Ergod. Th. & Dynam. Sys. (2008), **28**, 1145–1151
doi:10.1017/S0143385707000788

Square summability of variations and convergence of the transfer operator

ANDERS JOHANSSON† and ANDERS ÖBERG‡

† *Sub-department of Mathematics and Statistics, University of Gävle, SE-801 76 Gävle, Sweden*
(e-mail: ajj@hig.se)
‡ *Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden*
(e-mail: anders@math.uu.se)

(*Received* 5 *December* 2006 *and accepted in revised form* 20 *August* 2007)

Abstract. In this paper we study the one-sided shift operator on a state space defined by a finite alphabet. Using a scheme developed by Walters [P. Walters. *Trans. Amer. Math. Soc.* **353**(1) (2001), 327–347], we prove that the sequence of iterates of the transfer operator converges under square summability of variations of the g-function, a condition which gave uniqueness of a g-measure in our earlier work [A. Johansson and A. Öberg. *Math. Res. Lett.* **10**(5–6) (2003), 587–601]. We also prove uniqueness of the so-called G-measures, introduced by Brown and Dooley [G. Brown and A. H. Dooley. *Ergod. Th. & Dynam. Sys.* **11** (1991), 279–307], under square summability of variations.

1. *Introduction*

We consider the left one-sided shift map T on the state space $X^+ = S^{\mathbb{Z}_+}$, where S is a finite set. Thus T acts on elements x of X^+, $x = (x_0, x_1, x_2, \ldots)$, in the following way (each x_i belongs to S):

$$T(x_0, x_1, x_2, \ldots) = (x_1, x_2, \ldots).$$

A g-measure associated with a continuous function $g : S^{\mathbb{Z}_+} \to [0, 1]$ is a T-invariant measure μ in the space of Borel probability measures $\mathcal{P}X^+$, such that $g = d\mu/(d\mu \circ T)$, with $\sum_{y \in T^{-1}x} g(y) = 1$, for all $x \in X^+$. Such a function is referred to as a g-function and can be viewed as a probability transition function for the local inverses of T. Keane [**6**] introduced the concept of g-measures to ergodic theory.

Since the function g is continuous, existence of at least one g-measure follows if $g > 0$ and S is finite. If S is countably infinite, then existence is no longer automatic, as $S^{\mathbb{Z}_+}$ is not compact; however, a weak sufficient condition was given in [**8**].

Recent results concerning necessary conditions for the uniqueness of a g-measure have been given by Berger *et al* [**1**], and also by Hulse [**5**].

Necessary and sufficient conditions sometimes, but not always, focus on variations of the g-function over small parts of its domain. For $f : X^+ \to \mathbb{R}$, the variation is defined as

$$\mathrm{var}_n f = \sup_{x \sim_n y} |f(x) - f(y)|,$$

where $x \sim_n y$ if $x, y \in X^+$ coincide in their first n coordinates.

Walters [**12**] proved uniqueness and rates of convergence under summable variations of g-functions, or rather their logarithms $\log g$, which amounts to the same thing if $\inf g > 0$.

When S is finite, a sufficient condition for uniqueness of a g-measure, *square summability of variations* of the g-function, was found by the present authors in [**7**]. This result was extended to *countable state shifts* in Johansson *et al* [**8**], where square summability of variations of a g-function means that

$$\sum_{n=1}^{\infty} (\mathrm{var}_n f)^2 < \infty. \tag{1.1}$$

In this paper we prove convergence of the iterates of the transfer (transition) operator, which is defined by

$$\mathcal{L} f(x) = \sum_{y \in T^{-1}x} g(y) f(y) \tag{1.2}$$

for continuous functions f on X^+.

More specifically, we prove that with finitely many symbols in S and under square summability of variations of a strictly positive g-function, we have $\mathcal{L}^n f(x) \to \int f \, d\mu$ uniformly in x, where μ is the unique g-measure.

To accomplish this, we use a theory developed by Walters in [**13**] to study the sequence of iterates of an operator P, defined as

$$P^n f(x) = (\mathcal{L}^n f)(T^n x) \tag{1.3}$$

for $n \geq 1$.

