The Annals of Mathematical Statistics
1968, Vol. 39, No. 3, 864–876

LIMIT THEOREMS FOR THE MULTI-URN EHRENFEST MODEL[1]

BY DONALD L. IGLEHART

Cornell University and Stanford University

1. Introduction and summary. In the multi-urn Ehrenfest model N balls are distributed among $d + 1$ $(d \geqq 2)$ urns. If we label the urns $0, 1, \cdots, d$, then the system is said to be in state $\mathbf{i} = (i_1, i_2, \cdots, i_d)$ when there are i_j balls in urn j $(j = 1, 2, \cdots, d)$ and $N - \mathbf{1}\cdot\mathbf{i}$ balls in urn 0. (The vector $\mathbf{1}$ has all its components equal to 1 and $\mathbf{x}\cdot\mathbf{y}$ is the usual scalar product.) At discrete epochs a ball is chosen at random from one of the $d + 1$ urns; each of the N balls has probability $1/N$ of being selected. The ball chosen is removed from its urn and placed in urn i $(i = 0, 1, \cdots, d)$ with probability p^i, where the p^i's are elements of a given vector, $(p^0, \mathbf{p})$, satisfying $p^i > 0$ and $\sum_{i=0}^d p^i = 1$. We shall let $\mathbf{X}_N(k)$ denote the state of the system after the kth such rearrangement of balls. Our interest in this paper is to obtain limit theorems for the sequence of processes $\{\mathbf{X}_N(k): k = 0, \cdots, N\}$ as N tends to infinity.

For the classical Ehrenfest model $(d = 1, p^0 = p^1 = \frac{1}{2})$ Kac [7] showed that the distribution of $(X_N([Nt]) - N/2)/(N/2)^{\frac{1}{2}}$ converges as $N \to \infty$ to the distribution of the Ornstein-Uhlenbeck process at time t having started at y_0 at $t = 0$, provided $X_N(0) = [(N/2)^{\frac{1}{2}}y_0 + N/2]$. (The symbol $[x]$ denotes the integer part of x.) Recently, Karlin and McGregor [8] obtained a similar result for the continuous time version of the model with $d = 2$; in this version the random selection of balls is done at the occurrence of events of an independent Poisson process. In addition, they obtained a local limit theorem for the transition function. The proof in [7] depended on the continuity theorem of characteristic functions. On the other hand, the proof in [8] used the properties of the spectral representation of the n-step transition probabilities which is available for these processes. These results suggested the direction we shall follow in this paper.

A preliminary calculation indicates that the process $\{\mathbf{X}_N(k): k = 0, \cdots, N\}$ is attracted to the pseudo-equilibrium state $N\mathbf{p}$ and that states far from $N\mathbf{p}$ will only occur rarely. Thus it is natural to consider the fluctuations of $\mathbf{X}_N(k)$ about $N\mathbf{p}$ measured in an appropriate scale. For our purposes the appropriate processes to consider are $\{\mathbf{Y}_N(k): k = 0, \cdots, N\}$, where

$$\mathbf{Y}_N(k) = (\mathbf{X}_N(k) - N\mathbf{p})/N^{\frac{1}{2}}.$$

Next we define a sequence of stochastic processes $\{\mathbf{y}_N(t): 0 \leqq t \leqq 1\}$ which are continuous, linear on the intervals $((k-1)N^{-1}, kN^{-1})$, and satisfy $\mathbf{y}_N(kN^{-1}) = \mathbf{Y}_N(k)$ for $k = 0, 1, \cdots, N$. In other words we let

$$\mathbf{y}_N(t) = \mathbf{Y}_N(k) + (Nt - k)(\mathbf{Y}_N(k+1) - \mathbf{Y}_N(k))$$

Received 27 August 1967.
[1] This work was supported by Office of Naval Research Contract Nonr-401(55) at Cornell University.

if $kN^{-1} \leqq t \leqq (k + 1)N^{-1}$. Throughout this paper we shall let $X_N^i(0) = [N^{\frac{1}{2}}y_0^i + Np^i]$, where $\mathbf{y}_0 = (y_0^1, \cdots, y_0^d)$ is an arbitrary, but fixed, element of R^d. (It will always be understood that N is sufficiently large so that $0 \leqq X_N^i(0) \leqq N$ for all $i = 1, 2, \cdots, d$, where $X_N^i(\cdot)$ is the ith component of the vector $\mathbf{X}_N(\cdot)$. R^d is d-dimensional Euclidean space.) Observe that this initial condition implies that $|Y_N^i(0) - y_0^i| \leqq N^{-\frac{1}{2}}$. With this initial condition and the Markov structure of the model, the processes $\{\mathbf{X}_N(k):k = 0, \cdots, N\}$ for $N = 1, 2, \cdots$ can be defined on a probability triple $(\Omega_N, \mathfrak{F}_N, P_N)$. We shall let $C_d[0, 1]$ denote the product space of d copies of $C[0, 1]$, the space of continuous functions on $[0, 1]$ with the topology of uniform convergence, and endow $C_d[0, 1]$ with the product topology. The topological Borel field of $C_d[0, 1]$ will be denoted by $\mathcal{C}_d$. Clearly, the transformation taking the sequence $\{\mathbf{X}_N(k):k = 0, \cdots, N\}$ into $\{\mathbf{y}_N(t):0 \leqq t \leqq 1\}$ is measurable and induces a probability measure on $\mathcal{C}_d$. We shall denote this induced measure by $\mu_N(\cdot\,; \mathbf{y}_0)$.

The general notion of weak convergence of a sequence of probability measures is defined as follows. Let S be a metric space and $\mathcal{S}$ be the Borel field generated by the open sets of S. If ν_N and ν are probability measures on $\mathcal{S}$ and if the

$$\lim_{N\to\infty} \int_S f\, d\nu_N = \int_S f\, d\nu$$

for every bounded, continuous function f on S, then we say that ν_N *converges weakly* to ν and write $\nu_N \Rightarrow \nu$.

