Generalized Thue-Morse Sequences of Squares

Michael Drmota and Johannes F. Morgenbesser*

August 25, 2010

Abstract

We consider compact group generalizations $T(n)$ of the Thue-Morse sequence and prove that the subsequence $T(n^2)$ is uniformly distributed with respect to a measure ν that is absolutely continuous with respect to the Haar measure. The proof is based on a proper generalization of the Fourier based method of Mauduit and Rivat in their study of the sum-of-digits function of squares to group representations.

1 Introduction

The Thue-Morse sequence

$$(t_n)_{n\geqslant 0} = (01101001100101101001011001101001\ldots)$$

has been discovered several times in the literature. (For a survey see [12].) There are also several different definitions. For example, we have

$$t_n = s_2(n) \bmod 2,$$

where $s_2(n)$ denotes the number of 1's in the binary expansion of n. Alternatively, we can use recursive definitions like $t_0 = 0$, $t_{2k} = t_k$, $t_{2k+1} = 1 - t_k$ or identify it with a fixed point of the morphism $\mu : \{0,1\}^* \to \{0,1\}^*$ induced by $\mu(0) = 01$ and $\mu(1) = 10$ (see [7]). In any case, the binary expansion of n governs the behavior of t_n.

The Thue-Morse sequence has many interesting properties. For example, it is cube-free (that is, there is no subword of the form www) and every subword w that occurs once appears infinitely often with bounded gaps (although it is non-periodic). It is also an automatic sequence (see [4] and Section 3).

By definition it is clear that the Thue-Morse sequence has the property that the digits 0 and 1 appear with asymptotic frequency 1/2. Interestingly, this property persists for subsequences like linear progressions (see [8]). It has been a long standing conjecture (attributed to Gelfond [8]) that t_p, p prime, and t_{n^2} have the same property. Recently Mauduit and Rivat [13, 14] could settle these questions.

The purpose of this paper is to establish a distribution result for the quadratic subsequence $T(n^2)$, where $T(n)$ is a generalized Thue-Morse sequence of the following type. Let H be a compact group that satisfies the Hausdorff separation axiom, $q \geqslant 2$, and $g_0, g_1, \ldots, g_{q-1} \in H$ with $g_0 = e$ the identity element. Furthermore, let $G \leqslant H$ be the closure of the subgroup generated by $g_0, g_1, \ldots, g_{q-1}$, i.e., G is the smallest closed set

*The authors are supported by the Austrian Science Foundation FWF, grant S9604, that is part of the National Research Network "Analytic Combinatorics and Probabilistic Number Theory".

in H that contains the subgroup generated by $g_0, g_1, \ldots, g_{q-1}$ (note, that this is again a group). Suppose that

$$\begin{aligned} n &= \varepsilon_{\ell-1}(n)q^{\ell-1} + \varepsilon_{\ell-2}(n)q^{\ell-2} + \cdots + \varepsilon_1(n)q + \varepsilon_0(n) \\ &= (\varepsilon_{\ell-1}(n)\varepsilon_{\ell-2}(n)\ldots\varepsilon_1(n)\varepsilon_0(n))_q \end{aligned}$$

denotes the q-ary digital expansion of n and define

$$T(n) = g_{\varepsilon_0(n)} g_{\varepsilon_1(n)} \cdots g_{\varepsilon_{\ell-1}(n)}. \tag{1}$$

If $G = \mathbb{Z}/2\mathbb{Z}$ (with $+$ as the group operation), $q = 2$, and $g_0 = 0$, $g_1 = 1$, then $T(n) = s_2(n) \bmod 2 = t_n$. Thus, $T(n)$ is a proper generalization of the Thue-Morse sequence. Alternatively $T(n)$ can be seen as a completely q-multiplicative G-valued function which is defined by the property

$$T(j + qn) = T(j)T(n)$$

for $n \geqslant 0$ and $0 \leqslant j < q$. The sequence $T(n)$ is also an example of a so-called *chained sequence* with a transition matrix that is not contractive (see [1] and [2]).

It is relatively easy to show (see Theorem 2) that the sequence $(T(n))_{n\geqslant 0}$ is uniformly distributed in G, that is, the normalized counting measure induced by $T(n)$, $n < N$, converges weakly to the (normalized) Haar measure μ on G:[1]

$$\frac{1}{N}\sum_{n=0}^{N-1} \delta_{T(n)} \to \mu.$$

Our main result deals with the question whether this remains true if $T(n)$ is replaced by the subsequence of squares $T(n^2)$. Actually, this sequence is not necessarily uniformly distributed. Nevertheless, there is always a measure ν such that $T(n^2)$ is ν-uniformly distributed.

Theorem 1. *Let $T(n)$ be defined by (1). Then there exists a positive integer m depending on $g_0 = e, g_1, \ldots, g_{q-1}$ and q with $m \mid q-1$ such that the following holds.*

The group[2] $U = \mathrm{cl}(\{T(mn) : n \geqslant 0\})$ is a normal subgroup of G of index m with cosets $g_uU = \mathrm{cl}(\{T(mn+u) : n \geqslant 0\})$, $0 \leqslant u < m$. With the help of these cosets we define

$$\mathrm{d}\nu = \sum_{u=0}^{m-1} \mathbf{1}_{g_uU} \cdot Q(u,m)\,\mathrm{d}\mu,$$

where $Q(u,m) = \#\{0 \leqslant n < m : n^2 \equiv u \bmod m\}$ and μ denotes the Haar measure on G.

Then the sequence $(T(n^2))_{n\geqslant 0}$ is ν-uniformly distributed in G, that is,

$$\frac{1}{N}\sum_{n=0}^{N-1} \delta_{T(n^2)} \to \nu.$$

Remark 1. The integer m that we will call *characteristic integer of $g_0, \ldots, g_{q-1}$ and q* (see Section 2) is defined as the largest integer such that $m \mid q-1$ and such that there exists a one-dimensional representation D of G with

$$D(g_u) = \mathrm{e}\left(-\frac{u}{m}\right) \qquad \text{for all } u \in \{0, 1, \ldots, q-1\}.$$

[1] δ_x denotes the point measure concentrated at x.

[2] If $A \subseteq G$, then $\mathrm{cl}(A)$ denotes the topological closure of A in G.

Note also that if $m = 1$ or $m = 2$ then $\nu = \mu$. Hence, if $m \leqslant 2$ then $(T(n^2))_{n\geqslant 0}$ is uniformly distributed in G. In particular if $q = 2$ or $q = 3$ then $m \leqslant 2$. Furthermore it is easy to observe that $\nu \neq \mu$ for $m > 2$, that is, $T(n^2)$ is not uniformly distributed in these cases.

Remark 2. Theorem 1 is a generalization of the results of Mauduit and Rivat [14]. Suppose first that $H = \mathbb{Z}/r\mathbb{Z}$ and $g_j = j \bmod r$, $0 \leqslant j < q$. Then $T(n) = s_q(n) \bmod r$ ($s_q(n) = \varepsilon_{\ell-1}(n) + \varepsilon_{\ell-2}(n) + \cdots + \varepsilon_1(n) + \varepsilon_0(n)$ denotes the q-ary sum-of-digits function) and Theorem 1 translates into Théorème 3 from [14] on the distribution of $s_q(n^2)$ modulo r (the characteristic integer is given by $(q-1, r)$, see Section 3).

Similarly, if $H = \mathbb{R}/\mathbb{Z}$ and $g_j = \alpha j \bmod 1$, $0 \leqslant j < q$ for some irrational number α then $T(n) = \alpha s_q(n) \bmod 1$. It is easy to observe that $G = H = \mathbb{R}/\mathbb{Z}$. Hence, Theorem 1 implies that $(\alpha s_q(n^2))_{n\geqslant 0}$ is uniformly distributed modulo 1 (for example, one can use the fact that H is connected, see Remark 4); this is Théorème 2 from [14].

Remark 3. It is easy to derive some corollaries from Theorem 1. For example we have

$$\lim_{N\to\infty} \frac{1}{N} \#\{0 \leqslant n < N : T(n^2) \in g_u U\} = \frac{Q(u,m)}{m},$$

for all $0 \leqslant u < q$.

A similar idea applies to compact homogeneous spaces X. Let H be the group acting on X and suppose that $g_0, \ldots, g_{q-1}$ are chosen in a way that $G = H$. Then it follows that for every $x_0 \in X$ the sequence $x_n = T(n) \cdot x_0$ is uniformly distributed on X and the distribution behavior of x_{n^2} can be determined, too. For example, with the help of this approach we can construct uniformly distributed sequences on the sphere S^d.

Remark 4. It follows from the proof of Theorem 1 that the Radon-Nikodym derivative $f(g) = Q(u,m)$ (for $g \in g_u U$) is continuous, which implies that G cannot be connected if $(T(n^2))_{n\geqslant 0}$ is not uniformly distributed. Conversely if G is connected then $(T(n^2))_{n\geqslant 0}$ is definitely uniformly distributed. Similarly, if the commutator subgroup of G (that is, the subgroup generated by the elements $xyx^{-1}y^{-1}$) coincides with G, then $(T(n^2))_{n\geqslant 0}$ is also uniformly distributed (note, that the commutator subgroup is always a subgroup of U).

Remark 5. It would be also of interest to consider the subsequence $(T(p))$, where p runs over all primes. For example, an equidistribution result holds for t_p (see [13]). In order to handle this case one would need estimates of the form

$$\sum_{0\leqslant h<q^\lambda} \|F_\lambda(h)\| \ll q^{\eta\lambda} \tag{2}$$

for some $\eta < 1/2$, where $F_\lambda(h)$ is the Fourier term defined in Section 2. By using the Cauchy-Schwarz inequality it follows directly that (2) holds for $\eta = 1/2$. However, it is not clear how to derive such a general estimate for $\eta < 1/2$. Actually, this is one of the key estimates in [13], where the sum-of-digits function of primes is discussed.

Remark 6. If $g_0 \neq e$, then the sequence $(T(n))_{n\geqslant 0}$ is not q-multiplicative any more. This would not be essential for the proof of the main theorem since it is possible to reduce the function $T(n)$ to the function $T_\lambda(n)$ defined in Section 4.2, which is "almost" completely q-multiplicative even if $g_0 \neq e$ (it satisfies $T_\lambda(j + qn) = g_j T_{\lambda-1}(n)$ for all $0 \leqslant j < q$ and $n \geqslant 0$).

However, the condition $g_0 = e$ is important for the proof of Lemma 2 and Lemma 4. It is only possible to avoid this condition if one assumes instead that the group G is equal to the closure of the subgroup generated by $g_i^{-1}g_j, 0 \leqslant i, j < q$ and that there exists no one-dimensional representation D satisfying

$$D(g_u) = \mathrm{e}(-tu)D(g_0)$$

for all $0 \leqslant u < q$ with $t(q-1) \in \mathbb{Z}$ and $D(g_0) \neq 1$. However, for the sake of brevity we use the assumption $g_0 = e$ in the main theorem, since this yields a considerably simpler presentation of the proof.

The proof of Theorem 1 is based on a proper generalization of the Fourier-based method of Mauduit and Rivat [13, 14] to group representations. In Section 2 we use representation theory to prove uniform distribution of the sequence $(T(n))_{n \geqslant 0}$ and develop the theory to discuss the case of linear subsequences $(T(an+b))_{n \geqslant 0}$ (see Theorem 3). Although linear subsequences are not the main focus of this paper the analysis of them is also useful for the analysis of the quadratic subsequence $(T(n^2))_{n \geqslant 0}$. Interestingly, the characteristic integer m appears there in a quite natural way. In Section 3 we deal with finite groups G in more detail and also show that there is a close relation of $T(n)$ to so-called automatic sequences. Actually, this kind of application was the main motivation of the present study. We introduce the notion of invertible automatic sequences $(u_n)_{n \geqslant 0}$ and show that the frequencies of the letters for the quadratic subsequence $(u_{n^2})_{n \geqslant 0}$ exist. The technical part of the proof of Theorem 1 is presented in Sections 4 and 5, where we first establish some auxiliary results (like matrix generalizations of the techniques used in [14]) and then collect all necessary facts to complete the proof. A final Section 6 discusses possible generalizations and extensions.

Acknowledgment. The authors are very grateful to Christian Mauduit for several hints and comments to improve the presentation of the manuscript.

2 Group Representations

A unitary group representation D of a compact group is a continuous homomorphism $D : G \to U_d$ for some $d \geqslant 1$, where U_d denotes the group of unitary $d \times d$ matrices (over $\mathbb{C}$). A representation is irreducible if there is no proper subspace W of $\mathbb{C}^d$ with $D(x)W \subseteq W$ for all $x \in G$. The trivial representation that maps all elements to 1 will be denoted by D_0. The dimension d will be also called the dimension (or degree) of D.

Irreducible and unitary group representations can be used to prove uniform distribution of a sequence (x_n) in a compact group.

Lemma 1. *Let G be a compact group and ν a regular normed Borel measure in G. Then a sequence $(x_n)_{n \geqslant 0}$ is ν-uniformly distributed in G, that is, $\frac{1}{N}\sum_{n=0}^{N-1} \delta_{x_n} \to \nu$, if and only if*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} D(x_n) = \int_G D \, \mathrm{d}\nu$$

holds for all irreducible unitary representations D of G. In particular, $(x_n)_{n \geqslant 0}$ is uniformly distributed in G (with respect to the Haar measure μ) if and only if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} D(x_n) = 0$$

holds for all irreducible and unitary representations $D \neq D_0$.

