ISRAEL JOURNAL OF MATHEMATICS **139** (2004), 189–213

RANDOM HOMEOMORPHISMS AND FOURIER EXPANSIONS — THE POINTWISE BEHAVIOUR

BY

GADY KOZMA*

The Weizmann Institute of Science, Rehovot, 76100 Israel
e-mail: gadyk@wisdom.weizmann.ac.il, gadykozma@hotmail.com

ABSTRACT

Let φ be a Dubins–Freedman random homeomorphism on $[0,1]$ derived from the base measure uniform on $\{x=\frac{1}{2}\}$, and let f be a periodic function satisfying $|f(\delta)-f(0)| = o(\log\log\log\frac{1}{\delta})^{-1}$. Then the Fourier expansion of $f\circ\varphi$ converges at 0 with probability 1. In the condition on f, o cannot be replaced by O. Also we deduce some 0-1 laws for this kind of problem.

1. Introduction

This paper is a continuation of an earlier paper, [KO98], where a number of questions related to the Fourier expansions of $f\circ\varphi$ were discussed, most notably conditions under which $S_n(f\circ\varphi)$ converges uniformly for a set of φ's with probability 1, where S_n stands for the nth Fourier sum. It was proved that if

$$\omega_\delta(f) = o\Big(\log\log\frac{1}{\delta}\Big)^{-1}$$

then $S_n(f\circ\varphi)$ converges uniformly almost surely, where $\omega_\delta(f)$ stands as usual for the modulus of continuity of f, i.e.

$$\omega_\delta(f) := \sup_{|x-y|\le\delta} |f(x)-f(y)|,$$

and that this result is sharp (Theorems 4 and 6 ibid).

---

* The research was supported by The Israel Science Foundation (grant no. 4/01).
Received June 25, 2002

In sections 3 and 4 we address the question of convergence at a specific point. The most obvious formulation might be "under what conditions does $S_n(f\circ\varphi)(x)$ converge?" However, in this formulation it is impossible to get local conditions on f since φ smooths out all the points. A better formulation uses conditional probability, and reads "under what local conditions on f near y do we have that

$$S_n(f\circ\varphi)(x)|\varphi(x)=y$$

converges?" Essentially, the answer would be the same, i.e. a triple log condition, but this formulation incurs a number of technical problems, so we simplify the proof making use of the fact that $\varphi(0)=0$. Thus we reached the formulation of the result in the abstract, i.e.

THEOREM 1: *Suppose f is a continuous function on the circle satisfying*

$$|f(\delta)-f(0)|=o\Big(\log\log\log\frac{1}{\delta}\Big)^{-1}.$$

Then the Fourier expansion of $f\circ\varphi$ converges at 0 with probability 1.

and this condition is sharp in the following sense:

THEOREM 2: *There exists a continuous function f satisfying*

$$|f(\delta)-f(0)|=O\Big(\log\log\log\frac{1}{\delta}\Big)^{-1}$$

for which the Fourier expansion of $f\circ\varphi$ diverges at 0 with probability 1.

Actually, f may be constructed to satisfy this condition globally, i.e., $\omega_f(\delta)=O(\log\log\log\frac{1}{\delta})^{-1}$.

It is instructive to contrast these results with the non-stochastic case. The results of [KO98] are analogues of the Dini–Lipschitz test [Z59, 2.71] which gives a sufficient sharp condition for uniform convergence of $S_n(f)$, $\omega_f(\delta)=o(\log\frac{1}{\delta})^{-1}$; for convergence at a point we have the Dini test [Z59, 2.4] which gives a sufficient condition $\int\frac{1}{\delta}\omega_f(\delta;x)<\infty$ (again, sharp) where $\omega_f(\delta;x)$ is the modulus of continuity of f at the point x. Thus in the classical case the condition for pointwise convergence is slightly **stronger**, or in other words, a global estimate of ω_f gives better information about convergence at a specific point than an estimate only at that point. This behavior, as remarked, does not happen in our probabilistic settings. Of course, we also get a much wider gap, an additional log factor. We also wish to reiterate remark 4.4i from [KO98]: there exist functions f satisfying $\omega_f(\delta)=O(\log\log\frac{1}{\delta})^{-1}$ such that the Fourier expansion of $f\circ\varphi$ diverges at a (random) point. This result has no non-probabilistic equivalent. For a discussion of properties of $S_n(f \circ \varphi)$ where φ is non-probabilistic, e.g. problems such as when $S_n(f \circ \varphi)$ might satisfy certain properties for some φ, all φ or a second category set of φ see [K83], [O81] or [O85].

Of course, the discussion above does not make much sense without specifying the probabilistic model for picking φ, and the group of homeomorphisms has no Haar measure. We shall be using a model suggested by Dubins and Freedman [DF65] which uses a base measure ν on $[0,1]^2$. Roughly, a point (x, y) on the graph of φ is chosen at random using ν, then this process is repeated for the rectangles extending from $(0,0)$ to (x, y) and from (x, y) to $(1,1)$ with rescaled versions of ν. Repeating this over and over we get a sequence of points which can, with probability 1, be closed to a graph of a homeomorphism $[0,1] \to [0,1]$ with $\varphi(0) = 0$ and $\varphi(1) = 1$. A proper, though restricted, definition is provided in section 2.2. It must be noted, though, that Dubins and Freedman were not interested in homeomorphisms but in measures, and considered the Lebesgue–Stieltjes measures $d\varphi$ as random probability measures on $[0,1]$ and studied conditions under which a typical $d\varphi$ might be singular, atomic and so on.

Not all Dubins–Freedman measures are born equal, and the most natural ones are the ones with base measure uniform on $\{x = \frac{1}{2}\}$, $\{y = \frac{1}{2}\}$ and on $[0,1]^2$. See e.g. [GMW86] for a specific discussion of these three measures — they studied the properties of the set $\varphi(x) = x$ and other interesting facts about a typical φ. This paper will be using the first one. Of course, measures centered on a vertical line are easier to analyse because one can have an explicit formula for the distributions of $\varphi(x)$ for dyadic x, and sometimes for other x's too; for example, in the uniform case, $\varphi(\frac{1}{3})$ has the density function $1 - x$ [KO98, Lemma 1.6]. What might be less clear is that I really need the distribution to be uniform. Indeed, generalizing the results of [KO98] for measures on $\{x = \frac{1}{2}\}$ which are non-uniform is an open problem. Such a result could be interesting, for example, in order to play around with the almost-sure Hölder constant of φ.

In the last section we discuss the 0-1 law. It turns out that for this kind of problems, the 0-1 law is not self-evident. We shall reduce the problem to a functional-integral equation (15) which can be solved by elementary manipulations. This general technique allows one to get 0-1 laws for many problems related to $S_n(f \circ \varphi)$: uniform convergence, pointwise convergence, boundedness of partial sums etc.

I wish to end this introduction with a question I wasn't even able to formulate properly. If $I \subset [0,1]$ is a dyadic interval then the conditional restricted homeomorphism $\psi := \varphi|_I|\varphi(\partial I)$ is similar to the original φ — this is the "scaling invariance", see (1) below. If, however, I is not dyadic then this is no longer true, but ψ still seems to be very similar to φ. Many of the results of this paper and of [KO98] can be reproved for ψ. It could be very interesting (and useful) to prove that for "infinitesimal" problems, ψ and φ are equivalent.

2. Preliminaries

2.1 NOTATIONS. We denote by $\mathbb{T}$ the circle group, which we identify with the interval $[0,1]$; $\mathbf{m}$ denotes the Lebesgue measure on $[0,1]$; C and c denote absolute positive constants, possibly different, with C usually pertaining to constants large enough and c to constants small enough. For a continuous function f, $||f||$ denotes its supremum and $\operatorname{supp} f$ its support.

$\mathbb{P}$ denotes the probability of some event (with the measure on the random homeomorphisms defined in the next section); $\mathbb{E}$ denotes the expectation of a variable, and $\mathbb{V}$ its variance. The notation $X \sim Y$ for two variables means "X and Y have the same distribution".

Dyadic rationals are numbers of the type $k2^{-n}$, k and n integers, and dyadic intervals are intervals of the type $[k2^{-n},(k+1)2^{-n}]$. For an interval $I := [a,b]$ the boundary ∂I is the set $\{a,b\}$; $\lfloor x \rfloor$ denotes the largest integer $\leq x$ and $\lceil x \rceil$ the smallest integer $\geq x$.