The principal result of this paper is that $\mu_N(\cdot\,; \mathbf{y}_0) \Rightarrow \mu(\cdot\,; \mathbf{y}_0)$ as $N \to \infty$, where $\mu(\cdot\,; y_0)$ is the probability measure on $\mathcal{C}_d$ of a d-dimensional diffusion process, $\mathbf{y}(\cdot)$, starting at the point $\mathbf{y}_0$. The limit process $\mathbf{y}(\cdot)$ is a d-dimensional analog of the Ornstein-Uhlenbeck process whose distribution at time t is a multi-variate normal with mean vector $e^{-t}\mathbf{y}_0$ and covariance matrix $\mathbf{\Sigma}$, where the elements of $\mathbf{\Sigma}$ are

$$\begin{aligned} \sigma_{ij} &= (1 - e^{-2t})p^i(1 - p^i), & i &= j, \\ &= -(1 - e^{-2t})p^ip^j, & i &\neq j. \end{aligned}$$

For applications it is useful to note that $\mu_N(\cdot\,; \mathbf{y}_0) \Rightarrow \mu(\cdot\,; \mathbf{y}_0)$ is equivalent to the statement that[2]

$$\lim_{N\to\infty} \mu_N(\{f(\mathbf{y}(\cdot)) \leqq \alpha\}; \mathbf{y}_0) = \mu(\{f(\mathbf{y}(\cdot)) \leqq \alpha\}; \mathbf{y}_0)$$

for all functionals f on $C_d[0, 1]$ which are continuous almost everywhere with respect to $\mu(\cdot\,; \mathbf{y}_0)$.

To establish weak convergence two steps are usually required. First, the convergence of the finite-dimensional distributions (fdd) of the approximating processes, $\{\mathbf{y}_N(t):t \geqq 0, N = 1, 2, \cdots\}$ in our case, to the corresponding fdd of the limiting process must be obtained. Second, the probability that the approximating processes can have large fluctuations between points at which these processes are determined by their fdd must be shown to be small. The notion of

[2] The convergence here is in the ordinary sense of convergence of distributions; i.e., for all α for which the right-hand side is continuous.

weak convergence is intimately related to the so-called invariance principles. An invariance principle was first introduced for the case of sums of independent, identically distributed random variables by Erdös and Kac [4] and generalized by Donsker [3]. The method of proof used by Erdös-Kac and Donsker was later modified by Billingsley [1] for dependent random variables. In carrying out the second step outlined above we shall follow Billingsley's argument in Theorems 2.3 and 3.1 of [1]. A program similar in nature to ours was carried out by Lamperti [10] for a particular class of Markov processes.

This paper is organized into the following sections: Section 2 is devoted to our analog of the central limit theorem (clt), namely, that the distribution of $\mathbf{y}_N(t)$ converges to the distribution of $\mathbf{y}(t)$ at a single fixed value of t. In Section 3 the limit process $\mathbf{y}(t)$ is identified and the properties of the process needed here are discussed. Section 4 completes the proof of the convergence of the fdd of $\{\mathbf{y}_N(t)\}$ to those of $\{\mathbf{y}(t)\}$. The proofs in both Sections 2 and 4 are carried out using the Lévy continuity theorem for characteristic functions. Section 5 provides the proof required to show weak convergence. The main tool here, in addition to Billingsley's theorems mentioned above, is the result of Stone (1961) on the weak convergence of random walks. Finally, in Section 6 applications are mentioned along with a suggestion as to how the multi-urn Ehrenfest model might be used to study certain problems in statistical mechanics, networks of queues, and epidemic theory.

2. Analog of the central limit theorem. We begin by selecting an arbitrary vector $\mathbf{y}_0$ in R^d. Next we set $X_N^i(0) = [N^{\frac{1}{2}}y_0^i + Np^i]$ and only consider those values of N large enough to insure that $0 \leqq X_N^i(0) \leqq N$ for $i = 1, 2, \cdots, d$. For those values of N let the Markov process $\{\mathbf{X}_N(k): k = 0, 1, \cdots, N\}$, which characterizes the state of our multi-urn scheme, be defined on a probability triple $(\Omega_N, \mathfrak{F}_N, P_N)$. The construction of this probability triple from the initial distribution, one-step transition probabilities, and state space is standard; see Chung [2]; Theorem 1, page 7. In this section we shall prove for the multi-urn Ehrenfest model the analog of the clt for sums of independent, identically distributed random variables. The result is

THEOREM 1. *For any vector* $\mathbf{y}_0$ *in* R^d *and initial condition* $X_N^i(0) = [N^{\frac{1}{2}}y_0^i + Np^i]$ *the*

$$\lim_{N\to\infty} P_N\{\mathbf{Y}_N([Nt]) \leqq \mathbf{x}\} = (2\pi)^{-d/2}|\mathbf{\Sigma}|^{-\frac{1}{2}} \int_{\{\mathbf{y}:\mathbf{y}\leqq\mathbf{x}\}} \exp\{-\tfrac{1}{2}(\mathbf{y}-\boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})\}\, dy$$

for all $\mathbf{x} \in R^d$ *and* $t \in [0, 1]$, *where* $\boldsymbol{\mu} = e^{-t}\mathbf{y}_0$, *and* $\mathbf{\Sigma}$ *has elements*

$$\begin{aligned} \sigma_{ij} &= (1 - e^{-2t})p^i(1 - p^i), & i = j, \\ &= -(1 - e^{-2t})p^ip^j, & i \neq j. \end{aligned}$$

The proof of this theorem will depend on the Lévy continuity theorem for characteristic functions and the following two lemmas.

LEMMA 1. *If we let*[3] $\psi_N(\mathbf{s}; k) = E_N\{\exp\{i\mathbf{s}\cdot\mathbf{Y}_N(k)\}\}$, *then*

$$\psi_N(\mathbf{s}, k+1) = g_N(\mathbf{s})\psi_N(\mathbf{h}_N(\mathbf{s}, 1), k)$$

for $k = 0, 1, \cdots, N-1$, *where*

$$g_N(\mathbf{s}) = \exp\{-N^{-1}\mathbf{s}'\mathbf{A}\mathbf{s} + o(N^{-1})\},$$

$$\mathbf{A} = (1 - e^{-2t})^{-1}\mathbf{\Sigma}, \quad \textit{and}$$

$$\mathbf{h}_N(\mathbf{s}, 1) = (1 - N^{-1} + o(N^{-1}))\mathbf{s}$$

as $N \to \infty$, *and the terms* $o(N^{-1})$ *are uniform for* $\mathbf{s}$ *in a compact set of* R^d *and independent of* k.

