AN INVARIANCE PRINCIPLE FOR A CLASS OF d-DIMENSIONAL POLYGONAL RANDOM FUNCTIONS([1])

BY

LUIS G. GOROSTIZA

ABSTRACT. A class of random functions is formulated, which represent the motion of a point in d-dimensional Euclidean space $(d \geq 1)$ undergoing random changes of direction at random times while maintaining constant speed. The changes of direction are determined by random orthogonal matrices that are irreducible in the sense of not having an almost surely invariant nontrivial subspace if $d \geq 2$, and not being almost surely nonnegative if $d = 1$. An invariance principle stating that under certain conditions a sequence of such random functions converges weakly to a Gaussian process with stationary and independent increments is proved. The limit process has mean zero and its covariance matrix function is given explicitly. It is shown that when the random changes of direction satisfy an appropriate condition the limit process is Brownian motion. This invariance principle includes central limit theorems for the plane, with special distributions of the random times and direction changes, that have been proved by M. Kac, V. N. Tutubalin and T. Watanabe by methods different from ours. The proof makes use of standard methods of the theory of weak convergence of probability measures, and special results due to P. Billingsley and B. Rosén, the main problem being how to apply them. For this, renewal theoretic techniques are developed, and limit theorems for sums of products of independent identically distributed irreducible random orthogonal matrices are obtained.

1. **Introduction and notation.** In this work we present an invariance principle which is an extension of results obtained by M. Kac, V. N. Tutubalin and T. Watanabe. Let n be a positive integer and consider a point moving in the plane, starting from the origin at time zero, maintaining constant speed c_n, and changing direction in the following way: it stays on the ith direction for a random amount of time $\tau_{n,i}$ and then changes direction by a random angle α_i. Suppose the α_i are independent with common distribution μ. M. Kac has shown in [5] (although not stating the problem this way) that if $c_n = n$, $\tau_{n,i} = 1/n$ for all i, and μ is symmetric (i.e. $\mu(A) = \mu(-A)$), then at time $t = 1$, that is, after the nth step, the two coordinates of the position of the point, suitably normalized, jointly have an asymptotic

---

distribution as $n \to \infty$ which is that of two independent Gaussian random variables. V. N. Tutubalin has the same result in [8] when α_i is θ or $-\theta$, each with probability $1/2$, and θ is incommensurable with respect to 2π (and other minor differences), as a special case of a three-dimensional theorem of a type different from ours. T. Watanabe in [9] has proved that when the $\tau_{n,i}$ are independent and exponentially distributed with parameter c_n^2, and the α_i are uniformly distributed on $[-\pi, \pi)$ and independent of the $\tau_{n,i}$, the process (the space of trajectories) converges weakly to planar normalized Brownian motion when $c_n \to \infty$ as $n \to \infty$. Our purpose is to extend such results both in the type of the distributions involved and in the dimension of the space where the motion takes place. Since the complete definition of our random functions is long, we shall give now a rather nonstringent formulation of them and their invariance principle. Let a point move in a d-dimensional Euclidean space $(d \geq 1)$, changing direction at random times as above, with the direction changes being given by independent (and independent of the previous direction) and identically distributed random orthogonal matrices; assume these matrices are irreducible in the sense of not having an a.s. invariant nontrivial subspace if $d \geq 2$, and not being a.s. nonnegative if $d = 1$, and moreover let them depend on n, as do the random times; suppose the random times tend to zero in some adequate way, and $c_n \to \infty$ as $n \to \infty$; then under certain more technical than qualitative conditions the process determined by the trajectory of the point, suitably normalized, converges weakly as $n \to \infty$ to a Gaussian process with stationary and independent increments, mean zero and covariance matrix function which we can exhibit. The limit process is Brownian motion if the change of direction matrices satisfy a certain property (Condition (A)). Important parts of the proof are: a theorem of B. Rosén [6], used to prove the weak convergence of the one(time)-dimensional distributions, which turns out to imply the convergence of all the finite-dimensional distributions; the use of inequalities of P. Billingsley [1] (§12) to prove the basic Lemma 3.4, which implies the tightness of the sequence of random functions and is used in applying Rosén's theorem; renewal theoretic techniques that are employed throughout the proof; and limit theorems for irreducible random orthogonal matrices, which enter into the application of Rosén's theorem and are the tool that made the proof possible for any dimension ≥ 3. Kac's three-dimensional result in [5] is not included in our theorem.

In the remainder of this section we establish the notation. In §2 we define the random functions, state the invariance principle and some observations about it, and notice that the results previously cited are special cases. The proofs of the invariance principle and the observations form §3. §4 consists of the renewal theoretic results, and §5 contains some facts and the limit theorems regarding irreducible random orthogonal matrices. These two sections, although they are part of the proof, have been separated due to their independent nature, and for the same

reason the results therein are sometimes more general than is required by the proof, but only to the extent that seems natural without much extra effort; also, they will be used for further work.

Notation. Vectors are column-vectors, $'$ denotes matrix transposition, (,) stands for inner product of vectors and tr for the trace of a matrix. The vector norm $\|x\| = (\Sigma_{i=1}^{d} x_i^2)^{1/2}$, with $x = (x_1, \cdots, x_d)'$, and the matrix norm $\|A\| = \sup\{\|Ax\|: \|x\| = 1\}$ are used. E stands for expectation, $E[\mid \mathcal{F}]$ for conditional expectation with respect to the Borel field $\mathcal{F}$, and $E[\mid X_\alpha, \ \alpha \in \mathcal{C}]$ for conditional expectation with respect to $\mathcal{F}\{X_\alpha, \ \alpha \in \mathcal{C}\}$, the last symbol denoting the Borel field generated by the random elements $X_\alpha, \ \alpha \in \mathcal{C}$. With the above norms we have for a random matrix A of order d and a Borel field $\mathcal{F}$ the basic inequality

$$\|E[A\mid\mathcal{F}]\| \leq \sqrt{d}\, E[\|A\| \mid \mathcal{F}] \quad \text{a.s.,}$$

which is a consequence of the elementary inequalities $(\Sigma_{i=1}^{d} x_i^2)^{1/2} \leq \Sigma_{i=1}^{d} |x_i| \leq \sqrt{d}(\Sigma_{i=1}^{d} x_i^2)^{1/2}$. The distribution of the random element X is denoted μ_X, and δ_x is the probability measure concentrated at the point x. The symbols $\mathcal{D}$, P and a.s. are abbreviations for "in distribution", "in probability" and "almost surely". [ ] denotes the integral part of a real number, and when a real number appears in a position where there obviously should be an integer it is interpreted as its integral part. The symbol $^+$ stands for positive part. The indicator function of the set S is denoted 1_S. The identity matrix is I.

The spaces $D[0, T]^d$ and $D[0, \infty)^d$ of functions from $[0, T]$ and $[0, \infty)$, respectively, into the *d*-dimensional Euclidean space R^d $(d \geq 1)$ which are right-continuous with left limits everywhere are the same sets as the *d*-fold products of $D[0, T]^1$ with itself and $D[0, \infty)^1$ with itself, respectively. $C[0, \infty)^d$ is the subspace of $D[0, \infty)^d$ consisting of functions continuous everywhere. A metric on $D[0, \infty)^d$ exists which is separable and complete, and for each finite $T > 0$ its restriction to the set $\{f1_{[0,T]} : f \in D[0, \infty)^d\}$ is equivalent to the Skorokhod metric ρ on $D[0, T]^d$, namely $\rho(x, y) = \max_{1 \leq i \leq d} \delta(x_i, y_i)$, for $x = (x_1, \cdots, x_d)'$ and $y = (y_1, \cdots, y_d)'$, with

$$\delta(x_i, y_i) = \inf_{\lambda \in \Lambda} \left\{ \sup_{0 \leq t \leq T} |x_i(t) - y_i(\lambda(t))| + \sup_{0 \leq s, t \leq T; \ s \neq t} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right| \right\},$$

where Λ is the set of all strictly increasing and continuous functions of $[0, T]$ onto itself (see [7], [11]). $D[0, \infty)^d$ will denote this metric space. Convergence of random elements of this space is convergence in distribution, or weak convergence of their distributions. Our basic source for the weak convergence theory used here is P. Billingsley's book [1].

2. The random functions and the invariance principle.

2.1. *The random functions.* The random functions we shall study arise from the following physical process. A point moves in a d-dimensional Euclidean space R^d $(d \geq 1)$, starting from the origin at time $t = 0$ in a random direction (unit vector) ξ, with constant speed c, for a random amount of time τ_1, at which time it takes on a new direction, $A_1\xi$, where A_1 is a random orthogonal matrix, and the motion continues on by repetition of this pattern; thus, departing from the position $c\tau_1\xi$ the point travels in the random direction $A_1\xi$, with constant speed c, for a random amount of time τ_2, when it takes on the new direction $A_2A_1\xi$, where A_2 is a random orthogonal matrix; et cetera. We consider a sequence of such random motions, and define the nth random function $X_n(t)$, $t \geq 0$, to be the position at time t of the point in the nth motion. Thus, for $n = 1,2,\cdots$ (with $A_{n,0} = I$)

$$
X_n(t) = c_n \sum_{i=1}^{N_n(t)} \tau_{n,i} A_{n,i-1} \cdots A_{n,0}\xi_n
$$

(1)

$$
+ c_n \left(t - \sum_{i=1}^{N_n(t)} \tau_{n,i} \right) A_{n,N_n(t)} \cdots A_{n,0}\xi_n, \qquad t \geq 0,
$$

where $N_n(t) = \max\{k: \sum_{i=1}^{k} \tau_{n,i} \leq t\}$, $t \geq 0$, and c_n is a positive constant (the speed). The function

(2)
$$
\hat{X}_n(t) = c_n \sum_{i=1}^{N_n(t)} \tau_{n,i} A_{n,i-1} \cdots A_{n,0}\xi_n, \qquad t \geq 0,
$$

describes the pure jumping motion which consists in staying at each turning position until the time of the next turn. We make the following assumptions. For each n, the nonnegative random variables (times) $\tau_{n,i}$, $i \geq 1$, are not identically zero, are independent, and have the same distribution (possibly with an atom at 0), the random matrices $A_{n,i}$, $i \geq 1$, are independent, identically distributed and jointly independent of $\{\tau_{n,i}\}$, and the initial random direction ξ_n is independent of $\{\tau_{n,i}, A_{n,i}\}$. So defined, the $\hat{X}_n$ are random elements of $D[0, \infty)^d$ and the X_n are random elements of the subspace $C[0, \infty)^d$ of $D[0, \infty)^d$. A probability space $(\Omega_n, \mathcal{F}_n, P_n)$ on which X_n and $\hat{X}_n$ are defined of course exists, and we follow the usual practice of omitting the index n from the notation of the probability measure.

We wish to obtain an invariance principle for random functions of the types (1) or (2), that is, given a sequence of such functions satisfying certain conditions, we want to prove the existence of, and exhibit, a weak limit of the (possibly normalized) sequence, which limit, expectedly a diffusion, depends on the distributions of the random times and random matrices involved in the definitions of the elements of the sequence only through some of their moments. It will be seen

that under the conditions of the invariance principle weak limits of the sequences $\{X_n\}$ and $\{\hat{X}_n\}$ are the same, and therefore the principle need only be stated for $\{X_n\}$.

As they are at this point, our random functions are too general to yield the invariance theorem we want, and therefore we make further assumptions. The times $\tau_{n,i}$ should somehow tend to zero as $n \to \infty$. We shall assume that there is a sequence $\{\tau_i\}$ of random variables and a sequence $\{b_n\}$ of positive real numbers converging to infinity such that $\tau_{n,i} = \tau_i/b_n$. Letting $N(t) = \max\{k: \Sigma_{i=1}^{k} \tau_i \le t\}, t \ge 0$, we have $N_n(t) = N(b_n t), t \ge 0$. Regarding the orthogonal matrices $A_{n,i}$, in order to prevent the motion from remaining in a proper subspace of R^d, $d \ge 2$, or from not being able to change direction in R^1, we require that they be irreducible, the definition of an irreducible (as opposed to reducible) random matrix A (of order d) being that $P[AV \subseteq V] < 1$ for every nontrivial subspace V of R^d, for $d \ge 2$, and that the single entry of A not be a.s. nonnegative for $d = 1$. We mention at this point that if a random orthogonal matrix A is irreducible the matrix $I - EA$ is nonsingular (see § 5). We do not require the orthogonal matrices $A_{n,i}$ to be proper (i.e. with determinant $+ 1$). These qualitative conditions complete the definition of the class of random functions that concerns us here.

2.2. *The invariance principle.*

Theorem. *If*

(i) $E\tau_1^\rho < \infty$ *for some* $\rho > 3$,

(ii) $\{A_{n,1}, n \ge 1\}$ *has no reducible weak limit points and* $\lim_{n\to\infty} (I - EA_{n,1})^{-1}$ *exists, and*

(iii) $a_1, a_2, \cdots$ *are positive real numbers such that* $\lim_{n\to\infty}(c_n/a_n b_n^{1/2}) = 1$,

then as $n \to \infty$ *the sequence of random functions* $\{X_n/a_n\}$ *converges in distribution on* $D[0, \infty)^d$ *to the Gaussian random function* X *with stationary and independent increments, mean* 0 *and covariance matrix function* $EX(t)X(t)' = tC/dE\tau_1$, *where*

$$C = \text{Var}\,\tau_1 I + (E\tau_1)^2 \left\{ \lim_{n\to\infty} \, [(I - EA_{n,1})^{-1} + (I - EA'_{n,1})^{-1}] - I \right\} .$$

Let us make a few observations, some of which we will verify at the end of § 3.

