sciendo

uniform distribution theory

DOI: 10.2478/UDT-2022-0013
Unif. Distrib. Theory **17** (2022), no.2, 77–100

A NOTE ON THE DISTRIBUTIONS OF $(\log n) \bmod 1$

Arno Berger

University of Alberta, Edmonton, CANADA

To the memory of Reinhard Winkler – humanist and friend

ABSTRACT. For sequences sufficiently close to $(a \log n)$, with an arbitrary real constant a, this note describes the precise asymptotics of the associated empirical distributions modulo one, with respect to the Kantorovich metric as well as a discrepancy-style metric. In particular, the note demonstrates how these asymptotics depend on a in a delicate, discontinuous way. The results strengthen and complement known facts in the literature.

Communicated by Friedrich Pillichshammer

1. Introduction

Consider an increasing sequence (x_n) in $\mathbb{R}$, with $x_n \to \infty$ and $x_{n+1} - x_n \downarrow 0$ as $n \to \infty$. It is well known [7,15] that if $\liminf_{n\to\infty} n(x_{n+1} - x_n) < \infty$, then (x_n) does not have a (unique) distribution modulo one (mod 1, for short). Arguably the simplest, most natural example of such a sequence is $(x_n) = (a \log n)$ with a real constant $a > 0$, where $\lim_{n\to\infty} n(x_{n+1} - x_n) = a$. The importance of $(\log n)$ for uniform distribution theory goes well beyond that of an anecdotal example, though. For instance, a classical theorem by Niederreiter [17] states that if (x_n) is uniformly distributed (u.d., for short) mod 1, then $x_n / \log n \to \infty$. By contrast, $\big(\log n \,(\log \log n)^{\varepsilon}\big)$, for example, is u.d. mod 1 for every $\varepsilon > 0$. Thus $(\log n)$ marks a threshold for increasing sequences, separating those that are u.d. mod 1

2020 Mathematics Subject Classification: 11K06, 11K31, 60B10, 60F05.
Keywords: Slowly changing sequence, distribution modulo one, Kantorovich metric, discrepancy.

from those that are not. For this and many other reasons, sequence that are, in one way or another, close to $(a \log n)$ are classical objects of study [3, 7, 14, 15, 17, 24, 25].

The fact that sequences (x_n) close to $(a \log n)$ do not have a (unique) distribution mod 1 naturally raises further, quantitative questions. On the one hand, (x_n) may have a (w_n)-weighted distribution [12, 16], with appropriate (positive) weights w_n. For instance, $(a \log n)$ is u.d. with weights (n^{-1}) and $(n^{-1} \log n)$; see, e.g., [15, sec.3]. On the other hand, in the space of distributions mod 1, that is, among probability measures on $\mathbb{R}/\mathbb{Z}$, the empirical averages along (x_n) have a continuum of accumulation points. It is natural to analyze the rate of convergence to, as well as the precise structure of these accumulation points [15, 23]. The present note provides such an analysis.

Traditionally, the distributional behaviour mod 1 of sequences like $(a \log n)$ has been studied mainly in terms of (asymptotic) distribution functions [12, 16]. Though quite elementary, this approach can lead to "remarkable results" [16, p.57]. It will, however, also lead to unavoidable artifacts. For instance, the set of distribution functions of $(x_n) = (\log \log n)$, where $\lim_{n \to \infty} n(x_{n+1} - x_n) = 0$, contains the constant function c for *every* $0 \le c \le 1$. When understood as distributions on $\mathbb{R}/\mathbb{Z}$, clearly all these functions are but different representations of one and the same object, namely the Dirac distribution (or unit point mass) concentrated on $0 + \mathbb{Z}$; see, e.g., [22] for similar artifacts.

Motivated by the classical treatise [15], this note avoids the potentially artifactual asymptotic distribution functions mod 1. Instead, it studies the distributions mod 1 of (x_n) directly as (Borel) probability measures on $\mathbb{R}/\mathbb{Z}$. Though very direct, this approach relies heavily on particular metrics (on the set of all probability measures on $\mathbb{R}/\mathbb{Z}$), the precise definition of which may in turn require a fair amount of analytic preliminaries. Specifically, this note employs the Kantorovich (or 1-Wasserstein) metric $d_{\mathbb{T}}$, that is, the minimal L^1-norm of the difference between (appropriately shifted) distribution functions, as well as the discrepancy-style metric d_∞, that is, the L^∞-norm of the difference between distribution functions. For the convenience of the reader, the definitions and all pertinent properties of these two metrics are reviewed in detail in Section 2. For the purpose of this introduction, it suffices to notice that $d_{\mathbb{T}}$ and d_∞ both are bona fide metrics with many endearing features. On the one hand, $d_{\mathbb{T}}$ may be viewed as a minimal "transport cost" from one distribution to the other, via the Kantorovich–Rubinstein theorem [8, Sec. 11.8]. Also, $d_{\mathbb{T}}$ represents the largest difference between integrals (w.r.t. one probability measure or the other) incurred over all real-valued 1-Lipschitz functions on $\mathbb{R}/\mathbb{Z}$. On the other hand, d_∞ is a natural, immediate generalization of star discrepancy, a familiar concept used widely throughout uniform distribution theory [3, 16, 25].

To state but one simple, illustrative special case of the main results, Theorems 3.1 and 4.3 below, for every integer $N \geq 1$ denote by μ_N the discrete uniform distribution on the N points $\frac{1}{3}\log 1, \frac{1}{3}\log 2, \ldots, \frac{1}{3}\log N \bmod 1$, and by μ the distribution on $\mathbb{R}/\mathbb{Z}$ with distribution function $F_\mu(s) = (e^{3s}-1)/(e^3-1)$, an exponential distribution "wrapped up mod 1"; see Section 2 for precise definitions and notations. Now, the sequence (μ_N) of distributions is divergent, as alluded to earlier, but has many convergent subsequences. For instance, it is well known [15,16] that $\mu_{N_j} \to \mu$ for an increasing sequence (N_j) of positive integers if and only if $\mathrm{dist}(\frac{1}{3}\log N_j, \mathbb{Z}) \to 0$ as $j \to \infty$; see also Corollary 3.2 below. Theorem 3.1 greatly refines this by providing a precise rate of convergence: If $\sup_j N_j \mathrm{dist}(\frac{1}{3}\log N_j, \mathbb{Z}) < \infty$, as is the case, e.g., for $(N_j) = (\lfloor e^{3j} \rfloor)$, then

$$\lim_{j\to\infty} \frac{N_j}{\sqrt{\log N_j}} d_{\mathbb{T}}(\mu_{N_j}, \mu) = \frac{1}{3\sqrt{2\pi}} . \tag{1.1}$$

Furthermore, if, for instance, $(N_j) = (\lfloor 2e^{3j} \rfloor)$ then (1.1) remains valid, provided that μ is replaced by the appropriate "rotated" copy of μ; see Figure 1.

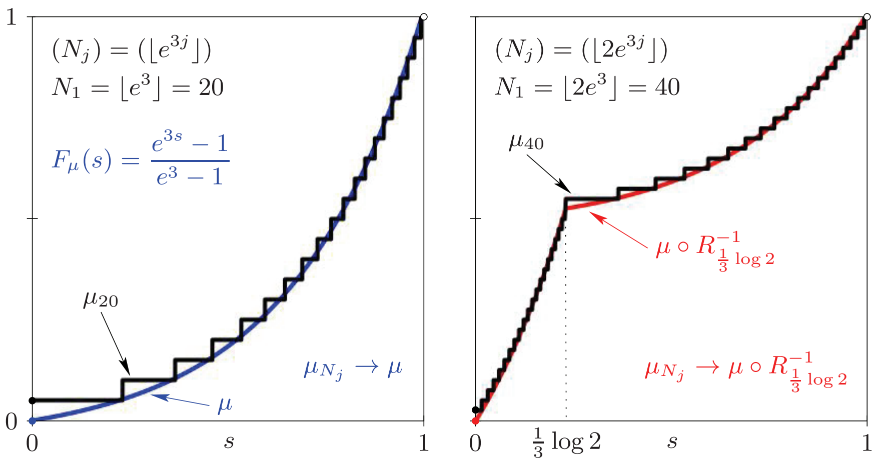

FIGURE 1. Illustrating (1.1) and Theorem 3.1: With μ_N denoting the discrete uniform distribution on $\frac{1}{3}\log 1, \frac{1}{3}\log 2, \ldots, \frac{1}{3}\log N \bmod 1$, for $(N_j) = (\lfloor e^{3j} \rfloor)$ the sequence (μ_{N_j}) converges to an exponential distribution μ "wrapped up mod 1" as $j \to \infty$ (left; see also Figure 2 below); for $(N_j) = (\lfloor 2e^{3j} \rfloor)$, the sequence (μ_{N_j}) converges to μ "rotated" by $\frac{1}{3}\log 2$ (right). By Theorem 3.1, the $d_{\mathbb{T}}$-rate of convergence is precisely $(\sqrt{\log N}/N)$ in either case.

To help put relations like (1.1) in perspective, a similar analysis is carried out using d_∞. In contrast to(1.1), Theorem 4.3 asserts that

$$\limsup_{j\to\infty} \frac{N_j}{\log N_j} d_\infty(\mu_{N_j}, \mu) \le \frac{1}{3},$$

as well as

$$\liminf_{j\to\infty} \frac{N_j}{\sqrt{\log N_j}} d_\infty(\mu_{N_j}, \mu) = \infty;$$

see Figure 2 which may also help explain why, as far as the author has been able to ascertain, the precise asymptotics of $\big(d_\infty(\mu_{N_j}, \mu)\big)$ remains unknown.

In a nutshell, then, this note illustrates how usage of $d_{\mathbb{T}}$ may lead, with little effort, to conclusions that are more robust and conclusive, compared to d_∞.

