The Annals of Probability
1982, Vol. 10, No. 2, 346–373

INVARIANCE PRINCIPLES IN PROBABILITY FOR TRIANGULAR ARRAYS OF B-VALUED RANDOM VECTORS AND SOME APPLICATIONS[1]

BY ALEJANDRO DE ACOSTA

Instituto Venezolano de Investigaciones Científicas

If μ_n, ν are probability measures on a separable Banach space, $j_n \to \infty$ and $\mu_n^{j_n} \to_w \nu$ (so ν is necessarily infinitely divisible), then it is possible to construct two row-wise independent triangular arrays $\{X_{nj}\}$, $\{Y_{nj}\}$ such that $\mathscr{L}(X_{nj}) = \mu_n$, $\mathscr{L}(Y_{nj}) = \nu^{1/j_n}$ and $\max_{k \leq j_n} \|S_{nk} - T_{nk}\| \to_P 0$, where S_{nk} and T_{nk} are the respective partial row sums. Several refinements are proved. These results are applied to establish the weak convergence of the distributions of certain functionals of the partial row sums, improving well-known results of Skorohod. As concrete applications, we prove an arc-sine law for triangular arrays generalizing the Erdös-Kac law and an arc-sine law for strictly stable processes generalizing P. Lévy's law for Brownian Motion.

0. Introduction. Let B be a separable Banach space, $\{\mu_n\}$ a sequence of probability measures on B such that $\mu_n^{j_n} \to_w \nu (j_n \in N, j_n \to \infty)$ (see Section 1 for the notation); the measure ν is then necessarily infinitely divisible. One of the main results of this work is that it is possible to choose two row-wise independent triangular arrays whose nth rows have common distribution μ_n and ν^{1/j_n}, respectively, and whose corresponding row partial sums are close in probability for n large, in the sense of the following.

Invariance principle in probability. In the situation described above, there exists a probability space $(\Omega, \mathscr{A}, P)$ and two row-wise independent triangular arrays

$$\{X_{nj} : j = 1, \cdots, j_n; \quad n \in N\}, \{Y_{nj} : j = 1, \cdots, j_n; \quad n \in N\}$$

such that

(1) $\mathscr{L}(X_{nj}) = \mu_n (j = 1, \cdots, j_n)$,

(2) $\mathscr{L}(Y_{nj}) = \nu^{1/j_n} (j = 1, \cdots, j_n)$,

(3) $\max_{k \leq j_n} \|S_{nk} - T_{nk}\| \to_P 0$,

where $S_{nk} = \sum_{j=1}^k X_{nj}$, $T_{nk} = \sum_{j=1}^k Y_{nj}$. Stronger results are proved in Section 3; see also the Addendum.

When $B = R^1$, $\mu_n = \mathscr{L}(X/n^{1/2})(EX^2 = 1, EX = 0)$, $j_n = n$ (so $\nu = N(0, 1)$), the statement is implicit in Breiman [6] (also in Freedman [10]); it has been explicitly stated in Major [13], [14]. It implies Donsker's invariance principle. In Major's paper [13] the triangular arrays are constructed from two independent identically distributed sequences, one with common law $\mathscr{L}(X)$ and the other with common law $N(0, 1)$.

For a separable Banach space B, $\mu_n = \mathscr{L}(X/a_n - b_n/n)$, $j_n = n$, $\mathscr{L}(X)$ belonging to the domain of attraction of a stable measure ν with norming constants $\{a_n\}$ and shifts $\{b_n\}$, a generalization of Major's statement appears in Theorem 1 of a recent paper of Philipp

Received November 17, 1980.

[1] The main results of this paper were presented at the C.N.R.S. Colloque International sur les Processus Gaussiens, St. Flour, France, June 1980 (see [1]) and at the 3rd International Conference on Probability in Banach spaces, Medford, Mass., August 1980.

AMS 1980 *subject classifications.* Primary 60F17, 60B12; secondary 60J30.

Key words and phrases. Infinitely divisible measures, triangular arrays, invariance principle in probability, functionals of partial row sums, arc-sine laws.

[16]. However, the proof of Theorem 1 of [16] contains an error and it is not clear at present if the result is valid.

Although our invariance principle in probability is the natural extension of these statements to triangular arrays, it is not a strict generalization because, as mentioned above, in Major's (and in Philipp's) work the triangular arrays are constructed from two independent identically distributed sequences. However, our theorem does generalize these results as far as the application to the proof of weak convergence of the distributions of functionals of the partial row sums goes. More explicitly, from the above statement one may deduce the following (somewhat unprecisely stated, but suggestive).

Invariance principle in distribution. Let $j_n \in N$, $j_n \to \infty$ and let $f_n : B^{j_n} \to S$ (a Polish space) be a "well behaved" sequence of functions.

Suppose that for some sequence $\{\mu_n\}$ of probability measures on B such that $\mu_n^{j_n} \to_w \nu$, one has

$$\mathscr{L}(f_n(S_{n1}, \cdots, S_{nj_n})) \to_w \beta,$$

where $S_{nk} = \sum_{j=1}^{k} X_{nj}$ and $\{X_{nj}\}$ is a row-wise independent triangular array with $\mathscr{L}(X_{nj}) = \mu_n$ $(j = 1, \cdots, j_n)$. Then for every sequence $\{\lambda_n\}$ of probability measures on B such that $\lambda_n^{j_n} \to_w \nu$, one has

$$\mathscr{L}(f_n(T_{n1}, \cdots, T_{nj_n})) \to_w \beta,$$

where $T_{nk} = \sum_{j=1}^{k} Y_{nj}$ and $\{Y_{nj}\}$ is a row-wise independent triangular array with $\mathscr{L}(Y_{nj}) = \lambda_n$ $(j = 1, \cdots, j_n)$.

Theorem 4.1 gives a precise statement of this form. Section 4 contains other weak convergence theorems for functionals of the partial row-sums.

As a concrete application, we prove a new arc-sine law for real valued triangular arrays and an arc-sine law for strictly stable processes.

We shall describe next the contents of each section.

Section 2 contains some preparatory results for the construction of the triangular arrays. Theorems 2.1 and 2.7 are the main results in this section.

The construction of the triangular arrays is carried out in Section 3. We will give a rough description of the proof of the main part of Theorem 3.1.

For simplicity we shall assume that $\mu_n(B_r^c) = 0$ $(B_r = \{x \in B : \|x\| \le r\})$ for some $r > 0$ and all n. Let $\nu = \delta_{z_\tau} * \gamma * c_\tau \mathrm{Pois}\, \mu$ be the Lévy-Khinchine decomposition of ν (see [2]; here $z_\tau \in B$, γ is a centered Gaussian measure, μ is a Lévy measure and $(c_\tau \mathrm{Pois}\, \mu)\hat{}(f) = \exp\{\int (e^{if(x)} - 1 - if(x) I_{B_\tau}(x))\, d\mu(x)\}$ for all $f \in B'$). By [2] Theorem 2.10, if $\mu_n^{j_n} \to_w \nu$ then for every continuity radius τ of the Lévy measure μ we have

$$\mu_{n,\tau}^{j_n} \to_w \nu_\tau \quad \text{and also} \quad (\mu_n^\tau)_n^j \to_w \nu^{(\tau)},$$

where $\mu_{n,\tau} = (\mu_n(B_\tau^c))\delta_0 + \mu_n | B_\tau$, $\nu_\tau = \delta_{z_\tau} * \gamma * c_\tau \mathrm{Pois}(\mu | B_\tau)$, $\mu_n^\tau = (\mu_n(B_\tau))\delta_0 + \mu_n | B_\tau^c$, $\nu^{(\tau)} = \mathrm{Pois}(\mu | B_\tau^c)$ (here $(\mu | A)(E) = \mu(A \cap E)$). For τ small, one might call ν_τ "almost Gaussian" and $\nu^{(\tau)}$ "incomplete Poisson".

The construction is performed separately for the pair of families $\{\mu_{n,\tau}\}$ and $\{\nu_\tau^{1/j_n}\}$ and the pair of families $\{\mu_n^\tau\}$ and $\{(\nu^{(\tau)})^{1/j_n}\}$, yielding triangular arrays $\{X'_{nj}\}$, $\{Y'_{nj}\}$ (respectively, $\{X''_{nj}\}$, $\{Y''_{nj}\}$); then one defines $\{X_{nj}\}$ and $\{Y_{nj}\}$ by addition (actually, a little extra care must be taken in the definition of X_{nj}, due to the fact that $\mu_{n,\tau} * \mu_n^\tau \neq \mu_n$).

Here is a sketch of the construction of $\{X'_{nj}\}$, $\{Y'_{nj}\}$. We divide the interval of integers $[1, j_n]$ into p successive blocks of roughly equal cardinal numbers. The measure $\mu_{n,\tau}^{j_n}$ is accordingly divided convolution-wise into p measures, each of the form $\mu_{n,\tau}^{l_n}$, with $l_n \sim j_n/p$; similarly for ν_τ. For fixed p and large n, $\mu_{n,\tau}^{l_n}$ is close in Prohorov distance to $\nu_\tau^{1/p}$ (Theorem 2.1); then one may apply Strassen's theorem (see e.g. [8]) to construct joint distributions $\lambda_{n,p,k}$ on $B \times B$ concentrated near the diagonal having as marginals the kth factor of $\mu_{n,\tau}^{j_n}$ and the kth factor of $\nu_\tau (k = 1, \cdots, p)$. Now by an existence theorem (Theorem A.1) one defines *indirectly* $\{X'_{nj}\}$, $\{Y'_{nj}\}$ with the correct individual laws and so that the joint law of $\sum_{\mathrm{block}\, k} X'_{nj}$ and $\sum_{\mathrm{block}\, k} Y'_{nj}$ is $\lambda_{n,p,k}$. Then if τ is small, p large and $n \geq$

$n(\tau, p)$, one shows that $\max_{k \le j_n} \| S'_{nk} - T'_{nk} \|$ is small in probability; in proving this, we exploit the fact that for τ small, the Lévy measure of ν_τ, though not equal to zero, is concentrated on a small ball.

The arrays $\{X''_{nj}\}$ and $\{Y''_{nj}\}$ are constructed so that they are close in a stronger sense. One can show that $j_n \| \mu_n^\tau - (\nu^{(\tau)})^{1/j_n} \|_L^* \to 0$ (Theorem 2.7; here $\| \lambda \|_L^* = \sup\{ | \int f \, d\lambda | : f \in \mathrm{Lip}(S), \| f \|_L \le 1 \}$; see Appendix B for more details on this notation). Then one may apply the Kantorovich-Rubinstein theorem (see Appendix B) in order to construct a joint distribution $\theta_{n,\tau}$ of μ_n^τ and $(\nu^{(\tau)})^{1/j_n}$ so that $j_n \int \| x - y \| \, d\theta_{n,\tau}(x, y)$ is small. The arrays $\{X''_{nj}\}$, $\{Y''_{nj}\}$ are now constructed *directly* so that $\mathscr{L}(X''_{nj}, Y''_{nj}) = \theta_{n,\tau}$. Then if $n \ge n(\tau)$, $\sum_j E \| X''_{nj} - Y''_{nj} \|$ is small. This stronger form of approximation which is possible in the "Poisson case" is isolated in Corollary 3.3.

Theorems 3.4 and 3.5 give more refined constructions; in Theorem 3.5 we construct a stochastic process with stationary independent increments and regular paths associated to ν, approximating the partial row sums in probability. In Theorem 3.7 we show that under the appropriate uniform integrability conditions, one may prove

$$E \sup_{k \le j_n} | \phi(S_{nk} - T_{nk}) - \phi(0) | \to 0$$

for a large class of functions ϕ, including in particular powers of the norm—the case of L^p convergence.

Section 4 contains applications to the weak convergence of functionals of the partial row sums of a triangular array with identically distributed rows. We improve and unify several results of Skorohod [18] (see also [11], Chapter 9, Section 6). We do not use the Skorohod metric on $D[0, 1]$; in fact, from the invariance principles in probability we obtain stronger results than those available for the weak convergence associated to the Skorohod metric (in the context of triangular arrays with identically distributed rows). The section closes with a weak convergence theorem for a special class of functionals, generalizing the approach to the Erdös-Kac arc-sine law in [11], Chapter 9.

Section 5 contains the arc-sine laws. Theorem 5.1 generalizes the Erdös-Kac arc-sine law to triangular arrays with identically distributed rows. Theorem 5.2 generalizes P. Lévy's arc-sine law for Brownian Motion to the case of strictly stable processes.

In Appendix A we prove an existence theorem for probability measures which induce prescribed measures via certain maps. This result is applied several times in the paper. The result is useful for a variety of constructions and appears to be of independent interest. Particular cases of Theorem A.1 have been used more or less implicitly in the literature.

Appendix B is devoted to proving the Kantorovich-Rubinstein theorem for Polish spaces by completing the arguments in Dudley [8]. Although the theorem is sometimes mentioned in the Soviet literature, we are not aware of any reference in which a complete proof is given. Since we use the theorem, we believe that it is of interest to present a proof.

1. Notation. Throughout the paper, B will denote a separable Banach space, $\mathscr{B}$ its Borel σ-algebra, B' its dual space, $B_r = \{x \in B : \| x \| \le r\}$, $B'_r = \{f \in B' : \| f \| \le r\}$. The space of probability measures (p.m.'s) on B will be denoted $\mathscr{P}(B)$. The nth convolution power of $\mu \in \mathscr{P}(B)$ will be noted μ^n; if ν is infinitely divisible, $\nu^{1/p}$ is its pth root. The Prohorov metric on $\mathscr{P}(B)$ will be denoted ρ; the Prohorov metric on the p.m.'s on a metric space (S, d) will be denoted ρ_d or, when no confusion may arise, ρ. The total variation norm on the space of finite signed measures on a metric space will be denoted $\| \cdot \|$.

To a given p.m. $\mu \in \mathscr{P}(B)$ we associate its *truncation at a set* A ($A \in \mathscr{B}$), $\tau(A)\mu = \mu(A^c)\delta_0 + \mu | A$, and its *conditioning at* A, $c(A)\mu = (\mu(A))^{-1}\mu | A$ if $\mu(A) \ne 0$, $= \delta_0$ if $\mu(A) = 0$; here $(\mu | A)(E) = \mu(A \cap E)$ for $E \in \mathscr{B}$. We shall use the abbreviation $\mu_{n,r} = \tau(B_r)\mu_n$.

We refer to [2] for results in connection to the Lévy-Khinchine decomposition of an infinitely divisible p.m. ν, $\nu = \delta_{z_\tau} * \gamma * c_\tau \mathrm{Pois}\, \mu$; as mentioned in the introduction, here $z_\tau \in B$, γ is a centered Gaussian measure and $(c_\tau \mathrm{Pois}\, \mu)\hat{}(f) = \exp\{ \int (e^{if(x)} - 1 - if(x) I_{B_\tau}(x)) \, d\mu(x) \}$ for $f \in B'$, with μ a Lévy measure. We shall write $\nu_\tau = \delta_{z_\tau} * \gamma * c_\tau \mathrm{Pois}(\mu | B_\tau)$. The set of continuity radii of the Lévy measure μ will be denoted $C(\mu)$.

Given a B-valued random vector (r.v.) X, and a set $A \subset B$, we define $X_A = XI_A(X)$; if $S_n = \sum_{j=1}^n X_j$, we write $S_{n,A} = \sum_{j=1}^n X_{jA}$.

The integer part of a number $t \in R$ will be denoted $[t]$.

Appendix B contains some additional notational conventions used occasionally in the paper.

2. Preparatory results. The first two results—Theorem 2.1 and Lemma 2.2—will be useful for the construction corresponding to the "almost Gaussian" part of the limiting distribution in Theorem 3.1.

Theorem 2.1 is of independent interest. Statement (2) will also be useful in Theorems 3.7 and 4.1. Given an infinitely divisible measure ν on B, we denote by $\{\nu^t : t \geq 0\}$ the unique weakly continuous convolution semigroup of probability measures on B such that $\nu^1 = \nu$.

