Journal of Integer Sequences, Vol. 24 (2021),
Article 21.10.8

The Ulam Sequence of Linear Integer Polynomials

Arseniy (Senia) Sheydvasser
Israel Institute of Technology
Department of Mathematics
Haifa 3200003
Israel
sheydvasser@campus.technion.ac.il

Abstract

The Ulam sequence $U(1,n)$ is defined as the sequence starting with integers $1,n$ such that $n>1$, and such that every subsequent term is the smallest integer that can be written as the sum of distinct previous terms in exactly one way. This family of sequences is notable for being the subject of several remarkable rigidity conjectures. We introduce an analogous notion of an Ulam sequence inside the integer polynomial ring, and use it both to give new, constructive proofs of old results as well as to produce a new conjecture that implies many of the other existing conjectures.

1 Introduction and main results

Given integers $1 \leq a < b$, define the *Ulam sequence* $U(a,b)$ to be the sequence starting with a,b, and such that every subsequent term is the smallest integer that can be written as the sum of two distinct prior terms in exactly one way. The sequence

$$U(1,2) = 1,2,3,4,6,8,11,13,16,18,26,28,36,38\ldots,$$

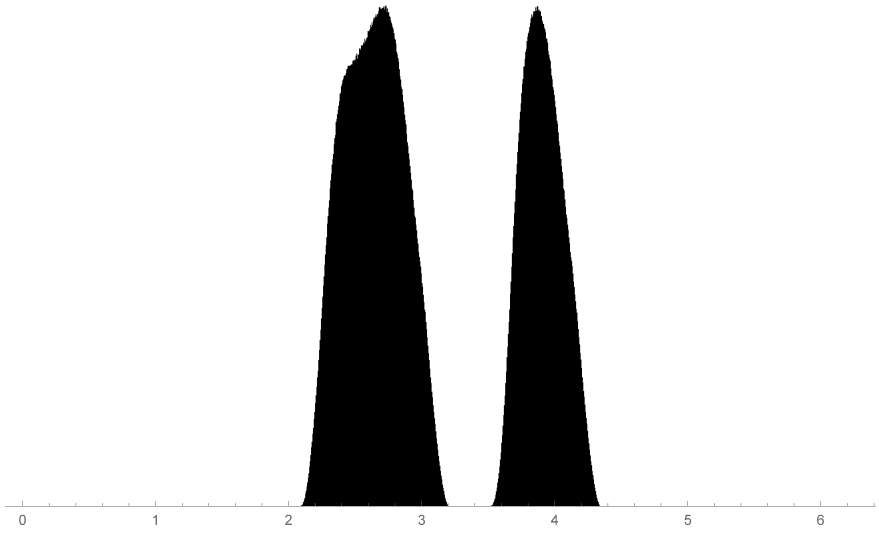

Figure 1: A histogram of the first billion terms of $U(1,2)$ mod λ_2, re-scaled into the interval $[0, 2\pi]$.

which is A002858 in the *On-Line Encyclopedia of Integer Sequences* (OEIS), was originally introduced in 1964 by Ulam [13], who posed the question of determining the growth rate of this sequence, which remains opens to this day. It is conjectured that $U(1,2)$ grows linearly, and it has positive density of about 0.079. The growth rate of certain other families of Ulam sequences was confirmed to be linear by proving the stronger result that they are eventually periodic—this was done for Ulam sequences $U(2, 2n+1)$ by Schmerl and Spiegel [11], using previous work of Finch [2, 4, 3] and Queneau [10]. Similarly, Cassaigne and Finch [1] proved that sequences $U(4,n)$ are eventually periodic if $n \equiv 1 \pmod 4$, and Hinman, Kuca, Schlesinger, and Sheydvasser [7] found a finite set of sequences $U(4,n)$ with $n \equiv 3 \pmod 4$ that are also eventually periodic. In contrast, none of the sequences of the form $U(1,n)$ seem to be eventually periodic, and virtually nothing was known about them until very recently, when Steinerberger [12] gave numerical evidence that there exists a real number $\lambda_2 \approx 2.44344296778$ with the curious property that $U(1,2)$ mod λ_2 is concentrated in the middle third of the interval. To be more precise, we have the following conjecture, formulated by Gibbs [5].