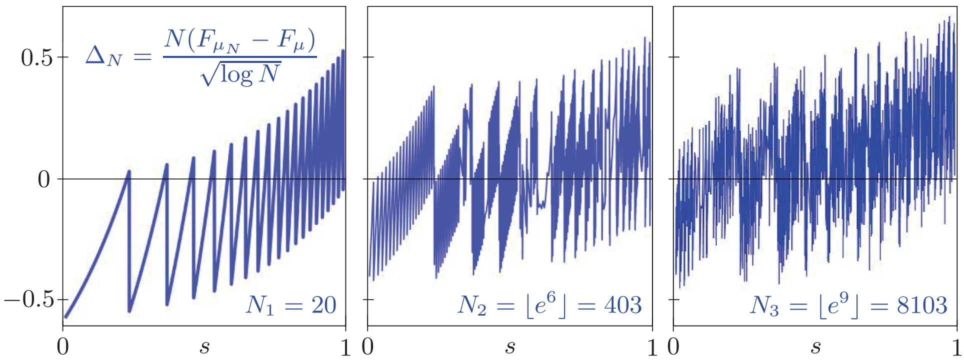

FIGURE 2. Why the asymptotics of $\big(d_\infty(\mu_N, \mu)\big)$ may be more delicate than the asymptotics of $\big(d_{\mathbb{T}}(\mu_N, \mu)\big)$: For $(N_j) = (\lfloor e^{3j} \rfloor)$ the (scaled) difference Δ_{N_j} between distribution functions oscillates wildly, and is in fact unbounded as $j \to \infty$, by Theorem 4.3. By contrast, Theorem 3.1 implies that $\lim_{j\to\infty} \min_{t\in\mathbb{R}} \int_0^1 |\Delta_{N_j}(s) - t|\, ds = 1/(3\sqrt{2\pi}) \approx 0.1329$.

2. Analytic preliminaries

Throughout, the sets of positive integers, integers, rational, positive real, real, and complex numbers are denoted by $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}^+$, $\mathbb{R}$, and $\mathbb{C}$, respectively. For every $a \in \mathbb{R}$, let $\lfloor a \rfloor$ and $\lceil a \rceil$ be the largest integer $\le a$ and the smallest integer $\ge a$, respectively; also, denote by $\langle a \rangle = a - \lfloor a \rfloor$ the fractional part of a. Recall that $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is a compact metrizable space when endowed with the usual (quotient) topology. For all $a, b \in \mathbb{R}$, let $d_{\mathbb{T}}(a,b) = \min_{k\in\mathbb{Z}} |a-b+k|$; when restricted to $\mathbb{T}\times\mathbb{T}$, this semi-metric yields a metric that induces the topology of $\mathbb{T}$. One may interpret $\mathbb{T}$ as the unit circle in $\mathbb{C}$, via the bi-Lipschitz bijection $\iota_{\mathbb{C}} : \mathbb{T} \to \{z \in \mathbb{C} : |z| = 1\}$ given by $\iota_{\mathbb{C}}(x) = e^{2\pi i x}$, for which

$$4d_{\mathbb{T}}(x,y) \le |\iota_{\mathbb{C}}(x) - \iota_{\mathbb{C}}(y)| \le 2\pi d_{\mathbb{T}}(x,y) \qquad \forall x, y \in \mathbb{T}\,.$$

By this interpretation, $d_{\mathbb{T}}(x,y)$ is the (normalized) arclength of the shortest arc joining $\iota_{\mathbb{C}}(x)$ and $\iota_{\mathbb{C}}(y)$. For every $x \in \mathbb{T}$, let $Q(x) = -x$ and $R_t(x) = x + t$ for all $t \in \mathbb{R}$. Thus the maps $\iota_{\mathbb{C}} \circ Q \circ \iota_{\mathbb{C}}^{-1}$ and $\iota_{\mathbb{C}} \circ R_t \circ \iota_{\mathbb{C}}^{-1}$ simply are the reflection of the unit circle about the real axis, and its counter-clockwise rotation by $2\pi t$, respectively. Clearly, Q and R_t are isometries of $\mathbb{T}$, with $R_t^{-1} = R_{-t}$ and $R_t \circ Q = Q \circ R_{-t}$ for every t.

Denote by $\mathcal{P}$ the set of all (Borel) probability measures on $\mathbb{T}$, endowed with the topology of weak convergence. Recall that $\mathcal{P}$ is compact and metrizable, with $\mu_n \to \mu$ in $\mathcal{P}$ if and only if $\int_{\mathbb{T}} h \, \mathrm{d}\mu_n \to \int_{\mathbb{T}} h \, \mathrm{d}\mu$ for every $h \in C(\mathbb{T})$. For explicit calculations later on, associate with each $\mu \in \mathcal{P}$ its distribution function F_μ as follows: The bijection $\iota_{\mathbb{R}} : \mathbb{T} \to [0,1[$ given by $\iota_{\mathbb{R}}(x) = \langle x \rangle$ is bi-measurable. (While $\iota_{\mathbb{R}}^{-1}$ is continuous, clearly $\iota_{\mathbb{R}}$ is discontinuous at $x = 0$.) With this, let

$$F_\mu(s) = \mu\bigl(\iota_{\mathbb{R}}^{-1}([0,s])\bigr) \qquad \forall 0 \le s < 1\,,$$

so the function F_μ is non-decreasing and right-continuous, with $F_\mu(0) \ge 0$ and $\lim_{\varepsilon\downarrow 0} F_\mu(1-\varepsilon) = 1$. Every function with these properties is the distribution function of one and only one $\mu \in \mathcal{P}$. Examples of elements of $\mathcal{P}$ relevant for this note include the Dirac measure δ_x at $x \in \mathbb{T}$, for which $\delta_x(B) = 1$ or $= 0$ for every Borel set $B \subset \mathbb{T}$ depending on whether $x \in B$ or $x \notin B$, and the Lebesgue measure $\lambda_{\mathbb{T}}$ characterized by $F_{\lambda_{\mathbb{T}}}(s) \equiv s$.

A familiar metric inducing the topology of weak convergence on $\mathcal{P}$, henceforth also denoted $d_{\mathbb{T}}$ for simplicity and often referred to as *Kantorovich* (or 1-*Wasserstein*) *metric*, is given by

$$d_{\mathbb{T}}(\mu,\nu) = \sup_{h\in L(\mathbb{T})} \left(\int_{\mathbb{T}} h \, \mathrm{d}\mu - \int_{\mathbb{T}} h \, \mathrm{d}\nu \right) \qquad \forall \mu, \nu \in \mathcal{P}\,,$$

where $L(\mathbb{T}) = \bigl\{ h \in C(\mathbb{T}) : h(0) = 0, |h(x) - h(y)| \le d_{\mathbb{T}}(x,y) \, \forall x, y \in \mathbb{T} \bigr\}$. As is well-known [6, 8, 26],

$$d_{\mathbb{T}}(\mu,\nu) = \min_{t\in\mathbb{R}} \int_0^1 |F_\mu(s) - F_\nu(s) - t| \, \mathrm{d}s \qquad \forall \mu, \nu \in \mathcal{P}\,, \tag{2.1}$$

and hence $d_{\mathbb{T}}(\mu,\nu) = \int_0^1 |F_\mu(s) - F_\nu(s) - t| \, \mathrm{d}s$ precisely if $t \in \mathbb{R}$ is a median of $F_\mu - F_\nu$, i.e., if

$$\lambda\bigl(\bigl\{0 \le s < 1 : F_\mu(s) - F_\nu(s) \le t\bigr\}\bigr) \ge \tfrac{1}{2}$$

as well as

$$\lambda\big(\{0 \le s < 1 : F_\mu(s) - F_\nu(s) \ge t\}\big) \ge \tfrac{1}{2}\,,$$

where λ denotes Lebesgue measure (on $\mathbb{R}$).

Given any (Borel measurable) map $S : \mathbb{T} \to \mathbb{T}$ and $\mu \in \mathcal{P}$, recall that $\mu \circ S^{-1}$, sometimes referred to as the *push-forward* of μ under S, is the unique $\nu \in \mathcal{P}$ with $\nu(B) = \mu\big(S^{-1}(B)\big)$ for every Borel set $B \subset \mathbb{T}$. For example, $\delta_x \circ R_t^{-1} = \delta_{x+t}$ as well as $\lambda_\mathbb{T} \circ R_t^{-1} = \lambda_\mathbb{T}$ for all $x \in \mathbb{T}$, $t \in \mathbb{R}$. Note specifically that $d_\mathbb{T}(\mu \circ S^{-1}, \nu \circ S^{-1}) = d_\mathbb{T}(\mu, \nu)$ for all $\mu, \nu \in \mathcal{P}$ and every isometry S of $\mathbb{T}$; in particular, $d_\mathbb{T}$ is invariant under all reflections and rotations. Using (2.1), it is readily deduced that $d_\mathbb{T}(\mu, \lambda_\mathbb{T}) \le \frac{1}{4}$, and this inequality is strict unless $\mu = \delta_x$ for some $x \in \mathbb{T}$. As a consequence, $d_\mathbb{T}(\mu, \nu) \le \frac{1}{2}$ for all $\mu, \nu \in \mathcal{P}$, and this inequality is strict unless $\mu = \delta_x$, $\nu = \delta_{x+1/2}$ for some $x \in \mathbb{T}$. Note that $d_\mathbb{T}(\delta_x, \delta_y) = d_\mathbb{T}(x, y)$ for all $x, y \in \mathbb{T}$, so $\{\delta_x : x \in \mathbb{T}\}$ is an isometric copy of $\mathbb{T}$ inside $\mathcal{P}$.

A one-parameter family in $\mathcal{P}$ particularly relevant for this note is the following: For every $a \in \mathbb{R} \setminus \{0\}$ let $\eta_a \in \mathcal{P}$ be defined by its distribution function

$$F_{\eta_a}(s) = \frac{e^{s/a} - 1}{e^{1/a} - 1} \qquad \forall 0 \le s < 1\,.$$

In probabilistic parlance, η_a simply is the distribution of $\langle -aX \rangle$, where the random variable X is standard exponential. A straightforward calculation using (2.1) yields

$$d_\mathbb{T}(\eta_a, \lambda_\mathbb{T}) = |a| \log\cosh\frac{1}{4a}\,, \quad d_\mathbb{T}(\eta_a, \delta_0) = a \tanh\frac{1}{4a} \qquad \forall a \in \mathbb{R} \setminus \{0\}\,, \quad (2.2)$$

which in turn suggests setting $\eta_0 = \delta_0$ and $\eta_{-\infty} = \eta_\infty = \lambda_\mathbb{T}$. With this, the set $\{\eta_a : -\infty \le a \le \infty\}$ is a homeomorphic copy of $\mathbb{T}$ inside $\mathcal{P}$. Note that $\eta_a \circ Q^{-1} = \eta_{-a}$ for every $a \in \mathbb{R} \cup \{-\infty, \infty\}$. Rotated versions of η_a are going to play a prominent role later on. For convenience, for every $a \in \mathbb{R} \cup \{-\infty, \infty\}$ and $t \in \mathbb{R}$, write $\eta_a \circ R_t^{-1}$ simply as $\eta_{a,t}$. For $a \in \mathbb{R} \setminus \{0\}$, the distribution function of $\eta_{a,t}$ equals

$$\begin{aligned} F_{\eta_{a,t}}(s) &= e^{\min\{s - \langle t \rangle, 0\}/a} - e^{-\langle t \rangle/a}\frac{e^{1/a} - e^{s/a}}{e^{1/a} - 1} \qquad (2.3) \\ &= \begin{cases} e^{(1-\langle t \rangle)/a}\dfrac{e^{s/a} - 1}{e^{1/a} - 1} & \text{if} \quad 0 \le s < \langle t \rangle\,, \\ 1 - e^{-\langle t \rangle/a}\dfrac{e^{1/a} - e^{s/a}}{e^{1/a} - 1} & \text{if} \quad \langle t \rangle \le s < 1\,; \end{cases} \end{aligned}$$

it is in this form that the probability measures $\eta_{a,t}$ traditionally appear in the literature [12,16]. Notice that $\eta_{0,t} = \delta_t$ for every $t \in \mathbb{R}$, whereas $\eta_{-\infty,t} = \eta_{\infty,t} = \lambda_\mathbb{T}$.