D_n denotes the Dirichlet kernel on $[0,1]$, i.e., $\sin((2n+1)\pi x)/\sin(\pi x)$, so

$$S_n(f;x) = \int_0^1 D_n(x-t)\cdot f(t)dt.$$

The pointwise modulus of continuity of f at x is defined by

$$\omega_f(x;\delta) := \sup_{0<|\mu|<\delta} |f(x+\mu) - f(x)|,$$

where for $\delta = 0$ we define $\omega_f(x;\delta)$ assuming the function to be periodic.

2.2 RANDOM HOMEOMORPHISMS. Let's start with the following definition of the particular Dubins–Freedman measure we will be using, which will be easy to work with. Let $X_{n,k}$ be independent uniform variables in $[0,1]$ for any $n \in \mathbb{N}$ and any odd $0 < k < 2^n$. We define an increasing function φ on the dyadic rational using the following procedure: Start by taking $\varphi(0) = 0$, $\varphi(1) = 1$, and

$$\varphi(\tfrac{1}{2}) = X_{1,1}.$$

On the second step, define

$$\varphi(\tfrac{1}{4}) = \varphi(\tfrac{1}{2}) \cdot X_{2,1}, \quad \varphi(\tfrac{3}{4}) = \varphi(\tfrac{1}{2}) + (1 - \varphi(\tfrac{1}{2})) \cdot X_{2,3},$$

i.e. $\varphi(\frac{1}{4})$ and $\varphi(\frac{3}{4})$ are distributed uniformly on $[0, \varphi(\frac{1}{2})]$ and $[\varphi(\frac{1}{2}), 1]$ respectively, and are otherwise independent. We continue this process, at the nth step taking

$$\varphi(k2^{-n}) := \varphi((k-1)2^{-n}) + X_{n,k} \cdot (\varphi((k+1)2^{-n}) - \varphi((k-1)2^{-n}).$$

This defines φ on all dyadic fractions. With probability 1, φ can be extended to a homeomorphism of $[0, 1]$ [DF65, Theorem 4.1]. We denote this measure by $\mathbb{P}$, and by φ the random change of variable.

The most useful property of φ is "scaling invariance", which roughly says that for any dyadic interval I, $\varphi|_I$ behaves like a small copy of φ. To be more precise,

Lemma 1: *If $I = [k2^{-n}, (k+1)2^{-n}]$ is a dyadic interval, then*

$$(\varphi|\varphi(\partial I) = \{a, b\})|_I \sim (\varphi \circ L) \cdot (b - a) + a \tag{1}$$

where L is a linear increasing map of I onto $[0, 1]$.

The proof may be found in [GMW86], Theorem 4.6.

Finally, we need the following simple calculation, which can be found in [KO98] in Lemma 1.4 and the remark that follows. For some constants K_1 and K_2 we have

$$\mathbb{P}\{r^{K_1} < \varphi(r) < r^{K_2}\} > 1 - Cr^2 \tag{2}$$

for any $r > 0$. It will be convenient to assume $K_2 < 1 < K_1$.

2.3 And Fourier expansions. We need the following lemmas, which are deeply related to (though unfortunately not direct consequences of) Theorem 2 from [KO98]:

Lemma 2: *For any continuous f, n, $r > 2/n$ and $K > 0$,*

$$\mathbb{P}\Big(\Big|\int_r^{1-r} (f \circ \varphi) \cdot D_n\Big| > K\|f\|\Big) < C\exp\Big(-c\frac{\sqrt{nr}}{\log nr}K\Big).$$

Lemma 3: *For any continuous f, n, interval I and $K > 0$,*

$$\mathbb{P}\Big(\Big|\int_I (f \circ \varphi) \cdot D_n\Big| > K\|f\|\Big) < Ce^{-e^{cK}}.$$

LEMMA 4: *For some constant β, the same c as above, and any constant K,*

$$\mathbb{P}\left(\left|\int_{-r}^{r}(f\circ\varphi)\cdot D_n\right| > 2K\omega_f(0;Cr^{\beta})\right) < Cr^2 + Ce^{-e^{cK}}.$$

In other words, if Theorem 2 of [KO98] gave an estimate of $\int_0^1(f\circ\varphi)\cdot D_n$ then these lemmas give split estimates for the head and the tail. The proof of Lemma 4 is an easy corollary to Lemma 3, so let's start with it:

Proof: Clearly, we may assume $r = 2^{-k}$. For each of the segments $[-r,0]$ and $[0,r]$, we use (2) ($\beta \equiv K_2$), apply the scaling invariance of φ and finally use Lemma 3 for a scaled version of f. ∎

As for the proofs of Lemmas 2 and 3, they follow quite closely the proof of the aforementioned Theorem 2, so the rest of this section must be read parallel to it. For Lemma 2, start from page 1029 ibid. There $||f|| = 1$ (which we can also assume here, of course), I_k denotes an arc of $\mathbb{T}$ symmetric around 0 containing $2k-1$ peaks of the Dirichlet kernel D_n and $Y_k := \int_{I_k^C} D_n\cdot(f\circ\varphi)$; I_k and Y_k are connected by the inequality

$$\mathbb{E}(Y_k^2 \mid \varphi|_{I_k}) \le C\frac{\log^2 k}{k},$$

which is Lemma 2.6 ibid. We define

$$j := \lfloor nr \rfloor \quad \text{and} \quad \mu := C\frac{\log j}{\sqrt{j}}$$

with C chosen to satisfy

$$\mathbb{P}(|Y_s| > \mu \mid \varphi|_{I_s}) \le \frac{1}{4} \quad \forall s \ge j.$$

With this μ we get

LEMMA 5: *If, for a given $\epsilon > 0$ and $\nu \ge 1$, the inequality*

$$\mathbb{P}(|Y_s| > \nu\mu \mid \varphi|_{I_s}) \le \epsilon \quad \forall s \ge j,$$

then

$$\mathbb{P}(|Y_s| > (2\nu+2)\mu \mid \varphi|_{I_s}) \le \frac{4}{3}\epsilon^2.$$

The proof is word-for-word identical to the proof of Lemma 2.8 ibid. Now, starting from the definition of μ we apply Lemma 5 inductively l times and get that

$$\mathbb{P}(|Y_s| > \mu d_l \mid \varphi|_{I_s}) \le \left(\frac{4}{3}\right)^{2^l-1}\left(\frac{1}{4}\right)^{2^l} < \left(\frac{1}{3}\right)^{2^l}$$

where the d_l's are defined recursively by $d_1 = 1$, $d_l = 2d_{l-1}+2$. Clearly $d_l < C2^l$. Picking a maximal l such that $\mu d_l < K$ we get that $2^l > cK/\mu$ and Lemma 2 follows. ∎

The proof of Lemma 3 is even more similar to that of Theorem 2 from [KO98], and we shall omit it.

3. Pointwise convergence

Proof of Theorem 1: Throughout the proof we shall assume that $f \in C(\mathbb{T})$ is some fixed function, that $||f|| \le 1$ and that $f(0) = 0$. We fix n sufficiently large for the rest of the proof. Define $r = (\log^5 n)/n$. Lemma 2 will ensure that

$$\text{(3)} \qquad \mathbb{P}\Big(\Big|\int_r^{1-r} (f\circ\varphi)\cdot D_n\Big| > \frac{1}{\log n}\Big) \le C\exp\Big(-c\frac{\log^{1.5} n}{\log\log n}\Big) < \frac{C}{n^3}.$$

Let us now assume that some m_1 and m_2 satisfy $m_1 - m_2 < n/\log^6 n$. A simple calculation will show

$$|D_{m_1} - D_{m_2}| = \Big|\frac{2\cos((m_1+m_2+1)\pi x)\sin((m_1-m_2)\pi x)}{\sin(\pi x)}\Big| < C\frac{n}{\log^6 n}$$

so

$$\int_{-r}^{r} |D_{m_1} - D_{m_2}| < C/\log n$$

which, combined with (3), gives

$$\text{(4)} \qquad \mathbb{P}\Big(\Big|\int_0^1 (f\circ\varphi)\cdot(D_{m_1} - D_{m_2})\Big| > \frac{C}{\log n}\Big) < \frac{C}{n^3}.$$

