Concepts Behind Divergent Square Averages

Zoltán Buczolich[*], Department of Analysis, Eötvös Loránd University, Pázmány Péter Sétány 1/c, 1117 Budapest, Hungary
email: buczo@cs.elte.hu
www.cs.elte.hu/∼buczo

and

R. Daniel Mauldin[†], Department of Mathematics, University of North Texas, Denton, Texas 76203-1430, USA
email: mauldin@unt.edu
www.math.unt.edu/∼mauldin

April 18, 2006

Abstract

In this paper we discuss the heuristic background of our result about divergent squares which states that the sequence $\{k^2\}_{k=1}^{\infty}$ is L^1-universally bad. Our goal is to present the main ideas of the proof without getting distracted by some of the technical details.

1 Introduction

Throughout this paper, a dynamical system is a quadruple (X, Σ, μ, T) where (X, Σ, μ) is a separable non-atomic separable probability space and T is an

[*]Supported by the Hungarian National Foundation for Scientific Research T049727.
[†]Supported in part by NSF grant DMS 0400481.
2000 Mathematics Subject Classification: Primary 37A05; Secondary 28D05, 47A35.
Keywords: ergodic theorem, quadratic residue, maximal inequality

invertible measure-preserving transformation on X. By Birkhoff's Ergodic Theorem, given a dynamical system (X, Σ, μ, T) if $f \in L^1(X)$, then the ergodic averages

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^n(x)) \tag{1}$$

converge almost everywhere.

Results of Bourgain [2], [3], [4] imply that if $f \in L^p(\mu)$, for some $p > 1$, the ergodic means

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^{n^2}(x)) \tag{2}$$

converge almost everywhere. Bourgain also asked in [2], [5] whether this result is true for $p = 1$, that is for L^1 functions. This problem has received quite a bit of attention, see for example [8], [13]. In Section 6 of [1], V. Bergelson writes the following about it: "The case p = 1 is still open and is perhaps one of the central open problems in that branch of ergodic theory which deals with almost everywhere convergence".

In [6] we gave a negative answer to this question. Here we present some ways in which we attacked the problem and illustrate some points in our proof. Let us recall some concepts related to this problem.

Definition 1. [12] A sequence $\{n_k\}_{k=1}^{\infty}$ is L^1-*universally bad* if for all ergodic dynamical systems there is some $f \in L^1$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} f(T^{n_k} x)$$

fails to exist for all x in a set of positive measure.

To show that the squares are universally bad for L^1 we used a known equivalence. By the Conze principle and the Banach principle of Sawyer (see [7], [14], or [13]), a sequence $\{n_k\}_{k=1}^{\infty}$ is *not* L^1-universally bad if and only if there exists a constant $C < \infty$ such that for all systems (X, Σ, μ, T) and all $f \in L^1(\mu)$ we have the following weak $(1, 1)$ inequality for all $t > 0$

$$\mu\left(\left\{x : \sup_{N\geq 1} \left|\frac{1}{N} \sum_{k=1}^{N} f(T^{n_k} x)\right| > t\right\}\right) \leq \frac{C}{t} \int |f| d\mu. \tag{3}$$

(By $\int f d\mu$ we always mean the integral on the entire domain of f, that is, $\int_X f d\mu$.)

In [6] we proved that the sequence $\{k^2\}_{k=1}^{\infty}$ is L^1-universally bad by showing that for each given constant C there is a positive integer τ_0 such for the rational rotation, $T(x) = x + \frac{1}{\tau_0}$ modulo 1, the weak $(1,1)$ inequality fails to hold with the given C for some nonnegative integrable f and number t. This is certainly what one would expect if the sequence of squares is bad. But, how to produce such τ_0's and f's? One first basic idea is the following. Suppose for each $0 < \delta$ and $K, M \in \mathbb{N}$ we can construct (and up to some details we can do so):

(i) a nonnegative function $f \in L^1(\lambda)$,

(ii) $E \subset [0,1)$ with $\lambda(E) < \delta$,

(iii) pairwise independent, identically distributed random variables $X_1, ..., X_K$ with mean approximately $M2^{-M-1}$ and variance less than 2^{-M} such that if $x \notin E$, then there is some N_x such that

$$A_{N_x} f(x) := \frac{1}{N_x} \sum_{k=1}^{N_x} f(T^{k^2} x) > \sum_{h=1}^{K} X_h(x)$$

(iv) $\int f \, d\lambda < K \cdot 2^{-M+2}$.

The leakage process and the properties of the distribution of squares mod n for particular n's allows us to (more or less) carry out such a construction.

Before discussing these, let us recall how the proof is completed. For each $p \in \mathbb{N}$, set $M = M_p = 4^p$ and take $\delta = \frac{1}{p}$. For each K, the expected value of $\sum_{h=1}^{K} X_h$ is approximately $\frac{KM}{2^{M+1}}$ and the variance is less than $\frac{K}{2^M}$. Thus, by the weak law of large numbers,

$$\lambda\left(\left\{x : \left|\sum_{h=1}^{K} X_h(x) - \frac{KM}{2^{M+1}}\right| \geq \frac{KM}{2 \cdot 2^{M+1}}\right\}\right) \leq \frac{4 \cdot 2^{M+2}}{KM^2}.$$

So, for each p, take $K = K_p$ so large that for $M = M_p$,

$$\lambda\left(\left\{x : \sum_{h=1}^{K} X_h(x) \geq \frac{KM}{2 \cdot 2^{M+1}}\right\}\right) \geq 1 - \frac{1}{p}. \tag{4}$$

Take $t = t_p = \frac{KM}{4\cdot 2^{M+1}}$. From (4) and $\lambda(E) < \delta = 1/p$ it follows that $\lambda(\{x : A_{N_x}(f)(x) > t\}) \geq 1 - \frac{2}{p}$. Using this inequality with $f = f_p$ and $t = t_p$ as $p \to \infty$ the left hand side of (3) converges to 1 and on the right hand side we have

$$C\frac{\int f\ d\ \lambda}{t} < C\frac{K2^{-M+2}}{KM2^{-M-3}} = C\frac{32}{M},$$

which converges to 0 as p increases. So, there can be no such inequality.

The leakage and asymptotic distributional properties of squares will tell us that we might possibly take the random variables, X, to be M distributed by which we mean $X(x) \in \{0, 1, \frac{1}{2}, ..., 2^{-M+1}\}$ and $X = \frac{1}{2^{-l}}$ with probability 2^{-M+l-1}, for $l = 0, 1, ..., M-1$. But, because we use the asymptotic distributional properties of squares taken along square free n, we cannot exactly get M distributed random variables. In our paper we make the X's M-0.99 distributed, i.e., $X(x) \in \{0, 0.99, 0.99\cdot 2^{-1}, ..., 0.99\cdot 2^{-M+1}\}$ and $X = 0.99\cdot 2^{-l}$ with probability $0.99 \cdot 2^{-M+l-1}$.

