Stochastics and Dynamics
Vol. 13, No. 1 (2013) 1250008 (16 pages)

DOI: 10.1142/S0219493712500086

A LIMIT THEOREM FOR OCCUPATION MEASURES OF LÉVY PROCESSES IN COMPACT GROUPS

ARNO BERGER
Mathematical and Statistical Sciences, University of Alberta,
Edmonton AB, T6G 2G1, Canada
aberger@math.ualberta.ca

STEVEN N. EVANS
Department of Statistics #3860, 367 Evans Hall,
University of California, Berkeley, CA 94720-3860, USA
evans@stat.berkeley.edu

Received 14 November 2011
Accepted 10 January 2012
Published 2 August 2012

A short proof utilizing dynamical systems techniques is given of a necessary and sufficient condition for the normalized occupation measure of a Lévy process in a metrizable compact group to be asymptotically uniform with probability one.

Keywords: Uniform distribution; continuous uniform distribution; ergodic theorem; recurrence; Haar measure; Benford function.

AMS Subject Classification: Primary: 60F15, 60G51, 37A50; Secondary: 60B15, 60G50

1. Introduction

Processes with stationary and independent increments are the continuous-time analogues of sums of independent, identically distributed random variables. They constitute one of the simplest yet also most fundamental classes of stochastic processes. With an additional assumption on the regularity of sample paths, such processes are referred to as *Lévy processes*, see Definition 2.2 below. The class of Lévy processes has been studied extensively and in many different state spaces, ranging from the classical case of $\mathbb{R}^d$ (where it contains both the Wiener and Poisson processes as extremely important examples) to more general topological groups, including non-Abelian ones. Due to their capability of incorporating both diffusion-type continuous evolution and jumps, Lévy processes are now widely used as basic stochastic models in applied mathematics, notably in mathematical finance and quantum physics, see, for example, [1] which also contains ample references to the vast literature on the subject.

The aim of this note is to provide an easily accessible, dynamical systems oriented proof of a fundamental fact concerning the convergence to Haar measure of the normalized occupation measures of any Lévy process taking values in a compact group. Various special cases of the main results, Theorem 3.1 and Corollary 3.1, have been (re-)discovered repeatedly over the years, as have some related facts, see Remark 4.5. However, arguments geared towards these special cases tend to obscure rather than elucidate the underlying general dynamical principle. As this note ventures to demonstrate, the latter is most easily understood when stripped of all superfluous particulars.

2. Basic Definitions and Notations

Throughout, G denotes a metrizable compact group, with the group operation written multiplicatively, with neutral element e_G and with Borel σ-algebra $\mathcal{B}_G$. When written without a subscript, the symbol $\mathcal{B}$ stands for the Borel σ-algebra on $\mathbb{R}$, or on some (Borel) subset thereof. The sets gB and Bg are the images of $B \in \mathcal{B}_G$ under, respectively, the left- and right-translation by g; that is, $gB = \{gb : b \in B\}$ and $Bg = \{bg : b \in B\}$. Write λ_G for the (normalized) Haar measure on G; that is, λ_G is the unique probability measure on $(G, \mathcal{B}_G)$ that is invariant under all left-translations. (Equivalently, λ_G is the unique probability measure on $(G, \mathcal{B}_G)$ invariant under all right-translations.) For any closed (and hence compact) subgroup H of G it will be understood that the corresponding Haar measure λ_H is defined on all of $(G, \mathcal{B}_G)$, rather than merely on $(H, \mathcal{B}_H)$ and $\lambda_H(G\backslash H) = 0$. For any $g \in G$, denote by ϵ_g the Dirac probability measure concentrated at g.

Denote by $C(G)$ the separable Banach space of continuous, complex-valued functions on G, equipped with the supremum norm. The dual of $C(G)$ is the space of finite, complex-valued measures on $(G, \mathcal{B}_G)$; from now on this space will always be equipped with the corresponding weak* topology.

Definition 2.1. A measurable function $\gamma : [0, +\infty) \to G$ is *continuously uniformly distributed* in G, abbreviated henceforth as *c.u.d.*, if

$$\lim_{T\to+\infty} \frac{1}{T} \int_0^T \varphi(\gamma(t))\mathrm{d}t = \int_G \varphi \,\mathrm{d}\lambda_G \quad \forall\, \varphi \in C(G). \tag{2.1}$$

Similarly, with $\lfloor y \rfloor$ denoting, as usual, the largest integer not larger than $y \in \mathbb{R}$, a sequence $(g_n)_{n\in\mathbb{N}}$ is *uniformly distributed* (*u.d.*) in G whenever the function $\gamma : t \mapsto g_{\lfloor t\rfloor+1}$ is c.u.d., see [19].

Note that γ is c.u.d. if and only if the normalized occupation measures Λ_T converge to λ_G as $T \to +\infty$; here the probability measure Λ_T is, for every $T > 0$, defined by

$$\Lambda_T(B) = \frac{1}{T} \int_0^T \mathbf{1}_B(\gamma(t))\mathrm{d}t \quad \forall\, B \in \mathcal{B}_G,$$

with $\mathbf{1}_B$ denoting the indicator function of any set $B \in \mathcal{B}_G$.

Definition 2.2. A *Lévy process* in G is a family $X = (X_t)_{t\geq 0}$ of G-valued random variables, defined on some underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with the following properties:

(i) For any $0 \leq t_1 < t_2$ the distribution of the increment $X_{t_1}^{-1} X_{t_2}$ is the same as the distribution of $X_0^{-1} X_{t_2 - t_1}$.
(ii) The random variables $X_{t_1}, X_{t_1}^{-1} X_{t_2}, X_{t_2}^{-1} X_{t_3}, \ldots, X_{t_{n-1}}^{-1} X_{t_n}$ are independent whenever $n \geq 2$ and $0 \leq t_1 < t_2 < \cdots < t_n$.
(iii) For $\mathbb{P}$-almost all $\omega \in \Omega$, the G-valued function $t \mapsto X_t(\omega)$ is right-continuous with left-limits (or *rcll* for short); that is, $\lim_{\varepsilon\downarrow 0} X_{t+\varepsilon}(\omega) = X_t(\omega)$ for all $t \geq 0$, and $\lim_{\varepsilon\downarrow 0} X_{t-\varepsilon}(\omega) =: X_{t-}(\omega)$ exists for every $t > 0$.

For fixed $\omega \in \Omega$, the rcll function $t \mapsto X_t(\omega)$ is referred to as a *path* of X.

Write D for the set of all rcll functions from $[0, \infty)$ to G. Note that the equivalent French acronym *càdlàg* is often used instead of rcll, and D is called the *Skorohod space* associated with G. It is possible to equip D with a complete, separable metric such that the corresponding Borel σ-algebra $\mathcal{B}_D$ coincides with the σ-algebra generated by the sets of the form

$$\{\gamma \in D : \gamma(t_j) \in B_j \text{ for } n \in \mathbb{N}; j = 1, \ldots, n; 0 \leq t_1 < \cdots < t_n; B_1, \ldots, B_n \in \mathcal{B}_G\},$$

see, for example, Secs. 3.5 and 3.7 of [10].

