

"High Dimensional Probability II",
Eds. Gine et al., p.221–243,
Birkhauser, Boston, 2000.

A NEW WAY TO OBTAIN ESTIMATES IN THE INVARIANCE PRINCIPLE

ALEXANDER I. SAKHANENKO

ABSTRACT This paper presents a new simple method of obtaining estimates on rates of convergence in the invariance principle, which may be used in arbitrary separable linear spaces. This method is applied to one-dimensional and infinite dimensional martingales, among other examples.

1. Introduction

Let Y_1, \dots, Y_n be a sequence of random variables taking their values in a separable linear space \mathcal{X} with the norm $\|\cdot\|$. Given the joint distribution of the random variables X_1, \dots, X_n with the values in \mathcal{X} , suppose that this joint distribution is close, in some sense, to the joint distribution of the given random variables Y_1, \dots, Y_n .

On the same probability space, where the random variables Y_1, \dots, Y_n are defined, we want to construct such random variables X_1, \dots, X_n that the distance

$$(1.1) \quad \Delta_n = \max_{1 \leq k \leq n} \left\| \sum_{m=1}^k X_m - \sum_{m=1}^k Y_m \right\|$$

would be small in some sense. The latter means that the value $\mathbf{P}(\Delta_n > x)$ is to be as small as possible for all $x > 0$ or for a specially chosen $x > 0$.

Solutions to this problem are usually the most difficult parts in deriving rates of convergence in the invariance principle. The main goal of this paper is to describe a new simple method to construct the random variables X_1, \dots, X_n and Y_1, \dots, Y_n as in (1.1) with a good rate of approximation.

This method is based upon the following simple idea. Suppose we have found smoothing random variables V_0, V_1, \dots, V_n that have specially chosen distributions and such that the differences

$$(1.2) \quad \delta_k = (X_k + V_k) - (Y_k + V_{k-1}), \quad k = 1, \dots, n,$$

are small in some sense. In this case

$$(1.3) \quad \forall k \quad \sum_{m=1}^k X_m - \sum_{m=1}^k Y_m = \sum_{m=1}^k \delta_m - V_k + V_0.$$

The representation (1.3) plays the key role in the method. For example, it follows from (1.3) that

$$(1.4) \quad \Delta_n \leq \|V_0\| + \max_{1 \leq k \leq n} \|V_k\| + \max_{1 \leq k \leq n} \|\delta_1 + \cdots + \delta_k\|.$$

In particular,

$$(1.5) \quad \mathbf{P}(\Delta_n > x) \leq \mathbf{P}(\max_{0 \leq k \leq n} \|V_k\| > x/2) + \sum_{k=1}^n \mathbf{P}(\delta_k \neq 0).$$

We describe our construction in Section 2 and list the properties of the constructed random elements in Theorem 2.1. In Corollary 2.2, we estimate the probabilities $\mathbf{P}(\delta_m \neq 0)$ in (1.5) and, thus, find a very simple, but general, inequality for obtaining estimates in the invariance principles. Corollary 2.4 contains another inequality which may give better estimates but under additional assumptions.

In Section 3 we illustrate the possibilities of our method for one-dimensional and for infinite dimensional martingales. The remainder of the paper is devoted to proofs of the results stated in Sections 2 and 3.

About conventions and notations. Later on we will assume that the linear space \mathcal{X} is a complete separable metric space with respect to a specially chosen metric. We denote by \mathcal{B} the corresponding σ -algebra of the Borel subsets of the space \mathcal{X} and suppose that the function $(x, y) \mapsto \alpha x + \beta y$ is measurable with respect to $\mathcal{B} \times \mathcal{B}$ for all $\alpha, \beta \in (-\infty, \infty)$. The function $\|\cdot\|$ is treated as a real functional on \mathcal{X} measurable with respect to \mathcal{B} and such that

$$0 \leq \|\alpha x + \beta y\| \leq |\alpha| \|x\| + |\beta| \|y\| \quad \forall x, y \in \mathcal{X}, \quad \forall \alpha, \beta \in (-\infty, \infty).$$

We assume that all the distributions in this paper are probability measures defined on $(\mathcal{X}, \mathcal{B})$ or on a product of at most a countable number of $(\mathcal{X}, \mathcal{B})$ and intervals $[0, 1]$ (see, for example, Loève (1978), for details). The symbols $\mathcal{L}(Z)$, $\mathcal{L}(Z|Y)$ and $\mathcal{L}(Z, Y)$ correspond, respectively, to the (law of) distribution of the random element Z , the conditional (law of) distribution of Z with respect to the random element Y and the joint (law of) distribution of random elements Z and Y . In particular, we will write

$$\mathcal{L}(\tilde{Z}) = \mathcal{L}(Z), \quad \mathcal{L}(\tilde{Z}|Y) = \mathcal{L}(Z|Y) \quad \text{and} \quad \mathcal{L}(\tilde{Z}, \tilde{Y}) = \mathcal{L}(Z, Y)$$

whenever the corresponding distributions coincide.

Later on we denote by $\mathcal{L}(Z_1) \times \mathcal{L}(Z_2)$ the product of the distributions $\mathcal{L}(Z_1)$ and $\mathcal{L}(Z_2)$, where Z_1 and Z_2 may be random elements taking values in *arbitrary* spaces. Whereas, we denote by $\mathcal{L}(X) * \mathcal{L}(Y)$ the convolution

of the distributions of the random variables X and Y taking their values in \mathcal{X} . Thus,

$$(1.6) \quad \mathcal{L}(Z_1, Z_2) = \mathcal{L}(Z_1) \times \mathcal{L}(Z_2) \quad \text{iff} \quad Z_1 \text{ and } Z_2 \text{ are independent};$$

$$(1.7) \quad \mathcal{L}(X + Y) = \mathcal{L}(X) * \mathcal{L}(Y) \quad \text{if} \quad X \text{ and } Y \text{ are independent.}$$

Denote by $\mathbf{P}(\cdot | Y)$ and $\mathbf{E}(\cdot | Y)$ the conditional probability and the conditional expectation with respect to the random element Y . For brevity, sometimes we will write

$$\mathbf{P}_k(\cdot) = \mathbf{P}(\cdot | \vec{X}_k) \quad \text{and} \quad \mathbf{E}_k(\cdot) = \mathbf{E}(\cdot | \vec{X}_k) \quad \text{for} \quad \vec{X}_k = (X_1, \dots, X_k).$$

Let us point out, that all the conditional distributions in the paper exist (see, for example, Loève (1978) for proofs) because \mathcal{X}^k are separable for all $k \geq 1$.

Finally, denote by $\rho(\cdot, \cdot)$ the distance in variation between distributions on $(\mathcal{X}, \mathcal{B})$:

$$\rho(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{B \in \mathcal{B}} |\mathbf{P}(X \in B) - \mathbf{P}(Y \in B)|.$$

2. Description of the Method

Given a sequence of mutually independent random variables Y_1, Y_2, \dots , there are many possibilities to construct random variables X_1, X_2, \dots and V_0, V_1, \dots which will satisfy the following three conditions:

- i) for all $k \geq 1$ the constructed random variables X_1, \dots, X_k have the given distribution $\mathcal{L}(X_1, \dots, X_k)$;
- ii) for all $k \geq 0$ the constructed random variable V_k has the given distribution $\mathcal{L}(V_k)$;
- iii) the random variables V_0, Y_1, Y_2, \dots are mutually independent and for all $k \geq 1$ the random elements $\vec{X}_{k-1}, V_k, Y_{k+1}, Y_{k+2}, \dots$ are also mutually independent.

For example, we may take independent sequences $\{Y_1, Y_2, \dots\}$, $\{V_0, V_1, \dots\}$ and $\{X_1, X_2, \dots\}$. But such simply constructed X_1, X_2, \dots and V_0, V_1, \dots may not be appropriate for our purposes.

It is important to remark, that conditions i), ii), iii) and definitions (1.6) and (1.7) allow us to define the distributions

$$(2.1) \quad \mathcal{L}(Y_k + V_{k-1} | \vec{X}_{k-1}) = \mathcal{L}(Y_k) * \mathcal{L}(V_{k-1})$$

and

$$\mathcal{L}(X_k + V_k | \vec{X}_{k-1}) = \mathcal{L}(X_k | \vec{X}_{k-1}) * \mathcal{L}(V_k).$$

On the other hand, these conditions, in general, do not define the *joint* distribution $\mathcal{L}(Y_k + V_{k-1}, X_k + V_k | \vec{X}_{k-1})$ uniquely. This fact makes us wish to construct the random variables $Y_k + V_{k-1}$ and $X_k + V_k$ with a specially chosen conditional joint distribution.