1. Even with the provisions of irreducibility of the $A_{n,1}$ and their weak limit points, the limit X may be degenerate. Example: In R^3, with τ_1 constant and

$$A_{n,i} = A_i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

each with probability ½ for all n, we have rank $C = 2$, so that X concentrates on a

two-dimensional subspace. What happens in this example is that, if x, y, z are the coordinate axes, the y-component of X_n/a_n is bounded by $\sqrt{3}\, c_n \tau_1/a_n b_n$, which converges to zero as $n \to \infty$. To see why this is so let u_x, u_y, u_z be the unit vectors along the corresponding coordinate directions and let $U = \{u_x, -u_x, u_y, -u_y, u_z, -u_z\}$; since for each i, $A_{i-1} \cdots A_1 U \subseteq U$, it suffices to take u_y or $-u_y$ as the initial direction; now, it can easily be seen that with probability one $A_1 u_y$ lies in the xz-plane, $A_{i-1} \cdots A_1 u_y$ does not lie in the xz-plane for some $i > 2$, and for the first such i, $A_{i-1} \cdots A_1 u_y = -u_y$, and that the same holds with $-u_y$ in place of u_y; hence the y-component of $(c_n/a_n) \sum_{i=1}^{N_n(t)} \tau_{n,i} A_{n,i-1} \cdots A_{n,0} u_y$ is 0, $c_n \tau_1/a_n b_n$ or $-c_n \tau_1/a_n b_n$. In this example we also have that in the xz-plane the component processes of the limit X are dependent. In general we can say, however, that the limit X is degenerate only if τ_1 is a.s. constant, that rank $C \geq 2$ when $d \geq 3$, and that, for $d = 1$ or 2, X can be totally degenerate (i.e. $C = 0$), which occurs if and only if both τ_1 and all weak limit points of $\{A_{n,1}\}$ are a.s. constant. For $d \geq 3$ every (constant) orthogonal matrix is reducible and hence cannot be a weak limit point of $\{A_{n,1}\}$ (this is why X cannot be totally degenerate), but, for $d = 1$ or 2, $\{A_{n,1}\}$ can have a.s. constant weak limit points (-1 for $d = 1$).

2. Of special interest are random motions of our class whose changes of direction satisfy the following property. If A_1, A_2 and A_3 are independent and identically distributed random orthogonal matrices (of order d), then, as functions of the direction x, the distributions of the projection of $A_2 x$ on x and of the projection of $A_3' A_2 x$ on x are invariant under the transformation A_1, i.e. for any direction x

$$(A_2 A_1 x, A_1 x) \overset{\mathcal{D}}{=} (A_2 x, x) \quad \text{and} \quad (A_3' A_2 A_1 x, A_1 x) \overset{\mathcal{D}}{=} (A_3' A_2 x, x).$$

It follows from this that the random matrix $A = A_1$ satisfies the condition

$$E(A'(EA + EA')A) = EA + EA' \quad \text{and} \quad E(A'(EA'EA)A) = EA'EA,$$

or equivalently, by Lemma 5.2,

Condition (A). $EA + EA' = \alpha I$ and $EA'EA = \beta I$ for some constants α and β.

This condition holds automatically for $d = 1$, and when A is proper, also for $d = 2$. If the matrices $A_{n,1} = (a_{ij}^{(n)})$ satisfy Condition (A), then

$$C = \left(\operatorname{Var} \tau_1 + (E\tau_1)^2 \lim_{n \to \infty} \frac{1 - \sum_{j=1}^{d} (Ea_{ji}^{(n)})^2}{1 - 2Ea_{ii}^{(n)} + \sum_{j=1}^{d} (Ea_{ji}^{(n)})^2} \right) I,$$

where i is arbitrary. Therefore the limit X is a d-dimensional Brownian motion process (understood here as a d-dimensional vector process whose components are independent (one-dimensional) Brownian motion processes with the same variance parameter), and hence for $d \geq 3$ it is nondegenerate. We further remark that under

Condition (A), for $d \geq 3$ if A is a weak limit point of $\{A_{n,1}\}$ then $\|EA\| < 1$, because $\|EA\| = 1$ is equivalent to A being a.s. constant and hence reducible. We note that the matrix $A_{n,1}$ in the above example does not satisfy Condition (A).

3. For $d = 2$ we represent the proper rotation $A_{n,1}$ in the form

$$A_{n,1} = \begin{pmatrix} \cos \alpha_n & -\sin \alpha_n \\ \sin \alpha_n & \cos \alpha_n \end{pmatrix},$$

where α_n is a random variable taking values in the interval $[-\pi, \pi)$. The condition of irreducibility takes the form $\mu_{\alpha_n} \neq p\delta_{-\pi} + (1-p)\delta_0$, $0 \leq p \leq 1$. The limit X is a two-dimensional (or totally degenerate) Brownian motion process with variance parameter

$$\frac{1}{2E\tau_1}\left(\operatorname{Var} \tau_1 + (E\tau_1)^2 \lim_{n \to \infty} \frac{1 - (E \cos \alpha_n)^2 - (E \sin \alpha_n)^2}{(1 - E \cos \alpha_n)^2 + (E \sin \alpha_n)^2}\right).$$

Suppose that α is a weak limit point of $\{\alpha_n\}$ $(\mu_\alpha \neq p\delta_{-\pi} + (1-p)\delta_0, \ 0 \leq p \leq 1)$. If α is a.s. constant X has the minimum possible variance parameter, $\operatorname{Var} \tau_1/2E\tau_1$; but if α is not a.s. constant the α_n can be shifted (mod the interval $[-\pi, \pi)$) so that X has any variance parameter in the interval

$$\left[(2E\tau_1)^{-1} (\operatorname{Var} \tau_1 + (E\tau_1)^2\beta), \ (2E\tau_1)^{-1}\left(\operatorname{Var} \tau_1 + (E\tau_1)^2\beta^{-1}\right) \right],$$

where

$$\beta = \frac{1 = ((E \cos \alpha)^2 + (E \sin \alpha)^2)^{1/2}}{1 + ((E \cos \alpha)^2 + (E \sin \alpha)^2)^{1/2}}.$$

In particular, if $E \cos \alpha = E \sin \alpha = 0$ we have the unique limit variance parameter $E\tau_1^2/2E\tau_1$, and we note that in this case the limits as $n \to \infty$ of the mean and variance parameter of X_n/a_n are equal to the limits as $n \to \infty$ of the mean and variance parameter of the random vector $\xi\sqrt{\tau_{N(n)}}$, where the unit vector ξ has the Haar distribution and is independent of $\{\tau_i\}$.

2.3. *Special cases*. We will include here only special cases that yield results obtained by other authors, and other means, for $d = 2$ and proper rotations.

Let $\alpha_n \overset{\mathcal{D}}{=} \alpha$ for all n, with α symmetrically distributed (i.e. $\mu_\alpha(A) = \mu_\alpha(-A)$ for each Borel set A); then the limit Brownian motion process has variance parameter

$$\sigma^2 = \frac{1}{2E\tau_1}\left(\operatorname{Var} \tau_1 + (E\tau_1)^2 \frac{1 + E \cos \alpha}{1 - E \cos \alpha}\right).$$

With $\tau_1 \equiv 1$, $b_n = c_n = n$, $a_n = \sqrt{n}$, $\xi_n = (\cos \alpha_0, \sin \alpha_0)'$ and $t = 1$ we obtain that the random vectors

$$\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \cos \sum_{j=0}^{i-1} \alpha_j, \frac{1}{\sqrt{n}} \sum_{i=0}^{n} \sin \sum_{j=0}^{i-1} \alpha_j \right)',$$

where α_j, $j \geq 0$, are independent and α_j, $j \geq 1$, are distributed as α, converge in distribution as $n \rightarrow \infty$ to $(N_1, N_2)'$, where N_1 and N_2 are independent normal distributions with mean 0 and variance $\sigma^2 = (1 + E \cos \alpha)/2(1 - E \cos \alpha)$. This central limit theorem appears in M. Kac's book [5], with α_0 also distributed as α and without an explicit limit variance, and in V. N. Tutubalin's paper [8], with $\mu_\alpha = (\delta_\theta + \delta_{-\theta})/2$ and the requirement that θ be incommensurable with respect to 2π, which, as we have seen, is unnecessary.

Now let r_1 have the gamma distribution $\Gamma(\beta, \lambda)$, $b_n = c_n^2$, $a_n = 1$ and $\xi_n = (\cos \dot{\alpha}_{n,0}, \sin \alpha_{n,0})'$, and let $\alpha_{n,j}$, $j \geq 1$, be distributed as α_n. Then the corresponding vector processes (which we might call generalized scattering transport processes) converge weakly as $n \rightarrow \infty$ to the planar Brownian motion process with variance parameter

$$\sigma^2 = \frac{1}{2} \left(\frac{1}{\lambda} + \frac{\beta}{\lambda} \lim_{n \to \infty} \frac{1 - (E \cos \alpha_n)^2 - (E \sin \alpha_n)^2}{(1 - E \cos \alpha_n)^2 + (E \sin \alpha_n)^2} \right).$$

If all $\alpha_{n,j}$ are uniformly distributed on $[-\pi, \pi)$ and $\beta = \lambda = 1$ (so that $r_{n,1}$ is exponentially distributed with parameter c_n^2), then $\sigma^2 = 1$, which coincides with T. Watanabe's result in his paper [9].

3. The proof. The hypothesis of the invariance theorem and the basic inequality (§1) will frequently be used without mention. The proof employs the results of §§4 and 5. To simplify the notation we will assume that $a_n = 1$, which implies no loss of generality, and we can obviously assume $\rho \leq 5$.

3.1. *Preparatory lemmas.* We will show first that weak limits of the random functions X_n and their associated jumping processes $\hat{X}_n$ given by (1) and (2), respectively, are the same.

Lemma 3.1. *If $Er_1^\alpha < \infty$ for some $\alpha > 0$ and the sequence $\{c_n/b_n^{(\alpha-1)/\alpha}\}$ is bounded, then $\{X_n\}$ and $\{\hat{X}_n\}$ both converge weakly to the same limit or neither converges.*

Proof. Since the Skorokhod metric ρ on $D[0, T]^d$ is dominated by the sup norm metric we have from (1) and (2) that

$$\rho(X_n, \hat{X}_n) \leq c_n \sup_{0 \leq t \leq T} \left(t - \sum_{i=1}^{N_n(t)} r_{n,i} \right) = \frac{c_n}{b_n} \sup_{0 \leq t \leq b_n T} \left(t - \sum_{i=1}^{N(t)} r_i \right),$$

and the last random variable converges to zero in probability as $n \rightarrow \infty$ by Lemma 4.1 (put $b_n = d_n^\alpha$, then $c_n/b_n = (c_n/b_n^{(\alpha-1)/\alpha})(1/d_n)$). This implies that the con-

clusion is true for the random functions restricted to the time interval $[0, T]$ for each finite $T > 0$ (see $[1$, Theorem 4.1]), but this is sufficient (see $[11]$). □

We remark that this lemma is valid if the matrices $A_{n,i}$ are only contractive.

From the hypothesis of the invariance theorem we see that the conditions of Lemma 3.1 are satisfied with $\alpha = 2$. We may then work with X_n or $\hat{X}_n$ for investigating weak convergence, whichever is more convenient, and therefore we make the following *notational convention*: both (1) and (2) shall be represented by X_n throughout the proof, and where appropriate we will point out which one we are using.

Next we will prove three preparatory lemmas.

Lemma 3.2. *Let t and h be real numbers such that* $0 \leq t \leq h$, h_n *real numbers such that* $h_n \to h$ *as* $n \to \infty$, *and* T_n *random variables that are independent of all the random elements which determine* X_n, *such that* $T_n \leq h_n$ *a.s. and* $T_n \overset{a.s.}{\to} t$ *as* $n \to \infty$. *Let* $0 < \alpha \leq 2$. *Then if* $h = t$

$$\lim_{n \to \infty} E \| X_n(h_n) - X_n(T_n^+) \|^\alpha = 0,$$

and there is a constant K such that if $h > t$ *there is an integer* n_0 *depending only on* $h - t$ *such that, for all* $n \geq n_0$,

$$E \| X_n(h_n) - X_n(T_n^+) \|^\alpha \leq K(h - t)^{\alpha/2},$$

and if h_n *and* t_n *are real numbers such that* $0 \leq t_n < h_n$ *and* $(h_n - t_n)^{-1} = O(h_n)$ *as* $n \to \infty$ *there is an integer* n_0 *such that, for all* $n \geq n_0$,

$$E \| X_n(h_n) - X_n(t_n) \|^\alpha \leq K(h_n - t_n)^{\alpha/2}.$$

Here X_n *is* (1) *or* (2).