PROOF. Using a standard conditional probability argument we see that

$$(1) \quad \psi_N(\mathbf{s}, k+1) = E_N\{\exp\{i\mathbf{s}\cdot\mathbf{Y}_N(k)\} \cdot E_N\{\exp\{iN^{-\frac{1}{2}}\mathbf{s}\cdot(\mathbf{X}_N(k+1) - \mathbf{X}_N(k))\} \mid \mathbf{X}_N(k)\}\}.$$

The transitions allowed for $\mathbf{X}_N(k)$ in one step are as follows, where the conditional probability of the transition is indicated at the right. (The vector $\mathbf{e}_i$ has components $e_i^j = \delta_{ij}$.)

$$\begin{aligned}
\mathbf{X}_N(k+1) - \mathbf{X}_N(k) &= 0, & & p^0(1 - N^{-1}\mathbf{1}\cdot\mathbf{X}_N(k)) + N^{-1}\mathbf{X}_N(k)\cdot\mathbf{p}, \\
&= -\mathbf{e}_i, & & p^0 N^{-1} X_N^i(k), \\
&= \mathbf{e}_i, & & p^i(1 - N^{-1}\mathbf{1}\cdot\mathbf{X}_N(k)), \\
&= \mathbf{e}_j - \mathbf{e}_i, & & p^j N^{-1} X_N^i(k) \qquad \text{for} \quad i \neq j.
\end{aligned}$$

Hence one version of the conditional expectation indicated in (1) is

$$\begin{aligned}
p^0(1 - N^{-1}\mathbf{1}\cdot\mathbf{X}_N(k)) &+ N^{-1}\mathbf{X}_N(k)\cdot\mathbf{p} + p^0 N^{-1}\mathbf{X}_N(k)\cdot\exp(-iN^{-\frac{1}{2}}\mathbf{s}) \\
&+ (1 - N^{-1}\mathbf{1}\cdot\mathbf{X}_N(k))\mathbf{p}\cdot\exp(iN^{-\frac{1}{2}}\mathbf{s}) \\
&+ N^{-1}\textstyle\sum_{\{(j,m):j\neq m\}} \exp\{iN^{-\frac{1}{2}}(s^j - s^m)\}p^j X_N^m(k).
\end{aligned}$$

Here and in the future we shall use the notation $f(\mathbf{s})$ as shorthand for the vector $(f(s_1), f(s_2), \cdots, f(s_d))$. A routine calculation allows us to write this conditional expectation in the following simpler form

$$(p^0 + \mathbf{p}\cdot\exp(iN^{-\frac{1}{2}}\mathbf{s}))\{1 - N^{-1}\mathbf{X}_N(k)\cdot(\mathbf{1} - \exp(-iN^{-\frac{1}{2}}\mathbf{s}))\}.$$

Using this result plus the relation $\mathbf{X}_N(k) = N^{\frac{1}{2}}\mathbf{Y}_N(k) + N\mathbf{p}$ in (1) yields

$$(2) \quad \psi_N(\mathbf{s}, k+1) = (p^0 + \mathbf{p}\cdot\exp(iN^{-\frac{1}{2}}\mathbf{s}))E_N\{\exp(i\mathbf{s}\cdot\mathbf{Y}_N(k)) \cdot [1 - (N^{-\frac{1}{2}}\mathbf{Y}_N(k) + \mathbf{p})\cdot(\mathbf{1} - \exp(-iN^{-\frac{1}{2}}\mathbf{s}))]\}.$$

By expanding the exponentials, the term in square brackets inside the expectation sign can be written as

$$(3) \qquad 1 - iN^{-1}\mathbf{Y}_N(k)\cdot\mathbf{s} - iN^{-\frac{1}{2}}\mathbf{p}\cdot\mathbf{s} - (2N)^{-1}\mathbf{p}\cdot\mathbf{s}^2 - (2N^{3/2})^{-1}\mathbf{Y}_N(k)\cdot\mathbf{s}^2 + o(N^{-1}).$$

[3] The symbol $E_N\{\cdot\}$ denotes expectation with respect to P_N.

For $\mathbf{s}$ in a compact set of R^d, the term $(2N^{3/2})^{-1}\mathbf{Y}_N(k)\cdot\mathbf{s}^2$ can clearly be written as $o(N^{-1})\mathbf{Y}_N(k)\cdot\mathbf{s}$. Thus the logarithm of the expression in (3) is

$$(4)\qquad -i(N^{-1}+o(N^{-1}))\mathbf{Y}_N(k)\cdot\mathbf{s} - iN^{-\frac{1}{2}}\mathbf{p}\cdot\mathbf{s} - (2N)^{-1}\mathbf{p}\cdot\mathbf{s}^2 + (2N)^{-1}(\mathbf{p}\cdot\mathbf{s})^2 + o(N^{-1}).$$

On the other hand, the coefficient of the expectation in (2) can be written as

$$1 + iN^{-\frac{1}{2}}\mathbf{p}\cdot\mathbf{s} - (2N)^{-1}\mathbf{p}\cdot\mathbf{s}^2 + o(N^{-1})$$

and its logarithm as

$$(5)\qquad iN^{-\frac{1}{2}}\mathbf{p}\cdot\mathbf{s} - (2N)^{-1}\mathbf{p}\cdot\mathbf{s}^2 + (2N)^{-1}(\mathbf{p}\cdot\mathbf{s})^2 + o(N^{-1}).$$

Combining the expressions in (4) and (5) with (2) we obtain

$$\psi_N(\mathbf{s}, k+1) = \exp\left(-N^{-1}\mathbf{p}\cdot\mathbf{s}^2 + N^{-1}(\mathbf{p}\cdot\mathbf{s})^2 + o(N^{-1})\right) \cdot E_N\{\exp\left(i\mathbf{s}\cdot\mathbf{Y}_N(k)(1 - N^{-1} + o(N^{-1}))\right)\}.$$

A trivial calculation shows that $\mathbf{p}\cdot\mathbf{s}^2 - (\mathbf{p}\cdot\mathbf{s})^2 = \mathbf{s}'\mathbf{A}\mathbf{s}$ which completes the proof of the lemma.

LEMMA 2. *The characteristic function*

$$\psi_N(\mathbf{s}, k) = \textstyle\prod_{j=0}^{k-1} g_N[\mathbf{h}_N(\mathbf{s}, j)]\psi_N[\mathbf{h}_N(\mathbf{s}, k), 0]$$

for $k = 1, 2, \cdots, N$, *where* $\mathbf{h}_N(\mathbf{s}, 0) = \mathbf{s}$ *and* $\mathbf{h}_N(\mathbf{s}, j) = \mathbf{h}_N[\mathbf{h}_N(\mathbf{s}, j-1), 1]$ *for* $j \geqq 1$.

PROOF. This result follows immediately by induction using Lemma 1 and the fact that $\mathbf{h}_N(\mathbf{s}, j) = \mathbf{h}_N[\mathbf{h}_N(\mathbf{s}, 1), j-1]$.