For each $k \geq 1$ put

$$\begin{aligned} \rho_k = \rho_k(\vec{X}_{k-1}) &\equiv \rho(\mathcal{L}(X_k | \vec{X}_{k-1}) * \mathcal{L}(V_k), \mathcal{L}(Y_k) * \mathcal{L}(V_{k-1})) \\ (2.2) \quad &= \sup_{B \in \mathcal{B}} |\mathbf{P}(X_k + V_k \in B | \vec{X}_{k-1}) - \mathbf{P}(Y_k + V_{k-1} \in B)|. \end{aligned}$$

Other useful representations for the values ρ_k may be found in Remark 2.6 at the end of the section.

Theorem 2.1. *Given a sequence of mutually independent random variables Y_1, Y_2, \dots , it is possible to construct random variables $X_1, X_2, \dots, V_0, V_1, \dots$ which satisfy the conditions i), ii), iii) and such that the following assertion is fulfilled:*

iv) for each $k \geq 1$

$$(2.3) \quad \mathbf{P}(\delta_k \neq 0 | \vec{X}_{k-1}) = \rho_k(\vec{X}_{k-1}) \quad a.s.$$

The next simple estimate for the invariance principle follows immediately from (1.5) and (2.3).

Corollary 2.2. *The random variables constructed in Theorem 2.1 satisfy the inequalities*

$$(2.4) \quad \mathbf{P}(\Delta_n > x) \leq \mathbf{P}(\max_{0 \leq k \leq n} \|V_k\| > x/2) + \sum_{k=1}^n \mathbf{E} \rho_k \quad \forall n \geq 1,$$

$$(2.5) \quad \mathbf{P}(\Delta_\infty > x) \leq \mathbf{P}(\sup_{k \geq 0} \|V_k\| > x/2) + \sum_{k \geq 1} \mathbf{E} \rho_k,$$

where

$$\Delta_\infty = \sup_{n \geq 1} \Delta_n = \sup_{k \geq 1} \left\| \sum_{m=1}^k X_m - \sum_{m=1}^k Y_m \right\|.$$

Remark 2.3. It is evident from (1.4), (2.4) and (2.5) that if

$$\forall k \geq 0 \quad \mathbf{P}(\|V_k\| \leq x/2) = 1,$$

then

$$\forall n \leq \infty \quad \mathbf{P}(\Delta_n > x) \leq \sum_{k=1}^n \mathbf{E} \rho_k.$$

This particular case of Corollary 2.2 was proved earlier in Sakhanenko (1987) by using a more complicated analytical method.

The next proposition may give better estimates than Corollary 2.2.

Corollary 2.4. *Suppose that for some non-random real numbers π_1, π_2, \dots the following conditions are fulfilled*

$$(2.6) \quad \forall k \geq 1 \quad \rho_k \leq \pi_k \quad a.s.$$

Then for all integers $m, n \geq 1$ and for all real $x, y, z > 0$ the random variables constructed in Theorem 2.1 satisfy the inequality

$$(2.7) \quad \mathbf{P}(\Delta_n > mx + (m-1)(y+z)) \leq \left(\sum_{k=1}^n \pi_k \right)^m / m! \\ + \mathbf{P}(\max_{0 \leq k \leq n} \|V_k\| > x/2) + \mathbf{P}(\max_{1 \leq k \leq n} \|Y_k\| > y) + \mathbf{P}(\max_{1 \leq k \leq n} \|X_k\| > z).$$

Moreover, inequality (2.7) holds also for $n = \infty$ if we only replace in it all symbols $\max_{l \leq k \leq n}$ by $\sup_{k \geq l}$.

Remark 2.5. If we want to apply inequalities, obtained in Corollaries 2.2 and 2.4, we need first to find appropriate distributions $\{\mathcal{L}(V_k)\}$ and estimate the values $\{\rho_k\}$ and $\{\mathbf{P}(\|V_k\| > x/2)\}$. In the one-dimensional case it is easy to do using the well-known idea of Lindeberg (see Example 1 in Section 3 and the proof of Lemma 3.1 in Section 6 for details). If the space \mathcal{X} is infinite dimensional, then it is not easy to find appropriate distributions $\{\mathcal{L}(V_k)\}$ and estimate the $\{\rho_k\}$. (See, for example, Bentkus and Račkauskas (1984) or Paulauskas and Račkauskas (1989) where the analogous problems in CLT are investigated.) We consider below an infinite dimensional example (see Example 2 in Section 3) using an idea due to Bentkus (1985).

Remark 2.6. Assume now that the distributions $\mathcal{L}(X_k | \vec{X}_{k-1}) * \mathcal{L}(V_k)$ and $\mathcal{L}(Y_k) * \mathcal{L}(V_{k-1})$ have the densities $f_k(\cdot | \vec{X}_{k-1})$ and $f_k(\cdot)$ with respect to the same measure $\mu(\cdot)$ on $(\mathcal{X}, \mathcal{B})$. In this case (see, for example, Lindvall (1992))

$$\rho_k = \frac{1}{2} \int_{\mathcal{X}} |f_k(x | \vec{X}_{k-1}) - f_k(x)| \mu(dx).$$

By the Radon–Nikodim Theorem such densities exist if we take the measure

$$\mu(B) = \mu_k(B | \vec{X}_{k-1}) = \mathbf{P}(X_k + V_k \in B | \vec{X}_{k-1}) + \mathbf{P}(Y_k + V_{k-1} \in B).$$

Remark 2.7. In this section we investigated infinite sequences Y_1, Y_2, \dots and X_1, X_2, \dots . We may always assume that we are in this situation, since in the case of finite sequences Y_1, \dots, Y_N and X_1, \dots, X_N we can take $X_k = Y_k = 0$ for all $k > N$.

3. Examples

Given $x > 0$, we will assume in this section that the distributions $\{\mathcal{L}(V_k)\}$ are such that

$$(3.1) \quad \forall k \geq 0 \quad \mathcal{L}(V_k/x) = \mathcal{L}(V)$$

for a specially chosen distribution $\mathcal{L}(V)$.

Example 1. Let $\mathcal{X} = (-\infty, \infty)$ and assume we are given a sequence of mutually independent (and one-dimensional) random variables Y_1, Y_2, \dots . Suppose that the given joint distribution $\mathcal{L}(X_1, X_2, \dots)$ satisfies the following conditions with the probability 1

$$(3.2) \quad \mathbf{E}_{k-1} X_k = \mathbf{E} Y_k \quad \text{and} \quad \mathbf{E}_{k-1} X_k^2 = \mathbf{E} Y_k^2 < \infty \quad \forall k \geq 1.$$

In Section 6 we prove

Lemma 3.1. *There exists a distribution $\mathcal{L}(V)$ such that $\mathbf{P}(|V| \leq 1/2) = 1$ and*

$$\rho_k \leq 16 \mathbf{E}_{k-1} \min \left\{ \frac{|X_k|^3}{x^3}, \frac{X_k^2}{x^2} \right\} + 16 \mathbf{E} \min \left\{ \frac{|Y_k|^3}{x^3}, \frac{Y_k^2}{x^2} \right\} \quad \forall k \geq 1$$

if conditions (3.1) and (3.2) hold.

Using Remark 2.3 and Lemma 3.1, we immediately obtain the following simple estimate in the invariance principle with an explicit bound for an absolute constant.

Corollary 3.1. *If conditions (3.2) hold then for all $x > 0$ there exists a sequence of random variables X_1, X_2, \dots with the given joint distribution and such that for all $n \leq \infty$*

$$\begin{aligned} \mathbf{P}(\Delta_n > x) &\leq 16 \sum_{k=1}^n \left(\mathbf{E} \min \left\{ \frac{|X_k|^3}{x^3}, \frac{X_k^2}{x^2} \right\} + \mathbf{E} \min \left\{ \frac{|Y_k|^3}{x^3}, \frac{Y_k^2}{x^2} \right\} \right) \\ &\leq 16 x^{-r} \sum_{k=1}^n (\mathbf{E} |X_k|^r + \mathbf{E} |Y_k|^r) \quad \text{for all } r \in [2, 3]. \end{aligned}$$

Example 2. Now let \mathcal{X} be the space of all infinite dimensional vectors $x = (x^{(1)}, x^{(2)}, \dots)$ with the norm

$$\|x\| = \sup_{i \geq 1} |x^{(i)}| < \infty.$$

As above, assume that we are given a sequence of mutually independent random variables Y_1, Y_2, \dots defined on some probability space and taking their values in \mathcal{X} . On the same probability space, we want to construct random variables X_1, X_2, \dots with the given joint distribution $\mathcal{L}(X_1, X_2, \dots)$ and such that the distance Δ_n , introduced in (1.1), will be small in some sense.