3. Main Result and Applications

Let $X = (X_t)_{t\geq 0}$ be a Lévy process in G. For every $t \geq 0$, write μ_t for the distribution of the increment $X_0^{-1} X_t$. Note that the family of probability measures $(\mu_t)_{t\geq 0}$ is a *convolution semigroup*; that is, $\mu_{t_1} * \mu_{t_2} = \mu_{t_1+t_2}$ for all $t_1, t_2 \geq 0$, where $*$ denotes the convolution of probability measures on $(G, \mathcal{B}_G)$. It follows that each probability measure μ_t is *infinitely divisible*, though no explicit use of this property will be made here. For any probability measure ν on $(G, \mathcal{B}_G)$, recall that the *support* of ν is the smallest closed set $F \subset G$ with $\nu(F) = 1$. For every $t \geq 0$ write S_t for the support of μ_t. The following characterization of the almost sure continuous uniform distribution for the paths of X is the main content of this note.

Theorem 3.1. *Let $X = (X_t)_{t\geq 0}$ be a Lévy process in the metrizable compact group G. Then the following statements are equivalent*:

(i) *The set $\bigcup_{t\geq 0} S_t$ is dense in G.*
(ii) *The paths of X are, with probability one, c.u.d. in G.*

Proof. In order not to interrupt the main thread, two auxiliary facts of a technical nature are deferred to the subsequent Lemmas 3.1 and 3.2.

It will be convenient to formulate the main part of the proof using the terminology of ergodic theory. To this end, consider the probability measure ρ defined on $(G \times D, \mathcal{B}_G \otimes \mathcal{B}_D)$ by setting, for any $n \in \mathbb{N}$, $0 = t_0 < t_1 < \cdots < t_n$ and $B_0, B_1, \ldots, B_n \in \mathcal{B}_G$,

$$\rho(B_0 \times \{\gamma \in D : \gamma(t_j) \in B_j \,\forall\, j = 1, \ldots, n\}) \\ := \mathbb{P}\{\xi \in B_0, X_0^{-1} X_{t_j} \in B_j \,\forall\, j = 1, \ldots, n\},$$

where ξ is a random variable, also defined on $(\Omega, \mathcal{F}, \mathbb{P})$, that is independent of the process X and has distribution λ_G. For every $t \geq 0$, define a map R_t of $G \times D$ into itself by

$$R_t(g, \gamma(\bullet)) = (g\gamma(0)^{-1}\gamma(t), \gamma(t)^{-1}\gamma(t + \bullet)) \quad \forall\, (g, \gamma) \in G \times D.$$

Clearly, R_t is measurable, and

$$\begin{aligned} R_{t_1} \circ R_{t_2}(g, \gamma(\bullet)) &= R_{t_1}(g\gamma(0)^{-1}\gamma(t_2), \gamma(t_2)^{-1}\gamma(t_2 + \bullet)) \\ &= (g\gamma(0)^{-1}\gamma(t_2)\gamma(t_2)^{-1}\gamma(t_2 + t_1), \\ &\qquad (\gamma(t_2)^{-1}\gamma(t_2 + t_1))^{-1}\gamma(t_2)^{-1}\gamma(t_2 + t_1 + \bullet)) \\ &= (g\gamma(0)^{-1}\gamma(t_1 + t_2), \gamma(t_1 + t_2)^{-1}\gamma(t_1 + t_2 + \bullet)) \\ &= R_{t_1 + t_2}(g, \gamma(\bullet)) \end{aligned}$$

holds for all $t_1, t_2 \geq 0$. Moreover, since $\xi X_0^{-1} X_t$ has distribution λ_G for all $t \geq 0$, it follows from the stationarity and independence of increments of X that for any $n \in \mathbb{N}$, $0 = t_0 < t_1 < \cdots < t_n$ and $B_0, B_1, \ldots, B_n \in \mathcal{B}_G$,

$$\begin{aligned} &\rho \circ R_t^{-1}(B_0 \times \{\gamma : \gamma(t_j) \in B_j \,\forall\, j = 1, \ldots, n\}) \\ &\quad = \mathbb{P}\{\xi X_0^{-1} X_t \in B_0, X_t^{-1} X_{t+t_j} \in B_j \,\forall\, j = 1, \ldots, n\} \\ &\quad = \mathbb{P}\{\xi \in B_0, X_0^{-1} X_{t_j} \in B_j \,\forall\, j = 1, \ldots, n\} \\ &\quad = \rho(B_0 \times \{\gamma : \gamma(t_j) \in B_j \,\forall\, j = 1, \ldots, n\}). \end{aligned}$$

Thus, $(R_t)_{t \geq 0}$ is a ρ-preserving semi-flow. In particular, the stochastic process $(\xi X_0^{-1} X_t)_{t \geq 0}$ is *stationary* [7, 18]. Recall that $(R_t)_{t \geq 0}$ is said to be *ergodic* if $\rho(R_t^{-1}(A) \triangle A) = 0$ for $A \in \mathcal{B}_G \otimes \mathcal{B}_D$ and all $t \geq 0$ implies that $\rho(A) \in \{0, 1\}$; here, as usual, $\triangle$ denotes the symmetric difference of two sets. From Lemma 3.1 below, it follows that $(R_t)_{t \geq 0}$ is ergodic if and only if

$$\mathbb{P}\{\lambda_G(B X_0^{-1} X_t \triangle B) = 0\} = 1 \quad \forall\, t \geq 0 \tag{3.1}$$

for some set $B \in \mathcal{B}_G$ implies that $\lambda_G(B) \in \{0, 1\}$. (Note that Lemma 3.1 is required only for the "if" part; the "only if" part is straightforward, cf. Theorem 3 in [16] and Theorem 1 in [25].)

With these preparations, the asserted implication (i)⇒(ii) will now be proved. Thus assume (i); that is, $\overline{\bigcup_{t \geq 0} S_t} = G$. The key step in establishing (ii) is to check that the semi-flow $(R_t)_{t\geq 0}$ is ergodic in this case. Assume, therefore, that (3.1) holds for some set $B \in \mathcal{B}_G$. Note that then

$$\mathbb{P}\{\lambda_G(BX_0^{-1}X_{t_n}\triangle B) = 0 \,\forall\, n \in \mathbb{N}\} = 1$$

holds for every sequence $(t_n)_{n\in\mathbb{N}}$ in $[0, +\infty)$. Specifically, choose $(t_n)_{n\in\mathbb{N}}$ such that $\bigcup_{n\in\mathbb{N}} S_{t_n}$ is dense in G. (This is possible due to the separability of G.) By Fubini's Theorem, $\lambda_G(Bh\triangle B) = 0$ holds for all h in a dense subset H_0 of G. To establish the ergodicity of $(R_t)_{t\geq 0}$, it remains to demonstrate how this last conclusion implies that $\lambda_G(B) \in \{0, 1\}$. Assume, therefore, that $\lambda_G(B) > 0$ and let ν be the normalized restriction of λ_G to B; that is,

$$\nu(A) := \frac{\lambda_G(B \cap A)}{\lambda_G(B)} \quad \forall\, A \in \mathcal{B}_G.$$