Proof. A proof of this lemma can be found for example in [10, Theorem 1.3 and Section 4.3]. □

We will first present a proof that a sequence $T(n)$ of the form (1) is uniformly distributed in G. Let D be a representation of G and let

$$\Psi_D = \sum_{0\leqslant u<q} D(g_u)$$

denote the sum of the matrices $D(g_0), \ldots, D(g_{q-1})$.

We use the following notations for matrices. If A is a matrix, then A^H is the Hermitian transpose, $\rho(A)$ denotes the spectral radius and $\mathrm{tr}(A)$ the trace of A. We use $\|\,.\,\|_2$ for the spectral norm ($\|A\|_2 = \sqrt{\rho(AA^H)}$) and $\|A\|_{\mathbb{F}}$ for the Frobenius norm (i.e. $\|A\|_{\mathbb{F}}^2 = \sum_{i,j} |a_{i\,j}|^2 = \mathrm{tr}(AA^H)$).

Lemma 2. *Let G be a compact group that is the closure of the subgroup generated by the elements $g_0, g_1, \ldots, g_{q-1}$, where $g_0 = e$. Suppose that $D \neq D_0$ is an irreducible and unitary representation of G. Then*

$$\|\Psi_D\|_2 < q. \tag{3}$$

Proof. Suppose that $\mathbf{y} \in \mathbb{C}^d$ is a non-zero vector. Then $\|D(x)\mathbf{y}\|_2 = \|\mathbf{y}\|_2$ for all $x \in G$, and consequently

$$\|\Psi_D\mathbf{y}\|_2 = \left\| \sum_{0\leqslant u<q} D(g_u)\mathbf{y} \right\|_2 \leqslant \sum_{0\leqslant u<q} \|D(g_u)\mathbf{y}\|_2 = q\, \|\mathbf{y}\|_2\,. \tag{4}$$

Hence $\|\Psi_D\|_2 \leqslant q$. Suppose now that $\|\Psi_D\|_2 = q$, that is, there exists a non-zero vector $\mathbf{y}$ with $\|\Psi_D\mathbf{y}\|_2 = q\,\|\mathbf{y}\|_2$. We have

$$\|\Psi_D\mathbf{y}\|_2^2 = \sum_{0\leqslant u,v<q} \langle D(g_u)\mathbf{y}, D(g_v)\mathbf{y}\rangle = q^2 \langle \mathbf{y}, \mathbf{y}\rangle\,.$$

The Cauchy-Schwarz inequality implies

$$|\langle D(g_u)\mathbf{y}, D(g_v)\mathbf{y}\rangle| \leqslant \|D(g_u)\mathbf{y}\|_2 \,\|D(g_v)\mathbf{y}\|_2 = \|\mathbf{y}\|_2^2\,. \tag{5}$$

Since $\langle D(g_0)\mathbf{y}, D(g_0)\mathbf{y}\rangle = \langle \mathbf{y}, \mathbf{y}\rangle$, we have that $\langle D(g_u)\mathbf{y}, D(g_v)\mathbf{y}\rangle = \langle y, y\rangle$ for all $0 \leqslant u, v < q$ and there has to be equality in (5). It follows that $D(g_u)\mathbf{y}$ and $D(g_v)\mathbf{y}$ have to be linear dependent. Since $D(g_0)\mathbf{y} = \mathbf{y}$, we obtain

$$D(g_u)\mathbf{y} = \mathbf{y} \tag{6}$$

for all $u = 0, 1, \ldots, q-1$. Consequently the one-dimensional space $W = \mathrm{span}(\mathbf{y})$ satisfies $D(x)W \subseteq W$ for all $x \in G$. (Recall that G is the closure of the subgroup generated by $g_0, \ldots, g_{q-1}$.) This contradicts the assumption that D is irreducible provided that $d \geqslant 2$. Thus, for all irreducible representations of dimension $d \geqslant 2$ we actually have $\|\Psi_D\|_2 < q$.

Finally suppose that $d = 1$, that is, we are considering characters. Then (6) says that $D(g_u) = 1$ for all $u = 0, 1, \ldots, q-1$. Since G is the closure of the subgroup generated by the elements $g_0, g_1, \ldots, g_{q-1}$ this would imply $D(x) = 1$ for all $x \in G$ which contradicts the assumption $D \neq D_0$. □

With the help of Lemma 2 it is easy to prove that the sequence $(T(n))_{n\geqslant 0}$ is uniformly distributed in G (compare also with [2]).

Theorem 2. *Let $q \geqslant 2$ and $g_0, g_1, \ldots, g_{q-1}$ with $g_0 = e$ be elements of a compact group and G the closure of the subgroup generated by $g_0, g_1, \ldots, g_{q-1}$. Then the sequence $(T(n))_{n\geqslant 0}$ defined by (1) is uniformly distributed in G with respect to its Haar measure μ.*

Proof. Let $D \neq D_0$ be an irreducible and unitary representation of G and recall that $T(n) = g_{\varepsilon_0} g_{\varepsilon_1} \cdots g_{\varepsilon_{\ell-1}}$, where $\varepsilon_{\ell-1}\varepsilon_{\ell-2}\ldots\varepsilon_1\varepsilon_0$ denotes the q-ary digital expansion of n. Let $N \geqslant 1$ be defined by $N = \sum_{\nu=0}^{\lambda} n_\nu q^\nu$ with $n_\lambda \neq 0$. We begin with the following identity

$$\begin{aligned}\sum_{0\leqslant n<N} D(T(n)) &= \sum_{\nu=0}^{\lambda} \sum_{n=0}^{q^\nu-1} \sum_{\varepsilon_\nu=0}^{n_\nu-1} D(T(n+\varepsilon_\nu q^\nu + n_{\nu+1}q^{\nu+1} + \cdots + n_\lambda q^\lambda)) \\ &= \sum_{\nu=0}^{\lambda} \left(\sum_{n=0}^{q^\nu-1} D(T(n)) \right) \left(\sum_{\varepsilon_\nu=0}^{n_\nu-1} D(T(\varepsilon_\nu)) \right) D(g_{n_{\nu+1}}) \cdots D(g_{n_\lambda}).\end{aligned}$$

Since D is a unitary representation and the 2-norm is submultiplicative, we obtain

$$\begin{aligned}\left\| \sum_{0\leqslant n<N} D(T(n)) \right\|_2 &\leqslant \sum_{\nu=0}^{\lambda} \sum_{\varepsilon_\nu=0}^{n_\nu-1} \|D(T(\varepsilon_\nu))\|_2 \left\| \sum_{n=0}^{q^\nu-1} D(T(n)) \right\|_2 \\ &\leqslant q \sum_{\nu=0}^{\lambda} \left\| \sum_{n=0}^{q^\nu-1} D(T(n)) \right\|_2 .\end{aligned}$$

By induction it follows from the definition of $T(n)$ that $\sum_{n=0}^{q^\nu-1} D(T(n)) = (\Psi_D)^\nu$. Lemma 2 implies that we have $\|\Psi_D\|_2 = q^\sigma$, where $\sigma < 1$. We finally get[3]

$$\left\| \frac{1}{N} \sum_{0\leqslant n<N} D(T(n)) \right\|_2 \leqslant \frac{q}{N} \sum_{\nu=0}^{\lambda} \|\Psi_D\|_2^\nu \ll \frac{q^{\sigma\lambda}}{N} \to 0$$

as $N \to \infty$. Applying Lemma 1, this proves the theorem. □

It is an interesting problem to generalize Theorem 2 to special subsequences of $T(n)$, for example to linear subsequences $T(an+b)$ or (as it is the main goal of this paper) to the subsequence $T(n^2)$ of squares. First we present a result for linear subsequences. Let D be an irreducible representation of G with degree d. In what follows, we need the function

$$\Psi_D(t) = \sum_{0\leqslant u<q} \mathrm{e}(tu) D(g_u),$$

and the *Fourier terms* (defined for $\lambda \geqslant 0$)

$$F_\lambda(h) = \frac{1}{q^\lambda} \sum_{0\leqslant u<q^\lambda} e\left(-\frac{hu}{q^\lambda}\right) D(T(u)) = \frac{1}{q^\lambda} \Psi_D\left(-\frac{h}{q^\lambda}\right) \Psi_D\left(-\frac{h}{q^{\lambda-1}}\right) \cdots \Psi_D\left(-\frac{h}{q}\right).$$

[3]The symbol $f \ll g$ means that there exists a constant c such that $f \leqslant cg$.

The product representation follows from the fact that

$$\begin{aligned} F_\lambda(h) &= \frac{1}{q^\lambda} \sum_{0\leqslant j<q} \sum_{0\leqslant u<q^{\lambda-1}} \mathrm{e}\left(-\frac{h(uq+j)}{q^\lambda}\right) D(T(uq+j)) \\ &= \frac{1}{q^\lambda} \sum_{0\leqslant j<q} \mathrm{e}\left(-\frac{hj}{q^\lambda}\right) D(g_j) \sum_{0\leqslant u<q^{\lambda-1}} \mathrm{e}\left(-\frac{hu}{q^{\lambda-1}}\right) D(T(u)) \\ &= \frac{1}{q}\,\Psi_D\left(-\frac{h}{q^\lambda}\right) F_{\lambda-1}(h), \end{aligned} \tag{7}$$

and that $F_0(h) = I_d$, where I_d is the identity matrix of dimension d. (Note, that $T(uq+j) = g_j T(u)$ for all $0 \leqslant j < q$ and that g_0 is the identity element.)

If the representation D is one-dimensional and satisfies

$$D(g_u) = \mathrm{e}\left(-\frac{ur}{q-1}\right)$$

for $0 \leqslant u < q$, then we see that D acts on the set $\{T(n) : n \geqslant 0\}$ in a special way. Indeed, since $s_q(n) \equiv n \bmod q-1$, we have

$$D(T(n)) = \mathrm{e}\left(-\frac{s_q(n)r}{q-1}\right) = \mathrm{e}\left(-\frac{nr}{q-1}\right).$$

Actually, these kinds of representations are crucial for the description of the distribution of $T(an+b)$ and $T(n^2)$.

Definition 1. Let m be the largest integer such that $m \mid q-1$ and such that there exists a one-dimensional representation D of G with

$$D(g_u) = \mathrm{e}\left(-\frac{u}{m}\right) \qquad \text{for all } u \in \{0, 1, \ldots, q-1\}. \tag{8}$$

We will call this integer *characteristic integer of* $g_0, \ldots, g_{q-1}$ *and* q.

Observe that this characteristic integer m always exists since the trivial representation fulfills (8) with $m = 1$.

The next Lemma collects some facts of this characteristic integer.

Lemma 3. *Let m be the characteristic integer of $g_0, \ldots, g_{q-1}$ and q. Then there exist m representations $D_0, \ldots, D_{m-1}$ of G with the following properties:*

(i) *Let $0 \leqslant k < m$. Then*

$$D_k(g_u) = \mathrm{e}\left(-\frac{k}{m}u\right) \qquad \text{for all } u \in \{0, 1, \ldots, q-1\}.$$

(ii) *All other representations of G do not satisfy $D(g_1) = \mathrm{e}(-t)$ and $D(g_u) = D(g_1)^u$ for all $0 \leqslant u < q$ with $(q-1)\,t \in \mathbb{Z}$.*

(iii) *The kernel $\ker D_1 = \{g \in G : D_1(g) = 1\}$ is a normal subgroup of G and the index of $\ker D_1$ in G is equal to m.*

(iv) *The m cosets of $\ker D_1$ are given by*

$$g_v \ker D_1 = \mathrm{cl}\left(\{T(mn+v) : n \geqslant 0\}\right) \qquad \text{for all } v \in \{0, 1, \ldots, m-1\}.$$

Proof. Let D be a one-dimensional representation of G that satisfies (8) and set

$$D_k(g) = D(g)^k$$

for all $g \in G$ and for all $0 \leqslant k < m$. Then D_0 (consistent with the already defined notation) is the trivial representation and the representations D_k satisfy the relation stated in (i) for all $0 \leqslant k < m$. (Note, that the functions D_k are indeed representations coming from the (iterated) tensor product of the representation D.) Next, we show that there are no other representations of G with this property. Assume, that there exists a representation $\tilde{D} \neq D_k$ for all $0 \leqslant k < m$ such that

$$\tilde{D}(g_u) = \mathrm{e}\left(-\frac{r}{m'}h\right) \qquad \text{for all} \quad 0 \leqslant u < q,$$

and for some integers $m' \geqslant 1$, $r \geqslant 0$ with $(r, m') = 1$ and $m' \mid q-1$. Then $(m, m') < m'$ and there exist non-negative integers x and y such that $xm' + yrm \equiv (m, m') \bmod mm'$ (note, that $(rm, m') = (m, m')$). If we set

$$\chi(g) = D_1^x(g)\tilde{D}^y(g),$$

then χ is a representation satisfying

$$\chi(g_u) = \mathrm{e}\left(-u\left(\frac{x}{m} + \frac{yr}{m'}\right)\right) = \mathrm{e}\left(-u\frac{m'x + yrm}{mm'}\right) = \mathrm{e}\left(-\frac{u}{\bar{m}}\right),$$

for all $0 \leqslant u < q$, where $\bar{m} = \mathrm{lcm}(m, m')$. Since $\bar{m} = mm'/(m, m') > m$, this is impossible by the definition of m. Thus, we have shown (ii). The kernel of a representation is clearly a normal subgroup and the factor group $G/\ker D_1$ is isomorph to the image of D_1. Let $0 \leqslant v < m$. Then we have

$$D_1(T(mn+v)) = \mathrm{e}\left(-\frac{s_q(mn+v)}{m}\right) = \mathrm{e}\left(-\frac{mn+v}{m}\right) = \mathrm{e}\left(-\frac{v}{m}\right).$$