Proof. X_n is (2). By straightforward computation one obtains

$$E \| X_n(h_n) - X_n(T_n^+) \|^2 = c_n^2 \sum_{r=0}^{\infty} \sum_{s=r+1}^{\infty} E1_{[N_n(T_n^+) = r, N_n(h_n) = s]}$$

$$\cdot \left\{ \sum_{i=r+1}^{s} \tau_{n,i}^2 + 2 \sum_{i=r+1}^{s-1} \sum_{j=i+1}^{s} \tau_{n,i} \tau_{n,j} E(A_{n,i-1} \cdots A_{n,0} \xi_n, \right.$$

$$\left. (E A_{n,1})^{j-i} A_{n,i-1} \cdots A_{n,0} \xi_n) \right\}.$$

By Lemma 4.3 the indices of the $\tau_{n,i}$ can be changed to obtain, performing the summation on j,

$$E\|X_n(b_n) - X_n(T_n^+)\|^2 \;=\; c_n^2 \sum_{r=0}^{\infty} \sum_{s=r+1}^{\infty} E1_{[N_n(T_n^+) = r,\; N_n(b_n) = s]}$$

$$\cdot \left\{ (s-r)\tau_{n,s}^2 + 2\tau_{n,s-1}\tau_{n,s} \sum_{i=r+1}^{s-1} E(A_{n,i-1}\cdots A_{n,0}\xi_n, \right.$$

$$\left. \theta_{n,i,s} A_{n,i-1}\cdots A_{n,0}\xi_n) \right\},$$

where $\theta_{n,i,s} = (EA_{n,1} - (EA_{n,1})^{s-i+1})(I - EA_{n,1})^{-1}$, and hence, because $\|\theta_{n,i,s}\| \le 2\|(I - EA_{n,1})^{-1}\|$,

$$E\|X_n(b_n) - X_n(T_n^+)\|^2$$

$$\le c_n^2 E(N_n(b_n) - N_n(T_n^+))(\tau_{N_n(b_n)}^2 + 4\|(I - EA_{n,1})^{-1}\|\tau_{N_n(b_n)-1}\tau_{N_n(b_n)}).$$

Using the assumption on $\tau_{n,i}$, and the Hölder inequality with $1/p + 1/q = 1$ and p an integer large enough so that $E\tau_1^{2q+1} < \infty$, we obtain

$$E\|X_n(b_n) - X_n(T_n^+)\|^2$$

$$\le \frac{c_n^2}{b_n} \left(E\left(\frac{N(b_n b_n) - N(b_n T_n^+)}{b_n} \right)^p \right)^{1/p}$$

$$\cdot \{ (E\tau_{N(b_n b_n)}^{2q})^{1/q} + 4\|(I - EA_{n,1})^{-1}\|(E\tau_{N(b_n b_n)-1}^{2q} E\tau_{N(b_n b_n)}^{2q})^{1/2q} \},$$

whence the conclusions follow by Lemmas 4.4 and 4.6 and Liapunov's inequality, for X_n given by (2).

It is easy to show that there is a constant M such that, with the notation of (1) and (2),

$$(E\|X_n(b_n) - X_n(T_n^+)\|^2)^{1/2} \le (E\|\hat{X}_n(b_n) - \hat{X}_n(T_n^+)\|^2)^{1/2} + M/b_n^{1/2},$$

which yields the lemma for X_n given by (1).   □

In what follows it will sometimes be convenient to introduce the initial direction in the notation of X_n, and to view X_n as a family of random functions, one member of the family for each initial direction; thus $X_n(x)$ is the member

$$X_n(x, t) = \hat{X}_n(x, t) + c_n \left(t - \sum_{i=1}^{N_n(t)} \tau_{n,i} \right) A_{n,N_n(t)}\cdots A_{n,0}x,$$

where

$$\hat{X}_n(x, t) = c_n \sum_{i=1}^{N_n(t)} \tau_{n,i} A_{n,i-1}\cdots A_{n,0}x, \qquad t \ge 0.$$

We say that two random functions of our class have the same distribution when they are identically distributed for each initial direction.

In the next lemma we express the fact that after a turning time the process X_n is influenced by its history up to that time only through the direction after the turn (which depends on all the previous directions).

Lemma 3.3. *Given* $t > 0$ *there is a random function* Y_n *in our class, defined on the same probability space as* X_n, *which is distributed as* X_n *and independent of* $\{X_n(r), r \le t\}$, *and such that for* $s > t$

$$X_n(\xi_n, s) - X_n(\xi_n, t) = Y_n(\zeta_n, (s - T_{n,N_n(t)+1})^+) + \phi_n,$$

and also

$$X_n(\xi_n, s) - X_n(\xi_n, t) = Y_n(\zeta_n, s - t) + \theta_n,$$

where

$$\zeta_n = A_{n,N_n(t)+1} \cdots A_{n,0}\xi_n, \qquad T_{n,N_n(t)+1} = \sum_{i=1}^{N_n(t)+1} \tau_{n,i},$$

and $E\|\phi_n\|^\alpha = O(b_n^{-\alpha/2})$ *for* $0 < \alpha \le \rho - 1$ *and* $E\|\theta_n\|^2 \to 0$ *as* $n \to \infty$. *Here* X_n *is* (1) *or* (2).

Proof. X_n is (2). It is easily verified that the first expression holds with Y_n defined by

$$Y_n(x, b) = c_n \sum_{i=1}^{\hat{N}_n(b)} \hat{\tau}_{n,i} \hat{A}_{n,i-1} \cdots \hat{A}_{n,0} x, \qquad b \ge 0,$$

with $\hat{\tau}_{n,i} = \tau_{n,N_n(t)+1+i}$, $\hat{N}_n(b) = \max\{k : \sum_{i=1}^{k} \hat{\tau}_{n,i} \le b\}$, $\hat{A}_{n,i} = A_{n,N_n(t)+1+i}$, $i \ge 1$, $\hat{A}_{n,0} = I$, and

$$\phi_n = c_n \tau_{n,N_n(t)+1} A_{n,N_n(t)} \cdots A_{n,0}\xi_n 1_{[T_{n,N_n(t)+1} \le s]},$$

and then it is clear that X_n and Y_n have the indicated distribution properties. From $E\|\phi_n\|^\alpha \le (c_n/b_n)^\alpha E\tau_{N(b_n t)+1}^\alpha$ and Lemma 4.4 it follows that $E\|\phi_n\|^\alpha = O(b_n^{-\alpha/2})$ as $n \to \infty$. For the second expression we put

$$\theta_n = \phi_n - (Y_n(\zeta_n, s - t) - Y_n(\zeta_n, (s - T_{n,N_n(t)+1})^+));$$

then by Minkowski's inequality

$$(E\|\theta_n\|^2)^{1/2} \le \frac{c_n}{b_n}(E\tau_{N(b_n t)+1}^2)^{1/2} + (E\|Y_n(s - t) - Y_n((s - T_{n,N_n(t)+1})^+)\|^2)^{1/2};$$

the first term on the right converges to zero by Lemma 4.4, and so does the second one by Lemma 3.2 because $s - t \geq s - T_{n,N_n(t)+1} \overset{\text{a.s.}}{\to} s - t$ as $n \to \infty$.

For X_n given by (1) the proof is similar. □

Lemma 3.4. *If b_n and t_n are real numbers such that $0 \leq t_n < b_n$, $\liminf_{n \to \infty}(b_n - t_n) > 0$ and $\limsup_{n \to \infty}(b_n - t_n) < \infty$, there is an integer n_0, depending only on the sequence $\{b_n - t_n\}$, and constants K and c such that*

$$P\left[\sup_{t_n \leq s \leq b_n} \|X_n(s) - X_n(t_n)\| \geq \epsilon\right] \leq \frac{K(b_n - t_n)^\gamma}{\epsilon^{2\gamma}}$$

for all $n \geq n_0$ and all $\epsilon > c \limsup_{n \to \infty}(b_n - t_n)$, where $\gamma = (\rho - 1)/2$ (note that $1 < \gamma \leq 2$). Here X_n is (1).

Proof. Let $\delta_n = b_n - t_n$ and

$$r_{n,k} = t_n + \delta_n k/[b_n^{1/2}], \qquad k = 0, \cdots, [b_n^{1/2}];$$

then

$$P\left[\sup_{t_n \leq s \leq b_n} \|X_n(s) - X_n(t_n)\| \geq \epsilon\right]$$

(a) $$\leq P\left[\max_{1 \leq k \leq b_n^{1/2}} \|X_n(r_{n,k}) - X_n(r_{n,0})\| \geq \frac{\epsilon}{2}\right]$$

$$+ P\left[\sup_{t_n \leq s \leq b_n} \|X_n(s) - X_n(t_n)\| - \max_{1 \leq k \leq b_n^{1/2}} \|X_n(r_{n,k}) - X_n(r_{n,0})\| \geq \frac{\epsilon}{2}\right].$$

To study the first term on the right of (a) we set $S_{n,k} = \sum_{i=1}^k (X_n(r_{n,i}) - X_n(r_{n,i-1}))$, $k = 1, \cdots, b_n^{1/2}$, and $S_{n,0} = 0$; then (as in [1, p. 88])

$$P\left[\max_{1 \leq k \leq b_n^{1/2}} \|X_n(r_{n,k}) - X_n(r_{n,0})\| \geq \frac{\epsilon}{2}\right]$$

(b) $$\leq P\left[\max_{1 \leq k \leq b_n^{1/2}} \min\{\|S_{n,k}\|, \|S_{n,b_n^{1/2}} - S_{n,k}\|\} \geq \frac{\epsilon}{8}\right]$$

$$+ P\left[\max_{1 \leq k \leq b_n^{1/2}} \|X_n(r_{n,k}) - X_n(r_{n,k-1})\| \geq \frac{\epsilon}{8}\right].$$

We will now obtain an upper bound for the first term on the right of (b). Let i, j

and k be integers such that $0 \le i < j < k \le b_n^{1/2}$. By and in the notation of Lemma 3.3, applied to $S_{n,k} - S_{n,j}$, we have for $\lambda > 0$

(c)
$$P[\|S_{n,j} - S_{n,i}\| \ge \lambda, \|S_{n,k} - S_{n,j}\| \ge \lambda]$$
$$= P[\|X_n(r_{n,j}) - X_n(r_{n,i})\| \ge \lambda, \|X_n(r_{n,k}) - X_n(r_{n,j})\| \ge \lambda]$$
$$\le P[\|\phi_n\| \ge \lambda/2]$$
$$+ P[\|X_n(r_{n,j}) - X_n(r_{n,i})\| \ge \lambda, \|Y_n(\zeta_n, (r_{n,k} - T_{n,N_n(r_{n,j})+1})^+)\| \ge \lambda/2].$$

For the first term on the right of (c) we see by Chebyshev's inequality and Lemma 3.3 that, as $n \to \infty$, $P[\|\phi_n\| \ge \lambda/2] \le \lambda^{-2\gamma} O(b_n^{-\gamma})$, and since $\liminf_{n \to \infty} \delta_n > 0$, $k - i \ge 1$, and $b_n \to \infty$, then, for all sufficiently large n,

(d)
$$P[\|\phi_n\| \ge \lambda/2] \le \lambda^{-2\gamma} (\delta_n(k - i)/[b_n^{1/2}])^\gamma.$$

For the second term of (c) we have

$$P[\|X_n(r_{n,j}) - X_n(r_{n,i})\| \ge \lambda, \|Y_n(\zeta_n, (r_{n,k} - T_{n,N_n(r_{n,j})+1})^+)\| \ge \lambda/2]$$

$$= \int P[\|X_n(r_{n,j}) - X_n(r_{n,i})\| \ge \lambda, \|Y_n(x, r)\| \ge \lambda/2 |$$

$$\zeta_n = x, (r_{n,k} - T_{n,N_n(r_{n,j})+1})^+ = r]\mu_n(dx, dr),$$

where μ_n is the joint distribution of ζ_n and $(r_{n,k} - T_{n,N_n(r_{n,j})+1})^+$. We will study this integral separately over the sets $b_n^{-1} < r \le r_{n,k} - r_{n,j}$ and $0 \le r \le b_n^{-1}$. Under the condition $\zeta_n = x$ and $(r_{n,k} - T_{n,N_n(r_{n,j})+1})^+ = r$ the random vectors $X_n(r_{n,j}) - X_n(r_{n,i})$ and $Y_n(x, r)$ are independent, and $Y_n(x, r)$ is independent of the random elements appearing in the condition; using these facts together with Chebyshev's inequality and the last part of Lemma 3.2 it follows that there are constants K_0 and K_1 such that, for all sufficiently large n,

$$\int_{b_n^{-1} < r \le r_{n,k} - r_{n,j}} \le \frac{1}{\lambda^\gamma} K_0 (r_{n,j} - r_{n,i})^{\gamma/2} \frac{2^\gamma}{\lambda^\gamma} K_0 (r_{n,k} - r_{n,j})^{\gamma/2}$$

$$\le \frac{1}{\lambda^{2\gamma}} \left(K_1 \frac{\delta_n(k-i)}{[b_n^{1/2}]} \right)^\gamma.$$

The integrand over $0 \le r \le b_n^{-1}$ is bounded above by $1_{[c_n/b_n \ge \lambda/2]}$ because $\|Y_n(r)\| \le c_n r$, and hence $\int_{0 \le r \le b_n^{-1}} \le 1_{[c_n/b_n \ge \lambda/2]}$, which equals zero for each $\lambda > 0$ and all sufficiently large n because $c_n/b_n \to 0$ as $n \to \infty$, so that, for each $\lambda > 0$ and all sufficiently large n,

$$\int_{0 \le r \le b_n^{-1}} \le \frac{1}{\lambda^{2\gamma}} \left(\frac{\delta_n(k-i)}{[b_n^{1/2}]} \right)^\gamma .$$

In conclusion, there is a constant K_2 such that, for each $\lambda > 0$ and all sufficiently large n,

$$P[\|X_n(r_{n,j}) - X_n(r_{n,i})\| \ge \lambda, \|Y_n(\zeta_n, (r_{n,k} - T_{n,N_n(r_{n,j})+1})^+)\| \ge \lambda/2]$$

(e)

$$\le \lambda^{-2\gamma} (K_2 \delta_n(k-i)/[b_n^{1/2}])^\gamma .$$

From (c), (d) and (e) we see that there is a constant K_3 such that, for each $\lambda > 0$ and all sufficiently large n,

$$P[\|S_{n,j} - S_{n,i}\| \ge \lambda, \|S_{n,k} - S_{n,j}\| \ge \lambda] \le \lambda^{-2\gamma} (K_3 \delta_n(k-i)/[b_n^{1/2}])^\gamma ,$$

whence follows, by Theorem 12.1 of [1], that there is a constant K_4 such that, for each $\epsilon > 0$ and all sufficiently large n,