Walters introduces an adjoint sequence of operators $P^{(n)^*} : \mathcal{P}X^+ \to \mathcal{P}X^+$, $n \geq 1$. Define $K_g^n = \{\nu \in \mathcal{P}X^+ \mid P^{(n)^*}\nu = \nu\}$. Then [**13**, Theorem 2.1(v)] implies that $K_g^n = P^{(n)^*}\mathcal{P}X^+$ and that $K_g^1 \supset K_g^2 \supset \cdots$. Walters defined $K_g = \bigcap_{n=1}^{\infty} K_g^n$, which corresponds to the set of G-measures in the terminology of Brown and Dooley [**2**]; see also Fan [**4**] and Dooley and Stenflo [**3**] for some recent contributions to establishing uniqueness of G-measures. One of the motivations for studying G-measures has been to obtain a more general framework for studying Riesz products; see [**4**], for example.

Our strategy is to prove that there exists a *unique* measure in K_g (it is trivial that any g-measure must be an element of K_g). We then use [**13**, Theorem 2.9] to conclude that $\mathcal{L}^n f(x) \to \int f \, d\mu$, where μ is the unique element of K_g. This follows from a compactness argument; therefore, to prove the convergence of $\mathcal{L}^n f$, we have to assume that the symbol space S is finite.

We need to define a probability measure, a Markov chain on X^+ which we call a g-chain, in order to reverse the dynamics in the absence of stationarity; it is not known whether or when the measures in K_g are invariant under T.

To conclude the proof, we use the same method as in [**7**] to show that any two extremal measures in K_g must be absolutely continuous, and hence that there can exist only one such measure. This method was developed further in [**8**] to cover the case of countable state shifts for g-measures (thus improving on the results of [**7**]), using more sophisticated methods to prove absolute continuity of measures. We note that these methods had been employed much earlier by Shiryaev and co-authors; see, for example, the survey by Shiryaev [**11**].

We end the paper by stating two open questions.

2. *Results and proofs*

THEOREM 2.1. *Assume that we have a finite symbol set S. Let $g > 0$ satisfy*

$$\sum_{n=1}^{\infty}(\mathrm{var}_n \log g)^2 \asymp \sum_{n=1}^{\infty}(\mathrm{var}_n\, g)^2 < \infty$$

and let μ be the corresponding unique g-measure. Then $K_g = \{\mu\}$ or, in other words, there exists a unique G-measure which is the same as the unique g-measure.

The following theorem is, in view of [**13**], a corollary of Theorem 2.1. It is based on [**13**, Theorem 2.9], and we reproduce part of the proof here to explain why we need to assume that there are only finitely many symbols: this is because one has to have a compact state space, or at least some kind of tightness property, so that a certain convergent subsequence will exist for a sequence of measures.

THEOREM 2.2. *Assume that we have a finite symbol set S. Let $g > 0$ be a g-function with square summable variation and let μ be the corresponding unique g-measure. Then* $\sup_x |\mathcal{L}^n f(x) - \int f\, d\mu| \to 0$.

Proof. (Borrowed from [**13**, Theorem 2.9]) Under these same assumptions, Theorem 2.1 gives us a unique measure in K_g. From this we can show that $\sup_x |P^n f(x) - \int f\, d\mu| \to 0$, because otherwise there would exist a continuous function f, an $\epsilon > 0$ and sequences $\{n_k\}$ and $\{x_k\}$ such that

$$\left| \int f\, d(P^{(n_k)^*}\delta_{x_k}) - \int f\, d\mu \right| \geq \epsilon \quad \text{for all } k \geq 1.$$

Owing to compactness, we can then pick a subsequence $\{k_j\}$ of $\{k\}$ such that the sequence of measures $\{P^{(n_{k_j})^*}\delta_{x_{k_j}}\}$ converges to $\nu \in \mathcal{P}X^+$; however, since $K_g^{(n)} = P^{(n)^*}\mathcal{P}X^+$, we would also have $\nu \in K_g$, which is a contradiction. Since $P^{(n)} f(x) = (\mathcal{L}^n f)(T^n x)$, we have $\sup_x |\mathcal{L}^n f(x) - \int f\, d\mu| \to 0$. □

2.1. *Proof of Theorem 2.1.*

We begin with some preliminary terminology; see [**8**] for a more thorough exposition. For a pair of probability measures $\nu, \tilde{\nu} \in \mathcal{P}X$ and some filtration $\{\mathcal{F}_n\}$, let $Z_n(x) = Z_n(x; \tilde{\nu}, \nu, \mathcal{F}_n)$ be the likelihood-ratio martingale

$$Z_n(x; \tilde{\nu}, \nu, \{\mathcal{F}_n\}) = \frac{d\tilde{\nu}|_{\mathcal{F}_n}}{d\nu|_{\mathcal{F}_n}}(x)$$

on $\mathcal{F}_n$, where we assume that $\tilde{\nu}$ and ν are locally absolutely continuous on $\mathcal{F}_n$.

