Probab. Th. Rel. Fields 77, 211–229 (1988)

Probability Theory and Related Fields

Ergodicity of a Measure-valued Markov Chain Induced by Random Transformations

R.W.R. Darling

Department of Mathematics, University of South Florida, Tampa, FL, 33620-5700, USA

Abstract. Let $Z_1, Z_2, \ldots$ be a sequence of i.i.d. random transformations (possibly discontinuous) of a compact metric space M, and let E denote the space of normalized mass distributions on M. Given μ in E, let μ_n denote the random measure $\mu \circ (Z_n \circ \ldots \circ Z_1)^{-1}$ (when well-defined). We construct the transition probability P of the E-valued Markov chain (μ_n), and give a necessary and sufficient condition for P to have a unique invariant measure concentrated on the degenerate mass distributions. Convergence to 'statistical equilibrium' of the associated discrete-time stochastic flow is investigated.

Introduction

The Physical Model

Imagine a bounded region T of space occupied by infinitesimal particles, with some normalized mass distribution η at time 0. Suppose that all particles undergo simultaneous random movement, such that the motion of any finite collection of k particles has a known probability law (motions are generally correlated). This probability law allows the trajectories of two particles to collide, in which case they coalesce to a single trajectory thereafter. However 'birth' of new particles is not considered in this model. If the k-particle motions are diffusions, this could be regarded as a probabilistic model of reaction-diffusion (see Fife (1969), and references therein).

In the present article we look at such a system at discrete times $0, 1, 2, \ldots$. The evolution of the system over a single time interval is represented by a probability measure Q on the set of all functions (possibly non-measurable) f from the region of space to itself; thus $f(x)$ represents the location at time 1 of a particle which is at x at time 0. The system is assumed to be time-homogeneous, with 'independent increments', so the evolution of the system

Research supported by NSF grants INT-8420360 and DMS-8502802. This work was done while the author was visiting the Universities of Warwick and Minnesota

from time 0 to time n is represented by a random transformation X_n which is the composition of n independent and identically distributed random transformations each with law Q.

The random transformations X_0 = identity, $X_1, X_2, \ldots$ take values in a space of functions which does not have a countable base. However the quantity of interest is the evolution of the normalized mass distribution η, i.e., the random measure valued process $(\mu_n, n \geqq 0)$, where $\mu_n = \eta \circ X_n^{-1}$ (when well-defined). This is easier to study because the space of normalized mass distributions is a separable, metrisable space under the weak topology. The goal of this research program is to discover all the "statistical equilibria" of the particle system, and to describe the convergence of the random measure valued process $(\mu_n, n \geqq 0)$ to some specific equilibrium. In the language of Markov processes, this is the problem of determining the structure of the set of invariant measures for the Markov chain $(\mu_n, n \geqq 0)$. The present paper is directed in particular to answering the following question:

(*) Under what conditions is there just one statistical equilibrium, in which all the particles have coalesced to a single, randomly moving point?

Background

The main inspiration for the project came from the work of Le Jan (1984, 1985) and Baxendale (1986) on stochastic flows of diffeomorphisms. If T is a differentiable manifold (usually compact), a stochastic process $(X_{st}, 0 \leqq s \leqq t < \infty)$ with values in the diffeomorphism group of T is called a stochastic flow of diffeomorphisms if for all $s \leqq t \leqq u$, $X_{tu} \circ X_{st} = X_{su}$. (For x in T, interpret $X_{st}(x)$ as the position at time t of a particle which is at x at time s.) Given a measure μ on T, the flow induces a measure-valued process $(\mu_t, t \geqq 0)$ where $\mu_t = \mu \circ X_{0t}^{-1}$. Le Jan (1984) considered isotropic stochastic flows on the d-dimensional torus; when $d=1$ or 2, the equilibrium state is a random Dirac measure with uniform distribution on the torus, whereas when $d \geqq 3$, it is a singular diffuse random measure. Le Jan (1985b) and Darling and Le Jan (1988) also studied equilibrium states for isotropic stochastic flows on $\mathbb{R}^d$. Baxendale (1986) investigated a stochastic flow of diffeomorphisms on the sphere where the equilibrium state is almost surely a Dirac measure. Le Jan (1985a) obtained an upper bound for the Hausdorff dimension of the equilibrium measure, in terms of the Lyapounov exponents of a stochastic flow of diffeomorphisms of an arbitrary compact manifold. The subject of Lyapounov exponents of stochastic flows has been discussed by many authors; the reader is referred to the conference proceedings edited by Arnold (1986), and Kifer (1985).

Overview of This Paper

The situation discussed in this paper is far more general, since X_{st} need not even be a homeomorphism. We assume only that the mappings $x \to X_{st}(x)$ are continuous in probability for all $s \leqq t$.

The motivation to study the non-homeomorphic case arises from the physical considerations discussed above (see [*]), and from work on coalescing stochastic flows by Arratia (1979, 1984), Harris (1984), and Darling (1987).

In order to concentrate on essentials, we discretize time, and take as our initial data a probability measure Q on the space of functions from M to M, where M is a compactification of T. The idea is to take an initial normalized mass distribution η, and push it forward under a sequence of independent, identically distributed random transformations with distribution Q.

Sect. 1 is devoted to the construction of a transition probability P on the space of normalized mass distributions by elementary functional-analytic means. In Sect. 4 we answer (*) by characterizing the case where P is ergodic, in terms of the transition probability of the two-point motion.

In Sect. 3 we obtain a realization $(\mu_n,\ n \geqq 0)$ of a measure-valued Markov chain with transition probability P, by the means described in 'The Physical Model' above. After disposing of some measure-theoretic technicalities, we obtain a "random ergodic theorem" (Proposition 3.3) showing the convergence in law of $(\mu_n,\ n \geqq 0)$, at least when the initial value is suitably chosen.

1. Construction of a Measure-valued Markov Chain

Notation. For every Hausdorff space Y discussed in the sequel, let $\mathscr{B}(Y)$ and $\mathscr{B}_0(Y)$ denote the Borel and Baire σ-algebras respectively; and let $b(Y)$ and $C(Y)$ denote the Banach spaces of bounded, Borel measurable (resp. continuous) functions from Y to $\mathbb{R}$ with the supremum norm. For a transition probability $P\colon Y \times \mathscr{B}(Y) \to [0, 1]$, the expressions $P\phi$ (for ϕ in $b(Y)$) and μP (for μ a probability measure on $\mathscr{B}(Y)$), and $\langle \phi, Y \rangle$, have the same measuring as in Revuz (1984).

Henceforward T will denote a locally compact space with a countable base. If T is compact, take $M = T$; if T is non-compact, let M denote a suitable compactification of T.

M is metrizable, and ρ will denote a metric on M compatible with the topology of M. Let Γ denote M^M, considered as the space of functions from M to M with the topology of pointwise convergence (i.e., product topology). Γ is compact by Tychonoff's theorem, and a basis for the topology of Γ is given by sets of the form $\{f\colon f(x_1) \in G_1, \ldots, f(x_n) \in G_n\}$, where $x_1, \ldots, x_n$ are points in M and $G_1, \ldots, G_n$ are open subsets of M. The Borel sets and Baire sets of Γ were studied by Nelson (1959).

The *canonical Γ-valued random field* is the function:

$$Z\colon M \times \Gamma \to M, \qquad Z(x, F) = f(x). \tag{1.1}$$

For each x in M, $Z(x, \cdot)\colon \Gamma \to M$ is automatically $(\mathscr{B}_0(\Gamma), \mathscr{B}(M))$-measurable.

The starting-point of our study is a probability measure Q on $\mathscr{B}_0(\Gamma)$. Our goal will be to construct in a natural way a transition probability on the space of normalized mass distributions on M.