With the help of Lemmas 1 and 2 we now return to the

PROOF OF THEOREM 1. From Lemma 2 we have

$$(6)\qquad \ln \psi_N(\mathbf{s}, [Nt]) = \textstyle\sum_{j=0}^{[Nt]-1} \ln g_N[\mathbf{h}_N(\mathbf{s}, j)] + i\mathbf{h}_N(\mathbf{s}, [Nt])\cdot\mathbf{Y}_N(0).$$

From the definition of $\mathbf{h}_N(\mathbf{s}, 1)$ and $\mathbf{h}_N(\mathbf{s}, j)$ it is easy to verify that $\mathbf{h}_N(\mathbf{s}, j) = (1 - N^{-1} + o(N^{-1}))^j\mathbf{s}$. Thus from Lemma 1

$$(7)\qquad \textstyle\sum_{j=0}^{[Nt]-1} \ln g_N[\mathbf{h}_N(\mathbf{s}, j)] = -N^{-1}\mathbf{s}'\mathbf{A}\mathbf{s}\sum_{j=0}^{[Nt]-1} (1 - N^{-1} + o(N^{-1}))^{2j} + o(1).$$

Furthermore,

$$(8)\qquad \textstyle\sum_{j=0}^{[Nt]-1} (1 - N^{-1} + o(N^{-1}))^{2j} \sim N(1 - e^{-2t})/2$$

and

$$(9)\qquad h_N(\mathbf{s}, [Nt]) \to e^{-t}\mathbf{s} \quad \text{as} \quad N \to \infty,$$

where the second relation is meant component-wise. Thus combining (7), (8), and (9) with (6) yields the fact that the

$$(10)\qquad \lim_{N\to\infty} \ln \psi_N(\mathbf{s}, [Nt]) = -\tfrac{1}{2}\mathbf{s}'\mathbf{\Sigma}\mathbf{s} + ie^{-t}\mathbf{y}_0\cdot\mathbf{s}.$$

Finally, appealing to the Lévy continuity theorem for characteristic functions the limit in (10) establishes the theorem.

3. The multi-dimensional Ornstein-Uhlenbeck process. In this section it will be convenient to assemble the few facts we shall need about our limit process, $\mathbf{y}(t)$. The multi-dimensional Ornstein-Uhlenbeck process is a diffusion process (strong Markov process with continuous paths) governed by the transition probability kernel

$$g(t; \mathbf{x}, \mathbf{y}) = (2\pi)^{-d/2} |\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\{-\tfrac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})\}$$

for $t \varepsilon [0, 1]$ and $\mathbf{x}, \mathbf{y} \varepsilon R^d$, where $\boldsymbol{\mu} = e^{-t}\mathbf{x}$ and $\mathbf{\Sigma} = (1 - e^{-2t})\mathbf{A}$ with $\mathbf{A}$ being a $d \times d$ positive definite symmetric matrix. It is easy to verify that this kernel is the elementary solution of the parabolic partial differential equation

$$(\partial/\partial t)g(t; \mathbf{x}, \mathbf{y}) = \textstyle\sum_{i,j} a_{ij}(\partial^2 g(t; \mathbf{x}, \mathbf{y})/\partial x_i \partial x_j) - \sum_i x_i(\partial g(t; \mathbf{x}, \mathbf{y})/\partial x_i)$$

where $t > 0$, $\mathbf{x}, \mathbf{y} \varepsilon R^d$.

For any finite number of times $0 < t_1 < t_2 < \cdots < t_k < 1$ it is easy to calculate the joint characteristic function of the vectors $\mathbf{y}(t_1), \mathbf{y}(t_2), \cdots, \mathbf{y}(t_k)$. In fact, if we set $\mathbf{y}(0) = \mathbf{y}_0$ with probability one and let

$$\phi(\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_k; t_1, t_2, \cdots, t_k) = E\{\exp[i(\mathbf{s}_1 \cdot \mathbf{y}(t_1) + \cdots + \mathbf{s}_k \cdot \mathbf{y}(t_k))]\},$$

then it is easy to show that

$$\begin{aligned}\phi(\mathbf{s}_1, \cdots, \mathbf{s}_k; t_1, \cdots, t_k) = \exp\{&-\tfrac{1}{2}\mathbf{s}_k'(b_k\mathbf{A})\mathbf{s}_k \\ &- \tfrac{1}{2}(\mathbf{s}_{k-1} + a_k\mathbf{s}_k)'(b_{k-1}\mathbf{A})(\mathbf{s}_{k-1} + a_k\mathbf{s}_k) \cdots -\tfrac{1}{2}(\mathbf{s}_1 + a_2\mathbf{s}_2 \\ &+ \cdots + \textstyle\prod_{j=2}^k a_j\mathbf{s}_k)'(b_1\mathbf{A})(\mathbf{s}_1 + a_2\mathbf{s}_2 + \cdots + \prod_{j=2}^k a_j\mathbf{s}_k) \\ &+ i(\mathbf{s}_1 + a_2\mathbf{s}_2 + \cdots + \textstyle\prod_{j=2}^k a_j\mathbf{s}_k)\cdot(a_1\mathbf{y}_0)\},\end{aligned}$$

where $a_i = \exp\{-(t_i - t_{i-1})\}$, $b_i = (1 - \exp\{-2(t_i - t_{i-1})\})$, and $t_0 = 0$.

The stochastic process $\{\mathbf{y}(t): 0 \leqq t \leqq 1\}$ can be constructed on a probability triple as follows. Let $C[1, e^2]$ be the space of continuous functions on the closed interval $[1, e^2]$ and endow it with the topology of uniform convergence. Denote by $\mathcal{C}[1, e^2]$ the topological Borel field of $C[1, e^2]$. Construct on $\mathcal{C}[1, e^2]$ Wiener measure corresponding to Brownian motion, $x(\cdot)$, on $[1, e^2]$ with $x(1)$ a constant with probability one; see Itô and McKean [6], section 1.4. Now take d copies of $C[1, e^2]$ and form the product space $C_d[1, e^2]$ with the product topology and the product Borel field $\mathcal{C}_d[1, e^2]$. Take for the measure on $\mathcal{C}_d[1, e^2]$ the product measure with the initial condition for the ith Brownian motion being $x^i(1) = \sum_{j=1}^d b_{ij}^{-1} y_0^j$, where the matrix $B^{-1} = \{b_{ij}^{-1}\}$ is the inverse of the symmetric positive definite root of A; i.e., $\mathbf{A} = \mathbf{BB}'$. Now let $\{\mathbf{x}(t): 1 \leqq t \leqq e^2\}$ be this d-dimensional Brownian motion and introduce the transformation

$$\mathbf{y}(t) = e^{-t}\mathbf{B}\mathbf{x}(e^{2t}), \qquad 0 \leqq t \leqq 1.$$

This transformation is a measurable mapping of $C_d[1, e^2]$ onto $C_d[0, 1]$ and induces on $\mathcal{C}_d$ the measure $\mu(\cdot; \mathbf{y}_0)$. The measure μ is the one generated by the kernel $g(t; \mathbf{x}, \mathbf{y})$ and the one referred to in Section 1.