(f) $$P\left[\max_{1 \le k \le b_n^{1/2}} \min\{\|S_{n,k}\|, \|S_{n,b_n^{1/2}} - S_{n,k}\|\} \ge \frac{\epsilon}{8} \right] \le K_4 \frac{\delta_n^\gamma}{\epsilon^{2\gamma}} .$$

We can treat simultaneously the second term on the right of (a) and the second term on the right of (b) because they have a common bound; indeed, the following inequalities are easily verified.

$$P\left[\sup_{t_n \le s \le b_n} \|X_n(s) - X_n(t_n)\| - \max_{1 \le k \le b_n^{1/2}} \|X_n(r_{n,k}) - X_n(r_{n,0})\| \ge \lambda \right]$$

$$\le P\left[\max_{1 \le k \le b_n^{1/2}} \left(\sup_{r_{n,k-1} \le s \le r_{n,k}} \|X_n(s) - X_n(t_n)\| - \|X_n(r_{n,k}) - X_n(t_n)\| \right) \ge \lambda \right]$$

(g)

$$\le P\left[\max_{1 \le k \le b_n^{1/2}} \sup_{r_{n,k-1} \le s \le r_{n,k}} \|X_n(r_{n,k}) - X_n(s)\| \ge \lambda \right]$$

$$\le 1_{[c_n \delta_n / [b_n^{1/2}] \ge \lambda]} ,$$

and

(h) $$P\left[\max_{1 \le k \le b_n^{1/2}} \|X_n(r_{n,k}) - X_n(r_{n,k-1})\| \ge \lambda \right] \le 1_{[c_n \delta_n / [b_n^{1/2}] \ge \lambda]} ;$$

and clearly if $\lim \sup_{n \to \infty} \delta_n < \epsilon/8$ then, for all sufficiently large n,

(i) $$1_{[c_n \delta_n / [b_n^{1/2}] \ge \epsilon/8]} = 0.$$

(a), (b), (f), (g), (h) and (i) establish the lemma.   □

3.2. *Weak convergence.*

Lemma 3.5. $\{X_n(t),\ 0 \le t \le T;\ n = 1,2,\cdots\}$ *is tight for each finite positive* T. *Here* X_n *is* (1).

Proof. It is enough to show that for each $\epsilon > 0$ and $\eta > 0$ there is a δ, $0 < \delta < 1$, and an integer n_0 such that

$$P\left[\sup_{t \le s \le t+\delta} \|X_n(s) - X_n(t)\| \ge \epsilon\right] \le \delta\eta$$

for all $n \ge n_0$ (n_0 independent of t) and all t, $0 \le t \le T - \delta$ (see [1, p. 58]). In Lemma 3.4 let $t_n = t$, $b_n = t + \delta$ and choose δ so that $c\delta < \epsilon$ and $K\delta^{\gamma-1}/\epsilon^{2\gamma} \le \eta$.   □

In the next two lemmas we will prove that the finite-dimensional distributions of the sequence $\{X_n\}$ converge weakly to the desired limiting distributions. By the finite-dimensional distributions of a random element X of $D[0,\infty)^d$ we mean all the k-dimensional distributions of X, for $k = 1,2,\cdots$, the k-dimensional distributions of X being the (kd-dimensional) distributions of the random vectors $(X(t_1),\cdots, X(t_k))$ for all $t_1 < t_2 < \cdots < t_k$.

Lemma 3.6. *Let* X *be a random element of* $D[0,\infty)^d$ *with stationary and independent increments. If the one-dimensional distributions of* $X_n(\xi_n)$ *converge weakly to those of* X, *independently of* $\{\xi_n\}$, *as* $n \to \infty$, *then all the finite-dimensional distributions of* X_n *converge weakly to those of* X *as* $n \to \infty$. *Here* X_n *is* (1) *or* (2).

Proof. The proof is by induction on the dimension of the finite-dimensional distributions, the initial step of the induction being the hypothesis. Let us assume the truth of the result for the k-dimensional distributions and consider arbitrary times $t_1 < t_2 < \cdots < t_k = t < t_{k+1} = s$. The convergence of the $(k+1)$-dimensional distributions associated with these times will obtain if we show that for an arbitrary function $f: R^{(k+1)d} \to R$, continuous with compact support, it holds that

$$Ef(X_n(t_1),\cdots, X_n(t_k), X_n(t_{k+1}) - X_n(t_k))$$
$$\to Ef(X(t_1),\cdots, X(t_k), X(t_{k+1}) - X(t_k)) \quad \text{as } n \to \infty,$$

or, since X has stationary independent increments and $X(0) \overset{\text{a.s.}}{=} 0$,

$$\int E[f(x_1,\cdots, x_k, X_n(s) - X_n(t))| X_n(t_i) = x_i,\ i = 1,\cdots, k]\mu_n(dx_1,\cdots, dx_k)$$
$$\to \int Ef(x_1,\cdots, x_k, X(s-t))\mu(dx_1,\cdots, dx_k) \quad \text{as } n \to \infty,$$

where μ_n is the distribution of $(X_n(t_1),\cdots, X_n(t_k))$ and μ is that of $(X(t_1),\cdots, X(t_k))$. By and in the notation of Lemma 3.3 we have

$$X_n(\xi_n, s) - X_n(\xi_n, t) = \theta_n + Y_n(\zeta_n, s - t),$$

and hence, since f is uniformly continuous, given arbitrary $\epsilon > 0$ there is a $\delta > 0$ such that

$$\left| \int E[f(x_1, \cdots, x_k, X_n(s) - X_n(t))| X_n(t_i) = x_i, \ i = 1, \cdots, k]\mu_n(dx_1, \cdots, dx_k) \right.$$

$$\left. - \int E[f(x_1, \cdots, x_k, Y_n(s - t))| X_n(t_i) = x_i, \ i = 1, \cdots, k]\mu_n(dx_1, \cdots, dx_k) \right|$$

$$\leq \epsilon + 2 \sup_{z \in R^{(k+1)d}} |f(z)| P[\|\theta_n\| \geq \delta],$$

and therefore, since we know by Lemma 3.3 that $\|\theta_n\| \xrightarrow{P} 0$ as $n \rightarrow \infty$, it suffices to show that

$$E[f(x_1, \cdots, x_k, Y_n(s - t))| X_n(t_i) = x_i, \ i = 1, \cdots, k]$$

$$\rightarrow Ef(x_1, \cdots, x_k, X(s - t)) \quad \text{as } n \rightarrow \infty$$

uniformly for $(x_1, \cdots, x_k)$ in compact sets of R^{kd} (it is easy to see that $Ef(x_1, \cdots, x_k, X(s - t))$ is continuous on R^{kd}), or equivalently that

$$E[f(x_{n,1}, \cdots, x_{n,k}, Y_n(s - t))| X_n(t_i) = x_{n,i}, \ i = 1, \cdots, k]$$

$$\rightarrow Ef(x_1, \cdots, x_k, X(s - t)) \quad \text{as } n \rightarrow \infty$$

when $(x_{n,1}, \cdots, x_{n,k}) \rightarrow (x_1, \cdots, x_k)$ as $n \rightarrow \infty$. Let us demonstrate that this is true. Let

$$g_n(\cdot) = f(x_{n,1}, \cdots, x_{n,k}, \cdot) \quad \text{and} \quad g(\cdot) = f(x_1, \cdots, x_k, \cdot);$$

hence $g_n(x_n) \rightarrow g(x)$ when $x_n \rightarrow x$ as $n \rightarrow \infty$, and since the condition $X_n(t_i) = x_{n,i}$, $i = 1, \cdots, k$, affects $Y_n(\zeta_n, s - t)$ only through ζ_n, and by hypothesis $Y_n(\zeta_n, s - t)$ converges weakly to $X(s - t)$ independently of $\{\zeta_n\}$, it follows that subject to the condition $X_n(t_i) = x_{n,i}$, $i = 1, \cdots, k$, for each n,

$$g_n(Y_n(s - t)) \xrightarrow{\mathcal{D}} g(X(s - t)) \quad \text{as } n \rightarrow \infty;$$

finally, since the g_n are uniformly bounded we have that

$$E[g_n(Y_n(s - t))| X_n(t_i) = x_{n,i}, \ i = 1, \cdots, k] \rightarrow Eg(X(s - t)) \quad \text{as } n \rightarrow \infty,$$

the desired conclusion. $\square$

Lemma 3.7. *The one-dimensional distributions of* $X_n(\xi_n)$ *converge weakly, independently of* $\{\xi_n\}$, *to those of the Gaussian random function* X *with stationary and independent increments, mean 0 and covariance matrix function* $EX(t)X(t)' = tC/dE\tau_1$, *where*

$$C = \operatorname{Var} r_1 I + (E r_1)^2 \left\{ \lim_{n \to \infty} \left[(I - EA_{n,1})^{-1} + (I - EA'_{n,1})^{-1} \right] - I \right\} \quad as\ n \to \infty.$$

Here X_n *is* (1) *or* (2).

Proof. X_n is (1). We note first that C can be written as

$$C = E r_1^2 I + (E r_1)^2 \lim_{n \to \infty} \left[EA_{n,1}(I - EA_{n,1})^{-1} + EA'_{n,1}(I - EA'_{n,1})^{-1} \right].$$

The conclusion follows from a theorem of B. Rosén [6], and we need only to verify that its conditions are satisfied.

Let us fix $t > 0$ and denote $s_i = t\,i/n,\ i = 0,1,\cdots, n,$

$$Y_{n,i} = X_n(s_i) - X_n(s_{i-1}), \quad i = 1,\cdots, n,$$

and

$$S_n(\alpha) = X_n(t[\alpha n]/n), \quad 0 \le \alpha \le 1.$$

We have

$$S_n(\alpha) = \sum_{i=1}^{\alpha n} Y_{n,i}, \quad 0 \le \alpha \le 1,$$

and we wish to show that $S_n(1) \xrightarrow{\ \mathfrak{D}\ } X(t)$ as $n \to \infty$. In our case sufficient conditions are

(i) there is a positive constant K such that

$$\limsup_{n \to \infty} E \left\| \sum_{i=\beta n+1}^{\alpha n} Y_{n,i} \right\|^2 \le K(\alpha - \beta) \quad \text{for } 0 \le \beta < \alpha \le 1,$$

(ii)

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\delta} E 1_{\left[\left\| \sum_{i=\alpha n+1}^{(\alpha+\delta)n} Y_{n,i} \right\| \ge \epsilon \right]} \left\| \sum_{i=\alpha n+1}^{(\alpha+\delta)n} Y_{n,i} \right\|^2 = 0$$

$$\text{for } 0 \le \alpha < 1 \text{ and every } \epsilon > 0$$

(this implies condition (C4) of [6]),

(iii) $$\limsup_{n \to \infty} E \left\| E \left[\sum_{i=\alpha n+1}^{(\alpha+\delta)n} Y_{n,i} \middle| S_n(\alpha) \right] \right\| = 0 \quad \text{for } 0 \le \alpha < \alpha + \delta \le 1,$$

and

(iv) $$\limsup_{n \to \infty} E \left\| E \left[\sum_{i=\alpha n+1}^{(\alpha+\delta)n} Y_{n,i} \left(\sum_{i=\alpha n+1}^{(\alpha+\delta)n} Y_{n,i} \right)' \middle| S_n(\alpha) \right] - \delta \frac{t}{d E r_1} C \right\| = 0$$

$$\text{for } 0 \le \alpha < \alpha + \delta \le 1.$$

Let us see that these conditions are satisfied.

Condition (i). By Lemma 3.2,

$$\limsup_{n \to \infty} E \left\| \sum_{i=\beta n+1}^{\alpha n} Y_{n,i} \right\|^2 = \limsup_{n \to \infty} E \left\| X_n \left(\frac{t[\alpha n]}{n} \right) - X_n \left(\frac{t[\beta n]}{n} \right) \right\|^2 \le Mt(\alpha - \beta),$$

where M is a constant.

Condition (ii). We have the identity [1, p. 223],

$$E \frac{1}{[\| \sum_{i=\alpha n+1}^{(\alpha+\delta)n} Y_{n,i} \| \ge \epsilon]} \left\| \sum_{i=\alpha n+1}^{(\alpha+\delta)n} Y_{n,i} \right\|^2$$

$$= \epsilon^2 P \left[\left\| X_n \left(\frac{t[(\alpha+\delta)n]}{n} \right) - X_n \left(\frac{t[\alpha n]}{n} \right) \right\| \ge \epsilon \right]$$

$$+ \int_{\epsilon^2}^{\infty} P \left[\left\| X_n \left(\frac{t[(\alpha+\delta)n]}{n} \right) - X_n \left(\frac{t[\alpha n]}{n} \right) \right\| \ge s^{1/2} \right] ds,$$

whence the result follows by using Lemma 3.4.

We will verify Conditions (iii) and (iv) for X_n given by (2). It is easy to see that these conditions will then hold for X_n given by (1) also.

Condition (iii). We will omit the procedure since it is similar, but much simpler, to that for Condition (iv), which we will treat in some detail. The following obtains

$$E \left\| E \left[\sum_{i=\alpha n+1}^{(\alpha+\delta)n} Y_{n,i} \middle| S_n(\alpha) \right] \right\| \le \sqrt{d} \frac{c_n}{b_n} (E \tau_{N(b_n t_n)+1} + 2 \| (I - EA_{n,1})^{-1} \| E \tau_{N(b_n b_n)}),$$

where $t_n = t[\alpha n]/n$ and $b_n = t[(\alpha + \delta)n]/n$. The condition follows from the hypothesis of the theorem and Lemma 4.4.