Conjecture 1. There exists a real number $\lambda_2 \approx 2.44344296778$ such that for all $\epsilon > 0$, for K sufficiently large,

$$U(1,2) \cap [K, \infty) \bmod \lambda_2 \subset \left(\frac{\lambda_2}{3} - \epsilon, \frac{2\lambda_2}{3} + \epsilon \right).$$

This conjecture has since been confirmed for the first trillion terms of $U(1,2)$ by Gibbs and McCranie [6]. Numerical evidence suggests that for Ulam sequences $U(a,b)$ that are not

eventually periodic, similar behavior occurs—such "magic numbers" for Ulam sequences are referred to as *periods* in the literature. In particular, there is the following generalization of Gibbs' conjecture in the mathematical folklore, communicated to the author by Joshua Hinman, who did an extensive numerical study of the periods of various families of Ulam sequences.

Conjecture 2. For all $n \geq 2$, there exist real numbers λ_n, K_n such that for all $\epsilon > 0$,

$$U(1,n) \cap [K_n, \infty) \bmod \lambda_n \subset \left(\frac{\lambda_n}{3} - \epsilon, \frac{2\lambda_n}{3} + \epsilon \right).$$

Furthermore, for $n \geq 4$, we can take $\lambda_n = 3n + \lambda'$, where $\lambda' \approx 0.417031$.

The observed empirical fact that for $n \geq 4$, the periods λ_n grow linearly has been poorly understood up until now. Our goal is to show that this curious phenomenon is deeply tied to the following, seemingly unconnected, numerical observation of Hinman, Kuca, Schlesinger, and the present author [8]. Specifically, we noted that for $n \geq 4$ runs of consecutive elements of $U(1,n)$ group into blocks whose endpoints grow linearly.

$$\begin{array}{rlllllll}
U(1,4) = & \boxed{1}, & \boxed{4,5,6,7,8}, & \boxed{10}, & \boxed{16}, & \boxed{18,19}, & \boxed{21}\ldots \\
U(1,5) = & \boxed{1}, & \boxed{5,6,7,8,9,10}, & \boxed{12}, & \boxed{20}, & \boxed{22,23,24}, & \boxed{26}\ldots \\
U(1,6) = & \boxed{1}, & \boxed{6,7,8,9,10,11,12}, & \boxed{14}, & \boxed{24}, & \boxed{26,27,28,29}, & \boxed{31}\ldots \\
U(1,n) = & \boxed{1}, & \boxed{n, n+1, \ldots 2n}, & \boxed{2n+2}, & \boxed{4n}, & \boxed{4n+2 \ldots 5n-1}, & \boxed{5n+2}\ldots
\end{array}$$

This observation was made precise by the following conjecture.

Conjecture 3. There exist unique integer coefficients a_i, b_i, c_i, d_i such that for all $n \geq 4$,

$$U(1,n) = \bigcup_{i=0}^{\infty} [a_i n + b_i, c_i n + d_i],$$

such that $a_{i+1}n + b_{i+1} > c_i n + d_i n + 1$ for all i.

At present, this conjecture is still wide open, but there is a somewhat weaker result.