Suppose that the given joint distribution $\mathcal{L}(X_1, X_2, \dots)$ is such that the following conditions are fulfilled with the probability 1

$$(3.3) \quad \begin{aligned} \mathbf{E}_{k-1} X_k^{(i)} &= \mathbf{E} Y_k^{(i)}, & \mathbf{E}_{k-1} (X_k^{(i)})^2 &= \mathbf{E} (Y_k^{(i)})^2 < \infty \\ \text{and} \quad \mathbf{E}_{k-1} X_k^{(i)} X_k^{(j)} &= \mathbf{E} Y_k^{(i)} Y_k^{(j)} & \forall i, j, k \geq 1. \end{aligned}$$

If, in addition, $\mathbf{E}_{k-1} X_k^{(i)} = 0$ for all $i, k \geq 1$, then such a sequence X_1, X_2, \dots may be naturally considered as an ‘‘infinite dimensional martingale’’.

For every infinite dimensional vector $x = (x^{(1)}, x^{(2)}, \dots)$ put

$$\|x\|_2 = \left(\sum_{i \geq 1} (x^{(i)})^2 \right)^{1/2}.$$

The following statement is an infinite-dimensional analog of Lemma 3.1.

Lemma 3.2. *There exist a distribution $\mathcal{L}(V)$ and an absolute constant $C < \infty$ such that $\mathbf{P}(\|V\| \leq 1/2) = 1$ and*

$$(3.4) \quad \rho_k \leq C \mathbf{E}_{k-1} \min \left\{ \frac{\|X_k\|_2^3}{x^3}, \frac{\|X_k\|_2^2}{x^2} \right\} + C \mathbf{E} \min \left\{ \frac{\|Y_k\|_2^3}{x^3}, \frac{\|Y_k\|_2^2}{x^2} \right\}$$

for all $k \geq 1$, if conditions (3.1) and (3.3) hold.

As was shown in Sakhanenko (1987), Lemma 3.2 may be easily derived from a similar assertion (with $\|\cdot\|_2^3/x^3$ instead of $\min \{\|\cdot\|_2^3/x^3, \|\cdot\|_2^2/x^2\}$) proved in Bentkus (1985). But the latter work is difficult to access. For this reason and for the convenience of the readers in Section 7 we give a proof of Lemma 3.2.

If we substitute the estimate (3.4), obtained in Lemma 3.2, into Remark 2.3 and into Corollary 2.4, we immediately obtain the following two assertions.

Corollary 3.2. *If conditions (3.3) hold then for all $x > 0$ there exists a sequence of random variables X_1, X_2, \dots with the given joint distribution and such that for all $n \leq \infty$*

$$(3.5) \quad \begin{aligned} \mathbf{P}(\Delta_n > x) &\leq C \sum_{k=1}^n \left(\mathbf{E} \min \left\{ \frac{\|X_k\|_2^3}{x^3}, \frac{\|X_k\|_2^2}{x^2} \right\} + \mathbf{E} \min \left\{ \frac{\|Y_k\|_2^3}{x^3}, \frac{\|Y_k\|_2^2}{x^2} \right\} \right) \\ &\leq C x^{-r} \sum_{k=1}^n (\mathbf{E} \|X_k\|_2^r + \mathbf{E} \|Y_k\|_2^r) \quad \text{for all } r \in [2, 3], \end{aligned}$$

where C is the same absolute constant as in Lemma 3.2.

Corollary 3.3. *Let conditions (3.3) be fulfilled and suppose that there exist non-random numbers $\{\pi_k(x)\}$ such that with probability 1*

$$\mathbf{E}_{k-1} \min \left\{ \frac{\|X_k\|_2^3}{x^3}, \frac{\|X_k\|_2^2}{x^2} \right\} + \mathbf{E} \min \left\{ \frac{\|Y_k\|_2^3}{x^3}, \frac{\|Y_k\|_2^2}{x^2} \right\} \leq \pi_k(x)$$

for each $k \geq 1$. Then for all integers $m, n \geq 1$ and for all real $y, z > 0$ the random variables X_1, X_2, \dots , constructed in Corollary 3.2, satisfy the inequality

$$(3.6) \quad \mathbf{P}(\Delta_n > mx + (m-1)(y+z)) \leq \left(C \sum_{k=1}^n \pi_k(x)\right)^m / m! \\ + \mathbf{P}\left(\max_{1 \leq k \leq n} \|Y_k\| > y\right) + \mathbf{P}\left(\max_{1 \leq k \leq n} \|X_k\| > z\right),$$

where C is the same absolute constant as in Lemma 3.2. Moreover, inequality (3.6) holds also for $n = \infty$ if we replace in it all symbols $\max_{1 \leq k \leq n}$ by $\sup_{k \geq 1}$.

Remark 3.4. We treat the results obtained in Corollaries 3.2 and 3.3 as estimates in the invariance principle for the random variables taking values in the space \mathcal{X} with the norm $\|x\|$. But these estimates depend essentially on the values $\|X_k\|_2$ and $\|Y_k\|_2$. For this reason we may say that $\|x\|_2$ is “the second norm in the space \mathcal{X} ”. (See Bentkus and Račkauskas (1984) or Paulauskas and Račkauskas (1989) for other example of spaces in which it makes sense to introduce two different norms.)

Remark 3.5. Let us point out that the method, by which the random variables X_1, X_2, \dots were constructed in Corollaries 3.1–3.3, depends very crucially on the chosen real number $x > 0$. It follows from the fact (see (3.1)) that the distributions of the smoothing random variables $\{V_k\}$ depend on x .

Remark 3.6. Earlier, in Sakhanenko (1987), Corollary 3.2 was proved by a more complicated analytical method, mentioned in Remark 2.3. In the particular case of mutually independent random elements $\{X_k\}$, a similar estimate was obtained independently by the same method, also in Bentkus and Liubinskas (1987). But Corollary 3.3 is impossible to derive by this analytical method.

Remark 3.7. If the random variables $\{X_k\}$ and $\{Y_k\}$ have moments only of the order $r \leq 3$ (see (3.5)), then the estimates, obtained in Corollaries 3.1–3.3, are unimprovable in some sense, because they are unimprovable in the one-dimensional case (see, for example, Sakhanenko (1974)). But if these random variables have moments of the order $r > 3$, then the famous but complicated method of Komlós–Major–Tusnády and its more complicated generalizations (see Zaitsev (1998), for example) allow one to obtain incomparably better estimates for sequences of independent random variables. Moreover, the less complicated method of Einmahl (1987) also gives better estimates in case $r > 3$.

Remark 3.8. Due to the simplicity of our method, it works only under the additional assumptions (3.3). Whereas the more difficult method of Morrow and Philipp (1982) and Monrad and Philipp (1991) does not require such restricted assumptions.

4. Auxiliary lemmas

In this section we assume that the random variables V, X, Y, Z, α_0 and α are defined on the same probability space and that the variables V, X and Y take their values in the complete separable metric space \mathcal{X} . Whereas, α_0 and α are uniformly distributed on the interval $[0, 1]$. The latter will be denoted, for brevity, as $\alpha \sim \mathcal{U}_{0,1}$.

Lemmas 4.1 and 4.3–4.5 are based on an idea of Skorohod (1976).

Lemma 4.1. *Let $\alpha_0 \sim \mathcal{U}_{0,1}$ be independent of the vector (X, V, Y) . Then there exist functions $\psi_1(\cdot, \cdot)$, $\psi_2(\cdot, \cdot)$ and $\psi'(\cdot, \cdot, \cdot, \cdot)$ such that for the random variable $\alpha' = \psi'(X, V, Y, \alpha_0)$ the following statements hold:*

$$X = \psi_1(Y, \alpha'), \quad V = \psi_2(Y, \alpha'), \quad Y \text{ is independent of } \alpha' \sim \mathcal{U}_{0,1}.$$

Proof. We may treat the pairs (X, V) and (Y, Y) as random elements taking values in the complete separable metric space $\mathcal{X} \times \mathcal{X}$. Thus, Lemma 4.1 is a particular case of the main result from Skorohod (1976).

