National Research University Higher School of Economics

as a manuscript

Borzykh Dmitriy

Joint distributions of generalized integrable increasing processes and their generalized compensators

Summary of the PhD thesis
for the purpose of obtaining academic degree
Doctor of Philosophy in Mathematics

Academic supervisor
Doctor of Science, professor
Alexander Gushchin

Moscow – 2022

Introduction

The dissertation was prepared at the International Laboratory of Stochastic Analysis and its Applications of the National Research University "Higher School of Economics" (NRU HSE).

Relevance of the topic. An important area of stochastic analysis consists in searching for a set of joint distributions of stochastic processes and their components. The central objects of our dissertation are increasing processes and their compensators. The problem that we study is related to the characterization of the set of joint distributions of an increasing process and its compensator at two consecutive points in time. Since similar problems were considered separately for increasing processes and for martingales (the difference between an increasing process and its compensator is a martingale), let us briefly consider the previous history of studies in these problems, in which integral orders play an essential role.

Apparently, one of the first works where the ideas of stochastic ordering arose is the first edition of the book «Inequalities» by Hardy, Littlewood and Polya, 1934 (see [23]). Their idea of majorization of vectors in the space $\mathbb{R}^n$ was not formulated in terms of stochastic orders, but can be naturally reformulated into this language by interpreting a vector $x = (x_1, \dots, x_n)$ as a discrete probability measure $\sum_{j=1}^{n} \frac{1}{n}\delta_{\{x_j\}}$ on the real line concentrated at the points $x_1, \dots, x_n$ and having mass $1/n$ at each of these points. In section 2.18 of this book, Hardy, Littlewood, and Polya introduce the following order relation on the set of nonnegative n-dimensional real vectors. A vector $x = (x_1, \dots, x_n)$ is said to be *majorized* by a vector $y = (y_1, \dots, y_n)$ if $\sum_{j=1}^{k} x_{(j)} \leq \sum_{j=1}^{k} y_{(j)}$ for all $k = 1, \dots, n$ and $\sum_{j=1}^{n} x_{(j)} = \sum_{j=1}^{n} y_{(j)}$, where $(x_{(1)}, \dots, x_{(n)})$ means the vector x reordered in descending order. In section 3.17 of the same book, Hardy, Littlewood, and Polya obtained an interesting characterization of this order relation (see Proposition 108). It states that the following conditions are equivalent:

(i) a vector x is majorized by a vector y,

(ii) there exists a doubly stochastic matrix Π such that $x = \Pi y$ (here vectors x and y are treated as column vectors),

(iii) $\sum_{j=1}^{n} f(x_j) \le \sum_{j=1}^{n} f(y_j)$ for any convex function $f\colon \mathbb{R} \to \mathbb{R}$.

More information about the theory of majorization and the history of the development of this direction can be found in the classic book by Marshall and Olkin [7].

Now, consider more recent research in this direction. Let us recall some definitions.

Let $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$ be vectors from the space $\mathbb{R}^d$. In the space $\mathbb{R}^d$ we introduce the *partial order relation* $\preceq$ in the standard way. We say that $x \preceq y$ if $x_i \le y_i$ for all $i = 1, \dots, d$. A function $f\colon \mathbb{R}^d \to \mathbb{R}$ is called *nondecreasing* if it is nondecreasing with respect to the partial order $\preceq$, i.e. $f(x) \le f(y)$ if $x \preceq y$.

In what follows, we will repeatedly need the definition of *Markov kernel* (*transition kernel* or *transition probability*). Let two measurable spaces $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be given. A mapping $\mathrm{Q}\colon \Omega_1 \times \mathcal{F}_2 \to [0;1]$ is called a Markov kernel from Ω_1 to Ω_2 if

1) for any $\omega_1 \in \Omega_1$ the function $\mathrm{Q}(\omega_1; \cdot\,)$ is a probability measure on $(\Omega_2, \mathcal{F}_2)$;

2) for any $A_2 \in \mathcal{F}_2$ the function $\mathrm{Q}(\,\cdot\,; A_2)$ is measurable on $(\Omega_1, \mathcal{F}_1)$.

More detailed information about Markov kernels, including the Fubini theorem for Markov kernels, can be found, for example, in [8] (Ch. III, § III.2), [28] (Ch. 8, § 8.3 and Ch. 14, § 14.2), or [10] (Ch. 2, § 2.6).

As noted at the beginning of this section, the problem we are studying is closely related to stochastic orders. A fairly complete review of this topic is contained in the fundamental monograph by Müller and Stoyan [32]. In a concentrated form, a necessary information about stochastic orders can be found, for example, in the book [9] by Föllmer and Schied (see Ch. 2, §§ 2.4, 2.6). Here we will focus only on two stochastic orders that are directly related to our study. These are the ordinary stochastic order and the convex stochastic order.

Let μ_1 and μ_2 be two Borel probability measures on $\mathbb{R}^d$. We say that *a measure μ_2 stochastically dominates a measure μ_1 in sense of the usual stochastic order*, denote $\mu_1 \preceq_{\mathrm{st}} \mu_2$, if for all bounded Borel nondecreasing functions $f\colon \mathbb{R}^d \to \mathbb{R}$,

$$\int_{\mathbb{R}^d} f(x)\,\mu_1(dx) \le \int_{\mathbb{R}^d} f(x)\,\mu_2(dx). \tag{1}$$

Apparently, the stochastic order $\preceq_{\mathrm{st}}$ first appeared in the work of Mann and Whitney in 1947 (see [31]) and Lehman's paper in 1955 (see [29]) in hypothesis testing problems. The properties of the stochastic order $\preceq_{\mathrm{st}}$ have been studied in detail in the works [25, 26, 36, 34]. Various characterizations of the order $\preceq_{\mathrm{st}}$ were obtained. Below we will give one modern formulation of such characterizations. The following theorem holds.

