Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2012, Vol. 48, No. 2, 551–578
DOI: 10.1214/10-AIHP411


ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

# Dynamical attraction to stable processes[1]

Albert M. Fisher[a] and Marina Talet[b]

[a]*Dept Mat IME-USP, Caixa Postal 66281, CEP 05315-970 São Paulo, Brazil. E-mail: afisher@ime.usp.br; url: http://www.ime.usp.br/~afisher/*
[b]*C.M.I. Université de Provence LATP, CNRS-UMR 6632, F-13453 Marseille Cedex 13, France. E-mail: marina@cmi.univ-mrs.fr*

Received 22 March 2010; revised 27 December 2010; accepted 27 December 2010

**Abstract.** We apply dynamical ideas within probability theory, proving an almost-sure invariance principle in log density for stable processes. The familiar scaling property (self-similarity) of the stable process has a stronger expression, that the scaling flow on Skorokhod path space is a Bernoulli flow. We prove that typical paths of a random walk with i.i.d. increments in the domain of attraction of a stable law can be paired with paths of a stable process so that, after applying a non-random regularly varying time change to the walk, the two paths are forward asymptotic in the flow except for a set of times of density zero. This implies that a.e. time-changed random walk path is a generic point for the flow, i.e. it gives all the expected time averages. For the Brownian case, making use of known results in the literature, one has a stronger statement: the random walk and the Brownian paths are forward asymptotic under the scaling flow (now with no exceptional set of times), at an exponential rate given by the moment assumption.

**Résumé.** En appliquant des idées venues des systèmes dynamiques aux probabilités, nous prouvons un principe d'invariance presque sûr au sens de la densité logarithmique pour des processus stables. L'auto-similarité d'un processus stable revêt une expression plus forte, celle de la Bernoullicité du flot d'échelle agissant sur l'espace de Skorokhod des trajectoires. Nous montrons qu'il existe un couplage de la marche aléatoire à accroissements i.i.d. dans le domaine d'attraction d'une loi stable et d'un processus stable tel que presque sûrement, après un changement de temps déterministe et à variation régulière, sous l'action du flot d'échelle, les deux processus soient asymptotiques dans le futur sauf pour un ensemble de temps de densité nulle. Il en découle que presque toute marche (à un changement de temps près) est un point générique du flot. Dans le cas brownien, compte-tenu de résultats bien connus dans la littérature, nous avons un résultat plus fort : sous l'action du flot, les trajectoires de la marche et du brownien sont asymptotiques dans le futur avec une vitesse exponentielle donnée par l'hypothèse de moment.

*MSC:* 37A50; 60F17; 60G18; 60G52

*Keywords:* Brownian motion; Stable process; Almost-sure invariance principle in log density; Generic point; Pathwise Central Limit Theorem; Scaling flow

## 1. Introduction

In this paper we explore and bridge notions of attraction stemming from probability and dynamical systems theory.

Let $\nu$ be an invariant probability measure for a flow $\tau_t$ acting on a topological space $\Omega$. Hence, $\nu$ is a fixed point for the flow $\tau_t^*$, the induced action on the space of all probability measures on $\Omega$ defined by $\tau_t^*(\mu) = \mu \circ \tau_{-t}$.

The *stable manifold* of this fixed point $\nu$, written $W^s(\nu)$, is the set of probability measures $\mu$ such that $\tau_t^*(\mu)$ converges weakly (or in law) to $\nu$ as $t$ increases to infinity. We shall see that the stable manifold of the measure $\nu$ can be viewed as a dynamical counterpart of the domain of attraction of a law.

For a first example we take $\nu$ to be the Wiener measure; this is a probability measure on $\mathcal{C}$, the space of continuous functions from $\mathbb{R}^+$ to $\mathbb{R}$. With the topology of uniform convergence on compact subsets $\mathcal{C}$ is a Polish space, i.e.

[1]A. Fisher and M. Talet partially supported by CNPQ, FAPERGS, FAPESP, CNRS, Cooperation Franco-Brazil.

a separable topological space for which there exists a complete metric. By a theorem of Rochlin [30] therefore, the measure space $(\mathcal{C}, \nu)$ is a Lebesgue space (as it is measure-isomorphic to the unit interval with Lebesgue measure) which is a good situation for applying ideas from ergodic theory.

The Brownian self-similarity states that the Wiener measure $\nu$ is preserved by $\tau_t$, the scaling flow of index $1/2$ acting on $\Omega \equiv \mathcal{C}$ by:

$$(\tau_t f)(x) = \frac{f(\mathrm{e}^t x)}{\mathrm{e}^{t/2}}.$$

In fact, $\tau_t$ is an ergodic flow on $(\Omega, \nu)$ and is isomorphic to the translation (left-shift) flow on a stationary ergodic Gaussian process, the Ornstein–Uhlenbeck velocity process $V$ defined as $V(t) = B(\mathrm{e}^t)/\mathrm{e}^{t/2}$, where $B$ is a standard Brownian motion.

Now suppose $(X_i)$ is a sequence of independent and identically distributed (i.i.d.) random variables with finite second moment, centered and of variance unity. We embed the random walk path $S_n$ with $S_0 = 0$, $S_n = X_0 + \cdots + X_{n-1}$ for $n \geq 1$ in a continuous path $S(t) \in \mathcal{C}$ through polygonal interpolation between $S_n$ and $S_{n+1}$.

Donsker's Theorem, or the Functional Central Limit Theorem, see e.g. [20], p. 70, states that the rescaled random walk paths $(S(n\cdot)/\sqrt{n})$ converge weakly to a Brownian motion as $n \to \infty$. This also holds for rescaling continuously instead of by $n$.

Denoting by $\mu$ the measure on $\Omega = \mathcal{C}$ corresponding to the polygonal process $(S(\cdot))$, Donsker's Theorem says, precisely, that for each $\varphi \in \mathrm{CB}(\Omega)$, the continuous and bounded real-valued functions,

$$\big\langle \varphi, \tau_t^*(\mu) \big\rangle \equiv \int_\Omega \varphi \, \mathrm{d}\big(\tau_t^*(\mu)\big) = \int_\Omega \varphi(\tau_t x) \, \mathrm{d}\mu(x) \to \int_\Omega \varphi \, \mathrm{d}\nu = \langle \varphi, \nu \rangle \quad \text{as } t \to +\infty. \tag{1.1}$$

In dynamical terms, this says that the probability measures $\tau_t^*(\mu)$ converge weakly to $\nu$ as $t$ goes to infinity, written $\tau_t^*(\mu) \Rightarrow \nu$; in other words, the polygonal random walk measure $\mu$ is in the stable manifold $W^{\mathrm{s}}(\nu)$.

This notion of attraction is defined through convergence of space averages, which raises the question of time averages.

Space averages meet time averages in Birkhoff's ergodic theorem. A strong form of this is given by Fomin [15], who proved that for a continuous flow $\tau_t$ on a Polish space with ergodic invariant probability measure $\nu$, then for $\nu$-almost every $x$,

$$\frac{1}{T} \int_0^T \varphi(\tau_t x) \, \mathrm{d}t \to \langle \varphi, \nu \rangle \quad \text{as } T \to +\infty \text{ for all } \varphi \in \mathrm{CB}(\Omega) \tag{1.2}$$

and so equivalently in terms of weak convergence:

$$\frac{1}{T} \int_0^T \delta_{\tau_t x} \, \mathrm{d}t \Rightarrow \nu \quad \text{as } T \to \infty,$$

where $\delta$ denotes a Dirac mass.

In ergodic theory terminology, an element $x$ of $\Omega$ satisfying (1.2) is said to be a *generic point* for the measure $\nu$; so Fomin's theorem says exactly that $\nu$-almost every $x$ in $\Omega$ is a generic point for $\nu$.

By combining the previous notion of attraction with Cesáro time averaging we are led to a weaker notion of attraction. Writing $\mu_T \equiv \frac{1}{T}\int_0^T \tau_t^*(\mu) \, \mathrm{d}t$, and $\varphi_T = \frac{1}{T}\int_0^T \varphi \circ \tau_t \, \mathrm{d}t$, we say that $\tau_t^*(\mu)$ Cesáro-weakly converges to $\nu$, in short $\tau_t^*(\mu) \Rightarrow \nu$ (Cesáro), if and only if given any $\varphi \in \mathrm{CB}(\Omega)$,

$$\langle \varphi, \mu_T \rangle = \frac{1}{T} \int_0^T \big\langle \varphi, \tau_t^*(\mu) \big\rangle \mathrm{d}t = \frac{1}{T} \int_0^T \langle \varphi \circ \tau_t, \mu \rangle \, \mathrm{d}t = \langle \varphi_T, \mu \rangle \to \langle \varphi, \nu \rangle \quad \text{as } T \to \infty. \tag{1.3}$$

The set of all such $\mu$'s, the *Cesáro stable manifold* of $\nu$, written $W^{\mathrm{s}}_{\mathrm{Ces}}(\nu)$, contains $W^{\mathrm{s}}(\nu)$.

Now if $\mu$-almost every $x$ is a generic point for $\nu$, we have $\varphi_T \to \langle \varphi, \nu \rangle$ a.s., and so by the Lebesgue Dominated Convergence Theorem (henceforth LDCT), $\langle \varphi_T, \mu \rangle \to \langle \varphi, \nu \rangle$. Then from the above expression, $\mu_T \Rightarrow \nu$ and $\mu$ belongs to $W^{\mathrm{s}}_{\mathrm{Ces}}(\nu)$.

We recall that to show weak convergence it is equivalent to check this on *uniformly* continuous functions, since by [4], p. 12, having convergence for each $\varphi$ in $\mathrm{CB}(\Omega)$ is equivalent to convergence for each $\varphi$ in the space $\mathrm{UCB}(\Omega, d)$ of uniformly continuous and bounded functions for any chosen metric $d$ which gives the topology.

A link between metric and measure comes from the notion of the stable manifold of a point $x$, now in $\Omega$ rather than the space of measures on $\Omega$ and in general no longer a fixed point. By definition $W^{\mathrm{s},d}(x)$ is the set of all points $y \in \Omega$ which are $d$-forward asymptotic to $x$, i.e. such that $d(\tau_t x, \tau_t y) \to 0$ as $t \to \infty$; we observe that if $x$ is a generic point, then any such $y$ will have the same time averages for $\varphi \in \mathrm{UCB}(\Omega, d)$, and so as just remarked this will pass over to $\mathrm{CB}(\Omega)$. Hence $y \in W^{\mathrm{s},d}(x)$ also will be a generic point.

The use of time averages again leads us to a weaker notion: if there exists a set of times $\mathcal{B} = \mathcal{B}_{(x,y)}$ of Cesáro density zero such that

$$d(\tau_t x, \tau_t y) \to 0, \quad t \to \infty, t \notin \mathcal{B},$$

then we abbreviate this as $\lim_{t\to\infty} d(\tau_t x, \tau_t y) = 0$ (Cesáro), and write $W^{\mathrm{s},d}_{\mathrm{Ces}}(x)$ for the set of such points $y$. Note that $y \in W^{\mathrm{s},d}_{\mathrm{Ces}}(x)$ has the same ergodic averages as $x$ for any $\varphi \in \mathrm{UCB}(\Omega, d)$; this again passes to $\mathrm{CB}(\Omega)$, so the generic point property is true for all of $W^{\mathrm{s},d}_{\mathrm{Ces}}(x)$ as well.

The main focus of this paper will be on the stable manifolds $W^{\mathrm{s},d}(x)$ and the larger sets $W^{\mathrm{s},d}_{\mathrm{Ces}}(x)$ for certain flows arising in probability theory, and in particular on the link between stable manifolds and *almost sure invariance principles*, hereafter abbreviated *a.s.i.p.*'s.

We illustrate this again with the Brownian case, adopting the same notation as before.

In the case where the common law $F$ of the i.i.d. sequence $X_i$ has finite second moment (equal to one), Strassen's a.s.i.p. (see [35,36]) states that a standard Brownian motion $B$ and the polygonal random walk $S$ can be redefined to live on the same probability space, in such a way that

$$\left|S(n) - B(n)\right| = \mathrm{o}\big(\sqrt{n \log\log n}\big) \quad \text{a.s.,}$$

where $f(t) = \mathrm{o}(g(t))$ means that $f(t)/g(t) \to 0$ as $t$ goes to infinity.

Assuming that $F$ has finite $r$th moment for some $r > 2$ (and is centered with variance one), Strassen's bound was improved by Breiman [6] to $\mathrm{o}(n^{1/r}\sqrt{\log n})$, which is stronger than

$$\left|S(n) - B(n)\right| = \mathrm{o}\big(\sqrt{n}\big) \quad \text{a.s.}$$

This extends to continuous time:

$$\|S - B\|^{\infty}_{[0,T]} \stackrel{\text{def}}{=} \sup_{t\in[0,T]} \left|S(t) - B(t)\right| = \mathrm{o}\big(\sqrt{T}\big) \quad \text{a.s.} \tag{1.4}$$

Defining $d^u_1$ on path space $\mathcal{C}$ by $d^u_1(f, g) = \|f - g\|^{\infty}_{[0,1]}$, then one has this equivalent dynamical version of (1.4): there exists a joining (or coupling, see Definition 3.7) of the two processes $S$ and $B$ such that for almost every pair $(S, B)$ with respect to the joining measure,

$$\lim_{t\to+\infty} d^u_1(\tau_t S, \tau_t B) = 0, \tag{1.5}$$

where as before $\tau_t$ denotes the scaling flow of index 1/2. We would like to have the similar statement for an actual metric (rather than pseudometric) on $\mathcal{C}$ which gives the topology of uniform convergence on compacts; to do this we set first $d^u_t(f, g) = \|f - g\|^{\infty}_{[0,t]}$ and then define:

$$d^u_\infty(f, g) = \int_0^{+\infty} \mathrm{e}^{-t} \frac{d^u_t(f, g)}{1 + d^u_t(f, g)}\, \mathrm{d}t. \tag{1.6}$$

One can verify that $d^u_\infty$ is complete and that (1.5) holds also for this metric.

So dynamically speaking, (1.4) says that $S$ is in the $d^u_1$- and hence $d^u_\infty$-stable manifold of $B$.

Now by Fomin's theorem applied to the flow $\tau_t$, for almost every path $B$, the time average for any $\varphi$ in $\mathrm{CB}(\mathcal{C})$ equals the space average (the expected value) $\langle \varphi, \nu\rangle$; making use of $\mathrm{UCB}(\mathcal{C}, d^u_\infty)$, this passes to the rest of the stable manifold, in particular to $S$. We conclude that under the assumption of finite $r$th moment, $\mu$-almost every path $S$ is a generic point for the Wiener measure $\nu$.

In fact in this case of higher than second moments, Breiman's upper bound can be improved still further: Komlós, Major and Tusnády [21,22] and Major [26], see also [8], pp. 107 and 108, were able to demonstrate a bound of $o(n^{1/r})$. This yields the following dynamical statement: there exists a joining of the polygonal paths $S$ and a standard Brownian process $B$, such that for almost every pair $(S, B)$,

$$d_1^u(\tau_t S, \tau_t B) = o\big(e^{(1/r-1/2)t}\big)$$

and also for the metric $d_\infty^u$.

We mention that one can embed the random walk $S_n$ in a second (discontinuous) path by $\overline{S}(t) = S_{[t]}$ and that the previous results also hold for this step path extension, though the polygonal extension $S(t)$ is more appropriate in this context as it belongs to the space $\mathcal{C}$; step path extensions will be more natural below, when we deal with stable non-Gaussian processes.

There remains the intriguing question as to whether Strassen's bound $o(\sqrt{n \log\log n})$ can also be improved when $F$ has finite second moment but all higher moments are infinite. However counterexamples, first by Breiman [6] and then by Major in [25] showed that Strassen's upper bound is indeed sharp; see especially [8], p. 93. We draw the following dynamical conclusion from this result: there exists $F$ (centered and with variance 1) such that for any Brownian motion $B$ and any joining of $S$ and $B$, then for almost every pair $(S, B)$,

$$\limsup_{t\to+\infty} d_1^u\big(\tau_t S, \tau_t B\big) = +\infty \tag{1.7}$$

and similarly for $d_\infty^u$.

How, then, can we understand this apparent discrepancy between Donsker's theorem, which tells us that $\mu \in W^{\mathrm{s}}(\nu)$, and Strassen's sharp upper bound, which says that $S$ does not belong to the stable manifold of $B$, $W^{\mathrm{s},d_\infty^u}(B)$?

An explanation is given in [11], where we proved by way of Skorokhod's embedding that in the case where $F$ has finite second moment, there exists a joining of $B$ and $S$ such that for almost every pair $(S, B)$,

$$d_1^u(\tau_t S, \tau_t B) = \|\tau_t S - \tau_t B\|_{[0,1]}^\infty \to 0 \quad \text{(Cesáro)}, \tag{1.8}$$

that is, convergence takes place off a set of times $\mathcal{B}$ of Cesáro density zero. The same holds for the metric $d_\infty^u$ (see the proof of Lemma 3.6). So $S$ belongs to $W_{\mathrm{Ces}}^{\mathrm{s},d_\infty^u}(B)$.

Statement (1.8) gives, after an exponential change of variables:

$$\|S - B\|_{[0,T]}^\infty = o\big(\sqrt{T}\big) \quad \text{(log)}, \tag{1.9}$$

which means that the convergence holds off a set (this is just $\exp(\mathcal{B})$) which has log density zero. We call (1.9) an *almost-sure invariance principle in log density* or a.s.i.p. (log) for short.

In summary, for *most* times, in the sense of log density, Strassen's upper bound can be improved to $o(\sqrt{n})$; there are, however, exceptional times where $o((n \log\log n)^{1/2})$ is the best one can do.

We have addressed the situation when $F$ has finite second or higher moments, so now in the same line of thought, what can be said when the variance is infinite? This splits into two cases: the so-called *non-normal* domain of attraction of the Gaussian law, and the *stable non-Gaussian* case.

