

VISITING EARLY AT PRIME TIMES

TONY HADDAD, SUN-KAI LEUNG, AND CIHAN SABUNCU

ABSTRACT. Given an integer $m \geq 2$ and a sufficiently large q , we apply a variant of the Maynard–Tao sieve weight to establish the existence of an arithmetic progression with common difference q for which the m -th least prime in such progression is $\ll_m q$, which is best possible. As we vary over progressions instead of fixing a particular one, the nature of our result differs from others in the literature. Furthermore, we generalize our result to dynamical systems. The quality of the result depends crucially on the first return time, which we illustrate in the case of Diophantine approximation.

1. INTRODUCTION

Analytic Number Theory has witnessed many breakthroughs on fundamental problems in recent years, in particular the bounded gaps between primes, thanks to Zhang [Zha14], Maynard [May15] and Tao (see also [Pol14]), building upon the celebrated work of Goldston–Pintz–Yıldırım [GPY09]. Maynard [May16] further proved that any subset of primes which is “well distributed” in arithmetic progressions contains many primes that are close together.

In particular, for primes in a fixed progression, the best known result is due to Baker–Zhao [BZ16], who proved that given $m \geq 1$, let x be a sufficiently large real number, $q = x^\theta$ a suitably “powerful” modulus, and $(a, q) = 1$. Then there exist primes $p_1 < \cdots < p_m$ in $(x, 2x]$ such that $p_i \equiv a \pmod{q}$ for $1 \leq i \leq m$ and

$$p_m - p_1 \leq q \exp(cm)$$

for some constant $c = c(\theta) > 0$.

Apart from progressions, there are also infinitely many bounded gaps between primes in a fixed “Chebotarev set” [Tho14], “Sato–Tate interval” [Gil+20] or “nil-Bohr set” [ST21].

To motivate our work, we begin by asking: how big is the first prime in a given arithmetic progression? Let $p(q, a)$ denote the least prime in the arithmetic progression $a \pmod{q}$. Then recall Linnik’s celebrated theorem that $p(q, a) \ll q^L$ for some absolute constant $L > 0$. Assuming a uniform variant of the Hardy–Littlewood prime k -tuple conjecture, the second author [Leu24a] showed that as soon as $f(q) \rightarrow \infty$ with $q \rightarrow \infty$, the least prime $p(q, a)$ is $\leq f(q)\varphi(q) \log q$ for almost all $a \pmod{q}$. In fact, the least prime in arithmetic progressions to a common large modulus has an exponential limiting distribution. More precisely, if a reduced residue $a \pmod{q}$ is chosen uniformly at random, then we have the convergence in distribution

$$\frac{p(q, a)}{\varphi(q) \log q} \xrightarrow{d} \exp(1) \quad \text{as } q \rightarrow \infty.$$

Given $m \geq 1$ and $(a, q) = 1$, we are also interested in the m -th least prime $p \equiv a \pmod{q}$, which we denote as $p_m(q, a)$. As $a \pmod{q}$ varies uniformly with $q \rightarrow \infty$, the second author [Leu24b] further showed conditionally that the normalized primes in the

progression,

$$\frac{p_1(q, a)}{\varphi(q) \log q} < \frac{p_2(q, a)}{\varphi(q) \log q} < \dots$$

are located on the positive real number line $(0, \infty)$ as if they were randomly positioned points, in the sense of a Poisson point process.

In this paper, applying a variant of the Maynard–Tao sieve weight, we establish the existence of a residue class $a \pmod{q}$ for which the first few normalized primes in the progression form an early cluster, which can be viewed as the progression counterpart to bounded gaps between primes.

Theorem 1.1. *Let $m \geq 2$. Then for any sufficiently large modulus q (in terms of m), we have*

$$\min_{a \pmod{q}} p_m(q, a) \ll_m q.$$

Remark 1.1. When $m = 2$, following the proof in [Pol14], the implied constant can be taken as 270.

Aside from the dependence on m , the optimality here is clear, whereas by the pigeonhole principle, one can only show that $\min_{a \pmod{q}} p_m(q, a) \ll_m \varphi(q) \log q$. Also, by varying the progressions instead of fixing a particular one, we are able to obtain a prime cluster early,¹ which makes the nature of our result distinct from the existing literature, including those discussed in the beginning of the paper.

Furthermore, we generalize our result to dynamical systems, including circle rotations (Diophantine approximation) and Möbius transformations (see Example 2.2 and Example 2.3 respectively). It turns out that the quality of the result depends crucially on the first return time, which we explore in Section 7. In particular, in the case of circle rotations $T_\alpha : x \mapsto x + \alpha$ for almost all α or those α with bounded partial quotients, we establish tighter lower and upper bounds than those in the literature for the first return time (see Proposition 7.1).

Recurrence is a central theme in ergodic theory, i.e. how points in dynamical systems return close to themselves under iteration. In 1890, Poincaré proved a simple yet far-reaching recurrence theorem in the context of the “three-body” problem of planetary orbits (see [EW11, p. 21]). For our purposes, here we state a simple version (see [EW11, Exercise 2.2.1] for instance).

Lemma 1.1 (Poincaré recurrence theorem). *Given a measure-preserving system (X, \mathcal{B}, μ, T) , let $A \in \mathcal{B}$ with $\mu(A) > 0$. Then there exists an integer $1 \leq n \leq \mu(A)^{-1}$ such that*

$$\mu(A \cap T^{-n}A) > 0.$$

This is essentially the pigeonhole principle. A famous example is Dirichlet’s approximation theorem, corresponding to circle rotation. Inspired by the recurrence theorem, we shall visit early at prime times in *metric-measure-preserving systems* defined as follows.

Definition 1.1 (Metric-measure-preserving system). *A metric-measure-preserving system is a quintuple $(X, d, \mathcal{B}, \mu, T)$ where (X, d) is a metric space and (X, \mathcal{B}, μ, T) is a measure-preserving system satisfying the following compatibility conditions:*

- the Borel σ -algebra \mathcal{B} is induced by the metric d ;

¹See [Pom80] and a recent improvement [LPS17] for a lower bound for the least prime in a progression instead.

- the probability measure μ is *pseudo-doubling*, i.e. given a sufficiently small $\epsilon > 0$, there exists $\lambda \geq 1$ such that

$$\mu(B(x; 2\epsilon)) \leq \lambda \cdot \mu(B(x; \epsilon))$$

for any $x \in X$, where $B(x; \epsilon) := \{y \in X : d(y, x) < \epsilon\}$;

- the measure-preserving map T is an isometry.²

Let $p_m^T(x_0, x; \epsilon)$ denote the m -th least prime p for which $T^p x_0 \in B(x; \epsilon)$. Then, with the definition in hand, we state the dynamical generalization of Theorem 1.1.

Theorem 1.2. *Let $m \geq 2$ and $(X, d, \mathcal{B}, \mu, T)$ be a metric-measure-preserving system. Then for any sufficiently small $\epsilon > 0$ (in terms of m) and $x_0 \in X$, we have*

$$\min_{x \in X} p_m^T(x_0, x; \epsilon) \ll_{m, \lambda} \mu(B(x_0; \epsilon))^{-1}.$$

Notation. Throughout the paper, we use the standard big O and little o notations as well as the Vinogradov notations \ll, \gg and the Hardy notation \asymp , where the implied constants depend only on the subscripted parameters.

