MAXIMAL EVENNESS VIA DISCREPANCY THEORY

ADAM QUINN JAFFE

Abstract. The notion of *maximal evenness* was first defined in the context of music theory: For fixed integers n and k, how can one distribute n notes "as evenly as possible" in a scale $\mathbb{Z}_k$? Later work explored equivalent definitions of this intuitive idea using a variety of mathematical tools. Presently, we propose and explore a new definition based on discrepancy theory. First, we prove that this definition is equivalent to the original definition of Clough and Myerson. Then, we use this characterization to develop a fast algorithm (linear in time and space) for checking whether a given set is maximally even. Finally, we use the discrepancy characterization to propose a natural definition of maximal evenness for the spaces $\mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$ and $\mathbb{Z}$.

1. Introduction

Consider the following combinatorial problem: Given fixed integers k and n, which sets $A \subseteq \mathbb{Z}_k$ with $|A| = n$ achieve the "most even" distribution? Such sets are referred to as *maximally even* sets, and they are natural objects of study in mathematics, physics, and music theory [7].

The concept of maximal evenness was first introduced by Clough and Myerson in [6] in the context of tonal and atonal music theory, in which elements of $\mathbb{Z}_k$ are referred to as *pitch classes* (or *pc*s) and subsets of $\mathbb{Z}_k$ are referred to as *pitch class sets* (or *pc set*s). Their motivation for such work was the observation that many well-known musical scales can be viewed as maximally even subsets of the chromatic scale, $\mathbb{Z}_{12}$. In the subsequent decades, many music theorists and mathematicians began to formulate different precise notions of "even distribution", many of which turn out to be equivalent to the definition of Clough and Myerson; see Section 4, Theorem 7 of [7] for an outline of a few of these alternative definitions.

These equivalent definitions of maximal evenness are motivated by a variety of mathematical disciplines: combinatorics [5], optimization [3], discrete Fourier analysis [2], and, in our case, discrepancy theory. In accordance with these many connections accross mathematics, these ideas first came to me in the context of applied topology [1]: While studying the homotopy types and persistent homology of the regular polygons, it was observed that the edges of P_n which are touched by an inscribed equilateral star must form a maximally even set.

In this paper, we contribute another equivalent definition to this list. Our approach is based on *discrepancy theory*, as developed by Erdős and Turán in [9]. The language of discrepancy theory is very appealing when discussing maximal evenness, as it provides a useful means of quantifying how much a set of points fails to be equidstirbuted in the circle. Discrepancy theory has found a wide range of

2010 *Mathematics Subject Classification.* Primary 05B25, 68R15; Secondary 05A05.
Key words and phrases. maximal evenness, discrepancy theory, computational music theory.

Scale	Integer Set	Binary String
Ionian (Major)	$\{0,2,4,5,7,9,11\}$	101011010101
Aeolian (Minor)	$\{0,2,3,5,7,8,10\}$	101101011010
Harmonic Minor	$\{0,2,3,5,7,8,11\}$	101101011001
Whole Tone Scale	$\{0,2,4,6,8,10\}$	101010101010
Pentatonic Major	$\{0,2,4,7,9\}$	101010010100
Pentatonic Minor	$\{0,3,5,7,10\}$	100101010010

FIGURE 1. A few well-known musical scales, represented both as sets of integers and as binary strings.

applications across mathematics and computer science, including number theory [8], quasi-Monte-Carlo integration [13], and computational geometry [4].

Our main contributions are as follows: First, we prove that a set $A \subseteq \mathbb{Z}_k$ with $|A| = n$ is maximally even if and only if its discrepancy is bounded above by $1/n + 1/k$. Then we use this characterization to construct a fast algorithm (linear in time and in space) which checks whether a given set is maximally even. Our main tool for this algorithm is the construction of a "compression" map which preserves maximal evenness. Finally, we use the discrepancy characterization to generalize the notion of maximal evenness to subsets of spaces more general than $\mathbb{Z}_k$. In particular, we pose some open questions about maximal evenness for $\mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$ and $\mathbb{Z}$.

2. Definition

Consider any set $A \subseteq \mathbb{Z}_k$ with $|A| = n$. Let the tuple (k, n) be called the *type* of the set A. Now, let us be able to reimagine the set A in a variety of different ways, among which we will interchange freely throughout the paper. First, note that we may let A be a binary string by considering the indicator function $\mathbf{1}_A : \mathbb{Z}_k \to \{0, 1\}$. Towards this end, we may write such a string explicitly as $0^{g_0}10^{g_1}10^{g_2}\dots 10^{g_n}$, where $g_0, g_1, \dots g_n$ are nonnegative integers with $n + \sum_{e=0}^{n} g_e = k$. See Figure 1 for an example of some well-known scales represented in these forms.

Alternatively, we may want to study the geometry of A by embedding $\mathbb{Z}_k$ in $\mathbb{S}^1$ in the natural way. Precisely, we may think of A as the set $\{\omega_k^i : i \in A\}$, where ω_k is the kth root of unity. From this perspective, it is clear that we should not distinguish between binary strings which agree up to a cyclic permutation of their digits, and are hence motivated to form the following notion of equivalence:

Definition 2.1. Suppose that A_1 and A_2 are two binary strings of type (k, n). We say that A_1 and A_2 are *translation-equivalent*, denoted $A_1 \sim A_2$, if there exists some integer $r \in \mathbb{Z}_k$ such that $r + A_1 = \{r + a : a \in A_1\} = A_2$. A *binary necklace* or *necklace* $\tilde{A}$ is an equivalence class under $\sim$, and a set A in this equivalence class is called an *unfolding* of $\tilde{A}$.