Given δ, K, and M a general idea is to construct component functions f_h such that if $x \notin E$, then there is some N_x such that for all h

$$A_{N_x} f_h(x) := \frac{1}{N_x}\sum_{k=1}^{N_x} f_h(T^{k^2}x) > X_h(x) \tag{5}$$

and $\int f_h\ d\lambda < \Gamma' \cdot 2^{-M+2}$, where Γ' is some fixed constant. We can then take $f = f_1 + ... + f_K$. In Theorem 6 of [6] we prove the following result from which it follows as indicated earlier that the squares are bad.

Theorem 1. *Given $\delta > 0$, M and K there exist $\tau_0 \in \mathbb{N}$, $\overline{E}_\delta \subset [0, 1)$, a measurable transformation $T : [0, 1) \to [0, 1)$, $T(x) = x + \frac{1}{\tau_0}$ modulo 1, $f : [0, 1) \to [0, +\infty)$, $\overline{X}_h$, $h = 1, ..., K$ which are pairwise independent $M-0.99$-distributed random variables defined on $[0, 1)$ equipped with the Lebesgue measure, λ, such that $\lambda(\overline{E}_\delta) < \delta$, for all $x \in [0, 1) \setminus \overline{E}_\delta$ there exists N_x satisfying*

$$\frac{1}{N_x}\sum_{k=1}^{N_x} f(T^{k^2}(x)) > \sum_{h=1}^{K}\overline{X}_h(x), \tag{6}$$

and

$$\int_{[0,1)} f d\lambda < K \cdot 2^{-M+2}. \tag{7}$$

The proof of this theorem is quite complicated and in this note by omitting many technical details we try to help the reader to understand the heuristics behind it. Especially, we want to explain in detail the "leakage process", which is behind the most tricky part of the proof in [6]. This process is the simultaneous construction of the functions f_h and random variables X_h.

Since the sequence of integers $n_k = k - 1$ is not universally bad for this sequence there is C for which (3) holds for all dynamical systems.

Given $f : \mathbb{R} \to \mathbb{R}$, periodic by p we put

$$\overline{\int} f = \frac{1}{p} \int_0^p f(x) dx.$$

Given a Lebesgue measurable set A, periodic by p we put $\overline{\lambda}(A) = \frac{1}{p}\lambda(A \cap [0, p))$.

Remark 1. It is very easy to convert results obtained for periodic functions to results of dynamical systems. For example, if we have a function $f : \mathbb{R} \to \mathbb{R}$, periodic by $p \in \mathbb{N}$ such that for every $x \in \mathbb{R}$ there is N_x such that $(1/N_x) \sum_{k=1}^{N_x} f(x + k^2) > 1/2$ then one can consider the dynamical system $([0, 1), \mathcal{L}, \lambda, T)$, where λ is the Lebesgue measure, $\mathcal{L}$ is the sigma algebra of Lebesgue measurable sets, $T(x) = x + \frac{1}{p}$, modulo 1, and for the function $\overline{f}(x) = f(px)$ for every $x \in [0, 1)$ we have $(1/N_x) \sum_{k=1}^{N_x} \overline{f}(T^{k^2} x) > 1/2$, and $\int \overline{f} = \overline{\int} f$.

Similarly, after a suitable rescaling, results from probability theory will be applied to "random variables" periodic on $\mathbb{R}$.

The authors would like to thank Idris Assani, Christoph Thiele, and Andrew Yingst for their valuable comments and questions during the preparation of this article.

2 How to leak a function onto larger sets?

The technically most challenging part of the proof in [6] is a procedure called the leakage process. To understand it, first we discuss a leakage process when the sequence of the integers, $n_k = k - 1$ is considered instead of the squares $n_k = k^2$. Of course, this leakage will fail because the sequence $k - 1$ is good. But we will see where it fails and where the distribution of squares comes in. We suppose for the moment that $\tau_0 \in \mathbb{N}$ is given.

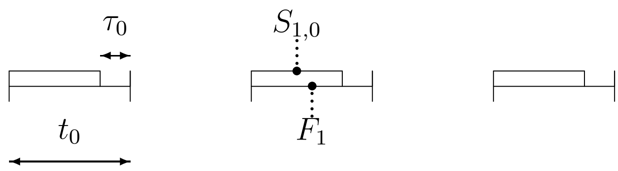

Figure 1: The sets appearing in Step 1

Step 0: Put $f_{1,0}(x) \equiv 1$ on $\mathbb{R}$, $S_{0,0} = \mathbb{R}$.

At step L of the leakage process we obtain a function $f_{1,L}$ and sets $S_{L,0}, ..., S_{L,L}$.

Step 1: Set $t_{-1} = 0$. Choose $\tau_0 \ll t_0 \in \mathbb{N}$. Set $F_1 = \cup_{k\in\mathbb{Z}}[2kt_0, (2k+1)t_0)$ and $f_{1,1}(x) = \chi_{F_1}(x)$. Then $\overline{\int} f_{1,1} = 1/2$. If $t_{-1} < n \leq \tau_0$ we have for most $x \in F_1$

$$\frac{1}{n}\sum_{k=0}^{n-1} f_{1,1}(x+k) = 1. \tag{8}$$

To be more precise, we have (8) for $x \in S_{1,0} \stackrel{\text{def}}{=} \cup_{k\in\mathbb{Z}} [2kt_0, (2k+1)t_0 - \tau_0)$. We also put $S_{1,1} = \mathbb{R}$. If $x \in S_{1,1}$, especially if $x \in S_{1,1} \setminus S_{1,0}$ and $n \gg t_0$, then

$$\frac{1}{n}\sum_{k=0}^{n-1} f_{1,1}(x+k) \approx \frac{1}{2}. \tag{9}$$

We say that in (9) the values of the function $f_{1,1}$ "leak" from its support to a larger set $S_{1,1}$, especially onto $S_{1,1} \setminus S_{1,0}$, when ergodic averages and the corresponding maximal operators are considered. In Step 2 we will explain why the word leakage is used here.

Step 2: Choose $\tau_1, t_1 \in \mathbb{N}$ such that $2t_0$ divides t_1 and $t_0 \ll \tau_1 \ll t_1$. Set $F_2 = F_1 \cap \cup_{k\in\mathbb{Z}}[2kt_1, (2k+1)t_1)$ and $f_{1,2}(x) = \chi_{F_2}(x)$. Then $\overline{\lambda}(F_2) = \overline{\int} f_{1,2} = \frac{1}{4}$. Set $S_{2,0} = S_{1,0} \cap \cup_{k\in\mathbb{Z}}[2kt_1, (2k+1)t_1)$, $S_{2,1} = \cup_{k\in\mathbb{Z}}[2kt_1, (2k+1)t_1 - \tau_1)$, and $S_{2,2} = \mathbb{R}$. Then $\overline{\lambda}(S_{2,0}) \approx 1/4$, $\overline{\lambda}(S_{2,1}) \approx 1/2$ and $\overline{\lambda}(S_{2,0}) = 1$. Setting $S_{2,-1} = \emptyset$ we also have $\overline{\lambda}(S_{2,0} \setminus S_{2,-1}) \approx 1/4 > 1/8$, $\overline{\lambda}(S_{2,1} \setminus S_{2,0}) \approx 1/4$ and $\overline{\lambda}(S_{2,2} \setminus S_{2,1}) \approx 1/2$.