Theorem 0.1. *For two Borel probability measures μ_1 and μ_2 given on $\mathbb{R}^d$, the following conditions are equivalent:*

(i) $\mu_1 \preceq_{st} \mu_2$*;*

(ii) *there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random vectors $X_i\colon \Omega \to \mathbb{R}^d$, $i = 1, 2$, such that* $\mathrm{Law}(X_i) = \mu_i$*, $i = 1, 2$, and $X_1 \preceq X_2$ $\mathbb{P}$-a. s.;*

(iii) *there exists a Markov kernel $\mathrm{Q}(x; B)$, where $x \in \mathbb{R}^d$, $B \in \mathcal{B}(\mathbb{R}^d)$, such that $\mu_2(B) = \int_{\mathbb{R}^d} \mathrm{Q}(x; B)\, \mu_1(dx)$ for all $B \in \mathcal{B}(\mathbb{R}^d)$, and $\mathrm{Q}\big(x; \{y\colon x \preceq y\}\big) = 1$ for any $x \in \mathbb{R}^d$.*

A modern proof of this theorem can be found, for example, in the book [9] by Föllmer and Schied (see Ch. 2, § 2.6, Theorem 2.95).

Now, let us consider convex stochastic order. Let two Borel probability measures μ_1 and μ_2 with finite expectations be given on $\mathbb{R}^d$, i.e. $\int_{\mathbb{R}^d} ||x||\, \mu_i(dx) < \infty$, $i = 1, 2$, where $||x||$ is the Euclidean norm of vector x. We say that *a measure μ_2 stochastically dominates a measure μ_1 in sense of the convex order*, denote $\mu_1 \preceq_{\mathrm{cx}} \mu_2$, if the inequality (1) holds for all convex functions $f\colon \mathbb{R}^d \to \mathbb{R}$ for which both integrals in (1) make sense.

Apparently, the concept of a convex stochastic order first arose in Blackwell's paper in 1953 (see [14]) in the problem of comparing statistical experiments. The properties of the convex stochastic order were studied in detail in the works [19, 18, 37, 33]. Below we give one of the modern formulations of the theorem, which contains a characterization of the convex stochastic order. The following theorem holds.

Theorem 0.2. *Let two Borel probability measures μ_1 and μ_2 with finite expectations be given on $\mathbb{R}^d$. Then the following conditions are equivalent:*

(i) $\mu_1 \preceq_{cx} \mu_2$*;*

(ii) *there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random vectors $X_i\colon \Omega \to \mathbb{R}^d$, $i = 1, 2$, such that* $\mathrm{Law}(X_i) = \mu_i$*, $i = 1, 2$, and $\mathbb{E}[X_2|X_1] = X_1$ $\mathbb{P}$-a. s.;*

(iii) *there exists a Markov kernel* $\mathrm{Q}(x; B)$*, where* $x \in \mathbb{R}^d$*,* $B \in \mathcal{B}(\mathbb{R}^d)$*, such that* $\mu_2(B) = \int_{\mathbb{R}^d} \mathrm{Q}(x; B)\, \mu_1(dx)$ *for all* $B \in \mathcal{B}(\mathbb{R}^d)$*, and* $\int_{\mathbb{R}^d} y\, \mathrm{Q}(x; dy) = x$ *for any* $x \in \mathbb{R}^d$*.*

A modern proof of this theorem can be found, for example, in the book [9] by Föllmer and Schied (see Ch. 2, § 2.6, Theorem 2.93, and Corollary 2.94).

Note that in the proofs of Theorems 0.1 and 0.2, the Strassen theorem [37] plays a key role, and now we turn to its formulation.

Let S be a Polish space. Consider an arbitrary continuous function $\psi\colon S \to [1; +\infty)$, which will be called the *gauge function*. Define a *class* $C_\psi(S)$ *of continuous test functions* $f\colon S \to \mathbb{R}$ such that

$$\forall f \in C_\psi(S) \;\; \exists c \in \mathbb{R} \;\; \forall x \in S \quad |f(x)| \le c \cdot \psi(x).$$

Denote by $\mathcal{M}_1^\psi(S)$ *the set of all Borel probability measures on* S for which $\int_S \psi(x)\, \mu(dx) < \infty$. The ψ*-weak topology* on $\mathcal{M}_1^\psi(S)$ is the coarest topology such that

$$\mathcal{M}_1^\psi(S) \ni \mu \mapsto \int_S f(x)\, \mu(dx)$$

is a continuous mapping for all $f \in C_\psi(S)$. It is easy to see that sets

$$U_\varepsilon^\psi(\mu;\, f_1, \dots, f_m) := \bigcap_{i=1}^m \left\{ \nu \in \mathcal{M}_1^\psi(S)\colon \left| \int_S f_i\, d\nu - \int_S f_i\, d\mu \right| < \varepsilon \right\},$$

where $\mu \in \mathcal{M}_1^\psi(S)$, $\varepsilon > 0$, $m \in \mathbb{N}$, and $f_1, \dots, f_m \in C_\psi(S)$, form the base of the ψ-weak topology on $\mathcal{M}_1^\psi(S)$. Note that the space $\mathcal{M}_1^\psi(S)$ is metrizable and separable (see [9], Corollary A.44). More details about the space $\mathcal{M}_1^\psi(S)$ and its properties can be found, for example, in [9] (§A.6, pp. 442–445).

On the product of the spaces $S \times S$, consider the following gauge function

$$\overline{\psi}(x_1, x_2) := \psi(x_1) + \psi(x_2).$$

We define the corresponding set of continuous test functions $C_{\overline{\psi}}(S \times S)$ and the space of probability measures $\mathcal{M}_1^{\overline{\psi}}(S \times S)$ endowed with the $\overline{\psi}$-weak topology. The following famous theorem holds.