We see that

$$D_1\left(\mathrm{cl}\left(\{T(mn+v) : n \geqslant 0\}\right)\right) = \mathrm{e}\left(-\frac{v}{m}\right).$$

Since D_1 is continuous and $(T(n))_{n\geqslant 0}$ is dense in G (recall, that the sequence $(T(n))_{n\geqslant 0}$ is uniformly distributed in G), the family $\mathrm{cl}(\{T(mn+v) : n \geqslant 0\})$, $0 \leqslant v < m$ is a partition of G. We obtain that $G/\ker D_1$ is isomorph to the m-*th* roots of unity (and hence to $\mathbb{Z}/m\mathbb{Z}$). Moreover, we see that $\ker D_1 = \mathrm{cl}(\{T(mn) : n \geqslant 0\})$ and m is the index of $\ker D_1$ in G. Since

$$D_1(g_v^{-1}\{T(mn+v) : n \geqslant 0\}) = 1,$$

we finally have that $g_v \ker D_1 = \mathrm{cl}(\{T(mn+v) : n \geqslant 0\})$, $v = 0, \ldots, m-1$ are the m different cosets of $\ker D_1$. $\square$

Lemma 4. *Let G be a compact group that is the closure of the subgroup generated by the elements $g_0 = e, g_1, \ldots, g_{q-1}$. Suppose that $D \notin \{D_0, \ldots, D_{m-1}\}$ is an irreducible and unitary representation of G. Then there exists a constant $c > 0$ such that*

$$\max_{h\in\mathbb{Z}} \|F_\lambda(h)\|_2 \ll q^{-c\lambda}.$$

Remark 7. If $D = D_k$ for some $0 \leqslant k < m$, then $|F_\lambda(h)| = q^{-\lambda}\varphi_{q^\lambda}(h/q^\lambda - k/m)$. (For the definition and the properties of φ_{q^λ} see Lemma 7.) In particular, this implies that

$$\lim_{\lambda\to\infty} \max_{h\in\mathbb{Z}} |F_\lambda(h)| \neq 0,$$

and Lemma 4 cannot hold true in this case.

Proof. The proof is very similar to the proof of Lemma 2. We begin with considering the function $\Psi_D(t)$ for a fixed real number t. Let us first assume that D has dimension d greater than 1. It is clear that $\|\Psi_D(t)\|_2 \leqslant q$. Moreover, if $\|\Psi_D(t)\|_2 = q$ then there exists a non-zero vector $\mathbf{y}$ with $\|\Psi_D(t)\mathbf{y}\|_2 = q\,\|\mathbf{y}\|_2$. By the same reasoning as in Lemma 2, we obtain that

$$e(ut)D(g_u)\mathbf{y} = D(g_0)\mathbf{y} = \mathbf{y} \tag{9}$$

for all $u = 0, 1, \ldots, q-1$. This contradicts the assumption that D is irreducible (we assumed that $n \geqslant 2$). Thus, for all irreducible representations of dimension $d \geqslant 2$ we have $\|\Psi_D(t)\|_2 < q$. If the dimension of D is equal to 1, then $\|\Psi_D(t)\|_2 = q$ means $\mathrm{e}(ut)D(g_u) = 1$ for all $u = 0, 1, \ldots, q-1$ which is equivalent to $D(g_1) = \mathrm{e}(-t)$ and $D(g_u) = D(g_1)^u$. Hence, if this is not true, then we also get $\|\Psi_D(t)\|_2 < q$. If $D(g_1) = \mathrm{e}(-t)$ and $D(g_u) = D(g_1)^u$, then $(q-1)t \notin \mathbb{Z}$ (recall that $D \notin \{D_0, \ldots, D_{m-1}\}$) and we obtain

$$\|\Psi_D(t)\Psi_D(qt)\|_2 < q^2.$$

Indeed, as in the considerations above, the condition $\|\Psi_D(t)\Psi_D(qt)\|_2 = q^2$ would imply $D(g_1) = e(-t) = e(-qt)$ which contradicts the assumption $(q-1)t \notin \mathbb{Z}$.

Now we can finish the proof of Lemma 4. Since the spectral norm $\|\Psi_D(t)\|_2$ is a submultiplicative norm and a continuous function in t, we obtain

$$\sup_{t\in\mathbb{R}} \|\Psi_D(t)\Psi_D(qt)\|_2 < q^2.$$

Using the product representation of F_λ (see (7)), this implies that there exists a constant $c > 0$ such that

$$\max_{h\in\mathbb{Z}} \|F_\lambda(h)\|_2 \ll q^{-c\lambda}.$$

□

Theorem 3. *Let $q \geqslant 2$, $a \geqslant 1$ and $b \geqslant 0$ be integers and m the characteristic integer of $g_0, \ldots, g_{q-1}$ and q. The sequence $(T(an+b))_{n\geqslant 0}$ is uniformly distributed in G (with respect to the Haar measure) if and only if $(a, m) = 1$.*

If $(a, m) > 1$, then there exists a normal subgroup U (with index (a, m) in G) such that $(T(an+b))_{n\geqslant 0}$ is ν-uniformly distributed on a coset of U, where ν is the (translated) Haar measure of U.

Proof. Let D be an irreducible and unitary representation of G with $D \neq D_k$ for all $0 \leqslant k < m$. Furthermore, let the integers λ and β be defined by $q^{\lambda-1} \leqslant N < q^\lambda$ and $a \leqslant q^{\beta-1}$. Then we have (for sufficiently large N) that $aN + b < q^{\lambda+\beta}$. We can write

$$\begin{aligned}
&\sum_{0\leqslant n<N} D(T(an+b)) \\
&\qquad = \sum_{0\leqslant u<q^{\nu+\beta}} \sum_{0\leqslant n<N} D(T(u)) \cdot \frac{1}{q^{\lambda+\beta}} \sum_{0\leqslant h<q^{\lambda+\beta}} \mathrm{e}\left(\frac{h(an+b-u)}{q^{\lambda+\beta}}\right) \\
&\qquad = \sum_{0\leqslant h<q^{\lambda+\beta}} F_{\lambda+\beta}(h) \sum_{0\leqslant n<N} \mathrm{e}\left(\frac{h(an+b)}{q^{\lambda+\beta}}\right).
\end{aligned}$$

The exponential sum can be easily calculated and we obtain

$$\left\| \sum_{0 \leqslant n < N} D(T(an+b)) \right\|_2 \ll \sum_{0 \leqslant h < q^{\lambda+\beta}} \|F_{\lambda+\beta}(h)\|_2 \cdot \min\left(N, \frac{1}{\left|\sin \frac{\pi h a}{q^{\lambda+\beta}}\right|}\right).$$

Since $D \neq D_k$, Lemma 4 implies that $\|F_{\lambda+\beta}(h)\|_2 \ll q^{-c(\lambda+\beta)}$ for some $c > 0$. We get

$$\left\| \sum_{0 \leqslant n < N} D(T(an+b)) \right\|_2 \ll q^{-c(\lambda+\beta)} \sum_{0 \leqslant h < q^{\lambda+\beta}} \min\left(N, \frac{1}{\left|\sin \frac{\pi h a}{q^{\lambda+\beta}}\right|}\right).$$

For the sum in the last expression, one can apply [13, Lemma 6] to obtain

$$\begin{aligned}\left\| \sum_{0 \leqslant n < N} D(T(an+b)) \right\|_2 &\ll q^{-c(\lambda+\beta)} \left(N + (\lambda+\beta) q^{\lambda+\beta}\right) \\ &\ll \lambda q^{(1-c)\lambda} \ll N^{1-\sigma}\end{aligned}$$

for an appropriately chosen constant $\sigma > 0$. Thus, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} D(T(an+b)) = 0.$$

Next, we consider the representations D_k, $0 \leqslant k < m$. We can use the fact that $D(T(n)) = \mathrm{e}(-nk/m)$ for all $n \geqslant 0$ and obtain

$$\frac{1}{N}\sum_{n=0}^{N-1} D_k(T(an+b)) = \frac{1}{N}\sum_{n=0}^{N-1} \mathrm{e}\left(-\frac{k(an+b)}{m}\right) = \mathrm{e}\left(-\frac{kb}{m}\right)\frac{1}{N}\sum_{n=0}^{N-1}\mathrm{e}\left(-\frac{kan}{m}\right).$$

Set $d = m/(a,m)$. If $k \equiv 0 \bmod d$ (this is equivalent to $m \mid ak$), then we have

$$\frac{1}{N}\sum_{n=0}^{N-1}\mathrm{e}\left(-\frac{kan}{m}\right) = 1.$$

If $k \not\equiv 0 \bmod d$, we obtain

$$\frac{1}{N}\left|\sum_{n=0}^{N-1} D_k(T(an+b))\right| = \frac{1}{N}\left|\sum_{n=0}^{N-1}\mathrm{e}\left(-\frac{kan}{m}\right)\right| = \frac{1}{N}\left|\frac{\sin(\pi N k a/m)}{\sin(\pi k a/m)}\right| \to 0,$$

as n goes to infinity. Thus we obtain for $0 \leqslant k < m$ that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} D_k(T(an+b)) = \begin{cases}\mathrm{e}\left(-\frac{kb}{m}\right), & \text{if } k \equiv 0 \bmod d,\\ 0, & \text{otherwise.}\end{cases}$$

Note, that $d = m$ if and only if $(a,m) = 1$. Thus, we just have shown that all irreducible and unitary representations satisfy the necessary and sufficient condition for $(T(an+b))_{n\geqslant 0}$ to be uniformly distributed in G if and only if $(a,m) = 1$ (see Lemma 1).

In order to complete the proof of Theorem 3, we have to deal with the case $(a,m) > 1$. We define the function f for all $g \in G$ by

$$f(g) = 1 + \mathrm{e}\left(\frac{db}{m}\right) D_d(g) + \mathrm{e}\left(\frac{2db}{m}\right) D_{2d}(g) + \cdots + \mathrm{e}\left(\frac{(m-d)b}{m}\right) D_{m-d}(g).$$

It is a real-valued positive function on G. To see this, choose an element $g \in \{T((m,a)n+b') : n \geqslant 0\}$, where $0 \leqslant b' < (m,a)$. We can write

$$\begin{aligned} f(g) &= \sum_{\ell=0}^{(m,a)-1} \mathrm{e}\left(\frac{\ell d b}{m}\right) D_{\ell d}(g) \\ &= \sum_{\ell=0}^{(m,a)-1} \mathrm{e}\left(\frac{\ell d b}{m} - \frac{\ell d((m,a)n+b')}{m}\right) \\ &= \sum_{k=0}^{(a,m)-1} \mathrm{e}\left(\frac{\ell(b-b')}{(a,m)}\right). \end{aligned}$$

Since $(T(n))_{n\geqslant 0}$ is uniformly distributed in G with respect to the Haar measure, we have that $(T(n))_{n\geqslant 0}$ is dense in G. Set $U_1 = \mathrm{cl}(\{T((m,a)n+b') : n \geqslant 0,\ 0 \leqslant b' < (a,m),\ b' \equiv b \bmod (m,a)\})$ and $U_2 = \mathrm{cl}(\{T((m,a)n+b') : n \geqslant 0,\ 0 \leqslant b' < (a,m),\ b' \not\equiv b \bmod (m,a)\})$. Then we have $G = U_1 \cup U_2$ and

$$f(g) = \begin{cases} (a,m), & \text{if } g \in U_1, \\ 0, & \text{otherwise.} \end{cases}$$

This proves the claim that f is positive. (Moreover, since f is continuous, we see that the group has to have more than one component in this case.) Using this function, we define the measure ν by[4]

$$\mathrm{d}\nu = f\mathrm{d}\mu.$$

In what follows, we show that the sequence $(T(an+b))_{n\geqslant 0}$ is ν-uniformly distributed in G. Let us consider a complete set of pairwise inequivalent irreducible unitary representations D^α, $\alpha \in \mathcal{A}$, where $\mathcal{A}$ is some index set. Put $e_{ij}^\alpha(g) = \sqrt{n_\alpha} d_{ij}^\alpha(g)$, where $D^\alpha = (d_{ij})_{1\leqslant i,j\leqslant n_\alpha}$. It follows from representation theory that the set $\{e_{ij}^\alpha\}$ forms a complete orthonormal system in the Hilbert space $L^2(G)$ with the scalar product $\langle f, g\rangle = \int_G f\overline{g}\mathrm{d}\mu$. We obtain for D_k, $k \equiv 0 \bmod d$ that

$$\int_G D_k f \mathrm{d}\mu = \sum_{\ell=0}^{(m,a)-1} \mathrm{e}\left(\frac{\ell d b}{m}\right) \langle D_k, \overline{D_{\ell d}}\rangle = \mathrm{e}\left(-\frac{kb}{m}\right).$$