2. MAIN RESULTS

We begin by stating the quantitative version of Theorem 1.1, which provides a lower bound for the number of residue classes with early prime clusters.

Theorem 2.1. *Let $m \geq 2$. Then for any sufficiently large modulus q (in terms of m), there exist absolute constants $C, C' > 0$ such that*

$$\#\{a \pmod{q} : p_m(q, a) \leq Cqm \exp(4m)\} \gg_m \varphi(q)(\log q)^{-C' \exp(4m)}.$$

Remark 2.1. Unsurprisingly, the constant 4 here can be reduced to 2 if one assumes the Elliott–Halberstam conjecture [EH70]. See [Pol14] and [Sta23] for potential unconditional improvements.

We also state the quantitative version of Theorem 1.2, which is the dynamical generalization of Theorem 2.1.

Theorem 2.2. *Let $m \geq 2$ and $(X, d, \mathcal{B}, \mu, T)$ be a metric-measure-preserving system. Then there exists constants $C = C(\lambda), C' = C'(\lambda) > 0$ such that for any $x_0 \in X$ and sufficiently small $\epsilon > 0$ (in terms of m), we have*

$$\begin{aligned} & \mu(\{x \in X : p_m^T(x_0, x; \epsilon) \leq Cqm \exp(4m)\}) \\ & \gg_{m, \lambda} \varphi(q)(1 + \log q)^{-C' \exp(4m)} \mu(B(x_0; \epsilon)) \end{aligned}$$

for some integer $1 \leq q \leq m^{\log_2 \lambda} \exp(4m \log_2 \lambda) \mu(B(x_0; \epsilon))^{-1}$.³

To illustrate our theorem, we provide several examples.

Example 2.1 (Right shift). *Consider the metric-measure-preserving system with $X = \mathbb{Z}/q\mathbb{Z}$, $d =$ discrete metric, $\mu =$ normalized counting measure and $T_q : a \mapsto a + 1 \pmod{q}$. Since the m -th least prime $p_m(q, a) \gg_m q$, we recover Theorem 2.1.*

²In fact, for our purposes, it suffices to assume that T is non-expanding, i.e. $d(Tx, Ty) \leq d(x, y)$ for any $x, y \in X$. In particular, both Theorem 1.2 and Theorem 2.2 hold for such T .

³See Section 7 for further discussion on the size of q , which depends crucially on the first return time.

Example 2.2 (Circle rotation). *Consider the metric-measure-preserving system with $X = [0, 1]$, $d = \|\cdot\|$, $\mu = \text{Lebesgue measure}$ and $T_\alpha : x \mapsto x + \alpha \pmod{1}$, where $\|\cdot\|$ is the distance to the nearest integer function and $\alpha \in [0, 1]$. Let $m \geq 2$. Then for any sufficiently small $\epsilon > 0$ (depending on m), there exists $\beta \in [0, 1]$ such that*

$$p_m^\alpha(\beta; \epsilon) \ll \epsilon^{-1} m^2 \exp(8m)$$

as $\epsilon \rightarrow 0^+$, where $p_m^\alpha(\beta; \epsilon)$ denotes the m -th least prime p for which $\|p\alpha - \beta\| < \epsilon$. See Corollary 7.1 for the quantitative version.

Example 2.3 (Möbius transformation). *Consider the metric-measure-preserving system with $X = SL_2(\mathbb{Z}) \backslash \mathbb{H}$, $d = d_{\text{hyp}}$, $\mu = \text{hyperbolic measure}$ and $T_g : z \mapsto gz \pmod{SL_2(\mathbb{Z})}$, where d_{hyp} is the hyperbolic distance and $g \in SL_2(\mathbb{R})$. Let $m \geq 2$. Then for any sufficiently small $\epsilon > 0$ (depending on m), there exists $\tau \in SL_2(\mathbb{Z}) \backslash \mathbb{H}$ such that*

$$p_m^g(\tau; \epsilon) \ll \epsilon^{-2} m^3 \exp(12m)$$

as $\epsilon \rightarrow 0^+$, where $p_m^g(\tau; \epsilon)$ denotes the m -th least prime p for which $d_{\text{hyp}}(g^p(i), \tau) < \epsilon$.

3. A VARIANT OF MAYNARD–TAO SIEVE WEIGHT

Following the style of [Pol14], we shall provide an outline of the proof of Theorem 2.1. Similar to the proof of [May16, Theorem 3.1], let $m \geq 2$ and an admissible k -tuple (h_1, \dots, h_k) depending on m , i.e. a tuple of k non-negative integers $h_1 < \dots < h_k$ satisfying $\#\{h_i \pmod{p} : 1 \leq i \leq k\} < p$ for any prime p . Then, we would like to construct a weight w_a (with size depending on the number of prime factors) over the reduced residue system $\{1 \leq a \leq q : (a, q) = 1\}$ for any $q = q(m)$ sufficiently large, so that

$$S := \sum_{\substack{a=1 \\ (a,q)=1}}^q \left(\sum_{i=1}^k 1_{\mathbb{P}}(a + qh_i) - (m-1) - k \sum_{i=1}^k \sum_{\substack{p|a+qh_i \\ p \leq q^\rho}} 1 \right) w_a \quad (3.1)$$

is positively large for some small $\rho = \rho(k) > 0$. By the pigeonhole principle, there exists a reduced residue $a \pmod{q}$ such that the expression in the parenthesis is positive, in which case at least m of the $a + qh_i$ are primes and none of them have prime factors $p \leq q^\rho$, so that w_a is bounded in terms of the number of prime factors $\Omega(a + qh_i) \leq \rho^{-1}$ for $1 \leq i \leq k$. Since the expression in the parenthesis is bounded by k from above, the expression (3.1) gives

$$\#\{a \pmod{q} : \text{at least } m \text{ of the } a + qh_i \text{ are primes}\} \geq \left(k \max_{a \pmod{q}} w_a \right)^{-1} S. \quad (3.2)$$

To construct our weight w_a , let $w := \log \log \log q$ and $W_q := \prod_{\substack{p \leq w \\ p \nmid q}} p$. By the admissibility of the k -tuple (h_1, \dots, h_k) , we can choose $b_0 \pmod{W_q}$ so that for those $a \equiv b_0 \pmod{W_q}$, we have $a + qh_i \equiv b_0 + qh_i \not\equiv 0 \pmod{p}$ for any prime $p \leq w$, $p \nmid q$. Then, analogous to the Maynard–Tao sieve weight (see [Pol14, p.11]), ours is

$$w_a := 1_{a \equiv b_0 \pmod{W_q}} \left(\sum_{\substack{d_i | a + qh_i \\ i=1, \dots, k}} \prod_{i=1}^k \mu(d_i) \times F \left(\frac{\log d_1}{\log q}, \dots, \frac{\log d_k}{\log q} \right) \right)^2$$

for some function $F : [0, \infty)^k \rightarrow \mathbb{R}$ such that its mixed derivative $\partial_{t_1 \dots t_k} F$ is square-integrable and is supported on the simplex

$$\Delta_k(\theta; \epsilon) := \{(t_1, \dots, t_k) \in [0, \infty)^k : t_1 + \dots + t_k \leq (\theta - \epsilon)/2\},$$

where $\theta > 0$ being the ‘‘level of distribution’’ of primes, and $\epsilon = \epsilon(k) = \frac{1}{\log k} > 0$ say. Thanks to the Bombieri–Vinogradov theorem, throughout the paper, we can simply take $\theta = 1/2$ unconditionally (or $\theta = 1$ under the Elliott–Halberstam conjecture). For future reference, the family of such functions is denoted as $\mathcal{F}_{k,\theta}$.