Theorem 0.3 (Strassen)**.** *Suppose that* $\Lambda \subseteq \mathcal{M}_1^{\overline{\psi}}(S \times S)$ *is convex and closed in the* $\overline{\psi}$*-weak topology, and that* μ_1*,* μ_2 *are probability measures in* $\mathcal{M}_1^\psi(S)$*. Then there exists some measure* $\overline{\mu} \in \Lambda$ *with marginal distributions* μ_1 *and* μ_2 *if and only if*

$$\int_S f_1(x_1)\, \mu_1(dx_1) + \int_S f_2(x_2)\, \mu_2(dx_2) \le \sup_{\lambda \in \Lambda} \int_{S \times S} \big(f_1(x_1) + f_2(x_2)\big)\, \lambda(dx_1, dx_2).$$

The idea behind the proof of the Strassen theorem is to use the Hahn–Banach theorem in the form of a separability theorem for sets in a locally convex topological space. However, to apply the Hahn–Banach theorem, a proper topological setting of the problem is required. This is the most difficult part of the proof. The proof of the Strassen theorem can be found in Strassen's original article [37] or, for example, in the book by Föllmer and Schied (see [9], Ch. 2, §2.6, Theorem 2.88).

We emphasize that the Strassen theorem plays a central role not only in substantiating Theorems 0.1 and 0.2, but also in proving the main proposition of our dissertation, Theorem 1.3.

Remark 0.1. Note that Theorems 0.1 and 0.2 can be reformulated in terms of the existence of d-dimensional stochastic process with certain properties. Indeed, let us fix two moments of time $0 \le a < b < \infty$. Let two Borel probability measures μ_1 and μ_2 be given on $\mathbb{R}^d$. Theorem 0.1 (see items (i) and (ii)) gives necessary and sufficient conditions on the measures μ_1 and μ_2 for the existence of a d-dimensional random process X_t, $t \in [a;b]$ having nondecreasing trajectories, for which $\mathrm{Law}(X_a) = \mu_1$ and $\mathrm{Law}(X_b) = \mu_2$. If the Borel measures μ_1 and μ_2 on $\mathbb{R}^d$ have finite expectations, then Theorem 0.2 (see items (i) and (ii)) contains necessary and sufficient conditions on the measures μ_1 and μ_2 for the existence of d-dimensional martingale X_t, $t \in [a;b]$, such that $\mathrm{Law}(X_a) = \mu_1$ and $\mathrm{Law}(X_b) = \mu_2$.

The two constructions described above refer to the situation when all the components of the d-dimensional process X have "the same nature". This is, to some extent, a simpler situation. A more complicated situation arises when the components of the process X are of "different nature". For example, when one of the components of the process X is built in some way by its other component. An example of such a situation is the case when the first component of a two-dimensional process X is a nonnegative submartingale of class (D) starting from zero, and the second component is its compensator, i.e., a predictable increasing process from the Doob–Meyer decomposition. Such a construction was considered in a recent paper in 2017 (see [4]). Other constructions of this kind are presented, for example, in the classical work of C. Rogers [35] and a number of other works [15, 21, 11, 12, 27, 38].

Now, let us turn directly to the objectives of our study. Let the stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in\mathbb{R}_+})$ be given. An adapted stochastic process $X = (X_t)_{t\in\mathbb{R}_+}$ is called a *increasing process* if all its trajectories are right-continuous, start from zero, and are nondecreasing functions. An increasing process $X = (X_t)_{t\in\mathbb{R}_+}$ is called an *integrable increasing process* if $\mathbb{E}[X_\infty] < \infty$. The class of all integrable increasing processes is denoted by $\mathcal{A}^+$.

Note that every integrable increasing process $X = (X_t)_{t\in\mathbb{R}_+}$ is a submartingale of class (D) (see [5], §1.46). Hence, by the Doob–Meyer decomposition (see [5], §3.15) there exists a unique (up to indistinguishability) increasing integrable predictable process A with $A_0 = 0$, such that the process $X - A$ is a uniformly integrable martingale. The process A in this decomposition will be called the *compensator* of the process X.

In the paper [4] mentioned above, a *class* $\mathbb{W}$ *of probability measures* was introduced. It includes all probability measures μ on $(\mathbb{R}^2_+,\, \mathcal{B}(\mathbb{R}^2_+))$ satisfying the following conditions:

1) $\int_{\mathbb{R}^2_+} (x+y)\,\mu(dx,\, dy) < \infty$,

2) $\int_{\mathbb{R}^2_+} x\,\mu(dx,\, dy) = \int_{\mathbb{R}^2_+} y\,\mu(dx,\, dy)$,

3) $\forall c \geq 0\ \int_{\{y\leq c\}} x\,\mu(dx,\, dy) \leq \int_{\mathbb{R}^2_+} (y\wedge c)\,\mu(dx,\, dy)$.

Let $T \in [0;\infty]$ be an arbitrary fixed moment of time. In [4] it is shown that a measure μ belongs to the class $\mathbb{W}$ if and only if there exists an integrable increasing process $(X_t)_{t\in\mathbb{R}_+}$ with a compensator $(A_t)_{t\in\mathbb{R}_+}$ such that $\mathrm{Law}(X_T,\, A_T) = \mu$.

In our work, we generalize the problem statement considered in [4]. To do this, we introduce the concept of a generalized integrable increasing process and its generalized compensator. An adapted process $X = (X_t)_{t\in\mathbb{R}_+}$ will be called a *generalized integrable increasing process* if it can be represented as $X_t = \xi_0 + X^\circ_t$, $t \in \mathbb{R}_+$, where ξ_0 is a $\mathcal{F}_0$-measurable integrable random variable, and $X^\circ = (X^\circ_t)_{t\in\mathbb{R}_+}$ is an integrable increasing process in the usual sense. By the Doob–Meyer theorem, the process X° has a compensator $A^\circ = (A^\circ_t)_{t\in\mathbb{R}_+}$. Then the *generalized compensator* of the generalized integrable increasing process X will be defined as a random process $A = (A_t)_{t\in\mathbb{R}_+}$ of the form $A_t = \eta_0 + A^\circ_t$, where η_0 is an arbitrary $\mathcal{F}_0$-measurable integrable random variable. Thus, according to this definition, the generalized compensator of a generalized integrable increasing process is uniquely defined up to the addition of a $\mathcal{F}_0$-measurable integrable random variable. Note that every generalized compensator of an integrable generalized increasing process is itself an integrable generalized increasing process.

