{\mp})^{-1}B\right)\quad\text{as}\quad k\to\infty.$$

The set of those $x\in X$ for which this convergence holds can depend on the choice of B, but it is possible to fix a countable collection of sets B_s, $s=1,\dots$ (a basis in $(Y,\mathfrak{B},\nu)$), such that the measures m_x are determined by the convergence of the values of the measures $m_{x,k}$ only on this collection, and, at the same time, the set of those $x\in X$ for which the conditional measures m_x are determined does not depend on s and has full μ-measure.

Thus, we have established a one-to-one correspondence between the class of doubly stochastic operators and the class of doubly stochastic measures (kernels of operators).

The simplest doubly stochastic operators, though they are not integral operators, are the operators I_T giving an isomorphism of the Banach algebras $L^\infty(X,\mathfrak{A},\mu)$ and $L^\infty(Y,\mathfrak{B},\nu)$, and then, automatically, an isomorphism of the algebras $S(X,\mathfrak{A},\mu)$ and $S(Y,\mathfrak{B},\nu)$. Each such isomorphism is the adjoint of some isomorphism $x=Ty$ of the complete measure spaces $(X,\mathfrak{A},\mu)$ and $(Y,\mathfrak{B},\nu)$: $(I_Tf)(y)=f(Ty)$. The kernel of such an isomorphism operator is a measure in $X\times Y$ that is concentrated on the graph of the isomorphism T, i.e., on the set $G_T=\{(x,y)\colon\ x=Ty\}\subset M$, and the canonical projections π_X and π_Y carry this measure into the measures μ and ν.

If we do not assume the completeness of $(X,\mathfrak{A},\mu)$ and $(Y,\mathfrak{B},\nu)$, then we can give an example in which some operator I_T does not have a kernel. Indeed, let $X=P$ and $Y=Q$ be two nonmeasurable subsets of $[0,1]$ with the class L of Lebesgue measurable sets such that $l^*P=l^*Q=1$ and $P\cap Q=\varnothing$, where l is Lebesgue measure (regarding the existence of such subsets see, for example, [38], p. 70). On each such subset the function l^*, considered on the classes L_P and L_Q of intersections of P and Q, respectively, with l-measurable subsets of the segment, is completely additive [38], and the spaces $S([0,1],L,l)$ and $S(P,L_P,l^*)$, as well as the spaces $S([0,1],L,l)$ and $S(Q,L_Q,l^*)$, are canonically isomorphic (to each function in $S([0,1],L,l)$ we assign its restriction to P and Q, respectively). This defines a canonical isomorphism I_T of $S(P,L_P,l^*)$ and $S(Q,L_Q,l^*)$ that, however, cannot be given by means of a kernel m on $(P\times Q, L_P\,\overline{\otimes}\,L_Q)$; the kernel of the identity mapping from $S([0,1],L,l)$ to $S([0,1],L,l)$ is a measure concentrated on the diagonal of the unit square, which does not intersect the set $P\times Q\subset[0,1]\times[0,1]$. (For other purposes a similar example occurs again below.)

The kernel m_T of an isomorphism I_T is a natural infinite-dimensional analogue of the permutation matrices P_g of a finite set. Such a matrix P_g can be identified with the kernel of an isomorphism operator when the measure spaces $(X,\mathfrak{A},\mu)$ and $(Y,\mathfrak{B},\nu)$ each consist of n atoms having the same weight $(1/n)$. We denote by P_g a discrete measure defined on the product of two identical finite sets that corresponds to a permutation matrix. In general, if B is a doubly stochastic matrix of dimension $n\times n$, then B signifies the kernel of the doubly stochastic operator $\mathbf{R}^n\to\mathbf{R}^n$ whose matrix is $n^{-1}\mathrm{B}$.

Similar to the fact that in the finite-dimensional case each doubly stochastic matrix can (by the Birkhoff-von Neumann theorem) be represented as a weighted mean of permutation matrices, it will be shown that each doubly stochastic integral operator can be represented as the barycenter of a measure σ on the set $\widetilde{J}$ of isomorphisms I_T; moreover, it will be shown (what is, generally speaking, not possible to get in the finite-dimensional case) that the measure σ on the set $\widetilde{J}$ of isomorphisms can be chosen in such a way that the barycenters of any two disjoint measurable subsets of $\widetilde{J}$ are doubly stochastic operators with mutually singular kernels (the precise concepts of barycenter, collection of measurable subsets of $\widetilde{J}$, etc., are discussed below).

The set $\widetilde{M}$ of all doubly stochastic operators with various topologies (weak, strong operator in various norms, normed) was studied by Peck [83], Brown [16], and Kim [59], where it was determined in what sense it is possible to approximate the operators in $\widetilde{M}$ or the doubly stochastic integral operators by means of the isomorphisms or linear combinations of them, and also by means of operators adjoint to measure-preserving mappings that are not necessarily invertible. We do not need these results in the following.

2. As we mentioned above and will discuss in detail in §11, the set M of all doubly stochastic measures on $X \times Y$ is in the natural topology a compact set equipped with an affine structure and lying in the linear space of all measures of bounded variation whose positive and negative components have projections under the canonical projections onto X and Y that are absolutely continuous with respect to the measures μ and ν, or are proportional to them. In the finite-dimensional case the extreme points of the compact convex set $M_n \subset \mathbf{R}^{(n-1)^2}$ of all doubly stochastic matrices (the "Hungarian polyhedron") are precisely the permutation matrices (that is, in essence, the Birkhoff-von Neumann theorem, since each point of a convex polyhedron is the barycenter of a mass distribution on the vertices of the polyhedron).

In the compact set M each measure m_T that is the kernel of some isomorphism is an extreme point; but, unfortunately, such measures do not exhaust the set of extreme points. The simplest example of an extreme point of M that is not the kernel of an isomorphism is the kernel of an operator B acting according to the formula $(Bf)(y) = f(Ty)$, where $T\colon\ X \longrightarrow Y$ is a measure-preserving noninvertible transformation. Indeed, the corresponding kernel m_T is concentrated on the graph of the mapping T, just as in the case when T is an isomorphism. The conditional measures of the distribution m_T on the elements of the coordinate decomposition ξ_Y are δ-measures, and if we have $m_T = (m_1 + m_2)/2$, where $m_1,\ m_2 \in M$, then m_1 and m_2 are absolutely continuous with respect to m_T; therefore, since the conditional measures of m_1 and m_2 are probability measures by definition, and both m_1 and m_2 are doubly stochastic, we have on each element of the decomposition ξ_Y the single possible decomposition of the conditional δ-measure: $\delta = (\delta + \delta)/2$; consequently $m_1 = m_2$, i.e., m_T is an extreme point of M.

Later we give examples of other extreme points e of the compact set M, including some for which all the conditional measures of e with respect to the decompositions ξ_X and ξ_Y are pairs of equal masses. These examples show that there are doubly stochastic measures (kernels, operators) that a fortiori cannot be represented as barycenters of measures concentrated on the set J of kernels of isomorphisms.

Therefore, to reach our goal we cannot make use of the usual and more natural (since we are dealing with integral representations) apparatus of the Choquet-Kreĭn-Mil′man theorem on representation of the points of a compact convex set (with affine structure) as barycenters of certain Borel measures on the set of extreme points. And if we consider only the extreme points of the compact set M corresponding to isomorphisms, then, although we can describe a topology on the space of operators in which these and only these points are extreme points of a certain closed convex set, the set itself ceases to be compact in this topology, so here too the Choquet-Kreĭn-Mil′man theorem turns out to be inapplicable. Our approach presented below, which is not based on Choquet's theorem, gives us more than we would get if we could apply this theorem. We show that each doubly stochastic integral operator is the barycenter of a measure that is concentrated on a subset having the following property: the barycenters of any two disjoint subsets of it are disjunct. This improvement, which is of theoretical importance from the point of view of measure theory, does not (as was noted) have an analogue in the case of matrices, finite or infinite.

3. We give some other reformulations of the problem. Let $(\Omega, \mathfrak{A}, m)$ be a sample space, and $f(\omega)$ and $g(\omega)$ two statistics having purely continuous distributions such that the pair $(f(\omega), g(\omega))$ determines a unique sample element ω. Does there exist a third statistic $h(\omega)$ that is independent of $f(\omega)$ and of $g(\omega)$ and such that the pair $(f(\omega), h(\omega))$ determines a point of the sample space and the pair $(g(\omega), h(\omega))$ also determines a point of the sample space (formulation in terms of mathematical statistics)?

Let $(\Omega, \mathfrak{A}, m)$ be a measure space, and ξ and η two measurable decompositions such that the measures m/ξ and m/η are purely continuous. Does there exist a third measurable decomposition ζ that is an independent complement of ξ and of η (formulation in terms of measure theory)?

In these formulations we do not have to assume in advance that the product $\xi\eta$ of the decompositions ξ and η (the coarsest decomposition that is finer than ξ and η) is the decomposition ϵ into points (in the statistical formulation this corresponds to the requirement that an element ω of the sample space is determined by the values at it of the statistics f and g). It is shown later that under slight additional assumptions (which cannot be entirely avoided) there exists an independent complement of a pair of decompositions ξ and η even without the assumption of the condition $\xi\eta = \epsilon$. The omission of this last condition, which is natural in the approach from the point of view of pure measure theory, also leads to a formal improvement of the formulation in comparison with the formulation of the usual version of the Birkhoff-von Neumann theorem. The case when the decomposition $\xi \wedge \eta$ is nontrivial is especially important for applications.

If the decomposition $\xi \wedge \eta$ contains a continuous component, the doubly stochastic measure m a fortiori does not correspond to an integral operator; but even in this case we can prove, under appropriate assumptions, the existence of an independent complement. It is just this case that plays a role in the discussion of the existence of an optimal one-to-one plan of transport in the Monge-Kantorovič problem.

We show that these two formulations of the problem (with the condition $\xi\eta = \epsilon$) are actually equivalent to the Birkhoff problem with the additional requirement of disjunctness of the barycenters of disjoint subsets, which was mentioned above.

PROPOSITION 4.2. *Let ξ and η be measurable decompositions of the Lebesgue space $(M, \mathfrak{A}, m)$ such that $\xi\eta = \epsilon$, and ζ an independent complement both of ξ and of η (i.e., $\xi\zeta = \epsilon$, $\eta\zeta = \epsilon$, and if β': $M \subset M/\xi \times M/\zeta$ and β'': $M \subset M/\eta \times M/\zeta$ are the canonical imbeddings, then $(m/\xi) \times (m/\zeta) = m\beta'^{-1}$ and $(m/\eta) \times (m/\zeta) = m\beta''^{-1}$). Then almost every element $C^{(\zeta)}$ of the decomposition ζ, equipped with its conditional (under the decomposition ζ) measure $m_{C^{(\zeta)}}$, has the property that under the canonical imbedding β: $M \subset M/\zeta \times M/\eta$ this element $C^{(\zeta)}$ is carried into the graph E_T of some isomorphism T of the measure spaces $(M/\xi, m/\xi)$ and $(M/\eta, m/\eta)$, and the measure $m_{C^{(\zeta)}}$ is carried into the measure $m_T = m_{C^{(\zeta)}}\beta^{-1}$ that is the kernel of the corresponding doubly stochastic operator acting from the space of functions on $(M/\xi, m/\xi)$ into the space of functions on $(M/\eta, m/\eta)$. Conversely, if $\xi\eta = \epsilon$ and the measurable decomposition ζ is such that for almost every element $C^{(\zeta)}$ of it the image $\beta C^{(\zeta)}$ is the graph in $M/\xi \times M/\eta$ of some isomorphisms, and the canonical projections of the image $m_{C^{(\zeta)}}\beta^{-1}$ of the conditional measure on this element onto the spaces M/ξ and M/η are the measures m/ξ and m/η, then ζ is an independent complement both of ξ and of η.*

PROOF. Behind this unwieldy formulation there is actually a very simple fact. Since $\xi\eta = \epsilon$, and the canonical mapping β: $M \longrightarrow M/\xi \times M/\eta$ is an imbedding, we can carry the decomposition ζ over to the space $M/\xi \times M/\eta$. If the projections of the conditional measures of some decomposition onto the quotient space with respect to another decomposition coincide with the quotient measure with respect to this other decomposition, then the two decompositions are independent, and conversely ("the conditional probability coincides with the unconditional"). But if each (mod 0) element of some decomposition intersects the elements of another decomposition in not more than one point (in particular, is the graph of some mapping), then this means precisely that the two decompositions are mutual complements (their product is ϵ). ●

In the following we identify, for convenience of terminology, the respective spaces $(M/\xi, m/\xi)$ and $(M/\eta, m/\eta)$ with two copies of a Lebesgue space with continuous measure, denoted by $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$, preserving thereby the continuity of notation with the original formulation of the problem connected with the Birkhoff-von Neumann theorem. The image $m\beta^{-1}$ of the measure m under the imbedding β: $M \subset M/\xi \times M/\eta$ will also be denoted by the letter m. In the measure-theoretic formulation of the problem the condition that the doubly stochastic operator B be an integral operator corresponds to the condition of quasi-independence of the decompositions ξ and η, that is, the condition of absolute continuity of the measure m (more precisely, of the image $m\beta^{-1}$) with respect to the product $\mu \times \nu$, i.e., with respect to the product $m/\xi \times m/\eta$. The corresponding density $d(m\beta^{-1})/d(m/\xi \times m/\eta)$ is denoted by $k(x, y)$. Thus, the symbols x and y used to denote elements of the spaces X and Y can be regarded as notation for elements of M/ξ and M/η.

4. Before proceeding to the proof of the basic theorem, we mention also that even the existence (under the same assumptions) of a notrivial decomposition that is

independent with respect to two given ones (and not necessarily an independent complement) is a nontrivial fact.(1) Furthermore, it is clear from the start that it is far from being the case that each doubly stochastic operator admits a representation by means of isomorphisms (even not necessarily disjunct); a fortiori not every pair of decompositions has an independent complement. Obviously, there does not exist an independent complement of the decomposition ϵ of (M, Σ, m) into points; this corresponds to the case when m is concentrated on the graph of some noninvertible, generally speaking, transformation T. In fact, if T is noninvertible, then not only is there no independent complement of the coordinate decompositions ξ and η, but in general such a measure m_T is an extreme point of the set $\mathcal{M}$ of all doubly stochastic measures (here we can use the same argument as on p. 75).

We give another example that emphasizes well the peculiarity of the situation in the infinite-dimensional (continuous) case (this idea is due to A. M. Veršik). We construct a doubly stochastic measure w on the square $X \times Y$ whose conditional measures on (almost all) elements of the coordinate decompositions are two-point measures with equal ($= \frac{1}{2}$) masses, although this measure cannot be represented in the form $w = \frac{1}{2}(m_{T_1} + m_{T_2})$, where T_1 and T_2 are isomorphisms of $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$. Such a measure is the direct continuous analogue of a (0, 1)-matrix with two ones in each column and each row, which can be trivially represented (by the Birkhoff-von Neumann theorem) as half the sum of two permutation matrices. Let M be the circle of length one with Lebesgue measure m, whose points we denote by the angles φ, and let the decompositions ξ and η be defined in the following way. The elements of ξ are all possible pairs $(\varphi, \pi - \varphi)$, $\varphi \in (-\pi/2, \pi/2)$, and the elements of η are all possible pairs $(\varphi, \alpha + \pi - \varphi)$, where α is incommensurable with π. We consider the automorphisms T_1 and T_2 of M acting in such a way that T_1 interchanges the weighted points of each element of ξ, and T_2 interchanges the weighted points of each element of η. The automorphism $T_2 T_1$ carries φ into $\varphi + \alpha$; consequently it is ergodic; on the other hand, if there existed an independent complement of ξ and of η, then its elements would be invariant with respect to $T_2 T_1$, which is incompatible with ergodicity.

In Figure 1 the measure w on the unit square $M/\xi \times M/\eta$ is uniformly distributed along the perimeter of the rectangle drawn with heavy lines. We show that w is an extreme point of the compact set $\mathcal{M}$ of doubly stochastic measures. (This does not follow from the nonexistence of an independent complement, as shown by the example of the measure $m = \frac{1}{2}(m_{T_1} + m_{T_2})$, where $T_1\colon X \longrightarrow Y$ and $T_2\colon X \longrightarrow Y$ act according to the formulas $T_1 x = 2x \pmod 1$ and $T_2 x = 2x + 1/2 \pmod 1$; with respect to this measure the coordinate decompositions ξ and η do not have a common independent complement, since the conditional discrete measures $m_{C(\xi)}$ and $m_{C(\eta)}$ are made up of a different number of masses.) Indeed, if $w = \frac{1}{2}(m_1 + m_2)$, where m_1 and m_2

(1)This assertion was proved in different terms by Kellerer [56], [57] and in a more general situation (vector measures, which are useful for statistical applications) by Romanovskiĭ and the author [93]. More general conditions than the existence of a density and sufficient for the existence of a nontrivial independent decomposition were given by the author in [117] (see Proposition 43 below). Recently, V. I. Arkin and V. L. Levin [5] generalized this theorem, considering general linear restrictions connecting the values of the probability measure.

are doubly stochastic measures, then, clearly, m_1 and m_2 are absolutely continuous with respect to w, with densities $dm_1/dw = p_1$ and $dm_2/dw = p_2$. The conditional measures for m_1 and m_2 are pairs of masses equal to the values of the functions $p_1(x, y)$ and $p_2(x, y)$ at the corresponding points. We consider the function $q_1 = \operatorname{sign}(p_1 - 1)$ ($q_1 = 0$ when $p_1 - 1 = 0$). The measurability of q_1 follows from that of p_1. Let A_+ be the set on which $q_1(x, y) > 0$, and A_- the set on which $q_1(x, y) < 0$. Since on each vertical and each horizontal segment the sum of two values of the function p_1 (the total mass of the conditional measure) is equal to 1, the sets A_+ and A_- are constructed in such a way that the automorphism T_1 carries A_+ into A_- and A_- into A_+, and the same is true for T_2. Therefore, A_+ and A_- are invariant with respect to the ergodic automorphism T_2T_1, i.e., they are empty. In other words, $m_1 = w$, and then also $m_2 = w$.

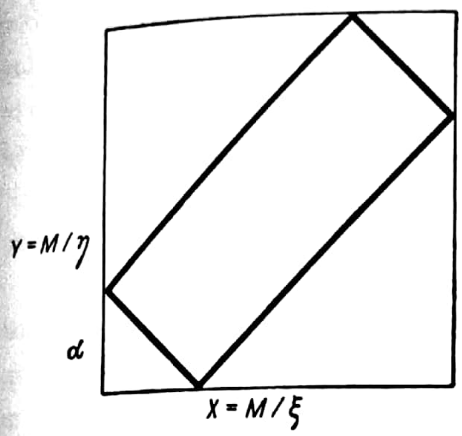

FIGURE 1

Veršik constructed for any n a pair of measurable decompositions of a Lebesgue space such that each conditional measure consists of n equal masses and there does not exist a nontrivial measurable decomposition that is independent with respect to each of these two. For this purpose Veršik considers the Bernoulli action of the free product of the cyclic group $\mathbf{Z}_n$ with itself, and takes the required decompositions to be the decompositions corresponding to the two factor subgroups. According to Veršik's communication, consideration of a projective limit of such groups enables one to get an example of a pair of measurable decompositions of a Lebesgue space with nonatomic conditional measures that does not admit a nontrivial independent complement. The existence of such a pair of decompositions is fundamental: it shows that our basic result on the existence of an independent complement under the assumption of quasi-independence of the given decompositions cannot be improved. The existence itself of such a pair, as observed by Veršik, follows from the fact that both the set of all extreme points of M and the set of measures in M with nonatomic conditional measures are dense G_δ sets in M.

Obviously, each extreme point of the compact set M does not admit a nontrivial decomposition that is independent with respect to both the coordinate decompositions. The problem of the study and clear description of the set of extreme points of M is complicated. Ryff, Shiflett, and others have dealt with the study of these extreme points.

5. We proceed, finally, to the proof of the basic assertion of this chapter. It will be shown that if $\xi\eta = \epsilon$ and the image of the measure m under the canonical mapping $M \to M/\xi \times M/\eta$ of the Lebesgue space (M, Σ, m) into the product of the spaces M/ξ and M/η, where the measures m/ξ and m/η are purely continuous, is absolutely continuous with respect to the product measure $m/\xi \times m/\eta$, then there exists a measurable

decomposition ζ of the space (M, Σ, m) that is an independent complement of the decompositions ξ and η. As explained, we can assume that $M = X \times Y$, where $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ are nonatomic Lebesgue spaces, ξ and η are the coordinate decompositions, and the Radon-Nikodým derivative $dm/d(\mu \times \nu)$ exists and is equal to $k(x, y)$.

We first sketch the plan of the proof. To begin with, we obtain simple conditions on the pair of measurable decompositions ξ and η of (M, Σ, m) sufficient for there to exist, for each subprobability measure m_1 satisfying the relation $m_1 \leqslant m$, a measure $\widetilde{m}_1$ such that $m_1/\xi = \widetilde{m}_1/\xi$, $m_1/\eta = \widetilde{m}_1/\eta$, and $d\widetilde{m}_1/dm$ takes the values 1 and 0. Under this condition the set of extreme points of the compact set $K = K(\xi, \eta, m_\xi, m_\eta)$ of measures m_1 on (M, Σ) satisfying the relations $m_1 \leqslant m$, $m_1/\xi = m_\xi$, and $m_1/\eta = m_\eta$, where m_ξ and m_η are fixed measures on M/ξ and M/η, has a simple description. In particular, these conditions are easily verified in the case of interest to us, when there is a density $k(x, y)$.

Then we obtain conditions in a form convenient for later use that are necessary and sufficient for the nonemptiness of the compact set $K(\xi, \eta, m_\xi, m_\eta)$.(2) The idea of the following argument is to construct a refining sequence of finite decompositions ζ_k, each of which is independent with respect to ξ and to η. The main difficulty here consists in getting convergence of the sequence ζ_k to a complement. Of course, the limit of a refining sequence of measurable decompositions always exists, and, together with the decompositions ζ_k, this limit is independent with respect to ξ and η. Furthermore, it is not hard to show by using Zorn's lemma that the limit decomposition can be so fine that the conditional measures on its elements are "indivisible", i.e., do not admit further independent subdecompositions. (If we do not speak of decompositions of the measure space (M, Σ, m), but of its "coverings", then we arrive at a "covering" made up of extreme points of the compact set M—a result that was discussed earlier in connection with Choquet's theorem.)

To make the limit decomposition an independent complement, we use the following method. We prove an approximation theorem on the possibility of a good approximation of the characteristic functions of subsets of M that are sufficiently "narrow" with respect to the decompositions ξ and η by means of characteristic functions of subsets having exactly constant width (i.e., by means of densities of subprobability measures whose marginal distributions with respect to ξ and η are proportional to the measures μ and ν). This theorem is proved under the assumption of boundedness of the density $k(x, y) = dm/d(\mu \times \nu)$. Then subsets are constructed such that the decompositions of M by them approximate a complement of ξ and η and which can be approximated, by the approximation theorem, by sets that are measurable with respect to the desired independent complement. The (transfinite) passage to the limit gives a solution to the

---

(2) Conditions for the existence of a measure that is majorized by a given one and has given marginal distributions for finite measure spaces (the problem of Fréchet) were obtained by Berge (unpublished) and Dall'Aglio [22]. For the case of a measure that is absolutely continuous with respect to a product measure, see Kellerer [56]; one particular case (m a product measure) was considered by Kantorovič and Romanovskiĭ [50]. Still another important case (measures on a product of spaces of which one is a Polish space) was investigated by Strassen [113]. Our brief proof contained in [117] and relating to a very general situation is based on another idea.

problem in the case of a density $k(x, y)$ that is bounded from above. Then for an arbitrary density $k(x, y)$ we consider the truncations $k_N(x, y)$, we construct finite decompositions close to an independent complement of the decompositions ξ and η with respect to the measure with density $k_N(x, y)$, and we prove that a certain approximation of them by decompositions that are independent of ξ and η with respect to m is close to the desired independent complement, whose construction is concluded by an appropriate limit passage.

In the sequel we discuss possible generalizations of the theorem, of which the main one deals with the case $\xi \vee \eta \neq \epsilon$.

We introduce some definitions and notation. Let ξ be a measurable decomposition of the space (Ω, Σ, m), and m_1 a measure that is absolutely continuous with respect to m and for which $dm_1/dm = f(\omega)$, $\omega \in \Omega$. The width of the measure m_1 on an element C of ξ is defined to be $d(m_1/\xi)/d(m/\xi)|_C = \int_C f(\omega)dm_C$, the width of the function $f_1(\omega)$ on the element C is defined to be the width on this element of the measure m_1 with density $dm_1/dm = f_1(\omega)$, and the width of a measurable subset $A \in \mathfrak{A}$ is defined to be the width of its characteristic function $\chi_A(\omega)$.

If the measure m_1 is not absolutely continuous with respect to m, but is in some sense correctly defined (for example, if $\Omega = X \times Y$, ξ and η are the coordinate decompositions, $m = \mu \times \nu$, and m_1 is an arbitrary doubly stochastic measure, i.e. a measure such that $m_1/\xi = \mu$ and $m_1/\eta = \nu$), then the expression $d(m_1/\xi)/d(m/\xi)|_C$ is also meaningful (for (m/ξ)-almost all C) and is also called the width of the measure m_1 on the element C of ξ.

Obviously, a decomposition that is independent with respect to ξ is a decomposition such that all the conditional measures on its elements have constant (unit) width on the elements $C^{(\xi)}$. If $\Omega_1 \subset \Omega$, $\Omega_1 \in \Sigma$ and $m\Omega_1 > 0$, then m_{Ω_1} denotes the measure defined by $m_{\Omega_1}A = (m\Omega_1)^{-1}mA$ (in accordance with the common notation for conditional measures). The absence of any indication of the set over which an integral is taken means that the integration is performed over the whole space. The canonical projections $X \times Y \longrightarrow X$ and $X \times Y \longrightarrow Y$ are denoted by π_X and π_Y, respectively. The expressions $(\pi_X f)(x)$ and $(\pi_Y f)(y)$, where $f(x, y)$ is an arbitrary m-integrable function on $M = X \times Y$, are used to denote the densities with respect to μ and ν, respectively, of the marginal distributions of the measure with density $f(x, y)$ with respect to m, and similarly for an arbitrary homomorphism π of (Ω, Σ, m) into some other Lebesgue space.

6. As before, let $M = X \times Y$. We use the notation

$$W(M) = \{f : |f(x, y)| \leqslant 1 \ \ m\text{-almost everywhere}\},$$
$$V(M) = \{f : 0 \leqslant f(x, y) \leqslant 1 \ \ m\text{-almost everywhere}\}.$$

Let $\hat{\xi}$ and $\hat{\eta}$ be certain coarsenings of the decompositions ξ and η consisting each of two sets, $\hat{\varphi}$ and $\hat{\psi}$ the respective canonical homomorphisms $X = M/\xi \longrightarrow M/\hat{\xi} = \hat{X}$ and $Y = M/\eta \longrightarrow M/\hat{\eta} = \hat{Y}$, and $\hat{\pi}_{\hat{X}}$ and $\hat{\pi}_{\hat{Y}}$ the respective canonical homomorphisms $M/\hat{\xi}\hat{\eta} = \hat{M} \longrightarrow \hat{M}/\hat{\xi} = \hat{X}$ and $\hat{M} \longrightarrow \hat{M}/\hat{\eta} = \hat{Y}$. The expressions $\hat{\pi}_{\hat{X}} f$ and $\hat{\pi}_{\hat{Y}} f$, where f is a function on $\hat{M}$, are interpreted, according to the preceding, as the densities with respect to the measures $\hat{\mu} = \mu\hat{\varphi}^{-1}$ and $\hat{\nu} = \nu\hat{\psi}^{-1}$ of the marginal distributions of the measure on $\hat{M}$ having density $f(\hat{x}, \hat{y})$ with respect to the measure $\hat{m} = m(\hat{\varphi} \times \hat{\psi})^{-1}$. (The measure space $(\hat{M}, \hat{m})$ consists of four masses, and the measure spaces $(\hat{X}, \hat{\mu})$ and $(\hat{Y}, \hat{\nu})$ each contain two masses.)

Further, for a measurable decomposition ξ of the space (M, Σ, m) let the expression $L^p_\xi(M)$ denote the subspace of the space $L^p(M)$ (of functions on (M, Σ, m) whose pth powers are integrable) consisting of all such functions that are measurable with respect to ξ; the expression $S_\xi(M)$ has an analogous sense. The restriction of a measurable decomposition ξ to a subset $\widetilde{M} \subset M$ will also be denoted by the letter ξ.

PROPOSITION 43. *On the probability measure space (M, Σ, m) let $\zeta_1, \ldots, \zeta_n$ be measurable decompositions such that for any subset $\widetilde{M} \in \Sigma$, $m\widetilde{M} > 0$, the space $L_{\zeta_1}(\widetilde{M}) + \cdots + L_{\zeta_n}(\widetilde{M})$ is not dense (in the norm) in $L(\widetilde{M})$. Then for any integrable functions $b_k(x_k)$, $k = 1, \ldots, n$, defined on the respective spaces $M/\zeta_1, \ldots, M/\zeta_n$ the subset B of $L^\infty(M, \Sigma, m)$ distinguished by the conditions*

1) $f \in B \Rightarrow 0 \leqslant f \leqslant 1$,

2) *f has width $b_k(x_k)$, $x_k \in M/\zeta_k$, with respect to the decomposition ζ_k, $k = 1, \ldots, n$,*

is convex, compact in the weak topology of the dual space, and has as its extreme points only functions f taking only the values 0 *and* 1.

In other words, each extreme point of the set of measures on M majorized by the measure m and having given marginal distributions has density with respect to m that takes only the values 0 and 1.

PROOF. Obviously, B is a convex bounded set in $L^\infty(M)$. We verify its closedness in the topology $\sigma(L^\infty, L)$. If $f \in B$, $f_n \in B$ $(n = 1, 2, \ldots)$, and $\int f_n g\,dm \rightarrow \int fg\,dm$ for any function $g \in L$, then the satisfaction of the last condition for $g \in L_{\zeta_k}$ means the (weak) convergence of the functions $\pi_{\zeta_k} f_n$ to the function $\pi_{\zeta_k} f$ as $n \rightarrow \infty$. But the condition $f_n \in B$ means that all $(\pi_{\zeta_k} f_n)(x_k)$, $n = 1, 2, \ldots$, coincide with the functions $b_k(x_k)$, and therefore also $\pi_{\zeta_k} f$ coincides with it, and condition 2) is satisfied for f. Moreover, it is clear that for the limit function f also the condition 1) is satisfied so that $f \in B$. The weak compactness of the subset follows from its weak closedness and its boundedness [12]; therefore, if the (obviously) convex set B is not empty, then it has extreme points. Our aim is to show that under the conditions of the theorem no function f that takes (essentially) values between 0 and 1 can be an extreme point of B.

Suppose that B is not empty and let f be an extreme point of it. We assume that on a set of positive measure f is different from both 0 and 1. We can assume that $0 < \delta \leqslant f \leqslant 1 - \delta$ on a set $\widetilde{M}$ of positive measure for some positive number δ. We now consider the Lebesgue space $(\widetilde{M}, \widetilde{\Sigma}, m_{\widetilde{M}})$ (a subspace of (M, Σ, m)). We prove that in $L^\infty(\widetilde{M})$ there is a nonzero element that is orthogonal at once to all the subspaces $L_{\zeta_K}(\widetilde{M})$ $(k = 1, \ldots, n)$. Since, by hypothesis, the space $L_{\zeta_1}(\widetilde{M}) + \cdots + L_{\zeta_n}(\widetilde{M})$ is not dense in $L(\widetilde{M})$, its annihilator N in $L^\infty(\widetilde{M})$ is different from zero. Any element

$\widetilde{f}\in N$ has the property that $\pi_{\zeta_k}\widetilde{f}=0$ for any $k=1,\ldots,n$; in fact, $\int\widetilde{f}g\,dm=0$ for any function $g\in L_{\zeta_k}$, and this is equivalent to saying that the function $\pi_{\zeta_k}\widetilde{f}\in L^\infty(\widetilde{M}/\zeta_k)$ is orthogonal to any function in $L(\widetilde{M}/\zeta_k)$, i.e., is identically zero. But if $\widetilde{f}\neq 0$, then $f=\frac{1}{2}(f+\widetilde{\widetilde{f}})+\frac{1}{2}(f-\widetilde{\widetilde{f}})$, where $\widetilde{\widetilde{f}}=\delta\|\widetilde{f}\|^{-1}_{L^\infty(\widetilde{M})}\widetilde{f}$ on the set $\widetilde{M}$ and $\widetilde{\widetilde{f}}=0$ outside $\widetilde{M}$. This relation implies that the element $f\in B$ is not extreme. ●

REMARK. Only slight changes are needed to prove a proposition generalizing Proposition 43 to the case of a vector-valued measure. We do not need this case in the following, but it proves to be very useful in certain statistical applications (see [144] and [70], Chapter X, §1); therefore we explain, if only briefly, those changes entailed by such a complication in the formulation and in the proof.

We can consider a vector-valued measure $m=(m_1,\ldots,m_s)$ such that the probability measures $m_1,\ldots,m_s$ are all pairwise mutually absolutely continuous, otherwise considering the "pieces" of the space M on which some of the measures m_k are pairwise mutually absolutely continuous, while the remaining ones vanish. The set $B\subset L^\infty$ is now distinguished by the following conditions:

1) $f\in B\Rightarrow 0\leqslant f\leqslant 1$.

2a) With respect to the measure m_i and the decomposition ζ_k, f has given width $b_k^{(i)}(x_k)$, where $x_k\in M/\zeta_k$, $i=1,\ldots,s$ and $k=1,\ldots,n$.

The space L^∞ can be regarded, in particular, as the dual space of $L(M,\Sigma,m_1\wedge\cdots\wedge m_s)$, where $m_1\wedge\cdots\wedge m_s$ is the infimum of the measures $m_1,\ldots,m_s$ in the sense of the Riesz space structure [13] on the set of measures on M (i.e., the measure having density with respect to a certain measure m_0 equal to $\min\{p_1,\ldots,p_s\}$, where $p_i=dm_i/dm_0$). For any $k=1,\ldots,n$ and $i=1,\ldots,s$ the following inclusions are obvious:

$$L_{\zeta_k}(M,\Sigma,m_i)\subset L(M,\Sigma,m_i)\subset L(M,\Sigma,m_1\wedge\cdots\wedge m_s);$$

therefore

$$\sum_{\substack{k=1,\ldots,n\\ i=1,\ldots,s}} L_{\zeta_k}(M,\Sigma,m_i)\subset \dot{L}(M,\Sigma,m_1\wedge\cdots\wedge m_s).$$

The smallness of the norm of the element g in $L(M,\Sigma,m_1\wedge\cdots\wedge m_s)$ means that it can be represented in the form $g=g_1+\cdots+g_s$, where all the norms $\|g_i\|_{L(M,\Sigma,m_i)}$ are small. Arguing as in the case of a scalar measure, we arrive at the following assertion.

PROPOSITION 43a. *Let (M,Σ,m) be a space with vector measure $m=(m_1,\ldots,m_s)$, where $m_1,\ldots,m_s$ are probability measures, and let $\epsilon_1,\ldots,\epsilon_n$ be measurable decompositions of $(M,\Sigma,s^{-1}(m_1+\cdots+m_s))$ such that, for any subset $\widetilde{M}\subset M$ on which all measures $m_{r_1},\ldots,m_{r_p}$ that do not vanish identically are pairwise equivalent, the subspace*

$$L'=\sum_{\substack{i=1,\ldots,p\\ k=1,\ldots,n}} L_{\zeta_k}(\widetilde{M},\Sigma,m_{r_i})$$

of $L(M, \Sigma, m_{r_1} \wedge \cdots \wedge m_{r_p})$ *is not norm dense (the imbedding of the spaces* $L_{\zeta_k}(\widetilde{M}, \Sigma, m_{r_p})$ *in* $L(\widetilde{M}, \Sigma, m_{r_1} \wedge \cdots \wedge m_{r_p})$ *is understood in the sense that elements that are identified coincide as measures on* $\widetilde{M}$). *Then for any integrable functions* $b_k^{(i)}(x_k)$, *where* $x_k \in M/\zeta_k$, $i = 1, \ldots, s$ *and* $k = 1, \ldots, n$, *the subset* B *of* $L^\infty(M, \Sigma, s^{-1}(m_1 + \cdots + m_s))$ *distinguished by the conditions* 1) *and* 2a) *in the Remark after Proposition* 43 *is convex, compact in the weak topology* $\sigma(L^\infty, L)$ *of the dual space, and has as extreme points only the functions* $f \in B$ *that do not take (essential) values different from* 0 *and* 1.

For the proof we assume the nonemptiness of the compact set $B \subset L^\infty$ and consider an arbitrary function $f \in B$. We assume that on some subset $M_\delta \in \Sigma$ for which $\Sigma_{j=1}^s m_j M_\delta > 0$ we have

$$0 < \delta \leqslant f(x) \leqslant 1 - \delta < 1, \quad x \in M_\delta,$$

and we show that the function f is not an extreme point of the compact set by proving the existence of a function $g(x) \not\equiv 0 \pmod{s^{-1}\Sigma_j m_j}$ for which $f + g \in B$ and $f - g \in B$. Assuming without loss of generality that on M_δ some of the measures m_j vanish while the rest, say $m_{r_1}, \ldots, m_{r_p}$, are equivalent, we find, as above, that we can take g to be a function that vanishes outside M_δ and coincides on M_δ with some sufficiently small function in the annihilator of the space

$$\sum_{\substack{k=1,\ldots,n \\ i=1,\ldots,p}} L_{\zeta_n}(\dot{M}_\delta, \Sigma, m_{r_i}). \ \bullet$$

Proposition 43a can be regarded also as a particular generalization of the well-known theorem of Ljapunov [72] on the convexity of the range of a vector-valued measure. Indeed, Proposition 43a shows that the image of the set of characteristic functions $\{\chi_A\}$ of subsets A of the space M with vector-valued measure $m = (m_1, \ldots, m_s)$ ($\overline{m} = (\Sigma_1^s m_i)/s$) on it, under the mapping

$$\begin{aligned} \pi: L(M, \overline{m}) \to\ & L(M/\zeta_1, m_1/\zeta_1) \times L(M/\zeta_1, m_2/\zeta_1) \times \ldots \times L(M/\zeta_1, m_s/\zeta_1) \\ & \times L(M/\zeta_2, m_1/\zeta_2) \times \ldots \times L(M/\zeta_2, m_s/\zeta_2) \\ & \cdots\cdots\cdots\cdots\cdots\cdots \\ & \cdots\cdots\cdots\cdots\cdots\cdots \\ & \times L(M/\zeta_n, m_1/\zeta_n) \times \ldots \times L(M/\zeta_n, m_s/\zeta_n), \end{aligned}$$

that assigns to a function $f \in L^\infty$ the tuple $\{\pi_{\zeta_k}^{(m_i)} f\}$ of marginal densities (under the decompositions $\zeta_1, \ldots, \zeta_n$) of the measures having densities equal to f with respect to the given measures $m_1, \ldots, m_s$, coincides with the image under this mapping of the closed (in L^∞) convex set of all nonnegative measurable functions not exceeding 1. In fact, taking for an arbitrary function $f \in L^\infty(M, \overline{m})$, $0 \leqslant f \leqslant 1$, the functions $b_k^{(i)}(x_k)$ in Proposition 43a to be the densities of the marginal distributions of the measures with densities equal to f, we get a compact convex set $B = B_f = \pi^{-1}\pi f$ that contains f and is hence not empty. Any extreme point of this compact set is, by Proposition 43a, the characteristic function of a subset of M.

In the following we limit ourselves to the case of a scalar measure and consider, for simplicity of notation, two decompositions ξ and η. To apply Proposition 43 (or 43a) it suffices to prove the existence of a function that cannot be approximated by elements of the space $L_\xi + L_\eta$.

On the square $[0, 1]^2$ we consider the measure uniformly distributed on the graphs $y = x$ and $y = x + \alpha$. When α is irrational, it can be shown that for this measure the sum $S_\xi + S_\eta$ is dense in S, but L_ξ and L_η are orthogonal to the function taking the values 1 and -1 on the respective graphs $y = x$ and $y = x + \alpha \pmod 0$; hence their sum is also orthogonal to this function, which thus cannot be approximated in the metric of L by functions of the form $f(x) + g(y)$. It is also possible to give examples of sets whose widths with respect to both decompositions differ uniformly by an arbitrarily small amount from constants that are neither 0 nor 1, and whose characteristic functions belong to the sum $L_\xi + L_\eta$. The mapping $L_\xi \times L_\eta \to L_\xi + L_\eta \subset L$ is not, generally speaking, a homomorphism.

We prove the existence of a function that cannot be approximated in the case of interest to us, when there exists a density $k(x, y)$.

PROPOSITION 44. *Let ξ and η be measurable decompositions of (M, Σ, m) such that $\xi\eta = \epsilon$, the measures $\mu = m/\xi$ and $\nu = m/\eta$ are purely continuous, and the image of m under the canonical mapping $\pi_X \times \pi_Y$: $M \to M/\xi \times M/\eta \equiv X \times Y$ is absolutely continuous, with density $k(x, y)$, with respect to the product measure $m^* = m/\xi \times m/\eta \equiv \mu \times \nu$. Suppose that $h(x, y)$ is a measurable function defined on $X \times Y$ and taking values in $[0, 1]$. Then there exists a measurable subset $A \subset \{(x, y)\colon\ h(x, y) > 0\}$ such that its characteristic function $\chi_A(x, y)$ satisfies*

$$\pi_X\chi_A = \pi_X h, \quad \pi_Y\chi_A = \pi_Y h.$$

PROOF. Without loss of generality, we can assume that $h(x, y) > 0$ for m-almost all points (x, y); in the opposite case we consider the subspace $\{(x, y)\colon\ h(x, y) > 0\}$. It is convenient to present the proof assuming that the spaces $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ are realized as unit intervals with Lebesgue measure, so that M is the square $[0, 1] \times [0, 1]$ with the doubly stochastic kernel $k(x, y) = dm/dm^*$ on it. To verify the hypotheses of Proposition 43 we consider an arbitrary measurable subspace $\widetilde{M} \subset M$ of positive m-measure and on which $k(x, y) > 0$ $(\mu \times \nu)$-almost everywhere, and we construct on it a function that cannot be approximated in L by functions in $L_\xi + L_\eta$. Let $\epsilon', \epsilon'' > 0$ be arbitrarily small positive numbers such that

$$\widetilde{M} \cap (M + (\varepsilon', 0)) \cap (M + (0, \varepsilon'')) \cap (M + (\varepsilon', \varepsilon'')) = P, \quad mP > 0,$$

and let $Q \subset P$ be a subset of positive measure such that the four sets Q, $Q - (\epsilon', 0)$, $Q - (0, \epsilon'')$ and $Q - (\epsilon', \epsilon'')$ are pairwise disjoint. We consider the set

$$R = Q \cup (Q - (\varepsilon', 0)) \cup (Q - (0, \varepsilon'')) \cup (Q - (\varepsilon', \varepsilon'')) \subset P.$$

The function χ_R is the desired one. Indeed, let θ be the measurable decomposition of this space, regarded as a subspace of (M, Σ, m), whose elements are the quadruples of points $((x, y) \in Q, (x - \epsilon', y), (x, y - \epsilon''), (x - \epsilon', y - \epsilon''))$. Under the decomposition θ the conditional measures of the points with respect to the Lebesgue measure $\mu \times \nu$ on M are each equal to ¼, and the conditional measures of these points with respect to the measure m with density $k(x, y)$ are proportional to the values of this density at each of the points, and therefore, since $k(x, y) > 0$ on $\widetilde{M}$ and a fortiori on R, all are different from zero. But on the discrete space consisting of four point masses t_1, t_2, t_3, t_4 the characteristic functions of the singleton sets cannot be represented in the form $f(t) + g(t)$, where $f(t_1) = f(t_2), f(t_3) = f(t_4), g(t_1) = g(t_3)$ and $g(t_2) = g(t_4)$, since the dimension of the set $\{f(t) + g(t)\}$ is equal to three, while the four characteristic functions of the singleton sets form a basis in the space of all functions on $\{t_1, t_2, t_3, t_4\}$. If it were possible to approximate χ_Q arbitrarily well by functions of the form $f(x) + g(y)$, then it would be possible to approximate the characteristic function of a point $(x, y) \in Q \subset R$ arbitrarily well on elements of the decomposition θ, which, as we have seen, is impossible. ●

The proof of the analogous proposition for the case of an arbitrary finite number of decompositions and an arbitrary vector measure would differ only in the fact that we would have to construct a decomposition θ not into quadruples of points, but into certain lattices containing a sufficiently large number of points such the characteristic functions of their points cannot be approximated by corresponding sums of functions, each depending only on one coordinate.

COROLLARY. *For any number λ, $0 \leqslant \lambda \leqslant 1$, there exists a subset $M_\lambda \subset M$ such that $\pi_X \chi_{M_\lambda}(x) \equiv \lambda$ and $\pi_Y \chi_{M_\lambda}(y) \equiv \lambda$. For any measurable decomposition ζ^N of the Lebesgue space (N, n) into subsets of positive measure $\{C_{\lambda_k}\}, nC_{\lambda_k} > 0, k = 1, \ldots$, there exists a decomposition ζ of (M, m) into subsets M_{λ_k} of constant width with respect to each of the decompositions ξ and η such that the discrete spaces M/ζ and N/ζ^N are isomorphic; in other words, the width $S_{M_{\lambda_k}} = \pi_X \chi_{M_{\lambda_k}}$ of M_{λ_k} is equal to the measure nC_{λ_k} of the corresponding element of ζ^N.*

We now prove a generalization (important for the sequel) of Proposition 44 to the case when $\xi \wedge \eta \neq \nu$.

PROPOSITION 44*. *Let ξ and η be measurable decompositions of the space (M, Σ, m) such that $\xi\eta = \epsilon$, and the image of m under the canonical mapping $\pi_X \times \pi_Y$: $M \longrightarrow M/\xi \times M/\eta \equiv X \times Y$ is absolutely continuous, with density $k(x, y)$, with respect to the measure m^* for which:*

1) the canonical projections $m^*\pi_X^{-1}$ and $m^*\pi_Y^{-1}$ coincide with the measures μ and ν;

2) $m/(\xi \wedge \eta)$ coincides with $m^*/(\xi \wedge \eta)$;

3) on almost every element of $\xi \wedge \eta$ the conditional measure coincides with the product of the conditional measures on the corresponding elements of the decompositions ξ_η and η_ξ of X and Y into the preimages of the elements of $M/(\xi \wedge \eta)$ under the canonical mappings $M/\xi \longrightarrow M/(\xi \wedge \eta)$ and $M/\eta \longrightarrow M/(\xi \wedge \eta)$, and the conditional measures on the elements of the decompositions ξ_η and η_ξ are purely continuous.