Here we recall that a distribution function $F$ is in the *domain of attraction* of $G$ if and only if there exists an i.i.d. sequence $(X_i)$ with common distribution function $F$, a *centering* sequence $(b_n)$ and a *normalizing* sequence $(a_n)$, with $a_n > 0$, such that the following convergence in law holds:

$$\frac{1}{a_n}(S_n - b_n) \xrightarrow{\text{law}} G \quad \text{as } n \to \infty. \tag{1.10}$$

As Lévy showed, the only possible non-trivial attracting laws $G$ are the $\alpha$-stable laws, $0 < \alpha \le 2$.

In this framework, Berkes and Dehling in [3], Theorems 4 and 5, p. 1658, proved the following, extending a result of [11] to laws with infinite variance:

**Theorem A.** *Let $(X_i)_{i\geq 0}$ be an i.i.d. sequence of random variables of distribution function $F$ in the domain of attraction of an $\alpha$-stable law with $0<\alpha\leq 2$. Then after enlarging the probability space there exist an i.i.d. sequence of $\alpha$-stable random variables $(Y_i)$ and a slowly varying sequence $(\lambda_i)$ such that*:

$$\sup_{1\leq k\leq n}\left|S_k-c_k-\sum_{i=0}^{k-1}\lambda_i Y_i\right| = \mathrm{o}(a_n) \quad a.s.\ (log), \tag{1.11}$$

*where $(a_n)$ is the normalizing sequence in* (1.10) *and $(c_k)$ a centering sequence*, *which can be taken equal to zero for $0<\alpha<1$ and to $k\mathbb{E}(X_0)$ for $\alpha>1$.*

*For $\alpha=2$, we replace $\sum_{i=0}^{k-1}\lambda_i Y_i$ $(k\geq 1)$ by $B(a_k^2)$ with $B$ a standard Brownian motion.*

However for our purposes, this statement lacks in several respects. First, the rescaling used for the general regularly varying case acts on the stable increments $Y_i$, and so does not exhibit the connection with dynamics of the scaling transformations. Second, one would like to unify the statements for the Gaussian and the stable non-Gaussian cases, replacing the weighted sum $\sum_{i=0}^{k-1}\lambda_i Y_i$ with $Z(a_k^\alpha)$ where $Z$ is the corresponding $\alpha$-stable process.

In our approach, we construct a joining by sampling via a specially chosen continuously differentiable and increasing time change (see Section 5.1) directly from the continuous-time stable process $Z$ (see (5.5)) rather than beginning with an i.i.d. stable sequence $(Y_i)$ as in [3]; this enables us to resolve both problems simultaneously. We transfer this time change to the random-walk path, and then can use the scaling flow on the stable paths. This gives the dynamical result we are really after.

To carry out this program we need first a dynamical framework: a Polish space, a topology, a flow $\tau_t$ and a flow-invariant probability measure.

As for $\alpha\neq 2$ the paths are highly discontinuous, we can no longer use the space $\mathcal{C}$. Thus we replace $\mathcal{C}$ by $D\equiv D_{\mathbb{R}^+}$, the collection of *càdlàg* (continuous from the right and such that the limits from the left exist) paths defined on $\mathbb{R}^+$.

We describe a topology which makes $D$ a Polish space; this was shown by Billingsley (Theorem 14.2 of [4]) for Skorokhod's space $D_I$, the càdlàg functions on the unit interval $I$, by defining a complete metric $d_1$ for Skorokhod's $J_1$ topology on $D_I$. We first extend $d_1$ to the space $D=D_{\mathbb{R}^+}$; we rescale that to obtain a pseudometric $d_t$ on the interval $[0,t]$ for each $t>0$ (see Lemma 3.4). What we would like to do now is to imitate the topology on the space $\mathcal{C}$ of uniform convergence on compact sets. One can define such a metric directly, as in Whitt [38], though we choose a slightly different definition than that given there. Integrating as done above for $d_\infty^u$, we define:

$$d_\infty(f,g)=\int_0^{+\infty}\mathrm{e}^{-t}\frac{d_t(f,g)}{1+d_t(f,g)}\,\mathrm{d}t. \tag{1.12}$$

By mimicking Whitt's argument, one can verify that this metric is complete; we give an alternative proof in Lemma 7.2.

For any chosen $\alpha\in(0,2]$, we define the *scaling flow* $\tau_t$ of index $1/\alpha$ on $D$ by:

$$(\tau_t f)(x)=\frac{f(\mathrm{e}^t x)}{\mathrm{e}^{t/\alpha}}. \tag{1.13}$$

We write $\nu$ for the stable measure on $D$, the law of $Z$ (the corresponding stable process). For $\alpha\neq 1$ this measure is $\tau_t$-invariant but the Cauchy case $\alpha=1$ requires special attention; for the non-symmetric case ($\xi\neq 0$, see Definition 2.1), $\nu$ is no longer invariant and we replace $Z$ by

$$\widetilde{Z}(t)=Z(t)-\xi t\log t, \tag{1.14}$$

giving a 1-self-similar process (see Lemma 3.2) with independent but not identically distributed increments. Writing $\widetilde{\nu}$ for the corresponding measure on path space, this is $\tau_t$-invariant.

We then show that for all $\alpha$ the flow is $d_\infty$-continuous (Proposition 7.3) and that just as for Brownian motion, the flow $\tau_t$ on the Lebesgue space $(D,\nu)$ (with $\widetilde{\nu}$ for $\alpha=1$) is ergodic (and indeed is a Bernoulli flow of infinite entropy), see Lemma 3.3.

We are now ready to state the main result of this paper:

**Theorem 1.1 (An a.s.i.p. (log) for stable processes).** *Let $(X_i)$ be an i.i.d. sequence of random variables of common distribution function $F$ in the domain of attraction of an $\alpha$-stable law with $\alpha \in (0,2]$. For $\alpha > 1$ assume also for simplicity that the $X_i$ are centered. Then there exists a $C^1$, strictly increasing, regularly varying function $a(\cdot)$ of index $1/\alpha$ with regularly varying derivative, which is explicitly defined from $F$ in Proposition 5.1 and for which $a(n)$ gives a normalizing sequence, such that there exists a joining of the process $\overline{S}$ with an $\alpha$-stable process $Z$ satisfying: for almost every pair $(\overline{S}, Z)$ with respect to this joining, then (for $\alpha \neq 1$),*

$$\lim_{t\to\infty} d_1\bigl(\tau_t\bigl(\overline{S}\circ\bigl(a^\alpha\bigr)^{-1}\bigr), \tau_t Z\bigr) = 0 \quad (\textit{Cesáro}) \tag{1.15}$$

*with $d_1$ Billingsley's complete metric on $D_{[0,1]}$ and $\tau_t$ the scaling flow of index $1/\alpha$. Equivalently, for $d_T$ this metric rescaled to $[0,T]$ (see Section 3.3), we have the a.s.i.p. (log)*

$$d_T\bigl(\overline{S}\circ\bigl(a^\alpha\bigr)^{-1}, Z\bigr) = \mathrm{o}\bigl(T^{1/\alpha}\bigr) \quad (\textit{log}). \tag{1.16}$$

*Statement (1.15) also holds with $d_\infty$, the metric on $D_{\mathbb{R}+}$ defined in (1.12), replacing $d_1$. As a consequence the time-changed path $\overline{S}\circ(a^\alpha)^{-1}$ is in the $\tau_t$-Cesáro stable manifold of the path $Z$:*

$$\overline{S}\circ\bigl(a^\alpha\bigr)^{-1} \in W^{\mathrm{s},d_\infty}_{\mathrm{Ces}}(Z).$$

*All the above stays valid for $\alpha = 1$ upon replacing $Z$ by $\widetilde{Z}$, defined in (1.14), and $S$ by $S - \varrho$ where $\varrho(t) = t\int_{-a(t)}^{a(t)} x \, \mathrm{d}F(x)$.*

*In the case where $\alpha = 2$, $d_1$ is replaced by $d_1^u$ (or $d_\infty^u$), and $B$ replaces $Z$.*

We start by proving statement (1.15) in several steps. First (Lemma 4.1) we find a step path approximation to self-similar processes; this result applies not only to stable processes for $\alpha \neq 1$, but for $\alpha = 1$ to the process $\widetilde{Z}$ which is not Lévy. Next we show, using the ergodicity of the scaling flow, that the step path $\overline{Z}_{a^\alpha(\mathcal{Q})}$ over the partition $a^\alpha(\mathcal{Q})$ with $\mathcal{Q} \equiv ([n; n+1))_{n\geq 0}$ is an element of $W^{\mathrm{s},d_1}_{\mathrm{Ces}}(Z)$, then we prove our extension of Theorem A, see Propositions 4.2 and 5.1. From this we deduce in Proposition 5.2 a step path version of the a.s.i.p. (log). Combining these results proves (1.15); we then derive from that the corresponding statement for $d_\infty$.

In Theorem 1.1 the time change $a(\cdot)$ is constructed from $F$; we are interested in how modifying $a(\cdot)$ for fixed $F$ or modifying the distribution $F$ itself might affect the previous results:

**Proposition 1.2 (Comparison of paths/alternate time changes).** *Under the assumptions and notation of Theorem 1.1, we have:*

(i) *If $\widetilde{a}(\cdot)$ is a $C^1$, strictly increasing function with regularly varying derivative which is asymptotically equivalent to the time change $a(\cdot)$ for $F$, then we can replace $a(\cdot)$ by $\widetilde{a}(\cdot)$ in the statements of Theorem 1.1. In particular this is true for a smoothed polygonal interpolation of a normalizing sequence for $F$, see Lemma 5.3.*

*Moreover, for $\alpha \neq 1$ there exists a joining of two copies $\overline{S}_{(1)}$, $\overline{S}_{(2)}$ of the random walk process $\overline{S}$ for $F$ so that for almost every pair $(\overline{S}_{(1)}, \overline{S}_{(2)})$, the paths $\overline{S}_{(1)}\circ(a^\alpha)^{-1}$ and $\overline{S}_{(2)}\circ(\widetilde{a}^\alpha)^{-1}$ are elements of the same Cesáro stable manifold $W^{\mathrm{s},d_\infty}_{\mathrm{Ces}}(\cdot)$.*

(ii) *Let $F$ and $\widetilde{F}$ be two distribution functions in the domain of attraction of an $\alpha$-stable law for $\alpha \neq 1$ such that they have equivalent truncated variances. Then there exist equivalent smooth time changes $a(\cdot)$ and $\widetilde{a}(\cdot)$ constructed from $F$ and $\widetilde{F}$ as in Proposition 5.1, and a joining of $\overline{S}$ and $\overline{\widetilde{S}}$ for $F$ and $\widetilde{F}$ respectively, so that for almost every $(\overline{S}, \overline{\widetilde{S}})$,*

$$\|\overline{S} - \overline{\widetilde{S}}\|^\infty_{[0,T]} = \mathrm{o}\bigl(a(T)\bigr) \quad \textit{a.s. (log)}, \tag{1.17}$$

*and furthermore there is a joining of the four processes $\overline{S}\circ(a^\alpha)^{-1}$, $\overline{\widetilde{S}}\circ(a^\alpha)^{-1}$, $\overline{S}\circ(\widetilde{a}^\alpha)^{-1}$ and $\overline{\widetilde{S}}\circ(\widetilde{a}^\alpha)^{-1}$ with $Z$ such that a.s. the four paths are all elements of the same Cesáro stable manifold $W^{\mathrm{s},d_\infty}_{\mathrm{Ces}}(Z)$. The above statements hold for $\alpha = 1$ upon centering $S$ and $\widetilde{S}$.*

Next we see how to derive pathwise limit theorems from flow ergodicity together with an a.s.i.p. (log), first in a general context of a self-similar process, then specializing to the case of Theorem 1.1.

**Proposition 1.3.** (i) *Let $\beta > 0$, $Y \in D$ an ergodic $\beta$-self-similar process with law $\rho$ and $U \in D$ another process with law $\widetilde{\rho}$. Assuming that there exists a joining $\widehat{\rho}$ of $Y$ and $U$ such that for $\widehat{\rho}$-a.e. pair $(Y, U)$, we have $U \in W^{\mathrm{s},d}_{\mathrm{Ces}}(Y)$ for $\tau_t$ the scaling flow with $\beta = 1/\alpha$ and $d$ some metric which gives $D$ the Skorokhod topology. Then*:

(a) *$\widetilde{\rho}$-a.e. $U$ is a generic point for $\tau_t$ and $\rho$, i.e. for all $\Phi \in \mathrm{CB}(D, d)$,*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \Phi\big(\tau_t(U)\big)\,\mathrm{d}t = \langle \Phi, \rho\rangle. \tag{1.18}$$

(b) *For $\widetilde{\rho}$-a.e. $U$, writing $\rho_1$ for the distribution of $Y(1)$, we have that for all $\psi \in \mathrm{CB}(\mathbb{R})$,*

$$\lim_{T\to\infty} \frac{1}{\log T} \int_1^T \psi\left(\frac{U(t)}{t^{1/\alpha}}\right)\frac{1}{t}\,\mathrm{d}t = \int_{\mathbb{R}} \psi\,\mathrm{d}\rho_1. \tag{1.19}$$

(ii) *Under the assumptions of Theorem 1.1, then for $\alpha \neq 1$: (1.18) and (1.19) hold for $\mu$-a.e. $\overline{S}$ with $Z$ replacing $Y$ and $\overline{S} \circ (a^\alpha)^{-1}$ replacing $U$, with $d$ any metric giving the Skorokhod topology. Moreover*: *defining $(\tau^a_t f)(x) = f(\mathrm{e}^t x)/a(\mathrm{e}^t)$, then for $\mu$-a.e. $\overline{S}$, we have $\tau^a_t(\overline{S}) \Rightarrow \nu$ (Cesáro); that is, for any $\Phi$ in $\mathrm{CB}(D, d_\infty) = \mathrm{CB}(D, d)$,*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \Phi\big(\tau^a_t(\overline{S})\big)\,\mathrm{d}t = \lim_{T\to\infty} \frac{1}{\log T} \int_1^T \Phi\left(\frac{\overline{S}(t\cdot)}{a(t)}\right)\frac{1}{t}\,\mathrm{d}t = \langle \Phi, \nu\rangle. \tag{1.20}$$

*For $\alpha = 1$, these statements hold true with $Z$ replaced by $\widetilde{Z}$ defined in (1.14) and with $S$ replaced by $S - \varrho$ where $\varrho(\cdot)$ is the centering function defined in Theorem 1.1.*

The proofs we give of the proposition bring up some special points we wish to emphasize. The argument for the proof of (1.19) highlights an important difference between the spaces $D$ and $\mathcal{C} = \mathcal{C}(\mathbb{R}^+)$. For $\mathcal{C}$ the projection to one-dimensional distributions is continuous, so (1.19) would follow automatically from (1.18). However this is no longer true for the space $D$. We circumvent this difficulty by "convolving along the flow $\tau_t$," as seen in Lemma 6.2. We mention that this step was inspired by a key idea in Ambrose and Kakutani's proof that any ergodic measurable flow can be represented by a flow built over a cross-section map [2].

Let us say that a process with paths in $D$ is *asymptotically self-similar* if a.e. path is a generic point for some self-similar process; part (ii) first tells us that this is true for the time-changed random walk process, then converts this into a statement for the random walk path $\overline{S}$ without the time change. However now the transformations $\tau^a_t$ form a non-stationary dynamical system, only giving an actual flow when $a(t)$ is $t^{1/\alpha}$ (in which case $\tau^a_t = \tau_t$).

Deriving (1.20) from (1.18) will involve not just the complete metric $d_\infty$ but also two non-complete metrics denoted by $d^0_\infty$ and $\widetilde{d}^0_\infty$, both of which give the same topology as $d_\infty$. We construct $d^0_\infty$ from Billingsley's non-complete metric $d^0_1$ on $D_{[0,1]}$ by integration as for $d_\infty$. The definition for $\widetilde{d}^0_\infty$ is quite different, and is inspired by Stone's original definition of the $J_1$ topology; see Section 8.

Taking $U = \overline{S} \circ (a^\alpha)^{-1}$, (1.19) (after a change of variables and Karamata's theorem) gives a continuous-time version of the pathwise CLT known for the Gaussian case, see [7,11,12,23,24,32], and for the stable case, Corollary 1 of [3]. We emphasize that a corresponding continuous-time statement does also follow from the discrete-time Corollary 1 of [3].

We note that Berkes and Dehling in Corollary 2 of [3] and Major in Theorem 3 of [28] (with part of the proof in [27]) give discrete-time versions of (1.20) and of the specialization to $\overline{S}$ of (1.18) respectively, in both cases for the metric $d_1$. Corollaries 1, 2 of [3] were proved in that paper not only for their own interest but as steps in the proofs of Theorems 4, 5 there.

We can picture the relationship between Proposition 1.3, Proposition 1.2 and Theorem 1.1 as follows. Write $\mathcal{M}$ for the collection of all probability measures on $D$, with the topology of weak convergence. Given a law $F$ in the domain of attraction of a stable distribution, with measure $\mu$ for its step-path process $\overline{S}$, write (temporarily) $S^a$ for the path $\overline{S} \circ (a^\alpha)^{-1}$ and $\mu^a \in \mathcal{M}$ for the measure on $D$ of these time-changed paths. Then by (1.18), using the same notation as in (1.3) above, we have that for all $\varphi \in \mathrm{CB}(D, d_\infty)$, $\varphi_T \to \langle \varphi, \nu\rangle$ $\mu^a$-a.s., so by the LDCT $\langle \varphi_T, \mu^a\rangle \to \langle \varphi, \nu\rangle$, which in turn says that for the flow $\tau^*_t$ acting on $\mathcal{M}$, then $\mu^a$ is in $W^{\mathrm{s}}_{\mathrm{Ces}}(\nu)$.