If $x \in S_{2,0}$ and $t_{-1} < n \leq \tau_0$, then

$$\frac{1}{n}\sum_{k=0}^{n-1} f_{1,2}(x+k) = 1. \tag{10}$$

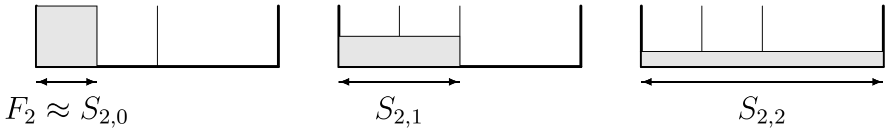

Figure 2: Leakage in two steps, the Reservoir Example

One should think of k and n as time. That is, we have a large value on a small set when averages computed for a small time.

If $x \in S_{2,1} \setminus S_{2,0}$ and $t_0 \ll n \leq \tau_1$, then

$$\frac{1}{n}\sum_{k=0}^{n-1} f_{1,2}(x+k) \approx \frac{1}{2}. \tag{11}$$

Now we have a medium value on a medium sized set, when medium time averages are computed.

If $x \in S_{2,2} \setminus S_{2,1}$ and $t_1 \ll n$, then

$$\frac{1}{n}\sum_{k=0}^{n-1} f_{1,2}(x+k) \approx \frac{1}{4}. \tag{12}$$

That is, here we have a small average on a large set when large time averages are computed.

If we set $X_1'(x) = 1$ on $S_{2,0}$, $X_1'(x) = 1/2$ on $S_{2,1} \setminus S_{2,0}$ and $X_1'(x) = 1/4$ on $S_{2,2} \setminus S_{2,1}$ then equations (10), (11) and (12) can be combined into one formula:

$$\frac{1}{n}\sum_{k=0}^{n-1} f_{1,2}(x+k) \approx X_1'(x), \tag{13}$$

for a suitable choice of n which depends on x. Also observe that $\overline{\lambda}\{x : X_1'(x) = 1\} \approx 1/4 > 1/8$, $\overline{\lambda}\{x : X_1'(x) = 1/2\} \approx 1/4$ and $\overline{\lambda}\{x : X_1'(x) = 1/4\} \approx 1/2$.

Next we discuss the Reservoir Example, which can help to understand and visualize why we talked about leakage in the above steps.

Suppose we have a reservoir filled with water which has volume $\overline{\int} f_{1,2} = 1/4$. This water is 1 unit deep and stands approximately over the set F_2

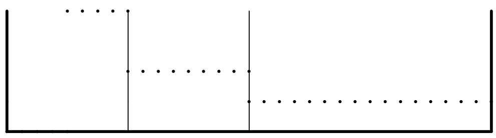

Figure 3: X_1 when $L = 3$

during the time interval $[t_{-1}, \tau_0]$, see the left side of Figure 2, (to simplify our example only the relative sizes of F_2, $S_{2,l}$, $l = 0, 1, 2$ are kept, but not the actual definitions of these sets given above). Now we have two dams, one in the middle of the reservoir, and another in the middle of the left half of the reservoir. First the one in the middle of the left half of the reservoir breaks, and the water leaks onto the left half of the reservoir as illustrated in the middle of Figure 2, here it stands 1/2 units high as long as time is much bigger than t_0 but less than τ_1. After some time the second dam breaks and the water leaks onto the whole reservoir and when time is much bigger than t_1 it stays 1/4 high everywhere. Time or the number of iterations in the construction of one function and one random variable progresses from left to right.

Continuing the "leakage," after M steps one can obtain sets F_M, $S_{M,-1} = \emptyset$, $S_{M,0}, ..., S_{M,M}$, $\tau_0 \ll t_0 \ll \tau_1 \ll t_1 \ll ... \ll \tau_{n-1} \ll t_{n-1}$ such that if $f_{1,M} = \chi_{F_M}$ then $\overline{\int} f_{1,M} = \overline{\lambda}(F_M) = 1/2^M$, $\overline{\lambda}(S_{M,l}) \approx 2^{-(M-l)}$ $(l = 0, ..., M)$. Moreover, if $x \in S_{M,l} \setminus S_{M,l-1}$, $(l = 0, ..., M)$ and $t_{l-1} \ll n \leq \tau_l$, $l = 0, ..., M-1$, for $l = M$ we use $\tau_M = \infty$ here), then we have

$$\frac{1}{n} \sum_{k=0}^{n-1} f_{1,M}(x+k) \approx \frac{1}{2^l}. \tag{14}$$

This also implies that if $X_1(x) \approx 1/2^l$ when $x \in S_{M,l} \setminus S_{M,l-1}$ (in fact, $X_1(x)$ should be a little less than $1/2^l$) then there exists n_x such that

$$\frac{1}{n_x} \sum_{k=0}^{n_x-1} f_{1,M}(x+k) > X_1(x). \tag{15}$$

Observe that $\overline{\lambda}(S_{M,l} \setminus S_{M,l-1}) \approx 2^{-M+l-1}$ for $l = 1, ..., M$ and $\overline{\lambda}(S_{M,0} \setminus$

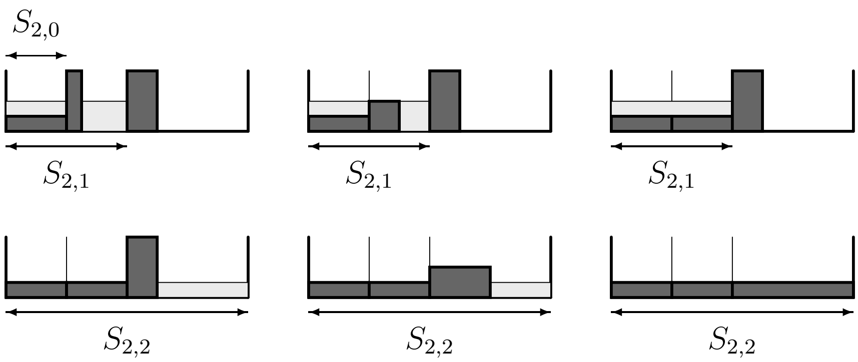

Figure 4: Leakage of two functions

$S_{M,-1}) \approx 2^{-M} > 2^{-M-1}$.

In order to keep our further computations simple assume that $X_1(x)$ is M-distributed. By this we mean that its range is $\{0, 2^{-M}, 2^{-M+1}, ..., 1\}$ and $\overline{\lambda}\{x : X_1(x) = 2^{-l}\} = 2^{-M+l-1}$, for $l = 0, ..., M$. To do so we need to redefine $X_1(x)$ to be 0 for approximately half of the x's from $S_{M,0}$.