Let $h(x, y)$ be a measurable function defined on $X \times Y$ and taking values in $[0, 1]$. Then there exists a measurable subset $A \subset \{(x, y): h(x, y) > 0\}$ such that its characteristic function $\chi_A(x, y)$ satisfies

$$\pi_X\chi_A = \pi_X h, \quad \pi_Y\chi_A = \pi_Y h.$$

PROOF. The changes that must be made in the proof of Proposition 44 consist in the following. First, using the results of [92] on the construction of measurable decompositions and the assumption that the conditional measures on the elements of the decompositions ξ_η and η_ξ are purely continuous, we represent the spaces $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ as the respective squares $\{(x, x'): x, x' \in [0, 1]\}$ and $\{(y, y'): y, y' \in [0, 1]\}$, equipped with Lebesgue measure, where ξ_η and η_ξ are the decompositions into the segments $x =$ const and $y =$ const. The space $(X \times Y, \Sigma, m^*)$ is then the four-dimensional unit cube, on the subset $D = \{(x, x', y, y'): x = y\}$ of which the three-dimensional Lebesgue measure m^* is given. This measure m^* plays the role of $\mu \times \nu$ in the proof of Proposition 44. The role of P is played by the set

$$P^* = \tilde{M} \cap (M + (0, \varepsilon', 0, 0)) \cap (M + (0, 0, 0, \varepsilon'')) \cap (M + (0, \varepsilon', 0, \varepsilon'')),$$

and the role of Q is played by a subset Q^* of positive measure for which the four sets

$$Q, \; Q - (0, \varepsilon', 0, 0), \; Q - (0, 0, 0, \varepsilon''), \; Q - (0, \varepsilon', 0, \varepsilon'')$$

are pairwise disjoint. As previously, it is proved that $\chi_R(x, x', y, y')$ cannot be approximated by functions in $L_\xi + L_\eta$, which concludes the proof. ●

7. We have shown that any pair of marginal distributions of a subprobability measure $m_1 \leqslant m$ coincides with the pair of marginal distributions of some measure $\tilde{m}_1$ such that $d\tilde{m}_1/dm$ is the characteristic function χ_A of some measurable subset $A \subset M$. We now determine, in general, the set of pairs of marginal distributions of the measures subject to the condition $m_1 \leqslant m$.

Let

$$K = \{h(x, y): h(x, y) = f(x) + g(y), \; f(x) \in L(X, \mu), \; g(y) \in L(Y, \nu), \\ \|h\|_L \leqslant 1\},$$

and let $\hat{K} \subset K$ be the subset of K consisting of the functions $\hat{h}(x, y)$ of unit L-norm that admit a representation

$$\hat{h}(x, y) = k(\hat{f}(x) + \hat{g}(y)),$$

where $\hat{f}(x)$ and $\hat{g}(y)$ are functions that each take only the values 1 and -1.

PROPOSITION 45. *The convex hull* conv $\hat{K}$ *of* $\hat{K}$ *is dense in* K *with respect to the norm of* $L(M, n)$.

We remark that K is not compact in any topology compatible with the linear structure. This follows, for example, from the fact that it does not have extreme points. In particular, the functions in $\hat{K}$ are not extreme points of K.

PROOF. We prove that for any function $h \in K$, $\|h\|_L = 1$, and any positive number ϵ we can find a convex combination $\Sigma_k p_k \hat{h}_k$ of functions $\hat{h}_k \in \hat{K}$ such that

$$\left\| h - \sum_k p_k \hat{h}_k \right\|_L < \varepsilon.$$

We consider the arbitrary function

$$h(x, y) = f(x) + g(y) \in K, \ \|h\|_L = 1$$

and we choose step functions $\tilde{f}$ and $\tilde{g}$, each with a finite number of values equal to $nd, n = -N, \ldots, N-1, N,$ for which

$$\|f(x) - \tilde{f}(x)\|_{L(X, \mu)} < \frac{\varepsilon}{2}, \quad \|g(y) - \tilde{g}(y)\|_{L(Y, \nu)} < \frac{\varepsilon}{2},$$

so that

$$\|h(x, y) - \tilde{h}(x, y)\|_{L(M, m)} < \varepsilon,$$

where $\tilde{h}(x, y) = \tilde{f}(x) + \tilde{g}(y)$. Furthermore, we can assume that $\|\tilde{h}\|_L \leqslant 1$.

We prove the assertion by representing $\tilde{h}$ in the form of the required convex combination of functions in $\hat{K}$. The functions $\tilde{f}$, $\tilde{g}$, and $\tilde{h}$ are considered as defined exactly on all points of the square, and not only as elements of $L(M, m)$. Let

$$A = \{x : \tilde{f}(x) > 0\} \subset X, \quad B = \{y : \tilde{g}(y) < 0\} \subset Y.$$

We consider the function $\hat{\hat{h}}_1(x, y)$ defined by

$$\hat{\hat{h}}_1(x, y) = \frac{1}{2} d\left((\chi_A(x) - \chi_{CA}(x)) + (\chi_{CB}(y) - \chi_B(y))\right)$$

(here CD is the complement of the set D). Obviously,

$$\hat{\hat{h}}_1(x, y) = \begin{cases} d & \text{for } (x, y) \in A \times CB, \\ 0 & \text{for } (x, y) \in A \times B \text{ and } (x, y) \in CA \times CB, \\ -d & \text{for } (x, y) \in CA \times B. \end{cases}$$

But, by the definition of the sets A and B,

$$\tilde{f}(x) + \tilde{g}(y) \geqslant d \quad \text{for} \quad (x, y) \in A \times CB,$$

and

$$\tilde{f}(x) + \tilde{g}(y) \leqslant -d \quad \text{for} \quad (x, y) \in CA \times B;$$

therefore

$$\|\tilde{f}(x) + \tilde{g}(y) - \hat{\hat{h}}_1(x, y)\|_{L(M, m)} = \|\tilde{f}(x) + \tilde{g}(y)\|_L - \|\hat{\hat{h}}_1(x, y)\|_L.$$

Moreover, the maximum of the function $\tilde{h}(x, y) = \tilde{f}(x) + \tilde{g}(y)$ is obviously attained on the set $A \times CB$, and the minimum on the set $CA \times B$; therefore the maximum of $\tilde{h}$ is decreased after subtraction of $\hat{\hat{h}}_1$ from it, and the minimum is increased by the quantity d. (Only one of the sets $A \times CB$ or $CA \times B$ can be empty, and then either the maximum or the minimum is not changed.)

If we now take $\tilde{h} - \hat{\hat{h}}_1$ instead of $\tilde{h}$ (obviously, $\tilde{h} - \hat{\hat{h}}_1 \in K$ and admits a representation as a difference of step functions $\tilde{f}(x) - \frac{1}{2} d(\chi_A(x) - \chi_{CA}(x))$ and $\tilde{g}(y) - \frac{1}{2} d(\chi_{CB}(y) - \chi_B(y))$, on repeating the above construction through a finite number of steps we arrive at the function identically equal to zero. Finally,

$$\tilde{f}(x) + \tilde{g}(y) = \sum_{k=1}^{n} \hat{h}_k(x, y)$$

and

$$1 \geqslant \|\tilde{f}(x) + \tilde{g}(y)\|_L = \sum_{k=1}^{N} \|\hat{h}_k(x, y)\|_L,$$

from which we get the required representation

$$\tilde{f}(x) + \tilde{g}(y) = \sum_{k=1}^{N} \|\hat{h}_k(x, y)\| \frac{\hat{h}_k(x, y)}{\|\hat{\hat{h}}_k(x, y)\|} = \sum_{k=1}^{N} p_k \hat{h}_k(x, y),$$

where $\Sigma p_k \leqslant 1$ and $\hat{h}_k(x, y) \in \hat{K}$. ●

REMARK. Although we have essentially proved the possibility of uniformly approximating on the square each bounded function h that is representable in the form

$$h(x, y) = f(x) + g(y) \tag{2}$$

by means of linear combinations of functions $\hat{h}$ admitting the same representation and taking only the values -1, 0, and 1 (when $f(x), g(x) = \pm \frac{1}{2}$), not every function in the subset K_∞ of the unit ball in $L^\infty(M, m)$ consisting of the functions representable in the form (2) can be represented in the form of a convex combination of such three-valued functions, as shown by the following finite-dimensional (matrix) example (Figure 2). Here, in determining the L^∞ norm of $\tilde{h}(x, y)$, where (x, y) is an element of a 3×3 matrix, only its values on those elements of the matrix on which these values are written are considered. In this example $\|\tilde{h}\|_{L^\infty(M,m)} = 1$, but $\tilde{h}$ cannot be represented in the form of a convex combination of "three-valued" functions: the functions $f(x)$ and $g(y)$ are determined by $h(x, y)$ to within an additive constant; hence it is easy to verify that the L^∞ norms of $\tilde{f}(x)$ and $\tilde{g}(y)$ in the representation $\tilde{h}(x, y) = \tilde{f}(x) + \tilde{g}(y)$ cannot both be made less than or equal to ½, and, consequently they cannot be represented as convex combinations of functions whose norms equal ½.

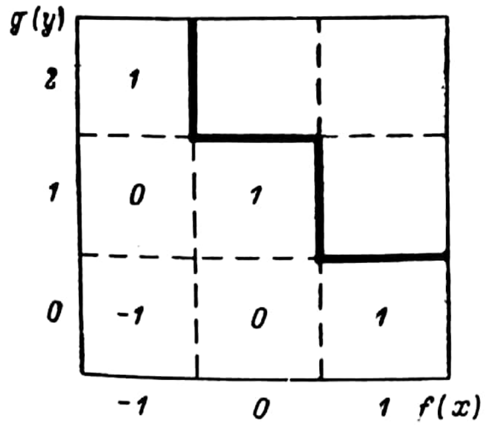

FIGURE 2

However, if the type of m is the type of a product measure $\mu \times \nu$, we can prove the following assertion.

PROPOSITION 46. *Suppose that the measure m is equivalent to the product*

$\mu \times \nu$. *Each function $h(x, y)$ in the set K_∞ can be represented as the barycenter of a measure concentrated on the set $\hat{K}_\infty$ of functions in K_∞ having L^∞ norm equal to 1 and taking only the values -1, 0, and 1, or only the values -1 and 1, and it can be uniformly approximated by convex combinations of such functions. Moreover, if $h(x, y) \geqslant 0$ on some set $Z \subset X \times Y$, then it can be assumed that the same holds for all functions in $\hat{K}_\infty$ used to make up the approximating convex combinations.*

PROPOSITION 47. *If the measure m is equivalent to the product $\mu \times \nu$, then the subspace $L_\xi^\infty + L_\eta^\infty$ is closed in L^∞ ($L_\xi^\infty = \{f\colon\ f = f(x), f \in L^\infty\}$; $L_\eta^\infty = \{g\colon\ g = g(y), g \in L^\infty\}$).*

PROOF. We prove that the continuous mapping

$$s: L_\xi^\infty \times L_\eta^\infty \to L^\infty, \quad s(f(x), g(y)) = f(x) + g(y)$$

is a homomorphism. If $h(x, y) \in L_\xi^\infty + L_\eta^\infty \subset L^\infty$, then

$$\inf_{f(x)+g(y)=h(x,\, y)} \left(\|f\|_{L_\xi^\infty} + \|g\|_{L_\eta^\infty}\right) = \|h\|.$$

Indeed, since

$$\operatorname{ess\,sup} f(x) + \operatorname{ess\,sup} g(y) = \operatorname{ess\,sup}(f(x) + g(y)),$$

and the same is true for ess inf, and since the space $N = s^{-1}(0) \subset L_\xi^\infty \times L_\eta^\infty$ is one-dimensional (consists of the constants), it follows that

$$\begin{aligned} &\operatorname{ess\,sup}(f(x) + g(y)) - \operatorname{ess\,inf}(f(x) + g(y)) \\ &= (\operatorname{ess\,sup} f(x) - \operatorname{ess\,inf} f(x)) + (\operatorname{ess\,sup} g(y) - \operatorname{ess\,inf} g(y)) \end{aligned}$$

and for a function $h(x, y) \in L_\xi^\infty + L_\eta^\infty$ it is possible to choose $f(x) \in L_\xi^\infty$ and $g(y) \in L_\eta^\infty$ such that

$$f(x) + g(y) = h(x, y), \quad \|f\|_{L_\xi^\infty} = \|g\|_{L_\eta^\infty} = \frac{1}{2}\|h\|_{L^\infty}.$$

Therefore, the subspace $L_\xi^\infty + L_\eta^\infty$ of L^∞ is linearly homeomorphic to the Banach space $(L_\xi^\infty \times L_\eta^\infty)/N$ and hence is itself Banach; in particular, closed in L^∞. ●

REMARK. Proposition 47 remains true also if we replace the condition $m \sim \mu \times \nu$ by the condition $m \sim m^*$, where m^* is a measure whose properties are described in Proposition 44 (this is not used in the following).

PROOF OF PROPOSITION 46. It follows from Proposition 47 that K_∞ is a closed subset of L^∞; consequently (the convexity is obvious) it is a compact convex set in the topology $\sigma(L^\infty, L)$. From this we get the first part of the assertion. To prove the second part we consider a function $h(x, y) \in K_\infty$. Let

$$h(x, y) = f(x) + g(y), \quad \|h\|_{L^\infty} = \|f\|_{L_\xi^\infty} + \|g\|_{L_\eta^\infty}.$$

We approximate the functions $f(x)$ and $g(y)$ to within $\epsilon/2$ by functions $\tilde{f}(x)$ and $\tilde{g}(y)$ taking only a finite number of values with step-size ϵ. The function $\tilde{h}(x, y) =$

$\tilde{f}(x) + \tilde{g}(y)$ then approximates $h(x, y)$ uniformly to within ϵ. We can assume that both $\tilde{f}(x)$ and $\tilde{g}(y)$ take more than one value (otherwise the argument is trivial). We now consider the function $\hat{\hat{h}}(x, y) = \hat{\hat{f}}(x) + \hat{\hat{g}}(y)$, where

$$\hat{\hat{f}}(x) = \begin{cases} 1 & \text{for } x \in \{x : \tilde{f}(x) > 0\}, \\ -1 & \text{for } x \in \{x : \tilde{f}(x) \leqslant 0\}, \end{cases}$$

$$\hat{\hat{g}}(y) = \begin{cases} 1 & \text{for } y \in \{y : \tilde{g}(y) \geqslant 0\}, \\ -1 & \text{for } y \in \{y : \tilde{g}(y) < 0\}. \end{cases}$$

Then, arguing as in the proof of Proposition 45, we get that $\hat{\hat{h}} \in K_\infty$, and for some $d > 0$, $d \geqslant \epsilon$,

$$\|\tilde{h}(x, y) - d\hat{\hat{h}}(x, y)\|_{L^\infty} = \|\tilde{h}\|_{L^\infty} - d.$$

Repeating the argument with the function $\tilde{h} - d\hat{\hat{h}}$ and considering that from the start we can assume that $\tilde{f}(x) \geqslant f(x)$ and $\tilde{g}(y) \geqslant g(y)$, whence

$$Z \subset \{(x, y) : \tilde{h}(x, y) \geqslant 0\} = \tilde{Z},$$

and that

$$\{(x, y) : \hat{\hat{f}}(x) + \hat{\hat{g}}(y) < 0\} \cap \tilde{Z} = \varnothing,$$

we arrive at the desired conclusion. ●

THEOREM 6. *Suppose that the subprobability measures $\bar{\mu} \leqslant \mu$ and $\bar{\nu} \leqslant \nu$ are given on $(X, \mathfrak{A})$ and $(Y, \mathfrak{B})$, respectively. For there to exist on (M, Σ) a subprobability measure $\bar{m} \leqslant m$ whose marginal distributions coincide with $\bar{\mu}$ and $\bar{\nu}$ it is necessary and sufficient that, for any measurable coarsenings $\hat{\xi}$ and $\hat{\eta}$ of the decompositions ξ and η consisting each of not more than two elements, there exists a subprobability measure $\hat{\bar{m}}$ on the space $\hat{M} = M/(\hat{\xi}\hat{\eta})$ (which is made up of four elements) having the measures $\bar{\mu}/\hat{\xi}$ and $\bar{\nu}/\hat{\eta}$ as its marginal distributions.*

In other words, the solution of the problem of the existence of a measure $\bar{m}$ with the given marginal distributions $\bar{\mu}$ and $\bar{\nu}$ reduces to clearing up the problem of the existence of a solution of each "coarsened 2×2 problem" arising in the decomposition of each of the spaces X and Y into two measurable subsets, i.e., the problem of the existence of a solution of a finite system of linear inequalities. If the m-measures of the elements of $\hat{\xi}$ are a_1 and a_2, the m-measures of the elements of $\hat{\eta}$ are b_1 and b_2 $(a_1 + a_2 = b_1 + b_2 = 1)$, the $\bar{\mu}$-measures of the subsets X_1 and X_2 into which X is decomposed are $\bar{a}_1$ and $\bar{a}_2$, and the $\bar{\nu}$-measures of the subsets Y_1 and Y_2 into which Y is decomposed are $\bar{b}_1$ and $\bar{b}_2$, and if m_{ik} $(i, k = 1, 2$; the first index relates to $\hat{\eta}$ and the second to $\hat{\xi})$ are the m-measures of the elements M_{ik} of $\hat{\xi}\hat{\eta}$, i.e. the masses of the elements of $(\hat{M}, \hat{m})$, where $\hat{m} = m/(\hat{\xi}\hat{\eta})$, then the aforementioned system of inequalities to be satisfied by the values $\hat{\bar{m}}_{ik} = \bar{m}M_{ik}$ is

$$\begin{aligned} m_{11} &\geq \hat{m}_{11} \geq 0, & \hat{m}_{11}+\hat{m}_{21} &= \bar{a}_1, \\ m_{21} &\geq \hat{m}_{21} \geq 0, & \hat{m}_{12}+\hat{m}_{22} &= \bar{a}_2, \\ m_{12} &\geq \hat{m}_{12} \geq 0, & \hat{m}_{11}+\hat{m}_{12} &= \bar{b}_1, \\ m_{22} &\geq \hat{m}_{22} \geq 0, & \hat{m}_{21}+\hat{m}_{22} &= \bar{b}_2. \end{aligned}$$

It can be shown that a necessary and sufficient condition for the solution of this system is the fulfillment of the relations $\bar{a}_i + \bar{b}_i \leqslant 1 + m_{ki}$, $i, k = 1, 2$; but here we do not need the concrete form of these conditions.

In the following it is convenient for us to work, not with the set of subprobability measures majorized by the measure m on M, or, what is the same, the set V of measurable functions taking values in $[0, 1]$ (the densities of the subprobability measures majorized by m), but with the set W of measurable functions taking values in $[-1, 1]$, i.e., with the unit ball of $L^\infty(M, m)$. Instead of marginal distributions we now deal with projections: the functions $\pi_X h$ and $\pi_Y h$, $h \in \mathrm{W}$. For the existence of a function $h \in \mathrm{W}$ with given projections $q(x)$ and $r(y)$ it is necessary and sufficient that there exists a measure that is bounded by the measure m and for which the densities of the marginal distributions with respect to μ and ν are equal to $(q(x) + 1)/2$ and $(r(y) + 1)/2$, respectively; moreover, each coarsened 2×2 problem is solvable or not solvable simultaneously both for the problem of finding a function in W with projections $q(x)$ and $r(y)$ and for that of finding a function in V with projections $(q(x) + 1)/2$ and $(r(y) + 1)/2$. Therefore, the following assertion is an equivalent reformulation of Theorem 6.

THEOREM 6 (second formulation). *Let the functions $q(x)$ and $r(y)$ be given. For there to exist a function $h(x, y) \in \mathrm{W}$ for which $\pi_X h = q$ and $\pi_Y h = r$ it is necessary and sufficient that for any $\hat{\xi}$ and $\hat{\eta}$ consisting each of two elements that is a function $\hat{h} \in \mathrm{W}(\hat{M})$ for which $\pi_{\hat{X}} \hat{h} = \hat{\varphi} q$ and $\pi_{\hat{Y}} \hat{h} = \hat{\psi} r$.*

(We recall that $\hat{\varphi}$ and $\hat{\psi}$ are the canonical homomorphisms $X \longrightarrow \hat{X} = M/\hat{\xi}$ and $Y \longrightarrow \hat{Y} = M/\hat{\eta}$, respectively, extended to integrable functions that are regarded as the Radon-Nikodým derivatives of certain distributions with respect to the measures μ and ν.)

PROOF OF THEOREM 6. We prove the theorem in the second formulation. The space $L^\infty(X, \mu) \times L^\infty(Y, \nu)$ is regarded as the dual space of $L(X, \mu) \times L(Y, \nu)$. Each functional $(u, v) \in L^\infty(X, \mu) \times L^\infty(Y, \nu)$ acts on an element $(f, g) \in L(X, \mu) \times L(Y, \nu)$ by the formula

$$\langle (f, g), (u, v) \rangle_{\mathrm{IV}} = \langle f, u \rangle_{\mathrm{II}} + \langle g, v \rangle_{\mathrm{III}} = \int_X f(x) u(x)\, d\mu + \int_Y g(y) v(y)\, d\nu.$$

Together with the mapping

$$\pi = \pi_X \times \pi_Y, \quad \pi : L^\infty(M, m) \to L^\infty(X, \mu) \times L^\infty(Y, \nu)$$

we consider the adjoint mapping π^*:

$$\begin{array}{ccc} W\subset L^\infty(M,\ m) & \xrightarrow{\pi} & L^\infty(X,\ \mu)\times L^\infty(Y,\ \nu)\supset \pi W, \\ \mathbf{R}\xleftarrow{\mathrm{I}} \langle & \rangle\xrightarrow{\mathrm{II}}\mathbf{R} & \rangle\xrightarrow{\mathrm{III}}\mathbf{R}\quad \rangle\xrightarrow{\mathrm{IV}}\mathbf{R} \\ W^\circ\subset L(M,\ m) & \xleftarrow{\pi^*} & L(X,\ \mu)\times L(Y,\ \nu)\supset(\pi W)^\circ. \end{array}$$

The mapping π is continuous from the topology $\sigma(L^\infty(M, m), L(M, m))$ into the topology $\sigma(L^\infty(X, \mu) \times L^\infty(Y, \nu), L(X, \mu) \times L(Y, \nu))$ (the notation is from [18]), since in these weak topologies the mappings π_X and π_Y are continuous. The unit ball W of the dual space is obviously compact [12] in the topology $\sigma(L^\infty(M, m), L(M, m))$; therefore its image in $L^\infty(X, \mu) \times L^\infty(Y, \nu)$ is weakly compact, and to verify the inclusion $(u, v) \in \pi W$ it suffices to verify that $|\langle (f, g), (u, v)\rangle_{\mathrm{IV}}| \leqslant 1$ for any element (f, g) of the polar $(\pi W)^\circ$ of the set πW. But

$$\langle (f,\ g),\ (u,\ v)\rangle_{\mathrm{IV}}=\langle \pi^*(f,\ g),\ h\rangle_{\mathrm{I}},$$

where h is an arbitrary element of $\pi^{-1}(u, v)$, and if $(u, v) \in \pi W$, then we can assume that $\pi^{-1}(u, v) \ni h \in W$; therefore the element $(f, g) \in L(X, \mu) \times L(Y, \nu)$ belongs to $(\pi W)^\circ$ if and only if its image $\pi^*(f, g)$ belongs to the polar W° of W, i.e., if and only if $\|\pi^*(f, g)\|_{L(M,m)} \leqslant 1$ (W is the unit ball of $L^\infty(M, m)$). In fact

$$\pi^*(L(X,\ \mu)\times L(Y,\ \nu))=(\pi^{-1}(0))^\circ.$$

Since

$$\langle \pi^*(f,\ g),\ h\rangle=\int_X f(x)\,\pi_X h\,d\mu+\int_Y g(y)\,\pi_Y h\,d\nu=\int_M (f(x)+g(y))\,h(x,\ y)\,dm,$$

we get that $(\pi W)^\circ$ consists of those pairs (f, g) for which

$$\|f(x)+g(y)\|_{L(M,\ m)}\leqslant 1.$$

To verify that the pair $(q(x), r(y))$ belongs to the image πW of W it suffices to verify that, for any two integrable functions $f(x)$ and $g(y)$ such that

$$\int_M |f(x)+g(y)|\,dm\leqslant 1,$$

we have

$$\left|\int_{M/\xi} f(x)\,q(x)\,d\mu+\int_{M/\eta} g(y)\,r(y)\,d\nu\right|\leqslant 1. \tag{3}$$

As is well known, the polar of any set coincides with the polar of its closed convex hull. By Proposition 45, a dense subset of $\pi^*(\pi W)^\circ = K$ is formed by the convex hull of the set $\hat{K}$ of functions $\hat{h}(x, y)$ that are representable in the form $\hat{h}(x, y) = \hat{f}(x) + \hat{g}(y)$, where $\|\hat{h}(x, y)\|_{L(M,m)} \leqslant 1$ and the functions $\hat{f}(x)$ and $\hat{g}(y)$ take two values. Norm density implies density in the weak topology, so it suffices to verify (3) only for pairs (f, g) for which $f(x) + g(y) \in \hat{K}$. But the consideration of all functions f and g, each taking two values on two fixed subsets of M/ξ and M/η, respectively, is equivalent to the consideration of all functions $\hat{f}$ and $\hat{g}$ in $L(\hat{X}, \hat{\mu})$ and $L(\hat{Y}, \hat{\nu})$ satisfying the condition

$$\|\hat{f}(\hat{x}) + \hat{g}(\hat{y})\|_{L(\hat{M}, \hat{m})} \leqslant 1,$$

and the satisfaction of (3) for such functions $\hat{f}(x)$ and $\hat{g}(y)$ is equivalent to the satisfaction of the condition

$$\left| \int_{\hat{X}} \hat{f}(\hat{x}) (\hat{\varphi} q)(\hat{x})\, d\hat{\mu} + \int_{\hat{Y}} \hat{g}(\hat{y}) (\hat{\psi} r)(\hat{y})\, d\hat{\nu} \right| \leqslant 1,$$

which, according to the above, is exactly equivalent to the positive solution of the coarsened 2×2 problem corresponding to the fixed decompositions $\hat{\xi}$ and $\hat{\eta}$. But the satisfaction of (3) for any $f(x)$ and $g(y)$ for which $f(x) + g(y) \in \hat{K}$ means the solvability of any coarsened 2×2 problem, i.e., it means the existence of a function $h(x, y) \in W$ for which $\pi_X h = q(x)$ and $\pi_Y h = r(y)$. ●

REMARK. The assertion of the theorem remains true, and the proof can be completely retained, when the requirement that m is a probability measure is replaced by the requirement that it is nonnegative and σ-finite, and that the corresponding marginal distributions $\mu = m\pi_X^{-1}$ and $\nu = m\pi_Y^{-1}$ are σ-finite (i.e., the requirement that any measurable subset of (X, μ) and (Y, ν) can be represented as the union of not more than countably many pairwise disjoint sets of finite measure). The term "subprobability" in such a strengthened formulation must be omitted. Below (Proposition 75 in §11.5), Theorem 6 is carried over to the case of a σ-finite measure (without the assumption of σ-finiteness of its marginal distributions).

8. Let $\{X_n, n = 1, \dots\}$ and $\{Y_m, m = 1, \dots\}$ be bases [92] for the respective measure spaces $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$. When necessary, we assume, without special mention and without introducing additional notation, that functions on X and on Y can also be regarded as functions on $X \times Y$ (depending only on one argument: $f(x) = f(x, y)$). For the proof of the fact that some decomposition that is independent of ξ and η is actually an independent complement, we shall use the following criterion.

PROPOSITION 48. *For the doubly stochastic measure m to be the kernel of an isomorphism between the spaces $(M/\xi, m/\xi)$ and $(M/\eta, m/\eta)$ (i.e., to be a typical element of the complement, equipped with its conditional measure) it is necessary and sufficient that for the subsets of the bases $\{X_n, n = 1, \dots\}$ and $\{Y_m, m = 1, \dots\}$ the functions $\pi_X[\chi_{X \times Y_m}(x, y)](x)$ and $\pi_Y[\chi_{X_n \times Y}(x, y)](y)$ coincide (mod 0) with the characteristic functions of some sets $\bar{X}_m \subset X$ and $\bar{Y}_n \subset Y$.*

PROOF. The necessity is clear, since if the element considered is the graph of the isomorphism T: $(X, \mu) \longrightarrow (Y, \nu)$, then

$$\pi_Y[\chi_{X_n}(x)](y) = \chi_{TX_n}(y) \quad \text{and} \quad \pi_X[\chi_{Y_n}(y)](x) = \chi_{T^{-1}Y_n}(x).$$

Conversely, if this condition holds, then the transformation $\pi_Y[\chi_{X_n}]$ establishes a measure-preserving correspondence between the basis $\{X_n\}$ and some system $\{Y'_n\}$

of subsets of $(Y, \mathfrak{B}, \nu)$, i.e., generates a measure-preserving measurable mapping γ: $(Y, \mathfrak{B}, \nu) \longrightarrow (X, \mathfrak{A}, \mu)$, and, similarly, the transformation $\pi_X[\chi_{Y_n}]$ generates a measure-preserving mapping β: $(X, \mathfrak{A}, \mu) \longrightarrow (Y, \mathfrak{B}, \nu)$. These transformations can be described in the following way. For definiteness, we consider the mappings $\pi_X[\chi_{Y_m}]$. Let η_k be the decomposition of Y into a finite number of sets representable in the form of an intersection $Y'_1 \cap \cdots \cap Y'_k$, where $Y'_m = Y_m$ or $Y'_m = Y \backslash Y_m$. Since $\{Y_m\}$ is a basis, $\eta_k \nearrow \epsilon_Y$ (the decomposition into points). The decomposition η corresponds to some decomposition ξ_k of X. Indeed,

$$\pi_X[\chi_{Y'_m}] = \bar{X}'_m,$$

where $\bar{X}'_m = X \backslash \bar{X}_m$ if $Y'_m = Y \backslash Y_m$, and

$$\pi_X[\chi_{Y'_1 \cap \ldots \cap Y'_k}] = \bar{X}'_1 \cap \ldots \cap \bar{X}'_k,$$

because the condition in Proposition 48 means that the set $\pi_Y^{-1} Y'_m$ is ξ-measurable and $\pi_X \pi_Y^{-1} Y'_m = \bar{X}'_m$; consequently, the intersection $\pi_Y^{-1} Y'_1 \cap \cdots \cap \pi_Y^{-1} Y'_k$ is also ξ-measurable, and all possible such ξ-subsets are measurable with respect to some coarsening ξ_k of ξ generated by all possible subsets of the form $\bar{X}'_1 \cap \cdots \cap \bar{X}'_k$. Letting k now go to infinity, we arrive at a limit decomposition $\bar{\xi} = \lim \nearrow \xi_k$, so that to each point $y \in Y$ there corresponds one (and only one) element of the decomposition $\bar{\xi}$. This correspondence is given by a graph on $X \times Y$ (assign to each point $x \in X$ the element $y = \beta(x) \in Y$ for which the point x belongs to the corresponding element of $\bar{\xi}$). We now equip the graph Γ of this mapping with the measure $\mu \pi_X^{-1}|_\Gamma$. It is easy to verify that the space $(\Gamma, \mu \pi_X^{-1}|_\Gamma)$ coincides (mod 0) with that space with doubly stochastic measure m with respect to which we began constructing the mapping β. Indeed, $(\Gamma, \mu \pi_X^{-1}|_\Gamma)$ satisfies the conditions of Proposition 48, and it is possible to begin the construction with it. But, since the measure $\mu \pi_X^{-1}$ is, by construction, concentrated on the graph of some mapping β: $X \longrightarrow Y$, it is clear that, by carrying out the same construction, we return to the measure $\mu \pi_X^{-1}|_\Gamma$. On the other hand, our construction determines the conditional measures of the original measure with respect to the decomposition η, i.e., it determines this original measure uniquely. Consequently, the original doubly stochastic measure m coincides with $\mu \pi_X^{-1}|_\Gamma$. The spaces X and Y are completely equivalent, from which it follows that the homomorphism β of the measure spaces X and Y is invertible, i.e., is an automorphism, and the measure m really is concentrated on the graph of an isomorphism. ●

We have another auxiliary result.

PROPOSITION 49. *Let $m \ll \mu \times \nu$ and $dm/d(\mu \times \nu) = k(x, y)$. If among the subsets of the set M of the form $A \times B$, where $A \in \mathfrak{A}$, $B \in \mathfrak{B}$, $\mu A + \nu B \geqslant c$ and μA, $\nu B \geqslant \epsilon_0$, there are sets of arbitrarily small m-measure, then there is a set $A_0 \times B_0$ with the same properties for which $m(A_0 \times B_0) = 0$.*

PROOF. Let A_n and B_n, $n = 1, \ldots,$ be such that $\mu A_n \geqslant \epsilon_0$, $\nu B_n \geqslant \epsilon_0$, $\mu A_n + \nu B_n = c$, and $m(A_n \times B_n) \longrightarrow 0$. The families of functions $\{\chi_{A_n}(x)\}$ and $\{\chi_{B_n}(y)\}$ are bounded, and hence precompact, in the topologies $\sigma(L^\infty(X, \mu), L(X, \mu))$ and

$\sigma(L^\infty(Y,\nu), L(Y,\nu))$, respectively. We therefore assume that in these topologies $\chi_{A_n}(\cdot) \to a(\cdot) \in L^\infty(X,\mu)$ and $\chi_{B_n}(\cdot) \to b(\cdot) \in L^\infty(Y,\nu)$. From the fact that for arbitrary measurable subsets $A \subset X$ and $B \subset Y$ we have the convergence

$$\langle \chi_{A_n}, \chi_A \rangle \to \langle a, \chi_A \rangle \text{ and } \langle \chi_{B_n}, \chi_B \rangle \to \langle b, \chi_B \rangle,$$

it follows that $0 \leqslant a(x) \leqslant 1$ and $0 \leqslant b(y) \leqslant 1$. We prove that for arbitrary measurable characteristic functions χ_A and χ_B

$$\int_M \chi_{A_n}(x)\chi_{B_n}(y)\chi_A(x)\chi_B(y)k(x,y)\,d(\mu\times\nu)$$
$$\to \int_M a(x)b(y)\chi_A(x)\chi_B(y)k(x,y)\,d(\mu\times\nu).$$

Indeed,

$$\left|\iint \chi_{A_n}\chi_{B_n}\chi_A\chi_B k(x,y)\,dxdy - \iint ab\chi_A\chi_B k(x,y)\,dxdy\right|$$
$$\leqslant \left|\int_A \chi_{A_n}(x)\,dx \int_B (\chi_{B_n}(y)-b(y))k(x,y)\,dy\right|$$
$$+\left|\int_B b(y)\,dy\int_A (\chi_{A_n}(x)-a(x))k(x,y)\,dx\right|.$$

Let

$$\psi_n(x) = \int_B (\chi_{B_n}(y)-b(y))k(x,y)\,dy.$$

For almost every fixed $x \in X$

$$\int_B k(x,y)\,dy \leqslant 1$$

(double stochasticity of the kernel k), i.e., $k(x,y) \in L(Y,\nu)$, and hence $\psi_n(x) \to 0$. Moreover,

$$|\psi_n(x)| \leqslant 2\int_B k(x,y)\,dy \leqslant 2.$$

From this we get that

$$\left|\int_A \chi_{A_n}(x)\psi_n(x)\,dx\right| \leqslant \int_A |\psi_n(x)|\,dx \to 0.$$

In an analogous way we consider also the second term. From the convergence

$$\langle \chi_{A_n}\chi_{B_n}, \chi_A\chi_B\rangle \to \langle ab, \chi_A\chi_B\rangle$$

just proved, we immediately get the convergence

$$\chi_{A_n}(x)\chi_{B_n}(y) \to a(x)b(y)$$

in the topology $\sigma(L^\infty(M,m), L(M,m))$, since the linear span of the set $\{\chi_A\chi_B\}$ is dense in $L(M, m)$, and the sequence $\chi_{A_n}\chi_{B_n}$ is bounded in $L^\infty(M, m)$. From the fact that $0 \leqslant a(x) \leqslant 1$, $0 \leqslant b(y) \leqslant 1$ and

$$\int a(x)\, d\mu = \lim \int \chi_{A_n}(x)\, d\mu, \quad \int b(y)\, d\nu = \lim \int \chi_{B_n}(y)\, d\nu,$$

it follows that

$$\mu\{x : a(x) > 0\} + \nu\{y : b(y) > 0\} \geqslant c.$$

The condition $m(A_n \times B_n) \to 0$ means that

$$a(x)\, b(y) = \lim \chi_{A_n \times B_n} = 0 \pmod m.$$

But since $a(x)b(y) \equiv 0$, we have $k(x, y) = 0$ on the set $\{(x, y)\colon\ a(x) > 0,\ b(y) > 0\}$, which proves Proposition 49. ●

REMARK 1. From the proof it is clear that $A_0 = \{x\colon\ a(x) > 0\}$ and $B_0 = \{y\colon\ b(y) > 0\}$, where $(a(x), b(y))$ is a limit point of the set $\{(\chi_{A_n}(x), \chi_{B_n}(y))\}$. In particular, if it is known beforehand that the condition $\mu A + \nu B > c$ implies $m(A \times B) > 0$, then, necessarily,

$$a(x) = \chi_{A_0}(x) \text{ and } b(y) = \chi_{B_0}(y);$$

this is because

$$\int a(x)\, d\mu + \int b(y)\, d\nu = \lim (\mu A_n + \nu B_n) \geqslant c, \ a(x) \leqslant 1, \ b(y) \leqslant 1,$$

and, consequently,

$$\mu A_0 \geqslant \int a(x)\, d\mu, \quad \nu B_0 \geqslant \int b(y)\, d\nu,$$

wherefore

$$c \geqslant \mu A_0 + \nu B_0 \geqslant \int a(x)\, d\mu + \int b(y)\, d\nu,$$

from which it follows that

$$\int a(x)\, d\mu = \mu A_0, \quad \int b(y)\, d\nu = \nu B_0,$$

i.e., $a(x) = \chi_{A_0}(x)$ and $b(y) = \chi_{B_0}(y)$.

REMARK 2. We mention an especially useful assertion: if $\chi_{A_n}(x) \to a(x)$ and $\chi_{B_n}(y) \to b(y)$, then

$$\chi_{A_n \times B_n}(x, y) \to a(x)\, b(y).$$

9. DEFINITION. We use the notation

$$\Pi = \Pi_m = \sup\{\mu A + \nu B : m(A \times B) = 0,\ \mu A > 0,\ \nu B > 0\}.$$

PROPOSITION 50. *Let $k(x, y) \leqslant K < \infty$ and $\Pi_m = 1$. Then there exist finite decompositions*

$$X = X_1 \cup X_2 \cup \ldots \cup X_n, \quad Y = Y_1 \cup Y_2 \cup \ldots \cup Y_n$$

of the spaces X and Y, where $2 \leqslant n \leqslant K$, $\mu X_1 = \nu Y_1, \ldots, \mu X_n = \nu Y_n$, *such that*

$$\sum_1^n m(X_k \times Y_k) = 1$$

and for each subset $X_k \times Y_k$, *regarded as the subspace* $(X_k \times Y_k, m_k)$ *of the space* (M, m), *we have*

$$\Pi_{m_k} < 1 \text{ and } \xi \wedge \eta = \nu$$

(the unique nonempty subset of $X_k \times Y_k$ *that is measurable both with respect to* ξ *and with respect to* η *coincides with the whole space* $X_k \times Y_k$*).*

PROOF. First of all, we observe that for doubly stochastic measures we always have $\Pi_m \leqslant 1$. We construct a sequence of sets $A_k \times B_k$ for which

$$\mu A_k + \nu B_k \to 1,\ m(A_k \times B_k) = 0.$$

Let $A'_k \supset A_k$ and $B'_k \supset B_k$ be sets such that $\mu A'_k + \nu B'_k = 1$. Obviously,

$$m(A'_k \times B'_k) \leqslant m(A_k \times B_k) + m(A'_k \setminus A_k \times Y) + m(X \times B'_k \setminus B_k)$$
$$= \mu(A'_k \setminus A) + \nu(B'_k \setminus B) \to 0.$$

Moreover, if $m(A_k \times B_k) = 0$, then $m(A_k \times CB_k) = mA$, from which it follows that the supremum of the density $k(x, y)$ on the set $A_k \times CB_k$ is not less than

$$\frac{\mu A^k}{\mu A_k \cdot \nu CB_k} = \frac{1}{\nu CB_k} = \frac{1}{1 - \nu B_k} = \frac{1}{\mu A_k + (1 - (\mu A_k + \nu B_k))}$$

and, therefore, for sufficiently large k it turns out that

$$\mu A'_k \geqslant \mu A_k \geqslant \frac{1}{K} - \varepsilon,$$

for any $\epsilon > 0$, and a similar inequality holds for $\nu B'_k$. Thus, the conditions of Proposition 49 have been verified, so there exist sets $A \subset X$, $\mu A \geqslant K^{-1}$, and $B \subset Y$, $\nu B \geqslant K^{-1}$, such that

$$\mu A + \nu B = 1,\ m(A \times B) = 0.$$

From this it follows immediately that

$$m(CA \times CB) = 0, \quad \text{i.e.} \quad m(A \times CB) + m(CA \times B) = 1.$$

Considering $A \times CB$ and $CA \times B$ now as subspaces of (M, m) in the case when $\Pi = 1$ for any of these subspaces, we repeat the argument, dividing it again into two subspaces, where each time the measures of the projections of each subspaces onto X and Y are not less than K^{-1}. After not more than $[K] - 1$ steps we arrive at the required decomposition of M. For each component of this decomposition (which is an element of the decomposition $\xi \wedge \eta$) we have that $\Pi < 1$, since otherwise we could continue the decompositions, and this is impossible by the condition $\mu A \geqslant K^{-1}$. ●

REMARK. In a completely analogous way it can be shown that *if m is an* *arbitrary measure that is absolutely continuous with respect to the measure $\mu \times \nu$, then the decomposition $\xi \wedge \eta$ of $(X \times Y, m)$ is not more than countable.* Indeed, if, for example, the marginal distribution under the decomposition ξ_η for which $\pi_X^{-1}\xi_\eta = \xi \wedge \eta$, i.e., the measure on $(X, \mu)/\xi_\eta$, could have a continuous component, then the type of the conditional measures on the elements of the decomposition η corresponding to those points $y \in Y$ for which $\pi_X \pi_Y^{-1} y$ coincides with points in X belonging to a "continuous" set of the decomposition ξ_η could not be subordinate to the type of the measure $\mu\pi_X$ (since it would coincide with the type of μ_C, where μ_C is the conditional measure on the corresponding element C of ξ_η).

PROPOSITION 51. *Suppose that the decompositions $\hat{\xi}$ and $\hat{\eta}$ of the space $(M, \mathfrak{A}, m)$ are such that $\hat{X} = M/\hat{\xi} = \{\hat{x}_1, \hat{x}_2\}$, $\hat{Y} = M/\hat{\eta} = \{\hat{y}_1, \hat{y}_2\}$, $\hat{M} = \hat{X} \times \hat{Y}$, $\hat{\mu}\{\hat{x}_1\} = a$, $\hat{\nu}\{\hat{y}_1\} = b$ (where $\hat{\mu} = \mu\pi_{\hat{X}}^{-1}$, $\hat{\nu} = \nu\pi_{\hat{Y}}^{-1}$), and $a + b = c < 1$. For any measure $\hat{n}$ defined on $\hat{M}$ and equal to zero on the set $\{(\hat{x}_1, \hat{y}_1), (\hat{x}_2, \hat{y}_2)\} \subset \hat{M}$, the sum of the distances (with respect to variation) of the canonical projections $\hat{n}\hat{\pi}_{\hat{X}}^{-1}$ and $\hat{n}\hat{\pi}_{\hat{Y}}^{-1}$ of this measure from the respective measures $\hat{\mu}$ and $\hat{\nu}$ is not less than $2(1-c)$:*

$$\inf_{\{\hat{n}:\, \hat{n}\,\{(\hat{x}_1, \hat{y}_1), (\hat{x}_2, \hat{y}_2)\} = 0\}} (\operatorname{Var}(\hat{\mu} - \hat{n}\hat{\pi}_{\hat{X}}^{-1}) + \operatorname{Var}(\hat{\nu} - \hat{n}\hat{\pi}_{\hat{Y}}^{-1})) = 2(1-c) > 0$$

(here $\pi_{\hat{X}}$ is the canonical projection $M \to \hat{X}$, $\hat{\pi}_{\hat{X}}$ is the canonical projection $\hat{M} \to \hat{X}$, and similarly for $\pi_{\hat{Y}}$ and $\hat{\pi}_{\hat{Y}}$).

PROOF. Let $\hat{n}\{(\hat{x}_2, \hat{y}_1)\} = p$ and $\hat{n}\{(\hat{x}_1, \hat{y}_2)\} = q$. The quantity $|(1-a) - p| + |a - q| + |(1-b) - q| + |b - p|$ is to be estimated. Obviously,

$$|1-a-p| + |b-p| = \begin{cases} 1-a+b-2p & \text{for } p < \min(1-a,\ b), \\ 1-a-b = 1-c & \text{for } p \in [\min(1-a,\ b),\ \max(1-a,\ b)], \\ 2p-1+a-b & \text{for } p > \max(1-a,\ b) \end{cases}$$

and analogously for $|a-q| + |(1-b)-q|$, from which we get the desired estimate, which is attained for

$$p \in [\min(1-a,\ b),\ \max(1-a,\ b)],$$
$$q \in [\min(1-b,\ a),\ \max(1-b,\ a)]. \quad \bullet$$

COROLLARY. *For any number $\lambda > 0$ and any measure $\hat{\nu}$*

$$\operatorname{Var}(\lambda\hat{\mu} - \hat{n}\hat{\pi}_{\hat{X}}^{-1}) + \operatorname{Var}(\lambda\hat{\nu} - \hat{n}\hat{\pi}_{\hat{Y}}^{-1}) \geqslant 2\lambda(1-c).$$

10. We proceed to the proof of the approximation theorem.

THEOREM 7. *Suppose that the density $k(x, y)$ of a doubly stochastic measure on $X \times Y$, where (X, μ) and (Y, ν) are measure spaces, is bounded: $k(x, y) \leqslant K < \infty$. There exists a number $s_0 > 0$ such that for each $\epsilon > 0$ there is a number $\delta_0 > 0$ for which, given a subset $\Lambda \subset M$ whose widths with respect to the decompositions ξ and η do not exceed some number $s \leqslant s_0$ and such that $\|\pi_X \chi_\Lambda - s\|_{L(X,\mu)} + \|\pi_Y \chi_\Lambda - s\|_{L(Y,\nu)} \leqslant \delta s$ for some $\delta \leqslant \delta_0$, there exists a subset $\Lambda_1 \supset \Lambda$ having constant width not exceeding $(1+\epsilon)s$.*

In other words, if a "strip" that is sufficiently narrow with respect to both decompositions has relatively almost constant width with respect to each of them, then it is contained in a strip whose width with respect to these decompositions is constant and does not much exceed the maximum of the widths of the original set.