Definition 1.1. A probability measure Q on $\mathscr{B}_0(\Gamma)$ will be called *stochastically continuous* on M if the mapping $x \to Z(x, \cdot)$ is stochastically continuous on M;

i.e., for all x in M, and for all $\varepsilon>0$,

$$Q(\{f\colon \rho(f(x),f(y))>\varepsilon\})\to 0 \quad \text{as } y\to x. \tag{1.2}$$

Remark 1.2. For each x and y in M and each $\varepsilon>0$, the set $A=\{f\colon \rho(f(x),f(y))>\varepsilon\}$ is indeed in $\mathscr{B}_0(\Gamma)$. Note first that A is open in Γ, since A is the union of the basic open sets $\{f\colon \rho(f(x),w)<\varepsilon/2,\ \rho(f(y),\mathfrak{z})<\varepsilon/2\}$ over the set of $(w,\mathfrak{z})$ in M^2 such that $\rho(w,\mathfrak{z})\geqq 2\varepsilon$. Hence A^c is closed in Γ and therefore compact. A^c is a G_δ since it is the intersection of the open sets $\{f\colon \rho(f(x),f(y))<\varepsilon+1/n\}$. Since A^c is a compact G_δ, it follows that A is a Baire set.

For $k\geqq 1$, $x_1,\dots,x_k$ in M, and u in $b(M^k)$, define

$$R^{(k)}u(x_1,\dots,x_k)=E^Q\,[u(Z(x_1),\dots,Z(x_k))], \tag{1.3}$$

where the Q on the right side means that we regard Z as a Γ-valued random variable on the probability space $(\Gamma,\mathscr{B}_0(\Gamma),Q)$. Thus $R^{(k)}$ is a linear map from $b(M^k)$ to $b(M^k)$.

Lemma 1.3. *If Q is stochastically continuous on M* (Definition 1.1), *then $R^{(k)}u$ is in $C(M^k)$ for all u in $C(M^k)$. Also $R^{(k)}$ is a positive contraction on $C(M^k)$ with $R^{(k)}1=1$; hence $R^{(k)}$ is associated with a unique Feller transition probability on M^k, also denoted $R^{(k)}$, namely*

$$R^{(k)}((x_1,\dots,x_k),G)=Q(\{f\in\Gamma\colon (f(x_1,\dots,f(x_k))\in G\}). \tag{1.4}$$

The proof, which is elementary, is omitted.

Definition 1.4. The transition probability $R^{(k)}$ defined in Lemma 1.3 will be called the *k-point transition probability.* Henceforward we shall deal mainly with the cases $k=1$ and $k=2$; so abbreviate $R^{(1)}$ to R and $R^{(2)}$ to S. The iterates of $R^{(k)}$ will be denoted $R_1^{(k)}=R^{(k)}, R_2^{(k)},\dots$.

Definition 1.5. Suppose Q is stochastically continuous on M. Define probability measures $Q_1,Q_2,\dots$ on $\mathscr{B}_0(\Gamma)$ as follows: $Q_1=Q$, and for $j\geqq 2$, Q_j is the unique Baire measure consistent with the system of finite dimensional distributions $\{Q_{j,\alpha},\ \alpha$ a finite subset of $M\}$, defined as follows:

If $\alpha=\{x_1,\dots,x_k\}$ and G is in $\mathscr{B}(M^k)$, then

$$Q_{j,\alpha}(\{f\colon (f(x_1),\dots,f(x_k))\in G\})=R_j^{(k)}1_G(x_1,\dots,x_k). \tag{1.5}$$

This is equivalent to saying that for each $k\geqq 1$ and all u in $C(M^k)$,

$$R_j^{(k)}u(x_1,\dots,x_k)=E^{Q_j}[u(Z(x_1),\dots,Z(x_k))]. \tag{1.6}$$

Lemma 1.6. *If Q is stochastically continuous on M, then so is Q_j for every $j\geqq 2$.*

Proof. To show stochastic continuity of Q_j, take $k=2$ and consider the distance function ρ on M^2. Suppose (x_n) is a sequence in M converging to x; then Lemma 1.3 and the continuity of ρ imply that $R_j^{(2)}\rho$ is continuous for every $j\geqq 2$, and so by (1.6),

$$\lim_n E^{Q_j}[\rho(Z(x),Z(x_n))]=\lim_n R_j^{(2)}\rho(x,x_n)=0.$$

Hence for fixed j, $(\rho(Z(x), Z(x_n)), n=1,2,\dots)$ is a sequence of non-negative random variables on $(\Gamma, \mathscr{B}_0(\Gamma), Q_j)$, converging in mean to zero. Therefore the sequence converges to zero in probability, proving that Q_j is stochastically continuous on M. $\square$

Let E denote the space of Borel probability measures on the space M, otherwise known as normalized mass distributions on M, with the weak topology, as discussed for example in Bauer (1981), Chap. 7.7.

Definition 1.7. Define sets of functions $V_0, V_1, V_2, \dots$ from E to $\mathbb{R}$ as follows. V_0 is the constant functions, and for $j \geqq 1$, V_j is the set of functions $\varphi: E \to \mathbb{R}$ of the form

$$\psi(\lambda) = \langle \varphi_1, \lambda \rangle \langle \varphi_2, \lambda \rangle \dots \langle \varphi_j, \lambda \rangle, \quad \lambda \in E, \tag{1.7}$$

for some $j \geqq 1$ and some $\varphi_1, \dots, \varphi_j$ in $C(M)$. Define V to be the linear span of $V_0, V_1, V_2, \dots$.

Lemma 1.8. *V is a dense subalgebra of $C(E)$.*

Proof. Use the Stone-Weierstrauss theorem (we omit the details). Similar results are noted by Fleming (1982) and Dawson and Kurtz (1982) for dealing with measure-valued processes.

Lemma 1.9. *For each $j \geqq 1$, and for each function u in $C(M^j)$, the mapping*

$$\lambda \to \langle u, \lambda \otimes \dots \otimes \lambda \rangle = \int_{M^j} u(y_1, \dots, y_j)\, \lambda(dy_1) \dots \lambda(dy_j) \tag{1.8}$$

is in $C(E)$.

Proof. Fix $j \geqq 1$, and let J denote the linear subspace of $C(M^j)$ spanned by the functions v of the form $v(y_1, \dots, y_j) = \varphi_1(y_1)\varphi_2(y_2) \dots \varphi_j(y_j)$, where $\varphi_1, \dots, \varphi_j$ are elements of $C(M)$. It is easy to check that J is a subalgebra of $C(M^j)$ which contains the constant functions and which separates the points of M^j. The Stone-Weierstrauss theorem implies that J is dense in $C(M^j)$. Moreover the weak continuity of $\lambda \to \langle \varphi, \lambda \rangle$ for φ in $C(M)$ implies that (1.8) is in $C(E)$ whenever u is in J.

Given a function u in $C(M^j)$, then denseness of J implies that there exists a sequence (v_n) in J converging to u in the supremum topology. Hence for each λ in E,

$$|\langle u, \lambda \otimes \dots \otimes \lambda \rangle - \langle v_n, \lambda \otimes \dots \otimes \lambda \rangle| \leqq \|u - v_n\|\, \lambda(M)^j = \|u - v_n\|.$$

Thus $\langle v_n, \lambda \otimes \dots \otimes \lambda \rangle$ converges to $\langle u, \lambda \otimes \dots \otimes \lambda \rangle$ uniformly over λ in E. Each v_n satisfies (1.8), and the uniform limit of continuous functions is continuous; hence u satisfies (1.8), as desired. $\square$

Now we will show how a stochastically continuous probability measure Q on function space gives rise in a natural way to a transition probability P on the space E of normalized mass distributions. The idea is that for μ in E and A in $\mathscr{B}(E)$, the probability $P(\mu, A)$ is $Q(\mu \circ Z^{-1} \in A)$. To construct P rigorously, we proceed as follows.

Suppose ψ is in V_j (Definition 1.7); thus

$$\psi(\lambda) = \langle \varphi_1, \lambda \rangle \langle \varphi_2, \lambda \rangle \dots \langle \varphi_j, \lambda \rangle, \quad \lambda \in E \tag{1.9}$$

for some $\varphi_1, \ldots, \varphi_j$ in $C(M)$. If $u(y_1, \ldots, y_j)$ denotes the product $\varphi_1(y_1)\varphi_2(y_2)\ldots\varphi_j(y_j)$, then u is an element of $C(M^j)$. Let $R^{(j)}$ be as in (1.3); thus

$$R^{(j)}u(x_1, \ldots, x_j) = E^Q[\varphi_1(Z(x_1)) \ldots \varphi_j(Z(x_j))]. \tag{1.10}$$

If Q is stochastically continuous on M, then $R^{(j)}u$ belongs to $C(M^j)$ by Lemma 1.3. Hence $\langle R^{(j)}u, \lambda\otimes\ldots\otimes\lambda\rangle$ is well-defined for λ in E.