With our choice of weight w_a , it turns out that not only is the sum S in (3.1) computable, but it is also positively large.

To simplify notation, given a square-integrable function $G : [0, \infty)^k \rightarrow \mathbb{R}$, let us denote the singular integrals

$$\mathfrak{I}(G) := \int_{[0,\infty)^k} G(t_1, \dots, t_k)^2 dt_1 \cdots dt_k$$

and

$$\mathfrak{J}_i(G) := \int_{[0,\infty)^{k-1}} \left(\int_{[0,\infty)} G(t_1, \dots, t_k) dt_i \right)^2 dt_1 \cdots dt_{i-1} dt_{i+1} \cdots dt_k$$

for $1 \leq i \leq k$. Then in view of (3.1), we are required to estimate the following sums.

Proposition 3.1 (Non-prime sum). *Let q be sufficiently large in terms of k . Then*

$$\sum_{\substack{a=1 \\ (a,q)=1}}^q w_a = (\mathfrak{I}(\partial_{t_1 \dots t_k} F) + o(1)) \frac{\varphi(q)}{W_q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \times \log q \right)^{-k}.$$

Proposition 3.2 (Prime sum). *Let $1 \leq i_0 \leq k$ and q be sufficiently large in terms of k . Then*

$$\sum_{\substack{a=1 \\ (a,q)=1}}^q 1_{\mathbb{P}}(a + qh_{i_0}) w_a = (\mathfrak{J}_{i_0}(\partial_{t_1 \dots t_k} F) + o(1)) \frac{\varphi(q)}{W_q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \times \log q \right)^{-k}.$$

Proposition 3.3 (Small prime factor sum). *Let $0 < \rho \leq \frac{1}{100k}$ and q be sufficiently large in terms of k . Then*

$$\sum_{\substack{a=1 \\ (a,q)=1}}^q \left(\sum_{\substack{p|a+qh_{i_0} \\ p \leq q^\rho}} 1 \right) w_a \leq C_1 \rho \times \frac{\varphi(q)}{W_q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \times \log q \right)^{-k}$$

for some absolute constant $C_1 > 0$.

4. PROOF OF THEOREM 2.1

In this section, we shall prove Theorem 2.1 assuming the aforementioned propositions. Combining Propositions 3.1, 3.2 and 3.3, the expression (3.1) gives

$$\begin{aligned} S \geq & \left(\sum_{i=1}^k \mathfrak{J}_i(\partial_{t_1 \dots t_k} F) - (m-1)\mathfrak{I}(\partial_{t_1 \dots t_k} F) - C_1 k^2 \rho - o(1) \right) \\ & \times \frac{\varphi(q)}{W_q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \times \log q \right)^{-k}. \end{aligned} \quad (4.1)$$

Applying [Pol14, Theorem 3.9] followed by suitable rescaling, the ratio of the singular series is

$$\sup_{G \in \mathcal{F}} \frac{\sum_{i=1}^k \mathfrak{J}_i(\partial_{t_1 \dots t_k} G)}{\mathfrak{I}(\partial_{t_1 \dots t_k} G)} \geq \frac{\sum_{i=1}^k \mathfrak{J}_i(\partial_{t_1 \dots t_k} F)}{\mathfrak{I}(\partial_{t_1 \dots t_k} F)} \geq \frac{\theta}{2} \log k - C_2$$

for some absolute constant $C_2 > 0$ and function F such that its mixed derivative

$$\partial_{t_1 \dots t_k} F(t_1, \dots, t_k) = 1_{\Delta_k(\theta; \epsilon)}(t_1, \dots, t_k) \prod_{i=1}^k \psi(t_i),$$

where $\psi(t) := \frac{1}{c+(k-1)t}$ with $c = \frac{1}{\log k} - \frac{1}{\log^2 k}$ (see [Pol14, Theorem 6.7]). Given such a choice of function F , since by definition

$$\mathfrak{I}(\partial_{t_1 \dots t_k} F) = \int_{\Delta_k(\theta; \epsilon)} \left(\prod_{i=1}^k \psi(t_i) \right)^2 dt_1 \dots dt_k \gg \left(\frac{\log k}{k} \right)^k,$$

it follows from (4.1) that

$$\begin{aligned} S &\geq \left(\left(\frac{\theta}{2} \log k - C_2 - (m-1) \right) \mathfrak{I}(\partial_{t_1 \dots t_k} F) - C_1 k^2 \rho - o(1) \right) \\ &\quad \times \frac{\varphi(q)}{W_q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \times \log q \right)^{-k} \\ &\gg \frac{\varphi(q)}{W_q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \times \log q \right)^{-k} \left(\frac{\log k}{k} \right)^k \end{aligned} \quad (4.2)$$

by taking $k = \lceil \exp\left(\frac{2}{\theta}(m + C_2)\right) \rceil$ and $\rho = k^{-k}$.

On the other hand, since $g(t) \leq \min\left\{\frac{1}{c}, \frac{1}{(k-1)t}\right\}$, we have

$$\begin{aligned} \max_{a \pmod{q}} w_a &\ll (k-1)^{-2k} (\log q)^{2k} \left(\prod_{i=1}^k \sum_{d_i | a+qh_i} 1 \right)^2 \\ &= (k-1)^{-2k} (\log q)^{2k} \prod_{i=1}^k \prod_{p|a+qh_i} 4. \end{aligned}$$

Since $\Omega(a + qh_i) \leq \rho^{-1}$ for $1 \leq i \leq k$ as discussed earlier, this is

$$\leq (\log q)^{2k} \exp(C_3 k / \rho) \quad (4.3)$$

for some absolute constant $C_3 > 0$.

Therefore, combining (4.2) and (4.3), it follows from (3.2) that

$$\begin{aligned} &\#\{a \pmod{q} : \text{at least } m \text{ of the } a + qh_i \text{ are primes}\} \\ &\gg_k \frac{\varphi(q)}{(\log q)^{3k}} \times \frac{1}{W_q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \right)^{-k} \gg_{k, \eta} \frac{\varphi(q)}{(\log q)^{(3+\eta)k}}. \end{aligned}$$

Finally, let (h_1, \dots, h_k) be the narrowest admissible k -tuple with $h_1 = 0$. Then $h_k \ll k \log k$ by [Pol14, Theorem 3.3], so that $p_m(q, a) \leq a + qh_k \ll qk \log k$ and the proof is completed.