For all other representations $D^\alpha = (d_{ij}^\alpha)_{1\leqslant i,j\leqslant n_\alpha}$, we get

$$\int_G d_{ij}^\alpha f \mathrm{d}\mu = \sum_{\ell=0}^{(m,a)-1} \mathrm{e}\left(\frac{\ell d b}{m}\right) \langle d_{ij}^\alpha, \overline{D_{\ell d}}\rangle = 0.$$

Lemma 1 implies that $(T(an+b))_{n\geqslant 0}$ is ν-uniformly distributed in G. If we set $U := \ker D_d$, then U is a normal subgroup of G (with index (m,a) in G). Similar to the proof of Lemma 3, one can show that

$$U = \mathrm{cl}\left(\{T((m,a)n) : n \geqslant 0\}\right)$$

and $U_1 = T(b)^{-1}U$, that is, U_1 is a coset of U. Since the support of ν is U_1, we have that $(T(an+b))_{n\geqslant 0}$ is dense in U_1. If we define the measure $\tilde{\nu}$ on U_1 by

$$\tilde{\nu}(B) = \int_B 1 \,\mathrm{d}\nu$$

[4] that is, $\nu(A) = \int_A f\mathrm{d}\mu$ for all Borel-sets A

for all Borel-sets B in U_1, we have that $(T(an+b))_{n\geqslant 0}$ is $\tilde{\nu}$-uniformly distributed in U_1. Moreover, $\tilde{\nu}$ is the translated normed Haar measure on U. Indeed, if we set

$$\tilde{\mu}(A) := \int_{T(b)^{-1}A} 1 \, \mathrm{d}\tilde{\nu}$$

for all Borel-sets A in U, then $\tilde{\mu}$ is translation invariant on U. This finishes the proof of Theorem 3. $\square$

3 Frequencies of Letters in Automatic Sequences

The special case of a finite group G is of particular interest, since there is an immediate relation to so-called automatic sequences. The main result of this section is an application of Theorem 1 to invertible automatic sequences (see Theorem 4).

Let G be a finite group of order $|G|$. Then G is also a compact topological group with respect to the discrete topology on G (every element is an open set). The Haar measure on G is the (normed) counting measure, that is

$$\mu(B) = \frac{1}{|G|} \#\{g : g \in B\}$$

for every $B \subseteq G$. Since the one-dimensional representations of a finite group have to be $|G|$-*th* roots of unity, we see that the characteristic integer m has to divide $(|G|, q-1)$. If we take for example $G = \mathbb{Z}/r\mathbb{Z}$ and $g_j = j \bmod r$, $0 \leqslant j < q$, then we have $m = (r, q-1)$ and $T(n) = s_q(n) \bmod r$, where s_q denotes the q-ary sum-of-digits function. As already mentioned, this is exactly the case considered by Mauduit and Rivat in [14] (see Remark 2).

It is convenient to work with permutation matrices instead of abstract group elements. If G is a finite group, then G is isomorphic to a subgroup of the symmetric group $S_{|G|}$ (Cayley's Theorem). Thus, we can assume that $g_0, \ldots, g_{q-1}$ are permutations in S_d for some integer $d \geqslant 1$ and G is a subgroup of S_d. The group G has a natural d dimensional representation χ, the so-called *permutation representation.* It is defined as follows: Let $\pi \in S_d$, then

$$\chi(\pi) = \left(e_{\pi(1)}, \ldots, e_{\pi(d)}\right),$$

where e_j denotes the j-*th* standard vector in $\mathbb{Z}^d$ (that is, all entries are 0 except the j-*th*, which is equal to 1). By definition it is clear that $\chi(\pi)$ is a permutation matrix, that is, $(x_1, \ldots, x_d)\, \chi(\pi) = (x_{\pi(1)}, \ldots, x_{\pi(d)})$. Obviously, permutation matrices are orthogonal (and unitary) matrices.

As already mentioned there is a close relation between sequences $(T(n))_{n\geqslant 0}$ of the form (1) and automatic sequences. Automatic sequences $(u_n)_{n\geqslant 0}$ are sequences that can be seen as the output sequence (or the image of the output sequence in an alphabet Δ) of a finite automaton when the input is the q-ary digital expansion of n (see for example [4, 11]). More precisely, a finite automaton has finitely many states. One starts in an initial state and then moves around the states depending on the input sequence (the q-ary digits of n). The moves are deterministic and can be encoded with the help of so-called *transition matrices* M_k, $k = 0, \ldots, q-1$, which have the property that the entry m_{ij}^k (the i-th row and j-th column of the matrix M_k) is 1 if there is a move from state j to state i when the input digit equals k. All other entries are zero. (Note that the dimension of the matrices M_k is equal to the number of states.) Such a finite automaton can be represented by a directed graph, where a directed edge is labeled with a number k (between 0 and $q-1$) which indicates the new state if the input digit equals k. For every n, the automaton terminates at some state $s(n)$. Automatic sequences are now computed with the help of an output function that is defined on the states: $s(n) \mapsto u_n$.

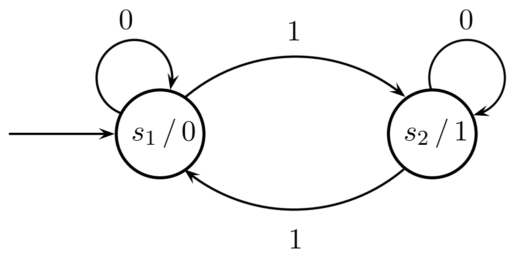

Figure 1: Automaton of the Thue-Morse sequence

In Figure 1 we see the automaton that creates the Thue-Morse sequence. The two states of the automaton are s_1 (initial state) and s_2, the output function maps s_1 to 0 and s_2 to 1 and the transition matrices are given by

$$M_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad M_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The automaton defined in Figure 2 has three states and the output function maps s_1 and s_2 to the letter a and s_3 to b. The transition matrices are given by

$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

The sequence starts with $aaaaabaabaabaaabba\ldots$, and, as we will see later, it is related to the symmetric group S_3.

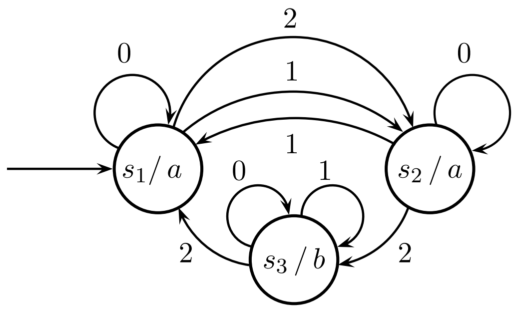

Figure 2: Automaton related to the S_3

There are several equivalent ways to describe automatic sequences (see [4], in particular Definition 5.1.1 and Theorem 5.2.1). For example, one can also describe q-automatic sequences in terms of morphisms (substitutions). Alternatively, by a theorem of Cobham a sequence $(u_n)_{n\geqslant 0}$ is q-automatic if and only if it is the image, under a coding, of a fixed point of a q-uniform morphism (compare with [4, Theorem 6.3.2]).

However, we will use the approach that is related to our sequence $(T(n))_{n\geqslant 0}$. Let $\mathcal{A}_d$ be the set of all $d \times d$ matrices with the property that in each column there is exactly one entry equal to 1 and all other entries are 0. If the automaton has d states, then $M_k \in \mathcal{A}_d$ for every $0 \leqslant k < q$. Set

$$S(n) = M_{\varepsilon_0} M_{\varepsilon_1} \cdots M_{\varepsilon_{\ell-1}},$$

where $(\varepsilon_{\ell-1}\varepsilon_{\ell-2}\ldots\varepsilon_1\varepsilon_0)_q$ denotes the q-ary digital expansion of n. Then we have that the last state reached is s_j (if the input is n) if and only if

$$S(n)e_1 = e_j.$$

Thus, a sequence $(u_n)_{n\geqslant 0}$ is a q-automatic sequence if and only if there exists q matrices in $\mathcal{A}_d$ (for some $d \geqslant 2$), such that u_n is given by the image of an output function (acting on $e_1, \ldots, e_d$) of $S(n)e_1$.

Definition 2. Let $(u_n)_{n\geqslant 0}$ be a q-automatic sequence. Then we call $(u_n)_{n\geqslant 0}$ an *invertible q-automatic sequence* if there exists an automaton such that all transition matrices are invertible and such that the transition matrix of zero is given by the identity matrix.

The set $\mathcal{A}_d$ is a monoid with respect to the matrix multiplication and all invertible matrices form a group (which is isomorphic to S_d). Taking this group as our group H, the matrices $M_0, \ldots, M_{q-1}$ generate a subgroup G and we can use Theorem 1 to analyze the subsequence $(u_{n^2})_{n\geqslant 0}$ of such invertible automatic sequences. As already indicated above, these matrices can be also seen as the permutation representation of the symmetric group S_d. Note, that the Thue-Morse sequence is an invertible 2-automatic sequence. The sequence induced by the automaton given in Figure 2 is an invertible 3-automatic sequence. (The transition matrices can be seen as the permutation representation of the identity element, the 2-cycle (12) and the 3-cycle (123) in S_3). If the matrices are interpreted as elements of the S_3, then they generate the whole group (that is, $G = H$ is isomorphic to S_3).

Next we give the definition of the frequency of a letter in a sequence. Let Δ be an alphabet and $(u_n)_{n\geqslant 0}$ a sequence in Δ. Furthermore, let $a \in \Delta$. If the limit

$$\lim_{N\to\infty} \frac{1}{N} \#\{0 \leqslant n < N : u_n = a\}$$

exists, then it is called the frequency of a in $(u_n)_{n\geqslant 0}$.

For arbitrary automatic sequences $(u_n)_{n\geqslant 0}$ there need not exist the frequency of each letter. Take for example

$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \text{and} \qquad M_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

(cf. [4, Example 8.1.2]). Nevertheless, if $(u_n)_{n\geqslant 0}$ is primitive (that is, the corresponding graph is strongly connected), then it is known that the frequencies of all letters exist (see [4, Theorem 8.4.7]). Furthermore, it is known that a subsequence of an automatic sequence of the form $(u_{an+b})_{n\geqslant 0}$ is again an automatic sequence (see [4, Theorem 6.8.1]). If we consider the subsequence $(u_{n^2})_{n\geqslant 0}$, then it is of a completely different nature. If one takes for example the Thue-Morse sequence $(t_n)_{n\geqslant 1}$, then it follows by a Theorem of Allouche and Salon [3] that $(t_{n^2})_{n\geqslant 0}$ is not 2-automatic. Moshe [15] showed in a recent work the much stronger result that the subword complexity of $(t_{n^2})_{n\geqslant 0}$ is maximal, that is, every finite word appears as a subword. (Note, that the subword complexity of an automatic sequence is $O(n)$, see [4, Theorem 10.3.1]). The following Theorem is a direct consequence of Theorem 1.

Theorem 4. *Let $q \geqslant 2$ and $(u_n)_{n\geqslant 0}$ be an invertible q-automatic sequence. Then the frequency of each letter of the subsequence $(u_{n^2})_{n\geqslant 0}$ exists.*

Remark 8. If the output function and the corresponding group generated by the transition matrices is known, then the exact frequencies of all letters can be given. Furthermore, one can show that if the output function is trivial (that is, every state is mapped to a different letter in the alphabet Δ) and the graph is strongly connected, then the frequencies are all equal if the characteristic integer of the underlying group is $\leqslant 2$.

If we take for example the automatic sequence generated by the automaton given in Figure 2, then we obtain that the underlying group is the symmetric group S_3 and $m = 2$. Thus, taking into account the special output function, the subsequence generated by the squares has the property, that the letter a has frequency $2/3$ and the letter b has frequency $1/3$.

Proof. Without loss of generality, we can assume that every state is mapped to a different letter in the alphabet Δ. Let the possible outcomes (the different states) be denoted by $s_1, \dots, s_d$. Since we consider invertible automatic sequences, the matrix sequence $(S(n))_{n\geqslant 0}$ defined above coincides with the sequence $(T(n))_{n\geqslant 0}$ defined in (1) (if we set $g_k = M_k$ for $0 \leqslant k < q$). Let the different elements in the subgroup generated by $M_0, \dots, M_{q-1}$ be denoted by $N_0, \dots, N_{s-1}$. Then

$$\sum_{n=0}^{N-1} T(n^2) = \sum_{r=0}^{s-1} \#\{0 \leqslant n < N : T(n^2) = N_r\} N_r.$$

Moreover, we have

$$e_i^T T(v) e_1 = \begin{cases} 1 & \text{if } u_v = s_i, \\ 0 & \text{otherwise.} \end{cases}$$

for every $v \geqslant 0$. Thus, we obtain

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \#\{0 \leqslant n < N : u_{n^2} = s_i\} &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} e_i^T T(n^2) e_1 \\ &= \sum_{r=0}^{s-1} \lim_{N\to\infty} \frac{1}{N} \#\{0 \leqslant n < N : T(n^2) = N_r\} e_i^T N_r e_1, \end{aligned}$$

for every $1 \leqslant i \leqslant d$. Theorem 1 implies that all limits exist, which in turn proves Theorem 4. □

4 Auxiliary Results

4.1 Van der Corput type inequalities

We begin with two van der Corput type inequality for matrices which enable us to "truncate" the sequence $T(n)$ twice (see Section 4.2 and Section 5). The presented lemmas are inspired by [14, Lemme 15 and 17]. We want to remark that van der Corput inequalities for matrices have been already considered by Hlawka [9] (see also [10, Chapter 4.2]).