Proposition 1.10. *Suppose that Q is stochastically continuous on M. Then there exists a positive linear contraction Λ: $V\to C(E)$ such that $\Lambda 1=1$, and such that for ψ as in* (1.9), *and $R^{(j)}u$ as in* (1.10),

$$\begin{aligned} \Lambda\psi(\lambda) &= \langle R^{(j)}u, \lambda\otimes\ldots\otimes\lambda\rangle \\ &= \int_{M^j} E^Q[\varphi_1(Z(x_1)) \ldots \varphi_j(Z(x_j))]\, \lambda(dx_1) \ldots \lambda(dx_j) \end{aligned} \tag{1.11}$$

for $\lambda\in E$.

Proof. For ψ in V_j, define $\Lambda\psi$ by (1.11); also define $\Lambda 1=1$. Lemmas 1.3 and 1.9 combined show that $\Lambda\psi$ is in $C(E)$ for each ψ in V_j. It is clear from (1.11) that Λ is linear on V_j. Hence there is a unique linear extension of Λ to V (see Definition 1.7).

Next we prove that $\Lambda\psi\geqq 0$ for all $\psi\geqq 0$ in V. Since $\Lambda\psi$ is continuous on E, and since the discrete probability measures are dense in E, it suffices to prove that $\Lambda\psi(\mu)\geqq 0$ for all discrete probability measures μ, for $\psi\geqq 0$ in V. For such a μ, and for ψ as in (1.9),

$$\begin{aligned} \Lambda\psi(\mu) &= \int_{M^j} \left\{\int_\Gamma Q(df)\, \varphi_1(f(x_1)) \ldots \varphi_j(f(x_j))\right\} \mu(dx_1) \ldots \mu(dx_j) \\ &= \int_\Gamma Q(df) \left\{\int_{M^j} \varphi_1(f(x_1)) \ldots \varphi_j(f(x_j))\, \mu(dx_1) \ldots \mu(dx_j)\right\}. \end{aligned}$$

(Interchanging the order of integration is allowed because the integral over M^j is really a finite sum, since μ is discrete.)

$$= \int_\Gamma Q(df) \left\{\int_{M^j} \varphi_1(y_1) \ldots \varphi_j(y_j)\, v_f(dy_1) \ldots v_f dy_j)\right\} \tag{1.12}$$

where for f in Γ, v_f is the discrete Borel measure defined as follows:

If $\mu(A) = \sum_{1\leqq i\leqq r} \beta_i 1_A(\mathfrak{z}_i)$, some $\beta_i, \ldots, \beta_r \geqq 0$ with

$$\sum \beta_i = 1, \quad \text{some } \mathfrak{z}_1, \ldots, \mathfrak{z}_r \text{ in } M,$$

then

$$v_f(A) = \mu(f^{-1}(A)) = \sum_{1\leqq i\leqq r} \beta_i 1_A(f(\mathfrak{z}_i)).$$

Equation (1.12) says that

$$\Lambda\psi(\mu) = \int_\Gamma \psi(v_f)\, Q(df) \geqq 0$$

since $\psi\geqq 0$. Then same conclusion holds for all ψ in V, by linearity.

To show that Λ is a contraction, suppose that ψ is in V and $\sup\{\psi(\lambda)\colon \lambda\in E\}\leqq 1$. Then $1-\psi\geqq 0$, and the positivity of Λ implies that $1-\Lambda\psi=\Lambda(1-\psi)\geqq 0$. Hence $\sup\{\Lambda\psi(\lambda)\colon \lambda\in E\}\leqq 1$, as desired. □

Corollary 1.11. *Λ has an extension to a positive linear contraction $\Lambda\colon C(E)\to C(E)$ such that $\Lambda 1=1$. Hence there corresponds a unique transition probability P on E, and P is Feller.*

Proof. Given ψ in $C(E)$, Lemma 1.8 shows that there is a sequence (ψ_m) in V which converges to ψ in the supremum topology. Hence (ψ_m) is a Cauchy sequence in $C(E)$. The fact that $\Lambda\colon V\to C(E)$ is a contraction shows that $(\Lambda\psi_m)$ is a Cauchy sequence in the complete, separable metric space $C(E)$ (note that E itself is compact, separable and metrizable). Hence $(\Lambda\psi_m)$ has a limit in $C(E)$, and we define $\Lambda\psi$ to be that limit. Evidently the mapping $\Lambda\colon C(E)\to C(E)$ inherits the property of being a positive linear operator with $\Lambda 1=1$. The last paragraph of the proof of Proposition 1.10 may be repeated to show that Λ is a contraction. Since E is a compact space with a countable base, we obtain a Feller transition probability P on E as desired. □

The iterates of the transition probability P will be denoted $P_1=P, P_2, \ldots$. In terms of the probability measures Q_m introduced in Definition 1.6, observe that for ψ as in (1.9), and u as in (1.10),

$$(1.13)\qquad \begin{aligned} P_m\psi(\lambda) &= \langle R_m^{(j)} u, \lambda\otimes\ldots\otimes\lambda\rangle \\ &= \int_{M^j} E^{Q_m}[\varphi_1(Z(x_1))\ldots\varphi_j(Z(x_j))]\,\lambda(dx_1)\ldots\lambda(dx_j). \end{aligned}$$

2. Constructing Measurable Versions of Sequences of Random Transformations

This section contains some technical measure-theoretic results; the reader may prefer to skip to Sect. 3, and refer to Sect. 2 as necessary.

Suppose η is a normalized mass distribution on M, and Q is a stochastically continuous probability measure on $\mathscr{B}_0(\Gamma)$. The goal of this section is to construct a sequence $\ldots, W_{-1}, W_0, W_1, \ldots$ of independent random transformations of M, each with law Q, such that all compositions of form $W_{pq}=W_q\circ W_{q-1}\circ\ldots\circ W_{p+1}$, for $p<q$, are jointly measurable, in the sense that

$$(2.1)\qquad (x,\omega)\to W_{pq}(x,\omega) \quad \text{is } \eta\otimes\mathrm{Pr}\text{-measurable into } \mathscr{B}(M),$$

where $(\Omega,\mathscr{F},\mathrm{Pr})$ is the underlying probability space. To achieve this, we need to place a mild condition on the normalized mass distribution η, namely we assume that

$$(2.2)\qquad \text{For } A \text{ in } \mathscr{B}(M), \quad \eta(A)=0 \Rightarrow \int Q(\{f\colon f(x)\in A\})\,\eta(dx)=0.$$

This is equivalent to saying that ηR is absolutely continuous with respect to η (see Definition 1.4), and implies that $\eta R_k \ll \eta$ for all $k\geqq 1$. An invariant measure for R satisfies (2.2), for example.

Let $(\Omega, \mathscr{F}, \Pr)$ denote the product of $\mathbb{Z}$ copies of $(\Gamma, \mathscr{B}_0(\Gamma), Q)$, where $\mathscr{F}$ is the completed product σ-algebra. A sample point in Ω may be written as $\omega = (\ldots, f_{-1}, f_0, f_1, \ldots)$ where each f_p is in Γ. On this probability space there exist independent $(\Gamma, \mathscr{B}_0(\Gamma))$-valued random variables $\ldots, Z_{-1}, Z_0, Z_1, \ldots$, each with law Q, namely $Z_j(\omega) = f_j$. The shift τ: $\Omega \to \Omega$ is characterized by $Z_j(\tau_p \omega) = Z_{j+p}(\omega)$, for p in $\mathbb{Z}$.