It is straightforward to show that for every $a \in \mathbb{R} \setminus \{0\}$,

$$4|a| \log\cosh \frac{1}{4a}\, d_{\mathbb{T}}(s,t) \le d_{\mathbb{T}}(\eta_{a,s}, \eta_{a,t}) \le \tanh \frac{1}{4|a|}\, d_{\mathbb{T}}(s,t) \qquad \forall s,t \in \mathbb{R}\,; \quad (2.4)$$

when interpreted with caution, (2.4) remains valid for $a = 0$ and $|a| = \infty$ also.

Given any sequence (x_n) in $\mathbb{R}$, for every $N \in \mathbb{N}$ denote by $\mu_N[x_n] \in \mathcal{P}$ the empirical distribution mod 1 of $\{x_1, \ldots, x_N\}$, that is, let

$$\mu_N[x_n] = \frac{1}{N} \sum\nolimits_{n=1}^{N} \delta_{x_n} \qquad \forall N \in \mathbb{N}\,.$$

Thus, $\mu_N[x_n]$ is the unique $\mu \in \mathcal{P}$ with $\int_{\mathbb{T}} h \,\mathrm{d}\mu = \frac{1}{N}\sum_{n=1}^{N} h(x_n)$ for every $h \in C(\mathbb{T})$. Also, denote by $\mathcal{M}[x_n]$ the set of all accumulation points of the sequence (μ_N) in $\mathcal{P}$, i.e.,

$$\mathcal{M}[x_n] = \bigl\{\mu \in \mathcal{P} : \exists\, N_1 < N_2 < \cdots \text{ s.t. } \mu_{N_j}[x_n] \to \mu \text{ as } j \to \infty\bigr\}\,.$$

The set $\mathcal{M}[x_n] \subset \mathcal{P}$ is non-empty, compact and connected [12, 23, 25]. Conversely, every non-empty, compact and connected subset of $\mathcal{P}$ equals $\mathcal{M}[x_n]$ for some sequence (x_n). This note focuses on sequences (x_n) for which $\mathcal{M}[x_n]$ is very small. An extreme situation obviously occurs when $\mathcal{M}[x_n] = \{\mu\}$ for some $\mu \in \mathcal{P}$. In this situation, (x_n) is said to *have distribution* μ mod 1, with the most classical special case being $\mathcal{M}[x_n] = \{\lambda_{\mathbb{T}}\}$, or equivalently, (x_n) being *u.d.* mod 1 [15, 16]. For the slowly changing sequences precisely defined in the next section, it is well known that $\mathcal{M}[x_n]$ is *not* a singleton. Still, $\mathcal{M}[x_n]$ has a very simple structure that can be described rather transparently. Sections 3 and 4 below provide such a description together with sharp rates of convergence.

Although the metric $d_{\mathbb{T}}$ on $\mathcal{P}$ (and $\mathbb{T}$) is the main tool used here, in order to relate the results to classical concepts and studies, the metric

$$d_{\infty}(\mu,\nu) := \sup\nolimits_{0 \le s < 1} |F_\mu(s) - F_\nu(s)| \qquad \forall \mu,\nu \in \mathcal{P}$$

is considered as well. Note that by (2.1),

$$d_{\mathbb{T}}(\mu,\nu) \le \int_0^1 |F_\mu(s) - F_\nu(s)|\,\mathrm{d}s \le d_\infty(\mu,\nu) \qquad \forall \mu,\nu \in \mathcal{P}\,,$$

and d_∞ induces a strictly finer topology on $\mathcal{P}$ than does $d_{\mathbb{T}}$. Also, d_∞ is not invariant under Q or any $R_t \ne \mathrm{id}_{\mathbb{T}}$, unlike $d_{\mathbb{T}}$. Recall, however, that if $\mu_n \to \mu$ and μ is continuous, i.e., if $\mu(\{x\}) = 0$ for every $x \in \mathbb{T}$, then $d_\infty(\mu_n,\mu) \to 0$ also. In this situation, therefore, $d_{\mathbb{T}}(\mu_n,\mu) \to 0$ if and only if $d_\infty(\mu_n,\mu) \to 0$. A classical case in point is $\mu = \lambda_{\mathbb{T}}$, where $d_\infty(\mu_N[x_n], \lambda_{\mathbb{T}})$ is traditionally referred to as *star discrepancy*, often denoted $D_N^*(x_n)$; see, e.g., [16]. Even in this familiar context, however, the rates of convergence for $\bigl(d_{\mathbb{T}}(\mu_N,\mu)\bigr)$ and $\bigl(d_\infty(\mu_N,\mu)\bigr)$ may differ considerably, as the reader is going to see shortly.

3. A Kantorovich rate of convergence

Throughout this note, let $f : \mathbb{R}^+ \to \mathbb{R}$ be a C^1-function with the property that
$$\int_1^\infty |tf'(t) - a|\, \mathrm{d}t < \infty \tag{3.1}$$
for some (necessarily unique) $a \in \mathbb{R}$. In a way, therefore, $f(t)$ differs but little from $a \log t$. For any f satisfying (3.1), consider the sequence $(x_n) = \big(f(n)\big)$. For convenience, henceforth write $\mu_N[f(n)]$ and $\mathcal{M}[f(n)]$ as μ_N^f and $\mathcal{M}^f$, respectively; in other words,
$$\mu_N^f = \frac{1}{N} \sum\nolimits_{n=1}^N \delta_{f(n)} \in \mathcal{P} \qquad \forall N \in \mathbb{N}\,,$$
$$\mathcal{M}^f = \big\{\mu \in \mathcal{P} : \exists\, N_1 < N_2 < \cdots \text{ s.t. } \mu_{N_j}^f \to \mu \text{ as } j \to \infty\big\} \subset \mathcal{P}\,.$$
Unless $a = 0$, clearly $|f(n)| \to \infty$ and $\lim_{n\to\infty} n\big(f(n+1) - f(n)\big) = a$. By [15, Thm.3], the set $\mathcal{M}^f$ is not a singleton. The main result of this section, Theorem 3.1 below, elucidates why $\mathcal{M}^f$ instead is a homeomorphic (in fact, bi-Lipschitz equivalent) copy of $\mathbb{T}$ inside $\mathcal{P}$; see Corollary 3.2. In preparation for the precise statement of the result, denote by $T : \mathbb{R}^+ \to \mathbb{Q}$ the classical Thomae function,
$$T(t) = \begin{cases} \dfrac{1}{q} & \text{if} \quad t = p/q \text{ with } p, q \in \mathbb{N} \text{ and } \gcd(p,q) = 1\,, \\ 0 & \text{if} \quad t \in \mathbb{R}^+ \setminus \mathbb{Q}\,, \end{cases}$$
and for every $t > 1$, let
$$V(t) = 1 + 2 \max_{n\in\mathbb{N}} \frac{T(t^n)^2}{t^n - T(t^n)^2}\,.$$
More explicitly, for every $t > 1$ and with $\{t^n : n \in \mathbb{N}\} \cap \mathbb{Q}$ abbreviated as B_t,
$$V(t) = \begin{cases} 1 & \text{if } B_t = \varnothing\,, \\ \dfrac{pq+1}{pq-1} & \text{if } B_t = \{(p/q)^j : j \in \mathbb{N}\} \text{ with } p, q \in \mathbb{N} \text{ and } \gcd(p,q) = 1. \end{cases}$$
Plainly, $1 \le V(t) \le 3$. Moreover, $V(t) = 1$ for all but countably many $t > 1$, and the function V is continuous at t if and only if $V(t) = 1$. Let $V(\infty) := \lim_{t\to\infty} V(t) = 1$. Utilizing the function V allows for a neat description of the asymptotics of (μ_N^f). The following theorem greatly strengthens [25, Thm.3.1], with the convention $e^{1/0} := \infty$ applied in case $a = 0$. To appreciate the statement of the result, recall that $\eta_{a,t}$, with $a, t \in \mathbb{R}$, denotes an exponential distribution on $\mathbb{T}$, "wrapped up mod 1 and rotated" as per (2.3).

Theorem 3.1. *Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a C^1-function. If f satisfies* (3.1) *for some $a \in \mathbb{R}$ then*

$$\lim_{N\to\infty} \frac{N}{\sqrt{\log N}}\, d_{\mathbb{T}}\big(\mu_N^f, \eta_{a,f(N)}\big) = \sqrt{\frac{|a|}{6\pi} V(e^{1/|a|})}\,. \tag{3.2}$$

Remark. As pointed out by one referee, the appearance of the factor $N/\sqrt{\log N}$ in (3.2) is reminiscent of recent studies on certain u.d. mod 1 sequences [4], [13, 21].

Before turning towards the proof of Theorem 3.1, two immediate corollaries are worth mentioning. On the one hand, (3.2) contains a simple description of $\mathcal{M}^f$ as the set of all rotated copies of a single element of $\mathcal{P}$.

Corollary 3.2. *Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a C^1-function. If f satisfies* (3.1) *for some $a \in \mathbb{R} \setminus \{0\}$, then $\mathcal{M}^f = \{\eta_{a,t} : t \in \mathbb{R}\}$. Moreover, $\mu_{N_j}^f \to \eta_{a,t}$ as $j \to \infty$ for some increasing sequence (N_j) of integers and some $t \in \mathbb{R}$ if and only if $d_{\mathbb{T}}(f(N_j), t) \to 0$.*

As a matter of fact, both conclusions of Corollary 3.2 are valid even if (3.1) fails, provided that $\lim_{t\to\infty} tf'(t) = a$; see [2]. For the (forbidden) case $a = 0$, the second conclusion of Corollary 3.2 is valid as well, but the first is not.