Let us fix on $[0;\infty]$ two moments of time a and b. Without loss of generality, we can assume that $a = 1$ and $b = 2$. Consider the *class of probability measures* Λ, which includes all joint distributions $\lambda := \mathrm{Law}\left(\left[\begin{smallmatrix} X_1 \\ A_1 \end{smallmatrix}\right],\, \left[\begin{smallmatrix} X_2 \\ A_2 \end{smallmatrix}\right]\right)$, where $(X_t)_{t\in[1;2]}$ is a generalized integrable increasing process, and $(A_t)_{t\in[1;2]}$ its generalized compensator. We are interested in how the class Λ of measures is arranged.

More specifically, in this work we solve the following two main problems. The first of them is to obtain necessary and sufficient conditions for a probability measure λ given on $\mathcal{B}(\mathbb{R}^2\times\mathbb{R}^2)$ to belong to the class Λ. If we take into account the definition of the class Λ, then the first problem can be reformulated as follows: it is required to find necessary and sufficient conditions on the measure λ for the existence of a generalized integrable increasing process $(X_t)_{t\in[1;2]}$ having a generalized compensator $(A_t)_{t\in[1;2]}$ such that $\operatorname{Law}\left(\left[\begin{smallmatrix}X_1\\A_1\end{smallmatrix}\right],\left[\begin{smallmatrix}X_2\\A_2\end{smallmatrix}\right]\right)=\lambda$.

The second task is posed as follows. Let two probability measures μ_1 and μ_2 be given on $(\mathbb{R}^2,\mathcal{B}(\mathbb{R}^2))$ and satisfy the conditions $\int(|x|+|y|)\,d\mu_i<\infty$, $i=1,2$. It is required to obtain necessary and sufficient conditions for the measures μ_1 and μ_2 in order to the set Λ to contain a measure λ for which μ_1 and μ_2 are marginal distributions, i.e., $\lambda(B\times\mathbb{R}^2)=\mu_1(B)$ and $\lambda(\mathbb{R}^2\times B)=\mu_2(B)$ for any $B\in\mathcal{B}(\mathbb{R}^2)$. In other words, it is required to find necessary and sufficient conditions on the measures μ_1 and μ_2 for the existence of a generalized integrable increasing process $(X_t)_{t\in[1;2]}$ having a generalized compensator $(A_t)_{t\in[1;2]}$, such that $\operatorname{Law}\left[\begin{smallmatrix}X_1\\A_1\end{smallmatrix}\right]=\mu_1$ и $\operatorname{Law}\left[\begin{smallmatrix}X_2\\A_2\end{smallmatrix}\right]=\mu_2$.

The study of the properties of increasing processes and their compensators is an important area of stochastic analysis. In particular, this is due to the fact that quadratic characteristics of martingales and random time changes are increasing processes. The study of the properties of increasing processes and their compensators is not only of purely theoretical interest, dictated by the internal needs of the development of stochastic analysis (see, e.g., [20], [24], [30], [6]). These objects also often arise in applied areas such as financial mathematics. For example, in [22], square integrable semimartingales are considered and convex-order relations between their quadratic and predictable quadratic variation, i.e. between an increasing process and its compensator, are investigated. The results of this work are used in the pricing of options in which the underlying asset is a realized variance. Another example is the article [13], related to credit risk models, which examines the increasing process generated by the moment of default of a company (or state) and its compensator.

It should be noted that the problems considered in the dissertation are of a theoretical nature, and issues related to specific applications require separate consideration. Apparently, this may become one of the directions of our further research.

The purpose of the study. The main goal of the dissertation is to study the properties of the set of measures Λ introduced above, as well as to obtain necessary and sufficient conditions for the existence of a measure $\lambda \in \Lambda$ with given marginal distributions μ_1 and μ_2.

Scientific novelty. All main results of the dissertation are new and are as follows.

1. Consider the set Λ of all boundary joint distributions $\mathrm{Law}([X_a, A_a], [X_b, A_b])$ at moments of time $t = a$ and $t = b$ of integrable increasing processes $(X_t)_{t\in[a;b]}$ and their compensators $(A_t)_{t\in[a;b]}$, which at the initial moment of time start from an arbitrary integrable initial condition $[X_a, A_a]$. We have established that the set Λ is convex and closed in the ψ-weak topology with a gauge function ψ of linear growth. We found necessary and sufficient conditions for a certain probability measure λ defined on $\mathcal{B}(\mathbb{R}^2 \times \mathbb{R}^2)$ to belong to the class of measures Λ. The main result of the dissertation is the following: for two measures μ_a and μ_b defined on $\mathcal{B}(\mathbb{R}^2)$, necessary and sufficient conditions are obtained in order to the set Λ to contain a measure λ whose marginals are μ_a and μ_b.

2. In [4], the class $\mathbb{W}$ of terminal distributions of integrable increasing processes and their compensators was introduced. We have shown that distributions with finite support lying in $\mathbb{W}$ form a dense subset in the set $\mathbb{W}$ in the ψ-weak topology with a gauge function of linear growth.

3. We have proved that the joint distribution of an arbitrary locally integrable increasing process and its compensator at a terminal time can be realized as a joint terminal distribution of some other locally integrable increasing process and its compensator, but the compensator is already continuous.

Research methods. Methods of probability theory, methods of the general theory of stochastic processes and, in particular, martingale theory, as well as methods of real and functional analysis are used in the work.