Thus, if we only calculate the behavior of $\int (f\circ\varphi)\cdot D_m$ on a sequence of m's from n to $2n$ with jumps $\lfloor n/\log^6 n\rfloor$, we will get a uniform estimate for all $m \in [n, 2n]$. Now is the time to use the log-log-log assumption on f. Let $\epsilon(n) \to 0$ be some sequence converging to 0 sufficiently slow as to satisfy

$$\frac{1}{\epsilon(n)}\omega_f(0; n^{-\beta}) = o(\log\log\log n)^{-1}.$$

Remembering Lemma 4 (from which we also take the β above), this gives

$$\mathbb{P}\Big(\Big|\int_{-r}^{r} (f\circ\varphi)\cdot D_n\Big| > \epsilon(n)\Big) < Cr^2 + Ce^{-e^{\Omega(\log\log\log n)}} < C/\log^8 n$$

(Ω, as usual, denoting the opposite of o). We use this inequality on a sequence of m's which has a length $< C\log^6 n$ and throw in (3) and (4) to get

$$\mathbb{P}\Big(\exists m \in [n, 2n], \ \Big|\int_0^1 (f\circ\varphi)\cdot D_m\Big| > \epsilon(n) + \frac{C}{\log n}\Big) < \frac{C}{\log^2 n} + n\frac{C}{n^3},$$

and summing these probabilities for $n = 2^k$ we get the desired result: that these events happen only for a finite number of n's for almost every φ. ∎

Remarks: 1. Actually, we never used the continuity of f. The theorem holds for any L^∞ function satisfying

$$|f(\delta) - f(0^+)| = o\Big(\log\log\log\frac{1}{\delta}\Big)^{-1}$$

and

$$|f(1-\delta) - f(1^-)| = o\Big(\log\log\log\frac{1}{\delta}\Big)^{-1}$$

(for an explanation why $S_n(f \circ \varphi)$ is even well defined for non-continuous f, see [KO98, Lemma 1.3]). If $f(0^+) \neq f(1^-)$, we can simply take $f - g$ where g is an appropriate linear function; $h := g \circ \varphi$ will be a monotone function, for which we always have that the Fourier expansion at x converges to $\frac{1}{2}(h^+(x) + h^-(x))$.

2. A similar proof shows that for any continuous function f,

$$S_n(f \circ \varphi; 0) = o(\log\log\log n),$$

and for any $f \in L^\infty$,

$$S_n(f \circ \varphi; 0) = O(\log\log\log n).$$

These results, too, are sharp.

4. Sharpness

This section will be devoted to the proof of Theorem 2. Ideologically, the essentials of the proof are contained in the following heuristics. Examine the following function:

$$f_n(t) = \begin{cases} \sin 2\pi(tn^k + \psi_k), & t \in [n^{-k}, n^{-k+1}],\ 1 \leq k \leq e^{n^4}, \\ 0, & t < n^{-e^{n^4}}. \end{cases}$$

with some phases $\psi_k \in [0,1]$ (usually chosen to make f_n continuous). On each interval $\varphi^{-1}([n^{-k}, n^{-k+1}])$, $f \circ \varphi$ has $n-1$ peaks, and with some small probability they will be "aligned" with the peaks of some Dirichlet kernel D_r. The probability to get a good alignment of $n-1$ peaks is approximately e^{-Cn}, and the variables

$$\varphi|_{\varphi^{-1}([n^{-k}, n^{-k+1}])}$$

are "approximately independent" so one would expect that in $> e^{Cn}$ such variables, with big probability this alignment will happen at least once. In this case, we will have

$$\int_{\varphi^{-1}([n^{-k},n^{-k+1}])} (f\circ\varphi)\cdot D_r \approx c\int |D_r| > c\log n.$$

So

$$\sup_{I\subset[0,1],\ r\in\mathbb{N}} \int_I (f\circ\varphi)\cdot D_r > c\log n > c\log\log\log n^{e^{n^4}}.$$

To make these calculations into a proper proof, we need to do the following:

(i) Explain what it means to "get a good alignment of $f\circ\varphi$ with D_r" and calculate the probability. The calculation will not give e^{-Cn} but a rather weaker estimate — hence the element n^4 in the definition of f_n.

(ii) Explain how to overcome the problem that these "approximately independent" variables are not properly independent.

(iii) Explain why it is enough to get a supremum of $\int_I$ for some $I\subset[0,1]$ rather than of $\int_{[0,1]}$.

(iv) Combine the f_n's into a single function f which will satisfy the requirements of the theorem.

We start with issue (iii).

Lemma 6: *Let $||f||\le 1$, $K>1$, p and $r_0<r_1$ be given with the condition*

$$\mathbb{P}\Bigg\{\sup_{\substack{y\in[0,1]\\ r\in[r_0,r_1]}}\left|\int_{[0,y]}(f\circ\varphi)\cdot D_r\right| > K\Bigg\} > p;$$

then

$$\mathbb{P}\Bigg\{\sup_{r\in[r_0,r_1]}|S_r(f\circ\varphi;0)| > \tfrac{1}{2}K\Bigg\} > p-\frac{C}{K}.$$

The proof is practically identical to the proof of Lemma 4.5 from [KO98], and we shall omit it. The reader might want to skip to the final steps of the proof of Theorem 2 to see how this lemma is used.

To investigate the independence properties of φ, i.e. to explain issue (ii), let us return to the variables $X_{n,k}$ defining the measure. For each $i>j\in\mathbb{N}$ define $\Omega_{i,j}$ to be the σ-field spanned by

$$\{X_{n,k}: 2^{-i} < k2^{-n} < 2^{-j}\}.$$

Clearly, $j>k$ imply that $\Omega_{i,j}$ and $\Omega_{k,l}$ are independent.

Lemma 7: *Let $i > 1$ be an integer, $0 < \epsilon < 1$ and $0 < y \le x \le 1 - \epsilon$. Then*

$$\mathbb{P}(\varphi(1/2) \in [x, x+\epsilon] \mid \varphi(2^{-i}) = y) > \left(\frac{\epsilon}{2\max(1, |\log y|)}\right)^i. \tag{5}$$

This is a somewhat tedious exercise in calculus. Let us work it out. A simple calculation (which may be found in [KO98], (1) page 1022) shows that the distribution function of $\varphi(2^{-i})$ is

$$\frac{1}{(i-1)!}\left|\log^{i-1} y\right| \tag{6}$$

which, using the scaling invariance of φ, gives the conditional distribution function

$$\operatorname{dist}(\varphi(1/2) = x \mid \varphi(2^{-i}) = y) = \begin{cases} \left|\frac{(i-1)\cdot\log^{i-2}(x/y)}{x\cdot\log^{i-1}(y)}\right|, & x > y, \\ 0, & x \le y. \end{cases} \tag{7}$$

This distribution (as a function of x) is increasing until $x_0 = ye^{i-2}$ and then decreasing — which clearly implies that the probability (5) as a function of x is increasing until some x_1 defined by the equality

$$\frac{\log^{i-2}(x_1/y)}{x_1} = \frac{\log^{i-2}((x_1+\epsilon)/y)}{x_1+\epsilon}$$

and then decreasing, so the minimum is achieved at $x = y$ or, if $x_1 < 1 - \epsilon$, possibly at $x = 1 - \epsilon$. At $x = y$ we have

$$\mathbb{P}(\varphi(1/2) \in [y, y+\epsilon] \mid \varphi(2^{-i}) = y) = \frac{\log^{i-1}(1+\epsilon/y)}{\log^{i-1}(y)}.$$

If $\epsilon/y < 2$ we estimate $\log(1+\epsilon/y) > \epsilon/2y > \epsilon/2$ and otherwise $\log(1+\epsilon/y) > 1 > \epsilon/2$, so in either case we get (5). At $x = 1 - \epsilon$,

$$\mathbb{P}(\varphi(1/2) \in [1-\epsilon, 1] \mid \varphi(2^{-i}) = y) \ge \epsilon \cdot \min_{1-\epsilon \le x \le 1}\left|\frac{(i-1)\cdot\log^{i-2}(x/y)}{x\cdot\log^{i-1}(y)}\right|$$

and the minimum, remembering $x_1 < 1 - \epsilon$, is achieved at $x = 1$, so it equals

$$\epsilon \cdot \frac{i-1}{|\log y|}$$

and again we get (5). ∎

In the following lemma and its proof the notation $\mathbb{P}(\cdot \mid \varphi(2^{-i}) = \varphi_i)$ is understood as a shorthand for $\lim_{\delta\to 0}\mathbb{P}(\cdot \mid |\varphi(2^{-i}) - \varphi_i| < \delta)$.

