Nonlinearity **16** (2003) 403–410 PII: S0951-7715(03)27158-9

Uniqueness in *g*-measures

Örjan Stenflo[1]

Department of Mathematics, Stockholm University, SE-10691 Stockholm, Sweden

E-mail: stenflo@math.su.se

Received 24 July 2001, in final form 29 October 2002
Published 18 December 2002
Online at stacks.iop.org/Non/16/403

Recommended by P Gaspard

Abstract
Using coupling techniques extending ideas from Harris (1955 *Pacific J. Math.* **5** 707–24), we prove uniqueness in g-measures and give estimates of the rates of convergence for the associated Markov chains, for strictly positive continuous g-functions under a weak regularity condition. Our regularity condition is weaker than the earlier weakest known conditions for uniqueness (Harris T E 1955 *Pacific J. Math.* **5** 707–24; Iosifescu M and Spătaru A 1973 *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **27** 195–214; Comets F *et al* 2002 *Ann. Appl. Probab.* **12** 921–43). As a consequence of our method, we obtain sharper bounds on the rates of convergence also in cases when more restrictive regularity conditions are satisfied, and thus in particular, we extend results by Bressaud *et al* (1999 *Electron. J. Probab.* **4** 19).

Mathematics Subject Classification: 28A80, 37H99, 60J05

1. Preliminaries and statements of the results

Let $\Sigma_N := \{1, 2, \ldots, N\}^{\mathbb{N}}$ and introduce a topology on Σ_N induced by the metric

$$\rho(x, y) := \begin{cases} 2^{-n}, & \text{if } x \text{ and } y \text{ differ for the first time in the } n\text{th digit} \\ 0, & \text{if } x = y. \end{cases}$$

The space (Σ_N, ρ) is a compact metric space.

For $j \in \{1, 2, \ldots, N\}$ and $x = x_1x_2\cdots \in \Sigma_N$, let $jx = jx_1x_2\cdots$. Consider a continuous function $g : \Sigma_N \to (0, \infty)$, and suppose that g is normalized in the sense that

$$\sum_{j=1}^{N} g(jx) = 1, \qquad \text{for any } x \in \Sigma_N. \tag{1}$$

(Such a function g is called a (strictly positive, continuous) g-function, see [13].)

[1] Also at: Mathematical Sciences Institute, Australian National University, Canberra, ACT 0200, Australia.

Define $p_j(x) = g(jx)$, and $w_j(x) = jx$. Then $\{(\Sigma_N, \rho);\ w_j(x),\ p_j(x), j \in \{1, 2, \ldots, N\}\}$ is an iterated function system (IFS) with place-dependent probabilities. Note that the maps w_j are contractions, and that this system can be represented by the function g. Place-dependent random iterations according to this system generates a Markov chain with transfer operator $T_g : C(\Sigma_N) \to C(\Sigma_N)$ defined for $f \in C(\Sigma_N)$ by

$$T_g f(x) = \sum_{j=1}^{N} p_j(x) f(w_j(x)),$$

where $C(\Sigma_N)$ denotes the set of real-valued continuous functions on Σ_N. These kinds of Markov chains have been studied under the name 'chains of infinite order'. (The reason for this name is that the projection on the first coordinate of the generated Markov chain on Σ_N typically forms a chain with complete connections, i.e. heuristically expressed, an 'n-step Markov chain with $n = \infty$'.) For accounts on the history of this topic (see [10, 12]).

Let $M(\Sigma_N)$ denote the set of Borel probability measures on Σ_N and let $T_g^\star : M(\Sigma_N) \to M(\Sigma_N)$ denote the adjoint operator of the linear operator T_g, i.e. $T_g^\star$ is defined by requiring that $\int_{\Sigma_N} f \,\mathrm{d}T_g^\star \nu = \int_{\Sigma_N} T_g f \,\mathrm{d}\nu$, for any $\nu \in M(\Sigma_N)$, and $f \in C(\Sigma_N)$.