In fact, as mentioned earlier, the X_1 we could use in (15) is only "approximately" M-distributed. One can visualize X_1 when $M = 2$ in the Reservoir Example on Figure 3, observe that on the first interval of length 1/8 we defined $X_1 = 0$ to keep it $M = 2$-distributed, though we could have defined $X_1 = 1$ on the whole first interval of length 1/4. We have more or less constructed an f and an X with the desired properties (i)-(iv) for $K = 1$.

Given the time windows for which (14) holds one might try to obtain other nonnegative functions $f_{h,M}$, and M-distributed X_h, $h = 1, ..., K$, all periodic by the same period, such that $\overline{\int} f_{h,M} = 1/2^M$, X_h is independent from $X_{h'}$ if $h \neq h'$ and for $x \in \mathbb{R}$ there exists n_x such that

$$\frac{1}{n_x} \sum_{k=0}^{n_x-1} f_{h,M}(x+k) > X_h(x) \tag{16}$$

for $h = 1, ..., K$.

For the L^1 good sequence of integers $n_k = k - 1$ the above choice of $f_{h,m}$ and X_h, as we have indicated earlier, would contradict the weak $(1, 1)$ inequality (3). However, if in (16) we use the square averages $\frac{1}{n_x} \sum_{k=1}^{n_x} f_{h,M}(x + k^2)$ then the analogous result holds.

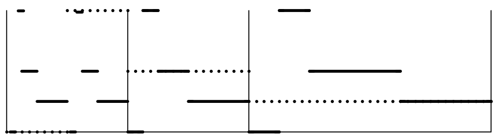

Figure 5: Independent X_1 and X_2

Next we want to illustrate by our Reservoir Example when $K = 2$ how can one think of the simultaneous leakage of two functions. See Figure 4. The passage of time or many iterations in the construction of our functions and variables is indicated from top to bottom, left to right. In the upper half we illustrate what happens during the time interval when the first leaking function, indicated by the lighter shaded area, stays in the left half of the reservoir. This corresponds with the period of time represented by the second frame of Figure 2. The darker shaded area corresponds to the leaking second function. One can see on the upper left image that the second function has already leaked over $S_{2,0}$, but has not started to leak on the sets $S_{2,1} \setminus S_{2,0}$ and $S_{2,2} \setminus S_{2,1}$. In the other two images on the upper half of Figure 4 one can see the second function gradually leaks over $S_{2,1} \setminus S_{2,0}$ while it stands still over $S_{2,2} \setminus S_{2,1}$. The leakage of the second function onto $S_{2,2} \setminus S_{2,1}$ is illustrated on the bottom half of Figure 4. We wait until the first function leaks over $S_{2,2} \setminus S_{2,1}$ and after this in three steps we allow the second function to leak also onto this set. Figures 4 and 5 illustrate the two component functions and two random variables generated in attempting to satisfy property (5). In Figure 5 we illustrate for the Reservoir Example independent X_1 (dotted line) and X_2 (thick line).

For the sequence of the integers, $n_k = k - 1$ this leakage procedure must break down since we know the the weak maximal inequality (3) holds. The main difficulty is that one cannot do a leakage process of the second (darker) function on $S_{2,1} \setminus S_{2,0}$ while keeping this function "unleaked" on $S_{2,2} \setminus S_{2,1}$ as it is illustrated on the upper half of Figure 4. On the other hand, for the sequence of the squares this is possible.

We can use this construction to show that no weak $(1, 1)$ inequality can hold for the sequence of the squares, that is, for $n_k = k^2$ the above argument can be used to show that this sequence is L^1-universally bad. This is exactly what happens at the end of Section 5, "Proof of the Main result" of [6].

3 $M - 0.99$ distributed variables and $K - M$ families

In the previous section for ease of computational details we dealt with M-distributed "random variables", X_h. Working with the squares we had to do some technical adjustments of this definition to obtain the definition of $M - 0.99$ distributed "random variables":

Definition 2. For a positive integer M we say that a periodic function or a "random variable", $X : \mathbb{R} \to \mathbb{R}$ is *conditionally* $M-0.99$ *distributed* on the set Λ, which is periodic by the same period, if

- $X(x) \in \{0, 0.99, 0.99 \cdot \frac{1}{2}, ..., 0.99 \cdot 2^{-M+1}\}$, and
- $\overline{\lambda}(\{x \in \Lambda : X(x) = 0.99 \cdot 2^{-l}\}) = 0.99 \cdot 2^{-M+l-1} \overline{\lambda}(\Lambda)$ for $l = 0, ..., M-1$.

(We regard $\mathbb{R}$ as being periodic by 1 with $\overline{\lambda}(\mathbb{R}) = 1$ and if $\Lambda = \mathbb{R}$ then we just simply say that X is $M-0.99$-distributed.)

By an obvious adjustment this definition can also be formulated for random variables $\overline{X}$ defined on $[0, 1)$ equipped with the Lebesgue measure λ. If we have two "random variables" X_1 and X_2, both conditionally $M-0.99$ distributed on Λ, then they are called pairwise independent (on Λ) if for any $y_1, y_2 \in \mathbb{R}$

$$\overline{\lambda}\{x \in \Lambda : X_1(x) = y_1 \text{ and } X_2(x) = y_2\}/\overline{\lambda}(\Lambda) = \qquad (17)$$
$$(\overline{\lambda}(\{x \in \Lambda : X_1(x) = y_1\})/\overline{\lambda}(\Lambda))(\overline{\lambda}(\{x \in \Lambda : X_2(x) = y_2\})/\overline{\lambda}(\Lambda)).$$

To construct the function f of Theorem 1 for which (6) and (7) hold for $K, M \in \mathbb{N}$, $\delta > 0$, $\Gamma > 1$ we had to find periodic functions $f_h : \mathbb{R} \to [0, \infty)$, and periodic pairwise independent $M - 0.99$ distributed random variables $X_h : \mathbb{R} \to \mathbb{R}$ for $h = 1, ..., K$ such that there is an exceptional set E_δ with $\overline{\lambda}(E_\delta) < \delta$ and for all $x \notin E_\delta$ there is a sufficiently large time interval $[\alpha(x), \omega(x)]$ and a number $\tau(x) \ll \omega(x) - \alpha(x)$ such that if $\alpha(x) \leq n < n + m \leq \omega(x)$ and $\tau(x)$ divides m then for all $h = 1, ..., K$

$$\frac{1}{m} \sum_{k=n}^{n+m-1} f_h(x + k^2) > X_h(x), \qquad (18)$$

and

$$\overline{\int} f_h < \Gamma \cdot 2^{-M+2}. \qquad (19)$$

In [6] in Definition 6 we define $K-M$ families living on sets $\Lambda \subset \mathbb{R}$ which are periodic by $\widetilde{q} \in \mathbb{N}$. The above paragraph contains the most important features of $K-M$ families living on $\Lambda = \mathbb{R}$. We omit here those technical details of the definition of $K-M$ families which are needed to accomodate the proof of existence of such families by induction on K, for the details we refer to [6]. In fact, the most difficult part of [6] is the proof of the existence of $K-M$ families, (Lemma 4 of [6]). Assuming that $K-M$ families exist on $\mathbb{R}$ in Lemma 5 of [6] we show that one can also obtain $K-M$ families which live on suitable periodic sets $\Lambda \subset \mathbb{R}$. To give a rough understanding what this means, one needs that the "random variables" X_h appearing in (18) are conditionally $M-0.99$ distributed on Λ and we have to replace (19) by

$$\overline{\int} f_h < \Gamma\gamma' \cdot 2^{-M+2}, \tag{20}$$

where the constant γ' approximately equals $\overline{\lambda}(\Lambda)$.