https://doi.org/10.1017/S0143385707000788 Published online by Cambridge University Press

Now let $X = S^{\mathbb{Z}}$ and $X^+ = S^{\mathbb{Z}_+}$, and extend the one-sided shift T on X^+ to a two-sided shift on X. For $a, b \in \mathbb{Z}$, let $\Pi_{a,b}$ be the mapping that takes $x \in X$ to $(x_a, x_{a+1}, \ldots, x_b) \in S^{b-a+1}$, and define the natural projection $\Pi_+ : X \to X^+$ taking bi-infinite sequences x in X to one-sided ones in X^+ by $\Pi_+((x_i)_{i=-\infty}^{i=\infty}) = (x_0, x_1, \ldots)$. Also let $\mathcal{F}_a^b$ be the algebra generated by $\Pi_{a,b}$ and define $\mathcal{F}_n^- := \mathcal{F}_{-n}^{-1}$, the 'forward' algebra, and $\mathcal{F}_n^+ := \mathcal{F}_0^{n-1}$, the 'backward' algebra. A *cylinder set* is a set of the form $[x]_a^b = \Pi_{a,b}^{-1}\Pi_{a,b}(x)$.

We now define a certain Markov chain that will enable us to go forward in time, since we may not assume that the measures in K_g defined by Walters are invariant under T.

Definition 1. A *g-chain* on $X = S^{\mathbb{Z}}$ is a probability measure $\nu \in \mathcal{P}X$ such that, for all $n \in \mathbb{Z}$,

$$\nu(x_n | x_{n+1}, x_{n+2}, \ldots) = g(x_n, x_{n+1}, x_{n+2}, \ldots). \tag{2.1}$$

A *forward g-chain* is a probability measure ν on X satisfying (2.1) for $n \leq -1$.

The distribution under a g-chain ν of the process $x^{(t)} \in X^+$, $t \in \mathbb{Z}$ defined by

$$x^{(t)} := \Pi_+(T^{-t}x) = (x_{-t}, x_{-t+1}, \ldots)$$

is that of a Markov chain such that the transition probabilities are given by g (and the transition operator of the chain is $\mathcal{L}$). That is, for all $t \in \mathbb{Z}$,

$$\nu(x^{(t)} | x^{(t-1)}) = g(x^{(t)}).$$

The same holds for $t \geq 1$ in the case of a forward g-chain ν.

LEMMA 2.3. *A probability measure $\nu \in \mathcal{P}X$ is a g-chain only if $\nu \circ \Pi_+^{-1} \in \mathcal{P}X^+$ is an element of the set K_g of eigen-measures as defined by Walters in* [**13**]. *Conversely, any ν in K_g corresponds, by extension, to a unique g-chain.*

Proof. Suppose that we have $P^*\nu = \nu$. This is equivalent to $\mathcal{L}^*(\nu \circ T^{-1}) = \nu$. By interpreting the conditional probability $\nu(x_0 | x_1, x_2, \ldots)$ as the transition probability $\nu(x^{(0)} | x^{(-1)})$ from the Markov chain defined above, we reach the conclusion that

$$\nu(x_0 | x_1, x_2, \ldots) = g(x_0, x_1, \ldots).$$

Thus, if $P^{(n)^*}\nu = \nu$, we have, for $0 \leq k \leq n-1$,

$$\nu(x_k | x_{k+1}, x_{k+2}, \ldots) = g(x_k, x_{k+1}, x_{k+2}, \ldots). \tag{2.2}$$

Since $K_g^n = P^{(n)^*}\mathcal{P}X^+$ and $K_g = \bigcap_{n=1}^{\infty} K_g^n$, we know (2.2) holds for all $k \geq 0$. Using measures in the non-empty set K_g as initial distributions, we may define the full g-chain by making a unique canonical extension (see Neveu [**9**, p. 83, the corollary]) so that (2.2) is true for $k \leq -1$. Hence (2.2) holds for all $k \in \mathbb{Z}$. □

From now on, we shall not distinguish between a g-chain ν and its one-sided restriction $\nu \circ \Pi_+^{-1}$.