It is a standard application of the Schauder–Tychonoff fixpoint theorem to prove that there exists at least one g-measure, i.e. probability measure μ such that $T_g^\star \mu = \mu$.

Note that a g-measure is the same as an invariant probability measure for the associated IFS or stationary probability measure for the associated Markov chain.

If we define $\phi(x) = \log g(x)$ and let θ denote the shift map (i.e. $\theta(x_1 x_2 \cdots) = x_2 x_3 \cdots$), we see that the transfer operator can be written as

$$T_g f(x) = \sum_{j=1}^{N} p_j(x) f(w_j(x)) = \sum_{j=1}^{N} \mathrm{e}^{\phi(jx)} f(jx) = \sum_{y \in \theta^{-1} x} \mathrm{e}^{\phi(y)} f(y). \tag{2}$$

The transfer operator, sometimes also called the Ruelle–Perron–Frobenius operator, occurs naturally in statistical physics but the normalization condition (1) is not so natural. The right-hand version of (2) is how this operator is most commonly expressed in the thermodynamic formalism literature.

If g is assumed to be Hölder-continuous then there is a unique g-measure, μ, and $T_g^{\star n} \nu$ converges with exponential rate to μ, for any $\nu \in M(\Sigma_N)$ (see [2]).

For a uniformly continuous function $g : \Sigma_N \to (0, \infty)$, define the modulus of uniform continuity

$$\Delta_g(t) = \sup\{g(x) - g(y) : \rho(x, y) < t\}.$$

Walters, [16], proved that if

$$\sum_{k=1}^{\infty} \Delta_g(2^{-k}) < \infty, \tag{3}$$

then there is a unique g-measure.

(The results in [2, 16] are more general and also covers non-necessarily normalized cases.)

Condition (3) means that g is Dini-continuous. This condition appeared for the first time in this context already in Doeblin and Fortet [7].

The Dini-condition (posed on ϕ) is usually referred to as 'summable variation' in the thermodynamic formalism literature. Observe that g is Dini-continuous iff ϕ is Dini-continuous since g is assumed to be strictly positive and continuous and thus bounded away from zero. Observe also that Hölder-continuity is more restrictive than Dini-continuity.

By using ideas from Harris [9], Berbee [1] proved uniqueness in g-measures for strictly positive continuous g-functions satisfying the condition

$$\sum_{n=1}^{\infty} \exp\left(-\sum_{m=1}^{n} \Delta_{\phi}(2^{-m})\right) = \infty. \tag{4}$$

Berbee's condition (4) is clearly weaker than the Dini-condition (3). Harris [9] considered the condition:

$$\sum_{n=l}^{\infty} \prod_{m=l}^{n} \left(1 - \frac{N}{2}\Delta_g(2^{-m})\right) = \infty, \qquad \text{for some } l \geqslant 1. \tag{5}$$

It is easy to check, in the case $N = 2$, that condition (5) is weaker than Berbee's condition (4). For general N, if $g(x) < 2/N$ in a neighbourhood of points where g has its maximal oscillation, then condition (5) is weaker than condition (4), and it is easy to construct strictly positive g-functions such that condition (5) holds but not condition (4). Conversely if $g(x) > 2/N$ in the most oscillating region then condition (4) is weaker than condition (5). (Note that $g(x) > 2/N$ is impossible in the case $N = 2$.)

Remark 1. In [9], Harris gave an incomplete proof of uniqueness in g-measures under condition (5). Harris proved uniqueness under this condition in the case $N = 2$.

It is not straightforward to extend the proof in [9] to general N (cf Iosifescu and Spătaru [11] and Iosifescu and Grigorescu [10], p 282.) A complete proof of uniqueness in g-measures under Harris condition (5) follows from theorem 1 and proposition 1.

Note added in proof. Recently it has come to the author's attention that this also follows from Coelho and Quas [5]. The author is grateful to Anders Öberg for pointing out this reference.