THEOREM 2.1. *Let $\mu_n \in \mathscr{P}(B)$, $\mu_n^{j_n} \to_w \nu$ $(j_n \in N, j_n \to \infty)$.*

(1) *Let $l_n \in N$ be such that $1 \leq l_n \leq j_n$ and $l_n/j_n \to t$ as $n \to \infty$. Then*

$$\mu_n^{l_n} \to_w \nu^t.$$

(2) *The set of* p.m.'s $\{\mu_n^k : n \in N,\ k \leq j_n\}$ *is relatively compact; in fact*

$$\lim_n \sup_{k \leq j_n} \rho(\mu_n^k, \nu^{k/j_n}) = 0.$$

PROOF. (1) Since $\mu_n^{l_n} * \mu_n^{j_n - l_n} = \mu_n^{j_n} \to_w \nu$, it follows from [15], page 59 that $\{\mu_n^{l_n}\}$ is relatively shift-compact. By [15], page 171 (which is valid for the Banach space case), the proof will be complete if we show:

$$(2.1) \qquad (\mu_n^{l_n})\hat{} \to (\nu^t)\hat{} \text{ uniformly over balls in } B'.$$

Fix $r > 0$. Then $\delta = \inf\{|\hat{\nu}(f)| : f \in B'_r\} > 0$ because $\hat{\nu}(f) \neq 0$ for all $f \in B'$ and $\hat{\nu}$ is sequentially w^*-continuous. Therefore there exists $n_0 \in N$ such that $n \geq n_0$ implies

$$\inf_{f \in B'_r} |\hat{\mu}_n(f)|^{j_n} \geq \delta/2 > 0.$$

For $f \in B'_r$, $n \geq n_0$, define $\Psi_f^{(n)}(t) = \hat{\mu}_n(tf)$ for $|t| \leq 1$. Then $\Psi_f^{(n)}(0) = 1$, $\Psi_f^{(n)}$ is continuous and does not vanish on $[-1, 1]$. By [7], page 241, there exists a unique function $\lambda_f^{(n)} : [-1, 1] \to C$ such that $\lambda_f^{(n)}(0) = 0$, $\lambda_f^{(n)}$ is continuous and $\Psi_f^{(n)}(t) = \exp\{\lambda_f^{(n)}(t)\}$ for $t \in [-1, 1]$. Now define

$$\phi_n(f) = \lambda_f^{(n)}(1) \quad \text{for each} \quad f \in B'_r.$$

Then for $n \geq n_0$, $\phi_n(0) = 0$, $\hat{\mu}_n(f) = \exp\{\phi_n(f)\}$ for all $f \in B'_r$ and ϕ_n is sequentially w^*-continuous (this is proved using [7], page 242).

On the other hand, it is well known that there exists a (unique) function $\phi : B' \to C$ such that $\phi(0) = 0$, ϕ is sequentially w^*-continuous and

$$(\nu^t)\hat{}(f) = \exp\{t\phi(f)\} \quad \text{for all} \quad f \in B'.$$

Since for $n \geq n_0 (\mu_n^{l_n})\hat{}(f) - (\nu^t)\hat{}(f) = \exp\{l_n\phi_n(f)\} - \exp\{t\phi(f)\}$, it is easily seen that (2.1) will follow if we can prove that

$$(2.2) \qquad \sup_{f \in B'_r} |l_n\phi_n(f) - t\phi(f)| \to 0 \quad \text{as} \quad n \to \infty;$$

since $l_n\phi_n(f) - t\phi(f) = (l_n/j_n)(j_n\phi_n(f) - \phi(f)) + (l_n/j_n - t)\phi(f)$, (2.2) follows from

$$(2.3) \qquad \sup_{f \in B'_r} |j_n\phi_n(f) - \phi(f)| \to 0 \quad \text{as} \quad n \to \infty.$$

In order to prove (2.3) we argue as in [7], page 242–243. Let $L(z) = \sum_{j=1}^\infty ((-1)^j/j) \cdot (z-1)^j$ for $|z - 1| < ½$. Choose $0 < \epsilon \leq ½$ so that

$$|z - 1| \leq \epsilon \quad \text{implies} \quad |L(z)| \leq 1.$$

Let $g_n(f) = \exp\{j_n\phi_n(f)\}$, $g(f) = \exp\{\phi(f)\}$. Since $g_n(f) = (\mu_n^{j_n})\hat{}(f) \to \hat{\nu}(f) = g(f)$ uniformly on B'_r and $|g(f)| \geq \delta$ on B'_r, there exists $n_1 \geq n_0$ such that $n \geq n_1$ implies

$$\sup_{f\in B'_r}\left|\frac{g_n(f)}{g(f)} - 1\right| \leq \epsilon$$

and consequently $\sup_{f\in B'_r}|L(g_n(f)/g(f))| \leq 1$. But $\exp\{j_n\phi_n(f) - \phi(f)\} = \exp L(g_n(f)/g(f))$ for all $f \in B'_r$; since the two exponents are continuous and vanish at 0, this implies

$$j_n\phi_n(f) - \phi(f) = L(g_n(f)/g(f)) \quad \text{for all} \quad f\in B'_r.$$

Since L is continuous at 1 and $g_n(f)/g(f) \to 1$ uniformly on B'_r, (2.3) follows.

(2) Suppose the limit formula is false. Then there exist $\epsilon > 0$ and two sequences $\{n(i)\}$, $\{k(i)\}(i \in N)$ such that $k(i) \leq j_{n(i)}$, $n(i) \to \infty$ and $\rho(\mu_{n(i)}^{k(i)}, \nu^{k(i)/j_{n(i)}}) \geq \epsilon$. Since $0 < k(i)/j_{n(i)} \leq 1$, $k(i')/j_{n(i')} \to s \in [0, 1]$ for a certain subsequence. By (1) this implies $\rho(\mu_{n(i')}^{k(i')}, \nu^s) \to 0$, contradiction.

Since $\{\nu^t : t \in [0, 1]\}$ is compact, the relative compactness of $\{\mu_n^k : n \in N, k \leq j_n\}$ follows. □

REMARK. It is also possible to prove (1) of Theorem 2.1 from Theorems 2.10 and 2.14 of [2].

LEMMA 2.2. *Let ν be an infinitely divisible* p.m. *and let μ be its associated Lévy measure. Then there exists a positive sequence $\tau_p \in C(\mu)$, $\tau_p \downarrow 0$ $(p \in N)$, such that for every $\epsilon > 0$,*

$$\lim_p \sup_{t\leq p^{-1}} p\,\nu_{\tau_p}^t(B_\epsilon^c) = 0.$$

PROOF. If μ is finite, choose any sequence $\tau_p \in C(\mu)$, $\tau_p \downarrow 0$. If μ is infinite, let $\beta_p = \inf\{\eta > 0 : \mu(B_\eta^c) \leq p^{1/2}\}$; then it is easily shown that $\mu(B_{\beta_p}^c) \leq p^{1/2}$ and $\beta_p \downarrow 0$. Now choose $\tau_p \downarrow 0$ so that $\beta_p \leq \tau_p$ and $\tau_p \in C(\mu)$.

Suppose the limit formula is false. Then there exists $\epsilon > 0$, $\delta > 0$ and two sequences $\{p_n\}$, $\{t_n\}$ such that $p_n \in N$, $p_n \to \infty$, $0 \leq t_n \leq p_n^{-1}$ and $p_n\nu_{\tau_{p_n}}^{t_n}(B_\epsilon^c) \geq \delta$ for all n. Since $0 \leq t_np_n \leq 1$, by passing to a subsequence (which we denote like the whole sequence) we may assume $t_np_n \to s$. Let $\{X_{nj} : 1 \leq j \leq p_n;\ n \in N\}$ be a triangular array such that $\mathscr{L}(X_{nj}) = \nu_{\tau_{p_n}}^{t_n}$. We show next that $\{X_{nj}\}$ is infinitesimal. Choose and fix $\tau > 0$, $\tau \geq \tau_p$ for all p. One may write

$$\nu_{\tau_{p_n}}^{t_n} = \delta_{t_nz_{\tau'}} * \alpha_n, \quad \text{where} \quad \alpha_n = \gamma^{t_n} * c_\tau \mathrm{Pois}(t_n\mu \mid B_{\tau'}) \quad \text{and} \quad \tau' = \tau_{p_n}.$$

Since $t_n \to 0$, it easily follows that $\alpha_n \to_w \delta_0$. Now (see [2], page 6) $z_{\tau'} = z_\tau - \int_{B_{\tau'}^c \cap B_\tau} x\, d\mu(x)$, which implies

$$\| t_nz_{\tau'} \| \leq t_n \| z_\tau \| + \tau t_n\mu(B_{\tau'}^c) \leq t_n \| z_\tau \| + \tau p_n^{-1}P_n^{1/2} \to 0.$$

Therefore $\nu_{\tau_{p_n}}^{t_n} \to_w \delta_0$.

By a standard argument, $\alpha_n^{p_n} \to_w \gamma^s$. Therefore $\mathscr{L}(\sum_j X_{nj}) = \nu_{\tau_{p_n}}^{t_np_n}$ is shift-convergent to a Gaussian measure. By [2], Corollary 2.11, it follows that $p_n\nu_{\tau_{p_n}}^{t_n}(B_\epsilon^c) = \sum_j P\{\|X_{nj}\| > \epsilon\} \to 0$, a contradiction. □

The rest of the section leads up to Theorem 2.7, which will be used in Theorem 3.1 for the construction corresponding to the "incomplete Poisson" part of the limiting distribution.

We refer to Appendix B for the definitions of $\|\cdot\|_L$, $\|\cdot\|_L^*$ and $\mathscr{M}_1^+(B)$. Also, let us recall the following definition (see e.g. [8]): for a finite measure μ on B, $\|\mu\|_{BL}^* = \sup\{|\int f d\mu| : \|f\|_\infty + \|f\|_L \leq 1\}$.

LEMMA 2.3. *Let μ_n, $\nu_n (n \in N)$, $\lambda \in \mathscr{M}_1^+(B)$. Assume*

(1) $\mu_n \to_w \lambda$, $\nu_n \to_w \lambda$,

(2) $\|\mu_n\| = \|\nu_n\|$ *for all* n,

(3) $\lim_{r\to\infty}\sup_n \int_{B_r^c} \|x\|\, d(\mu_n + \nu_n)(x) = 0.$

Then $\|\mu_n - \nu_n\|_L^* \to 0$ *as* $n \to \infty$.

PROOF. Let f be such that $\|f\|_L \le 1$, and define $\tilde f = f - f(0)$. Then for all $x \in B$, $|\tilde f(x)| = |f(x) - f(0)| \le \|x - 0\| = \|x\|$. Next, let us observe that by assumption (2)

$$\int fd(\mu_n - \nu_n) = \int \tilde f d(\mu_n - \nu_n) + f(0)(\mu_n(B) - \nu_n(B)) = \int \tilde f d(\mu_n - \nu_n).$$

It follows that $\|\mu_n - \nu_n\|_L^* = \sup\{|\int fd(\mu_n - \nu_n)| : \|f\|_L \le 1 \text{ and } |f| \le \|\cdot\|\}$.

Let $g(r) = \sup_n \int_{B_r^c} \|x\|\, d(\mu_n + \nu_n)(x)$. For $f: B \to R$ such that $\|f\|_L \le 1$ and $|f| \le \|\cdot\|$,

$$\left|\int fd(\mu_n - \nu_n)\right| \le \int_{B_r^c} |f|\, d(\mu_n + \nu_n) + \left|\int_{B_r} fd(\mu_n - \nu_n)\right|$$
$$\le g(r) + (1 + r)\|\mu_n | B_r - \nu_n | B_r\|_{BL}^*$$

which implies

$$\|\mu_n - \nu_n\|_L^* \le g(r) + (1 + r)\{\|\mu_n | B_r - \lambda | B_r\|_{BL}^* + \|\nu_n | B_r - \lambda | B_r\|_{BL}^*\}.$$

Choose now r so that $g(r) < \epsilon/2$ and $\lambda(\partial B_r) = 0$. Then $\mu_n | B_r \to_w \lambda | B_r$, $\nu_n | B_r \to_w \lambda | B_r$ and by [8], Theorem 8.3, one may choose n_0 so that the bracketed term is smaller than $\epsilon/2(1 + r)$ for $n \ge n_0$. □

For the definition of $\mathscr{M}_1(B)$ we refer to Appendix B.

LEMMA 2.4. *Let* $\mu_j, \nu_j \in \mathscr{M}_1^+(B)$, $j = 1, \cdots, n$. *Let* $M = \max\{\max_j \|\mu_j\|, \max_j \|\nu_j\|\}$. *Then*

$$\|\mu_1 * \cdots * \mu_n - \nu_1 * \cdots * \nu_n\|_L^* \le M^{n-1} \textstyle\sum_{j=1}^n \|\mu_j - \nu_j\|_L^*.$$

PROOF. Using the fact that $\|f\|_L = \|f(\cdot + y)\|_L$ for all $f \in \mathrm{Lip}(B)$ and all $y \in B$, one easily checks that for $\mu \in \mathscr{M}_1(B)$, $\nu \in \mathscr{M}_1^+(B)$,

$$\|\mu * \nu\|_L^* \le \|\mu\|_L^* \|\nu\|. \tag{2.4}$$

Then, denoting the convolution product by $\prod$, we have by (2.4)

$$\begin{aligned}\|\textstyle\prod_{j=1}^n \mu_j - \prod_{j=1}^n \nu_j\|_L^* &\le \|\textstyle\prod_{j=1}^n \mu_j - \prod_{j=1}^{n-1} \mu_j * \nu_n + \prod_{j=1}^{n-1} \mu_j * \nu_n - \prod_{j=1}^n \nu_j\|_L^* \\ &\le \|\textstyle\prod_{j=1}^{n-1} \mu_j\| \|\mu_n - \nu_n\|_L^* + \|\nu_n\| \|\prod_{j=1}^{n-1} \mu_j - \prod_{j=1}^{n-1} \nu_j\|_L^* \\ &\le M^{n-1}\|\mu_n - \nu_n\|_L^* + M\|\textstyle\prod_{j=1}^{n-1} \mu_j - \prod_{j=1}^{n-1} \nu_j\|_L^*.\end{aligned}$$

The result follows by induction. □

LEMMA 2.5. *Let* $\mu, \nu \in \mathscr{M}_1^+(B)$, *and assume* $\|\mu\| = \|\nu\|$. *Then* $\|\mathrm{Pois}\mu - \mathrm{Pois}\nu\|_L^* \le \|\mu - \nu\|_L^*$.

PROOF. We shall use the following remark: if $\lambda_1, \lambda_2 \in \mathscr{M}_1^+(B)$ and $\|\lambda_1\| = \|\lambda_2\|$, then

$$\|\lambda_1 - \lambda_2\|_L^* \le \int \|x\|\, d(\lambda_1 + \lambda_2)(x). \tag{2.5}$$

Let

$$\alpha_n = \exp(-\|\mu\|) \textstyle\sum_{k=0}^n \dfrac{\mu^k}{k!}, \qquad \beta_n = \exp(-\|\nu\|) \textstyle\sum_{k=0}^n \dfrac{\nu^k}{k!}.$$

Then by Lemma 2.4, for all n

$$
\begin{aligned}
\|\alpha_n - \beta_n\|_L^* &\le \exp(-\|\mu\|) \textstyle\sum_{k=1}^n \dfrac{\|\mu^k - \nu^k\|_L^*}{k!} \\
&\le \exp(-\|\mu\|) \textstyle\sum_{k=1}^n \dfrac{\|\mu\|^{k-1} k \|\mu - \nu\|_L^*}{k!} \\
&\le \|\mu - \nu\|_L^* \exp(-\|\mu\|) \textstyle\sum_{k=1}^\infty \dfrac{\|\mu\|^{k-1}}{(k-1)!} = \|\mu - \nu\|_L^*.
\end{aligned}
$$

Now by (2.5)

$$
\begin{aligned}
&|\, \|\mathrm{Pois}\mu - \mathrm{Pois}\nu\|_L^* - \|\alpha_n - \beta_n\|_L^* \,| \le \|(\mathrm{Pois}\mu - \alpha_n) - (\mathrm{Pois}\nu - \beta_n)\|_L^* \\
&\qquad \le \int \|x\|\, d(\mathrm{Pois}\mu - \alpha_n)(x) + \int \|x\|\, d(\mathrm{Pois}\nu - \beta_n)(x) \to 0
\end{aligned}
$$

using the easily checked inequality $\int \|x\|\, d(\mathrm{Pois}\mu)(x) \le \int \|x\|\, d\mu(x)$ (similarly for ν). The result follows. □

The following lemma gives an estimate of the dual Lipschitz distance between a p.m. μ and $\mathrm{Pois}\mu$. The proof is partly based on an idea implicit in Le Cam [12], page 186. We refer to Appendix B for the definition of $\mathscr{P}_1(B)$.