The following assertion is well-known as Dobrushin's lemma (see, for example, Dobrushin (1970) or Lindvall (1992)).

Lemma 4.2. *For each pair of distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ there exist random variables \widehat{X} and \widehat{Y} such that*

$$\mathcal{L}(\widehat{X}) = \mathcal{L}(X), \quad \mathcal{L}(\widehat{Y}) = \mathcal{L}(Y) \quad \text{and} \quad \mathbf{P}(\widehat{X} \neq \widehat{Y}) = \rho(\mathcal{L}(X), \mathcal{L}(Y)).$$

Using Lemma 4.1, we may rewrite Lemma 4.2 in another known form.

Lemma 4.3. *Given a random variable X and a distribution $\mathcal{L}(Y)$. Then there exists a random variable \widehat{Y} such that*

$$\mathcal{L}(\widehat{Y}) = \mathcal{L}(Y) \quad \text{and} \quad \mathbf{P}(X \neq \widehat{Y}) = \rho(\mathcal{L}(X), \mathcal{L}(Y)).$$

Now we prove the main assertion of the section.

Lemma 4.4. *There exist functions $\Psi_1(\cdot, \cdot, \cdot, \cdot, \cdot)$ and $\Psi_2(\cdot, \cdot, \cdot, \cdot, \cdot)$ such that the following statement is valid: if a random variable Y is independent of $\alpha \sim \mathcal{U}_{0,1}$, then*

$$(4.1) \quad \mathcal{L}(\widetilde{X}, \widetilde{V}) = \mathcal{L}(X) \times \mathcal{L}(V), \quad \mathbf{P}(\widetilde{X} + \widetilde{V} \neq Y) = \rho(\mathcal{L}(X) * \mathcal{L}(V), \mathcal{L}(Y))$$

for

$$(4.2) \quad \widetilde{X} = \Psi_1(Y, \alpha, \mathcal{L}(X), \mathcal{L}(V), \mathcal{L}(Y)), \quad \widetilde{V} = \Psi_2(Y, \alpha, \mathcal{L}(X), \mathcal{L}(V), \mathcal{L}(Y)).$$

Proof. Let random variables V and X be independent. Using Lemma 4.3, we may construct a random variable \widehat{Y} such that

$$(4.3) \quad \mathcal{L}(\widehat{Y}) = \mathcal{L}(Y) \quad \text{and} \quad \mathbf{P}(X + V \neq \widehat{Y}) = \rho(\mathcal{L}(X) * \mathcal{L}(V), \mathcal{L}(Y)).$$

Now let $\alpha_0 \sim \mathcal{U}_{0,1}$ be independent of the random vector (X, V, \widehat{Y}) . By Lemma 4.1, there exist functions $\psi_1(\cdot, \cdot)$, $\psi_2(\cdot, \cdot)$ and a random variable α' such that

$$(4.4) \quad X = \psi_1(\widehat{Y}, \alpha'), \quad V = \psi_2(\widehat{Y}, \alpha'), \quad \widehat{Y} \text{ is independent of } \alpha' \sim \mathcal{U}_{0,1}.$$

Put

$$(4.5) \quad \widetilde{X} = \psi_1(Y, \alpha) \quad \text{and} \quad \widetilde{V} = \psi_2(Y, \alpha).$$

Comparing (4.4) and (4.5), it is easy to see that the random vectors

$$(\widehat{Y}, \alpha', X, V) \quad \text{and} \quad (Y, \alpha, \widetilde{X}, \widetilde{V})$$

are identically distributed because the pairs (\widehat{Y}, α') and (Y, α) are identically distributed by assumption. This fact and (4.3) imply that all the statements in (4.1) are fulfilled for \widetilde{X} and \widetilde{V} defined by (4.5).

Note that we may construct such functions $\psi_1(\cdot, \cdot)$ and $\psi_2(\cdot, \cdot)$ for all distributions $\mathcal{L}(X)$, $\mathcal{L}(V)$ and $\mathcal{L}(Y)$. So, we may rewrite (4.5) in the form (4.2).

If we replace all distributions by conditional ones in Lemma 4.4, we obtain the following assertion.

Lemma 4.5. *Let a random element (Y, Z) be independent of $\alpha \sim \mathcal{U}_{0,1}$, where Z takes its values in \mathcal{X}^k for some $k \geq 1$. Put*

$$\widetilde{X} = \Psi_1(Y, \alpha, \mathcal{L}(X|Z), \mathcal{L}(V), \mathcal{L}(Y|Z)),$$

$$\widetilde{V} = \Psi_2(Y, \alpha, \mathcal{L}(X|Z), \mathcal{L}(V), \mathcal{L}(Y|Z)).$$

Then

$$\mathcal{L}((\widetilde{X}, \widetilde{V})|Z) = \mathcal{L}(X|Z) \times \mathcal{L}(V),$$

$$\mathbf{P}(\widetilde{X} + \widetilde{V} \neq Y|Z) = \rho(\mathcal{L}(X|Z) * \mathcal{L}(V), \mathcal{L}(Y|Z)).$$

5. Proofs of results of Section 2

Suppose that we are given a set of mutually independent random variables

$$(5.1) \quad V_0, Y_1, Y_2, \dots, \alpha_1, \alpha_2, \dots,$$

where V_0 has the desired distribution $\mathcal{L}(V_0)$ and $\alpha_k \sim \mathcal{U}_{0,1}$ for all $k \geq 1$.

Put

$$(5.2) \quad X_1 = \Psi_1(Y_1 + V_0, \alpha_1, \mathcal{L}(X_1), \mathcal{L}(V_1), \mathcal{L}(Y_1) * \mathcal{L}(V_0)),$$

$$(5.3) \quad V_1 = \Psi_2(Y_1 + V_0, \alpha_1, \mathcal{L}(X_1), \mathcal{L}(V_1), \mathcal{L}(Y_1) * \mathcal{L}(V_0)).$$

For $k > 1$ we define the random variables $\{X_k, V_k\}$ recursively, by setting $k = 2, 3, \dots$ in the following formulae

$$(5.4) \quad X_k = \Psi_1(Y_k + V_{k-1}, \alpha_k, \mathcal{L}(X_k|\vec{X}_{k-1}), \mathcal{L}(V_k), \mathcal{L}(Y_k) * \mathcal{L}(V_{k-1})),$$

$$(5.5) \quad V_k = \Psi_2(Y_k + V_{k-1}, \alpha_k, \mathcal{L}(X_k|\vec{X}_{k-1}), \mathcal{L}(V_k), \mathcal{L}(Y_k) * \mathcal{L}(V_{k-1})).$$

Now we shall investigate the properties of the random variables constructed in (5.2)–(5.5).

Lemma 5.1. *For all $k > 1$ the sequence $X_1, \dots, X_{k-1}, V_{k-1}$ is independent of the sequence Y_k, Y_{k+1}, \dots of mutually independent random variables.*

Proof. It follows from (5.2)–(5.5) that $X_1, \dots, X_{k-1}, V_{k-1}$ are the functions only of the variables $V_0, Y_1, \dots, Y_{k-1}, \alpha_1, \dots, \alpha_{k-1}$. This fact and the mutual independence of the random variables in (5.1) give us the assertion of Lemma 5.1.

Lemma 5.2. *If for some $k > 1$*

$$(5.6) \quad \mathcal{L}(X_1, \dots, X_{k-1}, V_{k-1}) = \mathcal{L}(X_1, \dots, X_{k-1}) \times \mathcal{L}(V_{k-1}),$$

then, for given k , condition (2.3) holds and

$$(5.7) \quad \mathcal{L}(X_1, \dots, X_k, V_k) = \mathcal{L}(X_1, \dots, X_k) \times \mathcal{L}(V_k).$$

Proof. Assumption (5.6) and Lemma 5.1 imply that condition (2.1) holds. Hence, we may rewrite (5.4) and (5.5) in the following way

$$(5.8) \quad X_k = \Psi_1(Y_k + V_{k-1}, \alpha_k, \mathcal{L}(X_k|\vec{X}_{k-1}), \mathcal{L}(V_k), \mathcal{L}(Y_k + V_{k-1}|\vec{X}_{k-1})),$$

$$(5.9) \quad V_k = \Psi_2(Y_k + V_{k-1}, \alpha_k, \mathcal{L}(X_k|\vec{X}_{k-1}), \mathcal{L}(V_k), \mathcal{L}(Y_k + V_{k-1}|\vec{X}_{k-1})).$$

Lemma 4.5 and equalities (5.8)–(5.9) yield all the assertions of Lemma 5.2.