Condition (iv). Let

$$\mathcal{G}_{n,i} = \mathcal{F}\{\xi_n, A_{n,0}, \cdots, A_{n,i-1}, \tau_{n,1}, \cdots, \tau_{n,i+1}\}, \quad i \ge 1,$$

$$\mathcal{G}_{n,0} = \mathcal{F}\{\tau_{n,1}\},$$

$$\mathcal{F}_{n,t_n} = \{A \in \mathcal{F}_n : A \cap [N_n(t_n) = i] \in \mathcal{G}_{n,i} \text{ for each } i \ge 0\},$$

$t_n = t[\alpha n]/n$ and $b_n = t[(\alpha + \delta)n]/n$. Then

$$\sum_{i=\alpha n+1}^{(\alpha+\delta)n} Y_{n,i} = X_n(b_n) - X_n(t_n), \quad S_n(\alpha) = X_n(t_n)$$

and $\mathcal{F}\{X_n(t_n)\} \subseteq \mathcal{F}_{n,t_n}$, so we have

$$E\left[\sum_{i=an+1}^{(a+\delta)n} Y_{n,i}\left(\sum_{i=an+1}^{(a+\delta)n} Y_{n,i}\right)'\bigg| S_n(\alpha)\right]$$

$$= c_n^2 E\left\{E\left[\sum_{i=N_n(t_n)+1}^{N_n(b_n)}\sum_{j=N_n(t_n)+1}^{N_n(b_n)} \tau_{n,i}\tau_{n,j}\right.\right.$$

$$\left.\left. \cdot (A_{n,i-1}\cdots A_{n,0}\xi_n)(A_{n,j-1}\cdots A_{n,0}\xi_n)'\big|\mathcal{F}_{n,t_n}\right]\bigg| X_n(t_n)\right\},$$

and

$$E\left\|E\left[\sum_{i=an+1}^{(a+\delta)n} Y_{n,i}\left(\sum_{i=an+1}^{(a+\delta)n} Y_{n,i}\right)'\bigg| S_n(\alpha)\right] - \delta\frac{t}{dE\tau_1} C\right\|$$

$$\leq \sqrt{d}\, E\left\|c_n^2 E\left[\sum_{i=N_n(t_n)+1}^{N_n(b_n)}\sum_{j=N_n(t_n)+1}^{N_n(b_n)} \tau_{n,i}\tau_{n,j}\right.\right.$$

$$\left.\left. \cdot (A_{n,i-1}\cdots A_{n,0}\xi_n)(A_{n,j-1}\cdots A_{n,0}\xi_n)'\big|\mathcal{F}_{n,t_n}\right] - \delta\frac{t}{dE\tau_1} C\right\|.$$

We will show that the last expression converges to zero. Denote

$$Y_n = E\left[\sum_{i=N_n(t_n)+1}^{N_n(b_n)}\sum_{j=N_n(t_n)+1}^{N_n(b_n)} \tau_{n,i}\tau_{n,j}\right.$$

$$\left. \cdot (A_{n,i-1}\cdots A_{n,0}\xi_n)(A_{n,j-1}\cdots A_{n,0}\xi_n)'\big|\mathcal{F}_{n,t_n}\right]$$

and $f_n(r,s) = 1_{[N_n(t_n)=r,\ N_n(b_n)=s]}$. Using Lemmas 4.2 and 4.3 we can write

$$Y_n \overset{a.s.}{=} \sum_{r=0}^{\infty}\sum_{s=r+1}^{\infty}\left\{E[f_n(r,s)\tau_{n,r+1}^2(A_{n,r}\cdots A_{n,0}\xi_n)(A_{n,r}\cdots A_{n,0}\xi_n)'|\mathcal{G}_{n,r}]\right.$$

$$+ \sum_{i=r+2}^{s} E[f_n(r,s)\tau_{n,r+1}\tau_{n,s}((A_{n,i-1}\cdots A_{n,0}\xi_n)(A_{n,r}\cdots A_{n,0}\xi_n)'$$

$$+ (A_{n,r}\cdots A_{n,0}\xi_n)(A_{n,i-1}\cdots A_{n,0}\xi_n)')|\mathcal{G}_{n,r}]$$

$$+ \sum_{i=r+2}^{s} E[f_n(r,s)\tau_{n,s}^2(A_{n,i-1}\cdots A_{n,0}\xi_n)(A_{n,i-1}\cdots A_{n,0}\xi_n)'|\mathcal{G}_{n,r}]$$

$$+ \sum_{i=r+2}^{s-1}\sum_{j=i+1}^{s} E[f_n(r,s)\tau_{n,s-1}\tau_{n,s}((A_{n,i-1}\cdots A_{n,0}\xi_n)(A_{n,j-1}\cdots A_{n,0}\xi_n)'$$

$$\left. + (A_{n,j-1}\cdots A_{n,0}\xi_n)(A_{n,i-1}\cdots A_{n,0}\xi_n)')|\mathcal{G}_{n,r}]\right\}.$$

Let us introduce random matrices $\{B_{n,1}, B_{n,2}, \cdots\}$ that are distributed as $\{A_{n,1}, A_{n,2}, \cdots\}$ and are independent of everything determining X_n, and let $\zeta_{n,r} = A_{n,r-1} \cdots A_{n,0}\xi_n$ ($\zeta_{n,r}$ is measurable with respect to $\mathcal{G}_{n,r}$). Then, using the independence assumptions, performing the summation on j and applying Lemma 4.2 again we obtain

$$Y_n \overset{a.s.}{=} E[(N_n(b_n) - N_n(t_n))\tau^2_{n,N_n(b_n)}U_n$$

$$+ (N_n(b_n) - N_n(t_n))\tau_{n,N_n(b_n)-1}\tau_{n,N_n(b_n)}(V_n\theta'_n + \theta_n V_n)| \mathcal{F}_{n,t_n}] + R_n,$$

where $\theta_n = EA_{n,1}(I - EA_{n,1})^{-1}$,

$$U_n = \frac{1}{N_n(b_n)-N_n(t_n)} \sum_{j=2}^{N_n(b_n)-N_n(t_n)} E[(B_{n,j}\cdots B_{n,1}\eta_n)(B_{n,j}\cdots B_{n,1}\eta_n)' | \eta_n],$$

$$V_n = \frac{1}{N_n(b_n)-N_n(t_n)} \sum_{j=2}^{N_n(b_n)-N_n(t_n)-1} E[(B_{n,j}\cdots B_{n,1}\eta_n)(B_{n,j}\cdots B_{n,1}\eta_n)'|\eta_n]$$

with $\eta_n = \zeta_{n,N_n(t_n)}$, and

$$R_n = E\left\{ \tau^2_{n,N_n(t_n)+1} E[(B_{n,1}\eta_n)(B_{n,1}\eta_n)'|\eta_n] \right.$$

$$+ \tau_{n,N_n(t_n)+1}\tau_{n,N_n(b_n)}((EA_{n,1} - (EA_{n,1})^{N_n(b_n)-N_n(t_n)})(I - EA_{n,1})^{-1}E[(B_{n,1}\eta_n)(B_{n,1}\eta_n)'|\eta_n]$$

$$+ E[(B_{n,1}\eta_n)(B_{n,1}\eta_n)'|\eta_n](EA'_{n,1} - (EA'_{n,1})^{N_n(b_n)-N_n(t_n)})(I - EA_{n,1})^{-1})$$

$$- \tau_{n,N_n(b_n)-1}\tau_{n,N_n(b_n)}$$

$$\sum_{k=2}^{N_n(b_n)-N_n(t_n)-1} (E[(B_{n,k}\cdots B_{n,1}\eta_n)(B_{n,k}\cdots B_{n,1}\eta_n)'|\eta_n](EA'_{n,1})^{N_n(b_n)-N_n(t_n)-k+1}(I - EA'_{n,1})^{-1}$$

$$\left. + (EA_{n,1})^{N_n(b_n)-N_n(t_n)-k+1}(I - EA_{n,1})^{-1} E[(B_{n,k}\cdots B_{n,1}\eta_n)(B_{n,k}\cdots B_{n,1}\eta_n)'|\eta_n])|\mathcal{F}_{n,t_n}\right\}.$$

Using the assumption on $r_{n,i}$ we have

$$c_n^2 Y_n - \delta \frac{t}{dE\tau_1} \left(E\tau_1^2 I + (E\tau_1)^2 \lim_n (\theta_n + \theta_n') \right)$$

$$= \frac{c_n^2}{b_n} E \left[\left(\frac{N(b_n b_n) - N(b_n t_n)}{b_n} - \frac{\delta t}{E\tau_1} \right) \right.$$

$$\left. \cdot \{\tau_{N(b_n b_n)}^2 U_n + \tau_{N(b_n b_n) - 1} \tau_{N(b_n b_n)} (V_n \theta_n' + \theta_n V_n)\} \Big| \mathcal{F}_{n, t_n} \right]$$

$$+ \frac{c_n^2}{b_n} \frac{\delta t}{E\tau_1} \{E[\tau_{N(b_n b_n)}^2 | \mathcal{F}_{n, t_n}] - E\tau_1^2$$

$$+ (\theta_n' + \theta_n)(E[\tau_{N(b_n b_n) - 1} \tau_{N(b_n b_n)} | \mathcal{F}_{n, t_n}] - (E\tau_1)^2)\} \frac{I}{d}$$

$$+ \frac{c_n^2}{b_n} \frac{\delta t}{E\tau_1} E \left[\tau_{N(b_n b_n)}^2 \left(U_n - \frac{I}{d} \right) \right.$$

$$\left. + \tau_{N(b_n b_n) - 1} \tau_{N(b_n b_n)} \left\{ \left(V_n - \frac{I}{d} \right) \theta_n' + \theta_n \left(V_n - \frac{I}{d} \right) \right\} \Big| \mathcal{F}_{n, t_n} \right]$$

$$+ \frac{c_n^2}{b_n} \frac{\delta t}{d} E\tau_1 \left\{ \theta_n' + \theta_n - \lim_n (\theta_n' + \theta_n) \right\}$$

$$+ \left(\frac{c_n^2}{b_n} - 1 \right) \frac{\delta t}{dE\tau_1} \{E\tau_1^2 I + (E\tau_1)^2 \lim_n (\theta_n' + \theta_n)\} + c_n^2 R_n;$$

hence

$$E \left\| c_n^2 Y_n - \delta \frac{t}{dE\tau_1} C \right\| \leq \sum_{i=1}^5 I_n^{(i)} + c_n^2 E \| R_n \|,$$

where (by using the basic inequality and the Hölder inequality with $1/p + 1/q = $ and q an integer large enough so that $E\tau_1^{2p+1} < \infty$, and the fact that $\| U_n \|$, $\| V_n \|$ and $\| \theta_n \|$ are bounded uniformly in n, say by a common bound K)

$$I_n^{(1)} = \frac{c_n^2}{b_n} \left(E \left| \frac{N(b_n b_n) - N(b_n t_n)}{b_n} - \frac{\delta t}{E\tau_1} \right|^q \right)^{1/q}$$

$$\cdot \{(E\tau_{N(b_n b_n)}^{2p})^{1/p} K + (E\tau_{N(b_n b_n) - 1}^{2p} E\tau_{N(b_n b_n)}^{2p})^{1/2p} 2K^2\} \sqrt{d},$$

$$I_n^{(2)} = \frac{c_n^2}{b_n} \frac{\delta t}{E\tau_1} \{E |E[\tau_{N(b_n b_n)}^2 | \mathcal{F}_{n, t_n}] - E\tau_1^2|$$

$$+ 2KE |E[\tau_{N(b_n b_n) - 1} \tau_{N(b_n b_n)} | \mathcal{F}_{n, t_n}] - (E\tau_1)^2|\} / d,$$

$$I_n^{(3)} = \frac{c_n^2}{b_n} \frac{\delta t}{E\tau_1} \left\{ (E\tau_{N(b_n b_n)}^{2p})^{1/p} \left(E \left\| U_n - \frac{I}{d} \right\|^q \right)^{1/q} \right.$$

$$\left. + 2K(E\tau_{N(b_n b_n) - 1}^{2p} E\tau_{N(b_n b_n)}^{2p})^{1/2p} \left(E \left\| V_n - \frac{I}{d} \right\|^q \right)^{1/q} \right\} \sqrt{d},$$

$$I_n^{(4)} = \frac{c_n^2}{b_n} \frac{\delta t}{d} E\tau_1 \| \theta_n' + \theta_n - \lim_n (\theta_n' + \theta_n) \|, \quad \text{and}$$

$$I_n^{(5)} = \left| \frac{c_n^2}{b_n} - 1 \right| \frac{\delta t}{dE\tau_1} \{E\tau_1^2 + (E\tau_1)^2 2K\}.$$

By the hypothesis of the theorem $I_n^{(4)} \to 0$ and $I_n^{(5)} \to 0$. By Lemmas 4.4 and 4.6, $I_n^{(1)} \to 0$, since $b_n \to t(\alpha + \delta)$ and $t_n \to t\alpha$. By Lemma 4.7, $I_n^{(2)} \to 0$, because $E\tau_1^{2p+1} < \infty$ for some $p > 1$ (notice what $\mathcal{F}_{n,t_n}$ is with $\tau_{n,i} = \tau_i/b_n$). By Lemmas 4.4 and 5.3, $I_n^{(3)} \to 0$, as $N(b_n b_n) - N(b_n t_n) \overset{a.s.}{\to} \infty$ and $\|\eta_n\| = 1$. Now, $c_n^2 E\|R_n\| \le J_n^{(1)} + J_n^{(2)}$, where

$$J_n^{(1)} = (c_n/b_n)^2 \ L\{E\tau_{N(b_n t_n)+1}^2 + (E\tau_{N(b_n t_n)+1}^2 \ E\tau_{N(b_n b_n)}^2)^{1/2}\}$$

and, using the Hölder inequality as above,

$$J_n^{(2)} = \frac{c_n^2}{b_n} \ L \left(E\tau_{N(b_n b_n)-1}^{2p} \ E\tau_{N(b_n b_n)}^{2p} \right)^{1/2p} \left(E \left(\frac{N(b_n b_n)-N(b_n t_n)-1}{b_n} \right)^{2q} \right)^{1/2q}$$

$$\cdot \left\{ \left(E \left\| \frac{1}{N(b_n b_n)-N(b_n t_n)-1} \sum_{k=2}^{N(b_n b_n)-N(b_n t_n)-1} I_{n,k}(EA'_{n,1})^{N(b_n b_n)-N(b_n t_n)+1-k} \right\|^{2q} \right)^{1/2q} \right.$$

$$\left. + \left(E \left\| \frac{1}{N(b_n b_n)-N(b_n t_n)-1} \sum_{k=2}^{N(b_n b_n)-N(b_n t_n)-1} (EA_{n,1})^{N(b_n b_n)-N(b_n t_n)+1-k} I_{n,k} \right\|^{2q} \right)^{1/2q} \right\},$$

where $I_{n,k} = E[(B_{n,k} \cdots B_{n,1}\eta_n)(B_{n,k} \cdots B_{n,1}\eta_n)'|\eta_n]$ and L is a constant. By the hypothesis of the theorem and Lemmas 4.4, 4.6 and 5.4, $J_n^{(1)} \to 0$ and $J_n^{(2)} \to 0$ as $n \to \infty$.