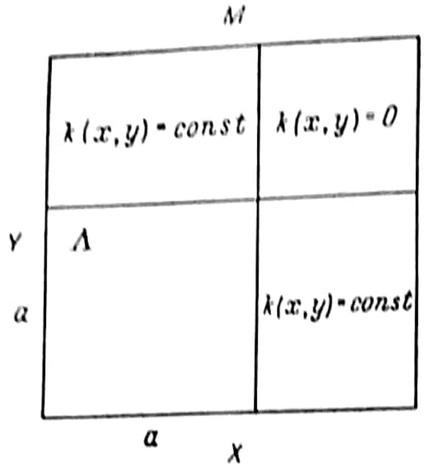

FIGURE 3

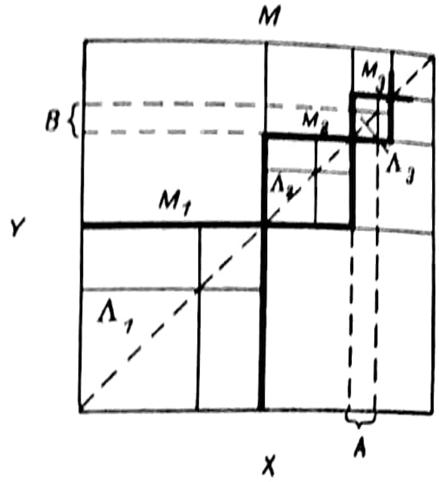

FIGURE 4

The example in Figure 3 of a measure on a square shows that there exist measure spaces and subsets Λ of them whose maximum width with respect to each of the coordinate decompositions ξ and η is arbitrarily small and that are not contained in any set of constant width besides the whole space M.

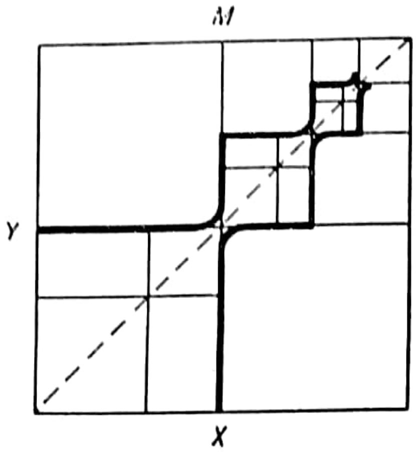

FIGURE 5

The essentialness of the hypothesis about boundedness of the density in the approximation theorem is illustrated by the measure on the square shown in Figure 4, where the squares M_n, regarded as subspaces of (M, m), are each constructed as the space shown in Figure 3, with widths of the subsets Λ_n that unboundedly approach zero. Also the assumption of boundedness for the density cannot be replaced by the condition $\xi \wedge \eta = \nu$ (which does not hold for the measure in Figure 4), as shown by the example of the measure in Figure 5, where $\xi \wedge \eta = \nu$, but the pathologically bad approximability of certain subsets of arbitrarily small width is preserved as before.

The example in Figure 4 shows also that without additional assumptions we cannot hope, generally speaking, to construct economically strips of constant width and subsets contained in the product of two sets of small measure. It suffices to consider the same subset Λ.

The conclusion of the approximation theorem is trivially not guaranteed for doubly stochastic measures that are not absolutely continuous with respect to the product $\mu \times \nu$.

PROOF. As Proposition 50 shows, it suffices to limit ourselves to the case when $\xi \wedge \eta = \nu$, i.e., when $\Pi_m < 1$; in the contrary case we apply the theorem to each of the finite number of subspaces of the form $A_k \times B_k$ for which $\Sigma_k m(A_k \times B_k) = 1$ and $\Pi_{m_k} < 1$. If $\Pi_m < 1$, then, by Proposition 49, all the nondegenerate sets of the form $A \times B$ for which $\mu A > \alpha$, $\nu B > \alpha$, and the "semiperimeter" $\mu A + \nu B$ is equal to a fixed number t have measure that is not less than some number $n_\alpha(t)$, where $n_\alpha(t) \geqslant m_0 > 0$ for $t > t_0 = \text{const} < 1$.

We show that the theorem is proved if we prove that for any decompositions $\hat{\xi}$ and $\hat{\eta}$, each containing two subsets, the image under the canonical mapping $L^\infty(M, m) \to L^\infty(\hat{M}, \hat{m})$ of the characteristic function χ_Λ of the set Λ having the properties indicated in the hypotheses of the theorem can be majorized by a function of constant width $s_1 = (1+\epsilon)s$ in the class $V(\hat{M}, \hat{m})$. Indeed, it suffices for us to show that among the functions in $V(M, m)$ vanishing on Λ there is a function of constant width equal to $1 - s_1$. Then, by Proposition 44, there is also a subset disjoint from Λ and having constant width equal to $1 - s_1$. Its complement is a strip of constant width s_1 containing Λ, and is the set whose existence is asserted in the theorem. By Theorem 6, to prove the existence of a function of width $1 - s_1$ that vanishes on Λ it suffices to prove the solvability of all 2×2 coarsenings of the problem on the existence of the required function on the space M, equipped with the measure that is the restriction of m to the subset $M \backslash \Lambda$. But if it is possible each time to majorize the image of the characteristic function $\chi_\Lambda(x, y)$ (under the coarsening of the problem corresponding to the choice of decompositions $\hat{\xi}$ and $\hat{\eta}$) by a function of constant width s_1 in $V(\hat{M}, \hat{m})$, then for the complement $C\Lambda$ of the strip Λ there exists for each of $\hat{\xi}$ and $\hat{\eta}$ a function of constant width $1 - s_1$.

We consider a pair of decompositions $\hat{\xi}$ and $\hat{\eta}$. Let m'_{ik}, $i, k = 1,2$, be the measures of the subsets of $\Lambda \subset M$ falling in the corresponding "cells" (elements) of $\hat{\xi}\hat{\eta}$ (we recall that the first index relates to $\hat{\eta}$ and the second to $\hat{\xi}$). We also use the notation

$$h_{ik} = \frac{m'_{ik}}{a_k b_i}, \quad k_{ik} = \frac{m_{ik}}{a_k b_i}.$$

The numbers $p_{ik} = m'_{ik}/m_{ik}$ are the values on $\hat{M}$ of the 2×2 coarsening $\hat{\chi}_\Lambda$ of the characteristic function of Λ (i.e., the measures of the elements of $\hat{\xi}\hat{\eta}$ relative to the measure whose Radon-Nikodým derivative with respect to the measure is equal to the characteristic function of Λ). If the widths of Λ with respect to ξ and η do not exceed the number s and are close to constants in the metric of $L(X, \mu)$ and of $L(Y, \nu)$, respectively, then the widths of the indicated 2×2 coarsenings of the function $\chi_\Lambda(x, y)$ with respect to the coordinate decompositions, which are equal, as is easily checked, to

$$(s_1, s_2) = \left(\frac{m'_{11} + m'_{12}}{m_{11} + m_{12}}, \frac{m'_{21} + m'_{22}}{m_{21} + m_{22}}\right) \text{ on the elements of } \hat{X} = M/\hat{\xi},$$

$$(s_3, s_4) = \left(\frac{m'_{11} + m'_{21}}{m_{11} + m_{21}}, \frac{m'_{12} + m'_{22}}{m_{12} + m_{22}}\right) \text{ on the elements of } \hat{Y} = M/\hat{\eta},$$

do not exceed (coordinatewise) the same number s and differ by an arbitrarily small amount (more precisely, not by a greater amount) from the vector $((s, s), (s, s))$ in the metric of the four-dimensional space L.

We now attempt to "complete" in a suitable way the function $\hat{\chi}_\Lambda$ defined on $\hat{M}$ to a function having width exactly equal to s with respect to each of the coordinate decompositions. With this aim, we increase some of the numbers h_{ik}, or, what is the same, the numbers m'_{ik}, being careful, however, not to disturb the four conditions

$$m'_{ik} \leqslant m_{ik}, \quad i, k = 1, 2,$$

and the conditions

$$s_1 = \frac{m'_{11} + m'_{12}}{m_{11} + m_{12}} \leqslant s, \quad s_2 = \frac{m'_{21} + m'_{22}}{m_{21} + m_{22}} \leqslant s,$$

$$s_3 = \frac{m'_{11} + m'_{21}}{m_{11} + m_{21}} \leqslant s, \quad s_4 = \frac{m'_{12} + m'_{22}}{m_{12} + m_{22}} \leqslant s.$$

If for some new values of the numbers m'_{ik} in the latter conditions the four inequalities pass into exact equalities (three suffice), then our goal is attained. It thus remains to consider the case when an increase of the numbers h_{ik} (or m'_{ik}) leads to the situation in which $h_{ik} = k_{ik}$ for some i and k, but the corresponding width of the increased function $\hat{\chi}_\Lambda$ on one of the elements of the coordinate decompositions containing the cell $X_k \times Y_i$ is still less than s. For definiteness, let $h_{12} = k_{12}$. Let g_{ik}, $i, k = 1, 2$, be the values of the elements $\hat{M}_{ik}$ of a function having constant width with respect to both decompositions equal to some number $\bar{s}$ and for which $g_{12} = k_{12}$ $(= h_{12})$. These conditions determine the values g_{ik} uniquely. We compute these values, considering only values of $\bar{s}$ for which the function $g_{ik} = g(\hat{M}_{ik})$ majorizes $h_{ik} = h(M_{ik})$:

$$1)\ h_{12} = g_{12} = k_{12},$$

$$2)\ h_{11} \leqslant g_{11} = \frac{1}{a_1}(\bar{s} - k_{12}a_2) \leqslant k_{11},$$

$$3)\ h_{22} \leqslant g_{22} = \frac{1}{b_1}(\bar{s} - k_{12}b_1) \leqslant k_{22},$$

$$4)\ h_{21} \leqslant g_{21} = \frac{1}{a_1 b_1}((a_1 - b_1)\bar{s} + m_{12}) \leqslant k_{21}.$$

Since all the numbers g_{ik} pass into the corresponding k_{ik} for $\bar{s} = 1$, all the above inequalities are satisfied for this value of $\bar{s}$. We begin to decrease $\bar{s}$ as long as possible, i.e., until one of the inequalities 2)–4) becomes an equality (the quantities g_{ik} depend linearly on $\bar{s}$). We consider all conceivable cases.

I. First let $a_1 - b_1 = 1 - (a_2 + b_1) = 1 - c > 0$. Then the quantities g_{ik} are monotonically decreasing with a decrease of $\bar{s}$, and therefore one of the left-hand inequalities in 2)–4) turns into an equality as the result of a decrease of $\bar{s}$.

Ia. If $h_{11} = (1/a_1)(\bar{s} - k_{12}a_2)$, then

$$\bar{s} = h_{11}a_1 + h_{12}a_2 = \frac{m'_{11}}{b_1} + \frac{m'_{21}}{b_1} = \frac{m'_{11} + m'_{21}}{m_{11} + m_{21}} = s_3 \leqslant s,$$

i.e., in this case we have the required function of width s_3 not greater than s.

Ib. If $h_{22} = g_{22}$, then, analogously to the preceding, there is a required function of width s_2 not greater than s.

Ic. Let

$$h_{21}=g_{21}=\frac{1}{a_1 b_1}((a_1-b_1)s+m_{12}).$$

It is required to show that the value $\bar{s}$ for which this equality is attained exceeds s by a small amount if the vector $((s_1, s_2), (s_3, s_4))$ is sufficiently close to the vector $((s, s), (s, s))$. Let

$$\sum_{k=1}^{4}|s_k-s|\leqslant\varepsilon_0. \tag{4}$$

We consider the function $n = g - h$; $n(\hat{M}_{12}) = n(\hat{M}_{21}) = 0$. The projections of n onto $\hat{X}$ and $\hat{Y}$ are equal to the respective vectors $(\bar{s} - s_1, \bar{s} - s_2)$ and $(\bar{s} - s_3, \bar{s} - s_4)$. By the corollary of Proposition 51, both of these vectors cannot be close to any constant $a \neq 0$:

$$\sum_{k=1}^{4}|(s-s_k)-a|\geqslant 2a(1-c). \tag{5}$$

In this inequality we set $a = \bar{s} - s$:

$$\sum_{k=1}^{4}|(s-s_k)-(s-s)|=\sum|s-\varepsilon_k|\geqslant 2(s-s)(1-c). \tag{6}$$

Comparing (4) and (6), we get $2(\bar{s} - s)(1 - c) \leqslant \epsilon_0$, i.e., $\bar{s} - s \leqslant \epsilon_0/2(1 - c)$, and if $\epsilon_0 \leqslant \delta s$, $\delta \leqslant \delta_0 = 2(1 - c)\epsilon$, then

$$s\leqslant s\left(1+\frac{\delta}{2(1-c)}\right)\leqslant s(1+\varepsilon)$$

in correspondence with the assertion of the theorem.

II. We now consider the case when the difference $a_1 - b_1$ is small (in particular, negative; see Figure 6). Namely, suppose that

$$a_2+b_1=c,\quad \text{i.e.}\quad a_1-b_1=1-c, \tag{7}$$

where

$$c>1-\frac{1}{8K} \tag{8}$$

and c is such that

$$n_{\frac{1}{8K}}(c)>0. \tag{9}$$

(For the definition of the function $n_\alpha(c)$ see p. 101.) And let

$$0<s_0<n_{\frac{1}{8K}}(c) \tag{10}$$

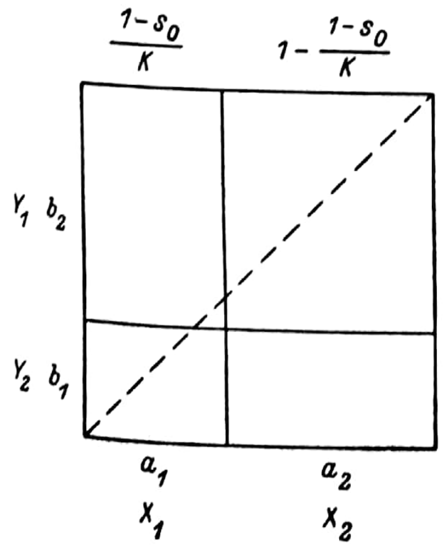

FIGURE 6

and

$$s_0 < \frac{1}{4}. \tag{11}$$

First of all we point out that for s_0 fixed and $s \leqslant s_0$ the quantities a_1 and b_2 cannot be arbitrarily close to 1. For, since the ξ-width of the cell $X_2 \times Y_2$ and the η-width of $X_1 \times Y_1$ are not less than $1 - s_0$ (in fact, we assume that $X_2 \times Y_1$ is entirely contained in Λ, since $h_{12} = k_{12}$, and therefore it has width not exceeding s_0 with respect to each of the coordinate decompositions), it follows that

$$Ka_1 \geqslant 1 - s_0, \quad Kb_2 \geqslant 1 - s_0. \tag{12}$$

We now prove that

$$b_1 \geqslant c - 1 + \frac{1 - s_0}{K}, \quad a_2 \geqslant c - 1 + \frac{1 - s_0}{K}.$$

Indeed, the first of the inequalities in (12) implies the condition

$$a_2 = 1 - a_1 \leqslant \frac{1 - s_0}{K},$$

from which, by (7), (8), and (11), we get

$$b_1 \geqslant c - 1 + \frac{1 - s_0}{K} > 1 - \frac{1}{8K} - 1 + \frac{1}{4K} = \frac{1}{8K}. \tag{13}$$

Completely analogously it follows from the second inequality in (12) that

$$a_2 \geqslant \frac{1}{8K}. \tag{14}$$

The inequalities (13) and (14) imply, by the definition of $n_\alpha(c)$, that

$$m_{12} \geqslant n_{\frac{1}{8K}}(c) > 0,$$

from which, in particular, it follows that the supremum with respect to x of the ξ-width of the cell $X_2 \times Y_1$ is not less than $n_{1/8K}(c)$ (and the same for the η-width), and this contradicts (10). Thus, the assumption that (7)–(11) hold is incompatible with the assumed equality $h_{12} = k_{12}$; in other words, under all the indicated restrictions case II cannot occur.

We summarize the above presentation. Let the number c_0 be such that

$$1 > c_0 > 1 - \frac{1}{8K}, \quad n_{\frac{1}{8K}}(c_0) > 0$$

(it was shown above that such a number c_0 always exists under the assumption that $\xi \wedge \eta = \nu$). The approximation theorem is proved for s_0 satisfying the inequalities

$$s_0 < \frac{1}{4}, \quad s_0 < n_{\frac{1}{8K}}(c_0),$$

and for $\delta_0 = 2(1 - c_0)\epsilon$. ●

We now establish an analogue of the approximation theorem for the case when the decomposition $\xi \wedge \eta$ is purely continuous.

THEOREM 7*. *Let m be a doubly stochastic measure such that for the decompositions ξ and η ($\xi\eta = \epsilon$) the measure $m/(\xi \wedge \eta)$ is purely continuous and the image of m under the canonical mapping*

$$\pi_X \times \pi_Y : M \to M/\xi \times M/\eta \equiv X \times Y$$

is absolutely continuous, with bounded density $k(x, y)$, with respect to the measure m^* for which

1) $m^*\pi_X^{-1} = \mu$ *and* $m^*\pi_Y^{-1} = \nu$.

2) $m/(\xi \wedge \eta) = m^*/(\xi \wedge \eta)$.

3) *On almost every element of $\xi \wedge \eta$ the conditional measure coincides with the product of the conditional measures on the corresponding elements of the decompositions ξ_η and η_ξ of (X, μ) and (Y, ν) into the preimages of the elements of the decomposition $M/(\xi \wedge \eta)$ under the canonical mappings $(X, \mu) \longrightarrow (M, m)/(\xi \wedge \eta)$ and $(Y, \nu) \longrightarrow (M, m)/(\xi \wedge \eta)$, respectively.*

4) *The conditional measures on the elements of ξ_η and η_ξ are purely continuous.*

5) *On each element of $\xi \wedge \eta$ the density $k(x, y)$ is bounded (by a constant depending on the element).*

Then for each $\epsilon > 0$ there exist numbers $\delta_0 > 0$ and $s_0 > 0$ such that if $\Lambda \subset M$ is a subset whose widths with respect to ξ and η do not exceed some number $s \leqslant s_0$ and

$$\|\pi_X\chi_\Lambda - s\|_{L(X,\mu)} + \|\pi_Y\chi_\Lambda - s\|_{L(Y,\nu)} \leqslant \delta s \quad \text{for } \delta \leqslant \delta_0, \tag{15}$$

then there is a subset Λ_1 of constant width s' not greater than $(1 + \epsilon)s$ for which $m(\Lambda_1 \Delta \Lambda) < 5\epsilon s$.

PROOF. We first point out the difference in the assertions of Theorems 7 and 7*. Theorem 7* does not assert that $\Lambda_1 \supset \Lambda$, and only guarantees the smallness of the measure of the symmetric difference $m(\Lambda_1 \Delta \Lambda)$. Another, perhaps more essential, weakening of the formulation consists in the fact that the assertion begins not with the quantifiers $\exists s_0 \forall \epsilon \exists \delta_0 \forall \Lambda$, but with the quantifiers $\forall \epsilon \exists (s_0, \delta_0) \forall \Lambda$. However, for the subsequent use of both theorems this weaker form suffices. It can be shown (see Figure 4) that under the hypotheses of Theorem 7* the conclusion of Theorem 7 does not hold.

Suppose that we are given an arbitrary $\epsilon > 0$. We choose a number $K > 0$ such that on a set $M_1 \subset M$ that is measurable with respect to $\xi \wedge \eta$ the density $k(x, y)$ is uniformly bounded by the constant K, and $mM_1 > 1 - \epsilon_1$. Then for each element of $\xi \wedge \eta$ we consider the corresponding function $n_\alpha(t)$ and choose numbers $c_0 < 1$ and $\kappa > 0$ such that $n_{1/8K}(c_0) \geqslant \kappa > 0$ for all elements of $\xi \wedge \eta$ appearing in M_1, except certain ones whose union M_2 has measure not exceeding ϵ_2. From (15) it follows that for any $\gamma > 1$ the total measure of the (measurable) set M_3 made up of the union of all elements C of $\xi \wedge \eta$ for which

$$\|\pi_X\chi_\Lambda - s\|_{L(X,\mu_C)} + \|\pi_Y\chi_\Lambda - s\|_{L(Y,\nu_C)} \leqslant \gamma\delta s, \tag{16}$$

is not less than $1 - 1/\gamma$, since the number $\|\pi_X\chi_\Lambda - s\|_{L(X,\mu)} + \|\pi_Y\chi_\Lambda - s\|_{L(Y,\nu)}$ is the barycenter of the numbers on the left-hand side of (16) (with respect to the measure $m/(\xi \wedge \eta)$). Thus, on each element of the set $M' = M_1 \cap (M\backslash M_2) \cap M_3$ the conditions of Theorem 7 hold with $s_0 < 1/4$, $s_0 < \kappa$ and $\delta_0 = 2(1 - c_0)\epsilon/\gamma$, and we can assume that

$$mM' \geqslant 1 - \varepsilon_1 - \varepsilon_2 - \frac{1}{\gamma} > 1 - \varepsilon$$

and $\delta \leqslant \epsilon$. Using Theorem 7, we observe that on the set M' we can complete its intersection with a set Λ satisfying the conditions of Theorem 7* to form a measurable set Λ_1' whose width does not exceed $(1 + \epsilon)s$. Indeed, since such a completion is possible on each element of $\xi \wedge \eta$ that is in M', each coarsened 2×2 complementation problem is solvable for each element, and hence each coarsened 2×2 problem is solvable for the whole set M', which means the existence of the required set Λ_1' of constant width. Now, to conclude the construction of Λ_1, we add an arbitrary set of the required constant width on $M\backslash M'$ to the set Λ_1' (such a set of constant width exists, by Proposition 44*). Finally, we get

$$\begin{aligned} m(\Lambda_1\Delta\Lambda) &= m((M'\cap(\Lambda_1\Delta\Lambda))\cup((M\setminus M')\cap(\Lambda_1\Delta\Lambda))) \\ &\leqslant m((M'\cap(\Lambda_1\setminus\Lambda))\cup(((M\setminus M')\cap\Lambda_1)\cup((M\setminus M')\cap\Lambda))) \\ &\leqslant (1+\varepsilon)s - (1-\delta)s + \varepsilon(1+\varepsilon)s + \varepsilon s < 5\varepsilon s. \end{aligned}$$

Theorem 7* is proved. ●

The approximation theorem shows the possibility of getting a good approximation of narrow strips of width that is close to being constant by means of strips of exactly constant width that can be assumed to be elements of some decomposition that is independent with respect to ξ and η. We now proceed to the construction of strips (sets $\Lambda \subset M$) that we approximate with sets that are measurable with respect to the desired independent complement.

11. DEFINITION. Let C be a measurable subset of $M = X \times Y$, n a measure on $X \times Y$. We use the notation

$$\Pi_n C = \Pi_{n,\mu,\nu} C = \sup\{\mu A + \nu B : \mu A + \nu B > 0,\ n((A \times B)\cap C) = 0\};$$
$$\Pi_n' C = \sup\{\mu A + \nu B : n((A \times B)\cap C) = 0\};$$
$$\Pi C = \Pi_{\mu\times\nu} C,\quad \Pi' C = \Pi_{\mu\times\nu}' C.$$

Thus, we always have $\Pi_n' C \geqslant 1$, and for a doubly stochastic measure m and a set C such that $mC = 1$ we always have $\Pi_m C = 1$ (Proposition 51).

PROPOSITION 52. *Let $X = X_1 \cup \cdots \cup X_m$ and $Y = Y_1 \cup \cdots \cup Y_m$ be decompositions of the spaces X and Y such that $\mu X_k = \nu Y_k = p_k$, $k = 1, \ldots, n$, and let the measurable set D be contained in the union $\bigcup_1^n X_k \times Y_k$. Let $\mu^{(k)}$ and $\nu^{(k)}$ be the normalized restrictions of μ and ν to the subsets X_k and Y_k, and let*

$$\Pi'^{(k)} D = \Pi'_{\mu^{(k)}\times\nu^{(k)}}(D \cap (X_k \times Y_k)).$$

Then

$$\Pi' D = \sum_{k=1}^{n} p_k \Pi'^{(k)} D.$$

PROOF. For some measurable subsets $A \subset X$ and $B \subset Y$ let

$$(\mu \times \nu)((A \times B) \cap D) = 0.$$

Then

$$(\mu^{(k)} \times \nu^{(k)})(((A \cap X_k) \times (B \cap Y_k)) \cap ((X_k \times Y_k) \cap D)) = 0,\ k = 1, \dots, m,$$

and

$$\sum_{k=1}^{n} p_k (\mu^{(k)}(A \cap X_k) + \nu^{(k)}(B \cap Y_k)) = \mu A + \nu B.$$

Conversely, if $A_k \subset X_k$ and $B_k \subset Y_k$ are such that

$$(\mu^{(k)} \times \nu^{(k)})((A_k \times B_k) \cap ((X \times Y) \cap D)) = 0,$$

then

$$(\mu \times \nu)\left(\left(\bigcup_{k=1}^{n} A_k \times \bigcup_{k=1}^{n} B_k\right) \cap D\right) = 0$$

and

$$\mu \bigcup_{k=1}^{n} A_k + \nu \bigcup_{k=1}^{n} B_k = \sum_{k=1}^{n} p_k (\mu^{(k)} A_k + \nu^{(k)} B_k).$$

The required assertion follows from this. ●

The following generalization of this proposition is just as obvious.

PROPOSITION 52*. *Let m be a doubly stochastic measure, and $\{m_C\}$, $\{\mu_C\}$, and $\{\nu_C\}$ the conditional measures on the elements C of the decomposition $\xi \wedge \eta$ and the corresponding elements of the decompositions ξ_η and η_ξ (determined by the canonical mappings $X \longrightarrow M/(\xi \wedge \eta)$ and $Y \longrightarrow M/(\xi \wedge \eta)$). Then for an arbitrary set D*

$$\Pi' D = \int \Pi'_{\mu_C \times \nu_C} D d(m/\xi \wedge \eta).$$

PROPOSITION 53. *For any refining sequence of finite measurable decompositions $\zeta_n \nearrow \epsilon$ of the Lebesgue space $(M, \mathfrak{A}, m)$ and any measurable subset $C \subset M$ there is a sequence of sets C_n satisfying the following conditions:*

1) *Each C_n is measurable with respect to the corresponding decomposition ζ_n.*

2) *There exists a numerical sequence $\alpha_n \nearrow 1$ such that for each n and each element $C^{(\zeta_n)}$ of ζ_n contained in C_n the measure of the part of C inside $C^{(\zeta_n)}$ is not less than $\alpha_n m C^{(\zeta_n)}$.*

3) *The characteristic functions χ_{C_n} of the sets C_n converge in measure as $n \longrightarrow \infty$ to the characteristic function of C.*

PROOF. For each n we define a ζ_n-measurable function $q_n(z)$, $z \in M$, by

$$q_n(z) = m(C \cap C^{(\zeta_n)}(z))(mC^{(\zeta_n)}(z))^{-1},$$

where $C^{(\zeta_n)}(z)$ is the element of ζ_n containing the point $z \in M$. By the well-known Lebesgue theorem (see, for example, [90]) on points of density of measurable sets, for almost all $z \in C$ we have $q(z) \longrightarrow 1$, and for almost all $z \notin C$ we have $q_n(z) \to 0$. (The Lebesgue theorem relates to the case when (M, m) is the unit interval with Lebesgue measure and ζ_n is a refining sequence of finite decompositions of the unit interval into smaller segments whose maximal lengths go to zero, but it is easy to see that our formally more general situation is isomorphic to this "classical" one.) Therefore, for any number $0 < \alpha < 1$ the sequence of sets $C_n^\alpha = \{z\colon\ q_n(z) > \alpha\}$ satisfies requirements 1) and 3) of Proposition 53.

Let $\rho(f, g)$ be a metric on $S(M, m)$ whose convergence is equivalent to convergence in measure. By what was proved above, for any number α, $0 < \alpha < 1$, the sequence of functions $\chi_{C_n^\alpha}$ converges in measure to the function χ_C, i.e., $\rho(\chi_{C_n^\alpha}, \chi_C) \longrightarrow 0$ for any such number α. Therefore, we can choose a sequence of numbers $\alpha_n \to 1$, $n = 1, \ldots,$ such that $\rho(\chi_{C_n^{\alpha_n}}, \chi_C) \longrightarrow 0$ for $n \longrightarrow \infty$. For example, let $\beta_k = 1 - 1/(k + 1)$, and let the numbers $n_1 < n_2 < \cdots$ be chosen successively so that n_1 is such that $\rho(\chi_{C_n^{\beta_2}}, \chi_C) < ½$ for $n > n_1, \ldots, n_k$ is such that $\rho(\chi_{C_n^{\beta_{k+1}}}, \chi_C) < 1/(k + 1)$ for $n > n_k$, etc. Then we can take α_n to be the nth term of the sequence

$$\underbrace{\beta_1, \beta_1, \ldots, \beta_1}_{n_1 \text{ times}}\ \underbrace{\beta_2, \beta_2, \ldots, \beta_2}_{(n_2 - n_1)\text{times}}\ \underbrace{\beta_3, \beta_3 \ldots, \beta_3}_{(n_3 - n_2 - n_1) \text{ times}}, \ldots$$

Moreover, it is easy to see that $\rho(\chi_{C_n^{\alpha_n}}, \chi_C) \longrightarrow 0$, i.e., $\chi_{C_n^{\alpha_n}} \longrightarrow \chi_C$ in measure. ●

PROPOSITION 54. *Let $X = X_1 \cup X_2, X_1 \cap X_2 = \varnothing, \mu X_1 = a,\ Y = Y_1 \cup Y_2, Y_1 \cap Y_2 = \varnothing, \nu Y_1 = b$, and $a + b = 1 + c,\ c > 0$. Then $m(X_1 \times Y_1) \geqslant c$.*

PROOF. Let $m_{ik} = m(X_k \times Y_i)$. Then $m_{12} + m_{22} = 1 - a$ and $m_{21} + m_{22} = 1 - b$, and so

$$m_{12} + m_{21} + 2m_{22} = 2 - (a + b),$$

$$1 - m_{11} + m_{22} = 1 - c,$$

$$m_{11} - m_{22} = c, \quad \text{i.e.} \quad m_{11} \geqslant c. \ ●$$

PROPOSITION 55. *As before, let the measure m on the space $M = X \times Y$ be given by a density $k(x, y)$ with respect to the product measure $\mu \times \nu$. For any measurable set $X_1 \subset X$ and any $\epsilon > 0$ there is a measurable set $Y^1 \subset Y$ for which $\nu Y^1 = 1 - \mu X_1$ and*

$$\Pi'_m (M \setminus (X_1 \times Y^1)) < 1 + \varepsilon.$$

PROOF. We can limit ourselves to the case when the number μX_1 is a dyadic rational. In the opposite case we can use the fact that if

$$\bar{X}_1 \subset X_1,\ \mu (X_1 \setminus \bar{X}_1) < \frac{\varepsilon}{2} \quad \text{and} \quad \Pi'_m (M \setminus (\bar{X}_1 \times \bar{Y}^1)) < 1 + \frac{\varepsilon}{2},$$

then
$$\Pi'_m(M \setminus (X_1 \times Y^1)) < 1 + \varepsilon,$$

and take $\bar{X}_1$ to be a set with dyadic rational measure and Y^1 an arbitrary subset of $\bar{Y}^1$ having measure equal to $1 - \mu X_1$. From the "double stochasticity" of m, i.e., from the conditions $\mu = m\pi_X^{-1}$ and $\nu = m\pi_Y^{-1}$, it follows that $\Pi_m \leqslant 1$ and $\Pi'_m M = 1$. Indeed, if the sets $A \subset X$ and $B \subset Y$ are such that $m(A \times B) = 0$, then it follows from Proposition 51 that $\mu A + \nu B = 1$ (a fact that is clear and immediate). Let $C = \{(x, y)\colon\ k(x, y) > 0\}$. We consider refining sequences ξ_n and η_n of finite coarsenings of the decompositions ξ and η that converge monotonically to ξ and η and consist each of 2^n subsets of equal measure. We can assume that X_1 is measurable with respect to ξ_n for some n. The characteristic function $\chi_C(x, y)$ of C can be arbitrarily well approximated in the measure $\mu \times \nu$ (and hence also in the measure m) by the characteristic functions χ_{C_n} of sets C_n that are measurable with respect to the decompositions $\xi_n \vee \eta_n$ of M.

For sufficiently large n we have $\Pi' C_n < 1 + \epsilon/2$. Indeed, otherwise we could find cells $A^{(n)} \times B^{(n)} \subset M$ for which $\mu A^{(n)} + \mu B^{(n)} \geqslant 1 + \epsilon/2$ and

$$(\mu \times \nu)((A^{(n)} \times B^{(n)}) \cap C_n) = 0.$$

But for such subsets $A^{(n)}$ and $B^{(n)}$ we have, by Proposition 54,

$$m(A^{(n)} \times B^{(n)}) = m((A^{(n)} \times B^{(n)}) \cap C) \geqslant \frac{\varepsilon}{2}.$$

By the convergence in measure $\chi_{C_n} \to \chi_C$ the $(\mu \times \nu)$-measures, and hence the m-measures, of the subsets of M on which these functions differ converge to zero with increasing n, and for sufficiently large n, therefore, the simultaneous satisfaction of the conditions $m(D \cap C_n) = 0$ and $m(D \cap C) \geqslant \epsilon/2$ is impossible for any measurable set D, in particular, for $D = A^{(n)} \times B^{(n)}$.

Further, by Proposition 53, we can assume that on each element of $\xi_n \vee \eta_n$ contained in C_n the measure of the part of C inside this element is not less than $2^{-2n}\alpha_n$, where $\alpha_n \to 1$. We now consider for each n the function $\hat{\chi}_{C_n}$ defined on the space

$$\hat{M}_n = M/\xi_n \vee \eta_n,$$

which consists of 2^{2n} elements, and taking the values 0 and 1 on those elements of $\hat{M}_n$ on whose preimages under the canonical projection $M \to M/(\xi_n \vee \eta_n)$ the function χ_{C_n} takes the respective values 0 and 1 (the latter function is measurable with respect to $\xi_n \vee \eta_n$, so our definition is correct). We regard the space $\hat{M}$ with the function $\hat{\chi}_{C_n}$ as a (0, 1)-matrix of dimension $2^n \times 2^n$. As is well known from the theory of (0, 1)-matrices (see, for example, [101]), it follows from the inequalities

$$\Pi' C_n < 1 + \frac{\varepsilon}{2} \leqslant 1 + \frac{d}{2^n}$$

(d a positive integer, which can be attained if n is sufficiently large; for simplicity we assume that $\epsilon/2 = d/2^n$) that for the matrix $(\hat{M}_n, \hat{\chi}_{C_n})$ there is a (0, 1)-matrix $(\hat{M}_n, \hat{\chi}^0_{C_n})$ of the same size that contains in each column and in each row not more than one 1, and such that the corresponding elements of the matrix $\hat{\chi}_{C_n}$ are also equal to 1, and if $C_n^0 \subset M$ denotes a set for which $\hat{\chi}_{C_n^0} = \hat{\chi}_{C_n}$, then $\Pi' C_n^0 = \Pi' C_n$; that is, not more than d rows (and columns) of the matrix $(\hat{M}_n, \hat{\chi}_{C_n}^0)$ do not contain ones.

We now consider each element of $\xi_n \vee \eta_n$ that is contained in C_n^0 as an independent space with measure defined as the normalized restriction of m to this element (a subspace of the Lebesgue space $(M, \mathfrak{A}, m)$). Besides the measure $\hat{m}$ determined by m, we consider also the measure $\widehat{\mu \times \nu}$ determined by $\mu \times \nu$ on this element. (The restriction of m, of course, is no longer "doubly stochastic" with respect to μ and ν.) From the preceding, on each such subspace the trace of C has $(\widehat{\mu \times \nu})$-measure not greater than α_n, from which it follows that for this subspace

$$\Pi_{\widehat{\mu\times\nu}} \leqslant 2 - \alpha_n$$

(since the maximum of the function $u + v$ under the condition $uv = 1 - \alpha_n$, $0 \leqslant u \leqslant 1$, $0 \leqslant v \leqslant 1$, is equal to $2 - \alpha_n$). Therefore, the subset $\bar{C}_n = C_n^0 \cap C$ of C is already such that

$$\Pi' \bar{C}_n \leqslant (2 - \alpha_n)\left(1 - \frac{\varepsilon}{2}\right) + \varepsilon \quad \text{(for sufficiently large } n\text{)},$$

i.e.,

$$\Pi' \bar{C}_n \leqslant 2 - \alpha_n - \varepsilon + \frac{1}{2}\alpha_n \varepsilon + \varepsilon < 1 + \varepsilon \quad \text{(for large } n\text{)},$$

and a fortiori $\Pi' C_n < 1 + \epsilon$ (for large n).

On the other hand, as mentioned, we can assume without loss of generality that the set X_1 is ξ_n-measurable. For some subset $Y^1 \subset Y$ that is measurable with respect to the decomposition η_n and for which $\nu Y^1 = 1 - \mu X_1$, we have

$$(X_1 \times Y^1) \cap \bar{C}_n = \varnothing. \tag{17}$$

Indeed, to construct such a set Y^1 it is sufficient to consider a $(\xi_n \vee \eta_n)$-measurable set $\tilde{C}_n^0 \supset C_n^0$ for which the matrix $(\hat{M}_n, \hat{\chi}_{\tilde{C}_n^0})$ is a permutation matrix (i.e., contains exactly one 1 in each row and in each column), and to define Y^1 to be the union of those elements $C^{(\eta_n)}$ of η_n for which there is an element $C^{(\xi_n)}$ of ξ_n, not contained in X_1, such that

$$C^{(\xi_n)} \times C^{(\eta_n)} \subset \tilde{C}_n^0.$$

In this case the product of $C^{(\eta_n)}$ with any element of ξ_n contained in X_1 is, on the contrary, not contained in $\tilde{C}_n^0$ (and does not intersect it), i.e., (17) holds.

Finally, we get

$$1 + \varepsilon > \Pi' \bar{C}_n = \Pi'_m \bar{C}_n = \Pi'_m (\bar{C}_n \setminus (X_1 \times Y^1)) \geqslant \Pi'_m (M \setminus (X_1 \times Y^1)). \quad \bullet$$

We give an analogue of Proposition 55 when the decomposition $\xi \wedge \eta$ is arbitrary.

PROPOSITION 55*. *Suppose that the hypotheses of Proposition 44* hold. For any measurable set $X_1 \subset X$ and any $\epsilon > 0$ there is a measurable set $Y^1 \subset Y$ for which $\nu Y^1 = 1 - \mu X_1$ and*

$$\Pi_m (M \setminus (X_1 \times Y^1)) < 1 + \varepsilon.$$

PROOF. By Proposition 55 it suffices to consider the case when $\xi \wedge \eta$ is purely continuous. As before, it suffices to assume that the number μX_1 is a dyadic rational. Let $C = \{(x, y)\colon\ k(x, y) > 0\}$. We consider a decomposition κ that is an independent complement of $\xi \wedge \eta$ with respect to the measure m^* (defined in Proposition 44*). The space (M, m^*) is now canonically isomorphic to the space

$$(M,\ m^*)/\xi \wedge \varkappa \times (M,\ m^*)/\eta \wedge \varkappa \times (M,\ m^*)/\xi \wedge \eta.$$

We consider refining sequences of measurable decompositions

$$(\xi \wedge \varkappa)_n \nearrow \xi \wedge \varkappa,\ \ (\eta \wedge \varkappa)_n \nearrow \eta \wedge \varkappa,\ \ (\xi \wedge \eta)_n \nearrow \xi \wedge \eta,$$

each containing 2^n subsets of equal measure. Let

$$\xi_n = (\xi \wedge \varkappa)_n \vee (\xi \wedge \eta)_n,\ \ \eta_n = (\eta \wedge \varkappa)_n \vee (\xi \wedge \eta)_n.$$

The space (M, m^*) can thus be represented as a block cut by planes parallel to the edges (the decompositions $(\xi \wedge \kappa)_n$, $(\eta \wedge \kappa)_n$ and $(\xi \wedge \eta)_n$) into 2^{3n} equal parts. We can assume that $\overline{X}_1$ is measurable with respect to some decomposition $\xi_n \vee \eta_n$. The characteristic function χ_C can be approximated arbitrarily well in the measure m^* by characteristic functions χ_{C_n} of sets C_n that are measurable with respect to $\xi_n \vee \eta_n$. For any $\delta > 0$ and sufficiently large n we have, as above, the inequality $\Pi' C_n < 1 + \delta\epsilon/2$. Furthermore, by Proposition 53 we can assume that on each element of $\xi_n \vee \eta_n$ contained in C_n the measure of the part of C contained in this element is not less than $2^{-3n}\alpha_n$, where $\alpha_n \nearrow 1$. Using Proposition 52, we find that $\Pi' C_n < 1 + \delta\epsilon/2$ implies the inequality

$$\Pi'_{m^*_{C^{(n)}},\ m^*_{C^{(n)}}\pi_X^{-1},\ m^*_{C^{(n)}}\pi_Y^{-1}} C_n < 1 + \frac{1}{2}\varepsilon, \tag{18}$$

where $m^*_{C^{(n)}}$ is the normalized restriction of m^* to the element $C^{(n)}$ of $\xi_n \vee \eta_n$, on the collection of elements of $\xi_n \wedge \eta_n$ of total measure not less than $1 - \delta$. Considering each such element and arguing as in the proof of Proposition 55, we construct for such an element a (0, 1)-matrix that is analogous to the matrix $(\hat{M}_n, \hat{\chi}_{C_n^0})$ and that contains in each column and in each row not more than one 1, and for those elements of $\xi_n \wedge \eta_n$ for which (18) does not hold we let such a (0, 1)-matrix consist only of zeros. Using Proposition 52 again, we find that

$$\Pi' \tilde{C}_n^0 < 2\delta + \left(1 + \frac{1}{2}\varepsilon\right)(1 - \delta) < 1 + \frac{1}{2}\varepsilon + \delta,$$

where $\tilde{C}_n^0$ is the union of those elements of $\xi_n \vee \eta_n \vee \kappa_n$ to which the ones in the matrices $(\hat{M}_n, \hat{\chi}_{C_n^0})$ correspond. Finally, using the fact that $\alpha_n \longrightarrow 1$ and choosing the number δ sufficiently small, we find, as above, that

$$\Pi'(\tilde{C}_n^\circ \cap C) < 1 + \varepsilon.$$

The construction of Y^1 reduces to a 2^n-step repetition of the construction of Y^1 in the proof of Proposition 55 (according to the number of elements of $\xi_n \wedge \eta_n$), and this concludes the proof. ●

12. Let us now consider a measure n that is absolutely continuous with respect to the product $\mu \times \nu$. We determine to what extent measures that are absolutely continuous with respect to this fixed measure n can be doubly stochastic.

We assume n is such that $\Pi_n M = 1 + a$, $a > 0$ (if $\Pi_n M \leqslant 1$, then among the measures that are absolutely continuous with respect to n there are necessarily measures that are arbitrarily close to being doubly stochastic; we shall not consider this case now).

Let

$$M = M_1 \cup M_2, \quad M_1 \cap M_2 = \varnothing, \quad \frac{dn}{d(\mu \times \nu)} > 0 \text{ on } M_1, \quad nM_2 = 0.$$

Let P denote the class of nonnegative n-integrable functions vanishing on M_2. We are interested in how well the pair $(1, 1) \in L(X, \mu) \times L(Y, \nu)$ can be approximated by the pairs $(\pi_X h, \pi_Y h)$ for $h \in C$.

We consider the space $E = (\pi_X \times \pi_Y) L(M, \mu \times \nu) \subset L(X) + L(Y)$, which is canonically isomorphic to $L(M)/\mathrm{Ker}(\pi_X \times \pi_Y)$, by the homomorphism theorem. The norm on E is also defined canonically as the norm on $L(M)/\mathrm{Ker}(\pi_X \times \pi_Y)$ of a normed space:

$$\|(f(x), g(y))\|_E = \inf_{\substack{f=\pi_X h \\ g=\pi_Y h}} \|h\|_{L(M)}. \tag{19}$$

On the one hand, if $f = \pi_X h$ and $g = \pi_Y h$, then

$$\int_X f(x)\, d\mu = \int_Y g(y)\, d\nu = \int_{X \times Y} h(x, y)\, d(\mu \times \nu)$$

and

$$\|f\|_{L(X)} = \int |f|\, d\mu = \int \left| \int h\, d\nu \right| d\mu \leqslant \int |h|\, d(\mu \times \nu) = \|h\|_{L(M)},$$

$$\|g\|_{L(Y)} \leqslant \|h\|_{L(M)}, \quad \|f\|_{L(X)} + \|g\|_{L(Y)} \leqslant 2\|h\|_{L(M)},$$

and, on the other hand, if we are given functions $f(x)$ and $g(y)$ such that $\int_X f(x)d\mu = \int_Y g(y)d\nu$, then $f = \pi_X h_1$ and $g = \pi_Y h_1$, where $h_1(x, y) = f(x) + g(y) - \int_X f(x)d\mu$, and, therefore,

$$\inf_{\substack{f=\pi_X h \\ g=\pi_Y h}} \|h\| \leqslant 2(\|f\|_{L(X)} + \|g\|_{L(Y)}).$$

Thus the norms $\|\cdot\|_E$ and $\|f(x), g(y)\|_1 = \|f\|_{L(X)} + \|g\|_{L(Y)}$ are equivalent:

$$\frac{1}{2}\|(f, g)\|_1 \leqslant \|(f, g)\|_E \leqslant 2\|(f, g)\|_1. \tag{20}$$

The space dual to $L(X) \times L(Y)$ is $L^\infty(X, \mu) \times L^\infty(Y, \nu)$, and the space dual to the subspace $E \subset L(X) \times L(Y)$ is the space

$$L^\infty(X) \times L^\infty(Y) / \operatorname{Ker}(\pi_X \times \pi_Y)^*,$$

where

$$(\pi_X \times \pi_Y)^* : L^\infty(X) \times L^\infty(Y) \to L^\infty(M, \mu \times \nu),$$
$$(\pi_X \times \pi_Y)^* (u(x), v(y)) = u(x) + v(y).$$

An equivalent point of view is that the space dual to E can be interpreted as the whole space $L^\infty(X) \times L^\infty(Y)$, equipped with the seminorm

$$p((u, v)) = \|u(x) + v(x)\|_{L^\infty(M)},$$

which can be regarded as a real norm on the quotient space by its kernel, i.e., on

$$L^\infty(X) \times L^\infty(Y) / \operatorname{Ker}(\pi_X \times \pi_Y)^*.$$

This norm agrees with the norm $\|\cdot\|_E$ introduced above.