Lemma 5. *Let N and B be positive integers satisfying $N \leqslant B$. Furthermore let $Z(n) \in \mathbb{C}^{d\times d}$ for all $n \in \mathbb{Z}$ satisfying $\|Z(n)\|_{\mathbb{F}} \leqslant f$. Then we have for any real number $R \geqslant 1$,*

$$\left\| \sum_{0\leqslant n<N} Z(n) \right\|_{\mathbb{F}} \leqslant \left(\frac{d^{1/2}N}{R} \sum_{|r|<R} \left(1-\frac{|r|}{R}\right) \left\| \sum_{0\leqslant n,n+r\leqslant B} Z(n+r)Z(n)^H \right\|_{\mathbb{F}} \right)^{1/2} + \frac{f}{2}R.$$

Proof. We take for convenience $Z(n) = 0$ (the $d\times d$ matrix consisting only of zeros) if $n \notin [0,B]$. We can write

$$\begin{aligned}
&\left\| R \sum_{0\leqslant n<N} Z\left(n\right) - \sum_{-R/2<r\leqslant R/2} \sum_{0\leqslant n<N} Z\left(n+r\right) \right\|_{\mathbb{F}} \\
&\qquad \leqslant \sum_{-R/2<r\leqslant R/2} \left\| \sum_{0\leqslant n<N} Z\left(n\right) - \sum_{0\leqslant n<N} Z\left(n+r\right) \right\|_{\mathbb{F}} \\
&\qquad \leqslant \sum_{-R/2<r\leqslant R/2} 2f|r| \leqslant \frac{f}{2}R^2.
\end{aligned}$$

Thus, we have

$$\left\| \sum_{0\leqslant n<N} Z\left(n\right) \right\|_{\mathbb{F}} \leqslant \frac{1}{R} \sum_{0\leqslant n<N} \left\| \sum_{-R/2<r\leqslant R/2} Z\left(n+r\right) \right\|_{\mathbb{F}} + \frac{f}{2}R.$$

Using the Cauchy-Schwarz inequality, we get

$$\begin{aligned}
&\left(\sum_{0\leqslant n<N} \left\| \sum_{-R/2<r\leqslant R/2} Z\left(n+r\right) \right\|_{\mathbb{F}} \right)^2 \\
&\qquad \leqslant N \sum_{0\leqslant n<N} \left\| \sum_{-R/2<r\leqslant R/2} Z\left(n+r\right) \right\|_{\mathbb{F}}^2 \\
&\qquad = N \sum_{0\leqslant n\leqslant B} \operatorname{tr} \left(\left(\sum_{-R/2<r_1\leqslant R/2} Z\left(n+r_1\right) \right) \left(\sum_{-R/2<r_2\leqslant R/2} Z\left(n+r_2\right) \right)^H \right) \\
&\qquad = N \sum_{-R/2<r_1\leqslant R/2} \sum_{-R/2<r_2\leqslant R/2} \operatorname{tr} \left(\sum_{0\leqslant n\leqslant B} Z\left(n+r_1\right) Z\left(n+r_2\right)^H \right).
\end{aligned}$$

Changing the index of summation, we obtain that the last expression is the same as

$$N \sum_{-R<r<R} \left(R-|r|\right) \operatorname{tr} \left(\sum_{0\leqslant n\leqslant B} Z\left(n+r\right) Z\left(n\right)^H \right).$$

The statement of the lemma follows from the fact that we have for all matrices $A = (a_{ij})_{1\leqslant i,j\leqslant d}$ (using the Cauchy-Schwarz inequality)

$$|\operatorname{tr}(A)| \leqslant \sum_{1\leqslant i\leqslant d} |a_{ii}| \leqslant \sqrt{d} \sqrt{\sum_{1\leqslant i\leqslant d} |a_{ii}|^2} \leqslant \sqrt{d} \sqrt{\sum_{1\leqslant i,j\leqslant d} |a_{ij}|^2} = \sqrt{d}\, \|A\|_{\mathbb{F}}. \tag{10}$$

□

Lemma 6. *Let N be a positive integer and $Z(n) \in \mathbb{C}^{d\times d}$ for all $0 \leqslant n \leqslant N$. Then we have for any real number $S \geqslant 1$ and any integer $k \geqslant 1$ the estimate*

$$\left\| \sum_{0\leqslant n\leqslant N} Z(n) \right\|_{\mathbb{F}}^2 \leqslant \frac{N+k(S-1)+1}{S} \sum_{|s|<S} \left(1 - \frac{|s|}{S}\right) \sum_{0\leqslant n, n+ks\leqslant N} \operatorname{tr}\left(Z(n+ks)Z(n)^H\right).$$

Proof. Again, we take for convenience $Z(n) = 0$ (the $d \times d$ matrix consisting only of zeros) if $n \notin [0, N]$. Then we can write

$$S \sum_{n\in\mathbb{Z}} Z(n) = \sum_{n\in\mathbb{Z}} \sum_{0\leqslant s<S} Z(n+ks).$$

If the last sum is not zero, then n satisfies $-k(S-1) \leqslant n \leqslant N$ and there are at most $N + k(S-1) + 1$ such values for n. Hence, applying the Cauchy-Schwarz inequality and changing the summation index yields (cf. Lemma 5)

$$\begin{aligned} S^2 \left\| \sum_{n\in\mathbb{Z}} Z(n) \right\|_{\mathbb{F}}^2 &\leqslant (N+k(S-1)+1) \sum_{n\in\mathbb{Z}} \left\| \sum_{0\leqslant s<S} Z(n+ks) \right\|_{\mathbb{F}}^2 \\ &\leqslant (N+k(S-1)+1) \sum_{0\leqslant s_1<S} \sum_{0\leqslant s_2<S} \sum_{n\in\mathbb{Z}} \operatorname{tr}\left(Z(n+ks_1)Z(n+ks_2)^H\right) \\ &\leqslant (N+k(S-1)+1) \sum_{|s|<S} (S-|s|) \sum_{n\in\mathbb{Z}} \operatorname{tr}\left(Z(n+ks)Z(n)^H\right). \end{aligned}$$

This proves the desired result. □

4.2 Fourier transform

Before we consider the Fourier transform, we recall the general setting. The compact group G is the closure of the subgroup generated by the elements $g_0, g_1, \ldots, g_{q-1}$ with $g_0 = e$ and D is an irreducible and unitary representation of dimension $d \geqslant 1$.

We start with defining truncated versions of the sequence $T(n)$. Let $\lambda \geqslant 1$ and $\mu < \lambda$ be positive integers. Set

$$T_\lambda(n) = g_{\varepsilon_0(n)} g_{\varepsilon_1(n)} \cdots g_{\varepsilon_{\lambda-1}(n)},$$

and

$$T_{\mu,\lambda}(n) = g_{\varepsilon_\mu(n)} g_{\varepsilon_{\mu+1}(n)} \cdots g_{\varepsilon_{\lambda-1}(n)},$$

where $\varepsilon_0(n), \ldots, \varepsilon_{\lambda-1}(n)$ are the λ lower placed digits in the q-ary digital expansion (with possibly leading zeros) of n. Recall, that the function F_λ (which depends on D) equals

$$F_\lambda(h) = \frac{1}{q^\lambda} \sum_{0\leqslant u<q^\lambda} \mathrm{e}\left(-\frac{hu}{q^\lambda}\right) D(T_\lambda(u)).$$

Additionally, we set

$$F_{\mu,\lambda}(h) := \frac{1}{q^\lambda} \sum_{0\leqslant u<q^\lambda} \mathrm{e}\left(-\frac{hu}{q^\lambda}\right) D(T_{\mu,\lambda}(u)).$$

We start our treatment on the Fourier terms with the definition of two related functions and some useful properties of them. Then we show a result on the second order average of $\|F_\lambda(h)\|_{\mathbb{F}}$ before we finally discuss the function $F_{\mu,\lambda}(h)$ in detail. The results we obtain here are very similar to the results in [14, Section 4.2]. Nevertheless, the proofs use many techniques from matrix theory.

Lemma 7. *Let the function φ_q be defined for all $q \geqslant 1$ by*

$$\varphi_q(t) = \left| \sum_{0 \leqslant u < q} \mathrm{e}(ut) \right| = \begin{cases} \frac{|\sin \pi q t|}{|\sin \pi t|}, & \text{if } t \in \mathbb{R} \setminus \mathbb{Z} \\ q, & \text{if } t \in \mathbb{Z}. \end{cases}$$

Then φ_q is periodic of period 1 and the function $\Phi(q)$ defined by

$$\Phi(q) := \max_{t \in \mathbb{R}} q^{-1} \sum_{0 \leqslant r < q} \varphi_q\left(t + \frac{r}{q}\right),$$

satisfies

$$\Phi(q) \leqslant \frac{2}{\pi} \log \frac{2 e^{\pi/\sqrt{2}} q}{\pi}.$$

Proof. See [14, Lemmas 1 and 2]. □

Lemma 8. *Let $q \geqslant 2$, $a \in \mathbb{Z}$ and $0 \leqslant \delta \leqslant \lambda$. Then we have*

$$\sum_{\substack{0 \leqslant h < q^\lambda \\ h \equiv a \bmod q^\delta}} \|F_\lambda(h)\|_{\mathbb{F}}^2 = \|F_\delta(a)\|_{\mathbb{F}}^2. \tag{11}$$

Proof. If $\delta = \lambda$, the assertion is trivial. Hence, we assume that $\delta < \lambda$. Let us denote the left hand side of (11) by S. Then we can write

$$\begin{aligned} S &= \sum_{\substack{0 \leqslant h < q^{\lambda-1} \\ h \equiv a \bmod q^\delta}} \sum_{0 \leqslant r < q} \left\| F_\lambda\left(h + r q^{\lambda-1}\right) \right\|_{\mathbb{F}}^2 \\ &= \sum_{\substack{0 \leqslant h < q^{\lambda-1} \\ h \equiv a \bmod q^\delta}} \sum_{0 \leqslant r < q} \operatorname{tr}\left(F_\lambda\left(h + r q^{\lambda-1}\right) F_\lambda\left(h + r q^{\lambda-1}\right)^H \right) \\ &= \sum_{\substack{0 \leqslant h < q^{\lambda-1} \\ h \equiv a \bmod q^\delta}} \sum_{0 \leqslant r < q} \operatorname{tr}\left(\frac{1}{q^2} \Psi_D\left(-\frac{h + r q^{\lambda-1}}{q^\lambda}\right) F_{\lambda-1}(h) F_{\lambda-1}(h)^H \Psi_D\left(-\frac{h + r q^{\lambda-1}}{q^\lambda}\right)^H \right). \end{aligned}$$

Here we used the definition of the Frobenius norm and the recursive description of F_λ (see (7)). The matrix trace is a linear operator with the property, that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ for two matrices A and B. Thus, we obtain

$$S = \sum_{\substack{0 \leqslant h < q^{\lambda-1} \\ h \equiv a \bmod q^\delta}} \operatorname{tr}\left(\frac{1}{q^2} \sum_{0 \leqslant r < q} \Psi_D\left(-\frac{h + r q^{\lambda-1}}{q^\lambda}\right)^H \Psi_D\left(-\frac{h + r q^{\lambda-1}}{q^\lambda}\right) F_{\lambda-1}(h) F_{\lambda-1}(h)^H \right). \tag{12}$$

Next we claim that for every $t \in \mathbb{R}$ we have

$$\frac{1}{q^2} \sum_{0 \leqslant r < q} \Psi_D\left(-t - \frac{r}{q}\right)^H \Psi_D\left(-t - \frac{r}{q}\right) = I_d. \tag{13}$$

Indeed, this holds true since the left hand side of (13) is equal to

$$\begin{aligned}
&\frac{1}{q^2} \sum_{0 \leqslant r < q} \sum_{0 \leqslant u < q} \mathrm{e}\left(tu + \frac{ru}{q}\right) D(g_u)^H \sum_{0 \leqslant v < q} \mathrm{e}\left(-tv - \frac{rv}{q}\right) D(g_v) \\
&\qquad = \frac{1}{q} \sum_{0 \leqslant u < q} \sum_{0 \leqslant v < q} \mathrm{e}\left(t(u-v)\right) D(g_u)^H D(g_v) \frac{1}{q} \sum_{0 \leqslant r < q} \mathrm{e}\left(r \frac{u-v}{q}\right) \\
&\qquad = \frac{1}{q} \sum_{0 \leqslant u < q} D(g_u)^H D(g_u),
\end{aligned}$$

and $D(g_u)^H D(g_u) = I_d$ for every $0 \leqslant u < q$. Using this result in (12), we obtain

$$\begin{aligned}
S &= \sum_{\substack{0 \leqslant h < q^{\lambda-1} \\ h \equiv a \bmod q^\delta}} \operatorname{tr}\left(F_{\lambda-1}(h) F_{\lambda-1}(h)^H\right) \\
&= \sum_{\substack{0 \leqslant h < q^{\lambda-1} \\ h \equiv a \bmod q^\delta}} \left\| F_{\lambda-1}(h) \right\|_{\mathbb{F}}^2 .
\end{aligned}$$

Applying this relation $\lambda - \delta$ times finally yields the desired result. □

Lemma 9. *Let $q \geqslant 2$ and $1 \leqslant \mu < \lambda$ be integers. Then we have*

$$\left\| F_{\mu,\lambda}(h) \right\| = \left\| F_{\lambda-\mu}(h) \right\| q^{-\mu} \varphi_{q^\mu}\left(h q^{-\lambda}\right), \tag{14}$$

where $\|\cdot\|$ is an arbitrary norm on $\mathbb{C}^{d \times d}$.