LEMMA 8: *Let $i \in \mathbb{N}$, let $\varphi_i \in [0, e^{-1}]$ and let τ be an increasing Lipschitz homeomorphism $[2^{-i}, 1] \to [\varphi_i, 1]$ with a constant K, i.e. $|\tau(x) - \tau(y)| < K|x-y|$, and let $0 < \epsilon < e^{-1}$. Then*

$$\mathbb{P}\big(\max_{2^{-i} \le x \le 1} |\varphi(x) - \tau(x)| < \epsilon \,\big|\, \varphi(2^{-i}) = \varphi_i\big) > c(K)^{1/\epsilon} \cdot \Big(\frac{c(K)\epsilon}{|\log \varphi_i|}\Big)^{C(K)i|\log \epsilon|}.$$

This lemma is a variation on Lemma 4.1 from [KO98], in which only the $c(K)^{1/\epsilon}$ factor appeared. Think of $c(K)^{1/\epsilon}$ as the "main term", with the other factor meaningful only for "unusual" cases where φ_i is very small or i is very large.

Proof: It is clearly enough to consider $\epsilon = (K+2)/2^q$ where q is some integer. For any $s \le q$ we denote

$$A_s := \{\varphi\colon |\varphi(j2^{-s}) - \tau(j2^{-s})| < 2^{-q},\ \forall 2^{s-i} < j < 2^s\}.$$

Let us estimate $\mathbb{P}(A_s | A_{s-1}, \varphi(2^{-i}) = \varphi_i)$. Denote j^* to be the minimal $j > 2^{s-i}$. If $\varphi \in A_{s-1}$ with some $s \le q$ then, for any odd $j \ne j^*$, the probability of the event

$$|\varphi(j2^{-s}) - \tau(j2^{-s})| < 2^{-q}$$

can be estimated from below by

$$\frac{2^{-q}}{|\varphi((j+1)2^{-s}) - \varphi((j-1)2^{-s})|} \ge \frac{2^{-q}}{K2^{-s+1} + 2^{-q+1}} \tag{8}$$

and for different j's these events are independent. This estimate also holds for $j = j^*$ if $s > i$, and if $s = i$, j^* is even and therefore irrelevant. Otherwise, for $j = j^*$ use Lemma 7 and the scaling invariance of φ to get

$$\begin{aligned} &\mathbb{P}(|\varphi(j^*2^{-s}) - \tau(j^*2^{-s})| < 2^{-q} \,\big|\, |\varphi(2^{-i}) = \varphi_i) \\ &> \Bigg(\frac{2^{-q}}{2\max(1, |\log(\frac{\varphi_i}{\varphi((j^*+1)2^{-s})})|)}\Bigg)^{i-s} > \Big(\frac{2^{-q}}{2|\log \varphi_i|}\Big)^i. \end{aligned} \tag{9}$$

Summing (8) and (9) (and replacing s with $s+1$) we get

$$\mathbb{P}(A_{s+1} \,|\, A_s, \varphi(2^{-i}) = \varphi_i) > (K2^{q-s} + 2)^{-2^s} \cdot \Big(\frac{2^{-q-1}}{|\log \varphi_i|}\Big)^i,$$

so

$$\begin{aligned} \mathbb{P}(A_q \,|\, \varphi(2^{-i}) = \varphi_i) &> \Big(\frac{2^{-q-1}}{|\log \varphi_i|}\Big)^{iq} \cdot \prod_{s=0}^{q-1} (K+2)^{-2^s} 2^{-(q-s)2^s} \\ &> \Big(\frac{2^{-q-1}}{|\log \varphi_i|}\Big)^{iq} \cdot \exp\Big(-2^q \log(K+2) - 2^q \log 2 \sum_{j=1}^{\infty} \frac{j}{2^j}\Big) \\ &> \Big(\frac{c(K)\epsilon}{|\log \varphi_i|}\Big)^{C(K)i|\log \epsilon|} \cdot \exp\Big(-\frac{C(K)}{\epsilon}\Big) \end{aligned}$$

and clearly A_q implies $||\varphi - \tau|| < \epsilon$. ∎

For the following lemma, we fix $n \in \mathbb{N}$ large enough and $s_0 \le 1$ and inspect the function

$$g(x) := \begin{cases} f_n(xs_0^{-1}), & x < s_0, \\ 0, & \text{otherwise;} \end{cases}$$

we define $s_1 = s_0 n^{-e^{n^4}}$ so that $\operatorname{supp} g = [s_1, s_0]$.

Lemma 9: *Let $i > j \in \mathbb{N}$ and $s_1 < \varphi_i < \varphi_j < s_0$ satisfy $4n < 2^{i-j} < n^{K_1}$ and $n^2 < \varphi_j/\varphi_i < n^{K_2}$ and let us define the event*

$$A_{i,j} := \{\varphi(2^{-i}) = \varphi_i\} \cap \{\varphi(2^{-j}) = \varphi_j\}.$$

Then

$$\mathbb{P}\left(\exists I \subset [2^{-i}, 2^{-j}],\ r \in \mathbb{N}\colon\ \int_I (g \circ \varphi) \cdot D_r > c \log n \,\Big|\, A_{i,j}\right) > e^{-C(K_1,K_2)n^3}.$$

Note that the above event is in $\Omega_{i,j}$.

Proof: The conditions on φ_i and φ_j imply that for at least one k,

$$[s_0 n^{-k}, s_0 n^{-k+1}] \subset [\varphi_i, \varphi_j]. \tag{10}$$

Define k to be the least one satisfying (10);

$$r = (2n)2^j; \quad \alpha = \frac{6}{2r+1}, \quad \beta = \frac{2n-4}{2r+1}, \quad I = [\alpha, \beta];$$

and let us consider the piece-linear homeomorphism $\tau\colon [2^{-i}, 2^{-j}] \to [\varphi_i, \varphi_j]$ defined by

$$\tau(2^{-i}) = \varphi_i, \quad \tau(\alpha) = (3 - \psi_k)s_0 n^{-k},$$
$$\tau(2^{-j}) = \varphi_j, \quad \tau(\beta) = (n - 1 - \psi_k)s_0 n^{-k}.$$

These values were, of course, chosen to ensure $g \circ \tau|_I \equiv \sin((2r+1)\pi x)$. Simple algebra shows

$$\tau' < C\frac{\varphi_j - \varphi_i}{2^{-j} - 2^{-i}} < C\frac{\varphi_j}{2^{-j}}.$$

Combining this, lemma 8 and the scaling invariance of φ gives

$$\mathbb{P}\Big(\max_{2^{-i} \le x \le 2^{-j}} |\varphi(x) - \tau(x)| > \epsilon \,\Big|\, A_{i,j}\Big) > c^{\varphi_j/\epsilon} \cdot \left(\frac{c\epsilon/\varphi_j}{|\log(\varphi_i/\varphi_j)|}\right)^{-C\cdot(i-j)|\log \epsilon/\varphi_j|}.$$

Taking $\epsilon = n^{-3}\varphi_j$ will give

$$\varphi(\alpha) > \tau(\alpha) - \epsilon > s_0 n^{-k}\left(2 - \frac{1}{n}\right)$$

and

$$\varphi(\beta) < \tau(\beta) + \epsilon < s_0 n^{-k}\Big(n - 1 + \frac{1}{n}\Big)$$

so

$$g \circ \varphi - \sin((2r+1)\pi x)|_I \le \epsilon \cdot \max_{\varphi I} g' = n^{-3}\varphi_j \cdot 2\pi s_0^{-1} n^k \le 2\pi n^{-1}. \tag{11}$$

The Dirichlet kernel D_r and $\sin((2r+1)\pi x)$ are aligned in the sense that

$$\int_I \sin((2r+1)\pi x) \cdot D_r > c \log n,$$

and with (11),

$$\int_I (g \circ \varphi) \cdot D_r > c \log n - \frac{C \log n}{n} > c \log n$$

and the probability is

$$> c^{n^3} \cdot \Big(\frac{cn^{-3}}{K_2 \log n}\Big)^{C \cdot K_1 \log n \cdot \log n} > e^{-Cn^3 - C(K_1, K_2)\log^3 n} > e^{-C(K_1,K_2)n^3}. \qquad \blacksquare$$

This lemma is the "local" component of the proof of Theorem 2. The complement, the "global" component, is to show that for typical φ, many pairs i, j satisfying the above conditions exist.