For $g \in G$, denote by T_g the right-translation by g. Notice that, for every $h \in H_0$,

$$\begin{aligned}\nu \circ T_h^{-1}(A) &= \frac{\lambda_G(B \cap Ah^{-1})}{\lambda_G(B)} = \frac{\lambda_G(Bh \cap A)}{\lambda_G(B)} \\ &= \nu(A) + \frac{\lambda_G(Bh \cap A) - \lambda_G(B \cap A)}{\lambda_G(B)},\end{aligned}$$

and hence

$$\begin{aligned}|\nu \circ T_h^{-1}(A) - \nu(A)| &= \frac{|\lambda_G(Bh \cap A) - \lambda_G(B \cap A)|}{\lambda_G(B)} \\ &\leq \frac{\lambda_G(Bh\triangle B)}{\lambda_G(B)} = 0,\end{aligned}$$

showing that $\nu \circ T_h^{-1} = \nu$. Given any $g \in G$ and $\varphi \in C(G)$, pick a sequence $(h_n)_{n\in\mathbb{N}}$ in H_0 such that $\lim_{n\to\infty} h_n = g$. Since $\varphi \circ T_{h_n} \to \varphi \circ T_g$ uniformly on G, it follows by dominated convergence that

$$\begin{aligned}\int_G \varphi(x)\mathrm{d}\nu \circ T_g^{-1}(x) &= \int_G \varphi(xg)\mathrm{d}\nu(x) = \lim_{n\to\infty} \int_G \varphi(xh_n)\mathrm{d}\nu(x) \\ &= \lim_{n\to\infty} \int_G \varphi(x)\mathrm{d}\nu(x) = \int_G \varphi(x)\mathrm{d}\nu(x).\end{aligned}$$

Thus, ν is invariant under *all* right-translations, and consequently $\nu = \lambda_G$. In particular, $\lambda_G(B) = 1$, as required. Hence, the semi-flow $(R_t)_{t\geq 0}$ is ergodic.

By the Birkhoff Ergodic Theorem, for every integrable $\varphi : G \to \mathbb{C}$,

$$\frac{1}{T}\int_0^T \varphi(g\gamma(0)^{-1}\gamma(t))\mathrm{d}t \xrightarrow{T\to+\infty} \int_{G\times D} \varphi(g')\mathrm{d}\rho(g', \gamma') = \int_G \varphi(g')\mathrm{d}\lambda_G(g')$$

holds for ρ-a.e. $(g, \gamma) \in G \times D$. In probabilistic terms, this means that

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T \varphi(\xi X_0^{-1}X_t)\mathrm{d}t = \int_G \varphi\, \mathrm{d}\lambda_G \tag{3.2}$$

holds with probability one. For any $\varphi \in C(G)$, $g \in G$ and $n \in \mathbb{N}$, denote the set

$$\left\{\omega \in \Omega : \limsup_{T\to+\infty}\left|\frac{1}{T}\int_0^T \varphi(gX_t(\omega))\mathrm{d}t - \int_G \varphi\,\mathrm{d}\lambda_G\right| < \frac{1}{n}\right\} \in \mathcal{F}$$

by $\Omega_{\varphi,g,n}$. As ξX_0^{-1} is Haar-distributed in G, $1 = \int_G \mathbb{P}(\Omega_{\varphi,g,n})\mathrm{d}\lambda_G(g)$ for every n by the above, and so $\mathbb{P}(\Omega_{\varphi,g,n}) = 1$ for λ_G-almost every $g \in G$. If $\lambda_G(\{e_G\}) > 0$ or, equivalently, if G is finite, then $\mathbb{P}(\Omega_{\varphi,e_G,n}) = 1$ for all n, and consequently $\mathbb{P}(\bigcap_n \Omega_{\varphi,e_G,n}) = 1$. If, on the other hand, e_G is not an atom of λ_G then, by the uniform continuity of φ, there exists a sequence $(g_n)_{n\in\mathbb{N}}$ in G with $\lim_{n\to\infty} g_n = e_G$ such that $\mathbb{P}(\Omega_{\varphi,g_n,2n}) = 1$ and $\Omega_{\varphi,g_n,2n} \subseteq \Omega_{\varphi,e_G,n}$ for all n. From

$$1 = \mathbb{P}\left(\bigcap_n \Omega_{\varphi,g_n,2n}\right) \le \mathbb{P}\left(\bigcap_n \Omega_{\varphi,e_G,n}\right) \le 1,$$

it is clear that also in this case

$$\lim_{T\to+\infty}\frac{1}{T}\int_0^T \varphi(X_t)\mathrm{d}t = \int_G \varphi\,\mathrm{d}\lambda_G \quad \text{with probability one.} \tag{3.3}$$

Finally, recall that $C(G)$ is separable, and hence taking the intersection of (3.3) over a dense family $\{\varphi_n : n \in \mathbb{N}\}$ in $C(G)$ yields that the paths $t \mapsto X_t$ are, with probability one, c.u.d. in G. Thus, (i) implies (ii).

To show the reverse implication (ii) $\Rightarrow$ (i), suppose that (i) does not hold. In this case, Lemma 3.2 shows that $H_X := \overline{\bigcup_{t\ge 0} S_t}$ is a proper (compact) subgroup of G. It then follows from the first part of the proof that the paths of $(X_0^{-1}X_t)_{t\ge 0}$ are c.u.d. in H_X. Thus,

$$\lim_{T\to+\infty}\frac{1}{T}\int_0^T \varphi(X_t)\mathrm{d}t = \int_{H_X} \varphi(X_0 h)\mathrm{d}\lambda_{H_X}(h) \quad \forall\, \varphi \in C(G) \tag{3.4}$$

holds with probability one. It is straightforward to see that, no matter what the distribution of X_0 is, for some suitable choice of $\varphi \in C(G)$ the integral on the right of (3.4) will not almost surely equal $\int_G \varphi(g)\mathrm{d}\lambda_G(g)$. □

The following two somewhat technical lemmas have been relied on in the proof of Theorem 3.1.

Lemma 3.1. *With the notation used in the proof of Theorem* 3.1, *let the* $\mathcal{B}_G \otimes \mathcal{B}_D$*-measurable function* $\psi : G \times D \to \mathbb{R}$ *be invariant under the semi-flow* $(R_t)_{t\ge 0}$; *that is, assume that* $\psi \circ R_t = \psi$ *holds* ρ*-a.e. for all* $t \ge 0$. *Then there exists a* $\mathcal{B}_G$*-measurable function* $\overline{\psi} : G \to \mathbb{C}$ *such that* $\psi(g,\gamma) = \overline{\psi}(g)$ *for* ρ*-a.e.* $(g,\gamma) \in G \times D$.