Theoretical and practical value. The work is of a theoretical nature. Its results can be useful in the theory of random processes, stochastic analysis, as well as in problems of financial mathematics.

Approbation of the work. The results related to the dissertation were presented at the following conferences and scientific seminars:

1. "5-th International Conference on Stochastic Methods (ICSM-5)", Moscow, 2020. Topic: "Locally integrable increasing processes with continuous compensators";

2. "LSA Autumn Meeting 2020", Moscow, 2020. Topic: "Locally integrable increasing processes with continuous compensators";

3. "LSA Autumn Meeting 2021", Moscow, 2021. Topic: "On the denseness of the subset of discrete distributions in a certain set of two-dimensional distributions";

4. Scientific seminar of CEMI "Probabilistic control problems and stochastic models in economics, finance and insurance" under the guidance of V. I. Arkin, T. A. Belkina, E. L. Presman. Topic: "Joint distributions of increasing processes and their compensators", (in Russian), Moscow, 2021.

Publications. The results of the dissertation are published in [1, 17, 16]. All articles are published in journals indexed in the abstract and citation database «Scopus»:

1) the article [1] was published without co-authors in the journal «Theory of Probability and Its Applications» (Q3);

2) the article [17] was published jointly with the supervisor in the journal «Modern Stochastics: Theory and Applications» (Q2–Q3);

3) the article [16] was published without co-authors in the journal «Theory of Stochastic Processes» (Q4).

Structure and volume of the work. The dissertation is presented on 95 pages and consists of a table of contents, a list of designations, an introduction, three chapters, conclusion and bibliography, including 47 titles.

The content of the work

Chapter 1 contains the main results of the dissertation. First of all, these are Theorems 1.1 and 1.3. In Theorem 1.1, necessary and sufficient conditions are obtained that a certain probability measure λ, given on $\mathcal{B}(\mathbb{R}^2 \times \mathbb{R}^2)$, belongs to the class of measures Λ introduced above. We give an exact formulation of this theorem.

Theorem 1.1. *The probability measure λ defined on $\mathcal{B}(\mathbb{R}^2 \times \mathbb{R}^2)$ belongs to the class Λ if and only if it satisfies the following conditions:*

$$\int_{\mathbb{R}^2\times\mathbb{R}^2} \big(|x_1| + |y_1| + |x_2| + |y_2|\big)\, \lambda\left(\left[\begin{smallmatrix} dx_1 \\ dy_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} dx_2 \\ dy_2 \end{smallmatrix}\right]\right) < \infty;$$

for all $B \in \mathcal{B}(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2\times\mathbb{R}^2} (x_2 - x_1) \cdot \mathbb{1}_{\left\{\left[\begin{smallmatrix} x_1 \\ y_1 \end{smallmatrix}\right] \in B\right\}}\, \lambda\left(\left[\begin{smallmatrix} dx_1 \\ dy_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} dx_2 \\ dy_2 \end{smallmatrix}\right]\right) =$$
$$= \int_{\mathbb{R}^2\times\mathbb{R}^2} (y_2 - y_1) \cdot \mathbb{1}_{\left\{\left[\begin{smallmatrix} x_1 \\ y_1 \end{smallmatrix}\right] \in B\right\}}\, \lambda\left(\left[\begin{smallmatrix} dx_1 \\ dy_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} dx_2 \\ dy_2 \end{smallmatrix}\right]\right);$$

for any $B \in \mathcal{B}(\mathbb{R}^2)$ and for all $c \geq 0$,

$$\int_{\mathbb{R}^2\times\mathbb{R}^2} (x_2 - x_1) \cdot \mathbb{1}_{\{y_2 - y_1 \leq c\}} \cdot \mathbb{1}_{\left\{\left[\begin{smallmatrix} x_1 \\ y_1 \end{smallmatrix}\right] \in B\right\}}\, \lambda\left(\left[\begin{smallmatrix} dx_1 \\ dy_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} dx_2 \\ dy_2 \end{smallmatrix}\right]\right) \leq$$
$$\leq \int_{\mathbb{R}^2\times\mathbb{R}^2} \big[(y_2 - y_1) \wedge c\big] \cdot \mathbb{1}_{\left\{\left[\begin{smallmatrix} x_1 \\ y_1 \end{smallmatrix}\right] \in B\right\}}\, \lambda\left(\left[\begin{smallmatrix} dx_1 \\ dy_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} dx_2 \\ dy_2 \end{smallmatrix}\right]\right).$$

If we take into account the definition of the class Λ, then it becomes clear that Theorem 1.1 gives necessary and sufficient conditions on the measure λ for the existence of a generalized integrable increasing process $(X_t)_{t\in[1;2]}$ having a generalized compensator $(A_t)_{t\in[1;2]}$ such that $\mathrm{Law}\left(\left[\begin{smallmatrix} X_1 \\ A_1 \end{smallmatrix}\right], \left[\begin{smallmatrix} X_2 \\ A_2 \end{smallmatrix}\right]\right) = \lambda$.

Now, we turn to Theorem 1.3. Let two probability measures μ_1 and μ_2 be given on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ and satisfy the conditions $\int (|x| + |y|)\, d\mu_i < \infty$, $i = 1, 2$. Theorem 1.3 contains necessary and sufficient conditions for the set Λ to contain a measure λ for which μ_1 and μ_2 are marginal distributions. In order to give an exact formulation of Theorem 1.3, we define the following class of test functions.