LEMMA 10: *Let $0 < x \le y < 1$. The probability that $\varphi(2^{-i}) \in [0, x]$, where i is the smallest integer satisfying $\varphi(2^{-i}) \in [0, y]$, is x/y.*

Proof: Denote this event by $A_{x,y}$. Then

$$\begin{aligned}
\mathbb{P}A_{x,y} &= \sum_i \mathbb{P}((\varphi(2^{-i}) \le x) \wedge (\varphi(2^{-i+1}) > y)) \\
&= \sum_i \int_y^1 \mathbb{P}(\varphi(2^{-i}) \le x \mid \varphi(2^{-i+1}) = t) d\nu_i(t) \\
&= \sum_i \int_y^1 \frac{x}{t} d\nu_i(t) \\
&= \sum_i \int_y^1 \frac{x}{y}\mathbb{P}(\varphi(2^{-i}) \le y \mid \varphi(2^{-i+1}) = t) d\nu_i(t) \\
&= \frac{x}{y} \sum_i \mathbb{P}((\varphi(2^{-i}) \le y) \wedge (\varphi(2^{-i+1}) > y)) \\
&= \frac{x}{y}\mathbb{P}\{\exists i\colon (\varphi(2^{-i}) \le y) \wedge (\varphi(2^{-i+1}) > y)\} = \frac{x}{y}
\end{aligned}$$

where the measure ν_i is the distribution of $\varphi(2^{-i+1})$. $\blacksquare$

LEMMA 11: *For n sufficiently large, for the same g as above,*

$$\mathbb{P}\left\{\exists I,\ r\in\mathbb{N}\colon \int_I D_r\cdot(g\circ\varphi) > c\log n\right\} > 1-e^{-n}.$$

Proof: We use (2) for 2^{-d} when d is defined by $d := \lceil \frac{2}{K_2}\log_2 n\rceil$ and get

$$\mathbb{P}\{n^{-K_3} < \varphi(2^{-d}) < n^{-2}\} > 1 - Cn^{-2}.$$

We need intervals $I_k := [2^{-d(k+1)}, 2^{-dk}]$ such that $\varphi I_k \subset [s_1, s_0]$, so the first point is to show that many do exist. Lemma 10 ensures that for the random variable i_0 defined by

$$(\varphi(2^{-i_0}) \le s_0) \wedge (\varphi(2^{-i_0+1}) > s_0)$$

one has

$$\mathbb{P}\{\varphi(2^{-i_0}) < e^{-2n}s_0\} = e^{-2n}.$$

Denote this event by R_1. Next, define

$$i_1 := i_0 + d\left\lfloor \frac{1}{dK_1}(\log_2 n\cdot e^{n^4} - 2n\log_2 e)\right\rfloor$$

(K_1 from (2)), and using (2) and the scaling invariance of φ get

$$\mathbb{P}\left\{\frac{\varphi(2^{-i_1})}{\varphi(2^{-i_0})} < n^{-e^{n^4}}e^{2n}\right\} < C2^{2(i_0-i_1)} < Ce^{-2n}.$$

Denote this event by R_2. Between i_0 and i_1 we have $> c_1e^{n^4}$ intervals I_k. For each k we define the event

$$r_k := \neg\left\{n^2 < \frac{\varphi(2^{-dk})}{\varphi(2^{-d(k+1)})} < n^{-K_3}\right\}$$

so that $\mathbb{P}r_k < Cn^{-2}$ and the r_k's are independent. With these r_k's define the variable

$$X := \#\{k\colon I_k\subset[i_0,i_1]\wedge r_k\}.$$

Clearly, $\mathbb{E}X < c_1n^{-2}e^{n^4}$ so

$$\mathbb{P}\left\{X > \tfrac{1}{2}c_1e^{n^4}\right\} < Ce^{-\frac{1}{2}n^2} < Ce^{-2n}.$$

Denote this event by R_3. Finally, we can calculate our probability. If none of the R_i's happen, we have $> \frac{1}{2}c_1e^{n^4}$ intervals I_k satisfying the conditions of lemma 9. For each I_k, the behavior of $\varphi|_{I_k}|\varphi(\partial I_k)$ is independent for each k and lemma 9 gives an estimate of the probability

$$\mathbb{P}\left\{\exists I\subset I_k, r\colon \int_I D_r\cdot(g\circ\varphi) > c\log n\right\} > e^{-Cn^3}$$

(C depends on our K_2 and K_3, but is still a constant). Totally we get

$$\mathbb{P}\neg\left\{\exists I,\, r\in\mathbb{N}\colon \int_I D_r\cdot(g\circ\varphi)>c\log n\right\}<$$
$$(1-e^{-Cn^3})^{(\frac{1}{2}c_1 e^{n^4})}+\mathbb{P}(R_1\cup R_2\cup R_3)<Ce^{-2n}$$

and the lemma is proved. ∎

Proof of Theorem 2: Define values s_n and functions g_n as follows:

$$s_{n+1}:=s_n\cdot\frac{1}{4}n^{-e^{n^4}},$$
$$g_n(x):=\begin{cases} f_n(xs_n^{-1}), & 4s_{n+1}<x<s_n,\\ \text{linear}, & x\in[2s_{n+1},4s_{n+1}]\cup[s_n,2s_n],\\ 0, & \text{otherwise}\end{cases}$$

(take $s_0=\frac{1}{4}$) with the relevant ψ's and the linear portions chosen to make g_n continuous. Now pick a sequence $n_k\to\infty$ fast enough as to satisfy, for all k,

$$\mathbb{P}\left\{\exists r\colon (|S_r(g_{n_k}\circ\varphi;0)|>1)\wedge\left(\sum_{l\neq k}|S_r(g_{n_l}\circ\varphi;0)|>\frac{1}{k}\right)\right\}<\frac{1}{k} \tag{12}$$

(this is possible since $S_r(g_n\circ\varphi;0)\to 0$ when $r\to\infty$ for any fixed n and when $n\to\infty$ for any fixed r). Now define

$$f:=\sum_k\frac{1}{\log n_k}g_{n_k}.$$

Clearly, $\omega_f(0;\delta)=O(\log\log\log\frac{1}{\delta})^{-1}$. On the other hand, lemma 11 ensures that for sufficiently large n,

$$\mathbb{P}\left\{\exists I\subset[s_{n+1},s_n],\ r\in\mathbb{N}\colon \int_I D_r\cdot(g_n\circ\varphi)>c_1\log n\right\}>1-e^{-n} \tag{13}$$

but $\int_x^y>c_1\log n$ implies that either $\int_0^x$ or $\int_0^y>\frac{1}{2}c_1\log n$. Pick any r_1 sufficiently large to allow the restriction $r\in[1,r_1]$ in (13), and combine this with lemma 6 to get

$$\mathbb{P}\{\exists r\in\mathbb{N}:|S_r(g_n\circ\varphi;0)|>\tfrac{1}{4}c_1\log n\}>1-e^{-n}-C/\log n,$$

and for $n=n_k$, again sufficiently large, using (12) this event implies

$$\mathbb{P}\left\{\sum_{l\neq k}|S_r(g_{n_l};0)|>1/k\right\}<1/k$$

so

$$\mathbb{P}\Big\{\exists r \in \mathbb{N}\colon |S_r(f\circ\varphi;0)| > \tfrac{1}{4}c_1 - \frac{1}{k}\Big\} > 1 - e^{-n_k} - C/\log n_k - 1/k$$

and taking $k \to \infty$ (which clearly forces $r \to \infty$) the theorem is done. ∎

Remark: Merely changing the $1/\log n_k$ factors in the proof above, one may get a number of other examples of divergence:

1. For every $\omega(\delta) = \Omega(\log\log\log\frac{1}{\delta})^{-1}$, a continuous function f which satisfies $\omega_f(\delta;0) = o(\omega(\delta))$, and $S_n(f\circ\varphi;0)$ is almost surely (i.e. with probability 1) unbounded.
2. For every $\omega(n) = o(\log\log\log n)$, a continuous function f for which one has $S_n(f\circ\varphi;0) > \omega(n)$ for infinitely many n's almost surely.
3. An L^∞ function f satisfying $S_n(f\circ\varphi;0) > \log\log\log n$ for infinitely many n's almost surely.