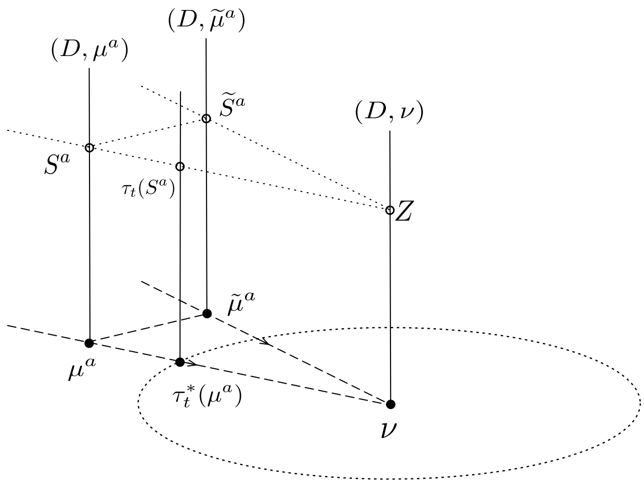


Fig. 1. Action of the scaling flow on the fiber bundle $\mathcal{M}_F \times D$: since $\nu$ is a fixed point, $\tau_t$ moves $Z$ within its fiber $(D, \nu)$, with the dotted lines indicating part of its Cesáro stable manifold.

We consider the product space $\mathcal{M} \times D$, thinking of it as a fiber bundle over $\mathcal{M}$, with the metric $d_\infty$ on the fibers and with each fiber $D$ carrying the corresponding measure; this is acted on by the product flow $(\tau_t^*, \tau_t)$. Now fixing a stable measure $\nu$ on $D$, we restrict attention to the collection $\mathcal{M}_\nu$ of all measures $\mu^a$ on $D$ coming from the domain of attraction of its law, together with their rescalings $\tau_t^*(\mu^a)$ for all $t \in \mathbb{R}$. Then Theorem 1.1 says that the statement just derived from Proposition 1.3 (that $\mu^a \in W^{\mathrm{s}}_{\mathrm{Ces}}(\nu)$) can be lifted to the fibers of $\mathcal{M}_\nu \times D$, via a joining. This is depicted in the first vertical rectangle of Fig. 1.

Now we fit the last statement of Proposition 1.2 into this picture: we partition $\mathcal{M}_\nu$ into equivalence classes such that the laws $F$ have equivalent truncated variances. As a consequence of the proposition, then for two equivalent laws $F, \widetilde{F}$ we have not only that $\mu^a, \widetilde{\mu}^a \in W^{\mathrm{s}}_{\mathrm{Ces}}(\nu)$, but that this statement also lifts to the fibers via a joining, with all three paths $S^a, \widetilde{S}^a, Z$ in the same Cesáro stable manifold, see Fig. 1; here $a(\cdot)$ in fact represents any of the equivalence class of time changes from the first part of the proposition.

The outline of the paper is as follows. In Section 2, we list known results on stable laws, their domains of attraction and log averaging which will be of use throughout the paper. In Section 3, we describe the dynamical setting, define $d_T$ and show how to pass the a.s.i.p. (log) from $d_1$ to $d_\infty$, following which we develop the needed background material on joinings. The main result is Theorem 1.1, proved in Section 5; two key steps in the proof are Proposition 4.2 and Proposition 5.1. At the end of Section 5 we give the proof of Proposition 1.2, and in Section 6 we prove Proposition 1.3. In Section 7 we present proofs of the completeness of $(D, d_\infty)$ and of the continuity of $\tau_t$ on that space. In Section 8 we focus on the non-complete metrics $d^0_\infty$ and $\tilde{d}^0_\infty$.

## 2. Preliminaries

In this section we first recall some properties of stable laws and of regularly varying functions which will be of use throughout the paper, after which we give the characterization of the domain of attraction of a stable law; we refer the reader to [5] and [10]. Then we consider how the log average behaves with respect to regular variation.

### 2.1. *Attraction to stable laws and regular variation*

In defining stable laws, we fix the specific conventions to be used throughout. Several different versions of these formulas appear in the literature, with other choices of signs and constants (and sometimes with errors! [18]).

***Definition 2.1 (See [10], p. 570).*** *A random variable $X$ has a stable law if there are parameters $\alpha \in (0,2]$, $\xi \in [-1,1]$, $b \in \mathbb{R}$, $c > 0$ such that its characteristic function has the following form*:

$$E\big(e^{itX}\big) = \begin{cases} \exp\big(ibt + c \cdot \frac{\Gamma(3-\alpha)}{\alpha(\alpha-1)} |t|^\alpha \big(\cos \frac{\pi\alpha}{2} - \operatorname{sign}(t) i\xi \sin \frac{\pi\alpha}{2}\big)\big) & \textit{for } \alpha \neq 1, \\ \exp\big(ibt - c \cdot |t| \big(\frac{\pi}{2} + \operatorname{sign}(t) i\xi \log |t|\big)\big) & \textit{for } \alpha = 1, \end{cases}$$

*where* $\operatorname{sign}(t) = t/|t|$ *with the convention* $\operatorname{sign}(0) = 0$. *The parameters $\alpha$, $\xi$, $c$ and $b$ are called the exponent or index*, *symmetry* (*or skewness*), *the scaling and the centering parameters respectively. We write $G_{\alpha,\xi,c,b}$ for the distribution function of $X$.*

We write $G_{\alpha,\xi}$ or simply $G_\alpha$ for $G_{\alpha,\xi,1,0}$, the $(\alpha,\xi)$-stable or just $\alpha$-stable when it is clear from the context which $\xi$ is intended.

Two functions $f, g : \mathbb{R}^+ \to \mathbb{R}$ are *asymptotically equivalent* at $+\infty$ (written $f \sim g$) iff they are eventually non-zero and $f(t)/g(t) \to 1$ as $t \to +\infty$. We make the similar definition for sequences.

An eventually positive and measurable function $l$ is *slowly varying* iff $\forall x > 0$, $l(xt) \sim l(t)$. It follows from p. 12 of [5] that $\log l(x) = o(\log x)$. A function $f$ is *regularly varying* with exponent (or index) $\gamma \in \mathbb{R}$ iff $f(x) = x^\gamma l(x)$ for $l$ some slowly varying function.

**Theorem 2.2 (See [5], pp. 12–28).** (i) (*Karamata's Theorem, first part*) *Let $f$ be regularly varying with exponent $\gamma$, with $\gamma > -1$. Then $\int_0^x f$ is regularly varying with exponent $\gamma + 1$*:

$$g(x) \equiv \int_0^x f(t)\,dt \sim \frac{1}{(\gamma+1)} x f(x). \tag{2.1}$$

(ii) *Let $f$ be an invertible and regularly varying function with exponent $\gamma > 0$. Then its inverse $f^{-1}$ is regularly varying with exponent $1/\gamma$.*

Lévy's characterization of the distributions $F$ which are in the domain of attraction of $G_{\alpha,\xi}$ is given in terms of the tail of $F$; see [10], pp. 312–315 (XVII.5, IX.8), and also [5], pp. 346–347:

**Theorem 2.3.** (i) *A distribution function $F$ is attracted to a non-normal stable law $G_{\alpha,\xi}$ with $0 < \alpha < 2$ and $\xi$ uniquely written as $p - q$ with $p, q \in [0,1]$ and $p + q = 1$ iff for a slowly varying function $L$*

$$\frac{1 - F(t)}{1 - F(t) + F(-t)} \to p \quad \textit{and} \quad V(t) \equiv \int_{-t}^{t} x^2\,dF(x) \sim t^{2-\alpha} L(t). \tag{2.2}$$

(ii) *$F$ is attracted to a normal law $G_2$ iff the truncated variance $V(\cdot)$ is slowly varying.*

*In all cases*, *the function $L$ and the normalizing sequence $a_n$ of* (1.10) *with $G = G_{\alpha,\xi}$ are related by*:

$$a_n^\alpha \sim n L(a_n). \tag{2.3}$$

## 2.2. *Log density and regularly varying changes of scale*

The *Cesáro average* of a locally integrable function $f$ is one's usual notion of time average, $\lim_{T\to+\infty} 1/T \int_0^T f(x)\,dx$. The *logarithmic average* of $f$ is

$$\log \operatorname{average}(f) = \lim_{T\to\infty} \frac{1}{\log T} \int_1^T \frac{f(x)}{x}\,dx. \tag{2.4}$$

The *logarithmic density* of a set $A$ in $\mathbb{R}$ is the log average of $\chi_A$, its indicator function.

We mention first a lemma regarding Cesáro averages which will be needed later.

**Lemma 2.4.** *Let $f:\mathbb{R}^+\to\mathbb{R}$ be a locally integrable function. Then these are equivalent*:

(a) $\forall\varepsilon>0$, $\{t\colon |f(t)|>\varepsilon\}$ *has Cesáro density zero*,
(b) *there exists a set* $B\subseteq\mathbb{R}$ *of Cesáro density* 0 *such that* $\lim_{t\to\infty,t\notin B} f(t)=0$.

See Theorem 1.20 of [37]. Next we prove that the log average is preserved by composition with a positive regularly varying parameter change, *if* in addition this has a regularly varying derivative:

**Proposition 2.5.** *Assume $\zeta:\mathbb{R}^+\to\mathbb{R}^+$ is regularly varying with exponent $\gamma>0$, strictly increasing and that it is differentiable with regularly varying derivative. Let $M$ be a subset of $\mathbb{R}^+$. Then $M$ has log density equal to $c$ iff the image $\zeta(M)$ does.*

(This easily follows from parts (i) and (ii) of Theorem 2.2.)

We do need here the strong hypothesis that $\zeta'$ (exists and) is regularly varying: even though that is always the case up to asymptotic equivalence, this will not be enough to prove invariance of log averages. Indeed, for a counterexample, let $M=\bigcup_{k\geq 0}[2k,2k+1]$; this has Cesáro hence log density $1/2$ in $\mathbb{R}^+$. We shall find $\zeta$ satisfying the assumptions of Proposition 2.5 with $\gamma=1$ and such that the log density of $\zeta(M)$ is different from $1/2$.

To this end, let $\varepsilon>0$, $\varepsilon\neq 1/2$, and let $g$ be a 2-periodic function equal to $2-\varepsilon$ on $[2k,2k+1]$ and $\varepsilon$ on $[2k+1, 2k+2]$, for all $k\geq 0$. Now let $f$ be a smoothed version of $g$.

Next, taking $\zeta(t)=\int_0^t f(x)\,\mathrm{d}x$, one can check that $\zeta$ is regularly varying of index 1 (as $\zeta(t)\sim t$), that its derivative is 2-periodic (and non-constant), so it cannot be slowly varying and that the Cesáro (hence the log) average of $\zeta(M)$ is $1-\varepsilon\neq 1/2$, as claimed.

## 3. Flows on Skorokhod space and the a.s.i.p. (log)

Our point of view will borrow both from ergodic theory and probability theory. For this purpose it is most convenient to use what we call the *path space model* for a stochastic process. To speak of a stochastic process $X$ with paths in $D$ means that we are given an underlying probability space $(\Omega,\mathbb{P})$ and a measurable function $X:\Omega\to D$. Choosing some $\omega\in\Omega$, then $X_\omega=X_\omega(\cdot)=X(\omega,\cdot)$ is a path of $X$. Let $\nu$ denote the measure on $D$ which is the push-forward of $\mathbb{P}$ via the measurable function $X$. For the path space model, we take for the underlying space $(D,\nu)$ itself, with the identity map $\mathcal{I}$; then a path is $\mathcal{I}(X)=\mathcal{I}_X=\mathcal{I}_X(\cdot)$ which we write simply as $X(\cdot)$. So now we can think of an element $X$ of $D$ interchangeably as a path $X(\cdot)$, as a *point* in a dynamical system ($D$ acted on by the scaling flow, for instance) or as the entire stochastic process.

### 3.1. $J_1$-*topology on path space*

The relevant choice for the present paper is Skorokhod's $J_1$-topology for $D_{[0,1]}$, and its extension to the domain $\mathbb{R}^+$ introduced by Stone [34] (which we shall also call the $J_1$-topology). We begin with the unit interval $I$, where we follow [4], pp. 112–116.

Let $\Lambda=\Lambda_1$ be the collection of strictly increasing continuous maps of $I$ onto itself (so in particular, $\lambda(0)=0$ and $\lambda(1)=1$). Billingsley in fact defines two equivalent metrics on $D_I$; we start with the simplest which however fails to be complete. For $f,g\in D_I$, we set:

$$d_1^0(f,g)=\inf\bigl\{\varepsilon\colon \text{ there exists } \lambda\in\Lambda \text{ with } \|\lambda-\mathcal{I}\|_{[0,1]}\leq\varepsilon \text{ and with } \|f-g\circ\lambda\|_{[0,1]}^\infty\leq\varepsilon\bigr\}. \tag{3.1}$$

Billingsley's complete metric makes use of elements of $\Lambda$ which are bounded with respect to the following measurement. For a function $\lambda\in\Lambda$, write:

$$|||\lambda|||_1=\sup_{0\leq s\neq t\leq 1}\left|\log\frac{\lambda(t)-\lambda(s)}{t-s}\right|. \tag{3.2}$$

Similarly now for $f,g\in D_I$, we set:

$$d_1(f,g)=\inf\bigl\{\varepsilon\colon \text{ there exists } \lambda\in\Lambda \text{ with } |||\lambda|||_1\leq\varepsilon \text{ and with } \|f-g\circ\lambda\|_{[0,1]}^\infty\leq\varepsilon\bigr\}. \tag{3.3}$$

Either metric can be extended to $\mathbb{R}^+$ as folllows. First we define the corresponding metrics on $D_{[0,A]}$; to define $d_A^0$ we simply replace [0, 1] by [0, $A$] in (3.1), while $d_A$ is defined by rescaling $d_1$ as we now explain. In both cases $d_\infty$ (respectively $d_\infty^0$) are then defined by integration as in (1.12).

Let $\Lambda_A$ be the collection of strictly increasing continuous maps of $[0, A]$ onto itself, with the notation $\Lambda \equiv \Lambda_1$ for $A = 1$. For a function $\lambda \in \Lambda_A$,

$$|||\lambda|||_A = \sup_{0 \le s \ne t \le A} \left| \log \frac{\lambda(t) - \lambda(s)}{t - s} \right|.$$

Fix $\beta > 0$; for a self-similar process $\beta$ will be the scaling exponent. For $f, g \in D$ we define

$$d_A^\beta(f, g) = \inf\{\varepsilon \colon \exists \lambda_A \in \Lambda_A \text{ with } |||\lambda_A|||_A \le \varepsilon \cdot A^{-\beta} \text{ and } \|f - g \circ \lambda\|_{[0,A]}^\infty \le \varepsilon\}. \tag{3.4}$$

3.2. *Stable flows*

Recalling from (1.13) the definition of the *scaling flow* $\tau_t$, of index $\beta$, and defining for each $t \ge 0$ the *increment (semi-)flow* $\theta$ on $D = D_{\mathbb{R}^+}$ by:

$$(\theta_t f)(x) = f(x + t) - f(t),$$

we have, expressing a basic fact about stable processes, in more dynamical terms:

**Proposition 3.1.** *For each choice of* $\alpha \in (0, 1) \cup (1, 2]$ *and* $\xi \in [-1, 1]$, *there is a unique Borel probability measure* $\nu$ *on* $D$ *satisfying*:

(i) $\nu$ *is invariant for the scaling flow* $\tau_t$ *of index* $\beta = 1/\alpha$, *and for the increment semiflow* $\theta_t$;
(ii) *the process* $Z$ *has independent increments*; *and for* $\nu$*-a.e. path* $Z$, $Z(0) = 0$ *and* $Z(1)$ *has distribution* $G_{\alpha,\xi}$.

The non-symmetric Cauchy process $Z$, i.e. the $(\alpha, \xi)$ stable process with $\alpha = 1$ and $\xi \ne 0$, needs to be treated as a special case as it is not self-similar but rather *self-affine* in a sense we now explain. We define the *affine scaling flow* with parameter $\xi$ on $D$ by:

$$\tau_t^\xi \colon f(\cdot) \mapsto f(e^t \cdot)/e^t - \xi t \cdot$$

One checks that this is a flow. The reason for the name is that, in its action on the space of functions from $\mathbb{R}$ to $\mathbb{R}$, the maps $\tau_t^\xi$ are indeed affine, and only linear for the symmetric case ($\xi = 0$).

Recalling from (1.14) the definition of $\widetilde{Z}$, we have:

**Lemma 3.2.** *The affine flow of index one* $\tau_t^\xi$ *on* $D$ *preserves the Cauchy stable measure* $\nu$. *Equivalently*, *the index-one scaling flow preserves the corresponding measure* $\widetilde{\nu}$ *for* $\widetilde{Z}$, *that is* $\widetilde{Z}$ *is* 1*-self-similar. The correspondence* $Z \mapsto \widetilde{Z}$ *gives a flow isomorphism. The process* $\widetilde{Z}$ *vanishes at* 0 *and has independent non-stationary increments. The measure* $\nu$ *is the unique* $\tau_t^\xi$*-invariant measure with i.i.d. increments*, *such that a.s.* $Z(0) = 0$, *and with distribution* $G_{1,\xi}$ *for* $Z(1)$.

In our proofs below, rather than use the affine flow on $Z$, we use $\widetilde{Z}$ with the scaling flow of index one, as this allows us to give a unified treatment for all $\alpha$.

Next we prove that the scaling flow $(D, \nu, \tau_t)$ (or $(D, \widetilde{\nu}, \tau_t)$ for $\alpha = 1$) is *ergodic*, and indeed *Bernoulli*. The ergodicity (i.e. that all invariant sets have either zero or full measure) is all we actually use in this paper; it provides a key ingredient for our proof of the a.s.i.p. (log).

A Bernoulli flow is, by definition, a measure-preserving flow of a Lebesgue space whose time-one map is measure-theoretically isomorphic to a Bernoulli shift. As Ornstein showed, two Bernoulli flows are isomorphic if and only if they have the same entropy; this can be a strictly positive number or $+\infty$. Ornstein then came up with a sufficient condition for the Bernoullicity of transformations or flows, *very weak Bernoulli*, which is easily verified in many examples. See [29,33].