Lemma 2.6. *Let $\mu \in \mathscr{P}_1(B)$. Then*

$$\|\mu - \mathrm{Pois}\mu\|_L^* \le 2\mu(\{0\}^c) \int \|x\|\, d\mu(x).$$

In particular, if $\mu(B_r^c) = 0$, then

$$\|\mu - \mathrm{Pois}\mu\|_L^* \le 2r\mu(\{0\}^c)^2.$$

Proof. Let $\alpha = \mu(\{0\}^c)$. One may write $\mu = (1-\alpha)\delta_0 + \alpha\nu$, for a certain p.m. ν. Define $\lambda = (1-\alpha)\delta_0 + \alpha\exp(-\alpha)\nu$; then $\mu - \lambda = \alpha(1 - \exp(-\alpha))\nu$ and $(\mathrm{Pois}\mu - \lambda)$ is a non-negative measure. Let U, V, W, ξ be independent r.v.'s with

$$\mathscr{L}(U) = a^{-1}\lambda, \qquad \mathscr{L}(V) = (1-a)^{-1}(\mu - \lambda), \qquad \mathscr{L}(W) = (1-a)^{-1}(\mathrm{Pois}\mu - \lambda),$$

where $a = \|\lambda\|$, and $\mathscr{L}(\xi) = (1-a)\delta_0 + a\delta_1$. Define

$$X = \xi U + (1-\xi)V, \qquad Y = \xi U + (1-\xi)W;$$

then $\mathscr{L}(X) = \mu$, $\mathscr{L}(Y) = \mathrm{Pois}\mu$. Now let $\|f\|_L \le 1$. Then

$$
\begin{aligned}
\left| \int f d(\mu - \mathrm{Pois}\mu) \right| &= |E(f(X) - f(Y))| \\
&\le E\|X - Y\| \\
&= E|(1-\xi)|\,\|V - W\| \\
&\le E(1-\xi)(E\|V\| + E\|W\|). \\
&\le 2\int \|x\|\, d(\mu - \lambda)(x) \\
&= 2(1 - \exp(-\alpha)) \int \|x\|\, d\mu(x) \\
&\le 2\alpha \int \|x\|\, d\mu(x),
\end{aligned}
$$

since $E(1-\xi) = 1 - a$ and $\int \|x\|\, d\,\mathrm{Pois}\mu(x) \le \int \|x\|\, d\mu(x)$. The assertion follows. □

THEOREM 2.7. *Let* $\lambda_n, \lambda \in \mathscr{M}^+(B)$, $\|\lambda_n\| \le 1$ *and let* $j_n \to \infty$, $j_n \in N$. *Assume* $\lambda_n(B_r^c) = \lambda(B_r^c) = 0$ *for some* $r > 0$ *and* $j_n\lambda_n \to_w \lambda$. *Then*

$$\lim_n j_n \|\{\lambda_n + (1 - \|\lambda_n\|)\delta_0\} - \mathrm{Pois}(\lambda/j_n)\|_L^* = 0.$$

PROOF. Define

$$\alpha_n = \begin{cases} j_n\lambda_n + (\|\lambda\| - j_n\|\lambda_n\|)\delta_0 & \text{if } \|\lambda\| \ge j_n\|\lambda_n\| \\ j_n\lambda_n & \text{if } \|\lambda\| < j_n\|\lambda_n\|, \end{cases}$$

$$\beta_n = \begin{cases} \lambda & \text{if } \|\lambda\| \ge j_n\|\lambda_n\| \\ \lambda + (j_n\|\lambda_n\| - \|\lambda\|)\delta_0 & \text{if } \|\lambda\| < j_n\|\lambda_n\|. \end{cases}$$

Then $\|\alpha_n\| = \|\beta_n\|$ and $\alpha_n \to_w \lambda$, $\beta_n \to_w \lambda$. Let $\epsilon > 0$. By Lemma 2.3, there exists n_0 such that $n \ge n_0$ implies

$$\|\alpha_n - \beta_n\|_L^* < \epsilon/2,$$

$$\|\alpha_n/j_n - \beta_n/j_n\|_L^* < \epsilon/2j_n,$$

and hence by Lemma 2.5.,

$$j_n\|\mathrm{Pois}\lambda_n - \mathrm{Pois}(\lambda/j_n)\|_L^* < \epsilon/2.$$

On the other hand, by Lemma 2.6.,

$$\|\{\lambda_n + (1 - \|\lambda_n\|)\delta_0\} - \mathrm{Pois}\lambda_n\|_L^* \le 2r\|\lambda_n\|^2;$$

since $j_n\|\lambda_n\|^2 = \|\lambda_n\|(j_n\|\lambda_n\|) \to 0$ as $n \to \infty$, there exists n_1 such that $n \ge n_1$ implies

$$j_n\|\{\lambda_n + (1 - \|\lambda_n\|)\delta_0\} - \mathrm{Pois}\,\lambda_n\|_L^* < \epsilon/2$$

and therefore

$$j_n\|\{\lambda_n + (1 - \|\lambda_n\|)\delta_0\} - \mathrm{Pois}\,(\lambda/j_n)\|_L^* < \epsilon$$

for $n \ge \max\{n_0, n_1\}$. □

3. The main theorems. In the next theorem we write $S_{nk} = \sum_{j=1}^k X_{nj}$, $T_{nk} = \sum_{j=1}^k Y_{nj}$.

THEOREM 3.1. (see also the Addendum). *Let* $\mu_n, \lambda_n \in \mathscr{P}(B)$ *be such that* $\mu_n^{j_n} \to_w \nu$, $\lambda_n^{j_n} \to_w \nu$ $(j_n \in N, j_n \to \infty)$. *Then there exist a probability space and two row-wise independent triangular arrays of B-valued random vectors*

$$\{X_{nj}: j = 1, \cdots, j_n;\, n \in N\}, \qquad \{Y_{nj}: j = 1, \cdots, j_n;\, n \in N\}$$

such that:

(1) $\mathscr{L}(X_{nj}) = \mu_n \quad (j = 1, \cdots, j_n)$,

(2) $\mathscr{L}(Y_{nj}) = \lambda_n \quad (j = 1, \cdots, j_n)$,

(3) $\max_{k \le j_n}\|S_{nk} - T_{nk}\| \to_P 0$.

Also, if $A \in \mathscr{B}$, $0 \notin \partial A$ *and* $\mu(\partial A) = 0$, *where* μ *is the Lévy measure associated with* ν, *then*

$$\max_{k \le j_n}\|S_{nk,A} - T_{nk,A}\| \to_P 0;$$

if furthermore $d(0, A) > 0$, *then* $\sum_{j=1}^{j_n}\|X_{njA} - Y_{njA}\| \to_P 0$.

PROOF. I. *Reduction to the case* $\lambda_n = \nu^{1/j_n}$. Suppose that the first part of the statement has been proved for the case $\lambda_n = \nu^{1/j_n}$. We shall show that then it holds in the general case.

Let $\{X_{nj}\}$, $\{Y_{nj}\}$ be triangular arrays satisfying the conclusion for $\{\mu_n\}$ and $\{\nu^{1/j_n}\}$; similarly, let $\{Y'_{nj}\}$, $\{Z_{nj}\}$ be triangular arrays satisfying the conclusion for $\{\nu^{1/j_n}\}$ and $\{\lambda_n\}$. Let

$$\alpha_n = \mathscr{L}(\{X_{nj}\}, \{Y_{nj}\}), \qquad \beta_n = \mathscr{L}(\{Y'_{nj}\}, \{Z_{nj}\}).$$

Since $\mathscr{L}(\{Y_{nj}\}) = (\nu^{1/j_n})^{\otimes j_n} = \mathscr{L}(\{Y'_{nj}\})$, by Corollary A.2 there exists a p.m. γ_n on B^{3j_n} such that $\pi_{12}\gamma_n = \alpha_n$, $\pi_{23}\gamma_n = \beta_n$. Let $\sigma_n = \pi_{13}\gamma_n$, and define $P = \otimes_{n=1}^{\infty} \sigma_n$ on the product space $(\Omega, \mathscr{A}) = (\prod_{n=1}^{\infty} B^{2j_n}, \otimes_{n=1}^{\infty} \mathscr{B}^{2j_n})$. Let

$$U_{nj} = j\text{th component of the canonical map } \pi_n : \Omega \to B^{2j_n}\ (1 \le j \le j_n),$$

$$V_{nj} = (j_n + j)\text{th component of } \pi_n\ (j = 1, \cdots, j_n).$$

Denote the elements of B^{j_n} by $x = (x_j) = (x_1, \cdots, x_{j_n})$. Now $\mathscr{L}(U_{nj}) = \mu_n$, $\mathscr{L}(V_{nj}) = \lambda_n$ and

$$\begin{aligned}
&P\{\max_{k\le j_n} \|\textstyle\sum_{j=1}^k U_{nj} - \sum_{j=1}^k V_{nj}\| > \epsilon\}\\
&= \sigma_n\{(x, z) \in B^{2j_n} : \max_{k\le j_n} \|\textstyle\sum_{j=1}^k x_j - \sum_{j=1}^k z_j\| > \epsilon\}\\
&= \gamma_n\{(x, y, z) \in B^{3j_n} : \max_{k\le j_n} \|\textstyle\sum_{j=1}^k x_j - \sum_{j=1}^k z_j\| > \epsilon\}\\
&\le \gamma_n\{\max_{k\le j_n} \|\textstyle\sum_{j=1}^k x_j - \sum_{j=1}^k y_j\| > \epsilon/2\}\\
&\quad + \gamma_n\{\max_{k\le j_n} \|\textstyle\sum_{j=1}^k y_j - \sum_{j=1}^k z_j\| > \epsilon/2\}\\
&= \alpha_n\{\max_{k\le j_n} \|\textstyle\sum_{j=1}^k x_j - \sum_{j=1}^k y_j\| > \epsilon/2\}\\
&\quad + \beta_n\{\max_{k\le j_n} \|\textstyle\sum_{j=1}^k y_j - \sum_{j=1}^k z_j\| > \epsilon/2\} \to 0 \quad \text{as} \quad n \to \infty.
\end{aligned}$$

II. *Choice of the sequence* $\{n_p\}$. Choose and fix two positive real sequences $\{\tau_p\}$, $\{r_p\}$ such that

$$\tau_p \text{ is as in Lemma 2.2, } r_p \uparrow \infty,\ r_p \in C(\mu).$$

Let $\{n_p\} \subset N$ be a sequence such that $n_p \uparrow \infty$ and for each $p \in N$, $n \ge n_p$ implies (the possibility of choosing n_p is justified below)

(i) $\rho(\mu_{n,\tau_p}^{l_n}, \nu_{\tau_p}^{1/p}) < p^{-2}$,

$\rho(\nu_{\tau_p}^{l_n/j_n}, \nu_{\tau_p}^{1/p}) < p^{-2}$,

for $l_n = [j_n/p]$ and $l_n = [j_n/p] + 1$.

(ii) $\sup_{k\le l_n} \rho(\mu_{n,\tau_p}^{k}, \nu_{\tau_p}^{k/pl_n}) < p^{-1}$,

where $l_n = [j_n/p] + 1$.

(iii) $j_n\mu_n(B_{\tau_p}^c) \displaystyle\int_{B_{\tau_p}} \|x\|\, d\mu_n(x) < p^{-1}$.

(iv) $\|\tau(B_{r_p} \cap B_{\tau_p}^c)\mu_n - \operatorname{Pois}(\mu/j_n \,|\, B_{r_p} \cap B_{\tau_p}^c)\|_L^* < (pj_n)^{-1}$.

By [2], Theorem 2.10, $\mu_{n,\tau_p}^{j_n} \to_w \nu_{\tau_p}$; then Theorem 2.1(1) gives $\mu_{n,\tau_p}^{l_n} \to_w \nu_{\tau_p}^{1/p}$. This shows that condition (i) may be fulfilled. Theorem 2.1 justifies condition (ii). Condition (iii) may be fulfilled because $\sup_n j_n\mu_n(B_{\tau_p}^c) < \infty$ (Theorem 2.2 of [2]) and $\int_{B_{\tau_p}} \|x\|\, d\mu_n(x) \to 0$ as $n \to \infty$.

By [2], Theorem 2.10, $j_n\mu_n | B_{r_p} \cap B_{\tau_p}^c \to_w \mu | B_{r_p} \cap B_{\tau_p}^c$; condition (iv) is now justified by Theorem 2.7.

III. *Construction of certain measures on* $B \times B$ *with mass concentrated near the diagonal.* We will need to consider a partition of the interval of integers $[1, j_n]$ into p intervals of roughly equal cardinal numbers. For $p \in N$, $k = 0, \cdots, p - 1$, let

$$I(n, p, k) = \{j \in N : kj_np^{-1} < j \le (k + 1)j_np^{-1}\}.$$

Let $c(n, p, k) = \text{card } I(n, p, k)$; then $c(n, p, k) = [j_n p^{-1}]$ or $c(n, p, k) = [j_n p^{-1}] + 1$, and therefore $c(n, p, k)/j_n \to p^{-1}$ for each k.

Let $n \geq n_p$. By condition (i) and Strassen's theorem, there exists a p.m. $\lambda_{n,p,k}$ on $B \times B$ such that

$$\lambda_{n,p,k}\{(x, y) \in B \times B; \| x - y \| > 1/p^2\} < 1/p^2$$

and

$$\pi_1 \lambda_{n,p,k} = \mu_{n,\tau_p}^{c(n,p,k)}, \qquad \pi_2 \lambda_{n,p,k} = \nu_{\tau_p}^{c(n,p,k)/j_n}.$$

Define

$$\lambda_{n,p} = (\otimes_{k=0}^{p-1} \lambda_{n,p,k}) \circ f^{-1} \quad \text{on} \quad B^{2p},$$

where $f: B^{2p} \to B^{2p}$ is defined by

$$f(x_1, \cdots, x_{2p}) = (x_1, x_3, \cdots, x_{2p-1}; x_2, \cdots, x_{2p}).$$

By the Kantorovich-Rubinstein theorem and condition (iv), there exists a p.m. $\theta_{n,p}$ on $B \times B$ such that

$$\int \| x - y \| \, d\theta_{n,p}(x, y) < (pj_n)^{-1}$$

and

$$\pi_1 \theta_{n,p} = \tau(B_{r_p} \cap B_{\tau_p}^c)\mu_n, \ \pi_2 \theta_{n,p} = \text{Pois}(\mu/j_n \,|\, B_{r_p} \cap B_{\tau_p}^c).$$

Define

$$\sigma_{n,p} = (\theta_{n,p}^{\otimes j_n}) \circ g^{-1} \quad \text{on} \quad B^{2j_n},$$

where $g: B^{2j_n} \to B^{2j_n}$ is defined by

$$g(x_1, \cdots, x_{2j_n}) = (x_1, x_3, \cdots, x_{2j_{n}-1}; x_2, \cdots, x_{2j_n}).$$

IV. *Construction of the triangular arrays* $\{X_{nj}\}$ *and* $\{Y_{nj}\}$. We will apply III and Theorem A.1 to the construction of the r.v.'s. In order to do this, we introduce several objects. Let $T = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. For fixed $p \in N$, $n \geq n_p$, let

$$\kappa_n = \mu_n(B_{\tau_p})\delta_{(1,0,0)} + \mu_n(B_{\tau_p}^c \cap B_{r_p})\delta_{(0,1,0)} + \mu_n(B_{r_p}^c)\delta_{(0,0,1)},$$

$$\epsilon_n = \mu_n(B_{\tau_p}^c)\delta_0 + \mu_n(B_{\tau_p})\delta_1,$$

$$\alpha = (\epsilon_n \otimes \kappa_n \otimes c(B_{\tau_p})\mu_n \otimes c(B_{\tau_p}^c \cap B_{r_p})\mu_n \otimes c(B_{r_p}^c)\mu_n)^{\otimes j_n} \quad \text{(a p.m. on } (\{0, 1\} \times T \times B^3)^{j_n} = S_1),$$

$$\beta = (\nu_{\tau_p}^{1/j_n})^{\otimes j_n} \otimes (\text{Pois}(\mu/j_n \,|\, B_{\tau_p}^c \cap B_{r_p}))^{\otimes j_n} \otimes (\text{Pois}(\mu/j_n \,|\, B_{r_p}^c))^{\otimes j_n} \quad \text{(a p.m. on } B^{3j_n} = S_2).$$

In order to simplify the notation, we temporarily establish the following convention: the index j will run through the integers in $[1, j_n]$, the index k through the integers in $[0, p - 1]$. Let $\phi: S_1 \to T_1 = B^p \times B^{j_n}$ be defined by

$$\phi((a_j, b_j, x_j, y_j, z_j)_j) = ((\textstyle\sum_{j \in I(n,p,k)} a_j x_j)_k; (b_j^{(2)} y_j)_j)$$

for $b_j = (b_j^{(1)}, b_j^{(2)}, b_j^{(3)}) \in T$ and $(a_j, b_j, x_j, y_j, z_j)_j \in S_1$. Define $\psi: S_2 \to T_2 = B^p \times B^{j_n}$ by

$$\psi((u_j, v_j, w_j)_j) = ((\textstyle\sum_{j \in I(n,p,k)} u_j)_k; (v_j)_j)$$

for $(u_j, v_j, w_j)_j \in S_2$.

On the space $T_1 \times T_2 = (B^p \times B^{j_n})^2$ we define the p.m.

$$\lambda = (\lambda_{n,p} \otimes \sigma_{n,p}) \circ h^{-1},$$

where $h: T_1 \times T_2 \to T_1 \times T_2$ is defined by

$$h((x_k)_k;\ (x'_k)_k;\ (y_j)_j;\ (y'_j)_j) = ((x_k)_k;\ (y_j)_j;\ (x'_k)_k;\ (y'_j)_j).$$

Then

$$\alpha \circ \phi^{-1} = (\otimes_{k=0}^{p-1} \mu_{n,\tau_p}^{c(n,p,k)}) \otimes (\tau(B_{r_p} \cap B_{\tau_p}^c)\mu_n)^{\otimes j_n}$$
$$= \pi_1 \lambda$$

and

$$\beta \circ \psi^{-1} = (\otimes_{k=0}^{p-1} \nu_{\tau_p}^{c(n,p,k)/j_n}) \otimes \mathrm{Pois}(\mu/j_n \,|\, B_{r_p} \cap B_{\tau_p}^c)^{\otimes j_n}$$
$$= \pi_2 \lambda,$$

as one may verify by direct application of the definitions.