On the other hand, Theorem 3.1 provides a quantitative form of the familiar fact that sequences $\big(f(n)\big)$ with f satisfying (3.1), and indeed also if merely $\lim_{t\to\infty} tf'(t) = a$, do not have a distribution mod 1, i.e., (μ_N^f) is divergent.

Corollary 3.3. *Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a C^1-function. If f satisfies 3.1 for some $a \in \mathbb{R}$ then*

$$\lim_{N\to\infty} \Big| d_{\mathbb{T}}\big(\mu_N^f, \mu\big) - d_{\mathbb{T}}\big(\eta_{a,f(N)}, \mu\big)\Big| = 0 \qquad \forall \mu \in \mathcal{P}\,.$$

Note that for every $\mu \in \mathcal{P}$ the function $t \mapsto d_{\mathbb{T}}(\eta_{a,t}, \mu)$ is continuous, 1-periodic, and attains a positive maximal value; for $\mu = \lambda_{\mathbb{T}}$ this function is constant, its (positive) value being displayed in (2.2).

Theorem 3.1 hinges on a central limit theorem for Riesz-Raikov sums, of which a simple special case, tailor-made for subsequent use, is stated here for the reader's convenience; further background and details on this classical subject can be found, e.g., in [10, 11]. That a central limit theorem is key to the proof of Theorem 3.1 may seem surprising, but is consistent with a well-established paradigm: Even completely deterministic systems may exhibit statistics that are best understood in terms of appropriately scaled random walks [1, 5].

Proposition 3.4. *For every $\beta \in \mathbb{R}^+$ with $\beta > 1$,*

$$\lim_{M\to\infty} \frac{1}{M} \int_0^1 \left(\sum\nolimits_{j=1}^M \left(\langle \beta^j s\rangle - \frac{1}{2}\right)\right)^2 \mathrm{d}s = \frac{V(\beta)}{12}\,. \tag{3.3}$$

Moreover, for every Borel probability measure ν on $\mathbb{R}$ that is absolutely continuous w.r.t. λ,

$$\lim_{M\to\infty} \nu\left(\left\{t \in \mathbb{R} : \frac{1}{\sqrt{M}} \sum\nolimits_{j=1}^{M} \left(\langle \beta^j t\rangle - \frac{1}{2}\right) \le b\right\}\right) = \Phi\left(2b\sqrt{\frac{3}{V(\beta)}}\right) \quad (3.4)$$

for every $b \in \mathbb{R}$, where Φ denotes the standard normal distribution function.

P r o o f o f T h e o r e m 3.1. Since $\mu_N^f \circ Q^{-1} = \mu_N^{-f}$ and $\eta_a \circ Q^{-1} = \eta_{-a}$, it can be assumed that $a \ge 0$. For the time being, assume that in fact $a > 0$, and fix any $h \in L(\mathbb{T})$. It will first be shown that it suffices to establish (3.2) solely for the case of f being a constant multiple of log. To this end, for every $t \in \mathbb{R}^+$ let $\Delta(t) = tf'(t) - a$,

$$\widetilde{f}(t) = a \log t + f(1) + \int_1^\infty \frac{\Delta(u)}{u}\,\mathrm{d}u \qquad \forall t \in \mathbb{R}^+ ,$$

and correspondingly $\widetilde{\mu_N} = \mu_N[\widetilde{f}(n)]$. Observe that

$$\begin{aligned}\left| N\left(\int_{\mathbb{T}} h\,\mathrm{d}\mu_N^f - \int_{\mathbb{T}} h\,\mathrm{d}\widetilde{\mu_N}\right)\right| &= \left|\sum\nolimits_{n=1}^{N} \big(h\circ f(n) - h \circ \widetilde{f}(n)\big)\right| \\ &\le \sum\nolimits_{n=1}^{N} |f(n) - \widetilde{f}(n)| = \sum\nolimits_{n=1}^{N} \left| -\int_n^\infty \frac{\Delta(u)}{u}\,\mathrm{d}u\right| \\ &\le \int_1^{N+1}\int_{\max\{v-1,1\}}^\infty \frac{|\Delta(u)|}{u}\,\mathrm{d}u\,\mathrm{d}v \le \int_1^\infty |\Delta(u)|\,\mathrm{d}u ,\end{aligned}$$

so with the definition of $d_{\mathbb{T}}$ and by virtue of (3.1),

$$N d_{\mathbb{T}}\big(\mu_N^f, \widetilde{\mu_N}\big) \le \int_1^\infty |\Delta(u)|\,\mathrm{d}u \qquad \forall N \in \mathbb{N}. \quad (3.5)$$

Similarly, with (2.4) and since every rotation R_t is an isometry of $\mathbb{T}$,

$$\begin{aligned} N d_{\mathbb{T}}\big(\eta_{a,f(N)}, \eta_{a,\widetilde{f}(N)}\big) &\le N d_{\mathbb{T}}\big(f(N), \widetilde{f}(N)\big) \\ &\le N\int_N^\infty \frac{|\Delta(u)|}{u}\,\mathrm{d}u \le \int_N^\infty |\Delta(u)|\,\mathrm{d}u . \end{aligned} \quad (3.6)$$

From (3.5) and (3.6) it is clear that

$$\lim_{N\to\infty} \frac{N}{\sqrt{\log N}} \left| d_{\mathbb{T}}\big(\mu_N^f, \eta_{a,f(N)}\big) - d_{\mathbb{T}}\big(\widetilde{\mu_N}, \eta_{a,\widetilde{f}(N)}\big)\right| = 0 .$$

Thus, it suffices to establish (3.2) for $f(t) = a\log t$, and this will now be done. For convenience, let $L_N = \max\{1, \lfloor a \log N\rfloor\}$ and $s_N = \langle a \log N\rangle$ for every $N \in \mathbb{N}$.

By a direct calculation using Euler summation, for every $N \ge e^{1/a}$,

$$
\begin{aligned}
N\int_{\mathbb{T}} h \,\mathrm{d}\mu_N^f = {} & \frac{h(0)+h(a\log N)}{2} + \int_1^N h(a\log u)\,\mathrm{d}u \\
& + \int_0^N h'(a\log u)\frac{a}{u}\left(\langle u\rangle - \frac{1}{2}\right)\,\mathrm{d}u \\
= {} & C_{1,N} + \frac{1}{a}\int_0^{a\log N} h(v)e^{v/a}\mathrm{d}v + \int_0^{a\log N} h'(v)\left(\langle e^{v/a}\rangle - \frac{1}{2}\right)\mathrm{d}v \\
= {} & C_{2,N} + \frac{1}{a}\int_0^1 h(v)\sum\nolimits_{j=1}^{L_N} e^{(v+j-1)/a}\mathrm{d}v + \frac{1}{a}\int_0^{s_N} h(v)e^{(v+L_N)/a}\mathrm{d}v \\
& + \int_0^1 h'(v)\sum\nolimits_{j=1}^{L_N}\left(\langle e^{(v+j-1)/a}\rangle - \frac{1}{2}\right)\mathrm{d}v \\
= {} & C_{3,N} + \frac{N}{a(e^{1/a}-1)}\int_0^{s_N} h(v)e^{(v+1-s_N)/a}\mathrm{d}v \\
& + \frac{N}{a(e^{1/a}-1)}\int_{s_N}^1 h(v)e^{(v-s_N)/a}\mathrm{d}v \\
& + \int_0^1 h'(v)\sum\nolimits_{j=1}^{L_N}\left(\langle e^{(v+j-1)/a}\rangle - \frac{1}{2}\right)\mathrm{d}v \\
= {} & C_{3,N} + N\int_{\mathbb{T}} h\,\mathrm{d}\eta_{a,s_N} + \int_0^1 h'(v)\sum\nolimits_{j=1}^{L_N}\left(\langle e^{(v+j-1)/a}\rangle - \frac{1}{2}\right)\mathrm{d}v\,;
\end{aligned}
$$

here the numbers $C_{1,N}, C_{2,N}, C_{3,N}$ depend on h and a but are uniformly bounded in that

$$
|C_{1,N}| + |C_{2,N}| + |C_{3,N}| \le 3 \qquad \forall N \ge e^{1/a}\,.
$$

In other words, for every $N \ge e^{1/a}$,

$$
\begin{aligned}
& \frac{N}{\sqrt{L_N}}\left(\int_{\mathbb{T}} h\ \mathrm{d}\mu_N^f - \int_{\mathbb{T}} h\,\mathrm{d}\eta_{a,f(N)}\right) = \\
& \qquad \frac{C_{3,N}}{\sqrt{L_N}} + \int_0^1 h'(v)\frac{1}{\sqrt{L_N}}\sum\nolimits_{j=1}^{L_N}\left(\langle e^{(v+j-1)/a}\rangle - \frac{1}{2}\right)\mathrm{d}v\,. \qquad (3.7)
\end{aligned}
$$

To make the remainder of the argument easy to grasp, let $\beta = e^{1/a} > 1$, and $S(v) = e^{(v-1)/a}$ for every $v \in \mathbb{R}$. The Borel probability measure $\nu = \lambda|_{[0,1]}\circ S^{-1}$ is absolutely continuous w.r.t. λ, its density being equal to a/v on $[\beta^{-1}, 1] = S[0,1]$, and vanishing outside this interval. Also, for every $N \in \mathbb{N}$ let

$$
g_N(v) = \frac{1}{\sqrt{L_N}}\sum\nolimits_{j=1}^{L_N}\left(\langle \beta^j v\rangle - \frac{1}{2}\right) \qquad \forall v \in \mathbb{R}\,.
$$

Utilizing these abbreviations, deduce from (3.7) that for every $N \ge \beta$,

$$\begin{aligned}\frac{N}{\sqrt{L_N}}\, d_{\mathbb{T}}\big(\mu_N^f, \eta_{a,f(N)}\big) &\le \frac{3}{\sqrt{L_N}} + \int_0^1 |g_N \circ S(v)|\, \mathrm{d}v \\ &= \frac{3}{\sqrt{L_N}} + \int_{[\beta^{-1},1]} |g_N(w)|\, \mathrm{d}\nu(w)\,. \qquad (3.8)\end{aligned}$$