Definition 1.1. Let us introduce a *class $\mathcal{K}$ of upper semicontinuous test functions* $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ that satisfy the following two conditions:

1) $\forall \varphi \in \mathcal{K}\ \ \exists c \in \mathbb{R}\ \ \forall x, y \in \mathbb{R}\quad \varphi(x, y) \le c \cdot \psi(x, y)$, where $\psi(x, y) := 1 + |x| + |y|$;

2) for all $x, y \in \mathbb{R}$ and any probability measure $\mu \in \mathbb{W}$,

$$\varphi(x, y) \le \int_{\mathbb{R}^2_+} \varphi(x + u, y + v)\, \mu(du, dv).$$

Theorem 1.3. *Let two probability measures μ_1 and μ_2 be given on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ and satisfy the conditions $\int (|x| + |y|)\, d\mu_i < \infty$, $i = 1, 2$. Then the following conditions are equivalent*:

(a) *on some stochastic basis there is a generalized integrable increasing process $X = (X_t)_{t\in[1;2]}$ with generalized compensator $A = (A_t)_{t\in[1;2]}$ such that* $\mathrm{Law}\left[\begin{smallmatrix} X_1 \\ A_1 \end{smallmatrix}\right] = \mu_1$ *and* $\mathrm{Law}\left[\begin{smallmatrix} X_2 \\ A_2 \end{smallmatrix}\right] = \mu_2$;

(b) *there exists a measure $\lambda \in \Lambda$ whose marginal distributions are μ_1 and μ_2, i.e., $\lambda(B \times \mathbb{R}^2) = \mu_1(B)$ and $\lambda(\mathbb{R}^2 \times B) = \mu_2(B)$ for any B in$\mathcal{B}(\mathbb{R}^2)$*;

(c) $\int \varphi(x, y)\, d\mu_1 \le \int \varphi(x, y)\, d\mu_2$ *for any test function $\varphi \in \mathcal{K}$*.

Let us note that both processes $(X_t)_{t\in[1;2]}$ and $(A_t)_{t\in[1;2]}$ from Theorem 1.3 have nondecreasing trajectories. Thus, from the first part of Remark 0.1, the class of test functions $\mathcal{K}$ must contain all bounded Borel functions $f\colon \mathbb{R}^2 \to \mathbb{R}$ such that $f(x_1, y_1) \le f(x_2, y_2)$ if $x_1 \le x_2$ and $y_1 \le y_2$. Moreover, as $X_t - A_t$, $t \in [1; 2]$, is a martingale, according to the second part of Remark 0.1, the class $\mathcal{K}$ contains all functions $g\colon \mathbb{R}^2 \to \mathbb{R}$ of the form $g(x, y) = h(x - y)$, where $h\colon \mathbb{R} \to \mathbb{R}$ is a convex function. Therefore, we come to the conclusion that the class $\mathcal{K}$ contains the cone $\mathcal{C}$ generated by all the above functions f and g. At the same time, it turns out that there exists a function φ (see the example from Paragraph 1.6 of Chapter 1 of the dissertation) that belongs to the class $\mathcal{K}$, but does not belong to the cone $\mathcal{C}$. Hence, the class of test functions $\mathcal{K}$ is wider than the cone $\mathcal{C}$. This remark explains the nontriviality of the problem being solved in the dissertation.

Despite the fact that Definition 1.1 contains a rather non-constructive description of the class of test functions $\mathcal{K}$, the definition turned out to be very convenient in the proof of Theorem 1.3. Nevertheless, it is quite difficult to use this definition when checking that a concrete function belongs to the class $\mathcal{K}$. Therefore, we would like to have a description of the class $\mathcal{K}$ with more easily verifiable characterizing conditions. The following theorem gives such a description.

Theorem 1.4. *Let an upper semicontinuous function $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ satisfy condition 1) from the definition of the class $\mathcal{K}$. Then the following conditions are equivalent:*

(i) *function φ satisfies condition 2) from the definition of class $\mathcal{K}$;*

(ii) *for all points $x, y \in \mathbb{R}$, any real number $k > 0$, and any real number $h \geq k$, we have the inequality*

$$0 \leq \frac{\varphi(x, y+k) - \varphi(x, y)}{k} + \frac{\varphi(x+h, y+k) - \varphi(x, y+k)}{h};$$

(iii) *for any points $x, y \in \mathbb{R}$, any real number $k > 0$, and arbitrary $x_j \geq 0$, $p_j \geq 0$, such that $\sum_{j=1}^n p_j = 1$ and $\sum_{j=1}^n x_j p_j = k$, the inequality holds*

$$\varphi(x, y) \leq \sum_{j=1}^{n} p_j \varphi(x + x_j, y + k).$$

In other words, in the definition of the class $\mathcal{K}$, condition 2) can be replaced by any of the conditions (ii) or (iii) from Theorem 1.4. Note that condition (ii) means that the sum of the tangents of the angles in the corresponding right triangles must be nonnegative. In particular, it follows from condition (ii) for $h = k$ that the function φ grows «along the diagonals», i.e. for any $x, y \in \mathbb{R}$ and any real $k > 0$ we have the inequality $\varphi(x, y) \leq \varphi(x + k, y + k)$.

Chapters 2 and 3 contain auxiliary statements necessary to prove the main results of the dissertation. At the same time, it should be noted that some of these auxiliary propositions are of independent interest.

Chapter 2 contains technical Theorem 2.1 which is very important in proving the main Theorem 1.3.

Before giving the exact formulation of Theorem 2.1, we will give some definitions. In the set of measures $\mathbb{W}$, consider a subset of *simple measures* $\mathbb{W}_{\mathrm{simp}}$ and a subset of *discrete measures* $\mathbb{W}_{\mathrm{disc}}$. We say that $\mu \in \mathbb{W}_{\mathrm{simp}}$ (correspondingly $\mu \in \mathbb{W}_{\mathrm{disc}}$) if $\mu \in \mathbb{W}$, and the measure μ has the form

$$\mu(dx,\, dy) = \sum_{j \in J} p_j \cdot \delta_{\left[\begin{smallmatrix} x_j \\ a_j \end{smallmatrix}\right]}(dx,\, dy),$$

where J is a finite set (correspondingly, J is at most countable), $p_j \geq 0$, $\sum_{j\in J} p_j = 1$, and $\delta_{\left[\begin{smallmatrix} x_j \\ a_j \end{smallmatrix}\right]}(dx,\, dy)$ is the Dirac measure at point $\left[\begin{smallmatrix} x_j \\ a_j \end{smallmatrix}\right] \in \mathbb{R}^2$.