PROPOSITION 56. *For any $\epsilon > 0$ a number $a > 0$ can be found such that, for any measure n that is absolutely continuous with respect to the measure $\mu \times \nu$ and for which $\Pi_n M \leqslant 1 + a$, there is a nonnegative bounded function $h(x, y)$ such that*

$$h(x, y) = 0 \quad \text{on the set} \quad \left\{(x, y): \frac{dn}{d(\mu \times \nu)} = 0\right\}, \tag{21}$$
$$(\pi_X h)(x) \leqslant 1, \quad (\pi_Y h)(y) \leqslant 1, \tag{22}$$
$$\|\pi_X h - 1\|_{L(X, \mu)} + \|\pi_Y h - 1\|_{L(Y, \nu)} < \epsilon. \tag{23}$$

PROOF. We first prove that for $a < \epsilon/2$ there exists a nonnegative function $h(x, y)$ that satisfies (21) and (23). We consider the space

$$E = (\pi_X \times \pi_Y) L(M, \mu \times \nu)$$

and introduce on it the norm $\|(f, g)\|_E$ in (19). We estimate the distance in this norm from the element $(1, 1) \in E$ to the image $K^+ \subset E$ under the mapping $\pi_X \times \pi_Y$ of the cone $P \subset L(M, \mu \times \nu)$ of nonnegative functions that vanish outside the set on which the density $dn/d(\mu \times \nu)$ is positive. For this purpose it suffices to get a lower estimate of the values on the element $(1, 1)$ of the functionals

$$(u(x), v(y)) \in L^\infty(X) \times L^\infty(Y) = (L(X, \mu) \times L(Y, \nu))^*$$

that take nonnegative values on the cone K^+ and have norm equal to 1. In fact, by the well-known theorem on separation of a convex open set from an arbitrary convex set by means of a hyperplane (see [12], Chapter II, §3, no. 2, Proposition 1), for any open ball $V((1, 1), r)$ in E with center at $(1, 1)$ and radius r that is disjoint from the convex cone K^+ there is a functional w in $L^\infty(X) \times L^\infty(Y)$ (on which the seminorm p is considered) taking values at elements of this ball that are not in the set $w(K^+)$. Since for any functional $w \in L^\infty(X) \times L^\infty(Y)$ the set $w(K^+)$ coincides with the whole line, or with the ray $(0, \infty]$, or with the ray $(-\infty, 0]$, we get that there is a functional that takes nonnegative values on K^+ ("positive with respect to the cone K^+"), has p-norm 1, and takes a value on the element $(1, 1)$ that is arbitrarily close in absolute value to the distance from $(1, 1)$ to K^+. (For any $\epsilon > 0$ it suffices to consider a functional of unit norm separating K^+ from the ball $V((1, 1), r_1 - \epsilon)$, where r_1 is the distance in question.)

Let $K_\infty^+ \subset L^\infty(X) \times L^\infty(Y)$ be the set of functionals that are positive with respect to K^+, and $\bar{K}_\infty^+$ the subset of K_∞^+ consisting of functionals of p-norm not exceeding 1. We give a lower estimate for the values taken on this set of functionals by the element $(1, 1) \in E$. The set K_∞^+ is compact in the weak topology

$$\sigma(L^\infty(X) \times L^\infty(Y), L(X) \times L(Y)),$$

since it is closed in norm and bounded. Therefore, the infimum of the values of $(1, 1)$ on this set is attained at some functional $(u_0(x), v_0(y)) \in \bar{K}_\infty^+$.

The positivity of this functional means that $u_0(x) + v_0(y) \geqslant 0$ for almost all $(x, y) \in (M, n)$.

By Proposition 46, the function $u_0(x) + v_0(y)$ can be arbitrarily well approximated in norm by convex combinations of elements in the set $\hat{K}_\infty$ of functions that can be represented in the form $u(x) + v(y)$ by means of functions $u(x)$ and $v(y)$ taking only the values $-½$ and $½$ or the values -1 and 1 (one of them) and 0 (the other). Moreover, as mentioned, it can be assumed that $\hat{K}_\infty \subset \bar{K}_\infty^+$, i.e., that the functions $\hat{u}(x) + \hat{v}(y)$ appearing in the approximating convex combinations are also nonnegative for n-almost all points of the set $X \times Y$. From this it follows that

$$\begin{aligned} \langle (u_0, v_0), (1, 1) \rangle &= \inf_{\substack{(\hat{u}, \hat{v}) \in \hat{K}_\infty \\ \hat{u}(x) + \hat{v}(y) \geqslant 0 (n)}} \langle (\hat{u}, \hat{v}), (1, 1) \rangle \\ &= \inf \left(\int_X \hat{u}(x)\, d\mu + \int_Y \hat{v}(y)\, d\nu \right) = \frac{1}{2} \inf \Big[\mu \left\{ x : \hat{u}(x) = \frac{1}{2} \right\} \\ &- \mu \left\{ x : \hat{u}(x) = -\frac{1}{2} \right\} + \nu \left\{ y : \hat{v}(y) = \frac{1}{2} \right\} - \nu \left\{ y : \hat{v}(y) = -\frac{1}{2} \right\} \Big] \\ &= 1 - \Pi_n M = -a, \end{aligned}$$

from which we get that the distance in the norm $\|\cdot\|_E$ from $(1, 1)$ to K^+ is equal to a. By (20) it follows from this that in the norm $\|(f, g)\|_1$ this distance does not exceed $2a$. Since, by assumption, $a < \epsilon/2$, we have proved the existence of a function $h(x, y)$ such that $h(x, y) \geqslant 0$ and the conditions (21) and (23) hold.

To conclude the proof of Proposition 56, we show how to construct, for a given nonnegative function $h(x, y)$ for which

$$\|\pi_X h - 1\|_{L(X)} + \|\pi_Y h - 1\|_{L(Y)} = b > 0,$$

a function $h_0(x, y)$ such that $h_0(x, y) \geqslant 0$, $h(x', y') > 0$ if $h_0(x', y') > 0$, and

$$\|(\pi_X h_0, \pi_Y h_0)\|_1 \leqslant b + 4\sqrt{b},$$
$$(\pi_X h_0)(x) \leqslant 1, \quad (\pi_Y h_0)(y) \leqslant 1.$$

With this purpose we set $\lambda = b^{½}$ and

$$h_1(x, y)=\frac{1}{1+\lambda}h(x, y),$$

$$h_0=\begin{cases} h_1(x, y) & \text{on the set } \{(x, y):(\pi_X h_1)(x)\leqslant 1 \text{ and } (\pi_Y h_1)(y)\leqslant 1\}, \\ 0 & \text{on the set } \{(x, y):(\pi_X h_1)(x)>1 \text{ or } (\pi_Y h_1)(y)>1\}. \end{cases}$$

Let

$$E_\lambda^X=\{x:(\pi_X h)(x)>1+\lambda\}, \quad E_\lambda^Y=\{y:(\pi_Y h)(y)>1+\lambda\}.$$

Since $\|\pi_X h - 1\|_{L(X)} \leqslant b$ and $\|\pi_Y h - 1\|_{L(Y)} \leqslant b$, we have the inequalities $\mu E_\lambda^X \leqslant b/\lambda$ and $\nu E_\lambda^Y \leqslant b/\lambda$. Further,

$$\|\pi_X h_1 - 1\|_{L(X)}=\left\|\frac{1}{1+\lambda}(\pi_X h-1)+\frac{\lambda}{1+\lambda}\right\|\leqslant\frac{1}{1+\lambda}\|\pi_X h-1\|_{L(X)}+\frac{\lambda}{1+\lambda},$$

$$\|\pi_Y h_1 - 1\|_{L(Y)}\leqslant\frac{1}{1+\lambda}\|\pi_Y h-1\|_{L(Y)}+\frac{\lambda}{1+\lambda},$$

$$\|\pi_X h_0-1\|+\|\pi_Y h_0-1\|\leqslant\|\pi_X h_1-1\|+\|\pi_Y h_1-1\|+\mu E_\lambda^X+\nu E_\lambda^Y$$

$$\leqslant\frac{b}{1+\lambda}+\frac{2\lambda}{1+\lambda}+\frac{2b}{\lambda}<b+2\left(\lambda+\frac{b}{\lambda}\right)=b+4\sqrt{b}.$$

Now, considering the first part of the proof, we find that if for given $\epsilon > 0$ the constant a is chosen so that $2a + 4(2a)^{1/2} = \epsilon$, then there exists a function $h(x, y)$ such that

$$h(x, y)\geqslant 0, \quad \|(\pi_X h, \pi_Y h)\|_1=b\leqslant 2a$$

and also (21) holds. By what has been shown, there then exists a function $h_0(x, y) \geqslant 0$ satisfying (21) and the conditions

$$(\pi_X h_0)(x)\leqslant 1, \quad (\pi_Y h_0)(y)\leqslant 1,$$

$$\|(\pi_X h_0, \pi_Y h_0)\|_1\leqslant b+4\sqrt{b}\leqslant 2a+4\sqrt{2a}=\varepsilon. \bullet$$

It is not hard to see that the condition $n \ll \mu \times \nu$ is really used only for the formulation and for simplification of the presentation. The following more general assertion has actually been proved.

PROPOSITION 56*. *For any $\epsilon > 0$ a number $a > 0$ can be found such that for any nonnegative measure n on $M = X \times Y$ whose marginal distributions are absolutely continuous with respect to the measures μ and ν and for which $\Pi_n M \leqslant 1 + a$, there is a nonnegative measure n_1 such that*

$$n_1\ll n, \tag{24}$$

$$\frac{d(n_1\pi_X^{-1})}{d\mu}\leqslant 1, \quad \frac{d(n_1\pi_Y^{-1})}{d\nu}\leqslant 1, \tag{25}$$

$$\left\|\frac{d(n_1\pi_X^{-1})}{d\mu}-1\right\|_{L(X,\mu)}+\left\|\frac{d(n_1\pi_Y^{-1})}{d\nu}-1\right\|_{L(Y,\nu)}<\varepsilon. \tag{26}$$

Also, in the proof it is sufficient to replace $\pi_X h$ and $\pi_Y h$ everywhere by $d(n_1\pi_X^{-1})/d\mu$ and $d(n_1\pi_Y^{-1})/d\nu$, respectively.

PROPOSITION 57. *For any $\epsilon > 0$ a number $a > 0$ can be found such that, for any measure n that is absolutely continuous with respect to the measure $\mu \times \nu$ and for which $\Pi_n M \leqslant 1 + a$, there is a number $\bar{s} > 0$ such that for $s < \bar{s}$ there exists a subset $\Lambda \subset M$ for which*

$$\Lambda \cap \left\{(x, y) : \frac{dn}{d(\mu \times \nu)} = 0\right\} = \varnothing, \tag{27}$$

$$(\pi_X \chi_\Lambda)(x) \leqslant s, \quad (\pi_Y \chi_\Lambda)(y) \leqslant s, \tag{28}$$

$$\|\pi_X \chi_\Lambda - s\| + \|\pi_Y \chi_\Lambda - s\| \leqslant s\varepsilon. \tag{29}$$

PROOF. Let the function $h(x, y)$ whose existence is asserted in Proposition 56 be such that $h(x, y) \leqslant A < \infty$. We set $\bar{s} = 1/A$. For any $s < 1/A$ the function $sh(x, y)$ satisfies the inequalities $0 \leqslant sh(x, y) \leqslant 1$; therefore, by Proposition 44, there is a subset

$$\Lambda \subset \{(x, y) : h(x, y) > 0\} \subset \left\{(x, y) : \frac{dn}{d(\mu \times \nu)} > 0\right\},$$

such that $\pi_X \chi_\Lambda = \pi_X(sh)$ and $\pi_Y(\chi_\Lambda) = \pi_Y(sh)$. Since (21)–(23) hold for $h(x, y)$, conditions (27)–(29) hold for the set Λ. ●

In a similar way, using Proposition 56* and 44* instead of 56 and 44, we get the following assertion.

PROPOSITION 57*. *For any $\epsilon > 0$ a number $a > 0$ can be found such that, for any measure n that is absolutely continuous with respect to the measure m^* in Proposition 44* and for which $\Pi_n M \leqslant 1 + a$, there is a number $\bar{s} > 0$ such that for $s < \bar{s}$ there exists a subset $\Lambda \subset M$ for which the measure n_1 whose density with respect to m^* is the characteristic function χ_Λ has the following properties:*

$$\frac{dn_1\pi_X^{-1}}{d\mu} \leqslant s, \quad \frac{dn_1\pi_Y^{-1}}{d\nu} \leqslant s, \qquad \left\|\frac{dn_1\pi_X^{-1}}{d\mu} - s\right\|_1 + \left\|\frac{dn_1\pi_Y^{-1}}{d\nu} - s\right\|_1 < s\varepsilon.$$

PROPOSITION 58. *Let $\Lambda \subset (M, m)$ be a set of constant width s_Λ with respect to each of the decompositions ξ and η, and let $X_1 \subset X$ and $Y^1 \subset Y$ be subsets such that*

$$\mu X_1 + \nu Y^1 = 1, \quad m(\Lambda \cap (X_1 \times Y^1)) \leqslant \varepsilon s_\Lambda. \tag{30}$$

Then

$$\|\pi_Y(\chi_{(X_1 \times Y) \cap \Lambda}(x, y))(y) - s\chi_{CY^1}(y)\|_{L(Y, \nu)} \leqslant 2\varepsilon s_\Lambda. \tag{31}$$

PROOF. From the double stochasticity of the normalized restriction of m to the subset Λ it follows that

$$m(\Lambda \cap (X_1 \times Y^1)) = m(\Lambda \cap (CX_1 \times CY^1)). \tag{32}$$

Therefore

$$\begin{aligned}\|\pi_Y(\chi_{(X_1\times Y)\cap\Lambda})&-s\chi_{CY'}\|_{L(Y)}\\&=\|\pi_Y(\chi_{(X_1\times CY')\cap\Lambda})-s\chi_{CY'}+\pi_Y\chi_{(X_1\times Y')\cap\Lambda}\|_{L(Y)}\\&\leqslant\|\pi_Y(\chi_{(X_1\times CY')\cap\Lambda})-s\chi_{CY'}\|_{L(Y)}+\|\pi_Y\chi_{(X_1\times Y')\cap\Lambda}\|_{L(Y)}.\end{aligned}\tag{33}$$

By assumption, the second term of this sum does not exceed ϵs. To estimate the first term we observe that $s\chi_{CY'}=\pi_Y(\chi_{(X\times CY')\cap\Lambda})=\pi_Y(\chi_{(X_1\times CY')\cap\Lambda})+\pi_Y(\chi_{(CX_1\times CY')\cap\Lambda})$, from which, by (30) and (32), we get

$$\begin{aligned}\|\pi_Y(\chi_{(X_1\times CY')\cap\Lambda})-s\chi_{CY'}\|_{L(Y)}&=\|\pi_Y\chi_{(CX_1\times CY')\cap\Lambda}\|_{L(Y)}\\&=\|\chi_{CX_1\times CY')\cap\Lambda}\|_{L(M,\,m)}<\varepsilon s.\end{aligned}\tag{34}$$

The inequality (31) follows immediately from (33) and (34). ●

13. We now proceed to the proof of the existence of an independent complement of ξ and η when there exists a bounded density $k(x, y)$.

We say that a subset $\Lambda\subset(M, m)$ of constant width s_Λ with respect to each of ξ and η satisfies the condition $\mathcal{Y}(\epsilon; X_1, \dots, X_k; Y_1, \dots, Y_l)$, where $\epsilon>0$ and $X_1, \dots, X_k\subset X$, $Y_1, \dots, Y_l\subset Y$, if for each $i=1, \dots, k$ there is a subset $Y^i_\Lambda\subset Y$ for which $\nu Y^i_\Lambda=1-\mu X_i$ and

$$\|\pi_Y(\chi_{(X_i\times Y)\cap\Lambda})-s_\Lambda\chi_{CY^i_\Lambda}\|_{L(Y,\,\nu)}<\varepsilon s_\Lambda,\tag{35}$$

and for each $j=1, \dots, l$ there is a subset $X^j_\Lambda\subset X$ for which $\mu X^j_\Lambda=1-\nu Y_j$ and

$$\|\pi_X(\chi_{(X\times Y_j)\cap\Lambda})-s_\Lambda\chi_{CX^j_\Lambda}\|_{L(X,\,\mu)}<\varepsilon s_\Lambda.\tag{36}$$

We say that a measurable decomposition ζ of (M, m) into subsets of positive measure and constant width with respect to both the decompositions ξ and η satisfies the condition $\mathcal{Y}(\epsilon; X_1, \dots, X_k; Y_1, \dots, Y_l)$ if each of its elements (and there are not more than a countable number of them) satisfies the condition $\mathcal{Y}(\epsilon; X_1, \dots, X_k; Y_1, \dots, Y_l)$.

PROPOSITION 59. *Let $k(x, y)=dm/d(\mu\times\nu)<K<\infty$. Then for any $\epsilon>0$ and any collections of subsets $X_1, \dots, X_k\subset X$ and $Y_1, \dots, Y_k\subset Y$ there is a decomposition $\zeta(\epsilon; X_1, \dots, X_k; Y_1, \dots, Y_k)$ that is not more than countable and that satisfies the condition $\mathcal{Y}(\epsilon; X_1, \dots, X_k; Y_1, \dots, Y_k)$.*

PROOF. We carry out the proof by the method of complete induction.

I. $k=1$. We show that there exists a decomposition $\zeta(\epsilon; X_1; Y_1)$ with the required properties. Let $\epsilon_1=\epsilon/4>0$. By Theorem 7, we choose for the density $k(x, y)$ a constant s_0, and for ϵ_1 a corresponding $\delta>0$; we can assume that $\delta\leqslant\epsilon_1$.

By Proposition 57, for this δ (which plays the role of the ϵ in the formulation of Proposition 57) we find a corresponding $a>0$. Further, by Proposition 55, we construct for a (which plays the role of the ϵ in the formulation of Proposition 55) and the set X_1 a set $Y^1\subset Y$, $\nu Y^1=1-\mu X_1$, such that

$$\Pi_m (M \setminus (X_1 \times Y^1)) < 1 + a.$$

Then, by Proposition 57, we find a subset Λ that does not intersect $X_1 \times Y^1$ and whose widths with respect to ξ and η do not exceed some number $s > 0$, $s < \bar{s}$, where $\bar{s} < s_0$, and for which

$$\|\pi_X \chi_\Lambda - s\|_{L(X,\,\mu)} + \|\pi_Y \chi_\Lambda - s\|_{L(Y,\,\nu)} < \delta s.$$

By Theorem 7 there is a set $\Lambda_1 \supset \Lambda$ of constant width not exceeding $s(1+\epsilon_1)$ with respect to each of ξ and η.

The set Λ_1 intersects the set $X_1 \times Y^1$ in a subset whose measure does not exceed the difference of the measures of Λ_1 and Λ:

$$m(\Lambda_1 \cap (X_1 \times Y^1)) < s(1+\varepsilon_1) - (s - \delta s) = s(\varepsilon_1 + \delta) \leqslant 2\varepsilon_1 s = \frac{1}{2}\varepsilon s,$$
$$(m\Lambda_1 \leqslant s(1+\varepsilon_1),\ \ m\Lambda \geqslant s - \|\pi_X \chi_\Lambda - s\| > s - \delta s). \tag{37}$$

Then, by Proposition 58,

$$\|\pi_Y \chi_{(X_1 \times Y) \cap \Lambda_1} - s\chi_{CY^1}\| < \varepsilon s.$$

Next, we consider the set $\Lambda_1 \subset (M, m)$ as a subspace of the measure space (M, m), i.e., as an independent space with measure that is the normalized restriction m_{Λ_1} of m to the subset Λ_1. With this subspace we can now repeat the above argument, interchanging the sets X_1 and Y_1. We prove that there is a subset $\bar{\Lambda}_1 \subset \Lambda_1 \subset M$ satisfying the condition $\mathcal{Y}(\epsilon; X_1; Y_1)$. With this aim we show that there is a measurable decomposition θ of the space $(\Lambda_1, m_{\Lambda_1})$ into not more than countably many subsets $\tilde{\Lambda}$ of constant width with respect to each of ξ and η (more precisely, with respect to the traces of ξ and η on Λ_1) that satisfies the condition

$$\left\|\pi_X \chi_{(X \times Y_1) \cap \tilde{\Lambda}} - s_{\tilde{\Lambda}} \chi_{CX^1_{\tilde{\Lambda}}}\right\|_{L(X,\,\mu)} < \varepsilon s_{\tilde{\Lambda}}, \tag{38}$$

where, for each $\tilde{\Lambda} \in \theta$, $X^1_{\tilde{\Lambda}}$ is a subset of X such that $\mu X^1_{\tilde{\Lambda}} = 1 - \nu Y_1$.

We consider the set $\mathfrak{M}$ whose elements are pairs (κ, θ_κ), where κ is a subset of $(\Lambda_1, m_{\Lambda_1})$ of constant positive width with respect to both ξ and η, and θ_κ is a measurable decomposition of Λ_1 into subsets of positive measure and constant width with respect to ξ and η and whose elements Λ satisfy (38). On the set $\mathfrak{M}$ we consider a partial order structure, setting $(\kappa_1, \theta^{(1)}_{\kappa_1}) < (\kappa_2, \theta^{(2)}_{\kappa_2})$ if $\kappa_1 \subset \kappa_2$, the set κ_1 is $\theta^{(2)}_{\kappa_2}$-measurable, and $\theta^{(1)}_{\kappa_1}$ is the restriction of $\theta^{(2)}_{\kappa_2}$ to κ_1. The partially ordered set $\mathfrak{M}$ is inductive ([14], Chapter III, §2, no. 4, Definition 3), i.e., each of its totally ordered subsets has a majorant. Indeed, the majorant of a totally ordered family $\{(\kappa_\alpha, \theta^{(\alpha)}_{\kappa_\alpha}),\ \alpha \in A\}$ can be taken to be the element $(\bar{\kappa}, \bar{\theta}_{\bar{\kappa}})$, where $\bar{\kappa} = \bigcup_{\alpha \in A} \kappa_\alpha$, and $\bar{\theta}_{\bar{\kappa}}$ consists of all possible subsets in the various $\theta^{(\alpha)}_{\kappa_\alpha}$ (in view of the positivity of their measures, there cannot be more than a countable number of such subsets). By Zorn's lemma, $\mathfrak{M}$ has at least one maximal element $(\kappa_0, \theta^{(0)}_{\kappa_0})$. But in this maximal element the set κ_0 cannot be different from the whole space Λ_1, because, otherwise, considering the space $(\Lambda_1 \setminus \kappa_0, m_{\Lambda_1 \setminus \kappa_0})$ and arguing as above, it would be possible to select in it a subset

$\widetilde{\Lambda}\subset\Lambda_1\backslash\kappa_0$ of constant positive width with respect to each of ξ and η that satisfies (38), and this contradicts the maximality of $(\kappa_0, \theta^{(0)}_{\kappa_0})$: the set $\widetilde{\Lambda}$ can be added to κ_0, and on $\kappa_0 \cup \widetilde{\Lambda}$ we consider the coarsest decomposition that contains $\widetilde{\Lambda}$ and whose restriction to κ_0 coincides with $\theta^{(0)}_{\kappa_0}$.

We now prove that among the (not more than countably many) elements of the decomposition $\theta = \theta^{(0)}_{\kappa_0}$ there is a subset $\widetilde{\Lambda}$ for which

$$\left\| \pi_Y \chi_{(X_1\times Y)\cap\widetilde{\Lambda}} - s_{\widetilde{\Lambda}}\chi_{CY^1_{\widetilde{\Lambda}}} \right\| < \varepsilon s_{\widetilde{\Lambda}} \tag{39}$$

($s_{\widetilde{\Lambda}}$ is the width of $\widetilde{\Lambda}$, and $Y^1_{\widetilde{\Lambda}} \subset Y$ is a subset depending on $\widetilde{\Lambda}$ for which $\nu Y^1_{\widetilde{\Lambda}} = 1 - \mu X_1$). With this aim, we prove that for some element $\widetilde{\Lambda}$ of θ we have

$$m(\widetilde{\Lambda} \cap (X_1 \times Y^1)) < \frac{1}{2}\varepsilon s_{\widetilde{\Lambda}}. \tag{40}$$

If for each $\widetilde{\Lambda} \in \theta$ we had the opposite inequality

$$m(\widetilde{\Lambda} \cap (X_1 \times Y^1)) \geqslant \frac{1}{2}\varepsilon s_{\widetilde{\Lambda}},$$

then for their union $\Lambda_1 = \bigcup_{\widetilde{\Lambda}\in\theta}\widetilde{\Lambda}$ we would have

$$m(\bigcup_{\widetilde{\Lambda}\in\theta}\widetilde{\Lambda} \cap (X_1\times Y^1)) = \sum_{\widetilde{\Lambda}\in\theta} m(\widetilde{\Lambda}\cap(X_1\times Y^1)) \geqslant \frac{1}{2}\varepsilon\sum_{\widetilde{\Lambda}\in\theta} s_{\widetilde{\Lambda}} = \frac{1}{2}\varepsilon s,$$

and this contradicts (37). Consequently, there is a set $\widetilde{\Lambda} \subset M$ of constant positive width with respect to each of ξ and η and for which both (38) and (39) hold. To conclude the proof of the assertion that forms the basis of our induction argument it suffices to use Zorn's lemma a second time, repeating word for word the previous argument based on this lemma, with the single change that the decomposition θ_κ must now consist of subsets $\widetilde{\Lambda}$ of constant positive width and satisfying simultaneously both (38) and (39).

II. *Induction step.* We assume that Proposition 59 is proved for any $\epsilon > 0$ and for $k = 1, \ldots, n$, and we prove it for $k = n + 1$. Let $\epsilon_1 = \epsilon/4n$. We first prove that for each element $\widetilde{\Lambda}$ of the decomposition $\zeta(\epsilon; X_1, \ldots, X_n; Y_1, \ldots, Y_n)$, whose existence is assumed, we have

$$m(\widetilde{\Lambda}\cap(X_k\times Y^k_{\widetilde{\Lambda}})) < \varepsilon_1 s_{\widetilde{\Lambda}}, \quad k=1,\ldots,n, \tag{41}$$

$$m(\widetilde{\Lambda}\cap(X^k_{\widetilde{\Lambda}}\times Y_k)) < \varepsilon_1 s_{\widetilde{\Lambda}}, \quad k=1,\ldots,n. \tag{42}$$

Indeed, by hypothesis, for each element $\widetilde{\Lambda} = \Lambda$ we have (35) and (36), from which it follows that

$$m(\widetilde{\Lambda}\cap(X_k\times Y^k_{\widetilde{\Lambda}})) = m(\widetilde{\Lambda}\cap(CX_k\times CY^k_{\widetilde{\Lambda}})) < \varepsilon_1 s_{\widetilde{\Lambda}}, \quad k=1,\ldots,n,$$

$$m(\widetilde{\Lambda}\cap(X^k_{\widetilde{\Lambda}}\times Y_k)) = m(\widetilde{\Lambda}\cap(CX^k_{\widetilde{\Lambda}}\times CY_k)) < \varepsilon_1 s_{\widetilde{\Lambda}}, \quad k=1,\ldots,n.$$

We now choose some element $\widetilde{\Lambda}_0 \in \zeta(\epsilon_1; X_1, \ldots, X_n; Y_1, \ldots, Y_n)$ of positive measure, and for the subspace $(\widetilde{\Lambda}_0, m_{\widetilde{\Lambda}_0})$ construct a decomposition $\zeta_{\widetilde{\Lambda}_0} = \zeta_{\widetilde{\Lambda}_0}(\epsilon; X_{n+1}; Y_{n+1})$, just as in part I, such that for each element $\widetilde{\widetilde{\Lambda}} \in \zeta_{\widetilde{\Lambda}_0}$ we have

$$\left\| \pi_Y (\chi_{(X_{n+1}\times Y)\cap\tilde{\Lambda}}) - s_{\tilde{\Lambda}}\chi_{C Y^{n+1}_{\tilde{\Lambda}}} \right\|_{L(Y,\,\nu)} < \varepsilon s_{\tilde{\Lambda}}, \tag{43}$$

$$\left\| \pi_X (\chi_{(X\times Y_{n+1})\cap\tilde{\Lambda}}) - s_{\tilde{\Lambda}}\chi_{C X^{n+1}_{\tilde{\Lambda}}} \right\|_{L(X,\,\mu)} < \varepsilon s_{\tilde{\Lambda}}. \tag{44}$$

We prove that among the elements of this decomposition there is at least one $\tilde{\tilde{\Lambda}} \in \zeta_{\tilde{\Lambda}_0}$ of positive measure that satisfies the condition $\mathcal{Y}(\epsilon; X_1, \ldots, X_{n+1}; Y_1, \ldots, Y_{n+1})$. It suffices to show that there is a $\tilde{\tilde{\Lambda}} \in \zeta_{\tilde{\Lambda}_0}$ satisfying the condition $\mathcal{Y}(\epsilon; X_1, \ldots, X_n; Y_1, \ldots, Y_n)$. For this we show that there is a $\tilde{\tilde{\Lambda}} \in \zeta_{\tilde{\Lambda}_0}$ for which

$$m(\tilde{\tilde{\Lambda}} \cap (X_k \times Y^k_{\tilde{\Lambda}})) < \frac{1}{2}\varepsilon s_{\Lambda}, \quad k = 1, \ldots, n, \tag{45}$$

$$m(\tilde{\tilde{\Lambda}} \cap (X^k_{\tilde{\Lambda}} \times Y_k)) < \frac{1}{2}\varepsilon s_{\Lambda}, \quad k = 1, \ldots, n, \tag{46}$$

where $s_{\tilde{\tilde{\Lambda}}}$ is the width of $\tilde{\tilde{\Lambda}}$ (with respect to the measure m). Assume, on the contrary, that for each $\tilde{\tilde{\Lambda}} \in \zeta_{\tilde{\Lambda}_0}$ we can find an inequality opposite to one of the $2n$ inequalities (45) and (46):

$$m(\tilde{\tilde{\Lambda}} \cap (X_{\bar{k}} \times Y^{\bar{k}}_{\tilde{\Lambda}})) \geqslant \frac{1}{2}\varepsilon s_{\Lambda} \tag{47}$$

or

$$m(\tilde{\tilde{\Lambda}} \cap (X^{\bar{k}}_{\tilde{\Lambda}} \times Y_{\bar{k}})) \geqslant \frac{1}{2}\varepsilon s_{\Lambda}. \tag{48}$$

Let

$$Z = \bigcup_{k=1}^{n} (X_k \times Y^k_{\tilde{\Lambda}}) \cup \bigcup_{k=1}^{n} (X^k_{\tilde{\Lambda}} \times Y_k).$$

By (41) and (42),

$$m(\tilde{\Lambda}_0 \cap Z) < 2n\varepsilon_1 s_{\tilde{\Lambda}_0} = \frac{1}{2}\varepsilon s_{\tilde{\Lambda}_0}. \tag{49}$$

On the other hand, from (47) and (48) we get for each $\tilde{\tilde{\Lambda}} \in \zeta_{\tilde{\Lambda}_0}$ that

$$m(\tilde{\tilde{\Lambda}} \cap Z) \geqslant \frac{1}{2}\varepsilon s_{\Lambda}. \tag{50}$$

Combining all the inequalities (50) and using the fact that

$$\bigcup_{\tilde{\tilde{\Lambda}} \in \zeta_{\tilde{\Lambda}_0}} \tilde{\tilde{\Lambda}} = \tilde{\Lambda}_0, \quad \sum_{\tilde{\tilde{\Lambda}} \in \zeta_{\tilde{\Lambda}_0}} s_{\tilde{\tilde{\Lambda}}} = s_{\tilde{\Lambda}_0},$$

we find

$$\sum_{\tilde{\tilde{\Lambda}} \in \zeta_{\tilde{\Lambda}_0}} m(\tilde{\tilde{\Lambda}} \cap Z) = m(\tilde{\Lambda}_0 \cap Z) \geqslant \frac{1}{2}\varepsilon s_{\tilde{\Lambda}_0},$$

which contradicts (49); by the same token we have proved the existence of a $\tilde{\tilde{\Lambda}} \in \zeta_{\tilde{\Lambda}_0}$ of positive measure for which the $2n$ inequalities (45) and (46) and the inequalities (43) and (44) hold, and with this, by Proposition 58, we have also the condition $\mathcal{Y}(\epsilon; X_1, \ldots, X_{n+1}; Y_1, \ldots, Y_{n+1})$.

To finish the proof of the proposition it now suffices to remark that the naturally ordered collection of subsets of M of constant width with respect to ξ and η and for which there exist decompositions having the property $\mathcal{Y}(\epsilon; X_1, \ldots, X_{n+1}; Y_1, \ldots, Y_{n+1})$ is, by what was proved, not empty, and it is obviously inductive. A maximal element (which exists, by Zorn's lemma) cannot fail, by the above, to be the whole space M, equipped with the required decomposition having the property $\mathcal{Y}(\epsilon; X_1, \ldots, X_{n+1}; Y_1, \ldots, Y_{n+1})$. (The necessary argument by means of Zorn's lemma was carried out in more detail in part I: from the existence of some subset of positive measure having some property we get the existence of a decomposition of the whole space into subsets of positive measure, each element of which also has this property.) ●

If in the proof of Proposition 59 we use Propositions 55*, 57*, and Theorem 7* instead of Propositions 55, 57, and Theorem 7, we get the following improvement of Proposition 59.

PROPOSITION 59*. *Let the measure m be absolutely continuous with respect to the measure m^* in Proposition 44*, and let the density $k(x, y) = dm/dm^*$ be bounded on each element of $\xi \wedge \eta$. Then for any $\epsilon > 0$ and any collection of subsets $X_1, \ldots, X_k \subset X$ and $Y_1, \ldots, Y_k \subset Y$ there is a decomposition $\zeta(\epsilon; X_1, \ldots, X_k; Y_1, \ldots, Y_k)$ that is not more than countable and satisfies the condition $\mathcal{Y}(\epsilon; X_1, \ldots, X_k; Y_1, \ldots, Y_k)$.*

Instead of (37) the right-hand side of the inequality must contain the measure of the symmetric difference of the sets Λ_1 and Λ estimated in Theorem 7*, so we should set $\epsilon_1 = \epsilon/10$. Part II of the proof of Proposition 59 can be carried over without change. ●

We now introduce the following notation for the truncation of a function. We set

$$k_N(x, y) = \begin{cases} k(x, y), & \text{when } k(x, y) \leqslant N, \\ 0, & \text{when } k(x, y) > N, \end{cases}$$

and $k^N(x, y) = k(x, y) - k_N(x, y)$.

PROPOSITION 60. *For any doubly stochastic density $k(x, y)$ and any $\epsilon > 0$ there is a measurable subset $M_1 \subset M$ such that $mM_1 > 1 - \epsilon$, and on the subspace (M_1, m_{M_1}) the density of the measure m_{M_1} with respect to the product measure $m_{M_1}/\xi \times m_{M_1}/\eta$ is bounded (m_{M_1}/ξ and m_{M_1}/η are the canonical projections of m_{M_1} onto $(M_1, m_{M_1})/\xi$ and $(M_1, m_{M_1})/\eta$; it is not required that M_1 is a set of constant width).*

PROOF. We choose a number N so that

$$\int_M k_N(x, y)\, dm > 1 - \frac{\epsilon}{4}.$$

We consider the functions

$$\alpha(x) = \int k_N(x, y)\, d\nu = \pi_X \chi_{\{(x, y): k(x, y) \leqslant N\}}(x, y)$$

and

$$\beta(y) = \int k_N(x, y)\, d\mu = \pi_Y \chi_{\{(x, y): k(x, y) \leqslant N\}}(x, y).$$

Let $X' \subset X$ be a subset such that for some number $c' > 0$ we have $\alpha(x) > c'$ for $x \in X'$, and $\mu X' > 1 - \epsilon/4$. We choose a subset $Y' \subset Y$ and a positive number c'' such that

$$\int_{X'} k_N(x, y)\, d\mu > c'' \quad \text{for} \quad y \in Y',$$

$$\nu Y' > 1 - \frac{\varepsilon}{2} \quad \text{and} \quad c''N < c'.$$

We see that we can take M_1 to be the set

$$M_1 = (X' \times Y') \cap \{(x, y): k(x, y) \leqslant N\}.$$

Indeed,

$$mM_1 > 1 - \frac{\varepsilon}{4} - \left(1 - \left(1 - \frac{\varepsilon}{4}\right)\left(1 - \frac{\varepsilon}{2}\right)\right) > 1 - \varepsilon.$$

Further,

$$\int_{X'} k_N(x, y)\, d\mu \geqslant c'' \quad \text{for} \quad y \in Y',$$

$$\int_{Y'} k_N(x, y)\, d\nu = \int k_N(x, y)\, d\nu - \int_{Y \setminus Y'} k_N(x, y)\, d\nu > c' - c''N > 0 \quad \text{for } x \in X'.$$

Consequently, for the density $k_1(x, y)$ of the measure m_{M_1} with respect to the product $m_{M_1}/\xi \times m_{M_1}/\eta$ on $M_1/\xi \times M_1/\eta \subset X' \times Y'$ we have

$$k_1(x, y) = \frac{k_N(x, y)}{\int_{X'} k_N(x, y)\, d\mu \int_{Y'} k_N(x, y)\, d\nu} < \frac{N}{c'(c' - c''N)} < \infty. \bullet$$

PROPOSITION 60.* *For any doubly stochastic measure m that is absolutely continuous with respect to the measure m^* (defined in Proposition 44*) and any $\epsilon > 0$ there is a measurable subset $M_1 \subset M$ such that $mM_1 > 1 - \epsilon$, and on the subset (M_1, m_{M_1}) the density of m_{M_1} with respect to $m^*_{M_1}$ is bounded on each element of the decomposition $\xi \wedge \eta$ of $(M_1, m^*_{M_1})$.*

The proof is easily obtained from the proof of Proposition 60, replacing $\int_{X'} k_N(x, y) d\mu$ and $\int_{Y'} k_N(x, y) d\nu$ by the expressions

$$\pi_X \chi_{\{(x, y): k(x, y) \leqslant N(C(x, y)),\, x \in X'\}} \quad \text{and} \quad \pi_Y \chi_{\{(x, y): k(x, y) \leqslant N(C(x, y)),\, y \in Y'\}},$$

where $C(x, y)$ is the element of $\xi \wedge \eta$ containing the point (x, y); instead of the constant N we use the function $N(C)$, which is measurable with respect to $\xi \wedge \eta$. It is also possible to obtain the uniform boundedness of the density $dm_{M_1}/dm^*_{M_1}$. $\bullet$

14. By means of Proposition 59 it is already possible to prove the existence of an independent complement of the coordinate decompositions ξ and η under the assumption of a bounded density $k(x, y)$. However, our immediate goal is to dispense with this assumption.

If we take a subset M_1 such that on the subspace (M_1, m_{M_1}) the density with respect to the product measure $m_{M_1}/\xi \times m_{M_1}/\eta$ is bounded (see Proposition 60) and construct a decomposition that is an independent complement of the coordinate decompositions ξ and η with respect to m_{M_1}, then this decomposition will not be independent with respect to ξ and η and the original measure (without the specified refinements it will not even be a decomposition of the whole space (M, m)). However, if we are concerned not about the complementation property, but only about the independence, then, as we know, such independent and sufficiently fine decompositions exist even without the assumption of a bounded density. Therefore, we can hope to construct the required independent complement in the case of an arbitrary density by approximating it with decompositions of the whole space M into subsets of positive measure that are made up of two parts: a more massive part that approximates an element of some decomposition that is an independent complement of the pair ξ, η with respect to the measure $m_{M_1(\epsilon)}$, and a less massive part whose contribution goes to 0 in the course of the approximation and that is independent of ξ and η with respect to a "truncated" part of the measure and that makes the relevant subset independent with respect to the coordinate decompositions. We proceed to the rigorous realization of this plan.

PROPOSITION 61. *Suppose that the doubly stochastic measure m is absolutely continuous with respect to the measure $\mu \times \nu$. Then for any $\epsilon > 0$ and any collection of subsets $X_1, \dots, X_n \subset X$, $Y_1, \dots, Y_n \subset Y$ there is a decomposition $\zeta(\epsilon; X_1, \dots, X_n; Y_1, \dots, Y_n)$ that is not more than countable and that satisfies the condition $\mathcal{Y}(\epsilon; X_1, \dots, X_n; Y_1, \dots, Y_n)$.*

PROOF. We set $\epsilon_1 = \epsilon/4$ and construct, using Proposition 60, a subset $M_1 \subset M = X \times Y$ such that $mM_1 > 1 - \epsilon_1$ and the measure m_{M_1} on the space (M, m_{M_1}) is such that its density $k_1(x, y)$ with respect to the product of its projections m_{M_1}/ξ and m_{M_1}/η onto X and Y is bounded by some constant $K_1 < \infty$. Then, using Proposition 59, we construct a decomposition $\tilde{\zeta}$ of (M_1, m_{M_1}) that is independent with respect to ξ and η and satisfies the condition $\mathcal{Y}(\epsilon; X_1, \dots, X_n; Y_1, \dots, Y_n)$ in this space. Let $M_2 = M \setminus M_1$. We consider the space (M_2, m_{M_2}), or (M, m_{M_2}). The measure m_{M_2} on $M = X \times Y$ is absolutely continuous with respect to the product $\mu \times \nu$, and consequently also with respect to the product of its projections m_{M_2}/ξ and m_{M_2}/η; therefore, by the corollary of Proposition 44, there exists a decomposition $\tilde{\tilde{\zeta}}$ of (M_2, m_{M_2}) into subsets of constant width with respect to ξ and η such that the discrete measure space $(M_2, m_{M_2})/\tilde{\tilde{\zeta}}$ is isomorphic to $(M_1, m_{M_1})/\tilde{\zeta}$.

Now let $\tilde{\Lambda} \subset M_1$ be an element of $\tilde{\zeta}$ and $\tilde{\tilde{\Lambda}} \subset M_2$ the element of $\tilde{\tilde{\zeta}}$ that corresponds to it under this isomorphism. Let $\Lambda = \tilde{\Lambda} \cup \tilde{\tilde{\Lambda}}$, and let ζ be the measurable decomposition of (M, m) into all possible subsets of this form. From the fact that $\tilde{\Lambda}$

has width $s_{\tilde{\Lambda}}^{M_1} = s$ with respect to ξ and η and m_{M_1}, while $\tilde{\tilde{\Lambda}}$ has constant width $s_{\tilde{\tilde{\Lambda}}}^{M_2}$ equal to the same number $\bar{s} > 0$ (by the isomorphism) with respect to m_{M_2}, and from the relation $\tilde{\Lambda} \cap \tilde{\tilde{\Lambda}} \subset M_1 \cap M_2 = \varnothing$ it follows the set $\Lambda = \tilde{\Lambda} \cup \tilde{\tilde{\Lambda}}$ has constant width equal to s with respect to any convex combination of m_{M_1} and m_{M_2}; in particular, with respect to m.

For some k, $1 \leqslant k \leqslant n$, we consider the set X_k and an arbitrary element $\Lambda \in \zeta$, $\Lambda = \tilde{\Lambda} \cup \tilde{\tilde{\Lambda}}$. Since $\tilde{\zeta}$ satisfies the condition $\mathcal{Y}(\epsilon_1; X_k)$, it follows from (41) that

$$m_{M_1}(\tilde{\Lambda} \cap (X_k \times Y_{\tilde{\Lambda}}^k)) < \varepsilon_1 s_{\tilde{\Lambda}}^{M_1},$$

and hence, by the choice of the set,

$$m(\tilde{\Lambda} \cap (X_k \times Y_{\tilde{\Lambda}}^k)) < (1 - \varepsilon_1)\, \varepsilon_1 s_{\tilde{\Lambda}}^{M_1} < \varepsilon_1 s_{\tilde{\Lambda}}^{M_1}. \tag{51}$$

Furthermore, for the same reason,

$$m(\tilde{\tilde{\Lambda}} \cap (X_k \times Y_{\tilde{\Lambda}}^k)) \leqslant m\tilde{\tilde{\Lambda}} \leqslant \varepsilon_1 m_{M_2}\tilde{\tilde{\Lambda}} = \varepsilon_1 s_{\tilde{\tilde{\Lambda}}}^{M_2}, \tag{52}$$

and, combining (51) and (52), we get

$$m(\Lambda \cap (X_k \times Y_{\tilde{\Lambda}}^k)) < 2\varepsilon_1 s_{\Lambda}. \tag{53}$$

Instead of $Y_{\tilde{\Lambda}}^k$ it is now better to write Y_{Λ}^k. By Proposition 58, the set Λ of constant width s satisfies, by (53), the condition $\mathcal{Y}(4\epsilon_1; X_k) = \mathcal{Y}(\epsilon; X_k)$. Repeating the same argument with each of the remaining subsets $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$, we conclude the proof of the assertion. ●

We mention that the assumption of absolute continuity of m with respect to $\mu \times \nu$ was used only so that we could use Proposition 60. As is shown by Proposition 60*, the absolutely continuity of m with respect to $\mu \times \nu$ can be replaced by the absolute continuity of m with respect to m^*. Therefore, the following generalization of Proposition 61 holds.

PROPOSITION 61*. *Suppose that the doubly stochastic measure m is absolutely continuous with respect to the measure m^* (defined in Proposition 44*). Then for any $\epsilon > 0$ and any collection of subsets $X_1, \ldots, X_n \subset X$, $Y_1, \ldots, Y_n \subset Y$ there is a decomposition $\zeta(\epsilon; X_1, \ldots, X_n; Y_1, \ldots, Y_n)$.*

We can prove the central theorem of this chapter.

THEOREM 8. *Let (M, m) be a Lebesgue space with nonatomic measure, and ξ and η measurable decompositions of (M, m) such that $\xi \vee \eta = \epsilon$ and the image of m under the canonical imbedding $M \to M/\xi \times M/\eta$ is absolutely continuous with respect to the product of the canonically defined measures $\mu = m/\xi$ and $\nu = m/\eta$ on M/ξ and M/η. Then there exists a measurable decomposition ζ that is an independent complement of ξ and of η.*

PROOF. Let $\{X_k, k = 1, \ldots\}$ and $\{Y_k, k = 1, \ldots\}$ be bases of the spaces $(X, \mu) = (M/\xi, m/\xi)$ and $(Y, \nu) = (M/\eta, m/\eta)$. We construct a sequence of decompositions ζ_n, $n = 1, \ldots$, in the following way. Let $\epsilon_k \searrow 0$ be a sequence of positive numbers. Let $\zeta_1 = \zeta(\epsilon; X_1; Y_1)$ be a decomposition satisfying the condition $\mathcal{Y}(\epsilon_1; X_1; Y_1)$ (Proposition 61). We now consider each element Λ of ζ_1 as an independent subspace (Λ, m_Λ) of (M, m) and construct a decomposition ζ_2^Λ of this subspace satisfying the condition $\mathcal{Y}(\epsilon_2; X_1, X_2; Y_1, Y_2)$.

The collection of all elements of ζ_2^Λ for all $\Lambda \in \zeta_1$ forms a measurable decomposition of (M, m), which we denote by ζ_2. We continue this construction ad infinitum, using Proposition 61 each time, decreasing ϵ and adding a new pair of sets from the given bases $\{X_k\}$ and $\{Y_k\}$ each time. We get a refining sequence of decompositions ζ_k, $k = 1, \ldots, \zeta_1 < \zeta_2 < \cdots$, each of which is not more than countable and is independent with respect to each of the coordinate decompositions ξ and η. For each $k = 1, \ldots$, the decomposition ζ_k satisfies the condition $\mathcal{Y}(\epsilon_k; X_1, \ldots, X_k; Y_1, \ldots, Y_k)$.

We now prove that the decomposition $\zeta = \bigvee \zeta_k$ is the required one.