By Theorem A.1, there exist a r.v. $X: \Omega \to S_1$,

$$X = ((\xi_{n1}, \eta_{n1}, U_{n1}, V_{n1}, W_{n1}), \cdots, (\xi_{nj_n}, \eta_{nj_n}, U_{nj_n}, V_{nj_n}, W_{nj_n})),$$

with $\eta_{nj} = (\eta_{nj}^{(1)}, \eta_{nj}^{(2)}, \eta_{nj}^{(3)}): \Omega \to T$, such that $\mathscr{L}(X) = \alpha$ and a r.v. $Y: \Omega \to S_2$,

$$Y = ((K_{n1}, \cdots, K_{nj_n}), \qquad (L_{n1}, \cdots, L_{nj_n}), \qquad (M_{n1}, \cdots, M_{nj_n}))$$

such that $\mathscr{L}(Y) = \beta$ and $\mathscr{L}(\phi(X), \psi(Y)) = \lambda$.

We define now the triangular arrays.

Let

$$E_{nj} = \eta_{nj}^{(1)} U_{nj}, \qquad F_{nj} = \eta_{nj}^{(2)} V_{nj}, \qquad G_{nj} = \eta_{nj}^{(3)} W_{nj}.$$

For $n_p \le n < n_{p+1}$,

$$j = 1, \cdots, j_n, \qquad X_{nj} = E_{nj} + F_{nj} + G_{nj}, \qquad Y_{nj} = K_{nj} + L_{nj} + M_{nj};$$

one explicit way of achieving the construction of the total arrays $\{X_{nj}\}$, $\{Y_{nj}\}$ is to choose the rows for different n's so that they are independent.

It is easily checked that for all $n, j = 1, \cdots, j_n$,

$$\mathscr{L}(X_{nj}) = \mu_n, \qquad \mathscr{L}(Y_{nj}) = \nu^{1/j_n}.$$

Let $S_{nk} = \sum_{j=1}^k X_{nj}$, $T_{nk} = \sum_{j=1}^k Y_{nj}$. We must prove:

$$\max_{k \le j_n} \| S_{nk} - T_{nk} \| \to_P 0.$$

For $n_p \le n < n_{p+1}$ let

$$A_{nk} = \| \textstyle\sum_{j=1}^k \xi_{nj} U_{nj} - \sum_{j=1}^k K_{nj} \|,$$
$$B_{nk} = \| \textstyle\sum_{j=1}^k (F_{nj} + G_{nj}) - \sum_{j=1}^k (L_{nj} + M_{nj}) \|,$$
$$C_{nk} = \| \textstyle\sum_{j=1}^k (\eta_{nj}^{(1)} - \xi_{nj}) U_{nj} \|.$$

Then

$$\| S_{nk} - T_{nk} \| \le A_{nk} + B_{nk} + C_{nk};$$

we will prove that $\max_{k \le j_n} A_{nk} \to_P 0$, $\max_{k \le j_n} B_{nk} \to_P 0$, etc.

V. *Proof that* $\max_{k \le j_n} A_{nk} \to 0$. Let us write $H_{nj} = \xi_{nj} U_{nj}$. Let $\epsilon > 0$ be given. For $n_p \le n < n_{p+1}$,

$$P\{\max_{k \le j_n} \| \textstyle\sum_{j=1}^k H_{nj} - \sum_{j=1}^k K_{nj} \| > \epsilon\}$$
$$\le P\{\max_{0 \le i \le p-1} \| \textstyle\sum_{k=0}^i (\sum_{j \in I(n,p,k)} H_{nj} - \sum_{j \in I(n,p,k)} K_{nj}) \| > \epsilon/3\}$$
$$+ P\{\max_{0 \le k \le p-1} \max_{d \le \sum_{l=0}^k c(n,p,l)} \| \textstyle\sum_{j \in I(n,p,k), j \le d} H_{nj} \| > \epsilon/3\}$$
$$+ P\{\max_{0 \le k \le p-1} \max_{d \le \sum_{l=0}^k c(n,p,l)} \| \textstyle\sum_{j \in I(n,p,k), j \le d} K_{J_n} \| > \epsilon/3\}$$
$$= \text{(I)} + \text{(II)} + \text{(III)}.$$

We shall deal with each term separately. Given $\eta > 0$, choose $q_1 > \max\{3\eta^{-1}, 3\epsilon^{-1}\}$; since for $n_p \le n < n_{p+1}$

$$\mathscr{L}(\textstyle\sum_{j\in I(n,p,k)} H_{nj}, \quad \sum_{j\in I(n,p,k)} K_{nj}) = \lambda_{n,p,k}$$

we have by III above: for $n \ge n_{q_1}$, say $n_p \le n < n_{p+1}$ for some $p \ge q_1$,

$$\begin{aligned}
\text{(I)} &\le P\{\max_{0\le i\le p-1} \textstyle\sum_{k=0}^{i} \|\sum_{j\in I(n,p,k)} H_{nj} - \sum_{j\in I(n,p,k)} K_{nj}\| > \epsilon/3\} \\
&\le P\{\textstyle\sum_{k=0}^{p-1} \|\sum_{j\in I(n,p,k)} H_{nj} - \sum_{j\in I(n,p,k)} K_{nj}\| > \epsilon/3\} \\
&\le \textstyle\sum_{k=0}^{p-1} P\{\|\sum_{j\in I(n,p,k)} H_{nj} - \sum_{j\in I(n,p,k)} K_{nj}\| > \epsilon/3p\} \\
&= \textstyle\sum_{k=0}^{p-1} \lambda_{n,p,k}\{(x, y) \in B \times B: \|x - y\| > \epsilon/3p\} \\
&\le \textstyle\sum_{k=0}^{p-1} \lambda_{n,p,k}\{(x, y) \in B \times B: \|x - y\| > 1/p^2\} \\
&\le p \cdot 1/p^2 = 1/p \le 1/q_1 < \eta/3.
\end{aligned}$$

Next, choose q_2 so that $q_2 > \max\{4, 12\eta^{-1}, 12\epsilon^{-1}\}$ and for $p \ge q_2$,

$$p\nu_{\tau_p}^{1/p}(B_{\epsilon/12}^c) < \eta/12, \qquad \sup_{t\le p^{-1}} \nu_{\tau_p}^{t}(B_{\epsilon/12}^c) < 1/4;$$

the choice of q_2 is possible by Lemma 2.2. Now by condition (ii), if $n \ge n_{q_2}$, say $n_p \le n < n_{p+1}$ for some $p \ge q_2$,

$$\sup_{k\le l_n} \mu_{n,\tau_p}^{k}(B_{\epsilon/6}^c) < 1/4 + p^{-1} \le 1/4 + q_2^{-1} < 1/2. \tag{3.1}$$

By the Ottaviani inequality and (3.1),

$$\begin{aligned}
\text{(II)} &\le \textstyle\sum_{k=0}^{p-1} P\{\max_{d\le \sum_{l=0}^{k} c(n,p,l)} \|\sum_{j\in I(n,p,k), j\le d} H_{nj}\| > \epsilon/3\} \\
&\le \textstyle\sum_{k=0}^{p-1} P\{\|\sum_{j\in I(n,p,k)} H_{nj}\| > \epsilon/6\}(1 - \sup_{k\le l_n} \mu_{n,\tau_p}^{k}(B_{\epsilon/6}^c))^{-1} \\
&\le 2\textstyle\sum_{k=0}^{p-1} P\{\|\sum_{j\in I(n,p,k)} H_{nj}\| > \epsilon/6\}.
\end{aligned}$$

Since $\mathscr{L}(\sum_{j\in I(n,p,k)} H_{nj}) = \mu_{n,\tau_p}^{c(n,p,k)}$, by condition (i) we have

$$\begin{aligned}
\text{(II)} &\le 2\textstyle\sum_{k=0}^{p-1} \mu_{n,\tau_p}^{c(n,p,k)}(B_{\epsilon/6}^c) \\
&< 2\textstyle\sum_{k=0}^{p-1} (\nu_{\tau_p}^{1/p}(B_{\epsilon/12}^c) + p^{-2}) \\
&= 2(p\nu_{\tau_p}^{1/p}(B_{\epsilon/12}^c) + p^{-1}) \\
&< 2(\eta/12 + \eta/12) = \eta/3.
\end{aligned}$$

Term (III) may be treated in a similar (but simpler) fashion, using condition (i) for $\{n_p\}$ and the fact that

$$\mathscr{L}(\textstyle\sum_{j\in I(n,p,k)} K_{nj}) = \nu_{\tau_p}^{c(n,p,k)/j_n}.$$

Following the procedure used for (II), we will obtain a number q_3 such that (III) $< \eta/3$ for $n \ge n_{q_3}$. Finally we may conclude: for n large enough,

$$P\{\max_{k\le j_n} A_{nk} > \epsilon\} \le \text{(I)} + \text{(II)} + \text{(III)} < 3(\eta/3) = \eta.$$

VI. *Proof that* $\max_{k\le j_n} B_{nk} \to_P 0$. Since $\mathscr{L}(G_{nj}) = \tau(B_{r_p}^c)\mu_n$ for $n_p \le n < n_{p+1}$, we have

$$P\{G_{nj} \ne 0 \text{ for some } j\} \le \textstyle\sum_{j=1}^{j_n} P\{G_{nj} \ne 0\} = j_n \mu_n(B_{r_p}^c).$$

By [2], Theorem 2.2, one may choose $q_1 \in N$ so that $\sup_n j_n \mu_n(B_{r_{q_1}}^c) < \eta/3$. Therefore, for $n \ge n_{q_1}$ and any $\epsilon > 0$,

$$P\{\textstyle\sum_j \|G_{nj}\| > \epsilon/3\} < \eta/3.$$

Since $\mathscr{L}(M_{nj}) = \text{Pois}(\mu/j_n \mid B^c_{r_p})$ for $n_p \le n < n_{p+1}$, we have

$$P\{M_{nj} \ne 0\} \le 1 - \exp\{-\mu(B^c_{r_p})/j_n\} \le \mu(B^c_{r_p})/j_n.$$

Therefore,

$$P\{M_{nj} \ne 0 \text{ for some } j\} \le \textstyle\sum_{j=1}^{j_n} P\{M_{nj} \ne 0\} \le j_n\mu(B^c_{r_p})/j_n = \mu(B^c_{r_p}).$$

Choose q_2 so that $\mu(B^c_{r_{q_2}}) < \eta/3$. Then for $n \ge n_{q_2}$ and any $\epsilon > 0$,

$$P\{\textstyle\sum_j \|M_{nj}\| > \epsilon/3\} < \eta/3.$$

Now let $q > \max\{q_1, q_2, 9(\epsilon\eta)^{-1}\}$. Then for any $n \ge n_q$, say $n_p \le n < n_{p+1}$ for some $p \ge q$,

$$\begin{aligned} &P\{\textstyle\sum_j \|(F_{nj} + G_{nj}) - (L_{nj} + M_{nj})\| > \epsilon\} \\ &\quad \le P\{\textstyle\sum_j \|G_{nj}\| > \epsilon/3\} + P\{\textstyle\sum_j \|M_{nj}\| > \epsilon/3\} + P\{\textstyle\sum_j \|F_{nj} - L_{nj}\| > \epsilon/3\}. \end{aligned}$$

But $\mathscr{L}(F_{nj}, L_{nj}) = \theta_{n,p}$; hence by III above,

$$E\|F_{nj} - L_{nj}\| = \int \|x - y\| \, d\theta_{n,p}(x, y) < (pj_n)^{-1}$$

and by Chebyshev's inequality,

$$P\{\textstyle\sum_j \|(F_{nj} + G_{nj}) - (L_{nj} + M_{nj})\| > \epsilon\} < 2\eta/3 + 3\,\epsilon^{-1} j_n q^{-1} j_n^{-1} < \eta.$$

VII. *Proof that* $\max_{k \le j_n} C_{nk} \to_P 0$. Let $n_p \le n < n_{p+1}$. We have

$$\begin{aligned} \textstyle\sum_j E|\eta^{(1)}_{nj} - \xi_{nj}|\,\|U_{nj}\| &= j_n E|\eta^{(1)}_{n1} - \xi_{n1}|\,\|U_{n1}\| \\ &= j_n 2\mu_n(B^c_{\tau_p})\mu_n(B_{\tau_p})\mu_n(B_{\tau_p})^{-1} \int_{B_{\tau_p}} \|x\| \, d\mu_n(x) \\ &= 2j_n\mu_n(B^c_{\tau_p}) \int_{B_{\tau_p}} \|x\| \, d\mu_n(x). \end{aligned}$$

Given $\epsilon > 0$, $\eta > 0$, choose $q > 2\,\epsilon^{-1}\,\eta^{-1}$. Then for $n \ge n_q$, say $n_p \le n < n_{p+1}$ for some $p \ge q$, we have by condition (iii) for n_p:

$$P\{\max_{k \le j_n} C_{nk} > \epsilon\} \le \epsilon^{-1} E(\textstyle\sum_j |\eta^{(1)}_{nj} - \xi_{nj}|\,\|U_{nj}\|) < \epsilon^{-1} 2p^{-1} < \eta.$$

VIII. *Proof that* $\max_{k \le j_n} \|S_{nk,A} - T_{nk,A}\| \to_P 0$. Let $\{X_{nj}\}$, $\{Y_{nj}\}$ be two triangular arrays as in the statement of the theorem, satisfying (1)–(3). In order to prove VIII, we need the following.

LEMMA 3.2. *Let X, Y be B-valued* r.v.'s. *Let A be a subset of B. Then for every $\epsilon > 0$,*

$$\|X_A - Y_A\| \le \|X - Y\| I_A(X) I_A(Y) + \max(\|X\|, \|Y\|)\{I_D(X) I_D(Y) + I_{B^c_\epsilon}(X - Y)\},$$

where $D = (\partial A)^{\epsilon]}$.

PROOF. Let $E = \{X \in A, Y \in A^c\}$, $F = \{X \in A^c, Y \in A\}$. Then

$$\begin{aligned} E &= E \cap (\{X \in (A^c)^{\epsilon]}, Y \in A^{\epsilon]}\} \cup \{X \notin (A^c)^{\epsilon]}\} \cup \{X \in (A^c)^{\epsilon]}, Y \notin A^{\epsilon]}\}) \\ &\subset \{X \in D, Y \in D\} \cup \{X - Y \in B^c_\epsilon\} = G \end{aligned}$$

(observe that since B is a normed linear space, $d(x, A) = d(x, \partial A)$ for $x \in A^c$). Analogously, $F \subset G$.

Since E and F are disjoint,

$$\begin{aligned} \|XI_E - YI_F\| &\le \max(\|X\|, \|Y\|)\{I_E + I_F\} \\ &\le \max(\|X\|, \|Y\|) I_G. \end{aligned}$$

Finally, since

$$XI_A(X) - YI_A(Y) = (X - Y) I_A(X) I_A(Y) + \|XI_E - YI_F\|,$$

we have

$$\|X_A - Y_A\| \le \|X - Y\| I_A(X) I_A(Y) + \|XI_E - YI_F\|$$
$$\le \|X - Y\| I_A(X) I_A(Y) + \max(\|X\|, \|Y\|) I_G,$$

which yields the inequality. □

We return now to assertion VIII.

Let $A \in, \mathscr{B}, 0 \notin \partial A, \mu(\partial A) = 0$. Then either $d(0, A) > 0$ or $d(0, A^c) > 0$. We may assume that $d(0, A) > 0$, since

$$S_{nk,A^c} - T_{nk,A^c} = (S_{nk} - T_{nk}) - (S_{nk,A} - T_{nk,A})$$

and we have already proved that $\max_{k\le j_n} \|S_{nk} - T_{nk}\| \to_P 0$.