The proof is complete. For X_n given by (2) it follows from Lemma 3.1. $\square$

Lemmas 3.1, 3.5, 3.6 and 3.7 prove that for every finite positive T the sequence of random functions $\{X_n(t), \ 0 \le t \le T; \ n = 1, 2, \cdots\}$ converges weakly on $D[0, T]^d$ to the Gaussian random function $X(t)$, $0 \le t \le T$, described in the invariance principle (see [3]). The works of Stone [7] and Whitt ([10], [11]) show that this implies the conclusion of the invariance principle.

3.3. *Completion of the proof.* Here we will verify some of the observations made after the invariance principle.

1. We have $C = \text{Var} \, \tau_1 I + (E\tau_1)^2 M$, where M is the matrix of the theorem in §5. Therefore C is singular (X is degenerate) if and only if τ_1 is a.s. constant and M is singular, and the statements about C follow from the properties of M at the end of §5.

2. This follows directly from the results at the end of §5.

3. The variance parameter may be written as

$$\frac{1}{2E\tau_1} \left(\text{Var} \, \tau_1 + (E\tau_1)^2 \lim_{n \to \infty} \text{Re} \frac{1 + Ee^{i\alpha_n}}{1 - Ee^{i\alpha_n}} \right)$$

Therefore, since the mapping $(1 + z)/(1 - z)$ of the complex plane carries the circumference $|z| = a$, $0 \le a \le 1$, onto the (generalized) circumference with center on the real axis and crossings of the real axis at the points $(1 - a)/(1 + a)$ and $(1 + a)/(1 - a)$, and if θ is a fixed angle

$$\mathrm{Re}(1 + Ee^{i(a_n + \theta)})/(1 - Ee^{i(a_n + \theta)}) = \mathrm{Re}(1 + e^{i\theta}Ee^{ia_n})/(1 - e^{i\theta}Ee^{ia_n}),$$

then by shifting a_n (mod the interval $[-\pi, \pi)$),
$(2E\tau_1)^{-1}(\mathrm{Var}\,\tau_1 + (E\tau_1)^2\mathrm{Re}(1 + Ee^{ia_n})/(1 - Ee^{ia_n}))$ can be made to take any value in the interval

$$\left[\frac{1}{2E\tau_1}(\mathrm{Var}\,\tau_1 + (E\tau_1)^2\beta_n), \frac{1}{2E\tau_1}\left(\mathrm{Var}\,\tau_1 + (E\tau_1)^2\frac{1}{\beta_n}\right)\right],$$

where $\beta_n = (1 - |Ee^{ia_n}|)/(1 + |Ee^{ia_n}|)$, which is positive since a_n is implicitly assumed not to be a.s. constant. The desired results follow from this.

4. **Renewal theoretic lemmas.** All the statements in this section refer to a sequence $\tau_1, \tau_2, \cdots$ of nonnegative, nonidentically zero, independent and identically distributed random variables, and to the random variable (the number of renewals up to t) $N(t) = \max\{k : \sum_{i=1}^{k} \tau_i \le t\}$, $t \ge 0$. We define $\tau_i = 0$ for $i \le 0$. The underlying probability space is denoted $(\Omega, \mathcal{F}, P)$. Some indications will be given of only those proofs that are less elementary; other proofs and further results appear in [4].

Lemma 4.1. *If* $E\tau_1^\alpha < \infty$ *for given* $\alpha > 0$, *then*

$$\frac{1}{t}\sup_{0 \le s \le t^\alpha}\left(s - \sum_{i=1}^{N(s)}\tau_i\right) \xrightarrow{P} 0 \quad as \ t \to \infty.$$

Let $\eta, \theta_0, \theta_1, \cdots$ be a sequence of random elements (of some topological space), with θ_0 constant, which is independent of the sequence $\tau_1, \tau_2, \cdots$, and consider the Borel fields $\mathcal{G}_n = \mathcal{F}\{\eta, \theta_0, \cdots, \theta_{n-1}, \tau_1, \cdots, \tau_{n+1}\}$ for each $n \ge 1$, and $\mathcal{G}_0 = \mathcal{F}\{\tau_1\}$, and the Borel field $\mathcal{F}_t$ generated by $\eta, \theta_0, \cdots, \theta_{n-1}, \tau_1, \cdots, \tau_{n+1}$ up to the random index $N(t)$, that is $\mathcal{F}_t = \{A \in \mathcal{F} : A \cap [N(t) = n] \in \mathcal{G}_n$ for each $n \ge 0\}$, for each $t > 0$. If $X(t)$ is a measurable function (with values in an arbitrary measurable space) of $\eta, \theta_0, \cdots, \theta_{N(t)-1}, \tau_1, \cdots, \tau_{N(t)+1}$, then $\mathcal{F}\{X(t)\} \subseteq \mathcal{F}_t$.

Lemma 4.2. *For any random vector X and for each* $n \ge 0$ *and* $t > 0$,

$$E[1_{[N(t)=n]}X|\mathcal{F}_t] \overset{a.s.}{=} E[1_{[N(t)=n]}X|\mathcal{G}_n].$$

Lemma 4.3. *For each* $n \ge 0$ *and* $k \ge 1$, *for any permutation* σ *of* $\{1, \cdots, k\}$ *and any integers* $i_1, \cdots, i_j$ *such that* $1 \le i_1 < \cdots < i_j \le k$, *and for any measurable vector X defined on* R^∞ *or* R^m *for appropriate m,*

$$E[1_{[N(t)=n+k]}X(\tau_1, \cdots, \tau_n, \tau_{n+i_1}, \cdots, \tau_{n+i_j}, \tau_{n+k+1}, \cdots)|\mathcal{G}_{n-1}]$$

$$\overset{a.s.}{=} E[1_{[N(t)=n+k]}X(\tau_1, \cdots, \tau_n, \tau_{n+\sigma(i_1)}, \cdots, \tau_{n+\sigma(i_j)}, \tau_{n+k+1}, \cdots)|\mathcal{G}_{n-1}],$$

and in particular (with $\mathcal{G}_{-1} = \{\phi, \Omega\}$)

$$\int_{[N(t)=k]} X(\tau_{i_1}, \cdots, \tau_{i_j}, \tau_{k+1}, \cdots)\,dP$$

$$= \int_{[N(t)=k]} X(\tau_{\sigma(i_1)}, \cdots, \tau_{\sigma(i_j)}, \tau_{k+1}, \cdots)\,dP.$$

Lemma 4.4. *For all* $\alpha > 0$, *if* $E\tau_1^{\alpha+1} < \infty$, *then, for every integer* $n \geq 0$,

$$\sup_t \; E\tau_{N(t)+1-n}^{\alpha} < \infty.$$

Lemma 4.5. *For every integer* $m \geq 1$ *there is a constant* K *such that, for all* $t \geq 0$,

$$EN(t)^m \leq K \max\{1, t^m\}.$$

Lemma 4.6. *Let* t *and* h *be real numbers such that* $0 \leq t \leq h$, T_n *and* H_n *non-negative random variables that are jointly independent of* $\{\tau_i, i \geq 1\}$ *and such that* $T_n \leq H_n$ *a.s.,* $T_n \overset{a.s.}{\rightarrow} t$ *and* $H_n \overset{a.s.}{\rightarrow} h$ *as* $n \rightarrow \infty$, *and for a positive integer* m *and some* $\delta > 0$, $\sup_n EH_n^{m+\delta} < \infty$, *and let* $b_1, b_2, \cdots$ *be a sequence of positive real numbers converging to infinity. Then*

$$\lim_{n \rightarrow \infty} \; E\left|\frac{N(b_n H_n) - N(b_n T_n)}{b_n} - \frac{h-t}{E\tau_1}\right|^m = 0,$$

and there is a constant K *such that if* $h - t > 0$ *there is an integer* n_0 *depending only on* $h - t$ *such that for all* $n \geq n_0$

$$E((N(b_n H_n) - N(b_n T_n))/b_n)^m \leq K(h - t)^m,$$

and if t_n *and* h_n *are real numbers such that* $0 \leq t_n < h_n$ *and* $(h_n - t_n)^{-1} = O(b_n)$ *as* $n \rightarrow \infty$ *there is an integer* n_0 *such that for all* $n \geq n_0$

$$E((N(b_n H_n) - N(b_n T_n))/b_n)^m \leq K(h_n - t_n)^m.$$

(The last assertion is valid for every positive integer m, *with* K *depending on* m.)

Proof. For $t > 0$ the first assertion is obtained from the following result. If $F(s)$, $s > 0$, are positive random variables that are independent of $N(s)$ for each s and such that $F(s)/s \overset{a.s.}{\rightarrow} 1$ as $s \;\;\; \infty$, and for a positive integer m and some $\delta > 0$, $\sup_{s \geq 1} E(F(s)/s)^{m+\delta} < \infty$, then

$$\lim_{s \rightarrow \infty} \; E\left|\frac{N(F(s))}{s} - \frac{1}{E\tau_1}\right|^m = 0.$$

Let us prove this. From the strong law of large numbers we have $N(F(s))/s \overset{\text{a. s.}}{\longrightarrow}$ $1/E\tau_1$ as $s \to \infty$, and hence it is enough to show that $\{(N(F(s))/s)^m, \ s \geq 1\}$ is uniformly integrable (see [2, Theorem 4.5.4]), which is true if $\sup_{s \geq 1} E(N(F(s))/s)^{m + \delta}$ $< \infty$, but this follows from the inequality

$$E(N(F(s))/s)^{m + \delta} \leq K(1/s^{m + \delta} + E(F(s)/s)^{m + \delta}),$$

which is obtained by using the independence hypothesis and Lemma 4.5 (K is the constant of Lemma 4.5). For $t = 0$ the proof is similar.

We now prove the second assertion. It can be shown by elementary methods that

$$E(N(b_n H_n) - N(b_n T_n))^m \leq E(1 + N(b_n(H_n - T_n)))^m.$$

Using this inequality together with Lemma 4.5 we obtain that there is a constant K_1 such that

$$E\left(\frac{N(b_n H_n) - N(b_n T_n)}{b_n} \right)^m \leq K_1 \sum_{k = 0}^{m} \frac{E(H_n - T_n)^k}{b_n^{m-k}}.$$

Now, since $E(H_n - T_n)^k \to (b - t)^k$ for $k = 0, \cdots, m$, and $b_n \to \infty$ as $n \to \infty$, there is an integer n_0, depending only on $b - t$, and a constant K_2 such that, for $k = 0, \cdots, m$ and all $n \geq n_0$, $E(H_n - T_n)^k / b_n^{m-k} \leq K_2(b - t)^m$. The conclusion then follows.

The last result has a similar proof. □

Lemma 4.7. *If given nonnegative constants* $c_1, \cdots, c_{m+1}$,

$$E\tau_1^{p(m+1) \max\{c_1, \cdots, c_{m+1}\} + 1} < \infty$$

for some $p > 1$, *and if* α_t *and* β_t, $t \geq 0$, *are real numbers such that* $\lim \inf_{t \to \infty} \alpha_t > 0$, $\lim \sup_{t \to \infty} \beta_t < \infty$ *and* $\lim \inf_{t \to \infty} (\beta_t - \alpha_t) > 0$, *then*

$$E\left| E\left[\tau_{N(\beta_t t) - m}^{c_1} \cdots \tau_{N(\beta_t t)}^{c_{m+1}} \Big| \mathcal{F}_{\alpha_t t} \right] - E\tau_1^{c_1} \cdots E\tau_1^{c_{m+1}} \right| \to 0$$

as $t \to \infty$ ($\mathcal{F}_{\alpha_t t}$ *is as defined after Lemma* 4.1).

Proof. It is simpler and more general to prove that

$$E\left| E\left[f(\tau_{N(\beta_t t) - m}, \cdots, \tau_{N(\beta_t t)}) \Big| \mathcal{F}_{\alpha_t t} \right] - Ef(\tau_1, \cdots, \tau_{m+1}) \right| \to 0$$

as $t \to \infty$ for a measurable function $f: R^{m+1} \to R$ such that $E|f(\tau_1, \cdots, \tau_{m+1})|^p$

$< \infty$ and $\sup_t E |f(\tau_{N(t)-m}, \cdots, \tau_{N(t)})|^p < \infty$, when $E\tau_1 < \infty$. The conclusion will then follow by using the generalized Hölder inequality and Lemma 4.4.