5. PROOF OF PROPOSITIONS 3.1, 3.2 & 3.3

We follow the treatment in [Pol14]. Without loss of generality, by the Stone–Weierstrass theorem, we may assume F is a linear combination of tensor products, i.e. there are constants $c_j \in \mathbb{R}$ and $f_{i,j} \in \mathcal{C}_c^\infty([0, \infty))$ for $1 \leq i \leq k$ and $1 \leq j \leq J$ such that

$$F(t_1, \dots, t_k) = \sum_{j=1}^J c_j \prod_{i=1}^k f_{i,j}(t_i)$$

(see [Pol14, Section 5.1] for details). Also, let us denote

$$\lambda_f(n) := \sum_{d|n} \mu(d) f\left(\frac{\log d}{\log q}\right).$$

Then our weight becomes

$$\begin{aligned} w_a &= 1_{a \equiv b_0 \pmod{W_q}} \left(\sum_{j=1}^J c_j \prod_{i=1}^k \lambda_{f_{i,j}}(a + qh_i) \right)^2 \\ &= 1_{a \equiv b_0 \pmod{W_q}} \sum_{j=1}^J \sum_{j'=1}^J c_j c_{j'} \prod_{i=1}^k \lambda_{f_{i,j}}(a + qh_j) \lambda_{f_{i,j'}}(a + qh_{j'}). \end{aligned}$$

Before embarking the proof of the propositions, we first show that reduced residues are well-distributed in progressions.

Lemma 5.1 (Level of distribution of reduced residues). *Let $\epsilon \in (0, 1/2)$. Then as $q \rightarrow \infty$, we have*

$$\sum_{\substack{r \leq q^{1-2\epsilon} \\ (r,q)=1}} \max_{c \pmod{r}} \left| \sum_{a=1}^q 1_{(a,q)=1} \left(1_{a \equiv c \pmod{r}} - \frac{1}{r} \right) \right| \ll_{\epsilon} q^{1-\epsilon}.$$

Proof. By Möbius inversion, we have

$$\sum_{a=1}^q 1_{(a,q)=1} 1_{a \equiv c \pmod{r}} = \sum_{d|q} \mu(d) \sum_{\substack{a=1 \\ d|a}}^q 1_{a \equiv c \pmod{r}}. \quad (5.1)$$

Since $(r, q) = 1$ and $d | q$, the Chinese remainder theorem gives

$$\sum_{\substack{a=1 \\ d|a}}^q 1_{a \equiv c \pmod{r}} = \frac{q}{rd} + O(1),$$

so that (5.1) is

$$\frac{q}{r} \sum_{d|q} \frac{\mu(d)}{d} + O(\tau(q)) = \frac{\varphi(q)}{r} + O(\tau(q)).$$

Therefore, we have

$$\begin{aligned} \sum_{\substack{r \leq q^{1-2\epsilon} \\ (r,q)=1}} \max_{c \pmod{r}} \left| \sum_{a=1}^q 1_{(a,q)=1} \left(1_{a \equiv c \pmod{r}} - \frac{1}{r} \right) \right| &\ll q^{1-2\epsilon} \tau(q) \\ &\ll_{\epsilon} q^{1-\epsilon}, \end{aligned}$$

and the lemma follows. \square

Proof of Proposition 3.1. By linearity, it suffices to estimate

$$\sum_{\substack{a=1 \\ (a,q)=1}}^q 1_{a \equiv b_0 \pmod{W_q}} \prod_{i=1}^k \lambda_{f_i}(a + qh_i) \lambda_{g_i}(a + qh_i) \quad (5.2)$$

for all $f_i, g_i \in \mathcal{C}_c^\infty([0, \infty))$ with $1 \leq i \leq k$. Then, after interchanging the order of summation, this becomes

$$\sum_{d_1, d'_1, \dots, d_k, d'_k \geq 1} \cdots \sum_{i=1}^k \left(\prod_{i=1}^k \mu(d_i) \mu(d'_i) f_i \left(\frac{\log d_i}{\log q} \right) g_i \left(\frac{\log d'_i}{\log q} \right) \right) S(\mathbf{d}, \mathbf{d}'),$$

where $\mathbf{d} := (d_1, \dots, d_k)$, $\mathbf{d}' := (d'_1, \dots, d'_k)$ and

$$S(\mathbf{d}, \mathbf{d}') := \sum_{\substack{a=1 \\ [d_i, d'_i] \mid a+qh_i, 1 \leq i \leq k \\ a \equiv b_0 \pmod{W_q}}}^q 1_{(a,q)=1}.$$

Since by assumption $|h_i - h_j| < w$ and $(a + qh_i, W_q) = 1$, the sum $S(\mathbf{d}, \mathbf{d}')$ vanishes unless $q, W_q, [d_1, d'_1], \dots, [d_k, d'_k]$ are pairwise coprime. Let $r_{W_q, \mathbf{d}, \mathbf{d}'} := W_q \prod_{i=1}^k [d_i, d'_i]$. Then, it follows from the Chinese remainder theorem that there exists a unique reduced residue $c_0 \pmod{r_{W_q, \mathbf{d}, \mathbf{d}'}}$ such that

$$S(\mathbf{d}, \mathbf{d}') = \frac{\varphi(q)}{r_{W_q, \mathbf{d}, \mathbf{d}'}} + \sum_{a=1}^q 1_{(a,q)=1} \left(1_{a \equiv c_0 \pmod{r_{W_q, \mathbf{d}, \mathbf{d}'}}} - \frac{1}{r_{W_q, \mathbf{d}, \mathbf{d}'}} \right).$$

Therefore, the main term of (5.2) is

$$\frac{\varphi(q)}{W_q} \sum_{\substack{d_1, d'_1, \dots, d_k, d'_k \geq 1 \\ q, W_q, [d_i, d'_i] \text{ pairwise coprime}}} \prod_{i=1}^k \frac{\mu(d_i) \mu(d'_i)}{[d_i, d'_i]} f_i \left(\frac{\log d_i}{\log q} \right) g_i \left(\frac{\log d'_i}{\log q} \right),$$

which is also

$$\frac{\varphi(q)}{W_q} \sum_{\substack{d_1, d'_1, \dots, d_k, d'_k \geq 1 \\ q_W, W, [d_i, d'_i] \text{ pairwise coprime}}} \prod_{i=1}^k \frac{\mu(d_i) \mu(d'_i)}{[d_i, d'_i]} f_i \left(\frac{\log d_i}{\log q} \right) g_i \left(\frac{\log d'_i}{\log q} \right),$$

where $q_W := q/(q, W)$. Then, applying [Pol14, Lemma 4.1], this becomes

$$\begin{aligned} & (\mathfrak{J}(\mathbf{f}, \mathbf{g}) + o(1)) \frac{\varphi(q)}{W_q} \left(\frac{\varphi(W)}{W} \times \frac{\varphi(q_W)}{q_W} \times \log q \right)^{-k} \\ &= (\mathfrak{J}(\mathbf{f}, \mathbf{g}) + o(1)) \frac{\varphi(q)}{W_q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \times \log q \right)^{-k}, \end{aligned}$$

where $\mathbf{f} := (f_1, \dots, f_k)$, $\mathbf{g} := (g_1, \dots, g_k)$ and

$$\mathfrak{J}(\mathbf{f}, \mathbf{g}) := \prod_{i=1}^k \int_0^\infty f'_i(t_i) g'_i(t_i) dt_i.$$