4. Convergence of finite dimensional distributions. We proceed now to show that the fdd of the sequence of processes $\{\mathbf{y}_N(t):0 \leqq t \leqq 1\}$ for $N = 1, 2, \cdots$ converge to the corresponding distributions of the multi-dimensional Ornstein-Uhlenbeck process, $\{\mathbf{y}(t):0 \leqq t \leqq 1\}$. The convergence of the one-dimensional distributions has already been established for the sequence $\{\mathbf{Y}_N([Nt]):0 \leqq t \leqq 1\}$ for $N = 1, 2, \cdots$. But since $|Y_N^i([Nt]) - y_N^i(t)| \leqq N^{-\frac{1}{2}}$ for $i = 1, 2, \cdots, d$ with probability one, it is clear that we also have the convergence of the one-dimensional distributions of $\{\mathbf{y}_N(t):0 \leqq t \leqq 1\}$ as $N \to \infty$. More generally, we have

THEOREM 2. *The*

$$\lim_{N\to\infty} P_N\{\mathbf{Y}_N([Nt_1]) \leqq \mathbf{x}_1, \cdots, \mathbf{Y}_N([Nt_k]) \leqq \mathbf{x}_k\}$$

exists and is the same as the joint distribution of $\{\mathbf{y}(t):0 \leqq t \leqq 1\}$.

PROOF. Since the argument used for the convergence of the two-dimensional distributions embodies that for the general case, we shall restrict our attention to the case $k = 2$. Let $\psi_N(\mathbf{s}_1 . \mathbf{s}_2 ; \mathbf{k}) = E_N\{\exp\{i(\mathbf{s}_1\cdot\mathbf{Y}_N(k_1) + \mathbf{s}_2\cdot\mathbf{Y}_N(k_2))\}\}$, where $\mathbf{k} = (k_1, k_2)$ and $k_1 < k_2$. Then we can write

$$\psi_N(\mathbf{s}_1, \mathbf{s}_2 ; \mathbf{k}) = E_N\{\exp(i\mathbf{s}_1\cdot\mathbf{Y}_N(k_1))E_N\{\exp(i\mathbf{s}_2\cdot\mathbf{Y}_N(k_2))| \mathbf{X}_N(k_1)\}\}.$$

Now using the notation and arguments of Lemmas 1 and 2 we obtain the relation

$$E_N\{\exp(i\mathbf{s}_2\cdot\mathbf{Y}_N(k_2))| \mathbf{X}_N(k_1)\} = \prod_{j=0}^{k_2-k_1-1} g_N[\mathbf{h}_N(\mathbf{s}_2, j)] \exp\{i\mathbf{h}_N(\mathbf{s}_2, k_2 - k_1)\cdot\mathbf{Y}_N(k_1)\}.$$

Hence

$$\psi_N(\mathbf{s}_1, \mathbf{s}_2 ; \mathbf{k}) = \prod_{j=0}^{k_2-k_1-1} g_N[\mathbf{h}_N(\mathbf{s}_2, j)]\psi_N(\mathbf{s}_1 + \mathbf{h}_N(\mathbf{s}_2, k_2 - k_1), k_1).$$

If we now set $k_1 = [Nt_1]$ and $k_2 = [Nt_2]$ and let $N \to \infty$, we have

$$\begin{aligned}\lim_{N\to\infty} \psi_N(\mathbf{s}_1, \mathbf{s}_2 ; ([Nt_1], [Nt_2])) = \exp\{&-\tfrac{1}{2}\mathbf{s}_2'(b_2\mathbf{A})\mathbf{s}_2 \\ &- \tfrac{1}{2}(\mathbf{s}_1 + a_2\mathbf{s}_2)'(b_1\mathbf{A})(\mathbf{s}_1 + a_2\mathbf{s}_2) \\ &+ i(\mathbf{s}_1 + a_2\mathbf{s}_2)\cdot a_1\mathbf{y}_0\}.\end{aligned}$$

But this limit is exactly the two-dimensional characteristic function of $\{\mathbf{y}(t):0 \leqq t \leqq 1\}$ given in Section 3. This completes the proof of Theorem 2.

Again the convergence of the fdd of $\{\mathbf{Y}_N[Nt]:0 \leqq t \leqq 1\}$ is equivalent to the convergence of the fdd of $\{\mathbf{y}_N(t):0 \leqq t \leqq 1\}$.

5. Weak convergence. To complete the proof of weak convergence we shall follow the methods of Donsker [3] and Billingsley [1]. We shall need a theorem which is a straightforward generalization of Billingsley's Theorem 2.3.

Let A_N be the set of functions $\mathbf{y} \varepsilon C_d[0, 1]$ which are linear on each of the intervals $((i - 1)N^{-1}, iN^{-1})$ for $i = 1, 2, \cdots, N$ and satisfy $\mathbf{y}^j(0) =$

$([N^{\frac{1}{2}}y_0^j + Np^j] - Np^j)N^{-\frac{1}{2}}$ for $j = 1, 2, \cdots, d$. Then from the definition of the measures $\mu_N(\cdot\,; \mathbf{y}_0)$ given in the introduction, we have $\mu_N(A_N; \mathbf{y}_0) = 1$. For any positive integer c and vectors $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_c, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_c \in R^d$ we shall define sets in $C_d[0, 1]$ such as

$$(11)\quad E = \{\mathbf{x} \in C_d[0, 1]; \boldsymbol{\alpha}_j \leqq \mathbf{x}(t) \leqq \boldsymbol{\beta}_j, (j-1)c^{-1} \leqq t \leqq jc^{-1}, j = 1, 2, \cdots . c\}.$$

Then by generalizing Billingsley's set-up for measures on $\mathfrak{C}$ (Borel sets of $C[0, 1]$) to measures on $\mathfrak{C}_d$ we have

THEOREM 3. *If the* $\lim_{N\to\infty} P_N\{G_N\} = \mu(E; \mathbf{y}_0)$ *for any set* E *of the form* (11) *where*

$$(12)\quad G_N = \{\omega: \boldsymbol{\alpha}_j \leqq \mathbf{Y}_N(i, \omega) \leqq \boldsymbol{\beta}_j;\ (j-1)c^{-1} \leqq iN^{-1} \leqq jc^{-1}, j = 1, \cdots, c\} \subset \Omega_N,$$

then $\mu_N(\cdot\,; \mathbf{y}_0) \Rightarrow \mu(\cdot\,; \mathbf{y}_0)$ *as* $N \to \infty$.