**Lemma 3.3.** *For every $\alpha \in (0,2]$, $\alpha \neq 1$ and $\xi \in [-1,1]$, the scaling flow $\tau$ of the $(\alpha,\xi)$-stable process $Z$ is ergodic, and indeed is Bernoulli of infinite entropy. For $\alpha = 1$ this holds (equivalently) for the flows $\tau^\xi$ on $Z$ and $\tau$ on $\widetilde{Z}$.*

**Proof.** We follow the proof for Brownian motion, the case $\alpha = 2$, given in [12]. We claim first that for $\alpha \neq 1$:

$$\lim_{t\to\infty} \nu\big(\big\{Z\colon a < Z(1) < b; c < (\tau_t Z)(1) < d\big\}\big) = \nu\big(\big\{Z\colon a < Z(1) < b\big\}\big)\nu\big(\big\{Z\colon c < Z(1) < d\big\}\big)$$

for all $a < b$ and $c < d$ and that the same holds with $\widetilde{\nu}$, and $\widetilde{Z}$ replacing $\nu$ and $Z$.

This follows from $Z$ having independent increments together with the fact that $\tau$ preserves $\nu$ for $\alpha \neq 1$ and that it preserves $\widetilde{\nu}$ for $\alpha = 1$.

The above claim shows mixing of this process for one-cylinders; that extends by the same reasoning to finite-dimensional cylinder sets and thus proves mixing, from which ergodicity follows.

Ornstein's property of very weak Bernoulli follows from the same observation: since $D$ is Polish, $(D,\nu)$ is a Lebesgue space, and hence one has Bernoullicity of the flow by Ornstein's theorem. □

### 3.3. *Flow approximation in the $d_\infty$-metric*

We now show how to pass from an a.s.i.p. for $d_1$ on $D$ to an a.s.i.p. for the complete metric $d_\infty$, as needed for the proof of the main theorem.

For each $r > 0$, define $\Delta_r$ the *scaling transformation of order $\beta$*, on path space by

$$(\Delta_r f)(x) \equiv \frac{f(rx)}{r^\beta}. \tag{3.5}$$

Since $\tau_t = \Delta_{e^t}$, the maps $(\Delta_r)_{r>0}$ give an action of the multiplicative group of positive real numbers, and so a multiplicative version of the scaling flow of index $\beta$.

We have defined $d_A$ (see (3.4)) so as to have the following scaling property:

**Lemma 3.4.** *For $f, g \in D$, we have for all $A > 0$,*

$$d_A(f,g) = A^\beta d_1(\Delta_A f, \Delta_A g). \tag{3.6}$$

**Lemma 3.5.** *Given $f$ and $g$ in $D$, the following are equivalent*:

$$\begin{aligned} &\text{(i)} \quad d_1(\tau_t f, \tau_t g) \to 0 \quad (\textit{Cesáro}),\\ &\text{(ii)} \quad d_1(\Delta_T f, \Delta_T g) \to 0 \quad (\textit{log}),\\ &\text{(iii)} \quad d_T(f,g) = o\big(T^\beta\big) \quad (\textit{log}). \end{aligned} \tag{3.7}$$

Next we consider a consequence of the convergence described in Lemma 3.5.

**Lemma 3.6.** *The equivalent statements in Lemma 3.5 imply*:

$$d_\infty(\tau_t f, \tau_t g) \to 0 \quad (\textit{Cesáro}). \tag{3.8}$$

We remark that were the convergence in (i) (and hence in (3.9) below) true for all $t$ then (3.8) (without the "Cesáro") would be a direct consequence via the LDCT. But since it only holds off a set of $t$ of density zero, we need the more careful argument which follows.

**Proof of Lemma 3.6.** We begin with (i), so there exists $\mathcal{B}_1 \subset \mathbb{R}^+$ of Cesáro density zero such that $d_1(\tau_t f, \tau_t g)$ goes to zero as $t \to \infty$, for $t \notin \mathcal{B}_1$. We claim that then in fact:

$$\forall A > 0, \quad d_A(\tau_t f, \tau_t g) \to 0 \quad (\text{Cesáro}). \tag{3.9}$$

Indeed, from Lemma 3.4 we have $d_A(\tau_t f, \tau_t g) = A^{1/\alpha} d_1(\tau_{t+\log A}(f), \tau_{t+\log A}(g))$, for all $A > 0$. Taking $t \in \mathcal{B}_A \equiv \mathcal{B}_1 - \log A$ implies that $d_A(\tau_t f, \tau_t g) \to 0$ off $\mathcal{B}_A$. But since the Cesáro density of a set is invariant under the action of a translation, $\mathcal{B}_A$ is of Cesáro density zero, proving (3.9).

Next we reword (3.9) with the help of Lemma 2.4 as follows: for all $A > 0$ and for all $\varepsilon > 0$,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \chi_{d_A(\tau_t f, \tau_t g) > \varepsilon} \, \mathrm{d}t = 0. \tag{3.10}$$

Recalling from (1.12) the definition of $d_\infty$, then again by Lemma 2.4, (3.8) is equivalent to showing that for all $\varepsilon > 0$, the Cesáro average of $\chi_{d_\infty(\tau_t f, \tau_t g) > \varepsilon}$ a.s. goes to zero. Since $d_\infty$ is a metric bounded by 1, $\varepsilon \cdot \chi_{d_\infty > \varepsilon} \le d_\infty \le \varepsilon + \chi_{d_\infty > \varepsilon}$; thus it is equivalent to prove that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T d_\infty(\tau_t f, \tau_t g) \, \mathrm{d}t = 0.$$

On the other hand, again from (1.12) it is immediate that for all $\varepsilon > 0$,

$$d_\infty(\tau_t f, \tau_t g) \le \int_0^\infty \mathrm{e}^{-A} \chi_{d_A(\tau_t f, \tau_t g) > \varepsilon} \, \mathrm{d}A + \varepsilon. \tag{3.11}$$

So we are done so long as we prove that

$$\frac{1}{T} \int_0^T \left( \int_0^\infty \mathrm{e}^{-A} \chi_{d_A(\tau_t f, \tau_t g) > \varepsilon} \, \mathrm{d}A \right) \mathrm{d}t = \int_0^\infty \mathrm{e}^{-A} \left( \frac{1}{T} \int_0^T \chi_{d_A(\tau_t f, \tau_t g) > \varepsilon} \, \mathrm{d}t \right) \mathrm{d}A$$

approaches zero, as $T \to \infty$. From (3.10) and then LDCT, it does, finishing the proof of (3.8). □

Later in the paper we shall encounter measure-theoretic versions of the equations in Lemmas 3.5 and 3.6, for two stochastic processes $f$ and $g$ which are paired together by their joint distributions having been specified in some consistent way. We discuss these pairings in the next section, first in the general setting. See e.g. [17] for further background and references.

### 3.4. *Underlying probability spaces and the composition of joinings*

***Definition 3.7.*** *Given two measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$, then a joining (or coupling) of the two spaces is a measure $\hat{\nu}$ on $X \times Y$ which has marginals $\mu, \nu$, i.e. which projects to those measures.*

We recall that a measure space $(Y, \mathcal{B}, \nu)$ is a *factor* of $(X, \mathcal{A}, \mu)$ when there is a measure-preserving map $f$ from $X$ onto $Y$; in this case one also says that $(X, \mathcal{A}, \mu)$ is an *extension* of $(Y, \mathcal{B}, \nu)$.

Thus, a joining gives a common extension of the two spaces. There is a converse; the proof follows directly from the definitions:

**Lemma 3.8.** *Suppose $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ have $(Z, \mathcal{D}, \rho)$ as common extension via maps $\alpha : Z \to X$, $\beta : Z \to Y$; defining $\varphi : Z \to X \times Y$ by $\varphi(z) = (\alpha(z), \beta(z))$ and $\hat{\nu}$ on $(X \times Y, \mathcal{A} \times \mathcal{B})$ to be the pushed-forward measure $\hat{\nu} = \rho \circ \varphi^{-1}$, then $\hat{\nu}$ is a joining measure.*

For an example, the probability idea of "redefining two processes so as to live on a common probability space" (this just means the two path space models are given a common extension) is equivalent to defining a joining of these two measure spaces.

Now suppose that rather than having a common extension, our spaces have a common factor. In this case, there is a unique joining, the *relatively independent joining*, which exhibits the maximum possible independence while respecting the common factor. Thus, given two probability Polish spaces $(X, \mathcal{A}, \mu)$, $(Z, \mathcal{D}, \rho)$ and measure preserving maps $\alpha : X \to Y$, $\beta : Z \to Y$ to the factor space $(Y, \mathcal{B}, \nu)$, we define the relatively independent joining $\hat{\nu} = \mu \times_Y \nu$ by the formula

$$\hat{\nu}(A \times B) = \int_Y \mu_y(A)\rho_y(B)\,\mathrm{d}\nu(y), \quad A \in \mathcal{A}, B \in \mathcal{B}, \tag{3.12}$$

where $\mu_y$, $\rho_y$ are the *disintegrations* of the measures with respect to the factor map, the existence of which is guaranteed by the Disintegration Theorem, see Section 1 of [1]. This respects the common factor in that $\alpha \circ \pi_X = \beta \circ \pi_Z$ almost surely, where $\pi_X, \pi_Z$ are the coordinate projections, see Proposition 5.11 of [16].

We now arrive at the composition of two joinings; see Definition 6.9 of [17]. We shall need the next result in the proof of Propositions 1.2 and 5.1.

**Proposition 3.9.** *Suppose that we are given a measurable equivalence relation $R$ on a Polish space $X$ with Borel $\sigma$-algebra $\mathcal{A}$. We define a relation $\widehat{R}$ on $\mathcal{M}_1(X)$, the collection of probability Borel measures on $X$, by $\mu$ is related to $\nu$ (written $\mu\widehat{R}\nu$) iff there exists a joining $\hat{\nu}$ of $\mu$ with $\nu$ such that $\hat{\nu}(R) = 1$. Then $\widehat{R}$ is an equivalence relation on $\mathcal{M}_1(X)$.*

This is an immediate consequence of the following lemma. We leave out the $\sigma$-algebras for simplicity of notation:

**Lemma 3.10.** *Suppose we have three probability measures $\mu, \nu, \rho$ on the Polish measure space $(X, \mathcal{A})$, and are given joinings $\hat{\nu}_1$ of $(X, \mu)$ with $(X, \nu)$ and $\hat{\nu}_2$ of $(X, \nu)$ with $(X, \rho)$. Then there exists a joining $\hat{\nu}_3$ of $(X, \mu)$ with $(X, \rho)$ called the composition of joinings $\hat{\nu}_3 = \hat{\nu}_2 \circ \hat{\nu}_1$ and defined below, such that, assuming that $\hat{\nu}_1(R) = \hat{\nu}_2(R) = 1$ then $\hat{\nu}_3(R) = 1$.*

**Proof.** Noting that $(X \times X, \hat{\nu}_1)$ and $(X \times X, \hat{\nu}_2)$ have as a common factor $(X, \nu)$, we denote by $(X \times X \times X \times X, \hat{\nu})$ their relatively independent joining.

This is a common extension of $(X, \mu)$ and $(X, \rho)$, and so by Lemma 3.8 it determines a joining $\hat{\nu}_3$ of $(X, \mu)$ and $(X, \rho)$. This by definition is the composition $\hat{\nu}_3 = \hat{\nu}_2 \circ \hat{\nu}_1$ of the joinings. We observe that the relatively independent joining provides a common probability space for all three joinings, $\hat{\nu}_1$, $\hat{\nu}_2$ and $\hat{\nu}_3$.

Now, denote by $\pi_i$ the projection to the $i$th coordinate of $X \times X$ and $\pi_{i,j}$ the projection to the product of the $i$th and $j$th coordinates of $X \times X \times X \times X$. By assumption, $\hat{\nu}_1(R) = \hat{\nu}_2(R) = 1$. So defining $\mathcal{G}_1 = \pi_{1,2}^{-1}(R)$ and $\mathcal{G}_2 = \pi_{3,4}^{-1}(R)$ we have $\hat{\nu}(\mathcal{G}_i) = \hat{\nu}_i(R) = 1$ for $i = 1, 2$. On the other hand, recall that the common factor is respected by $\hat{\nu}$; indeed the two factor maps $\alpha = \pi_2$ and $\beta = \pi_1$ satisfy $\alpha \circ \pi_{1,2} = \beta \circ \pi_{3,4}$. Accordingly, since the equivalence relation $R$ is transitive, $\mathcal{G}_1 \cap \mathcal{G}_2 \subseteq \pi_{1,4}^{-1}(R)$. Therefore $\hat{\nu}(\pi_{1,4}^{-1}(R)) = \hat{\nu}_3(R) = 1$ as well, finishing the proof. □

***Remark 3.1.*** *We note that the notion of relatively independent joinings extends naturally to a finite sequence of $n$ spaces which are joined two-by-two; we proceed inductively to adjoin the next space. The result is a common extension of all $n$ spaces, which projects to a joining of the spaces, i.e. a measure on their product which has the correct marginals.*

For a concrete example of the proposition, and of the remark just made, see the proof of Proposition 1.2; there we take the equivalence relations on the space $D$ defined in Lemmas 3.5 or 3.6, and used in stating an a.s.i.p. (log) for $d_1$ or for $d_\infty$.

## 4. A step path approximation in the space $D$

Here we develop a key tool needed for the proof of our main theorem; this is valid in a general context of self-similar processes with paths in $D$, and shows that the paths can be $J_1$-approximated by step paths. The statement is in Proposition 4.2; first we need the following lemma.

Let $Z \in D$ and let $\mathcal{P}$ be a locally finite partition of $\mathbb{R}^+$ (i.e. it is finite on any bounded interval) with endpoints $0 = x_0 < x_1 < \cdots$. We write $|\mathcal{P}| \equiv \sup_{i \geq 0}(x_{i+1} - x_i)$ for the mesh of the partition and we define $\overline{Z}_{\mathcal{P}}$ to be the step function over that partition, so

$$\overline{Z}_{\mathcal{P}}(t) \equiv Z(x_i) \quad \text{for } x_i \leq t < x_{i+1}. \tag{4.1}$$

In the following lemma $d_1$ denotes the pseudometric on $D = D_{\mathbb{R}^+}$ which comes from the complete metric $d_1$ on $D_I$. This lemma will be applied to $\alpha$-stable processes for $\alpha \neq 1$ and to the 1-self-similar modified Cauchy process $\widetilde{Z}$.

**Lemma 4.1.** *Let $\nu$ be a probability measure on $D$ which has zero mass on the set $D_\star \stackrel{\text{def}}{=} \{Z \in D\colon Z \text{ has a jump at } 1\}$. Then for all $\varepsilon > 0$, we have*:

$$\nu\Big\{Z \in D\colon \forall \mathcal{P} \text{ with } |\mathcal{P}| < \delta \text{ then } d_1(\overline{Z}_{\mathcal{P}}, Z) < \varepsilon\Big\} \to 1 \quad \text{as } \delta \to 0.$$

**Proof.** The first task is to prove the pointwise statement:

$$\forall Z \in D \setminus D_\star, \quad d_1(\overline{Z}_{\mathcal{P}}, Z) \to 0 \quad \text{as } |\mathcal{P}| \to 0 \tag{4.2}$$

uniformly over all partitions of $\mathbb{R}^+$. The measure statement then will follow from the monotonicity property of $\nu$ (i.e. that $\nu(\bigcup A_i) = \lim \nu(A_i)$ for nested increasing sets).

Note that for the case when the measure $\nu$ happens to be supported on the continuous paths then (4.2) stated for the sup norm instead of for $d_1$, follows immediately from uniform continuity of continuous paths on compact intervals.

So we need only prove (4.2) for the case where $Z$ has jumps. The proof will be carried out in several steps. We begin by showing that a type of uniformity still holds away from a finite set where $Z$ has "big" jumps.

Choose $Z \in D \setminus D_\star$ and fix $1/2 > \varepsilon > 0$. Since there can be at most finitely many points $t$ in any compact interval at which the jump $|Z(t) - Z(t^-)|$ exceeds $\varepsilon$, it follows that

$$F_\varepsilon \equiv \Big\{t \in \mathbb{R}^+\colon \big|Z(t) - Z\big(t^-\big)\big| \geq \varepsilon\Big\}$$

can be written as $(b_i)$ with $0 = b_0 < b_1 < b_2 < \cdots$, such that $b_q < 1 < b_{q+1}$ for some non-negative finite $q = q(\varepsilon)$. Let $J_x^-$ (resp. $J_x^+$) denote the largest open interval immediately to the left (resp. right) of $x$ which does not intersect $F_\varepsilon$.

On our road to (4.2), we begin by showing there exists $0 < \delta_0 = \delta_0(\varepsilon)$ such that $\forall x \in I = [0, 1]$:

$$\big|Z(t) - Z\big(t'\big)\big| < \varepsilon \quad \text{for all } t, t' \in B(x, \delta_0) \cap J_x^\pm \tag{4.3}$$

with $B(x, \delta)$ denoting the open $\delta$-ball centered at $x$ and $J_x^\pm$ meaning that the previous statement holds for both $J_x^+$ and $J_x^-$. Since $Z \in D$ then for all $x \in I$, (4.3) holds for $\delta(x) = \delta(x, \varepsilon)$ replacing $\delta_0$ and (4.3) follows from a compactness argument.

Now suppose $\mathcal{P}$ is a locally finite partition of $\mathbb{R}^+$ with mesh $|\mathcal{P}| < \delta_{00} < \delta_0$, with $\delta_{00}$ to be further specified in what follows; we write $0 = x_0 < x_1 < \cdots < x_p < 1$ for the endpoints of $\mathcal{P}$ in $[0, 1)$.

Recalling from (3.3) and (3.2) the definitions of $|||\cdot|||_1$ and of Billingsley's metric, proving (4.2) reduces to defining $\lambda$, a continuous, strictly increasing parameter change onto $[0, 1]$ satisfying both

$$|||\lambda|||_1 < \varepsilon \quad \text{and} \quad \|\overline{Z}_{\mathcal{P}} \circ \lambda - Z\|_{[0,1]}^\infty < \varepsilon.$$

Now as long as $[x_j, x_{j+1}]$ contains no element of $F_\varepsilon$, so $[x_j, x_{j+1}] \subseteq (b_k, b_{k+1})$ for some $k$, then $[x_j, x_{j+1}] \subseteq B(x_j, \delta_0) \cap J_{x_j}^+$ and so by (4.3) this gives

$$\big|\overline{Z}_{\mathcal{P}}(t) - Z(t)\big| < \varepsilon \quad \forall t \in [x_j, x_{j+1}]. \tag{4.4}$$

Hence on these intervals we can simply take $\lambda(t) = t$. The idea therefore is to begin with $\lambda(t) = t$ on $[0, 1]$ and then modify it near each point in $F_\varepsilon$, in the following fashion.