By the lemma, putting $D = (\partial A)^{\epsilon]}$, we have for every $\epsilon > 0$

$$\sum_j \|X_{nj} I_A(X_{nj}) - Y_{nj} I_A(Y_{nj})\| \le \sum_j \|X_{nj} - Y_{nj}\| I_A(X_{nj})$$
$$+ \sum_j \max(\|X_{nj}\|, \|Y_{nj}\|)\{I_D(X_{nj}) + I_{B_\epsilon^c}(X_{nj} - Y_{nj})\}.$$

Now for any $\delta > 0$, $\epsilon > 0$,

$$P\{\sum_j \|X_{nj} - Y_{nj}\| I_A(X_{nj}) > \delta\}$$
$$\le P\{\sum_j \|X_{nj} - Y_{nj}\| I_A(X_{nj}) > \delta, \max_{k\le j_n} \|X_{nk} - Y_{nk}\| \le \epsilon\}$$
$$+ P\{\max_{k\le j_n} \|X_{nk} - Y_{nk}\| > \epsilon\}$$
$$\le \delta^{-1}\epsilon\, E(\sum_j I_A(X_{nj})) + P\{\max_{k\le j_n} \|S_{nk} - T_{nk}\| > \epsilon/2\}$$
$$= \delta^{-1}\epsilon\, j_n \mu_n(A) + P\{\max_{k\le j_n} \|S_{nk} - T_{nk}\| > \epsilon/2\}.$$

On the other hand,

$$P\{\sum_j \max(\|X_{nj}\|, \|Y_{nj}\|)\{I_D(X_{nj}) + I_{B_\epsilon^c}(X_{nj} - Y_{nj})\} > 0\}$$
$$\le \sum_j P\{X_{nj} \in D\} + P\{\max_{k\le j_n} \|X_{nk} - Y_{nk}\| > \epsilon\}$$
$$\le j_n \mu_n(D) + P\{\max_{k\le j_n} \|S_{nk} - T_{nk}\| > \epsilon/2\}.$$

Since $d(0, A) > 0$, $s = \sup_n j_n \mu_n(A) < \infty$ by [2], Theorem 2.2. Given $\eta > 0$, choose $\epsilon > 0$ so that $\delta^{-1}\epsilon s \le \eta/3$, $0 \notin D$ and $\mu(D) < \eta/3$. By [2], Theorem 2.10, there exists n_0 such that $n \ge n_0$ implies $j_n \mu_n(D) < \eta/3$. Let n_1 be such that $n \ge n_1$ implies

$$P\{\max_{k\le j_n} \|S_{nk} - T_{nk}\| > \epsilon/2\} > \eta/6.$$

Then for $n \ge \max\{n_0, n_1\}$,

$$P\{\max_{k\le j_n} \|S_{nk,A} - T_{nk,A}\| > \delta\} \le P\{\sum_j \|X_{nj} I_A(X_{nj}) - Y_{nj} I_A(Y_{nj})\| > \delta\}$$
$$\le \delta^{-1}\epsilon s + j_n \mu_n(D) + 2P\{\max_{k\le j_n} \|S_{nk} - T_{nk}\| > \epsilon/2\} \le \eta/3 + \eta/3 + \eta/3 = \eta.$$

The proof of the second statement concerning sums of truncated vectors is contained in the above argument. □

For certain special sequences $\{\mu_n\}$ one may strengthen the main conclusion of Theorem 3.1 (it may be worth remarking that the classical weak convergence of binomial distributions to the Poisson distribution is covered by the corollary).

COROLLARY. 3.3. *Let $t > 0$, and for each $n \in N$ let σ_n be a non-negative measure on B such that $\|\sigma_n\| \le 1$ and $\sigma_n(B_t) = 0$. Let $\mu_n = (1 - \|\sigma_n\|)\,\delta_0 + \sigma_n$, and assume $\mu_n^{j_n} \to_w \nu (j_n \in N, j_n \to \infty)$. Then there exist a probability space and two row-wise independent triangular arrays*

$$\{X_{nj} : j = 1, \cdots, j_n; n \in N\}, \qquad \{Y_{nj} : j = 1, \cdots, j_n; n \in N\}$$

such that:

(1) $\mathscr{L}(X_{nj}) = \mu_n (j = 1, \cdots, j_n)$,
(2) $\mathscr{L}(Y_{nj}) = \nu^{1/j_n}(j = 1, \cdots, j_n)$,
(3) $\sum_{j=1}^{j_n} \|X_{nj} - Y_{nj}\| \to_P 0$.

If $A \in \beta$, $0 \notin \partial A$ *and* $\mu(\partial A) = 0$, *where* μ *is the Lévy measure associated with* ν, *then*

$$\sum_{j=1}^{j_n} \|X_{njA} - Y_{njA}\| \to_P 0.$$

Furthermore, if $\sigma_n(B_r^c) = 0$ *for some* $r > 0$ *and all* n, *then the triangular arrays may be chosen so that*

$$\sum_{j=1}^{j_n} E\|X_{nj} - Y_{nj}\| \to 0.$$

PROOF. Let $0 < s < t$. From [2], Theorem 2.10, it follows that $\gamma = \delta_0$, $z_s = 0$ and $\mu(B_s) = 0$. Taking p so that $\tau_p < s$, we have in step VI of the proof of Theorem 3.1, for $n \geq n_p$:

$$\mathscr{L}(F_{nj} + G_{nj}) = \mu_n, \qquad \mathscr{L}(L_{nj} + M_{nj}) = \text{Pois}(\mu/j_n)(=c_s\ \text{Pois}(\mu/j_n))$$

proving the first statement.

The proof of the second statement follows from (3), the last statement in Theorem 3.1 and the fact that either $d(0, A) > 0$ or $d(0, A^c) > 0$.

The third statement also follows from step VI by observing that for p large enough, $G_{nj} = M_{nj} = 0$ for $n \geq n_p$. □

Under the assumptions of Theorem 3.1 it is possible to refine its conclusion by making a more precise choice of one of the triangular arrays, say $\{Y_{nj}\}$; in fact, one may construct it so that it has any prescribed global distribution subject only to the obvious constraints. We denote by π_t the canonical projection of $\Pi_{k\in N} B^{j_k}$ onto B^{j_t}

THEOREM 3.4. *Let* $\mu_n, \lambda_n, j_n (n \in N)$, ν *be as in Theorem* 3.1. *Let* λ *be a* p.m. *on* $\Pi_{k\in N} B^{j_k}$ *such that for all* n, $\pi_n\lambda = \lambda_n^{\otimes j_n}$. *Then there exist a probability space* $(\Omega, \mathscr{A}, P)$ *and two triangular arrays*

$$\{X_{nj} : j = 1, \cdots, j_n; n \in N\}, \qquad \{Y_{nj} : j = 1, \cdots, j_n; n \in N\}$$

such that

(1) $\mathscr{L}(X_{nj}) = \mu_n (j = 1, \cdots, j_n)$,
(2) $\mathscr{L}(\{Y_{nj} : j = 1, \cdots, j_n; n \in N\}) = \lambda$,
(3) $\max_{k \leq j_n} \|S_{nk} - T_{nk}\| \to_P 0$.

PROOF. Let $\{X'_{nj}\}$, $\{Y'_{nj}\}$ be two triangular arrays such that $\mathscr{L}(X'_{nj}) = \mu_n$, $\mathscr{L}(Y'_{nj}) = \lambda_n (j = 1, \cdots, j_n, n \in N)$ and satisfying (3) of Theorem 3.1. Let $\alpha_n = \mathscr{L}(\{X'_{nj}\}_j : \{Y'_{nj}\}_j)$.

We will apply Theorem A.1. For clarity, let us introduce two copies B_1 and B_2 of B. Put $J = N$, $S_k = B_1^{j_k} \times B_2^{j_k}$, $T_k = B_2^{j_k}$, $\phi_k =$ canonical projection of S_k onto $T_k (k \in N)$. Also, let $\mu_k = \alpha_k$ and let λ be the prescribed measure on T. Now $\mu_k \circ \phi_k^{-1} = \lambda_k^{\otimes j_k} = q_k\lambda$ for all $k \in N$. Let σ be the measure on S given by Theorem A.1. Define $(\Omega, P) = (S, \sigma)$, and for $j = 1, \cdots, j_n$, $n \in N$, $X_{nj} = j$th coordinate of the canonical projection of S onto $B_1^{j_n}$, $Y_{nj} = j$th coordinate of the canonical projection of S onto $B_2^{j_n}$. Clearly $\{X_{nj}\}$ and $\{Y_{nj}\}$ satisfy (1)–(3). □

In the case where $\lambda_n = \nu^{1/j_n}$, there is a natural interesting choice of the array $\{Y_{nj}\}$; this is the content of the next result. Let $I = [0, 1]$ and let $D(I, B)$ be the space of maps of I into B which are right-continuous on $[0, 1)$ and have left limits on $(0, 1]$. Let $\mathscr{D}$ be the σ-algebra of subsets of $D(I, B)$ generated by the coordinate maps $\{\pi_t : t \in I\}$ (here $\pi_t(x) = x(t)$); we recall that $(D(I, B), \mathscr{D})$ is the measurable space generated by a Polish topology on $D(I, B)$ (see e.g. [5], Chapter 3). Given an infinitely divisible measure ν on B, we shall denote by P_ν the distribution on $(D(I, B), \mathscr{D})$ of a stochastic process $\{Z(t) : t \in I\}$ with stationary independent increments, $Z(0) = 0$ a.s., sample paths in $D(I, B)$ and such that $\mathscr{L}(Z(1)) = \nu$.

THEOREM 3.5. *Let $\mu_n \in \mathscr{P}(B)$ be such that $\mu_n^{j_n} \to_w \nu$ ($j_n \in N, j_n \to \infty$). Then there exist a probability space $(\Omega, \mathscr{A}, P)$, a triangular array $\{X_{nj}\}$ and a stochastic process $Z = \{Z(t) : t \in I\} : \Omega \to D(I, B)$, such that:*

$$(1)\ \mathscr{L}(X_{nj}) = \mu_n (j = 1, \cdots, j_n),$$

$$(2)\ \mathscr{L}(Z) = P_\nu,$$

$$(3)\ \max_{k \le j_n} \| S_{nk} - Z(k/j_n) \| \to_P 0.$$

REMARK. In this context one may also state and prove (using the argument in step VIII of the proof of Theorem 3.1) results for sums of truncated random vectors similar to those in Theorem 3.1. We omit these statements.

PROOF. Let B_1, B_2 be as in the proof of Theorem 3.4. Let $\psi : D(I, B) \to \Pi_{i \in N} B_2^{j_i}$ be defined by

$$\psi(x) = (\{x(j/j_i) - x(j - 1/j_i)\}_{1 \le j \le j_i})_{i \in N}.$$

Let σ be the p.m. constructed in Theorem 3.4 for $\lambda = P_\nu \circ \psi^{-1}$.

We apply Theorem A.1 again with $J = \{1, 2\}$ and

$$S_1 = \prod_{i \in N} B_1^{j_i}, \qquad T_1 = S_1, \qquad \phi_1 = Id_{S_1};$$

$$S_2 = D(I, B), \qquad T_2 = \prod_{i \in N} B_2^{j_i}, \qquad \phi_2 = \psi;$$

$$\mu_1 = q_1 \sigma, \qquad \mu_2 = P_\nu, \quad \lambda = \sigma.$$

By Theorem A.1 there exists a p.m. β on $S_1 \times S_2$ such that

$$p_2 \beta = P_\nu, \qquad \beta \circ (p_1, \psi \circ p_2)^{-1} = \sigma.$$

Define now on $(\Omega, P) = (S_1 \times S_2, \beta)$ the map Z and the triangular array $\{X_{nj}\}$ by $Z = p_2$, $X_{nj} = j$th coordinate of the canonical projection of $S_1 \times S_2$ onto $B_1^{j_n} (j = 1, \cdots, j_n;\ n \in N)$. □

Let Φ_+ be the class of continuous functions $\phi : B \to R$ such that

(1) $\phi \ge 0$
(2) There exists a constant $a > 0$ such that $\phi(x + y) \le a\{\phi(x) + \phi(y)\}$ for all x, y in B.

This class is considered in [3], page 216.

Let us observe that if $p > 0$, q is any continuous seminorm on B and $\phi = q^p$, then $\phi \in \Phi_+$.

The following version of Ottaviani's inequality is valid for the functions of the class Φ_+. We omit the proof, which follows the usual lines.

LEMMA 3.6. *Let X_j, $j = 1, \cdots, n$ be independent B-valued* r.v.'s, *$S_k = \sum_{j=1}^k X_j (1 \le k \le n)$. Let $\phi \in \Phi_+$. Then for every $t > 0$,*

$$P\{\sup_{k \le n} \phi(S_k) > 2at\} \le (1 - c)^{-1} P\{\phi(S_n) > t\},$$

where $c = \sup_{k \le n} P\{\phi(S_k - S_n) > t\}$.

THEOREM 3.7. *Let $\mu_n, \lambda_n, j_n (n \in N)$, ν, $\{X_{nj}\}$, $\{Y_{nj}\}$ be as in Theorem* 3.1. *Let $\phi \in \Phi_+$ and assume*:

(1) $\lim_{t \to \infty} \sup_n j_n \int_{B_t'} \phi \, d\mu_n = 0$,
(2) $\lim_{t \to \infty} \sup_n j_n \int_{B_t'} \phi \, d\lambda_n = 0$.

Then $E\{\sup_{k \le j_n} | \phi(S_{nk} - T_{nk}) - \phi(0) | \} \to 0$ as $n \to \infty$.

If $\lambda_n = \nu^{1/j_n}$, then assumption (2) *is superfluous.*

PROOF. Given $\epsilon > 0$, let $\delta > 0$ be such that $\|x\| \leq \delta$ implies $|\phi(x) - \phi(0)| \leq \epsilon$. Then $\{\sup_{k \leq j_n} |\phi(S_{nk} - T_{nk}) - \phi(0)| > \epsilon\} \subset \{\max_{k \leq j_n} \|S_{nk} - T_{nk}\| > \delta\}$; therefore $V_n = \sup_{k \leq j_n} |\phi(S_{nk} - T_{nk}) - \phi(0)| \rightarrow_P 0$. In order to complete the proof, it is enough to show: $\{V_n\}$ is uniformly integrable.

Now $V_n \leq a \sup_{k \leq j_n} \phi(S_{nk}) + a \sup_{k \leq j_n} \phi(T_{nk}) + \phi(0)$. Thus it is enough to prove: $\{\sup_{k \leq j_n} \phi(S_{nk})\}$ is uniformly integrable (the argument is the same for the Y_{nj}'s).

As observed in [3], page 216, there exist $\alpha > 0$, $\beta > 0$ such that $\phi(x) \leq \exp(\beta\|x\|)(x \in B)$. Therefore for any $t > 0$

$$\begin{aligned} \sup_{k \leq j_n} P\{\phi(S_{nk} - S_{nj_n}) > t\} &\leq \sup_{k \leq j_n} P\{\alpha \exp(\beta \| S_{nk} - S_{nj_n} \|) > t\} \\ &= \sup_{k \leq j_n} \mu_n^k(\mathrm{B}_{t'}^c), \end{aligned}$$

where $t' = \beta^{-1} \log(t/\alpha)$. By Theorem 2.1, there exists events $t_0 > 0$ such that $t \geq t_0$ implies $\sup_{k \leq j_n} \mu_n^k(B_{t'}^c) \leq 1/2$. By Lemma 3.6, for all n, all $t \geq t_0$,

$$P\{\sup_{k \leq j_n} \phi(S_{nk}) > 2at\} \leq 2\, P\{\phi(S_{nj_n}) > t\}.$$

By [3], Theorem 3.2, $\{\phi(S_{nj_n})\}$ is uniformly integrable; hence so is $\{\sup_{k \leq j_n} \phi(S_{nk})\}$. Finally, if $\lambda_n = \nu^{1/j_n}$, then for all n

$$\int_{\{\phi(T_{nj_n})>t\}} \phi(T_{nj_n})\, dP = \int_{B_t^c} \phi d\nu \rightarrow 0 \text{ as } t \rightarrow \infty,$$

since $\int \phi d\nu < \infty$ by [3], Theorem 3.3. □

COROLLARY 3.8. *Lt $\mu_n, \lambda_n, j_n (n \in N)$, ν, $\{X_{nj}\}$, $\{Y_{nj}\}$ be as in Theorem* 3.1. *Let $p > 0$, q a continuous seminorm on B and assume*

(1) $\lim_{t \to \infty} \sup_n j_n \int_{B_t^c} q^p d\mu_n = 0$,

(2) $\lim_{t \to \infty} \sup_n j_n \int_{B_t^c} q^p d\lambda_n = 0$.

Then $E \sup_{k \leq j_n} q_p(S_{nk} - T_{nk}) \rightarrow 0$ as $n \rightarrow \infty$.

If $\lambda_n = \nu^{1/j_n}$, then assumption (2) *is superfluous.*

4. Applications to weak convergence: invariance principles in distribution. The first two results deal with asymptotic equivalence in law.

THEOREM 4.1. *Let $\{X_{nj}\}$, $\{Y_{nj}\}$ be row-wise independent triangular arrays of B-valued* r.v.'*s, with $j_n \rightarrow \infty$. Assume that $\mathscr{L}(X_{nj})$* (resp., *$\mathscr{L}(Y_{nj})$*) *does not depend on j, $\mathscr{L}(S_{nj_n}) \rightarrow_w \nu$, $\mathscr{L}(T_{nj_n}) \rightarrow_w \nu$.*

Let (S, d) be a separable metric space, and for each $n \in N$ let $f_n: B^{j_n} \rightarrow S$ be a measurable map such that: for every $\eta > 0$ and for every compact convex symmetric set K in B, there exist $\delta > 0$ and $n_0 \in N$ such that: $n \geq n_0$, $x_j, y_j \in K$ for $1 \leq j \leq j_n$, $\max_{ij \leq j_n} \|x_j - y_j\| \leq \delta$ imply

$$d(f_n(x_1, \cdots, x_{j_n}), \quad f_n(y_1, \cdots, y_{j_n})) < \eta.$$

Then $\lim_n \rho_d(\mathscr{L}(f_n(S_{n1}, \cdots, S_{nj_n})), \mathscr{L}(f_n(T_{n1}, \cdots, T_{nj_n}))) = 0$.

PROOF. Let $\mu_n = \mathscr{L}(X_{n1})$, $\lambda_n = \mathscr{L}(Y_{n1})$. By Theorem 3.1, one may construct two triangular arrays $\{X'_{nj}\}$, $\{Y'_{nj}\}$ such that $\mathscr{L}(X'_{nj}) = \mu_n$, $\mathscr{L}(Y'_{nj}) = \lambda_n$ $(j = 1, \cdots, j_n)$ and $\max_{k \leq j_n} \|S'_{nk} - T'_{nk}\| \rightarrow_p 0$. We shall prove:

$$d(f_n(S'_{n1}, \cdots, S'_{nj_n}), f_n(T'_{nj}, \cdots, T'_{nj_n})) \rightarrow_p 0. \tag{4.1}$$

Clearly this implies the conclusion.