On the Reservoir Example on Figure 4 the set Λ may correspond to the set $S_{2,1} \setminus S_{2,0}$ and by Lemma 5 of [6] we are able to do the leakage of the second, darker function only onto this set as it is pictured on the upper half of Figure 4.

4 Definition of the sets $\overline{\Lambda}'(q)$

Next we need some definitions and facts related to number theory:

For each $q \in \mathbb{N}$ and $n \in \mathbb{Z}$ set $\varepsilon(n,q) = 1$ if n is congruent to a square modulo q, and let $\varepsilon(n,q) = 0$ if not. We denote by σ_q the number of squares modulo q. If p is an odd prime, then $\sigma_p = \frac{p+1}{2}$. If $q = p_1 \cdots p_\kappa$ where $p_1, ..., p_\kappa$ are distinct odd primes, then (by the fact that something is a square modulo q if and only if it is a square modulo each p_i plus by using the Chinese remainder theorem) $\sigma_q = \prod_{i=1}^{\kappa} \frac{p_i+1}{2}$. Since $\varepsilon(n,q) = 1$ when n is congruent to 0 modulo q we will regard 0 a quadratic residue (or square) in this paper. For the purpose of our proof it is sufficient to use square free numbers q, however some of the number theoretical background results hold for other numbers as well, see [9] and [10].

Put $\Lambda_0(q) = \{n : \varepsilon(n,q) = 1\}$, the set of quadratic residues modulo q. Clearly,

$$\#(\Lambda_0(q) \cap [0,q)) = \sigma_q > \frac{q}{2^\kappa}, \quad \text{with } \sigma_q \approx \frac{q}{2^\kappa}. \tag{21}$$

For later arguments we now introduce some parameters. We will need a suitably small "leakage constant" $\gamma \in (0,1)$ of the form $\gamma = 2^{-c_\gamma}$ where $c_\gamma \in \mathbb{N}$. Then we work with $\kappa > c_\gamma$.

Definition 3. Assume $q = p_1 \cdots p_\kappa$, where $p_1 < ... < p_\kappa$ are odd primes. Set $\Lambda_\gamma(q) = -\Lambda_0(q) + \{j \in \mathbb{Z} : 0 \le j < \gamma 2^\kappa\}$, $\overline{\Lambda}_\gamma(q) = \Lambda_\gamma(q) + [0,1) = -\Lambda_0(q) + \{x : 0 \le x < \gamma 2^\kappa\}$.

Moreover, we put

$$\Lambda'_0(q) = \{n \in \Lambda_0(q) : p_j \not| n, \text{ for all } j = 1, ..., \kappa\},$$

$$\Lambda'_\gamma(q) = -\Lambda'_0(q) + \{j \in \mathbb{Z} : 0 \le j < \gamma 2^\kappa\},$$

$$\overline{\Lambda}'_\gamma(q) = \Lambda'_\gamma(q) + [0,1) = -\Lambda'_0(q) + \{x : 0 \le x < \gamma 2^\kappa\}.$$

The main motivation for the definition of $\Lambda'_0(q)$ and other objects marked by $'$ is that if $n \in \Lambda'_0(q)$, then there are exactly 2^κ many solutions of $x^2 \equiv n \mod q$ and this makes some computations much easier in [6].

For fixed κ

$$\text{if } p_1 \to \infty, \text{ then } \frac{\#(\Lambda_0(q) \cap [0,q))}{\#(\Lambda'_0(q) \cap [0,q))} \to 1. \tag{22}$$

This means that we do not make too much mistake by using $\Lambda'_0(q)$ instead of $\Lambda_0(q)$ when p_1 is large.

For ease of notation in the sequel if we have a fixed γ and we do not want to emphasize the dependence on γ we will write $\Lambda(q)$, $\overline{\Lambda}(q)$, $\Lambda'(q)$ and $\overline{\Lambda}'(q)$ instead of $\Lambda_\gamma(q)$, $\overline{\Lambda}_\gamma(q)$, $\Lambda'_\gamma(q)$ and $\overline{\Lambda}'_\gamma(q)$, respectively.

The big difference between the sequence of the squares and of the integers is that by Lemma 5 of [6] if $\widetilde{q}$ is a suitably chosen highly composite square free number then there are $K-M$ families living on $\Lambda = \overline{\Lambda}'(\widetilde{q})$ while if in (18) the sequence $n_k = k$ is used instead of $n_k = k^2$ then an analogous statement concerning suitable Λ's does not hold.

5 The leakage process of one function onto sets of $\overline{\Lambda}'(q)$

Next we want to illustrate/explain how to do leakage on a residue class.

Assume a leakage constant $\gamma \in (0,1)$ is fixed.

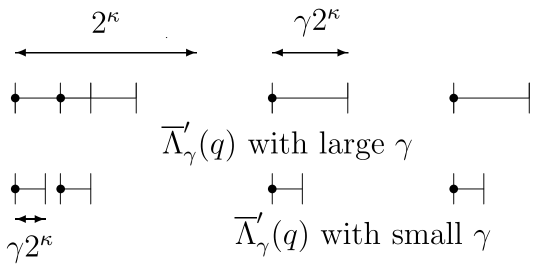

Figure 6: $\overline{\Lambda}'(\gamma)$ for different γ's

Take $f_{0,0}(x) \equiv 1$ on $\mathbb{R}$. Set $F_0 = S_{0,0} = \mathbb{R}$.

Suppose we have τ_0 given. Choose κ_1 such that $\gamma 2^{\kappa_1} \gg \tau_0$. Suppose $q_1 = p_{1,1} \cdots p_{\kappa_1,1}$ with $p_{1,1} < \ldots < p_{\kappa_1,1}$. If κ_1 and p_1 are sufficiently large then the average gap length between consecutive points of $\Lambda_0(q_1)$ is approximately $2^{\kappa_1} \gg \tau_0$.

By results of [9] we have property

(†) *the distribution of the normalized difference between consecutive elements of* $\Lambda_0(q_1)$ *approximates a Poisson distribution, with exponential limiting distribution as* $\kappa_1 \to \infty$.

From our point of view the most important feature of the above property is that the limiting distribution is continuous.