First, let $r$ be such that

$$0 < r < \inf\left(\frac{\delta_0}{2}, \frac{\min_{0 \leq i \leq q}(b_{i+1} - b_i) - \delta_0}{2}, 1 - b_q\right)$$

with $\delta_0$ given in (4.3); we note that by construction, $\delta_0 < \min_{0 \leq i \leq q}(b_{i+1} - b_i)$. And for that choice of $r$ pick $\delta_{00}$ such that

$$\delta_{00} < \frac{r}{3}\big(1 - \mathrm{e}^{-\varepsilon}\big).$$

For a chosen point $a = b_i \in F_\varepsilon$, $a$ is in an interval $[x_m, x_{m+1})$ for some $m \le p$. Let $l$ and $n$ be such that $x_m - r < x_l < x_m - r + \delta_{00}$ and $x_m + r - \delta_{00} \le x_n < x_m + r$. We define $\lambda$ to be linear on the intervals $[x_l, a]$ and $[a, x_n]$ connecting the points $(x_l, x_l)$, $(a, x_{m+1})$ and $(x_n, x_n)$. We define $\lambda$ in this way at the vicinity of each $a \in F_\varepsilon$, and on the remaining intervals keep the definition $\lambda(t) = t$. Thus $\lambda$ is continuous and strictly increasing and, from the way we have chosen $\varepsilon$, $\delta_0$, $r$ and $\delta_{00}$, it is easily checked that

$$|||\lambda|||_1 \le \max\left(\log\frac{r - 3\delta_{00}}{r - 2\delta_{00}}, \log\frac{r}{r - \delta_{00}}\right) < \varepsilon.$$

The effect of $\lambda$ is to move the jumps in $\overline{Z}_{\mathcal{P}}$ so that they line up exactly with the big jumps in $Z$.

We check the resulting spatial error: we have $\overline{Z}_{\mathcal{P}} \circ \lambda(t) = Z(x_i)$ for all $t \in [x_l, a)$ and some $i \in [l, m]$, and for all $t \in [a, x_n)$, and some $i \in [m+1, n)$. Next, one can check that for all $t \in [x_l, x_n)$, by (4.3), $|\overline{Z}_{\mathcal{P}} \circ \lambda(t) - Z(t)| = |Z(t') - Z(t)| < \varepsilon$, where $t' = x_i$ is defined as above for the two cases.

We claim that in fact

$$\|\overline{Z}_{\mathcal{P}} \circ \lambda - Z\|_{[0,1]}^{\infty} < \varepsilon. \tag{4.5}$$

First we note that for $x_l, x_n, x_m$ assigned to $b_i$ and $x_{\overline{l}}, x_{\overline{n}}, x_{\overline{m}}$ assigned to $b_{i+1}$, then $x_m + r < x_{\overline{m}} - r$ and so $x_n < x_{\overline{l}}$. Thus the modifications made to $\lambda$ near $b_i$ and near $b_{i+1}$ do not interfere with each other, as $\lambda(t) = t$ on $[x_n, x_{\overline{l}}]$. Also, $r < 1 - b_q$ so for the point $x_{\hat{n}}$ assigned to $b_q$, $x_{\hat{n}} < b_q + r < 1$; thus the rightmost modification does not interfere with the definition of $\lambda(t) = t$ near 1.

Next by (4.4), (4.5) holds on all the intervals $[x_j, x_{j+1}]$ where we still have $\lambda(t) = t$.

Furthermore, for all $t \in [x_l, a)$, $\overline{Z}_{\mathcal{P}} \circ \lambda(t) = Z(x_i)$ for some $l \le i \le m$, and $|t - x_i| < r < \delta_0$ by the previous construction and the choice of $r$. Hence $|\overline{Z}_{\mathcal{P}} \circ \lambda(t) - Z(t)| = |Z(t') - Z(t)| < \varepsilon$, by the estimate (4.3). For $t \in [a, x_n)$, the same reasoning holds. This completes the proof of (4.2).

Lastly we show how to get the measure statement from this. By (4.2) for fixed $\varepsilon > 0$,

$$G_\delta^\varepsilon = \{Z \in D \setminus D_\star\colon \text{ for each } \mathcal{P} \text{ with } |\mathcal{P}| < \delta, d_1(\overline{Z}_{\mathcal{P}}, Z) < \varepsilon\},$$

increases to $D \setminus D_\star$ as $\delta \downarrow 0$. Hence, due to the monotonicity of measure, as $\delta \downarrow 0$, $\nu(G_\delta^\varepsilon) \to \nu(D \setminus D_\star) = \nu(D) = 1$, since $\nu$ gives no mass to $D_\star$. This finishes the proof of Lemma 4.1. □

The argument in the next proposition is where the dynamics first comes in, and it is also here that one clearly sees the interplay between density and measure.

**Proposition 4.2.** *Let $Z$ be an ergodic self-similar process of index $\beta > 0$ with paths in $D$; equivalently, assume we are given an invariant ergodic probability measure $\nu$ on $D$ for the scaling flow $\tau_t$ of index $\beta$. Let $\eta$ be a positive strictly increasing regularly varying function of index* 1 *and define $\mathcal{Q}$ to be the partition of $\mathbb{R}^+$ with integer endpoints. Then for $\nu$-a.e. $Z \in D$,*

(i) $d_1(\tau_t \overline{Z}_{\eta(\mathcal{Q})}, \tau_t Z) \to 0$ *(Cesáro)*

*and equivalently,*

(ii) $d_T(\overline{Z}_{\eta(\mathcal{Q})}, Z) = \mathrm{o}(T^\beta)$ *(log).*

**Proof.** By Lemma 3.5, (i) and (ii) are equivalent; we prove (i). Since $\eta$ is a positive and strictly increasing function, we have

$$\overline{Z}_{\eta(\mathcal{Q})} = (\overline{Z \circ \eta})_{\mathcal{Q}} \circ \eta^{-1}; \tag{4.6}$$

in fact, this is true for $\mathcal{Q}$ *any* partition of $\mathbb{R}^+$. Now as one checks from the definitions,

$$\tau_t(\overline{Z}_{\eta(\mathcal{Q})}) = \overline{(\tau_t Z)}_{\mathrm{e}^{-t}(\eta(\mathcal{Q}))}. \tag{4.7}$$

Accordingly, taking $f(t) \equiv d_1(\tau_t \overline{Z}_{\eta(\mathcal{Q})}, \tau_t Z) = d_1(\overline{(\tau_t Z)}_{e^{-t}(\eta(\mathcal{Q}))}, \tau_t Z)$ in Lemma 2.4 tells us that demonstrating (i) will reduce to showing that for all $\varepsilon > 0$, the Cesáro average

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \chi_{\{t: f(t)>\varepsilon\}} \,\mathrm{d}t = \lim_{T\to\infty} \frac{1}{T} \int_0^T \chi_{\{U: d_1(\overline{U}_{\mathcal{Q}_t}, U)>\varepsilon\}}(\tau_t Z)\,\mathrm{d}t = 0$$

with $\mathcal{Q}_t = e^{-t}(\eta(\mathcal{Q}))$ and where $\chi_A$ denotes the indicator function of a set $A$.

Now, since $\eta$ is regularly varying of index 1, it follows that $\eta(n) \sim \eta(n-1)$; therefore the diameter of the rescaled partition $\mathcal{Q}_t \equiv e^{-t}(\eta(\mathcal{Q}))$ on [0, 1], written $|\mathcal{Q}_t|_{[0,1]}$, vanishes as $t \to \infty$.

Hence, for any $\delta > 0$, there exists some large $T_0$ such that for all $t > T_0$, we have $|\mathcal{Q}_t|_{[0,1]} < \delta$. This implies that the set of paths

$$\{U: d_1(\overline{U}_{\mathcal{Q}_t}, U) > \varepsilon\} \subset \{U: \exists \mathcal{P} \text{ such that } |\mathcal{P}| < \delta \text{ and } d_1(\overline{U}_{\mathcal{P}}, U) > \varepsilon\} \equiv A_{\varepsilon,\delta}.$$

Thus, for $Z$ in a set of full $\nu$-measure $\mathcal{G}_{\varepsilon,\delta}$, we have for $T > T_0$

$$0 \le \frac{1}{T} \int_{T_0}^T \chi_{\{t: f(t)>\varepsilon\}} \,\mathrm{d}t \le \frac{1}{T} \int_{T_0}^T \chi_{A_{\varepsilon,\delta}}(\tau_t Z)\,\mathrm{d}t \to \nu(A_{\varepsilon,\delta}) \quad \text{as } T \to \infty$$

with convergence given by the Birkhoff ergodic theorem. From Lemma 4.1 we know that for all $\varepsilon > 0$, $\nu(A_{\varepsilon,\delta}) \to 0$ as $\delta \to 0$, which implies (i) and finishes the proof of Proposition 4.2. □

## 5. Proof of Theorem 1.1

### 5.1. *Defining the time change*

The statement of Berkes and Dehling (Theorem A) for the general stable case differs from the Gaussian case in that the weighted sum $\sum_{i\le n-1} \lambda_i Y_i$ replaces the time-changed Brownian path $B(a^2(n))$. For our dynamical theorem, we wish to use a time-changed sequence for the non-Gaussian case as well. A unified statement is given in the following:

**Proposition 5.1.** *Let $(X_i)$ be an i.i.d. sequence of common distribution function $F$ in the domain of attraction of $G_{\alpha,\xi}$, $\alpha$-stable with $0 < \alpha \le 2$. In the case where $\alpha > 1$, assume without loss of generality that the $X_i$'s are centered. Then there exists a $C^1$ normalizing function $a(t)$, explicitly constructed from $F$, strictly increasing, regularly varying of index $1/\alpha$ and with regularly varying derivative, such that $a(n)$ gives a normalizing sequence for $F$, and such that there is a joining between $S$ and the $(\alpha, \xi)$-stable process $Z$, so that*

$$\text{for } \alpha \neq 1: \quad \sup_{0\le k\le n} \bigl|S(k) - Z\bigl(a^\alpha(k)\bigr)\bigr| = o\bigl(a(n)\bigr) \quad a.s.\ (log), \tag{5.1}$$

$$\text{for } \alpha = 1: \quad \sup_{0\le k\le n} \bigl|S(k) - kv\bigl(a(k)\bigr) - \widetilde{Z}\bigl(a(k)\bigr)\bigr| = o\bigl(a(n)\bigr) \quad a.s.\ (log), \tag{5.2}$$

*where $v(x) = \int_{-x}^x t\,\mathrm{d}F(t)$, and $\widetilde{Z}$ is the centered Cauchy process, see* (1.14).

**Proof.** For $\alpha = 2$, in the finite variance case we choose $a(t) = \sigma\sqrt{t}$ with $\mathrm{var}(X_0) = \sigma^2$, and statement (5.1) is proved in [11].

In all other cases, to define the time change $a(t)$ from $F$, we first improve the distribution function $F$ by convolution and then show that proving Proposition 5.1 for the smoothed law will be sufficient. To define the smoothing, we begin with the independent joining of the process $(X_i)$ with a sequence of i.i.d. standard normal variables $(X_i^*)$; that is, writing $\Pi = \prod_0^\infty \mathbb{R}$, with $(\Pi, \mu)$ the path space model of the process $(X_i)$ and $(\Pi, \mu^*)$ that for $(X_i^*)$, we let $\hat{\mu}$ denote the product measure on $\Pi \times \Pi$.

Set $\widetilde{X}_i = X_i + X_i^*$. Since $(\Pi \times \Pi, \hat{\mu})$ serves as an underlying space for both $(X_i)_{i\ge0}$ and the smoothed process $(\widetilde{X}_i)_{i\ge0}$, by Lemma 3.8 this determines a joining of $(X_i)_{i\ge0}$ and $(\widetilde{X}_i)_{i\ge0}$.

Since a process $(X_i)_{i\geq 0}$ and its partial sums $(S_k)_{k\geq 0}$ (both given by measures on $\Pi$) are measure-theoretically isomorphic via the map $(x_i)_{i\geq 0} \mapsto (\sum_{i=0}^{k-1} x_i)_{k\geq 1}$, the joining of $(X_i)_{i\geq 0}$ and $(\widetilde{X}_i)_{i\geq 0}$ equivalently gives one of $(S_k)_{k\geq 0}$ and $(\widetilde{S}_k)_{k\geq 0}$.

Now the relation $R$ on the Polish space $\prod_0^\infty \mathbb{R}$ defined by $(f, g) \in R$ iff

$$\sup_{0\leq k\leq n} \left|f(k) - g(k)\right| = \mathrm{o}\big(a(n)\big) \quad \text{a.s. (log)} \tag{5.3}$$

is an equivalence relation. We first show that the joining of $(S_k)_{k\geq 0}$ with $(\widetilde{S}_k)_{k\geq 0}$ satisfies (5.3); then from Proposition 3.9 it will follow that if $(\widetilde{S}_k)$ satisfies (5.1) or (5.2) then so does $(S_k)$.

To this end, let $(a_n)$ be any regularly varying sequence of order $1/\alpha$. First we consider the case $\alpha = 2$ with infinite variance; then $S_n^* \stackrel{\text{def}}{=} X_0^* + \cdots + X_{n-1}^* = \mathrm{o}(a_n)$ in probability. From Corollary 4 of [3], this tells us that $\max_{k\leq n} |S_n^*| = \mathrm{o}(a_n)$ a.s. (log). Then since for a.e. pair with respect to the joining measure we have $|S_n - \widetilde{S}_n| = |S_n^*|$, the relation (5.3) holds for that case.

For the case $\alpha < 2$, the law of the iterated logarithm delivers that $S_n^* = \mathrm{o}(a_n)$ a.s., and the same reasoning holds a fortiori.

Thus all we have to prove is that Proposition 5.1 holds true for the smoothed distribution $\widetilde{F}$; then there exists a joining of $(\widetilde{S}_k)_{k\geq 0}$ with the process $Z$ and hence with $Z(a^\alpha(k))$ (resp. $(-kv(a(k)) - \widetilde{Z}(a(k))))$ such that (5.1) (resp. (5.2)) holds. By Proposition 3.9 the composition of the joining of $(S_k)_{k\geq 0}$ with $(\widetilde{S}_k)_{k\geq 0}$ and that of $(\widetilde{S}_k)_{k\geq 0}$ with $Z$ will give the desired joining of $(S_k)_{k\geq 0}$ with $Z$.

The smoothed $\widetilde{F}$ has the following properties: it is still in the domain of attraction of $G_{\alpha,\xi}$, by construction is $C^1$, and has a strictly positive density on the reals hence is strictly increasing.

***Remark 5.1.*** *We note that if $(a_n)$ is a normalizing sequence for $F$, then it is also a normalizing sequence for $\widetilde{F}$, the smoothed version of $F$. Moreover, the centering sequences differ up to $\mathrm{o}(a_n)$. This follows from convergence of types,* [5], *p.* 328.

Accordingly, we shall assume without loss of generality that we begin with $F$ already smoothed, so it is a $C^1$ function with continuous, positive density on the reals. We start by constructing $a(t)$ from this $F$. Defining

$$\widehat{L}(t) \stackrel{\text{def}}{=} t^{\alpha-2} \int_{-t}^{t} x^2\, \mathrm{d}F(x), \quad t \geq 0, \tag{5.4}$$

we see from (2.2) that $\widehat{L} \sim L$, some slowly varying function, hence is slowly varying as well.

Next, we set

$$\widehat{a}(t) \stackrel{\text{def}}{=} \alpha \int_0^t \frac{u^{\alpha-1}}{\widehat{L}(u+1)}\, \mathrm{d}u \quad \text{and} \quad a(t) \stackrel{\text{def}}{=} \widehat{a}^{-1}(t) \quad \text{for } t \geq 0.$$

By (i) then (ii) of Theorem 2.2, $a(\cdot)$ is regularly varying of index $1/\alpha$. Moreover, $a(\cdot)$ is a strictly increasing function on $\mathbb{R}^+$, with $a(0) = 0$; it is easily checked that $a'(t) = \frac{\widehat{L}(a(t)+1)}{\alpha a^{\alpha-1}(t)}$. So $a'(\cdot)$ has a regularly varying derivative and that $a^\alpha(t) \sim t\widehat{L}(a(t))$.

We observe that $(a(n))$ satisfies (2.3) and is in fact also a normalizing sequence for $F$. To see this, first note that by definition of $\widehat{L}$, we have $\widehat{L} \sim L$, with $L$ as in (2.2). It follows that $a(n)$ satisfies the condition $nV(a(n)x) \sim a(n)^2 x^{2-\alpha}$ for all $x > 0$ for $V$ the truncated variance for $F$. From [5], (8.3.7) on p. 346 and top of p. 347, it then follows that $a(n)$ is a normalizing sequence for $F$.

For $\alpha = 2$ with infinite variance, we make use of Berkes and Dehling's joining, which is produced via Skorokhod embedding, but with our $a(n)$ as the normalizing sequence.

