Symmetries of the Three-Gap Theorem

Aneesh Dasgupta and Roland Roeder

Abstract. The Three-Gap Theorem states that for any $\alpha \in \mathbb{R}$ and $N \in \mathbb{N}$, the fractional parts of $\{0, 1\alpha, \ldots, (N-1)\alpha\}$ partition the unit circle into gaps of at most three distinct lengths. It is also of interest to find patterns in how the order of different gap sizes appear as one goes counterclockwise around the circle. This note is devoted to proving a result about symmetries in this ordering.

1. INTRODUCTION. Choose an irrational angle $\alpha$ measured in “turns,” where one turn corresponds to $2\pi$ radians, and plot the points on the circle at angles

$$0, \quad \alpha, \quad 2\alpha, \quad 3\alpha, \ldots, (N-1)\alpha.$$  

For $\alpha = \sqrt{2}$ and $N = 27$, one obtains Figure 1.

![Figure 1](image-url)

Figure 1. Illustration of the 3-gap theorem for $\alpha = \sqrt{2}$ and $N = 27$. The short gaps are labeled with $a$, the medium gaps with $b$, and the longest gaps with $c$. One example of the reflectional symmetry that is proved in the Symmetry Theorem is indicated with the center of symmetry being the red $c$ between $10\alpha$ and $15\alpha$, and the symmetric letters shown in blue. The length of the symmetry is indicated by the next closest $c$ gap, which is also shown in red (between $11\alpha$ and $16\alpha$).

A surprising observation is that for any choice of $N$ and $\alpha$, the distances (gaps) between consecutive points on the circle attain only three values. This is the content of the famous Three-Gap Theorem, proved by Sós, Surányi, and Świerczkowski in the 1950s, and it can be seen in the special case of Figure 1. It is also of interest to find patterns in how the order of different gap sizes appear on the circle, and we invite the
reader to find their own in Figure 1. This note is devoted to describing one such pattern:
a curious symmetry in how the sizes of the gaps are distributed on the circle, illustrated
in Figures 1 and 3. For a precise statement, see the Symmetry Theorem below.

**Setup.** In order to work more carefully, it is convenient to represent the circle as the
interval \([0, 1]\) with the endpoints identified. We will now rephrase the setup in this
context, and state the Three-Gap Theorem more precisely.

Let \(\alpha \in \mathbb{R} \setminus \mathbb{Q}\) and \(N \in \mathbb{N}\), and for any real number \(x\), denote the fractional part
as \(\{x\}\). We order the numbers \(\{m\alpha\}\), where \(0 \leq m < N\), into the sequence

\[0 = y_0(N) < y_1(N) < \cdots < y_{N-1}(N) < 1.\]

We then consider the differences between consecutive numbers in the sequence, called
gaps (or spacings),

\[\delta_j(N) = y_{j+1}(N) - y_j(N),\]

for \(j = 0, \ldots, N-2\) and \(\delta_{N-1}(N) = 1 - y_{N-1}(N)\). Now, let \(D(N)\) be the number of
distinct gaps and let \(\Delta_j(N)\) be the ordered sequence of distinct gaps from the \(\delta_j(N)\),
so that

\[0 < \Delta_1(N) < \cdots < \Delta_{D(N)}(N) < 1.\]

**Figure 2.** Illustration of the definitions of \(y_j(N), \delta_j(N),\) and \(\Delta_k(N)\) in the case of \(\alpha = \sqrt{2}\) and \(N = 27\) (these
are the same values as in Figure 1). Only the left part of the interval is shown, and the dependence on \(N\)
is dropped from the notation.

**Three-Gap Theorem** (Sós, Surányi, and Świerczkowski). For all choices of \(\alpha\) and \(N\),
we have \(D(N) \leq 3\).

See \([12–14]\) for the original references and their proofs. Many other proofs and
interpretations of the Three-Gap Theorem have been given since; see, e.g., \([9–11]\). Also, many higher dimensional versions of the three-gap theorem have recently been
studied \([3, 4, 6, 7]\).