Proof of Theorem 2.1. Lemma 4.4 and definitions (5.2) and (5.3) imply that conditions (5.7) and (2.3) hold for $k = 1$. By induction on k , from Lemma 5.2 we have that properties (5.7) and (2.3) are satisfied for all $k \geq 1$. Hence, conditions i), ii) and iv) are fulfilled. After that, from (5.7) and Lemma 5.1 we obtain iii).

Remark 5.1. In the proof of Theorem 2.1 we have introduced a set (5.1) of mutually independent random variables. It means that we have enlarged the initial probability space on which the given random variables Y_1, Y_2, \dots were defined. It is well-known (see, for example, Skorohod (1976)) that we do not need to enlarge the initial probability space in the case when it is “rich enough” in the following sense: on this probability space it is possible to define a random variable uniformly distributed on the interval $[0, 1]$, which is independent of all given random variables.

Proof of Corollary 2.4. Denote by I_k the indicator of the event $\{\delta_k \neq 0\}$. It follows from (1.4) that

$$(5.10) \quad \Delta_n \leq \|V_0\| + \max_{1 \leq k \leq n} \|V_k\| + \max_{1 \leq k \leq n} \|\delta_k\| \cdot \sum_{k=1}^n I_k.$$

Thus, from (5.10) and (1.2) we obtain the inequality

$$(5.11) \quad \mathbf{P}(\Delta_n > x + (m-1)(x+y+z)) \leq \mathbf{P}\left(\sum_{k=1}^n I_k > m-1\right) \\ + \mathbf{P}\left(\max_{0 \leq k \leq n} \|V_k\| > x/2\right) + \mathbf{P}\left(\max_{1 \leq k \leq n} \|Y_k\| > y\right) + \mathbf{P}\left(\max_{1 \leq k \leq n} \|X_k\| > z\right).$$

Corollary 2.4 follows immediately from (5.11) and from the following auxiliary Lemma 5.3 for $l = 1$.

Lemma 5.3. *If conditions (2.6) are fulfilled then for all integers $m \geq 1$ and for all $1 \leq l \leq n$*

$$(5.12) \quad \mathbf{P}_{l-1}\left(\sum_{k=l}^n I_k > m-1\right) \leq \left(\sum_{k=l}^n \pi_k\right)^m / m!.$$

Proof. The law of total probability gives

$$(5.13) \quad \mathbf{P}_{l-1}\left(\sum_{k=l}^n I_k > m-1\right) \leq \sum_{k=l}^n \mathbf{P}_{l-1}\left(\delta_k \neq 0, \sum_{i=k+1}^n I_i > m-2\right)$$

for all integers $m \geq 1$ and $1 \leq l \leq n$. On the other hand

$$(5.14) \quad \sum_{k=l}^n \pi_k \left(\sum_{i=k+1}^n \pi_i\right)^{m-1} / (m-1)! \leq \left(\sum_{k=l}^n \pi_k\right)^m / m!.$$

Inequalities (5.13) and (5.14) allow us to prove (5.12) using induction on m .

6. Proof of Lemma 3.1

Introduce the density

$$(6.1) \quad g(x) = 4((1-2|x|)^+)^2 - 2((1-4|x|)^+)^2.$$

We need the following auxiliary

Lemma 6.1. *Let random variables X and Y be such that*

$$(6.2) \quad \mathbf{E}X = \mathbf{E}Y \quad \text{and} \quad \mathbf{E}X^2 = \mathbf{E}Y^2 < \infty.$$

Suppose that a random variable V has the density $g(\cdot)$ and is independent of the random variables X and Y . Then $\mathbf{P}(|V| < 1/2) = 1$ and

$$(6.3) \quad \rho \equiv \rho(\mathcal{L}(X) * \mathcal{L}(V), \mathcal{L}(Y) * \mathcal{L}(V)) \\ \leq 16 \mathbf{E} \min\{|X|^3, X^2\} + 16 \mathbf{E} \min\{|Y|^3, Y^2\}.$$

Proof. The assertion $\mathbf{P}(|V| < 1/2) = 1$ follows from the definition (6.1) of the density $g(\cdot)$ concentrated on the interval $(-1/2, 1/2)$. So we need only establish (6.3).

Remark 2.6 and the convolution formula yield

$$(6.4) \quad \rho = \frac{1}{2} \int_{-\infty}^{\infty} |\mathbf{E}g(x-X) - \mathbf{E}g(x-Y)| dx.$$

For

$$(6.5) \quad \varepsilon_0(x, y) = g(x-y) - g(x) + yg'(x) - y^2g''(x)/2$$

we immediately obtain

$$(6.6) \quad \mathbf{E}g(x-Y) = g(x) - g'(x)\mathbf{E}Y + g''(x)\mathbf{E}Y^2/2 + \mathbf{E}\varepsilon_0(x, Y),$$

$$\mathbf{E}g(x-X) = g(x) - g'(x)\mathbf{E}X + g''(x)\mathbf{E}X^2/2 + \mathbf{E}\varepsilon_0(x, X)$$

$$(6.7) \quad = g(x) - g'(x)\mathbf{E}Y + g''(x)\mathbf{E}Y^2/2 + \mathbf{E}\varepsilon_0(x, X).$$

The last equality in (6.7) is a corollary of conditions (6.2). Using (6.4), (6.6) and (6.7), we have

$$(6.8) \quad \rho = \frac{1}{2} \int_{-\infty}^{\infty} |\mathbf{E}\varepsilon_0(x, X) - \mathbf{E}\varepsilon_0(x, Y)| dx \leq \mathbf{E}\varepsilon_1(X) + \mathbf{E}\varepsilon_1(Y),$$

where

$$(6.9) \quad \varepsilon_1(y) = \frac{1}{2} \int_{-\infty}^{\infty} |\varepsilon_0(x, y)| dx.$$

Assertion (6.3) follows now from (6.8) and the inequality

$$(6.10) \quad \varepsilon_1(y) \leq 16 \min\{|y|^3, y^2\},$$

which will be proved below.

Lemma 6.2. *For the function $g(\cdot)$ defined in (6.1) inequality (6.10) holds.*

Proof. By definition (6.1), the function $g(x)$ has continuous derivative $g'(x)$ and its second derivative $g''(x)$ has the following simple form

$$(6.11) \quad g''(x) = 32\mathbf{I}_{[-1/2, 1/2]}(x) - 64\mathbf{I}_{[-1/4, 1/4]}(x),$$

where $\mathbf{I}_A(\cdot)$ denotes the indicator of the set A . Put

$$(6.12) \quad \varepsilon_2(x, y) = g''(x-y) - g''(x), \quad \varepsilon_3(y) = \int_{-\infty}^{\infty} |\varepsilon_2(x, y)| dx.$$

Definitions (6.5), (6.12) and a Taylor expansion yield

$$(6.13) \quad \begin{aligned} \varepsilon_0(x, y) &= \int_0^1 (1-t) y^2 g''(x-ty) dt - y^2 g''(x)/2 \\ &= \int_0^1 (1-t) y^2 \varepsilon_2(x, ty) dt. \end{aligned}$$

Equalities (6.9), (6.12) and (6.13) imply the inequality

$$(6.14) \quad \varepsilon_1(y) \leq \frac{1}{2} \int_0^1 (1-t) y^2 \varepsilon_3(ty) dt.$$

Using definitions (6.12) and (6.11), we obtain

$$\begin{aligned} \varepsilon_3(y) &\leq \int_{-\infty}^{\infty} |g''(x-y)| dx + \int_{-\infty}^{\infty} |g''(x)| dx \\ &= 2 \int_{-\infty}^{\infty} |g''(x)| dx = 2 \int_{-1/2}^{1/2} 32 dx = 64. \end{aligned}$$

So, by (6.14)

$$(6.15) \quad \varepsilon_1(y) \leq \frac{1}{2} \int_0^1 (1-t) \cdot 64 y^2 dt = 16 y^2.$$

On the other hand, the explicit formulae (6.11) and (6.12) give us the following representation

$$(6.16) \quad \begin{aligned} \varepsilon_2(x, y) &= 32 (\mathbf{I}_{[-1/2+y, 1/2+y]}(x) - \mathbf{I}_{[-1/2, 1/2]}(x)) \\ &\quad - 64 (\mathbf{I}_{[-1/4+y, 1/4+y]}(x) - \mathbf{I}_{[-1/4, 1/4]}(x)). \end{aligned}$$

Using (6.16) and definition (6.12) it is not difficult to see that

$$\varepsilon_3(y) \leq \int_{-1/2}^{-1/2+y} 32 dx + \int_{1/2}^{1/2+y} 32 dx + \int_{-1/4}^{-1/4+y} 64 dx + \int_{1/4}^{1/4+y} 64 dx = 192 y$$

for any $y > 0$. Hence, by (6.14), we obtain that

$$\varepsilon_1(y) \leq \frac{1}{2} \int_0^1 (1-t) t \cdot 192 y^3 dt = 16 y^3$$

for $y > 0$. Similarly, $\varepsilon_1(y) \leq 16 |y|^3$ for $y < 0$ and, thus, for all y . The latter fact, together with (6.15), imply (6.10).