Since each ζ_k is independent with respect to ξ and η, the same is true for their limit ζ. It remains to verify that $\xi \vee \zeta = \epsilon$ and $\eta \vee \zeta = \epsilon$. For this, we show that for a typical element Λ of ζ the complementation criterion in Proposition 48 is satisfied. Let $\Lambda = \bigcap_1^\infty \Lambda_k$, where $\Lambda_k = \Lambda_k(\Lambda) \in \zeta_k$, and let $m_k = m_{\Lambda_k}$ be the conditional measure on Λ_k. We consider an arbitrary element X_n of the chosen basis. By the definition of ζ_k, for $k \geqslant n$ we have

$$\| \pi_Y \chi_{(X_n \times Y) \cap \Lambda_k} - s_{\Lambda_k} \chi_{CY^n_{\Lambda_k}} \|_{L(Y,\nu)} < \epsilon_k s_{\Lambda_k}, \tag{54}$$

where $Y^n_{\Lambda_k} \subset Y$ is some subset for which $\mu X_n + \nu Y^n_{\Lambda_k} = 1$. We verify that the following convergence holds in the norm of L:

$$\frac{1}{\mu X_n s_{\Lambda_k}} (\pi_Y \chi_{(X_n \times Y) \cap \Lambda_k})(y) \xrightarrow[k \to \infty]{} (\pi_Y m_{(X_n \times Y) \cap \Lambda})(y), \tag{55}$$

where $m_{(X_n \times Y) \cap \Lambda}$ is the measure on the subspace $(X_n \times Y) \cap \Lambda$ of (Λ, m_Λ). Indeed, we consider the subset $X_n \times Y \subset M$ as an independent subspace with the measure $m_{X_n \times Y}$, and consider its quotient space $(X_n \times Y, m_{X_n \times Y})/(\eta \vee \xi)$. Let the elements of this quotient space be assigned the coordinates $(\dot{y}, \Lambda)$, where $\dot{y}$ denotes the corresponding element $X_n \times y$ of the decomposition η, and Λ is the corresponding element of ζ. Let $\widetilde{k}(\dot{y}, \Lambda)$ be the density of the image of $m_{X_n \times Y}$ under this homomorphism with respect to the product of the measures $m^\eta_{X_n \times Y}$ and $m^\zeta_{X_n \times Y}$ that are canonically determined on $(X_n \times Y)/\eta$ and $(X_n \times Y)/\zeta$. We define a sequence of functions $\widetilde{k}_k(\dot{y}, \Lambda)$, setting $\widetilde{k}_k(y, \Lambda)$ for fixed $\dot{y}$ equal to the mean value of the function $k(\dot{y}, \Lambda)$ over the set of values for Λ contained in the corresponding element $\Lambda_k \supset \Lambda$ of ζ_k. It is well known that for an integrable function the sequence of such averages (i.e., conditional mathematical expectations with respect to an infinite refining sequence of σ-algebras) converges to the function itself in the mean (in the norm of L) (see, for example, [24], Chapter VII, §4, Theorem 4.1).

$$\widetilde{k}(\dot{y}, \Lambda) = (\pi_Y m_{(X_n \times Y) \cap \Lambda})(y),$$

and

$$k_k(y, \Lambda) = \frac{1}{\mu X_n s_{\Lambda_k}} (\pi_Y \chi_{(X_n \times Y) \cap \Lambda_k})(y),$$

where $\Lambda \subset \Lambda_k$, from which (55) follows.

In a completely analogous way it is shown that for each $n = 1, \ldots$ we have convergence in the norm of $L(X, \mu)$:

$$\frac{1}{\nu Y_n s_{\Lambda_k}} (\pi_X \chi_{(X \times Y_n) \cap \Lambda_k})(x) \xrightarrow[k \to \infty]{} (\pi_X m_{(X \times Y_n) \cap \Lambda})(x), \tag{56}$$

where $m_{(X \times Y_n) \cap \Lambda}$ is the measure on $(X \times Y_n) \cap \Lambda$, regarded as a subspace of (Λ, m_Λ).

From (54) and (55) it follows that in the metric of $L(Y, \nu)$ and for each $n = 1, \ldots$

$$\chi_{CY^n_{\Lambda_k}}(y) \xrightarrow[k \to \infty]{} \mu X_n (\pi_Y m_{(X_n \times Y) \cap \Lambda})(y),$$

and from (56) and the analogue of (54) it follows that in the metric of $L(X, \nu)$ and for each $n = 1, \ldots$

$$\chi_{CX^n_{\Lambda_k}}(x) \xrightarrow[k \to \infty]{} \nu Y_n (\pi_X m_{(X \times Y_n) \cap \Lambda})(x).$$

But a sequence of characteristic functions of subsets of constant measure cannot converge in the norm of L to anything other than the characteristic function of some subset of the same measure; consequently, for each $n = 1, \ldots$ the functions

$$\nu Y_n \pi_X m_{(X \times Y_n) \cap \Lambda}(x) \quad \text{and} \quad \mu X_n \pi_Y m_{(X_n \times Y) \cap \Lambda}(y)$$

are all characteristic functions of some subsets. This means that for the conditional measure m_Λ the criterion formulated in Proposition 48 holds; consequently m_Λ is the kernel of some isomorphism of the spaces (X, μ) and (Y, ν), and the decomposition ζ really is an independent complement of the decompositions ξ and η (see Proposition 42). ●

If instead of Proposition 61, which is the basis of the proof of Theorem 8, we use the more general Proposition 61*, we arrive at the following modification of Theorem 8, which is useful for subsequent applications.

THEOREM 8*. *Let (M, m) be a Lebesgue space with nonatomic measure, and ξ and η measurable decompositions of this space satisfying the following conditions:*

1) $\xi \vee \eta = \epsilon$.

2) *If ξ_η and η_ξ denote the measurable decompositions of the spaces $X = M/\xi$ and $Y = M/\eta$ induced by the canonical mapping $M/\xi \longrightarrow M/(\xi \wedge \eta)$ and $M/\eta \longrightarrow M/(\xi \wedge \eta)$, then the conditional measure on each element of the decomposition $\xi \wedge \eta$ is absolutely continuous with respect to the product of the conditional measures on those elements of ξ_η and η_ξ that are carried under these canonical mappings into this element of $\xi \wedge \eta$ (i.e., $m \ll m^*$).*

3) *The conditional measures on the elements of ξ_η and η_ξ are purely continuous.*

Then there exists a decomposition ζ that is an independent complement of ξ and η.

We emphasize that when the decomposition $\xi \wedge \eta$ contains a continuous component (i.e., the measure $m/(\xi \wedge \eta)$ is not purely atomic), then the measure m is a fortiori not absolutely continuous with respect to the product of the measures $\mu = m/\xi$ and $\nu = m/\eta$.

15. Theorem 8* permits the description of many situations in which a space with a doubly stochastic measure provided with certain other structures has an independent complement of a pair of given decompositions. For example, when the measure m is given on an affine subspace of some finite-dimensional space and is absolutely continuous with respect to Lebesgue measure on this subspace, we have the following assertion.

PROPOSITION 62. *Let $E_1 = \mathbf{R}^{n_1}$ and $E_2 = \mathbf{R}^{n_2}$, and let $L \subset E_1 \times E_2$ be an affine subspace of dimension n on which a measure m is defined and absolutely continuous with respect to Lebesgue measure on L. Then the space $(E_1 \times E_2, m)$ admits a complement ζ that is independent with respect to the coordinate decompositions, if*

$$\dim (E_1 \times \{0\}) \cap L > 0 \quad \textit{and} \quad \dim (\{0\} \times E_2) \cap L > 0.$$

PROOF. It suffices to limit ourselves to the case when $\pi_{E_1} L = E_1$ and $\pi_{E_2} L = E_2$, otherwise taking the space $\pi_{E_1} L \times \pi_{E_2} L$ instead of $E_1 \times E_2$. For convenience, we can also assume that L is a linear subspace, otherwise translating L parallel to itself to the origin. Let

$$L_1 = (E_1 \times \{0\}) \cap L, \quad L_2 = (\{0\} \times E_2) \cap L.$$

If m^* is Lebesgue measure on L, and ξ and η are the coordinate decompositions, then $\xi \wedge \eta$ is the decomposition of L into cosets with respect to the subspace $L_1 + L_2$, and the decompositions ξ_η and η_ξ are the decompositions of E_1 and E_2, identified with $E_1 \times \{0\}$ and $\{0\} \times E_2$, into cosets with respect to L_1 and L_2, respectively. Since $\dim(L_1 + L_2) = \dim L_1 + \dim L_2$, the conditional measure on the elements of $\xi \wedge \eta$, i.e., the Lebesgue measure on each of the affine subspaces of L parallel to $L_1 + L_2$, is the product of the conditional measures on the elements of ξ_η and η_ξ. With that, we are in the setting of Theorem 8*, and this concludes the proof of Proposition 62. ●

With a view to formal completeness we say something about the cases when $\xi \vee \eta \neq \epsilon$ and when the number of factors in the product space is greater than two.

PROPOSITION 63. *Suppose that the decompositions ξ and η of the space (M, m) are such that the measure $m(\pi_X \times \pi_Y)^{-1}$ defined on $X \times Y$ $(X = M/\xi;\ Y = M/\eta)$ is absolutely continuous with respect to $m/\xi \times m/\eta$, and the measures m/ξ and m/η are purely continuous. For there to exist a decomposition ζ that is an independent complement of ξ and η it is sufficient that $\xi \vee \eta$ admits an independent complement.*

PROOF. Let ζ_1 be an independent complement of the coordinate decompositions of the space $(M/\xi \times M/\eta, m(\pi_X \times \pi_Y)^{-1})$, which exists, by Theorem 8. If ζ_2 is an independent complement of $\xi \vee \eta$, and ζ_3 is the decomposition of (M, m) that is the preimage of ζ_1 under the canonical mapping $M \longrightarrow M/\xi \times M/\eta$, then the decomposition $\zeta = \zeta_2 \vee \zeta_3$ is an independent complement of ξ and η. Indeed, $\zeta \vee \xi = \zeta_2 \vee \zeta_3 \vee \zeta = \zeta_2 \vee (\zeta_3 \vee \xi) = \zeta_2 \vee (\xi \vee \eta) = \epsilon$, and, similarly, $\zeta \vee \eta = \epsilon$. The independence of ζ and ξ follows from the fact that ζ_3 is independent with respect to ξ, and the conditional measure on each element C of ζ_3, identified with the product of an element of ζ_1 and an element of ζ_2, is the product of the conditional measures on these elements (the independence of ζ_2 and $\xi \vee \eta$). Therefore, under the canonical projection $M \longrightarrow M/\xi$ the conditional measure on each element of $\zeta_2 \vee \zeta_3$ is projected into the same measure as the conditional measure of the whole element of ζ_3. The desired conclusion follows from the independence of ζ_3 and ξ. The independence of ζ and η is proved in a similar way. ●

We remark that the existence of an independent complement of the decomposition $\xi \vee \eta$ is not a necessary condition for the existence of an independent complement of ξ and η. We give a "discrete" example that clarifies this assertion; an example with purely continuous m/ξ and m/η can also be obtained from it without difficulty.

EXAMPLE. The measure space (M, m) consists of 27 points a_{ikl}, $i, k, l = 1, 2, 3$, $m(\{a_{ikl}\}) = p_{ikl}$, where

$$\begin{aligned} &p_{111}=p_{222}=p_{333}, \quad p_{121}=p_{332}=p_{213}, \quad p_{131}=p_{312}=p_{223},\\ &p_{211}=p_{322}=p_{133}, \quad p_{221}=p_{132}=p_{313}, \quad p_{231}=p_{112}=p_{323},\\ &p_{311}=p_{122}=p_{233}, \quad p_{321}=p_{232}=p_{113}, \quad p_{331}=p_{212}=p_{123}. \end{aligned} \tag{57}$$

Moreover, $\Sigma_{i,l} p_{ikl} = \Sigma_{k,l} p_{ikl} = 1/3$, $p_{111} = p_{112} = p_{113}$, and the numbers p_{121}, p_{122}, and p_{123} are not all equal to one another. It is easy to see that if ξ and η are the decompositions generated by the mappings $a_{ikl} \longmapsto i$ and $a_{ikl} \longmapsto k$, then the decomposition of the whole space into the nine subsets consisting each of the three points a_{ikl} whose indices coincide with those of the probabilities p_{ikl} in each of the nine groups of equalities (57) is an independent complement of ξ and η, but the decomposition $\xi \vee \eta$, which is generated by the mapping $a_{ikl} \longmapsto l$, trivially does not admit an independent complement, because of the conditions $p_{111} = p_{112} = p_{113}$ and $|p_{121} - p_{122}| + |p_{121} - p_{123}| + |p_{122} - p_{123}| > 0$, which imply that the conditional measures on two elements of $\xi \vee \eta$ are not isomorphic.

Finally, we mention the case when an independent complement of a number of decompositions $\xi_1, \ldots, \xi_n$ is to be found. Assuming the condition $\xi_1 \vee \cdots \vee \xi_n = \epsilon$, we can suppose that we are dealing with a "multiply stochastic" measure defined on a subset of the unit cube in $\mathbf{R}^n$. In contrast to the case $n = 2$, the analogue of the Birkhoff-von Neumann theorem does not hold even for $n = 3$, for cubic matrices of dimension $2 \times 2 \times 2$. As a counterexample we consider the "triply stochastic" $2 \times 2 \times 2$ matrix (a_{ijk}) consisting of zeros and ones, for which $a_{000} = a_{100} = a_{101} =$

$a_{011} = \frac{1}{4}$ and $a_{100} = a_{010} = a_{001} = a_{111} = 0$. It is immediately clear that such a matrix is itself an extreme point of the set of all "triply stochastic" $2 \times 2 \times 2$ matrices. However, we consider the "triply stochastic" $4 \times 4 \times 4$ matrix (b_{ijk}) for which

$$b_{ijk} = \frac{1}{32} \quad \text{if} \quad a_{\left[\frac{i}{3}\right]\left[\frac{j}{3}\right]\left[\frac{k}{3}\right]} = \frac{1}{4}$$

and $b_{ijk} = 0$ otherwise, i.e., the matrix obtained from (a_{ijk}) by refinement of each of its elements into eight equal parts distributed in the cells of the $2 \times 2 \times 2$ matrix, which now takes the place of the original element (a_{ijk}). It can be shown that such a matrix can be represented in the form of a convex combination of triply stochastic $4 \times 4 \times 4$ (0, 1)-matrices. From this, in particular, it follows that if on the unit cube $M = \{0 \leqslant x, y, z \leqslant 1\} \subset \mathbf{R}^3$ we consider the triply stochastic measure m that is absolutely continuous with respect to Lebesgue measure with density

$$p(x, y, z) = \begin{cases} 2, \text{ when } x < \frac{1}{2}, \ y < \frac{1}{2}, \ z < \frac{1}{2}, \text{ or} \\ \qquad\quad x > \frac{1}{2}, \ y > \frac{1}{2}, \ z < \frac{1}{2}, \text{ or} \\ \qquad\quad x > \frac{1}{2}, \ y < \frac{1}{2}, \ z > \frac{1}{2}, \text{ or} \\ \qquad\quad x < \frac{1}{2}, \ y > \frac{1}{2}, \ z > \frac{1}{2}; \\ 0 \text{ at the remaining points} \end{cases}$$

(the measure "similar" to the above matrix (a_{ijk}), in that the matrix (a_{ijk}) can be regarded as a measure on the space $M/(\hat{\xi}_1 \vee \hat{\xi}_2 \vee \hat{\xi}_3)$, where $\hat{\xi}_1$, $\hat{\xi}_2$ and $\hat{\xi}_3$ are certain decompositions into two subsets each), then there exists a decomposition of (M, m) that is an independent complement to each of the three coordinate decompositions.

The consideration of other examples of cubic triply stochastic matrices for which the Birkhoff-von Neumann theorem is false does not lead in a similar way to the construction of a doubly stochastic density on the cube that does not admit an independent complement to the three coordinate decompositions. It would be interesting to clear up the situation, if only in the matrix case: is it possible to prove that for any multiply stochastic "matrix" we can construct a "refinement" (in the sense described above) for which the analogue of the Birkhoff-von Neumann theorem holds?

§11. Probability measures on subsets of direct products

0. In this section we isolate a quite broad class $\mathfrak{F}$ of subsets of the product of two Lebesgue spaces for which there is a simple criterion for the existence on the particular subset of a doubly stochastic probability measure, i.e., a probability measure having given marginal distributions. Our approach enables us to get, in particular, a criterion for the case when among such doubly stochastic measures there are measures that are absolutely continuous with respect to the products of their marginal distributions ("doubly stochastic densities"). The criterion for the existence of doubly stochastic densities is used in an essential way in the subsequent applications. The methods and auxiliary propositions of the preceding section are used for the derivation of this criterion.

1. Let (Ω, Σ) be a set with a distinguished σ-algebra of subsets, and $p\colon \Omega \longrightarrow X$ and $q\colon \Omega \longrightarrow Y$ measurable mappings of (Ω, Σ) into the probability spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$.

PROBLEM A. When is there a probability measure μ on (Ω, Σ) whose projections under the mappings p and q, i.e., the measures μp^{-1} and μq^{-1}, coincide with the measures μ_X and μ_Y?

DEFINITION 1. Let $(X, \mathfrak{A})$ and $(Y, \mathfrak{B})$ be two sets with distinguished σ-algebras of subsets. Then $\mathfrak{A} \otimes \mathfrak{B}$ denotes the smallest σ-algebra of subsets of $X \times Y$ with respect to which the canonical projections $\pi_X\colon X \times Y \longrightarrow X$ and $\pi_Y\colon X \times Y \longrightarrow Y$ are measurable.

DEFINITION 2. Let $(X, \mathfrak{A}, \mu)$ be a measure space. Then $\mathfrak{A}/\mu$ denotes the Boolean algebra of classes of μ-equivalent sets.

DEFINITION 3. $\mathfrak{N}$ denotes the σ-ring of all subsets $N \subset X \times Y$ that can be represented in the form $N = N_X \cup N_Y$, where $\mu_X(\pi_X N_X) = \mu_Y(\pi_Y N_Y) = 0$.

DEFINITION 4. $\mathfrak{A} \mathbin{\bar{\bar{\otimes}}} \mathfrak{B}$ denotes the smallest σ-algebra of subsets of the product $X \times Y$ containing the σ-ring $\mathfrak{N}$ and with respect to which the canonical projections π_X and π_Y are measurable.

We consider the mapping $p \times q\colon \Omega \longrightarrow X \times Y$, $(p \times q)(\omega) = (p(\omega), q(\omega))$, and let $Q = (p \times q)\Omega \subset X \times Y$.

If there is a measure μ on Ω with the required properties, then the measure $\mu(p \times q)^{-1}$, which is concentrated on the subset $Q \subset X \times Y$, equipped with the trace of the σ-algebra $\mathfrak{A} \mathbin{\bar{\otimes}} \mathfrak{B}$, has similar properties with respect to the canonical projections $\pi_X\colon X \times Y \longrightarrow X$ and $\pi_Y\colon X \times Y \longrightarrow Y$.

Conversely, if there exists a measure $\tilde{\mu}$ on Q for which $\mu_X = \tilde{\mu}\pi_X^{-1}$ and $\mu_Y = \tilde{\mu}\pi_Y^{-1}$, then the measure $\tilde{\mu}(p \times q)$ is defined on the σ-subalgebra $(p \times q)^{-1}(\mathfrak{A} \mathbin{\bar{\otimes}} \mathfrak{B})$ of the σ-algebra Σ, and the problem reduces to the extension of this measure from such a σ-subalgebra to the whole of Σ. It is not a great restriction to require beforehand the measurability of the set $Q \subset X \times Y$. To within the solution of this separate extension problem, the original Problem A reduces, then, to the case when Ω is a measurable subset of the direct product $X \times Y$. However, the concept of measurability requires some comments here.

For any measure space $(X, \mathfrak{A}, \mu)$ the class of measures defined on $\mathfrak{A}$ and absolutely continuous with respect to μ is completely determined by the Boolean algebra $\mathfrak{A}/\mu$, or, what is the same (see Chapter I, §1), by the ring S of classes of measurable functions on $(X, \mathfrak{A}, \mu)$ that are μ-almost everywhere finite. Therefore, in all arguments, when we speak of the existence of some measure on X, the μ-equivalent sets are identified; in fact, the problem is solved in terms of the Boolean algebra $\mathfrak{A}/\mu$. In the case of the product of two spaces it would be desirable to use only the Boolean algebras $\mathfrak{A}/\mu_X$ and $\mathfrak{B}/\mu_Y$, and not the spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ themselves.

We recall that an isomorphism of the Boolean algebras of equivalence classes of sets does not imply, generally speaking, an isomorphism of the measure spaces, even in the case of countably generated σ-algebras, if there is no additional requirement that the measure spaces be complete. However, in the case of the product of two spaces $(X, \mathfrak{A})$

and $(Y, \mathfrak{B})$ we see that the class of measures defined on subsets of the product (on the σ-algebra $\mathfrak{A} \bar{\otimes} \mathfrak{B}$) is not determined by the pair of corresponding Boolean algebras $\mathfrak{A}/\mu_X$ and $\mathfrak{B}/\mu_Y$ of equivalence classes of subsets of X and Y, respectively. On the face of it, this circumstance affirms the sometimes stated opinion that it is in some sense correct to consider on the space $(X \times Y, \ \mathfrak{A} \bar{\otimes} \mathfrak{B})$ only measures that are absolutely continuous with respect to the product $\mu_X \times \mu_Y$. But in reality the Boolean algebras $\mathfrak{A}/\mu_X$ and $\mathfrak{B}/\mu_Y$ determine a maximal Boolean algebra $\mathfrak{Q}$ that can be realized as the quotient algebra of $\mathfrak{A} \bar{\otimes} \mathfrak{B}$ by the ideal $\mathfrak{N}$ of subsets of $X \times Y$, where $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ are complete measure spaces. Each countably additive nonnegative function on $\mathfrak{Q}$ extends to a measure on $\mathfrak{A} \bar{\otimes} \mathfrak{B}$ whose projections onto X and Y are absolutely continuous with respect to μ_X and μ_Y, and conversely.

The Boolean σ-algebra $\mathfrak{Q}$, for which it is convenient to introduce the notation

$$\mathfrak{Q} = \mathfrak{A}/\mu_X \bar{\otimes} \mathfrak{B}/\mu_Y,$$

is now defined as the algebra $(\mathfrak{A} \bar{\otimes} \mathfrak{B})/(\mathfrak{N} \cap (\mathfrak{A} \bar{\otimes} \mathfrak{B}))$, or, what is the same,

$$\mathfrak{Q} = \mathfrak{A} \bar{\otimes} \mathfrak{B}/\mathfrak{N}.$$

It is clear that, since the complete measure spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ are uniquely (to within an isomorphism of measure spaces) determined by their Boolean algebras $\mathfrak{A}/\mu_X$ and $\mathfrak{B}/\mu_Y$, the notation $\mathfrak{Q} = \mathfrak{A}/\mu_X \bar{\otimes} \mathfrak{B}/\mu_Y$ is correct. If the completeness of $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ is not assumed, then, repeating the construction just described, we arrive at the Boolean algebra of $\mathfrak{N}$-equivalence classes of $\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$-measurable subsets, and the inclusions $X \subset \hat{X}$ and $Y \subset \hat{Y}$, which imbed the spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ canonically in the complete measure spaces $(\hat{X}, \hat{\mathfrak{A}}, \mu_{\hat{X}})$ and $(\hat{Y}, \hat{\mathfrak{B}}, \mu_{\hat{Y}})$, generate an epimorphism j: $\hat{\mathfrak{A}} \bar{\bar{\otimes}} \hat{\mathfrak{B}} \longrightarrow \hat{\mathfrak{A}} \bar{\bar{\otimes}} \mathfrak{B}$ (to each subset $C \subset \hat{X} \times \hat{Y}$, $C \in \hat{\mathfrak{A}} \bar{\bar{\otimes}} \hat{\mathfrak{B}}$, we assign its trace $C \cap (X \times Y) \in \mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$, which commutes with the corresponding relations of equivalence modulo $\mathfrak{N}$). Therefore, each such Boolean σ-algebra of classes of $\mathfrak{N}$-equivalent $\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$-measurable subsets of $X \times Y$ is the image of a Boolean algebra $\mathfrak{Q}$. It is proved below that the epimorphism j does not necessarily carry a quasi-measure on $\hat{\mathfrak{A}}/\mu_X \bar{\otimes} \hat{\mathfrak{B}}/\mu_Y$ into a quasi-measure (a quasi-measure on a Boolean algebra is defined to be a nonnegative (not necessarily strictly positive) countably additive normalized function on it [28]).

Finally, we show that the Boolean σ-algebra $(\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B})/\mathfrak{N}$ constructed for incomplete spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ can really be essentially poorer than the Boolean σ-algebra $\mathfrak{Q} = \mathfrak{A}/\mu_X \bar{\otimes} \mathfrak{B}/\mu_Y$. For this, we consider two subspaces X and Y of a Lebesgue space (M, Σ, m) such that $X \cup Y = M$, $X \cap Y = \varnothing$, and $m^*X = m^*Y = 1$ (and, consequently, $m_*X = m_*Y = 0$), where m^* is the outer measure. Obviously, the measure spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ (where $\mathfrak{A}$ and $\mathfrak{B}$ are the σ-algebras of subsets of X and Y, respectively, of the form $C \cap X$ and $C \cap Y$, $C \in \Sigma$, and $\mu_X(C \cap X) = mC$ and $\mu_Y(C \cap Y) = mC$) are such that the Boolean σ-algebras $\mathfrak{A}/\mu_X$ and $\mathfrak{B}/\mu_Y$ are isomorphic to the Boolean σ-algebra Σ/m. On $(M \times M, \Sigma \bar{\bar{\otimes}} \Sigma)$ we construct a measure μ such that $\mu C = 1$ for some $C \in \Sigma \bar{\bar{\otimes}} \Sigma$ and, moreover, $\mu\pi_M^{-1} = m$ for both projections π_M: $M \times M \longrightarrow M$, and $C \cap (X \times Y) = \varnothing$ ($X \times Y$ is regarded as being canonically imbedded in $M \times M$). The set C with the required properties can be taken to be, for example, the diagonal D of $M \times M$, and the measure μ can be taken to be the image of m under the mapping $f\colon\ M \to D$, $M \ni x \to (x, x) \in D$. It follows at once from the construction of X and Y that the diagonal D does not intersect the set $X \times Y \subset M \times M$; hence it is meaningless to consider the measure μ on the Boolean algebra $(\mathfrak{A} \mathbin{\bar{\bar{\otimes}}} \mathfrak{B})/\mathfrak{N}$.

Thus, in the following (and this turns out to be very essential) we consider only Lebesgue spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$. In this connection the constructions can be carried out in terms of the Boolean algebras $\mathfrak{A}/\mu_X$ and $\mathfrak{B}/\mu_Y$. The problem thus consists of determining conditions for the existence, on a given element of the Boolean σ-algebra $\mathfrak{A}/\mu_X \mathbin{\bar{\otimes}} \mathfrak{B}/\mu_Y$, of a quasi-measure whose restrictions to the σ-subalgebras that are canonically isomorphic to $\mathfrak{A}$ and $\mathfrak{B}$ coincide with μ_X and μ_Y.

The Boolean σ-algebra $\mathfrak{Q} = \mathfrak{A}/\mu_X \mathbin{\bar{\otimes}} \mathfrak{B}/\mu_Y$ described above is far from having all the useful properties enjoyed by the Boolean σ-algebra Σ/m of m-equivalence classes of subsets of (m, Σ, M) with countably generated σ-algebra Σ.

PROPOSITION 64. *For measure spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ with measures containing continuous components the Boolean σ-algebra $\mathfrak{Q} = \mathfrak{A}/\mu_X \mathbin{\bar{\otimes}} \mathfrak{B}/\mu_Y$ is not a Boolean algebra of countable type and is not complete.*

PROOF. A Boolean algebra is said to be of countable type (see [143]) if any subset of pairwise disjoint elements of it is not more than countable. It suffices to consider measure spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ with countably generated σ-algebras $\mathfrak{A}$ and $\mathfrak{B}$ and purely continuous measures μ_X and μ_Y. The Boolean algebras of equivalence classes of measurable subsets of all such spaces are isomorphic; therefore, it suffices to consider any concrete realization. Let $(X, \mathfrak{A}, \mu_X) = (Y, \mathfrak{B}, \mu_Y)$ be the group $\mathbf{R}/\mathbf{Z}$ of rotations of the circle of unit length with Haar measure. Our uncountable subset $\mathfrak{U}$ of pairwise disjoint nonzero elements of the Boolean algebra $\mathfrak{Q}$ can now be defined as $\mathfrak{U} = \mathfrak{U}_A = \{\dot{U}_\alpha, \alpha \in A\}$, where $U_\alpha = \{(x, y)\colon\ x, y \in \mathbf{R}/\mathbf{Z}, x - y = \alpha\}$, and $A \in \mathbf{R}/\mathbf{Z}$ is an arbitrary uncountable set, while $\dot{U}$ denotes the element of $\mathfrak{Q}$ generated by the subsets $U \subset X \times Y$. The measurability of the U and their pairwise disjointness are obvious, and it is proved that $\mathfrak{Q}$ is of uncountable type.

Now, to prove the incompleteness of $\mathfrak{Q}$, i.e., to show that not every subset of it has a supremum, it suffices to consider the same type of subset $\mathfrak{U}_A$, where A is a (Lebesgue) nonmeasurable subset of $\mathbf{R}/\mathbf{Z}$. Indeed, suppose that $\mathrm{B} = \sup \mathfrak{U}_A$ exists. Let

$$B = \{\beta : \{(x,\ y) : x - y = \beta\}^{\boldsymbol{\cdot}} \wedge \mathrm{B} \neq 0\}.$$

Obviously, $B \supset A$.

Each element $\dot{U}_A$ of the Boolean algebra $\mathfrak{Q}$ is invariant with respect to the automorphisms of $\mathfrak{Q}$ generated by the transformations of the set $X \times Y$ of the form $(x, y) \mapsto (x + \gamma, y + \gamma)$; consequently, by uniqueness, the supremum B of the set of such elements must also be invariant with respect to all such transformations. From this it follows that if $\beta \in B$, then not only is it true that $\{(x, y)\colon\ x - y = \beta\}^{\boldsymbol{\cdot}} \wedge \mathrm{B} \neq 0$, but also $\{(x, y)\colon\ x - y = \beta\}^{\boldsymbol{\cdot}} < \mathrm{B}$. If there were a $\beta_0 \in B \backslash A$, then it would happen that

$$B \wedge C\{(x, y): x - y = \beta_0\}^{\cdot} < B,$$

but, as before,

$$B \wedge C\{(x, y): x - y = \beta_0\}^{\cdot} > \dot{U}_\alpha \quad \text{for} \quad \alpha \in A,$$

and this would contradict the equation $B = \sup_{\alpha \in A}\{\dot{U}_\alpha\}$.

Thus, $B = A$, i.e.

$$B = \{(x, y): x - y = \alpha, \ \alpha \in A\}.$$

Now we show that the set $\{(x, y)\colon\ x - y = \alpha, \alpha \in A\}$ is not in the σ-algebra $\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$ when A is a subset of the circle that is not Lebesgue measurable. Indeed, the product measure $\mu_X \times \mu_Y$ is defined on all subsets in $\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$ and is obviously invariant with respect to all the transformations $(x, y) \longmapsto (x + \gamma, y - \gamma)$, $\gamma \in \mathbf{R}/\mathbf{Z}$, of $X \times Y$. The mapping $t\colon\ X \times Y \longmapsto \mathbf{R}/\mathbf{Z}$, $t(x, y) = x + y$, carries the measure $\mu_X \times \mu_Y$ into an invariant measure on $\mathbf{R}/\mathbf{Z}$, and the collection of measurable subsets of $\mathbf{R}/\mathbf{Z}$ coincides with the collection of $\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$-measurable subsets of $X \times Y$ of the form $\bigcup_{\alpha \in A} U_\alpha$. Since [81] on $\mathbf{R}/\mathbf{Z}$ there does not exist an invariant measure defined on the σ-algebra of all subsets, for some $A \subset \mathbf{R}/\mathbf{Z}$ (which is obviously not Lebesgue measurable) the set $\{(x, y)\colon\ x - y = \alpha, \alpha \in A\}$ is not in $\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$. ●

In the following, admitting abuse of language, we frequently do not distinguish among a subset $M \subset X \times Y$, the class of subsets that are $\mathfrak{N}$-equivalent to this subset, and the corresponding element of the Boolean σ-algebra $\mathfrak{A}/\mu_X \bar{\otimes} \mathfrak{B}/\mu_Y$.

2. We proceed now to the basic problem.

Problem B. The Lebesgue spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ are given, and an $\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$-measurable subset K of the product $X \times Y$ is selected. We consider the question of the existence of a probability measure μ on $\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$ having the following properties:

$$1)\ \mu K = 1; \quad 2)\ \mu\pi_X^{-1} = \mu_X; \quad 3)\ \mu\pi_Y^{-1} = \mu_Y.$$

The solution of Problem B given below reduces to the determination of a class $\mathfrak{F}$ of subsets of $\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$ (or a class of elements of the Boolean algebra $\mathfrak{Q} = \mathfrak{A}/\mu_X \bar{\otimes} \mathfrak{B}/\mu_Y$) for which a simple criterion is stated and proved for the positive solution of the problem; it also turns out not to be complicated to test whether or not a set K is in $\mathfrak{F}$.

In its properties the class $\mathfrak{F}$ turns out to be similar to the class of compact subsets of a product of compact sets, but it is defined in terms of pure measure theory. Before introducing it, we state some definitions.

Let ξ_k and η_k be two refining sequences of measurable decompositions of the respective nonatomic spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$, each of which converges to the decomposition into points ($\xi_k \nearrow \epsilon_X$ and $\eta_k \nearrow \epsilon_Y$), and such that ξ_k and η_k are each decompositions of the spaces into 2^k subsets of equal measure.

Definition 5. P is the class of subsets of $X \times Y$ of the form $A \times B$, $A \in \mathfrak{A}$, $B \in \mathfrak{B}$; P_k is the class of subsets of $X \times Y$ of the form $A \times B$, where $A \subset X$ is

ξ_k-measurable and $B \subset Y$ is η_k-measurable; αP is the smallest algebra of sets containing P; αP_k is the smallest algebra of sets containing P_k.

Thus, $C \in \alpha P$ if and only if C admits a representation $C = \bigcup_1^n A_k \times B_k$, where $A_k \in \mathfrak{A}$, $B_k \in \mathfrak{B}$, $n < \infty$, and the sets $A_k \times B_k$ are pairwise disjoint.

DEFINITION 6. Let $C \in \alpha P$. Then

$$\Pi(C) = \max_{\substack{A \in \mathfrak{A},\ B \in \mathfrak{B} \\ (A \times B) \cap C = \varnothing}} (\mu_X A + \mu_Y B)$$

(here it is not excluded that $\mu_X A = 0$ or $\mu_Y B = 0$, so that $\Pi(C) \geqslant 1$; this maximum is attained, by the condition $C \in \alpha P$).

DEFINITION 7. Let $C, D \in \alpha P$. Then $\rho(C, D) = 2 - \Pi(C \Delta D)$.

PROPOSITION 65. *$\rho(C, D)$ is a metric on αP.*

PROOF. It suffices to verify the triangle inequality. We set $\rho(U) = \rho(U, \varnothing)$, so that

$$\rho(C, D) = \rho(C \Delta D, \varnothing) = \rho(C \Delta D).$$

Further, $\rho(C, E) \leqslant \rho(C, D) + \rho(D, E)$ is equivalent to

$$\rho(C \Delta E) \leqslant \rho(C \Delta D) + \rho(D \Delta E).$$

Since, as is easily checked, $C \Delta E \subset (C \Delta D) \cup (D \Delta E)$, and since the function $\rho(U)$ is monotonically nondecreasing with respect to inclusion, i.e., $W \supset V$ implies $\rho(W) \geqslant \rho(V)$, it suffices to verify that $\rho(V \cup W) \leqslant \rho(V) + \rho(W)$ for $V, W \in \alpha P$. Indeed, if

$$(A_V \times B_V) \cap V = \varnothing, \quad (A_W \times B_W) \cap W \neq \varnothing,$$

where $A_V, A_W \in \mathfrak{A}$ and $B_V, B_W \in \mathfrak{B}$, then

$$((A_V \cap A_W) \times (B_V \cap B_W)) \cap (V \cup W) = \varnothing,$$

and

$$\mu_X(A_V \cap A_W) \geqslant \mu_X A_V + \mu_X A_W - 1,$$
$$\mu_Y(B_V \cap B_W) \geqslant \mu_Y B_V + \mu_Y B_W - 1,$$

from which we obtain

$$\Pi(V \cup W) \geqslant \mu_X(A_V \cap A_W) + \mu_Y(B_V \cap B_W) \geqslant \Pi(V) + \Pi(W) - 2$$

or

$$2 - \Pi(V \cup W) \leqslant (2 - \Pi(V)) + (2 - \Pi(W)),$$

i.e., $\rho(V \cup W) \leqslant \rho(V) + \rho(W)$. ●

It is useful to observe that the quantity $\rho(V)$ is analogous to the term *rank* of a (0, 1)-matrix [101].

We remark that if in the definition of the functional Π we take the maximum only over the subsets $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$ of positive measure, then ρ ceases to be a metric, as is clear from the example of the sets C, D, E (see Figure 7, where $X = Y = [0, 1]$):

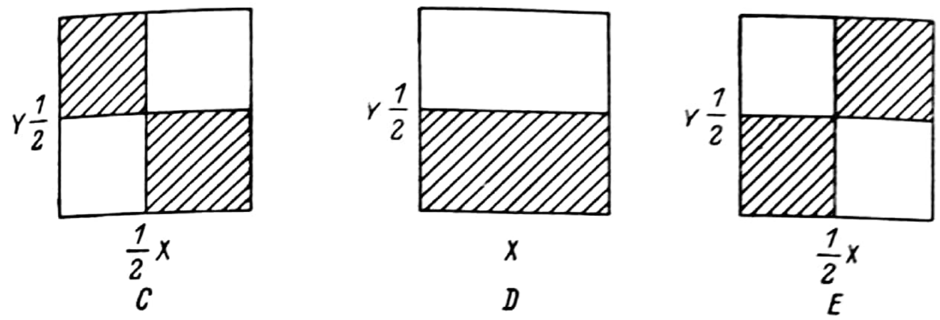

FIGURE 7

$$\rho_1(C, E) = 2, \; \rho_1(C, D) = \rho_1(D, E) = \frac{1}{2},$$

where

$$\rho_1(F, G) = 2 - \max\{\mu_X A + \mu_Y B : \mu_X A > 0, \; \mu_Y B > 0, \; (A \times B) \cap (F \Delta G) = \varnothing\},$$
$$\rho_1(C, E) > \rho_1(C, D) + \rho_1(D, E).$$

PROPOSITION 66. *$(\alpha P, \rho)$ is a separate metric space.*

PROOF. Clearly, we can take the set $\bigcup_k \alpha P_k \subset \alpha P$ as a countable dense subset in αP. ●

Of course, the metric ρ can also be defined on the whole σ-algebra $\mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B}$; in the definition of the functional Π it would be natural to require only the condition $(A \times B) \cap C \in \mathfrak{N}$ instead of the disjointness of $A \times B$ and C. However, the metric thus defined on the Boolean algebra $\mathfrak{A}/\mu_X \bar{\otimes} \mathfrak{B}/\mu_Y$ turns it into a nonseparable metric space. An example of a continuum subset, any two elements of which are at a distance of not less than ½ from each other, is given by the subset of this Boolean algebra considered on p. 174 for the proof of the noncountability of its type. Nevertheless, the metric is very natural: the less the number $\rho(C)$, the more the set C "resembles" a set in the ideal $\mathfrak{N}$. Obviously, if $N \in \mathfrak{N}$, then $\rho(N) = 0$. We prove the converse assertion.

PROPOSITION 67. *Let $V \subset X \times Y$ and* $\sup_{(A \times B) \cap V \in \mathfrak{N}} (\mu_X A + \mu_Y B) = 2$. *Then $V \in \mathfrak{N}$.*

PROOF. Let $\{A_n\}$ and $\{B_n\}$ be sequences of subsets of $\mathfrak{A}$ and $\mathfrak{B}$ respectively (whose existence follows from the hypothesis) such that

$$(A_n \times B_n) \cap V \in \mathfrak{N},$$
$$\sum_{n=1}^{\infty} (1 - \mu_X A_n) < \infty, \; \sum_{n=1}^{\infty} (1 - \mu_Y B_n) < \infty.$$

It is easy to see that

$$V\subset(\overline{\lim}\, CA_n\times Y)\cup(X\times\overline{\lim}\, CB_n)$$

to within a subset in $\mathfrak{N}$. But

$$\mu_X\,\overline{\lim}\, CA_n=\mu_Y\,\overline{\lim}\, CB_n=0,$$

since

$$\sum_{n=1}^{\infty}\mu_X CA_n<\infty,\quad \sum_{n=1}^{\infty}\mu_Y CB_n<\infty,$$

from which we get the required conclusion. ●

Proposition 68. *Each family $S\subset \mathcal{P}$ has a supremum in the sense of the structure of the Boolean algebra $\mathfrak{A}/\mu_X \overline{\otimes}\, \mathfrak{B}/\mu_Y$ (i.e.,* mod $\mathfrak{N}$ *).*

Proof. Let $\{P_n, n = 1, \ldots\}$ be a countable subset that is dense in S in the sense of the metric ρ. We consider the set $Q = \bigcup_1^\infty P_n$ and prove that $Q = \sup S$. Let $P\in S$; we show that $P\subset Q$ (mod $\mathfrak{N}$). Let $P_{n_k}\longrightarrow P$, i.e., $\rho(P, P_{n_k})\longrightarrow 0$. Then

$$\rho\Big(P\setminus\bigcup_{k=1}^{m}P_{n_k},\varnothing\Big)\to 0 \quad\text{and}\quad P\setminus\bigcup_{k=1}^{\infty}P_{n_k}\underset{m}{\to}\varnothing,$$

i.e., by Proposition 61, we get that $P\setminus\bigcup_{k=1}^{\infty}P_{n_k}\in\mathfrak{N}$.

Conversely, it is obvious that no subset of $X\times Y$ not containing the set P mod $\mathfrak{N}$ can be an upper bound for $\{P_n\}$, and a fortiori for all S. ●

We now introduce the class of subsets of $X\times Y$ that is basic for us.

Definition 8. $\mathfrak{F}$ denotes the class of subsets of $X\times Y$ whose complements are suprema mod $\mathfrak{N}$ (in the sense of the Boolean algebra $\mathfrak{A}/\mu_X \overline{\otimes}\, \mathfrak{B}/\mu_Y$) of families of subsets in $\mathcal{P}$.

The class $\mathfrak{F}$ is quite broad. In one particular case when the spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ are complete separable metric spaces with Borel measures μ_X and μ_Y, the class $\mathfrak{F}$ contains all the closed subsets of the topological product $X\times Y$: since a basis of open subsets in $X\times Y$ can be taken to be the class of products of open subsets of X and Y, the closed subsets are the complements of countable unions of open subsets in this basis. Informally, we can think of the family $\mathcal{P}$ as analogous to a base of open-closed subsets of a compact space, the suprema of subsets of $\mathcal{P}$ as analogous to the open sets, and the class $\mathfrak{F}$ as analogous to the class of closed subsets of a compact space.

In spite of the uncountable type of the Boolean algebra $\mathfrak{A}/\mu_X \overline{\otimes}\, \mathfrak{B}/\mu_Y$, the test for whether or not a given set F belongs to $\mathfrak{F}$ bears an essentially countable character.

Proposition 69. $F\in\mathfrak{F}$ *if and only if for some* P_k, $k = 1, \ldots,$

$$F=(X\times Y)\setminus\bigcup_{k=1}^{\infty}P_k \ (\mathrm{mod}\,\mathfrak{N}),\ \textit{where}\ P_k\in \mathcal{P},\ P_k\cap F\in\mathfrak{N}.$$

The proof follows at once from Proposition 68.

3. We now consider the space of probability measures μ on $(X \times Y, \mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B})$ that are compatible with the structure of the Boolean algebra $\mathfrak{A}/\mu_X \otimes \mathfrak{B}/\mu_Y$.

DEFINITION 9. If $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ are Lebesgue spaces, then $M(X \times Y)$ denotes the space of probability measures μ on $(X \times Y, \mathfrak{A} \bar{\bar{\otimes}} \mathfrak{B})$ for which $\mu\pi_X^{-1} = \mu_X$ and $\mu\pi_Y^{-1} = \mu_Y$, equipped with the topology τ determined by specifying the following fundamental system of neighborhoods for an element $\mu \in M(X \times Y)$:

$$V(\mu;\ A_1,\ \dots,\ A_n;\ B_1,\ \dots,\ B_n;\ \varepsilon)$$
$$=\{\mu' : \mu' \in M,\ |\mu(A_k \times B_k) - \mu'(A_k \times B_k)| < \varepsilon,\ k = 1,\ \dots,\ n\},$$

where $A_k \in \mathfrak{A}$, $B_k \in \mathfrak{B}$, $k = 1, \dots, n$, and $\epsilon > 0$.

In other words, the topology τ is defined as the trace of the weak topology of the space dual to the tensor product $L^\infty(X) \otimes L^\infty(Y)$ under the natural imbedding in it of the set M. The space $M(F)$ for an arbitrary set $F \in \mathfrak{F}$ is defined similarly; our main goal is to determine conditions under which $M(F)$ is not empty.

PROPOSITION 70. *The space* $(M(X \times Y), \tau)$ *is a compact metrizable space.*

The proof makes essential use of the completeness ("Lebesgueness", since it is assumed that the measure spaces are of countable type) of $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$. However, the assertion that $(M\tau)$ is metrizable is, of course, true even without the assumption of completeness.

We first prove that the topology τ can be given also by the neighborhoods of the form $V(\mu; A_1, \dots, A_n; B_1, \dots, B_n; \epsilon)$, where the sets $A_1, \dots, A_n$ are ξ_k-measurable and $B_1, \dots, B_n$ are η_k-measurable for some $k = k(V)$. Indeed, let $P = A \times B$, where $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$, and let $\mu \in M$ and $\epsilon > 0$. We construct a neighborhood $V(\mu; \bar{A}; \bar{B}; \bar{\epsilon})$ of μ, where $\bar{A}$ is ξ_k-measurable, $\bar{B}$ is η_k-measurable, and $\bar{\epsilon} > 0$, that is contained in the neighborhood $V(\mu; A; B; \epsilon)$ of μ. For this, we choose k and corresponding ξ_k-measurable and η_k-measurable sets $\bar{A} \in \mathfrak{A}$ and $\bar{B} \in \mathfrak{B}$ such that

$$\mu_X(A \Delta \bar{A}) < \frac{\varepsilon}{8},\quad \mu_Y(B \Delta \bar{B}) < \frac{\varepsilon}{8}.$$

It is not hard to see that for any measure $\mu' \in M$

$$\mu'((A \times B) \Delta (\bar{A} \times \bar{B})) < \frac{\varepsilon}{8} + \frac{\varepsilon}{8} = \frac{\varepsilon}{4}.$$

Therefore, if $\mu' \in V(\mu; \bar{A}; \bar{B}; \epsilon/2)$, i.e., if

$$|\mu'(\bar{A} \times \bar{B}) - \mu(\bar{A} \times \bar{B})| < \frac{\varepsilon}{2},$$

then, since we always have $|\mu(V) - \mu(W)| \leqslant \mu(V \Delta W)$, we obtain

$$|\mu'(A \times B) - \mu(A \times B)|$$
$$= |\mu'(A \times B) - \mu'(\bar{A} \times \bar{B}) + \mu'(\bar{A} \times \bar{B}) - \mu(\bar{A} \times \bar{B}) + \mu(\bar{A} \times \bar{B})$$
$$- \mu(A \times B)| \leqslant \mu'((A \times B) \Delta (\bar{A} \times \bar{B})) + \frac{\varepsilon}{2} + \mu((\bar{A} \times \bar{B}) \Delta (A \times B))$$
$$\leqslant \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon,$$

which proves the desired inclusion, and, together with it, the countability of a basis of open sets, i.e., the metrizability.