Proof. Since $T_{\mu,\lambda}(uq^\mu + v) = T_{\lambda-\mu}(u)$ for $0 \leqslant u < q^{\lambda-\mu}$ and $0 \leqslant v < q^\mu$, we get

$$\begin{aligned}
F_{\mu,\lambda}(h) &= q^{-\lambda} \sum_{0 \leqslant u < q^{\lambda-\mu}} \sum_{0 \leqslant v < q^\mu} \mathrm{e}\left(-\frac{h(uq^\mu + v)}{q^\lambda}\right) D(T_{\mu,\lambda}(uq^\mu + v)) \\
&= q^{-(\lambda-\mu)} \sum_{0 \leqslant u < q^{\lambda-\mu}} \mathrm{e}\left(-\frac{hu}{q^{\lambda-\mu}}\right) D(T_{\lambda-\mu}(u)) q^{-\mu} \sum_{0 \leqslant v < q^\mu} \mathrm{e}\left(-\frac{hv}{q^\lambda}\right).
\end{aligned}$$

Hence, we obtain (see Lemma 7)

$$\left\| F_{\mu,\lambda}(h) \right\| = \left\| F_{\lambda-\mu}(h) \right\| q^{-\mu} \varphi_{q^\mu}\left(h q^{-\lambda}\right).$$

□

Lemma 10. *Let $q \geqslant 2$ and $1 \leqslant \mu < \lambda$ be integers. Then we have*

$$\sum_{0 \leqslant h < q^\lambda} \left\| F_{\mu,\lambda}(h) \right\|_2 \leqslant \Phi(q^\mu) q^{\lambda-\mu}. \tag{15}$$

Proof. In a first step, we can write

$$\sum_{0\leqslant h<q^\lambda} \|F_{\mu,\lambda}(h)\|_2 = \sum_{0\leqslant u<q^{\lambda-\mu}} \sum_{0\leqslant v<q^\mu} \left\|F_{\lambda-\mu}(u+vq^{\lambda-\mu})\right\|_2 q^{-\mu}\varphi_{q^\mu}\left(\frac{u+vq^{\lambda-\mu}}{q^\lambda}\right)$$

$$= \sum_{0\leqslant u<q^{\lambda-\mu}} \|F_{\lambda-\mu}(u)\|_2\, q^{-\mu} \sum_{0\leqslant v<q^\mu} \varphi_{q^\mu}\left(\frac{u}{q^\lambda}+\frac{v}{q^\mu}\right).$$

Next, note that $\|F_\lambda(h)\|_2 \leqslant 1$ since D is a unitary representation. This observation (together with Lemma 7) finally implies the result. ☐

Lemma 11. *Let $q\geqslant 2$, $a\in\mathbb{Z}$ and $1\leqslant \lambda-\mu\leqslant\delta\leqslant\lambda$ be positive integers. Then we have*

$$\sum_{\substack{0\leqslant h<q^\lambda\\ h\equiv a \bmod q^\delta}} \|F_{\mu,\lambda}(h)\| \leqslant \Phi\left(q^{\lambda-\delta}\right) q^{-\mu+\lambda-\delta}\varphi_{q^{\mu-\lambda+\delta}}\left(aq^{-\delta}\right)\|F_{\lambda-\mu}(a)\|,$$

where $\|\cdot\|$ is an arbitrary norm on $\mathbb{C}^{d\times d}$.

Proof. Since by assumption $\lambda-\mu\leqslant\delta$, we have $F_{\lambda-\mu}(a+\ell q^\delta)=F_{\lambda-\mu}(a)$. We obtain

$$\sum_{\substack{0\leqslant h<q^\lambda\\ h\equiv a \bmod q^\delta}} \|F_{\mu,\lambda}(h)\| = \sum_{0\leqslant \ell<q^{\lambda-\delta}} \left\|F_{\lambda-\mu}\left(a+\ell q^\delta\right)\right\| q^{-\mu}\varphi_{q^\mu}\left(\frac{a+\ell q^\delta}{q^\lambda}\right)$$

$$= \|F_{\lambda-\mu}(a)\|\, q^{-\mu} \sum_{0\leqslant \ell<q^{\lambda-\delta}} \varphi_{q^\mu}\left(\frac{a}{q^\lambda}+\frac{\ell}{q^{\lambda-\delta}}\right).$$

A short calculation shows (cf. [14, Proof of Lemma 12]), that

$$q^{-\mu}\sum_{0\leqslant \ell<q^{\lambda-\delta}} \varphi_{q^\mu}\left(\frac{a}{q^\lambda}+\frac{\ell}{q^{\lambda-\delta}}\right) \leqslant \Phi\left(q^{\lambda-\delta}\right) q^{-\mu+\lambda-\delta}\varphi_{q^{\mu-\lambda+\delta}}\left(aq^{-\delta}\right).$$

This proves Lemma 11. ☐

Lemma 12. *Let $q\geqslant 2$, $a\in\mathbb{Z}$, $\ell\in\mathbb{Z}$, $1\leqslant\mu<\lambda$ and $0\leqslant\delta\leqslant\lambda-\mu$. Then we have*

$$\sum_{\substack{0\leqslant h<q^\lambda\\ h\equiv a \bmod q^\delta}} \|F_{\mu,\lambda}(h)\|_{\mathbb{F}} \|F_{\lambda-\mu}(h+\ell)\|_{\mathbb{F}} \leqslant \Phi\left(q^\mu\right)\|F_\delta(a)\|_{\mathbb{F}}\|F_\delta(a+\ell)\|_{\mathbb{F}}. \tag{16}$$

Proof. If we write $h=uq^{\lambda-\mu}+v$, where $0\leqslant u<q^\mu$ and $0\leqslant v<q^{\lambda-\mu}$, then $h\equiv a \bmod q^\delta$ is equivalent to $v\equiv a \bmod q^\delta$ (we have $\delta\leqslant\lambda-\mu$). Denote the left hand side of (16) by S. Then we obtain

$$S = \sum_{\substack{0\leqslant v<q^{\lambda-\mu}\\ v\equiv a \bmod q^\delta}} \sum_{0\leqslant u<q^\mu} \left\|F_{\mu,\lambda}(uq^{\lambda-\mu}+v)\right\|_{\mathbb{F}} \left\|F_{\lambda-\mu}(uq^{\lambda-\mu}+v+\ell)\right\|_{\mathbb{F}}$$

$$= \sum_{\substack{0\leqslant v<q^{\lambda-\mu}\\ v\equiv a \bmod q^\delta}} \|F_{\lambda-\mu}(v)\|_{\mathbb{F}}\|F_{\lambda-\mu}(v+\ell)\|_{\mathbb{F}}\, q^{-\mu}\sum_{0\leqslant u<q^\mu}\varphi_{q^\mu}\left(\frac{uq^{\lambda-\mu}+v}{q^\lambda}\right)$$

$$\leqslant \sum_{\substack{0\leqslant v<q^{\lambda-\mu}\\ v\equiv a \bmod q^\delta}} \|F_{\lambda-\mu}(v)\|_{\mathbb{F}}\|F_{\lambda-\mu}(v+\ell)\|_{\mathbb{F}}\,\Phi\left(q^\mu\right).$$

Applying the Cauchy-Schwarz inequality and Lemma 8 yields

$$\begin{aligned} S &\leqslant \Phi\left(q^{\mu}\right) \left(\sum_{\substack{0 \leqslant v < q^{\lambda-\mu} \\ v \equiv a \bmod q^{\delta}}} \left\| F_{\lambda-\mu}(v) \right\|_{\mathbb{F}}^2 \right)^{1/2} \left(\sum_{\substack{0 \leqslant v < q^{\lambda-\mu} \\ v \equiv a \bmod q^{\delta}}} \left\| F_{\lambda-\mu}(v+\ell) \right\|_{\mathbb{F}}^2 \right)^{1/2} \\ &= \Phi\left(q^{\mu}\right) \left\| F_{\delta}(a) \right\|_{\mathbb{F}} \left\| F_{\delta}(a+\ell) \right\|_{\mathbb{F}}. \end{aligned}$$

□

Lemma 13. *Let $q \geqslant 2$, $a \in \mathbb{Z}$, $1 \leqslant \mu < \lambda$ and $\lambda - \mu \leqslant \delta \leqslant \lambda$. Then we have*

$$\sum_{\substack{0 \leqslant h_1, h_2 < q^{\lambda} \\ h_1 + h_2 \equiv a \bmod q^{\delta}}} \left\| F_{\mu,\lambda}(h_1) \right\|_{\mathbb{F}} \left\| F_{\mu,\lambda}(-h_2) \right\|_{\mathbb{F}} \leqslant d\, \Phi(q^{\lambda-\delta})\, \Phi(q^{\mu}). \tag{17}$$

Proof. We have

$$\begin{aligned} &\sum_{\substack{0 \leqslant h_1, h_2 < q^{\lambda} \\ h_1 + h_2 \equiv a \bmod q^{\delta}}} \left\| F_{\mu,\lambda}(h_1) \right\|_{\mathbb{F}} \left\| F_{\mu,\lambda}(-h_2) \right\|_{\mathbb{F}} \\ &\qquad = \sum_{0 \leqslant h_2 < q^{\lambda}} \left\| F_{\mu,\lambda}(-h_2) \right\|_{\mathbb{F}} \sum_{\substack{0 \leqslant h_1 < q^{\lambda} \\ h_1 \equiv -h_2 + a \bmod q^{\delta}}} \left\| F_{\mu,\lambda}(h_1) \right\|_{\mathbb{F}} \\ &\qquad \leqslant \Phi(q^{\lambda-\delta}) \sum_{0 \leqslant h_2 < q^{\lambda}} \left\| F_{\mu,\lambda}(-h_2) \right\|_{\mathbb{F}} \left\| F_{\lambda-\mu}(-h_2 + a) \right\|_{\mathbb{F}}. \end{aligned}$$

To obtain the last inequality, we employed Lemma 11 (note, that $q^{-\mu+\lambda-\delta}\varphi_{q^{\mu-\lambda+\delta}}(\,\cdot\,) \leqslant 1$). Since we have

$$\left\| F_0(u) \right\|_{\mathbb{F}} = \left\| I_d \right\|_{\mathbb{F}} = \sqrt{d}$$

for all $u \in \mathbb{Z}$, Lemma 12 (with $\delta = 0$) yields the desired result. □

5 Proof of Theorem 1

In the next two sections we show that

$$\frac{1}{N} \sum_{0 \leqslant n < N} D(T(n^2))$$

converges for every irreducible and unitary representation D (for $N \to \infty$). With the help of this result, we will finish the proof of Theorem 1 in Section 5.3.

5.1 Irreducible representations of the form D_k

In this section we consider the m representations $D_0, \ldots, D_{m-1}$. Recall that $D_k(T(v)) = \mathrm{e}(-kv/m)$ for all $v \geqslant 0$. We have

$$\begin{aligned} \frac{1}{N} \sum_{n=0}^{N-1} D_k(T(n^2)) &= \frac{1}{N} \sum_{n=0}^{N-1} \mathrm{e}\left(-\frac{kn^2}{m}\right) \\ &= \frac{1}{N} \left\lfloor \frac{N}{m} \right\rfloor \sum_{n=0}^{m-1} \mathrm{e}\left(-\frac{kn^2}{m}\right) + \frac{1}{N} \sum_{n=m\left\lfloor \frac{N}{m} \right\rfloor}^{N-1} \mathrm{e}\left(-\frac{kn^2}{m}\right). \end{aligned}$$

It follows that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} D_k(T(n^2)) = \frac{1}{m} G(-k,m), \tag{18}$$

where $G(a,c) = G(a,0;c)$ and $G(a,b;c)$ denotes the quadratic Gauss sum

$$G(a,b;c) = \sum_{n=0}^{c-1} \mathrm{e}\left(\frac{an^2+bn}{c}\right). \tag{19}$$

This sum is well studied in the literature (see for example [5]).