Proof of Lemma 3.1. It follows immediately from (2.2) and (3.1) that

$$(6.17) \quad \begin{aligned} \rho_k &= \rho(\mathcal{L}((X_k/x)|\vec{X}_{k-1}) * \mathcal{L}(V_k/x), \mathcal{L}(Y_k/x) * \mathcal{L}(V_{k-1}/x)) \\ &= \rho(\mathcal{L}((X_k/x)|\vec{X}_{k-1}) * \mathcal{L}(V), \mathcal{L}(Y_k/x) * \mathcal{L}(V)). \end{aligned}$$

Thus, to prove Lemma 3.1, we need only to put $\mathcal{L}(X) = \mathcal{L}((X_k/x)|\vec{X}_{k-1})$ and $\mathcal{L}(Y) = \mathcal{L}(Y_k/x)$ in Lemma 6.1 and to use conditions (3.2) instead of (6.2).

7. Proof of Lemma 3.2

Later on we suppose that $V^{(1)}, V^{(2)}, \dots$ are independent and identically distributed random variables with the following density:

$$(7.1) \quad f(x) = C_0((1/2 - |x|)^+)^4 - 8((1/4 - |x|)^+)^4), \quad C_0 = 5 \cdot 2^6/3.$$

In particular, $f(x) = 0$ for $|x| \geq 1/2$ and

$$(7.2) \quad f(x) = C_0(1/32 - 3x^2/2 + 6|x|^3 - 7x^4) \quad \text{for } |x| \leq 1/4.$$

It is easy to see that the derivatives $f'(x)$ and $f''(x)$ exist and they are continuous for all $x \in (-\infty, +\infty)$, and that $f'''(x)$ exists and it is continuous for all $x \neq 0$. Put

$$(7.3) \quad \lambda_1(x) = (\ln f(x))', \quad \lambda_2(x) = (\ln f(x))'', \quad \lambda_3(x) = (\ln f(x))'''$$

for $|x| < 1/2$ and let

$$(7.4) \quad \lambda_1(x) = \lambda_2(x) = \lambda_3(x) = f(x) = 0 \quad \text{for } |x| \geq 1/2.$$

Using (7.1)–(7.4), it is not difficult to verify that the following expectations are finite:

$$(7.5) \quad C_1 = \mathbf{E} \lambda_1^2(V^{(1)}) = \int_{-1/4}^{1/4} \lambda_1^2(x) f(x) dx + 2 \cdot 4^2 C_0 \int_{1/4}^{1/2} (1/2 - x)^2 dx,$$

$$(7.6) \quad C_2 = \mathbf{E} \lambda_2^2(V^{(1)}) = \int_{-1/4}^{1/4} \lambda_2^2(x) f(x) dx + 2 \cdot 4^2 C_0 \int_{1/4}^{1/2} (1/2 - x)^0 dx,$$

$$(7.7) \quad C_3 = \mathbf{E} |\lambda_3(V^{(1)})| = \int_{-1/4}^{1/4} |\lambda_3(x)| f(x) dx + 2 \cdot 8 C_0 \int_{1/4}^{1/2} (1/2 - x)^1 dx,$$

$$(7.8) \quad C_4 = \mathbf{E} \lambda_1^4(V^{(1)}) = \int_{-1/4}^{1/4} \lambda_1^4(x) f(x) dx + 2 \cdot 4^4 C_0 \int_{1/4}^{1/2} (1/2 - x)^0 dx.$$

For n fixed later on, let R^n denote the n -dimensional Euclidean space, $\mu(\cdot)$ the n -dimensional Lebesgue measure and set

$$z = (z^{(1)}, \dots, z^{(n)}), \quad |z|_n = \left(\sum_{i=1}^n (z^{(i)})^2 \right)^{1/2}, \quad \text{for } z \in R^n.$$

For $z, u \in R^n$ introduce the following notation:

$$(7.9) \quad L_m(z, u) = \sum_{k=1}^n (-u^{(k)})^m \lambda_m(z^{(k)}), \quad m = 1, 2, 3,$$

$$(7.10) \quad p(z) = \prod_{k=1}^n f(z^{(k)}), \quad p(t, z, u) = p(z - tu),$$

where $t \in [0, 1]$. Let $p'(\cdot, \cdot, \cdot)$, $p''(\cdot, \cdot, \cdot)$ and $p'''(\cdot, \cdot, \cdot)$ be the derivatives of the function $p(\cdot, \cdot, \cdot)$ with respect to its first argument. It is easy to obtain that

$$(7.11) \quad p'(t, z, u) = L_1(z - tu, u) p(z - tu),$$

$$(7.12) \quad p''(t, z, u) = (L_2(z - tu, u) + L_1^2(z - tu, u)) p(z - tu),$$

$$(7.13) \quad \begin{aligned} p'''(t, z, u) = & (L_3(z - tu, u) + 3L_2(z - tu, u)L_1(z - tu, u) \\ & + L_1^3(z - tu, u)) p(z - tu). \end{aligned}$$

Put

$$(7.14) \quad \varepsilon_0(z, u) = p(z - u) - p(z) - p'(0, z, u) - p''(0, z, u)/2,$$

$$(7.15) \quad \varepsilon_1(u) = \frac{1}{2} \int_{R^n} |\varepsilon_0(z, u)| \mu(dz).$$

Lemma 7.1. *There exists an absolute constant $C < \infty$ such that*

$$(7.16) \quad \varepsilon_1(u) \leq C \min \{ |u|_n^3, |u|_n^2 \} \quad \text{for all } u \in R^n.$$

Proof. We will use the following two partial cases of the Taylor expansion formula:

$$\begin{aligned} p(1) - p(0) - p'(0) - p''(0)/2 &= \\ &= \int_0^1 (1-t) p''(t) dt - p''(0)/2 = \int_0^1 (1-t)^2 p'''(t) dt/2. \end{aligned}$$

Hence, from definition (7.14), we have

$$(7.17) \quad \varepsilon_0(z, u) = \int_0^1 (1-t) p''(t, z, u) dt - p''(0, z, u)/2$$

$$(7.18) \quad = \int_0^1 (1-t)^2 p'''(t, z, u) dt/2.$$

Put

$$(7.19) \quad \beta(u) = \int_{R^n} |p''(t, z, u)| \mu(dz), \quad \gamma(u) = \int_{R^n} |p'''(t, z, u)| \mu(dz).$$

Equalities (7.12) and (7.13) allow us to rewrite (7.19) in the following way

$$(7.20) \quad \beta(u) = \int_{R^n} |p''(0, z - tu, u)| \mu(dz) = \int_{R^n} |p''(0, z, u)| \mu(dz),$$

$$(7.21) \quad \gamma(u) = \int_{R^n} |p'''(0, z - tu, u)| \mu(dz) = \int_{R^n} |p'''(0, z, u)| \mu(dz).$$

Thus, from (7.15), (7.18) (7.19) and (7.21) we immediately obtain that

$$(7.22) \quad \varepsilon_1(u) \leq \int_0^1 (1-t)^2 \gamma(u) dt/4 = \gamma(u)/12,$$

and, in a similar way, (7.15), (7.17), (7.19) and (7.20) yield

$$(7.23) \quad \varepsilon_1(u) \leq \int_0^1 (1-t) \beta(u) dt/2 + \beta(u)/4 = \beta(u)/2.$$

Equalities (7.20) and (7.12) give us the formula

$$(7.24) \quad \beta(u) = \int_{R^n} |L_2(z, u) + L_1^2(z, u)| p(z) \mu(dz).$$

Similarly, from (7.21) and (7.13) we have

$$(7.25) \quad \gamma(u) = \int_{R^n} |L_3(z, u) + 3 L_2(z, u) L_1(z, u) + L_1^3(z, u)| p(z) \mu(dz).$$

