Without checking the conditions of the type presented in the paper, the recommendation of such a decision rule as described above is based on act of faith only.

Acknowledgements. The authors wish to thank their colleagues, especially Dr Jacek Koronacki and Dr Andrzej Kozek for many helpful discussions and suggestions which led to substantial improvements of the paper.

References

[1] J. R. Barra, *Notions fondamentales de statistique mathématique*, Dunod, Paris 1971.
[2] B. Bednarek-Kozek and A. Kozek, *Stability, sensitivity and sensitivity of characterizations*, Preprint No 99 of the Math. Inst. of the Polish Academy of Sciences; also this volume pp. 39–64.
[3] G. E. P. Box, G. M. Jenkins, *Time series analysis, forecasting and control*, Holden Day 1970.
[4] F. R. Hampel, *A general qualitative definition of robustness*, Ann. Math. Statist. 42 (1971). pp. 1887–1896.
[5] E. L. Lehmann, *Testing statistical hypotheses*, J. Wiley, New York 1959.
[6] Ju. Prohorov, *Convergence of random processes and limit theorems in probability theory* (in Russian), Teor. Veroyatnost. i Primenen. 1 (1956), pp. 177–238.
[7] V. Strassen, *The existence of probability measures with given marginals*, Ann. Math. Statist. 36 (1965), pp. 423–439.
[8] V. M. Zolotarev, *On continuity of stochastic sequences generated by recurrent procedures* (in Russian), Teor. Veroyatnost. i Primenen. 20 (1976), pp. 834–847.

Presented to the semester
MATHEMATICAL STATISTICS
September 15–December 18, 1976

MATHEMATICAL STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 6
PWN—POLISH SCIENTIFIC PUBLISHERS
WARSAW 1980

STABILITY, SENSITIVITY AND SENSITIVITY OF CHARACTERIZATIONS

BOGUSŁAWA BEDNAREK-KOZEK

*Institute of Mathematics, Wrocław University, Wrocław, Poland**

and

ANDRZEJ KOZEK

*Institute of Mathematics, Polish Academy of Sciences, Wrocław Branch, Wrocław, Poland***

0. Introduction

The term "stability" has a long history. It has been used by Lagrange, Poisson, Poincaré, and Liapunov in problems of mechanics. Ulam [14] discussed the notion of the stability of mathematical theorems from a rather general point of view: "When is it true that by changing 'a little' the hypothesis of a theorem one can still assert that the thesis of the theorem remains true or 'approximately' true?". Ulam decided not to formulate a generally applicable definition of stability and we do not try to do it here, either. However, a review of theorems on the stability shows that there are groups of problems in which the stability can be treated from the same point of view. Here we restrict ourselves to three types of "stabilities" which are related to some properties of transformations of metric spaces. We call them $\delta(\varepsilon)$-*stability*, $\delta^{-1}(\varepsilon)$-*sensitivity* and $\gamma(\varepsilon)$-*sensitivity of characterizations*, respectively. The present paper is strongly inspired by lectures of V. M. Zolotarev given in 1976 in Varna and Warsaw on his approach published in papers [15], [16] and [17]. In particular we use the set-theoretical model of $\gamma(\varepsilon)$-sensitivity of characterizations given in [16] and [17].

Let (X, ϱ_X) and (Y, ϱ_Y) be metric spaces and let f be a function from X into Y. We are concerned with functions which have one of the following properties:

I. $\delta(\varepsilon)$-*stability*

$$f(C^{\delta(\varepsilon)}) \subset f(C)^{\varepsilon]}$$

(intuitively: *similar reasons have similar consequences*);

* Plac Grunwaldzki 2/4, 50-384 Wrocław.
** Kopernika 18, 51-617 Wrocław.

II. $\delta^{-1}(\varepsilon)$-*sensitivity*

$$f^{-1}(C^{\delta(\varepsilon)}) \subset f^{-1}(C)^{\varepsilon]}$$

(intuitively: *similar consequences must have similar reasons*);

III. $\delta^{-1}(\varepsilon)$-*sensitivity of characterizations*

$$A^{\delta(\varepsilon)} \cap f^{-1}(B^{\delta(\varepsilon)}) \subset C^{\varepsilon]},$$

where

$$C = A \cap f^{-1}(B)$$

(intuitively: *a slight deviation from "assumptions A" and, simultaneously, a slight deviation from a considered "property B" may occur only on a neighbourhood of the characterized set C*).

We are interested in a quantitative approach to investigations of these properties. We present necessary and sufficient conditions for f to be $\delta(\varepsilon)$-stable, for f to be $\delta^{-1}(\varepsilon)$-sensitive and for a characterization of C by A, B and f to be $\delta^{-1}(\varepsilon)$-sensitive. Next, we apply the obtained results in cases when f is a transformation from a set of probability measures into another set of probability measures. Finally, we give some applications of the obtained results to statistics (cf. [5], [6]).

As is easy to see (cf. Proposition 2.1) $\delta(\varepsilon)$-stability of transformation f with respect to a family of subsets $\boldsymbol{C}_X$ becomes equivalent to $\delta(\varepsilon)$-uniform continuity of f in the extremal case when $\boldsymbol{C}_X$ consists of all subsets of X. Another and more important extremal case is when $\boldsymbol{C}_X$ consists of one subset of X. For example in mechanics, $\boldsymbol{C}_X$ may consist of the set of all stable points whereas in the theory of probability or in statistics $\boldsymbol{C}_X$ may consist of the set of all gaussian or all infinitely divisible distributions (cf. [3], p. 1572). In order that f be $\delta(\varepsilon)$-stable with respect to a given set it is enough that f be "$\delta(\varepsilon)$-uniformly continuous on the neighbourhood of this set", only. More precisely this is formulated in Proposition 2.4. Clearly, if f is $\delta(\varepsilon)$-uniformly continuous, it is $\delta(\varepsilon)$-stable with respect to any family of subsets $\boldsymbol{C}_X$ (Proposition 2.1a). On the other hand, the continuity of f does not, in general, imply $\delta(\varepsilon)$-stability of f with respect to a given set even, $\delta(\varepsilon)$-uniform continuity (at least local) of f is essential for this.

For example, Zolotarev's perfect or ideal metrics yield uniformities especially convenient for investigations of $\delta(\varepsilon)$-stabilities in the case where $\boldsymbol{C}_X$ consists of sets of distributions of random vectors which are invariant with respect to translations of random vectors and its multiplications by numbers. Normal and infinitely divisible distributions are of this type. So, this and ε-uniform continuity of convolution with respect to an ideal metric are the most important reasons of very elegant results of Zolotarev on approximation of distributions of sums of independent random variables ([17], Section 2.2).

In some cases it is possible to attain $\delta(\varepsilon)$-continuity of f by changing the metric on X (more generally, by changing uniformity $U_{\rho X}$ on X). Therefore, we formulate in Proposition 2.3 a necessary and sufficient condition for f to be $\delta(\varepsilon)$-uniformly continuous on X for a given metric ϱ_X. Moreover, Proposition 2.3 implies that there always exists a pseudometric ϱ_X^f on X such that f is ε-uniformly continuous. Section 2 and the discussion above supplement the program of investigations of stochastic continuity formulated by Zolotarev in [15] and [17]. Proposition 2.3 supplements Zolotarev's discussion on the intuitive choice of a metric essential for a given problem.

Moreover, the use of $\delta(\varepsilon)$-stability inclusion and the definition of the Prokhorov distance immediately yield Theorem 5.2 which was obtained by Zolotarev ([15], Theorems 1 and 15; [17], Theorem 3) by much more deep considerations on minimal metrics. Similarly, Theorem 5.5, which is another important but particular case of Zolotarev's Theorem 3 in [17], admits a trivial proof. More detailed discussion of it is given below.

It seems that there are some consequences of Theorems 5.2 and 5.3 interesting for statisticians:

(1) ε-uniform continuity of the transformation generated by the sample mean in the spaces of distributions (see the end of Example 4 in Section 7) should be compared with the non-robust results of Hampel ([5], [6]);

(2) ε-uniform continuity of transformation generated by the standard deviation and, in view of Theorem 5.3, the lack of this property in the case of transformations generated by S^2 (see the end of Example 4, Section 7) should be compared with the old discussion of statisticians on the "good" properties of standard deviation and "bad" properties of S^2.

$\delta^{-1}(\varepsilon)$-sensitivity is convenient for a formulation of "inverse problems of stability" and in the simplest case where f is bijective it reduces to $\delta^{-1}(\varepsilon)$-stability of f^{-1}. Moreover, the use of $\delta^{-1}(\varepsilon)$-sensitivity of the transformation f leads to a necessary and sufficient condition for the $\gamma(\varepsilon)$-sensitivity of characterizations (Theorem 4.4). This last notion has firstly been formulated and investigated in [16] and [17] by Zolotarev who, outside of the incorrect Theorem 1 in [16], investigated continuous transformations, only. The use of $\delta^{-1}(\varepsilon)$-sensitivity of f instead of its continuity leads to a better understanding of the problem and shows the role of "shapes" of the sets A and $f^{-1}(B)$.

In Section 1 we recall the definitions of uniform spaces ([4]), proper regular conditional probabilities ([9], [10]) and, moreover, we introduce some terms and symbols which will be used in the sequel.

In Section 2 we give a precise definition of $\delta(\varepsilon)$-stability and we show that uniformities yield the most convenient tool for a quantitative theory of stability of transformations. This is not surprising because it is well known that uniform spaces are more appropriate for investigations of neighbourhoods of sets than topology which deals with neighbourhoods of points.

In Section 3 we give a precise definition of $\delta^{-1}(\varepsilon)$-sensitivity. We show that the sensitivity is related to the uniformity of an inverse transformation into a quotient space. Theorems given in Sections 2 and 3 show relations of $\delta(\varepsilon)$-stability and $\delta^{-1}(\varepsilon)$-sensitivity to another concepts.

Section 4 contains a precise definition of $\gamma(\varepsilon)$-sensitivity of characterizations.

This is a reformulation of the definition given by Zolotarev in [16]. Theorem 4.1 is a reformulation of Theorem 2 in [16] in a somewhat more general framework. Theorems 4.2 and 4.3 may be considered as substitutes for incorrect Theorem 1 in [16]. Moreover, Theorem 4.3 yields a decomposition of the sensitivity of characterizations into two parts: into the sensitivity of a transformation and into a part which depends on the "shapes" and mutual positions of "the set of assumptions A" and "the set of properties B". Finally, in Theorem 4.4 we present a necessary and sufficient condition for a characterization to be $\gamma(\varepsilon)$-sensitive.

The aim of Section 5 is to investigate the stability of transformations of spaces of measures. In view of Proposition 2.1 we restrict ourselves to investigations of uniformities of transformations of spaces of measures. Thus, Theorem 5.2 yields sufficient conditions for $\delta(\varepsilon)$-uniformity of a transformation of spaces of measures endowed with uniformities of the Prokhorov distances. Theorem 5.3 shows that if the domain of the transformation is large enough, then this condition is also necessary. Theorems 5.4–5.6 yield sufficient conditions for the $\delta(\varepsilon)$-uniformity of transformations of spaces of measures in the cases of metrics ϱ_{BL*}, ϱ_{B*}, and ϱ_M, respectively (cf. Section 1 for the definitions of these metrics).

In view of Strassen's theorem on a characterization of the Prokhorov distance [13], Theorem 5.2 follows from Theorems 1 and 12 announced in [15] and proved in [17], provided metric spaces are Polish (cf. [3] in the separable case). In this paper we need no assumptions on the metric spaces. Moreover, our proof of Theorem 5.2 is a direct one, short and simpler than, e.g., the proofs of Strassen or Dudley's theorems. Similarly, in the case of metric spaces Theorem 5.5 follows from a theorem of Dobrushyn [2] and Theorems 1 and 12 in [15]. Our proof of Theorem 5.5 is trivial and leads to the best estimation of $\delta(\varepsilon)$. A theorem similar to Theorem 5.2 has been independently proved in [1] by Bartoszyński and Pleszczyńska.

In Section 6 we note that if a transformation f of metric spaces is an isomorphism, then the corresponding theorems on $\gamma(\varepsilon)$-sensitivity follows from theorems of Section 5 applied to f^{-1}. Theorem 6.1 establishes the ε-sensitivity of transformations of spaces of measures for arbitrary measurable f.

Section 7 contains simple examples which illustrate the use of the results obtained in Sections 2 and 5. Similarly, Section 8 illustrate the use of the results obtained in Sections 3, 4 and 6.

Acknowledgements. The authors would like to thank:

Professor R. Bartoszyński and Dr E. Pleszczyńska for rendering a manuscript of their's unpublished paper accessible for us and for a discussion,

Dr W. Klonecki for his remarks and suggestions,

Dr J. Koronacki for a discussion,

Dr R. Zieliński who called our attention to the approach presented in Propositions 2.4 and 3.3,

Professor V. M. Zolotarev who called our attention to the relations between Theorems 1 and 12 in [15] and Theorems 5.2 and 5.5 and who rendered his unpublished papers accessible for us.

1. Preliminaries

In this section we remind some definitions and simple properties of uniform spaces and proper regular conditional probabilities.

Let X be a non-empty set. A vicinity of the diagonal Δ is a set $V \subset X\times X$ such that $\Delta \subset V$ and $(x_1, x_2) \in V$ implies $(x_2, x_1) \in V$. Let $\mathscr{D}_X$ be the family of all vicinities of the diagonal. If $V \in \mathscr{D}_X$, then $2V$ is an element of $\mathscr{D}_X$ given by

$$2V = \{(x_1, x_2)\colon \text{there exists an } x_3 \in X \text{ such that } (x_1, x_3) \in V \text{ and } (x_3, x_2) \in V\}.$$

U_X is called a *uniformity* on X if $U_X \subset \mathscr{D}_X$ and if the following conditions are fulfilled

U1. *if* $V \in U_X$ *and* $V \subset W \in \mathscr{D}_X$, *then* $W \in U_X$;

U2. *if* $V_1, V_2 \in U_X$, *then* $V_1 \cap V_2 \in U_X$;

U3. *for each* $V \in U_X$ *there exists a* $W \in U_X$ *such that* $2W \subset V$;

U4. $\bigcap_{V \in U_X} V = \Delta$.