Proposition 2.1. *Suppose Q is a stochastically continuous probability measure on $\mathscr{B}_0(\Gamma)$, and η is a normalized mass distribution on M, satisfying* (2.2). *Then, on the probability space $(\Omega, \mathscr{F}, \Pr)$ above, there exist independent $(\Gamma, \mathscr{B}_0(\Gamma))$-valued random variables $W_1, W_2, \ldots$, with the following properties: let*

$$W_{pq}(\omega) = W_q(\omega) \circ W_{q-1}(\omega) \circ \ldots \circ W_{p+1}(\omega), \quad \text{for } 0 \leqq p < q, \ \omega \text{ in } \Omega; \tag{2.3}$$

then (2.1) *holds for all* $0 \leqq p < q$, *and*

$$W_{pq} \text{ has law } Q_{q-p} \text{ (see Definition 1.6).} \tag{2.4}$$

Proof. Step I. The proof is related to that of Doob's result on the existence of measurable versions, as presented by M.M. Rao (1979), p. 179. Since (M, ρ) is compact, there exists for each $n \geqq 1$ an open covering $\{G_{n,i}, i = 1, 2, \ldots, k(n)\}$ of M such that $\operatorname{diam}(G_{n,i}) \leqq \frac{1}{n}$ for each i. Let $H_{n,1} = G_{n,1}$, and let $H_{n,i} = G_{n,i} - \bigcup_{1 \leqq j \leqq i-1} G_{n,i}$ for $i = 2, \ldots, k(n)$. We may assume $H_{n,i}$ is non-empty for each i, and so we select an arbitrary point $y_{n,i}$ in $H_{n,i}$ for each i.

For $n \geqq 1$, define a function L_n: $M \times \Omega \to M$ by

$$L_n(x, \omega) = Z_1(y_{n,i}, \omega) = f_1(y_{n,i}) \quad \text{if } x \in H_{n,i}. \tag{2.5}$$

Observe that L_n is $\mathscr{B}(M) \times \mathscr{F}$-measurable into $\mathscr{B}(M)$. Moreover for any $\varepsilon > 0$,

$$\begin{aligned} &\Pr(\{\omega: \rho(L_n(x, \omega), Z_1(x, \omega)) \geqq \varepsilon\}) \\ &= \sum_i 1_{H_{n,i}}(x) \Pr(\{\omega: \rho(Z_1(y_{n,i}, \omega), Z_1(x, \omega)) \geqq \varepsilon\}) \to 0 \quad \text{as } n \to \infty \end{aligned} \tag{2.6}$$

by the stochastic continuity of Z_1. Hence for every x in M $(L_n(x), n \geqq 1)$ is a Cauchy sequence in probability. Since ρ is a bounded metric, Fubini's theorem implies that

$$\begin{aligned} &\int_{M \times \Omega} \rho((L_n(x, \omega), L_m(x, \omega)) \, \eta \otimes \Pr(dx, d\omega) \\ &= \int_M [\rho(L_n(x), L_m(x))] \, \eta(dx) \end{aligned} \tag{2.7}$$

and the right side tends to zero as $n, m \to \infty$ by (2.6) and the bounded convergence theorem. Thus $(L_n, n \geqq 0)$ is Cauchy in $L^1(\eta \otimes \Pr)$, and hence it converges in $\mu \otimes \Pr$-measure. Therefore there is a subsequence $(L_{n(j)}, j \geqq 1)$ which converges pointwise to a limit, denoted W_1, except on a set J_1 in $M \times \Omega$ such that

$\eta\otimes\Pr(J_1)=0$; moreover J_1 is in $\mathscr{B}(M)\times\mathscr{F}$, since each L_n is $\mathscr{B}(M)\times\mathscr{F}$-measurable. Let $W_1(x,\omega)=Z_1(x,\omega)$ for (x,ω) in J_1.

Using the shift τ: $\Omega\to\Omega$ defined above, define $W_p(\omega)=W_1(\tau_{p-1}\omega)$, for $p=2,3,\ldots$, and define J_p in $M\times\Omega$ by: $J_p=\{(x,\omega)\colon (x,\tau_{p-1}\omega)\in J_1\}$. Then J_p is in $\mathscr{B}(M)\times\mathscr{F}$ and $\eta\otimes\Pr(J_p)=0$, for every $p\geqq 1$. Since each L_n is $\eta\otimes\Pr$-measurable, it follows that W_1, and hence every W_p, is $\eta\otimes\Pr$-measurable into $\mathscr{B}(M)$. Also (2.6) shows that $\Pr(W_p(x)=Z_p(x))=1$ for all x, so $W_p(\cdot,\omega)$ induces the law Q on $\mathscr{B}_0(\Gamma)$.

Step II. Let $H(q)$ be the inductive hypothesis that (2.1) and (2.4) hold for all p such that $0\leqq p<q$. The previous paragraph shows that $H(1)$ is true. Suppose that $H(q)$ is true for some $q\geqq 1$. Observe that

$$\begin{aligned}(2.8)\quad W_{p,q+1}(x,\omega)&=W_{q+1}(W_{pq}(x,\omega),\omega)\\ &=\lim_j L_{n(j)}(W_{pq}(x,\omega),\tau_q\omega),\qquad \text{if } (W_{pq}(x,\omega),\omega)\in J^c_{q+1}.\end{aligned}$$

Now

$$(2.9)\quad L_n(W_{pq}(x,\omega),\tau_q\omega)=\sum_i Z_{q+1}(y_{n,i},\omega)\,1_{H_{n,i}}(W_{pq}(x,\omega))\quad \text{(formal sum)}.$$

By the inductive hypothesis, $\{(x,\omega)\colon W_{pq}(x,\omega)\in H_{n,i}\}$ is $\eta\otimes\Pr$-measurable; also Z_{q+1} is $\mathscr{F}$-measurable, so the left side of (2.9) is an $\eta\otimes\Pr$-measurable function of (x,ω).

To prove that $W_{p,q+1}$ is $\eta\otimes\Pr$-measurable for $p=0,1,\ldots,q$, it suffices by (2.8) to prove that:

$$(2.10)\qquad \eta\otimes\Pr(\{(x,\omega)\colon (W_{pq}(x,\omega),\omega)\in J_{q+1}\})=0.$$

We noted in Step I that J_{q+1} is in $\mathscr{B}(M)\times\mathscr{F}$; moreover the inductive hypothesis for (2.1) shows that the set $\{\ldots\}$ in (2.10) is $\eta\otimes\Pr$-measurable. Hence by Fubini's theorem, the left side is equal to:

$$\begin{aligned}&\int\eta(dx)\Pr(\{\omega\colon W_{pq}(x,\omega)\in\{y\colon \lim_j L_{n(j)}(y,\tau_q\omega)\text{ does not exist}\}\})\\ &=\int\eta(dx)\,[Q_{q-p}\otimes Q(\{(g,h)\colon g(x)\in A(h)\})]\end{aligned}$$

by the inductive hypothesis for (2.4) and the independence of W_{pq} and Z_{q+1}, where for h in Γ

$$A(h)=\{y\colon \lim_j\sum_i h(y_{n(j),i})\,1_{H_{n(j),i}}(y)\text{ does not exist}\}$$

(see [2.5]). Now $A(h)$ is in $\mathscr{B}(M)$ for each h in Γ, and $\eta(A(h))=0$ for $[Q]$-almost all h since J_{q+1} is an $\eta\otimes\Pr$-nullset. So the integral (2.10) is

$$=\int\eta(dx)\int Q(dh)\,Q_{q-p}(\{g\colon g(x)\in A(h)\})=0,$$

since $Q_{q-p}(\{g\colon g(x)\in C\})=0$ when $\eta(C)=0$, by (2.2). This proves that $W_{p,q+1}$ is $\eta\otimes\Pr$-measurable for $p=0,1,\ldots,q$.