Proof. The argument given here mimics the proof of Theorem 1 in [25]. Write expectations and conditional expectations with respect to ρ as $\rho[\bullet]$ and $\rho[\bullet \,|\, \bullet]$, respectively. Put $\mathcal{G} := \mathcal{B}_G \otimes \{\emptyset, D\}$ and $\mathcal{H} := \{\emptyset, G\} \otimes \mathcal{B}_D$. Since there exists, for any bounded $\mathcal{B}_G \otimes \mathcal{B}_D$-measurable function ψ, a $\mathcal{B}_G$-measurable function $\overline{\psi} : G \to \mathbb{R}$ such that $\rho[\psi \,|\, \mathcal{G}](g, \gamma) = \overline{\psi}(g)$ for ρ-a.e. $(g, \gamma) \in G \times D$, it suffices to show for any bounded invariant function ψ that $\psi = \rho[\psi \,|\, \mathcal{G}]$ ρ-a.e. By a monotone class argument, this is equivalent to showing for any bounded invariant function ψ, any bounded $\mathcal{G}$-measurable function α, and any bounded $\mathcal{H}$-measurable function β that

$$\rho[\psi\alpha\beta] = \rho[\rho[\psi \,|\, \mathcal{G}]\alpha\beta]. \tag{3.5}$$

Note that due to the independence of the sub-σ-algebras $\mathcal{G}$ and $\mathcal{H}$ under ρ the right-hand side of (3.5) is

$$\rho[\rho[\psi \,|\, \mathcal{G}]\alpha]\rho[\beta] = \rho[\psi\alpha]\rho[\beta],$$

and so it further suffices to show for any bounded $\mathcal{H}$-measurable function β with $\rho[\beta] = 0$ that $\rho[\psi\alpha\beta] = 0$ for any bounded $\mathcal{G}$-measurable function α. Since

$$\rho[\psi\alpha\beta] = \rho[\rho[\psi\beta \,|\, \mathcal{G}]\alpha],$$

this is equivalent to establishing

$$\rho[|\rho[\psi\beta \,|\, \mathcal{G}]|] = 0. \tag{3.6}$$

Note also that $\rho[|\rho[\psi\beta \,|\, \mathcal{G}]|] = \rho[|\rho[\psi\beta \,|\, \mathcal{G}] \circ R_t|]$ for all $t \geq 0$ because R_t preserves the measure ρ. Moreover, observe that

$$\rho[\psi\beta \,|\, \mathcal{G}](g, \gamma) = \mathbb{E}[\psi(g, X_0^{-1}X_\bullet)\overline{\beta}(X_0^{-1}X_\bullet)],$$

where $\overline{\beta}$ is a $\mathcal{B}_D$-measurable function such that $\beta(g, \gamma) = \overline{\beta}(g)$ for ρ-a.e. $(g, \gamma) \in B \times D$. By definition of R_t and the stationary increments property of X,

$$\begin{aligned} \rho[\psi\beta \,|\, \mathcal{G}] \circ R_t(g, \gamma) &= \mathbb{E}[\psi(g\gamma(0)^{-1}\gamma(t), X_0^{-1}X_\bullet)\overline{\beta}(X_0^{-1}X_\bullet)] \\ &= \mathbb{E}[\psi(g\gamma(0)^{-1}\gamma(t), X_t^{-1}X_{t+\bullet})\overline{\beta}(X_t^{-1}X_{t+\bullet})]. \end{aligned}$$

Since ψ is invariant, $\psi(g, \gamma) = \psi(g\gamma(0)^{-1}\gamma(t), \gamma(t)^{-1}\gamma(t + \bullet))$, and hence

$$\rho[\psi\beta \,|\, \mathcal{G}] \circ R_t(g, \gamma) = \mathbb{E}[\psi(g, X_\bullet^{(\gamma,t)})\overline{\beta}(X_t^{-1}X_{t+\bullet})],$$

where

$$X_s^{(\gamma,t)} := \begin{cases} \gamma(s) & \text{if } 0 \leq s < t, \\ \gamma(t)X_t^{-1}X_s & \text{if } s \geq t. \end{cases}$$

For every $t \geq 0$, let $\mathcal{H}_t$ be the sub-σ-algebra of $\mathcal{B}_G \otimes \mathcal{B}_D$ generated by the maps $(g, \gamma) \mapsto \gamma(s)$, $0 \leq s \leq t$, and denote, as usual, by $\mathcal{G} \vee \mathcal{H}_t$ the σ-algebra generated by $\mathcal{G} \cup \mathcal{H}_t$. Since $\rho[\psi \,|\, \mathcal{G} \vee \mathcal{H}_t](g, X_\bullet^{(\gamma,t)}) = \rho[\psi \,|\, \mathcal{G} \vee \mathcal{H}_t](g, \widetilde{\gamma})$ for any $\widetilde{\gamma} \in D$

such that $\widetilde{\gamma}(s) = \gamma(s)$ for $0 \le s \le t$, it follows that

$$\begin{aligned}\mathbb{E}[\rho[\psi \,|\, \mathcal{G} \vee \mathcal{H}_t](g, X_\bullet^{(\gamma,t)})\overline{\beta}(X_t^{-1}X_{t+\bullet})] &= \rho[\psi \,|\, \mathcal{G} \vee \mathcal{H}_t](g,\gamma)\mathbb{E}[\overline{\beta}(X_t^{-1}X_{t+\bullet})] \\ &= \rho[\psi \,|\, \mathcal{G} \vee \mathcal{H}_t](g,\gamma)\rho[\beta] \\ &= 0.\end{aligned}$$

Also, by the independent increments property of X and the Martingale Convergence Theorem,

$$\begin{aligned}&\int_{G\times D} \mathbb{E}[|\rho[\psi \,|\, \mathcal{G} \vee \mathcal{H}_t](g, X_\bullet^{(\gamma,t)}) - \psi(g, X_\bullet^{(\gamma,t)})|]\mathrm{d}\rho(g,\gamma) \\ &\quad = \int_{G\times D} |\rho[\psi \,|\, \mathcal{G} \vee \mathcal{H}_t](g,\gamma) - \psi(g,\gamma)|\mathrm{d}\rho(g,\gamma) \xrightarrow{t\to+\infty} 0.\end{aligned}$$

Hence,

$$\begin{aligned}\rho[|\rho[\psi\beta \,|\, \mathcal{G}]|] &= \rho[|\rho[\psi\beta \,|\, \mathcal{G}] \circ R_t|] \\ &= \int_{G\times D} |\mathbb{E}[\psi(g, X_\bullet^{(\gamma,t)})\overline{\beta}(X_t^{-1}X_{t+\bullet})]|\mathrm{d}\rho(g,\gamma) \\ &= \int_{G\times D} |\mathbb{E}[\psi(g, X_\bullet^{(\gamma,t)})\overline{\beta}(X_t^{-1}X_{t+\bullet})] \\ &\qquad - \mathbb{E}[\rho[\psi \,|\, \mathcal{G} \vee \mathcal{H}_t](g, X_\bullet^{(\gamma,t)})\overline{\beta}(X_t^{-1}X_{t+\bullet})]|\mathrm{d}\rho(g,\gamma) \\ &\le \|\beta\|_\infty \int_{G\times D} \mathbb{E}[|\psi(g, X_\bullet^{(\gamma,t)}) \\ &\qquad - \rho[\psi \,|\, \mathcal{G} \vee \mathcal{H}_t](g, X_\bullet^{(\gamma,t)})|]\mathrm{d}\rho(g,\gamma) \xrightarrow{t\to+\infty} 0,\end{aligned}$$

which in turn shows that (3.6) holds. Therefore, for each invariant function ψ there does indeed exist a $\mathcal{B}_G$-measurable function $\overline{\psi} : G \to \mathbb{R}$ such that $\psi(g,\gamma) = \overline{\psi}(g)$ for ρ-a.e. $(g,\gamma) \in G \times D$. □

Lemma 3.2. *For every Lévy process in the metrizable compact group G, the set $\overline{\bigcup_{t\ge0} S_t}$ is a subgroup of G.*