Now we move to the case $\alpha < 2$. Taking $\mu_i = (a^\alpha(i+1) - a^\alpha(i))^{1/\alpha}$, then since the process $Z$ has independent increments, writing $\delta_{1,\alpha} = 1$ if $\alpha = 1$, and $\delta_{1,\alpha} = 0$ if $\alpha \neq 1$,

$$Y_i = \frac{1}{\mu_i}\big(Z\big(a^\alpha(i+1)\big) - Z\big(a^\alpha(i)\big) - \delta_{1,\alpha}\xi\mu_i \log\mu_i\big) \quad \text{for } i \geq 0 \tag{5.5}$$

defines an independent sequence of random variables; by the scaling property of $Z$ for $\alpha \neq 1$ and of $\widetilde{Z}$ for $\alpha = 1$, each $Y_i$ has distribution $G_{\alpha,\xi}$. Hence the above equation defines a measure-preserving function from the path space of $Z$ with stable measure to the path space of $(Y_i)$, an i.i.d. sequence of $\alpha$-stable variables. Thus, with the convention $\sum_0^{-1} = 0$,

$$Z\big(a^\alpha(n)\big) = \sum_{0 \le i \le n-1} \mu_i Y_i + \delta_{1,\alpha}\xi \sum_{0 \le i \le n-1} \mu_i \log \mu_i \quad \text{for all } n \ge 0.$$

Next we define an i.i.d. sequence $(X_i)$ of distribution $F$ via the quantile transform

$$X_i = F^{-1}\big(G_{\alpha,\xi}(Y_i)\big) \quad \text{for } i \ge 0.$$

This is a measure-preserving map from the path space of $(Y_i)$ to the path space of $(X_i)$. In fact this map is one-to-one: indeed, from [31], we have that $G_{\alpha,\xi}$ is an invertible function from $\mathbb{R}$ to $(0, 1)$ for $1 \le \alpha < 2$ (and $\xi$ arbitrary in $[-1, 1]$) and also for $0 < \alpha < 1$ when $\xi \neq \pm 1$. From the invertibility of $F$, therefore, $G_{\alpha,\xi}^{-1}(F(X_i)) = Y_i$, giving a bijection as claimed. Note that this last identity remains true for $0 < \alpha < 1$ with $\xi = \pm 1$. Indeed, for $\xi = 1$ for instance, the case with $\xi = -1$ being similar, the $Y_i$'s are a.s. strictly positive, while $G_{\alpha,1}$ *is* invertible from $(0, \infty)$ to $(0, 1)$.

Now the measure isomorphism from $(Y_i)$ to $(X_i)$ gives a joining of the two path spaces. Berkes and Dehling's result (1.11) then holds with $Y_i = G_{\alpha,\xi}^{-1}(F(X_i))$, $c_k$ some centering sequence we shall describe later and with $\lambda_i = L(a_{i+1})^{1/\alpha}$. (Since they use the same joining, that given by the quantile transform, we can make use of their result here.) We remark that (1.11) was proved for *any* normalizing sequence $a_n$ for $F$, and *any* $L(t)$ slowly varying such that (2.2) holds. So (1.11) also holds with $a_n$ replaced by $a(n)$.

We note that a key ingredient in the proof of (1.11) in [3] was to first show that as $n \to \infty$

$$\frac{S(n) - \sum_{i \le n-1} \lambda_i Y_i}{a(n)} - \frac{c_n}{a(n)} \xrightarrow{\mathbb{P}} 0, \tag{5.6}$$

where $\xrightarrow{\mathbb{P}}$ denotes convergence in probability, then to apply Corollary 4. We prove (5.1), following a similar strategy. Here $\alpha \neq 1$ so we have $\sum_{i \le n-1} \mu_i Y_i = Z(a^\alpha(n))$ and thus (5.6) is equivalent to:

$$\frac{S(n) - Z(a^\alpha(n))}{a(n)} + \frac{\sum_{i \le n-1} (\mu_i - \lambda_i) Y_i}{a(n)} \xrightarrow{\mathbb{P}} 0.$$

The second term above is a normalized stable sum; computing its parameters gives $G_{\alpha,\xi_n,d_n,0}$ where

$$d_n = d_n(\alpha) \equiv \frac{1}{a^\alpha(n)} \sum_{0 \le i \le n-1} |\lambda_i - \mu_i|^\alpha = \frac{\sum_{0 \le i \le n-1} \mu_i^\alpha |1 - \lambda_i/\mu_i|^\alpha}{\sum_{0 \le i \le n-1} \mu_i^\alpha}.$$

We claim here that $\mu_i \sim \lambda_i$ which implies that $d_n \to 0$ as $n \to \infty$. As we have seen, $a(\cdot)$ has a regularly varying derivative of exponent $> -1$ and so it follows that

$$\mu_i^\alpha = \int_i^{i+1} \frac{\mathrm{d}}{\mathrm{d}s} a^\alpha(s)\,\mathrm{d}s = \int_i^{i+1} \alpha \frac{s a'(s)}{a(s)} \frac{a^\alpha(s)}{s}\,\mathrm{d}s \sim \int_i^{i+1} \frac{a^\alpha(s)}{s}\,\mathrm{d}s \sim \frac{a^\alpha(i)}{i} \sim L\big(a(i)\big) = \lambda_i^\alpha,$$

where we have used Theorem 2.2(ii) in deriving the first equivalence. Hence, indeed, $\mu_i \sim \lambda_i$ and $G_{\alpha,\xi_n,d_n,0} \xrightarrow{\mathbb{P}} 0$. Putting all this together says indeed that for $\alpha \neq 1$, $(S(n) - Z(a^\alpha(n)))/a(n) \xrightarrow{\mathbb{P}} 0$, which by Corollary 4 of [3] gives (5.1).

We move on to the case $\alpha = 1$. Here we know from [9], p. 305, that since $F$ lies in the domain of attraction of $G_{1,\xi}$, one has $(S(n) - nv(a(n)))/a(n) \xrightarrow{\text{law}} G_{1,\xi}$, where $v(x) = \int_{-x}^{x} t\, \mathrm{d}F(t)$ is the truncated mean of $F$. Having in mind the definition of $\widetilde{Z}$ (1.14) and then recalling that $\sum_{i \le n-1} \mu_i Y_i = Z(a(n)) - \xi \sum_{i \le n-1} \mu_i \log \mu_i$, (5.6) is equivalent to:

$$\frac{S(n) - nv(a(n)) - \widetilde{Z}(a(n))}{a(n)} + \frac{\sum_{i \le n-1} (\mu_i - \lambda_i) Y_i}{a(n)} + \xi \sum_{i \le n-1} \frac{\mu_i}{a(n)} \log \frac{\mu_i}{a(n)} + \frac{nv(a(n)) - c_n}{a(n)} \xrightarrow{\mathbb{P}} 0.$$

As for the case $\alpha \neq 1$, the law of the second term of the above sum is $G_{1,\xi_n,d_n,b_n} = b_n + G_{1,\xi_n,d_n,0}$, where $d_n = d_n(1) \to 0$ as $n \to \infty$. So $G_{1,\xi_n,d_n,0}$ goes to 0 in probability and denoting by $u_n$ the last two terms of the above sum plus $b_n$ we have

$$\frac{S(n) - nv(a(n))}{a(n)} - \frac{\widetilde{Z}(a(n))}{a(n)} - u_n \xrightarrow{\mathbb{P}} 0, \quad n \to \infty. \tag{5.7}$$

But since the first term in (5.7) converges in law to $G_{1,\xi}$, while the second $\widetilde{Z}(a(n))/a(n)$ has a constant law $G_{1,\xi}$, one can check that $u_n$ converges to 0 and so could be omitted in (5.7). This in conjunction with Corollary 4 of [3] yields (5.2) and completes the proof of Proposition 5.1. □

## 5.2. *From discrete time to step paths*

Proceeding toward Theorem 1.1, we next prove:

**Proposition 5.2.** *Under the conditions of Theorem 1.1, there exists a joining of $\overline{S}$ and $Z$ such that, for $\alpha \neq 1$:*

$$\left\| \overline{S} \circ \left(a^\alpha\right)^{-1} - \overline{Z}_{a^\alpha(\mathcal{Q})} \right\|_{[0,T]}^{\infty} = \mathrm{o}\left(T^{1/\alpha}\right) \quad \text{a.s. (log)} \tag{5.8}$$

*and equivalently*

$$\left\| \tau_t \left(\overline{S} \circ \left(a^\alpha\right)^{-1}\right) - \tau_t (\overline{Z}_{a^\alpha(\mathcal{Q})}) \right\|_{[0,1]}^{\infty} \to 0 \quad \text{a.s. (Cesáro)}, \tag{5.9}$$

*where $\mathcal{Q} = ([n, n+1))_{n \ge 0}$ and $\overline{Z}_{a^\alpha(\mathcal{Q})}$ denotes the step path over the partition $a^\alpha(\mathcal{Q})$, see (4.1).*

*The previous results still hold true for $\alpha = 1$ with $S - \varrho$ replacing $S$ and $\widetilde{Z}$ replacing $Z$.*

**Proof.** From Proposition 5.1 we know that there exists a set $\mathcal{B} \subseteq \mathbb{N}$ of times $n$ of integer log density zero such that for a.e. pair $(S, Z)$ (or equivalently $(\overline{S}, Z)$) with respect to the joining measure,

$$\sup_{0 \le j \le n} \left| S(j) - Z\left(a^\alpha(j)\right) \right| = \mathrm{o}\left(a(n)\right) \quad (n \notin \mathcal{B}).$$

Therefore for $\overline{\mathcal{B}} \equiv \{t \in \mathbb{R}^+ \colon [t] \in \mathcal{B}\}$, which has real log density zero,

$$\sup_{0 \le t \le R} \left| \overline{S}(t) - \overline{Z \circ a^\alpha}_{\mathcal{Q}}(t) \right| = \mathrm{o}\left(a(R)\right) \quad (R \notin \overline{\mathcal{B}}).$$

We observe using (4.6) that

$$\left\| \overline{S} \circ \left(a^\alpha\right)^{-1} - \overline{Z}_{a^\alpha(\mathcal{Q})} \right\|_{[0,a^\alpha(R)]}^{\infty} = \left\| \overline{S} - \overline{Z \circ a^\alpha}_{\mathcal{Q}} \right\|_{[0,R]}^{\infty} = \mathrm{o}\left(a(R)\right) \quad (R \notin \overline{\mathcal{B}}).$$

Since $a^\alpha(\cdot)$ is regularly varying of index 1, invertible and with regularly varying derivative, by Proposition 2.5, $\widetilde{\mathcal{B}} \stackrel{\text{def}}{=} a^\alpha(\overline{\mathcal{B}})$ also has log density zero. This proves (5.8). Lastly, it is easily checked that (5.9) holds off $\log(\widetilde{\mathcal{B}})$, which has Cesáro density zero. This finishes the proof of Proposition 5.2. □

We have done all the preparatory work, and are now ready to put these pieces together.

5.3. *End of proof of Theorem* 1.1

For $\alpha \neq 1$, by Lemma 3.3 the scaling flow of the stable process $Z$ is an ergodic flow. Since $a^{\alpha}(\cdot)$ is positive, strictly increasing and regularly varying of index one, we can apply Proposition 4.2(i) and we have for a.e. pair $(S, Z)$ with respect to the joining of Proposition 5.2:

$$d_1\big(\tau_t\big(\overline{S}\circ\big(a^{\alpha}\big)^{-1}\big),\tau_t(Z)\big)\le d_1\big(\tau_t\big(\overline{S}\circ\big(a^{\alpha}\big)^{-1}\big),\tau_t(\overline{Z}_{a^{\alpha}(Q)})\big)+d_1\big(\tau_t(\overline{Z}_{a^{\alpha}(Q)}),\tau_t(Z)\big)$$
$$\le \big\|\tau_t\big(\overline{S}\circ\big(a^{\alpha}\big)^{-1}\big)-\tau_t(\overline{Z}_{a^{\alpha}(Q)})\big\|_{[0,1]}^{\infty}+d_1\big(\tau_t(\overline{Z}_{a^{\alpha}(Q)}),\tau_t(Z)\big)\to 0 \quad \text{a.s. (Cesáro)},$$

where the set of zero Cesáro density is the union of the two sets from Propositions 5.2(ii) and 4.2(i).

Concluding, this gives (1.15) which by Lemma 3.5 is equivalent to (1.16).

Replacing $Z$ by $\widetilde{Z}$ and $S$ by $S-\varrho$, the previous proof runs exactly in the same way for $\alpha = 1$ since the scaling flow of $\widetilde{Z}$ is ergodic, while Proposition 4.2 holds for *any* ergodic self-similar process.

This proves (1.16) and (1.15) of Theorem 1.1, and together with Lemma 3.6 (which extends the result to the metric $d_{\infty}$) completes the proof of the theorem.

5.4. *Comparing paths, and alternative time changes*

First we give an alternate definition of time change, as promised in Proposition 1.2.

**Lemma 5.3.** *Let $F$ be an element of the domain of attraction of $G_{\alpha}$ and let $(a_n)$ be a normalizing sequence for $F$ (satisfying (2.3)). We denote by $\overline{a}(t)$ the polygonal interpolation of $(a_n)$. Then setting $\widetilde{a}(t) = \alpha\int_0^t \frac{\overline{a}(s)}{s}\,\mathrm{d}s$, this defines a $C^1$, strictly increasing function with regularly varying derivative such that $\widetilde{a}(n)$ is a normalizing sequence for $F$.*

Next we come to:

**Proof of Proposition 1.2.** We begin by proving (i). Let $a(\cdot)$ be the smooth time change constructed explicitly from $F$ in Proposition 5.1, that is, after first smoothing the distribution if necessary. By assumption $\widetilde{a}(t) \sim a(t)$ so $\widetilde{a}(n)$ is also a normalizing sequence for $F$. Then an examination of the proof of Proposition 5.1 shows that its conclusion holds with $\widetilde{a}(n)$ taking the place of $a(n)$. Now $\widetilde{a}(\cdot)$ is invertible and has a regularly varying derivative, which are the only additional properties of the time change needed for the rest of the proof of Theorem 1.1 to go through; hence the statements of the theorem are also true for $\widetilde{a}(\cdot)$.

Next we consider two copies $\overline{S}_{(1)}, \overline{S}_{(2)}$ of the random walk process $\overline{S}$ for $F$. From Theorem 1.1 there exists a joining of $\overline{S}_{(1)}$ with $Z$, and a joining of $Z$ with $\overline{S}_{(2)}$, such that almost every pair $(\overline{S}_{(1)}\circ(a^{\alpha})^{-1}, Z)$ lies in the same Cesáro stable manifold and similarly for almost every pair $(Z, \overline{S}_{(2)}\circ(\widetilde{a}^{\alpha})^{-1})$. The composition of these two joinings therefore gives a joining of the processes $\overline{S}_{(1)}$ and $\overline{S}_{(2)}$ for which the last part of (i) holds.

We move to the proof of (ii). We associate to $F$ and $\widetilde{F}$ (again these may be non-smoothed distributions!) two smooth time changes $a(\cdot)$ and $\widetilde{a}(\cdot)$; this can be any time change satisfying the conditions of (i) in the proposition, so it can for instance be explicitly constructed as in the proof of Proposition 5.1 after first smoothing the distribution, or as in Lemma 5.3.

By assumption, the slowly varying functions $L$ and $\widetilde{L}$ associated to $F$ and $\widetilde{F}$ given in (i) of Theorem 2.3 are equivalent. This together with (2.3) written for $a(\cdot)$ and $\widetilde{a}(\cdot)$ easily gives that $a(t) \sim \widetilde{a}(t)$.

We now turn to the last part of (ii). First by part (i) for $F$ we join $\overline{S}\circ(\widetilde{a}^{\alpha})^{-1}$ with $\overline{S}\circ(a^{\alpha})^{-1}$, then by Theorem 1.1 we join $\overline{S}\circ(a^{\alpha})^{-1}$ with $Z$, and $Z$ with $\overline{\widetilde{S}}\circ(\widetilde{a}^{\alpha})^{-1}$; then once more by part (i) but now for $\widetilde{F}$, we join $\overline{\widetilde{S}}\circ(\widetilde{a}^{\alpha})^{-1}$ with $\overline{\widetilde{S}}\circ(a^{\alpha})^{-1}$. From Remark 3.1, we have common underlying space for these five processes, and hence four joinings, such that by Proposition 3.9 they are a.s. all simultaneously in the same Cesáro stable manifold.

We keep with the notation of the proof of Proposition 5.1, taking first the case $\alpha \neq 1$. We recall that using the quantile transform, we joined $(S_n)$ for $F$ with $(\sum_{i\le n}\mu_i Y_i)$ where $(Y_i)$ is an i.i.d. sequence of $\alpha$-stable variables and $\mu_i^{\alpha} = a^{\alpha}(i+1) - a^{\alpha}(i)$, and then proved the a.s.i.p. (log) of (5.1).

By the same scheme, there exists a joining of $(\widetilde{S}_n)$ for $\widetilde{F}$ with $(\sum_{i\le n}\widetilde{\mu}_i\widetilde{Y}_i)$ and a corresponding a.s.i.p. (log). Now the two processes $(Y_i)$ and $(\widetilde{Y}_i)$ have the same law, hence so do $(\sum_{i\le n}\mu_i Y_i)$ and $(\sum_{i\le n}\mu_i\widetilde{Y}_i)$; this correspondence defines a third joining. On the other hand, since $\widetilde{\mu}_i \sim \mu_i$ we have $\sum_{i\le N}(\widetilde{\mu}_i - \mu_i)\widetilde{Y}_i = o(a(N))$ a.s. (log), as shown at the end of the proof of Proposition 5.1. As a result, taking the composition of these joinings produces a joining of $(\overline{S}_n)$ with $(\overline{\overline{S}}_n)$ such that (1.17) is satisfied. This finishes the proof for $\alpha \neq 1$; all of the above then holds for $\alpha = 1$ upon centering. □

## 6. Proof of Proposition 1.3: Generic points and pathwise limit theorems

In the passage from the a.s.i.p. (log) of our main theorem to the pathwise limit theorems of Proposition 1.3, we shall need the two lemmas which follow. First, as we saw in the Introduction, the Cesáro average of a continuous bounded observable is constant on an equivalence class $W^{\mathrm{s},d}_{\mathrm{Ces}}(g)$ for $g \in D$. We now come to this related statement:

**Lemma 6.1.** *Let $\nu$ be an ergodic invariant probability measure for the flow $\tau_t$ on $D$. Suppose that $\mu$ is a probability measure on $D$, with $\hat{\nu}$ a joining of $\mu$ and $\nu$ such that for $\hat{\nu}$-a.e. pair $(f, g)$ we have $f \in W^{\mathrm{s},d_\infty}_{\mathrm{Ces}}(g)$; then $\mu$-a.e. path $f$ is a generic point for the flow $(D, \nu, \tau_t)$. This statement holds with $d_\infty$ replaced by any metric $\mathrm{d}$ which gives the same topology.*

**Proof.** By assumption there is a set $\widehat{G} \subseteq D \times D$ of $\hat{\nu}$-measure one, such that for every $(f, g) \in \widehat{G}$, then there exists a set $\mathcal{B}$ of Cesáro density zero such that $d_\infty(\tau_t f, \tau_t g) \to 0$ as $t \to \infty$ for $t \notin \mathcal{B}$.