Step III. To prove $H(q+1)$, it only remains to prove that for any $x_1, \ldots, x_j$ in M, and any u in $C(M^j)$,

$$E[u(W_{p,q+1}(x_1), \ldots, W_{p,q+1}(x_j))] = R^{(j)}_{q+1-p} u(x_1, \ldots, x_j)$$

(see Definitions 1.4, 1.5). The left side is

$$\begin{aligned} &E[E[u(W_{q+1}(W_{pq}(x_1)), \ldots, W_{q+1}(W_{pq}(x_j))) \mid W_{pq}]] \\ &\quad = E[R^{(j)} u(W_{pq}(x_1), \ldots, W_{pq}(x_j))]. \end{aligned}$$

since W_{q+1} has law Q. By the second part of the inductive hypothesis, this is

$$= R^{(j)}_{q-p}(R^{(j)} u)(x_1, \ldots, x_j) = R^{(j)}_{q+1-p} u(x_1, \ldots, x_j), \qquad \text{as desired.}$$

This completes the induction and the proof. □

Definition 2.2. Suppose Q is stochastically continuous, η satisfies (2.2), and $(\Omega, \mathscr{F}, \Pr)$ is the probability space described above. For $k=0,1,2,\ldots$, define $\eta\otimes\Pr$-measurable maps X_k: $M \times \Omega \to (M, \mathscr{B}(M))$ as follows:

$$\begin{aligned} X_0(x,\omega) &= x, \qquad X_k(x,\omega) = W_{0k}(x,\omega) \\ &= W_k(\omega) \circ W_{k-1}(\omega) \circ \ldots \circ W_1(\omega)(x), \end{aligned} \tag{2.11}$$

$$\begin{aligned} Y_0(x,\omega) &= x, \qquad Y_k(x,\omega) = W_{0k}(x, \tau_{-k-1}\omega) \\ &= X_k(x, \tau_{-k-1}\omega), \qquad \text{for } k \geqq 1. \end{aligned} \tag{2.12}$$

By Proposition 2.1, each X_k and each Y_k can be regarded as an $(\Gamma, \mathscr{B}_0(\Gamma)$-valued random variable with law Q_k.

Lemma 2.3. *For $k \geqq 1$, $X_k(\omega) = W_k(\omega) \circ X_{k-1}(\omega)$: $M \to M$, and*

$$Y_k(\omega) = Y_{k-1}(\omega) \circ W_{-k}(\omega): \ M \to M,$$

where by definition $W_p(\omega) = W_1(\tau_{p-1}\omega)$ for all p in $\mathbb{Z}$.

Proof. The assertion about X is obvious. As for the other,

$$\begin{aligned} Y_k(\tau_{k+1}\omega) &= W_{0k}(\omega) = W_k(\omega) \circ \ldots \circ W_1(\omega) \\ &= W_{k-1}(\tau\omega) \circ \ldots \circ W_1(\tau\omega) \circ W_0(\tau\omega), \end{aligned}$$

since

$$\begin{aligned} W_p(\omega) &= W_1(\tau_{p-1}\omega), \\ &= W_{0,k-1}(\tau\omega) \circ W_0(\tau\omega), \\ &= Y_{k-1}(\tau_{k+1}\omega) \circ W_1(\omega) \\ &= Y_{k-1}(\tau_{k+1}\omega) \circ W_{-k}(\tau_{k+1}\omega). \quad \square \end{aligned}$$

3. The Measure-valued Process and Its Statistical Equilibrium

The first goal of this section is the construction of a measure-valued process $(\mu_n,\ n\geqq 0)$, which is a Markov chain with transition probability P as in Sect. 1. The idea is to take $\mu_0=\eta$, where η satisfies (2.2), and $\mu_n=\eta\circ X_n^{-1}$ (see Definition 2.2). To study the asymptotic behaviour of $(\mu_n,\ n\geqq 0)$, it is helpful to consider another measure-valued process $(\nu_n,\ n\geqq 0)$, Where $\nu_0=\eta$ and $\nu_n=\eta\circ Y_n^{-1}$ (i.e., apply successive random transformations on the right, rather than the left). A convenient and elegant way to specify the random measures μ_n and ν_n will be the following.

Suppose that Q is stochastically continuous, and η satisfies (2.2). Using Definition 2.2 and Fubini's theorem, we see that for all functions φ in $b(M)$, and all $n\geqq 0$, the functions

(3.1) $$\omega\to\int_M \varphi(X_n(x,\omega))\,\eta(dx)=M_n(\varphi,\omega),$$

and

(3.2) $$\omega\to\int_M \varphi(Y_n(x,\omega))\,\eta(dx)=N_n(\varphi,\omega)$$

are $\mathscr{F}$-measurable (recall that $\mathscr{F}$ is complete). For each ω in Ω, the mappings $\varphi\to M_n(\varphi,\omega)$ and $\varphi\to N_n(\varphi,\omega)$ are positive linear functionals on $C(M)$. According to the Riesz representation theorem, there exist (random) probability measures $\mu_n(\cdot,\omega)$ and $\nu_n(\cdot,\omega)$ on $\mathscr{B}(M)$ such that for all φ in $C(M)$,

(3.3) $$\langle\varphi,\mu_n((\cdot,\omega)\rangle=M_n(\varphi,\omega),\qquad \langle\varphi,\nu_n(\cdot,\omega)\rangle=N_n(\varphi,\omega).$$

It follows from this construction that each μ_n and ν_n is $\mathscr{F}$-measurable into the weak topology on E (see Sect. 1); i.e., $(\mu_n,\ n\geqq 0)$ and $(\nu_n,\ n\geqq 0)$ are E-valued random processes on $(\Omega,\mathscr{F},\Pr)$, determined by the equations:

(3.4) $$\begin{aligned}&\langle\varphi,\mu_n\rangle=\langle\varphi\circ X_n(\cdot),\eta\rangle,\\ &\langle\varphi,\nu_n\rangle=\langle\varphi\circ Y_n(\cdot),\eta\rangle,\qquad \varphi\in C(M).\end{aligned}$$

Proposition 3.1. *The process* $(\mu_n,\ n\geqq 0)$ *specified by* (3.1) *and* (3.3) *is a homogeneous Markov chain with respect to the σ-algebras* $\mathscr{F}_0=\{\emptyset,\Omega\}$, $\mathscr{F}_n=\sigma\{W_1,\dots,W_n\}$, $n\geqq 1$, *and with the transition probability P defined in Corollary 1.11.*

Remark 3.2. There is no reason why $(\nu_n,\ n\geqq 0)$ should be Markov. However $(\mu_n,\ n\geqq 0)$ and $(\nu_n,\ n\geqq 0)$ are related by means of the shift $\tau\colon\Omega\to\Omega$; (2.12) and (3.4) show that $\nu_n(\omega,\cdot)=\mu_n(\tau_{-n-1}\omega,\cdot)$. For each fixed n, μ_0 and ν_n have the same law.

Proof. It suffices to prove that for any integers $n>m\geqq 0$ and any ψ in $b(E)$, we have

(3.5) $$E[\psi(\mu_n)\,|\,\mathscr{F}_m]=P_{n-m}\psi(\mu_m),\qquad [\Pr]\text{-a.s.}$$

Actually it suffices to prove (3.5) for all ψ in V of the form (1.7), since the extension to $C(E)$ and to $b(E)$ is routine. In this case, for $m\geqq 1$,

$$\begin{aligned}E[\psi(\mu_n)|\mathscr{F}_m] &= E[\int \varphi_1\, d\mu_n \ldots \int \varphi_k\, d\mu_n | W_1, \ldots, W_m] \\ &= E[\int \varphi_1(X_n(\cdot))\, d\eta \ldots \int \varphi_k(X_n(\cdot))\, d\eta | W_1, \ldots, W_m] \\ &= E[\int \varphi_1 \circ W_n \circ \ldots \circ W_{m+1} \circ X_m(\cdot)\, d\eta \ldots \int \varphi_k \\ &\quad \circ W_n \circ \ldots \circ W_{m+1} \circ X_m(\cdot)\, d\eta | W_1, \ldots, W_m] \\ &= E[\int \varphi_1 \circ W_n \circ \ldots \circ W_{m+1}(\cdot)\, d\mu_m \ldots \int \varphi_k \\ &\quad \circ W_n \circ \ldots \circ W_{m+1}(\cdot)\, d\mu_m | W_1, \ldots, W_m]\end{aligned}$$

using the measurability assertions of Proposition 2.1, and the definition of μ_m. Since $W_{m+1}, \ldots, W_n$ are independent of $W_1, \ldots, W_m$, (1.13) implies that this is $P_{n-m}\psi(\mu_m)$ [Pr]-a.s. □