Now, fix $\varepsilon > 0$, and with $D_\varepsilon = a\beta V(\beta)/(12\varepsilon)$ let $A_{N,\varepsilon} = \big\{w \in [\beta^{-1}, 1] : |g_N(w)| < D_\varepsilon\big\}$. For the integral in (3.8), observe that

$$\begin{aligned}&\int_{[\beta^{-1},1]} |g_N(w)|\, \mathrm{d}\nu(w) \\ &= D_\varepsilon \int_{[\beta^{-1},1]\setminus A_{N,\varepsilon}} \frac{|g_n(w)|}{D_\varepsilon}\, \mathrm{d}\nu(w) + \int_{A_{N,\varepsilon}} \min\{|g_N(w)|, D_\varepsilon\}\, \mathrm{d}\nu(w) \\ &\le \frac{1}{D_\varepsilon} \int_{\beta^{-1}}^1 g_N(w)^2 \frac{a}{w}\, \mathrm{d}w + \int_{[\beta^{-1},1]} \min\{|g_N(w)|, D_\varepsilon\}\, \mathrm{d}\nu(w) \\ &\le \frac{a\beta}{D_\varepsilon} \int_0^1 g_N(w)^2\, \mathrm{d}w + \int_0^\infty \nu\left(\big\{w \in [\beta^{-1}, 1] : \min\{|g_N(w)|, D_\varepsilon\} \ge b\big\}\right) \mathrm{d}b \\ &= \frac{12\varepsilon}{V(\beta)} \int_0^1 g_N(w)^2\, \mathrm{d}w + \int_0^{D_\varepsilon} \nu\big(\{w \in \mathbb{R} : b \le |g_N(w)| \le D_\varepsilon\}\big) \mathrm{d}b\,.\end{aligned}$$

After plugging this bound back into the right-hand side of (3.8), an application of (3.3) and (3.4), together with Fatou's lemma, yields

$$\begin{aligned}&\limsup_{N\to\infty} \frac{N}{\sqrt{L_N}}\, d_{\mathbb{T}}\big(\mu_N^f, \eta_{a,f(N)}\big) \\ &\le \varepsilon + \int_0^{D_\varepsilon} \left(\Phi\left(2D_\varepsilon\sqrt{\frac{3}{V(\beta)}}\right) - \Phi\left(2b\sqrt{\frac{3}{V(\beta)}}\right)\right. \\ &\qquad\qquad \left. + \Phi\left(-2b\sqrt{\frac{3}{V(\beta)}}\right) \Phi\left(-2D_\varepsilon\sqrt{\frac{3}{V(\beta)}}\right)\right) \mathrm{d}b \\ &\le \varepsilon + \int_0^\infty \left(1 - \Phi\left(2b\sqrt{\frac{3}{V(\beta)}}\right) + \Phi\left(-2b\sqrt{\frac{3}{V(\beta)}}\right)\right) \mathrm{d}b \\ &= \varepsilon + \sqrt{\frac{V(\beta)}{3}} \int_0^\infty \big(1 - \Phi(t)\big)\, \mathrm{d}t \;=\; \varepsilon + \sqrt{\frac{V(\beta)}{6\pi}}\,.\end{aligned}$$

Since $\varepsilon > 0$ has been arbitrary,

$$\limsup_{N\to\infty} \frac{N}{\sqrt{L_N}}\, d_{\mathbb{T}}\big(\mu_N^f, \eta_{a,f(N)}\big) \le \sqrt{\frac{V(\beta)}{6\pi}}\,. \tag{3.9}$$

Note that $\lim_{N\to\infty} L_N/\log N = a$. Thus to complete the proof of (3.2) for $a > 0$, it only remains to establish (3.9) with $\liminf$ instead of $\limsup$, and with the inequality reversed. To this end, let $m_N \in \mathbb{R}^+$ be the unique number with

$$\lambda\left(\left\{v \in [0,1] : \sum\nolimits_{j=1}^{L_N} \langle \beta^j S(v)\rangle \le m_N\right\}\right) = \frac{1}{2}\,.$$

In probabilistic parlance, m_N is the median of $\sum_{j=1}^{L_N}\langle \beta^j S(X)\rangle$, where the random variable X is uniform on $[0,1]$. Since $\langle \beta^j S(X)\rangle$ is very close to being uniform for large j, the probabilistic view suggests that m_N should be close to $\frac{1}{2}L_N$. To see that indeed it is, notice that

$$\frac{1}{2} = \nu\left(\left\{w \in \mathbb{R} : \frac{1}{\sqrt{L_N}}\sum\nolimits_{j=1}^{L_N}\left(\langle \beta^j w\rangle - \frac{1}{2}\right) \le \frac{2m_N - L_N}{2\sqrt{L_N}}\right\}\right) \qquad \forall N \in \mathbb{N}\,.$$

Thus, if m_N were larger than $\frac{1}{2}(L_N + \varepsilon\sqrt{L_N})$ for $\varepsilon > 0$ and infinitely many N then, by (3.4),

$$\begin{aligned}\frac{1}{2} &\ge \liminf_{N\to\infty} \nu\left(\left\{w \in \mathbb{R} : \frac{1}{\sqrt{L_N}}\sum\nolimits_{j=1}^{L_N}\left(\langle \beta^j w\rangle - \frac{1}{2}\right) \le \frac{\varepsilon}{2}\right\}\right)\\ &= \Phi\left(\varepsilon\sqrt{\frac{3}{V(\beta)}}\right) > \frac{1}{2}\,,\end{aligned}$$

an obvious contradiction. It follows that

$$\limsup_{N\to\infty} \frac{2m_N - L_N}{\sqrt{L_N}} \le \varepsilon\,,$$

and similarly for the corresponding lower bound. Since $\varepsilon > 0$ has been arbitrary,

$$\lim_{N\to\infty} \frac{2m_N - L_N}{\sqrt{L_N}} = 0\,.$$

Returning now to (3.7), recall that $\int_0^1 h'(v)\,\mathrm{d}v = 0$, and hence the term $-\frac{1}{2}$ on the far right in that equation may be replaced by $-m_N/L_N$. Doing so yields

$$\begin{aligned}\frac{N}{\sqrt{L_N}}\, d_{\mathbb{T}}\big(\mu_N^f, \eta_{a,f(N)}\big) &\ge -\frac{3}{\sqrt{L_N}} + \int_0^1 \left|\frac{1}{\sqrt{L_N}}\sum\nolimits_{j=1}^{L_N}\langle \beta^j S(v)\rangle - m_N\right| \mathrm{d}v\\ &\ge -\frac{3}{\sqrt{L_N}} - \frac{|L_N - 2m_N|}{2\sqrt{L_N}} + \int_0^1 |g_N \circ S(v)|\,\mathrm{d}v\,.\end{aligned}$$

As in (3.8), observe that

$$\int_0^1 |g_N \circ S(v)| \, \mathrm{d}v \;=\; \int_{[\beta^{-1},1]} |g_N(w)| \, \mathrm{d}\nu(w) \;=\; \int_0^\infty \nu\bigl(\{w \in \mathbb{R} : |g_N(w)| \ge b\}\bigr) \, \mathrm{d}b \,,$$

and so, with (3.4) and Fatou's lemma,

$$\begin{aligned} \liminf_{N\to\infty} \int_0^1 |g_N \circ S(v)| \, \mathrm{d}v \;&\ge\; \int_0^\infty \nu\bigl(\{w \in \mathbb{R} : |g_N(w)| \ge b\}\bigr) \, \mathrm{d}b \\ &= \int_0^\infty \left(1 - \Phi\left(2b\sqrt{\frac{3}{V(\beta)}}\right) + \Phi\left(-2b\sqrt{\frac{3}{V(\beta)}}\right)\right) \mathrm{d}b = \sqrt{\frac{V(\beta)}{6\pi}} \,. \end{aligned}$$

In summary, therefore,

$$\liminf_{N\to\infty} \frac{N}{\sqrt{L_N}} \, d_{\mathbb{T}}\bigl(\mu_N^f, \eta_{a,f(N)}\bigr) \ge \sqrt{\frac{V(\beta)}{6\pi}} \,.$$

As noted earlier, this, together with (3.9), completes the proof for $a > 0$.

It remains to consider the case $a = 0$, for which (3.2) merely asserts that

$$\lim_{N\to\infty} \frac{N}{\sqrt{\log N}} \, d_{\mathbb{T}}\bigl(\mu_N^f, \delta_{f(N)}\bigr) = 0 \,. \tag{3.10}$$

Here, (3.1) reads $\int_1^\infty t|f'(t)| \, \mathrm{d}t < \infty$. It is easy to see that this in fact entails $\sup_{N\in\mathbb{N}} N d_{\mathbb{T}}\bigl(\mu_N^f, \delta_{f(N)}\bigr) < \infty$, which in turn implies (3.10): Indeed, by Euler summation,

$$\begin{aligned} N \int_{\mathbb{T}} h \, \mathrm{d}\mu_N^f \;&=\; \frac{h \circ f(1) + h \circ f(N)}{2} + \int_1^N h \circ f(u) \, \mathrm{d}u \\ &\quad + \int_1^N h' \circ f(u) f'(u) \left(\langle u \rangle - \frac{1}{2}\right) \mathrm{d}u \\ &=\; N h \circ f(N) + \frac{h \circ f(N) - h \circ f(1)}{2} \\ &\quad - \int_1^N h' \circ f(u) f'(u) \left(\lfloor u \rfloor + \frac{1}{2}\right) \mathrm{d}u \,, \end{aligned}$$

and hence for every $N \in \mathbb{N}$,

$$N d_{\mathbb{T}}\bigl(\mu_N^f, \delta_0 \circ R_{f(N)}^{-1}\bigr) \le 1 + \int_1^N \left(u + \frac{1}{2}\right) |f'(u)| \, \mathrm{d}u \le 1 + 2\pi \int_1^\infty u|f'(u)| \, \mathrm{d}u \,.$$

As seen at the outset, this completes the proof. □

Arguably the most prominent concrete example within the scope of Theorem 3.1 is $f(t) = a \log t$ with $a \in \mathbb{R}$, where (3.2), with $\mu_N[a \log n]$ written simply as μ_N, reads

$$\lim_{N\to\infty} \frac{N}{\sqrt{\log N}}\, d_{\mathbb{T}}\big(\mu_N, \eta_{a, a\log N}\big) = \sqrt{\frac{|a|}{6\pi}} \tag{3.11}$$

for all but countably many a; for instance, (3.11) holds whenever a is an algebraic number. By contrast, for every integer $q \ge 2$, and with $\log_q$ denoting the logarithm base q,

$$\lim_{N\to\infty} \frac{N}{\sqrt{\log N}}\, d_{\mathbb{T}}\big(\mu_N, \eta_{\log_q e, \log_q N}\big) = \sqrt{\frac{q+1}{6\pi(q-1)\log q}}\,;$$

cf. [25, Thm.3.3]. To relate (3.2) to the example given in the Introduction, notably to (1.1), observe that for every increasing sequence (N_j) in $\mathbb{N}$ and $t \in \mathbb{R}$, by the triangle inequality for $d_{\mathbb{T}}$ as well as (2.4),

$$\begin{aligned}\frac{N_j}{\sqrt{\log N_j}}\big|d_{\mathbb{T}}(\mu_{N_j}, \eta_{a, a\log N_j}) - d_{\mathbb{T}}(\mu_{N_j}, \eta_{a,t})\big| &\le \frac{N_j}{\sqrt{\log N_j}}\, d_{\mathbb{T}}(\eta_{a, a\log N_j}, \eta_{a,t})\\ &\le \frac{N_j}{\sqrt{\log N_j}}\, d_{\mathbb{T}}(a\log N_j, t)\end{aligned}$$

for every $j \in \mathbb{N}$. Thus, if $\sup_j N_j d_{\mathbb{T}}(a\log N_j, t) < \infty$ then

$$\lim_{j\to\infty} \frac{N_j}{\sqrt{\log N_j}}\, d_{\mathbb{T}}\big(\mu_{N_j}, \eta_{a,t}\big) = \sqrt{\frac{|a|}{6\pi} V(e^{1/|a|})}\,.$$

Specifically, with $a = \frac{1}{3}$ this yields ((1.1) for $t = 0$; Figure 1 illustrates how $\mu_{N_j} \to \eta_{1/3,t}$ for $t = 0$ (left) and $t = \frac{1}{3}\log 2$ (right), respectively.