In Figure 2 we see a few examples of binary necklaces represented in combinatorial form and in geometric form. Note in particular that these necklaces correspond exactly to the musical scales of Figure 1. Such a quotient structure is instrumental for the results of this paper, as it will turn out to be the case that many important properties hold for all binary strings within an equivalence class. In particular, type and maximal evenness are both translation-invariant properties. In this way,

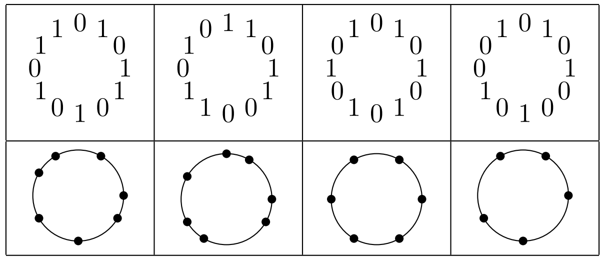

FIGURE 2. The musical scales of Figure 1, represented both as binary necklaces and as collections of points in $\mathbb{S}^1$.

we will freely interchange between saying that a binary string has a property and that the associated binary necklace has that property.

Next, we invest in some new notation that will be useful for the remainder of the paper. For indices $i, j \in \mathbb{Z}_k$, let $\vec{d}_k(i,j)$ represent the *oriented distance* from i to j in $\mathbb{Z}_k$. That is, $\vec{d}_k(i,j) = j - i$ if $i \leq j$, and $\vec{d}_k(i,j) = k - j + i$ if $i > j$. Moreover, we define $[i,j)_k$ to be the interval $[\omega_k^i, \omega_k^j) \subseteq \mathbb{S}^1$ and likewise for other types of intervals. We also let λ be the Lebesgue (probability) measure on $\mathbb{S}^1$. Notably, these definitions give $\vec{d}_k(i,j)/k = \lambda([i,j)_k)$.

This in mind, we can recall the definition of maximal evenness given by Clough and Myerson in [6]. Observe, in particular, that maximal evenness is a translation-invariant property of binary strings. That is, we may discuss the maximal evenness of binary strings and binary necklaces alike.

Definition 2.2. A binary string A of type (k, n) is called *maximally even* if any elements $i, j \in A$ give

$$\left| \vec{d}_k(i,j) - \frac{k}{n} \#(A \cap [i,j)_k) \right| < 1. \tag{1}$$

We call a binary necklace maximally even if any (equivalently, all) of its unfoldings are maximally even.

What Clough and Myerson identified is that the values $\vec{d}_k(i,j)$ ("length") and $A \cap [i,j)_k$ ("mass") should be dual to each other, in the sense that the value of one should almost determine the value of the other. In this way, it is clear that a set A which fails to be maximally even must have some interval whose length and mass are not in accordance. But observe that $\vec{d}_k(i,j)$ is "too long" compared to $A \cap [i,j)_k$ if and only if $\vec{d}_k(j,i)$ is "too short" compared to $A \cap [j,i)_k$. Using this fact, we can reduce the amoung of case-checking that may be required when working with sets that are known to not be maximally even.

Likewise, let us recall the definition of discrepancy given in [9]. Intuitively, this quantity represents the degree to which length and mass fail to be in accordance.

Definition 2.3. For X a finite set of points in $\mathbb{S}^1$, the *discrepancy* of X is

$$D(X) = \sup_{I \subseteq \mathbb{S}^1} \left| \frac{\#(X \cap I)}{\#X} - \lambda(I) \right| \tag{2}$$

where the supremum is taken over all intervals (connected subsets) of $\mathbb{S}^1$.

We observe that there seems to be some relation between (1) and (2): The value $\lambda(I)$ corresponds to $\vec{d}_k(i,j)/k$ and the value $\#(A \cap [i,j)_k)/n$ corresponds to $\#(X \cap I)/n$. Hence, it appears we may be able to cast maximal evenness in terms of discrepancy by making this correspondence precise. Indeed, we do this by the following combinatorial lemma:

Lemma 2.4. A set A of type (k,n) is maximally even if and only if and any indices $i,j \in \mathbb{Z}_k$ give

$$\left| \#(A \cap [i,j)_k) - n\frac{\vec{d}_k(i,j)}{k} \right| < 1. \tag{$*$}$$

Proof. First suppose that A of type (k,n) is maximally even, and let $i,j \in \mathbb{Z}_k$ be arbitrary indices. We may then choose indices $i', j' \in A$ such that $\#(A \cap [i,j)_k) = \#(A \cap [i',j')_k)$ and $|\vec{d}_k(i',j') - \vec{d}_k(i,j)| \le \frac{k}{n} - 1$ both hold. Then we can bound:

$$\begin{aligned} \left| \#(A \cap [i,j)_k) - \frac{n}{k}\vec{d}_k(i,j) \right| &\le \frac{n}{k}\left| \vec{d}_k(i',j') - \frac{k}{n}\#(A \cap [i',j')_k) \right| + \frac{n}{k}\left| \vec{d}_k(i',j') - \vec{d}_k(i,j) \right| \\ &< \frac{n}{k} + \frac{n}{k}\left(\frac{k}{n} - 1\right) = 1. \end{aligned}$$