A pair (X, U_X) is called a *uniform space*.

Let (X, ϱ_X) be a metric space and $V_\varepsilon = \{(x_1, x_2) \in X\times X\colon \varrho_X(x_1, x_2) < \varepsilon\}$. The family $\{V_\varepsilon\}$ is a base of a uniformity $U(\varrho_X)$, where

$$U(\varrho_X) = \{V \in \mathscr{D}_X\colon \text{there is a } V_\varepsilon \text{ such that } V_\varepsilon \subset V\}.$$

If ϱ_X is a pseudo-metric only and $U(\varrho_X)$ is defined as above, then condition U4 need not be fulfilled. In such a case we shall use terms "pseudo-uniformity" and "pseudo-uniform space" instead of "uniformity" and "uniform space", respectively.

Let (X, U_X) and (Y, U_Y) be pseudo-uniform spaces and let $f\colon X \to Y$. The function f is said to be *uniform* if for each $V \in U_Y$ there exists a $U \in U_X$ such that $(f(x_1), f(x_2)) \in V$ whenever $(x_1, x_2) \in U$. If $U_X = U(\varrho_X)$ and $V = V(\varrho_Y)$, then f is uniform if and only if for each $\varepsilon > 0$ there is a $\delta > 0$ such that $\varrho_X(x_1, x_2) < \delta$ implies $\varrho_Y(f(x_1), f(x_2)) < \varepsilon$.

If U_X and U'_X are two pseudo-uniformities on X, then U_X is said to be *weaker than* U'_X whenever $U_X \subset U'_X$. The identical embedding of (X, U'_X) into (X, U_X) is uniform if and only if U_X is weaker than U'_X. Hence $U(\varrho_X)$ is weaker than $U(\varrho'_X)$ if and only if there exists a non-decreasing function $\delta(\varepsilon)$ such that $\delta(\varrho_X) \leqslant \varrho'_X$ ([4]).

Let $(X, \mathscr{A}, \mu)$ be a probability space. If $\mathscr{A}_1$ and $\mathscr{A}_2$ are two sub-σ-algebras of $\mathscr{A}$, then $\mathscr{A}_1 \sim \mathscr{A}_2$ means that for each $A_1 \in \mathscr{A}_1$ there exists an $A_2 \in \mathscr{A}_2$ such that $\mu(A_1 \triangle A_2) = 0$ and for each $A_2 \in \mathscr{A}_2$ there exists $A_1 \in \mathscr{A}_1$ such that

$\mu(A_1 \triangle A_2) = 0$. A proper regular conditional probability (p.r.c.p.) on a σ-algebra $\mathscr{A}_0$ given a σ-algebra $\mathscr{A}_1$ and probability measure μ is a function $\pi(\cdot, \cdot; \mathscr{A}_1, \mu)$: $X \times \mathscr{A}_0 \to [0, 1]$ such that the following conditions are fulfilled:

CP1. *$\pi(x, \cdot; \mathscr{A}_1, \mu)$ is a probability measure on $\mathscr{A}_0$ for all $x \in X$;*

CP2. *$\pi(\cdot, A; \mathscr{A}_1, \mu)$ is $\mathscr{A}_1$-measurable for all $A \in \mathscr{A}_0$;*

CP3. *$\int_{A_1} \pi(A, x; \mathscr{A}_1, \mu)\mu(dx) = \mu(A \cap A_1)$ for all $A \in \mathscr{A}_0$ and $A_1 \in \mathscr{A}_1$;*

CP4. *$\pi(x, A_1; \mathscr{A}_1, \mu) = \chi_{A_1}(x)$ for all $A_1 \in \mathscr{A}_1 \cap \mathscr{A}_0$ and $x \in X$.*

In the sequel we shall use the expression "$\pi(x, A; \mathscr{B}, \mu)$ is a p.r.c.p. given $\mathscr{B}$ and μ" provided there exist $\mathscr{A}_0 \sim \mathscr{A}$ and $\mathscr{A}_1 \sim \mathscr{B}$ such that $\pi(x, A; \mathscr{B}, \mu)$ is a p.r.c.p. on $\mathscr{A}_0$ given $\mathscr{A}_1$ and μ.

The existence of p.r.c.p.'s has been investigated in papers of Musiał [9] and Pfanzagl [10]. It is known, e.g., that if $\mathscr{A}$ is the μ-completion of a separable σ-algebra, $\mathscr{A}_1$ is a σ-algebra containing all μ-null sets from $\mathscr{A}$ and μ is either purely atomic or $\mathscr{A}$ contains a μ-null set of the cardinality c, then a p.r.c.p. given $\mathscr{B}$ and μ exists ([9], [10]).

Finally, we recall definitions of metrics on sets of probability (or, respectively, finite non-negative) measures which will be used in the sequel. Let (X, ϱ_X) be a metric space and μ, ν finite non-negative measures on Borel subsets of (X, ϱ_X).

The *Prokhorov metric* is given by

$$\varrho_P(\mu, \nu) = \max\big(\sigma(\mu, \nu), \sigma(\nu, \mu)\big), \tag{1.1}$$

where

$$\sigma(\mu, \nu) = \inf\{\varepsilon > 0\colon \mu(C) \leqslant \mu(C^\varepsilon) + \varepsilon \text{ for all closed } C \subset X\}. \tag{1.2}$$

We may replace C^ε by $C^{\varepsilon]}$ in the definition of σ without changing its value. Also, we may replace "all closed C" by "all Borel sets B" since if C is a closure of B, $C^\varepsilon = B^\varepsilon$ and $C^{\varepsilon]} = B^{\varepsilon]}$. Here

$$C^\varepsilon = \{x \in X\colon \varrho_X(x, C) < \varepsilon\} \quad \text{and} \quad C^{\varepsilon]} = \{x \in X\colon \varrho_X(x, C) \leqslant \varepsilon\}.$$

The Prokhorov metric metirizes the weak-star topology on the space of all probability measures on X for X separable ([3]; cf. also [11]).

Let $\mathrm{BL}(X, \varrho_X)$ denote the set of all bounded real-valued functions f on S which are Lipschitzian, i.e.

$$\|f\|_{\mathrm{BL}} = \sup\{|f(x)|;\ x \in X\} + \sup\{|f(x_1) - f(x_2)|/\varrho_X(x_1, x_2);\ x_1, x_2 \in X,\ \varrho_X(x_1, x_2) \neq 0\} < \infty.$$

Then

$$\|\mu\|_{\mathrm{BL}^*} = \sup\{|\textstyle\int f\mu(dx)|\colon \|f\|_{\mathrm{BL}} \leqslant 1\}.$$

The metrics

$$\varrho_{\mathrm{BL}^*}(\mu, \nu) = \|\mu - \nu\|_{\mathrm{BL}^*} \tag{1.3}$$

and ϱ_P define the same uniformity on the space of probability measures for X metric and separable. Hence, both the metrics define the same topology. We have

$$\varrho_{\mathrm{BL}^*}(\mu, \nu) \leqslant 2\varrho_P(\mu, \nu), \tag{1.4}$$

$$\varrho_P(\mu, \nu) \leqslant \left(\frac{3}{2}\varrho_{\mathrm{BL}^*}(\mu, \nu)\right)^{1/2} \tag{1.5}$$

(see [3]).

Now, let $(X, \mathscr{B}(X))$ be an arbitrary measure space and let μ, ν be probability measures on $\mathscr{B}(X)$. Then

$$\begin{aligned} \varrho_{B^*}(\mu, \nu) &= \max\{(\mu-\nu)^+(X), (\mu-\nu)^-(X)\} \\ &= \sup\{|(\mu-\nu)(A)|\colon A \in \mathscr{B}(X)\}. \end{aligned} \tag{1.6}$$

Now, let $X = \boldsymbol{R}^n$ and let $\mathscr{B} = \mathscr{B}(\boldsymbol{R}^n)$ be the σ-algebra of Borel sets. A *Meshalkin metric* ϱ_M is given by

$$\varrho_M(\mu, \nu) = \sup\{|(\mu-\nu)(A)|\colon A \text{ is an intersection of at most } n \text{ half-spaces}\},$$

where the "half-space" is an arbitrary set of points in $\boldsymbol{R}^n$ which is of the form $\{(x_1, \ldots, x_n) \in \boldsymbol{R}^n\colon \sum_{i=1}^n a_i x_i < b\}$, $a_i, b \in \boldsymbol{R}$ (see [8]).

The *Levy distance of probability measures* on $\mathscr{B}(\boldsymbol{R}^n)$ is given by

$$\varrho_L(\mu, \nu) = \max\big(\sigma_L(\mu, \nu), \sigma_L(\nu, \mu)\big), \tag{1.7}$$

where

$$\sigma_L(\mu, \nu) = \inf\{\varepsilon > 0\colon \mu\big((-\infty, x]\big) \leqslant \nu\big((-\infty, x]^\varepsilon\big) + \varepsilon \text{ for all } x \in \boldsymbol{R}^n\}.$$

The *Kolmogorov distance of probability measures* on $\mathscr{B}(\boldsymbol{R}^n)$ is given by

$$\varrho_K(\mu, \nu) = \sup\{|\mu(A) - \nu(A)|\colon A = (-\infty, x],\ x \in \boldsymbol{R}^n\}. \tag{1.8}$$

If $n = 1$, then the *Sibley metric* ϱ_S is given by

$$\varrho_S(\mu, \nu) = \max\big(\sigma_S(\mu, \nu), \sigma_S(\nu, \mu)\big), \tag{1.9}$$

where

$$\sigma_S(\mu, \nu) = \inf\left\{\varepsilon > 0\colon \mu\big((-\infty, x]\big) \leqslant \nu\big((-\infty, x+\varepsilon]\big) + \varepsilon \text{ for all } x \in \left(-\varepsilon - \frac{1}{\varepsilon}, \frac{1}{\varepsilon}\right)\right\}. \tag{1.10}$$

ϱ_S metrizes the pointwise convergence of distribution functions of sub-probability measures at each continuity point of the limit distribution function (see [12]).

In the sequel, whenever we deal with one of the metrics listed above, we assume tacitly that the assumptions on $(X, \mathscr{B}(X))$ indicated in an appropriate definition are fulfilled.

2. Stable functions

Let (X, ϱ_X) and (Y, ϱ_Y) be metric spaces with metrics ϱ_X and ϱ_Y, respectively, and let f be a function from X into Y. Denote by $\boldsymbol{C}_X$ a family of non-empty subsets of X. If $x \in X$ and $C \subset X$, then

$$\varrho_X(x, C) = \inf\{\varrho_X(x, x_1);\ x_1 \in C\},$$
$$C^{\varepsilon} = \{x \in X\colon\ \varrho_X(x, C) < \varepsilon\},$$
$$C^{\varepsilon]} = \{x \in X\colon\ \varrho_X(x, C) \leqslant \varepsilon\}$$

and

$$\operatorname{diam} \varrho_X(C) = \sup\{\varrho_X(x_1, x_2);\ x_1, x_2 \in C\}.$$

An analogous notation will also be used for metric spaces different from X. Let $\delta(\varepsilon) > 0$.

DEFINITION 1. A function f is called *$\delta(\varepsilon)$-stable with respect to $\boldsymbol{C}_X$* if

$$f(C^{\delta(\varepsilon)}) \subset f(C)^{\varepsilon]} \big(= (f(C))^{\varepsilon]}\big)$$

holds for every $C \in \boldsymbol{C}_X$ and $\varepsilon > 0$. If $\boldsymbol{C}_X$ is the family of all subsets of X and f is $\delta(\varepsilon)$-stable with respect to $\boldsymbol{C}_X$, then f is called *$\delta(\varepsilon)$-stable.*

Roughly speaking, property "to be stable" means that "similar reasons" lead to "similar consequences", where arguments of f are interpreted as "reasons" and values of f are interpreted as "consequences".

A function f from $(X, U(\varrho_X))$ into $(Y, U(\varrho_Y))$ is called *$\delta(\varepsilon)$-uniform* if $\varrho_X(x_1, x_2) < \delta(\varepsilon)$ implies $\varrho_Y(f(x_1), f(x_2)) \leqslant \varepsilon$.

PROPOSITION 2.1.a. *If f is a $\delta(\varepsilon)$-uniform transformation from $(X, U(\varrho_X))$ into $(Y, U(\varrho_Y))$, then f is $\delta(\varepsilon)$-stable with respect to each non-empty family $\boldsymbol{C}_X$.*

b. *If f is $\delta(\varepsilon)$-stable with respect to $\boldsymbol{C}_X$ satisfying the following condition*:

C1. *for each $\alpha > 0$ and each $x \in X$ there is a $C \in \boldsymbol{C}_X$ such that $x \in C$ and* $\operatorname{diam} \varrho_Y(f(C)) < \alpha$,

then f is $\delta(\varepsilon)$-uniform.

Proof. Suppose $C \in \boldsymbol{C}_X$ and $x \in C^{\delta(\varepsilon)}$. Then there is an $x_0 \in C$ such that $\varrho_X(x, x_0) < \delta(\varepsilon)$ and the $\delta(\varepsilon)$-uniformity of f yields $\varrho_Y(f(x), f(x_0)) \leqslant \varepsilon$. Hence $f(x) \in f(C)^{\varepsilon]}$.