Proposition 3.3. *Suppose that the one-point transition probability R (see Definition 1.4) has an invariant probability measure π.*

(i) *If $\nu_0 = \eta = \pi$ in* (3.2) *and* (3.3), *then* $(\nu_n, n \geqq 0)$ *converges almost surely in the weak topology to a random measure ν_∞, as n tends to infinity.*

(ii) *If $\mu_0 = \eta = \pi$ in* (3.1) *and* (3.3), *then* $(\mu_n, n \geqq 0)$ *converges in law to ν_∞ in the weak topology, as n tends to infinity, meaning that* $E[\psi(\mu_n)] \to E[\psi(\nu_\infty)]$ *for all ψ in $C(E)$.*

(iii) *Let q be the probability measure on $\mathscr{B}(E)$ given by:* $q(A) = \Pr(\{\omega: \nu_\infty(\omega, \cdot) \in A\})$ *for A in $\mathscr{B}(E)$. Then $qP = q$, i.e., q is a P invariant measure.*

Remarks. The random measure ν_∞ is called the *statistical equilibrium* for the discrete-time stochastic flow $\{X_k, k \geqq 0\}$. Assertion (i) was noted by Le Jan (1984) in a special case. In effect it is a statement about the limit of the nonlinear random transformations (Y_n), and thus is a form of random ergodic theorem; compare Kifer (1985), Chap. I, Corollary 2.2.

A kind of invariance similar to (iii) was studied by Liggett (1978) in the case where the motions of any k distinct points are independent.

Proof. Notice first that π satisfies (2.2), so the results of Sect. 2 apply.

(i) Take an arbitrary φ in $C(M)$, and consider the process $(N_n(\varphi), n \geqq 0)$ defined in (3.2). Evidently $|N_n(\varphi)| \leqq \|\varphi\|$, so the process is bounded. Define an increasing sequence of σ-algebras contained in $\mathscr{F}$ as follows: $\mathscr{G}_0 = \{\emptyset, \Omega\}$, and $\mathscr{G}_n = \sigma\{W_{-1}, \ldots, W_{-n}\}$ for $n \geqq 1$. Then for $n \geqq 0$, (3.2) and (3.3) say that

$$\begin{aligned}E[N_{n+1}(\varphi)|\mathscr{G}_n] &= E[\langle \varphi \circ Y_n \circ W_{-n-1}(\cdot), \pi\rangle | \mathscr{G}_n] \\ &= E[\langle \varphi \circ Y_n(\cdot), \pi R\rangle | \mathscr{G}_n]\end{aligned}$$

using the Lemma 2.3, Proposition 2.1, and the independence of Y_n and W_{-n-1}. This equals $\langle \varphi \circ Y_n(\cdot), \pi\rangle = N_n(\varphi)$, [Pr]-a.s. Thus $(N_n(\varphi), n \geqq 0)$ is a bounded martingale, and therefore has a limit almost surely as n tends to infinity, denoted $N_\infty(\varphi)$.

Since M has a countable base, there exists a countable collection of functions $\{\varphi_1, \varphi_2, \ldots\}$ such that a sequence (λ_n) in E converges weakly to λ in E if and only if $\langle \varphi_i, \lambda_n\rangle$ converges to $\langle \varphi_i, \lambda\rangle$ for every $i \geqq 1$. We see that for each $i \geqq 1$, there is a nullset K_i in Ω such that

$$\lim_n \langle \varphi, \nu_n(\omega, \cdot)\rangle = N_\infty(\varphi_i, \omega), \quad \text{for } \omega \text{ outside } K_i.$$

Hence $\lim_n \nu_n(\omega, \cdot)$ exists in E for all ω outside the nullset $K = \bigcup_{i \geq 1} K_i$, which verifies (i).

(ii) We saw in Remark 3.2 that for each fixed n, μ_n and ν_n have the same law. However $(\nu_n, n \geq 0)$ converges in law as n tends to infinity, since it converges almost surely by part (i). This proves (ii).

(iii) First we shall verify that

$$\langle \psi, q \rangle = E[\psi(\nu_\infty)] \quad \text{for all } \psi \text{ in } b(E). \tag{3.6}$$

Observe that if $\psi = 1_A$ for A in $\mathscr{B}(E)$, then both sides of (3.6) equal $q(A)$. By linearity, (3.6) holds for all simple measurable functions. However every ψ is $b(E)$ is the pointwise limit of an increasing sequence of simple measurable functions, and Lebesgue's Monotone Convergence Theorem establishes (3.6) for the limit.

To show that $qP = q$, it suffices to show that $\langle P\psi, q \rangle = \langle \psi, q \rangle$ for all ψ in $C(E)$, since $\langle \psi, qP \rangle = \langle P\psi, q \rangle$. By (3.6) it suffices to show that

$$E[\psi(\nu_\infty)] = E[P\psi(\nu_\infty)], \quad \psi \in C(E). \tag{3.7}$$

Now

$$
\begin{aligned}
E[\psi(\nu_\infty)] &= E[\psi(\lim_n \nu_n)], \\
&= E[\lim_n \psi(\nu_n)], \text{ by continuity of } \psi. \\
&= \lim_n E[\psi(\nu_n)], \text{ by dominated convergence}, \\
&= \lim_n E[\psi(\mu_n)], \text{ since } \mu_n \text{ and } \nu_n \text{ have the same law}, \\
&= \lim_n E[\psi(\mu_{n+1})], \\
&= \lim_n E[E[\psi(\mu_{n+1}) \mid \mu_n]], \\
&= \lim_n E[P\psi(\mu_n)] \text{ by (3.5)}, \\
&= \lim_n E[P\psi(\nu_n)], \text{ since } \mu_n \text{ and } \nu_n \text{ have the same law}, \\
&= E[P\psi(\nu_\infty)] \text{ by the reasoning above.}
\end{aligned}
$$

This verifies (3.7) and completes the proof. □

Remark 3.4. The convergence described in Proposition 3.3, (i) and (ii), may still occur even if the initial state η is not R-invariant. For example, suppose that η satisfies (2.2) (i.e., $\eta R \leq \eta$), and that for some $c > 0$ and $\beta > 0$, $\|((\eta - \pi) \otimes (\eta - \pi)) S_n\| \leq c n^{-1-\beta}$, for all n (absolute variation norm) where π is an invariant measure for the one-point motion (i.e. $\pi R = \pi$), and S is the two-point transition probability (Definition 1.4). (Stochastic continuity of Q remains in force.) If $\nu_0 = \eta = \mu_0$, then ν_n converges almost surely, and μ_n converges in law, to the same random measure ν_∞, as when $\nu_0 = \pi = \mu_0$.

The preceding condition will hold, for example if S satisfies a Doeblin condition (see Doob, 1953, p. 197). The author thanks P. Baxendale for giving a proof of this result (which is omitted here).

4. Conditions for the Measure-valued Process to be Ergodic

Recall the transition probabilities R and S from Definition 1.4, and P from Corollary 1.11. A transition probability is called ergodic if it has a unique invariant probability measure.

Theorem 4.1. *Suppose Q is a probability measure on $\mathscr{B}_0(\Gamma)$ such that*

(a) *Q is stochastically continuous on M (Definition 1.1), and*

(b) *the one-point transition probability R is ergodic, with a unique invariant probability measure π.*

Then (i), (iii) *and* (iii) *are equivalent:*

(i) *S is ergodic.*

(ii) $\lim_n \frac{1}{n} \sum_{j=0}^{n-1} S_j \rho(x, y) = 0$, *$x$, y in M. (ρ is the distance function).*

(iii) *P is ergodic.*

If (i)–(iii) *hold, then the unique invariant probability measures θ and m, of S and P respectively, are as follows:*

θ is concentrated on the diagonal, and is determined by:

$$\theta(G \times H) = \pi(G \cap H) \quad \text{for } G, H \text{ in } \mathscr{B}(M). \tag{4.1}$$

m is concentrated on E_δ, the set of degenerate normalized mass distributions, and (denoting by δ_x the point mass at x),

$$m(\{\delta_x \colon x \in G\}) = \pi(G), \quad \text{for } G \text{ in } \mathscr{B}(M). \tag{4.2}$$

Remarks. 1. Of course, the measures θ and m are invariant, for S and P respectively, even if (i)–(iii) do not hold. The calculation is omitted.