Conversely, suppose that A is not maximally even. That is, there are indices $i,j \in A$, such that we have

$$\vec{d}_k(i,j) \ge \frac{k}{n}\#(A \cap [i,j)_k) + 1. \tag{3}$$

In particular, let us assume that i,j are such that they minimize the value $\vec{d}_k(i,j)$ over all pairs $i,j \in A$ satisfying (3). Then the interval $[i+1,j)_k$ must satisfy $\#(A \cap [i+1,j)_k) = \#(A \cap [i,j)_k) - 1$ and $\vec{d}_k(i+1,j) = \vec{d}_k(i,j) - 1$. Plugging this into (3) and rearranging gives

$$\#(A \cap [i+1,j)_k) \le \frac{n}{k}\vec{d}_k(i+1,j) - 1. \tag{4}$$

Hence, the indices $i+1, j \in \mathbb{Z}_k$ violate $(*)$. $\square$

Finally, we can state our first important result:

Theorem 2.5. A set A of type (k,n) is maximally even if and only if it satisfies the discrepancy bound $D(A) < 1/n + 1/k$.

Proof. First, let A be a set of type (k,n), and suppose that A is maximally even. Then note that, for any inteval $I \subseteq \mathbb{S}^1$, we may choose indices $i,j \in \mathbb{Z}_k$ such that we get $\#(A \cap I) = \#(A \cap [i,j))$ and $|\vec{d}_k(i,j)/k - \lambda(I)| < 1/k$. Then it follows that we get

$$\begin{aligned} \left| \frac{\#(A \cap I)}{n} - \lambda(I) \right| &\le \frac{1}{n}\left| \#(A \cap I) - n\frac{\vec{d}_k(i,j)}{k} \right| + \left| \frac{\vec{d}_k(i,j)}{k} - \lambda(I) \right| \\ &< \frac{1}{n} + \frac{1}{k}, \end{aligned} \tag{5}$$

where the second inequality follows from Lemma 2.4. But note that the transformation $I \mapsto [i,j)_k$ maps the set of all intervals $I \subseteq \mathbb{S}^1$ onto the set of all pairs of indices in $\mathbb{Z}_k$, which is a finite set. Hence, the supremum in Definition 2.3 is taken over a finite set of values which are all strictly less than $1/n + 1/k$. The resulting value must then be strictly less than $1/n + 1/k$, so we get $D(A) < 1/n + 1/k$.

Conversely, suppose that A is not maximally even. Then by Lemma 2.4, there exist indices $i, j \in \mathbb{Z}_k$ such that $(*)$ fails. In particular, we have

$$\#(A \cap [i,j)_k) - n\frac{\vec{d}_k(i,j)}{k} \geq 1. \tag{6}$$

Then we consider the interval $I = [i, j-1]_k$. Observe that $\#(A \cap I) = \#(A \cap [i,j)_k)$ and $\lambda(I) = \vec{d}_k(i,j)/k - 1/k$. Plugging this in to $(*)$ and rearranging, we get

$$\frac{\#(A \cap I)}{n} - \lambda(I) \geq \frac{1}{n} + \frac{1}{k}. \tag{7}$$

Hence, A fails the discrepancy bound. □

We now explore the prospect of letting Theorem 2.5 serve as the definition of maximal evenness. Indeed, the following advantages are clear:

First, the definition implies that maximal evenness coincides with a notion of equidistribution, at least in the following asymptotic sense: Let $\{(k_i, n_i)\}_{i=1}^{\infty}$ be a sequence of types with $n_i \to \infty$ (hence $k_i \to \infty$) as $i \to \infty$, let A_i be a maximally even set of type (k_i, n_i) for each i, and let $I \subseteq \mathbb{S}^1$ be some fixed interval. Then, we get $\#(A_i \cap I)/n_i \to \lambda(I)$ as $i \to \infty$. That is, the proportion of $\{A_i\}_{i=1}^{\infty}$ contained in I is determined only by the size of I.

Second, a few key results about maximal evenness are very easy to prove from this starting point: A is maximally even if and only if its bitwise complement $A \oplus 1^k$ is maximally even; A is maximally even if and only if any translate $r + A$ is maximally even; and, there is a unique maximally even binary necklace for each type (Theorem 3.8). None of these properties is obvious from any prior definition of maximal evenness; in fact, they took quite a bit of effort to prove in the original work.

Finally, the well-understood theory of discrepancy puts maximal evenness on a deeper mathematical footing compared to previous definitions. As we will later see via the generalizations of Section 4, maximal evenness may have deep connections to ergodic theory and number theory.