Iosifescu and Spătaru [11] (see also [10]) presented a condition for uniqueness in g-measures equivalent to

$$\sum_{n=1}^{\infty} \prod_{m=1}^{n} isc_g(2^{-m}) = \infty, \tag{6}$$

where

$$isc_g(2^{-m}) = \inf_{\rho(x,y)\leqslant 2^{-m}} \sum_{j=1}^{N} \left(\min\left(\sum_{i=1}^{j} g(ix), \sum_{i=1}^{j} g(iy)\right) - \max\left(\sum_{i=1}^{j-1} g(ix), \sum_{i=1}^{j-1} g(iy)\right)\right)^{+}.$$

Comets *et al* [6] considered the condition

$$\sum_{n=1}^{\infty} \prod_{m=1}^{n} cff_g(2^{-m}) = \infty, \tag{7}$$

where

$$cff_g(2^{-m}) = \inf_{1\leqslant j_l\leqslant N, 1\leqslant l\leqslant m-1} \sum_{i=1}^{N} \inf_{y\in\Sigma_N} g(ij_1\cdots j_{m-1}y).$$

In this paper, we are going to prove uniqueness in g-measures under the condition:

$$\sum_{n=1}^{\infty} \prod_{m=1}^{n} osc_g(2^{-m}) = \infty, \tag{8}$$

where

$$osc_g(2^{-m}) = \inf_{\rho(x,y)\leqslant 2^{-m}} \sum_{j=1}^{N} \min(g(jx), g(jy)),$$

($= 1 - \frac{1}{2}\sup_{\rho(x,y)\leqslant 2^{-m}} \sum_{j=1}^{N} |g(jy) - g(jx)|$, see the proof of proposition 1).

The fact that condition (8) is indeed weaker than the other conditions for uniqueness in g-measures is a consequence of the following proposition.

Proposition 1. *We have*

$$1 - \frac{N}{2}\Delta_g(2^{-n}) \leqslant osc_g(2^{-n}), \tag{9}$$

$$isc_g(2^{-n}) \leqslant osc_g(2^{-n}), \tag{10}$$

$$cff_g(2^{-n}) \leqslant osc_g(2^{-n}), \tag{11}$$

for all $n \geqslant 1$, *and*

$$\exp(-\Delta_\phi(2^{-n})) \leqslant cff_g(2^{-n}), \tag{12}$$

for large n.

Consequently, (4) $\Rightarrow$ (7), (5) $\Rightarrow$ (8), (6) $\Rightarrow$ (8) *and* (7) $\Rightarrow$ (8).

Remark 2. The inequalities (9)–(11) are typically strict if $N > 2$. If $N = 2$ then we have equality in (9)–(11).

It was proved by Bramson and Kalikow [3] that merely assuming continuity and strict positivity on g is not sufficient for a unique g-measure. Their constructive counterexample is discussed in the overview paper [15].

For Borel probability measures ν_1 and ν_2 on Σ_N, we define the Monge–Kantorovich metric

$$d_w(\nu_1, \nu_2) = \sup\left\{\left|\int_{\Sigma_N} f\,\mathrm{d}(\nu_1 - \nu_2)\right|;\ f : \Sigma_N \to \mathbb{R},\ |f(x) - f(y)| \leqslant \rho(x, y)\right\}.$$

It is well known that this metric metrizes the weak star topology on $M(\Sigma_N)$ (see, e.g. Dudley [8]).

Theorem 1. *Let* $g : \Sigma_N \to (0, 1)$ *be a continuous strictly positive* g*-function satisfying condition (8).*

Then there exists a unique g*-measure,* μ*, and*

$$\sup_{x\in\Sigma_N} d_w(T_g^{\star n}\delta_x, \mu) \leqslant \psi(n), \tag{13}$$

where δ_x *denotes the Dirac probability measure concentrated in* $x \in \Sigma_N$*, and*

$$\psi(n) = \sum_{j=1}^{n+1}\left(P(Y_{n-j+1} = 1)\prod_{m=1}^{j-1} osc_g(2^{-m})\right)2^{-j},$$

where $\{Y_n\}_{n=0}^\infty$ *is a Markov chain with state space* $\mathbb{N}$ *starting at* $Y_0 := 1$ *with* $P(Y_{n+1} = k + 1 \mid Y_n = k) = osc_g(2^{-k})$*, and* $P(Y_{n+1} = 1 \mid Y_n = k) = 1 - osc_g(2^{-k})$*, for any* $k \geqslant 1$.