Proof. For convenience, denote $\overline{\bigcup_{t\ge0} S_t}$ by H_X. Note that $\mu_{t_1} * \mu_{t_2} = \mu_{t_1+t_2}$ implies

$$S_{t_1}S_{t_2} := \{g_1g_2 : g_j \in S_{t_j} \text{ for } j = 1,2\} = S_{t_1+t_2} \quad \forall\, t_1, t_2 \ge 0.$$

It follows that $h_1h_2 \in H_X$ whenever $\{h_1, h_2\} \subset H_X$. Given any $h \in H_X$, therefore, $hH_X \subset H_X$. To see that $hH_X = H_X$, first choose a metric d on G that is invariant under all left- as well as right-translations. (Such a metric exists, see, for example, §0.6 of [33].) If $g \in H_X \setminus hH_X$, then $d(g, h^ng) \ge \min\{d(g, hh') : h' \in H_X\} > 0$ for every $n \in \mathbb{N}$, and in particular $d(h^mg, h^ng) = d(g, h^{n-m}g)$ is bounded away from zero for $m, n \in \mathbb{N}$, $n > m$. Consequently, the sequence $(h^ng)_{n\in\mathbb{N}}$ does not contain any convergent subsequence, contradicting the compactness of G. Hence $hH_X = H_X$, and it is clear that $\{e_G, h^{-1}\} \subset H_X$. Since $h \in H_X$ was arbitrary, the set H_X is indeed a subgroup. □

If $(X_t)_{t\geq 0}$ is a Lévy process with $X_0 = e_G$ and $h > 0$, then X_{nh} is, for every $n \in \mathbb{N}$, the product of n independent random variables, all of which have the same distribution as X_h. Conversely, let $(\xi_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence in the metrizable compact group G. Denote by S the support of the common distribution of the ξ_n, $n \geq 1$, and, for every $n \in \mathbb{N}$, let $S^n = \{g_1 \cdots g_n : g_j \in S \text{ for } j = 1, \ldots, n\}$. If $(\tau_n)_{n\in\mathbb{N}}$ is a sequence of i.i.d. exponential random variables that is independent of $(\xi_n)_{n\in\mathbb{N}}$, then the process $(X_t)_{t\geq 0}$ defined by $X_t = e_G$ for $0 \leq t < \tau_1$ and $X_t = \xi_1 \cdots \xi_n$ for $\tau_1 + \cdots + \tau_n \leq t < \tau_1 + \cdots + \tau_{n+1}$ is a Lévy process. The following discrete-time analogue is, by the latter observation, immediate from Theorem 3.1 and the Strong Law of Large Numbers applied to the random variables $(\tau_n)_{n\in\mathbb{N}}$.

Corollary 3.1. *Let $(\xi_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence in the metrizable compact group G with the common distribution having support S. Then the sequence $(\xi_1 \cdots \xi_n)_{n\in\mathbb{N}}$ is, with probability one, u.d. in G if and only if $\overline{\bigcup_{n\in\mathbb{N}} S^n} = G$.*

Example 3.1. Let $Y = (Y_t)_{t\geq 0}$ be a Lévy process in the (non-compact, Abelian) group $\mathbb{R}$ with the usual topology. According to the classical Lévy–Khintchine formula (see, for example, Theorem 1.2.14 in [2]), $\mathbb{E}[e^{iyY_t}] = e^{t\eta(y)}$ for all $t \geq 0$ and $y \in \mathbb{R}$, where

$$\eta(y) = i\beta y - \frac{1}{2}\sigma^2 y^2 + \int_{\mathbb{R}} (e^{ixy} - 1 - ixy\mathbf{1}_{(-1,1)}(x))\mathrm{d}\nu(x), \tag{3.7}$$

with $\beta \in \mathbb{R}$, $\sigma^2 \geq 0$, and ν a Borel measure on $\mathbb{R}$ that satisfies $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}} y^2 \wedge 1\, \mathrm{d}\nu(y) < +\infty$. The triple (β, σ^2, ν) uniquely determines Y.

For any $y \in \mathbb{R}$, denote by $\langle y \rangle$ the fractional part of y, that is, $\langle y \rangle = y - \lfloor y \rfloor$. Set $X_t = \langle Y_t \rangle$ for all $t \geq 0$. Clearly, X is a Lévy process in the compact (Abelian) group $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. From (3.7) and Theorem 3.1 it is readily deduced that the paths of X are, with probability one, c.u.d. in $\mathbb{T}$ unless simultaneously $\sigma^2 = 0$ (i.e. Y has no Gaussian component), $\nu(\mathbb{R} \setminus \frac{1}{m}\mathbb{Z}) = 0$ for some $m \in \mathbb{N}$ (i.e. ν is concentrated on the lattice $\frac{1}{m}\mathbb{Z} = \{\frac{k}{m} : k \in \mathbb{Z}\}$), and $\beta = \sum_{|k|<m} \frac{k}{m}\nu(\{\frac{k}{m}\})$. In the latter case, assuming that m is chosen to be minimal, the paths of X are c.u.d. in the closed subgroup $\langle \frac{1}{m}\mathbb{Z} \rangle$ of $\mathbb{T}$, that is, with probability one,

$$\lim_{T\to+\infty} \frac{1}{T} \int_0^T \varphi(X_t)\mathrm{d}t = \frac{1}{m} \sum_{j=0}^{m-1} \varphi\left(\frac{j}{m}\right) \quad \forall \varphi \in C(\mathbb{T}).$$

Note that with $\vartheta \in \mathbb{R}$ and $\mathbb{P}\{\xi_1 = \vartheta\} = 1$, Corollary 3.1 contains the well-known fact that $(\langle n\vartheta \rangle)_{n\in\mathbb{N}}$ is u.d. in $\mathbb{T}$ if and only if ϑ is irrational.

Example 3.2. If G is merely *locally* compact then (2.1) will usually not hold for almost all paths of a Lévy process X in G, even when both sides exist, perhaps for some appropriate subspace of $C(G)$. However, Example 3.1 can be extended in a way that not only highlights the role played by the compactness of G, but also provides a new perspective on Theorem 1 in [22].

Let $Y = (Y_t)_{t\geq 0}$ again be a Lévy process in $\mathbb{R}$, with characteristic triple (β, σ^2, ν), and fix a bounded continuous function $f : \mathbb{R} \to \mathbb{C}$. To avoid trivialities, assume f is non-constant. Theorem 3.1 can be used to show that $\lim_{T\to+\infty} \frac{1}{T}\int_0^T f(Y_t)\mathrm{d}t$ does exist with probability one, provided that f is *almost periodic* (or *a.p.* for short). Recall that f is a.p. if, for every $\varepsilon > 0$, there exist a set $P_\varepsilon \subset \mathbb{R}$ which is relatively dense (i.e. P_ε "has bounded gaps") such that

$$\sup_{y\in\mathbb{R}} |f(y+p) - f(y)| < \varepsilon \quad \forall\, p \in P_\varepsilon.$$