3. Algorithm

As an application of Lemma 2.4/Theorem 2.5, we create a fast algorithm for checking whether a given set $A \subseteq \mathbb{Z}_k$ is maximally even. To do this, we first consider the procedure which converts A from its representation as a set to its representation as a binary string; it is clear that such an algorithm runs in $\Theta(k)$ time and $\Theta(k)$ space. For the remainder of this section, it will be assumed that A has been stored as a binary string.

Next, we develop some machinery that will help us define the algorithm that will check if a given binary necklace is maximally even or not.

Definition 3.1. We say that a binary necklace $\tilde{A}$ of type (k,n) with $n>0$ has *balanced gaps* if any unfolding A of $\tilde{A}$ of the form $A = 10^{g_1}10^{g_2}\dots 10^{g_n}$ has, for all $e \in \{1,\dots n\}$, either $g_e = \lfloor \frac{k}{n}-1 \rfloor$ or $\lceil \frac{k}{n}-1 \rceil$.

We now seek an algorithm which checks whether a given binary string has balanced gaps. Evidently, we can compute and store the size of the gaps in a binary string using a method GetGaps that runs in $\Theta(k)$ time and $\Theta(n)$ space. Using this method, we can pass through the resulting array and check whether each gap is of the desired size. This gives rise to an algoritm HasBalancedGaps that runs in $\Theta(k)+O(n)=\Theta(k)$ time and $\Theta(n)$ space. We also note the following property of balanced gap necklaces: If A has balanced gaps, then the number of e with $g_e = \lfloor \frac{k}{n}-1 \rfloor$ is $n-(k \mod n)$ and the number of e with $g_e = \lceil \frac{k}{n}-1 \rceil$ is $k \mod n$. Next, we prove the following basic properties of balanced gap necklaces.

Lemma 3.2. If n divides k, then there is only one binary necklace with balanced gaps, and it is maximally even.

Proof. If n divides k, then $\lfloor \frac{k}{n}-1 \rfloor = \lceil \frac{k}{n}-1 \rceil = \frac{k}{n}-1$. Hence, the only possible necklace with bounded gaps is the $\tilde{A}$ which corresponds to the unfolding $A = (10^{\frac{k}{n}-1})^n$. Now take any indices $i,j \in \mathbb{Z}_k$ and define $a = \#(A \cap [i,j))$. Then let us write $A \cap [i,j) = 0^{g_0}10^{g_1}\dots 10^{g_a}$ so that we get $g_e = \frac{k}{n}-1$ for all $e \in \{1,\dots a-1\}$ and $0 \le g_e \le \frac{k}{n}-1$ for $e = 0,a$. Then we may bound:

$$\left| a - \frac{n}{k}\vec{d}_k(i,j) \right| = \left| a - \frac{n}{k}(a + \sum_{e=0}^{a} g_e) \right| \tag{8}$$
$$= \frac{1}{k}\left| k-n-n(g_0+g_a) \right|$$

But note that that $0 \le g_0, g_a \le \frac{k}{n}-1$ implies $|k-n-n(g_0+g_a)| \le k-n$, so we can further bound (8) by:

$$\left| a - \frac{k_1}{k}\vec{d}_k(i,j) \right| = 1-\frac{n}{k} < 1 \tag{9}$$

Therefore, $\tilde{A}$ is maximally even. □

Lemma 3.3. If $\tilde{A}$ is maximally even, then it has balanced gaps.

Proof. Suppose that $\tilde{A}$ is maximally even, and let A be any unfolding of $\tilde{A}$. Then any substring of the form $10^{g_e}1$ at indices $[i,j]$ in A, has, by Definition 2.2, $|g_e - (\frac{k}{n}-1)| < 1$. This forces $g_e = \lfloor \frac{k}{n}-1 \rfloor$ or $\lceil \frac{k}{n}-1 \rceil$, as desired. □

This discussion of balanced gap necklaces leads us to the following construction, which should be seen as the key step in the development of this theory.

Definition 3.4. Let $\tilde{A}$ be a binary necklace with balanced gaps whose type is (k,n) with $n>0$. If n does not divide k, then we can construct a new binary necklace by substituting

$$10^{\lfloor \frac{k}{n}-1 \rfloor} \mapsto 0 \text{ and } 10^{\lceil \frac{k}{n}-1 \rceil} \mapsto 1. \tag{10}$$

The resulting necklace is called the *compression* of $\tilde{A}$, denoted compress($\tilde{A}$). If it is the case that n divides k, then we use the convention that compress($\tilde{A}$) is the binary necklace of length n which consists only of 0's.

Remark 3.5. If $\tilde{A}$ is any binary necklace of type (k, n), then compress($\tilde{A}$) has type $(n, k \mod n)$. In particular, we see that the number of 1's in the compressed necklace is strictly less than the number of 1's in the original necklace.

It should be noted that, to be fully precise with this definition, the type of the binary necklace should parameterize a family of compress maps. This abuse of notation will be convenient in our mathematical notation, but it will be addressed more carefully in our algorithms. Algorithm 1 contains an implementation of the compress map, and it is clear that it runs in $\Theta(k)$ time and $\Theta(n)$ space.