LEMMA 2.4. *The set K_g is a non-empty convex subset with mutually singular extreme points.*

Proof. See Walters [**13**, Theorem 2.11 and its proof]. □

https://doi.org/10.1017/S0143385707000788 Published online by Cambridge University Press

We work under the assumption that $g > 0$ and hence that any two g-chains will be locally absolutely continuous on any of the algebras $\mathcal{F}_a^b$, where a and b are finite. Given two g-chains $\nu, \tilde{\nu} \in K_g$, let

$$\xi_n(x) := Z_n(x; \tilde{\nu}, \nu, \{\mathcal{F}_n^+\}). \tag{2.3}$$

Note that ξ_n is the likelihood-ratio martingale and, since it is a positive $L_1(\nu)$-bounded ν-martingale, we know that it converges ν-almost surely. We want to show that it is uniformly integrable (*UI*) with respect to ν, which gives L_1-density between $\tilde{\nu}$ and ν on $\mathcal{F}^+ = \lim \mathcal{F}_n^+$ and thus, according to Lemma 2.4, a contradiction if ν and $\tilde{\nu}$ are chosen to be distinct and extremal in K_g.

It was shown in [**7**, p. 595] that verifying the *UI* property for a likelihood-ratio martingale such as ξ_n amounts to showing that

$$\lim_{K\to\infty} \sup_n \tilde{\nu}(\log \xi_n > K) = 0. \tag{2.4}$$

To see (2.4), we note that, for a fixed value of m, a translation by m steps to the left of both the point and the measure gives

$$\xi_m(x) = Z_m(T^m x; \tilde{\nu} \circ T^{-m}, \nu \circ T^{-m}, \{\mathcal{F}_m^-\}). \tag{2.5}$$

Hence the law of ξ_m under $\tilde{\nu}$, namely $\tilde{\nu} \circ \xi_m^{-1}$, equals the distribution $(\tilde{\nu} \circ T^{-m}) \circ \zeta_{m,m}^{-1}$ where $\{\zeta_{m,n}(x) : n \in \mathbb{Z}_+\}$ is the forward likelihood-ratio martingale

$$\zeta_{m,n}(x) = Z_n(x; \tilde{\nu} \circ T^{-m}, \nu \circ T^{-m}, \{\mathcal{F}_n^-\}).$$

This means that we start with two extremal measures $\nu, \tilde{\nu} \in K_g$ translated m times to the left, i.e. $\nu \circ T^{-m}$ and $\tilde{\nu} \circ T^{-m}$; these can be extended to well-defined g-chains and we may then go forward along $\mathcal{F}_n^-$.

Thus, it is enough to prove the following lemma which states that the forward likelihood ratios are uniformly tight in a strong sense. The estimates have to be uniform in all g-chains, since our substitutions $\mu = \nu \circ T^{-m}$ and $\tilde{\mu} = \tilde{\nu} \circ T^{-m}$ are valid only for a fixed m.

LEMMA 2.5. *Assume that* $\mathrm{var}_n\, g$ *is square summable and that* g *is bounded away from zero. For all* $\epsilon > 0$*, there exists a* $K = K(\epsilon)$ *such that*

$$\sup_{\tilde{\mu},\mu} \sup_n \tilde{\mu}(\log Z_n(x; \tilde{\mu}, \mu, \{\mathcal{F}_n^-\}) > K) < \epsilon, \tag{2.6}$$

where $\tilde{\mu}$ *and* μ *are chosen among all pairs of forward* g*-chains.*

Proof. We adapt part of the proof in [**7**, pp. 597–598]. One could also use the more advanced theory of Shiryaev [**11**] and his co-authors, as was done in [**8**].

Given μ and $\tilde{\mu}$, write (using the notation from [**7**]) $M_n(x) = Z_n(x; \tilde{\mu}, \mu, \{\mathcal{F}_n^-\})$. We show that $\log M_n$ has a Doob decomposition

$$\log M_n = A_n + \eta_n$$

where A_n is pre-visible with the uniform bound $A_n \leq C_1 \sum(\mathrm{var}_n\, g)^2$ and, moreover, η_n is a $\tilde{\mu}$-martingale and uniformly bounded in $L_2(\tilde{\mu})$ with $\tilde{\mu}(\eta_n^2) \leq C_2 \sum(\mathrm{var}_n\, g)^2$ (see, for example, [**14**, pp. 120–121]). It is the uniformity of the estimates that makes it possible to conclude the strong formulation of (2.6). We define