Set $F_1 = \mathbb{R} \setminus \overline{\Lambda}'(q_1)$ and $f_{1,1}(x) = \chi_{F_1}(x)$. Then $\overline{\Lambda}'(q_1)$ is the union of, possibly overlapping, intervals of length $\gamma 2^{\kappa_1}$. Due to these overlaps $\overline{\lambda}(\overline{\Lambda}'(q_1))$ turns out to be a little less than γ, however, given the fact that we have property (†), if γ is small, κ_1 and $p_{1,1}$ are large then $\overline{\lambda}(\overline{\Lambda}'(q_1))/\gamma$ is close to one. This is illustrated on Figure 6 where the thick dots correspond to elements of $-\Lambda_0'(q)$. The fact that for larger γ's we have some unwanted loss due to overlaps forces us to use smaller γ's in our leaking process over residue classes.

If $[x, x+\tau_0] \subset F_1$ and $n^2 < \tau_0$ then

$$\frac{1}{n} \sum_{k=1}^{n} f_{1,1}(x+k^2) = 1. \tag{23}$$

We want to obtain an estimate analogous to (9). In order to obtain this estimate we need that, at least for most $x \in \overline{\Lambda}'(q_1)$, we have for large n's

$$\frac{1}{n}\sum_{k=1}^{n} f_{1,1}(x+k^2) = \frac{1}{n}\sum_{k=1}^{n} \chi_{F_1}(x+k^2) \approx \tag{24}$$
$$\overline{\lambda}(F_1) \approx \overline{\lambda}(\mathbb{R}\setminus\overline{\Lambda}'(q_1)) \approx (1-\gamma),$$

when γ is small, κ_1 and $p_{1,1}$ are large. This is the motivation for Lemma 2 in [6].

Set $S_{1,0} = \mathbb{R}\setminus\overline{\Lambda}'(q_1) = F_1$ and $S_{1,1} = \mathbb{R}$. We can find a small exceptional set E_1 such that for each $x \in \mathbb{R}\setminus E_1$ there is a time interval $[\alpha_1(x), \omega_1(x)]$ and a number $\tau_1(x) \ll \omega_1(x) - \alpha_1(x)$ for which if $x \in S_{1,0}\setminus E_1$, $n, n+m \in [\alpha_1(x), \omega_1(x)]$ and $\tau_1(x)$ divides m then

$$\frac{1}{m}\sum_{k=n}^{n+m-1} f_{1,1}(x+k^2) \approx 1, \tag{25}$$

moreover, if $x \in S_{1,1}\setminus(S_{1,0}\cup E_1)$, $n, n+m \in [\alpha_1(x), \omega_1(x)]$ and $\tau_1(x)$ divides m then

$$\frac{1}{m}\sum_{k=n}^{n+m-1} f_{1,1}(x+k^2) \approx 1-\gamma. \tag{26}$$

We can also assume that our functions and sets are periodic by τ_1 and $\tau_1 \gg \omega_1^2(x)$, for all $x \in \mathbb{R}\setminus E_1$.

For the next step of the leakage we choose a large κ_2 such that $2^{\kappa_2} \gg \tau_1$ and $q_2 = p_{1,2}\cdots p_{\kappa_2,2}$ with $\tau_1 < p_{1,2} < \ldots < p_{\kappa_2,2}$.

Set $F_2 = F_1\setminus\overline{\Lambda}'(q_2)$, $f_{1,2}(x) = \chi_{F_2}(x)$, $S_{2,0} = S_{1,0}\setminus\overline{\Lambda}'(q_2) = \mathbb{R}\setminus(\overline{\Lambda}'(q_1)\cup\overline{\Lambda}'(q_2))$, $S_{2,1} = \mathbb{R}\setminus\overline{\Lambda}'(q_2)$, $S_{2,2} = \mathbb{R}$. Then $\overline{\lambda}(F_2) \approx (1-\gamma)^2$. We again have a small exceptional set E_2 and for each $x \in \mathbb{R}\setminus E_2$ there is a time interval $[\alpha_2(x), \omega_2(x)]$ and a number $\tau_2(x) \ll \omega_2(x) - \alpha_2(x)$. For $x \in S_{2,1}\setminus E_2$ we can use the "old time interval and $\tau_1(x)$", that is, $\alpha_2(x) = \alpha_1(x)$, $\omega_2(x) = \omega_1(x)$, $\tau_2(x) = \tau_1(x)$. If $x \in S_{2,0}\setminus E_2 \subset S_{2,1}\setminus E_2$, $n, n+m \in [\alpha_2(x), \omega_2(x)]$ and $\tau_2(x)$ divides m then we want to have

$$\frac{1}{m}\sum_{k=n}^{n+m-1} f_{1,2}(x+k^2) \approx 1, \tag{27}$$

if $x \in S_{2,1}\setminus(S_{2,0}\cup E_1)$, $n, n+m \in [\alpha_2(x), \omega_2(x)]$ and $\tau_2(x)$ divides m then we want to have

$$\frac{1}{m}\sum_{k=n}^{n+m-1} f_{1,2}(x+k^2) \approx 1-\gamma, \tag{28}$$

moreover, when $x \in S_{2,2} \setminus (S_{2,1} \cup E_1) \approx \overline{\Lambda}'(q_2)$, $n, n+m \in [\alpha_2(x), \omega_2(x)]$ and $\tau_2(x)$ divides m our desire is to have

$$\frac{1}{m} \sum_{k=n}^{n+m-1} f_{1,2}(x+k^2) = \frac{1}{m} \sum_{k=n}^{n+m-1} \chi_{F_2}(x+k^2) \approx \overline{\lambda}(F_2) \approx (1-\gamma)^2. \tag{29}$$

To have (29) we need that for "most elements", x of $\overline{\Lambda}'(q_2) \approx S_{2,1} \setminus (S_{2,0} \cup E_1)$ we have

$$\frac{1}{q_2} \#\{k : 0 \le k < q_2,\ x+k^2 \in \mathbb{R} \setminus \overline{\Lambda}'(q_2)\} \approx 1-\gamma \approx \overline{\lambda}(\mathbb{R} \setminus \overline{\Lambda}'(q_2)).$$

This will be ensured by Lemma 2 of [6]. Moreover, we also need that when $x+k^2 \in \mathbb{R} \setminus \overline{\Lambda}'(q_2)$ then it hits $F_2 = F_1 \setminus \overline{\Lambda}'(q_2)$ for sufficiently many k's in order to have

$$\frac{1}{m} \sum_{k=n}^{n+m-1} \chi_{F_2}(x+k^2) \approx \overline{\lambda}(F_1)\overline{\lambda}(\mathbb{R} \setminus \overline{\Lambda}'(q_2)) \approx (1-\gamma)^2.$$

To ensure this we need to modify F_1 slightly by a procedure called τ-periodic rearrangement, this is discussed in Lemma 3 of Section 3 of [6].