Remark 3. Note that $\{Y_n\}_{n=0}^\infty$ is a non-ergodic Markov chain under condition (8) (see, e.g. [14], p 80, ex 18), so $\psi(n) \to 0$.

If $\sum_{n=1}^\infty(1 - osc_g(2^{-n})) < \infty$, then we can use estimates from [4] to obtain bounds on the convergence rates for $\psi(n)$.

2. Proofs

Proof of proposition 1. From (1) we obtain

$$1 - \sum_{j=1}^N \min(g(jx), g(jy)) = \sum_{j=1}^N (g(jx) - \min(g(jx), g(jy)))$$

and

$$1 - \sum_{j=1}^N \min(g(jx), g(jy)) = \sum_{j=1}^N (g(jy) - \min(g(jx), g(jy))).$$

By summing these two equations, we get

$$2\left(1-\sum_{j=1}^{N}\min(g(jx),g(jy))\right)=\sum_{j=1}^{N}|g(jy)-g(jx)|.$$

Thus

$$osc_g(2^{-n})=\inf_{\rho(x,y)\leqslant 2^{-n}}\sum_{j=1}^{N}\min(g(jx),g(jy))=\inf_{\rho(x,y)\leqslant 2^{-n}}\left(1-\frac{1}{2}\sum_{j=1}^{N}|g(jy)-g(jx)|\right)$$
$$=1-\frac{1}{2}\sup_{\rho(x,y)\leqslant 2^{-n}}\sum_{j=1}^{N}|g(jy)-g(jx)|$$

and since

$$\sup_{\rho(x,y)\leqslant 2^{-n}}\sum_{j=1}^{N}|g(jy)-g(jx)|\leqslant N\Delta_g(2^{-n}),$$

it follows that (9) holds.

The proofs of (10) and (11), follows immediately from the definitions.

To prove (12) first note that

$$\begin{aligned}
1-cff(2^{-n}) &= \sup_{1\leqslant j_l\leqslant N,1\leqslant l\leqslant n-1}\left(1-\sum_{i=1}^{N}\inf_{y\in\Sigma_N}g(ij_1\cdots j_{n-1}y)\right)\\
&\leqslant \sup_{1\leqslant j_l\leqslant N,1\leqslant l\leqslant n-1}\left(\inf_{y_0\in\Sigma_N}\left(\sum_{i=1}^{N}(g(ij_1\cdots j_{n-1}y_0)-\inf_{y\in\Sigma_N}g(ij_1\cdots j_{n-1}y))\right)\right)\\
&\leqslant \sup_{1\leqslant j_l\leqslant N,1\leqslant l\leqslant n-1}\min_{j\in\{1,\ldots,N\}}\left(\sum_{i=1,\ i\neq j}^{N}\Big(\sup_{y\in\Sigma_N}g(ij_1\cdots j_{n-1}y)\right.\\
&\qquad\left.-\inf_{y\in\Sigma_N}g(ij_1\cdots j_{n-1}y)\Big)\right)\\
&\leqslant \sup_{x\in\Sigma_N}\min_{j\in\{1,\ldots,N\}}\left(\sum_{i=1,\ i\neq j}^{N}\sup_{\{y:\rho(x,y)\leqslant 2^{-n}\}}|g(ix)-g(iy)|\right).
\end{aligned}\tag{14}$$

(The second from last inequality of (14) can be seen by considering a point $y_0^\star$ such that $g(j^\star j_1\cdots j_{n-1}y_0^\star)=\inf_{y\in\Sigma_N}g(j^\star j_1\cdots j_{n-1}y)$, where $j^\star$ is an index such that the minimum in the third line of (14) is attained.)