On the other hand, note that for any $r \geq 1$ with $(r, q) = 1$ we have

$$\#\{(\mathbf{d}, \mathbf{d}') : r = r_{W_q, \mathbf{d}, \mathbf{d}'}\} \leq \tau_{3k}(r),$$

and the product $\prod_{i=1}^k f_i \left(\frac{\log d_i}{\log q} \right) g_i \left(\frac{\log d'_i}{\log q} \right)$ vanishes unless

$$\begin{aligned} r_{W_q, \mathbf{d}, \mathbf{d}'} &\leq W_q \times q \left(\sum_{i=1}^k \frac{\log d_i}{\log q} + \sum_{i'=1}^k \frac{\log d'_{i'}}{\log q} \right) \\ &\leq q^{\frac{1}{2}-\epsilon}. \end{aligned}$$

Therefore, the error term of (5.2) is

$$\leq \sum_{\substack{r \leq q^{1/2-\epsilon} \\ (r,q)=1}} \tau_{3k}(r) \max_{(c,r)=1} \left| \sum_{a=1}^q 1_{(a,q)=1} \left(1_{a \equiv c \pmod{r}} - \frac{1}{r} \right) \right|. \quad (5.3)$$

Since

$$\max_{(c,r)=1} \left| \sum_{a=1}^q 1_{(a,q)=1} \left(1_{a \equiv c \pmod{r}} - \frac{1}{r} \right) \right| \ll \frac{q}{r},$$

it follows from the Cauchy–Schwarz inequality that (5.3) is

$$\begin{aligned} &\ll q^{\frac{1}{2}} \sum_{\substack{r \leq q^{1/2-\epsilon} \\ (r,q)=1}} \frac{\tau_{3k}(r)}{r^{1/2}} \max_{(c,r)=1} \left(\left| \sum_{a=1}^q 1_{(a,q)=1} \left(1_{a \equiv c \pmod{r}} - \frac{1}{r} \right) \right| \right)^{1/2} \\ &\ll q^{\frac{1}{2}} \left(\sum_{r \leq q^{1/2-\epsilon}} \frac{\tau_{3k}^2(r)}{r} \right)^{1/2} \left(\sum_{\substack{r \leq q^{1/2-\epsilon} \\ (r,q)=1}} \max_{(c,r)=1} \left| \sum_{a=1}^q 1_{(a,q)=1} \left(1_{a \equiv c \pmod{r}} - \frac{1}{r} \right) \right| \right)^{1/2} \\ &\ll_k q^{\frac{1}{2}} (\log q)^{\frac{9k^2}{2}} \left(\sum_{\substack{r \leq q^{1/2-\epsilon} \\ (r,q)=1}} \max_{(c,r)=1} \left| \sum_{a=1}^q 1_{(a,q)=1} \left(1_{a \equiv c \pmod{r}} - \frac{1}{r} \right) \right| \right)^{1/2}. \end{aligned}$$

Applying Lemma 5.1, this is $\ll_k q^{7/8}$, and thus the proposition follows. \square

Proof of Proposition 3.2. Similarly, it suffices to estimate

$$\sum_{\substack{a=1 \\ (a,q)=1}}^q 1_{\mathbb{P}}(a + qh_k) 1_{a \equiv b_0 \pmod{W_q}} \prod_{i=1}^{k-1} \lambda_{f_i}(a + qh_i) \lambda_{g_i}(a + qh_i). \quad (5.4)$$

Then, after interchanging the order of summation, this becomes

$$\sum_{d_1, d'_1, \dots, d_k, d'_k \geq 1} \left(\prod_{i=1}^k \mu(d_i) \mu(d'_i) f_i \left(\frac{\log d_i}{\log q} \right) g_i \left(\frac{\log d'_i}{\log q} \right) \right) \tilde{S}(\mathbf{d}, \mathbf{d}'),$$

where

$$\tilde{S}(\mathbf{d}, \mathbf{d}') := \sum_{\substack{a=1 \\ [d_i, d'_i] \mid a + qh_i, 1 \leq i \leq k \\ a \equiv b_0 \pmod{W_q}}}^q 1_{\mathbb{P}}(a + qh_k).$$

Arguing as before, there exists a unique reduced residue $c_0 \pmod{r_{W_q, \mathbf{d}, \mathbf{d}'}}$ such that

$$\begin{aligned} \tilde{S}(\mathbf{d}, \mathbf{d}') &= \frac{1}{\varphi(r_{W_q, \mathbf{d}, \mathbf{d}'})} \sum_{a=1}^q 1_{\mathbb{P}}(a + qh_k) \\ &\quad + \sum_{a=1}^q 1_{\mathbb{P}}(a + qh_k) \left(1_{a \equiv c_0 \pmod{r_{W_q, \mathbf{d}, \mathbf{d}'}}} - \frac{1}{\varphi(r_{W_q, \mathbf{d}, \mathbf{d}'})} \right). \end{aligned}$$

Therefore, by the prime number theorem, the main term of (5.4) is

$$\frac{(1+o(1))q}{\varphi(W_q)\log q} \sum_{\substack{d_1, d'_1, \dots, d_k, d'_k \geq 1 \\ q, W_q, [d_i, d'_i] \text{ pairwise coprime}}} \cdots \sum_{\substack{d_1, d'_1, \dots, d_k, d'_k \geq 1 \\ q, W_q, [d_i, d'_i] \text{ pairwise coprime}}} \prod_{i=1}^{k-1} \frac{\mu(d_i)\mu(d'_i)}{\varphi([d_i, d'_i])} f_i\left(\frac{\log d_i}{\log q}\right) g_i\left(\frac{\log d'_i}{\log q}\right),$$

which is also

$$\frac{(1+o(1))q}{\varphi(W_q)\log q} \sum_{\substack{d_1, d'_1, \dots, d_k, d'_k \geq 1 \\ q_W, W, [d_i, d'_i] \text{ pairwise coprime}}} \cdots \sum_{\substack{d_1, d'_1, \dots, d_k, d'_k \geq 1 \\ q_W, W, [d_i, d'_i] \text{ pairwise coprime}}} \prod_{i=1}^{k-1} \frac{\mu(d_i)\mu(d'_i)}{\varphi([d_i, d'_i])} f_i\left(\frac{\log d_i}{\log q}\right) g_i\left(\frac{\log d'_i}{\log q}\right),$$

where recall that $q_W = q/(q, W)$. Then, applying [Pol14, Lemma 4.1], this becomes

$$\begin{aligned} & (\mathfrak{J}_k(\mathbf{f}, \mathbf{g}) + o(1)) \frac{q}{\varphi(W_q)\log q} \left(\frac{\varphi(W)}{W} \times \frac{\varphi(q_W)}{q_W} \times \log q \right)^{-(k-1)} \\ &= (\mathfrak{J}_k(\mathbf{f}, \mathbf{g}) + o(1)) \frac{q}{\varphi(W_q)\log q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \times \log q \right)^{-(k-1)}, \end{aligned}$$

where

$$\mathfrak{J}_k(\mathbf{f}, \mathbf{g}) := \prod_{i=1}^{k-1} \int_0^\infty f'_i(t_i) g'_i(t_i) dt_i.$$

On the other hand, note that for any $r \geq 1$ with $(r, q) = 1$ we have

$$\#\{(\mathbf{d}, \mathbf{d}') : r = r_{W_q, \mathbf{d}, \mathbf{d}'}\} \leq \tau_{3k-3}(r),$$

and the product $\prod_{i=1}^{k-1} f_i\left(\frac{\log d_i}{\log q}\right) g_i\left(\frac{\log d'_i}{\log q}\right)$ vanishes unless

$$\begin{aligned} r_{W_q, \mathbf{d}, \mathbf{d}'} &\leq W_q \times q \left(\sum_{i=1}^{k-1} \frac{\log d_i}{\log q} + \sum_{i'=1}^{k-1} \frac{\log d'_{i'}}{\log q} \right) \\ &\leq q^{\frac{1}{2}-\epsilon}. \end{aligned}$$