5. The 0-1 law

Our aim in this section is to prove claims of the type "For any f, the probability that the Fourier expansion of $f\circ\varphi$ converges uniformly (or pointwise, or in 0, or ...) is either 0 or 1". As hinted in [KO98] on page 1037, the first step is to transform the desired property into an "interval property", for example, to remark that probabilistically, the property

$$\sup_{\substack{n>0\\ t\in[0,1]}} \left| \int_{[0,1]} D_n(t-x)f(x)dx \right| < C$$

(i.e. $f \in U_0$, the set of functions with uniformly bounded Fourier partial sums) is equivalent to

$$\sup_{\substack{I\subset[0,1]\\ n>0\\ t\in[0,1]}} \left| \int_I D_n(t-x)f(x)dx \right| < C. \tag{14}$$

Denote this set of functions with uniformly bounded "interval Fourier partial sums" by $\tilde{U}$. That $\mathbb{P}(f\circ\varphi \in U_0) = \mathbb{P}(f\circ\varphi \in \tilde{U})$ was shown in [KO98] in the corollary to Lemma 4.5, and $f \in \tilde{U}$ is an interval property, in the following sense:

Definition 1: A map $T(f;I) \to \{0,1\}$, where $f \in C(\mathbb{T})$ is a function and $I \subset [0,1]$ is an interval, is called an interval property if the following conditions hold:

(i) T considered as a map $C(\mathbb{T}) \times [0,1]^2 \to \{0,1\}$ is Borel measurable.
(ii) $f|_I = g|_I$ a.e. $\Rightarrow T(f;I) = T(g;I)$.
(iii) $T(f;[x,y]) = T(f;[x,t])T(f;[t,y])$ whenever $x \le t \le y$; $T(f;[x,x]) = 1$.

(iv) $T(f\circ L; L^{-1}(I)) = T(f;I)$ for any linear map L.

Denote $T(f) := T(f;[0,1])$. When we say that $f\in\tilde{U}$ is an interval property we mean that the map defined by

$$T(f;I)=1 \Leftrightarrow \sup_{\substack{J\subset I\\ n>0\\ t\in[0,1]}} \left|\int_J D_n(t-x)f(x)dx\right| < C$$

is an interval property. Property (iii) of this T is clear (here the difference between the classes U_0 and $\tilde{U}$ is crucial). For property (iv), standard arguments* show that the above is equivalent to

$$\sup_{\substack{J\subset I\\ \alpha\in[0,\infty)\\ t\in\mathbb{R}}} \left|\int_J \frac{\sin\alpha(t-x)}{t-x} f(x)dx\right| < C$$

for which (iv) is clear.

THEOREM 3: *If $T(f;I)$ is an interval property and f is any function, then*

$$\mathbb{P}(T(f\circ\varphi)=1)\in\{0,1\}.$$

Proof: Let us discuss the following function, defined on $\{0\le x\le y\le 1\}$:

$$p(x,y) := \mathbb{E}(T(f\circ L_{[x,y]}\circ\varphi))$$

where L_I is the linear increasing mapping of $[0,1]$ onto I. The analysis of p will be based on one equality, (15) below, which we will now prove.

$$\begin{aligned} p(x,y) &= \mathbb{E}\mathbb{E}(T(f\circ L_{[x,y]}\circ\varphi;[0,\tfrac12])T(f\circ L_{[x,y]}\circ\varphi;[\tfrac12,1])\mid \varphi(\tfrac12)=t)\\ &= \int_0^1 \mathbb{E}(T(f\circ L_{[x,y]}\circ\varphi;[0,\tfrac12])\mid \varphi(\tfrac12)=t)\cdot\\ &\qquad\cdot\mathbb{E}(T(f\circ L_{[x,y]}\circ\varphi;[\tfrac12,1])\mid \varphi(\tfrac12)=t)dt. \end{aligned}$$

We now note that

$$\begin{aligned} T(f\circ L_{[x,y]}\circ\varphi;[0,\tfrac12]) &= T(f\circ L_{[x,y]}\circ\varphi\circ L_{[0,\frac12]})\\ &\sim T(f\circ L_{[x,y]}\circ(t\varphi))\\ &= T(f\circ L_{[x,x+t(y-x)]}\circ\varphi) \end{aligned}$$

* For example, one might show that the difference between the two kernels (where $n=\lfloor\alpha\rfloor$) is uniformly bounded.

and similarly

$$T(f\circ L_{[x,y]}\circ\varphi;[\tfrac{1}{2},1])\sim T(f\circ L_{[x+t(y-x),y]}\circ\varphi),$$

so

$$p(x,y)=\int_0^1 p(x,x+t(y-x))p(x+t(y-x),y)dt$$

or, after a change of variable,

$$p(x,y)=\frac{1}{y-x}\int_x^y p(x,t)p(t,y)dt. \tag{15}$$

It might be worth noting that the measurability requirement on T is used only to ensure that p is well defined and measurable on $[0,1]^2$. Thus weaker properties might also do.

First, a technical lemma.

LEMMA 12: *If $h(x)$ is a bounded function satisfying, for every $x<y_0$,*

$$h(x)\le\frac{1}{y_0-x}\int_x^{y_0}h(s)ds,$$

then for every $x<t<y_0$,

$$h(x)\le\frac{1}{y_0-t}\int_t^{y_0}h(s)ds.$$

Proof: If not, define

$$s_0:=\sup\{s:s<t,\ h(s)\ge h(x)\}$$

and let $s_n\to s_0$ be a series satisfying $h(s_n)\ge h(x)$ (not necessarily different from s_0). We have, for n sufficiently large,

$$h(x)\le h(s_n)\le\frac{1}{y_0-s_n}\int_{s_n}^{y_0}h(s)ds$$

so

$$h(x)\le\frac{1}{y_0-t+s_0-s_n}\left(\int_{s_n}^{s_0}+\int_t^{y_0}\right)h(s)ds$$

and, taking $n\to\infty$, the lemma is proved. ∎

LEMMA 13: *A measurable function $0 \le p(x,y) \le 1$ satisfying (15) is decreasing in y almost everywhere.*

Proof: For $x < y < z$, denote

$$\Delta(x,y,z) := p(x,z) - p(x,y),$$
$$\Delta(x,y) := \operatorname*{ess\,sup}_{z \ge y} \Delta(x,y,z),$$

and define

$$\mu := \operatorname{ess\,sup} \Delta(x,y)$$

and assume to the contrary that $\mu > 0$. Δ satisfies the following:

$$\begin{aligned}\Delta(x,y,z) &= \frac{1}{z-x}\left(\int_x^y p(x,t)\Delta(t,y,z)dt + \int_y^z p(x,t)p(t,z) - p(x,y)dt\right)\\ &\le \frac{1}{z-x}\left(\int_x^y p(x,t)\Delta(t,y)dt + \int_y^z \Delta(x,y,t)dt\right).\end{aligned}$$

We now iterate this inequality. The second iteration looks like

$$\begin{aligned}\Delta(x,y,z) \le&\frac{1}{z-x}\left(\int_x^y p(x,t)\Delta(t,y)dt\right.\\ &\left.+\int_y^z \frac{1}{t-x}\left(\int_x^y p(x,s)\Delta(s,y)ds + \int_y^t \Delta(x,y,s)ds\right)dt\right)\\ =&\frac{1}{z-x}\left(\int_x^y p(x,t)\Delta(t,y)\left(1+\int_y^z \frac{ds}{s-x}\right)dt\right.\\ &\left.+\int_y^z \Delta(x,y,t)\left(\int_t^z \frac{ds}{s-x}\right)dt\right)\\ =&\frac{1}{z-x}\left(\int_x^y p(x,t)\Delta(t,y)\left(1+\ln\left(\frac{z-x}{y-x}\right)\right)dt\right.\\ &\left.+\int_y^z \Delta(x,y,t)\ln\left(\frac{z-x}{t-x}\right)dt\right)\end{aligned}$$