Given $\epsilon > 0$, by Theorem 2.1 there exists a compact convex symmetric set K in B such that: for all n,

$$\begin{aligned} &\mu_n^{j_n}(K^c) < \epsilon/8, \qquad \sup_{k \leq j_n} \mu_n^k(K^c) < 1/2 \\ &\lambda_n^{j_n}(K^c) < \epsilon/8, \qquad \sup_{k \leq j_n} \lambda_n^k(K^c) < 1/2. \end{aligned} \tag{4.2}$$

By the Ottaviani inequality, applied to the Minkowski functional q of K,

$$P\{S'_{nk} \notin 2K \text{ for some } k \le j_n\} = P\{\sup_{k\le j_n} q(S'_{nk}) > 2\}$$
$$\le (1-c)^{-1}P\{q(S'_{nj_n}) > 1\},$$

where $c = \sup_{k\le j_n} P\{q(S'_{nj_n} - S'_{nk}) > 1\}$. By (4.2), we have: for all n,

$$P\{S'_{nk} \notin 2K \quad \text{for some} \quad k \le j_n\} < \epsilon/4;$$

analogously, for all n,

$$P\{T'_{nk} \notin 2K \quad \text{for some} \quad k \le j_n\} < \epsilon/4.$$

Given $\eta > 0$, choose $\delta > 0$ and $n_0 \in N$ as in the statement of the theorem, with K replaced by $2K$. Next, choose n_1 so that $n \ge n_1$ implies $P\{\max_{k\le j_n} \|S'_{nk} - T'_{nk}\| > \delta\} < \epsilon/2$. Then for $n \ge \max\{n_0, n_1\}$, if $A_n = \{d(f_n(S'_{n1}, \cdots, S'_{nj_n}), f_n(T'_{n1}, \cdots, T'_{nj_n})) > \eta\}$,

$$P(A_n) \le P(A_n \cap \{S'_{nk} \in 2K \text{ for } k \le j_n\} \cap \{T'_{nk} \in 2K \text{ for } k \le j_n\}) + \epsilon/4 + \epsilon/4$$
$$\le P\{\max_{k\le j_n} \|S'_{nk} - T'_{nk}\| > \delta\} + \epsilon/2 < \epsilon. \qquad \square$$

DEFINITION 4.2. Let $D(I, B)$ be as in the paragraph preceding Theorem 3.6. The maps $\ell_n: B^n \to D(I, B)$ and $r_n: B^n \to D(I, B)$ are defined by:

$$\ell_n(x_1, \cdots, x_n)(t) = x_{[nt]}, \qquad r_n(x_1, \cdots, x_n)(t) = x_{[nt]+1}$$

with the conventions $x_0 = 0$, $x_{n+1} = x_n$.

We shall write $U_r = \{x \in D(I, B): \|x\|_\infty \le r\}$ for $r > 0$.

THEOREM 4.3. *Let $\{X_{nj}\}$, $\{Y_{nj}\}$, (S, d) be as in Theorem 4.1. Let $f: D(I, B) \to S$ be such that*

(1) *f is $\mathscr{D}$-measurable,*

(2) *for every $r > 0$, $f \mid U_r$ is $\|\cdot\|_\infty$—uniformly continuous.*

Let $\xi_n^{(\ell)} = \ell_{j_n}(S_{n1}, \cdots, S_{nj_n})$, $\xi_n^{(r)} = r_{j_n}(T_{n1}, \cdots, T_{nj_n})$, $\eta_n^{(\ell)} = \ell_{j_n}(T_{n1}, \cdots, T_{nj_n})$, $\eta_n^{(r)} = r_{j_n}(T_{n1}, \cdots, T_{lj_n})$. Then

$$\lim_n \rho_d(\mathscr{L}(f(\xi_n^{(\ell)})), \mathscr{L}(f(\eta_n^{(\ell)}))) = \lim_n \rho_d(\mathscr{L}(f(\xi_n^{(r)})), \mathscr{L}(f(\eta_n^{(r)}))) = 0.$$

PROOF. Let $f_n^{(\ell)} = f\circ\ell_{j_n}$, $f_n^{(r)} = f\circ r_{j_n}$. Then $\{f_n^{(\ell)}\}$ and $\{f_n^{(r)}\}$ satisfy the condition in Theorem 4.1. In fact, given $\epsilon > 0$ and a compact convex symmetric set K in B, let $r = \sup_{x\in K} \|x\|$ and let $\delta > 0$ be such that $x, y \in U_r$, $\|x - y\|_\infty < \delta$ imply $d(f(x), f(y)) < \epsilon$.

Now assume $x_j, y_j \in K$, $\max_{j\le j_n} \|x_j - y_j\|_\infty < \delta$. Then $\|\ell_{j_n}(x_1, \cdots, x_{j_n})\|_\infty \le r$ and $\|\ell_{j_n}(y_j, \cdots, y_{j_n})\|_\infty \le r$ because $\|x_j\| \le r$ and $\|y_j\| \le r$ for $1 \le j \le j_n$ and $\|l_{j_n}(x_1, \cdots, x_{j_n}) - l_{j_n}(y_1, \cdots, y_{j_n})\|_\infty = \max_{j\le j_n} \|x_j - y_j\| < \delta$, implying

$$d(f_n^{(\ell)}(x_1, \cdots, x_{j_n}), f_n^{(\ell)}(y_1, \cdots, y_{j_n})) = d(f(\ell_{j_n}(x_1, \cdots, x_{j_n})), f(\ell_{j_n}(y_1, \cdots, y_{j_n}))) < \epsilon.$$

Similarly for $\{f_n^{(r)}\}$.

Now $f(\xi_n^{(\ell)}) = f_n^{(\ell)}(S_{n1}, \cdots, S_{nj_n})$, etc., so the statement follows from Theorem 4.1. $\square$

REMARK. Theorem 4.3 improves Theorem 3.4 of [18].

DEFINITION 4.4. The maps $L_n: D(I, B) \to D(I, B)$ and $R_n: D(I, B) \to D(I, B)$ are defined by:

$$L_n x(t) = x([nt]/n), \qquad R_n x(t) = x(([nt]+1)/n) \quad \text{for} \quad t \in [0, 1), \qquad R_n x(1) = x(1).$$

DEFINITION 4.5. Let ν be an infinitely divisible measure on B, P_ν the p.m. on $(D(I, B), \mathscr{D})$ associated to ν as in the paragraph preceding Theorem 3.5.

$\mathscr{F}_\ell(\nu)$(resp., $\mathscr{F}_r(\nu)$) is the set of all maps $f:D(I, B) \to R$ such that
(1) f is $\mathscr{D}$-measurable,
(2) for every $r > 0$, $f \mid U_r$ is $\|\cdot\|_\infty$—uniformly continuous,
(3) $\lim_n f(L_n x) = f(x)$(resp., $\lim_n f(R_n x) = f(x)$) for P_ν—almost all x.

$\bar{\mathscr{F}}_\ell(\nu)$(resp., $\bar{\mathscr{F}}_r(\nu)$) is the set of all maps $f:D(I, B) \to R$ such that
(i) f is $\mathscr{D}$-measurable.
(ii) for every $\epsilon > 0$ and $r > 0$, there exist $g, h \in \mathscr{F}_\ell(\nu)$(resp., $\mathscr{F}_r(\nu)$) such that $g(x) \le f(x) \le h(x)$ for all $x \in U_r$ and $P_\nu\{x \in U_r : h(x) - g(x) > \epsilon\} < \epsilon$.

The next result is the main convergence theorem for functionals of the step processes associated with the partial row sums of a triangular array.

THEOREM 4.6. *Let* $\{X_{nj}\}$ *be a triangular array of B-valued* r.v.'s *with* $j_n \to \infty$. *Assume that* $\mathscr{L}(X_{nj})$ *does not depend on* j *and* $\mathscr{L}(S_{nj_n}) \to_w \nu$. *Let* $\xi_n^{(\ell)}$ *and* $\xi_n^{(r)}$ *be as in Theorem* 4.3. *Then for all* $f \in \bar{\mathscr{F}}_\ell(\nu)$(resp., $f \in \bar{\mathscr{F}}_r(\nu)$),

$$\mathscr{L}(f(\xi_n^{(\ell)})) \to_w P_\nu \circ f^{-1} \qquad (\text{resp., } \mathscr{L}(f(\xi_n^{(r)})) \to_w P_\nu \circ f^{-1}).$$

PROOF. (1) Let $f \in \mathscr{F}_\ell(\nu)$. Let $Z = \{Z(t) : t \in I\}$ be as in the paragraph preceding Theorem 3.5, and let $Y_{nj} = Z(j/j_n) - Z(j - 1/j_n)$, $j = 1, \cdots, j_n$, $n \in N$. By Theorem 4.3, putting, $\xi_n = \xi_n^{(\ell)}$,

$$\rho(\mathscr{L}(f(\xi_n)), \mathscr{L}(f(L_n Z)) \to 0 \qquad \text{as } n \to \infty, \tag{4.2}$$

where ρ is the Prohorov distance on $\mathscr{P}(R^1)$.

Since $P_\nu(\{x : f(L_n x) \not\to f(x)\}) = 0$ and $\mathscr{L}(Z) = \mathrm{P}_\nu$, we have $f(L_n Z) \to f(Z)$ a.s. and hence $\mathscr{L}(f(L_n Z)) \to \mathscr{L}(f(Z)) = P_\nu \circ f^{-1}$; this fact and (4.2) imply $\mathscr{L}(f(\xi_n)) \to_w P_\nu \circ f^{-1}$.

(2) Let $f \in \bar{\mathscr{F}}_\ell(\nu)$. In order to prove the conclusion, it is enough to show: for all $\phi : R \to [0, 1]$ continuous and nondecreasing,

$$\int \phi(f(\xi_n))\, dP \to \int (\phi \circ f)\, dP_\nu. \tag{4.3}$$

Let $r > 0$ be such that $P\{\|\xi_n\|_\infty > r\} < \epsilon$ for all n and $P_\nu(U_r^c) < \epsilon$. For each $k \in N$, let g_k, $h_k \in \mathscr{F}_\ell(\nu)$ be such that $g_k(x) \le f(x) \le h_k(x)$ for all $x \in U_r$ and $P_\nu\{x \in U_r : h(x) - g(x) > 1/k\} < 1/k$. For all k, n,

$$\int \phi(f(\xi_n))\, dP \le \int_{\{\xi_n \in U_r\}} \phi(f(\xi_n))\, dP + \epsilon \le \int \phi(h_k(\xi_n))\, dP + \epsilon$$

and similarly $\int \phi(f(\xi_n))\, dP \ge \int \phi(g_k(\xi_n))\, dP - 2\epsilon$. Letting $n \to \infty$, we obtain by (1)

$$\begin{aligned} \int \phi(g_k)\, dP_\nu - 2\epsilon &\le \liminf_n \int \phi(f(\xi_n))\, dP \\ &\le \limsup_n \int \phi(f(\xi_n))\, dP \le \int \phi(h_k)\, dP_\nu + \epsilon. \end{aligned}$$

Let $\delta > 0$ be such that $|s - t| < \delta$ implies $|\phi(s) - \phi(t)| < \epsilon$ and let $k > \max\{\delta^{-1}, \epsilon^{-1}\}$. Then

$$\int \{\phi(h_k) - \phi(g_k)\}\, dP_\nu \le \epsilon + P_\nu\{h_k - g_k > \delta\} < 3\epsilon.$$

Since $\int \phi(g_k)\, dP_\nu \le \int \phi(f)\, dP_\nu \le \int \phi(h_k)\, dP_\nu$ for all k, (4.3) follows.

The argument for $f \in \bar{\mathscr{F}}_r(\nu)$ and $\{\xi_n^{(r)}\}$ is totally similar. □

DEFINITION 4.7. A metric σ on $D(I, B)$ in an *ℓ-metric* (resp., an *r-metric*) if
(1) $Id : (D(I, B), \|\cdot\|_\infty) \to (D(I, B), \sigma)$ is uniformly continuous,
(2) the Borel σ-algebra generated by σ is $\mathscr{D}$,
(3) for all $x \in D(I, B)$, $\sigma(L_n x, x) \to 0$ (resp., $\sigma(R_n x, x) \to 0$).

It is easily shown that if σ is an ℓ-metric or an r-metric, then $D(I, B)$ is σ-separable and σ is $\mathscr{D} \otimes \mathscr{D}$ measurable.

Skorohod's metric on $D(I, B)$ (see e.g. [5], Chapter 3) is both an ℓ-metric and an r-metric; (1) is immediate from its definition, (2) is proved in [5] and (3) may be proved using Lemma 1, Chapter 3 of [5].

COROLLARY 4.8. *Let $\{X_{nj}\}$ be as in Theorem 4.6. Let σ be an ℓ-metric* (resp., an r-metric) *on $D(I, B)$. Then*

$$\mathscr{L}(\xi_n^{(\ell)}) \to_{w_\sigma} P_\nu (\text{resp., } \mathscr{L}(\xi_n^{(r)}) \to_{w_\sigma} P_\nu)$$

where $\to_{w_\sigma}$ stands for convergence in the weak topology of measures associated with σ.

Taking σ to be Skorohod's metric, we obtain Theorem 2.7 of [18] (see also [11], Theorem 2, page 480). This well-known result of Skorohod also follows more directly from Theorem 3.5.

The next result gives the limiting distribution for a sequence of functionals of a special form. It generalizes the approach to the Erdös-Kac arc-sine law in [11], Chapter 9. Let $\{\nu^t : t \geq 0\}$ be as in Section 2; for a map $\phi : B \to R$, D_ϕ will denote the set of points at which ϕ is discontinuous.

THEOREM 4.9. *Let ν be an infinitely divisible* p.m. *on B. Let λ be a finite measure on* $[0, 1]$ *without atoms, g its distribution function. Let $\phi : B \to R$ be measurable and bounded on spheres, and assume*: $\lambda(\{t \in [0, 1] : \nu^t(D_\phi) > 0\}) = 0$.

Let $\{X_{nj}\}$ be a triangular array of B-valued r.v.'s *such that $\mathscr{L}(X_{nj})$ does not depend on j, and assume $\mathscr{L}(S_{nj_n}) \to_w \nu$. Let $\{Z(t) : t \in [0, 1]\}$ be as in the paragraph preceding Theorem* 3.5. *Then*

$$\mathscr{L}\big(\textstyle\sum_{k=1}^{j_n} \phi(S_{nk})\{g(k/j_n) - g((k-1)/j_n)\}\big) \to_w \mathscr{L}\left(\displaystyle\int_0^1 \phi(Z(t))\, d\lambda(t)\right),$$

$$\mathscr{L}\big(\textstyle\sum_{k=1}^{j_n-1} \phi(S_{nk})\{g((k+1)/j_n) - g(k/j_n)\}\big) \to_w \mathscr{L}\left(\displaystyle\int_0^1 \phi(Z(t))\, d\lambda(t)\right).$$

PROOF. Let $f(x) = \int_0^1 \phi(x(t))\, d\lambda(t)$. To show that f is $\mathscr{D}$-measurable, let us remark that the map $\psi : [0, 1] \times D(I, B) \to B$ defined by $\psi(t, x) = x(t)$ is easily seen to be measurable. Hence so is $\phi \circ \psi$ and this implies that f is $\mathscr{D}$-measurable.

We shall prove: $f \in \bar{\mathscr{F}}_\ell(\nu)$, $f \in \bar{\mathscr{F}}_r(\nu)$. As a first step, let $u : B \to R$ be a bounded uniformly continuous function and define $g(x) = \int u(x(t))\, d\lambda(t)$. Then it is easily seen that $g \in \mathscr{F}_\ell(\nu)$ and $g \in \mathscr{F}_r(\nu)$.

Let $\beta(A) = \int_0^1 \nu^t(A)\, d\lambda(t)$ for $A \in \mathscr{B}$; then $\beta(D_\phi) = 0$. Given $\epsilon > 0$, $r > 0$, let $s > r$ be such that $\beta(\partial B_s) = 0$. Let $\phi_1 = \phi\, I_{B_s}$. Let u, v be bounded uniformly continuous functions on B such that $u \leq \phi_1 \leq v$ and $\int (v - u)\, d\beta < \epsilon^2$ (to show that such functions exist one may use, for example, Lemma (126), page 80 in [10] and take into account the fact that $\beta(D_{\phi_1}) = 0$). Define

$$g(x) = \int u(x(t))\, d\lambda(t), \qquad h(x) = \int v(x(t))\, d\lambda(t).$$

Then g, $h \in \mathscr{F}_\ell(\nu)$ (also to $\mathscr{F}_r(\nu)$), $g(x) \leq f(x) \leq h(x)$ for $x \in U_r$ and

$$\begin{aligned} P_\nu(\{h - g > \epsilon\}) &\leq \epsilon^{-1} \int (h - g)\, dP_\nu = \epsilon^{-1} \int dP_\nu(x) \int \{v(x(t)) - u(x(t))\}\, d\lambda(t) \\ &= \epsilon^{-1} \int d\lambda(t) \int \{v(x(t)) - u(x(t))\}\, dP_\nu(x) \\ &= \epsilon^{-1} \int d\lambda(t) \int (v - u)\, d\nu^t \\ &= \int (v - u)\, d\beta < \epsilon^{-1}\epsilon^2 = \epsilon. \end{aligned}$$

This shows that $f \in \bar{\mathscr{F}}_l(\nu)$, $f \in \bar{\mathscr{F}}_r(\nu)$.