A related example of enduring interest [14, 18–20, 22, 24] is $(\log p_n)$, where p_n denotes the n-th prime number. Recall that by the prime number theorem, $|\log p_n - f(n)| \to 0$ with $f(t) = \log t + \log\log t$. Evidently, this f does *not* lie within the scope of Theorem 3.1, as $\lim_{t\to\infty}(tf'(t) - 1)\log t = 1$, and so (3.1) fails. However, $\lim_{t\to\infty} tf'(t) = 1$, and it can be shown [2] that

$$\limsup_{N\to\infty} \log N d_{\mathbb{T}}\big(\mu_N[\log p_n], \eta_{1,\log p_N}\big) < \infty\,, \tag{3.12}$$

which is consistent with [19, Thm.6]. As of this writing, the precise rate of convergence for $d_{\mathbb{T}}\big(\mu_N[\log p_n], \eta_{1,\log p_N}\big) \to 0$ is unknown to the author, though (3.12) and the finer analysis in [20] suggest that it is much slower than $\big(\sqrt{\log N}/N\big)$, the rate observed for $(\log n)$ in (3.11).

4. A discrepancy rate of convergence

This section considers exactly the same real-valued sequences as the previous section, i.e., $\bigl(f(n)\bigr)$ with f satisfying (3.1), but employs the discrepancy-style metric d_∞ instead of $d_{\mathbb{T}}$. Again, $\mu_N[f(n)]$ is abbreviated as μ_N^f throughout. The following simple calculus fact is going to be useful.

Proposition 4.1. *Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a C^1-function. If f satisfies* (3.1) *for some $a \in \mathbb{R}$ then there exists a C^1-function $g : \mathbb{R}^+ \to \mathbb{R}$ such that $g(n) = f(n)$ for all $n \in \mathbb{N}$, with*

$$\int_1^\infty t|g'(t) - f'(t)|\,\mathrm{d}t < \infty \qquad \textit{and} \qquad \lim\nolimits_{t\to\infty} tg'(t) = a\,.$$

In addition, it will be convenient to use a straightforward generalization of (a special case of) the central limit theorem (3.4).

Proposition 4.2. *Let $I \subset \mathbb{R}$ be an interval with $\lambda(I) < \infty$, and for every $j \in \mathbb{N}$ let $\epsilon_j : I \to \mathbb{R}$ be measurable. If $\sum_{j=1}^\infty \sup_I |\epsilon_j| < \infty$ then for every $\beta \in \mathbb{R}^+$ with $\beta > 1$,*

$$\lim\nolimits_{M\to\infty} \lambda\left(\left\{t \in I : \frac{1}{\sqrt{M}} \sum\nolimits_{j=1}^M \left(\langle \beta^j t + \epsilon_j(t)\rangle - \frac{1}{2}\right) \le b\right\}\right) = \lambda(I)\,\Phi\left(2b\sqrt{\frac{3}{V(\beta)}}\right) \qquad \forall b \in \mathbb{R}\,.$$

With the usage of μ_N^f and $\eta_{a,t}$ exactly as in Theorem 3.1, the following analogue of that theorem for d_∞ hints at an asymptotics of (μ_N^f) quite different from what has been observed earlier. The result naturally strengthens and complements [19, 25].

Theorem 4.3. *Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a C^1-function. If f satisfies* (3.1) *for some $a \in \mathbb{R} \setminus \{0\}$, then*

$$\limsup\nolimits_{N\to\infty} \frac{N}{\log N}\, d_\infty\bigl(\mu_N^f, \eta_{a,f(N)}\bigr) \le |a|\,, \tag{4.1}$$

as well as

$$\liminf\nolimits_{N\to\infty} \frac{N}{\sqrt{\log N}}\, d_\infty\bigl(\mu_N^f, \eta_{a,f(N)}\bigr) = \infty\,. \tag{4.2}$$

Proof. Assume first that $a > 0$, and also that $f'(t) > 0$ for all sufficiently large t, with the latter assumption being justified by Proposition 4.1. Pick N_0

so large that $f(t) > 1$ and $f'(t) > 0$ whenever $t \ge N_0$, and denote by $g = g(t)$ the inverse of f, defined for all $t \ge f(N_0)$; for convenience, let $g(t) = N_0$ for all $t \le f(N_0)$. As in the proof of Theorem 3.1, define $\Delta(t) = tf'(t) - a$ for all $t \in \mathbb{R}^+$, but also $E(t) = \int_t^\infty \Delta(u)/u \, \mathrm{d}u$ and $E^*(t) = \int_t^\infty |\Delta(u)|/u \, \mathrm{d}u$. By (3.1), the functions E and E^* are well-defined, E^* is non-increasing, with $0 \le t|E(t)| \le tE^*(t) \to 0$ as $t \to \infty$, and $\int_1^\infty tE^*(t) \, \mathrm{d}t < \infty$. Observe that

$$f(t) = a \log t + D - E(t) \qquad \forall t \in \mathbb{R}^+,$$

where D is an appropriate constant, and hence also

$$g(t) = e^{(t-D)/a} e^{E \circ g(t)/a} \qquad \forall t \ge f(N_0) .$$

Both assertions of the theorem will follow easily once a sufficiently accurate asymptotic representation of $N(F_{\mu_N^f} - F_{\eta_{a,f(N)}})$, in the form of (4.8) below, has been obtained. To this end, with $L_N = \lfloor f(N) \rfloor$ and $s_N = \langle f(N) \rangle$ for convenience, straightforward counting yields for every $N \ge N_0$ and $0 \le s < 1$,

$$\begin{aligned} NF_{\mu_N^f}(s) &= \#\{1 \le n \le N : 0 \le \langle f(n) \rangle \le s\} \\ &= N_0 F_{\mu_{N_0}^f}(s) + \lfloor g(\max\{s + \lfloor f(N_0) \rfloor, f(N_0)\}) \rfloor - N_0 \\ &\quad + \sum\nolimits_{j=\lceil f(N_0) \rceil}^{L_N - 1} \big(\lfloor g(j+s) \rfloor - \lceil g(j) \rceil + 1 \big) \\ &\quad + \min\{ \lfloor g(L_N + s) \rfloor, N\} - \lceil g(L_N) \rceil + 1 \\ &= D_1(s) + L_N + \sum\nolimits_{j=\lceil f(N_0) \rceil}^{L_N - 1} \big(\lfloor g(j+s) \rfloor - \lceil g(j) \rceil \big) \\ &\quad + \min\{ \lfloor g(L_N + s) \rfloor, N\} - \lceil g(L_N) \rceil ; \end{aligned} \tag{4.3}$$

here $D_1(s)$ is bounded,

$$|D_1(s)| \le 1 + \lceil f(N_0) \rceil + N_0 e^{\big(1 + E^*(N_0)\big)/a} \qquad \forall 0 \le s < 1 .$$

To further analyze the remaining terms in the representation of $NF_{\mu_N^f}(s)$ in (4.3), observe first that for every $\lceil f(N_0) \rceil \le j \le L_N - 1$ and $0 \le s < 1$,

$$\begin{aligned} \lfloor g(j+s) \rfloor - \lceil g(j) \rceil &= g(j+s) - g(j) - \big(\langle g(j+s) \rangle + \langle -g(j) \rangle \big) \\ &= e^{(j+s-D)/a} e^{E \circ g(j+s)/a} \\ &\quad - e^{(j-D)/a} e^{E \circ g(j)/a} - \big(\langle g(j+s) \rangle + \langle -g(j) \rangle \big) \\ &= e^{(j+s-D)/a} - e^{(j-D)/a} - \big(\langle g(j+s) \rangle + \langle -g(j) \rangle \big) \\ &\quad + e^{(j+s-D)/a} \big(e^{E \circ g(j+s)/a} - 1 \big) - e^{(j-D)/a} \big(e^{E \circ g(j)/a} - 1 \big) . \end{aligned}$$

With regard to the last two terms, notice that for all $0 \le s < 1$,

$$\begin{aligned}
&\left|\sum\nolimits_{j=\lceil f(N_0)\rceil}^{L_N-1} e^{(j+s-D)/a}\bigl(e^{E\circ g(s+j)/a}-1\bigr)\right| \\
&\le \sum\nolimits_{j=\lceil f(N_0)\rceil}^{L_N-1} e^{(j+s-D)/a}\frac{E^*\circ g(j+s)}{a}e^{E^*\circ g(j+s)/a} \\
&\le \frac{e^{E^*(N_0)/a}}{a}\sum\nolimits_{j=\lceil f(N_0)\rceil}^{L_N-1} e^{(j+s-D)/a}\int_{g(j+s)}^{\infty}\frac{|\Delta(u)|}{u}\,\mathrm{d}u \\
&\le \frac{e^{E^*(N_0)/a}}{a}\int_1^{L_N} e^{(v+s-D)/a}\int_{g(j+s)}^{\infty}\frac{|\Delta(u)|}{u}\,\mathrm{d}u\,\mathrm{d}v \\
&= \frac{e^{E^*(N_0)/a}}{a}\int_{g(s)}^{\infty}\frac{|\Delta(u)|}{u}\int_1^{f(u)+1-s} e^{(v+s-D)/a}\mathrm{d}v\,\mathrm{d}u \\
&\le e^{E^*(N_0)/a}e^{1/a}\int_{N_0}^{\infty}|\Delta(u)|e^{E^*(u)/a}\mathrm{d}u \;\le\; e^{\bigl((1+2E^*(N_0)\bigr)/a}\int_{N_0}^{\infty}|\Delta(u)|\,\mathrm{d}u\,.
\end{aligned}$$