Algorithm 1 Compressing a binary string

procedure COMPRESS(A, k, n)
    $C \leftarrow$ LIST(n)
    $G \leftarrow$ GETGAPS(A)
    for $e \in \{0, \dots n-1\}$ **do**
        if $G[e] = \lfloor \frac{k}{n} - 1 \rfloor$ **then**
            $C[e] \leftarrow 0$
        else
            $C[e] \leftarrow 1$
    return C

Now we can proceed to the main lemma of this section:

Lemma 3.6. A binary necklace $\tilde{A}$ with balanced gaps is maximally even if and only if compress($\tilde{A}$) is maximally even.

Proof. Suppose that $\tilde{A}$ has balanced gaps and is of type (k, n). The claim is easy when n divides k, since the following must hold: $\tilde{A}$ is maximally even because of Lemma 3.2, and compress($\tilde{A}$) is maximally even because its type is $(n, 0)$. Now we proceed assuming that n does not divide k, which allow us to use the useful identity:

$$\left\lceil \frac{k}{n} - 1 \right\rceil - \left\lfloor \frac{k}{n} - 1 \right\rfloor = 1 \tag{11}$$

First suppose that $\tilde{A}$ of type (k, n) with balanced gaps is not maximally even. Then by Lemma 2.4, there exists some unfolding A of $\tilde{A}$ and some indices $i, j \in \mathbb{Z}_k$ that satisfy

$$\#(A \cap [i, j)_k) \geq \frac{n}{k} \vec{d_k}(i, j) + 1. \tag{12}$$

Let us assume that $i, j \in \mathbb{Z}_k$ are such that they minimize the value of $\vec{d_k}(i, j)$. Then taking $a = \#(A \cap [i, j)_k)$, we can write $A \cap [i, j)_k = 10^{g_1}1 \dots 10^{g_a}1$ with $g_e = \lfloor \frac{k}{n} - 1 \rfloor$ or $\lceil \frac{k}{n} - 1 \rceil$ for all $e = 1, \dots a$. Then (12) becomes

$$a+1 \geq \frac{n}{k}\left(a+1+\sum_{e=1}^{a} g_e\right)+1. \tag{13}$$

Letting r denote the number of $e \in \{1, \dots a\}$ for which $g_e = \lceil \frac{k}{n} - 1 \rceil$, we see that (13) can be expressed as

$$\begin{aligned} a+1 &\geq \frac{n}{k}\left(a+1+(a-r)\left\lfloor \frac{k}{n}-1 \right\rfloor + r\left\lceil \frac{k}{n}-1 \right\rceil\right)+1 \\ &= \frac{n}{k}\left(a+1+a\left\lfloor \frac{k}{n}-1 \right\rfloor + r\right)+1, \end{aligned} \tag{14}$$

where the equality follows from identity (11). Then check that (14) rearranges to

$$r \leq \left(\frac{k}{n}-1-\left\lfloor \frac{k}{n}-1 \right\rfloor\right)a-1 = \left(\frac{k \mod n}{n}\right)a-1. \tag{15}$$

But observe that, by definition, r is exactly the number of 1's in the image of $A \cap [i,j)_k$ under the compress map. Therefore, compress$(\tilde{A})$ is not maximally even.

For the second direction, suppose that $\tilde{A}$ of type (k,n) with balanced gaps is such that $\tilde{C} = \text{compress}(\tilde{A})$ is not maximally even. Then there must exist an unfolding C of $\tilde{C}$ and indices $i,j \in \mathbb{Z}_n$ such that, for $r = \#(C \cap [i,j)_k)$ and $a = \vec{d}_k(i,j)$, we get

$$r \leq \left(\frac{k \mod n}{n}\right)a-1 = \left(\frac{k}{n}-1-\left\lfloor \frac{k}{n}-1 \right\rfloor\right)a-1. \tag{16}$$

But observe that (16) is equivalent to (15), so it rearranges to (14). Then note that there exists an unfolding A of $\tilde{A}$ and indices $i',j' \in \mathbb{Z}_k$ such that $C \cap [i,j)_k$ is the image of $A \cap [i',j')_k = 10^{g_1} \dots 10^{g_a}$ under the compress map. This implies $a+(a-r)\lfloor \frac{k}{n}-1 \rfloor + r\lceil \frac{k}{n}-1 \rceil = \vec{d}_k(i,j)$. Then we have

$$a+1 \geq \frac{n}{k}(\vec{d}_k(i',j')+1)+1. \tag{17}$$

Let us consider this inequality with respect to the string $A \cap [i', j'+1) = 10^{g_1} \dots 10^{g_a}1$. On the left side, $a+1$ is equal to $\#(A \cap [i',j'+1))$, and, on the right side $\vec{d}_k(i',j')+1$, is equal to the length of the interval in question. That is, (17) implies that $\tilde{A}$ is not maximally even, which completes the proof. □

Theorem 3.7. A binary necklace $\tilde{A}$ is maximally even if and only if compress$^m(\tilde{A})$ has type $(k',0)$ for some nonnegative integer m and positive integer k'.