2. The diagonal in M^2 is an absorbing set for the Markov chain with transition probability S, and E_δ is an absorbing set for the measure-valued Markov chain.

3. In the case where Q is concentrated on the smooth diffeomorphisms of a compact manifold M, S is ergodic if the maximum Lyapounov exponent is negative; see Le Jan (1985a).

Proof. "(i)$\Rightarrow$(iii)". Assume S is ergodic. P is a Feller transition probability on the compact metrizable space E; according to Revuz (1984), Ch. 4, Ex. 3.14, ergodicity of P is equivalent to the following assertion:

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} P_j \psi(\lambda) \quad \text{is a constant, for each } \psi \text{ in } C(E). \tag{4.3}$$

Since the mapping $\psi \to \frac{1}{n}\sum_{j=0}^{n-1} P_j \psi$ is a contraction for every n, it is clearly sufficient to prove (4.3) for every ψ in V; in fact, by linearity, we need only consider ψ in V_k, for $k=1,2,\ldots$.

The case $k \geqq 1$. Suppose ψ is in V_1, i.e., $\psi(\lambda)=\int \varphi \, d\lambda$ for some φ in $C(M)$. Define real numbers $(a_n(\lambda), n \geqq 0)$ by

$$a_n(\lambda)=\left\langle \frac{1}{n}\sum_{j=0}^{n-1} R_j \varphi, \lambda \right\rangle = \frac{1}{n}\sum_{j=0}^{n-1} P_j \psi(\lambda). \tag{4.4}$$

Assumption (b) and the mean ergodic theorem for R implies that for all x in M,

$$\lim_n a_n(\lambda)=\int \varphi \, d\pi \qquad \text{for all } \lambda \text{ in } E; \tag{4.5}$$

this verifies (4.3) for ψ in V_1.

The case $k \geqq 2$. Suppose $k=m \geqq 2$, and ψ is in V_m, i.e., of the form $\psi(\lambda) = \langle \varphi_1, \lambda\rangle \ldots \langle \varphi_m, \lambda\rangle$ where $\varphi_1, \ldots, \varphi_m$ are in $C(M)$. By (1.13), for discrete measures λ and μ in E,

$$\left|\frac{1}{n}\sum_{j=0}^{n-1} P_j \psi(\lambda) - \frac{1}{n}\sum_{j=0}^{n-1} P_j \psi(\mu)\right| \tag{4.6}$$

$$= \left|\frac{1}{n}\sum_{j=0}^{n-1} E^{Q_j}[c_1(\lambda)\ldots c_m(\lambda) - c_1(\mu)\ldots c_m(\mu)]\right|,$$

where the random variables $(c_i(\lambda), i=1,\ldots,m)$ are given by $c_i(\lambda)=\langle \varphi_i \circ Z(\cdot), \lambda\rangle = \int \varphi_i(Z(x))\,\lambda(dx)$. (Interchanging the order of integration in (4.6) is allowed because λ and μ are discrete, and so the integral over M^m is really a finite sum.) Notice that $|c_i(\lambda)| \leqq \|\varphi_i\|$. The expression (4.6) is equal to

$$\Big|\frac{1}{n}\sum_{j=0}^{n-1} E^{Q_j}[c_1(\lambda)\,c_2(\lambda)\ldots c_m(\lambda) - c_1(\mu)\,c_2(\lambda)\ldots c_m(\lambda)$$
$$+c_1(\mu)\,c_2(\lambda)\ldots c_m(\lambda) - c_1(\mu)\,c_2(\mu)\,c_3(\lambda)\ldots c_m(\lambda) + \ldots$$
$$+c_1(\mu)\ldots c_{m-1}(\mu)\,c_m(\lambda) - c_1(\mu)\ldots c_m(\mu)]\Big|.$$

For $i=1,2,\ldots,m$, let $L_i = \|\varphi_1\| \ldots \|\varphi_{i-1}\| \, \|\varphi_{i+1}\| \ldots \|\varphi_m\|$ (i.e., product of the norms, except the i^{th}). The previous expression is

$$\leqq L_1 \frac{1}{n}\sum_{j=0}^{n-1} E^{Q_j}|c_1(\lambda)-c_1(\mu)| + \ldots + L_m \frac{1}{n} E^{Q_j}|c_m(\lambda)-c_m(\mu)|.$$

Now

$$\frac{1}{n}\sum_{j=0}^{n-1} E^{Q_j}|c_i(\lambda)-c_i(\mu)|$$

$$=\frac{1}{n}\sum_{j=0}^{n-1} E^{Q_j}|\int[\varphi_i(Z(x))-\varphi_i(Z(y))]\,\lambda(dx)\,\mu(dy)|$$

$$\leqq \int\left\{\frac{1}{n}\sum_{j=0}^{n-1} E^{Q_j}[u_i(Z(x),Z(y))]\right\}\lambda(dx)\,\mu(dy)$$

where $u_i(x,y)=|\varphi_i(x)-\varphi_i(y)|$. (Interchanging the order of integration is allowed because λ and μ are discrete.) Returning to (4.6), we have now shown that

$$\left|\frac{1}{n}\sum_{j=0}^{n-1} P_j\psi(\lambda)-\frac{1}{n}\sum_{j=0}^{n-1} P_j\psi(\mu)\right| \qquad (4.7)$$

$$\leqq \sum_{i=1}^{m} L_i\int\left\{\frac{1}{n}\sum_{j=0}^{n-1} S_j u_i(x,y)\right\}\lambda(dx)\,\mu(dy).$$

Both sides of this inequality are continuous in λ and μ, and discrete measures are dense in E; hence (4.7) holds for all λ and μ in E.

The assumption (i) and the Mean Ergodic Theorem imply that

$$\lim_n \frac{1}{n}\sum_{j=0}^{n-1} S_j u_i(x,y)=\int u_i(x,y)\,\theta(dx,dy)=\int u_i(x,x)\,\pi(dx)=0.$$

Using Dominated Convergence, we see that every term on the right side of (4.7) converges to zero as n tends to infinity. Hence the expression (4.6) converges to zero for all λ and μ in E, which verifies (4.3) for all ψ in V_m, as desired.

"(iii)$\Rightarrow$(ii)". Assume that P is ergodic. Define ψ in $C(E)$ as follows:

$$\psi(\lambda)=\int\rho(\mathfrak{z},w)\,\lambda(d\mathfrak{z})\,\lambda(dw). \qquad (4.8)$$

(Continuity of ψ follows from Lemma 1.9.) Fix x and y in M, and define $v=(\delta_x+\delta_y)/2$ in E. Then $\psi(v)=\rho(x,y)/2$, since $\rho(x,x)=\rho(y,y)=0$. Moreover $S_j\rho(x,y)=2P_j\psi(v)$ for all $j\geqq 1$, by (1.13). Using the Mean Ergodic Theorem for P, and (4.8), we have

$$\lim_n \frac{1}{n}\sum_{j=0}^{n-1} S_j\rho(x,y)=2\lim_n \frac{1}{n}\sum_{j=0}^{n-1} P_j\psi(v)$$

$$=2\int\psi(\lambda)\,m(d\lambda)=2\int\psi(\delta_x)\,\pi(dx)=0,$$

which verifies (ii).

"(ii)$\Rightarrow$(i)". Assume that (ii) holds. According to the criterion mentioned in (4.3), S is ergodic if:

(4.9) $$\lim_n \frac{1}{n}\sum_{j=0}^{n-1} S_j u(x,y) \quad \text{is a constant,}$$

for each u in $C(M^2)$.