5.2 Irreducible representations different from D_k

Let D be an irreducible and unitary representation of G of dimension $d \geqslant 1$ such that $D \neq D_k$ for all $0 \leqslant k < m$. Furthermore, let $\nu \in \mathbb{Z}^+$ be defined by $q^{\nu-1} < N \leqslant q^\nu$ and set

$$S_1 = \sum_{0\leqslant n<N} D(T(n^2)). \tag{20}$$

Then we will show that there exists a constant $\sigma > 0$, such that $\|S_1\|_{\mathbb{F}} \ll q^{(1-\sigma)\nu}$. In particular, this implies that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} D(T(n^2)) = 0.$$

We begin with applying Lemma 5 with $B = q^\nu$, $R = q^\rho$ and $Z(n) = D(T(n^2))$, where ρ is an integer satisfying $1 \leqslant \rho \leqslant \nu/2$ (note, that $\|D(T(n^2))\|_{\mathbb{F}} = \sqrt{d}$ for all $n \geqslant 0$). We can write

$$\begin{aligned}
\|S_1\|_{\mathbb{F}} &\ll \left(\frac{N}{q^\rho} \sum_{|r|<q^\rho} \left(1 - \frac{|r|}{q^\rho}\right) \left\| \sum_{0\leqslant n,n+r\leqslant q^\nu} D(T((n+r)^2)) D(T(n^2))^H \right\|_{\mathbb{F}} \right)^{1/2} + q^\rho \\
&\ll \left(q^{2\nu-\rho} + q^\nu \max_{1\leqslant|r|<q^\rho} \left\| \sum_{0\leqslant n,n+r\leqslant q^\nu} D(T((n+r)^2)) D(T(n^2))^H \right\|_{\mathbb{F}} \right)^{1/2} + q^\rho.
\end{aligned}$$

In the last step, we separated the case $r = 0$ and $r \neq 0$. Additionally, we get an error term $O(q^{\nu+\rho})$ (inside the square root) when removing the summation condition $0 \leqslant n+r \leqslant q^\nu$. Since we have assumed $\rho \leqslant \nu/2$, this term can be neglected. Hence, we obtain

$$\|S_1\|_{\mathbb{F}} \ll q^{\nu-\rho/2} + q^{\nu/2} \max_{1\leqslant|r|<q^\rho} \left\| \sum_{0\leqslant n\leqslant q^\nu} D(T((n+r)^2)) D(T(n^2))^H \right\|_{\mathbb{F}}^{1/2}.$$

We set

$$\lambda := \nu + 2\rho + 1.$$

Recall that we have (using the fact that $g_0 = e$)

$$T_\lambda(n) := g_{\varepsilon_0(n)} g_{\varepsilon_1(n)} \cdots g_{\varepsilon_{\lambda-1}(n)},$$

where $\varepsilon_0(n), \ldots, \varepsilon_{\lambda-1}(n)$ are the λ lower placed digits in the q-ary digital expansion (with possibly leading zeros) of n.

Lemma 14. *For all integers ν and ρ with $\nu \geqslant 2$ and $1 \leqslant \rho \leqslant \nu/2$ and for all $r \in \mathbb{Z}$ with $|r| < q^\rho$, we denote by $E(r,\nu,\rho)$ the number of integers n such that $0 \leqslant n \leqslant q^\nu$ and*

$$D(T_\lambda((n+r)^2))D(T_\lambda(n^2))^H \neq D(T((n+r)^2))D(T(n^2))^H.$$

Then we have

$$E(r,\nu,\rho) \ll q^{\nu-\rho}.$$

Proof. In order to prove this lemma we recall what Mauduit and Rivat actually proved in [14, Lemme 16]. Let us fix an integer r satisfying $|r| < q^\rho$ and denote by $F(r,\nu,\rho)$ the number of integers $0 \leqslant n \leqslant q^\nu$ such that not all digits of n^2 and $(n+r)^2$ which are higher placed than $\lambda - 1$ are equal. Then it follows from their reasoning that $F(r,\nu,\rho) \ll q^{\nu-\rho}$. (Note, that the additional condition $q^{\nu-1} < n$ in their statement is not needed.) Using this fact, we can complete our proof. Let $n^2 = \varepsilon_{\ell-1}\varepsilon_{\ell-2}\ldots\varepsilon_\lambda\ldots\varepsilon_0$, where $\ell > \lambda$ and with possibly leading zeros. If all digits of n^2 and $(n+r)^2$ which are higher placed than $\lambda - 1$ are equal, we have

$$\begin{aligned}
&D(T((n+r)^2))D(T(n^2))^H \\
&\qquad = D(T_\lambda((n+r)^2))D(g_{\varepsilon_\lambda}\cdots g_{\varepsilon_{\ell-1}})D(g_{\varepsilon_\lambda}\cdots g_{\varepsilon_{\ell-1}})^H D(T_\lambda(n^2))^H \\
&\qquad = D(T_\lambda((n+r)^2))D(T_\lambda(n^2))^H.
\end{aligned}$$

Here we used that D is a unitary representation. All matrices that come from the higher placed digits "cancel" out. We can bound $E(r,\nu,\rho)$ by $F(r,\nu,\rho)$ which proves the desired result. □

This lemma enables us to replace $D(T(u^2))$ by $D(T_\lambda(u^2))$. We obtain

$$\|S_1\|_{\mathbb{F}} \ll q^{\nu-\rho/2} + q^{\nu/2} \max_{1\leqslant |r| < q^\rho} \|S_2\|_{\mathbb{F}}^{1/2}, \tag{21}$$

where

$$S_2 := \sum_{0\leqslant n\leqslant q^\nu} D(T_\lambda((n+r)^2))D(T_\lambda(n^2))^H.$$

In view of Lemma 6, we set $Z(n) = D(T_\lambda((n+r)^2))D(T_\lambda(n^2))^H$, $N = q^\nu$, $S = q^{2\rho}$ and $k = q^\mu$, where μ is an integer satisfying

$$1 \leqslant \mu \leqslant \nu - 2\rho - 1.$$

Then we have to consider expressions of the form

$$\operatorname{tr}\left(D(T_\lambda((n+r+sq^\mu)^2))D(T_\lambda((n+sq^\mu)^2))^H D(T_\lambda(n^2))D(T_\lambda((n+r)^2))^H\right). \tag{22}$$

Using that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ for two matrices A and B, this is the same as

$$\operatorname{tr}\left(D(T_\lambda((n+r)^2))^H D(T_\lambda((n+r+sq^\mu)^2))D(T_\lambda((n+sq^\mu)^2))^H D(T_\lambda(n^2))\right).$$

Next, we recall that

$$T_{\mu,\lambda}(n) = g_{\varepsilon_\mu(n)}g_{\varepsilon_{\mu+1}(n)}\cdots g_{\varepsilon_{\lambda-1}(n)}.$$

Since

$$(n+r+sq^\mu)^2 = (n+r)^2 + q^\mu(2s(n+r) + s^2q^\mu),$$

we see that $(n+r+sq^\mu)^2$ and $(n+r)^2$ have the same μ lower placed digits. Using that D is a unitary representation (compare also with the proof of Lemma 14), we see that

$$D(T_\lambda((n+r)^2))^H D(T_\lambda((n+r+sq^\mu)^2)) = D(T_{\mu,\lambda}((n+r)^2))^H D(T_{\mu,\lambda}((n+r+sq^\mu)^2)),$$

since the terms coming from the μ lower placed digits cancel out. The same argument works for $(n+sq^\mu)^2$ and n^2 and we obtain that (22) can be written as

$$\operatorname{tr}\left(D(T_{\mu,\lambda}((n+r)^2))^H D(T_{\mu,\lambda}((n+r+sq^\mu)^2)) D(T_{\mu,\lambda}((n+sq^\mu)^2))^H D(T_{\mu,\lambda}(n^2))\right).$$

We set $I_{\nu,s,\mu} := \{0 \leqslant n \leqslant q^\nu : 0 \leqslant n+sq^\mu \leqslant q^\nu\}$. Then we finally have (cf. (10))

$$\begin{aligned}
&\left|\sum_{n\in I_{\nu,s,\mu}} \operatorname{tr}\left(D(T_\lambda((n+r)^2))^H D(T_\lambda((n+r+sq^\mu)^2)) D(T_\lambda((n+sq^\mu)^2))^H D(T_\lambda(n^2))\right)\right| \\
&\quad = \left|\operatorname{tr}\left(\sum_{n\in I_{\nu,s,\mu}}\left(D(T_{\mu,\lambda}((n+r)^2))^H D(T_{\mu,\lambda}((n+r+sq^\mu)^2))\right.\right.\right. \\
&\qquad\qquad \left.\left.\left.\cdot D(T_{\mu,\lambda}((n+sq^\mu)^2))^H D(T_{\mu,\lambda}(n^2))\right)\right)\right| \\
&\quad \leqslant \sqrt{d}\left\| \sum_{n\in I_{\nu,s,\mu}} D(T_{\mu,\lambda}((n+r)^2))^H D(T_{\mu,\lambda}((n+r+sq^\mu)^2))\right. \\
&\qquad\qquad \left.\cdot D(T_{\mu,\lambda}((n+sq^\mu)^2))^H D(T_{\mu,\lambda}(n^2))\right\|_{\mathbb{F}}.
\end{aligned}$$

Hence Lemma 6 gives

$$\begin{aligned}
\|S_2\|_{\mathbb{F}}^2 &\ll q^{\nu-2\rho} \sum_{|s|<q^{2\rho}} \left(1 - \frac{|s|}{q^{2\rho}}\right) \|S_3\|_{\mathbb{F}} \\
&\ll q^{2\nu-2\rho} + q^\nu \max_{1\leqslant |s| < q^{2\rho}} \|S_3\|_{\mathbb{F}},
\end{aligned} \tag{23}$$

where S_3 denotes the sum

$$\sum_{n\in I_{\nu,s,\mu}} D(T_{\mu,\lambda}((n+r)^2))^H D(T_{\mu,\lambda}((n+r+sq^\mu)^2)) D(T_{\mu,\lambda}((n+sq^\mu)^2))^H D(T_{\mu,\lambda}(n^2)).$$

The inverse of the Fourier term $F_{\mu,\lambda}$ (defined in Section 4.2) is given by

$$D(T_{\mu,\lambda}(u)) = \sum_{0\leqslant h<q^\lambda} F_{\mu,\lambda}(h)\, \mathrm{e}\left(\frac{uh}{q^\lambda}\right).$$

Hence we obtain

$$\begin{aligned}
S_3 = &\sum_{0\leqslant h_1,h_2,h_3,h_4<q^\lambda} F_{\mu,\lambda}(-h_1)^H F_{\mu,\lambda}(h_2) F_{\mu,\lambda}(-h_3)^H F_{\mu,\lambda}(h_4) \\
&\cdot \sum_{n\in I_{\nu,s,\mu}} \mathrm{e}\left(\frac{h_1(n+r)^2 + h_2(n+r+sq^\mu)^2 + h_3(n+sq^\mu)^2 + h_4 n^2}{q^\lambda}\right).
\end{aligned} \tag{24}$$

Assume that $c \geqslant 2$ is an integer and $(z_n)_{n\in\mathbb{Z}}$ is a sequence of complex numbers that is periodic of period c. It is shown in [14, Lemme 18], that one has for all M_1, $M_2 \in \mathbb{Z}$ with $1 \leqslant M_2 \leqslant c$ the estimate

$$\left| \sum_{M_1 < n \leqslant M_1 + M_2} z_n \right| \leqslant \frac{2}{\pi} \log\left(\frac{4e^{\pi/2}c}{\pi} \right) \max_{0 \leqslant \ell < c} \left| \sum_{0 \leqslant n < c} z_n \, \mathrm{e}\left(\frac{\ell n}{c} \right) \right|.$$

If we apply this to (24) with $c = q^\lambda$ and M_1 and M_2 chosen appropriately, we get

$$\begin{aligned} \|S_3\|_{\mathbb{F}} \leqslant \frac{2}{\pi} \log\left(\frac{4e^{\pi/2}q^\lambda}{\pi} \right) & \max_{0 \leqslant \ell < q^\lambda} \sum_{d | q^\lambda} \\ & \cdot \sum_{\substack{0 \leqslant h_1, h_2, h_3, h_4 < q^\lambda \\ (h_1+h_2+h_3+h_4, q^\lambda) = d}} \|F_{\mu,\lambda}(h_1)\|_{\mathbb{F}} \, \|F_{\mu,\lambda}(h_2)\|_{\mathbb{F}} \, \|F_{\mu,\lambda}(h_3)\|_{\mathbb{F}} \, \|F_{\mu,\lambda}(h_4)\|_{\mathbb{F}} \\ & \qquad \cdot |G(h_1 + h_2 + h_3 + h_4, 2r(h_1 + h_2) + 2sq^\mu(h_2 + h_3) + l; q^\lambda)|, \end{aligned}$$

where $G(a, b; c)$ denotes the quadratic Gauss sum already defined in (19). This sum has the following properties (see for example [14, Proposition 1 and 2]). Let $d = (a, c)$. Then $G(a, b; c) = 0$ if $d \nmid b$ and $|G(a, b, c)| \leqslant \sqrt{2dc}$ in any case. Thus we get

$$\begin{aligned} \|S_3\|_{\mathbb{F}} \leqslant \frac{2}{\pi} \log\left(\frac{4e^{\pi/2}q^\lambda}{\pi} \right) q^{\lambda/2} & \max_{0 \leqslant \ell < q^\lambda} \sum_{d | q^\lambda} d^{1/2} \\ & \cdot \sum_{\substack{0 \leqslant h_1, h_2, h_3, h_4 < q^\lambda \\ (h_1+h_2+h_3+h_4, q^\lambda) = d \\ d | 2r(h_1+h_2) + 2sq^\mu(h_2+h_3) + \ell}} \|F_{\mu,\lambda}(h_1)\|_{\mathbb{F}} \, \|F_{\mu,\lambda}(h_2)\|_{\mathbb{F}} \, \|F_{\mu,\lambda}(h_3)\|_{\mathbb{F}} \, \|F_{\mu,\lambda}(h_4)\|_{\mathbb{F}}. \end{aligned} \tag{25}$$

Exactly the same way as in the work of Mauduit and Rivat [14, Section 5.5 - 5.8], one can now show that the integer μ can be chosen in such a way that

$$\|S_3\|_{\mathbb{F}} \ll \nu^{\omega(q)+6} q^{\nu - 2\rho},$$

whenever ρ is an integer smaller than ν times some constant only depending on the chosen representation D. (The constant $\omega(q)$ denotes the number of distinct prime divisors of q). In order to do so, one has to use the results on the Fourier terms we have proved in Section 4.2 on the one hand and the L^∞ norm estimate of $\|F_\lambda(\,\cdot\,)\|$ proved in Section 2 (see Lemma 4) on the other hand. Note, that in order to apply Lemma 4 we need that $D \neq D_k$ for $0 \leqslant k < m$.