Now we prove the second part of the proposition. Let $\varrho_X(x_1, x_2) < \delta(\varepsilon)$ and $\gamma > 0$. Take a C, $C \in \boldsymbol{C}_X$, such that $x_2 \in C$ and $\operatorname{diam} \varrho_Y(f(C)) < \gamma$. By the $\delta(\varepsilon)$-stability of f there is an $x_0 \in C$ such that $\varrho_Y(f(x_1), f(x_0)) \leqslant \varepsilon$. Thus, we have

$$\varrho_Y(f(x_1), f(x_2)) \leqslant \varrho_Y(f(x_1), f(x_0)) + \varrho_Y(f(x_2), f(x_0)) \leqslant \varepsilon + \gamma.$$

Since γ is an arbitrary positive number, we have $\varrho_Y(f(x_1), f(x_2)) \leqslant \varepsilon$, i.e., f is $\delta(\varepsilon)$-uniform.

Remark 1. Clearly, if $\boldsymbol{C}_X$ contains all one-point subsets of X, then Condition C1 is fulfilled. In particular, if $\boldsymbol{C}_X$ is the family of all subsets of X, then the $\delta(\varepsilon)$-uniformity and $\delta(\varepsilon)$-stability of f are equivalent.

Remark 2. We can replace "$<$" by "$\leqslant$" or "$\leqslant$" by "$<$" in the definition of the $\delta(\varepsilon)$-uniform transformation. However, in such a case, we should replace in Proposition 2.1 "open neighbourhood" by "closed neighbourhood" or, respectively, "closed neighbourhood" by "open" one. In the sequel, the term "$\delta(\varepsilon)$-uniform transformation" will always be used in accordance with the definition given at the beginning of this section.

It is convenient to state as a proposition the following obvious result. Let (X_1, ϱ_{X_1}) and (Y_1, ϱ_{Y_1}) be metric spaces. Denote by g and h functions from X_1 into X and from Y into Y_1, respectively. Finally, let $t = h \circ f \circ g$.

PROPOSITION 2.2. *Let f be $\delta(\varepsilon)$-uniform. If*

(2.1) $$\varrho_{X_1}(x_1', x_1'') < \delta_X(\varepsilon) \quad \textit{implies} \quad \varrho_X(g(x_1'), g(x_1'')) < \varepsilon$$

and

(2.2) $$\varrho_Y(y', y'') \leqslant \delta_Y(\varepsilon) \quad \textit{implies} \quad \varrho_{Y_1}(h(y'), h(y'')) \leqslant \varepsilon$$

for each $\varepsilon > 0$, $x_1', x_1'' \in X_1$ and $y', y'' \in Y$, then t is $\gamma(\varepsilon)$-uniform, where $\gamma = \delta_X \circ \circ\, \delta \circ \delta_Y$.

This proposition gives possibilities for obtaining new results on $\delta(\varepsilon)$-stability of a transformation provided one such result is known. For example, when $X_1 = X$, $Y_1 = Y$ and the metrics ϱ_X and ϱ_Y are changed to ϱ_{X_1} and ϱ_{Y_1}, respectively, then Proposition 2.2 yields at once the "new" $\delta(\varepsilon)$. If $X = X_1$, then condition (2.1) is fulfilled for a δ_X if and only if $U(\varrho_X) \subset U(\varrho_{X_1})$ or, equivalently, if for a non-decreasing function δ_X we have $\delta_X(\varrho_X(x', x'')) \leqslant \varrho_{X_1}(x', x'')$.

Similarly, if $Y = Y_1$, then (2.2) is fulfilled for a δ_Y if and only if $U(\varrho_{Y_1}) \subset U(\varrho_Y)$ or, equivalently, if

$$\delta_Y(\varrho_{Y_1}(y', y'')) \leqslant \varrho_Y(y', y'') \quad \text{for a non-decreasing function } \delta_Y.$$

Denote by $\boldsymbol{X}$ a quotient space $X/(f)$, where (f) is an equivalence relation given by

(2.3) $$x_1 (f) x_2 \quad \text{if and only if} \quad f(x_1) = f(x_2).$$

Let $\boldsymbol{f}$ be a function from $\boldsymbol{X}$ into Y given by

(2.4) $$\boldsymbol{f}(x) = f(x) \quad \text{for any } x \in \boldsymbol{x}.$$

Let $\varrho_{\boldsymbol{X}}^f$ be a metric in $\boldsymbol{X}$ given by

$$\varrho_{\boldsymbol{X}}^f(\boldsymbol{x}_1, \boldsymbol{x}_2) = \varrho_Y(\boldsymbol{f}(\boldsymbol{x}_1), \boldsymbol{f}(\boldsymbol{x}_2)).$$

Clearly, we have

PROPOSITION 2.3. *A function f is a $\delta(\varepsilon)$-uniform transformation from $(X, U(\varrho_X))$ into $(Y, U(\varrho_Y))$ if and only if the canonical surjection $i_X\colon X \to \boldsymbol{X}$ is a $\delta(\varepsilon)$-uniform transformation from $(X, U(\varrho_X))$ into $(\boldsymbol{X}, U(\varrho_{\boldsymbol{X}}^f))$.*

Proof. Indeed, it is enough to note that $f = \boldsymbol{f} \circ i_X$, $i_X = \boldsymbol{f}^{-1} \circ f$ and use the fact that $\boldsymbol{f}$ is an isometry from $(\boldsymbol{X}, \varrho_{\boldsymbol{X}}^f)$ onto $(f(X), \varrho_Y)$.

COROLLARY 2.3.1. *$U(\varrho_X^f)$ is the weakest uniformity on $\boldsymbol{X}$ such that $\boldsymbol{f}$ is a uniform function from $\boldsymbol{X}$ into Y.*

Let 2^X be the family of all subsets of X. Denote by ϱ_{2^X} the Hausdorff pseudo-metric in 2^X given by

$$\varrho_{2^X}(C_1, C_2) = \max(\alpha_1, \alpha_2),$$

where

$$\alpha_i = \sup\{\varrho_X(x, C_i);\ x \in C_j,\ i \neq j\}, \quad i, j = 1, 2 \text{ (cf. [7]).}$$

Denote by $\hat{f}$ a function from 2^X into 2^Y given by $\hat{f}(C) = f(C)$, $C \subset X$. The function $\hat{f}$ is said to be *$\delta(\varepsilon)$-uniform from the right* on $\boldsymbol{C}_X$ if

$$C \in \boldsymbol{C}_X, \quad C_1 \supset C, \quad \varrho_{2^X}(C_1, C) < \delta(\varepsilon) \quad \text{implies} \quad \varrho_{2^Y}(\hat{f}(C_1), \hat{f}(C)) \leqslant \varepsilon.$$

PROPOSITION 2.4. *The function f is $\delta(\varepsilon)$-stable with respect to $\boldsymbol{C}_X$ if and only if $\hat{f}$ is $\delta(\varepsilon)$-uniform from the right on $\boldsymbol{C}_X$.*

Proof. Let f be $\delta(\varepsilon)$-stable. Then $C_1 \supset C$ and $\varrho_{2^X}(C_1, C) < \delta(\varepsilon)$ yield $C \subset C_1 \subset C^{\delta(\varepsilon)}$. By the $\delta(\varepsilon)$-stability of f we obtain

$$f(C) \subset f(C_1) \subset f(C^{\delta(\varepsilon)}) \subset f(C)^{\varepsilon]},$$

i.e., $\varrho_{2^Y}(\hat{f}(C_1), \hat{f}(C)) \leqslant \varepsilon$ provided $C \in \boldsymbol{C}_X$.

If $\hat{f}$ is $\delta(\varepsilon)$-uniform from the right on $\boldsymbol{C}_X$, then $\varrho_{2^Y}(\hat{f}(C^{\delta(\varepsilon)}), \hat{f}(C)) \leqslant \varepsilon$, i.e., $f(C^{\delta(\varepsilon)}) \subset f(C)^{\varepsilon]}$ provided $C \in \boldsymbol{C}_X$.

3. Sensitive functions

Let us start with the following definition. Let $\boldsymbol{C}_Y$ be a non-empty family of non-empty subsets of Y and let $\delta(\varepsilon)$ stand for a non-decreasing left-continuous positive function. By $\delta^{-1}(\varepsilon)$ we denote a function given by

$$\delta^{-1}(\varepsilon) = \inf\{\alpha > 0\colon\ \delta(\alpha) > \varepsilon\},$$

i.e., $\delta^{-1}(\varepsilon)$ is the right-continuous inverse of $\delta(\varepsilon)$.

DEFINITION 2. A function f from a metric space (X, ϱ_X) onto a metric space (Y, ϱ_Y) is called *$\delta^{-1}(\varepsilon)$-sensitive with respect to a family $\boldsymbol{C}_Y$* if

$$f^{-1}(C^{\delta(\varepsilon)}) \subset f^{-1}(C)^{\varepsilon]} \left(= (f^{-1}(C))^{\varepsilon]}\right)$$

holds for every set $C \in \boldsymbol{C}_Y$ and $\varepsilon > 0$. If $\boldsymbol{C}_Y$ is the family of all subsets of Y and f is $\delta^{-1}(\varepsilon)$-sensitive with respect to $\boldsymbol{C}_Y$, then f is called *$\delta^{-1}(\varepsilon)$-sensitive.*

It is easy to see that if f is $\delta^{-1}(\varepsilon)$-sensitive with respect to $\boldsymbol{C}_Y$, then $f^{-1}(C^{\varepsilon}) \subset f^{-1}(C)^{\delta^{-1}(\varepsilon)]}$ holds for every $C \in \boldsymbol{C}_Y$.

In order to avoid some pathological cases we assume in Definition 2 that f is onto Y. For example, let f be the identity embedding of $X = [0, 1)$ into $Y = \boldsymbol{R}$. If $C = \{0, 1\}$, then $f^{-1}(C^{\delta})$ is not a subset of $f^{-1}(C)^{\varepsilon]}$ whatever be $\delta > 0$ and $\varepsilon < 1$. Thus the identity embedding would not be sensitive. Let us note that if we change the assumption "f is onto Y" for "$\boldsymbol{C}_Y$ is a family of subsets of $f(X)$"; then we obtain an equivalent definition of the $\gamma(\varepsilon)$-sensitivity.

Roughly speaking, property "to be sensitive" means that "similar consequences" must be caused by "similar reasons", provided arguments of f are interpreted as "reasons" and values of f are interpreted as "consequences".

Let $\boldsymbol{X}$ be the quotient space $X/(f)$ defined in Section 2. We define in $\boldsymbol{X}$ a pseudo-metric $\varrho_{\boldsymbol{X}}$ given by

$$\varrho_{\boldsymbol{X}}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \max(\alpha_1, \alpha_2), \tag{3.1}$$

where

$$\alpha_i = \sup\{\varrho_X(x, \boldsymbol{x}_i);\ x \in \boldsymbol{x}_j,\ i \neq j\}, \quad i, j = 1, 2. \tag{3.2}$$

Clearly, $\varrho_{\boldsymbol{X}}$ is the Hausdorff pseudo-metric restricted from the space of all subsets of X into $\boldsymbol{X}$. In the sequel a pseudo-metric in $\boldsymbol{X}$ will always be $\varrho_{\boldsymbol{X}}$ given by (3.1)–(3.2).

Let $\boldsymbol{f}$ be given by (2.4). Clearly, $\boldsymbol{f}$ is a one-to-one function from $\boldsymbol{X}$ onto Y. $\boldsymbol{f}^{-1}$ is the inverse function to $\boldsymbol{f}$, i.e., $\boldsymbol{f}^{-1}\colon Y = f(X) \to \boldsymbol{X}$ and $\boldsymbol{f}^{-1}(y) = \boldsymbol{x}$ if and only if $\boldsymbol{f}(\boldsymbol{x}) = y$. $U(\varrho_{\boldsymbol{X}})$ stands for the pseudo-uniformity on $\boldsymbol{X}$ generated by $\varrho_{\boldsymbol{X}}$.

PROPOSITION 3.1. a. *If $\boldsymbol{f}^{-1}$ is a $\delta(\varepsilon)$-uniform transformation from $(Y, U(\varrho_Y))$ onto $(\boldsymbol{X}, U(\varrho_{\boldsymbol{X}}))$, then f is $\delta^{-1}(\varepsilon)$-sensitive with respect to each non-empty family $\boldsymbol{C}_Y$*

b. *If f is $\delta^{-1}(\varepsilon)$-sensitive with respect to $\boldsymbol{C}_Y$ satisfying the following condition.*

C2. *for each $\alpha > 0$ and each $y \in Y$ there exists a $C \in \boldsymbol{C}_Y$ such that $y \in C$ and $\operatorname{diam}_{\varrho_{\boldsymbol{X}}}(\boldsymbol{f}^{-1}(C)) < \alpha$, then $\boldsymbol{f}^{-1}$ is $\delta(\varepsilon)$-uniform.*

Proof. Let $C \in \boldsymbol{C}_Y$. Proposition 2.1.a applied to $\boldsymbol{f}^{-1}$ yields $\boldsymbol{f}^{-1}(C^{\delta(\varepsilon)}) \subset \boldsymbol{f}^{-1}(C)^{\varepsilon]}$. Since $i_{\boldsymbol{X}}^{-1} \circ \boldsymbol{f}^{-1}(C^{\delta(\varepsilon)}) = f^{-1}(C^{\delta(\varepsilon)})$ and $i_{\boldsymbol{X}}^{-1}(\boldsymbol{f}^{-1}(C)^{\varepsilon]}) \subset f^{-1}(C)^{\varepsilon]}$ holds, f is $\delta^{-1}(\varepsilon)$-sensitive with respect to $\boldsymbol{C}_Y$. Here i_X is the canonical surjection of X onto $\boldsymbol{X}$.