$$P_n(x) = \frac{\mu[x_{-n}, x_{-n+1}, \ldots, x_{-1}]}{\mu[x_{-n+1}, \ldots, x_{-1}]},$$
$$\tilde{P}_n(x) = \frac{\tilde{\mu}[x_{-n}, x_{-n+1}, \ldots, x_{-1}]}{\tilde{\mu}[x_{-n+1}, \ldots, x_{-1}]},$$

and note that $|\tilde{P}_n - P_n| \leq \mathrm{var}_n\, g$.

The proof then proceeds exactly as in [**7**, pp. 597–598]. □

3. *Open questions*

In this section we present two questions which the authors find interesting, as well as challenging, in the light of the present investigation.

3.1. *Question 1.*

Does uniqueness of a g-measure imply

$$\mathcal{L}^n f \to \int f \, d\mu \ ?$$

We assume that we have finitely many symbols for the left-shift map T and a continuous and strictly positive g-function.

In this paper we used the fact that a unique measure in K_g (in Walters' notation) or, equivalently, a unique G-measure (as defined by Brown and Dooley [**2**], Fan [**4**] and others) implies that $\mathcal{L}^n f(x) \to \int f \, d\mu$ uniformly in $x \in X^+$. It is natural to ask: if K_g only contains the g-measure, then would this also be the unique member of K_g? (Recall that a g-measure is always a member of K_g.) It would then follow that uniqueness of a g-measure (under the assumptions given in our question) implies convergence of the iterates of $\mathcal{L}$, without imposing any stronger regularity on g than continuity.

3.2. *Question 2.*

Under square summability of variations of the g-function, what is the rate of convergence of $\mathcal{L}^n f(x)$ in the supremum-norm? Rates of convergence in the case of *summable variations* are available; see Pollicott [**10**].

Acknowledgements. The authors are grateful for the hospitality shown by the University of Warwick and, in particular, for the conversations with Thomas Jordan, Mark Pollicott and Peter Walters.

References

[1] N. Berger, C. Hoffman and V. Sidoravicius. Non-uniqueness for specifications in $l^{2+\epsilon}$. *Preprint*, 2005. Available at www.arxiv.org (PR/0312344).

[2] G. Brown and A. H. Dooley. Odometer actions on G-measures. *Ergod. Th. & Dynam. Sys.* **11** (1991), 279–307.

[3] A. H. Dooley and Ö. Stenflo. A criterion for uniqueness in G-measures and perfect sampling. *Math. Proc. Camb. Philos. Soc.* **140** (2006), 545–551.
[4] A. H. Fan. On uniqueness of G-measures and g-measures. *Studia Math.* **119**(3) (1996), 255–269.
[5] P. Hulse. An example of non-unique g-measures. *Ergod. Th. & Dynam. Sys.* **26** (2006), 439–445.
[6] M. Keane. Strongly mixing g-measures. *Invent. Math.* **16** (1972), 309–324.
[7] A. Johansson and A. Öberg. Square summability of variations of g-functions and uniqueness of g-measures. *Math. Res. Lett.* **10**(5–6) (2003), 587–601.
[8] A. Johansson, A. Öberg and M. Pollicott. Countable state shifts and uniqueness of g-measures. *Amer. J. Math.* **129**(6) (2007), 1501–1511.
[9] J. Neveu. *Mathematical Foundations of the Calculus of Probability*. Holden-Day, San Francisco, 1965.
[10] M. Pollicott. Rates of mixing for potentials of summable variation. *Trans. Amer. Math. Soc.* **352** (2000), 843–853.
[11] A. N. Shiryaev. Absolute continuity and singularity of probability measures in functional spaces. *Proceedings of the International Congress of Mathematicians (Helsinki, 1978)*. Vol. 1. Acad. Sci. Fennica, Helsinki, 1980, pp. 209–225.
[12] P. Walters. Ruelle's operator theorem and g-measures. *Trans. Amer. Math. Soc.* **214** (1975), 375–387.
[13] P. Walters. Convergence of the Ruelle operator for a function satisfying Bowen's condition. *Trans. Amer. Math. Soc.* **353**(1) (2001), 327–347.
[14] D. Williams. *Probability with Martingales*. Cambridge University Press, Cambridge, 1991.