and, similarly, the nth iterate looks like

$$\begin{aligned}\Delta(x,y,z) \le&\frac{1}{z-x}\left(\int_x^y p(x,t)\Delta(t,y)\sum_{k=0}^{n-1}\frac{1}{k!}\ln^k\left(\frac{z-x}{y-x}\right)dt\right.\\ &\left.+\frac{1}{(n-1)!}\int_y^z \Delta(x,y,t)\ln^{n-1}\left(\frac{z-x}{t-x}\right)dt\right)\end{aligned}$$

and, when n tends to infinity, the second term vanishes ($|\Delta| \le 1$) and we are left with

$$\Delta(x,y,z) \le \frac{1}{y-x}\int_x^y p(x,t)\Delta(t,y)dt,$$

which is true for all z. So

$$\Delta(x,y) \leq \frac{1}{y-x}\int_x^y p(x,t)\Delta(t,y)dt. \tag{16}$$

Next, fix some small $\epsilon > 0$ and get from Lebesgue's density theorem the existence of a square $[x_0, x_0+\delta]\times[y_0, y_0+\delta]$ where $\Delta > \mu - \epsilon$ on a set of measure $> 0.9\delta^2$; and we may also assume that $x_0 + \delta < y_0$ and that $\Delta(x_0, y_0+\delta) > \mu - \epsilon$. Our contradiction will follow by examining the triangle

$$T := \{(t,y) \colon y_0 \leq t < y < y_0 + \delta\}.$$

Now, on one hand, we have a set $Y \subset [y_0, y_0+\delta]$, $\mathbf{m}Y > 0.9\delta$ of y's such that for each $y \in Y$ there exists an $x \in [x_0, x_0+\delta]$ satisfying $\Delta(x,y) > \mu-\epsilon$ and therefore using lemma 12 for $h(x) := \max\{0,\ \Delta(x,y)\}$ (ignoring, for the moment, the p in inequality (16)) gives

$$\frac{1}{y-y_0}\int_{y_0}^y \max\{0, \Delta(t,y)\}dt > \mu - \epsilon \quad \forall y \in Y.$$

This inequality for the average gives a simple measure estimate (assume $\epsilon < 0.1\mu$)

$$\mathbf{m}\{t \colon y_0 \leq t \leq y,\ \Delta(t,y) > \mu - 10\epsilon\} > 0.9(y-y_0) \quad \forall y \in Y \tag{17}$$

and on all of T

$$\mathbf{m}\{(t,y) \in T \colon \Delta(t,y) > \mu - 10\epsilon\} > 0.7\mathbf{m}T. \tag{18}$$

On the other hand, returning to (16) and inspecting p we get that $\Delta(t,y) > \mu-10\epsilon$ implies

$$\frac{1}{y-t}\int_t^y p(t,s)ds > 1 - \frac{10\epsilon}{\mu}$$

and, as before,

$$\mathbf{m}\left\{s \colon t \leq s \leq y,\ p(t,s) > 1 - \frac{100\epsilon}{\mu}\right\} > 0.9(y-t) \tag{19}$$

but $y_0 + \delta \in Y$, which can be combined with (17) and (19) to get

$$\mathbf{m}\left\{(t,s) \in T \colon p(t,s) > 1 - \frac{100\epsilon}{\mu}\right\} > 0.7\mathbf{m}T. \tag{20}$$

Finally, we return to the definition of Δ and note that $p(x,y) > 1 - c$ implies $\Delta(x,y) < c$, so we can combine (18) and (20) to conclude that for some $(x,y) \in T$,

$$\mu - 10\epsilon < \Delta(x,y) < 100\epsilon/\mu,$$

and since ϵ was arbitrary, the lemma is proved. ∎

Remark: The function

$$p(x,y) = \begin{cases} 0, & y = \frac{1}{2} \\ 1, & \text{otherwise} \end{cases}$$

satisfies (15) but is not monotone everywhere. Thus the "almost everywhere" in lemma 13 is not an artifact of the proof but a property of (15).

LEMMA 14: *A measurable function $0 \le p(x,y) \le 1$ satisfying (15) is increasing in x almost everywhere.*

Proof: Use Lemma 13 for $p'(x,y) := p(1-y, 1-x)$. ∎

We wish to avoid the complexities arising from the fact that p is monotone only almost everywhere. Luckily, all further operations will be pickings of certain values out of sets of positive measure. Thus, we can ignore the non-monotone triplets by simply redefining the notion of picking. Let us call a triplet $x < y < z$ good when $p(x,y) \le p(x,z)$ and $p(y,z) \le p(x,z)$; and a triplet x, y, z is good when it is good in the right order.

Definition 2: We say that we pick an x if x satisfies:

(i) For almost all x_2 and x_3, the triplet x, x_2, x_3 is good.
(ii) If x_2 has already been picked, then for almost all x_3, the triplet x_2, x, x_3 is good.
(iii) If x_2 and x_3 have already been picked, then the triplet x_2, x_3, x is good.

An induction on Lemmas 13 and 14 ensures that we can always pick out of every set of positive measure.

LEMMA 15: *For almost every $x < y < z$, $p(x,z) = p(x,y)p(y,z)$.*

Proof: Define

$$\Delta(x,y,z) := |p(x,z) - p(x,y)p(y,z)| \tag{21}$$

and assume to the contrary that

$$\mu := \operatorname{ess\,sup} \Delta(x,y,z) > 0. \tag{22}$$

Again, let $\epsilon > 0$ be arbitrary, and let

$$[x_0 - \tfrac{1}{2}\delta, x_0 + \tfrac{1}{2}\delta] \times [y_0 - \tfrac{1}{2}\delta, y_0 + \tfrac{1}{2}\delta] \times [z_0 - \tfrac{1}{2}\delta, z_0 + \tfrac{1}{2}\delta]$$

be a cube where $\Delta > \mu - \epsilon$ on a set of measure $> 0.99\delta^3$; and also assume $x_0 + \delta < y_0$, $y_0 + \delta < z_0$ and $\Delta(x_0, y_0, z_0) > \mu - \epsilon$. As before, we need a method to "push" x and z toward y. We start from the simple

$$\Delta(x,y,z) \leq \frac{1}{z-x}\left(\int_x^y p(x,t)\Delta(t,y,z)dt + \int_y^z \Delta(x,y,t)p(t,z)dt\right) \tag{23}$$

from which we can deduce

SUBLEMMA: *Assume* $\Delta(x_1, y_1, z_1) > \mu - \epsilon$ *with* $|x_1 - y_1| > 2\nu$ *and* $|z_1 - y_1| > 2\nu$. *Then there exist* x_2 *and* z_2 *such that*

(i) $x_2 < y_1 < z_2$;
(ii) $\Delta(x_2, y_1, z_2) > \mu - 4\epsilon$
(iii) $y_1 - x_2 < 2\nu$;
(iv) $z_2 - y_1 < 2\nu$;
(v) *either* $\nu < y_1 - x_2$ *or* $\nu < z_2 - y_1$.

Furthermore, if x_1, y_1 *and* z_1 *are picked in the sense of Definition 2 above then* x_2 *and* z_2 *are also picked.*

Proof of sublemma: Denoting

$$R_1 := \{(x, y_1, z_1) \colon x \in (x_1, y_1)\}, \quad R_2 := \{(x_1, y_1, z) \colon z \in (y_1, z_1)\}$$

and using (23) we get

$$\operatorname*{ess\,sup}_{(x,y,z)\in R_1\cup R_2} \Delta(x,y,z) > \mu - \epsilon.$$

Let us assume that $\operatorname{ess\,sup}_{R_2} > \mu - \epsilon$. The proof of the other case will be identical. We denote

$$M(x) := \operatorname*{ess\,sup}_{z\in[y_1,z_1]} \Delta(x, y_1, z)$$

(so that $M(x_1) > \mu - \epsilon$) and, using (23) again, we have

$$\Delta(x, y_1, z) \leq \frac{1}{z-x}\left(\int_x^{y_1} M(t)dt + \int_{y_1}^z M(x)dt\right),$$

so

$$M(x) \leq \frac{1}{y_1 - x}\int_x^{y_1} M(t)dt$$

and we can use lemma 12 for M to obtain

$$\frac{1}{2\nu}\int_{y_1-2\nu}^{y_1} M(t)dt > \mu - \epsilon.$$

We get a set $X_1 \subset (y_1 - 2\nu, y_1 - \nu)$ of positive measure with $x \in X_1$ satisfying $M(x) > \mu - 2\epsilon$, which implies a set of positive measure $X_2 \subset (y_1 - 2\nu, y_1 - \nu) \times (y_1, z_1)$ with $(x, z) \in X_2$ satisfying $(x, y_1, z) > \mu - 2\epsilon$. Let us pick a z_3 such that $X_3 := \{x\colon (x, z_3) \in X_2\}$ has a positive measure. If $z_3 < y_1 + 2\nu$, the lemma is proved — we denote $z_2 := z_3$, pick an x_2 out of X_3 and finish. Otherwise, we define

$$M_2(z) := \operatorname*{ess\,sup}_{x \in [y_1 - 2\nu, y_1]} \Delta(x, y_1, z)$$

and again use Lemma 12, this time for M_2, to get

$$\int_{y_1}^{y_1 + 2\nu} M_2(t)dt > \mu - 2\epsilon.$$

We complete the proof by picking $z_2 \in (y_1 + \nu, y_1 + 2\nu)$ with $M_2(z_2) > \mu - 4\epsilon$ and then picking an $x_2 \in (y_1 - 2\nu, y_1)$ satisfying $\Delta(x_2, y_1, z_2) > \mu - 4\epsilon$. ∎