Now, we prove the second part of the proposition. Let $y_1, y_2 \in Y$ be such that $\varrho_Y(y_1, y_2) < \delta(\varepsilon)$. Take $\gamma > 0$ and $C \in \boldsymbol{C}_Y$ such that $y_2 \in C$ and $\operatorname{diam}_{\varrho_{\boldsymbol{X}}} \boldsymbol{f}^{-1}(C) < \gamma$. If $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ stand for $\boldsymbol{f}^{-1}(y_1)$ and $\boldsymbol{f}^{-1}(y_2)$, respectively, then $\delta^{-1}(\varepsilon)$-sensitivity of f yields $\varrho_X(x, f^{-1}(C)) \leqslant \varepsilon$ for every $x \in \boldsymbol{x}_1$. Let us fix $x \in \boldsymbol{x}_1$. There exists an $x_0 \in f^{-1}(C)$ such that $\varrho_X(x, x_0) < \varepsilon + \gamma$. Condition $\operatorname{diam}_{\varrho_{\boldsymbol{X}}} \boldsymbol{f}^{-1}(C) < \gamma$ implies the existence of $x_2 \in \boldsymbol{x}_2$ such that $\varrho_X(x_0, x_2) < \gamma$. Hence

$$\varrho_X(x_1, x_2) \leqslant \varrho_X(x, x_0) + \varrho_X(x_0, x_2) < \varepsilon + 2\gamma,$$

i.e. $\varrho_X(x, \boldsymbol{x}_2) < \varepsilon + 2\gamma$. Since the argument is valid for every $x \in \boldsymbol{x}_1$, we have

$$\sup\{\varrho_X(x, \boldsymbol{x}_2);\ x \in \boldsymbol{x}_1\} \leqslant \varepsilon + 2\gamma.$$

By symmetry we have

$$\sup\{\varrho_X(\boldsymbol{x}_1, x);\ x \in \boldsymbol{x}_2\} \leqslant \varepsilon + 2\gamma$$

and hence $\varrho_{\boldsymbol{X}}(\boldsymbol{x}_1, \boldsymbol{x}_2) \leqslant \varepsilon + 2\gamma$. The left-hand side of the last inequality does not depend on $\gamma > 0$. Thus, we obtain $\varrho_{\boldsymbol{X}}(\boldsymbol{x}_1, \boldsymbol{x}_2) \leqslant \varepsilon$, i.e., $\boldsymbol{f}^{-1}$ is $\delta(\varepsilon)$-uniform.

Clearly, if f is a one-to-one function, then Proposition 3.1 follows from Proposition 2.1 applied to f^{-1}.

Remark 3. One can use in $\boldsymbol{X}$ metrics different from $\varrho_{\boldsymbol{X}}$. Metrics $\varrho_{\boldsymbol{X},W}$ given by

$$\varrho_{\boldsymbol{X},W}(\boldsymbol{x}_1, \boldsymbol{x}_2) = W(\alpha_1, \alpha_2),$$

where α_i are given by (3.2) are examples of such metrics. The function W is here supposed to be defined on $R^+ \times R^+$ and satisfies the following conditions:

W1. $W(\alpha_1, \alpha_2) = 0$ *if and only if* $\alpha_1 = \alpha_2 = 0$,

W2. $W(\alpha_1, \alpha_2) = W(\alpha_2, \alpha_1)$,

W3. $W(\alpha_1, \alpha_2) \leqslant W(\beta_1, \beta_2)$ *if* $\alpha_1 \leqslant \beta_1$ *and* $\alpha_2 \leqslant \beta_2$,

W4. $W(\alpha_1+\beta_1, \alpha_2+\beta_2) \leqslant W(\alpha_1, \alpha_2)+W(\beta_1, \beta_2)$,

W5. *W is continuous at zero on* $R^+ \times R^+$.

It is easy to see that if W satisfies W1–W5, then it is continuous on $R^+ \times R^+$. Let W_1 and W_2 be arbitrary two functions satisfying conditions W1–W5. Given $\varepsilon > 0$, let us take an ε_1, $0 < \varepsilon_1 < \varepsilon$, such that $W_1^{-1}(\varepsilon_1) \neq \emptyset$. Let

$$\delta(\varepsilon) = \tfrac{1}{2}\inf\{W_2(\alpha_1, \alpha_2)\colon \alpha_1 \geqslant 0, \alpha_2 \geqslant 0, W_1(\alpha_1, \alpha_2) = \varepsilon_1(\varepsilon)\}.$$

Clearly, $\delta(\varepsilon) > 0$ and $W_2(\alpha_1, \alpha_2) < \delta(\varepsilon)$ implies $W_1(\alpha_1, \alpha_2) < \varepsilon$. Similarly, there exists a $\delta'(\varepsilon) > 0$ such that $W_2(\alpha_1, \alpha_2) < \varepsilon$ whenever $W_1(\alpha_1, \alpha_2) < \delta'(\varepsilon)$. This implies that arbitrary two functions satisfying conditions W1–W5 give equivalent metrics, i.e., $U(\varrho_{X,W_1}) = U(\varrho_{X,W_2})$ (cf. [15], formula (2)). The Hausdorff metric is also of such a type and is given by $W(\alpha_1, \alpha_2) = \max(\alpha_1, \alpha_2)$.

The family of functions W satisfying W1–W5 has been used in [15] in order to construct metrics on product spaces. All such metrics $W(\varrho_1, \varrho_2)$, where ϱ_i, $i = 1, 2$, are fixed metrics, yield the same uniformity on the product space.

Suppose that W is a function satisfying W1–W5. Let $\gamma(\varepsilon)$ be such a function that $\max(\alpha_1, \alpha_2) \leqslant \alpha$ whenever $W(\alpha_1, \alpha_2) \leqslant \gamma(\alpha)$. Let $W^{-1}(\alpha) = \inf\{\beta > 0\colon W(\beta, \beta) = \alpha\}$.

If the metric $\varrho_{X,W}$ is used instead of ϱ_X, then slight changes in the proof of Proposition 3.1 yield the following (cf. also Proposition 3.4 given below):

PROPOSITION 3.2. a. *If* f^{-1} *is a* $\delta(\varepsilon)$*-uniform transformation from* $(Y, U(\varrho_Y))$ *onto* $(X, U(\varrho_{X,W}))$, *then* f *is* $(\delta \circ \gamma)^{-1}(\varepsilon)$*-sensitive with respect to each non-empty family* C_Y.

b. *If* f *is* $\delta^{-1}(\varepsilon)$*-sensitive with respect to* C_Y *and condition* C2 *is fulfilled, then* f^{-1} *is* $\delta \circ W^{-1}(\varepsilon)$*-uniform.*

Denote by $\hat{f}^{-1}$ the transformation from $(2^Y, U(\varrho_{2^Y}))$ into $(2^X, U(\varrho_{2^X}))$ given by $\hat{f}^{-1}(C) = f^{-1}(C)$, where $C \subset Y$.

PROPOSITION 3.3. f *is* $\delta^{-1}(\varepsilon)$*-sensitive with respect to* C_Y *if and only if* $\hat{f}^{-1}$ *is* $\delta(\varepsilon)$*-uniform from the right on* C_Y.

Proof. Let f be $\delta^{-1}(\varepsilon)$-sensitive. Take C_1 such that $C \subset C_1 \subset C^{\delta(\varepsilon)}$. Then we have

$$f^{-1}(C) \subset f^{-1}(C_1) \subset f^{-1}(C^{\delta(\varepsilon)}) \subset f^{-1}(C)^{\varepsilon]},$$

i.e., $\varrho_{2^X}(f^{-1}(C_1), f^{-1}(C)) \leqslant \varepsilon$.

If $\hat{f}^{-1}$ is $\delta(\varepsilon)$-uniform from the right on C_Y, then $\varrho_{2^X}(f^{-1}(C^{\delta(\varepsilon)}), f^{-1}(C)) \leqslant \varepsilon$, i.e., $f^{-1}(C^{\delta(\varepsilon)}) \subset f^{-1}(C)^{\varepsilon]}$ provided $C \in C_Y$.

Now, let (X_1, ϱ_{X_1}), (X, ϱ_X), (Y, ϱ_Y), and (Y_1, ϱ_{Y_1}) be metric spaces. Assume that g maps X_1 onto X, f maps X onto Y, and h maps Y onto Y_1. Let $t = h \circ f \circ g$. Let $\boldsymbol{X}_1 = X_1/(g)$, $\boldsymbol{X} = X/(f)$, $\boldsymbol{Y} = Y/(h)$, and $\hat{\boldsymbol{X}}_1 = X_1/(t)$, where the equivalence relations are given by (2.3). Denote by $\varrho_{\boldsymbol{X}_1}$, $\varrho_{\boldsymbol{X}}$, $\varrho_{\boldsymbol{Y}}$, and $\varrho_{\hat{\boldsymbol{X}}_1}$ the corresponding Hausdorff pseudo-metrics given by (3.1)–(3.2) in $\boldsymbol{X}_1$, $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\hat{\boldsymbol{X}}_1$, respectively. Finally, let $\boldsymbol{g}$, $\boldsymbol{f}$, $\boldsymbol{h}$, and $\boldsymbol{t}$ be given by (2.4), respectively.

PROPOSITION 3.4. *If* $\boldsymbol{f}^{-1}$ *is* $\delta(\varepsilon)$*-uniform and*

$$\varrho_{\boldsymbol{X}}(x', x'') \leqslant \delta_X(\varepsilon) \Rightarrow \varrho_{\boldsymbol{X}_1}(\boldsymbol{g}^{-1}(x'), \boldsymbol{g}^{-1}(x'')) \leqslant \varepsilon \tag{3.3}$$

and

$$\varrho_{Y_1}(y_1', y_1'') < \delta_Y(\varepsilon) \Rightarrow \varrho_{\boldsymbol{Y}}(\boldsymbol{h}^{-1}(y_1'), \boldsymbol{h}^{-1}(y_1'')) < \varepsilon \tag{3.4}$$

hold for each $x', x'' \in \boldsymbol{X}$ *and* $y_1', y_1'' \in Y_1$, *then* $\boldsymbol{t}^{-1}$ *is* $\gamma(\varepsilon)$*-uniform, where* $\gamma = \delta_Y \circ \circ \delta \circ \delta_X$.

Proof. By Proposition 3.1.a applied to g, f, and h we obtain that

$$t^{-1}(C^{\gamma(\varepsilon)}) \subset t^{-1}(C)^{\varepsilon]}$$

holds for every $C \subset Y_1$. Thus, by Proposition 3.1.b, the function t is $\gamma(\varepsilon)$-uniform.

Let us note that if $X_1 = X$ and g is the identity embedding, then (3.3) is fulfilled if and only if for some non-decreasing function δ_X

$$\delta_X(\varrho_{X_1}(x', x'')) \leqslant \varrho_X(x', x'')$$

holds for $x', x'' \in X$. Similarly, if $Y_1 = Y$ and h is the identity embedding, then (3.4) is fulfilled if and only if for a non-decreasing function δ_Y

$$\delta_Y(\varrho_Y(y', y'')) \leqslant \varrho_{Y_1}(y', y'')$$

holds for $y', y'' \in Y$.

4. Sensitivity of characterizations

Let C be a non-empty subset of the metric space (X, ϱ_X). C is said to be *characterized by sets* $A \subset X$ *and* $B \subset Y$ whenever

$$C = A \cap f^{-1}(B), \tag{4.1}$$

where f is a given function from (X, ϱ_X) into (Y, ϱ_Y).

DEFINITION 3. A characterization of $C \subset X$ by $A \subset X$, $B \subset Y$ and f is called $\gamma(\varepsilon)$*-sensitive* if

$$A^{\varepsilon} \cap f^{-1}(B^{\varepsilon}) \subset (A \cap f^{-1}(B))^{\gamma(\varepsilon)]} \tag{4.2}$$

holds for every $\varepsilon > 0$, where $\gamma(\varepsilon) > 0$ and

$$\lim_{\varepsilon \to 0} \gamma(\varepsilon) = 0. \tag{4.3}$$

Roughly speaking, the property "to be a $\gamma(\varepsilon)$-sensitive characterization" means that "slight deviations from assumptions A" and "slight deviations from the considered properties B" may happen simultaneously only when "slight deviations from the characterized set $C = A \cap f^{-1}(B)$" occur.

Let us note that if $A = X$, then (4.2) means that f is $\delta^{-1}(\varepsilon)$-sensitive with respect to a one-element family C_X, namely $C_X = \{B\}$ (cf. Example 7a in Section 8 in that case where the set of distributions of random vectors with independent components is the domain of F). However, if $A \neq X$, then even the $\delta^{-1}(\varepsilon)$-sensitivity of f does not yield (4.2). The reason is that B does not need to be a subset of $f(A^{\delta(\varepsilon)})$. Moreover, if f is not one-to-one, then the restriction of f to a subset of X does not need to be sensitive at all.

Let $\gamma(\varepsilon)$ be given by

$$\gamma(\varepsilon) = \sup\{\varrho_X(x, A \cap f^{-1}(B)) \colon x \in (A^\varepsilon \cap f^{-1}(B^\varepsilon))\}. \tag{4.4}$$

Clearly, we have

$$A^\varepsilon \cap f^{-1}(B^\varepsilon) \subset (A \cap f^{-1}(B))^{\gamma(\varepsilon)]} = C^{\gamma(\varepsilon)]}.$$

Now, we give theorems on sensitivity of characterizations. Theorem 4.1 is a slightly extended version of Theorem 2 in [16]. Theorems 4.2–4.4 deal with the case of sensitive f. Since $\delta^{-1}(\varepsilon)$-sensitive functions need not be continuous, Theorems 4.2 and 4.3 may be considered as substitutes for incorrect Theorem 1 in [16] (see the counter-example given below). In Theorem 4.4 we give a necessary and sufficient condition for the sensitivity of characterizations in the case of $\delta^{-1}(\varepsilon)$-sensitive f.