By Theorem 4.6, $\mathscr{L}(f(\xi_n^{(l)})) \to_w \mathscr{L}(f(Z))$, $\mathscr{L}(f(\xi_n^{(r)})) \to_w \mathscr{L}(f(Z))$. But

$$f(\xi_n^{(l)}) = \textstyle\sum_{k=1}^{j_n-1} \phi(S_{nk})\{g(k+1)/j_n) - g(k/j_n)\},$$

$$f(\xi_n^{(r)}) = \textstyle\sum_{k=1}^{j_n} \phi(S_{nk})\{g(k/j_n) - g((k-1)/j_n)\}. \qquad \square$$

REMARK. It is possible to give a different proof of Theorem 4.9. One may show (1) if $M = \{x \in D(I, B):\lambda(x^{-1}(D_\phi)) > 0\}$, then $M \in \mathscr{D}$ and $P_\nu(M) = 0$; (2) if $x \notin M$ and $x_n \in D(I, B)$, $x_n(t) \to x(t)$ for $t \notin D_x$ and $\sup_n \|x_n\|_\infty < \infty$, then $f(x_n) \to f(x)$. This proves that f is P_ν - a.s. continuous for the Skorohod metric and the result follows from Corollary 4.7 (in fact, f is P_ν — a.s. continuous for any metric σ on $D(I, B)$ such that $\sigma(x_n, x) \to 0$ implies $\sup_n \|x_n\|_\infty < \infty$ and $x_n(t) \to x(t)$ for $t \notin D_x$).

5. Some arc-sine laws. In this section we combine Theorem 4.9 and classical arguments of random walk theory to obtain arc-sine laws for triangular arrays and certain processes with stationary independent increments.

Let us remark that in the first result no assumptions are made on the triangular array beyond having identically distributed rows; in particular, there are no moment or symmetry assumptions on $\{X_{nj}\}$. It is the limiting measure ν that is subject to restrictions.

THEOREM 5.1. *Let $B = R^1$ and let $\{X_{nj}\}$ and ν be as in Theorem 4.9. Assume that ν satisfies*

(1) $\nu^t(\{0\}) = 0$ *for all* $t > 0$,

(2) $\nu^t(0, \infty) = \alpha > 0$ *for all* $t > 0$.

Let $L_n = \operatorname{card}\{k \le j_n : S_{nk} > 0\}$. *Then*

$$\mathscr{L}(L_n/j_n) \to_w \sigma_\alpha,$$

where

$$\frac{d\sigma_\alpha}{dx}(x) = \pi^{-1}\sin(\pi\alpha)\, x^{-(1-\alpha)}(1-x)^{-\alpha} I_{(0,1)}(x).$$

PROOF. In Theorem 4.9 let $\phi = I_{(0,\infty)}$, λ = Lebesgue measure on [0, 1]. Using assumption (1), we have by Theorem 4.9:

$$\mathscr{L}(L_n/j_n) = \mathscr{L}(\textstyle\sum_{k=1}^{j_n} \phi(S_{nk})/j_n) \to_w \mathscr{L}\left(\displaystyle\int_0^1 \phi(Z(t))\, d\lambda(t)\right)$$

$$= \mathscr{L}(\lambda\{t \in [0, 1]: Z(t) > 0\}).$$

In order to determine the limiting distribution, let $\{Y_{nj}\}$ be a triangular array with $\mathscr{L}(Y_{nj}) = \nu^{1/j_n}$, $j = 1, \cdots, j_n$ and let $M_n = \operatorname{card}\{k \le j_n : T_{nk} > 0\}$. By assumption (2) and Theorem 2, page 419 and (8.12), page 419 of Feller [9], we have for $k = 0, \cdots, j_n$:

$$P\{M_n = k\} = (-1)^{j_n}\tbinom{-\alpha}{k}\tbinom{\alpha-1}{j_n-k}.$$

Again by [9], page 419, $\mathscr{L}(M_n/j_n) \to_w \sigma_\alpha$. This implies

$$\sigma_\alpha = \mathscr{L}(\lambda\{t \in [0, 1]: Z(t) > 0\}). \qquad \square$$

Theorem 5.1 generalizes the Erdös-Kac arc-sine law.

The assumptions on ν are satisfied if ν is a strictly stable measure, as is easily verified. One may obtain from Theorem 5.1 the following corollary for stable processes, which generalizes P. Lévy's arc-sine law for Brownian Motion.

THEOREM 5.2. *Let ν be a strictly stable* p.m. *on R^1 and let $\{Z(t): t \in [0, 1]\}$ be a stochastic process with stationary independent increments, $Z(0) = 0$* a.s., *sample paths in $D[0, 1]$ and $\mathscr{L}(Z(1)) = \nu$. Let λ be Lebesgue measure on* [0, 1]. *Then*

$$\mathscr{L}(\lambda\{t \in [0, 1]: Z(t) > 0\}) = \sigma_\alpha,$$

where $\alpha = \nu(0, \infty)$.

REMARK. By way of contrast to Theorem 5.1, we shall exhibit a situation in which the behavior of $\mathscr{L}(L_n/j_n)$ is rather different. Let $\{S_{nj}:j = 1, \cdots, n;\, n \in N\}$ be a real-valued triangular array with $\mathscr{L}(X_{nj}) = (1 - p_n)\delta_0 + p_n\delta_1$, $j = 1, \cdots, n$. Assume $np_n \to \lambda > 0$ as $n \to \infty$. By Poisson's limit theorem, $\mathscr{L}(S_{nn}) \to_w \mathrm{Pois}(\lambda\delta_1)$. Then one may prove (for example, using Theorem 4.9 with $\phi = I_{(\epsilon,\infty)}$, $0 < \epsilon < 1$)

$$\mathscr{L}(L_n/n) \to_w \sigma,$$

where $d\sigma = e^{-\lambda}\, d\delta_0 + \{\lambda e^{-\lambda(1-x)} I_{(0,1)}(x)\}\, dx$.

Appendix A. An existence theorem for probability measures with certain prescribed induced measures. The following general result is useful for a variety of constructions. To state it, let J be an arbitrary index set and for each $j \in J$, let S_j and T_j be Polish spaces. We shall write $S = \prod_{j\in J} S_j$, $S_F = \prod_{j\in F} S_j$ for $F \subset J$; T and T_F are defined similarly. In each of these product spaces we take the product σ-algebra. Also, let $p_j : S \to S_j$, $p_F : S \to S_F$ be the canonical projections (when no confusion may arise, the canonical projection from S_F onto S_j ($j \in F$) will also be denoted p_j); q_j and q_F are defined similarly for the T_j's.

THEOREM A.1. *For each $j \in J$, let $\phi_j : S_j \to T_j$ be a measurable map and let μ_j be a* p.m. *on S_j. Let λ be a* p.m. *on T and assume: for all $j \in J$,*

$$\mu_j \circ \phi_j^{-1} = q_j\lambda.$$

Then there exists a p.m. *σ on S such that:*

$$p_j\sigma = \mu_j \quad \text{for all} \quad j \in J \quad \text{and} \quad \sigma\circ((\phi_j\circ p_j)_{j\in J})^{-1} = \lambda.$$

PROOF. Let $\mathscr{S}_j$ be the Borel σ-algebra of S_j ($j \in J$). For each $j \in J$, consider the p.s. $(S_j, \mathscr{S}_j, \mu_j)$ and let $\xi_j : T_j \times \mathscr{S}_j \to [0, 1]$ be a proper conditional distribution of Id_{S_j} given ϕ_j. Let F be a finite subset of J. We define a p.m. σ_F on S_F by

$$\sigma_F(A) = \int_{T_F} d\lambda_F(y) \int_{S_F} I_A((x_j)_{j\in F}) \otimes_{j\in F} \xi_j(q_j(y);\, dx_j)$$

where $\lambda_F = q_F\lambda$ and A is a measurable set in S_F. Then σ_F satisfies:

$$\begin{aligned} &(1) \qquad p_j\sigma_F = \mu_j \quad \text{for all} \quad j \in F,\\ &(2) \qquad \sigma_F\circ((\phi_j\circ p_j)_{j\in F})^{-1} = \lambda_F. \end{aligned}$$

In fact, for E measurable in S_j,

$$\begin{aligned} \sigma_F(p_j^{-1}(E)) &= \int d\lambda_F(y) \int I_E(x_j)\xi_j(q_j(y);\, dx_j)\\ &= \int d\lambda_F(y)\xi_j(q_j(y);\, E)\\ &= \int d(q_j\lambda)(v)\xi_j(v;\, E)\\ &= \int d(\mu_j\circ\phi_j^{-1})(v)\xi_j(v;\, E) = \mu_j(E), \end{aligned}$$

proving (1).

Now let F_j be measurable in T_j ($j \in F$). Then

$$\begin{aligned} \sigma_F\circ((\phi_j\circ p_j)_{j\in F})^{-1}\left(\textstyle\prod_{j\in F} F_j\right) &= \sigma_F\left(\textstyle\prod_{j\in F} \phi_j^{-1}(F_j)\right)\\ &= \int d\lambda_F(y) \textstyle\prod_{j\in F} \xi_j(q_j(y);\, \phi_j^{-1}(F_j)). \end{aligned}$$

But for $\mu_j\circ\phi_j^{-1}$—almost all v, $\xi_j(v; \phi_j^{-1}(F_j)) = I_{F_j}(v)$; since $\mu_j\circ\phi_j^{-1} = q_j\lambda_F$ for $j \in F$, it follows that for λ_F—almost all y, $\xi_j(q_j(y); \phi_j^{-1}(F_j)) = I_{F_j}(q_j(y))$. Therefore

$$\phi_F\circ((\phi_j\circ p_j)_{j\in F})^{-1}(\textstyle\prod_{j\in F} F_j) = \lambda_F(\prod_{j\in F} F_j),$$

which implies (2).

The family of p.m.'s $\{\sigma_F: F \text{ finite}, F \subset J\}$ is easily shown to be consistent and thus by Kolmogorov's theorem there exists a p.m. σ on S such that $p_F\sigma = \sigma_F$ for all F finite, $F \subset J$. The desired properties of σ follow now from (1) and (2). □

The following corollary appears in Berkes and Philipp [4]. To simplify the notation, when a space V_j (resp., $V_i \times V_j$) is a factor of product space V, the canonical projection of V onto V_j (resp., $V_i \times V_j$) will be devoted π_j (resp., π_{ij}).

COROLLARY A.2. *Let V_j be a Polish space ($j = 1, 2, 3$). Let $\alpha \in \mathscr{P}(V_1 \times V_2)$, $\beta \in \mathscr{P}(V_2 \times V_3)$ and assume $\pi_2\alpha = \pi_2\beta$. Then there exists $\gamma \in \mathscr{P}(V_1 \times V_2 \times V_3)$ such that $\pi_{12}\gamma = \alpha$, $\pi_{23}\gamma = \beta$.*

PROOF. Let $S_1 = V_1 \times V_2$, $T_1 = V_2$, $\phi_1 = \pi_2$, $S_2 = V_3$, $T_2 = V_3$, $\phi_2 = Id_{V_3}$. Let $\mu_1 = \alpha$, $\mu_2 = \pi_3\beta$, $\lambda = \beta$. Applying Theorem A.1 for $J = \{1, 2\}$, there exists a p.m. γ on $S_1 \times S_2 = V_1 \times V_2 \times V_3$ such that

$$\pi_{12}\gamma = p_1\gamma = \mu_1 = \alpha, \qquad \pi_{23}\gamma = \gamma\circ(\phi_1\circ p_1, \phi_2\circ p_2)^{-1} = \lambda = \beta. \qquad \square$$

Appendix B. On the Kantorovich—Rubinstein theorem. The object of this section is to prove Theorem B.1. This result, together with Theorem B.2 (Theorem 20.4 in [8]), yield Theorem B.3, which is the extension of the Kantorovich-Rubinstein theorem (see reference in [8]) to Polish spaces. Theorem B.3 is stated (for separable metric spaces) in [8] but there is a gap in the proof; Theorem B.1 fills this gap (at least in the Polish case).[2]

The main idea of the proof of Lemma 3 was communicated to us by R. Dudley (St. Flour, June 1980). Some private notes which B. Simon sent to R. Dudley and the present author contain proofs of Lemmas 1 and 4 for the compact metric case.

Throughout the present section, (S, d) will be a Polish space. We define

$$\begin{aligned}
\mathscr{M}(S) &= \text{the space of all finite signed measures on } S,\\
\mathscr{M}_0(S) &= \{\mu \in \mathscr{M}(S): \mu(S) = 0\},\\
\mathscr{M}^+(S) &= \{\mu \in \mathscr{M}(S): \mu \text{ is non-negative}\},\\
\mathscr{M}_1^+(S) &= \{\mu \in \mathscr{M}^+(S): \int d(x, y)\, d\mu(x) < \infty \text{ for some } y \in S\},\\
\mathscr{M}_1(S) &= \{\mu \in \mathscr{M}(S): |\mu| \in \mathscr{M}_1^+(S)\},\\
\mathscr{P}_1(S) &= \mathscr{P}(S) \cap \mathscr{M}_1^+(S),\\
\mathscr{M}_{01}(S) &= \{\mu \in \mathscr{M}_0(S): |\mu| \in \mathscr{M}_1^+(S)\}.
\end{aligned}$$

DEFINITION. Let $\mu, \nu \in \mathscr{M}_1^+(S)$, $\|\mu\| = \|\nu\|$. The *Wasserstein distance* $W(\mu, \nu)$ is defined by

$$W(\mu, \nu) = \inf\left\{\int d\,d\gamma: \gamma \in \mathscr{M}^+(S \times S),\ \pi_1\gamma = \mu,\ \pi_2\gamma = \nu\right\}.$$

[2] After this paper had been completed, R. Dudley kindly showed us a preliminary draft of a note by J. Neveu and himself in which the Kantorovich-Rubinstein theorem is proved by completing the arguments in [8] in a manner partially similar but somewhat different from the present one.

Let $m \in \mathscr{M}_{01}(S)$. The *Wasserstein seminorm* $\| m \|_W$ is defined by

$$\| m \|_W = \inf\left\{ \int d\,d\gamma : \gamma \in \mathscr{M}^+(S \times S),\ \pi_1\gamma - \pi_2\gamma = m \right\}.$$

It is proved in [8] that $\| \cdot \|_W$ is in fact a norm on $\mathscr{M}_{01}(S)$.

THEOREM B.1. *Let $\mu, \nu \in \mathscr{M}_1^+(S)$, $\| \mu \| = \| \nu \|$. Then*

$$W(\mu, \nu) = \| \mu - \nu \|_W.$$

Before proceeding to prove Theorem B.1, we state Theorem B.2 and prove Theorem B.3.

For $f: S \to R$, let us define

$$\| f \|_L = \sup\{ | f(x) - f(y) | / d(x, y) : x, y \in S, x \neq y \},$$
$$\mathrm{Lip}(S) = \{ f: S \to R : \| f \|_L < \infty \}.$$

For $\mu \in \mathscr{M}_1(S)$, define

$$\| \mu \|_L^* = \sup\left\{ \left| \int f\, d\mu \right| : \| f \|_L \leq 1 \right\}.$$

For brevity we shall write $\mathscr{M}_{01}$ instead of $\mathscr{M}_{01}(S)$.

THEOREM B.2 (Theorem 20.4 in [8]). *Let $u: \mathrm{Lip}(S) \to \mathscr{M}_{01}^*$ be defined by $u_f(m) = \int f\, dm$ for $f \in \mathrm{Lip}(S)$, $m \in \mathscr{M}_{01}$. Then u is an isometry of* $(\mathrm{Lip}(S), \| \cdot \|_L)$ *onto* $(\mathscr{M}_{01}^*, \| \cdot \|_W^*)$.