Theorem 4 ([8, Theorem 1.1])**.** *There exist integer coefficients a_i, b_i, c_i, d_i such that for every $C > 0$, there exists a positive integer N such that for all integers $n \geq N$,*

$$U(1,n) \cap [1, Cn] = \left(\bigcup_{i=0}^{\infty} [a_i n + b_i, c_i n + d_i] \right) \cap [1, Cn].$$

The original proof of this theorem was nonconstructive and in fact model-theoretic in nature, although we showed [8] that assuming Theorem 4, one can prove that there exists an algorithm that will both find the coefficients a_i, b_i, c_i, d_i and the minimal integer N_0 that will

$$
\begin{aligned}
U(1,X) = \{1\} &\cup [X, 2X] && \cup \{2X+2\} && \cup \{4X\} \\
&\cup [4X+2, 5X-1] && \cup \{5X+1\} && \cup [7X+3, 8X+1] \\
&\cup \{10X+2\} && \cup \{11X+2\} && \cup [13X+4, 14X+1] \\
&\cup \{16X+2\} && \cup \{17X+2\} && \cup \{19X+3\} \\
&\cup \{20X+2\} && \cup \{22X+3\} && \cup \{23X+4\} \\
&\cup [25X+4, 25X+5] && \cup \{26X+3\} && \cup \{28X+4\} \\
&\cup [31X+5, 32X+3] && \cup \{34X+5\} && \cup \{38X+6\} \\
&\cup \{40X+5\} && \cup [40X+8, 41X+4] && \cup [43X+7, 44X+4] \\
&\cup \{44X+6\} && \cup \{46X+7\} && \cup [49X+8, 50X+6] \\
&\cup [52X+8, 53X+7] && \cup [55X+9, 56X+6] \dots
\end{aligned}
$$

Table 1: The first 30 intervals of $U(1,X)$—equivalently, replacing X with n, the first $56n+6$ terms of $U(1,n)$ when $n \geq 4$.

satisfy the conditions of Theorem 4, given $C > 0$. Our present goal is to give an alternate, constructive proof of Theorem 4 by considering a variant of the Ulam sequence that sits inside the polynomial ring $\mathbb{Z}[X]$. This ring can be given the structure of an ordered ring via the lexicographical ordering—that is, $p(X) > q(X)$ if and only if the leading term of $p(X) - q(X)$ has a positive coefficient. In Section 2, we define a set

$$U(1,X) = \bigcup_{i=0}^{\infty} [a_i X + b_i, c_i X + d_i] \subset \mathbb{Z}[X]$$

which should be viewed as an analog of an Ulam sequence inside this ordered ring; specifically, in the additive subgroup of linear polynomials. Here $[x, y]$ has the usual meaning that it is the set of all elements $z \in \mathbb{Z}[X]$ such that $x \leq z \leq y$. We call the sequences $(a_i)_{i\in\mathbb{N}}$, $(b_i)_{i\in\mathbb{N}}$, $(c_i)_{i\in\mathbb{N}}$, and $(d_i)_{i\in\mathbb{N}}$ the *coefficients* of $U(1,X)$—as we shall see, they are uniquely determined and, better yet, there is an algorithm to compute them.

Theorem 5. *There exists a polynomial-time algorithm $\mathcal{A}_{Ulam}$ such that on an input of $k \in \mathbb{N}$, it returns the first k coefficients of $U(1,X)$—that is, $(a_i)_{i=0}^k$, $(b_i)_{i=0}^k$, $(c_i)_{i=0}^k$, and $(d_i)_{i=0}^k$—and an integer N such that for all $n \geq N$,*

$$U(1,n) \cap [1, c_k n + d_k + 1] = \bigcup_{i=0}^{k} [a_i n + b_i, c_i n + d_i].$$

Theorem 4 is an immediate corollary of Theorem 5. We construct an explicit example of such an algorithm in Section 5 and show that there should be many other such algorithms, including ones that are likely much more efficient. Studying the output of such algorithms

raises the possibility of proving that

$$U(1,n) = \bigcup_{i=0}^{\infty} [a_i n + b_i, c_i n + d_i]$$

for all $n \geq N$, for some natural number N. To the best of the author's knowledge, this is the first proposed method of attacking Conjecture 3. However, this is not the only benefit of introducing the set $U(1,X)$—it also makes it convenient to state a conjecture for which the author found ample numerical evidence, which is presented in Section 3.