Let $K = 1/E\tau_1$. Choose ϵ so that $0 < \epsilon < K$. Then for t sufficiently large so that $(K - \epsilon)\beta_t t > m$ we have

$$E[f(\tau_{N(\beta_t t)-m}, \cdots, \tau_{N(\beta_t t)})|\mathcal{F}_{\alpha_t t}]$$

$$= T + E[1_{[|N(\beta_t t)/\beta_t t - K| \geq \epsilon] \cup [|N(\alpha_t t)/\alpha_t t - K| \geq \epsilon]}$$
$$\cdot f(\tau_{N(\beta_t t)-m}, \cdots, \tau_{N(\beta_t t)})|\mathcal{F}_{\alpha_t t}],$$

where

$$T = E[1_{[|N(\beta_t t)/\beta_t t - K| < \epsilon, |N(\alpha_t t)/\alpha_t t - K| < \epsilon]} f(\tau_{N(\beta_t t)-m}, \cdots, \tau_{N(\beta_t t)})|\mathcal{F}_{\alpha_t t}],$$

which using Lemma 4.2 can be written as

$$T \stackrel{\text{a.s.}}{=} \sum{}' E[1_{[N(\alpha_t t)=s, N(\beta_t t)=r]} f(\tau_{r-m}, \cdots, \tau_r)|\mathcal{G}_s],$$

where $\sum'$ indicates summation over $|r - K\beta_t t| < \epsilon \beta_t t$, $|s - K\alpha_t t| < \epsilon \alpha_t t$. Let ϵ be so that $(K + \epsilon)\alpha_t < (K - \epsilon)\beta_t$, with t so large that $(K - \epsilon)\beta_t t - m > (K + \epsilon)\alpha_t t + 1$; then in the above sum the indices of summation r and s satisfy $s + 1 < r - m$ in their ranges, and therefore by Lemmas 4.3 and 4.2 we obtain

$$T \stackrel{\text{a.s.}}{=} 1_{[|N(\alpha_t t)/\alpha_t t - K| \geq \epsilon]} Ef(\tau_1, \cdots, \tau_{m+1})$$

$$- E[1_{[|N(\alpha_t t)/\alpha_t t - K| < \epsilon, |N(\beta_t t)/\beta_t t - K| \geq \epsilon]} f(\tau_{N(\alpha_t t)+2}, \cdots, \tau_{N(\alpha_t t)+m+2})|\mathcal{F}_{\alpha_t t}].$$

Hence

$$E \left| E\left[f(\tau_{N(\beta_t t)-m}, \cdots, \tau_{N(\beta_t t)})\Big|\mathcal{F}_{\alpha_t t} \right] - Ef(\tau_1, \cdots, \tau_{m+1}) \right|$$

$$\leq E|f(\tau_1, \cdots, \tau_{m+1})| P[|N(\alpha_t t)/\alpha_t t - K| \geq \epsilon]$$

$$+ E1_{[|N(\beta_t t)/\beta_t t - K| \geq \epsilon]} \{|f(\tau_{N(\alpha_t t)+2}, \cdots, \tau_{N(\alpha_t t)+m+2})|$$
$$+ |f(\tau_{N(\beta_t t)-m}, \cdots, \tau_{N(\beta_t t)})|\}$$

$$+ E1_{[|N(\alpha_t t)/\alpha_t t - K| \geq \epsilon]} |f(\tau_{N(\beta_t t)-m}, \cdots, \tau_{N(\beta_t t)})|.$$

The proof is completed by using the Hölder inequality with appropriate parameter, and the facts that $(\tau_{N(\alpha_t t)+2}, \cdots, \tau_{N(\alpha_t t)+m+2}) \stackrel{\mathcal{D}}{=} (\tau_1, \cdots, \tau_{m+1})$ and $N(t)/t \stackrel{\text{a.s.}}{\to} K$ as $t \to \infty$, to show that each of the terms on the right of the last expression converges to zero as $t \to \infty$. $\square$

5. **Irreducible random orthogonal matrices. Covariance limit theorems.** Matrices are real, of order d, and vectors are d-dimensional column vectors ($d \geq 1$). An underlying probability space (Ω, $\mathcal{F}$, P) where random elements are defined is assumed to exist.

For $d \geq 2$ we define a random matrix A to be irreducible when $P[AV \subseteq V] < 1$ for every nontrivial subspace V of R^d; otherwise, i.e. when A has an a.s. invariant nontrivial subspace, we say that A is reducible; for $d = 1$ we define A to be irreducible if its single entry is not a.s. nonnegative. Some elementary properties of this definition are: when A is constant it implies the definition of irreducibility in the usual sense (for $d \geq 2$ if $A = (a_{ij})$ is the matrix of an operator relative to a basis of R^d then none of the nontrivial subspaces generated by basis vectors is invariant under the operator, or equivalently, there is no decomposition $\{M, N\}$ of $\{1, \cdots, d\}$ such that $a_{ij} = 0$ for all $i \in M$ and $j \in N$, and for $d = 1$ the single entry of A is nonzero) but not conversely; if A is irreducible and T is a random nonsingular matrix independent of A (in particular if T is constant) then $T^{-1}AT$ is irreducible; if A is nonsingular then A is irreducible if and only if A^{-1} is; and for $d \geq 3$ an a.s. constant orthogonal matrix is reducible. The verifications of these properties are straightforward and we will omit them. Less elementary is the following property:

Lemma 5.1. *If A is an irreducible random contractive matrix then the matrix $I - EA$ is nonsingular.*

Proof. The result is obvious for $d = 1$. For $d \geq 2$ if $I - EA$ is singular there is a nonzero vector x such that $EAx = x$, which implies that $Ax \overset{a.s.}{=} x$ because the point EAx is an average of points that lie in the convex set $\{z: \|z\| \leq \|x\|\}$ and hence can be x only if A fixes x with probability one (one can also see this by showing that if T is the projection onto the subspace spanned by x then $TAx \overset{a.s.}{=} x$); but then the subspace spanned by x is a.s. invariant under A. □

Lemma 5.2. *If A is an irreducible random orthogonal matrix and S is a random symmetric matrix satisfying the equation $E[ASA'|S] \overset{a.s.}{=} S$, then $S \overset{a.s.}{=} ((\operatorname{tr} S)/d)I$.*

Proof. The result is obvious for $d = 1$. For $d \geq 2$ it suffices to obtain the conclusion for each point $\omega \in \Omega$ for which the condition holds; thus we may assume that S is a constant matrix satisfying the equation $E(ASA') = S$. There is an orthogonal matrix Q and a diagonal matrix L such that $QSQ' = L$, and therefore L satisfies $E(BLB') = L$, where $B = QAQ'$ is an irreducible random orthogonal matrix. Denoting $L = (l_i\delta_{ij})$, where $(\delta_{ij}) = I$, and $B = (b_{ij})$, we see that the numbers l_i satisfy the equations $l_i\delta_{ij} = \Sigma_k\, l_k E(b_{ik}b_{jk})$ and in particular, for $j = i$, $l_i = \Sigma_k\, l_k Eb_{ik}^2$, $i = 1, \cdots, d$, where we may suppose $l_1 \leq l_2 \leq \cdots \leq l_d$. Assuming that $l_1 = \cdots = l_u < l_{u+1} \leq \cdots \leq l_d$ with $u < d$ we have $l_i\Sigma_{k=u+1}^d\, Eb_{ik}^2 = \Sigma_{k=u+1}^d l_k Eb_{ik}^2$,

$i = 1, \cdots, u$, because $\Sigma_k E b_{ik}^2 = 1$; this implies that $E b_{ik}^2 = 0$, $i = 1, \cdots, u$, $k = u + 1, \cdots, d$, and hence $b_{ik} \overset{a.s.}{=} 0$, $i = 1, \cdots, u$, $k = u + 1, \cdots, d$; but B is irreducible. Therefore $l_i = l$, $i = 1, \cdots, d$ for some l, and we conclude that this is the solution to $l_i \delta_{ij} = \Sigma_k l_k E(b_{ik} b_{jk})$. We have shown that $L = lI$ for some constant l, whence the conclusion follows. $\quad\square$

Lemma 5.3. *For each $n \geq 1$ let $A_{n,i}$, $i \geq 1$, be independent and identically distributed irreducible random orthogonal matrices, and let ξ_n, $n \geq 1$, be random vectors such that, for each n, $\{A_{n,i}, i \geq 1\}$ and ξ_n are independent. Assume that $\{A_{n,1}, n \geq 1\}$ has no reducible weak limit points, and that $\{\|\xi_n\|^2, n \geq 1\}$ is uniformly integrable and there is a random variable ν such that $\|\xi_n\|^2 \overset{\mathfrak{D}}{\to} \nu$ as $n \to \infty$. Then*

$$\frac{1}{m} \sum_{i=1}^m E[(A_{n,i} \cdots A_{n,1} \xi_n)(A_{n,i} \cdots A_{n,1} \xi_n)' | \xi_n] \overset{\mathfrak{D}}{\to} \frac{\nu}{d} I \quad as\ n,\ m \to \infty,$$

and if ν is constant and $\{\|\xi_n\|^{2p}, n \geq 1\}$ is uniformly integrable for given $p \geq 1$, then

$$E \left\| \frac{1}{m} \sum_{i=1}^m E[(A_{n,i} \cdots A_{n,1} \xi_n)(A_{n,i} \cdots A_{n,1} \xi_n)' | \xi_n] - \frac{\nu}{d} I \right\|^p \to 0 \quad as\ n,\ m \to \infty.$$

Proof. Let us denote

$$\phi(m, n) = \frac{1}{m} \sum_{i=1}^m E[(A_{n,i} \cdots A_{n,1} \xi_n)(A_{n,i} \cdots A_{n,1} \xi_n)' | \xi_n].$$

$\phi(m, n)$ exists because $\|(A_{n,i} \cdots A_{n,1} \xi_n)(A_{n,i} \cdots A_{n,1} \xi_n)'\| \leq \|\xi_n\|^2$, and has the properties

(1) $\phi(m, n)$ is a random symmetric matrix bounded a.s. by $\sqrt{d} \|\xi_n\|^2$, with $\operatorname{tr} \phi(m, n) = \|\xi_n\|^2$, and

(2) $\phi(m, n) - E[A_{n,1} \phi(m, n) A_{n,1}' | \phi(m, n)] \overset{\mathfrak{D}}{\to} 0$ as $m, n \to \infty$.

The verification of (1) is elementary. We will prove (2): By the independence of $A_{n,1}, \cdots, A_{n,i}, \xi_n$ we have (with $A_{n,0} = I$)

$$\phi(m, n) = \frac{1}{m} \sum_{i=1}^m E[A_{n,1} E[A_{n,i-1} \cdots A_{n,0} \xi_n \xi_n' A_{n,0}' \cdots A_{n,i-1}' | \xi_n] A_{n,1}' | \xi_n]$$

$$= E[A_{n,1} \phi(m, n) A_{n,1}' | \xi_n] + \psi(m, n),$$

where

$$\psi(m, n) = \frac{1}{m} E[(A_{n,1} \xi_n)(A_{n,1} \xi_n)' - (A_{n,m+1} \cdots A_{n,1} \xi_n)(A_{n,m+1} \cdots A_{n,1} \xi_n)' | \xi_n];$$

and·since $\|\psi(m, n)\| \leq (2\sqrt{d}/m)\|\xi_n\|^2$ a.s. and $\{\|\xi_n\|^2\}$ is uniformly integrable,

$E\|\psi(m, n)\| \to 0$ as $m, n \to \infty$, and therefore $E\|E[\psi(m, n)|\phi(m, n)]\| \to 0$ as m, $n \to \infty$, so that $E[\psi(m, n)|\phi(m, n)] \xrightarrow{\mathcal{D}} 0$ as $m, n \to \infty$; but since $\phi(m, n)$ is a measurable function of (the elements of) ξ_n,

$$E[\psi(m, n)|\phi(m, n)] \stackrel{a.s.}{=} \phi(m, n) - E[A_{n,1}\phi(m, n)A'_{n,1}|\phi(m, n)],$$

and the conclusion follows.

The first assertion of the lemma will be proved if we show that for every pair of sequences $\{r_n\}$ and $\{s_n\}$ of positive integers converging to infinity there is a subsequence $\{n_k\}$ of $\{n\}$ such that $\phi(r_{n_k}, s_{n_k}) \xrightarrow{\mathcal{D}} \nu I/d$ as $k \to \infty$. Both $\{\phi(m, n)\}$ and $\{A_{n,1}\}$ are relatively compact; hence for each pair of sequences $\{r_n\}$ and $\{s_n\}$ as above there is a subsequence $\{n_k\}$ of $\{n\}$ such that $\phi(r_{n_k}, s_{n_k}) \xrightarrow{\mathcal{D}} L$ and $A_{s_{n_k},1} \xrightarrow{\mathcal{D}} A$ as $k \to \infty$, where the random matrix L is symmetric with $\operatorname{tr} L \stackrel{\mathcal{D}}{=} \nu$, because of (1) and the fact that the trace is a continuous function, and the random orthogonal matrix A is irreducible by hypothesis. Since $A_{s_{n_k},1}$ is independent of $\phi(r_{n_k}, s_{n_k})$ and $\{A_{s_{n_k},1}\}$ is uniformly bounded, the elements of the matrix

$$\phi(r_{n_k}, s_{n_k}) - E[A_{s_{n_k},1}\phi(r_{n_k}, s_{n_k})A'_{s_{n_k},1}|\phi(r_{n_k}, s_{n_k})]$$

are linear combinations of the elements of $\phi(r_{n_k}, s_{n_k})$ with coefficients which converge as $k \to \infty$, and therefore converge jointly in distribution as $k \to \infty$ to the same linear combinations of the elements of L with the corresponding limits of the coefficients as coefficients; thus

$$\phi(r_{n_k}, s_{n_k}) - E[A_{s_{n_k},1}\phi(r_{n_k}, s_{n_k})A'_{s_{n_k},1}|\phi(r_{n_k}, s_{n_k})] \xrightarrow{\mathcal{D}} L - E[ALA'|L] \quad \text{as } k \to \infty.$$

From property (2) it follows that $E[ALA'|L] \stackrel{a.s.}{=} L$, so that, by Lemma 5.2, $L \stackrel{\mathcal{D}}{=} \nu I/d$, as desired.