6 Termination of the leakage process, existence of $K-M$ families

After L many leakage steps we end up with sets F_L, E_L, $S_{L,0}, \dots, S_{L,L}$ and with the function $f_{1,L} = \chi_{F_L}$ such that $\overline{\lambda}(F_L) = \overline{\int} f_{1,L} \approx (1-\gamma)^L \approx 2^{-M}$, $\overline{\lambda}(S_{L,0}) \approx (1-\gamma)^L$ and for $x \in S_{L,l} \setminus (S_{L,l-1} \cup E_L)$, where $l = 0, ..., L$ and $S_{L,-1} = \emptyset$, there exist $\alpha_L(x) \ll \omega_L(x)$, $\tau_L(x) \ll \omega_L(x) - \alpha_L(x)$ such that if $n, n+m \in [\alpha_L(x), \omega_L(x)]$ and $\tau_L(x)$ divides m then

$$\frac{1}{m} \sum_{k=m}^{n+m-1} f_{1,L}(x+k^2) \approx (1-\gamma)^l. \tag{30}$$

Set $X_{1,L}(x) = (1-\gamma)^l$ if $x \in S_{L,l} \setminus S_{L,l-1}$. As we remarked earlier we need to use small values of γ. Then $X_{1,L}(x)$ is not $M-0.99$ distributed but (see Figure 7) can be estimated from below by an $M-0.99$ distributed "random variable" X_1. This is done in Section "Finishing the leakage" (4.2.8) of [6].

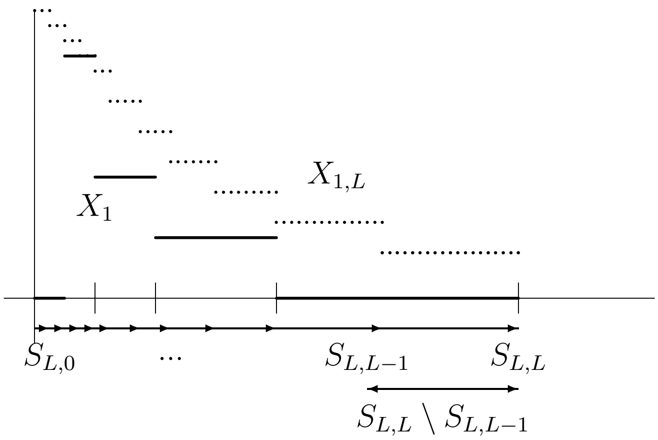

Figure 7: Schematic view of $X_{1,L}$ and X_1

By using "independence" properties of residue classes with respect to relatively prime bases, one can "put $K-M$ families" on residue classes see Sections "Putting $K-M$ families on quadratic residue classes" (4.1) (especially Lemma 5) and "Putting $K-M$ families on $\overline{\Lambda}'(q_L)$" (4.2.6) in [6].

The existence of $K-M$ families on $\mathbb{R}$ (Lemma 4 of Section "$K-M$ families" in [6]) is proved by induction on K.

If we know that $K-M$ families exist on $\mathbb{R}$ then by Lemma 5 of [6] we can "put $K-M$ families" on residue classes. This will provide us functions $f_1, \dots, f_K$ and $X_1, \dots, X_K$. By the leakage procedure a function f_{K+1} and the $M-0.99$ distributed "random variable" X_{K+1} can be defined so that X_{K+1} is independent from any function in $\{X_1, \dots X_K\}$ and we have (18) and (19) for $h = 1, \dots, K+1$.

7 How to build a $2-2$ family

To have a better understanding of how to build $K-M$ families next we give more details of how to build a $2-2$ family.

First we say a few words about Lemma 5 of [6].

Suppose first that an odd prime, q is given. Set $f(x) = q$ if $[x]$ is divisible by q and $f(x) = 0$ otherwise. Then $\overline{\int} f = 1$. Suppose $x \in \mathbb{R}$ and there exists

$k \not\equiv 0$ modulo q such that $[x+k^2] = [x]+k^2 \equiv 0 \bmod q$, that is, $[x] \equiv -k^2 \bmod q$. Then there are $k_1 \not\equiv k_2 \bmod q$ such that $k_1^2 = k_2^2 = k^2$. Therefore, for large N we have

$$\frac{1}{N}\sum_{k=1}^{N} f(x+k^2) \approx 2 \text{ when } [x] \in -\Lambda_0(q),\ [x] \not\equiv 0 \bmod q$$

and

$$\frac{1}{N}\sum_{k=1}^{N} f(x+k^2) \approx 0 \text{ when } [x] \notin -\Lambda_0(q).$$

If $q = p_1 \cdots p_\kappa$, $p_1 < ... < p_\kappa$ then the average spacing between consecutive quadratic residues mod q is about 2^κ. Set $f(x) = q/2^\kappa$ if $[x]$ is divisible by q and $f(x) = 0$, otherwise. Then $\overline{\int} f = 1/2^\kappa$. Suppose $x \in \mathbb{R}$ and there exists $k_x \not\equiv 0 \bmod q$ such that $[x+k_x^2] = [x]+k_x^2 \equiv 0 \bmod q$, that is, $[x] \in -\Lambda_0(q)$. If we assume in addition that none of the p_j, $(j = 1, ..., \kappa)$ divide $n_x = k_x^2$ then there are exactly 2^κ many solutions of $x^2 \equiv n_x \bmod q$. This implies that if $[x] \in -\overline{\Lambda}_0'(q)$ then for large N we have

$$\frac{1}{N}\sum_{k=1}^{N} f(x+k^2) \approx 1. \tag{31}$$

Suppose next that we set $f(x) = q/2^\kappa$ if $x \in q\mathbb{Z} + [0, \gamma 2^\kappa)$ and $f(x) = 0$ otherwise. Then we have $\overline{\int} f = \gamma$ and if $[x] \in \Lambda'(q) \overset{\text{def}}{=} \Lambda_\gamma'(q)$, that is, $x \in \overline{\Lambda}'(q)$ then for large N's we have (31). The set $\overline{\Lambda}'(q)$ consists of intervals of length $\gamma 2^\kappa$. By Property (†) if γ is small, κ and p_1 are large then most of these intervals are disjoint. This means that in this case if we take into consideration (22) as well, we have $\overline{\lambda}(\overline{\Lambda}'(q)) = \overline{\lambda}(\overline{\Lambda}_\gamma'(q)) \approx \gamma$. Therefore, we have obtained a function f with $\overline{\int} f \approx \gamma$ and for x's belonging to a set of approximate size γ we have (31).

In Lemma 5 of [6] we need to put a $K - M$ family onto a residue class. Here we just outline how to put a $1-2$ family onto a residue class. Suppose that we have a $1-2$ family on $\mathbb{R}$. Then we have $f_1 : \mathbb{R} \to [0, \infty)$ and a suitable $2-0.99$ distributed X_1, both periodic by τ, such that (18) and (19) hold for $h = 1$, $M = 2$ and $n, n+m$ with m divisible by a suitably chosen $\tau(x)$. For further reference we restate (19) in this special case

$$\overline{\int} f_1 < \Gamma \cdot 2^{-2+2} = \Gamma. \tag{32}$$

We also suppose that q is given so that

$$(\tau, q) = 1 \text{ and } (\tau(x), q) = 1 \tag{33}$$

for all $x \in \mathbb{R}$ which do not belong to a small exceptional set.