Conjecture 6. Let a_i, b_i, c_i, d_i be the coefficients of $U(1,X)$. There exist real numbers $\lambda' \approx 0.417031, \sigma_1 \approx 1.86, \sigma_2 \approx -1.3$ such that for every $\epsilon > 0$, if i is sufficiently large, then

$$a_i X + b_i, c_i X + d_i \bmod 3X + \lambda' \in \left(X + \frac{\lambda'}{3} - \epsilon, X + \sigma_1 + \epsilon\right) \cup \left(2X + \sigma_2 - \epsilon, 2X + \frac{2\lambda'}{3} + \epsilon\right).$$

The precise definition of taking a modulus in $\mathbb{Z}[X]$ is given in Section 3, but the intuitive idea is as follows: the residues of the endpoints of the intervals defining $U(1,X)$ are

1. constrained (to ϵ-neighborhoods of) the middle third of the interval $[0, 3X + \lambda']$—that is, between $X + \lambda'/3$ and $2X + 2\lambda'/3$, and
2. there also seems to be a gap inside that middle third interval, namely $(X+\sigma_1, 2X+\sigma_2)$, where these residues also do not appear.

This conjecture should be seen as an analog of Conjecture 2 for the ordered ring $\mathbb{Z}[X]$. In Section 4, we also demonstrate that Conjecture 6 has a number of remarkable consequences—for example, it implies that b_i grows linearly with respect to a_i, and similarly d_i grows linearly with respect to c_i; furthermore, we prove that Conjectures 6 and 3 together imply Conjecture 2, for $n \geq 4$.

2 An Ulam-like sequence

We begin by defining what we mean by Ulam sequences in $\langle 1, X\rangle$, the abelian group generated by $1, X$ in $\mathbb{Z}[X]$; equivalently, this is the set of linear integer polynomials. We always take this to be an ordered group via the lexicographical ordering—i.e., $aX + b > cX + d$ if and only if $a > c$ or $a = c$ and $b > d$.

Definition 7. Given $0 < a < b$ in $\langle 1, X\rangle$, an *Ulam sequence* starting with a, b is a set $U \subset \langle 1, X\rangle$ such that

1. $U \cap (-\infty, b] = \{a, b\}$,

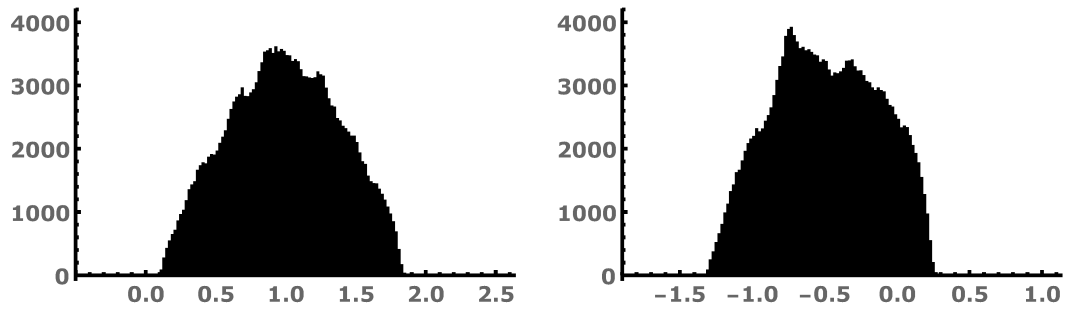

Figure 2: A histogram of $U(1, X)$ mod $\lambda(X)$ for the first 200000 intervals. The plot on the left represents elements in $(X + \frac{\lambda'}{3} - \epsilon, X + \sigma_1 + \epsilon)$. The plot on the right represents elements in $(2X + \sigma_2 - \epsilon, 2X + \frac{2\lambda'}{3} + \epsilon)$.