This theorem will be our main tool in proving weak convergence. In addition, however, we shall need the following lemmas. Recalling that $\{\mathbf{X}_N(k): k = 0, 1, \cdots, N\}$ is defined on the probability triple $(\Omega_N, \mathfrak{F}_N, P_N)$ mentioned in Section 2 we have

LEMMA 3. *For every* $\epsilon > 0$, *there exists a* $\lambda_0 > 0$ *and a positive integer* N_0 *such that the*

$$(13)\qquad P_N\{\max_{0\leqq k\leqq N} |Y_N^i(k)| \leqq \lambda_0, i = 1, \cdots, d\} \geqq 1 - \epsilon$$

for $N \geqq N_0$.

PROOF. Choose $\epsilon > 0$. Suppose we can show that for some N_0 and λ_0 that

$$(14)\qquad P_N\{\max_{0\leqq k\leqq N} |Y_N^i(k)| > \lambda_0\} \leqq \epsilon/d \qquad \text{when} \qquad N \geqq N_0$$

for $i = 1, \cdots, d$. If we denote the complement of the set in (14) by A_i, then

$$P_N\{\textstyle\bigcup_{i=1}^d A_i^c\} \leqq \sum_{i=1}^d P_N\{A_i^c\} \leqq \epsilon$$

for $N \geqq N_0$, or equivalently

$$(15)\qquad P_N\{\textstyle\bigcap_{i=1}^d A_i\} \geqq 1 - \epsilon,$$

which is the desired inequality. Hence it will suffice to prove (14). Furthermore, since there are only a finite number of values for i, we need only prove (14) for a fixed but arbitrary value of i.

Consider, therefore, the inequality (14) for $i = 1$. The random variable $X_N^1(k)$ can be thought of as describing the state of an Ehrenfest urn model consisting of two urns. One of the urns is urn 1 and the other urn represents an aggregation of urns 0, 2, 3, $\cdots$, d. Again a ball is chosen at random and placed in urn 1 with probability p^1 and in the other urn with probability $1 - p^1$. The random variable $Y_N^1(k) \equiv (X_N^1(k) - Np^1)/N^{\frac{1}{2}}$ is the appropriate one to consider for a limit theorem in this modified urn model. Appealing to Theorem 4.2 of Stone (1961) it is easy to check that the sequence of processes $\{Y_N^1([Nt]); t \in [0, 1]\}$

converges weakly to the Ornstein-Uhlenbeck process with infinitesimal mean $-y$ and infinitesimal variance $2p^1(1 - p^1)$. For this limit diffusion $\pm\infty$ are natural boundaries and hence inaccessible. We shall define the first exit time of the process $X(\cdot)$ from the interval (x_1, x_2) as

$$T_{x_1,x_2}(X(\cdot)) = \inf\{t \geqq 0: X(t) \geqq x_2 \quad \text{or} \quad X(t) \leqq x_1\},$$

then applying Stone's Lemma 3.7 we obtain

$$P_N\{T_{-\lambda,\lambda}(Y_N([N\cdot])) \leqq 1\} \leqq \epsilon/d \tag{16}$$

for N and λ sufficiently large. But this is precisely the desired inequality (14), which completes the proof of the lemma.

The second lemma is taken directly from Stone (1961).

LEMMA 4. *For every $\epsilon > 0$ and $\delta > 0$, there exists a positive integer N_1 and a positive number γ such that the*

$$P_N\{\max_{0\leqq\tau\leqq t} |Y_N^i([N\tau]) - Y_N^i(0)| > \epsilon \mid Y_N^i(0) = y_0\} \leqq \delta$$

for $i = 1, \cdots, d$, $N \geqq N_1$, and $t \leqq \gamma$, uniformly for $|y_0| \leqq \lambda_0$.

PROOF. The proof of this result is a direct consequence of Lemma 3.1 and Theorem 4.2 of Stone (1961). The only hypothesis we need check is that the infinitesimal mean and variance of the $Y_N([N\cdot])$ process converge uniformly in compact intervals to those of the Ornstein-Uhlenbeck process. But this fact is immediate and therefore omitted.

With the help of Lemmas 3 and 4 we shall apply Theorem 3 to obtain our principal result. The proof given here follows that of Billingsley [1], Theorem 3.1. The result is

THEOREM 4. $\mu_N(\cdot, \mathbf{y}_0) \Rightarrow \mu(\cdot, \mathbf{y}_0)$ *as $N \to \infty$, for all $\mathbf{y}_0 \in R^d$.*

PROOF. Select a positive integer c, vectors $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_c, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_c \in R^d$, and form the sets E and G_N defined in (11) and (12). Let $G_N(\epsilon)$ and $E(\epsilon)$ be sets similar to G_N and E where the $\boldsymbol{\alpha}_j$'s and $\boldsymbol{\beta}_j$'s are replaced by $\boldsymbol{\alpha}_j + \epsilon\mathbf{1}$ and $\boldsymbol{\beta}_j - \epsilon\mathbf{1}$ with $\epsilon > 0$. Choose a positive integer ν and let

$$N_{j,u} = [(j-1)c^{-1}N + Nc^{-1}\cdot u\nu^{-1}], \qquad j = 1, \cdots, c; \quad u = 1, \cdots, \nu.$$

Now let $G_{N,\nu}$ be a set similar to G_N, however where the inequalities on $\mathbf{Y}_N(i)$ are required to hold only when i is an integer among the $N_{j,u}$. The set $G_{N,\nu}(\epsilon)$ has the same relationship to $G_N(\epsilon)$. Clearly,

$$P_N\{G_{N,\nu}\} \geqq P_N\{G_N\}, \qquad \text{since} \qquad G_{N,\nu} \supset G_N. \tag{17}$$

Furthermore,

$$P_N\{G_N^c\} = 1 - P_N\{G_N\} = \sum_{l=0}^N \sum_{i=1}^d P_N\{A_{N,l,i}\},$$

where

$$A_{N,l,i} = \{\omega: \mathbf{Y}_N(k) \in [\boldsymbol{\alpha}_j, \boldsymbol{\beta}_j], k < l, (j-1)c^{-1} \leqq kN^{-1} \leqq jc^{-1}, \ j = 1, \cdots c;$$
$$Y_N^m(l) \in [\alpha_{j_0}^m, \beta_{j_0}^m], \quad m \leqq i-1, \quad (j_0 - 1)c^{-1} \leqq lN^{-1} \leqq j_0c^{-1};$$
$$Y_N^i(l) \notin [\alpha_{j_0}^i, \beta_{j_0}^i]\} \subset \Omega_N.$$