It is well known that f is a.p. if and only if the closure H_f of the family $\{f(y+\bullet) : y \in \mathbb{R}\}$ is compact in $C_b(\mathbb{R})$, the Banach space of bounded continuous complex-valued functions on $\mathbb{R}$ equipped with the supremum norm. (Usually, H_f is referred to as the *hull* of f, see, for example, [11].) Moreover, the average

$$A(f) := \lim_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} f(y)\mathrm{d}y$$

exists for every a.p. function f. The addition

$$f(y_1+\bullet) + f(y_2+\bullet) := f(y_1+y_2+\bullet) \quad \forall\, y_1, y_2 \in \mathbb{R},$$

extends continuously to H_f, turning the latter into a metrizable compact (Abelian) group. Clearly, $e_{H_f} = f$, and the Haar measure on H_f is uniquely determined by the requirement that

$$\int_{H_f} \varphi\, \mathrm{d}\lambda_{H_f} = A(\varphi \diamond f) \quad \forall\, \varphi \in C(H_f),$$

where $\varphi \diamond f$ denotes the a.p. function $y \mapsto \varphi\big(f(y+\bullet)\big)$. With these preparations, define a process X in H_f by simply setting $X_t = f(Y_t + \bullet)$ for all $t \geq 0$. It is readily confirmed that X is a Lévy process. (Note that Example 3.1 simply corresponds to the special case of f being periodic with period 1, in which case H_f is homeomorphic and isomorphic to $\mathbb{T}$, and $A(f) = \int_0^1 f(y)\mathrm{d}y$.) Observe that $H_X = H_f$ unless $\sigma^2 = 0$ and $\nu(\mathbb{R}\backslash a\mathbb{Z}) = 0$ for some $a > 0$. When $H_X = H_f$, Theorem 3.1 implies that, with probability one,

$$\frac{1}{T}\int_0^T \varphi(X_t)\mathrm{d}t = \frac{1}{T}\int_0^T \varphi(f(Y_t+\bullet))\mathrm{d}t \xrightarrow{T\to+\infty} \int_{H_f} \varphi\, \mathrm{d}\lambda_{H_f}$$
$$= A(\varphi \diamond f) \quad \forall\, \varphi \in C(H_f).$$

In particular, choosing $\varphi(g) := g(0)$ for all $g \in H_f$ yields $\varphi \diamond f = f$ and consequently

$$\frac{1}{T}\int_0^T f(Y_t)\mathrm{d}t \xrightarrow{T\to+\infty} A(f) \quad \text{with probability one.} \tag{3.8}$$

For every a.p. function f, therefore, (3.8) holds for any Lévy process Y on $\mathbb{R}$ provided that Y either has a nonzero Gaussian component or else the associated measure ν is not concentrated on a lattice.

Example 3.3. As an application of Theorem 3.1 and Corollary 3.1, let $b \geq 2$ be a positive integer and recall that a measurable function $f : [0, +\infty) \to \mathbb{R}$ is b-*Benford* if $\log_b |f|$ is c.u.d. in $\mathbb{T}$, where $\log_b$ denotes the base-b logarithm and the convention $\log_b 0 := 0$ is adopted for convenience, see [4] for background information and further details on the Benford property as well as its ramifications. Equivalently, the function f is b-Benford if

$$\lim_{T\to+\infty} \frac{\mathrm{Leb}\{t \in [0, T) : S_b(f(t)) \leq s\}}{T} = \log_b s \quad \forall\, s \in [1, b);$$

here $S_b(y)$, the base-b *significant* of $y \in \mathbb{R}$ is, by definition, the unique number in $\{0\} \cup [1, b)$ such that $|y| = S_b(y)b^k$ for some integer k. Similarly, a sequence $(y_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ is called b-Benford whenever the function $t \mapsto y_{\lfloor t\rfloor+1}$ is b-Benford.

Let $Y = (Y_t)_{t\geq0}$ be a Lévy process in $\mathbb{R}$ with characteristic triple (β, σ^2, ν) and, for any real constants $a \neq 0$, $c \neq 0$, and d, consider

$$X_t = ae^{cY_t+dt}, \quad t \geq 0.$$

Since $(cY_t + dt)_{t\geq0}$ is again a Lévy process, it follows from Theorem 3.1 that $t \mapsto X_t$ is b-Benford with probability one unless

$$\sigma^2 = 0 \quad \text{and} \quad \nu\left(\mathbb{R} \Big\backslash \frac{\ln b}{|c|m}\mathbb{Z}\right) = 0 \quad \text{for some } m \in \mathbb{N}. \tag{3.9}$$

Note that (3.9) does not hold if ν is non-atomic. In particular, the paths of any *geometric Brownian motion* (also referred to as a *Black–Scholes process*), corresponding to the case where Y is a standard Brownian motion, are almost surely b-Benford for all b. Similarly, if Y is a Poisson process then the paths $t \mapsto ae^{cY_t+dt}$ are, with probability one, b-Benford unless c is a rational multiple of $\ln b$, cf. [27].

For a discrete-time analogue of these observations, let $(\xi_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence in $\mathbb{R}$. By Corollary 3.1, the sequence $(\prod_{j=1}^n \xi_j)_{n\in\mathbb{N}}$ is b-Benford with probability zero or one, depending on whether or not the support of the distribution of $(\log_b |\xi_1|)$ is contained in $\frac{1}{m}\mathbb{Z}$ for some $m \in \mathbb{N}$, cf. [24].

4. Further Remarks and Observations

The following remarks aim at providing some background information that may help the reader putting the main results of this note, Theorem 3.1 and Corollary 3.1, in perspective.

Remark 4.1. The notion of continuous uniform distribution applies to all measurable functions (paths), not only to those that are rcll. Hence it may seem natural to consider the class of stochastic processes that arises by replacing assumption (iii) in Definition 2.2 with the weaker requirement that the map $(t, \omega) \mapsto X_t(\omega)$ be jointly measurable. As the following argument shows, nothing is gained from this seemingly greater generality.

By Theorem 1 in [9], joint measurability implies the existence of closed sets $F \subset [0,1]$ with Lebesgue measure arbitrarily close to 1 such that for all $\varepsilon > 0$

$$\lim_{\delta \downarrow 0} \sup_{|t_2 - t_1| \le \delta, t_1, t_2 \in F} \mathbb{P}\{d(X_{t_1}, X_{t_2}) > \varepsilon\} = 0,$$

where d is a translation-invariant metric on G. Observe that if $t_2 > t_1$, then

$$\begin{aligned}\mathbb{P}\{d(X_{t_1}, X_{t_2}) > \varepsilon\} &= \mathbb{P}\{d(e_G, X_{t_1}^{-1} X_{t_2}) > \varepsilon\} \\ &= \mu_{t_2 - t_1}(\{g \in G : d(e_G, g) > \varepsilon\}).\end{aligned}$$

A celebrated theorem of Steinhaus (Théorème VIII in [31], see also [5, 17]), asserts that the set $\{t_2 - t_1 : t_1, t_2 \in F\}$ contains an open neighborhood of 0 whenever F has positive Lebesgue measure. Thus, μ_t converges to ϵ_{e_G} as $t \downarrow 0$. Now, define a strongly continuous contraction semigroup of operators $(P_t)_{t \ge 0}$ on $C(G)$ by setting $P_0 \varphi = \varphi$ and

$$P_t \varphi = \int_G \varphi(\bullet g) \mathrm{d}\mu_t(g) \quad \forall\, t > 0 \text{ and } \varphi \in C(G).$$

It follows from the theory of Feller semigroups (see, for example, Sec. III.7 of [23]) that by modifying each random variable X_t on a $\mathbb{P}$-null set it is possible to produce a stochastic process with rcll paths. By Fubini's Theorem, the value of $\int_0^T \varphi(X_t)\mathrm{d}t$, $T \ge 0$, remains unchanged if X is replaced by such an rcll modification.