It is obvious that the function $p(\cdot)$, defined in (7.10), is the density of the random vector $V = (V^{(1)}, \dots, V^{(n)})$. Hence, we may rewrite (7.24) and (7.25) in the following way

$$(7.26) \quad \beta(u) = \mathbf{E} |L_2(V, u) + L_1^2(V, u)|,$$

$$(7.27) \quad \gamma(u) = \mathbf{E} |L_3(V, u) + 3 L_2(V, u) L_1(V, u) + L_1^3(V, u)|.$$

The random variables $\lambda_1(V^{(1)}), \dots, \lambda_1(V^{(n)})$ are independent and identically distributed with zero means. This fact, (7.5), (7.8) and (7.9) yield

$$(7.28) \quad \mathbf{E} L_1^2(V, u) = \sum_{k=1}^n (u^{(k)})^2 \mathbf{E} \lambda_1^2(V^{(k)}) = C_1 |u|_n^2,$$

$$(7.29) \quad \begin{aligned} \mathbf{E} L_1^4(V, u) &= 3(\mathbf{E} L_1^2(V, u))^2 + \sum_{k=1}^n (u^{(k)})^4 (\mathbf{E} \lambda_1^4(V^{(k)}) - 3(\mathbf{E} \lambda_1^2(V^{(k)}))^2) \\ &\leq (3C_1^2 + (C_4 - 3C_1^2)^+) |u|_n^4 \equiv C_5 |u|_n^4. \end{aligned}$$

The triangle inequality, (7.6) and (7.9) give

$$(7.30) \quad (\mathbf{E} L_2^2(V, u))^{1/2} \leq \sum_{k=1}^n (\mathbf{E} ((-u^{(k)})^2 \lambda_2(V^{(k)}))^2)^{1/2} = C_2^{1/2} |u|_n^2.$$

And it is evident from (7.7) and (7.9) that

$$(7.31) \quad \mathbf{E} |L_3(V, u)| \leq \sum_{k=1}^n |u^{(k)}|^3 \mathbf{E} |\lambda_3(V^{(k)})| \leq C_3 |u|_n^3.$$

Using (7.27) and (7.29)–(7.31), we obtain

$$(7.32) \quad \begin{aligned} \gamma(u) &\leq \mathbf{E} |L_3(V, u)| + 3(\mathbf{E} L_2^2(V, u))^{1/2} (\mathbf{E} L_1^2(V, u))^{1/2} \\ &+ (\mathbf{E} L_1^4(V, u))^{3/4} \leq C_6 |u|_n^3, \quad C_6 = C_3 + 3C_2^{1/2} C_1^{1/2} + C_5^{3/4}. \end{aligned}$$

Similarly, by (7.26), (7.28) and (7.30) we have

$$(7.33) \quad \beta(u) \leq (\mathbf{E} L_2^2(V, u))^{1/2} + \mathbf{E} L_1^2(V, u) \leq (C_1 + C_2^{1/2}) |u|_n^2.$$

Finally, formulae (7.22), (7.23), (7.32) and (7.33) yield the desired inequality

$$\varepsilon_1(u) \leq \min \{ \beta(u)/2, \gamma(u)/12 \} \leq C \min \{ |u|_n^3, |u|_n^2 \}$$

with $C = \max \{ (C_1 + C_2^{1/2})/2, C_6/12 \}$.

Let $X = (X^{(1)}, \dots, X^{(n)})$ and $Y = (Y^{(1)}, \dots, Y^{(n)})$ be n -dimensional random vectors, independent of the given vector $V = (V^{(1)}, \dots, V^{(n)})$.

Lemma 7.2. *If*

$$(7.34) \quad \begin{aligned} \mathbf{E} X^{(i)} &= \mathbf{E} Y^{(i)}, & \mathbf{E} (X^{(i)})^2 &= \mathbf{E} (Y^{(i)})^2 < \infty \\ \text{and} & & \mathbf{E} X^{(i)} X^{(j)} &= \mathbf{E} Y^{(i)} Y^{(j)} \end{aligned}$$

for all $i, j = 1, \dots, n$, then

$$\begin{aligned} \rho(n) &\equiv \rho(\mathcal{L}(X) * \mathcal{L}(V), \mathcal{L}(Y) * \mathcal{L}(V)) \leq \\ &\leq C \mathbf{E} \min \{ |X|_n^3, |X|_n^2 \} + C \mathbf{E} \min \{ |Y|_n^3, |Y|_n^2 \}, \end{aligned}$$

where C is the constant from Lemma 7.1.

Proof. Conditions (7.34) and definitions (7.9) imply the equalities

$$\mathbf{E}L_1(z, X) = \mathbf{E}L_1(z, Y) \quad \text{and} \quad \mathbf{E}L_2(z, X) = \mathbf{E}L_2(z, Y)$$

for all non-random z . Hence, using (7.11), (7.12) and (7.14) we obtain

$$\mathbf{E}p'(0, z, X) + \mathbf{E}p''(0, z, X)/2 = \mathbf{E}p'(0, z, Y) + \mathbf{E}p''(0, z, Y)/2,$$

$$(7.35) \quad \mathbf{E}\varepsilon_0(z, X) - \mathbf{E}\varepsilon_0(z, Y) = \mathbf{E}p(z - X) - \mathbf{E}p(z - Y).$$

By the convolution formula, the random vectors $X + V$ and $Y + V$ have the densities $\mathbf{E}p(z - X)$ and $\mathbf{E}p(z - Y)$, respectively. This fact and Remark 2.6 give us the equality

$$(7.36) \quad \rho(n) = \frac{1}{2} \int_{R^n} |\mathbf{E}p(z - X) - \mathbf{E}p(z - Y)| \mu(dz).$$

Using (7.15), (7.35) and (7.36) we immediately obtain

$$(7.37) \quad \begin{aligned} \rho(n) &\leq \frac{1}{2} \int_{R^n} |\mathbf{E}\varepsilon_0(z, X)| \mu(dz) + \frac{1}{2} \int_{R^n} |\mathbf{E}\varepsilon_0(z, Y)| \mu(dz) \\ &\leq \frac{1}{2} \mathbf{E} \int_{R^n} |\varepsilon_0(z, X)| \mu(dz) + \frac{1}{2} \mathbf{E} \int_{R^n} |\varepsilon_0(z, Y)| \mu(dz) \\ &= \mathbf{E}\varepsilon_1(X) + \mathbf{E}\varepsilon_1(Y). \end{aligned}$$

Lemma 7.2 follows from (7.16) and (7.37).

Denote by R^∞ the set of all infinite dimensional vectors $z = (z^{(1)}, z^{(2)}, \dots)$. Let

$$z = (z^{(1)}, z^{(2)}, \dots), \quad z_{(n)} = (z^{(1)}, \dots, z^{(n)}), \quad |z|_n = |z_{(n)}|_n, \quad \text{for } z \in R^\infty.$$

We will consider R^∞ as a metric space with the metric

$$d(z, u) = \sum_{k=1}^{\infty} 2^{-k} \frac{|z^{(k)} - u^{(k)}|}{1 + |z^{(k)} - u^{(k)}|}.$$

It is well known, that in this case R^∞ is a complete separable metric space.

Let $X = (X^{(1)}, X^{(2)}, \dots)$ and $Y = (Y^{(1)}, Y^{(2)}, \dots)$ be infinite dimensional random vectors, independent of the given vector $V = (V^{(1)}, V^{(2)}, \dots)$. Later on we will treat X , Y and V as random elements taking values in R^∞ .