The set $A_{N,l,i}$ can be described as the set of all $\mathbf{Y}_N(\cdot)$ paths for which $\mathbf{Y}_N(l)$ is the first random variable to break out of the tunnel and $Y_N^i(l)$ is the first component of $\mathbf{Y}_N(l)$ outside the required interval. For the remainder of the proof we let λ_0 be a positive constant which is large enough to satisfy the requirement of (13) and the inequality

$$(18) \qquad \lambda_0 > \max_{1 \leq i \leq d} \max_{1 \leq j \leq c} \{\max (|\alpha_j^i|, |\beta_j^i|)\}.$$

(This requirement is placed on λ_0 so that whenever a path $\{\mathbf{Y}_N(k)\}$ is in an appropriate tunnel set, the condition $\{\max_{0 \leq k \leq N} |Y_N^i(k)| \leq \lambda_0, i = 1, \cdots, d\}$ will not be operative.) Letting $B_N = \{\omega : \max_{0 \leq k \leq N} |Y_N^i(k)| \leq \lambda_0, i = 1, \cdots, d\}$ we obtain

$$(19) \qquad \begin{aligned} P_N\{G_N^c\} &= \textstyle\sum_{l=0}^N \sum_{i=1}^d [P_N\{A_{N,l,i}, B_N\} + P_N\{A_{N,l,i}, B_N^c\}] \\ &\leq \textstyle\sum_{l=0}^N \sum_{i=1}^d P_N\{A_{N,l,i}, B_N\} + \epsilon, \quad \text{for} \quad N \geq N_0, \end{aligned}$$

where N_0 is defined in Lemma 3.

If the index l is such that $N_{j,u-1} < l \leq N_{j,u}$, we can make the further decomposition

$$(20) \qquad \begin{aligned} P_N\{A_{N,l,i}, B_N\} &= P_N\{A_{N,l,i}, B_N, |Y_N^i(N_{j,u}) - Y_N^i(l)| > \epsilon\} \\ &\quad + P_N\{A_{N,l,i}, B_N, |Y_N^i(N_{j,u}) - Y_N^i(l)| \leq \epsilon\}. \end{aligned}$$

Using the Markov property of the $\{\mathbf{X}_N(k) : k = 0, \cdots, N\}$ we obtain the relation

$$(21) \qquad \begin{aligned} &P_N\{A_{N,l,i}, B_N, |Y_N^i(N_{j,u}) - Y_N^i(l)| > \epsilon\} \\ &\quad \leq \textstyle\sum_{\mathbf{j} \varepsilon S} P_N\{A_{N,l-1}, \mathbf{X}_N(l) = \mathbf{j}\} \\ &\qquad \cdot P_N\{|Y_N^i(N_{j,u}) - Y_N^i(l)| > \epsilon \mid \mathbf{X}_N(l) = \mathbf{j}\}, \end{aligned}$$

where,

$$\begin{aligned} A_{N,l-1} = \{\omega : \mathbf{Y}_N(k) \varepsilon [\boldsymbol{\alpha}_j, \boldsymbol{\beta}_j], k \leq l - 1, (j - 1)c^{-1} \\ \leq kN^{-1} \leq jc^{-1}, j = 1, \cdots, c\} \end{aligned}$$

and

$$\begin{aligned} S = \{\mathbf{j} : j^r \varepsilon \{0, 1, \cdots, N\}, \quad r = 1, \cdots, d; \\ (j^m - Np^m)N^{-\frac{1}{2}} \varepsilon [\alpha_{j_0}^m, \beta_{j_0}^m], \\ m = 1, \cdots, i - 1, (j_0 - 1)c^{-1} \leq lN^{-1} \leq j_0 c^{-1}; \\ (j^i - Np^i)N^{-\frac{1}{2}} \varepsilon [\alpha_{j_0}^i, \beta_{j_0}^i]^c \cap [-\lambda_0, \lambda_0]; \\ (j^r - Np^r)N^{-\frac{1}{2}} \varepsilon [-\lambda_0, \lambda_0], r = i + 1, \cdots, d\}. \end{aligned}$$

Since the process $\{\mathbf{X}_N(k) : k = 0, \cdots, N\}$ is temporally homogeneous, the second factor in the sum on the right-hand side of (21) is equal to

$$(22) \qquad P_N\{|Y_N^i(N_{j,u} - l) - Y_N^i(0)| > \epsilon \mid \mathbf{X}_N(0) = \mathbf{j}\}, \qquad \mathbf{j} \varepsilon S.$$

By choice of $N_{j,u}$, we have $N_{j,u} - l \leqq N/c\nu$. Thus we can let $N_{j,u} - l = [N\tau]$, where $\tau \leqq 1/c\nu$. Furthermore, for all $\mathbf{j} \varepsilon S$, j^r can be written as $[N^{\frac{1}{2}} y_0^r + Np^r]$ with $|y_0^r| \leqq \lambda_0$ for $r = 1, \cdots, d$. Hence appealing to Lemma 4 the probability in (22) is less than or equal to δ (arbitrary) for $N \geqq N_1$ and $1/c\nu \leqq \gamma$.

Returning now to (21) we have

$$(23) \qquad P_N\{A_{N,l,i}, B_N, |Y_N^i(N_{j,u}) - Y_N^i(l)| > \epsilon\} \leqq \delta P_N\{A_{N,l,i}\},$$

provided $N \geqq \max(N_0, N_1)$ and $1/c\nu \leqq \gamma$.

Pulling together (19), (20), and (23) yields

$$(24) \quad P_N\{G_N^c\} \leqq \sum_{j,u} \sum_{l=N_{j,u-1}+1}^{N_{j,u}} \sum_{i=1}^{d} P_N\{A_{N,l,i}, |Y_N^i(N_{j,u}) - Y_N^i(l)| \leqq \epsilon\} + \delta \sum_{l=0}^{N} \sum_{i=1}^{d} P_N\{A_{N,l,i}\} + \epsilon$$

for $N \geqq \max(N_0, N_1)$ and $1/c\nu \leqq \gamma$.