By the mean-value theorem, we have for any $x\in\Sigma_N$, and $i\in\{1,\ldots,N\}$,

$$\sup_{\{y:\rho(x,y)\leqslant 2^{-n}\}}|\log g(ix)-\log g(iy)|\geqslant\frac{1}{g(ix)+\Delta_g(2^{-n})}\sup_{\{y:\rho(x,y)\leqslant 2^{-n}\}}|g(ix)-g(iy)|$$

and thus, by taking summations and using (1) we obtain

$$\begin{aligned}
&(1+N\Delta_g(2^{-n}))\max_{i\in\{1,\ldots,N\}}\sup_{\{y:\rho(x,y)\leqslant 2^{-n}\}}|\log g(ix)-\log g(iy)|\\
&\qquad\geqslant\sum_{i=1}^{N}\sup_{\{y:\rho(x,y)\leqslant 2^{-n}\}}|g(ix)-g(iy)|.
\end{aligned}\tag{15}$$

From (14) and (15), we see that there exists a constant $c < 1$ such that

$$\begin{aligned}
1 - cff_g(2^{-n}) &\leqslant \sup_{x\in\Sigma_N} \min_{j\in\{1,\dots,N\}} \left(\sum_{i=1,\ i\neq j}^{N} \sup_{\{y:\rho(x,y)\leqslant 2^{-n}\}} |g(ix) - g(iy)| \right) \\
&\leqslant \sup_{x\in\Sigma_N} \frac{N-1}{N} \sum_{i=1}^{N} \sup_{\{y:\rho(x,y)\leqslant 2^{-n}\}} |g(ix) - g(iy)| \\
&\leqslant \sup_{x\in\Sigma_N} \frac{N-1}{N} (1 + N\Delta_g(2^{-n})) \max_{i\in\{1,\dots,N\}} \sup_{\{y:\rho(x,y)\leqslant 2^{-n}\}} |\log g(ix) - \log g(iy)| \\
&\leqslant c \sup_{x\in\Sigma_N} \max_{i\in\{1,\dots,N\}} \sup_{\{y:\rho(x,y)\leqslant 2^{-n}\}} |\log g(ix) - \log g(iy)| = c\Delta_\phi(2^{-n}),
\end{aligned}$$

if n is sufficiently large. Since $\mathrm{e}^{-x} \leqslant 1 - cx$ for small $x \geqslant 0$ it follows that

$$cff_g(2^{-n}) \geqslant 1 - c\Delta_\phi(2^{-n}) \geqslant \exp(-\Delta_\phi(2^{-n})),$$

for all n sufficiently large. This completes the proof of (12) and proposition 1. □

Proof of theorem 1. Let P denote the product Lebesgue measure on $((0,1)^{\mathbb{N}}, B)$, where B is the product Borel σ-field on $(0,1)^{\mathbb{N}}$, and let x_0 and y_0 be two fixed arbitrary elements of Σ_N. We are first going to construct random variables, $X_n^{x_0,y_0}(x_0) : (0,1)^{\mathbb{N}} \to \Sigma_N$ and $X_n^{x_0,y_0}(y_0) : (0,1)^{\mathbb{N}} \to \Sigma_N$, $n \geqslant 0$, with $P(X_n^{x_0,y_0}(x_0) \in \cdot) = T_g^{\star n}\delta_{x_0}(\cdot)$, $P(X_n^{x_0,y_0}(y_0) \in \cdot) = T_g^{\star n}\delta_{y_0}(\cdot)$, such that the distance $E\rho(X_n^{x_0,y_0}(x_0), X_n^{x_0,y_0}(y_0))$ is as small as possible. (This will be used when we are estimating the convergence rates below.) Let $X_0^{x_0,y_0}(x_0) = x_0$, $X_0^{x_0,y_0}(y_0) = y_0$.