Remark 4.2. Lévy processes are a special class of Markov processes, and so it is natural to inquire whether Theorem 3.1 is a consequence of more general results in the vast Markov process literature. The proof given above certainly uses the extra Lévy structure: The fact that the distribution of $(X_t)_{t \ge 0}$ when $X_0 = g$ is equal to the distribution of $(gX_t)_{t \ge 0}$ when $X_0 = e_G$ reduces checking the ergodicity of $(\xi X_0^{-1} X_t)_{t \ge 0}$ to verifying a criterion involving only the right-translations by $h \in H_X$. Moreover, the fact that the state space is a compact group permits the latter verification to be reduced to the uniqueness of normalized Haar measure on such a group.

A discussion of limit theorems for the occupation measures of discrete-time Markov processes is given in Chap. 17 of [21] under the assumption that the process is *Harris recurrent*. Similar results for continuous-time processes are obtained in Paragraph II of [3] by the device of sampling the process at the arrival times of a Poisson process to obtain a discrete-time process. The condition $H_X = G$ in Theorem 3.1 is easily seen to be equivalent to the condition that $\int_0^{+\infty} \mu_t(U)\mathrm{d}t > 0$ for all non-empty open sets $U \subset G$. If this condition is replaced by the stronger assumption that $\int_0^{+\infty} \mu_t(B)\mathrm{d}t > 0$ for all $B \in \mathcal{B}_G$ with $\lambda_G(B) > 0$, then it is possible to conclude from the results in [3, 21] that $\lim_{T \to +\infty} \frac{1}{T} \int_0^T \varphi(X_t)\mathrm{d}t = \int_G \varphi\, \mathrm{d}\lambda_G$

almost surely for any bounded measurable function φ. Note that if $Y = (Y_t)_{t\geq 0}$ is as in Example 3.1 with $(\beta, \sigma^2, \nu) = (0, 0, \epsilon_\vartheta)$ for some irrational $\vartheta \in \mathbb{R}$, then $X = (X_t)_{t\geq 0}$ defined by $X_t = \langle Y_t \rangle$ satisfies the condition $H_X = \mathbb{T}$, yet $\int_0^\infty \mathbb{P}\{X_t \in B\}\mathrm{d}t = 0$ when B is the complement of $\{\langle n\vartheta \rangle : n \in \mathbb{N}\} \cup \{0\}$ in $\mathbb{T}$, a set with full $\lambda_\mathbb{T}$-measure.

Remark 4.3. The ergodicity of the stationary process $(\xi X_0^{-1} X_t)_{t\geq 0}$ appearing in the proof of Theorem 3.1 and that of its discrete-time analogue $(\xi\xi_1 \cdots \xi_n)_{n\in\mathbb{N}}$ can be established more easily if one assumes that, respectively, S_t for some $t > 0$ and the support S of the common distribution of the ξ_n, $n \geq 1$, are not contained in the coset of any proper closed normal subgroup of G. Under this additional assumption, it follows from the Itô–Kawada Theorem (see, for example, Theorem 2.1.4 in [13]) that, for any $t \geq 0$, the random variables $X_t^{-1} X_{t+T}$ converge in distribution to λ_G as $T \to +\infty$, and analogously, $\xi_n \xi_{n+1} \cdots \xi_{n+N}$ converges in distribution to λ_G as $N \to \infty$. As a consequence, for any $A, B \in \mathcal{B}_G$,

$$\mathbb{P}\{\xi X_0^{-1} X_t \in A, \xi X_0^{-1} X_{t+T} \in B\} \xrightarrow{T\to+\infty} \mathbb{P}\{\xi X_0^{-1} X_t \in A\}\mathbb{P}\{\xi X_0^{-1} X_t \in B\},$$

showing that the process $(\xi X_0^{-1} X_t)_{t\geq 0}$ is actually *mixing* in this case, and thus *a fortiori* ergodic, cf. [18, 33]. Mixing properties of the semi-flow $(R_t)_{t\geq 0}$ have been studied in [12]. By contrast, the proof of Theorem 3.1 presented above only uses the ergodicity of $(\xi X_0^{-1} X_t)_{t\geq 0}$. Ergodicity is all that can be hoped for in general: For example, take $G = \mathbb{T}$ and let $(\xi_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence with $\mathbb{P}\{\xi_1 = \langle \sqrt{2} \rangle\} = 1$. In this case, the process $(\xi + \xi_1 + \cdots + \xi_n) = (\langle \xi + n\sqrt{2} \rangle)$, though stationary and ergodic, is *not* mixing, since for $B = \{\langle t \rangle : 0 \leq t \leq \frac{1}{2}\} \in \mathcal{B}_\mathbb{T}$,

$$\begin{aligned}\limsup_{N\to\infty} \mathbb{P}\{\langle \xi + n\sqrt{2} \rangle \in B, \langle \xi + (n+N)\sqrt{2} \rangle \in B\} \\ = \frac{1}{2} \neq \frac{1}{4} = \mathbb{P}\{\langle \xi + n\sqrt{2} \rangle \in B\}^2\end{aligned}$$

holds for every $n \in \mathbb{N}$.

Remark 4.4. Every Hausdorff compact group G carries a unique normalized Haar measure. Hence the statement of Theorem 3.1 (respectively, Corollary 3.1) makes sense for every jointly measurable stochastic process X (respectively, sequence) in G. Even though it makes sense, however, it is not generally true: The asserted equivalence may break down whenever the increments of X are non-stationary or dependent, or if G fails to be metrizable. While the former two observations are quite obvious, to see what might go wrong when G is not metrizable, recall that G is metrizable if and only if $C(G)$ is separable. The proof of Theorem 3.1 given here uses the metrizability of G, the separability of $C(G)$, and the consequent separability of G, and so there is no hope that this proof will extend. As shown by Corollary 4.5.4 in [19], if a Hausdorff compact Abelian group G is not separable, then no sequence is u.d. in G. Any extension of Corollary 3.1, therefore, must require the separability of G (at least in the Abelian case).

Remark 4.5. As the following short, non-exhaustive compilation illustrates, special cases of Theorem 3.1 and Corollary 3.1 as well as related results have repeatedly appeared in the literature.

The earliest pertinent references the authors have been able to identify are the announcement of a *Random Ergodic Theorem* in [32] and its subsequent significant generalization in [16, 25]. In a purely probabilistic setting, [22] focuses on discrete-time processes taking values in $\mathbb{R}$ but also considers the case $G = \mathbb{T}$. In addition, extensions to compact groups and general continuous-time processes $X = (X_t)_{t\geq 0}$ in $\mathbb{R}$ are discussed briefly. For the latter, a sufficient condition for the almost sure continuous uniform distribution of paths is given under the assumption that

$$\mathbb{E}[e^{i\lambda(X_t - X_0)}] = \mathcal{O}(t^{-\delta}) \quad \text{as } t \to +\infty \tag{4.1}$$

for every real $\lambda \neq 0$ and the appropriate $\delta = \delta(\lambda) > 0$. Note that (4.1) holds for every nondegenerate Brownian motion in $\mathbb{R}$, but it does not hold if X is, for instance, a Poisson process, since in this case $|\mathbb{E}[e^{i\lambda(X_t - X_0)}]| = 1$ whenever $\lambda \in 2\pi\mathbb{Z}$. Theorem 3.1 replaces (4.1) with a necessary and sufficient condition.