The Three-Gap Theorem is intimately related to the topic of circle rotations from
dynamical systems. Let us highlight two of the most classical theorems in the subject:
Kronecker’s Theorem \([8]\) and Weyl’s Theorem \([15]\). For a pleasant discussion of these
important theorems, we encourage the reader to see Section 6.3 of \([5]\). The Three-Gap
Theorem is also connected to the quantum harmonic oscillator \([1, 2]\).

**Words in the Gap Lengths.** As one goes counterclockwise around the circle, we
will describe the order with which the gap sizes occur with a word \(W\) in the letters
\(a, b,\) and \(c\), or just in the letters \(a\) and \(b\) if \(D(N) = 2\). More specifically, we define
the \(j\)th letter \(W_j\), of the word \(W\), to be \(a, b\) or \(c\) corresponding to the gap \(\delta_j(N)\), with \(a\)
corresponding to the smallest gap, \( \hat{b} \) the medium-sized gap, and \( c \) the largest gap. We interpret the word cyclically so that \( W_j = W_j \mod N \). When it is necessary to indicate the dependence on \( N \), we will denote the word \( W \) as \( W(N) \).

**Symmetry Theorem.** Fix any \( \alpha \in \mathbb{R} \setminus \mathbb{Q} \) and \( N \in \mathbb{N} \), and let \( W \) be the word generated by the corresponding gaps on the circle. Then for any \( c \) in \( W \), the \( k \)th letter to the right of it is always the same as the \( k \)th letter to the left of it, so long as the index \( k \) is smaller than the index of the first \( c \) occurrence on either side.

More precisely, if \( W_j = c \), then \( W_{j-k} = W_{j+k} \) for \( k = 0, \ldots, \ell \) where \( \ell + 1 \) is the smallest index such that \( W_{j-(\ell+1)} = c \) or \( W_{j+(\ell+1)} = c \).

![Figure 3](image-url) Illustration of the Symmetry Theorem in the case of \( \alpha = \sqrt{2} \) and \( N = 27 \) (the same values as in Figures 1 and 2). The word in \( \alpha, \hat{b}, \) and \( c \) is symmetric about either \( \alpha \). In the case of the leftmost one, the part of the word that is symmetric is shown in blue.

The symmetry becomes more impressive for larger values of \( N \). For example, the following is the word generated when \( \alpha = \sqrt{2} \) and \( N = 67 \), and the “limiting” \( c \) happens to occur on both sides:

\[
aababaababacababaababababaababaababaababaababaababaababaababaababaababaababaababa
\]

2. **PROOF OF THE SYMMETRY THEOREM.** Our proof closely follows the ideas and notation of van Ravenstein [11]. Let \( (u_0, \ldots, u_{N-1}) \) be the permutation of \((0, 1, 2, \ldots, N - 1)\) such that \( \{u_j \alpha\} = y_j(N) \). Note that we are deviating from the conventions of [11] where Ravenstein uses \((u_1, \ldots, u_N)\) instead. We interpret the \( u_j \) cyclically so that \( u_j = u_j \mod N \).

We remark that at certain choices of \( N \) there will be only two sizes of gap corresponding to the symbols \( \alpha \) and \( \hat{b} \). When one more point is added, it will result in a gap of a new (smaller) size, say \( \hat{d} \), and the labeling will have to be updated. For instance, \( \hat{d} \) becomes \( \alpha \), \( \alpha \) becomes \( \hat{b} \), and \( \hat{b} \) becomes \( c \). Therefore, we will call such “times” \( N \) where there are only two sizes of gap, the relabeling times. We denote the ordered sequence of relabeling times as \((R_k)_{k \geq 1}\).