Let us now complete the proof of Lemma 15. First we use the sublemma for x_0, y_0, z_0 and $\nu = \frac{1}{4}\delta$. Denote the resulting values by x_1 and z_1. Let us assume that $y_0 - x_1 > \frac{1}{4}\delta$ — it will be easy to verify that the same proof works in the second case. We return to (23) and observe that $\Delta(x_1, y_0, z_1) > \mu - 4\epsilon$ implies

$$\frac{1}{z_1 - x_1}\left(\int_{x_1}^{y_0} p(x_1, t)dt + \int_{y_0}^{z_1} p(t, z_1)dt\right) > 1 - \frac{4\epsilon}{\mu},$$

and thus we can pick a $t_1 \in [\frac{1}{2}x_1 + \frac{1}{2}y_0, y_0]$ satisfying $p(x_1, t_1) > 1 - 32\epsilon/\mu$. Denote now $I := (0.6x_1 + 0.4t_1,\ 0.4x_1 + 0.6t_1)$. Since

$$|I| = 0.2(t_1 - x_1) \geq 0.1(y_0 - x_1) > 0.025\delta$$

and since $I \subset [y_0 - \frac{1}{2}\delta, y_0 + \frac{1}{2}\delta]$, we can pick $y_1 \in I$ such that

$$\mathbf{m}\{(x, z) \in [x_0, x_0 + \delta] \times [z_0, z_0 + \delta]\colon \Delta(x, y_1, z) > \mu - \epsilon\} > 0.$$

This allows us to proceed and pick x_2 and z_2 satisfying

$$\Delta(x_2, y_1, z_2) > \mu - \epsilon.$$

We use the sublemma again, for x_2, y_1, z_2 and $\nu = 0.05\delta$. Denoting the output of the claim by x_3 and z_3 we are finally faced with the following situation:

$$\begin{aligned}
&x_1 < x_3 < y_1 < z_3 < t_1 < y_0 < z_1,\\
&p(x_1, t_1) > 1 - 32\epsilon/\mu,\\
&\Delta(x_3, y_1, z_3) > \mu - 4\epsilon.
\end{aligned}$$

This, however, is a contradiction to the assumption $\mu > 0$ since the monotonicity of p gives

$$p(x_3, y_1),\ p(y_1, z_3) > 1 - 32\epsilon/\mu,$$

so

$$\Delta(x_3, y_1, z_3) < 1 - (1 - 32\epsilon/\mu)^2 < 64\epsilon/\mu$$

and, since ϵ was arbitrary, μ must be zero. ∎

The fact that $p(x, y)$ is multiplicative only almost everywhere requires us to use a variation on the standard 0-1 law. The formulation follows:

LEMMA 16: *Let $\Omega = \prod \Omega_n$ be a (product) probability space and X a random variable defined on Ω such that for* almost every $\omega_1, \omega_1' \in \Omega_1, \ldots, \omega_n, \omega_n' \in \Omega_n$,

$$\mathbb{E}(X \mid \omega_1, \ldots, \omega_n) = \mathbb{E}(X \mid \omega_1', \ldots, \omega_n').$$

Then X is similar to a constant. In particular, if $X = \mathbf{1}_A$, then $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$.

The proof is identical to the proof of the standard 0-1 law — see, e.g., [K85, page 7].

Proof of Theorem 3: We want to use Lemma 16 with the independent variables $X_{n,k}$. Clearly, we may assume that the number of variables in the lemma is $2^N - 1$. Now, taking $\mathbb{E}(\cdot \mid \{X_{n,k} = \omega_{n,k}\})$ for $n = 1, \ldots, N$ and $1 < k < 2^n$ is identical to taking

$$\mathbb{E}(\cdot \mid \{\varphi(k2^{-N}) = s_k\}_{k=0}^{2^N})$$

(write $s_0 = 0$ and $s_{2^N} = 1$) and then

$$\begin{aligned}
\mathbb{E}(T(f \circ \varphi;&[0,1]) \mid \{\varphi(k2^{-N}) = s_k\}_{k=0}^{2^N}) \\
&= \mathbb{E}\Big(\prod_{k=0}^{2^N-1} T(f \circ \varphi; [k2^{-N}, (k+1)2^{-N}]) \mid \{\varphi(k2^{-N}) = s_k\}_{k=0}^{2^N} \Big) \\
&= \prod_k \mathbb{E}(T(f \circ \varphi; [k2^{-N}, (k+1)2^{-N}]) \mid \{\varphi(l2^{-N}) = s_l\}_{l=k,k+1}) \\
&= \prod_k \mathbb{E}(T(f \circ L_{[s_k, s_{k+1}]} \circ \varphi)) \\
&= \prod_k p(s_k, s_{k+1}) = p(0,1) \quad \text{for a.e. } \{s_k\}
\end{aligned}$$

and the theorem is proved. ∎

Theorem 3 can be applied to a number of harmonic properties of $f \circ \varphi$. Let us name a few, without proofs:

- Uniform convergence of $S_n(f \circ \varphi) \to f \circ \varphi$. One possible corresponding interval property is

$$T(f;I) = 1 \Leftrightarrow \forall J \subset I, \lim_{n\to\infty} \int_J D_n(x-t) \cdot (f(\varphi(t)) - f(\varphi(x)))dt = 0$$

  uniformly in x. Showing that this is probabilistically equivalent to $f \in U(\mathbb{T})$ is similar to the proof that U_0 is equivalent to $\tilde{U}$.
- Pointwise divergence on an infinite/uncountable/dense/second category set. All these properties (or their complements) are interval properties to begin with, so Theorem 3 applies directly.
- Pointwise convergence everywhere.
- Pointwise bounded Fourier partial sums.
- For any $\psi(n) \nearrow \infty$,

$$S_n(f \circ \psi) = o(\psi(n))$$

  uniformly or pointwise everywhere.

References

[DF65] L. E. Dubins and D. A. Freedman, *Random distribution functions*, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol II, part I (1965–6), 183–214.

[GMW86] S. Graf, R. D. Mauldin and S. C. Williams, *Random homeomorphisms*, Advances in Mathematics **60** (1986), 239–359.

[K83] J.-P. Kahane, *Quatre leçons sur les homéomorphismes du cercle et les séries de Fourier*, Proceedings of the International Seminar on Topics in Modern Harmonic Analysis (Torino/Milano 1982), pp. 955–990.

[K85] J.-P. Kahane, *Some Random Series of Functions*, second edition, Cambridge Studies in Advanced Mathematics **5**, Cambridge University Press, 1985.

[KO98] G. Kozma and A. M. Olevskiĭ, *Random homeomorphisms and Fourier expansions*, Geometric and Functional Analysis **8** (1998), 1016–1042.

[O81] A. M. Olevskiĭ, *Change of variable and absolute convergence of Fourier series*, Doklady Akademii Nauk SSSR **256** (1981), 284–288 (in Russian); Soviet Mathematics Doklady **23** (1981), 76–79 (English translation).

[O85] A. M. Olevskiĭ, *Modification of functions and Fourier series*, Uspekhi Matematicheskikh Nauk **40** (1985), 157–193 (in Russian); Russian Mathematical Surveys **40** (1985), 187–224 (English translation).

[Z59] A. Zygmund, *Trigonometric Series*, second edition, Cambridge University Press, 1959.