For $s = s_1 s_2 \cdots \in (0,1)^{\mathbb{N}}$, $x \in \{x_0, y_0\}$, and $j \in \{1,\dots,N\}$, define $X_{n+1}^{x_0,y_0}(x) = jX_n^{x_0,y_0}(x)$, if

$$\sum_{i=1}^{j-1} \min(g(iX_n^{x_0,y_0}(x_0)), g(iX_n^{x_0,y_0}(y_0))) \leqslant s_{n+1} < \sum_{i=1}^{j} \min(g(iX_n^{x_0,y_0}(x_0)), g(iX_n^{x_0,y_0}(y_0)))$$

or

$$\begin{aligned}
&\sum_{i=1}^{N} \min(g(iX_n^{x_0,y_0}(x_0)), g(iX_n^{x_0,y_0}(y_0))) \\
&\qquad + \sum_{i=1}^{j-1} (g(iX_n^{x_0,y_0}(x)) - \min(g(iX_n^{x_0,y_0}(x_0)), g(iX_n^{x_0,y_0}(y_0)))) \\
&\qquad \leqslant s_{n+1} \\
&\qquad < \sum_{i=1}^{N} \min(g(iX_n^{x_0,y_0}(x_0)), g(iX_n^{x_0,y_0}(y_0))) \\
&\qquad + \sum_{i=1}^{j} (g(iX_n^{x_0,y_0}(x)) - \min(g(iX_n^{x_0,y_0}(x_0)), g(iX_n^{x_0,y_0}(y_0)))).
\end{aligned} \tag{16}$$

We obtain

$$\begin{aligned}
&P(\rho(X_{n+1}^{x_0,y_0}(x_0), X_{n+1}^{x_0,y_0}(y_0)) = 2^{-(m+1)} \mid \rho(X_n^{x_0,y_0}(x_0), X_n^{x_0,y_0}(y_0)) = 2^{-m}) \\
&\qquad \geqslant \inf_{\rho(x,y)\leqslant 2^{-m}} \sum_{i=1}^{N} \min(g(ix), g(iy)) := osc_g(2^{-m}), \qquad 1 \leqslant m \leqslant n+1.
\end{aligned}$$

Let $\{Y_n\}_{n=0}^{\infty}$, be a stochastic sequence with $Y_n : (0,1)^{\mathbb{N}} \to \mathbb{N}, n \geqslant 0$, defined inductively in the following way. Let $Y_0(s) = 1$, for all $s \in (0,1)^{\mathbb{N}}$. Suppose $Y_n(s) = m$. Let $Y_{n+1}(s) = m+1$ if $s_{n+1} < osc_g(2^{-m})$ and $Y_{n+1}(s) = 1$ otherwise.

Then $\{Y_n\}_{n=0}^{\infty}$ is a homogeneous Markov chain with $Y_0 = 1$ and

$$P(Y_{n+1} = m+1 \mid Y_n = m) = osc_g(2^{-m})$$

and

$$P(Y_{n+1} = 1 \mid Y_n = m) = 1 - osc_g(2^{-m}), \qquad m \geqslant 1,$$

such that

$$\rho(X_n^{x_0,y_0}(x_0), X_n^{x_0,y_0}(y_0)) \leqslant 2^{-Y_n}, \qquad n \geqslant 0.$$

Let μ be a g-measure and let δ_x denote the Dirac measure concentrated in $x \in \Sigma_N$, i.e. $\delta_x(A) = 1$ if $x \in A$ and $\delta_x = 0$ otherwise for any Borel set A in Σ_N. We have

$$\begin{aligned}
d_w(T_g^{\star n}\delta_x, \mu) &= \sup\left\{\left|T_g^n f(x) - \int_{\Sigma_N} T_g^n f(y)\,\mathrm{d}\mu(y)\right|; f : \Sigma_N \to \mathbb{R}, |f(x) - f(y)| \leqslant \rho(x,y)\right\} \\
&\leqslant \sup\left\{\sup_{x,y\in\Sigma_N} |T_g^n f(x) - T_g^n f(y)|;\ f : \Sigma_N \to \mathbb{R},\ |f(x) - f(y)| \leqslant \rho(x,y)\right\} \\
&\leqslant \sup\left\{\sup_{x,y\in\Sigma_N}\left|\int_{(0,1)^{\mathbb{N}}} (f(X_n^{x,y}(x)) - f(X_n^{x,y}(y)))\,\mathrm{d}P(s)\right|;\right. \\
&\qquad \left. f : \Sigma_N \to \mathbb{R},\ |f(x) - f(y)| \leqslant \rho(x,y)\right\} \\
&\leqslant \sup_{x,y\in\Sigma_N} \int_{(0,1)^{\mathbb{N}}} \rho(X_n^{x,y}(x), X_n^{x,y}(y))\,\mathrm{d}P(s) \leqslant \sum_{j=1}^{n+1} P(Y_n = j)2^{-j}.
\end{aligned} \tag{17}$$