On the set $S = \mathbb{R}^2$, we define the gauge function $\psi(x, y) = 1 + |x| + |y|$ and consider the measure space $\mathcal{M}_1^{\psi}(\mathbb{R}^2)$, which definition is given above before the formulation of Strassen theorem. We have found the following fact.

Theorem 2.1. (a) *For any probability measure $\mu \in \mathbb{W}$, there is a sequence of discrete probability measures $(\mu_n)_{n=1}^{\infty} \subseteq \mathbb{W}_{\text{disc}}$ that converges to the measure μ in the ψ-weak topology of the space $\mathcal{M}_1^{\psi}(\mathbb{R}^2)$, i.e., for any test function $f \in C_{\psi}(\mathbb{R}^2)$ we have*

$$\int_{\mathbb{R}^2} f d\mu_n \to \int_{\mathbb{R}^2} f d\mu \quad \text{as } n \to \infty.$$

(b) *For any probability measure $\mu \in \mathbb{W}$, there is a sequence of simple probability measures $(\mu_n)_{n=1}^{\infty} \subseteq \mathbb{W}_{\text{simp}}$ that converges to the measure μ in the ψ-weak topology of the space $\mathcal{M}_1^{\psi}(\mathbb{R}^2)$.*

Let us explain why Theorem 2.1 is not trivial. It can be seen from the arguments given in the proof of Theorem 2.1 that the case where the relation 3) (from the definition of the class $\mathbb{W}$) is not equality for all $c \geq 0$ can be reduced to the situation where there is equality in 3) for all $c \geq 0$. However, a discretization of measures from $\mathbb{W}$, for which there is equality in 3) for all $c \geq 0$, may fail to belong to $\mathbb{W}$. Moreover, it is easy to see directly that there are no discrete measures in $\mathbb{W}$, for which equality in 3) holds for all $c \geq 0$, except for the measure $\delta_{(0,0)}$. For example, let $V = 1$. Then equality in 3) for all $c \geq 0$ implies that W has an exponential distribution with the parameter 1. However, the simplest discretization of W in the form

$$\hat{W} := \mathbb{E}[W|W \leq a]\mathbb{1}\{W \leq a\} + \mathbb{E}[W|W > a]\mathbb{1}\{W > a\},$$

where $a > 0$, leads to the distribution of $(V, \hat{W})$ does not belong to the class $\mathbb{W}$. This remark explains why the statement of Theorem 2.1 is not trivial.

In **Chapter 3**, we proved the theorem that the joint distribution of an arbitrary locally integrable increasing process and its compensator at a terminal moment of time can be realized as a joint terminal distribution of some other locally integrable increasing process and its compensator, while the compensator being continuous. We give an exact formulation of this result.

Theorem 3.1. *For any locally integrable process $X^{\circ} = (X_t^{\circ})_{t\in[0;\infty)}$ with a compensator $A^{\circ} = (A_t^{\circ})_{t\in[0;\infty)}$ on some stochastic basis there exists another locally in-*

tegrable increasing process $X^\star = (X^\star_t)_{t\in[0;\infty)}$ *with a compensator* $A^\star = (A^\star_t)_{t\in[0;\infty)}$, *such that*

$$\mathrm{Law}\begin{bmatrix} X^\star_\infty \\ A^\star_\infty \end{bmatrix} = \mathrm{Law}\begin{bmatrix} X^\circ_\infty \\ A^\circ_\infty \end{bmatrix},$$

while the compensator $A^\star$ *is continuous.*

Let us note that Theorem 3.1 itself is not used in proving the main theorems of the dissertation, Theorems 1.1 and 1.3. Nevertheless, this theorem is of independent interest, as will be discussed below. In addition, when substantiating Theorem 3.1, we have obtained an important Lemma 3.3, which, in its turn, is used in proving Theorem 1.1.

Now, let us say a few words about Theorem 3.1. Note that at the moment the description of the class $\mathbb{W}_{\mathrm{loc}}$ of possible distributions of a random vector $(X^\circ_\infty,\, A^\circ_\infty)$ from Theorem 3.1 is unknown. In the integrable case, such a description is available in [4], and it is the class $\mathbb{W}$. Condition 3) from the definition of the class $\mathbb{W}$, as in the integrable case, remains necessary for the measure μ to belong to the class $\mathbb{W}_{\mathrm{loc}}$. The sufficiency of condition 3) is proved under an additional assumption (see condition (3.10) from Proposition 3.6 in [4]), which replaces conditions 1) and 2) from the definition of the class $\mathbb{W}$. But it turns out that this additional assumption is not necessary (see Theorem 1 [2]). We hope that our Theorem 3.1 will be helpful to approach the solution of the problem of describing the class of possible distributions of a random vector $(X^\circ_\infty,\, A^\circ_\infty)$ in the locally integrable case, or at least can help to simplify the solution.

Acknowledgements. The work was written under the supervision of Doctor of Physical and Mathematical Sciences Alexander Alexandrovich Gushchin, to whom the author expresses sincere gratitude for the formulation of the problem and constant support during the writing of the dissertation. The author would also like to thank Valentin Dmitrievich Konakov and other employees of the International Laboratory of Stochastic Analysis and its applications of the Higher School of Economics for useful discussions of the interim results of the dissertation research at scientific seminars and conferences held by the laboratory.