THEOREM B.3. *Let $\mu, \nu \in \mathscr{M}_1^+(S)$, $\| \mu \| = \| \nu \|$. Then*

$$W(\mu, \nu) = \| \mu - \nu \|_L^*.$$

PROOF. By Theorem B.2 and the Hahn-Banach theorem,

$$\| \mu - \nu \|_W = \sup\left\{ \left| \int f\, d(\mu - \nu) \right| : \| f \|_L \leq 1 \right\}$$
$$= \| \mu - \nu \|_L^*.$$

But $W(\mu, \nu) = \| \mu - \nu \|_W$ by Theorem B.1. □

LEMMA 1. *Let $r > 0$. Then W is a pseudo-metric on $\{ \mu \in \mathscr{M}_1^+(S) : \| \mu \| = r \}$.*

PROOF. Clearly it is enough to prove the statement for $r = 1$. Let $\mu, \nu, \lambda \in \mathscr{P}_1(S)$. Given $\epsilon > 0$, choose $\alpha, \beta \in \mathscr{P}(S \times S)$ such that

$$\pi_1\alpha = \mu, \qquad \pi_2\alpha = \nu, \qquad \int d\,d\alpha \leq W(\mu, \nu) + \epsilon,$$
$$\pi_1\beta = \nu, \qquad \pi_2\beta = \lambda, \qquad \int d\,d\beta \leq W(\nu, \lambda) + \epsilon.$$

Let γ be as in Corollary A.2. Observing that the triangle inequality may be rewritten

$$d(\pi_{13}(x, y, z)) \leq d(\pi_{12}(x, y, z)) + d(\pi_{23}(x, y, z)) \qquad x, y, z \in S$$

we have

$$\begin{aligned}\int dd(\pi_{13}\gamma) &= \int (d\circ\pi_{13})\, d\gamma \\ &\le \int (d\circ\pi_{12})\, d\gamma + \int (d\circ\pi_{23})\, d\gamma \\ &= \int dd\alpha + \int dd\beta \\ &\le W(\mu, \nu) + W(\nu, \lambda) + 2\epsilon.\end{aligned}$$

Since $\pi_1(\pi_{13}\gamma) = \mu$, $\pi_2(\pi_{13}\gamma) = \lambda$, this implies

$$W(\mu, \lambda) \le W(\mu, \nu) + W(\nu, \lambda) + 2\epsilon;$$

since ϵ is arbitrary, the result follows. □

LEMMA 2. *Let $\mu, \nu, \in \mathscr{M}_1^+(S)$, $\|\mu\| = \|\nu\|$. Then*

$$W(\mu + \lambda, \nu + \lambda) \le W(\mu, \nu).$$

PROOF. Let $h: S \to S \times S$ be defined by $h(x) = (x, x)$. Given $\epsilon > 0$, choose $\gamma \in \mathscr{M}^+(S \times S)$ so that $\pi_1\gamma = \mu$, $\pi_2\gamma = \nu$,

$$\int dd\gamma \le W(\mu, \nu) + \epsilon.$$

Define $\tilde{\gamma} = \gamma + \lambda\circ h^{-1}$. Then $\pi_1\tilde{\gamma} = \mu + \lambda$, $\pi_2\tilde{\gamma} = \nu + \lambda$,

$$\int dd\tilde{\gamma} = \int dd\gamma + \int dd(\lambda\circ h^{-1}) = \int dd\gamma$$

since $\lambda\circ h^{-1}\{(x, y): d(x, y) > 0\} = 0$. Therefore

$$W(\mu + \lambda, \nu + \lambda) \le \int dd\tilde{\gamma} \le W(\mu, \nu) + \epsilon.$$

Since ϵ is arbitrary, the assertion follows. □

LEMMA 3. *Let $\mu, \nu \in \mathscr{M}_1^+(S)$, $\|\mu\| = \|\nu\|$. Let $c > 0$, $x \in S$. Then*

$$W(\mu, \nu) = W(\mu + c\delta_x, \nu + c\delta_x).$$

PROOF. Given $\epsilon > 0$, choose $\gamma \in \mathscr{M}^+(S \times S)$ such that $\pi_1\gamma = \mu + c\delta_x$, $\pi_2\gamma = \nu + c\delta_x$,

$$\int dd\gamma \le W(\mu + c\delta_x, \nu + c\delta_x) + \epsilon.$$

We will construct a measure $\tilde{\gamma} \in \mathscr{M}^+(S \times S)$ such that $\pi_1\tilde{\gamma} = \mu$, $\pi_2\tilde{\gamma} = \nu$ and $\int dd\tilde{\gamma} \le \int dd\gamma$; this will prove the assertion by arguing as in the previous lemma.

Let $p_x = (x, x)$, $m = \gamma\{p_x\}$. We distinguish two cases:

CASE I. $c \le m$. This case is very simple. Just define

$$\tilde{\gamma} = \gamma - c\delta_{p_x}.$$

Clearly $\pi_1\tilde{\gamma} = \mu$, $\pi_2\tilde{\gamma} = \nu$ and $\int dd\tilde{\gamma} = \int dd\gamma$.

CASE II. $c > m$. This case is more subtle. Define

$$I_x = \pi_1^{-1}(\{x\})\backslash\{p_x\}, \qquad I^x = \pi_2^{-1}(\{x\})\backslash\{p_x\},$$
$$\beta_2 = \pi_2(\gamma(I_x)^{-1}\gamma\,|\,I_x), \qquad \beta_1 = \pi_1(\gamma(I^x)^{-1}\gamma\,|\,I^x)$$

(observe that $\gamma(I_x) = \gamma(\pi_1^{-1}(\{x\})) - m \geq c - m > 0$; similarly $\gamma(I^x) > 0$). It is checked at once that

$$\gamma\,|\,I_x = \gamma(I_x)(\delta_x \otimes \beta_2), \qquad \gamma\,|\,I^x = \gamma(I^x)(\beta_1 \otimes \delta_x).$$

Define now, putting $q = c - m$,

$$\tilde{\gamma} = \gamma - q\gamma(I^x)^{-1}\gamma\,|\,I_x - q\gamma(I^x)^{-1}\gamma\,|\,I^x - m\delta_{p_x} + q(\beta_1 \otimes \beta_2).$$

Then $\tilde{\gamma} \in \mathscr{M}^+(S \times S)$ and

$$\pi_1\tilde{\gamma} = \pi_1\gamma - q\delta_x - q\beta_1 - m\delta_x + q\beta_1 = \pi_1\gamma - c\delta_x = \mu$$

and similarly $\pi_2\tilde{\gamma} = \nu$.

We will show now that $\int d\,d\tilde{\gamma} \leq \int d\,d\gamma$. By the triangle inequality, for any $z, y \in S$,

$$d(z, y) \leq d(z, x) + d(x, y)$$

and hence

$$\begin{aligned}\iint d(z, y)\, d(\beta_1 \otimes \beta_2)(z, y) &\leq \iint d(z, x)\, d(\beta_1 \otimes \beta_2)(z, y)\\ &\quad + \iint d(x, y)\, d(\beta_1 \otimes \beta_2)(z, y)\\ &= \int d(z, x)\, d\beta_1(z) + \int d(x, y)\, d\beta_2(y)\\ &= \iint d(z, y)\, d(\beta_1 \otimes \delta_x)(z, y)\\ &\quad + \iint d(z, y)\, d(\delta_x \otimes \beta_2)(z, y).\end{aligned}$$

Consequently

$$\begin{aligned}\int d\,d\tilde{\gamma} &= \int d\,d\gamma - q\int d\,d(\delta_x \otimes \beta_2) - q\int d\,d(\beta_1 \otimes \delta_x) + q\int d\,d(\beta_1 \otimes \beta_2)\\ &\leq \int d\,d\gamma.\end{aligned}$$

□

LEMMA 4. *The set $\{\mu \in \mathscr{P}_1(S) : \mu$ has finite support$\}$ is W-dense in $\mathscr{P}_1(S)$.*

PROOF. By the separability of S, given a fixed $y_0 \in S$ and $\epsilon > 0$, there exist disjoint Borel sets $A_1, \cdots, A_n$ such that diam $A_j \leq \epsilon$ $(j = 1, \cdots, n)$ and

$$\int_D d(x, y_0)\, d\mu(x) < \epsilon,$$

where $D = (\cup_{j=1}^n A_j)^c$.

For $y \in S$, we define $h_y(x) = (x, y)$ for $x \in S$. For $j = 1, \cdots, n$, choose $y_j \in A_j$ and define $\gamma \in \mathscr{P}(S \times S)$ by

$$\gamma(E) = \textstyle\sum_{j=1}^n \mu(A_j \cap h_{y_j}^{-1}(E)) + \mu(D \cap h_{y_0}^{-1}(E))$$

for $E \in \mathscr{B} \otimes \mathscr{B}$. Then, since $\pi_1 \circ h_y = Id_S$ for all $y \in S$, we have for $A \in \mathscr{B}$

$$\gamma(\pi_1^{-1}(A)) = \textstyle\sum_{j=1}^n \mu(A_j \cap A) + \mu(D \cap A) = \mu(A);$$

and since $(\pi_2 \circ h_y)(x) = y$ for all $x \in S$,

$$\gamma(\pi_2^{-1}(A)) = \textstyle\sum_{j=1}^n \mu(A_j)\, \delta_{y_j}(A) + \mu(D)\delta_{y_0}(A).$$

Now

$$\int d\,d\gamma = \textstyle\sum_{j=1}^n \displaystyle\int_{A_j} d(x, y_j)\, d\mu(x) + \int_D d(x, y_0)\, d\mu(x)$$

$$\leq \epsilon \textstyle\sum_{j=1}^n \mu(A_j) + \epsilon$$

$$\leq 2\epsilon.$$

Therefore

$$W(\mu, \textstyle\sum_{j=1}^n \mu(A_j)\delta_{y_j} + \mu(D)\delta_{y_0}) \leq 2\epsilon. \qquad \square$$

PROOF OF THEOREM B.1. We prove first: for all $\mu, \nu, \lambda \in \mathscr{M}_1^+(S)$, if $\|\mu\| = \|\nu\|$ then

$$\text{(I)} \qquad W(\mu + \lambda, \nu + \lambda) = W(\mu, \nu).$$

By Lemma 4, which is obviously valid for the class of measures in $\mathscr{M}_1^+(S)$ with a fixed total mass, there exists a sequence $\{\lambda_n\}$, $\lambda_n \in \mathscr{M}_1^+(S)$, λ_n has finite support, $\|\lambda_n\| = \|\lambda\|$ and $W(\lambda, \lambda_n) \to 0$. By Lemma 3 and induction,

$$W(\mu + \lambda_n, \nu + \lambda_n) = W(\mu, \nu).$$

Hence

$$W(\mu, \nu) \leq W(\mu + \lambda_n, \mu + \lambda) + W(\mu + \lambda, \nu + \lambda) + W(\nu + \lambda, \nu + \lambda_n)$$

$$\leq W(\mu + \lambda, \nu + \lambda) + 2W(\lambda, \lambda_n),$$

since by Lemma 2 we have: $W(\mu + \lambda_n, \mu + \lambda) \leq W(\lambda_n, \lambda)$, $W(\nu + \lambda, \nu + \lambda_n) \leq W(\lambda, \lambda_n)$. Therefore, letting $n \to \infty$,

$$W(\mu, \nu) \leq W(\mu + \lambda, \nu + \lambda);$$

this inequality and Lemma 2 yield (I).

Obviously $\|\mu - \nu\|_W \leq W(\mu, \nu)$. Given $\epsilon > 0$, let $\gamma \in \mathscr{M}^+(S \times S)$ be such that $\pi_1\gamma - \pi_2\gamma = \mu - \nu$ and

$$\int d\,d\gamma \leq \|\mu - \nu\|_W + \epsilon.$$

Let $\xi = (\mu - \nu)^+$, $\eta = (\mu - \nu)^-$. Then there exist $\sigma, \sigma' \in \mathscr{M}_1^+(S)$ such that

$$\mu = \xi + \sigma, \qquad \nu = \eta + \sigma,$$

$$\pi_1\gamma = \xi + \sigma', \qquad \pi_2\gamma = \eta + \sigma',$$

and by (I),

$$W(\mu, \nu) = W(\xi, \eta) = W(\pi_1\gamma, \pi_2\gamma) \leq \int d\,d\gamma \leq \|\mu - \nu\|_W + \epsilon.$$

Since ϵ is arbitrary, the proof is complete. $\square$

ADDENDUM. H. Dehling and W. Philipp have observed in a very recent preprint that taking Theorem 3.1 as a starting point, its main part (in the case $\lambda_n = \nu^{1/j_n}$) may be improved by replacing convergence in probability by a.s. convergence. We shall show that this fact follows easily from Theorem A.1 and a well-known result of Skorohod [17] (see also [8]).

PROPOSITION. *Under the conditions of Theorem* 3.1, (3) *may be strengthened as follows*:

(3′) $$\max_{k\le j_n} \| S_{nk} - T_{nk} \| \to 0 \qquad a.s.$$

PROOF. Let $\{X'_{nj}\}$, $\{Y'_{nj}\}$ be triangular arrays satisfying (1)–(3) of Theorem 3.1, and let $\alpha_n = \mathscr{L}(\max_{k\le j_n} \| S'_{nk} - T'_{nk} \|)$; then $\alpha_n \to_w \delta_0$. By Skorohod's lemma ([17]; also [8]), there exist random variables $\eta_n \to 0$ a.s. with $\mathscr{L}(\eta_n) = \alpha_n$. Let $\lambda = \mathscr{L}(\{\eta_n\}_{n\in N})$, $\beta_n = \mathscr{L}(\{X'_{nj}\}_{j=1,\dots,j_n};\{Y'_{nj}\}_{j=1,\dots,j_n})$ and define $\phi_n : B^{j_n} \times B^{j_n} \to R$ by $\phi_n((x_j);(y_j)) = \max_{k\le j_n} \| s_k - t_k \|$, where $s_k = \sum_{j=1}^k x_j$, $t_k = \sum_{j=1}^k y_j$. Then $\beta_n \circ \phi_n^{-1} = \alpha_n = q_n \lambda$. By Theorem A.1, there exists a p.m. σ on $\Omega = \prod_{n=1}^\infty (B^{j_n} \times B^{j_n})$ such that $p_n \sigma = \beta_n$ for all $n \in N$ and $\sigma \circ ((\phi_n \circ p_n)_{n\in N})^{-1} = \lambda$. Define now on (Ω, σ):

X_{nj} = jth coordinate of the canonical projection of Ω onto the first factor of $B^{j_n} \times B^{j_n}$,

Y_{nj} = jth coordinate of the canonical projection of Ω onto the second factor of $B^{j_n} \times B^{j_n}$. Then $\mathscr{L}(\{\max_{k\le j_n} \| S_{nk} - T_{nk} \| \}_{n\in N}) = \lambda$. □

Acknowledgment. I thank J. Samur for many stimulating conversations; in particular, for a remark in connection to the use of Lemma 2.1 (2) in the proof of Theorem 3.1. I also thank M. Cotlar for the discussion of a measure-theoretic point in a previous version of the proof of Theorem 4.9 and R. Dudley and B. Simon for some exchanges on the Kantorovich-Rubinstein theorem.

REFERENCES

[1] DE ACOSTA, A. (1981). An invariance principle in probability for triangular arrays of B-valued random vectors. Aspects Statistiques et Aspects physiques des Processus Gaussiens. Colloques International du C.N.R.S. No. 307. Editions du C.N.R.S., Paris.
[2] DE ACOSTA, A., ARAUJO, A. and GINÉ, E. (1978). On Poisson measures, Gaussian measures and the central limit theorem in Banach spaces. *Adv. in Probability* 4 1–68. Dekker, New York.
[3] DE ACOSTA, A. and GINÉ, E. (1979). Convergence of moments and related functionals in the general central limit theorem in Banach spaces. *Z. Wahrsch. verw. Gebiete* **48** 213–231.
[4] BERKES, I. and PHILIPP, W. (1979). Approximation theorems for independent and weakly dependent random vectors. *Ann. Probability* **7** 29–54.
[5] BILLINGSLEY, P. (1968). *Convergence of probability measures.* Wiley, New York.
[6] BREIMAN, L.(1968). *Probability.* Addison-Wesley, Reading, Mass.
[7] CHUNG, K. L. (1974). *A course in Probability Theory* (2nd edition). Academic, New York.
[8] DUDLEY, R. M. (1976). Probabilities and metrics. Lecture Notes Series No. 45. Matematisk Institut, Aarhus University.
[9] FELLER, W. (1971). *An Introduction to Probability Theory and its Applications*, Vol. II (2nd edition). Wiley, New York.
[10] FREEDMAN, D. (1971). *Brownian Motion and Diffusion.* Holden Day, San Francisco.
[11] GIKHMAN, I. I. and SKOROHOD, A. V. (1969). *Introduction to the theory of random processes.* Saunders, Philadelphia.
[12] LE CAM, L. (1965). On the distribution of sums of independent random variables. Bernoulli, Bayes, Laplace (Proceedings of a seminar) 179–202. Springer-Verlag, Berlin and New York.
[13] MAJOR, P. (1976). Approximation of partial sums of i.i.d. r.v.'s when the summands have only two moments. *Z. Wahrsch. verw. Gebiete* **35** 221–229.
[14] MAJOR, P. (1978). On the Invariance Principle for sums of independent identically distributed random variables. *J. Multivariate Anal.* **8** 487–571.
[15] PARTHASARATHY, K. R. (1967). Probability measures on metric spaces. Academic, New York.
[16] PHILIPP, W. (1980). Weak and L^p-invariance principles for sums of B-valued random variables. *Ann. Probability* **8** 68–82.
[17] SKOROHOD, A. V. (1956). Limit theorems for stochastic processes. *Theor. Probability Appl.* **1** 261–290.
[18] SKOROHOD, A. V. (1957). Limit theorems for stochastic processes with independent increments. *Theor. Probability Appl.* **2** 138–171.

DEPARTAMENTO DE MATHEMATICAS
IVIC
APARTADO 1827. CARACAS 1010-A, VENEZUELA