Proof. First, let $\tilde{A}$ be a maximally even binary necklace. Then use Proposition 3.6 inductively to see that every necklace that is an element of the sequence $\{\text{compress}^0(\tilde{A}), \text{compress}^1(\tilde{A}), \text{compress}^2(\tilde{A}), \dots\}$ is maximally even. Denoting the respective types of these necklaces by $\{(k_0,n_0),(k_1,n_1),((k_2,n_2),\dots\}$, Remark 3.5 guarantees that we must have $n_0 > n_1 > n_2 > \dots$. As this is a strictly decreasing sequence of positive integers, there must be some nonnegative integer m such that $n_m = 0$ holds. Then we see that compress$^m(\tilde{A})$ has type $(k_m, 0)$, as desired.

Conversely, suppose that $\tilde{A}$ of type (k,n) is such that $\text{compress}^m(\tilde{A})$ is of type $(k',0)$ for some nonnegative integer m and some positive integer k'. Use Proposition 3.6 inductively to see that $\{\text{compress}^m(\tilde{A}), \text{compress}^{m-1}(\tilde{A}), \ldots \text{compress}^0(\tilde{A})\}$ is a finite sequence of maximally even necklaces. In particular, we see that the element $\text{compress}^0(\tilde{A}) = \tilde{A}$ is maximally even, as desired. □

We would be remiss not to state and prove a uniqueness result for maximal evenness in this paper. While such a result is not novel by any means (see Section 4, Proposition 5A of [7]), the theory developed in this section reveals an elegant proof of the well-known fact:

Theorem 3.8. There is a unique maximally even binary necklace for each type. That is, if A_1 and A_2 are maximally even binary strings of the same type, then they are translation-equivalent: $A_1 \sim A_2$.

Proof. Let $\tilde{A}_1$ and $\tilde{A}_2$ be maximally even binary necklaces of the same type (k,n). By Theorem 3.7, there exist nonnegative integers m_1, m_2 and positive integers k_1', k_2' such that $\text{compress}^{m_1}(\tilde{A}_1)$ is of type $(k_1', 0)$ and $\text{compress}^{m_2}(\tilde{A}_2)$ is of type $(k_2', 0)$. However, Remark 3.5 guarantees that $m_1 = m_2$ and $k_1' = k_2'$. Then observe that every compress map is invertible (its inverse can be explicitly constructed if the type of the domain necklace is known) hence injective. Since the composition of injective functions is injective, we conclude $\tilde{A}_1 = \tilde{A}_2$. □

Finally, we are able to complete the task of computationally checking whether a given binary necklace is maximally even; Theorem 3.7 suggests a recursive algorithm for doing this. While it is slow in its current form, its runtime can be improved dramatically by making the following modification: If necessary, flip the bits of $\text{compress}(\tilde{A})$ in order to maintain that $n \leq \frac{k}{2}$ holds at every function call. In particular, types will be mapped under compress as follows: If $k \mod n < \frac{n}{2}$, then $(k,n) \mapsto (n, n-(k \mod n))$; if $k \mod n \geq \frac{n}{2}$, then $(k,n) \mapsto (n, n-(k \mod n))$. We record this procedure precisely in Algorithm 2.

Theorem 3.9. For a binary string A of type (k,n), Algorithm 2 returns whether or not A is maximally even. It runs in $\Theta(k)$ time and in $\Theta(n)$ space. The depth of its function call stack is $O(\log_2 n)$.

Proof. Theorem 3.7 proves the correctness of the unmodified algorithm. To see that the modification does not alter the correctness, simply recall that A is maximally even if and only if $A \oplus 1^k$ is maximally even.

To find the time complexity, recall that line 14 runs in $\Theta(k)$ time, and that all other lines run in $O(1)$ time. Since the modification guarantees that k drops by a factor of at least 2 at every function call, we see that the total time complexity is $\Theta(k) + \Theta(\frac{k}{2}) + \Theta(\frac{k}{4}) + \cdots = \Theta(k)$. Likewise, line 14 runs in $\Theta(n)$ space and all other lines are $O(1)$ space. Since the modification guarantees that n also drops by a factor of at least 2 at every function call, we see that the total space complexity is $\Theta(n) + \Theta(\frac{n}{2}) + \Theta(\frac{n}{4}) + \cdots = \Theta(n)$. The exponential decay of n also guarantees that the function call stack has a depth of $O(\log_2 n)$. □

Algorithm 2 Checking if a binary string is maximally even, recursively

```
procedure IsMaxEven(A)
    k ← length(A)
    n ← sum(A)
    return IsMaxEvenHelper(A, k, n)

procedure IsMaxEvenHelper(A, k, n)
    if n = 0 then
        return true
    else if not HasBalancedGaps(A, k, n) then
        return false
    else if n divides k then
        return true
    else
        C ← Compress(A, k, n)
        if k mod n < n/2 then
            return IsMaxEvenHelper(C, n, k mod n)
        else
            C ← C ⊕ 1^n                                  ▷ Flip all of the bits
            return IsMaxEvenHelper(C, n, n − (k mod n))
```

4. Generalizations

This section addresses two natural discrepancy-theoretic generalizations of maximal evenness to spaces more general than $\mathbb{Z}_k$: the higher-dimensional case of $\mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$ and the limiting object of $\mathbb{Z}$.