All the mutually exclusive events making up the first sum in (24) are contained in $G_{N,\nu}^c(\epsilon)$. Hence we have

$$(25) \qquad P_N\{G_N^c\} \leqq P_N\{G_{N,\nu}^c(\epsilon)\} + \delta + 2\epsilon.$$

Combining (17) and (25) we obtain

$$(26) \qquad P_N\{G_{N,\nu}(\epsilon)\} - \delta - 2\epsilon \leqq P_N\{G_N\} \leqq P_N\{G_{N,\nu}\}.$$

To complete the proof we must show that $P_N\{G_N\} \to \mu(E; \mathbf{y}_0)$ in order to apply Theorem 3. Using Theorem 2 we have for fixed c, ν, and ϵ the

$$(27) \quad \lim_{N\to\infty} P_N\{G_{N,\nu}\} = \mu(\{\mathbf{y} \,\varepsilon\, C_d(0,1) : \mathbf{y}((j-1)c^{-1} + u(c\nu)^{-1}) \,\varepsilon\, [\boldsymbol{\alpha}_j, \boldsymbol{\beta}_j], \; j = 1, \cdots, c, u = 1, \cdots, \nu\}; \mathbf{y}_0)$$

and a similar limit for $P_N\{G_{N,\nu}(\epsilon)\}$. If we call the set on the right-hand side of (27) D_ν and the corresponding set for the limit of $P_N\{G_{N,\nu}(\epsilon)\}$, $D_\nu(\epsilon)$, then from (26)

$$(28) \quad \mu(D_\nu(\epsilon); \mathbf{y}_0) - \delta - 2\epsilon \leqq \liminf_{N\to\infty} P_N\{G_N\} \leqq \limsup_{N\to\infty} P_N\{G_N\} \leqq \mu(D_\nu; \mathbf{y}_0),$$

for $1/c\nu \leqq \gamma$. Now if we let $\nu \to \infty$, we have $\mu(D_\nu; \mathbf{y}_0) \to \mu(E; \mathbf{y}_0)$ and $\mu(D_\nu(\epsilon); \mathbf{y}_0) \to \mu(E(\epsilon); \mathbf{y}_0)$. Since $E(\epsilon) \nearrow \operatorname{int}(E)$ as $\epsilon \searrow 0$ and $\mu(\partial E; \mathbf{y}_0) = 0$, we obtain from (28) by first letting $\nu \to \infty$ and then letting $\epsilon \searrow 0$ that

$$\lim_{N\to\infty} P_N\{G_N\} = \mu(E; \mathbf{y}_0)$$

for every positive integer c and vectors $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_c, \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_c \,\varepsilon\, R^d$. Thus applying Theorem 3 completes the proof.

6. Some applications. In the introduction it was pointed out that the weak convergence of μ_N to μ implied that the

$$\lim_{N\to\infty} \mu_N(\{f(\mathbf{y}_N(\cdot)) \leqq \alpha\}; \mathbf{y}_0) = \mu(\{f(\mathbf{y}(\cdot)) \leqq \alpha\} : \mathbf{y}_0)$$

for all functionals f on $C_d[0, 1]$ which are continuous almost everywhere with respect to μ. Thus for applications we can obtain an approximation for the distribution of $f(\mathbf{y}_N(\cdot))$ [or in turn of a corresponding functional on $\mathbf{X}_N(k)$] by studying only the distribution of $f(\mathbf{y}(\cdot))$. Examples of functionals f which might be of interest in particular applications are

(i) $f_1(\mathbf{y}(\cdot)) = \max_{0 \leqq t \leqq 1} y^i(t)$,

(ii) $f_2(\mathbf{y}(\cdot)) = \max_{0 \leqq t \leqq 1} |y^i(t)|$,

(iii) $f_3(\mathbf{y}(\cdot)) = \int_0^1 |y^i(t)|\, dt$, and

(iv) $f_4(\mathbf{y}(\cdot)) = \int_0^1 (y^i(t))^2\, dt$.

The calculation of the distribution of $f(\mathbf{y}(\cdot))$ is in general very difficult, if not impossible, at the present time. However, for a practical problem requiring an answer, this distribution could be approximated by carrying out a simulation.

There are at least three areas of application in which a multi-urn Ehrenfest model might be useful. The first is statistical mechanics in which the urns would be identified with cells in momentum space and the process $\{\mathbf{X}_N(k): k = 0, \cdots, N\}$ would measure the occupation numbers of elements (e.g., atoms or molecules) in the various cells. The probabilistic structure governing the rearrangement of balls would correspond to the *Stosszahlansatz* describing the Rayleigh model of a gas. For a more detailed description and further references on this subject the reader is referred to Siegert [12]. Actually the physical model in this case would probably be better described by a continuous time analog of $\{\mathbf{X}_N(k)\}$ in which the "collisions" took place according to events of an independent Poisson process, as was done by Karlin and McGregor [8]. However, for the continuous case similar results can be obtained.

The second area of potential application is that of networks of queues. We have in mind here a job-shop with d work centers which correspond to the urns $1, \cdots, d$. Urn 0 corresponds to potential customers outside the shop. The process $\{\mathbf{X}_N(k)\}$ would measure the number of jobs at the various work centers and the transitions in state would correspond to the arrival and/or completion of service of jobs.

In epidemic theory one can envision a population of size N whose individuals are each in one of $(d + 1)$ states of health. Urn 0 might refer to well individuals and urns $1, \cdots, d$ might represent d different diseases. The process would measure the state of the population and the transitions would relate to the probabilities of individuals contracting the various diseases. The model would not be completely satisfactory, however, because of an absence of a contagion effect.

While it is easy to conjure up applications of the above sort, one would undoubtedly find in many specific applications that the probabilities of transfer differ from those assumed for the multi-urn Ehrenfest model. However, there is a distinct possibility that analyses similar to that carried out here would be feasible for such models. As further examples of such work the reader should consult Karlin and McGregor [8], and [9].

REFERENCES

[1] Billingsley, P. (1956). The invariance principle for dependent random variables. *Trans. Amer. Math. Soc.* **83** 250–268.

[2] Chung, K. (1960). *Markov Chains with Stationary Transition Probabilities*. Springer-Verlag, Berlin.
[3] Donsker, M. (1951). An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* No. 6.
[4] Erdös, P. and Kac, M. (1946). On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc.* **52** 292–302.
[5] Iglehart, D. (1965). Limit diffusion approximations for the many server queue and the repairman problem. *J. Appl. Prob.* **2** 429–441.
[6] Itô, K. and McKean, H. P., Jr. (1965). *Diffusion Processes and their Sample Paths*. Springer-Verlag, Berlin.
[7] Kac, M. (1947). Random walk and the theory of Brownian motion. *Amer. Math. Monthly* **54** 369–391.
[8] Karlin, S. and McGregor, J. L. (1965). Ehrenfest urn models. *J. Appl. Prob.* **2** 352–376.
[9] Karlin, S. and McGregor, J. (1964). On some stochastic models in genetics. *Stochastic Models in Medicine and Biology*. Univ. of Wisconsin Press, 245–279.
[10] Lamperti, J. (1962). A new class of probability limit theorems. *J. Math. Mech.* **11** 749–772.
[11] Siegert, A. (1950). On the approach to statistical equilibrium. *Phys. Rev.* **76** 1708–1714.
[12] Stone, C. (1961). Limit theorems for birth and death processes and diffusion processes. Ph.D. thesis. Stanford Univ.