We will rely on three basic facts from van Ravenstein [11]:

**Fact 1:** \( N \) is a relabeling time if and only if \( N = u_1 + u_{N-1} \). Indeed, in this case, adding the point \( \{N \alpha\} \) results in a point closer to 0 than either \( \{u_1 \alpha\} \) or \( \{u_{N-1} \alpha\} \), and thus a gap of a new size appears. Note that \( u_1 \) and \( u_{N-1} \) depend upon \( N \), and this is a somewhat subtle condition which can be expressed nicely in terms of the continued fraction of \( \alpha \), but it is not needed in our paper.

**Fact 2:** At a relabeling time \( N \), we have:

\[
u_j = j \cdot u_1 \mod N \quad j = 0, 1, \ldots, N - 1. \tag{1}\]

**Fact 3:** If \( N \) is not a relabeling time, then adding the point \( \{N \alpha\} \) results in splitting a gap labeled by \( c \) into a gap labeled by \( \alpha \) and a gap labeled by \( \hat{b} \) in either possible order. This is illustrated in Figure 4.

Facts 1 and 2 are found in [11][Lemma 2.1] and Fact 3 is found in [11][Theorem 2.2].

**Proposition 1.** Let \( W \) be the word of \( \alpha \)'s and \( \hat{b} \)'s at a relabeling time. Then the word satisfies the following symmetry. Let \( J \) be such that \( u_j = N - 1 \). Then we have \( W_{j-1}W_j = “ab” \) or \( “ba” \) and \( W_{j-1-k} = W_{j+k} \) for \( k = 1, \ldots, N - 2 \).
Figure 4. Illustration of Fact 3. Top: Removal of the point \(\{26\alpha\}\) from the circle shown in Figure 1 leads to a new \(c\) between \(\{9\alpha\}\) and \(\{14\alpha\}\). The addition of this new \(c\) did not shorten the length (relative to the length in Figure 1) of the symmetry in blue. Bottom: Removal of the point \(\{25\alpha\}\) from the circle shown in Top leads to a new \(c\) between \(\{13\alpha\}\) and \(\{8\alpha\}\). Here however, the addition of this new \(c\) did shorten the length of the symmetry in blue. On the other hand, the addition of both \(c\)'s leads to further symmetries that weren’t required previously, though these are not indicated by color in the figure.

**Proof.** Note that \(u_J \equiv (u_{J-1} + u_1) \mod N\) because of (Fact 2, equation (1)), and the choice of \(J\) gives \(u_J = u_{J-1} + u_1\), and similarly \(u_{J+1} = u_J + (u_1 - N)\). This implies \(W_{J-1} \neq W_J\).

We will now inductively prove for \(k = 1, \ldots, N-2\) that \(u_{J-k} + u_{J+k} = N - 2\). For \(k = 1\), this statement immediately follows from the formula in the previous paragraph. Now, assume the equality holds for some \(1 \leq k < N - 2\). Then, \(u_{J-k} - u_1 \geq 0\) if and only if \(u_{J+k} + u_1 \leq N - 2\). Therefore, \(u_{J-(k+1)} = u_{J-k} - u_1\) if and only if \(u_{J+(k+1)} = u_{J+k} + u_1\). (Note that it is impossible to have either \(u_{J-(k+1)} = N - 1\) or \(u_{J+(k+1)} = N - 1\) since \(k < N - 1\).) If both sides of the if-and-only-if are false, we have \(u_{J-(k+1)} = u_{J-k} - (u_1 - N)\) and \(u_{J+(k+1)} = u_{J+k} + (u_1 - N)\). In either case, the
sum is still preserved. Now, we remark that at each step of the induction,

\[ u_{J-k} - u_{J-(k+1)} = u_{J+(k+1)} - u_{J+k}, \]

thus the gap sizes are the same and hence \( W_{J-1-k} = W_{J+k} \).

Now we describe the symmetry about other gaps in the word at relabeling times.