Given u in $C(M^2)$, define φ in $C(M)$ by: $\varphi(\mathfrak{z})=u(\mathfrak{z},\mathfrak{z})$. By uniform continuity, there exists a constant α such that $|u(\mathfrak{z},w)-u(\mathfrak{z},\mathfrak{z})|\leqq\alpha\,\rho(w,\mathfrak{z})$, for all $w,\mathfrak{z}$ in M. Hence

$$\left|\frac{1}{n}\sum_{j=0}^{n-1} S_j u(x,y)-\frac{1}{n}\sum_{j=0}^{n-1} R_j\varphi(x)\right|$$
$$\leqq\frac{1}{n}\sum_{j=0}^{n-1} E^{Q_j}|u(Z(x),Z(y))-u(Z(x),Z(x))|$$
$$\leqq\frac{\alpha}{n}\sum_{j=0}^{n-1} S_j\rho(x,y).$$

By assumption (ii), the last expression converges to zero as n tends to infinity. Hence

$$\lim_n \frac{1}{n}\sum_{j=0}^{n-1} S_j u(x,y)=\lim_n \frac{1}{n}\sum_{j=0}^{n-1} R_j\varphi(x)=\langle\varphi,\pi\rangle$$

by the Mean Ergodic Theorem, which verifies (4.9). Hence (i) holds. □

Corollary 4.2. *Suppose the two-point transition probability S has an invariant measure concentrated off the diagonal of M^2, and that (a) and (b) hold. Then the measure-valued chain $\{\mu_n, n\geqq 0\}$ introduced in Sect. 3, with $\mu_0=\pi$, has a statistical equilibrium ν_∞ which is not a random Dirac measure.*

Example 4.3. Let M denote the one-point compactification of $\mathbb{R}^3$, and let i: $M\to S^3$ be the inverse of the stereographic projection onto the unit 3-sphere.

The metric ρ on M we take to be the pullback of the usual Riemannian metric on S^3; thus the restriction of ρ to $\mathbb{R}^3$ is equivalent to the Euclidean metric on $\mathbb{R}^3$, while $\rho(x_n,y_n)\to 0$ if $|x_n|\to\infty$ and $|y_n|\to\infty$.

Let $\{X_{st},\ 0\leqq s\leqq t<\infty\}$ be any time-homogeneous pure stochastic flow on $\mathbb{R}^3$ with the following properties:

(4.10) For each fixed x in $\mathbb{R}^3$, $\{X_{0t}(x)-x,\ t\geqq 0\}$ is transient in $\mathbb{R}^3$.

(4.11) For each fixed $t>0$, and $\varepsilon>0$, and x in M $P(\rho(X_{0t}(x),X_{0t}(y))\geqq\varepsilon)\to 0$ as $y\to x$ in M where we define $X_{0t}(\infty,\omega)=\infty$ for all $t\geqq 0$ and ω in Ω.

(4.12) For each fixed x and y in $\mathbb{R}^3$, with probability one $|X_{0t}(x)-X_{0t}(y)|\to 0$ or to ∞, as $t\to\infty$ (typically $P(\lim_{t\to\infty}|X_{0t}(x)-X_{0t}(y)|=0)$ is a function of x and y).

In Darling (1988, Appendix B) an example is given of a class of isotropic stochastic flows in $\mathbb{R}^3$ which satisfy these conditions, and where almost all the mappings $X_{0t}(\omega)$: $M\to M$ are discontinuous, although the stochastic continuity condition (4.11) holds (see Darling, 1987a, Sect. 14, for the proof). Indeed for distinct x and y, there is a positive probability that the trajectories from x and y meet and coalesce in finite time.

The apparatus of this paper may be used to determine the possible normalized statistical equilibria of such a stochastic flow.

Let Q denote the law of X_{01}. Since the one-point motions are transient in $\mathbb{R}^3$ by (4.10), it follows that the only invariant measure for the one-point motion is $\pi=\delta_\infty$ (unit mass at infinity); thus the one-point motion on M is ergodic. Assumption (4.12) implies that $\rho(X_{0t}(x), X_{0t}(y))\to 0$ almost surely as t tends to infinity (through the integers), for each x, y in M; this verifies that condition (ii) of Theorem 4.1 holds.

Theorem 4.1 proves that the only random normalized mass distribution whose distribution is invariant under the action of the discrete-time stochastic flow on M is the degenerate normalized mass distribution δ_∞. By the homogeneity of the original continuous-time stochastic flow, it follows that there does not exist any *normalized* statistical equilibrium for the original stochastic flow $\{X_{st}, 0\leqq s\leqq t<\infty\}$ on $\mathbb{R}^3$. (The non-existence of *unnormalized* statistical equilibria for certain isotropic stochastic flows on $\mathbb{R}^3$ is shown in Darling and Le Jan, 1988.)

Acknowledgements. I thank Richard Arratia, Peter Baxendale, and Yuri Kifer for their assistance. I also thank the National Science Foundation for its financial support.

References

Arnold, L.: In: Arnold, L. (ed.) Workshop on Lyapounov exponents, Bremen 1984. (Lect. Notes Math., vol. 1186, Berlin Heidelberg New York: Springer 1986

Arratia, R.: Coalescing Brownian motions on the line. Ph.D. thesis, University of Wisconsin-Madison (1979)

Arratia, R.: Coalescing Brownian motions, and the voter model on $\mathbb{Z}$. U.S.C. Preprint 1984

Bauer, H.: Probability theory and the elements of measure theory. 2nd edn. London: Academic Press 1981

Baxendale, P.: Asymptotic behaviour of stochastic flows of diffeomorphisms. Probab. Th. Rel. Fields **73**, 51–85 (1986)

Darling, R.W.R.: Constructing non-homeomorphic stochastic flows. Mem. Am. Math. Soc. **376** (in press)

Darling, R.W.R.: Rate of growth of the coalescent set. Stochastics (to appear) (1988)

Darling, R.W.R., Le Jan, Y.: The statistical equilibrium of an isotropic stochastic flow with negative Lyapounov exponents is trivial. (Séminaire de Probabilités XXII) Berlin Heidelberg New York: Springer 1988

Dawson, D.A., Kurtz, T.G.: Applications of duality to measure-valued diffusion processes. In: Fleming, W.H., Gorostiza, L.G., (eds.) (Lecture Notes Cont. Inf. Sci., vol. 42, pp. 91–105) Berlin Heidelberg New York: Springer 1982

Doob, J.L.: Stochastic processes. New York: Wiley 1953

Fife, P.C.: Mathematical aspects of reacting and diffusing systems. (Lect. Notes Biomath. 28) Berlin Heidelberg New York: Springer 1979

Fleming, W.H.: Nonlinear semigroup for controlled partially observed diffusions. SIAM J. Control Optimization **20**, 286–301 (1982)

Harris, T.E.: Coalescing and non-coalescing stochastic flows in $\mathbb{R}^1$. Stochastic Processes Appl. **17**, 187–210 (1984)

Kifer, Y.: Ergodic theory of random transformations. Basel: Birkhauser 1985

Le Jan, Y.: Equilibre et exposants de Lyapunov de certains flots browniens. C.R. Acad. Sci. Paris Ser. I 298, 361–364 (1984)

Le Jan, Y.: Hausdorff dimension for the statistical equilibrium of stochastic flows. (Lect. Notes Math., vol. 1158) Berlin Heidelberg New York: Springer 1985a

Le Jan, Y.: On isotropic Brownian motions. Z. Wahrscheinlichkeitstheor. Verw. Geb. **70**, 609–620 (1985b)

Le Jan, Y.: Equilibre statistique pour les produits de difféomorphismes aléatoires indépendants. Ann. Inst. Henri Poincaire, Nouv. Ser. Sect. B **23**, 111–120 (1987)

Liggett, T.M.: Random invariant measures for Markov chains, and independent particle systems. Z. Wahrscheinlichkeitstheor. Verw. Geb. **45**, 297–313 (1978)

Nelson, E.: Regular probability measures on function spaces. Ann. Math. **69**, 630–643 (1959)

Rao, MM.: Stochastic processes and integration. Alphen aan den Rijn: Sijthoff and Noordhoff 1979

Revuz, D.: Markov chains. 2nd edn. Amsterdam New York: North-Holland 1984

Received August 4, 1986; received in revised form October 19, 1987