THEOREM 4.1 (Zolotarev). *Let f be continuous and A and B closed. If the following condition is fulfilled*

C3. *each sequence $\{x_n\}$ such that $x_n \in (A^{1/n} \cap f^{-1}(B^{1/n})) \setminus C$ contains a subsequence convergent to an element of X,*

then the characterization of C by A, B, and f is $\gamma(\varepsilon)$-sensitive, where $\gamma(\varepsilon)$ is given by (4.4), *i.e.,* (4.2) *and* (4.3) *hold.*

Proof. Inclusion (4.2) is evident. Suppose that (4.3) does not hold. Then there exists a sequence $\{x_n\}$ such that

$$x_n \in A^{\varepsilon_n} \cap f^{-1}(B^{\varepsilon_n}), \quad \varepsilon_n \searrow 0, \quad \text{and} \quad \varrho_X(x_n, C) > \alpha.$$

By condition C3 there is x_0, an accumulation point of $\{x_n\}$. Suppose that $\lim x_n = x_0$. Since $\varrho_X(x_0, A) \leqslant \varrho_X(x_n, A) + \varrho_X(x_n, x_0)$ (cf. [7], p. 105), we have that $x_0 \in A$. Moreover, $f(x_n) \in B^{\varepsilon_n}$. By the continuity of f, $\lim f(x_n) = f(x_0)$. Because $\varrho_Y(f(x_0), B) \leqslant \varrho_Y(f(x_n), B) + \varrho_Y(f(x_n), f(x_0))$ holds we infer that $f(x_0) \in B$. Thus, $x_0 \in C$ and hence $\lim \varrho_X(x_n, C) = 0$. This yields a contradiction.

THEOREM 4.2. *Let f be $\delta^{-1}(\varepsilon)$-sensitive $(\delta(\varepsilon) > 0)$ and A and $f^{-1}(B)$ closed. If condition* C3 *is fulfilled and $\gamma(\varepsilon)$ is given by* (4.4), *then the characterization of C by A, B and f is $\gamma(\varepsilon)$-sensitive, i.e.,* (4.2) *and* (4.3) hold.

Proof. Clearly, (4.2) holds. Suppose that (4.3) is not fulfilled. Then there is a sequence $\{x_n\}$ such that $x_n \in A^{\varepsilon_n} \cap f^{-1}(B^{\varepsilon_n}) \setminus C$, $\varrho_X(x_n, C) > \alpha$ and $\varepsilon_n \searrow 0$. We can assume with no loss of generality that $0 < \delta(\varepsilon) \leqslant \varepsilon$. Then we have $\delta^{-1}(\varepsilon) \geqslant \varepsilon$

and $\lim_{\varepsilon \to 0} \delta^{-1}(\varepsilon) = 0$. Now, by the $\delta^{-1}(\varepsilon)$-sensitivity of f, we obtain

$$A^{\varepsilon_n} \cap f^{-1}(B^{\varepsilon_n}) \subset A^{\delta^{-1}(\varepsilon)]} \cap f^{-1}(B)^{\delta^{-1}(\varepsilon_n)]}.$$

Condition C3 yields the existence of a subsequence $\{x_n'\}$ convergent to an $x_0 \in X$. Since $\lim \delta^{-1}(\varepsilon_n) = 0$, we infer that $x_0 \in A \cap f^{-1}(B) = C$, i.e., $\lim \varrho_X(x_n', C) = 0$. This yields a contradiction.

THEOREM 4.3. *Let f be $\delta^{-1}(\varepsilon)$-sensitive, $0 < \delta(\varepsilon) \leqslant \varepsilon$, and A and $f^{-1}(B)$ closed. Let $\gamma_1(\varepsilon)$ be given by*

$$\gamma_1(\varepsilon) = \sup\{\varrho_X(x, C);\ x \in (A^{\varepsilon]} \cap f^{-1}(B)^{\varepsilon]})\}. \tag{4.5}$$

If the following condition is fulfilled

C4. *each sequence $\{x_n\}$ such that $x_n \in A^{1/n} \cap (f^{-1}(B))^{1/n} \setminus C$ contains a subsequence convergent to an element of X,*

then the characterization of C by A, B and f is $\gamma_1 \circ \delta^{-1}(\varepsilon)$-sensitive, i.e., (4.2) *and* (4.3) *hold with $\gamma(\varepsilon) = \gamma_1 \circ \delta^{-1}(\varepsilon)$.*

Proof. We have

$$A^\varepsilon \cap f^{-1}(B^\varepsilon) \subset A^{\delta^{-1}(\varepsilon)]} \cap f^{-1}(B)^{\delta^{-1}(\varepsilon)]} \subset C^{\gamma(\delta^{-1}(\varepsilon))}. \tag{4.6}$$

Since $\lim_{\varepsilon \to 0} \delta^{-1}(\varepsilon) = 0$, it is enough to prove that $\lim_{\varepsilon \to 0} \gamma(\varepsilon) = 0$. If it is not the case, then there exists a sequence $\{x_n\}$, $x_n \in A^{\varepsilon_n]} \cap f^{-1}(B)^{\varepsilon_n]} \setminus C$ such that $\varrho_X(x_n, C) > \alpha$ for some $\alpha > 0$ and $\varepsilon_n \searrow 0$. Then condition C4 yields the existence of a subsequence $\{x_n'\}$ convergent to an $x_0 \in X$. Clearly, $x_0 \in C$, and hence $\lim \varrho_X(x_n', C) = 0$. This yields a contradiction.

Let us note that Theorem 4.3 yields a decomposition of the sensitivity of the characterization into two parts $\gamma_1(\varepsilon)$ and $\delta^{-1}(\varepsilon)$. The first function depends on "shapes" and the mutual position of the sets A and $f^{-1}(B)$, i.e., on "geometrical" properties of these sets. The second part depends on the properties of function f only. The following theorem makes more evident this "geometrical character" of $\gamma_1(\varepsilon)$.

THEOREM 4.4. *Let f be $\delta^{-1}(\varepsilon)$-sensitive, $0 < \delta(\varepsilon) \leqslant \varepsilon$, and let $\gamma_1(\varepsilon)$ be given by* (4.5). *Then the characterization of C by A, B, and f is $(\gamma_1 \circ \delta^{-1})(\varepsilon)$-sensitive, i.e.,* (4.2) *and* (4.3) *hold if and only if condition*

C5. $x_n' \in A$, $x_n'' \in f^{-1}(B)$, $\lim \varrho_X(x_n', x_n'') = 0 \Rightarrow \lim \varrho_X(x_n', C) = 0$

is fulfilled.

Condition C5 means that subsets of A which are far from C are not in proximity to $f^{-1}(B)$.

Proof. First, let us note that (4.6) and hence (4.2) with γ_1 instead of γ hold. Thus, it is enough to prove that condition C5 is equivalent to (4.3) with γ_1 instead of γ.

Suppose that C5 is fulfilled and (4.3) does not hold. Then there exist $\alpha > 0$ and $x_n \in A^{\varepsilon_n]} \cap f^{-1}(B)^{\varepsilon_n]}$, $\varepsilon_n \searrow 0$, such that $\varrho_X(x_n, C) > \alpha$. Take $x_n' \in A$ such that

$\varrho_X(x_n', x_n) < 2\varepsilon_n$ and $x_n'' \in f^{-1}(B)$ such that $\varrho_X(x_n'', x_n) < 2\varepsilon_n$. Thus, $\varrho_X(x_n', x_n'') < 4\varepsilon_n$ and $\varrho_X(x_n', C) > \alpha - 2\varepsilon_n > \alpha/2$ for n sufficiently large. This yields a contradiction.

Now, let $x_n' \in A$, $x_n'' \in f^{-1}(B)$ and let $\varrho_X(x_n', x_n'')$ converge to zero. Then for each $\varepsilon > 0$ there is an N such that $x_n' \in f^{-1}(B)^{\varepsilon]}$ provided $n > N$. Hence $x_n' \in A^{\varepsilon]} \cap f^{-1}(B)^{\varepsilon]}$ and $\varrho_X(x_n', C) \leqslant \gamma_1(\varepsilon)$. Since $\lim_{\varepsilon \to 0} \gamma_1(\varepsilon) = 0$, we obtain $\lim_{n\to\infty} \varrho_X(x_n', C) = 0$. The theorem is proved.

Clearly, Theorem 4.4 implies Theorem 4.3 and Theorem 4.3 implies Theorem 4.2.

Let us make several remarks on the assumptions used in Theorems 4.1–4.4. We have

(a) if A is compact, then condition C3 is fulfilled;

(b) the assumptions of Theorem 4.1 imply the compactness of the boundary of C;

(c) if A or $f^{-1}(B)$ is compact, then condition C4 is fulfilled;

(d) the assumptions of Theorem 4.3 imply both the compactness of the boundary of C and condition C5.

Finally, let us note that the assumption on the $\delta^{-1}(\varepsilon)$-sensitivity of f in Theorem 4.2 cannot be replaced even by the compactness of B. Indeed, let $X = Y = A = [0, 1]$ be endowed with the usual metric topology. If

$$f(x) = \begin{cases} 1 & \text{for} \quad x = 0, \\ x & \text{for} \quad x \in (0, 1), \\ 0 & \text{for} \quad x = 1, \end{cases}$$

and $B = \{1\}$, then $f^{-1}(B^\varepsilon) = \{0\} \cup (1-\varepsilon, 1)$. Since $A^\varepsilon = [0, 1]$, we have $C = \{0\}$ and $\gamma(\varepsilon) = 1$ for each $\varepsilon > 0$.

Essentially, this yields a contradiction to Theorem 1 in [16]. However, to be just in the framework of this theorem we should: (1) identify each point $x \in [0, 1]$ with a constant random variable identically equal to x, (2) endow the set of such degenerated r.v.'s with a metric equal to the Prokhorov distance between the distributions of these r.v.'s (clearly, this metric here corresponds to the usual metric in $[0, 1]$).

5. Stable transformations of spaces of measures

Let $\mathscr{B}(X)$ be the σ-algebra of Borel subsets of X. $\mathscr{M}(X)$ is the set of all finite, real-valued, non-negative, countably additive set functions on $\mathscr{B}(X)$. The set of all probability measures on $\mathscr{B}(X)$ will be denoted by $\mathscr{P}(X)$. The symbols μ and ν, may be with indices, will stand for elements of $\mathscr{M}(X)$.

If f is a $(\mathscr{B}(X), \mathscr{B}(Y))$-measurable function, then it generates a transformation F from $\mathscr{M}(X)$ into $\mathscr{M}(Y)$, where

$$F\mu = \mu \circ f^{-1},$$

i.e.,

$$F\mu(C) = \mu(f^{-1}(C)) \tag{5.1}$$

for $C \in \mathscr{B}(Y)$.

In this section we will consider property "to be stable" of transformation F given by (5.1). This property depends on properties of f and on the choice of metrics in $\mathscr{M}(X)$ and $\mathscr{M}(Y)$. In the subsequent sub-sections we will consider the cases where the sets of measures are endowed with the Prokhorov, BL*, B*, Meshalkin, Levy, Sibley and Kolmogorov metrics. In view of Proposition 2.1 we restrict ourselves to investigations on the $\delta(\varepsilon)$-uniformities of F.

5.1. LEMMA 5.1. *A function f is $\delta(\varepsilon)$-uniform if and only if*

$$f^{-1}(C)^{\delta(\varepsilon)} \subset f^{-1}(C^{\varepsilon]}) \tag{5.2}$$

holds for every closed set C, $C \subset Y$.

Proof. Let f be $\delta(\varepsilon)$-uniform. If $x \in f^{-1}(C)^{\delta(\varepsilon)}$, then there is an $x_0 \in f^{-1}(C)$ such that $\varrho_X(x, x_0) < \delta(\varepsilon)$. Hence $\varrho_Y(f(x), f(x_0)) \leqslant \varepsilon$, i.e., $f(x) \in C^{\varepsilon]}$. This yields $x \in f^{-1}(C^{\varepsilon]})$.

Now, let $\varrho_X(x_1, x_2) < \delta(\varepsilon)$. Take $C = \{f(x_1)\}$. Then $x_2 \in f^{-1}(C)^{\delta(\varepsilon)}$. Hence by (5.2), $x_2 \in f^{-1}(C^{\varepsilon]})$, i.e., f is $\delta(\varepsilon)$-uniform.

THEOREM 5.2[1] (Zolotarev). *Let f be $\delta(\varepsilon)$-uniform, $0 < \delta(\varepsilon) \leqslant \varepsilon$. Then F, given by* (5.1), *is a $\delta(\varepsilon)$-uniform and hence $\delta(\varepsilon)$-stable transformation from $\mathscr{M}(X)$ into $\mathscr{M}(Y)$ which are endowed with the corresponding uniformities $U(\varrho_{P,X})$ and $U(\varrho_{P,Y})$.*

Proof. Let $\mu_i \in \mathscr{M}(X)$, $i = 1, 2$, and $\varrho_{P,X}(\mu_1, \mu_2) < \delta(\varepsilon)$. We will show that $\varrho_{P,Y}(F_{\mu_1}, F_{\mu_2}) \leqslant \varepsilon$. Indeed, by Lemma 5.1, we have

$$\begin{aligned} \sigma(F_{\mu_1}, F_{\mu_2}) &= \inf\{\alpha > 0;\ \mu_1(f^{-1}(C)) \leqslant \mu_2(f^{-1}(C^{\alpha]}))+\alpha \\ &\qquad\qquad \text{for all closed } C \subset Y\} \\ &\leqslant \inf\{\alpha > 0;\ \mu_1(f^{-1}(C)) \leqslant \mu_2(f^{-1}(C)^{\delta(\alpha)})+\delta(\alpha) \\ &\qquad\qquad \text{for all closed } C \subset Y\} \\ &\leqslant \inf\{\alpha > 0;\ \mu_1(C) \leqslant \mu_2(C^{\delta(\alpha)})+\delta(\alpha) \text{ for all closed } C \subset X\} \leqslant \varepsilon. \end{aligned}$$

Similarly, we have that $\sigma(F\mu_2, F\mu_1) \leqslant \varepsilon$. Hence $\varrho_{P,Y}(F\mu_1, F\mu_2) \leqslant \varepsilon$. The $\delta(\varepsilon)$-stability of F follows now from Proposition 2.1.a.