Now

$$P(Y_n = j) = P(Y_{n-j+1} = 1)\prod_{m=1}^{j-1} osc_g(2^{-m})$$

and thus it follows from (17), that

$$\sup_{x\in\Sigma_N} d_w(T_g^{\star n}\delta_x, \mu) \leqslant \sum_{j=1}^{n+1}\left(P(Y_{n-j+1} = 1)\prod_{m=1}^{j-1} osc_g(2^{-m})\right)2^{-j}. \tag{18}$$

By assumption $\{Y_n\}$ is a non-ergodic Markov chain (see, e.g. [14], p 80, ex 18) and therefore it follows from (18), that

$$\sup_{x\in\Sigma_N} d_w(T_g^{\star n}\delta_x, \mu) \to 0.$$

This shows in particular that μ must be the only invariant probability measure and therefore completes the proof of theorem 1. □

Acknowledgments

This paper was written and revised during postdoctoral visits at the Departments of Mathematics at the Chinese University of Hong Kong, and the Australian National University, respectively. The author is grateful to John Hutchinson, Ka-Sing Lau and Yang Wang for their help in making this research possible. He also would like to thank one of the anonymous referees for pointing out [6], and for useful suggestions that improved the paper. This work was supported in part by Hellmuth Hertz' Foundation, a HK RGC grant, and an ARC grant.

References

[1] Berbee H 1987 Chains with infinite connections: uniqueness and Markov representation *Probab. Theory Related Fields* **76** 243–53

[2] Bowen R 1975 *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms (Lecture Notes in Math. vol 470)* (Berlin: Springer)

[3] Bramson M and Kalikow S 1993 Nonuniqueness in g-functions *Israel J. Math.* **84** 153–60

[4] Bressaud X, Fernández R and Galves A 1999 Decay of correlations for non-Hölderian dynamics. A coupling approach *Electron. J. Probab.* **4** 19

[5] Coelho Z and Quas A N 1998 Criteria for $\overline{d}$-continuity *Trans. Am. Math. Soc.* **350** 3257–68

[6] Comets F, Fernández R and Ferrari P A 2002 Processes with long memory: Regenerative construction and perfect simulation *Ann. Appl. Probab.* **12** 921–43

[7] Doeblin W and Fortet R 1937 Sur des chaines a liaisons completès *Bull. Soc. Math. France* **65** 132–48

[8] Dudley R M 1989 *Real Analysis and Probability* (New York: Wadsworth)

[9] Harris T E 1955 On chains of infinite order *Pacific J. Math.* **5** 707–24

[10] Iosifescu M and Grigorescu S 1990 *Dependence with Complete Connections and its Applications* (Cambridge: Cambridge University Press)

[11] Iosifescu M and Spătaru A 1973 On denumerable chains of infinite order *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **27** 195–214

[12] Kaijser T 1981 On a new contraction condition for random systems with complete connections *Rev. Roum. Math. Pures Appl.* **26** 1075–117

[13] Keane M 1972 Strongly mixing g-measures *Invent. Math.* **16** 309–24

[14] Prabhu N U 1965 *Stochastic Processes. Basic Theory and its Applications* (New York: Macmillan)

[15] Stenflo Ö 2002 Uniqueness of invariant measures for place-dependent random iterations of functions *IMA Vol. Math. Appl.* **132** 13–32

[16] Walters P 1975 Ruelle's operator theorem and g-measures *Trans. Am. Math. Soc.* **214** 375–87