**Proposition 2.** Let \( W \) be the word of \( a \)'s and \( b \)'s at some relabeling time \( N = R_q \). Then the word satisfies the following symmetry. Let \( J \) be such that \( u_J = N - p \) for some \( 1 \leq p \leq R_q - R_{q-1} \). Then we have \( W_{J-1}W_J = \text{“}ab\text{”} \) or \( \text{“}ba\text{”} \) and \( W_{J-1-k} = W_{J+k} \) for \( k = 1, \ldots, \ell \), where \( \ell + 1 \) is the smallest index such that \( \max\{u_{J-(\ell+1)}, u_{J+(\ell+1)}\} \geq u_J \).

**Proof.** By Facts 1 and 3, since \( 1 \leq p \leq R_q - R_{q-1} \) and since removing the point \( \{(N - p)\alpha\} \) corresponds to combining an \( \text{“}ab\text{”} \) or \( \text{“}ba\text{”} \) into a \( c \), we have that \( W_{J-1} \neq W_J \).

We now prove inductively that, for \( k = 1, \ldots, \ell \),

\[ u_{J-k} + u_{J+k} = N - 2p. \tag{2} \]

For \( k = 1 \), this statement immediately follows from the fact that \( W_{J-1} \neq W_J \) as in the proof of Proposition 1. Now, assume the equality holds for some \( 1 \leq k < \ell \).

We claim that the following four statements are equivalent:

1. \( u_{J-(k+1)} = u_{J-k} - u_1 \),
2. \( u_{J+(k+1)} = u_{J+k} + u_1 \),
3. \( u_{J-k} - u_1 \geq 0 \), and
4. \( u_{J+k} + u_1 \leq N - 2p \).

First, note that (i) is equivalent to (iii) by Fact 2 (equation (1)). Moreover, (iii) is equivalent to (iv) by the induction hypothesis (2). Finally, we show (ii) is equivalent to (iv). For the forward direction, note that \( u_{J+(k+1)} \geq N - p \) is impossible due to the choice of \( \ell \). Now suppose that \( N - 2p < u_{J+k} + u_1 < N - p \). By the induction hypothesis (2) we have \( u_{J-k} - u_1 = N - 2p - (u_{J+k} + u_1) \), hence

\[ N - 2p - (N - p) < N - 2p - (u_{J+k} + u_1) < N - 2p - (N - 2p), \]

or equivalently, \(-p < u_{J-k} - u_1 < 0\). This means that \( u_{J-(k+1)} > N - p \), which is again impossible by the choice of \( \ell \). Meanwhile, the reverse direction follows immediately from Fact 2 (equation (1)).

Therefore, \( u_{J-(k+1)} = u_{J-k} - u_1 \) if and only if \( u_{J+(k+1)} = u_{J+k} + u_1 \), and hence (2) holds when \( k \) is replaced by \( k + 1 \). Now the proof follows exactly as in the previous proposition.

Even though Proposition 1 is a special case of Proposition 2, we have included Proposition 1 to make the exposition clearer.

**Proof of the Symmetry Theorem.** Let \( (R_j)_{j \geq 1} \) be the ordered increasing sequence of relabeling times. It is clear that the word of length \( R_j \) satisfies the theorem: there are no \( c \)'s to center the symmetry around. Now, we remark that in moving from word \( W(R_j) \) to \( W(R_j - 1) \), the \( ab \) or \( ba \) centered at \( \{(R_j - 1)\alpha\} \) turns into a \( c \), and the symmetry centered at this \( c \) must span the entire word, which it indeed does by Proposition 1.
Now, consider $W(R_j - k)$ where $1 < k < R_j - R_{j-1}$. This word is obtained from $W(R_j)$ by removing $\{(N - i)\alpha\}$ for $1 \leq i \leq k$. As each point is removed, either an $ab$ or $ba$ turns into a $c$, and we must prove the asserted symmetry about each $c$. However, this corresponds directly to the symmetry proved in Proposition 2. Note that the condition $\max\{u_{j-(\ell+1)}, u_{j+(\ell+1)}\} \geq u_j$ corresponds to stopping the symmetry at the closest occurring $c$ to the left or right of the given one.