Remark 4. If metrics different from $\varrho_{P,X}$ and $\varrho_{P,Y}$ are used in $\mathscr{M}(X)$ and $\mathscr{M}(Y)$, respectively, then, in view of Proposition 2.2, we can deduce from Theorem 5.2 several new information on uniformities of F. Thus, the following metrics can be used in Proposition 2.2 instead of ϱ_{X_1} provided ϱ_X is the Prokhorov metric in $\mathscr{M}(X)$:

1. ϱ_{BL^*} $(\delta_X(\varepsilon) = \varepsilon)$, 2. ϱ_{B^*} $(\delta_X(\varepsilon) = \varepsilon)$.

Similarly, we may use the following metrics instead of metric ϱ_{Y_1} used in Proposition 2.2 provided ϱ_Y is the Prokhorov metric:

1. ϱ_{BL^*} $(\delta_Y(\varepsilon) = \frac{1}{2}\varepsilon)$, 2. ϱ_L $(\delta_Y(\varepsilon) = \varepsilon)$, 3. ϱ_S $(\delta_Y(\varepsilon) = \varepsilon)$.

---

[1] We refer the reader to Section 0 for a discussion on relationships of Theorem 5.2 to theorems given in [1], [15], and [17].

Proposition 2.1.b shows that, in general, in order to ensure the $\delta(\varepsilon)$-stability of F we have to ensure the $\delta(\varepsilon)$-uniformity of F. By Theorem 5.2 the $\delta(\varepsilon)$-uniformity of F is implied by the $\delta(\varepsilon)$-uniformity of f. Now, we show that, in general, the assumption on $\delta(\varepsilon)$-uniformity of f in Theorem 5.2 cannot be weakened. Let $K(z, \gamma)$ denote a ball centered at a point z and of a radius γ.

THEOREM 5.3. *Let F, given by* (5.1), *be a $\delta(\varepsilon)$-uniform transformation from* $(\mathscr{A}, U(\varrho_{P,X}))$, $\mathscr{A} \subset \mathscr{M}(X)$, *into* $(\mathscr{M}(Y), U(\varrho_{P,Y}))$. *If $\mathscr{A}$, the domain of F, satisfies the condition*

C6. *for each $\gamma > 0$ and each $x \in X$ there exists a $\mu \in \mathscr{A} \cap \mathscr{P}(X)$ such that*

$$\mu\big(f^{-1}(K(f(x), \gamma)) \cap K(x, \gamma)\big) > 1-\gamma,$$

then f is a $\delta(\varepsilon)$-uniform transformation from $(X, U(\varrho_X))$ *into* $(Y, U(\varrho_Y))$.

Proof. Let $\varrho_X(x_1, x_2) = \delta_1$, $\delta_1 < \delta(\varepsilon)$ for a given ε. Take $\gamma > 0$ such that $\delta_1 + 2\gamma < \delta(\varepsilon)$. Let μ_i be probability measures satisfying condition C6 for γ and x_i, $i = 1, 2$. We have

$$\mu_i(C) \leqslant \mu_2(C^{\delta_1+2\gamma]}) + \delta_1 + 2\gamma \tag{5.3}$$

for each closed $C \subset X$. Indeed, if $C \cap K(x_1, \gamma) = \emptyset$, then $\mu_1(C) < \gamma$ and (5.3) is fulfilled. If there is an $x_0 \in C \cap K(x_1, \gamma)$, then $C^{\delta_1+2\gamma]} \supset K(x_2, \gamma)$; hence

$$\mu_2(C^{\delta_1+2\gamma]}) + \delta_1 + 2\gamma > 1.$$

Thus, $\sigma(\mu_1, \mu_2) < \delta(\varepsilon)$. Similarly, we obtain $\sigma(\mu_2, \mu_1) < \delta(\varepsilon)$. Consequently, $\varrho_{P,X}(\mu_1, \mu_2) < \delta(\varepsilon)$. By the $\delta(\varepsilon)$-uniformity of F we obtain $\varrho_{P,Y}(F\mu_1, F\mu_2) \leqslant \varepsilon$. This means that

$$\mu_1\big(f^{-1}(K(f(x_1), \gamma))\big) \leqslant \mu_2\big(f^{-1}(K(f(x_1), \varepsilon+2\gamma))\big) + \varepsilon + \gamma.$$

If $\gamma < \frac{1}{2}(1-\varepsilon)$, then $K(f(x_1), \varepsilon+2\gamma) \cap K(f(x_2), \gamma) \neq \emptyset$. Hence $\varrho_Y(f(x_1), f(x_2)) \leqslant \varepsilon + 3\gamma$. Since the left-hand side of this inequality does not depend on γ, we obtain that $\varrho_Y(f(x_1), f(x_2)) \leqslant \varepsilon$, i.e., f is a $\delta(\varepsilon)$-uniform transformation.

Let us point out that condition C6 means that for each $x \in X$ and $\gamma > 0$ there is such a probability measure μ which is "almost" concentrated on the ball $K(x, \gamma)$ and that f transports "almost whole mass" from $K(x, \gamma)$ into $K(f(x), \gamma)$. Thus, this condition is fulfilled if, e.g., $\mathscr{A}$ contains all probability measures concentrated at points $x \in X$.

5.2. In this sub-section we show that in some cases, direct considerations on the transformations from $(\mathscr{M}(X), U(\varrho_{BL^*,X}))$ into $(\mathscr{M}(Y), U(\varrho_{BL^*,Y}))$ lead to stronger results than those which follow from Theorem 5.2 and Proposition 2.2 (cf. Remark 4).

THEOREM 5.4. *If f is $\alpha\varepsilon$-uniform ($\alpha \leqslant 1$), one-to-one and onto Y, then F is an $\alpha\varepsilon$-uniform and hence $\alpha\varepsilon$-stable transformation from* $(\mathscr{M}(X), U(\varrho_{BL^*,X}))$ *into* $(\mathscr{M}(Y), U(\varrho_{BL^*,Y}))$.

Proof. $\alpha\varepsilon$-uniformity of f implies that

$$\varrho_Y(f(x_1), f(x_2))/\varrho_X(x_1, x_2) \leqslant \alpha^{-1}$$

holds for each $x_1, x_2 \in X$. Thus, we have

$$\|g \circ f\|_{BL,X} \leqslant \alpha^{-1} \|g\|_{BL,Y},$$

where

$$\|g \circ f\|_{BL,X} = \sup\{|g \circ f(x)|;\ x \in X\} + \sup\{|g \circ f(x_1) - g \circ f(x_2)|/\varrho_X(x_1, x_2);\ x_1, x_2 \in X\},$$

and

$$\|g\|_{BL,Y} = \sup\{|g(y)|;\ y \in Y\} + \sup\{|g(y_1) - g(y_2)|/\varrho_Y(y_1, y_2);\ y_1 y_2 \in Y\}.$$

Hence we obtain

$$\begin{aligned}\varrho_{BL^*,Y}(\mu_1 \circ f^{-1}, \mu_2 \circ f^{-1}) &= \sup\Big\{\Big|\int g\,d((\mu_1 - \mu_2) \circ f^{-1})\Big|;\ \|g\|_{BL,Y} \leqslant 1\Big\} \\ &\leqslant \sup\Big\{\Big|\int h\,d(\mu_1 - \mu_2);\ \|h\|_{LB,X} \leqslant \alpha^{-1}\Big\} = \alpha^{-1} \varrho_{BL^*,X}(\mu_1, \mu_2),\end{aligned}$$

i.e., F is $\alpha\varepsilon$-uniform. By Proposition 2.1, function F is $\alpha\varepsilon$-stable.

Remark 5. If metrics different from $\varrho_{BL^*,X}$ and $\varrho_{BL^*,Y}$ are used in $\mathscr{M}(X)$ and $\mathscr{M}(Y)$, respectively, then, in view of Proposition 2.2, we can deduce from Theorem 5.4 several new information on uniformities of F. Thus, the following metrics can be used in Proposition 2.2 instead of ϱ_{X_1} provided ϱ_X is the Dudley ϱ_{BL^*} metric in $\mathscr{M}(X)$:

1. ϱ_P $(\delta_X(\varepsilon) = \frac{1}{2}\varepsilon)$, 2. ϱ_{B^*} $(\delta_X(\varepsilon) = \frac{1}{2}\varepsilon)$.

Similarly, we can use the following metrics instead of metric ϱ_{Y_1} used in Proposition 2.2 provided ϱ_Y is the metric $\varrho_{BL^*,Y}$:

1. ϱ_P $(\delta_Y(\varepsilon) = \frac{2}{3}\varepsilon^2)$, 2. ϱ_L $(\delta_Y(\varepsilon) = \frac{2}{3}\varepsilon^2)$,
3. ϱ_S $(\delta_Y(\varepsilon) = \frac{2}{3}\varepsilon^2)$.

5.3. In the case of metrics $\varrho_{B^*,X}$ and $\varrho_{B^*,Y}$ we have

THEOREM 5.5.(2) (Zolotarev). *Let $(X, \mathscr{B}(X))$ and $(Y, \mathscr{B}(Y))$ be measure spaces with σ-algebras $\mathscr{B}(X)$ and $\mathscr{B}(Y)$ of subsets of X and Y, respectively. If f is a $(\mathscr{B}(X), \mathscr{B}(Y))$-measurable function from X into Y, then F given by* (5.1) *is ε-uniform and hence ε-stable transformation from* $(\mathscr{M}(X), U(\varrho_{B^*,X}))$ *into* $(\mathscr{M}(Y), U(\varrho_{B^*,Y}))$.

Proof. We have

$$\begin{aligned}\varrho_{B^*,Y}(F\mu_1, F\mu_2) &= \sup\{|\mu_1(f^{-1}(C)) - \mu_2(f^{-1}(C))|;\ C \in \mathscr{B}(Y)\} \\ &\leqslant \sup\{|\mu_1(C) - \mu_2(C)|;\ C \in \mathscr{B}(X)\} = \varrho_{B^*,X}(\mu_1, \mu_2),\end{aligned}$$

i.e., F is ε-uniform. Proposition 2.1 yields the ε-stability of F.

Remark 6. If metrics different from $\varrho_{B^*,X}$ and $\varrho_{B^*,Y}$ are used in $\mathscr{M}(X)$ and $\mathscr{M}(Y)$, respectively, then in view of Proposition 2.2 we can deduce from Theorem

(2) In this theorem we do not assume that X and Y are metric spaces. See Section 0 for a discussion of relationships of Theorem 5.5 with theorems given in [15] and [17].

5.5 other information on uniformities of F. Thus, the following metrics can be used in Proposition 2.2 instead of ϱ_{Y_1} provided ϱ_Y is the $\varrho_{B^*,Y}$ metric in $\mathscr{M}(Y)$:

1. ϱ_P $(\delta_Y(\varepsilon) \doteq \varepsilon)$, 2. ϱ_{BL^*} $(\delta_Y(\varepsilon) = \frac{1}{2}\varepsilon)$, 3. ϱ_L $(\delta_Y(\varepsilon) = \varepsilon)$,
4. ϱ_S $(\delta_Y(\varepsilon) = \varepsilon)$, 5. ϱ_K $(\delta_Y(\varepsilon) = \varepsilon)$, 6. ϱ_M $(\delta_{\bar{Y}}(\varepsilon) = \varepsilon)$.

5.4. Now, we consider the case where the Meshalkin metrics are used.

THEOREM 5.6. *Let f be a linear transformation from $X = \mathbf{R}^n$ into $Y = \mathbf{R}^m$. Then F given by* (5.1) *is an ε-uniform and hence ε-stable transformation from $(\mathscr{M}(X), U(\varrho_{M,X}))$ into $(\mathscr{M}(Y), U(\varrho_{M,Y}))$.*

Proof. Indeed, we have

$$\begin{aligned}&\varrho_{M,Y}(F\mu_1, F\mu_2)\\ &= \sup\{|\mu_1(f^{-1}(C))-\mu_2(f^{-1}(C))|;\ C \text{ is an intersection of at most } m \text{ half-spaces}\}\\ &\leqslant \sup\{|\mu_1(A)-\mu_2(A)|;\ A \text{ is an intersection of at most } n \text{ half-spaces}\}\\ &= \varrho_{M,X}(\mu_1, \mu_2).\end{aligned}$$

Thus, F is ε-uniform and hence, by Proposition 2.1.a, ε-stable.

Remark 7. If metrics different from $\varrho_{M,X}$ and $\varrho_{M,Y}$ are used in $\mathscr{M}(X)$ and $\mathscr{M}(Y)$, then in view of Proposition 2.2 we can deduce from Theorem 5.6 other information on uniformities of F. Thus, if $\varrho_X = \varrho_{M,X}$, then we may use $\varrho_{B^*,X}$ instead of ϱ_{X_1} in Proposition 2.2. In this case $\delta_X(\varepsilon) = \varepsilon$.

Similarly, we can use the following metrics instead of metric ϱ_{Y_1}, used in Proposition 2.2 provided ϱ_Y is the metric $\varrho_{M,Y}$:

1. ϱ_K $(\delta_Y(\varepsilon) = \varepsilon)$, 2. ϱ_L $(\delta_Y(\varepsilon) = \varepsilon)$, 3. ϱ_S $(\delta_Y(\varepsilon) = \varepsilon)$.

6. Sensitive transformations of spaces of measures

In this section we consider $\gamma(\varepsilon)$-sensitive transformations from one set of measures, say $\mathscr{X} \subset \mathscr{P}(X)$, onto another one, say $\mathscr{Y} \subset \mathscr{P}(Y)$. We are concerned with transformations F given by (5.1), only.

In view of Proposition 3.1 it is clear that we should find conditions implying the $\delta(\varepsilon)$-uniformity of F^{-1}.