Combining (21), (23) and the last estimate, we are done since we can choose ρ such that

$$\|S_1\|_{\mathbb{F}} \ll q^{(1-\sigma)\nu}$$

for some constant $\sigma > 0$.

5.3 Final steps in the proof of Theorem 1

Let m be the characteristic integer of $g_0, \ldots, g_{q-1}$ and q. Lemma 3 implies that the set $U := \mathrm{cl}(\{T(mn) : n \geqslant 0\})$ is a normal subgroup of G (of index m) with cosets $g_u U = \mathrm{cl}(\{T(mn + u) : n \geqslant 0\})$, $0 \leqslant u < m$.

Let us define the function f by

$$f(g) = 1 + \frac{1}{m} G(1,m) D_1(g) + \cdots + \frac{1}{m} G(m-1,m) D_{m-1}(g).$$

Since the representations are continuous, f is a continuous function. Moreover, it is a positive function and it satisfies

$$f(v) = \sum_{u=0}^{m-1} \mathbf{1}_{g_u U}(v) \cdot Q(u,m). \tag{26}$$

In order to see this, we state and prove the following lemma:

Lemma 15. *Let $c \geqslant 1$, $a \in \mathbb{Z}$ and set $Q(a,c) = \#\{0 \leqslant n < c : n^2 \equiv a \bmod c\}$. Then we have*

$$Q(a,c) = \frac{1}{c} \sum_{k=0}^{c-1} G(k,c) \, \mathrm{e}\left(-\frac{ka}{c}\right).$$

Proof. Let us consider the group $\mathbb{Z}/c\mathbb{Z}$ (with $+$ as the group operation). The representations of $\mathbb{Z}/c\mathbb{Z}$ are given by

$$\chi_k(u) = \mathrm{e}\left(\frac{ku}{c}\right)$$

for all $u \in \mathbb{Z}/c\mathbb{Z}$ and all $0 \leqslant k < c$. For any $n \in \mathbb{N}$ let x_n be the element of $\mathbb{Z}/c\mathbb{Z}$ defined by $n \equiv x_n \bmod c$. Then we obtain

$$\lim_{N\to\infty} \frac{1}{N} \#\{0 \leqslant n < N : x_{n^2} \equiv a \bmod c\} = \frac{1}{c} Q(a,c).$$

Thus, the sequence $(x_{n^2})_{n \geqslant 0}$ is $\nu_{\mathbb{Z}/c\mathbb{Z}}$-uniformly distributed in $\mathbb{Z}/c\mathbb{Z}$, where the measure $\nu_{\mathbb{Z}/c\mathbb{Z}}$ is defined by

$$\nu_{\mathbb{Z}/c\mathbb{Z}}(v) = \frac{1}{c} Q(v,c).$$

It follows from Lemma 1 that for every $0 \leqslant k < c$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \chi_k(x_{n^2}) = \int_{\mathbb{Z}/c\mathbb{Z}} \chi_k \mathrm{d}\nu_{\mathbb{Z}/c\mathbb{Z}}. \tag{27}$$

As in Section 5.1, we obtain

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \chi_k(x_{n^2}) = \frac{1}{c} G(k,c).$$

On the other hand, we have

$$\int_{\mathbb{Z}/c\mathbb{Z}} \chi_k \mathrm{d}\nu_{\mathbb{Z}/c\mathbb{Z}} = \frac{1}{c} \sum_{v=0}^{c-1} \mathrm{e}\left(-\frac{vk}{c}\right) Q(v,c).$$

Summing up the left and right hand side of (27) from $k=0$ to $k=c-1$ (weighted with $\mathrm{e}(-ka/c)$), we obtain

$$\frac{1}{c} \sum_{k=0}^{c-1} G(k,c) \, \mathrm{e}\left(-\frac{ka}{c}\right) = \frac{1}{c} \sum_{k=0}^{c-1} \mathrm{e}\left(-\frac{ka}{c}\right) \sum_{v=0}^{c-1} \mathrm{e}\left(\frac{vk}{c}\right) Q(v,c). \tag{28}$$

The right hand side of (28) is equal to

$$\sum_{v=0}^{c-1} Q(v,c)\frac{1}{c}\sum_{k=0}^{c-1} \mathrm{e}\left(\frac{k(v-a)}{c}\right) = Q(a,c).$$

This shows the desired result. $\square$

If n and v are integers $\geqslant 0$, then

$$f(T(nm+v)) = \frac{1}{m}\sum_{k=0}^{m-1} G(k,m)D_k(T(nm+v)) = \frac{1}{m}\sum_{k=0}^{m-1} G(k,m)\,\mathrm{e}\left(-\frac{vk}{m}\right).$$

Employing Lemma 15, we obtain that

$$f(T(nm+v)) = Q(v,m).$$

Since $(T(n))_{n\geqslant 0}$ is dense in G, equation (26) holds true, indeed. This allows us to define the measure

$$\mathrm{d}\nu = f\mathrm{d}\mu.$$

We proceed as in the linear case and show that $(T(n^2))_{n\geqslant 0}$ is ν-uniformly distributed in G. Let $\{D^\alpha,\, \alpha \in \mathcal{A}\}$ be again a complete set of pairwise inequivalent irreducible unitary representations and set $e_{ij}^\alpha(g) = \sqrt{n_\alpha}d_{ij}^\alpha(g)$, where $D^\alpha = (d_{ij})_{1\leqslant i,j\leqslant n_\alpha}$ (recall, that the set $\{e_{ij}^\alpha\}$ forms a complete orthonormal system in the Hilbert space $L^2(G)$). We obtain for D_k, $k = 0,\ldots,m-1$ that

$$\int_G D_k f\mathrm{d}\mu = \sum_{\ell=0}^{m-1}\frac{1}{m}G(-\ell,m)\left\langle D_k,\overline{D_\ell}\right\rangle = \frac{1}{m}G(-k,m).$$

For all other representations $D^\alpha = (d_{ij}^\alpha)_{1\leqslant i,j\leqslant n_\alpha}$, we get

$$\int_G d_{ij}^\alpha f\mathrm{d}\mu = \sum_{\ell=0}^{m-1}\frac{1}{m}G(-\ell,m)\left\langle d_{ij}^\alpha,\overline{D_\ell}\right\rangle = 0.$$

Finally, this proves Theorem 1.

6 Possible Generalizations and Extensions

Similarly to completely q-multiplicative functions (which are precisely sequences of the form (1)) we can consider so-called block-multiplicative functions (cf. [6]). Let $L \geqslant 1$ be given and $B = (b_1\cdots b_L)$ be a block of length L of q-ary digits $b_j \in \{0,1,\ldots,q-1\}$. Furthermore, let $(g_B)_{B\in\{0,1,\ldots,q-1\}^L}$ be elements of a compact group H with the property $g_{(00\cdots 0)} = e$. Similarly to the definition (1) we set

$$\begin{aligned} T(n) =& g_{(0\cdots 00\varepsilon_0(n))}g_{(0\cdots 0\varepsilon_0(n)\varepsilon_1(n))}\cdots g_{(\varepsilon_0(n)\varepsilon_1(n)\cdots\varepsilon_{L-1}(n))}g_{(\varepsilon_1(n)\varepsilon_2(n)\cdots\varepsilon_L(n))}\cdots \\ &\cdots g_{(\varepsilon_{\ell-L}(n)\varepsilon_{\ell-L+1}(n)\cdots\varepsilon_{\ell-1}(n))}g_{(\varepsilon_{\ell-L+1}(n)\varepsilon_{\ell-L+2}(n)\cdots\varepsilon_{\ell-1}(n)0)}\cdots g_{(\varepsilon_{\ell-1}(n)00\cdots 0)}. \end{aligned}$$

A famous example of block-multiplicative functions is the Rudin-Shapiro sequence (see [2, 4]) which is defined by $g_{(00)} = g_{(01)} = g_{(10)} = 1$ and $g_{(11)} = -1$ in the multiplicative group $H = \{-1,1\}$. There is also a close relation to chained sequences, see [2].

As in the case $L = 1$, we can introduce the Fourier term

$$F_\lambda(h) = \frac{1}{q^\lambda} \sum_{0 \leqslant u < q^\lambda} \mathrm{e}\left(-\frac{hu}{q^\lambda}\right) D(T(u)).$$

The only (but essential) difference to the case $L = 1$ is that F_λ has a more involved representation if $L > 1$. We have to introduce a $q^{L-1} \times q^{L-1}$ matrix that is indexed by blocks $B', C' \in \{0, 1, \ldots, q-1\}^{L-1}$ and has unitary $d \times d$ matrices as entries. We set

$$G_D(t) = \left(A_{B',C'}\right)_{B',C' \in \{0,1,\ldots,q-1\}^{L-1}},$$

where

$$A_{(b_1 \cdots b_{L-1}),(c_1 \cdots c_{L-1})} = \begin{cases} D(g_{(b_1 \cdots b_{L-1} c_{L-1})})\, e(tc_{L-1}) & \text{if } (b_2 \cdots b_{L-1}) = (c_1 \cdots c_{L-2}), \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$F_\lambda(h) = \frac{1}{q^\lambda} \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} G_D\left(-\frac{h}{q^\lambda}\right) G_D\left(-\frac{h}{q^{\lambda-1}}\right) \cdots G_D\left(-\frac{h}{q}\right) G_D(0)^{L-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

It is expected that all results of this paper generalize to the case $L > 1$. For example, if one can derive that $\|G_D(t)\| < q$ (under suitable conditions) one obtains $\|F_\lambda(h)\| \ll q^{-c\lambda}$.

References

[1] Isabelle Abou and Pierre Liardet. Flots chaînées. *Proceedings of the sixth Congress of Romanian Mathematicians, Bucharest, 2007, Scientific Contributions. Editors L. Beznea et al. Editura Academiei Române*, 1:401–432, 2009.

[2] Jean-Paul Allouche and Pierre Liardet. Generalized Rudin-Shapiro sequences. *Acta Arith.*, 60(1):1–27, 1991.

[3] Jean-Paul Allouche and Olivier Salon. Sous-suites polynomiales de certaines suites automatiques. *J. Théor. Nombres Bordeaux*, 5(1):111–121, 1993.

[4] Jean-Paul Allouche and Jeffrey Shallit. *Automatic sequences*. Cambridge University Press, Cambridge, 2003. Theory, applications, generalizations.

[5] Bruce C. Berndt, Ronald J. Evans, and Kenneth S. Williams. *Gauss and Jacobi sums*. Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons Inc., New York, 1998. A Wiley-Interscience Publication.

[6] Michael Drmota, Peter J. Grabner, and Pierre Liardet. Block additive functions on the Gaussian integers. *Acta Arith.*, 135(4):299–332, 2008.

[7] N. Pytheas Fogg. *Substitutions in dynamics, arithmetics and combinatorics*, volume 1794 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.

[8] Aleksandr O. Gelfond. Sur les nombres qui ont des propriétés additives et multiplicatives données. *Acta Arithmetica*, 13:259–265, 1968.

[9] Edmund Hlawka. Zur formalen Theorie der Gleichverteilung in kompakten Gruppen. *Rend. Circ. Mat. Palermo (2)*, 4:33–47, 1955.

[10] Lauwerens Kuipers and Harald Niederreiter. *Uniform Distribution of Sequences.* Wiley-Interscience Publication, 1974.

[11] Pierre Liardet. Automata and generalized Rudin-Shapiro sequences. *Arbeitsbericht Mathematisches Institut der Universität Salzburg*, 3-4:21–52, 1990.

[12] Christian Mauduit. Multiplicative properties of the Thue-Morse sequence. *Period. Math. Hungar.*, 43(1-2):137–153, 2001.

[13] Christian Mauduit and Joël Rivat. Sur un problème de Gelfond: la somme des chiffres des nombres premiers. Annals of Mathematics, to appear.

[14] Christian Mauduit and Joël Rivat. La somme des chiffres des carrés. *Acta Math.*, 203(1):107–148, 2009.

[15] Yossi Moshe. On the subword complexity of Thue-Morse polynomial extractions. *Theoret. Comput. Sci.*, 389(1-2):318–329, 2007.