With this, it follows that

$$\begin{aligned}
&\sum\nolimits_{j=\lceil f(N_0)\rceil}^{L_N-1}\bigl(\lfloor g(j+s)\rfloor-\lceil g(j)\rceil\bigr) \\
&= (e^{s/a}-1)e^{-D/a}\frac{e^{L_N/a}-e^{\lceil f(N_0)\rceil/a}}{e^{1/a}-1} \\
&\quad -\sum\nolimits_{j=\lceil f(N_0)\rceil}^{L_N-1}\bigl(\langle g(j+s)\rangle+\langle -g(j)\rangle\bigr)+D_{2,N}(s) \\
&= Ne^{-s_N/a}\frac{e^{s/a}-1}{e^{1/a}-1} \\
&\quad -\sum\nolimits_{j=\lceil f(N_0)\rceil}^{L_N-1}\bigl(\langle g(j+s)\rangle+\langle -g(j)\rangle\bigr)+D_{3,N}(s)\,, \qquad (4.4)
\end{aligned}$$

where $D_{2,N}(s)$, $D_{3,N}(s)$ are uniformly bounded, in that

$$\sup\nolimits_{N\ge N_0,0\le s<1}\max\{|D_{2,N}(s)|,|D_{3,N}(s)|\}<\infty\,.$$

Next, a similar analysis for the final two terms in (4.3) yields

$$\begin{aligned}
&\min\{\lfloor g(L_N+s)\rfloor,N\}-\lceil g(L_N)\rceil = \lfloor g(L_N+\min\{s,s_N\})\rfloor+\lfloor -g(L_N)\rfloor \\
&= g(L_N+\min\{s,s_N\})-g(L_N)-\langle g(L_N+\min\{s,s_N\})\rangle-\langle -g(L_N)\rangle \\
&= g\bigl(f(N)+\min\{s-s_N,0\}\bigr)-g\bigl(f(N)-s_N\bigr) \\
&\quad -\bigl\langle g\bigl(L_N+\min\{s,s_N\}\bigr)\bigr\rangle-\bigl\langle -g(L_N)\bigr\rangle \\
&= N\left(e^{\min\{s-s_N,0\}/a}-e^{-s_N/a}\right)+D_{4,N}(s)\,, \qquad (4.5)
\end{aligned}$$

where $D_{4,N}(s)$ again is uniformly bounded.

Finally, plugging (4.4) and (4.5) into (4.3) yields

$$\begin{aligned} NF_{\mu_N^f}(s) &= D_1(s) + L_N + Ne^{-s_N/a}\frac{e^{s/a}-1}{e^{1/a}-1} \\ &\quad - \sum\nolimits_{j=\lceil f(N_0)\rceil}^{L_N-1}\big(\langle g(j+s)\rangle + \langle -g(j)\rangle\big) + D_{3,N}(s) \\ &\quad + N\big(e^{\min\{s-s_N,0\}/a} - e^{-s_N/a}\big) + D_{4,N}(s) \\ &= NF_{\eta_{a,f(N)}}(s) + L_N \\ &\quad - \sum\nolimits_{j=\lceil f(N_0)\rceil}^{L_N-1}\big(\langle g(j+s)\rangle + \langle -g(j)\rangle\big) \\ &\quad + D_1(s) + D_{3,N}(s) + D_{4,N}(s)\,. \end{aligned} \tag{4.6}$$

Now, for convenience define for every $N \ge N_0$ and $0 \le s < 1$,

$$\begin{aligned} h_N(s) &= 1 - \frac{1}{L_N}\sum\nolimits_{j=1}^{L_N}\big(\langle g(j+s)\rangle + \langle -g(j)\rangle\big) \\ &= -\frac{1}{L_N}\sum\nolimits_{j=1}^{L_N}\left(\left(\langle g(j+s)\rangle - \frac{1}{2}\right) + \left(\langle -g(j)\rangle - \frac{1}{2}\right)\right). \end{aligned} \tag{4.7}$$

With this, it is clear from (4.6) that

$$\sup\nolimits_{N\ge N_0, 0\le s<1}\left|N\big(F_{\mu_N^f}(s) - F_{\eta_{a,f(N)}}(s)\big) - L_N h_N(s)\right| < \infty\,. \tag{4.8}$$

As hinted at earlier, (4.1) now follows immediately, since $-1 < h_N(s) \le 1$ for all $N \ge N_0$ and $0 \le s < 1$, and consequently

$$\limsup\nolimits_{N\to\infty}\frac{N}{L_N}\,d_\infty\big(\mu_N^f, \eta_{a,f(N)}\big) \le 1\,.$$

Recalling that $\lim_{N\to\infty} L_N/\log N = a$ therefore completes the proof of (4.1) for $a > 0$.

To prove (4.2), notice first that for all $j \ge \lceil f(N_0)\rceil$ and $0 \le s < 1$,

$$g(j+s) = e^{(j+s-D)/a} + e^{(j+s-D)/a}\big(e^{E\circ g(j+s)/a} - 1\big) =: e^{(j+s-D)/a} + \epsilon_j(s)\,.$$

For convenience, let $\epsilon_j(s) = g(j+s) - e^{(j+s-D)/a}$ for all $1 \le j < \lceil f(N_0)\rceil$ as well, and $E_j := \sup_{0\le s<1}|\epsilon_j(s)|$. With this, for all $j \ge \lceil f(N_0)\rceil$,

$$E_j \le e^{(j+1-D)/a}\,\frac{E^*\circ g(j)}{a}\,e^{E^*(N_0)/a}\,,$$

and hence

$$\sum\nolimits_{j=\lceil f(N_0)\rceil}^{\infty} E_j \le \frac{e^{E^*(N_0)/a}}{a} \sum\nolimits_{j=\lceil f(N_0)\rceil}^{\infty} e^{(j+1-D)/a} \int_{g(j)}^{\infty} \frac{|\Delta(u)|}{u}\,\mathrm{d}u$$
$$\le \frac{e^{E^*(N_0)/a}}{a} \int_{f(N_0)}^{\infty} \int_{g(v-1)}^{\infty} e^{(v+1-D)/a} \frac{|\Delta(u)|}{u}\,\mathrm{d}u\,\mathrm{d}v$$
$$= \frac{e^{E^*(N_0)/a}}{a} \int_{N_0}^{\infty} \frac{|\Delta(u)|}{u} \int_{f(N_0)}^{f(u)-1} e^{(v+1-D)/a}\mathrm{d}v\,\mathrm{d}u$$
$$\le e^{2\left(1+E^*(N_0)\right)/a} \int_{N_0}^{\infty} |\Delta(u)|\,\mathrm{d}u < \infty\,.$$

Similarly to the proof of Theorem 3.1, let $S(v) = e^{(v-D)/a}$ for all $v \in \mathbb{R}$. Then, simply $g(j+s) = e^{j/a}S(s) + \epsilon_j(s)$ for all $j \in \mathbb{N}$, $0 \le s < 1$, and so, by Proposition 4.2 with $I = S[0,1]$ and $\beta = e^{1/a}$,

$$\lim\nolimits_{M\to\infty} \lambda\left(\left\{t \in I : \frac{1}{\sqrt{M}} \sum\nolimits_{j=1}^{M} \left(\langle e^{j/a}t + \epsilon_j \circ S^{-1}(t)\rangle - \frac{1}{2}\right) \le b\right\}\right) =$$
$$\lambda(I)\,\Phi\left(2b\sqrt{\frac{3}{V(\beta)}}\right) > 0 \qquad \forall b \in \mathbb{R}\,.$$

It follows that, given any $c \in \mathbb{R}^+$ and sufficiently large $N \in \mathbb{N}$, there exist $0 \le s_1, s_2 < 1$ such that

$$\frac{1}{\sqrt{L_N}} \sum\nolimits_{j=1}^{L_N} \left(\langle g(j+s_1)\rangle - \frac{1}{2}\right) \le -c\,, \quad \frac{1}{\sqrt{L_N}} \sum\nolimits_{j=1}^{L_N} \left(\langle g(j+s_2)\rangle - \frac{1}{2}\right) \ge c\,.$$

Thus, with

$$H_N := \frac{1}{L_N} \sum\nolimits_{j=1}^{L_N} \left(\langle -g(j)\rangle - \frac{1}{2}\right) \qquad \forall N \in \mathbb{N}$$

for convenience,

$$\sqrt{L_N}h_N(s_1) \ge c - \sqrt{L_N}H_N\,, \quad \sqrt{L_N}h_N(s_2) \le -c - \sqrt{L_N}H_N\,,$$

and consequently,

$$\sup\nolimits_{0\le s<1} \sqrt{L_N}|h_N(s)| \ge \max\left\{|c - \sqrt{L_N}H_N|, |c + \sqrt{L_N}H_N|\right\} \ge c\,.$$

Since $c \in \mathbb{R}^+$ has been arbitrary, $\liminf_{N\to\infty} \sqrt{L_N} \sup_{0\le s<1} |h_N(s)| = \infty$. Together with (4.8), this proves (4.2) for $a > 0$.

The case $a < 0$ is left to the reader, as the argument is completely analogous, notwithstanding the fact that d_∞ is not invariant under Q, unlike $d_{\mathbb{T}}$. $\square$

To see that the rate $(\log N/N)$ observed in (4.1) is sharp for *some* f, take for example $f(t) = a \log t$ with $a > 0$. Here (4.7) simply reads

$$h_N(s) = 1 - \frac{1}{L_N} \sum\nolimits_{j=1}^{L_N} \left(\left\langle e^{(j+s)/a} \right\rangle + \left\langle -e^{j/a} \right\rangle \right) .$$

For instance, if $e^{1/a}$ is an *integer*, then $\lim_{\varepsilon \downarrow 0} h_N(\varepsilon) = 1$, and consequently

$$\lim\nolimits_{N\to\infty} \frac{N}{\log N}\, d_\infty\big(\mu_N^f, \eta_{a,f(N)}\big) = a\,;$$

cf. [25, Cor.3.4]. Slightly more generally, it is not hard to see that

$$\liminf\nolimits_{N\to\infty} \frac{N}{\log N}\, d_\infty\big(\mu_N^f, \eta_{a,f(N)}\big) > 0 \tag{4.9}$$

whenever $e^{1/a}$ is a Pisot number for which the Galois conjugate with the second-largest modulus is non-negative. (For example, the larger root of $r^2 - 4r + 2 = 0$ is a Pisot number with this property.) In general, however, determining the precise asymptotics of $\big(d_\infty(\mu_N^f, \eta_{a,f(N)})\big)$ may be a challenging task, not least because

$$\lim\nolimits_{N\to\infty} \frac{N}{\log N} \int_0^1 |F_{\mu_N^f}(s) - F_{\eta_{a,f(N)}}(s)|\, \mathrm{d}s = 0$$

for Lebesgue almost every a. Nonetheless, the author suspects that (4.9) in fact always holds under the assumptions of Theorem 4.3. If indeed it does, then the precise value(s) of the lim sup and lim inf in (4.1) and (4.9), respectively, may depend on a in an exceedingly complicated manner. The reader may want to compare this delicate situation to the clear-cut conclusion of Theorem 3.1.