Recall that f is called an *isomorphism between measure spaces* $(X, \mathscr{B}(X))$ *and* $(Y, \mathscr{B}(Y))$ if f is one-to-one, onto, $(\mathscr{B}(X), \mathscr{B}(Y))$-measurable and f^{-1} is $(\mathscr{B}(Y), \mathscr{B}(X))$-measurable (if X and Y are Polish and $\mathscr{B}(X)$ and $\mathscr{B}(Y)$ are σ-algebras of Borel subsets, then, by a theorem of Kuratowski ([7], p. 397), the last condition is a consequence of the first ones). If f is an isomorphism, then F is a one-to-one transformation from $\mathscr{P}(X)$ onto $\mathscr{P}(Y)$. In such a case we can use Theorems 5.2–5.6 with f and F replaced by f^{-1} and F^{-1}, respectively, and no new theorems are here needed. Thus, essentially new cases arise only when $f^{-1}(\mathscr{B}(Y))$ is a proper sub-σ-algebra of $\mathscr{B}(X)$. Then we have to consider a proper quotient space $X = \mathscr{X}/(F)$ and we have to find conditions implying the $\delta(\varepsilon)$-uniformity of F^{-1}. In Theorem

6.1 given below no topology in X and Y is assumed. Thus, in this section $(X, \mathscr{B}(X))$ and $(Y, \mathscr{B}(Y))$ stand for measure spaces.

Denote by $\mathscr{X}(\mu_0)$, $\mu_0 \in \mathscr{P}(X)$, the set of all probability measures $\mu \in \mathscr{P}(X)$ such that μ and μ_0 are mutually absolutely continuous, i.e., $\mathscr{X}(\mu_0) = \{\mu \in \mathscr{P}(X)$: if $A \in \mathscr{B}(X)$, then $\mu(A) = 0$ if and only if $\mu_0(A) = 0\}$. We shall assume that $(X, \mathscr{B}(X), \mu_0)$ is a complete measure space. Let $\mathscr{D} = f^{-1}(\mathscr{B}(Y))$. Moreover, let $\varrho_{B^*,X}$ stand for the Hausdorff metric given by (3.1)–(3.2) in $\mathbf{X}(\mu_0)$ provided $\varrho_{B^*,X}$ is used in (3.2) instead of ϱ_X.

THEOREM 6.1. *Let f be a $(\mathscr{B}(X), \mathscr{B}(Y))$-measurable function from X into Y and $\mu_0 \in \mathscr{P}(X)$. Suppose that for each $\mu \in \mathscr{X}(\mu_0)$ there exists $\pi(x, A; \mathscr{D}, \mu)$ a* p.r.c.p. *on $\mathscr{B}(X)$ given $\mathscr{D} = f^{-1}(\mathscr{B}(Y))$.*

Let F, given by (5.1), *be a function from $\mathscr{X}(\mu_0)$ onto $\mathscr{Y} \subset \mathscr{P}(Y)$. Then, F^{-1} is an ε-uniform transformation from $(\mathscr{Y}, U(\varrho_{B^*,Y}))$ onto $(\mathbf{X}(\mu_0), U(\varrho_{B^*,X}))$.*

Proof. Let $\nu_1, \nu_2 \in \mathscr{Y}$ and let $\mu_1 \in \mathscr{X}(\mu_0)$ *be such that* $\mu_1 \in F^{-1}(\nu_1)$. Let $\pi(x, A; \mathscr{D}, \mu_1)$ be a version of proper regular conditional probability on $\mathscr{B}(X)$ given $\mathscr{D} = f^{-1}(\mathscr{B}(Y))$ provided μ_1 is a probability measure on $\mathscr{B}(X)$ (cf. Section 1). We shall show that there is $\mu_2 \in \mathscr{X}(\mu_0)$ such that $\mu_2 \in F^{-1}(\nu_2)$ and

$$\varrho_{B^*X}(\mu_1, \mu_2) = \varrho_{B^*,Y}(\nu_1, \nu_2) \tag{6.1}$$

holds. Indeed, let us take an arbitrary $\mu_2' \in \mathscr{X}(\mu_0) \cap F^{-1}(\nu_2)$.

We put

$$\mu_2(A) = \int \pi(x, A; \mathscr{D}, \mu_1)\mu_2'(dx) \quad \text{for} \quad A \in \tilde{\mathscr{B}}(X), \tag{6.2}$$

where $\tilde{\mathscr{B}}(X) \sim \mathscr{B}(X)$. By the regularity of π the set function μ_2 given by (6.2) is a probability measure on $\tilde{\mathscr{B}}(X)$. If $A' \in \mathscr{B}(X)$ and $\mu_0(A' \triangle A) = 0$, then we put

$$\mu_2(A') = \mu_2(A).$$

Clearly, μ_2 is a probability measure on $\mathscr{B}(X)$ and $\mu_2 \in \mathscr{X}(\mu_1) = \mathscr{X}(\mu_0)$. Since π is proper, we have

$$\mu_2\{f^{-1}(C)\} = \mu_2'\{f^{-1}(C)\} = \nu_2(C)$$

for every $C \in \mathscr{B}(Y)$, i.e., $\mu_2 \in F^{-1}(\nu_2)$. Finally, we have

$$\begin{aligned}&\sup\{|\mu_1(A)-\mu_2(A)|;\ A \in \mathscr{B}(X)\}\\ &\quad = \sup\Big\{\Big|\int \pi(x, A; \mathscr{D}, \mu_1)(\mu_1-\mu_2')(dx)\Big|;\ A \in \tilde{\mathscr{B}}(X)\Big\}\\ &\quad = \sup\Big\{\Big|\int_A (\mu_1-\mu_2')(dx)\Big|;\ A \in \mathscr{D}\Big\} = \sup\{|\nu_1(C)-\nu_2(C)|;\ C \in \mathscr{B}(Y)\},\end{aligned}$$

i.e., (6.1) holds.

Similarly, for each $\mu_2 \in \mathscr{X}(\mu_0) \cap F^{-1}(\nu_2)$ there exists a $\mu_1 \in \mathscr{X}(\mu_0) \cap F^{-1}(\nu_1)$ such that (6.1) holds. Consequently, $\varrho_{B^*,X}(\mu_1, \mu_2) = \varrho_{B^*,Y}(F\mu_1, F\mu_2)$. This yields the *$\varepsilon$-uniformity* of F^{-1}. By Proposition 3.1.a, transformation F is also ε-sensitive.

Remark 8. Suppose that $\mathscr{B}(X)$ is the μ_0-completion of the σ-algebra of Borel subsets of a metric space (X, ϱ_X). If a metric different from $\varrho_{B^*,X}$ is used in $\mathscr{M}(X)$, then, in view of Proposition 3.4, we can deduce from Theorem 6.1 other information on uniformities of F^{-1}. Thus, we can use the following metrics instead of metric ϱ_{X_1} used in Proposition 2.2 provided ϱ_X is metric $\varrho_{B*,X}$:

1. ϱ_P $(\delta_X(\varepsilon) = \varepsilon)$, 2. ϱ_{BL*} $(\delta_X(\varepsilon) = \frac{1}{2}\varepsilon)$, 3. ϱ_K $(\delta_X(\varepsilon) = \varepsilon)$,
4. ϱ_M $(\delta_X(\varepsilon) = \varepsilon)$, 5. ϱ_L $(\delta_X(\varepsilon) = \varepsilon)$, 6. ϱ_S $(\delta_X(\varepsilon) = \varepsilon)$.

7. Applications of stable functions

In this section we present some simple consequences of theorems given in Sections 2 and 5 on stable transformations. Examples 1–5 given below indicate the problems in which these theorems are useful, also.

EXAMPLE 1. Let f be a real $\delta(\varepsilon)$-stable function and let $f(x) \leqslant 0$ for each $x \in C$, $C \subset X$. If $\varrho_X(x, C) < \delta(\varepsilon)$, then $f(x) \leqslant \varepsilon$. Indeed, $x \in C^{\delta(\varepsilon)}$ and $\delta(\varepsilon)$-stability of f imply that $x \in f(C)^{\varepsilon]}$, i.e., $f(x) \leqslant \varepsilon$.

If $X = \boldsymbol{R}^n$, $\mathscr{B}(\boldsymbol{R}^n)$ is the σ-algebra of Borel subsets of X and $\mathscr{M}(X)$ is endowed with a metric ϱ, then "$\mu \in \mathscr{M}(X)$ is ε-gaussian" (ε]-gaussian) means that there exists a gaussian measure $\mu_0 \in \mathscr{M}(X)$ such that $\varrho(\mu, \mu_0) < \varepsilon$, $\varrho(\mu, \mu_0) \leqslant \varepsilon$. Similarly, "$\mu$ is ε-gaussian (ε]-gaussian) with independent components" means that there exists a measure $\mu_0 \in \mathscr{M}(\boldsymbol{R}^n)$ which is a product measure, i.e., $\mu_0 = \mu_1 \oplus \ldots \oplus \mu_n$, where $\mu_i \in \mathscr{M}(\boldsymbol{R}^1)$ are gaussian for $i = 1, 2, \ldots, n$ and $\varrho(\mu, \mu_0) < \varepsilon$ $(\varrho(\mu, \mu_0) \leqslant \varepsilon)$.

EXAMPLE 2. Let $X = Y = \boldsymbol{R}^n$ and let $f(x) = Cx$, where C is an orthonormal $n \times n$-dimensional matrix and x stands for a column vector. Let F be the transformation from $\mathscr{P}(\boldsymbol{R}^n)$ onto $\mathscr{P}(\boldsymbol{R}^n)$ given by (5.1). Then

(a) *if μ is ε-gaussian, then $F\mu$ is ε]-gaussian,*

(b) *if μ is ε-gaussian with independent components, then $F\mu$ is ε]-gaussian with independent components provided the Prokhorov metrics in $\mathscr{P}(X)$ and $\mathscr{P}(Y)$ are used.*

Clearly, f is ε-uniform and hence ε-stable. Thus, by Theorem 5.2 we obtain (a) and (b).

EXAMPLE 3. Let $\mathscr{P}(X)$ and $\mathscr{P}(Y)$ be metrized by Prokhorov metrics and $X = Y = \boldsymbol{R}^2$. Let $y_1 = x_1 + x_2$ and $y_2 = x_1 - x_2$. Then

(a) *if (X_1, X_2) has an ε-gaussian distribution, then (Y_1, Y_2) has $\sqrt{2}\varepsilon$]-gaussian distribution,*

(b) *if (X_1, X_2) has an ε-"gaussian with independent components" distribution, then (Y_1, Y_2) has $\sqrt{2}\varepsilon$]-"gaussian with independent components" distribution.*

Here $f(x_1, x_2) = (x_1 + x_2, x_1 - x_2)$ and therefore f is $(\varepsilon/\sqrt{2})$-uniform. Thus F given by (5.1) is $(\varepsilon/\sqrt{2})$-uniform, too (cf. Theorem 5.2). By Proposition 2.1.a, F is $(\varepsilon/\sqrt{2})$-stable. Now, if we take the gaussian measures on $\mathscr{B}(\boldsymbol{R}^n)$ instead of C, then the definition of $\delta(\varepsilon)$-stability yields (a). Similarly, if we take instead of C the gaussian measures on $\mathscr{B}(\boldsymbol{R}^n)$ which are products of gaussian measures on $\mathscr{B}(\boldsymbol{R}^1)$, then we obtain (b).

Let us note that Theorems 5.4–5.6 and Remarks 4–7 provide further possibilities of the use of the considered metrics.

EXAMPLE 4. Let $\mathscr{P}(X)$ and $\mathscr{P}(Y)$ be metrized by metrics $\varrho_{B*,X}$ and $\varrho_{B*,Y}$ respectively, $X = \boldsymbol{R}^n$, $Y = \boldsymbol{R}$, $x = (x_1, \ldots, x_n)$ and moreover, let

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad s = \left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{1/2}, \quad s^2 = \frac{1}{n-1}\sum_{n-1}^{n}(x_i - \bar{x})^2,$$

$$t = \frac{\bar{x}}{s}.$$

Let $\{N^n_{\theta_1, \ldots, \theta_n, \sigma}\}$ be the family of distributions of n-dimensional normal vectors with independent components. θ_i, $i = 1, \ldots, n$ and σ^2 are the mean values and variance of the ith components, respectively. If $\mu \in \mathscr{P}(\boldsymbol{R}^n)$ is a distribution of $(X_1, \ldots, X_n)$, then $\mathscr{D}(\bar{X})$, $\mathscr{D}(S)$, $\mathscr{D}(S^2)$, $\mathscr{D}(t)$ stand for the distributions of the corresponding statistics. If $N^n_{\theta_1, \ldots, \theta_n, \sigma}$ is the distribution of $(X_1, X_2, \ldots, X_n)$, then $N^1_{\bar{\theta}, \frac{1}{\sqrt{n}}\sigma}$, $\chi_{\sigma, \Sigma(\theta_i - \bar{\theta})^2}$, $\chi^2_{\sigma, \Sigma(\theta_i - \bar{\theta})^2}$, $t_{\bar{\theta}, n-1}$ stand for the distributions of $\bar{X}$, S, S^2, and t, respectively.

If $\varrho_{B*,\boldsymbol{R}^n}(\mu, \{N_{\theta_1, \ldots, \theta_n, \sigma}\}) < \varepsilon$, then

$$\varrho_{B*,\boldsymbol{R}}\big(\mathscr{D}(\bar{X}), \{N^1_{\bar{\theta}, \sigma/\sqrt{n}}\}\big) \leqslant \varepsilon, \tag{7.1}$$

$$\varrho_{B*,\boldsymbol{R}}\big(\mathscr{D}(S), \{\chi_{\sigma, \Sigma(\theta_i - \bar{\theta})^2}\}\big) \leqslant \varepsilon, \tag{7.2}$$

$$\varrho_{B*,\boldsymbol{R}}\big(\mathscr{D}(S^2), \{\chi^2_{\sigma, \Sigma(\theta_i - \bar{\theta})^2}\}\big) \leqslant \varepsilon, \tag{7.3}$$

$$\varrho_{B*,\boldsymbol{R}}\big(\mathscr{D}(t), \{t_{\bar{\theta}, n-1}\}\big) \leqslant \varepsilon. \tag{7.4}$$

Inequalities (7.1)–(7.4) folow immediately from Theorem 5.5. It is clear that similar inequalities hold when we put $\theta = \theta_1 = \ldots = \theta_n$ or $\sigma = \sigma_0$ or $\theta_1 = \ldots = \theta_n = 0$, etc. Moreover, Proposition 3.4 and Remark 6 indicate several variants of the use of other metrics. It is interesting, however, that if we wish to use here the Prokhorov metrics instead of ϱ_{B*}, then Theorem 5.2 implies inequalities (7.1) and (7.2), only. Indeed, since the functions $s^2(x)$ and $t(x)$ are not uniformly continuous, they do not lead to uniform transformations from $(\mathscr{P}(\boldsymbol{R}^n), U(\varrho_{P,\boldsymbol{R}^n}))$ into $(\mathscr{P}(\boldsymbol{R}), U(\varrho_{P,\boldsymbol{R}}))$. Thus, in the case of Prokhorov metrics, the commonly used transformations of spaces of measures induced by (5.1) and $s^2(x)$ and $t(x)$ are not stable (cf. [5]). This example illustrates the dependence of the property "to be $\delta(\varepsilon)$-stable" on the choice of metrics, too.