Therefore, the error term of (5.4) is

$$\leq \sum_{\substack{r \leq q^{1/2-\epsilon} \\ (r, q) = 1}} \tau_{3k-3}(r) \max_{(c, r) = 1} \left| \sum_{a=1}^q \mathbb{1}_{\mathbb{P}}(a + qh_k) \left(\mathbb{1}_{a \equiv c \pmod{r}} - \frac{1}{\varphi(r)} \right) \right|. \quad (5.5)$$

Since

$$\max_{(c, r) = 1} \left| \sum_{a=1}^q \mathbb{1}_{\mathbb{P}}(a + qh_k) \left(\mathbb{1}_{a \equiv c \pmod{r}} - \frac{1}{\varphi(r)} \right) \right| \ll \frac{q}{\varphi(r)},$$

it follows from the Cauchy–Schwarz inequality that (5.5) is

$$\begin{aligned}
&\ll q^{\frac{1}{2}} \sum_{\substack{r \leq q^{1/2-\epsilon} \\ (r,q)=1}} \frac{\tau_{3k-3}(r)}{\varphi(r)^{1/2}} \max_{(c,r)=1} \left(\left| \sum_{a=1}^q 1_{\mathbb{P}}(a+qh_k) \left(1_{a \equiv c \pmod{r}} - \frac{1}{\varphi(r)} \right) \right| \right)^{1/2} \\
&\ll q^{\frac{1}{2}} \left(\sum_{r \leq q^{1/2-\epsilon}} \frac{\tau_{3k-3}^2(r)}{\varphi(r)} \right)^{1/2} \left(\sum_{\substack{r \leq q^{1/2-\epsilon} \\ (r,q)=1}} \max_{(c,r)=1} \left| \sum_{a=1}^q 1_{\mathbb{P}}(a+qh_k) \left(1_{a \equiv c \pmod{r}} - \frac{1}{\varphi(r)} \right) \right| \right)^{1/2} \\
&\ll_k q^{\frac{1}{2}} (\log q)^{\frac{9(k-1)^2}{2}} \left(\sum_{\substack{r \leq q^{1/2-\epsilon} \\ (r,q)=1}} \max_{(c,r)=1} \left| \sum_{a=1}^q 1_{\mathbb{P}}(a+qh_k) \left(1_{a \equiv c \pmod{r}} - \frac{1}{\varphi(r)} \right) \right| \right)^{1/2},
\end{aligned}$$

which is $\ll_k q(\log q)^{-100k^2}$ by the Bombieri–Vinogradov theorem. Therefore, the proposition follows. \square

Proof of Proposition 3.3. Similarly, it suffices to estimate

$$\sum_{\substack{a=1 \\ (a,q)=1}}^q \left(\sum_{\substack{p|a+qh_k \\ p \leq q^{\rho}}} 1 \right) 1_{a \equiv b_0 \pmod{W_q}} \prod_{i=1}^k \lambda_{f_i}(a+qh_i) \lambda_{g_i}(a+qh_i). \quad (5.6)$$

Adapting the proof of [Pol14, Proposition 4.2], one can show that

$$\begin{aligned}
&\sum_{\substack{a=1 \\ (a,q)=1}}^q 1_{p|a+qh_k} 1_{a \equiv b_0 \pmod{W_q}} \prod_{i=1}^k \lambda_{f_i}(a+qh_i) \lambda_{g_i}(a+qh_i) \\
&\ll \frac{\log p}{p \log q} \times \frac{\varphi(q)}{W_q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \times \log q \right)^{-k}.
\end{aligned}$$

Then, by summing over primes p , the expression (5.6) is

$$\ll \rho \times \frac{\varphi(q)}{W_q} \left(\frac{\varphi(W_q)}{W_q} \times \frac{\varphi(q)}{q} \times \log q \right)^{-k},$$

and the proposition follows. \square

6. PROOF OF THEOREM 2.2

Let $h := Cm \exp(4m)$, where $C > 0$ is the absolute constant in Theorem 2.1. Then using Lemma 1.1 followed by the triangle inequality, there exists the least integer $1 \leq q \leq \mu(B(x_0; \epsilon/4h))^{-1}$ such that

$$d(T^q x_0, x_0) < \frac{\epsilon}{2h}.$$

Let $\mathcal{A} := \{1 \leq a \leq q : p_m(q, a) \leq qh, (a, q) = 1\}$ and $\mathcal{X}_{\mathcal{A}} := \bigcup_{a \in \mathcal{A}} B(T^a x_0; \epsilon/4h)$. Note that $\mathcal{X}_{\mathcal{A}}$ is in fact a disjoint union by the minimality of q . Then, for any $1 \leq i \leq m, a \pmod{q} \in \mathcal{A}$ and $x \in \mathcal{X}_{\mathcal{A}}$, we have

$$\begin{aligned}
d(T^{p_i(q,a)} x_0, x) &\leq d(T^{p_i(q,a)} x_0, T^a x_0) + d(T^a x_0, x) \\
&< d(T^{p_i(q,a)} x_0, T^a x_0) + \frac{\epsilon}{4h}.
\end{aligned} \quad (6.1)$$

Since T is an isometry, we have

$$\begin{aligned} d(T^{p_i(q,a)}x_0, T^a x_0) &= d(T^{p_i(q,a)-a}x_0, x_0) \\ &\leq \sum_{1 \leq j \leq \frac{p_i(q,a)-a}{q}} d(T^{jq}x_0, T^{(j-1)q}x_0) \\ &= \frac{p_i(q,a) - a}{q} \times d(T^q x_0, x_0) \\ &< \frac{\epsilon}{2}. \end{aligned}$$

Then, it follows from (6.1) that $d(T^{p_i(q,a)}x_0, x) < \epsilon$, and thus $p_m^T(x_0, x; \epsilon) \leq qh$.

On the other hand, since $T^{-a}(B(T^a x_0; \epsilon/4h)) = B(x_0; \epsilon/4h)$ and T is measure-preserving, we have $\mu(B(T^a x_0; \epsilon/4h)) = \mu(B(x_0; \epsilon/4h))$, which implies that

$$\begin{aligned} \mu(\mathcal{X}_{\mathcal{A}}) &= \#\mathcal{A} \times \mu(B(x_0; \epsilon/4h)) \\ &\gg_m \varphi(q)(1 + \log q)^{-C' \exp(4m)} \times \mu(B(x_0; \epsilon/4h)) \end{aligned}$$

by Theorem 2.1. Then the theorem follows from the pseudo-doubling property of the probability measure μ .