We now suppose that we are given an infinite subset of M and prove that we can extract from it a sequence that converges to some element in M. Indeed, it was just proved that convergence in M is convergence of the values of the measures on a countable number of sets. Since the values of the measures on each set do not exceed 1, by using a diagonal process we can get a sequence of measures μ_n, $n = 1, \ldots,$ whose values converge to a limit on each set in our countable system. It remains to show (here the completeness property is essential) that these limits are the values of a measure on $(X \times Y, \ \mathfrak{A} \mathbin{\overline{\overline{\otimes}}} \mathfrak{B})$.

For each k we choose a ξ_k- (respectively, η_k-) measurable set A_k (B_k) such that the system $\{A_k\}$ $(\{B_k\})$ forms a basis [92]. To each element of X (of Y) we assign a sequence of zeros and ones: its "coordinates" in this basis. Without loss of generality we can assume that a one-to-one correspondence is established in this way between the elements of X (of Y) and those of the compact set formed by the countable power of the doubleton $D = \{0, 1\}$, where the measure μ_X (μ_Y) corresponds here to the countable product of the measure with masses ½, and the σ-algebra of Lebesgue measurable (with respect to the aforementioned product measure on the compact set $D^{\aleph_0}$) subsets of the compact set. More briefly, fixing the bases compactifies (to within a subset of measure 0) the spaces X and Y. But this means that $X \times Y$ can be regarded as a compact set, and since the limit values of a sequence of measures μ_n clearly form an additive function on the set $\bigcup_k \alpha P_k$, this additive function uniquely determines a measure on $\mathfrak{A} \mathbin{\bar{\otimes}} \mathfrak{B}$ (and on $\mathfrak{A} \mathbin{\overline{\overline{\otimes}}} \mathfrak{B}$) that extends it. With this, the compactness of the space of measures $M(X \times Y)$ is proved. ●

PROPOSITION 71. *For any $F \in \mathfrak{F}$*

$$M(F) = \bigcap_{\substack{P:\, P \in \mathsf{P} \\ P \cap F \in \mathfrak{N}}} M((X \times Y) \setminus P). \tag{58}$$

PROOF. Equation (58) means that a measure vanishing on each of the uncountable number of sets P outside F vanishes in general outside F. We first mention that for any $P \in \mathsf{P}$ the set of measures $M((X \times Y)\setminus P)$ with the topology induced from $M(X \times Y)$ is a compact metric space (possibly empty). Indeed, $M((X \times Y)\setminus P)$, where $P = A \times B$, is a closed subset of M, since, by the definition of the topology in M, if $\mu_n(P) = 0$ and $\mu_n \longrightarrow \mu$, then $\mu(P) = 0$. Now suppose that $F = (X \times Y)\setminus \bigcup_1^\infty P_k$ (mod $\mathfrak{N}$), $P_k \in \mathsf{P}$ (Proposition 69). Each measure $\mu \in M(X \times Y)$ vanishing on each P_k, $k = 1, \ldots,$ has the property that $\mu F = 1$, i.e., $\mu \in M(F)$. By the same token, if a measure is in the right-hand side in (58), then it is automatically in the left-hand side. The reverse inclusion is trivial. ●

We can now prove the following theorem.

THEOREM 9. *Let $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$ be Lebesgue spaces, $F \subset X \times Y$, and $F \in \mathfrak{F}$. For there to be a measure μ on the space $(F, \mathfrak{A} \mathbin{\overline{\overline{\otimes}}} \mathfrak{B}\,|_F)$ for which $\mu\pi_X^{-1} = \mu_X$ and $\mu\pi_Y^{-1} = \mu_Y$, it is necessary and sufficient that*

$$\sup_{\substack{A \in \mathfrak{A},\ B \in \mathfrak{B} \\ (A \times B) \cap F \in \mathfrak{N}}} (\mu_X A + \mu_Y B) = 1. \tag{59}$$

PROOF. The necessity is obvious, by (58). Because of the compactness of the sets $M(X \times Y)\backslash P)$, and hence of $M(F)$ (for $F \in \mathfrak{F}$), it suffices to prove that finite intersections of the sets of the form $M((X \times Y)\backslash P)$, $P \in P$, $P \cap F \in \mathfrak{N}$, are not empty.

Here, however, we are in a situation where we can use Theorem 6. Let P_k, $k = 1, \ldots, n$, be subsets of $X \times Y$ having the form $P_k = A_k \times B_k$, and let $F \subset F' = (X \times Y)\backslash \bigcup_1^\infty P_k$. We show that there is a measure $\widetilde{\mu}\pi_X^{-1} = \mu_X$ and $\widetilde{\mu}\pi_Y^{-1} = \mu_Y$. Indeed, (59) means that $\Pi(F') = 1$; therefore each "coarsened 2×2 problem" is solvable if we take the marginal measures μ_X and μ_Y to be measures that are proportional to μ_X and μ_Y with a sufficiently small proportionality coefficient λ.

The "coarsened problem" singled out by fixing some two measurable sets $A \subset X$ and $B \subset Y$ is actually determined if the measures are given for the intersections of A and B with each of the (not more than) 2^n subsets of X and the 2^n subsets of Y into which these spaces are decomposed by all possible intersections of the sets A_k, B_k ($k = 1, \ldots, n$) and their complements. Thus, the aforementioned proportionality coefficient λ is a nonzero continuous function defined on some compact subset of $\mathbf{R}^{2n+1}$. By Theorem 6, it follows from this that there exists a subprobability measure on F' that is absolutely continuous with respect to the product measure $\mu_X \times \mu_Y$ and for which the marginal distributions are the respective measures $\lambda_{\min}\mu_X$ and $\lambda_{\min}\mu_Y$; therefore the required probability measure exists and has density not exceeding $\lambda_{\min}^{-1}$. The nonemptiness of the finite intersections of the compact sets $M((X \times Y)\backslash P)$, and with it the nonemptiness of the compact set $M(F)$, under the assumption of the condition (59) is proved. ●

We note the essentiality of the condition $F \in \mathfrak{F}$ for the criterion given by Theorem 9 to be correct. As the simplest counterexample we take $X = Y = [0, 1]$, μ_X and μ_Y each Lebesgue measure, and F the set $\{(x, y)\colon\ x > y\}$. Obviously, (59) holds, but on this set it is impossible to define a measure whose marginal distributions are the Lebesgue measures.

Theorem 9 can be regarded as an analogue of Theorem 6. Indeed, the condition $\mu_X A + \mu_Y B \leqslant 1$ for $(A \times B) \cap F \in \mathfrak{N}$, $(A \times CB) \cap F \notin \mathfrak{N}$, $(CA \times B) \cap F \in \mathfrak{N}$ and $(CA \times CB) \cap F \notin \mathfrak{N}$ can be considered to be a criterion for the solvability of a "coarsened 2×2 problem", understood as a problem on the existence of a nonnegative 2×2 matrix with given sums of elements in the rows and columns and one restriction: it is required that the element corresponding to the "cell" $A \times B$ is equal to 0.

We note in particular that it is not necessary to verify the condition in Theorem 9 for all possible pairs of measurable sets (A, B) for which $(A \times B) \cap F \in \mathfrak{N}$. We can limit ourselves to a family (which can be chosen to be countable) of pairs A_γ, B_γ, $\gamma \in \Gamma$, for which $(X \times Y)\backslash \bigcup_{\gamma \in \Gamma}(A_\gamma \times B_\gamma) \in \mathfrak{N}$. In particular, if F is a closed subset of the direct product of complete metric spaces X and Y with Borel measures μ_X

and μ_Y, then we can limit ourselves to the open subsets $A \subset X$ and $B \subset Y$.(3) As already mentioned, each closed subset $F \subset X \times Y$ belongs to the class $\mathfrak{F}$, but not conversely: the class $\mathfrak{F}$ is invariant with respect to the transformations $(x, y) \mapsto (Tx, Uy)$, where T and U are automorphisms of the measure spaces $(X, \mathfrak{A}, \mu_X)$ and $(Y, \mathfrak{B}, \mu_Y)$, which cannot be said of the class of closed subsets of $X \times Y$.

Theorem 9 does not carry over to the case of subsets of products of three or more measure spaces. An example given by Strassen [113], which shows that the criterion in the Strassen-Kellerer form (for closed subsets of complete metric spaces) does not carry over to the case of a product with more than two factors, serves also as an example of a subset of a product of finite spaces (2, 2, and 3 points) with given marginal measures and having the property that any "coarsened $2 \times 2 \times 2$ problem" is solvable, while there is no measure on this subset with the given marginal distributions.

4. We give another formulation of the theorem above, one whose assertion admits a useful extension.

DEFINITION 10. Let $F \in \mathfrak{F}$. Then we define the set $\mathrm{mi}\, F = \mathrm{mi}_{\mu,\nu} F$ by

$$\mathrm{mi}_{\mu,\nu} F = F \setminus \sup \{A \times B : \mu A + \nu B = 1,\ ((X \setminus A) \times (Y \setminus B)) \cap F \in \mathfrak{N}\},$$

and call it the *minimization* of the set F (with respect to μ and ν). Its existence is guaranteed by Proposition 68.

As mentioned, if m is a (μ, ν)-doubly stochastic measure, and the sets A and B are such that $\mu A + \nu B = 1$ and $m(A \times B) = 0$, then also $m(CA \times CB) = 0$. Therefore, if m is a (μ, ν)-doubly stochastic measure on the set $F \subset X \times Y$, then, taking Proposition 69 into account, $m(\mathrm{mi}_{\mu,\nu} F) = 1$. Thus, *the nonemptiness of the minimization of a set $F \in \mathfrak{F}$ is a necessary condition for the existence on this set of a doubly stochastic measure.*

We show that *the nonemptiness of the minimization of a set $F \in \mathfrak{F}$ is also a sufficient condition for the existence on this set of a doubly stochastic measure.*

Suppose that there is no doubly stochastic measure on the set $F \in \mathfrak{F}$. Then, by Theorem 9, there are sets $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$ such that

$$\mu A + \nu B > 1,\ (A \times B) \cap F \in \mathfrak{N}.$$

We show that we then have $\mathrm{mi}\, F = \varnothing$. Indeed, let $A' \in \mathfrak{A}$ and $B' \in \mathfrak{B}$ be arbitrary subsets for which $\mu A', \nu B' \leqslant \mu A + \nu B - 1$. We prove that

$$(A' \times B') \cap \mathrm{mi}\, F \in \mathfrak{N}. \tag{60}$$

For, if $A' \times B' \subset CA \times CB$, then the assertion is obvious, since, by the definition of minimization, $(CA \times CB) \cap \mathrm{mi}\, F \in \mathfrak{N}$. Now let $A' \times B' \subset A \times CB$. Then

$$\mu((A \setminus A') + \nu B) = \mu A + \nu B - \mu A' \geqslant 1$$

---

(3) In this form (with a different form of notation, apparently inspired by Kellerer [55], for the condition (59)), i.e., for the closed subsets of a product of complete metric spaces, the theorem on the existence of a measure with given marginal measures was first obtained by Strassen [113]. The criterion studied below for the existence of a doubly stochastic density does not follow directly from the results of Strassen.

and, since $A' \times B' \subset C(A \backslash A') \times CB$, we get, as before, that the sets $A' \times B'$ and mi F are disjoint. Now if the set $A' \times B'$ is in "general position" with respect to $A \times B$, then the verification of (60) reduces to its verification for four of its subsets having the properties just considered. ●

THEOREM 9 (second formulation). *For there to be a (μ, ν)-doubly stochastic measure on the subset $F \subset X \times Y$, $F \in \mathfrak{F}$, it is necessary and sufficient that* $\mathrm{mi}_{\mu,\nu} F \neq \varnothing$.

The operation $\mathrm{mi}_{\mu,\nu}$ is idempotent. In fact, let $\{(A_n, B_n), n = 1, \ldots\}$ be a sequence of pairs of subsets $A_n \in \mathfrak{A}$, $B_n \in \mathfrak{B}$ such that

$$\mu A_n + \nu B_n = 1, \ (A_n \times B_n) \cap F \in \mathfrak{N},$$
$$\sup\{A_n \times B_n, \ n = 1, \ldots\} = \sup\{A \times B : A \times B \cap F \in \mathfrak{N}\} \tag{61}$$

(see Proposition 69). We consider the measurable decompositions ξ_F and η_F of the spaces $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ generated by the systems of sets $\{A_n\}$ and $\{B_n\}$. Since $\mu A_n = \nu C B_n$, the relation $A_n \leftrightarrow CB_n$ establishes an isomorphism between the measure spaces $(X, \mu)/\xi_F$ and $(Y, \nu)/\eta_F$, and the set mi F lies on the "diagonal": on the union of the products of those elements of ξ_F and η_F that correspond by virtue of the isomorphism. If instead of μ and ν we now consider a pair μ', ν' of conditional measures on corresponding elements of ξ_F and η_F, then for these measures it is already true that the set $\mathrm{mi}_{\mu',\nu'}(\mathrm{mi}_{\mu,\nu} F)$ coincides (mod $\mathfrak{N}(\mu, \nu)$) with the set $\mathrm{mi}_{\mu',\nu'} F$, since, for each nontrivial pair of sets A', B' for which

$$\mu' A' + \nu' B' = 1, \ (A' \times B') \cap \mathrm{mi}_{\mu,\nu} F \in \mathfrak{N}(\mu', \nu'),$$

it is possible to find a pair of sets A, B such that

$$\mu A + \nu B = 1, \ (A \times B) \cap \mathrm{mi}_{\mu,\nu} F \in \mathfrak{N}(\mu, \nu),$$

and this contradicts (61). Since we do not use the idempotence of the operation mi in the following, we limit ourselves to this explanation.

It can be shown that it is impossible to remove a nonempty (mod $\mathfrak{N}$) subset of the form $(A \times B) \cap F$ from the set $\mathrm{mi}_{\mu,\nu} F$ without diminishing the collection of doubly stochastic measures on F.

5. We consider an extension of this result that is important for the sequel. Suppose that we want to establish not simply the existence of some doubly stochastic measure on a particular subset F of $X \times Y$, but the existence on F of a doubly stochastic measure that is absolutely continuous with respect to the product measure $\mu \times \nu$. It is trivial that the class of all such doubly stochastic probability densities on a subset $F \subset X \times Y$ is not compact in the topology τ (if it is not empty), so a direct application of the previous arguments does not lead to the goal. The subset $Q = \{(x, y): x \geqslant y\}$ of the unit square I^2, considered as the product of two segments I on which Lebesgue measure plays the role of the measures μ and ν, provides an example of a set in the class $\mathfrak{F}$ whose conditional measures under the coordinate decompositions of $(I^2, \mu \times \nu)$ are almost all positive, for which the conditions of Theorem 9 hold (i.e.,

there exist doubly stochastic measures on it), but for which there does not exist a doubly stochastic measure determined by a density with respect to the measure $(\mu \times \nu)_Q$: the probability measure having density $2\chi_Q(x, y)$ with respect to $\mu \times \nu$.

Let F be an arbitrary $\mathfrak{A} \bar{\otimes} \mathfrak{B}$-measurable set of positive $(\mu \times \nu)$-measure. We now do not assume that $F \in \mathfrak{F}$. Since we are interested only in measures that are absolutely continuous with respect to $\mu \times \nu$, the set F should be considered not to within subsets of the class $\mathfrak{N}$, but to within subsets of zero $(\mu \times \nu)$-measure.

In the above example of a subset Q of the unit square on which there is no doubly stochastic measure that is absolutely continuous with respect to two-dimensional Lebesgue measure, the unique doubly stochastic measure on this set is the one concentrated on the diagonal $\{(x, y): x = y\}$ of the square. In other words, the "greater" part of Q turns out to be useless in the problem of constructing a doubly stochastic density: for any doubly stochastic measure m on Q we have

$$m(Q \setminus \{(x, y): x = y\}) = 0.$$

Of course, in such a model situation the phrase "for any doubly stochastic measure" means a unique such measure; but, nevertheless, this example emphasizes the expediency of passing from the consideration of the whole set F to the consideration of its minimization $\mathrm{mi}_{\mu,\nu} F$.

If we are interested only in doubly stochastic densities on a set $F \in \mathfrak{F}$, then we can show that on such a set there exists a doubly stochastic density if and only if the conditional measures of the set $\mathrm{mi}_{\mu,\nu} F$ on the elements of the coordinate decompositions are almost all positive. However, we can establish a criterion for the existence of a doubly stochastic density on a subset $F \in \mathfrak{A} \bar{\otimes} \mathfrak{B}$ without assuming that it belongs to the class $\mathfrak{F}$. For an arbitrary set F it can happen that the set

$$\bar{F} = (X \times Y) \setminus \sup (A \times B : (\mu \times \nu)((A \times B) \cap F) = 0)$$

of the class $\mathfrak{F}$ is larger than F by a subset of positive $(\mu \times \nu)$-measure; therefore, the collection of doubly stochastic measures on F and on its "closure" $\bar{F}$ can be different. As an example we can take $(X, \mu) = (Y, \nu) = ([0, 1], l)$ and $F = \{(x, y): x + y \in T\}$. where l is Lebesgue measure and T is a subset of the real line whose complement has positive measure and whose intersection with any segment has positive measure (it is known [22] that if $lA > 0$ and $lB > 0$, then $A + B$ contains a segment and intersects F in a set of positive measure). Therefore, it is not always possible to choose in the class of sets that are $(\mu \times \nu)$-equivalent to a given one a set in the class $\mathfrak{F}$.

Nevertheless, we consider another form of "minimization" of a subset F, one that is determined to within a subset of $(\mu \times \nu)$-measure zero.

DEFINITION 11. The *strong minimization* of a set $F \in (\mathfrak{A} \bar{\otimes} \mathfrak{B})/(\mu \times \nu)$ is defined to be the set

$$\mathrm{mi}^*_{\mu,\nu} F = F \setminus \sup \{A \times B : (\mu \times \nu)((X \setminus A) \times (Y \setminus B)) \cap F = 0\},$$

where the sup is taken in the sense of the lattice order on $(\mathfrak{A} \bar{\otimes} \mathfrak{B})/(\mu \times \nu)$.

Unlike $\mathrm{mi}_{\mu,\nu}$, the operation $\mathrm{mi}^*_{\mu,\nu}$ is not idempotent, as shown by the following example (see Figure 8). In the example $\mathrm{mi}^*F \neq \varnothing \pmod{\mu \times \nu}$, while $\mathrm{mi}^*(\mathrm{mi}^*F) = \varnothing$. It can be shown that $(\mathrm{mi}^*)^3 = (\mathrm{mi}^*)^2$.

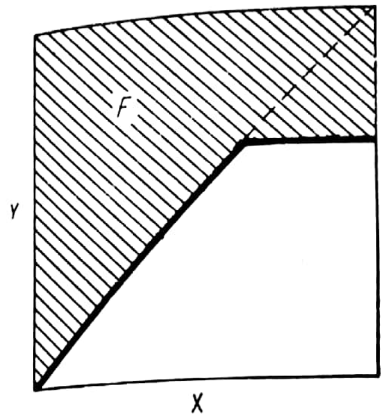

FIGURE 8

Obviously, the satisfaction of the conditions of Theorem 9 for the set mi^*F is necessary for the existence on it of a doubly stochastic measure. The necessity of the positivity of almost all conditional measures of the set mi^*F on the elements of the coordinate decompositions is just as obvious. We prove that the satisfaction of these two conditions, the second of which is equivalent to the condition

$$\mathrm{mi}^*_{\mu,\nu}(\mathrm{mi}^*_{\mu,\nu}F) \neq \varnothing \pmod{\mu \times \nu},$$

is also sufficient for the existence on F of a doubly stochastic density. In particular, to solve the problem of the existence of a doubly stochastic measure that is absolutely continuous with respect to some measure m that, in turn, is absolutely continuous with respect to the measure $\mu \times \nu$, it suffices to solve the problem of the existence of some (not necessarily doubly stochastic) probability measure m_1 that is absolutely continuous with respect to the measure m, satisfies the condition

$$A \in \mathfrak{A},\ B \in \mathfrak{B},\ \mu A + \nu B = 1,\ m_1(A \times B) = 0 \Rightarrow m_1((X \setminus A) \times (Y \setminus B)) = 0$$

and is such that its marginal distributions are equivalent to the measures μ and ν. (Indeed, the subset on which its Radon-Nikodým derivative with respect to $\mu \times \nu$ is positive is in this case contained in the set mi^*F.)

We precede the proof of this theorem by some auxiliary results.

PROPOSITION 72. *Let (m_{ik}) $(i, k = 1, 2)$ be a 2×2 matrix such that $m_{ik} \geqslant 0$, let b_i and a_k, $i, k = 1, 2$, be nonnegative numbers such that $a_1 + a_2 = 1$, $b_1 + b_2 = 1$, and let $0 \leqslant s \leqslant 1$. There exist numbers $m_{ik}^{(s)}$ for which*

$$0 \leqslant m_{ik}^{(s)} \leqslant m_{ik},\quad i, k = 1, 2, \tag{62}$$

$$\begin{aligned} m_{11}^{(s)} + m_{21}^{(s)} &= sa_1, \\ m_{12}^{(s)} + m_{22}^{(s)} &= sa_2, \\ m_{11}^{(s)} + m_{12}^{(s)} &= sb_1, \\ m_{21}^{(s)} + m_{22}^{(s)} &= sb_2, \end{aligned} \tag{63}$$

if and only if

$$\begin{aligned} m_{11} + m_{21} &\geqslant sa_1, \\ m_{12} + m_{22} &\geqslant sa_2, \\ m_{11} + m_{12} &\geqslant sb_1, \\ m_{21} + m_{22} &\geqslant sb_2, \end{aligned} \tag{64}$$

$$m_{ik} \geqslant s(b_i + a_k - 1),\quad i, k = 1, 2. \tag{65}$$

In other words, for the solvability of the stated "2×2 problem" we have, besides the trivially necessary conditions (64), the necessary and sufficient conditions (65), of which perhaps two are not completely trivial.

PROOF. For definiteness we assume that $a_2 + b_2 \geqslant 1$ and $a_1 + b_2 \geqslant 1$, so that $a_1 + b_1 \leqslant 1$ and $a_2 + b_1 \leqslant 1$. It suffices to limit ourselves to the case $s = 1$, replacing the matrix (m_{ik}) by $(s^{-1}m_{ik})$ for arbitrary $s > 0$. We show that the inequalities (64) and (65) (for $s = 1$) distinguish in the four-dimensional space of all matrices (m_{ik}) the set of all those matrices that majorize at least one "doubly stochastic" matrix, i.e., a matrix $(m_{ik}^{(1)})$ for which the equalities (63) hold. First we describe the set of all such doubly stochastic matrices. The conditions of double stochasticity (63) impose on the four elements m_{ik} three linearly independent constraints so that the subset of all doubly stochastic matrices is a one-dimensional and, clearly, convex subset of the four-dimensional linear space of all matrices, i.e., a segment. The endpoints of this segment are the points (matrices)

$$R = \begin{pmatrix} 0 & b_1 \\ a_1 & a_2 + b_2 - 1 \end{pmatrix} \text{ and } S = \begin{pmatrix} b_1 & 0 \\ a_1 + b_2 - 1 & a_2 \end{pmatrix}$$

(the presence of zero elements means that these matrices are extreme points of the segment of doubly stochastic matrices). We are interested in the unbounded convex polyhedron consisting of all matrices majorizing some point of the segment RS. This polyhedron can be regarded as the convex hull of two cones: the cone of matrices majorizing the matrix R and the cone of matrices majorizing S; or as the set of all points in such a cone when its vertex is translated along the segment RS. Each face of such a polyhedron belongs to at least one of the following two types: I) the faces belonging to the cones with vertices at R and S; II) the faces that are swept out by the two-dimensional faces of the cone when it is moved along RS. The faces of type I are given by those inequalities determining the faces of the cones with vertices at R and at S that are satisfied for all the points of both these cones. These are the following four inequalities:

$$\begin{aligned} &m_{11} \geqslant 0, \\ &m_{12} \geqslant 0, \\ &m_{21} \geqslant a_1 + b_2 - 1 \quad (\geqslant 0), \\ &m_{22} \geqslant a_2 + b_2 - 1 \quad (\geqslant 0) \end{aligned} \tag{66}$$

(since $a_1 + b_2 - 1 = a_1 - (1 - b_2) \leqslant a_1$, and $a_2 + b_2 - 1 \leqslant a_2$). Regarding the faces of type II, to obtain the corresponding inequalities we find all the faces of the infinite "prism" formed by the union of those points in the space of all matrices that majorize some matrix lying on RS. The linear inequalities distinguishing the faces of this prism are clearly the inequalities distinguishing the faces of type II.

The direction vector of RS is proportional to the matrix

$$\Delta = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

The linear functionals determining the faces of this prism must vanish on Δ. A basis in the three-dimensional space of such functionals is formed by any three of the following four functionals vanishing on Δ:

$$f_1((m_{ik})) = m_{11} + m_{21},$$
$$f_2((m_{ik})) = m_{12} + m_{22},$$
$$g_1((m_{ik})) = m_{11} + m_{12},$$
$$g_2((m_{ik})) = m_{21} + m_{22}.$$

In this connection, if the inequalities (64) hold (for $s = 1$), then we can make all the inequalities (64) pass into equalities by decreasing the numbers m_{ik}, and the new values of the m_{ik} (not necessarily nonnegative) form a matrix lying on RS. By the same token we have proved that the inequalities (64) describe all the faces of type II. Thus, to verify the existence of a matrix lying on RS and majorized by the matrix (m_{ik}) it suffices to verify the inequalities (66), i.e., (65) and (64). •

REMARK. It is clearly possible to assume that some of the quantities m_{ik} take the value $+\infty$.

PROPOSITION 73. *For some measure m on $X \times Y$ that is absolutely continuous with respect to the measure $\mu \times \nu$ and has bounded derivative $dm/d(\mu \times \nu) \leqslant K < \infty$ suppose that, for any $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$,*

$$m(A \times B) > \mu A + \nu B - 1. \tag{67}$$

Then for any measurable functions $a(x)$ and $b(y)$ satisfying $0 \leqslant a(x) \leqslant 1$ and $0 \leqslant b(y) \leqslant 1$ the following relation holds:

$$\int_{X \times Y} a(x)\, b(y)\, dm \geqslant \int a(x)\, d\mu + \int b(y)\, d\nu - 1. \tag{68}$$

PROOF. It is well known that the set of functions of the form $\chi_A(x)$, $A \in \mathfrak{A}$, is dense in the set of all measurable functions $a(x)$ satisfying the condition $0 \leqslant a(x) \leqslant 1$, and, similarly, the set $\{\chi_B(y), B \in \mathfrak{B}\}$ is dense in the set $\{b(y)\colon\ 0 \leqslant b(y) \leqslant 1\}$, in the topologies $\sigma(L^\infty(X, \mu), L(X, \mu))$ and $\sigma(L^\infty(Y, \nu), L(Y, \nu))$, respectively. As mentioned in the proof of Proposition 49, the convergence $\chi_{A_n}(x) \to a(x)$ and $\chi_{B_n}(y) \to b(y)$ implies the convergence

$$\chi_{A_n}(x)\, \chi_{B_n}(y) \to a(x)\, b(y)$$

in the topology $\sigma(L^\infty(X \times Y, m), L(X \times Y, m))$; therefore,

$$m(A_n \times B_n) = \int_{X \times Y} \chi_{A_n \times B_n}(x, y)\, dm \to \int_{X \times Y} a(x)\, b(y)\, dm. \tag{69}$$

Since, also,

$$\mu A_n = \int_X \chi_{A_n}(x)\, d\mu \to \int_X a(x)\, d\mu, \quad \nu B_n \to \int b(y)\, d\nu,$$

the required relation follows from (67) and (69). •

REMARK. We have actually proved that (68) holds if the restriction of the measure m to $\bigcup_n (A_n \times B_n)$ satisfies the conditions of Proposition 72, because on the complement of this set the function $a(x)b(y)$ is equal to zero, and the character of m on this complement does not play any role.

We now prove an assertion that we shall use to get a strengthening of Theorem 6 that is necessary for our purposes.

PROPOSITION 74. *Let* $(\tilde{m}_{ik})$ $(i, k = 1, \ldots, n)$ *be a square matrix whose elements are nonnegative numbers or* $+\infty$, *and let* (μ_k) *and* (ν_i) *be sets of nonnegative numbers such that* $\Sigma\mu_k = \Sigma\nu_i = 1$. *For there to exist a matrix* (m_{ik}) *for which*

$$\sum_{k=1}^{n} m_{ik} = \nu_i, \quad \sum_{i=1}^{n} m_{ik} = \mu_k, \quad 0 \leqslant m_{ik} \leqslant \tilde{m}_{ik},$$

it is necessary and sufficient that for the matrix $(\tilde{m}_{ik})$, *regarded as a nonnegative measure on the set* $X \times Y$ *of indices, each coarsened* 2×2 *problem for the marginal distributions* $\mu = (\mu_k, k = 1, \ldots, n)$, $\nu = (\nu_i, i = 1, \ldots, n)$ *is solvable, i.e., that the condition*

$$\tilde{m}(A \times B) \geqslant \mu A + \nu B - 1 \tag{70}$$

holds for any subsets A and B of the index sets X and Y.

PROOF. If any coarsened 2×2 problem is solvable for the matrix $(\tilde{m}_{ik})$, then, since there are only a finite number of such problems, we can replace those elements $\tilde{m}_{ik}$ that equal $+\infty$ by sufficiently large positive numbers, thereby reducing the proof of Proposition 74 to the proof of the analogous assertion for a matrix $(\tilde{m}_{ik})$ that can be regarded as a finite nonnegative measure on the set of indices $X \times Y$. Normalization reduces the proof to an application of Theorem 9. ●

We state and prove an interesting generalization of Theorem 6.

PROPOSITION 75. *Let* $\tilde{m}$ *be an arbitrary nonnegative σ-finite measure defined on the σ-algebra* $\mathfrak{A} \bar{\otimes} \mathfrak{B}$ *of the space* $X \times Y$, *and* μ *and* ν *probability measures on the spaces* $(X, \mathfrak{A})$ *and* $(Y, \mathfrak{B})$. *For there to exist a probability measure m that is absolutely continuous with respect to* $\tilde{m}$ *and has density* $dm/d\tilde{m}$ *not exceeding* 1 *and whose marginal distributions are* μ *and* ν, *it is sufficient (and, clearly, necessary) that any coarsened* 2×2 *problem is solvable, i.e., that for any subsets* $A \in \mathfrak{A}$ *and* $B \in \mathfrak{B}$ *the condition* (70) *holds.*

PROOF. The set of locally finite measures having bounded densities with respect to $\tilde{m}$ can be naturally identified with the space $L^\infty(X \times Y, \tilde{m})$, whose unit ball is weakly compact (in the dual space topology). If any coarsened 2×2 problem is solvable, then, by Proposition 74, any coarsened $n \times n$ problem is solvable. We now fix refining sequences of finite decompositions of the spaces X and Y such that $\xi_1 < \xi_2 < \cdots \to \xi$ and $\eta_1 < \eta_2 < \cdots \to \eta$. For each pair of decompositions (ξ_n, η_n) there is a probability measure m_n whose density with respect to $\tilde{m}$ does not exceed 1 and whose marginal distributions μ_n and ν_n are such that for any element X' of ξ_n we have $\mu_n X' = \mu X$, and, similarly, for any $Y' \in \eta_n$ we have $\nu_n Y' = \nu Y$. Because of the weak compactness of the unit ball of $L^\infty(X \times Y, \tilde{m})$ we can assume that the sequence m_n converges to a limit measure m, which is also absolutely continuous with respect to $\tilde{m}$ with density not exceeding 1. By the definition of weak convergence, for any set

$X' \in \xi_k$ and any $n \geqslant k$ we have

$$m_n(X' \times Y) \to m(X' \times Y), \quad m_n(X' \times Y) = \mu_n X' = \mu X',$$

i.e., $m(X' \times Y) = \mu X'$, and similarly for ν, from which it follows that μ and ν are the marginal distributions of this measure m. ●

We proceed to the proof of the assertion on the existence of doubly stochastic densities.

THEOREM 10. *Let $(X, \mathfrak{A}, \mu)$ and $(Y, \mathfrak{B}, \nu)$ be two separable probability measure spaces, and let $F \subset X \times Y$ be an $(\mathfrak{A} \bar{\otimes} \mathfrak{B})/(\mu \times \nu)$-measurable subset. For there to exist on F a probability measure m that is absolutely continuous with respect to $\mu \times \nu$ and for which the doubly stochastic condition $\mu = m\pi_X^{-1}$, $\nu = m\pi_Y^{-1}$ holds, it is necessary and sufficient that almost all conditional measures of the subset*

$$\begin{aligned}\mathrm{mi}^*_{\mu,\nu} F = F \setminus \sup \{A \times B : A \in \mathfrak{A}/\mu,\ B \in \mathfrak{B}/\nu,\ \mu A + \nu B = 1,\\ (\mu \times \nu)(F \cap ((X \setminus A) \times (Y \setminus B))) = 0\}\end{aligned}$$

on elements of the coordinate decompositions are positive.

We remarked above that this condition is equivalent to the condition

$$\mathrm{mi}^*_{\mu,\nu}(\mathrm{mi}^*_{\mu,\nu} F) \neq \varnothing \pmod{(\mu \times \nu)}. \tag{71}$$

PROOF. The necessity was discussed above and is completely clear. For a proof of the sufficiency we use Proposition 74 and construct on F a nonnegative σ-finite measure $\widetilde{m}$ that is absolutely continuous with respect to $\mu \times \nu$ and such that, for any measurable decompositions $\hat{\xi}$ and $\hat{\eta}$ of (X, μ) and (Y, ν) into two subsets each, there exists a 2×2 matrix (m_{ik}) $(i, k = 1, 2)$ majorized by the 2×2 matrix $(\widetilde{m}_{ik})$ of marginal measures of the space $(X \times Y, \widetilde{m})$ under the decomposition $\hat{\xi}\hat{\eta}$ and such that the marginal distributions for the matrix (m_{ik}), regarded as a measure on a four-point space, coincide with the corresponding marginal measures of the spaces (X, μ) and (Y, ν) on the elements of $\hat{\xi}$ and $\hat{\eta}$. In other words, we construct a nonnegative σ-finite measure $\widetilde{m}$ that is absolutely continuous with respect to $\mu \times \nu$, is such that $\widetilde{m}((X \times Y) \setminus F) = 0$, and for which each coarsened 2×2 problem associated with the marginal distributions μ and ν is solvable. Moreover, as shown by Proposition 72, it suffices to construct a measure $\widetilde{m}$ on F that is absolutely continuous with respect to $\mu \times \nu$ and such that for any $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$

$$\widetilde{m}(A \times B) \geqslant \mu A + \nu B - 1. \tag{72}$$

By assumption, $\mathrm{mi}^* F \neq \varnothing$, and the marginal distributions of the measure $(\mu \times \nu)_{\mathrm{mi}^* F}$ are mutually absolutely continuous with respect to μ and ν, i.e.,

$$\frac{1}{(\mu \times \nu)_{\mathrm{mi}^* F}} \geqslant u(x) \overset{\text{def}}{=} \frac{d\,[(\mu \times \nu)_{\mathrm{mi}^* F}\pi_X^{-1}]}{d\mu} > 0 \quad \mu\text{-almost everywhere},$$

and

$$\frac{1}{(\mu\times\nu)_{\mathrm{mi}^*F}} \geqslant v(y) \overset{\mathrm{def}}{=} \frac{d\left[(\mu\times\nu)_{\mathrm{mi}^*F}\pi_Y^{-1}\right]}{d\nu} > 0 \quad \nu\text{-almost everywhere.}$$

The measure $\widetilde{m}$ of interest to us is obtained as a sum of measures: $\widetilde{m} = \widetilde{m}_1 + \widetilde{m}_2 + \widetilde{m}_3$, where $d\widetilde{m}_1/d(\mu \times \nu)$ depends only on x on the set mi^*F, $d\widetilde{m}_2/d(\mu \times \nu)$ depends only on y on mi^*F, $d\widetilde{m}_3/d(\mu \times \nu) = \mathrm{const}$, and all three of these measures vanish outside mi^*F. Moreover, each of the measures $\widetilde{m}_1$, $\widetilde{m}_2$ and $\widetilde{m}_3$, and consequently their sum, is σ-finite. The inequality (72) is satisfied for $\widetilde{m}_1$ if $\mu A \leqslant 1/2$, for $\widetilde{m}_2$ if $\mu B \leqslant 1/2$, and for $\widetilde{m}_3$ if $\mu A > 1/2$ and $\mu B > 1/2$.

We prove the existence of a measure m_1 with the required properties, i.e., such that $\widetilde{m}_1(A \times B) > \mu A + \nu B - 1$ for $\mu A \leqslant 1/2$. For brevity we write

$$t = t(A, B) = \mu A + \nu B - 1.$$

Let $X = X_1 \cup X_2 \cup \cdots$ be a decomposition of X into pairwise disjoint subsets such that

$$\operatorname{ess\,inf}\{u(x) : x \in X_k\} = u_k > 0.$$

The density $d\widetilde{m}_1/d(\mu \times \nu)_{\mathrm{mi}^*F}$ on each of the sets $\mathrm{mi}^*F \cap (X_k \times Y)$ is equal to a constant p_k, and these constants are determined successively.

Let $l_1 > l_2 > \cdots > 1$ be a decreasing numerical sequence. We first determine the constant p_1. This constant is chosen so that for $\mu A \leqslant 1/2$ we have

$$\widetilde{m}_1^{(1)}(A \times B) > l_1(\mu A + \nu B - 1),$$

where $\widetilde{m}_1^{(1)}$ is the measure defined by

$$\frac{d\widetilde{m}_1^{(1)}}{d(\mu\times\nu)_{\mathrm{mi}^*F}} = \begin{cases} p_1 & \text{for those } (x, y) \in \mathrm{mi}^*F \text{ for which } x \in X_1, \\ +\infty & \text{for those } (x, y) \in \mathrm{mi}^*F \text{ for which } x \notin X_1. \end{cases}$$

We show that the required number p_1 exists. Let

$$n(t, \alpha) = \inf\left\{(\mu\times\nu)_{\mathrm{mi}^*F}(A \times B) : \mu A + \nu B - 1 = t, \tfrac{1}{2} \geqslant \mu A \geqslant \alpha > 0\right\}. \tag{73}$$

We can assume that for each $\alpha > 0$

$$\inf\{t : n(t, \alpha) > 0\} < 0. \tag{74}$$

Indeed, on the space $(X \times Y, (\mu \times \nu)_{\mathrm{mi}^*F})$ consider the decomposition $\xi \wedge \eta$. If $\xi \wedge \eta = \nu$ (the trivial decomposition), then there do not exist subsets $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$ such that

$$(\mu\times\nu)_{\mathrm{mi}^*F}(A \times B) = 0 \quad \text{and} \quad (\mu\times\nu)_{\mathrm{mi}^*F}((X\setminus A) \times (Y \setminus B)) = 0$$

simultaneously; but, as remarked, if $\mathrm{mi}^*F \neq \varnothing$, then these conditions imply that $\mu A + \nu B = 1$. Therefore, the condition $\xi \wedge \eta = \nu$ for $(\mu \times \nu)_{\mathrm{mi}^*F}$ is equivalent to the condition

$$\text{if } \mu A + \nu B = 1, \text{ then } (\mu\times\nu)_{\mathrm{mi}^*F}(A \times B) > 0.$$

For each $\alpha > 0$ we have here the condition (74), since, otherwise, we get from Proposition

49 that if $\inf\{t: n(t, \alpha_0) > 0\} = 0$ for some $\alpha_0 > 0$, then

$$(\mu \times \nu)_{\mathrm{mi}*F}(A_0 \times B_0) = 0$$

for certain $A_0 \in \mathfrak{A}$ and $B_0 \in \mathfrak{B}$ for which $\frac{1}{2} \geqslant \mu A_0 \geqslant \alpha_0$ and $\mu A_0 + \nu B_0 = 1$; and this contradicts the condition $\xi \wedge \eta = \nu$. But if $\xi \wedge \eta \neq \nu$, then, by the remark after Proposition 50, the absolute continuity of $(\mu \times \nu)_{\mathrm{mi}*F}$ with respect to $\mu \times \nu$ implies that the decomposition $\xi \wedge \eta$ is not more than countable, and it is possible as the set mi*F to consider separately the intersection of mi*F with each element of $\xi \wedge \eta$.

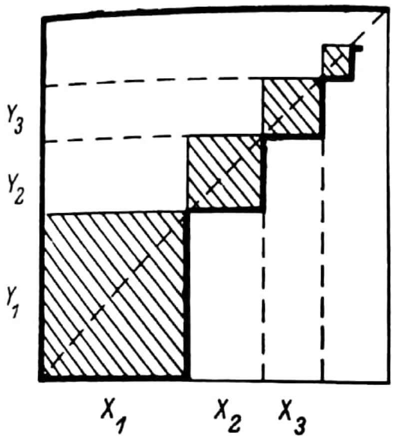

FIGURE 9

In Figure 9 the measures μ and ν are Lebesgue measures on the sides of the square, the set F has the thick contour, the set mi*F is shaded, ξ and η, as always, are the decompositions of the square into vertical and horizontal segments, and the decomposition $\xi \wedge \eta$ of $(X \times Y, (\mu \times \nu)_{\mathrm{mi}*F})$ consists (mod 0) of the sets $X_1 \times Y_1$, $X_2 \times Y_2, \ldots$, and $\mu X_1 = \nu Y_1$, $\mu X_2 = \nu Y_2$, etc.

On each of the subspaces $X_1 \times Y_1$, $X_2 \times Y_2$, etc., of $(X \times Y, (\mu \times \nu)_{\mathrm{mi}*F})$ the condition (74) holds for each $\alpha > 0$. It is clear that the minimization of $F \cap (X_k \times Y_k)$ coincides with the set $(\mathrm{mi}*F) \cap (X_k \times Y_k)$; therefore, it suffices to prove the theorem under the assumption that $\xi \wedge \eta = \nu$ on $(X \times Y, (\mu \times \nu)_{\mathrm{mi}*F})$, i.e., under the assumption that (74) holds. As a consequence of (74), we have for $\frac{1}{2} \geqslant \mu A \geqslant \alpha > 0$ and for $\mu A + \nu B \geqslant 1$ the condition

$$(\mu \times \nu)_{\mathrm{mi}*F}(A \times B) \geqslant n(\mu A + \nu B - 1, \alpha) \geqslant n(0, \alpha) \geqslant 0. \tag{75}$$

On the other hand, if $\mu A \leqslant \alpha$ and $\mu A + \nu B - 1 \geqslant 0$, then $\nu B \geqslant 1 - \alpha$. If

$$\tilde{m}_1^{(1)}(A \times B) < \infty,$$

then automatically

$$((A \cap (X \setminus X_1)) \times B) \cap \mathrm{mi}^* F = \varnothing,$$

from which it follows that

$$\mu(A \cap (X \setminus X_1)) + \nu B \leqslant 1$$

and, consequently,

$$\mu(A \cap X_1) = \mu A - \mu(A \cap (X \setminus X_1)) \geqslant \mu A + \nu B - 1.$$

We get for $\mu A \leqslant \alpha$ that

$$\begin{aligned}(\mu \times \nu)_{\mathrm{mi}*F}(A \times B) &\geqslant (\mu \times \nu)_{\mathrm{mi}*F}((A \cap X_1) \times B) \\ &\geqslant (\nu B + u_1 - 1)\,\mu(A \cap X_1)\,[(\mu \times \nu)_{\mathrm{mi}*F}]^{-1} \\ &\geqslant (u_1 - \alpha)\,[(\mu \times \nu)_{\mathrm{mi}*F}]^{-1}(\mu A + \nu B - 1).\end{aligned} \tag{76}$$

It follows from (75) and (76) that if $0 < \alpha < u_1$ and if $p_1 > l_1(u_1 - \alpha)$ and $p_1 > l_1/n(0, \alpha)$, then

$$\tilde{m}_1 (A \times B) > l_1 (\mu A + \nu B - 1)$$

for $t = \mu A + \nu B - 1 > 0$.