First, we study maximal evenness in $\mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$. To do this, let us embed $\mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$ inside the torus $(\mathbb{S}^1)^d$ in the same way as we did $\mathbb{Z}_k$ into $\mathbb{S}^1$: Elements of $\mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$ may be thought of as integer tuples $(i_1, \ldots i_d) \in \mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$ or as geometric points $(\omega_{k_1}^{i_1}, \ldots \omega_{k_d}^{i_d}) \in (\mathbb{S}^1)^d$. We may also sometimes think of a subset $A \subseteq \mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$ as a d-dimensional tensor $A = \{a_{j_1,\ldots j_d}\}_{j_1,\ldots j_d=1}^{k_1,\ldots k_d}$ where each element is either 0 or 1. In any case, these objects are referred to as *binary sheets*. For any binary sheet $A \subseteq \mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$ with $|A| = n$, we let its *type* be the tuple $(k_1, \ldots k_d, n)$.

Now we make precise what is meant by the generalization of discrepancy to these higher-dimensional tori:

Definition 4.1. For X a finite set of points in $(\mathbb{S}^1)^d$, the *discrepancy* of X is

$$D(X) = \sup_{R \subseteq (\mathbb{S}^1)^d} \left| \frac{\#(X \cap R)}{\#X} - \lambda(R) \right| \tag{18}$$

where the supremum is taken over all rectangles $R = I_1 \times \cdots \times I_d$ in $(\mathbb{S}^1)^d$.

This allows us to define the desired property.

Definition 4.2. A set A of type $(k_1, \ldots k_d, n)$ is called *maximally even* if it satisfies the discrepancy bound $D(A) < \frac{1}{n} + E(k_1, \ldots k_d)$, where E is the function

$$E(k_1, \dots k_d) = 1 - \prod_{i=1}^{d} \left(1 - \frac{1}{k_i}\right). \tag{19}$$

While this function E may seem mysterious at first, we remark that it in fact represents the error incurred by "snapping" a rectangle $R \subseteq (\mathbb{S}^1)^d$ to the grid $\mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$.

Generalizing the quotient structure of Definition 2.1 brings us to a subtlety of the theory: It appears that translation equivalence is too weak of an identification to guarantee a unique binary sheet of each type. For example there exist $A_1, A_2 \subseteq \mathbb{Z}_k \times \mathbb{Z}_k$ such that $r + A_1 \neq A_2$ for all $r \in \mathbb{Z}_k \times \mathbb{Z}_k$, but nonetheless with $\{(a_2, a_1) : (a_1, a_2) \in A_1\} = A_2$.

Question 4.3. Is there a notion of equivalence for binary sheets of type $(k_1, \dots k_d, n)$ such that there is a unique maximally even equivalence class of each type?

Another nuance is that it is not clear how to generalize the balanced gaps condition or the compress map to subsets $A \subseteq \mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_d}$, if such a generalization is even possible. One idea is to consider the projections of the tensor A onto its $(d-1)$-dimensional hyperplanes, thus reducing the d-dimensional problem to several problems in a lower dimension. Since the base case $d = 1$ is well-understood, such a recursive consideration may be helpful. At any rate, much care is needed in developing this theory.

Question 4.4. What is a fast algorithm for checking whether a given binary sheet is maximally even?

Second, we turn to studying maximal evenness in the integers. That is, which sequences $A \subseteq \mathbb{Z}$ achieve the "most even" distribution possible? As before, we may think of any A as a set of integers or equivalently as an infinite binary sequence, and we identify two binary sequences when they agree up to some translation. Then, we can think about maximal evenness for these objects by noticing that the equation

$$\left| \frac{\#(A \cap I)}{n} - \frac{\lambda(I)}{k} \right| < \frac{1}{n} + \frac{1}{k}$$

can be written as

$$\left| \#(A \cap I) - \frac{n}{k}\lambda(I) \right| < 1 + \frac{n}{k}.$$

In this way, the value $\frac{n}{k}$ can be seen as the "density" of A. This motivates us to form the following definition:

Definition 4.5. A sequence $A \in \mathbb{Z}$ is called *maximally even* if there exists a real number $\alpha \in [0, 1]$ such that the following bound holds:

$$\sup_{I \subseteq \mathbb{R}} \left| \#(A \cap I) - \alpha\lambda(I) \right| < 1 + \alpha, \tag{20}$$

where the supremum is taken over all intervals (connected subsets) of $\mathbb{R}$, and λ is the Lebesgue measure on $\mathbb{R}$.

We can justify the use of the name "density" given to α by observing that, if α is the maximal evenness density of A, then it is also the natural density of A. However, the converse is not true general. While natural density is only required to hold asymptotically, Definition 4.5 requires that a density condition holds on every interval of $\mathbb{R}$. As such, density in the sense of Definition 4.5 is an extremely strong requirement.

Also, we can define the balanced gaps conditions and the compress map analogously with the case of $\mathbb{Z}_k$:

Definition 4.6. We say that a binary sequence A of density $\alpha > 0$ has *balanced gaps* if any indices $i, j \in A$ with $A \cap (i, j) = 0$ satisfy $j - i = \lfloor \frac{1}{\alpha} - 1 \rfloor$ or $\lceil \frac{1}{\alpha} - 1 \rceil$.