Let $\Phi \in \mathrm{UCB}(D, d_\infty)$. By the uniform continuity of $\Phi$,

$$H(t) \equiv \left|\Phi(\tau_t f) - \Phi(\tau_t g)\right| \to 0 \quad \text{as } t \to \infty, \text{ for } t \notin \mathcal{B}.$$

Then since $\Phi$ is bounded, the Cesáro average of $H$ is zero and hence the Cesáro averages of $\Phi(\tau_t g)$ and $\Phi(\tau_t f)$ are the same.

By Lemma 7.2 $(D, d_\infty)$ is a Polish space, so by ergodicity of the scaling flow $(D, \tau_t, \nu)$ we can apply Fomin's theorem. This guarantees that there is a set $G_2 \subseteq D$ of measure one of generic points. Thus, for every $\Phi \in \mathrm{CB}(D, d_\infty)$, then for all $g \in G_2$,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \Phi(\tau_t g)\,\mathrm{d}t = \int_D \Phi\,\mathrm{d}\nu \equiv \langle \Phi, \nu\rangle. \tag{6.1}$$

We show how to pass this on to the paths $f$.

Since $\widehat{G} \cap (D \times G_2)$ has $\hat{\nu}$-measure one, therefore its projection $G_1$ to the $f$-coordinate has $\mu$-measure one. (A theorem of Rochlin [30] guarantees the forward image is a measurable set; this uses the fact that we have Lebesgue spaces.) For any $f \in G_1$ there exists $(f, g) \in \widehat{G}$; for any such pair, we know that $f \in W^{\mathrm{s},d_\infty}_{\mathrm{Ces}}(g)$ with $g$ a generic point for $\tau_t$.

We conclude that for a set of $\mu$-measure one of paths $f$, then for all $\Phi \in \mathrm{UCB}(D, d_\infty)$, (6.1) holds with $f$ replacing $g$. Then, by p. 12 of [4] for each such $f$ the same holds for $\Phi \in \mathrm{CB}(D, d_\infty)$, finishing the proof for $d_\infty$. Note lastly that the proof just given works for any equivalent metric. □

Next we see how to smooth a function in $\mathrm{UCB}(D, d_1)$ to one in $\mathrm{CB}(D, d_\infty)$, by *convolution along the flow*.

**Lemma 6.2.** *Let $\Phi \in \mathrm{UCB}(D, d_1)$ for $D = D_{\mathbb{R}^+}$. For all $b > 0$, and all $f$ in $D$, define*

$$\widehat{\Phi}(f) \equiv \Phi_b(f) = \frac{1}{b}\int_0^b \Phi(\tau_t f)\,\mathrm{d}t.$$

*Then $\widehat{\Phi}$ is in $\mathrm{CB}(D, d_\infty)$. The space averages of $\Phi$ and $\widehat{\Phi}$ agree, and moreover if the time average of $\Phi$ exists for a given $f \in D$ then the same is true for $\widehat{\Phi}$ (with the same value).*

**Proof.** We start by proving the $d_\infty$-continuity of $\widehat{\Phi}$. To this end, we prove its sequential continuity: for any sequence of elements $(f_n)$ of $D$ that $d_\infty$-converges to $f \in D$, $\widehat{\Phi}(f_n)$ converges to $\widehat{\Phi}(f)$.

From $d_\infty(f_n, f) \to 0$ one gets that for all $\delta > 0$, $\int_0^\infty e^{-s} \chi_{d_s(f_n,f)>\delta}\, ds \to 0$ as $n \to \infty$. Thus, for any $0 < b < c$ we have $\int_b^c \chi_{d_s(f_n,f)>\delta}\, ds$ goes to zero as $n \to \infty$.

Next, recalling that $\Delta_s = \tau_{\log s}$, then after a change of variables we have, for all $\delta > 0$,

$$\begin{aligned}\left|\widehat{\Phi}(f_n) - \widehat{\Phi}(f)\right| &\le \frac{1}{b}\int_1^{e^b} \left|\Phi(\Delta_s f_n) - \Phi(\Delta_s f)\right| \frac{ds}{s} \\ &\le \frac{2\|\Phi\|^\infty_{D_{\mathbb{R}^+}}}{b}\int_1^{e^b} \chi_{d_s(f_n,f)>\delta}\, ds + \frac{1}{b}\int_1^{e^b} \left|\Phi(\Delta_s f_n) - \Phi(\Delta_s f)\right| \chi_{d_s(f_n,f)\le\delta} \frac{ds}{s}.\end{aligned}$$

Hence, at fixed $b > 0$, the first term above goes to zero as $n \to \infty$. We turn to the second integral.

We know $\Phi$ is $d_1$-uniformly continuous: for all $\varepsilon > 0$, there exists $\hat{\delta} > 0$ such that for all $f, g$ in $D$ satisfying $d_1(f, g) < \hat{\delta}$, we have $|\Phi(f) - \Phi(g)| < \varepsilon$.

So, choosing $\delta < \hat{\delta}$, then remembering that $d_s(f_n, f) = s^{1/\alpha} d_1(\Delta_s f_n, \Delta_s f)$, we get that $d_s(f_n, f) \le \delta$ implies $d_1(\Delta_s f_n, \Delta_s f) < \hat{\delta}$ (for $s \ge 1$). By the uniform continuity of $\Phi$, the second integral above is less than $\varepsilon$. This proves that $\widehat{\Phi}(f_n)$ goes to $\Phi(f)$, as $n \to \infty$.

Hence $\widehat{\Phi}$ is bounded and continuous with respect to $d_\infty$. Next we compare the time and space averages of $\widehat{\Phi}$ with those of $\Phi$.

From the definition of $\widehat{\Phi}$, with the help of Fubini's theorem, we have

$$\frac{1}{T}\int_0^T \widehat{\Phi}(\tau_t f)\, dt = \frac{1}{bT}\int_{0\le u\le v\le b} \Phi(\tau_v f)\, du\, dv + \frac{1}{T}\int_b^T \Phi(\tau_v f)\, dv + \frac{1}{bT}\int_T^{T+b} \Phi(\tau_v f)(T + b - v)\, dv.$$

Since $\Phi$ is bounded, the first and the last terms of the above sum go to 0 as $T \to \infty$. Therefore the time averages of $\widehat{\Phi}$ and $\Phi$ for a given $f$ agree. Lastly, $\int_D \widehat{\Phi}\, d\nu = \int_D \Phi\, d\nu$ since $\nu$ is $\tau$-invariant. □

**Proof of Proposition 1.3.** The proof of (i)(a) follows from the a.s.i.p. for $Y$ and $U$, for the metric $d_\infty$, together with Fomin's theorem applied to the scaling flow and Lemma 6.1 for $d_\infty$.

*Proving* (i)(b): First we show that (i)(a) also holds true for $\Phi \in \mathrm{UCB}(D, d_1)$. To this end we construct a new function $\widehat{\Phi} \in \mathrm{CB}(D, d_\infty)$ by "convolving $\Phi$ along the flow" $\tau_t$, as carried out in Lemma 6.2; we then apply (i)(a) just proved to $\widehat{\Phi}$, and as shown in the lemma, the averages of $\Phi$ and $\widehat{\Phi}$ agree.

Lastly, we apply this to a specific $\Phi$. Starting with $\psi \in \mathrm{UCB}(\mathbb{R})$, define $\Phi : D \to \mathbb{R}$ by $\Phi(g) = \psi(g(1))$. From the definition of the pseudometric $d_1$, see (3.3), $\Phi$ is in $\mathrm{UCB}(D, d_1)$; indeed, all one needs to check is that for all $f, g \in D$ such that $d_1(f, g) < \delta$ then $|f(1) - g(1)| < \delta$ (this is so because in the definition of $d_1$, $\lambda(1) = 1$ with $\lambda$ the change of parameter). And (i)(b) is proved.

*Proof of* (ii): To this end, we first rewrite (i)(a), using for simplicity the notation $f = \overline{S}$, $h = a^\alpha$ and $\widehat{h} = h^{-1}$ and changing variables with $s = e^t$, and so for $\mu$-a.e. $f$, for any $\Phi \in \mathrm{CB}(D, d_\infty)$, we have

$$\lim_{T\to\infty} \frac{1}{\log T}\int_1^T \Phi\left(\frac{f(\widehat{h}(s\cdot))}{s^{1/\alpha}}\right)\frac{ds}{s} = \langle \Phi, \nu\rangle. \tag{6.2}$$

A key step will be proving that, for the non-complete metrics $d_1^0$ and $\tilde{d}_\infty^0$ (defined in Section 8),

$$\lim_{s\to\infty} d_1^0(f_s, g_s) = 0 \quad \text{and} \quad \lim_{s\to\infty} \tilde{d}_\infty^0(f_s, g_s) = 0 \tag{6.3}$$

with

$$f_s(x) = \frac{f(\widehat{h}(sx))}{s^{1/\alpha}} \quad \text{and} \quad g_s(x) = \frac{f(\widehat{h}(s)x)}{s^{1/\alpha}}.$$

We define $\lambda_s(x) = \widehat{h}(sx)/\widehat{h}(s)$ for $x \ge 0$. From the definitions of $f_s$ and $g_s$ one can see that $\lambda_s$ was chosen in such a way that $\|f_s(x) - g_s \circ \lambda_s(x)\|^\infty_{[0,T]} = 0$ for any $T > 0$. Thus proving (6.3) reduces to proving that $\lambda_s$ converges uniformly to the identity on $[0, T]$ as $s \to \infty$.

And indeed, it is easily checked that $\lambda_s$ is increasing and continuous, with $\lambda_s(0) = 0$ and that $\lambda_s$ converges uniformly to the identity on any compact interval $[\delta_0, T]$ with $0 < \delta_0 \le T$, see [5], p. 22. Now, from the increasingness of $\lambda_s$ we get that $\|\lambda_s(x) - x\|^{\infty}_{[0,\delta_0]} \le \lambda_s(\delta_0) + \delta_0$. This implies that $\|\lambda_s(x) - x\|^{\infty}_{[0,T]} \to 0$, as $s \to \infty$. Then (6.3) follows from the definition of $\tilde{d}^0_\infty$.

Next we see how to use that to deduce (1.20) from (6.2). Beginning with $\Phi \in \mathrm{UCB}(D, \tilde{d}^0_\infty)$, since $\tilde{d}^0_\infty(f_s, g_s) \to 0$, then from (6.3) $\Phi(g_s)$ has the same log average as $\Phi(f_s)$, which equals $\langle \Phi, \nu \rangle$ by (6.2). Changing variables ($t = \widehat{h}(s)$), then using Theorem 2.2, we have proved (1.20) for $\Phi \in \mathrm{UCB}(D, \tilde{d}^0_\infty)$. By p. 12 of [4] this holds also for $\Phi \in \mathrm{CB}(D, \tilde{d}^0_\infty) = \mathrm{CB}(D, d_\infty)$, since by Proposition 8.3 these metrics give the same topology. This completes the proof of (1.20).

Note that for the case $\alpha = 1$ we use the self-similar measure $\widetilde{\nu}$ rather than the Cauchy measure $\nu$ itself, and replace paths $\overline{S} \circ \widehat{h}$ by $\overline{(S - \varrho)} \circ \widehat{h}$ where $\varrho$ is the centering from Theorem 1.1; for this it is important that our lemmas were stated for general ergodic scaling flows. □

## 7. Completeness of *D* and continuity of *τ*

Having a Polish space is of crucial importance in this paper; this follows from [38], Theorem 2.6, where it is shown that $(D, d_\infty)$ is complete. We begin this section with an alternative proof of Whitt's result. Following that, we show that the scaling flow $\tau$ is $J_1$-continuous.

**Lemma 7.1.** *Let $f \in D$; then the set of continuity points of $f$ has a countable complement.*

**Proof.** For any $\varepsilon > 0$, the points where the jump is $>\varepsilon$ cannot have an accumulation point (otherwise a one-sided limit of $f$ will not exist contradicting the definition of $D$) so this is finite on any compact interval and the claim follows. □

**Lemma 7.2.** *$(D, d_\infty)$ is a complete metric space.*

**Proof.** Let $(f_n)_{n\ge1}$ be a Cauchy sequence for $d_\infty$; we shall find $f \in D$ to which $(f_n)$ converges. Let $(\widetilde{f_k})_{k\ge0}$ be a subsequence $\widetilde{f_k} = f_{n_k}$ for $n_0 < n_1 < \cdots$ such that $d_\infty(\widetilde{f_k}, \widetilde{f}_{k+1}) < 2^{-2k}$.

We set

$$E_k = \{A \in \mathbb{R}^+ \colon d_A(\widetilde{f_k}, \widetilde{f}_{k+1}) > 2^{-k}\}, \quad k \ge 0.$$

By Markov's inequality $\mathbb{P}(g > a) \le \frac{1}{a}\mathbb{E}(g)$; for $g(A) = d_A(\widetilde{f_k}, \widetilde{f}_{k+1})$ and $\mathbb{P}$ the exponential distribution on $\mathbb{R}^+$, we have

$$\mathbb{P}(E_k) \le 2^k \mathbb{E}(g) = 2^k d_\infty(\widetilde{f_k}, \widetilde{f}_{k+1}) \le 2^k 2^{-2k} = 2^{-k}.$$

By Borel–Cantelli, $\mathcal{G} = \liminf E_k^c$ is a set of full $\mathbb{P}$- (hence Lebesgue) measure. Thus for each $A \in \mathcal{G}$, $d_A(\widetilde{f_k}, \widetilde{f}_{k+1}) \le 2^{-k}$ for $k \ge k_0(A)$ and hence by the triangle inequality, $d_A(\widetilde{f}_{k+1}, \widetilde{f}_{k+l}) \le 2^{-k}$ for any $l \ge 1$. Thus $(\widetilde{f_k})_{k\ge1}$ is a $d_A$-Cauchy sequence for all $A$ in $\mathcal{G}$.

Since $d_1$ is a complete metric on $D_{[0,1]}$, so is $d_A$ on $D_{[0,A]}$. Thus for each $A \in \mathcal{G}$ there exists $f_A \in D$ to which $(\widetilde{f_k})$ converges for $d_A$. Now let $\mathcal{G}_A$ denote the intersection of $\mathcal{G}$ with the set of continuity points in $[0, A)$ of $f_A$, which by Lemma 7.1, is dense and of full Lebesgue measure in $[0, A)$.

Choose $A_0 < A$ with $A_0 \in \mathcal{G}_A$. We know that $d_A(\widetilde{f_k}, f_A) \to 0$ and we claim that $d_{A_0}(\widetilde{f_k}, f_A) \to 0$ as well. Indeed, letting $\lambda$ be the coordinate change for $d_A(\widetilde{f_k}, f_A)$, we define $\lambda_0$ on $[0, A_0]$, by modifying $\lambda$ on a small interval to the left of $A_0$; on this interval $\lambda_0$ is defined to be linear increasing with $\lambda_0(A_0) = A_0$. Since $|||\lambda|||_A$ is small and $A_0$ is a continuity point of $f_A$, we see that $d_{A_0}(\widetilde{f_k}, f_A)$ is small. Hence $f_A = f_{A_0}$ on the interval $[0, A_0]$.

Now consider $A, B \in \mathcal{G}$ with $A < B$. We repeat the argument just given for $A_0 \in \mathcal{G}_A \cap \mathcal{G}_B$, and have that $f_A = f_B$ on the interval $[0, A_0]$. It follows that $f_A = f_B$ on $[0, A)$, and hence, $\mathcal{G}_A = \mathcal{G}_B \cap [0, A)$. We let $\widetilde{\mathcal{G}}$ be the nested union of $\mathcal{G}_A$ over $A \in \mathcal{G}$; there is thus a unique $f \in D$ such that $d_A(\widetilde{f_k}, f) \to 0$ for all $A \in \widetilde{\mathcal{G}}$, with this set dense and of full measure in $\mathbb{R}^+$. Hence, by the LDCT, $d_\infty(\widetilde{f}_k, f) \to 0$ as $k \to \infty$ and by the triangle inequality indeed, $d_\infty(f_n, f) \to 0$. $\square$

**Proposition 7.3.** *The scaling flow $\tau$ is (jointly) continuous on $D$ for the $J_1$ topology on $D = D_{\mathbb{R}^+}$.*

**Proof.** We give the proof for the metric $d_\infty$, though by Lemma 8.1, $d_\infty^0$ could be used instead.

By definition for the flow $\tau$ to be continuous means it is continuous as a function $\tau : D \times \mathbb{R} \to D$. That is for all $t_0 \in \mathbb{R}$ and $f \in D$, if $t$ is close to $t_0$ and $g$ is $d_\infty$-close to $f$ then $\tau_t g$ is $d_\infty$-close to $\tau_{t_0}(f)$. By the triangle inequality,

$$d_\infty(\tau_t g, \tau_{t_0} f) \le d_\infty(\tau_t g, \tau_t f) + d_\infty\bigl(\tau_{t-t_0}(\tau_{t_0} f), \tau_{t_0} f\bigr). \tag{7.1}$$

We first show the time continuity of $\tau_t$ at 0: that for all $\widetilde{f} \in D$, $s \mapsto \Delta_s \widetilde{f} = \tau_{\log s} \widetilde{f}$ is continuous at 1 for $d_\infty$. By LDCT, it suffices to prove that for every continuity point of $\widetilde{f}$, that is almost every $A > 0$, $d_A(\Delta_s \widetilde{f}, \widetilde{f})$ goes to 0 as $s \to 1$.