ACKNOWLEDGMENTS. We thank the referees for their careful reading of the paper and helpful suggestions for how to improve the writing. The second author thanks Pavel Bleher for introducing him to this subject and for many interesting conversations about it. We also thank Valérie Berthé, Ethan Coven, Alan Haynes, and Ronnie Pavlov for their helpful comments. This work was supported by NSF grant DMS-1348589.

ORCID
Roland Roeder http://orcid.org/0000-0002-4623-2590

REFERENCES
[1] Bleher, P. M. (1990). The energy level spacing for two harmonic oscillators with golden mean ratio of frequencies. J. Stat. Phys. 61(3–4): 869–876. doi.org/10.1007/BF01027305
[2] Bleher, P. M. (1991). The energy level spacing for two harmonic oscillators with generic ratio of frequencies. J. Stat. Phys. 63(1–2): 261–283. doi.org/10.1007/BF01026603
[3] Bleher, P. M., Homma Y., Ji L. L., Roeder. R. K. W., Shen, J. (2012). Nearest neighbor distances on a circle: Multidimensional case. J. Stat. Phys. 146(2): 446–465. doi.org/10.1007/s10955-011-0367-8
[4] Dyson, F. J. (1992). Nearest neighbor distances on a circle. IAS Preprint IASSNS-HEP/92/27.
[5] Jacobs, K. (1992). Invitation to Mathematics. Princeton, NJ: Princeton University Press.
[6] Haynes, A., Marklof, J. (2020). Higher dimensional Steinhaus and Slater problems via homogeneous dynamics. Ann. Sci. Éc. Norm. Supér. 53(2): 537–557. doi.org/10.24033/asens.2427
[7] Haynes, A., Marklof, J. (2021). A five distance theorem for Kronecker sequences. Int. Math. Res. Not. doi.org/10.1093/imrn/rnab205
[8] Kronecker, L. (1884). Die Periodensysteme von Funktionen Reeller Variablen. Berliner Sitzungsberichte, pp. 1071–1080.
[9] Liang, F. M. (1979). A short proof of the 3d distance theorem. Discrete Math. 28(3): 325–326. doi.org/10.1016/0012-365X(79)90140-7
[10] Marklof, J., Strömbergsson, A. (2017). The three gap theorem and the space of lattices. Amer. Math. Monthly. 124(8): 741–745. doi.org/10.4169/amer.math.monthly.124.8.741
[11] van Ravenstein, T. (1988). The three gap theorem (Steinhaus Conjecture). J. Aust. Math. Soc. Ser. A. 45(3): 360–370. doi.org/10.1017/S1446788700031062
[12] Sós, V. T. (1957). On the theory of Diophantine approximations I. Acta Math. Acad. Sci. Hungar. 8: 461–472. doi.org/10.1007/BF02020329
[13] Surányi, J. (1958). Über die Anordnung der Vielfachen einer reellen Zahl mod 1. Ann. Univ. Sci. Budapest Eötvös Sect. Math. 1: 107–111.
[14] Świerczkowski, S. (1959). On successive settings of an arc on the circumference of a circle. Fund. Math. 46(2): 187–189. doi.org/10.4064/fm-46-2-187-189
[15] Weyl, H. (1916). Über die Gleichverteilung von Zahlen mod Eins. Eins. Math. Ann. 77: 313–352. doi.org/10.1007/BF01475864

ANEESH DASGUPTA is a recent high school graduate who intends to study mathematics at Harvard University. His interests include dynamical systems and algebraic number theory. He enjoys playing the tuba in his spare time.
Carmel High School, Carmel IN 46032
aneeshdg15@gmail.com

ROLAND ROEDER has been a faculty member in the Department of Mathematical Sciences at Indiana University–Purdue University Indianapolis since 2009 and is currently a Professor. His major interests are in dynamical systems, mathematical physics, and hyperbolic geometry. He enjoys leading students (high school through Ph.D. level) on research projects.
Department of Mathematical Sciences, Indiana University–Purdue University Indianapolis, Indianapolis IN 46202
roederr@iupui.edu