Definition 4.7. Let A be a binary sequence with balanced gaps whose density is $\alpha > 0$. Then we can define the compress map by sending

$$10^{\lfloor \frac{1}{\alpha} - 1 \rfloor} \mapsto 0 \text{ and } 10^{\lceil \frac{1}{\alpha} - 1 \rceil} \mapsto 1. \tag{21}$$

Let us now consider the effect of the compress map on the density of a given binary sequence. It can be shown that for any balanced gap sequence A of density $\alpha > 0$, the density of compress(A) is $T(\alpha) = 1/\alpha \mod 1$. Inductively, then, the density of compress$^m(A)$ must be $T^m(\alpha)$. Here we recognize that T is exactly the map that generates the continued fraction coefficients of α.

Next we make the following conjecture, which can be seen as link between the $\mathbb{Z}_k$ theory and the $\mathbb{Z}$ theory. Moreover, it provides some useful insight into how maximally even binary sequences are generally structured.

Conjecture 4.8. Suppose that $A \subseteq \mathbb{Z}$ is a maximally even with density α. Then, the following are equivalent:

(1) The set A is periodic
(2) The density α is rational
(3) There exists some nonnegative integer m such that compress$^m(A) = \{0\}$.

Dually, the following are equivalent:

(4) The set A is aperiodic
(5) The density α is irrational
(6) For all nonnegative integers m, we get compress$^m(A) \neq \{0\}$.

A few questions naturally arise from Conjecture 4.8. First, property (1) inspires us to ask the question of "how aperiodic" A really is. This vague question, it appears, may be made precise with the language of ergodic theory.

Question 4.9. How aperiodic is A?

Next, we recall that the key insight to proving the uniqueness of maximally even binary strings (Theorem 3.8) was that the iterated compress map is forced to eventually hit 0. However, (3) above guarantees that this map will never terminate for irrational α. Hence, we must seek an alternative method of proof to conclude a similar uniqueness result.

Question 4.10. Given an irrational density α, does there exist a unique maximally even sequence $A \subseteq \mathbb{Z}$, up to translation-equivalence?

Finally, we ask what is the common framework (algebraic, geometric, or topological) that allows us to develop these theories of maximal evenness via discrepancy theory. For example, the answer to the following simple question is not at all clear:

Question 4.11. How should one define maximal evenness for subsets $A \subseteq \mathbb{Z}_k \times \mathbb{Z}$?

5. Acknowledgements

This paper began during my residence at the Institute for Computational and Experimental Research in Mathematics (ICERM) in Providence, RI in the Summer of 2017, where I was supported by the National Science Foundation under Grant No. DMS-1439786. I am very grateful to the students and faculty who helped me formulate this problem. Specific thanks are due to Henry Adams, Samir Chowdhury, and Peter Eastwood. I also owe thanks to Jimmy He, Atticus Geiger, Prasanna Ramakrishnan, and Bora Uyumazturk at Stanford for many helpful conversations.

References

[1] Henry Adams, Samir Chowdhury, Adam Quinn Jaffe, and Bonginkosi Sibanda. Vietoris-Rips Complexes of Regular Polygons. Preprint: `https://arxiv.org/abs/1807.10971`.

[2] Emmanuel Amiot. David Lewin and Maximally Even Sets. *Journal of Mathematics and Music*, 1(3):120, 2007.

[3] Steven Block and Jack Douthett. Vector Products and Intervallic Weighting. *Journal of Music Theory*, 38(1):21–41, 1994.

[4] Bernard Chazelle. *Discrepancy Theory and Computational Geometry*. Spring Lecture Notes in Computer Science, 1997.

[5] John Clough and Jack Douthett. Maximally Even Sets. *Journal of Music Theory*, 35(1):93–173, 1991.

[6] John Clough and Gerald Myerson. Variety and Multiplicity in Diatonic Systems. *Journal of Music Theory*, 29(2):249–270, 1985.

[7] Jack Douthett and Richard Krantz. Maximally Even Sets and Configurations: Common Threads in Mathematics, Physics, and Music. *Journal of Combinatorial Optimization*, 14:385410, 2007.

[8] Paul Erdős and R. R. Hall. On the Angular Distribution of Gaussian Integers with Fixed Norm. *Discrete Mathematics*, 200(1):87–94, 1999.

[9] Paul Erdős and Paul Turán. On a Problem in the Theory of Uniform Distribution I, II. *Koninklijke Nederlandse Akademie van Wetenschappen*, 51:1146–1154,1262–1269, 1948.

[10] Ronald Graham, Donald Knuth, , and Oren Patashnik. *Concrete Mathematics: A Foundation in Computer Science*. Addison-Wesley Publishing Company, 1989.

[11] J. R. Hollos and S. Hollos. *Creating Rhythms*. Abrazol Publishing, 2014.

[12] Minghui Jiang. On the Sum of Distances Along a Circle. *Discrete Mathematics*, 308(10):2038–2045, 2008.

[13] Gunther Leobacher and Friedrich Pillichshammer. *Introduction to Quasi-Monte Carlo Integration and Applications*. Birkhuser Basel, 2014.

[14] Godfried T. Toussaint. *The Geometry of Musical Rhythm: What Makes a Good Rhythm Good?* Chapman and Hall/CRC, 2013.

Department of Mathematics, Stanford University, Stanford, CA 94305, United States
E-mail address: `aqjaffe@stanford.edu`