7. FIRST RETURN TIME

Given any $\epsilon > 0$, $x_0 \in X$ and $x \in B(x_0; \epsilon)$, let $\tau_{x_0; \epsilon}(x) := \min\{n \geq 1 : T^n x \in B(x_0; \epsilon)\}$ denote the *first return time*. For instance, the least integer $q \geq 1$ for which $d(T^q x_0, x_0) < \epsilon/2h$ in the last proof is the first return time $\tau_{x_0; \epsilon/2h}(x_0)$. Suppose the measure-preserving map T is ergodic. Then Kac's lemma [Kac47] states that

$$\frac{1}{\mu(B(x_0; \epsilon/2h))} \int_X \tau_{x_0; \epsilon/2h}(x) d\mu(x) = \mu(B(x_0; \epsilon/2h))^{-1},$$

i.e. the expected return time is $\mu(B(x_0; \epsilon/2h))^{-1}$. Therefore, given a ‘‘typical’’ measure-preserving map T , one should at least expect that

$$\mu(B(x_0; \epsilon/2h))^{-1} (\log \mu(B(x_0; \epsilon/2h))^{-1})^{-100} \ll \tau_{x_0; \epsilon/2h}(x) \ll \mu(B(x_0; \epsilon/2h))^{-1} \quad (7.1)$$

for some $x \in B(x_0; \epsilon/2h)$. In fact, this is indeed true provided that the map T is sufficiently mixing, so that the first return time follows the exponential distribution (see [GS97], [HSV99], [Sau09] and [Hay13] for details).

Suppose now (7.1) holds. Let $q' := \tau_{x_0; \epsilon/2h}(x)$. Then by following the last proof with additional triangle inequalities, one can show that

$$\begin{aligned} \mu(\{x \in X : p_m^T(x_0, x; \epsilon) \leq C \mu(B(x_0; \epsilon))^{-1} m^{\log_2(2\lambda)} \exp(4 \log_2(2\lambda)m)\}) \\ \gg_{m, \lambda} \varphi(q')(1 + \log q')^{-C' \exp(4m)} \mu(B(x_0; \epsilon)) \\ \gg_{m, \lambda} (\log \mu(B(x_0; \epsilon))^{-1})^{-C'' \exp(4m)} \end{aligned}$$

for some absolute constant $C'' > 0$.

Unfortunately, none of the circle rotations are even weakly mixing (see [EW11, p. 51] for instance). In fact, one can verify that the assumption (7.1) fails for those circle rotations $T_\alpha : x \mapsto x + \alpha$ with α of type $\eta_\alpha > 1$,⁴ where

$$\eta_\alpha := \sup\{\theta > 0 : \liminf_{n \rightarrow \infty} n^\theta \|n\alpha\| = 0\}.$$

⁴By the Borel–Cantelli lemma, the set of irrationals α of type $\eta_\alpha > 1$ has Lebesgue measure 0.

More precisely, let $\tau_\epsilon(\alpha) := \min\{n \geq 1 : \|n\alpha\| < \epsilon\}$. Then Choe–Seo [CS01] proved that

$$\liminf_{\epsilon \rightarrow 0^+} \frac{\log \tau_\epsilon(\alpha)}{\log \epsilon^{-1}} = \eta_\alpha^{-1}.$$

Nevertheless, we are able to establish tight lower and upper bounds for the first return time for most $\alpha \in (0, 1)$ of type $\eta_\alpha \leq 1$, thereby strengthening [KS03, Theorem 3].

Proposition 7.1. *Let $\alpha \in (0, 1)$, $\delta > 0$ and $\epsilon > 0$ sufficiently small (in terms of α). Then*

(1) *for almost all $\alpha \in (0, 1)$ with respect to the Lebesgue measure, we have*

$$\epsilon^{-1}(\log \epsilon^{-1})^{-2(1+\delta)} \ll_\delta \tau_\epsilon(\alpha) \ll \epsilon^{-1};$$

(2) *the irrational $\alpha \in (0, 1)$ has bounded partial quotients⁵ if and only if*

$$\tau_\epsilon(\alpha) \asymp_\alpha \epsilon^{-1}.$$

In particular, the equation (7.1) holds, allowing us to derive the following (metric) Diophantine approximation result on early visits at prime times.

Corollary 7.1. *Let $m \geq 2$ and $\alpha \in (0, 1)$. Then for any sufficiently small $\epsilon > 0$ (in terms of m, α), there exist absolute constants $C, C' > 0$ such that*

(1) *for almost all $\alpha \in (0, 1)$ with respect to the Lebesgue measure μ , we have*

$$\mu(\{\beta \in [0, 1] : p_m^\alpha(\beta; \epsilon) \leq C\epsilon^{-1}m^2 \exp(8m)\}) \gg_m (\log \epsilon^{-1})^{-C' \exp(4m)};$$

(2) *for $\alpha \in (0, 1)$ with bounded partial quotients, we also have*

$$\mu(\{\beta \in [0, 1] : p_m^\alpha(\beta; \epsilon) \leq C\epsilon^{-1}m^2 \exp(8m)\}) \gg_{m,\alpha} (\log \epsilon^{-1})^{-C' \exp(4m)}.$$

Proof of Proposition 7.1. We begin with the proof of the first part. The upper bound is an immediate consequence of the Dirichlet approximation theorem. For the lower bound, since the infinite series $\sum_{q=1}^{\infty} \frac{1}{s(\log s)^{1+\delta}}$ converges, a standard application of the Borel–Cantelli lemma yields for almost all (irrational) $\alpha \in [0, 1]$, there are only finitely many pairs of positive integers $(r_1, s_1), \dots, (r_k, s_k)$ satisfying

$$\left| \alpha - \frac{r_i}{s_i} \right| \leq \frac{1}{s_i^2 (\log s_i)^{1+\delta}}$$

for $i = 1, \dots, k$, and we denote $s_\alpha := \max_{i=1, \dots, k} s_i$. For each of these irrational α , let $p_1/q_1, p_2/q_2, \dots$ with $q_1 < q_2 < \dots$ be the convergents of the continued fraction for α and $q_\alpha := \min\{q_n : q_n > s_\alpha\}$. Suppose $0 < \epsilon < \frac{1}{q_\alpha (\log q_\alpha)^{1+\delta}}$. Then there exists $q_n \geq q_\alpha$ such that

$$\frac{1}{q_{n+1} (\log q_{n+1})^{1+\delta}} \leq \epsilon < \frac{1}{q_n (\log q_n)^{1+\delta}}. \quad (7.2)$$

Through our construction, we have

$$\left| \alpha - \frac{p_n}{q_n} \right| > \frac{1}{q_n^2 (\log q_n)^{1+\delta}} > \frac{\epsilon}{q_n}. \quad (7.3)$$

Therefore, it follows from the best rational approximation property of convergents (see [EW11, Proposition 3.3] for instance) that

$$\tau_\epsilon(\alpha) > q_n.$$

⁵An irrational α with continued fraction expansion $[a_0; a_1, a_2, \dots]$ is said to have bounded partial quotients if a_n is uniformly bounded for $n \geq 1$. They are badly approximable by rationals, and vice versa, i.e. there exists a constant $c > 0$ such that $|\alpha - \frac{p}{q}| > \frac{c}{q^2}$ for all $\frac{p}{q} \in \mathbb{Q}$. For instance, this includes quadratic irrationals but excludes the base of the natural logarithm, $e = [2; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, \dots]$.