We now define the measure $\tilde{m}_1^{(k)}$ inductively. Let $\tilde{m}_1^{(k)}$ be such that

$$\frac{d\tilde{m}_1^{(k)}}{d (\mu \times \nu)\ \mathrm{mi}^* F} = p_i$$

on the set $(\mathrm{mi}^*F) \cap (X_i \times Y)$ for $i = 1, \ldots, k$, and

$$\frac{d\tilde{m}_1^{(k)}}{d (\mu \times \nu)_{\mathrm{mi}^* F}} = +\infty$$

on the remaining part of mi^*F, and suppose that

$$\tilde{m}_1^{(k)} (A \times B) > l_k (\mu A + \nu B - 1)$$

for any $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$. We prove that there exists a constant p_{k+1} for which the measure $\tilde{m}_1^{(k+1)}$, that differs from $\tilde{m}_1^{(k)}$ only in the fact that on the set $X_{k+1} \times Y$ we have

$$\frac{d\tilde{m}_1^{(k+1)}}{d (\mu \times \nu)_{\mathrm{mi}^* F}} = p_{k+1} < +\infty,$$

satisfies for any $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$ the inequality

$$\tilde{m}_1^{(k+1)} (A \times B) > l_{k+1} (\mu A + \nu B - 1).$$

Suppose, on the contrary, that for any p_{k+1} there are sets $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$ such that the opposite inequality holds. Then let $p_{k+1}^{(n)} \underset{n}{\longrightarrow} +\infty$, let $\tilde{m}_1^{(k+1,n)}$ be a measure differing from $\tilde{m}_1^{(k)}$ only in the fact that on $X_{k+1} \times Y$ we have

$$\frac{d\tilde{m}_1^{(k+1,n)}}{d (\mu \times \nu)_{\mathrm{mi}^* F}} = p_{k+1}^{(n)},$$

and let $A_n \in \mathfrak{A}$ and $B_n \in \mathfrak{B}$ be such that $\mu A_n \leqslant ½$ and

$$\tilde{m}_1^{(k+1,n)} (A_n \times B_n) \leqslant l_{k+1} (\mu A_n + \nu B_n - 1) \quad (n = 1, \ldots). \tag{77}$$

We use the notation

$$A_n' = \bigcup_{i=1}^{k} (A_n \cap X_i), \quad A_n'' = A_n \cap X_{k+1},$$

$$A_n''' = \bigcup_{i=k+2}^{\infty} (A_n \cap X_i).$$

We first assume that for some $\alpha > 0$ and any n we have $\mu A_n \geqslant \alpha$. The sets $\{\chi_{A_n}(x)\}$ and $\{\chi_{B_n}(y)\}$ are weakly compact. Let

$$a (x) = \lim_i \chi_{A_{n_i}} (x) \quad \text{and} \quad b (y) = \lim_i \chi_{B_{n_i}} (y),$$

the convergence, as usual, being understood in the sense of the topology $\sigma(L^\infty, L)$. As follows from Remark 2 after Proposition 49, this implies the convergence

$$\chi_{A_{n_i}\times B_{n_i}}(x, y)=\chi_{A_{n_i}}(x)\chi_{B_{n_i}}(y)\underset{i}{\rightarrow} a(x)b(y)$$

in an appropriate weak topology—in

$$\sigma(L^\infty(X\times Y, (\mu\times\nu)_{\mathrm{mi}^*F}), L(X\times Y, (\mu\times\nu)_{\mathrm{mi}^*F}))$$

for example; therefore, for each m

$$\tilde{m}_1^{(k+1, m)}(A_{n_i}\times B_{n_i})\rightarrow\int_{X\times Y} a(x)b(y)\,d\tilde{m}_1^{(k+1, m)},$$

but for $n_i \geqslant m$ it follows from (77) that

$$\tilde{m}_1^{(k+1, m)}(A_{n_i}\times B_{n_i})=l_{k+1}(\mu A_{n_i}+\nu B_{n_i}-1)$$

and, moreover,

$$\mu A_{n_i}\rightarrow\int_X a(x)\,d\mu, \quad \nu B_{n_i}\rightarrow\int_Y b(y)\,d\nu,$$

from which we get

$$\int_{X\times Y} a(x)b(y)\,d\tilde{m}_1^{(k+1, m)}\leqslant\int_X a(x)\,d\mu+\int_Y b(y)\,d\nu-1, \quad m=1, 2, \ldots,$$

whence

$$\int_{X\times Y} a(x)b(y)\,d\tilde{m}_1^{(k)}\leqslant l_{k+1}\left(\int_X a(x)\,d\mu+\int_Y b(y)\,d\nu-1\right). \tag{78}$$

But, by Proposition 72 and the remark after it, for the same functions

$$a(x)=\lim_i\chi_{A_{n_i}}(x), \quad b(y)=\lim_i\chi_{B_{n_i}}(y)$$

we have

$$\int_{X\times Y} a(x)b(y)\,d\tilde{m}_1^{(k)}\geqslant l_k\left(\int_X a(x)\,d\mu+\int_Y b(y)\,d\nu-1\right). \tag{79}$$

Comparison of (78) and (79) shows the impossibility of our assumption that (77) holds and $\mu A_n \geqslant \alpha > 0$. Suppose, on the other hand, that (77) holds and $\mu A_n \longrightarrow 0$. For brevity we write

$$\mu A_n'=a_n', \quad \mu A_n''=a_n'', \quad \mu A_n'''=a_n''', \quad \nu B_n=b_n.$$

We can limit ourselves to a consideration only of the nontrivial case when $t = \mu A_n + \nu B_n - 1 > 0$, so that $b_n \longrightarrow 1$. Also, it suffices to assume that $\tilde{m}_1^{(k+1,n)}(A_n \times B_n) < \infty$, i.e., $(A_n''' \times B_n) \cap \mathrm{mi}^*F = \varnothing$. The following relations are obvious:

$$\begin{aligned}\tilde{m}_1^{(k+1, n)}(A_n\times B_n)&=\tilde{m}_1^{(k+1, n)}(A_n'\times B_n)+\tilde{m}_1^{(k+1, n)}(A_n''\times B_n)\\&=\tilde{m}_1^{(k)}((A_n'\cup A_n''')\times B_n)+\tilde{m}_1^{(k+1, n)}((A_n''\cup A_n''')\times B_n)\\&\geqslant l_k(a_n'+a_n'''+b_n-1)+p_{k+1}^{(n)}(u_{k+1}+b_n-1)a_n''.\end{aligned}$$

For sufficiently large n we have

$$p_{k+1}^{(n)} > \frac{l_k}{u_{k+1} + b_n - 1} > 0$$

and consequently

$$\tilde{m}_1^{(k+1,\, n)} (A_n \times B_n) > l_k (a'_n + a''_n + a'''_n + b_n - 1) > l_{k+1} (\mu A_n + \nu B_n - 1)$$

contrary to (77). This contradiction proves the existence of a number p_{k+1} with the required properties.

We now show that the measure $\tilde{m}_1$ constructed, which has density equal to the number p_k with respect to the measure $(\mu \times \nu)_{\mathrm{mi}^*F}$ on the set $(\mathrm{mi}^*F) \cap (X_k \times Y)$ $(k = 1, \dots)$, is indeed the required one, i.e., a σ-finite measure for which (72) holds for any $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$, provided that $\mu A \leqslant ½$. In fact, the σ-finiteness of $\tilde{m}_1$ is clear from the construction (we remark that the σ-finiteness of the projection of this measure onto Y is not guaranteed). Let $A \in \mathfrak{A}, B \in \mathfrak{B}$, and $\mu A \leqslant ½$. Then

$$\tilde{m}_1 (A \times B) = \lim \tilde{m}_1 (A'_n \times B) = \lim \tilde{m}_1^{(n)} (A'_n \times B) \geqslant \lim l_n (\mu A'_n + \nu B - 1).$$

It suffices to limit ourselves to the nontrivial case when $\mu A + \nu B - 1 \geqslant 0$. Under this assumption we get, finally,

$$\tilde{m}_1 (A \times B) \geqslant \mu A + \nu B - 1. \tag{80}$$

In a completely analogous way we construct a locally finite measure $\tilde{m}_2$ such that for any $A \in \mathfrak{A}$ and $B \in \mathfrak{B}$, $\nu B \leqslant ½$, we have

$$\tilde{m}_2 (A \times B) \geqslant \mu A + \nu B - 1. \tag{81}$$

But if $\mu A > ½$ and $\nu B > ½$, then $(\mu \times \nu)_{\mathrm{mi}^*F}(A \times B) \geqslant n(0, ½) > 0$; therefore, if we set $\tilde{m}_3 = (n(0, ½))^{-1}(\mu \times \nu)_{\mathrm{mi}^*F}$, it turns out that for $\mu A, \nu B > ½$

$$\tilde{m}_3 (A \times B) \geqslant 1 \geqslant \mu A + \nu B - 1. \tag{82}$$

Finally, (80)–(82) prove (72), from which, by Proposition 74, the theorem follows. •

REMARK 1. An assertion analogous to that of the above theorem is apparently true not only for the type of the product measure $\mu \times \nu$, but also for an arbitrary type; one should just replace the condition of nonemptiness of the minimization of the set modulo $\mu \times \nu$ by the condition of nonemptiness with respect to the corresponding type. However, in the following we need the theorem only in the given formulation. We emphasize that the principal difficulty that had to be overcome in the proof lay in the noncompactness of the set of doubly stochastic measures subordinate to a given type.

REMARK 2. By the same method it is possible to get the following slight extension of the theorem. Let θ_X and θ_Y be measurable decompositions of the spaces (X, μ) and (Y, ν) such that the spaces X/θ_X and Y/θ_Y are isomorphic, and let the set F be contained in the union of the direct products of the elements of θ_X and θ_Y that correspond under this isomorphism. For there to exist on F a doubly stochastic measure m having the property that under the decomposition of $(X \times Y, m)$ into the products of corresponding elements of θ_X and θ_Y (i.e., under the decomposition $\pi_X^{-1}\theta_X \wedge \pi_Y^{-1}\theta_Y$) its conditional measures on the elements of this decomposition are absolutely continuous with respect to the products of the conditional measures on the corresponding elements of θ_X and θ_Y, it is necessary and sufficient that the conditions of Theorem 10 hold on each typical element.

§12. Marginal sufficiency of statistics

We consider a problem connected with the concept of sufficiency, which is important to mathematical statistics. Suppose that we are given a family of probability measures $\{P_\theta, \theta \in \Theta\}$ on a space with distinguished σ-algebra $(X, \mathfrak{A})$. We assume that all the measures in this family are absolutely continuous with respect to some probability measure P. Let $y = f(x)$ be a measurable function (a statistic) defined on the space $(X, \mathfrak{A})$. The statistic f is said to be sufficient for the family P_θ if for any subset $A \in \mathfrak{A}$ the conditional probability $P_\theta(A \mid f(x))$ does not depend on the value of the parameter θ. When $(X, \mathfrak{A}, P)$ is a Lebesgue space and it is possible to speak of the conditional measures on the elements of the measurable decomposition ξ_f generated by the statistic $y = f(x)$, the sufficiency of the statistic means that these conditional measures on almost all elements of the decomposition ξ_f do not depend on θ. The words "almost all" are understood in the sense of the canonical measure on the set of elements of ξ_f, i.e., the measure on the quotient space $(X, \mathfrak{A}, P)/\xi_f$. As is clear from the definition, the property of sufficiency is, in essence, a property of the measurable decomposition generated by the statistic; therefore one frequently speaks of sufficient decompositions, and not of sufficient statistics, or, formally more general, of sufficient σ-algebras.

Suppose that, besides the statistic $y = f(x)$ on $(X, \mathfrak{A})$, we are given the statistics $y_1 = f_1(x), \ldots, y_n = f_n(x)$. A generalization of the concept of sufficiency is the following property of the statistic f: the conditional distributions of each of the statistics $f_1, \ldots, f_n$ for a fixed value of $f(x)$ do not depend on the parameter $\theta \in \Theta$. It is clear that if $y = f(x)$ is a sufficient statistic for the family $\{P_\theta, \theta \in \Theta\}$, then the conditional distribution of any other statistic $g(x)$ a fortiori does not depend on $\theta \in \Theta$. On the other hand, if the decomposition generated by a statistic $g(x)$ is coarser than the decomposition generated by $f(x)$, then from the fact that the conditional (for the condition $f(x)$) distribution of g does not depend on the parameter it does not follow that the statistic $f(x)$ is sufficient. The sufficiency of $f(x)$ also does not follow from the assumption that the conditional distributions of each of a number of statistics $f_1(x), \ldots, f_n(x)$ do not depend on the parameter, even if it is assumed that the product of the decompositions $\xi_{f_1} \vee \cdots \vee \xi_{f_n}$ is the decomposition ϵ of $(X, \mathfrak{A}, P)$ into points. The simplest example is the following:

$$X = \{x_1, x_2, x_3, x_4\};$$
$$f_1(x_1) = f_1(x_3) = f_2(x_1) = f_2(x_2) = 0;$$
$$f_1(x_2) = f_1(x_4) = f_2(x_3) = f_2(x_4) = 1;$$
$$f(x_1) = f(x_2) = f(x_3) = f(x_4) = 0;$$
$$P_\theta\{x_1\} = P_\theta\{x_4\} = \theta, \quad P_\theta\{x_2\} = P_\theta\{x_3\} = \frac{1}{2} - \theta, \quad \Theta = \left[0, \frac{1}{2}\right].$$

When $X = \mathbf{R}^n$, and $f_1(x), \ldots, f_n(x)$ are the coordinate functions (i.e., $f_k(x) = x_k$, $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$), a statistic f is said to be marginally sufficient if the conditional distributions of $f_1, \ldots, f_n$ for the condition $f(x)$ do not depend on $\theta \in \Theta$. In statistics, of course, the most frequently encountered case is that in which the distribution P_θ (for each value of θ) is a product of n measures (an independent sample). A number of years ago the Indian statistician V. S. Huzurbazar stated the conjecture that in the case of a repeated sample the marginal sufficiency of a statistic implies its sufficiency. In a 1968 preprint J. K. Ghosh [36] announced a proof of this conjecture; however, the present author and his colleagues (specialists in mathematical statistics) have encountered difficulties in reproducing the complete proof from Ghosh's outline. Below, the proof of a somewhat more general assertion is presented.

When all the probability measures of the family $\{P_\theta, \theta \in \Theta\}$ are absolutely continuous with respect to some fixed probability measure, the sufficiency of a statistic for any pair of distributions $(P_{\theta_1}, P_{\theta_2})$ in the family $\{P_\theta, \theta \in \Theta\}$ implies, as is well known (see, for example, [7]), its sufficiency with respect to the whole family. Therefore, we limit ourselves to the case when the set Θ consists of two elements.

THEOREM 11. *Let P and Q be two mutually absolutely continuous Borel probability measures in $\mathbf{R}^n$ that correspond to an independent sample. Let the statistic $y = f(x_1, \ldots, x_n)$ be marginally sufficient, i.e., for almost all (with respect to the distribution of f) values of y the conditions $(C_y, P_y)/\xi_k = (C_y, Q_y)/\xi_k$ $(k = 1, \ldots, n)$ hold, where the ξ_k are the coordinate decompositions, and P_y and Q_y are the conditional probability measures on the element $C_y \subset \mathbf{R}^n$ of the decomposition ξ_f. Then f is a sufficient statistic for the pair of distributions P and Q.*

In the case $n = 2$, and only in this case, the independence of the sample can be replaced by the requirement that $dQ/dP = q_1(x_1)q_2(x_2)$.

For arbitrary n the assumption of independence of the sample can be replaced by the assumption that for almost every (with respect to the distribution of f) number y there exist functions $p_1^y(x_1), \ldots, p_n^y(x_n)$ and numbers b_y and B_y such that

$$0 < b_y \leqslant p_k^y(x_1) \leqslant B_y < \infty, \quad k = 1, \ldots, n, \quad \text{and} \quad \frac{dQ_y}{dP_y} = p_1^y(x_1), \ \ldots, \ p_n^y(x_n).$$

We first prove two auxiliary statements.

PROPOSITION 76. *Let P and Q be two mutually absolutely continuous probability measures on the space $(\Omega, \mathfrak{A})$ and ξ a measurable decomposition for which there exist systems $\{P_C\}$ and $\{Q_C\}$ of conditional probability measures (C is an element of ξ). If $dQ/dP = q(\omega)$, then on almost every element $C \in \xi$ the measures P_C and Q_C are mutually absolutely continuous, and $dQ_C/dP_C = k(C)q(\omega)$ ($\omega \in C$; $k(C)$ is the value on the element C of the density $d(P/\xi)/d(Q/\xi)$).*

PROOF. For the proof it suffices to verify that the system of measures Q_C on the elements of ξ having densities with respect to the measures P_C equal to

$$\frac{dQ_C}{dP_C} = \frac{d(P/\xi)}{d(Q/\xi)} \frac{dQ}{dP}$$

really is the system of conditional measures for the measure Q under the decomposition ξ. We obtain

$$\int_{\Omega/\xi} d(Q/\xi) \int_{A\cap C} \frac{d(P/\xi)}{d(Q/\xi)} \frac{dQ}{dP} dP_C = \int_{\Omega/\xi} d(P/\xi) \int_{A\cap C} \frac{dQ}{dP} dP_C$$

$$= \int_{\Omega/\xi} d(P/\xi) \int_{A\cap C} q(\omega)\, dP_C = \int_A q(\omega)\, dP = \int_A dQ = Q(A). \bullet$$

PROPOSITION 77. *Let U and V be finite nonnegative Borel measures in $\mathbf{R}^n$ such that for some $l \geqslant 0$ the following conditions hold:*

1) $U\{x=(x_1, \ldots, x_n) \in \mathbf{R}^n : \sum x_k \leqslant l\}=0$;
2) $V\{x=(x_1, \ldots, x_n) \in \mathbf{R}^n : \sum x_k \geqslant l\}=0$;
3) $(\mathbf{R}^n, U)/\xi_k = (\mathbf{R}^n, V)/\xi_k$, $k=1, \ldots, n$;
4) $\int_{-\infty}^{\infty} |x_k|\, dU < \infty$, $k=1, \ldots, n$.

Then $U = V = 0$.

When $n = 2$ the conclusion remains true even without the last condition; when $n > 2$ the assumption of the existence of the first absolute moments cannot be omitted.

PROOF. We assume that $U \neq 0$ and $V \neq 0$. By the equality of the marginal distributions, the first absolute moments exist also for the measure V. For the same reason the point $b = (m_1, \ldots, m_n)$, where $m_k = |U|^{-1}\int_{-\infty}^{\infty} x_k dU$, $k = 1, \ldots, n$, is the barycenter both of U and of V. From 1) and 2) it follows that the barycenters of U and V are located on essentially different sides of the hyperplane defined by the equation $\Sigma x_k = l$. This contradiction proves the first part of the assertion.

When $n = 2$, for any nonnegative Borel measures U and V satisfying 1) and 2) we can select a function $f(x)$ such that the transformation F of $\mathbf{R}^2$ carrying (x, y) into $(f(x), l - f(l - y))$, and consequently carrying each of the half-spaces defined by the inequalities $x + y > l$ and $x + y < l$ into itself, transforms U and V into measures UF^{-1} and VF^{-1} having first moments. Since the coordinate decompositions are invariant with respect to the transformation F, the equality of the marginal distributions is preserved for the transformed measures, i.e., all the conditions 1)–4) hold; therefore the measures UF^{-1} and VF^{-1}, and with them U and V, must be zero.

Finally, we show that for $n > 2$ the conditions 1)–3) do not guarantee that $U = V = 0$. In $\mathbf{R}^3$ we construct two finite positive measures U and V concentrated on different sides of the hyperplane defined by the equation $x_1 + x_2 + x_3 = 0$ and having the same marginal distributions. We take U to be a purely atomic measure (of total mass ¾) concentrated on the hyperplane $x_1 + x_2 + x_3 = 1$:

three point masses of magnitude 1/8 at the points with coordinates $(1, 1, -1)$, $(1, -1, 1)$, $(-1, 1, 1)$;

three point masses of magnitude 1/16 at the points with coordinates $(-1, -1, 3)$, $(-1, 3, -1)$, $(3, -1, -1)$; . . . ;

three point masses of magnitude $2^{-(k+2)}$ at the points with coordinates $(2^k - 1, 1 - 2^{k-1}, 1 - 2^{k-1})$, $(1 - 2^{k-1}, 2^k - 1, 1 - 2^{k-1})$, $(1 - 2^{k-1}, 1 - 2^{k-1}, 2^k - 1)$, for $k = 2, 3, \ldots$.

As the measure V we take the image $U\varphi^{-1}$ of U under the mapping φ: $\mathbf{R}^3 \to \mathbf{R}^3$, $\varphi(x_1, x_2, x_3) = (-x_1, -x_2, -x_3)$. It is not hard to see that the marginal distributions of U and V coincide and are constructed in the following way: masses ¼ at the points -1 and $+1$, and masses $2^{-(k+2)}$ at the points $1 - 2^k$ and $2^k - 1$, $k = 2, 3, \ldots$. ●

PROOF OF THEOREM 11. We first consider the simpler case when the one-dimensional distributions μ_k and ν_k, $k = 1, \ldots, n$, whose products are the respective measures P and Q are such that $d\nu_k/d\mu_k \leqslant B$ and $d\mu_k/d\nu_k \leqslant B$. Let the mapping φ: $\mathbf{R}^n \to \mathbf{R}^n$ carry $(x_1, \ldots, x_n)$ into $(\ln q_1(x_1), \ldots, \ln q_n(x_n))$, where $q_k = d\nu_k/d\mu_k$. The mapping φ carries P and Q into the measures $\widetilde{P} = P\varphi^{-1}$ and $\widetilde{Q} = Q\varphi^{-1}$, while P_C and Q_C pass into measures $\widetilde{P}_C = P_C\varphi^{-1}$ and $\widetilde{Q}_C = Q_C\varphi^{-1}$, and $d\widetilde{Q}_C/d\widetilde{P}_C = k(C)\exp \Sigma_1^n x_k$, where $k(C) = d(P/\xi)/d(Q/\xi)$ (Proposition 71). Indeed, $\varphi^{-1}(x_1, \ldots, x_n)$ consists of those points $(x'_1, \ldots, x'_n)$ for which $\ln q_k(x'_k) = x_k$; consequently for them we have that

$$dQ_C/dP_C = k(C)\, q_1(x'_1) \ldots q_n(x'_n) = k(C)\, e^{\Sigma \ln q_k(x'_k)} = k(C)\, e^{\Sigma x_k},$$

and under the mapping φ, which is constant on the "level lines" of the density dQ/dP, the density of the images of the measures at some point is equal to the density of their preimages at any preimage of this point.

We note that the two-sided boundedness of the one-dimensional densities assumed here implies the two-sided boundedness of the logarithms of these densities, i.e., of the coordinates of the images (all to within subsets of measure zero). In other words, the measures $\widetilde{P}$ and $\widetilde{Q}$, and with them $\widetilde{P}_C$ and $\widetilde{Q}_C$, are concentrated in a bounded subset of $\mathbf{R}^n$; hence they have first moments.

Now let $\widetilde{Q}_C - \widetilde{P}_C = U - V$, where U and V are disjunct nonnegative measures. Obviously, U and V satisfy the conditions of Proposition 72. Indeed, 1) and 2) follow from the fact that

$$\frac{d\widetilde{Q}_C}{d\widetilde{P}_C} = k(C) \exp \sum x_k;$$

consequently, the signed measure $\widetilde{Q}_C - \widetilde{P}_C$ is nonnegative on the set where

$$k(C) \exp \sum x_k > 1,$$

i.e., for $\Sigma x_k > -\ln k(C)$, and it is nonpositive for $\Sigma x_k < -\ln k(C)$.

Thus, under the conditions 1) and 2) we should set $l = -\ln k(C)$.

The condition 3) means the marginal sufficiency, which is being assumed; and the condition 4) follows, as shown, from the assumption of boundedness of the one-dimensional densities. Consequently $U = V = 0$, i.e. $\widetilde{P}_C = \widetilde{Q}_C$, and the decomposition ξ is indeed sufficient for the pair of measures P and Q.

In the case $n = 2$ Proposition 77 is true without the requirement of the existence of first moments; therefore for $n = 2$ marginal sufficiency implies sufficiency, provided only that on each element C_y of ξ the density dQ_{C_y}/dP_{C_y} can be "factored", i.e., represented in the form $p_1^y(x_1)p_2^y(x_2)$. For arbitrary n our argument would not change if we required only the two-sided boundedness of the functions $p_1^y(x_1), \ldots, p_n^y(x_n)$ appearing in the factorization.

We now return to the general case and prove that U and V nevertheless have first absolute moments. Now, for the first time, we must make essential use of the fact that the distributions P and Q correspond to a repeated sample (up to this point we have only used the weaker assumption of factorization of the conditional densities).

First, let the statistic f be such that $P(C_{y_0}) > 0$ for some y_0. Obviously,

$$\frac{dQ_C}{dP_C} = \frac{P(C)}{Q(C)} q_1(x_1) \ldots q_n(x_n).$$

Further, let $s = \Sigma_1^n x_k$ and $R = P(C)/Q(C)$. We get

$$\frac{d\tilde{Q}_C}{d\tilde{P}_C} = R \exp s,$$

whence

$$\frac{dU}{d\tilde{P}_C} = \frac{d(\tilde{Q}_C - \tilde{P}_C)}{d\tilde{P}_C} = R \exp s - 1 \quad (\text{for } s > -\ln R)$$

and

$$\tilde{P}_C(dx) = (R \exp s - 1)^{-1} U(dx).$$

We now consider the decomposition of $(\mathbf{R}^n, U)$ into the hyperplanes

$$H_s = \{x = (x_1, \ldots, x_n) : \sum x_k = ns\},$$

and let $\{U_s\}$ be the family of conditional distributions (probability measures) on the hyperplanes $\{H_s\}$ and $\alpha(ds)$ the corresponding quotient measure on the set of such hyperplanes, i.e., on the set $[-\ln R, \infty)$ of values of the variable s, so that

$$U = \int_{-\ln R}^{\infty} U_s \alpha(ds) \qquad \text{and} \qquad \tilde{P}_C = \int_{-\ln R}^{\infty} \frac{U_s(dx)\,\alpha(ds)}{R \exp s - 1}.$$

Obviously $d\tilde{P}_C/d\tilde{P} \leqslant 1/P(C)$, and therefore

$$\int_{\mathbf{R}^n} \exp x_k d\tilde{P}_C \leqslant \int_{\mathbf{R}^n} \exp x_k d\tilde{P} \frac{1}{P(C)} = \int_{\mathbf{R}^n} dQ \frac{1}{P(C)} = \frac{1}{P(C)}$$

(here we use the assumption of an independent sample), i.e.,

$$\int_{\mathbf{R}^n} \exp x_k \int_{-\ln R}^{\infty} \frac{\alpha(ds)\, U_s(dx)}{R \exp s - 1} \leqslant \frac{1}{P(C)} < \infty.$$

Let us now consider the symmetrized measure $\tilde{\tilde{U}}$ obtained from U by averaging all $n!$ images of U under all possible permutations of coordinates of $\mathbf{R}^n$:

$$\tilde{\tilde{U}} = \frac{1}{n!} \sum_{g \in G_n} U g^{-1},$$

where G_n is the group of permutations of the n coordinates. Obviously, the averages $\tilde{\tilde{U}}_s$ of the conditional measures U_s coincide with the conditional measures for $\tilde{\tilde{U}}$. For $\tilde{\tilde{U}}$, just as for U, we have

$$\int_{\mathbf{R}^n} \exp x_1 \int_{-\ln R}^{\infty} \frac{\alpha(ds)\,\tilde{\tilde{U}}_s(dx)}{R\exp s - 1} = \int_{-\ln R}^{\infty} \frac{\alpha(ds)}{R\exp s - 1} \int_{\mathbf{R}^n} \exp x_1 \tilde{\tilde{U}}_s(dx) < \infty. \tag{83}$$

We now show that $\tilde{\tilde{U}}_s$ has first absolute moments, and its barycenter is the point $b_s = (s, \ldots, s)$. Indeed, from (83), in particular, it follows that

$$\int_{\mathbf{R}^n} \exp x_1 \tilde{\tilde{U}}_s(dx) < \infty$$

for almost all s (with respect to the measure $\alpha(ds)$), and consequently

$$\int_{x_1=0}^{\infty} x_1 \tilde{\tilde{U}}_s(dx) < \infty \qquad \text{and} \qquad \int_{x_1=s}^{\infty} (x_1 - s)\,\tilde{\tilde{U}}_s(dx) < \infty$$

(these are integrals over subsets of $\mathbf{R}^n$ of the forms $\{x\colon\ 0 \leqslant x_1 < \infty\}$ and $\{x\colon\ s \leqslant x_1 < \infty\}$). But U can be represented as the limit of its restrictions to the balls of radius $r \longrightarrow \infty$ and center at the point b_s. Each such restriction $\tilde{\tilde{U}}_s^{(r)}$ is a symmetric (invariant with respect to the group G_n) measure; therefore

$$\int_{s}^{\infty} (x_1 - s)\,\tilde{\tilde{U}}_s^{(r)}(dx) = \int_{-\infty}^{s} |x_1 - s|\,\tilde{\tilde{U}}_s^{(r)}(dx).$$

Passing to the limit with respect to r in this equation, we obtain

$$\int_{s}^{\infty} (x_1 - s)\,\tilde{\tilde{U}}_s(dx) = \int_{-\infty}^{s} |x_1 - s|\,\tilde{\tilde{U}}_s(dx),$$

and this equality describes the barycenter of $\tilde{\tilde{U}}_s$.

We now show that the measure $\tilde{\tilde{U}}$ itself has a left-sided absolute moment. Indeed,

$$\begin{aligned}
0 \leqslant \int_{x_1=-\infty}^{0} |x_1|\,\tilde{\tilde{U}}(dx) &= \int_{x_1=-\infty}^{0} |x_1| \int_{-\ln R}^{\infty} \alpha(ds)\,\tilde{\tilde{U}}_s(dx) \\
&= \int_{-\ln R}^{\infty} \alpha(ds) \int_{x_1=-\infty}^{0} |x_1|\,\tilde{\tilde{U}}_s(dx) \leqslant \int_{-\ln R}^{\infty} \alpha(ds) \int_{x_1=-\infty}^{s} |x_1 - s|\,\tilde{\tilde{U}}_s(dx) \\
&= \int_{-\ln R}^{\infty} \alpha(ds) \int_{x_1=s}^{\infty} (x_1 - s)\,\tilde{\tilde{U}}_s(dx) \leqslant \int_{-\ln R}^{\infty} \alpha(ds) \int_{x_1=s}^{\infty} \exp(x_1 - s)\,\tilde{\tilde{U}}_s(dx) \\
&\leqslant \int_{-\ln R}^{\infty} \alpha(ds) \int_{\mathbf{R}^n} \exp(x_1 - s)\,\tilde{\tilde{U}}_s(dx) \leqslant R \int_{-\ln R}^{\infty} \frac{\alpha(ds)}{R\exp s - 1} \int_{\mathbf{R}^n} \exp x_1 \tilde{\tilde{U}}_s(dx) < \infty
\end{aligned}$$

(by (83)).

If the marginal distributions of P_C and Q_C coincide, then the marginal distributions of U and V coincide; therefore the existence of the left-sided first moments of U (which follows from the existence of the left-sided first x_1-moment of $\widetilde{\widetilde{U}}$) and of the right-sided moments of V implies the existence of all first absolute moments and, consequently, the applicability of Proposition 77, i.e., $P_C = Q_C$ in this case.

But if $P(C_y) = 0$ for almost all y, then, as before, we get from Proposition 76 that for almost every y

$$\int_{-\ln R_y}^{\infty} \frac{\alpha_y (ds)\, U_{y,s}}{R_y \exp s - 1} = \tilde{P}_y = P_{C_y}\varphi^{-1},$$

where R_y is the value of the density of the distribution μ of the statistic f with respect to P relative to its distribution with respect to Q, at the point y. Since

$$\int_{-\infty}^{\infty} \tilde{P}_y \mu (dy) = \tilde{P}$$

and

$$\int_{\mathbf{R}^n} \exp x_k \tilde{P} (dx) = 1 \quad (k = 1, \ldots, n),$$

for μ-almost all values of y we have

$$\int_{\mathbf{R}^n} \exp x_k P_y (dx) < \infty$$

since

$$1 = \int_{\mathbf{R}^n} \exp x_k \int_{-\infty}^{\infty} P_y (dx)\, \mu (dy) = \int_{-\infty}^{\infty} \mu (dy) \int_{\mathbf{R}^n} \exp x_k P_y (dx),$$

from which, in quite the same way as before, we derive the existence of left-sided moments for U_y; consequently, by Proposition 77, we have proved that $P_C = Q_C$, i.e., the required sufficiency of the statistic f. ●

The above example of two distributions in $\mathbf{R}^3$ that are concentrated on different sides of the hyperplane given by the equation $s = 0$ and that have the same marginal distributions allows us to show that the condition

$$\frac{dQ_C}{dP_C} = \prod_{k=1}^{n} q_k (x_k)$$

does not guarantee the sufficiency of a marginally sufficient statistic f. For, if U and V are the measures in $\mathbf{R}^3$ with identical marginal distributions in the example on p. 155, then we consider in $\mathbf{R}^3$ the probability measures P and Q defined by

$$P = aU + bV, \quad Q = aeU + b\frac{1}{e}V, \quad \text{where } a = \frac{4}{3}\cdot\frac{1}{e+1}, \quad b = \frac{4}{3}\cdot\frac{e}{e+1}.$$

It is easy to see that

$$\frac{dQ}{dP} = \exp x_1 \cdot \exp x_2 \cdot \exp x_3$$

and the marginal distributions of P and Q coincide, i.e., the trivially insufficient statistic that is identically equal to a constant is marginally sufficient.

§13. Conditions for the existence of a one-to-one optimal plan in the problem of transport of mass in Minkowski spaces

1. We now consider a well-known problem (see [138] and [78]) connected with the so-called Monge problem on optimal transport: the problem of the existence of a one-to-one optimal plan of transport. In 1781 Gaspard Monge [80] studied the problem of the most rational transport of earth from an embankment into an excavation. The optimization problem arising in this way was formulated by Monge himself as follows: "Given two equivalent volumes, decompose them into infinitesimally small particles that correspond in such a way that the sum of the products of the paths traversed in carrying each pact to its correspondent by the volume of the part is a minimum" (quoted from [4], p. 1).* Monge obtained and stated, partly without proof, a number of important assertions about the character of an optimal plan of transport. In particular, he conjectured that for an optimal transport plan the directions of the displacements of the masses must form a vector field that is the gradient of some function $U(x)$ (more precisely, Monge suggested that the paths of the transported particles form the family of normals to some family of surfaces). Monge's conjecture was rigorously proved in 1884 by Appell (regarding this, see [14]); Dupin studied the same problem even before Appell.

In 1942 Kantorovič considered the problem on the most advantageous transport of masses given on an arbitrary compact metric space. Let (X, r) be a compact metric space, and μ and ν two Borel probability measures on it. Kantorovič proved [47], [48] that there is a most advantageous plan for the transport of the measure μ into the measure ν. This means that on the product $X \times X$ there exists a Borel probability measure m having μ and ν as marginal distributions (each such doubly stochastic measure is, by definition, a "plan of transport" of the measure μ into the measure ν) and such that in the class of all doubly stochastic measures m minimizes the "work" of transport, i.e., the magnitude of the integral

$$W(m) = \iint\limits_{X \times X} r(x, y)\, dm. \tag{84}$$

Kantorovič also proved that a necessary and sufficient condition for a doubly stochastic measure m on $X \times X$ to minimize (84) is the existence of a function $U(x)$ (a so-called potential) satisfying the conditions

---

*Editor's note. In Monge's own words the problem reads as follows (see [80], pp. 699–700): Étant donnés dans l'espace, deux volumes égaux entr'eux, & terminés chacune par une ou plusieurs surfaces courbes donnés; trouver dans le second volume le point où doit être transportée chaque molécule du premier, pour que la somme des produits des molecules multipliées chacune par l'espace parcouru soit un *minimum*!

$$1)\ U(x) - U(y) \leqslant r(x, y)\ \forall x \forall y \in X, \text{ i.e., } U(x) \in \mathrm{Lip}_1 1, \tag{85}$$

$$2)\ U(x) - U(y) = r(x, y) \text{ for } m\text{-almost all } (x, y) \in X \times X. \tag{86}$$

The "Kantorovič-Rubinšteĭn metric" arising in connection with this problem (see [5] and [52]) found numerous important applications. A very simple proof of Kantorovič's theorem on the potential, based on the duality theorem of linear programming, can be found in the survey article [138].

The statement of Kantorovič's problem clearly differs from the statement of Monge's in that Kantorovič's class of transport plans in which an extremum is sought is substantially broader than the class of plans of transport that Monge had in view and that correspond to one-to-one transport ("infinitesimally small particles that correspond"). Therefore, the problem of the existence of an optimal one-to-one plan in the Kantorovič sense, which is particularly nontrivial when the metric does not determine geodesics uniquely (for example, when X is a subset of a Banach space whose unit ball is not strictly convex), can be regarded as a bridge connecting the formulations of Monge and Kantorovič.(4)

It is not hard to give an example showing that even when the measures μ and ν are purely continuous and $X \subset \mathbf{R}^2$ with the Euclidean norm there may not exist an optimal plan of transport for which the corresponding doubly stochastic measure m is the kernel of an isomorphism of the measure spaces (X, μ) and (Y, ν) (to distinguish between the first and second copies of X we denote the space X, equipped with the measure ν, by the letter Y, and a typical element of it by the letter y). In other words, there may not exist an optimal one-to-one plan of transport, as is shown by the following example. Let $X = Y$ be the unit square in the plane $\mathbf{R}^2$, μ the one-dimensional Lebesgue measure on the segment $x_1 = 1/2$, and ν the measure concentrated on the segments $x_1 = 0$ and $x_1 = 1$ and proportional (with factor $1/2$) to the Lebesgue measures on these segments. It is not hard to see that the optimal plan of transport here is unique and consists of the doubly stochastic measure concentrated in $X \times Y = \{(x_1, x_2; y_1, y_2)\}$ on the segments $\{(1/2, t; 0, t), t \in [0, 1]\}$ and $\{(1/2, t; 1, t), t \in [0, 1]\}$. It is possible to modify this example somewhat and even obtain absolute continuity of ν with respect to Lebesgue measure.

2. Using our earlier results on conditions for the existence of an independent complement of a pair of given decompositions and the conditions for the existence of a doubly stochastic density, we can prove the existence of an optimal one-to-one plan when the compact set X is a subset of a finite-dimensional normed (not necessarily Euclidean) linear or affine space (a Minkowski space), and the measures μ and ν are absolutely continuous with respect to Lebesgue measure (on the possibility of weakening the last condition, see below). Of course, the concrete form of the compact set $X \subset (\mathbf{R}^n, \|\cdot\|)$ is not significant: it is sufficient to assume only that μ and ν have compact support (for convergence of the integral in (84)) or even only that the integrals $\int \|x\| d\mu$ and $\int \|y\| d\nu$ converge (conditions that are also sufficient for the validity of Kantorovič's theorem on the potential).

(4) M. I. Rvačev called the author's attention to this circumstance. The problem of the existence of one-to-one optimal plans in the Kantorovič sense was posed by Veršik (see [138]).

Henceforth let X and Y denote two copies of a finite-dimensional normed space, and μ and ν Borel probability measures that will serve as standards for the double stochasticity of a measure defined on the product $X \times Y$. A Borel decomposition of a linear or affine space (i.e., a decomposition into the "level sets" of a Borel mapping into a finite-dimensional space) is said to be locally affine if each element of this decomposition is a connected open subset of its affine span. If μ is a Borel measure on X defined by a density with respect to Lebesgue measure, and θ is a locally affine decomposition, then the conditional measures on the elements of θ are also absolutely continuous with respect to the Lebesgue measure on the affine subspaces spanned by the corresponding elements of θ.

PROPOSITION 78. *Let λ be a locally affine decomposition of the Lebesgue space (Q, m), where Q is a subset of a finite-dimensional affine space, and m is a measure proportional to the restriction of Lebesgue measure to Q. Then the conditional measures on the elements of λ are absolutely continuous with respect to Lebesgue measure.*

PROOF. We can limit ourselves to the case when the elements of λ are pairwise disjoint segments. Let $Q \subset \mathbf{R}^{n+1}$, and let $m^{(k)}$ denote k-dimensional Lebesgue measure.

Without loss of generality we can assume that the elements of the "ruled" decomposition λ are segments whose endpoints lie on two parallel hyperplanes L_0 and $L_1 \subset \mathbf{R}^{n+1}$, so that the set Q is located between them. The elements of λ determine a correspondence betwen the points of the set $M_0 = L_0 \cap Q$ and those of $M_1 = L_1 \cap Q$. Let this mapping $M_0 \longrightarrow M_1$ be denoted by T, let L_h be the hyperplane $(1-h)L_0 + hL_1$, and let h denote on each element of λ the relative distance of a point from L_0.

Let $m^{(n)}M_0 > 0$. We remark that if the mapping T is differentiable at almost every point of M_0, then the tangent mapping $T'(x)$ at almost every point $x \in M_0$ is a linear operator $\mathbf{R}^n \longrightarrow \mathbf{R}^n$ without large negative eigenvalues of odd multiplicity, while the conditional measures on each element of λ turn out to be absolutely continuous with respect to Lebesgue measure, and their densities $p_x(h)$ are polynomials in h of degree not greater than n: the density $p_x(h)$ is proportional to the quantity

$$\det\,(hT'(x)+(1-h)\,I),$$

whence the condition on the negative eigenvalues.

The uniform boundedness of the degrees of the polynomials giving the densities of the conditional measures suggests proving the polynomial character of the densities of the conditional measures of an arbitrary ruled decomposition by selecting a suitable smooth approximation of the decomposition λ. For our purposes, however, it suffices to prove a weaker assertion about the absolute continuity of the conditional measures.

Let $m^{(n)}M_0 > 0$, and let T: $M_0 \longrightarrow L_1$ be an arbitrary mapping that generates a ruled decomposition λ. It is assumed that the set-theoretic union Q of the segments $\{(1-h)x + hT(x),\ h \in [0, 1],\ x \in M_0\}$ is measurable. We show that $m^{(n+1)}Q > 0$ and that the function $m^{(n)}(L_h \cap Q) \circ [m^{(n+1)}Q]^{-1}$ is bounded as a function of $h \in [0, 1]$ by a constant depending only on n. By the arbitrariness of M_0, the analogous assertion is true also for the restriction of T to an arbitrary element of each decomposition in some sequence of measurable decompositions of M_0 that converges to the decomposition ϵ into points; hence the densities of the conditional measures on the elements of λ are bounded by the same constant.

We consider some sequence T_k of differentiable mappings generating ruled decompositions λ_k of the corresponding Lebesgue spaces (Q_k, m_k) and converging pointwise to T. By what was proved above, the function $m^{(n)}(L_h \cap Q_n)$ is a polynomial of degree not greater than n; therefore the densities of the measures m_k with respect to Lebesgue measure in $\mathbf{R}^{n+1}$ are uniformly bounded. Let $\{k_i\}$ be a subsequence such that the measures m_{k_i} converge weakly to a limit measure $\widetilde{m}$. From the uniform boundedness of the densities of the measures m_{k_i} it follows that $\widetilde{m}$ is absolutely continuous, and its density is bounded by the same constant ($\widetilde{m}$, generally speaking, is not proportional to the restriction of Lebesgue measure to Q). From the condition $\widetilde{m}Q = 1$ it follows that $m^{(n+1)}Q > 0$. For each k_i the function

$$m^{(n)}(L_h \cap Q_{k_i}) \cdot [m^{(n+1)}Q_{k_i}]^{-1},$$

as a function of h, is bounded by a constant depending only on n (and equal to the supremum on $[0, 1]$ of all the values of all the polynomials P of degree n for which $P(x) \geqslant 0$ for $x \in [0, 1]$ and $\int_0^1 P(x)dx = 1$), from which we get the boundedness of the values of the function $m^{(n)}(L_h \cap Q) \cdot [m^{(n+1)}Q]^{-1}$ by the same constant. ●

3. We first present briefly the basic idea of the proof of the existence of a one-to-one optimal plan. By the cited theorem of Kantorovič, there exists an optimal plan m, i.e., a doubly stochastic measure minimizing the quantity W in (84), and to it there corresponds a potential $U(x)$: a function on X satisfying a Lipshitz condition with exponent 1 and constant 1. Each doubly stochastic measure for which $U(x)$ is a potential is also an optimal plan of transport (the sufficiency in Kantorovič's theorem). Therefore, it suffices to construct the kernel of an isomorphism of the spaces (X, μ) and (Y, ν) on the closed set

$$V = V_U = \{(x, y) : U(x) - U(y) = \|x - y\|\}.$$

On V there exists at least one doubly stochastic measure: the measure m. If on V there exists a doubly stochastic density, then corresponding to it there is, by Theorem 8, an independent complement of the coordinate decompositions ξ and η, and each typical element of this decomposition, equipped with its conditional measure, can (by Proposition 42) serve as the kernel of the isomorphism of interest to us. However, such an extremal doubly stochastic density exists, generally speaking, very rarely (though it always exists in the one-dimensional case!): in fact, for almost all (with respect to such an extremal measure) points $(x, y) \in X \times Y$ the condition (86) holds, and therefore, if $A \subset X$ and $B \subset Y$ are subsets such that $\mu A > 0$ and $\nu B > 0$, and the measure $\mu_A \times \nu_B$ is absolutely continuous with respect to the extremal measure m, then for μ-almost all points $x \in A$ and ν-almost all points $y \in B$ (86) holds; but this is impossible if μ and ν are given by densities with respect to Lebesgue measure, and the unit sphere $\{x: \|x\| = 1\}$ is strictly convex or at least does not contain $(n-1)$-dimensional "faces".

If the potential function $U(x)$ is fixed, then for each typical point $x \in X$ we can consider the set of those points $y \in X$ for which $\|x - y\| = U(x) - U(y)$. For a strictly convex norm it is easily shown that all such points either coincide with x or lie on some ray going out from x, and that they form a set into which it is possible to transport mass from x for the given potential function U. If it is possible to carry a mass along several directions from the point x, then each such admissible direction is regarded as a generator of its segment or ray. Since two segments can have as a common point only an endpoint of one of them (they cannot cross), we get a Borel decomposition θ of the whole space X into open segments and points.

We regard θ as a measurable decomposition of the spaces (X, μ) and (Y, ν). An optimal plan m for transport of the measure μ into ν is constructed in such a way that displacements of mass take place only within the confines of the elements of θ. On the union of the elements consisting of a single point each plan of transport m is already one-to-one in a trivial way. For each of the segments I making up θ, as can be verified on the set $V \cap (I \times I)$, which is half of the square separated by the diagonal, it is possible to define, by Theorem 10, a measure that is doubly stochastic with respect to the conditional measures μ_I and ν_I and absolutely continuous with respect to the product $\mu_I \times \nu_I$; and then, by Theorem 8*, there exists an independent complement of the coordinate decompositions with respect to the measure "glued together" from the doubly stochastic (with respect to μ_I and ν_I) measures just constructed. A typical element of this independent complement, together with the union of the "set theoretic squares" of singleton subsets in θ generates the graph of the isomorphism of interest to us, which lies in V and, therefore, corresponds to an optimal plan of transport.

The situation is somewhat more complicated when the norm $\|\cdot\|$ on X is generated by a unit ball that is not strictly convex, though the basic idea of the argument is the same. Here we must construct, with respect to the set V, certain subdecompositions θ_X and θ_Y of θ that are determined by the minimization mi V of V with respect to the pair (μ, ν) and on each of the products of corresponding elements of which (as above) there exists a density that is doubly stochastic with respect to the pair of conditional measures; this allows us again to use Theorem 8* and prove the existence of a kernel of an isomorphism.

In the product $\mathbf{R}^n \times \mathbf{R}^n$ consider a closed convex body Q and a Borel probability measure m for which $mQ = 1$; on this set Q there does not necessarily exist a measure that is the kernel of an isomorphism of the marginal distributions of m.

For example, if Q is the triangle ABC in Figure 10, and the measure m is concentrated on the segments AB and AC, then the minimization mi Q of ABC with respect to the corresponding marginal distributions consists only of the segments AB and AC, from which it follows that it is impossible to give the kernel of an isomorphism on this triangle. On the other hand, it is easy to see that for any nonatomic measure m on the triangle ABC (Figure 11) there exists a kernel of an isomorphism of the marginal distributions of this measure that is concentrated on the same triangle.

We describe a quite broad class of convex subsets in the product space $X \times Y \equiv$

$\mathbf{R}^n \times \mathbf{R}^n$ having an analogous property. Although the assertion itself is not used in the following, its proof brings the reader closer to the more tedious proof of the corresponding step in Theorem 12.

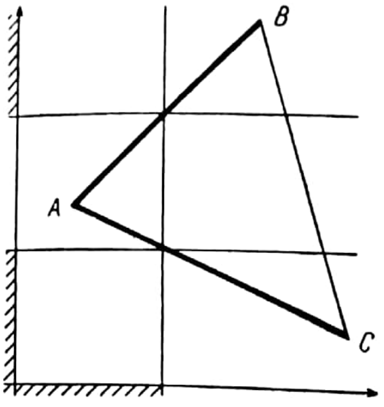

FIGURE 10

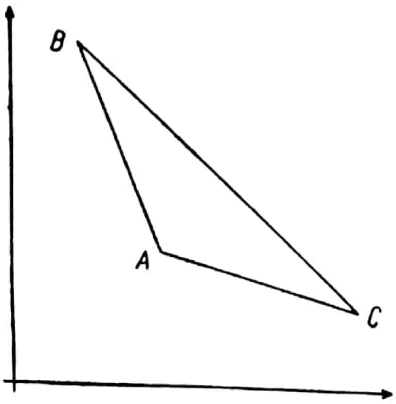

FIGURE 11

PROPOSITION 79. *Let $Q \subset X \times Y \equiv \mathbf{R}^n \times \mathbf{R}^n$ be a convex compact subset having the property that if $A \subset X$ and $B \subset Y$ are Borel subsets such that*

$$A \subset \pi_X Q, \quad B \subset \pi_Y Q, \quad (A \times B) \cap Q = \varnothing,$$

then also

$$(\operatorname{conv} A \times \operatorname{conv} B) \cap Q = \varnothing.$$

Let m be a Borel probability measure on Q whose marginal distributions are absolutely continuous with respect to the Lebesgue measures on X and Y. Then there exists a probability measure m_1 on Q with the same marginal distributions as m and such that its conditional measures on the elements of the decomposition $\xi \wedge \eta$ (where ξ and η are the coordinate decompositions generated by the canonical projections onto X and Y) are absolutely continuous with respect to the products of the conditional measures on those elements of the decompositions ξ_η and η_ξ that are carried, under the canonical mappings $Q/\xi \longrightarrow Q/(\xi \wedge \eta)$ and $Q/\eta \longrightarrow Q/(\xi \wedge \eta)$, into the element of $\xi \wedge \eta$ under consideration (ξ_η and η_ξ are the measurable decompositions of X and Y induced by these mappings).