EXAMPLE 5. Let $X = Y = C[0, 1]$, $\varrho_X(x_1(\cdot), x_2(\cdot)) = \sup(|x_1(t) - x_2(t)|; t \in [0, 1])$, $\varrho_X = \varrho_Y$. If $f(x(\cdot))(t) = x(t) - tx(1)$ and μ is an ε-Wiener measure, then $F\mu$ is a 2ε]-Brownian bridge provided F is given by (5.1) and the Prokhorov distances in $\mathscr{P}(X)$ and $\mathscr{P}(Y)$ are used. If we use ϱ_{B*} metrics in $\mathscr{P}(X)$ and $\mathscr{P}(Y)$, then $F\mu$ is an ε-Brownian bridge.

Indeed, since f is $(\varepsilon/2)$-uniform, the first statement follows from Theorem 5.2 and the known fact that F transforms Wiener measures into Brownian bridges. The second statement follows from Theorem 5.5.

8. Applications of sensitive functions

This section contains several simple consequences of theorems given in Sections 3 and 6 on sensitive transformations. Examples 6–10 indicate problems in which those theorems are useful.

EXAMPLE 6. Let f be a real $\gamma(\varepsilon)$-sensitive function and let C, $C \neq \emptyset$, be the set of all solutions of the equation $f(x) = 0$. If

$$|g(x)| < \varepsilon \tag{8.1}$$

for each $x \in X$, then the set of all solutions of the equation

$$f(x)+g(x) = 0 \tag{8.2}$$

is contained in $C^{\gamma(\varepsilon)]}$.

Indeed, if x_0 is a solution of equation (8.2), then $x_0 \in f^{-1}\{(-\varepsilon, \varepsilon)\}$. By $\gamma(\varepsilon)$-sensitivity of f we obtain

$$f^{-1}\{(-\varepsilon, \varepsilon)\} \subset f^{-1}(0)^{\gamma(\varepsilon)]} = C^{\gamma(\varepsilon)]}.$$

Let us borrow the notation from Section 7.

EXAMPLE 7. Let X, Y, f, and F be the same as in Example 2 and let $\mathscr{P}(X)$ and $\mathscr{P}(Y)$ be metrized by Prokhorov metrics. Then

(a) *if $F\mu$ is ε-gaussian, then μ is $\varepsilon]$-gaussian,*

(b) *if $F\mu$ is ε-"gaussian with independent components", then μ is $\varepsilon]$-"gaussian with independent components".*

Since f is a one-to-one function onto Y, function f^{-1} and F^{-1} are well-defined. Moreover, f^{-1} is an ε-uniform function from Y onto $X = X$. Thus, Theorem 5.2 yields (a) and (b) (cf. the remarks at the beginning of Section 6).

Clearly, in view of Theorems 5.4–5.6 we can use here metrics ϱ_{B^*}, ϱ_{BL^*} or ϱ_M instead of ϱ_P (cf. also Remarks 4–7).

Now, let us restrict the domain of F to the set A of distributions of random vectors with independently distributed components. Suppose that X from Section 4 is just $(A, \varrho_{P,X})$ and $B \subset \mathscr{P}(Y) \cap F(A)$ is the set of distributions of random vector with independently distributed components. Then, in view of a Skitovich theorem, $C = A \cap F^{-1}(B)$ is a set of gaussian measures. Since $C = F^{-1}(B)$ and F^{-1} is ε-sensitive in the case (e.g.) of Prokhorov or Meshalkin's metrics, we obtain

$$A^{\varepsilon} \cap F^{-1}(B^{\varepsilon}) = F^{-1}(B^{\varepsilon}) \subset (F^{-1}(B))^{\varepsilon} = C^{\varepsilon},$$

i.e., such a characterization of C by A, B, and F is ε-sensitive (cf. [8]).

The following example shows consequences and lacks of Theorem 6.1.

EXAMPLE 8. Let $X = \{(i,j)\colon i,j = 0, 1, \ldots, n\}$. Let S, $S \subset X$, be the support of a two-dimensional random vector $Z(i,j) = (i,j)$. Let $Y = \{0, 1, \ldots, 2n\}$ and $f(i,j) = i+j$. Suppose that

$$\varrho_{B^*,Y}(\mathscr{D}(f(Z)), \mathscr{B}(2n, p)) < \varepsilon \tag{8.3}$$

where $\mathscr{B}(2n, p)$ stands for a binomial distribution with given parameters $2n$ and $p \in (0, 1)$. By Theorem 6.1, F^{-1} is ε-sensitive, where F is given by (5.1). This yields the existence of numbers p_{ij}, $i,j = 0, 1, \ldots, n$, such that

$$p_{ij} > 0 \quad \text{for} \quad (i,j) \in S, \quad p_{ij} = 0 \quad \text{for} \quad (i,j) \notin S,$$

$$\sum_{i,j=0}^{n} p_{ij} = 1, \quad \sum_{i+j=k} p_{ij} = \binom{2n}{k} p^k (1-p)^{2n-k}, \quad k = 0, 1, \ldots, 2n$$

and

$$\sup\{|P(Z = (i,j)) - p_{i,j}|;\ i,j = 0, 1, \ldots, n\} \leqslant \varepsilon.$$

However, if it is known that Z has independently distributed components, then Theorem 6.1 yields no information on the distance of $\mathscr{D}(Z)$ from the set of distributions of random vectors $V = (V_1, V_2)$ with V_1 and V_2 independently distributed and $f(V) \sim \mathscr{B}(2n, p)$.

EXAMPLE 9. Suppose that inequalities (7.1)–(7.4) hold and $\mathscr{D}(X) \in \mathscr{X}(\lambda)$, where λ is the Lebesgue measure in $\boldsymbol{R}^n$. Then Theorem 6.1 implies

(a) $\varrho_{B^*,R^n}(\mathscr{D}(X), G_1) \leqslant \varepsilon$ *provided* (7.1) *holds and $G_1 \subset \mathscr{X}(\lambda)$ is the set of all distributions from $\mathscr{X}(\lambda)$ such that*

$$\mathscr{D}(\overline{X}) \in \{N_{\overline{\theta}, \sigma/\sqrt{n}}\};$$

(b) *if* (7.2) *holds, then $\varrho_{B^*,R^n}(\mathscr{D}(X), G_2) \leqslant \varepsilon$, where $G_2 \subset \mathscr{X}(\lambda)$ is the set of all distributions from $\mathscr{X}(\lambda)$ such that*

$$\mathscr{D}(S) \in \{\chi_{\sigma, \Sigma(\theta_i - \overline{\theta})^2}\};$$

(c) *if* (7.3) *holds, then $\varrho_{B^*R^n}(\mathscr{D}(X), G_3) \leqslant \varepsilon$, where $G_3 \subset \mathscr{X}(\lambda)$ is the set of all distributions from $\mathscr{X}(\lambda)$ such that*

$$\mathscr{D}(S^2) \in \{\chi^2_{\sigma, \Sigma(\theta_i - \overline{\theta})^2}\};$$

(d) *if* (7.4) *holds, then $\varrho_{B^*,R^n}(\mathscr{D}(X), G_4) \leqslant \varepsilon$, where $G_4 \subset \mathscr{X}(\lambda)$ is the set of all distributions from $\mathscr{X}(\lambda)$ such that*

$$\mathscr{D}(t) \in \{t_{\overline{\theta}, n-1}\}.$$

EXAMPLE 10. Let us consider the case described in Example 5. Suppose that the measure $F\mu$ is an ε-Brownian bridge provided the distance is expressed in metric ϱ_{B^*}. If $\mu \in \mathscr{X}(W)$, where W is a given Wiener measure on $C[0, 1]$, then Theorem 6.1 implies

$$\varrho_{B^*, C[0,1]}(\mu, G_5) \leqslant \varepsilon.$$

Here G_5 is the set of all probability measures ν from $\mathscr{X}(W)$ such that $F\nu$ is just a Brownian bridge.

Note added in proof: The authors are indebted to Dr W. Szczotka for pointing out the work of W. Whitt (Z. Wahrscheinlichkeitstheorie verw. Gebiete 29 (1974), pp. 39-44) where results similar to those of Section 5 have been obtained for Lipschitz mappings.

References

[1] R. Bartoszyński and E. Pleszczyńska, *Reducibility of statistical structures and decision problems*, Preprint No 93, Inst. Math. of Polish. Acad. Sci., Warszawa 1976; also this volume pp. 29–38.
[2] R. L. Dobrushyn, *Systems of random variables defined by conditional distributions*, Theory Prob. Appl. 15 (1970), pp. 469–497 (in Russian).
[3] R. M. Dudley, *Distances of probability measures and random variables*, Ann. Math. Statist. 39 (1968), pp. 1563–1572.
[4] R. Engelking, *Topologia ogólna*, PWN, Warszawa 1976 (in Polish); *Outline of general topology*, Wiley, New York 1968.
[5] F. R. Hampel, *A qualitative definition of robustness*, Ann. Math. Statist. 42 (1971), pp. 1887–1896.
[6] —, *Robust estimation: A condensed partial survey*, Z. Warscheinlichkeitstheorie verw. Geb. 27 (1973), pp. 87–104.
[7] K. Kuratowski, *Topologie I* (second ed.), Warszawa-Wrocław 1948.
[8] L. D. Meshalkin, *On the robustness of some characterizations of the normal distribution*, Ann. Math. Statist. 39 (1968), pp. 1747–1750.
[9] K. Musiał, *Existence of proper conditional probabilities*, Z. Wahrscheinlichkeitstheorie verw. Geb. 22 (1972), pp. 8–12.
[10] J. Pfanzagl, *On the existence of regular conditional probabilities*, ibid. 11 (1969), pp. 244–256.
[11] Yu. V. Prokhorov, *Convergence of random processes and limit theorems in probability theory*, Theory Prob. Appl. 1 (1956), pp. 157–214.
[12] D. A. Sibley, *A metric for weak convergence of distribution functions*, Rocky Mountain J. of Math. 1 (1971), pp. 427–430.
[13] V. Strassen, *The existence of probability measures with given marginals*, Ann. Math. Statist. 36 (1965), pp. 423–439.
[14] S. M. Ulam, *A collection of mathematical problems*, Interscience Publishers, No 8, Inc., New York 1960.
[15] V. M. Zolotarev, *On the continuity of stochastic sequences generated by recurrent procedures*, Theory Prob. Appl. 20.4 (1975), pp. 834–847 (in Russian).
[16] —, *Stability effect of characterizations of probability distributions*, Zapisky nautchnyh seminarov LOMI. 61 (1976), pp. 38–55 (in Russian).
[17] —, *Metric distances in spaces of random variables and their distributions*, Math. Sbornik 101 (143) (1976), pp. 416–454 (in Russian).

Presented to the semester
MATHEMATICAL STATISTICS
September 15–December 18, 1976

MATHEMATICAL STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 6
PWN—POLISH SCIENTIFIC PUBLISHERS
WARSAW 1980

ON WEAK CONVERGENCE OF SEQUENCES OF MEASURES

HARALD BERGSTRÖM

Chalmers University of Technology and University of Göteborg,
Department of Mathematics, Göteborg, Sweden

1. Introduction

In [1] A. D. Alexandroff has presented an exhaustive study of weak convergence of finite measures and finite-signed measures on normal spaces and completely normal spaces. In [2] I gave a survey of Alexandroff's theory, considering measures and I there also gave some applications to weak convergence of stochastic processes into the C-space and the D-space. Here, using these theorems, I shall give a more complete presentation of the theory of weak convergence of probability measures on the C-space and D-space.

Alexandroff's main tools are linear functionals. Since weak convergence means convergence of linear functionals, it seems natural to rely heavily upon linear functionals throughout in the theory.

Alexandroff has two main theorems which I here state for measures. Here a *measure* means a finitely additive non-negative set function on an algebra in a σ-topological space. A measure is called *σ-smooth* if it is σ-additive.

THEOREM 1.1 (Alexandroff's first theorem). *Let ψ be the Stone vector lattice of bounded continuous functions from a normal σ-topological space S into the real number field $\boldsymbol{R}$ and L a non-negative bounded linear functional from ψ into $\boldsymbol{R}$. Then L determines uniquely a regular measure μ on the algebra generated by the closed sets and μ satisfies the relation*

$$L(f) = \int_S f(x)\mu(dx), \quad f \in \psi. \tag{i}$$

COROLLARY. *For a metric σ-topological space Theorem* 1.1 *remains true if $\overline{\psi}$ is changed into the Stone vector lattice ψ of uniformly continuous functions from S into $\boldsymbol{R}$ and* (i) *still holds for $f \in \psi$.*

THEOREM 1.2 (Alexandroff's second theorem). *Let S be a completely normal space and $\mathscr{S}$ the algebra generated by the closed sets. Let $\{\mu_n\}_{n=1}^{+\infty}$ be a sequence of σ-smooth finite measures on S. If $\{\mu_n\}_{n=1}^{+\infty}$ converges weakly to a measure μ on $\mathscr{S}$, then μ is σ-smooth.*