It remains to give a lower bound for q_n . Since

$$\left| \alpha - \frac{p_n}{q_n} \right| \leq \frac{1}{q_n q_{n+1}}$$

(see [EW11, p. 72] for instance), combining with (7.3) gives

$$\frac{1}{q_n q_{n+1}} > \frac{1}{q_n^2 (\log q_n)^{1+\delta}} > \frac{\epsilon}{q_n},$$

which implies $q_{n+1} < q_n (\log q_n)^{1+\delta} < q_n (\log q_{n+1})^{1+\delta}$ and $q_n < q_{n+1} < \epsilon^{-1}$. Also, we have $q_{n+1} (\log q_{n+1})^{1+\delta} \geq \epsilon^{-1}$ by (7.2), so that

$$q_n > \frac{q_{n+1}}{(\log q_n)^{1+\delta}} = \frac{q_{n+1} (\log q_n)^{1+\delta}}{(\log q_n)^{2(1+\delta)}} > \epsilon^{-1} (\log \epsilon^{-1})^{-2(1+\delta)}.$$

Therefore, the first part of the proposition follows.

To prove the “if” part of (2), one can simply adapt the above argument with (7.2) and (7.3) replaced by

$$\frac{1}{q_{n+3}} \leq \epsilon < \frac{1}{q_{n+2}}$$

and

$$\left| \alpha - \frac{p_n}{q_n} \right| > \frac{1}{q_n q_{n+2}} > \frac{\epsilon}{q_n}$$

respectively. Then one can show that $q_n \geq (A+1)^{-3} \epsilon^{-1}$, where $A := \max_{n \geq 1} a_n$.

It remains to prove the “only if” part of (2). Let $\alpha \in (0, 1)$ be an irrational number with unbounded partial quotients and denote $q_{n_k} := \min\{q_n : q_{n+1} > kq_n\}$ for $k \geq 1$. Note that such q_{n_k} always exists. Take $\epsilon_k := q_{n_{k+1}}^{-1}$ for $k \geq 1$. Then by definition

$$\tau_{\epsilon_k}(\alpha) \leq q_{n_k} < \frac{q_{n_{k+1}}}{k} = \frac{\epsilon_k^{-1}}{k}.$$

Therefore, it is impossible that $\tau_\epsilon(\alpha) \gg_\alpha \epsilon^{-1}$ as $\epsilon \rightarrow 0^+$, and the proof is completed. \square

ACKNOWLEDGEMENTS

The authors are grateful to Andrew Granville and Dimitris Koukoulopoulos for their advice and encouragement. They would also like to thank Nikos Frantzikinakis and Bryna Kra for their helpful comments.

REFERENCES

- [BZ16] R. C. Baker and L. Zhao. “Gaps of smallest possible order between primes in an arithmetic progression”. In: *Int. Math. Res. Not. IMRN* 23 (2016), pp. 7341–7368.
- [CS01] G. H. Choe and B. K. Seo. “Recurrence speed of multiples of an irrational number”. In: *Proc. Japan Acad. Ser. A Math. Sci.* 77.7 (2001), pp. 134–137.
- [EH70] P. D. T. A. Elliott and H. Halberstam. “A conjecture in prime number theory”. In: *Symposia Mathematica, Vol. IV (INDAM, Rome, 1968/69)*. Academic Press, London-New York, 1970, pp. 59–72.
- [EW11] M. Einsiedler and T. Ward. *Ergodic theory with a view towards number theory*. Vol. 259. Graduate Texts in Mathematics. Springer-Verlag London, Ltd., London, 2011, pp. xviii+481.
- [Gil+20] N. Gillman et al. “Patterns of primes in the Sato-Tate conjecture”. In: *Res. Number Theory* 6.1 (2020), Paper No. 9, 30.

- [GPY09] D. A. Goldston, J. Pintz, and C. Y. Yıldırım. “Primes in tuples. I”. In: *Ann. of Math. (2)* 170.2 (2009), pp. 819–862.
- [GS97] A. Galves and B. Schmitt. “Inequalities for hitting times in mixing dynamical systems”. In: *Random Comput. Dynam.* 5.4 (1997), pp. 337–347.
- [Hay13] N. T.A. Haydn. “Entry and return times distribution”. In: *Dynamical Systems* 28.3 (2013), pp. 333–353.
- [HSV99] M. Hirata, B. Saussol, and S. Vaienti. “Statistics of return times: a general framework and new applications”. In: *Comm. Math. Phys.* 206.1 (1999), pp. 33–55.
- [Kac47] M. Kac. “On the notion of recurrence in discrete stochastic processes”. In: *Bull. Amer. Math. Soc.* 53 (1947), pp. 1002–1010.
- [KS03] D. H. Kim and B. K. Seo. “The waiting time for irrational rotations”. In: *Nonlinearity* 16.5 (2003), pp. 1861–1868.
- [Leu24a] S.-K. Leung. *Moments of primes in progressions to a large modulus*. 2024. arXiv: [2402.07941](https://arxiv.org/abs/2402.07941).
- [Leu24b] S.-K. Leung. *Pseudorandomness of primes at large scales*. 2024. arXiv: [2407.05380](https://arxiv.org/abs/2407.05380).
- [LPS17] J. Li, K. Pratt, and G. Shakan. “A lower bound for the least prime in an arithmetic progression”. In: *Q. J. Math.* 68.3 (2017), pp. 729–758.
- [May15] J. Maynard. “Small gaps between primes”. In: *Ann. of Math. (2)* 181.1 (2015), pp. 383–413.
- [May16] J. Maynard. “Dense clusters of primes in subsets”. In: *Compos. Math.* 152.7 (2016), pp. 1517–1554.
- [Pol14] D. H. J. Polymath. “Variants of the Selberg sieve, and bounded intervals containing many primes”. In: *Res. Math. Sci.* 1 (2014), Art. 12, 83.
- [Pom80] C. Pomerance. “A note on the least prime in an arithmetic progression”. In: *J. Number Theory* 12.2 (1980), pp. 218–223.
- [Sau09] B. Saussol. “An Introduction to Quantitative Poincaré Recurrence in Dynamical Systems”. In: *Reviews in Mathematical Physics* 21.08 (2009), pp. 949–979.
- [ST21] X. Shao and J. Teräväinen. “The Bombieri-Vinogradov theorem for nilsequences”. In: *Discrete Anal.* (2021), Paper No. 21, 55.
- [Sta23] J. Stadlmann. *On primes in arithmetic progressions and bounded gaps between many primes*. 2023. arXiv: [2309.00425](https://arxiv.org/abs/2309.00425).
- [Tho14] J. Thorner. “Bounded gaps between primes in Chebotarev sets”. In: *Res. Math. Sci.* 1 (2014), Art. 4, 16.
- [Zha14] Y. Zhang. “Bounded gaps between primes”. In: *Ann. of Math. (2)* 179.3 (2014), pp. 1121–1174.

DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, UNIVERSITÉ DE MONTRÉAL, CP 6128
 SUCC. CENTRE-VILLE, MONTRÉAL, QC H3C 3J7, CANADA
Email address: tony.haddad@umontreal.ca

DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, UNIVERSITÉ DE MONTRÉAL, CP 6128
 SUCC. CENTRE-VILLE, MONTRÉAL, QC H3C 3J7, CANADA
Email address: sun.kai.leung@umontreal.ca

DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE, UNIVERSITÉ DE MONTRÉAL, CP 6128
 SUCC. CENTRE-VILLE, MONTRÉAL, QC H3C 3J7, CANADA
Email address: cihan.sabuncu@umontreal.ca