PROOF. We consider the process of constructing the minimization of Q. It was shown (Proposition 69) that the minimization mi Q of Q is obtained by set-theoretic deletion from Q of a countable family of sets

$$(X \setminus A_n) \times (Y \setminus B_n), \quad n = 1, \ldots,$$

such that

$$(A_n \times B_n) \cap Q = \varnothing \tag{87}$$

and

$$\mu A_n + \nu B_n = 1. \tag{88}$$

Obviously, since the sets A_n and B_n are determined to within μ- and ν-equivalence, respectively, in the construction of the minimization mi Q it is possible to replace them by equivalent ones; in particular, to assume that $A_n \subset \pi_X Q$ and $B_n \subset \pi_Y Q$. By assumption, (88) implies the condition $(\operatorname{conv} A_n \times \operatorname{conv} B_n) \cap Q = \varnothing$. There is an affine hyperplane separating the disjoint convex set Q and $\operatorname{conv} A_n \times \operatorname{conv} B_n$, i.e., a linear functional F_n on the space $X \times Y$ such that

$$F_n(Q) \cap F_n(\operatorname{conv} A_n \times \operatorname{conv} B_n) = \varnothing.$$

Let $F_n^X(x) = F_n((x, 0))$ and $F_n^Y(y) = F_n((0, y))$. Then $F_n((x, y)) = F_n^X(x) + F_n^Y(y)$. For definiteness let (the meaning of the notation is obvious)

$$F_n(Q) \geqslant c_n > F_n(A_n \times B_n)$$

and

$$\alpha_n' = \sup_{x \in A_n} F_n^X(x), \quad \alpha_n'' = \sup_{y \in B_n} F_n^Y(y),$$

so that $\alpha_n' + \alpha_n'' < c_n$. Then for any pair of numbers β_n' and β_n'' such that $\beta_n' \geqslant \alpha_n'$, $\beta_n'' \geqslant \alpha_n''$ and $\beta_n' + \beta_n'' = c_n$ we have the condition

$$F_n(Q) \geqslant c_n > F_n(A_n^+ \times B_n^+),$$

where $A_n^+ = \{x\colon\ x \in X, F_n^X(x) < \beta_n'\}$ and $B_n^+ = \{y\colon\ y \in Y, F_n^Y(y) < \beta_n''\}$. From the fact that

$$(A_n^+ \times B_n^+) \cap Q = \varnothing,$$

it follows, by Proposition 49, that $\mu A_n^+ + \nu B_n^+ \leqslant 1$, and from this, the obvious inclusions $A_n^+ \supset A_n$ and $B_n^+ \supset B_n$, and (88) it follows that $\mu A_n^+ = \mu A_n$ and $\nu B_n^+ = \nu B_n$; i.e., the half-space A_n^+ is μ-equivalent to A_n and the half-space B_n^+ is ν-equivalent to B_n. Therefore, the subset $(X\backslash A_n) \times (Y\backslash B_n)$ "subtracted" from Q can be replaced by the product of half-spaces $(X\backslash A_n^+) \times (Y\backslash B_n^+)$. Next, the subtraction of $(X\backslash A_n^+) \times (Y\backslash B_n^+)$ can be replaced by the intersection with the union of the two convex sets $(X\backslash A_n^+) \times B_n^+$ and $A_n^+ \times (Y\backslash B_n^+)$. The sets A_n^+ and B_n^+ form bases of certain decompositions θ_X and θ_Y of the spaces X and Y, and these decompositions are locally affine, because A_n^+ and B_n^+ are half-spaces, hence convex, together with their complements, and any intersection of sets in the basis or sets complementary to basis sets is convex; consequently, the elements of the decompositions are convex sets. By means of the construction of the minimization mi Q we have established a canonical one-to-one correspondence between the elements of the decompositions θ_X and θ_Y that induces an isomorphism of the spaces X/θ_X and Y/θ_Y. A product of corresponding elements is a convex set on which the conditional measure (of the measure with respect to the decomposition $\pi_X^{-1}\theta_X \wedge \pi_Y^{-1}\theta_Y$) is doubly stochastic with respect to the conditional measures on the corresponding elements of θ_X and θ_Y and is concentrated on the intersection of Q with the element under consideration, i.e., also with a convex set.

We prove that the dimension of the intersection of Q with this element is equal to the dimension of the element. The assumption that the dimension of the part of Q

in some element of the decomposition is less than the dimension of the element itself contradicts the fact that θ_X and θ_Y are decompositions having the property that if $(A \times B) \cap Q = \varnothing$ and $\mu A + \nu B = 1$, then A is measurable with respect to θ_X, and B is measurable with respect to θ_Y (moreover, θ_X and θ_Y are the coarsest decompositions having this property). Indeed, for any doubly stochastic measure on an affine subspace of the product of two linear spaces it is clear that there are nontrivial sets A and B for which $\mu A + \nu B = 1$ and $A \times B$ does not intersect this subspace. Therefore, our assumption leads to the existence of sets with impossible properties and is thereby rejected. If we now consider some two corresponding elements of θ_X and θ_Y, which we regard as two affine subspaces, equipped with the corresponding conditional measures, then the trace of Q on the direct product of these spaces is a convex set of maximal dimension whose minimization with respect to the relevant conditional measures coincides with the set itself. Hence, on the trace of Q there is a doubly stochastic measure that is absolutely continuous with respect to the product of the marginal distributions, and, consequently, there is a doubly stochastic measure on Q such that for it $\pi_X^{-1}\theta_X = \pi_Y^{-1}\theta_Y$ (this condition is necessarily satisfied for any doubly stochastic measure), and on each element of the decomposition $\pi_X^{-1}\theta_X \wedge \pi_Y^{-1}\theta_Y$ $(= \pi_X^{-1}\theta_X = \pi_Y^{-1}\theta_Y)$ its conditional measure is absolutely continuous with respect to the product of the conditional measures on the corresponding elements of θ_X and θ_Y (Theorem 10 and Remark 2 after it).

The decomposition $\xi \wedge \eta$, considered for the measure m, is not coarser than $\pi_X^{-1}\theta_X \wedge \pi_Y^{-1}\theta_Y$. Since, for each measure on a product space that is absolutely continuous with respect to the product of its marginal distributions, the conditional measures on the elements of the infimum of the two coordinate decompositions have the same property, Proposition 79 is proved. ●

4. We now proceed to the proof of the basic result of this section, on the existence of an optimal one-to-one plan of transport.

THEOREM 12. *Suppose that on a bounded subset of a finite-dimensional Banach space* $(\mathbf{R}^n, \|\cdot\|)$ *two Borel probability measures* μ *and* ν *are given, each absolutely continuous with respect to Lebesgue measure on this space. Then there is an optimal one-to-one plan of transport of the measure* μ *into the measure* ν.

In other words, on the product $X \times Y$ of two copies of the Banach space there exists a kernel m_0 of an isomorphism of the spaces (X, μ) and (Y, ν) that supplies a minimum, in the class of all doubly stochastic measures, for the integral

$$W(m) = \int_{X \times Y} \|x - y\| dm. \tag{89}$$

PROOF. By the theorem of Kantorovič cited above, there exists at least one doubly stochastic measure m on $X \times Y$ that gives the integral in (89) a minimum. Moreover, there is a function $U(x)$ (called a potential) on the space such that

$$1)\ |U(x) - U(y)| \leqslant \|x - y\| \quad \forall x \forall y \in X, \tag{90}$$

$$2)\ U(x) - U(y) = \|x - y\| \quad \text{for } m\text{-almost all } (x, y) \in X \times Y. \tag{91}$$

To prove the theorem it suffices to show that there exists a measure m_0 that is the kernel of an isomorphism of (X, μ) and (Y, ν) and is concentrated on the set

$$V = V_U = \{(x, y) : U(x) - U(y) = \|x - y\|\}.$$

We consider the function $U(x)$. Let $\mathfrak{X}_1$ be the class of subsets $A \subset X$ having the property that A is a maximal (with respect to inclusion) set on which the function U is linear (more precisely, affine), with gradient in the given norm equal to 1. For each set A in the class $\mathfrak{X}_1$ we consider the maximal (with respect to inclusion) connected subsets of A that are open in A. These latter sets determine a locally affine decomposition into the subsets of the class $\mathfrak{X}$.

Another (equivalent) description of the sets in the class $\mathfrak{X}$ consists of the following. Let $x \in X$ be an arbitrary point, and let x^S be a largest (with respect to inclusion) affine subset, necessarily convex, of the sphere S, or one such subset, such that for some point $y \in X$ we have $U(x) - U(y) = \|x - y\|$ and

$$\frac{x - y}{\|x - y\|} \in x^S. \tag{92}$$

We consider the set of all those y for which (92) holds for a fixed set x^S. Then we consider the set of those points x_1 for which $(x_1 - y)/\|x_1 - y\| \in x^S$ for some of the indicated y's. Next, we consider the set of those y_1 for which

$$\frac{x_1 - y_1}{\|x_1 - y_1\|} \in x^S$$

for some of the x_1's. Repeating this procedure, taking the union of all the points appearing at some stage, and then passing to the closure, we obtain a set A in the class $\mathfrak{X}$.

Finally, we give one more description of the class $\mathfrak{X}$, which is, in essence, our working definition. In $\mathbf{R}^n \times \mathbf{R}$ we consider the graph of $U(x)$. By the Lipschitz property (90), for each point $(x, U(x)) \in \mathbf{R}^n \times \mathbf{R}$ the cone

$$K_x^+ = \{(x', z) \in \mathbf{R}^n \times \mathbf{R} : z \geqslant U(x) + \|x' - x\|\}$$

intersects the graph of U only on the surface of this cone, and the same is true for the cone

$$K_x^- = \{(x', z) \in \mathbf{R}^n \times \mathbf{R} : z \leqslant U(x) - \|x' - x\|\}.$$

For an arbitrary point x, if K_x^+ and K_x^- intersect the graph only at x, we put the singleton set $\{x\}$ in the class $\mathfrak{X}$. But if the intersection of the cone $K_x = K_x^+ \cup K_x^-$ with the graph of U contains more than one point, we consider the affine span of the set of all such points and the intersection of this affine subspace with the graph. We select the maximal connected relatively open subsets of this intersection and put the sets in X that lie "under" these subsets of the graph in the class $\mathfrak{X}$.

As follows from the Lipschitz condition satisfied by U, each set in the class $\mathfrak{X}$ is such that its boundary consists of two parts: the points $x \in X$ for which only the cone K_x^+ has nontrivial intersection with the graph over this set, and the points $x \in X$ for which only the cone K_x^- intersects the graph of $U(x)$ in more than just the vertex.

Both of these parts of the boundary also satisfy a certain Lipschitz condition; therefore the Lebesgue measure (of corresponding dimension) of the boundary is a fortiori equal to zero, and we can speak of a measurable decomposition.

We now turn our attention to the fact that (91) cannot be satisfied if the points x and y do not belong to some single subset of the family $\mathfrak{X}$. This means that there is transport of mass only within the confines of each subset of the decomposition $\mathfrak{X}$, i.e., that the measures $\mu/\mathfrak{X}$ and $\nu/\mathfrak{X}$ coincide. However, even within the confines of one subset in $\mathfrak{X}$ we do not have for arbitrary points x and y, generally speaking, the condition

$$|U(x) - U(y)| = \|x - y\|,$$

i.e., the admissibility of transport of an element of mass from one point into the other. Nevertheless, our construction allows us now to consider separately the problem of the existence of a one-to-one optimal plan of transport for the pair of conditional measures on each element of the decomposition $\mathfrak{X}$; and on each such set in the family $\mathfrak{X}$ the restriction of U is simply affine, by the construction of $\mathfrak{X}$. When we are considering a conditional measure, everything that happens outside the particular element of $\mathfrak{X}$ is insignificant for us, so we can assume that $U(x)$ is equal to a linear functional $L(x)$, since this holds almost everywhere. This now helps us to prove the existence of an optimal doubly stochastic measure m on $X \times Y$ such that on elements of the decomposition $\xi \wedge \eta$ (where ξ and η, as usual, are the coordinate decompositions) the conditional measures are absolutely continuous with respect to the products of the corresponding conditional measures on the elements of the decompositions ξ_η and η_ξ, and these conditional measures are either purely continuous or are δ-measures; Theorem 8* then leads us at once to the goal.

Thus, in the space $X \times Y \equiv \mathbf{R}^m \times \mathbf{R}^m$, $0 \leqslant m \leqslant n$, we consider the set

$$Q = \{(x, y) : L(x - y) = \|(x - y\|\},$$

where $L(x)$ is a linear functional. By assumption, on Q we are given a measure m whose marginal distributions μ and ν are absolutely continuous with respect to the Lebesgue measures on X and Y. We now prove that on Q there exists another measure m_1 with the same marginal distributions and with the following property: there exist measurable decompositions $\theta_X = \theta_Y$ of the space X (and, simultaneously, of its second copy, the space Y) such that the conditional measures corresponding to the measures μ and ν on elements of θ_X and θ_Y are purely continuous or are δ-measures; moreover, if π_X and π_Y are, as usual, the canonical projections in $(X \times Y, m_1)$, then the decomposition $\pi_X^{-1}\theta_X$ and $\pi_Y^{-1}\theta_Y$ coincide, and on each element of the decomposition $\pi_X^{-1}\theta_X \wedge \pi_Y^{-1}\theta_Y$ $(= \pi_X^{-1}\theta_X = \pi_Y^{-1}\theta_Y)$ the conditional measure corresponding to m_1 is absolutely continuous with respect to the product of the conditional measures on the corresponding elements of θ_X and θ_Y. With this aim, we begin to construct the minimization mi Q of Q. Let $\{(A_n, B_n), n = 1, \dots\}$ be a sequence of pairs of subsets $A_n \subset X$, $B_n \subset Y$ such that

$$(A_n \times B_n) \cap Q = \varnothing \text{ and } \operatorname{mi} Q = Q \setminus \bigcup_{n=1}^{\infty} ((X \setminus A_n) \times (Y \setminus B_n)).$$

Unfortunately, the present situation is not exactly the same as under the conditions of Proposition 79, because it is impossible to pass to convex hulls and from them to half-spaces; however, with certain complications an argument analogous to the one used in the proof of Proposition 79 works in this case.

We introduce the notation

$$C^+ = \{x \in X : L(x) = \|x\|\}, \quad C^- = \{x \in X : L(x) = -\|x\|\}.$$

The sets C^+ and C^- are convex closed cones in X. If, as assumed, only those doubly stochastic measures concentrated on Q are admissible, then only those transports are admissible for which an element of mass is moved from a point $x \in X$ without going outside the set $x + C^+$. If $(A \times B) \cap V = \varnothing$, then transport from A into B is prohibited. The set of all points to which mass can be moved from A is the set $A + C^+$, and the set of those points from which it is possible to transport mass into B is the set $B + C^-$. The condition $(A \times B) \cap V = \varnothing$ is thus equivalent to the condition

$$(A + C^+) \cap (B + C^-) = \varnothing \ (\operatorname{mod} \mu \text{ and } \operatorname{mod} \nu) \tag{93}$$

(here we regard both A and B as subsets of the same space X), or, what is the same, to the condition

$$((A + C^+) \times (B + C^-)) \cap Q = \varnothing. \tag{94}$$

We can assume that $A = A + C^+$ and $B = B + C^-$. Indeed, if $\mu A + \nu B = 1$, then, by Proposition 54 and (94), $\mu(A + C^+) = \mu A$ and $\nu(B + C^-) = \nu B$, i.e., the sets A and $A + C^+$, and also B and $B + C^-$, are respectively μ- and ν-equivalent. Further, it is clear that $\mu A = \nu A$ and $\mu B = \nu B$. It is thus possible to assume that each pair of sets (A_n, B_n) appearing in the minimization has the property that $A_n = A_n + C^+$ and $B_n = B_n + C^-$, and, moreover, $A_n \cup B_n = X$ (mod μ and mod ν).

Regarding the sets $\{A_n\}$ as a basis of a decomposition θ_X and the sets $\{B_n\}$ as a basis of a decomposition θ_Y, we find that $\theta_X = \theta_Y$ when X and Y are identified) and that each element of θ_X is obtained as the intersection of a countable sequence of sets of a special form. We now proceed to the study of these sets.

We call A a C^+-Lipschitz set if $A = A + C^+$; the set B is C^--Lipschitz if $B = B + C^-$, and F is C-Lipschitz if $F = A \cap B$, where A and B are, respectively, C^+- and C^--Lipschitz.

As shown, we can assume that each element of $\theta_X = \theta_Y$ is C-Lipschitz. We give another example. As is easy to see, each of the sets in the class $\mathfrak{X}$ mentioned at the beginning of the proof is a C-Lipschitz set, where C is the cone generated by some "face" A^S of the unit sphere S (some set that is the intersection of S with a supporting hyperplane). The sets that are elements of θ_X, however, do not simply belong to the class of C-Lipschitz sets. By construction,

$$Q \setminus \bigcup ((X \setminus A_n) \times (Y \setminus B_n)) = \operatorname{mi} Q;$$

therefore, for both conditional marginal measures on each element of θ_X $(= \theta_Y)$ the set Q already coincides with its minimization. The requirement that there exist on each element of θ_X $(= \theta_Y)$ two measures for which $Q = \mathrm{mi}\, Q$ restricts even more the class of subsets of X that can be elements of $\theta_X = \theta_Y$.

We list once more the properties that must be satisfied by a set $A \subset X$ that is an element of θ_X.

1. If $x_1, x_2 \in A$, then $((x_1 + C^+) \cap (x_2 + C^-)) \subset A$.
2. For each point $x \in A$ the set A is contained in the closure of the union of the sets $A_{x,k}$, where

$$A_{x,1} = (x + C^+) \cap A, \quad A_{x,2} = (A_{x,1} + C^-) \cap A, \ \dots,$$
$$A_{x,2k+1} = (A_{x,2k} + C^+) \cap A, \quad A_{x,2k+2} = (A_{x,2k+1} + C^-) \cap A, \ \dots .$$

From properties 1 and 2 it follows that θ_X is a locally affine decomposition (and the affine span of each element of this decompostion is parallel to some face of the convex cone C). From the absolute continuity of μ and ν with respect to Lebesgue measure it follows that the corresponding conditional measures on the elements of the locally affine decomposition θ_X are also absolutely continuous with respect to the Lebesgue measures on their affine spans; in particular, they are either purely continuous or are δ-measures. From the maximality (which has already been used) of the decomposition it follows that the dimension of the intersection of Q with the corresponding element of $\pi_X^{-1}\theta_X \wedge \pi_Y^{-1}\theta_Y$ is maximal (otherwise the decomposition θ_X could be refined; cf. the proof of Proposition 79). Thus, with respect to each pair of corresponding conditional measures on corresponding elements of the decomposition θ_X of (X, μ) and θ_Y of (Y, ν) the set Q is not minimizable and has maximal dimension. By Theorem 10 and the accompanying Remark 2, it follows from this that there exists on Q a doubly stochastic measure m_1 whose conditional measures on the elements of $\pi_X^{-1}\theta_X \wedge \pi_Y^{-1}\theta_Y$ are absolutely continuous with respect to the product of the corresponding conditional measures. And, by exactly the same arguments, there exists also a measure on the set V_U whose conditional distributions on the elements of $\xi \wedge \eta$ are absolutely continuous with respect to the product of the corresponding conditional measures on the corresponding elements of the decompositions ξ_η and η_ξ; moreover, these latter conditional measures are also either purely continuous or are (both) δ-measures.

If we now decompose the space X $(= Y)$ into two subsets, the first of which is the union of those elements of the decomposition ξ_η on which the conditional distributions of the measure μ (and ν) are purely continuous, and the second of which is the union of the singleton elements of ξ_η (and η_ξ), then, using Theorem 8* for the first subset and using trivial arguments for the second one, we get the existence of the desired measure m_0 that is the kernel of an isomorphism of the marginal measures μ and ν and is concentrated on the set V_U, thereby giving a minimum for the expression for $W(m)$ in (89). ●

REMARK. It is actually possible to weaken somewhat the requirement of absolute continuity of the marginal distributions μ and ν. Indeed, our hypothesis of absolute continuity for the measures μ and ν with respect to Lebesgue measure was used only to ensure the nonatomicity of the conditional measures on the elements of the locally affine decomposition that have nonzero dimension. This property is also enjoyed by measures μ such that at any point $x \in X$ the μ-measure of the ball of radius ϵ with center at x is an infinitesimally small quantity of higher order than ϵ^{n-1}, where $n =$ dim X. The above example of a situation in which there is no one-to-one plan among the optimal ones, and analogous examples for any dimension $n \geqslant 2$, show that these conditions are not improvable.

BIBLIOGRAPHY

[1] N. I. Ahiezer, *The classical moment problem and some related topics in analysis*, Fizmatgiz, Moscow, 1961; English transl., Hafner, New York, 1965.

[2] A. D. Aleksandrov, *Convex polyhedra*, GITTL, Moscow, 1950; German transl., Akademie-Verlag, Berlin, 1958.

[3] T. W. Anderson, *An introduction to multivariate statistical analysis*, Wiley, New York, 1958.

[4] Paul Appell, *Le problème géométrique des déblais et remblais*, Mémor. Sci. Math, Fasc. 27, Gauthier-Villars, Paris, 1928.

[5] V. I. Arkin and V. L. Levin, *Extreme points of a certain set of measurable vector-valued functions of several variables and convexity of the ranges of vector integrals*, Dokl. Akad. Nauk SSSR 199 (1971), 1223–1226 = Soviet Math. Dokl. 12 (1971), 1248–1252.

[6] Stefan Banach, *Théorie des opérations linéaires*, PWN, Warsaw, 1932; reprint, Chelsea, New York, 1955.

[7] Jean-René Barra, *Notions fondamentales de statistique mathématique*, Dunod, Paris, 1971.

[8] Ju. K. Beljaev, *Local properties of sample functions of stationary Gaussian processes*, Teor. Verojatnost. i Primenen. **5** (1960), 128–131 = Theor. Probability Appl. **5** (1960), 117–120.

[9] ———, *Unboundedness of the sample functions of Gaussian processes*, Teor. Verojatnost. i Primenen. **3** (1958), 351–354 = Theor. Probability Appl. **3** (1958), 327–329.

[10] Garrett Birkhoff, *Lattice theory*, rev. ed., Amer. Math. Soc., Providence, R. I., 1948.

[11] ———, *Three observations on linear algebra*, Univ. Nac. Tucumán Rev. Ser. A**5** (1946), 147–151.

[12] N. Bourbaki, *Espaces vectoriels topologiques*, Chaps. 1–5 and Fascicule de résultats, Actualitiés Sci. Indust., nos. 1189, 1229, 1230, Hermann, Paris, 1953, 1955.

[13] ———, *Intégration*, Chaps. 1–5, Actualitiés Sci. Indust., nos. 1175, 1244, Hermann, Paris, 1965, 1967.

[14] ———, *Théorie des ensembles*, Chaps. 1–4 and Fascicule de résultats, Actualités Sci. Indust., nos. 1141, 1212, 1243, 1258, Hermann, Paris, 1960, 1963, 1957; English transl., Hermann, Paris; Addison-Wesley, Reading, Mass., 1968.

[15] ———, *Intégration*. Chap. 9, Actualitiés Sci. Indust., no. 1340, Hermann, Paris, 1969.

[16] James R. Brown, *Doubly stochastic measures and Markov operators*, Michigan Math. J. **12** (1965), 367–375.

[17] ———, *Approximation theorems for Markov operators*, Pacific J. Math. **16** (1966), 13–23.

[18] Herbert Busemann, *Convex surfaces*, Interscience, New York, 1958.

[19] Simone Chevet, *G. B. ensembles, G. C. ensembles, épaisseur mixte et ϵ-entropie*, Séminaire A. Badrikian et S. Chevet, Dept. Math. Appl. Univ. Clermont-Ferrand, 1971/72; published as Exposé 8 in A. Badrikian and S. Chevet, *Mesures cylindriques, espaces de Wiener et fonctions aléatoires Gaussiennes*, Lecture Notes in Math., Vol. 379, Springer-Verlag, Berlin and New York, 1974.

[20] Harald Cramér, *On the theory of stationary random processes*, Ann. of Math. (2) **41** (1940), 215–230.

[21] ———, *Mathematical methods of statistics*, Princeton Univ. Press, Princeton, N. J., 1946.

[22] Giorgio Dall'Aglio, *Sulle distribuzioni doppie con margini assegnati soggete a delle limitazioni*, Giorn. Ist. Ital. Attuari **24** (1961), 94–108.

[23] R. L. Dobrušin, *Properties of the sample functions of a Gaussian stationary process*, Teor. Verojatnost. i Primenen. **5** (1960), 132–134 = Theor. Probability Appl. **5** (1960), 120–122.

[24] J. L. Doob, *Stochastic processes*, John Wiley, New York; Chapman & Hall, London, 1953.

[25] R. M. Dudley, *The sizes of compact subsets of Hilbert space and continuity of Gaussian processes*, J. Functional Analysis 1(1967), 290–330.

[26] ———, *Sample functions of the Gaussian process*, Ann. Probability 1 (1973), 66–103.

[27] Herbert Bristol Dwight, *Tables of integrals and other mathematical data*, Macmillan, New York, 1934.

[28] A. Dynin and B. Mitjagin, *Criterion for nuclearity in terms of approximative dimension*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **8** (1960), 535–540.

[29] Xavier Fernique, *Continuité des processus gaussiens*, C. R. Acad. Sci. Paris **258** (1964), 6058–6060.

[30] ———, *Intégrabilité des vectors gaussiens*, C. R. Acad. Sci. Sér. A **270** (1970), 1698–1699.

[31] ———, *Régularité de processus gaussiens*, Invent. Math. **12** (1971), 304–320.

[32] ———, *Des résultats nouveaux sur les processus gaussiens*, C. R. Acad. Sci. Sér. A **278** (1974), 363–365.

[33] ———, *Minorations des fonctions aléatoires gaussiennes*, Ann. Inst. Fourier (Grenoble) **24** (1974), fasc. 2, 61–66.

[34] I. M. Gel'fand, *Generalized random processes*, Dokl. Akad. Nauk SSSR **100** (1955), 853–856. (Russian)

[35] I. M. Gel'fand and N. Ja. Vilenkin, *Generalized functions*. Vol. 4: *Applications of harmonic analysis*, Fizmatgiz, Moscow, 1961; English transl., Academic Press, New York, 1964.

[36] J. K. Ghosh, *Marginal sufficiency implies sufficiency*, Indian Statist. Inst. Tech. Rep. no. 54/68, Calcutta, 1968.

[37] B. V. Gnedenko, *A course in the theory of probability*, 3rd ed., Fizmatgiz, Moscow, 1961; English transl., Chelsea, New York, 1962.

[38] Paul R. Halmos, *Measure theory*, Van Nostrand, Princeton, N. J., 1950.

[39] Felix Hausdorff, *Mengenlehre*, 3rd ed., de Gruyter, Berlin, 1937; reprint, Dover, New York, 1944; English transl., Chelsea, New York, 1957; 2nd ed., 1962.

[40] G. A. Hunt, *Random Fourier transforms*, Trans. Amer. Math. Soc. **71** (1951), 38–69.

[41] I. A. Ibragimov, *Distribution of the maximum of independent Gaussian variables*, Manuscript deposited at the Leningrad Branch, Steklov Inst. Math. Acad. Sci. USSR, Leningrad, 1972. (Russian)

[42] J. R. Isbell, *Birkhoff's problem* 111, Proc. Amer. Math. Soc. **6** (1955), 217–218.

[43] Kiyosi Itô, *Stationary random distributions*, Mem. Coll. Sci. Univ. Kyoto Ser. A Math. **28** (1954), 209–223.

[44] N. C. Jain and M. B. Marcus, *Sufficient conditions for the continuity of stationary Gaussian processes and applications to random series of functions*, Ann. Inst. Fourier (Grenoble) **24** (1974), fasc. 2, 117–141.

[45] A. M. Kagan and O. V. Šalaevskiĭ, *The Behrens-Fisher problem: the existence of similar regions in the algebra of sufficient statistics*, Dokl. Akad. Nauk SSSR **155** (1964), 1250–1252 = Soviet Math. Dokl. **5** (1964), 556–558.

[46] J. P. Kahane, *Propriétés locales des fonctions à séries de Fourier aléatoires*, Studia Math. **19** (1960), 1–25.

[47] L. V. Kantorovič, *On the translocation of masses*, C. R. (Dokl.) Acad. Sci. URSS **37** (1942), 199–201.

[48] ———, *On a problem of Monge*, Uspehi Mat. Nauk **3** (1948), no. 2 (24), 225–226. (Russian)

[49] L. V. Kantorovič and G. P. Akilov, *Functional analysis in normed spaces*, Fizmatgiz, Moscow, 1959; English transl., Macmillan, New York, 1964.

[50] L. V. Kantorovič and I. V. Romanovskiĭ, *The structure of depreciation assignments for a steady-state load in a machine yard*, Dokl. Akad. Nauk SSSR **166** (1966), 309–312; English transl. in Soviet Phys. Dokl. 11 (1966/67).

[51] L. V. Kantorovič and G. Š. Rubinšteĭn, *On a functional space and some extremal problems*, Dokl. Akad. Nauk SSSR **114** (1957), 1058–1061. (Russian)

[52] ———, *On a space of completely additive functions*, Vestnik Leningrad. Univ. Ser. Mat. Meh. Astronom. **13** (1958), no. 7, 52–59. (Russian)

[53] L. V. Kantorovič, B. Z. Vulih, and A. G. Pinsker, *Functional analysis in partially ordered spaces*, GITTL, Moscow, 1950. (Russian)

[54] Kari Karhunen, *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*, Ann. Acad. Sci. Fenn. Ser. A. I. Math. Phys. no. 37 (1947).

[55] Hans G. Kellerer, *Masstheoretische Marginalprobleme*, Math. Ann. **153** (1964), 168–198.

[56] ———, *Funktionen auf Produkträumen mit vorgegebenen Marginal-Funktionen*, Math. Ann. **144** (1961), 323–344.

[57] ———, *Die Schnittmass-Funktionen messbarer Teilmengen eines Produktraumes*, Math. Ann. **146** (1962), 103–128.

[58] David G. Kendall, *On infinite doubly stochastic matrices and Birkhoff's problem* 111, J. London Math. Soc. **35** (1960), 81–84.

[59] Choo-whan Kim, *Uniform approximation of doubly stochastic operators*, Pacific. J. Math. **26** (1968), 515–527.

[60] ———, *Approximation of doubly substochastic operators*, Michigan Math. J. **19** (1972), 83–95.

[61] A. N. Kolmogorov, *Sur l'interpolation et extrapolation des suites stationnaires*, C. R. Acad. Sci. Paris **208** (1939), 2043–2045.

[62] ———, *Kurven im Hilbertschen Raum, die gegenüber einer einparametrigen Gruppe von Bewegungen invariant sind*, C. R. (Dokl.) Acad. Sci. URSS **26** (1940), 6–9.

[63] ———, *Stationary sequences in Hilbert space*, Bjull. Moskov. Gos. Univ. Mat. **2** (1941), no. 6.

[64] ———, *Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum*, C. R. (Dokl.) Acad. Sci. URSS **26** (1940), 115–118.

[65] ———, *Le transformation de Laplace dans les espaces linéaires*, C. R. Acad. Sci. Paris **200** (1935), 1717–1718.

[66] ———, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer-Verlag, Berlin, 1933; reprint, 1973; English transl., Chelsea, New York, 1948; 2nd ed., 1956.

[67] ———, *A remark on some work of R. A. Minlos and V. Sazonov*, Teor. Verojatnost. i Primenen. **4** (1959), 237–239 = Theor. Probability Appl. **4** (1959), 221–223.

[68] A. N. Kolmogorov and V. M. Tihomirov, *ϵ-entropy and ϵ-capacity of sets in functional spaces*, Uspehi Mat. Nauk **14** (1959), no. 2 (86), 3–86; English transl., Amer. Math. Soc. Transl. (2) 17 (1961), 277–364.

[69] H. J. Landau and L. A. Shepp, *On the supremum of a Gaussian process*, Sankhyā Ser. A **32** (1970), 369–378.

[70] Ju. V. Linnik, *Statistical problems with nuisance parameters*, "Nauka", Moscow, 1966; English transl., Amer. Math. Soc., Providence, R. I., 1968.

[71] Ju. V. Linnik, V. I. Romanovskiĭ and V. N. Sudakov, *A non-randomized homogeneous test in the Behrens-Fisher problem*, Dokl. Akad. Nauk SSSR **155** (1964), 1262–1264 = Soviet Math. Dokl. **5** (1964), 570–572.

[72] A. A. Ljapunov, *On completely additive vector functions*, Izv. Akad. Nauk SSSR Ser. Mat. **4** (1940), 465–478. (Russian)

[73] B. M. Makarov, *A characteristic of invariantly order-bounded sets in the space $L^p(\Omega, \mu)$*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **47** (1974), 73–80 = J. Soviet Math. 9 (1978), 189–195.

[74] M. B. Marcus, *A comparison of continuity conditions for Gaussian processes*, Ann. Probability **1** (1973), 123–130.

[75] M. B. Marcus and L. A. Shepp, *Sample behavior of Gaussian processes*, Proc. Sixth Berkeley Sympos. Math. Statist. and Probability (1970/71). Vol. II: Probability Theory, Univ. of California Press, Berkeley, Calif., 1972, pp. 423–441.

[76] R. A. Minlos, *Generalized random processes and their extension to a measure*, Trudy Moskov. Mat. Obšč. **8** (1959), 497–518; English transl., Selected Transl. Math. Statist. and Probability, Vol. 3, Amer. Math. Soc., Providence, R. I., 1962, pp. 291–313.

[77] ———, *Extension of a generalized random process to a completely additive measure*, Dokl. Akad. Nauk SSSR **119** (1958), 439–442. (Russian)

[78] L. Mirsky, *Results and problems in the theory of doubly stochastic matrices*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **1** (1962/63), 319–334.

[79] B. S. Mitjagin, *Approximate dimension and bases in nuclear spaces*, Uspehi Mat. Nauk 16 (1961), no. 4 (100), 63–132 = Russian Math. Surveys **16** (1961), no. 4, 59–127.

[80] Gaspard Monge, *Mémoire sur la théorie des déblais et des remblais*, Mém. Math. Phys. Acad. Roy. Sci. Paris **1781**, 666–704 (1784).

[81] I. P. Natanson, *Theory of functions of a real variable*, rev. ed., GITTL, Moscow, 1957; English transl., Vols. I, II, Ungar, New York, 1955, 1960.

[82] Makiko Nisio, *On the continuity of stationary Gaussian processes*, Nagoya Math. J. **34** (1969), 89–104.

[83] J. E. L. Peck, *Doubly stochastic measures*, Michigan Math. J. **6** (1959), 217–220.

[84] Robert R. Phelps, *Lectures on Choquet's theorem*, Van Nostrand, Princeton, N. J., 1966.

[85] Ju. V. Prohorov, *Probability distributions in functional spaces*, Uspehi Mat. Nauk **8** (1953), no. 3 (55), 165–167. (Russian)

[86] ———, *Convergence of random processes and limit theorems in probability theory*, Teor. Verojatnost. i Primenen. **1** (1956), 177–238 = Theor. Probability Appl. **1** (1956), 157–214.

[87] Ju. V. Prohorov and V. V. Sazonov, *Some results connected with Bochner's theorem*, Teor. Verojatnost. i Primenen. **6** (1961), 87–93 = Theor. Probability Appl. **6** (1961), 82–87.

[88] B. A. Rattray and J. E. L. Peck, *Infinite doubly stochastic matrices*, Trans. Roy. Soc. Canada Sect. III (3) **49** (1955), 55–57.

[89] P. Révész, *A probabilistic solution of Problem* 111 *of G. Birkhoff*, Acta Math. Acad. Sci. Hungar. **13** (1962), 187–198.

[90] Frédéric Riesz and Béla Sz.-Nagy, *Leçons d'analyse fonctionnelle*, 2nd ed., Akad. Kiadó, Budapest, 1953; English transl., Ungar, New York, 1955.

[91] A. P. Robertson and W. J. Robertson, *Topological vector spaces*, Cambridge Univ. Press, New York, 1964.

[92] V. A. Rohlin, *On the fundamental concepts of measure theory*, Mat. Sb. **25** (67) (1949), 107–150; English transl., Amer. Math. Soc. Transl. (1) **10** (1962), 1–54.

[93] I. V. Romanovskiĭ and V. N. Sudakov, *On the existence of independent decompositions*, Trudy Mat. Inst. Steklov **79** (1965), 5–10 = Proc. Steklov Inst. Math. **79** (1966), 1–7.

[94] Ju. A. Rozanov, *Infinite-dimensional Gaussian distributions*, Trudy Mat. Inst. Steklov **108** (1968) = Proc. Steklov Inst. Math. **108** (1968) (1971).

[95] John V. Ryff, *On the representation of doubly stochastic operators*, Pacific J. Math. **13** (1963), 1379–1386.

[96] ———, *Orbits of L^1-functions under doubly stochastic transformations*, Trans. Amer. Math. Soc. **117** (1965), 92–100.

[97] ———, *Extreme points of some convex subsets of* $L^1(0, 1)$, Proc. Amer. Math. Soc. **18** (1967), 1026–1034.

[98] ———, *On Muirhead's theorem*, Pacific J. Math. **21** (1967), 567–576.

[99] ———, *Majorised functions and measures*, Nederl. Akad. Wetensch. Proc. Ser. A **71** = Indag. Math. **30** (1968), 431–437.

[100] ———, *Measure preserving transformations and rearrangements*, J. Math. Anal. Appl. **31** (1970), 449–458.

[101] Herbert John Ryser, *Combinatorial mathematics*, Math. Assoc. Amer.; Wiley, New York, 1963.

[102] Yūji Sakai and Tetsuya Shimogaki, *Equimeasurability of functions and doubly stochastic operators*, Kōdai Math. Sem. Rep. **24** (1972), 203–211.

[103] V. Sazonov, *A remark on characteristic functionals*, Teor. Verojatnost. i Primenen. **3** (1958), 201–205 = Theor. Probability Appl. **3** (1958), 188–192.

[104] L. Schlaefli, *On the multiple integral* $\int^n dx\,dy \ldots dz$, *whose limits are* $p_1 = a_1x + b_1y + \cdots + h_1z > 0$, $p_2 > 0, \ldots, p_n > 0$, *and* $x^2 + y^2 + \cdots + z^2 < 1$, Quart. J. Pure Appl. Math. **2** (1858), 269–301; **3** (1860), 54–68.

[105] Erhard Schmidt, *Beweis der isoperimetrischen Eigenschaft der Kugel im hyperbolischen und sphärischen Raum jeder Dimensionenzahl*, Math. Z. **49** (1943), 1–109.

[106] I. J. Schoenberg, *Metric spaces and positive definite functions*, Trans. Amer. Math. Soc. **44** (1938), 522–536.

[107] L. A. Shepp, *On the supremum of gaussian sequences*, Preprint, Bell Telephone Laboratories, 1970.

[108] Ray C. Shiflett, *Extreme Markov operators and the orbits of Ryff*, Pacific J. Math. **40** (1972), 201–206.

[109] A. V. Skorohod, *Remark on Gaussian measures on a Banach space*, Teor. Verojatnost. i Primenen. **15** (1970), 519–520 = Theor. Probability Appl. **15** (1970), 508.

[110] David Slepian, *The one-sided barrier problem for Gaussian noise*, Bell System Tech. J. **41** (1962), 463–501.

[111] W. Słowikowski, *Absolutely 2-summing mappings from and to Hilbert spaces and a Sudakov theorem*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **17** (1969), 381–386.

[112] V. I. Smirnov, editor, *Gaspard Monge: Collection of papers on the two hundredth anniversary of his birth*, Izdat. Akad. Nauk SSSR, Moscow, 1947. (Russian)

[113] V. Strassen, *The existence of probability measures with given marginals*, Ann. Math. Statist. **36** (1965), 423–439.

[114] V. N. Sudakov, *On the extension of measures from Baire sets in nonseparable metric spaces*, Sibirsk. Mat. Ž. **2** (1961), 946–948. (Russian)

[115] ———, *Some problems connected with distributions in infinite-dimensional linear spaces*, Candidate's Dissertation, Leningrad State Univ., Leningrad, 1962. (Russian)

[116] ———, *On a class of compact sets in Hilbert space*, Uspehi Mat. Nauk **18** (1963), no. 1 (109), 181–187; *On the characterization of the quasi-equivalence of measures in Hilbert space*, ibid., 188–190. (Russian)

[117] ———, *On the existence of random variables that are independent of several given ones*, Trans. Fourth Prague Conf. Information Theory, Statistical Decision Functions, Random Processes (Prague, 1965), Academia, Prague, 1967, pp. 589–594; English transl., Selected Transl. Math. Statist. and Probability, Vol. 8, Amer. Math. Soc., Providence, R. I., 1970, pp. 205–211.

[118] ———, *The connection between a theorem on expansion in generalized eigenelements and a theorem on the extension of a weak distribution to a measure in Hilbert space*, Abstracts of Brief Scientific Communications, Section 5, Internat. Congr. Math., Moscow, 1966, p. 75. (Russian)

[119] ———, *The Romanovskiĭ-Sudakov theorem*, §2 in Chapter X of [70].

[120] ———, *A general theory of measure in linear spaces and realizations of Gaussian processes*, Proc. Sixth Math. Summer School: Probability Theory and Math. Statist. (Kaciveli, 1968), Akad. Nauk Ukrain. SSR Inst. Mat., Kiev, 1969, pp. 211–227. (Russian)

[121] ———, *Gaussian and Cauchy measures and ϵ-entropy*, Dokl. Akad. Nauk SSSR **185** (1969), 51–53 = Soviet Math. Dokl. **10** (1969), 310–313.

[122] ———, *On an independent complement of two decompositions when there exists a bounded density*, Trudy Mat. Inst. Steklov **111** (1970), 3–11 (and errata insert) = Proc. Steklov Inst. Math. **111** (1972), 1–10.

[123] ———, *Gaussian random processes, convex sets, spherically invariant measures in Banach spaces*, Manuscript deposited at the Leningrad Branch, Steklov Math. Inst. Acad. Sci. USSR, Leningrad, 1970. (Russian)

[124] ———, *On the representation of doubly stochastic integral operators by means of disjunct isomorphisms*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **19** (1970), 196–208 = Sem. Math. V. A. Steklov Math. Inst. Leningrad. **19** (1970), 112–119.

[125] ———, *Gaussian random processes and the measures of solid angles in Hilbert space*, Dokl. Akad. Nauk SSSR **197** (1971), 43–45 = Soviet Math. Dokl. **12** (1971), 412–415.

[126] ———, *A remark on marginal sufficiency*, Dokl. Akad. Nauk SSSR **198** (1971), 301–302 = Soviet Math. Dokl. **12** (1971), 796–798.

[127] ———, *On a problem of G. Birkhoff*, Dokl. Akad. Nauk SSSR **196** (1971), 1280–1282 = Soviet Math. Dokl. **12** (1971), 362–365.

[128] ———, *On marginal sufficiency of statistics*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **29** (1972), 92–101 = J. Soviet Math. **3** (1975), 792–800.

[129] ———, *Measures on subsets of direct products*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **29** (1972), 113–128 = J. Soviet Math. **3** (1975), 812–824.

[130] ———, *Existence of doubly stochastic densities*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **29** (1972), 129–146 = J. Soviet Math. **3** (1975), 825–839.

[131] ———, *A remark on the criterion of continuity of Gaussian sample functions*, Proc. Second Japan-USSR Sympos. Probability Theory (Kyoto, 1972), Lecture Notes in Math., vol. 330, Springer-Verlag, Berlin and New York, 1973, pp. 444–454.

[132] ———, *Geometrical problems in the theory of infinite-dimensional probability distributions*, Doctoral Dissertation, Leningrad State Univ., Leningrad, 1972. (Russian)

[133] V. N. Sudakov and B. S. Cirel'son, *Extremal properties of half-spaces for spherically invariant measures*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **41** (1974), 14–24, 165 = J. Soviet Math. 9 (1978), 9–17.

[134] Jürgen Tippe, *Zur isoperimetrischen Eigenschaft der Kugel in Riemannischen Räumen konstanter positiver Krümmung*, Math. Ann. **152** (1963), 120–148.

[135] N. N. Vahanija, *Sur une propriété des représentations normales de probabilité dans les espaces l_p $(1 \leq p < \infty)$ et H*, C. R. Acad. Sci. Paris **260** (1965), 1334–1336.

[136] N. N. Vahanija, *Nondegenerate probability distributions in the spaces* l_p $(1 < p < \infty)$, Teor. Verojatnost. i Primenen. **11** (1966), 524–528 = Theor. Probability Appl. **11** (1966), 463–467.

[137] ———, *Probability distributions in linear spaces*, Sakharth. SSR Mecn. Akad. Gamothvl. Centr. Šrom. **10** (1971), no. 3. (Russian)

[138] A. M. Veršik, *Some remarks on infinite-dimensional problems in linear programming*, Uspehi Mat. Nauk **25** (1970), no. 5 (155), 117–124 = Russian Math. Surveys **25** (1970), no. 5, 117–124.

[139] ———, *Axiomatics of measure theory in linear spaces*, Dokl. Akad. Nauk SSSR **178** (1968), 278–281 = Soviet Math. Dokl. **9** (1968), 68–72.

[140] ———, *Duality in the theory of measure in linear spaces*, Dokl. Akad. Nauk SSSR **170** (1966), 497–500 = Soviet Math. Dokl. **7** (1966), 1210–1214.

[141] A. M. Veršik and V. N. Sudakov, *Probability measures in infinite-dimensional spaces*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **12** (1969), 7–67 = Sem. Math. V. A. Steklov Math. Inst. Leningrad. **12** (1971), 1–28.

[142] ———, *Topological questions in the theory of measure in linear spaces*, Uspehi Mat. Nauk **17** (1962), no. 4, 217–219. (Russian)

[143] D. A. Vladimirov, *Boolean algebras*, "Nauka", Moscow, 1969; German transl., Akademie-Verlag, Berlin, 1972.

[144] André Weil, *L'intégration dans les groupes topologiques et ses applications*, Actualités Sci. Indust., no. 869, Hermann, Paris, 1940; Princeton Univ. Press, Princeton, N. J., 1941.

[145] Samuel S. Wilks, *Mathematical statistics*, Wiley, New York, 1962.

[146] V. A. Zalgaller, *Mixed volumes and the probability of a hit in a convex region for a multi-dimensional normal distribution*, Mat. Zametki **2** (1967), 97–104 = Math. Notes **2** (1967), 542–545.