To this end, pick $\varepsilon > 0$ and assume that $s$ is close enough to 1 so that $A - \varepsilon < sA$. Set $\lambda_s(x) = sx$ on $[0, (A-\varepsilon)/s]$ and linear on $[(A-\varepsilon)/s, A]$ so that $\lambda_s(A) = A$. The aforedefined function is continuous, strictly increasing of $[0, A]$ onto itself and it is easily checked that $|||\lambda_s|||_A$ goes to 0 as $s \to 1$. The same holds for $\|\Delta_s \widetilde{f} - \widetilde{f} \circ \lambda_s\|_{[0,A]}^\infty$ since $\widetilde{f}$ is assumed to be continuous at $A$.

Next we show that the first term in (7.1), or equivalently $d_\infty(\Delta_s f, \Delta_s g)$ with $s = \mathrm{e}^t$, is small for $g$ $d_\infty$-close to $f$, and $s$ close to $\mathrm{e}^{t_0}$. Note that for all $N > 0$ and any sufficiently small $\delta$, we have

$$d_\infty(f, g) > \int_0^N \mathrm{e}^{-A} \frac{d_A(f, g)}{1 + d_A(f, g)} \,\mathrm{d}A > \frac{\delta}{2} \int_0^N \mathrm{e}^{-A} \chi_{d_A(f,g) > \delta} \,\mathrm{d}A, \tag{7.2}$$

where $\chi$ is the indicator function. From Lemma 3.4, we get $d_A(\Delta_s f, \Delta_s g) = s^{-1/\alpha} d_{As}(f, g)$, so

$$\begin{aligned} d_\infty(\Delta_s f, \Delta_s g) &= \left(\int_0^N + \int_N^\infty\right) \mathrm{e}^{-A} \frac{d_{As}(f, g)}{s^{1/\alpha} + d_{As}(f, g)} \,\mathrm{d}A \\ &\le \int_0^N \mathrm{e}^{-A} \chi_{d_{As}(f,g) > \delta} \,\mathrm{d}A + \frac{\delta}{s^{1/\alpha}} + \mathrm{e}^{-N} \end{aligned} \tag{7.3}$$

as $x \mapsto x/(s^{1/\alpha} + x)$ is strictly increasing and bounded by 1. Now, choose $N$ large enough.

For $s \le 1$, $\mathrm{e}^{-A} < \mathrm{e}^{-As}$ and we are done by first changing variables in (7.3) and then using (7.2). As for $s > 1$, by Hölder's inequality, one has

$$\int_0^N \mathrm{e}^{-A} \chi_{d_{As}(f,g) > \delta} \,\mathrm{d}A \le N^{1-1/s} \left(\int_0^N \mathrm{e}^{-As} \chi_{d_{As}(f,g) > \delta} \,\mathrm{d}A\right)^{1/s},$$

and we conclude just as for $s \le 1$. Thus, $d_\infty(\Delta_s f, \Delta_s g)$ is small, as claimed. $\square$

## 8. Non-complete metrics for the $J_1$ topology on $D_{\mathbb{R}^+}$

In this section we define two non-complete metrics on $D$, which both give the same topology as $d_\infty$. The first of these, $d_\infty^0$ (see Section 3.1) is easier to compare with $d_\infty$; the second $\tilde{d}_\infty^0$, which is closer to Stone's original definition of the topology on $D$, is better adapted for use in the proof of Proposition 1.3.

**Lemma 8.1.** *The complete and incomplete metrics $d_\infty$ and $d_\infty^0$ on $D$ are equivalent.*

**Proof.** It is sufficient to prove that sequential convergence corresponds for the two metrics. We start by showing that if $d_\infty(f_n, f) \to 0$ for $f_n, f \in D$ then $d_\infty^0(f_n, f) \to 0$.

Having the definition of $d_\infty$ in mind, we write $d_\infty(f_n, f) = \mathbb{E}(X_n)$, the expected value with respect to the exponential law of parameter one of the function $X_n(A) \equiv \frac{d_A(f_n,f)}{1+d_A(f_n,f)}$.

As a result, $d_\infty(f_n, f) \to 0$ implies that $X_n \to 0$ in $L^1$. Thus there exists a subsequence $\phi(n)$ such that for a.e. $A > 0$, $d_A(f_{\phi(n)}, f) \to 0$. Since by Theorem 14.1 of [4] the metrics $d_1$ and $d_1^0$ on $D_I$ give the same topology, then so do $d_A$ and $d_A^0$ on $D_{[0,A]}$. Hence for a.e. $A > 0$, $d_A^0(f_{\phi(n)}, f) \to 0$ which by LDCT implies that $d_\infty^0(f_{\phi(n)}, f) \to 0$. We claim that $d_\infty^0(f_n, f) \to 0$.

We have just proved that 0 is an accumulation point for the non-negative and bounded (by 1) sequence $(d_\infty^0(f_n, f))$. Suppose that $l \neq 0$ is another accumulation point. Then $d_\infty^0(f_{\psi(n)}, f) \to l$ for a subsequence $\psi(n)$. But since a fortiori $d_\infty(f_{\psi(n)}, f) \to 0$, running the above reasoning shows that there exists a subsequence $\widehat{\psi}(n)$ for which $d_\infty(f_{\psi(\widehat{\psi}(n))}, f) \to 0$. This contradicts $d_\infty^0(f_{\psi(n)}, f) \to l \neq 0$ and proves that $d_\infty^0(f_n, f) \to 0$, as desired. Reversing the argument, $d_\infty^0$-convergence implies $d_\infty$-convergence, so the two metrics give the same topology. □

In fact historically the first definition of a $J_1$ topology on $D$ was due to Stone [34]: for $\Lambda_\infty = \{\lambda : \mathbb{R}^+ \to \mathbb{R}^+ \text{continuous, increasing, onto}\}$ then $f_n \to f$ iff there exists $\lambda_n \in \Lambda_\infty$ such that

$$\forall T > 0, \quad \left\|\lambda_n(x) - x\right\|_{[0,T]}^\infty \to 0 \quad \text{and} \quad \|f_n - f \circ \lambda_n\|_{[0,T]}^\infty \to 0 \quad \text{as } n \to \infty.$$

We next see it is possible to define a metric on $D$ which is more closely based on Stone's idea, that is, allowing for parameter changes $\lambda$ which do not necessarily fix the endpoints of a compact interval $[0, A]$. This metric, denoted $\tilde{d}_\infty^0$, is used in the proof of part (ii) of Proposition 1.3.

The interesting point in finding an appropriate definition will be to somehow achieve the triangle inequality; in the case of $d_A$ or $d_A^0$ that was easy exactly because the $\lambda \in \Lambda_A$ leaves the endpoints fixed. Here we borrow a nice idea from Kalashnikov's presentation [19] of a complete metric on $D$, though things are simpler in the present case.

***Definition 8.2.*** *For fixed $f, g \in D$, then for a chosen $\lambda \in \Lambda_\infty$, we define*

$$t_\lambda = t_{\lambda,f,g} = \sup\left\{t \geq 0\colon \|f - g \circ \lambda\|_{[0,t]}^\infty \leq \frac{1}{t} \text{ and } \left\|\lambda(x) - x\right\|_{[0,t]}^\infty \leq \frac{1}{t}\right\}. \tag{8.1}$$

*We then define*

$$\rho(f, g) = \min\left\{\frac{1}{2}; \Big(\sup_{\lambda \in \Lambda_\infty} t_\lambda\Big)^{-1}\right\}$$

*and*

$$\tilde{d}_\infty^0(f, g) = \rho(f, g) + \rho(g, f).$$

**Proposition 8.3.** *$\tilde{d}_\infty^0$ defines a metric on $D$.*

**Proof.** We note that $\rho(f, g) = 0$ iff $f = g$; this passes on to $\tilde{d}_\infty^0$, which is defined so as to be symmetric. All that is left to do is to show the triangle equality for $\rho$, since this property will pass on to $\tilde{d}_\infty^0$ as well. So it suffices to show: for $f, g, h \in D$, $\rho(f, h) \leq \rho(f, g) + \rho(g, h)$.

We assume that both $\rho(f, g)$ and $\rho(g, h)$ are $<1/2$ (as it is trivial otherwise). This is equivalent to saying that there exist $\lambda, \mu \in \Lambda_\infty$, such that $t_\lambda = t_{\lambda,f,g}$ and $t_\mu = t_{\mu,g,h}$ are $>2$.

Fixing such a $\lambda$ and $\mu$, we define $\tilde{t}$ by

$$\frac{1}{\tilde{t}} = \frac{1}{t_\lambda} + \frac{1}{t_\mu},$$

and have $1 < \tilde{t} \leq \min\{t_\lambda, t_\mu\}$. We easily check that $\lambda(\tilde{t}) \leq t_\mu$.

Setting $\nu = \mu \circ \lambda$, we have:

$$\|f - h \circ \nu\|^{\infty}_{[0,\tilde{t}]} \le \|f - g \circ \lambda\|^{\infty}_{[0,\tilde{t}]} + \|g - h \circ \mu\|^{\infty}_{[0,\lambda(\tilde{t})]} \le \|f - g \circ \lambda\|^{\infty}_{[0,t_\lambda]} + \|g - h \circ \mu\|^{\infty}_{[0,t_\mu]}$$

and similarly

$$\|\nu(x) - x\|_{[0,\tilde{t}]} \le \|\lambda(x) - x\|_{[0,t_\lambda]} + \|\mu(x) - x\|_{[0,t_\mu]}.$$

So both $\|f - h \circ \nu\|^{\infty}_{[0,\tilde{t}]}$ and $\|\nu(x) - x\|_{[0,\tilde{t}]}$ are $\le \frac{1}{t_\lambda} + \frac{1}{t_\mu} = \frac{1}{\tilde{t}}$. Thus $\tilde{t} \le t_\nu = t_{\nu,f,h}$ and so $\frac{1}{t_\nu} \le \frac{1}{t_\mu} + \frac{1}{t_\lambda}$. This implies that $\rho(f,h) \le \rho(f,g) + \rho(g,h)$, completing the proof that $\tilde{d}^0_\infty$ is a metric. □

Next we relate $\tilde{d}^0_\infty$ to the other metrics, which were defined from integration of metrics on $D_{[0,A]}$.

**Proposition 8.4.** *The metrics $\tilde{d}^0_\infty$ and $d^0_\infty$ are equivalent; they give the same topology as Stone's.*

**Proof.** By considering sequential convergence, it is clear that Stone's topology is the same as that given by $\tilde{d}^0_\infty$.

Now assume that $d^0_\infty(f_n, f) \to 0$. Then for any $\varepsilon, T > 0$, $\exists A > T$ such that $d^0_A(f_n, f) \le \varepsilon$ for $n$ large enough, i.e. there exists $\mu_n \in \Lambda_A$ such that both $\|f_n - f \circ \mu_n\|^{\infty}_A$ and $\|\mu_n(x) - x\|^{\infty}_A$ are $\le \varepsilon$.

Defining $\lambda_n$ to be the extension of $\mu_n$ to $\mathbb{R}^+$ by taking $\lambda_n(x) = x$ for $x > A$, then a fortiori $\|f_n - f \circ \mu_n\|^{\infty}_T$ and $\|\mu_n(x) - x\|^{\infty}_T$ are $\le \varepsilon$. This shows that $f_n \to f$ in Stone's sense.

Conversely, if $f_n \to f$ in Stone's topology then there exists $\lambda_n$ on $\mathbb{R}^+$ such that $\forall T > 0$, $\|f_n - f \circ \lambda_n\|^{\infty}_{[0,T]}$ and $\|\lambda_n(x) - x\|^{\infty}_T \to 0$, as $n \to \infty$. By Lemma 7.1, a.e. $A \in \mathbb{R}^+$ is a point of continuity of $f$; let $A \in [0, T]$ be such a point. Then, as argued in the proof of Lemma 7.2, by making a small change in $\lambda_n$ near $A$ to get $\tilde{\lambda}_n$ so that $\tilde{\lambda}_n(A) = A$, we can achieve $d^0_A(f_n, f) < 2\varepsilon$. This works for Lebesgue-a.e. $A > 0$, hence $d^0_A(f_n, f) \to 0$ for a.e. $A$ and so by LDCT, $d^0_\infty(f_n, f) \to 0$. This finishes the proof of Proposition 8.4. □

## Acknowledgements

We thank Artur Lopes for conversations which led to a project regarding certain maps of the interval; this work is a first part for that study, see [13] and [14] where this circle of ideas is completed. We dedicate this paper to Walter Philipp, whom we warmly remember for conversations about these and related matters.

## References

[1] J. Aaronson. *An Introduction to Infinite Ergodic Theory*. *Mathematical Surveys and Monographs* **50**. Amer. Math. Soc., Providence, RI, 1997. MR1450400
[2] W. Ambrose and S. Kakutani. Structure and continuity of measurable flows. *Duke Math. J.* **9** (1942) 25–42. MR0005800
[3] I. Berkes and H. Dehling. Some limit theorems in log density. *Ann. Probab.* **21** (1993) 1640–1670. MR1235433
[4] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968. MR0233396
[5] N. H. Bingham, C. M. Goldie and J. L. Teugels. *Regular Variation*. Cambridge Univ. Press, Cambridge/New York, 1987. MR0898871
[6] L. Breiman. On the tail behavior of sums of independent random variables. *Z. Wahrsch. Verw. Gebiete* **9** (1967) 20–25. MR0226707
[7] G. Brosamler. An almost everywhere Central Limit Theorem. *Math. Proc. Cambridge Philos. Soc.* **104** (1988) 561–574. MR0957261
[8] M. Csörgo and P. Révész. *Strong Approximations in Probability and Statistics*. Academic Press, New York, 1981. MR0666546
[9] W. Feller. *Probability Theory*, Vol. II. Wiley, New York/Chichester/Brisbane/Toronto, 1966. MR0210154
[10] W. Feller. *Probability Theory*, Vol. II, 2nd edition. Wiley, New York/Chichester/Brisbane/Toronto, 1971. MR0270403
[11] A. M. Fisher. A Pathwise Central Limit Theorem for random walks. Preprint, Univ. Goettingen, 1989. *Ann. Probab.* To appear.
[12] A. M. Fisher. Convex-invariant means and a pathwise central limit theorem. *Adv. in Math.* **63** (1987) 213–246. MR0877784
[13] A. M. Fisher, A. Lopes and M. Talet. Self-similar returns in the transition from finite to infinite measure. Unpublished manuscript, 2010.
[14] A. M. Fisher and M. Talet. The self-similar dynamics of renewal processes. *Electron. J. Probab.* **16** (2011) 929–961. MR2801456
[15] S. Fomin. Finite invariant measures in the flows (Russian). *Rec. Math. (Mat. Sbornik) N.S.* **12** (1943) 99–108. MR0009097
[16] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton Univ. Press, Princeton, 1981. MR0603625
[17] E. Glasner. *Ergodic Theory via Joinings*. *Mathematicall Surveys and Monographs* **101**. Amer. Math. Soc., Providence, RI, 2003. MR1958753
[18] P. Hall. A comedy of errors: The canonical form for a stable characteristic function. *Bull. London Math. Soc.* **13** (1981) 23–27. MR0599635
[19] V. V. Kalashnikov. A complete metric in the space $D[0,\infty)$. *J. Math. Sci.* **47** (1989) 2725–2730. MR1040101

[20] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, Berlin/Heidelberg/New York, 1988. MR0917065
[21] J. Komlos, P. Major and G. Tusnady. An approximation of the partial sums of independent RV's, and the sample DF. I. *Z. Wahrsch. Verw. Gebiete* **32** (1975) 111–131. MR0375412
[22] J. Komlos, P. Major and G. Tusnady. An approximation of the partial sums of independent RV's, and the sample DF. II. *Z. Wahrsch. Verw. Gebiete* **34** (1976) 33–58. MR0402883
[23] M. Lacey and W. Philipp. A note on the almost sure central limit theorem. *Statist. Probab. Lett.* **9** (1990) 201–205. MR1045184
[24] P. Lévy. Sur le développement en fraction continue d'un nombre choisi au hasard. *Compositio Math.* **3** (1936) 286–303. MR1556945
[25] P. Major. Approximation of partial sums of i.i.d. r.v.s when the summands have only two moments. *Z. Wahrsch. Verw. Gebiete* **35** (1976) 221–229. MR0415744
[26] P. Major. The approximation of partial sums of independent r.v.s. *Z. Wahrsch. Verw. Gebiete* **35** (1976) 213–220. MR0415743
[27] P. Major. Almost sure functional limit theorems. Part I. The general case. *Studia Sci. Math. Hungar.* **34** (1998) 273–304. MR1645214
[28] P. Major. Almost sure functional limit theorems. Part II. The case of independent random variables. *Studia Sci. Math. Hungar.* **36** (2000) 231–273. MR1768228
[29] D. Ornstein. *Ergodic Theory, Randomness and Dynamical Systems*. *Yale Mathematical Monographs* **5**. Yale Univ. Press, New Haven, 1973. MR0447525
[30] V. A. Rohlin. On the fundamental ideas of measure theory. *Mat. Sb.* **25** (1949) 107–150. MR0030584
[31] G. Samorodnitsky and M. Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman & Hall, New York, 1994. MR1280932
[32] P. Schatte. On strong versions of the central limit theorem. *Math. Nachr.* **137** (1988) 249–256. MR0968997
[33] P. Shields. *The Theory of Bernoulli Shifts*. Univ. Chicago Press, Chicago, 1973. MR0442198
[34] C. Stone. Weak convergence of stochastic processes defined on semi-infinite time intervals. *Proc. Amer. Math. Soc.* **14** (1963) 694–696. MR0153046
[35] V. Strassen. An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. Verw. Gebiete* **3** (1964) 211–226. MR0175194
[36] V. Strassen. Almost-sure behavior of sums of independent random variables and martingales. In *Proc. 5th Berkeley Symp. Math. Stat. and Prob.*, Vol. 2 315–343. Univ. California Press, Berkeley, CA, 1965. MR0214118
[37] P. Walters. *An Introduction to Ergodic Theory*. Springer, New York/Berlin, 1982. MR0648108
[38] W. Whitt. Some useful functions for functional limit theorems. *Math. Oper. Res.* **5** (1980) 67–85. MR0561155