The second assertion follows from the fact that $\phi(m, n) \xrightarrow{P} \nu I/d$ as $n, m \to \infty$ and $\{\|\phi(m, n)\|^p\}$ is uniformly integrable. $\square$

Lemma 5.4. *Let* $A_{n,i}$ *and* ξ_n *be as in Lemma 5.3 but with* $\|\xi_n\|^2$ *not necessarily convergent in distribution. Then*

$$\frac{1}{m} \sum_{i=1}^{m} (EA_{n,1})^{m+1-i} E[(A_{n,i} \cdots A_{n,1}\xi_n)(A_{n,i} \cdots A_{n,1}\xi_n)'|\xi_n] \xrightarrow{P} 0 \quad \text{as } n, m \to \infty,$$

and if $\{\|\xi_n\|^{2p}, n \geq 1\}$ *is uniformly integrable for given* $p \geq 1$, *then*

$$E\left\|\frac{1}{m} \sum_{i=1}^{m} (EA_{n,1})^{m+1-i} E[(A_{n,i} \cdots A_{n,1}\xi_n)(A_{n,i} \cdots A_{n,1}\xi_n)'|\xi_n]\right\|^p \to 0 \quad \text{as } n, m \to \infty.$$

Proof. Since

$$\frac{1}{m} \sum_{i=1}^{m} (EA_{n,1})^{m+1-i} E[(A_{n,i} \cdots A_{n,1} \xi_n)(A_{n,i} \cdots A_{n,1} \xi_n)' | \xi_n]$$

$$= E\left[(A_{n,m+1} \cdots A_{n,1} \xi_n)(A_{n,m+1} \cdots A_{n,1} \xi_n)' \frac{1}{m} \sum_{i=1}^{m} A_{n,m+1} \cdots A_{n,i+1} \Big| \xi_n \right],$$

and $\|(A_{n,m+1} \cdots A_{n,1} \xi_n)(A_{n,m+1} \cdots A_{n,1} \xi_n)'\| \le \|\xi_n\|^2$, it suffices to show that $(1/m) \sum_{i=1}^{m} A_{n,m+1} \cdots A_{n,i+1} \xrightarrow{P} 0$ as $n, m \to \infty$, or equivalently, that $(1/m) \sum_{i=1}^{m} A_{n,1} \cdots A_{n,i} \xrightarrow{\mathfrak{D}} 0$ as $n, m \to \infty$. Since $\{(1/m) \sum_{i=1}^{m} A_{n,1} \cdots A_{n,i}\}$ is rela- · tively compact, for each pair of sequences $\{r_n\}$ and $\{s_n\}$ of positive integers converging to infinity there is a subsequence $\{n_k\}$ of $\{n\}$ and a random matrix M such that

$$\frac{1}{r_{n_k}} \sum_{i=1}^{r_{n_k}} A_{s_{n_k},1} \cdots A_{s_{n_k},i} \xrightarrow{\mathfrak{D}} M \quad \text{as } k \to \infty.$$

To obtain the desired result it is enough to show that $M \overset{a.s.}{=} 0$, and for this we need only to prove that $EMM' = 0$. By direct computation we obtain

$$E\left(\sum_{i=1}^{m} A_{n,1} \cdots A_{n,i} \right)\left(\sum_{i=1}^{m} A_{n,1} \cdots A_{n,i} \right)$$

$$= mI + \sum_{i=1}^{m-1} E(A_{n,1} \cdots A_{n,i}[(EA_{n,1} - (EA_{n,1})^{m-i+1})(I - EA_{n,1})^{-1}$$

$$+ (EA'_{n,1} - (EA'_{n,1})^{m-i+1})(I - EA'_{n,1})^{-1}]A'_{n,i} \cdots A'_{n,1}),$$

where $(I - EA_{n,1})^{-1}$ and $(I - EA'_{n,1})^{-1}$ exist by Lemma 5.1. Now, the terms of the sum are bounded uniformly in n, for if not it would be because $\{(I - EA_{n,1})^{-1}\}$ became unbounded, but if $\|(I - EA_{n,1})^{-1}\| = K_n$ can become arbitrarily large, since there are vectors x_n and y_n such that $\|x_n\| = 1$, $\|y_n\| = K_n$ and $(I - EA_{n,1})^{-1} x_n = y_n$, or $EA_{n,1} z_n = z_n - w_n$ where $z_n = y_n/K_n$ and $w_n = x_n/K_n$, and since there is a subsequence $\{n_k\}$ of $\{n\}$ such that, as $k \to \infty$, $A_{n_k,1} \xrightarrow{\mathfrak{D}} A$ and hence $EA_{n_k,1} \to EA$, and $z_{n_k} \to z$ with $\|z\| = 1$, it follows that $EAz = z$, so that $I - EA$ is singular, which by Lemma 5.1 contradicts the hypothesis that $\{A_{n,1}\}$ has no reducible weak limit points. Therefore

$$E\left(\sum_{i=1}^{m} A_{n,1} \cdots A_{n,i} \right)\left(\sum_{i=1}^{m} A_{n,1} \cdots A_{n,i} \right) = o(m^2) \quad \text{as } n, m \to \infty,$$

which implies the result we need because

$$EMM' = \lim_{k \to \infty} \cdot E\left(\frac{1}{r_{n_k}} \sum_{i=1}^{r_{n_k}} A_{s_{n_k},1} \cdots A_{s_{n_k},i}\right)\left(\frac{1}{r_{n_k}} \sum_{i=1}^{r_{n_k}} A_{s_{n_k},1} \cdots A_{s_{n_k},i}\right).$$

The first conclusion is thus proved.

The second conclusion is obtained as in Lemma 5.3. □

The following covariance limit theorem derives from Lemmas 5.3 and 5.4 by elementary computations (similar to part of the proof of Lemma 3.7), and therefore we omit the proof.

Theorem. *For each $n \geq 1$ let $A_{n,i}$, $i \geq 1$, be independent and identically distributed irreducible random orthogonal matrices, and let ξ_n, $n \geq 1$, be random vectors such that, for each n, $\{A_{n,i}, i \geq 1\}$ and ξ_n are independent. Assume that $\{A_{n,1}, n \geq 1\}$ has no reducible weak limit points and $\lim_{n \to \infty} (I - EA_{n,1})^{-1}$ exists, and that $\{\|\xi_n\|^2, n \geq 1\}$ is uniformly integrable and there is a random variable ν such that $\|\xi_n\|^2 \xrightarrow{D} \nu$ as $n \to \infty$. Then if*

$$X_{n,m} = (1/\sqrt{m}) \sum_{i=1}^{m} A_{n,i} \cdots A_{n,1} \xi_n$$

we have

$$E[(X_{n,m} - EX_{n,m})(X_{n,m} - EX_{n,m})'|\xi_n] \xrightarrow{D} \frac{\nu}{d}M \quad as\ n, m \to \infty,$$

where

$$M = \lim_{n \to \infty} [(I - EA_{n,1})^{-1} + (I - EA'_{n,1})^{-1}] - I,$$

and if ν is constant and $\{\|\xi_n\|^{2p}, n \geq 1\}$ is uniformly integrable for given $p \geq 1$, then

$$E\left\|E[(X_{n,m} - EX_{n,m})(X_{n,m} - EX_{n,m})'|\xi_n] - \frac{\nu}{d}M\right\|^p \to 0 \quad as\ n, m \to \infty.$$

The matrix M has the following properties. It is clearly symmetric and non-negative definite. If $d \geq 3$ then rank $M \geq 2$, but if $d = 1$ or 2 then M can be zero, which is possible only if all the weak limit points of $\{A_{n,1}\}$ are a.s. constant (-1 for $d = 1$) (in this case it is enough for the theorem that $\{\|\xi_n\|^2\}$ be tight). If the $A_{n,1} = (a_{ij}^{(n)})$ satisfy Condition (A) of §2, then

$$M = \lim_{n \to \infty} \frac{1 - \sum_{j=1}^{d} (Ea_{ii}^{(n)})^2}{1 - 2Ea_{ii}^{(n)} + \sum_{j=1}^{d} (Ea_{ii}^{(n)})^2} I,$$

where i is arbitrary.

Let us prove these properties. Let A be a weak limit point of $\{A_{n,1}\}$; then

$M = (I - EA)^{-1} + (I - EA')^{-1} - I$, or equivalently $I - EAEA' = (I - EA)M(I - EA')$, and from this expression we see that $M = 0$ if and only if $EAEA' = I$, that is $A \overset{\text{a.s.}}{=} EA$, and that $\text{rank}(I - EAEA') = \text{rank}\, M$; now if $\text{rank}(I - EAEA') = 1$ then $EAEA'$ has $d - 1$ characteristic values equal to 1, which can be shown to imply that the other characteristic value must also be 1, but then $\text{rank}(I - EAEA') = 0$, a contradiction; furthermore for $d \geq 3$, A cannot be a.s. constant because by assumption it is irreducible. If $d = 1$ or 2 then A can be a.s. constant. If the $A_{n,1}$ satisfy Condition (A) we have $EA'_{n,1}EA_{n,1} = \alpha_n I$ and $(I - EA'_{n,1})(I - EA_{n,1}) = \beta_n I$ with $\beta_n > 0$, and therefore

$$(I - EA_{n,1})^{-1} + (I - EA'_{n,1})^{-1} - I$$

$$= \beta_n^{-1}(I - EA'_{n,1} + I - EA_{n,1} - (I - EA'_{n,1})(I - EA_{n,1}))$$

$$= \beta_n^{-1}(I - EA'_{n,1}EA_{n,1}) = \beta_n^{-1}(1 - \alpha_n)I,$$

where

$$\alpha_n = \sum_{j=1}^{d} (Ea_{ji}^{(n)})^2 \quad \text{and} \quad \beta_n = 1 - 2Ea_{ii}^{(n)} + \sum_{j=1}^{d} (Ea_{ji}^{(n)})^2$$

for any $i = 1, \cdots, d$.

Finally, we remark that Lemmas 5.3 and 5.4 and the theorem admit the possibility that $\{\|EA_{n,1}\|\}$ have 1 as a limit point. However, if the $A_{n,1}$ satisfy Condition (A) this is possible only for $d = 1$ or 2, for if A is a weak limit point of $\{A_{n,1}\}$ and $\|EA\| = 1$, there are vectors x and y such that $\|x\| = \|y\| = 1$ and $EAx = y$, hence $Ax \overset{\text{a.s.}}{=} y$ and therefore $A'y \overset{\text{a.s.}}{=} x$ and $EA'y = x$, so that $EAEA'y = y$; on the other hand A satisfies Condition (A), in particular $EA'EA = \alpha I$; we conclude that $\alpha = 1$, so that $A \overset{\text{a.s.}}{=} EA$, which is possible only for $d = 1$ or 2.

Acknowledgement. The author would like to thank the referee for an improvement in the proof of Lemma 5.2.

Added in proof. S. Watanabe and T. Watanabe, *Convergence of isotropic scattering transport process to Brownian motion*, Nagoya Math. J. 40 (1970), 161–171, have extended the result and method of proof of [9] to any dimension, under the same hypothesis. This result is still a special case of our theorem.

REFERENCES

1. P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968. MR 38 #1718.

2. K. L. Chung, *A course in probability theory*, Harcourt, Brace & World, New York, 1968.   MR 37 #4842.

3. D. L. Iglehart, *Weak convergence of probability measures on product spaces with applications to sums of random vectors*, Technical Report No. 109, Dept. of Operations Research and Dept. of Statistics, Stanford University, Stanford, Calif., 1968.

4. L. G. Gorostiza, *Some limit theorems in renewal theory*, Bol. Soc. Mat. Mexicana 17 (1972), 32–39.

5. M. Kac, *Probability and related topics in physical sciences*, Lectures in Appl. Math., vol. 1, Interscience, New York, 1959.   MR 21 #1635.

6. B. Rosén, *On asymptotic normality of sums of dependent random vectors*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 7 (1967), 95–102.   MR 35 #1074.

7. C. J. Stone, *Weak convergence of stochastic processes defined on semi-infinite time intervals*, Proc. Amer. Math. Soc. 14 (1963), 694–696.   MR 27 #3015.

8. V. N. Tutubalin, *The central limit theorem for random motions of Euclidean space*, Vestnik Moskov. Univ. Ser. I Mat. Meh. 22 (1967), no. 6, 100–108. (Russian)   MR 39 #3566.

9. T. Watanabe, *Weak convergence of the isotropic scattering transport process with one speed in the plane to Brownian motion*, Proc. Japan Acad. 44 (1968), 677–680.   MR 38 #5289.

10. W. Whitt, *Weak convergence of probability measures on the function space $C[0, \infty)$*, Ann. Math. Statist. 41 (1970), 939–944.   MR 41 #6259.

11. ———, *Weak convergence of probability measures on the function space $D[0, \infty)$* (unpublished).

CENTRO DE INVESTIGACION Y DE ESTUDIOS AVANZADOS, INSTITUTO POLITECNICO NACIONAL, MEXICO 14, D. F., MEXICO