The $1-2$ family living on $\overline{\Lambda}'(q)$ will be periodic by τq. We define $\overline{f}_1(x) = f_1(x)q/2^\kappa$ if $x \in q\mathbb{Z} + [0, \gamma 2^\kappa)$ and otherwise we put $\overline{f}_1(x) = 0$. Then $\int \overline{f}_1 = \Gamma\gamma \approx \Gamma\overline{\lambda}(\overline{\Lambda}'(q))$, which yields (20) with $\gamma' = \gamma$. We set $\overline{X}_1(x) = X_1(x)$ for $x \in \overline{\Lambda}'(q)$ and otherwise we set $\overline{X}_1(x) = 0$. Then $\overline{f}_1$ and $\overline{X}_1$ will give us the $1-2$ family living on $\overline{\Lambda}'(q)$.

Next we outline how to build a $2-2$ family. To simplify this argument we suppose that $\gamma = 1/2$ and we also suppose (although it is not true) that $\overline{\Lambda}'(q) = \overline{\Lambda}'_{1/2}(q)$ consists of disjoint intervals if $q = p_1 \cdots p_\kappa$, $p_1 < \ldots < p_\kappa$, κ and p_1 are large. This simplification makes it possible that we need to discuss only one step of leakage, instead of L-many steps. This way we can skip discussing the procedure of obtaining X_2 from $X_{2,L}$, as it is discussed in Section 6 and is illustrated on Figure 7.

With the above simplifying assumptions one can quite easily build a $1-2$ family living on $\mathbb{R}$. So, first by a one step leakage process we choose a $1-2$ system periodic by τ_1, consisting of $f_{1,0}$ and $X_{1,0}$. Then we choose κ, $q = p_1 \cdots p_\kappa$, $p_1 < \ldots < p_\kappa$ such that $2^\kappa \gg \tau_1$, $(q, \tau_1) = 1$, $(q, \tau_{1,0}(x)) = 1$ for all $x \in \mathbb{R}$, where $\tau_{1,0}(x)$ is the auxiliary function for our $1-2$ system, see (33). We also suppose that τ_1 (not being necessarily the smallest period of $f_{1,0}$ and $X_{1,0}$) is so large that $\omega_{1,0}(x)$, which is now our $\omega(x)$ appearing above (18), satisfies $\omega_{1,0}^2(x) < \tau_1$ for all x which do not belong to a small exceptional set. This means that "all the needed action" of the $1-2$ system associated to $f_{1,0}$, $X_{1,0}$ is "done" in intervals of length τ_1.

We set $f_{2,0} \equiv 1$, $X_{2,0} \equiv 1$, $F_0 = S_{0,0} = \mathbb{R}$.

To obtain $f_{2,1}$ we do a one step leakage of $f_{2,0}$ onto $\overline{\Lambda}'(q) \stackrel{\text{def}}{=} \overline{\Lambda}'_{1/2}(q)$, that is, $f_2(x) = f_{2,1}(x) = 1$ if and only if $x \notin \overline{\Lambda}'(q)$, $F_1 = \mathbb{R} \setminus \overline{\Lambda}'(q)$, $S_{1,0} \approx \mathbb{R} \setminus \overline{\Lambda}'(q)$ and $S_{1,1} = \mathbb{R}$. We define $f_1(x) = f_{1,0}(x)$ if $x \in \mathbb{R} \setminus \overline{\Lambda}'(q)$. To define $f_1(x)$ on $\overline{\Lambda}'(q)$ we put a suitable $1-2$ family onto this residue class as described above. Similarly, $X_1(x)$ can be chosen to be equal to $X_{1,0}(x)$ for most points of $\mathbb{R} \setminus \overline{\Lambda}'(q)$ and for most points of $\overline{\Lambda}'(q)$ this will equal the function we obtain when we put a suitable $1-2$ family onto $\overline{\Lambda}'(q)$. This will mean that for our $2-2$ system "all the needed action" will take place for most $x \in \mathbb{R} \setminus \overline{\Lambda}'(q)$

quite “early”, before “time” τ_1, while for points in $\overline{\Lambda}'(q)$ the “needed action” takes place in a time interval $[\alpha(x), \omega(x)]$ with $\tau_1 \ll \alpha(x)$.

References

[1] V. Bergelson “Combinatorial and Diophantine Applications of Ergodic Theory” Handbook of dynamical systems. Vol. **1B**, Elsevier B. V., Amsterdam, (2006), 745–869,.

[2] J. Bourgain, “Pointwise ergodic theorems for arithmetic sets”, With an appendix by the author, Harry Fürstenberg, Yitzhak Katznelson and Donald S. Ornstein, *Inst. Hautes Études Sci. Publ. Math.* No. 69 (1989), 5–45.

[3] J. Bourgain, “An approach to pointwise ergodic theorems”, *Geometric aspects of functional analysis (1986/87)*, 204–223, Lecture Notes in Math., 1317, Springer, Berlin, 1988.

[4] J. Bourgain, “On the pointwise ergodic theorem on L^p for arithmetic sets”, *Israel J. Math.* **61** (1988), no. 1, 73–84.

[5] J. Bourgain, “Almost sure convergence in ergodic theory”, *Almost everywhere convergence. Proceedings of the International Conference on Almost Everywhere Convergence in Probability and Ergodic Theory held in Columbus, Ohio, June 11–14, 1988.* Edited by Gerald A. Edgar and Louis Sucheston. Academic Press, Inc., Boston, MA, (1989).

[6] Z. Buczolich and R. D. Mauldin “Divergent Square Averages”, *to appear.*

[7] J. P. Conze, “Convergence des moyennes ergodiques pour des sous suites”, *Bull. Soc. Math. France*, **35** (1973), 7-15.

[8] R. L. Jones, “Ergodic Theory and Connections with Analysis and Probability”, *New York J. Math.* **3A** (1997), 31–67.

[9] P. Kurlberg and Z. Rudnick, “The distribution of spacings between quadratic residues”, *Duke Math. J.* **100** (1999), no. 2, 211-242.

[10] P. Kurlberg, "The distribution of spacings between quadratic residues II", *Israel J. Math.* **120** (2000), part A, 205-224.

[11] J. Lamperti, *Probability: A survey of the Mathematical Theory*, W. A. Benjamin Inc., New York (1966).

[12] J. Rosenblatt, "Universally bad sequences in ergodic theory." *Almost everywhere convergence (II)* (eds., A. Bellow, R. Jones), Academic Press 1991, 227-245.

[13] J. Rosenblatt and M. Wierdl "Pointwise ergodic theorems via harmonic analysis", *Ergodic theory and its connections with harmonic analysis (Alexandria, 1993)*, 3–151, London Math. Soc. Lecture Note Ser., 205, Cambridge Univ. Press, Cambridge, 1995.

[14] S. Sawyer, "Maximal inequalities of weak type", *Ann. of Math.* **84** (1966) no. 4, 157-174.