Bibliography

[1] Borzykh D. A. Joint distributions of generalized integrable increasing processes and their generalized compensators. *Theory of Probability and its Applications.* (2022). (in Russian, in press)

[2] Gushchin A. A. On possible relations between an increasing process and its compensator in the non-integrable case. *Russian Mathematical Surveys* **73** (5) (2018). (in Russian)

[3] Gushchin A. A. The Joint Law of a Max-Continuous Local Submartingale and Its Maximum. *Theory of Probability and its Applications.* **65** (4), 693–709 (2020). (in Russian)

[4] Gushchin A. A. The Joint Law of Terminal Values of a Nonnegative Submartingale and Its Compensator. *Theory of Probability and its Applications.* **62** (2), 267–291 (2017). (in Russian)

[5] Jacod J., Shiryaev A. N. *Limit theorems for stochastic processes,* Vol. 1. Berlin., Springer-Verlag (1987).

[6] Kabanov Yu. M., Liptser R. Sh., Shiryaev A. N. Weak and Strong Convergence of the Distributions of Counting Processes. *Theory of Probability and its Applications.* **28** (2), 288–319 (1983). (in Russian)

[7] Marshall A.Ẇ., Olkin I. *Inequalities: theory of majorization and its applicatious.* New York, Academic Press (1979).

[8] Neveu J. *Mathematical foundations of the calculus of probability.* San Francisco., Holden-Day (1965).

[9] Föllmer H., Schied A. *Stochastic finance. An introduction in discrete time.* Berlin–New York., Walter de Gruyter & Co. (2004).

[10] Ash R. B., Doléans-Dade C. A. *Probability & Measure Theory.* Academic Press (1999).

[11] Azema J., Yor M. Une solution simple au problème de Skorokhod. *Séminaire des Probabilités, v. XIII (Univ. Strasbourg, Strasbourg, 1977/1978), Lecture Notes in Math.* **721**, Springer, Berlin, 90–115 (1979).

[12] Azema J., Yor. M. Le problème de Skorokhod: compléments à "Une solution simple au problème de Skorokhod". *Séminaire des Probabilités, v. XIII (Univ. Strasbourg, Strasbourg, 1977/1978), Lecture Notes in Math.* **721**, Springer, Berlin, 625–633 (1979).

[13] Bedini M. L., Buckdahn R., Engelbert H.-J. On the compensator of the default process in an information-based model. *Probability, Uncertainty and Quantitative Risk.* **2** (2017).

[14] Blackwell D. Equivalent comparisons of experiments. *Ann. Math. Statist.* **24**, 265–272 (1953).

[15] Blackwell D., Dubins L. E. A converse to the dominated convergence theorem. *Illinois J. Math.* **7**, 508–514 (1963).

[16] Borzykh D. A. On a property of joint terminal distributions of locally integrable increasing processes and their compensators. *Theory of Stochastic Processes.* **23(39)**:2, 7–20 (2018).

[17] Borzykh D. A., Gushchin A. A. On the denseness of the subset of discrete distributions in a certain set of two-dimensional distributions. *Modern Stochastics: Theory and Applications.* **9** (3), (2022).

[18] Cartier P., Fell J. M. G., Meyer P-A. Comparaison des mesures portées par un ensemble convexe compact. Bulletin de la Société Mathématique de France, T. 92, 435–445 (1964).

[19] Choquet G. Les cônes convexes faiblement complets dans l'analyse. In: Proc. Internat. Congr. Mathematicians (Stockholm 1962) 317–330. Inst. Mittag-Leffler 1963.

[20] Dellacherie C. Un exemple de la théorie générale des processus. In: Séminaire de Probabilités, IV. Lecture Notes in Math., vol. 124, pp. 60–70. Springer (1970).

[21] Dubins L. E., Gilat D. On the distribution of maxima of martingales. *Proc. Amer. Math. Soc.* **68** (3), 337–338 (1978).

[22] Griessler C., Keller-Ressel M. Convex order of discrete, continuous, and predictable quadratic variation and Applications to Options on Variance. *SIAM J. Financial Math.* **5**, 1–19 (2014).

[23] Hardy G. H., Littlewood J. E., Pólya G. *Inequalities.* Cambridge, Cambridge University Press (1934).

[24] Jacod J. Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales. Z. Wahrscheinlichkeitstheor. Verw. Geb. **31**, 235–253 (1975).

[25] Kamae T., Krengel U., O'Brien G. L. Stochastic Inequalities on Partially Ordered Spaces. *Ann. Probab.* **5** (6), 899–912 (1977).

[26] Kellerer H. G. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete* **67**, 399–432 (1984).

[27] Kertz R. P., Rösler U. Martingales with given maxima and terminal distributions. *Israel J. Math.* **69** (2), 173–192 (1990).

[28] Klenke A. *Probability Theory: A Comprehensive Course.* Springer (2014).

[29] Lehmann E. L. Ordered Families of Distributions. *Ann. Math. Statist.* **26** (3), 399–419 (1955).

[30] Lenglart É., Lépingle D., Pratelli M. "Présentation unifiée de certaines inégalités de la théorie des martingales", Séminaire des probabilités, v. XIV (Paris, 1978/1979), Lecture Notes in Math., **784**, Springer, Berlin, 1980, 26–48.

[31] Mann H. B., Whitney D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* **18**, 50–60 (1947).

[32] Müller A., Stoyan D. *Comparison Methods for Stochastic Models and Risks.* Wiley (2002).

[33] Müller A., Rüschendorf L. On the optimal stopping values induced by general dependence structures. *J. Appl. Probab.* **38**, 672–684 (2001).

[34] Ramachandran D., Rüschendorf L. A general duality theorem for marginal problems. *Probab. Theory Related Fields* **101**, 311–319 (1995).

[35] Rogers L. C. G. The joint law of the maximum and terminal value of a martingale. *Probab. Theory Related Fields* **95** (4), 451–466 (1993).

[36] Skala H. J. The existence of probability measures with given marginals. *Ann. Probab.* **21**, 136–142 (1993).

[37] Strassen V. The existence of probability measures with given marginals. *Ann. Math. Statist.* **36**, 423–439 (1965).

[38] Vallois P. Sur la loi du maximum et du temps local d'une martingale continue uniformément intégrable. *Proc. London Math. Soc. (3)* **69** (2), 399–427 (1994).