The uniform distribution of Brownian paths on $\mathbb{R}$ has been established in [8] and subsequently in [14]. Building on these, in the case of Brownian motion on $\mathbb{R}$, [29] proves a law of the iterated logarithm for the deviations of $\frac{1}{T}\int_0^T \varphi(X_t)\mathrm{d}t$ from its expected value, and [6] study the same problem on compact connected Riemannian manifolds, while [20, 26] consider more general processes on $\mathbb{R}$. A sufficient condition for sequences of real-valued random variables with stationary, but not necessarily independent increments to be u.d. in $\mathbb{T}$ is derived in [15], and in [30] sums of i.i.d. random variables are considered under the perspective of rotation invariance.

It appears that the Benford property for paths of (some) Lévy processes has been studied only rather recently. Utilizing large deviation results, [27] essentially establishes the almost sure c.u.d. property of the paths of X for $G = \mathbb{T}$ with $X = \langle Y \rangle$, where Y is a continuous local martingale plus a deterministic drift. The most important example of this type is the standard Brownian motion, and the test function φ in (2.1) may be taken to be merely *measurable* and *bounded* in this case, i.e. $\varphi \in L^\infty(G)$ instead of $\varphi \in C(G)$, cf. Remark 4.2. (Notice that $L^\infty(G)$ is non-separable whenever G is infinite.) In [28], a similar approach is extended to general Lévy processes in $\mathbb{R}$. In this more general setting, however, the desired conclusion — the almost sure Benford property of paths — is obtained only under an additional regularity condition on the characteristic function of Y_1, referred to as "standard condition". Many Lévy processes, most importantly perhaps any Poisson process, do not satisfy this condition and hence are not amenable to the techniques of [28]. Although this may look like a minor technicality, it is not: As shown in Remark 4.2, there are Lévy processes X on $\mathbb{T}$ for which $\lim_{T\to+\infty} \frac{1}{T}\int_0^T \varphi(X_t)\mathrm{d}t = \int_{\mathbb{T}} \varphi \,\mathrm{d}\lambda_{\mathbb{T}}$ does not almost surely hold for some $\varphi \in L^\infty(\mathbb{T})$.

Acknowledgments

The first author is much indebted to T. Hill, V. Losert, V. Runde and B. Schmuland for many stimulating conversations and helpful suggestions. He also wishes to acknowledge that, prior to this note and independently of it, the Benford property for Lévy processes has been studied in the as yet unpublished treatise [28].

A.B. was supported by an NSERC Discovery Grant; S.N.E. was supported in part by NSF grant DMS-0907630.

References

1. D. Applebaum, Lévy processes — From probability to finance and quantum groups, *Notices Amer. Math. Soc.* **51** (2004) 1336–1347.
2. D. Applebaum, *Lévy Processes and Stochastic Calculus*, 2nd edn. (Cambridge Univ. Press, 2009).
3. J. Azéma, M. Kaplan-Duflo and D. Revuz, Mesure invariante sur les classes récurrentes des processus de Markov, *Z. Wahr. Verw. Gebiete* **8** (1967) 157–181.
4. A. Berger and T. P. Hill, A basic theory of Benford's law, *Probab. Surv.* **8** (2011) 1–126.
5. N. H. Bingham and A. J. Ostaszewski, Dichotomy and infinite combinatorics: The theorems of Steinhaus and Ostrowski, *Math. Proc. Cambridge Philos. Soc.* **150** (2011) 1–22.
6. M. Blümlinger, M. Drmota and R. F. Tichy, A uniform law of the iterated logarithm for Brownian motion on compact Riemannian manifolds, *Math. Z.* **201** (1989) 495–507.
7. I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai, *Ergodic Theory* (Springer, 1982).
8. C. Derman, Ergodic property of the Brownian motion process, *Proc. Nat. Acad. Sci. USA* **40** (1954) 1155–1158.
9. G. Di Nunno and Yu. A. Rozanov, On measurable modification of stochastic functions, *Theory Probab. Appl.* **46** (2002) 122–127.
10. S. N. Ethier and T. G. Kurtz, *Markov Processes* (John Wiley & Sons, 1986).
11. A. M. Fink, *Almost Periodic Differential Equations*, Lecture Notes in Mathematics, Vol. 377 (Springer, 1974).
12. H.-O. Georgii, Mixing properties of induced random transformations, *Ergodic Th. Dynam. Syst.* **17** (1997) 839–847.
13. H. Heyer, *Probability Measures on Locally Compact Groups* (Springer, 1977).
14. E. Hlawka, Ein metrischer Satz in der Theorie der C-Gleichverteilung, *Monatsh. Math.* **74** (1970) 108–118.
15. P. J. Holewijn, On the uniform distribution of sequences of random variables, *Z. Wahr. Verw. Gebiete* **14** (1969) 89–92.
16. S. Kakutani, Random ergodic theorems and Markoff processes with a stable distribution, in *Proc. of the Second Berkeley Symposium on Mathematical Statistics and Probability* (Univ. of California Press, 1951), pp. 247–261.
17. H. Kestelman, On the functional equation $f(x+y) = f(x) + f(y)$, *Fund. Math.* **34** (1947) 144–147.
18. U. Krengel, *Ergodic Theorems* (de Gruyter, 1985).
19. L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences* (John Wiley & Sons, 1974).
20. R. M. Loynes, Some results in the probabilistic theory of asymptotic uniform distribution modulo 1, *Z. Wahr. Verw. Gebiete* **26** (1973) 33–41.

21. S. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, 2nd edn. (Cambridge Univ. Press, 2009).
22. H. Robbins, On the equidistribution of sums of independent random variables, *Proc. Amer. Math. Soc.* **4** (1953) 786–799.
23. L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes, and Martingales*, Vol. 1 (John Wiley & Sons, 1994).
24. K. A. Ross, Benford's Law, a growth industry, *Amer. Math. Monthly* **118** (2011) 571–583.
25. C. Ryll-Nardzewski, On the ergodic theorems (III). The random ergodic theorem, *Studia Math.* **14** (1954) 298–301.
26. P. Schatte, The asymptotic uniform distribution modulo 1 of cumulative processes, *Optimization* **16** (1985) 783–786.
27. K. Schürger, Extensions of Black–Scholes processes and Benford's law, *Stoch. Process. Appl.* **118** (2008) 1219–1243.
28. K. Schürger, Lévy processes and Benford's Law, preprint (2011).
29. O. Stackelberg, A uniform law of the iterated logarithm for functions C-uniformly distributed mod 1, *Indiana Univ. Math. J.* **21** (1971) 515–528.
30. W. Stadje, On the asymptotic equidistribution of sums of independent identically distributed random variables, *Ann. Inst. H. Poincaré Probab. Statist.* **25** (1989) 195–203.
31. H. Steinhaus, Sur les distances des points dans les ensembles de mesure positive, *Fund. Math.* **1** (1920) 93–104.
32. S. M. Ulam and J. von Neumann, Random ergodic theorem, *Bull. Amer. Math. Soc.* **51** (1945) 660.
33. P. Walters, *An Introduction to Ergodic Theory* (Springer, 1982).