Corollary 4.4. *Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a C^1-function. If f satisfies* (3.1) *for some $a \in \mathbb{R} \setminus \{0\}$, then*

$$\lim\nolimits_{N\to\infty} \left| d_\infty(\mu_N^f, \mu) - d_\infty(\eta_{a,f(N)}, \mu) \right| = 0 \qquad \forall \mu \in \mathcal{P}\,. \tag{4.10}$$

As with $d_\mathbb{T}$, the function $t \mapsto d_\infty(\eta_{a,t}, \mu)$ is continuous, 1-periodic, and attains a positive maximal value for every $\mu \in \mathcal{P}$. Unlike with $d_\mathbb{T}$, however, this function is *not* constant for $\mu = \lambda_\mathbb{T}$, as

$$0 < 2 \min\nolimits_{t\in\mathbb{R}} d_\infty(\eta_{a,t}, \lambda_\mathbb{T}) = \max\nolimits_{t\in\mathbb{R}} d_\infty(\eta_{a,t}, \lambda_\mathbb{T}) \qquad \forall a \in \mathbb{R}\,;$$

this fact has been recorded repeatedly, if often implicitly, in the literature [3, 9].

The reader will no doubt have noticed that, unlike in Theorem 3.1 and Corollary 3.3, the case $a = 0$ is forbidden in Theorem 4.3 and Corollary 4.4. To see that the conclusions of the latter results may indeed fail for $a = 0$, pick first an increasing sequence (N_j) in $\mathbb{N}$, as well as a decreasing sequence (a_j) in $\mathbb{R}^+$ with $a_1 < \frac{1}{2}$ and $\sum_{j=1}^{\infty} j a_j N_j < \infty$. With this, consider (x_n) given by

$$(x_n) = (\underbrace{0, \ldots, 0}_{N_1 \text{ zeros}}, a_1, \underbrace{0, \ldots, 0}_{N_2 \text{ zeros}}, a_2, \underbrace{0, \ldots}_{\text{etc.}}).$$

Clearly, (3.1) holds for an appropriate f with $(x_n) = \bigl(f(n)\bigr)$ and $a = 0$. Moreover, with $N = N_1 + \cdots + N_j + j$ for $j \in \mathbb{N}$,

$$d_\infty(\mu_N^f, \delta_{f(N)}) = \frac{N-j}{N} \geq 1 - \frac{2}{j+3},$$

but also

$$d_\infty(\mu_{N-1}^f, \delta_{f(N-1)}) = 1 - \frac{N-j}{N-1} \leq \frac{2}{j+3},$$

from which it is clear that

$$\limsup_{N\to\infty} d_\infty\bigl(\mu_N^f, \delta_{f(N)}\bigr) = 1 > 0 = \liminf_{N\to\infty} d_\infty\bigl(\mu_N^f, \delta_{f(N)}\bigr).$$

Moreover, taking for instance $N_j = j$ for all $j \in \mathbb{N}$ yields

$$\liminf_{N\to\infty} \sqrt{N} d_\infty(\mu_N^f, \delta_{f(N)}) = \sqrt{2}.$$

Notice that all this is inconsistent with (4.1) which for $a = 0$ would read $\lim_{N\to\infty} N d_\infty(\mu_N^f, \delta_{f(N)}) / \log N = 0$. By contrast, it is readily seen that

$$N d_{\mathbb{T}}\left(\mu_N^f, \delta_{f(N)}\right) \leq j a_j N_j + \sum_{k=1}^{j} a_k$$

for every $N \in \{N_1 + \cdots + N_j + j, \ldots, N_1 + \cdots + N_{j+1} + j\}$, so the sequence $\bigl(N d_{\mathbb{T}}(\mu_N^f, \delta_{f(N)})\bigr)$ is bounded, and (3.2) is valid for $a = 0$, as expected.

In a similar vein, to see that (4.10) may fail for $a = 0$, simply take $f(t) = e^{-t}\cos(\pi t)$. Here $\lim_{N\to\infty} F_{\mu_N^f}(s) = \frac{1}{2}$ for every $0 < s < 1$, and consequently $\lim_{N\to\infty} d_\infty(\mu_N^f, \lambda_{\mathbb{T}}) = \frac{1}{2}$, whereas clearly, $\lim_{N\to\infty} d_\infty(\delta_{f(N)}, \lambda_{\mathbb{T}}) = 1$. Again, contrast this with the fact that

$$\lim_{N\to\infty} d_{\mathbb{T}}(\mu_N^f, \lambda_{\mathbb{T}}) = \tfrac{1}{4} = \lim_{N\to\infty} d_{\mathbb{T}}(\delta_{f(N)}, \lambda_{\mathbb{T}}) = d_{\mathbb{T}}(\delta_0, \lambda_{\mathbb{T}}),$$

in perfect agreement with Corollary 3.3.

Examples like these illustrate once again how usage of $d_{\mathbb{T}}$ may lead to more robust and comprehensive statements, compared to d_∞.

Acknowledgement. The author was partially supported by an NSERC Discovery Grant. He wishes to thank K. Fukuyama for a helpful communication. He is most grateful to an anonymous referee who pointed out a serious oversight in the original manuscript, provided pertinent additional references, and made several helpful suggestions regarding the presentation overall.

REFERENCES

[1] BECK, J.: *Randomness of the square root of 2 and the giant leap, Part 1*, Period. Math. Hungar. **60** (2010), 137–242.

[2] BERGER, A.: *Circling the uniform distribution* (in preparation).

[3] BERGER, A.—HILL, T. P.: *An Introduction to Benford's Law*. Princeton University Press, Princeton, NJ, 2015.

[4] BROWN, L.—STEINERBERGER, S.: *On the Wasserstein distance between classical sequences and the Lebesgue measure*, Trans. Amer. Math. Soc. **373** (2020), 8943–8962.

[5] BURTON, R.—DENKER, M.: *On the central limit theorem for dynamical systems*, Trans. Amer. Math. Soc. **302** (1987), 715–726.

[6] CABRELLI, C. A.—MOLTER, U. M.: *The Kantorovich metric for probability measures on the circle*, J. Comput. Appl. Math. **57** (1995), 345–361.

[7] CIGLER, J.: *Asymptotische Verteilung reeller Zahlen* mod 1, Monatsh. Math. **64** (1960), 201–225.

[8] DUDLEY, R. M.: *Real Analysis and Probability*. Cambridge Studies in Advanced Mathematics Vol. 74, Cambridge University Press, Cambridge, 2002.

[9] FLEHINGER, B. J.: *On the probability that a random integer has initial digit A*, Amer. Math. Monthly **73** (1966), 1056–1061.

[10] FUKUYAMA, K.: *The central limit theorem for Riesz–Raikov sums*, Probab. Theory Related Fields **100** (1994), 57–75.

[11] FUKUYAMA, K.: *The central limit theorem for* $\sum f(\theta^n x)g(\theta^{n^2} x)$, Ergodic Theory Dynam. Systems **20** (2000), no. 5, 1335–1353.

[12] GIULIANO ANTONINI, R.—STRAUCH, O.: *On weighted distribution functions of sequences*, Unif. Distrib. Theory **3** (2008), 1–18.

[13] GRAHAM, C.: *Irregularity of distribution in Wasserstein distance*, J. Fourier Anal. Appl. **26** (2020), no. 5, Paper no. 75, 21 pp.

[14] HUTSON, H. L.: *On the distribution of* $\log(p)$, Int. J. Pure Appl. Math. **7** (2003), 499–508.

[15] KEMPERMAN, J. H. B.: *Distributions modulo* 1 *of slowly changing sequences*, Nieuw Arch. Wisk. **21** (1973), 138–163.

[16] KUIPERS, L.—NIEDERREITER, H.: *Uniform Distribution of Sequences*. Wiley--Interscience [John Wiley & Sons], New York-London-Sydney, 1974.

[17] NIEDERREITER, H.: *Distribution* mod 1 *of monotone sequences*, Nederl. Akad. Wetensch. Indag. Math. **46** (1984), 315–327.

[18] OHKUBO, Y.: *On sequences involving primes*, Unif. Distrib. Theory **6** (2011), 221–238.

[19] OHKUBO, Y.—STRAUCH, O.: *Distribution of leading digits of numbers*, Unif. Distrib. Theory **11** (2016), 23–45.

[20] OHKUBO, Y.—STRAUCH, O.: *Distribution of leading digits of numbers II*, Unif. Distrib. Theory **14** (2019), 19–42.

[21] STEINERBERGER, S.: *Wasserstein distance, Fourier series and applications*, Monatsh. Math. **194** (2021), 305–338.

[22] STRAUCH, O.— BLAŽEKOVÁ, O.: *Distribution of the sequence* $p_n/n \bmod 1$, Unif. Distrib. Theory **1** (2006), 45–63.

[23] WINKLER, R.: *On the distribution behaviour of sequences*, Math. Nachr. **186** (1997), 303–312.

[24] WINTNER, A.: *On the cyclical distribution of the logarithms of the prime numbers*, Q. J. Math., Oxf. Ser. 6 **1** (1935), 65–68.

[25] XU, C.: *The distributional asymptotics* mod1 *of* $(\log_b n)$, Unif. Distrib. Theory **14** (2019), 105–122.

[26] XU, C.—BERGER, A.: *Best finite constrained approximations of one-dimensional probabilities*, J. Approx. Theory **244** (2019), 1–36.

Received April 29, 2022
Accepted June 21, 2022

Arno Berger
Department of Mathematical and Statistical Sciences
University of Alberta
Edmonton T6G 2G1
CANADA
E-mail: berger@ualberta.ca