Lemma 7.3. *If conditions (7.34) hold for all $i, j = 1, 2, \dots$, then*

$$(7.38) \quad \begin{aligned} \rho &\equiv \rho(\mathcal{L}(X) * \mathcal{L}(V), \mathcal{L}(Y) * \mathcal{L}(V)) \leq \\ &\leq C \mathbf{E} \min \{ \|X\|_2^3, \|X\|_2^2 \} + C \mathbf{E} \min \{ \|Y\|_2^3, \|Y\|_2^2 \}. \end{aligned}$$

Proof. Put

$$(7.39) \quad \xi = X + V, \quad \eta = Y + V, \quad \rho(n) = \rho(\mathcal{L}(\xi_{(n)}), \mathcal{L}(\eta_{(n)})).$$

It follows from Lemma 7.2 that

$$(7.40) \quad \begin{aligned} \rho(n) &\leq C \mathbf{E} \min \{|X|_n^3, |X|_n^2\} + C \mathbf{E} \min \{|Y|_n^3, |Y|_n^2\} \\ &\leq C \mathbf{E} \min \{\|X\|_2^3, \|X\|_2^2\} + C \mathbf{E} \min \{\|Y\|_2^3, \|Y\|_2^2\} \equiv \delta. \end{aligned}$$

By Dobrushin's Lemma 4.3, for each n there exists an n -dimensional random vector $\tilde{\xi}_{(n)}$ such that

$$(7.41) \quad \mathcal{L}(\tilde{\xi}_{(n)}) = \mathcal{L}(\xi_{(n)}), \quad \mathbf{P}(\tilde{\xi}_{(n)} \neq \eta_{(n)}) = \rho(n).$$

Using Lemma A1 in Berkes and Philipp (1979), we obtain from (7.39)–(7.41) that for all n there exist infinite dimensional random elements $\xi(n)$ and $\eta(n)$ such that

$$(7.42) \quad \mathcal{L}(\xi(n)) = \mathcal{L}(\xi), \quad \mathcal{L}(\eta(n)) = \mathcal{L}(\eta),$$

$$(7.43) \quad \mathbf{P}(\xi_{(n)}(n) \neq \eta_{(n)}(n)) = \rho(n) \leq \delta.$$

Note that $\mathcal{L}(\xi)$ and $\mathcal{L}(\eta)$ are two fixed distributions in the complete separable metric space R^∞ . Hence, for every $\varepsilon > 0$ there exist compacts $K_1(\varepsilon)$ and $K_2(\varepsilon)$ such that

$$\mathbf{P}(\xi \notin K_1(\varepsilon)) < \varepsilon/2 \quad \text{and} \quad \mathbf{P}(\eta \notin K_2(\varepsilon)) < \varepsilon/2.$$

Thus,

$$\begin{aligned} \mathbf{P}((\xi(n), \eta(n)) \notin K_1(\varepsilon) \times K_2(\varepsilon)) &\leq \mathbf{P}(\xi(n) \notin K_1(\varepsilon)) + \mathbf{P}(\eta(n) \notin K_2(\varepsilon)) = \\ &= \mathbf{P}(\xi \notin K_1(\varepsilon)) + \mathbf{P}(\eta \notin K_2(\varepsilon)) < \varepsilon \quad \forall n \quad \forall \varepsilon > 0. \end{aligned}$$

So, the sequence of distributions $\mathcal{L}(\xi(n), \eta(n))$ is tight in the complete separable metric space $R^\infty \times R^\infty$. By Prokhorov's theorem, there exist a distribution $\mathcal{L}(\xi', \eta')$ and a subsequence $n' \rightarrow \infty$ such that

$$(7.44) \quad \mathcal{L}(\xi(n'), \eta(n')) \implies \mathcal{L}(\xi', \eta') \quad \text{as } n' \rightarrow \infty.$$

We shall now investigate the joint distribution of the resulting random variables ξ' and η' . First of all, note that

$$\forall n \quad \mathcal{L}(\xi) = \mathcal{L}(\xi(n)) = \mathcal{L}(\xi'), \quad \mathcal{L}(\eta) = \mathcal{L}(\eta(n)) = \mathcal{L}(\eta'),$$

by (7.42) and (7.44). Hence, from definitions (7.38) and (7.39) we have

$$(7.45) \quad \rho = \rho(\mathcal{L}(\xi), \mathcal{L}(\eta)) = \rho(\mathcal{L}(\xi'), \mathcal{L}(\eta')) \leq \mathbf{P}(\xi' \neq \eta').$$

The last inequality is valid for all random elements (see, for example, Dobrushin (1970) or Lindvall (1992)).

Consider the following sets

$$(7.46) \quad M(m) = \{x, y \in R^\infty : x_{(m)} \neq y_{(m)}\}.$$

It is obvious from (7.43) and (7.46) that

$$\begin{aligned} \forall n \geq m \quad \mathbf{P}((\xi(n), \eta(n)) \in M(m)) &= \mathbf{P}(\xi_{(m)}(n) \neq \eta_{(m)}(n)) \\ (7.47) \quad &\leq \mathbf{P}(\xi_{(n)}(n) \neq \eta_{(n)}(n)) = \rho(n) \leq \delta. \end{aligned}$$

But $M(m)$ is an open set in $R^\infty \times R^\infty$. This fact (7.44) and (7.47) yield

$$(7.48) \quad \mathbf{P}((\xi', \eta') \in M(m)) \leq \limsup_{n' \rightarrow \infty} \mathbf{P}((\xi(n'), \eta(n')) \in M(m)) \leq \delta.$$

Introduce the set

$$M = \{x, y \in R^\infty : x \neq y\} = \bigcup_{m=1}^{\infty} M(m).$$

It is evident that

$$(7.49) \quad \mathbf{P}(\xi' \neq \eta') = \mathbf{P}((\xi', \eta') \in M) = \lim_{m \rightarrow \infty} \mathbf{P}((\xi', \eta') \in M(m)).$$

Thus, from (7.45), (7.49) and (7.48) we obtain the inequality $\rho \leq \delta$ which is equivalent to the assertion (7.38) of Lemma 7.3.

Proof of Lemma 3.2. It follows immediately from (2.2), (3.1) and (6.17) that

$$\rho_k = \rho(\mathcal{L}((X_k/x)|\vec{X}_{k-1}) * \mathcal{L}(V), \mathcal{L}(Y_k/x) * \mathcal{L}(V)).$$

Thus, to prove Lemma 3.1, it is sufficient to put $\mathcal{L}(X) = \mathcal{L}((X_k/x)|\vec{X}_{k-1})$ and $\mathcal{L}(Y) = \mathcal{L}(Y_k/x)$ in Lemma 7.3 and to use conditions (3.3) instead of (7.34).

Acknowledgement. This research is partially supported by the Grant of RFFI: 99-01-00561 and by the Grant of INTAS: 98-1625. The author would like to thank the two referees, David Mason and Lyudmila Sakhanenko for comments and suggestions that greatly improved the paper.

REFERENCES

- Bentkus V. (1985) Asymptotic analysis of sums of independent random elements in Banach spaces. *Doctor Dissertation*, Vilnius.
- Bentkus V., Liubinskas K. (1987) Rates of convergence in the invariance principle in Banach spaces. *Lith. Math. J.*, **27**, 3, 205–213.
- Bentkus V., Račkauskas A. (1984) Estimates of the distances between sums of independent random elements in Banach spaces. *Theory Probab. Appl.*, **29**, 1, 50–65.
- Berkes I., Philipp W. (1979) Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.*, **7**, 1, 29–54.

- Dobrushin R.L. (1970) Prescribing a system of random variables by conditional distributions. *Theory Probab. Appl.*, **15**, 3, 458–486.
- Einmahl U. (1987) A useful estimate in the multidimensional invariance principle. *Probab. Th. Rel. Fields*, **76**, 81–101.
- Lindvall T. (1992) *Lectures on the Coupling Method*, Wiley, New York.
- Loève M. (1978) *Probability theory*, Springer-Verlag, New York.
- Morrow G., Philipp W. (1982) An almost sure invariance principle for Hilbert space valued martingales. *Trans. Amer. Math. Soc.*, **273**, 231–251.
- Monrad D., Philipp W. (1991) Nearby variables with nearby conditional laws and a strong approximation theorem for Hilbert space valued martingales. *Probab. Th. Rel. Fields*, **88**, 381–404.
- Paulauskas V. and Račkauskas A. (1989) *Approximation theory in the central limit theorem. Exact results in Banach spaces*, Kluwer Academic Publisher, Dordrecht.
- Sakhanenko A. I. (1974) Estimates of the rate of convergence in the invariance principle. *Soviet Math. Dokl.*, **15**, 6, 1752–1755.
- Sakhanenko A. I. (1987) Simple method of obtaining estimates in the invariance principle. *Lecture Notes in Math.*, Springer-Verlag, **1299**, 430–443.
- Skorohod A.V. (1976) On a representation of random variables. *Theory Probab. Appl.*, **21**, 3, 628–632.
- Zaitsev A.Yu. (1998) Multidimensional version of the results of Komlós, Major, Tusnády for vectors with finite exponential moments. *ESAIM: Probability and Statistics*, **2**, 41–108.

Departamento de Matemáticas
Universidad Autónoma Metropolitana – Iztapalapa
Av. Michoacán y la Purísima s/n
Col. Vicentina
09340 México, D.F.
MEXICO

E-mail: sakh@xanum.uam.mx; sakh@math.nsc.ru