2. for all $p < q \in \langle 1, X\rangle$, $U \cap [p, q]$ has both a minimum and a maximum, and
3. for every $p \in (b, \infty)$, $p \in U$ if and only if it is the smallest element in the set

$$\{q \in \langle 1, X\rangle | q > U \cap (-\infty, p) \text{ and } \exists! x \neq y \in U,\ q = x + y\}.$$

Remark 8. The idea of extending Ulam sequences to other ordered abelian groups than the integers is not new—Kravitz and Steinerberger [9] previously showed that one can define the notion of an Ulam subset of $\mathbb{R}^n$.

Why do we say that it is *an* Ulam sequence, rather than *the* Ulam sequence? This is because, in general, it is not unique! For example, it follows from earlier model-theoretic results [8] that there exist Ulam sequences U, U' starting with $2, X$ such that

$$U \cap [2, 2X + 1] = \{2\} \cup \{X + 2k | k \in \mathbb{N}\} \cup \{2X - 2k | k \in \mathbb{N}\}$$
$$U' \cap [2, 2X + 1] = \{2\} \cup \{X + 2k | k \in \mathbb{N}\} \cup \{2X - 2k + 1 | k \in \mathbb{N}\}.$$

Nevertheless, there are some specific cases where such an Ulam sequence is unique—in particular, we show that there is a unique Ulam sequence starting with $1, X$, which we denote by $U(1, X)$. This set is very special in that it is a disjoint union of intervals with an important finiteness property.

Definition 9. We say a set $S \subset \langle 1, X\rangle$ is a *union of intervals with finite intersections* if it is a disjoint union of intervals $[a, b] \subset \langle 1, X\rangle$ such that for every $a, b \in \langle 1, X\rangle$, there are only finitely many of these intervals between a and b.

Remark 10. The core idea is that while S might contain infinitely many intervals, every intersection with an interval $[a, b]$ can only contain finitely many.

X,$X+1, 2X+1, 3X+1, 3X+2, 4X+1, 4X+3, 5X+1, 5X+4, 6X+1,$
$6X+3, 6X+5, 7X+1, 7X+6, 8X+1, 8X+3, 8X+5, 8X+7, 9X+1,$
$9X+8, 10X+1, 10X+3, 10X+5, 10X+7, 10X+9, 11X+1, 11X+10,$
$12X+1, 12X+3, 12X+5, 12X+7, 12X+9, 12X+11, 13X+1, 13X+12,$
$14X+1, 14X+3, 14X+5, 14X+7, 14X+9, 14X+11, 14X+13, 15X+1,$
$15X+14, 16X+1, 16X+3, 16X+5, 16X+7, 16X+9, 16X+11, 16X+13,$
$16X+15, 17X+1, 17X+16, 18X+1, 18X+3, 18X+5, 18X+7, 18X+9,$
$18X+11, 18X+13, 18X+15, 18X+17, 19X+1, 19X+18, 20X+1, 20X+3,$
$20X+5, 20X+7, 20X+9, 20X+11, 20X+13, 20X+15, 20X+17, 20X+19,$
$21X+1, 21X+20, 22X+1, 22X+3, 22X+5, 22X+7, 22X+9, 22X+11,$
$22X+13, 22X+15, 22X+17, 22X+19, 22X+21, 23X+1, 23X+22,$
$24X+1, 24X+3, 24X+5, 24X+7, 24X+9, 24X+11, 24X+13,$
$24X+15, 24X+17, 24X+19, \ldots$

Table 2: The first 100 terms of the unique Ulam sequence starting with $X, X+1$.

It is easy to see that for every union of intervals with finite intersections S, there exist integer sequences a_i, b_i, c_i, d_i and $N \in \mathbb{Z}_{>0} \cup \{\infty\}$ such that

$$S = \bigcup_{i=0}^{N} [a_i X + b_i, c_i X + d_i]$$

and $c_i X + d_i + 1 < a_{i+1} + b_{i+1}$ for all $i < N$. This decomposition is clearly unique, which makes the following definition sensible.

Definition 11. Given a union of intervals with finite intersections S, we call the decomposition

$$S = \bigcup_{i=0}^{N} [a_i X + b_i, c_i X + d_i]$$

where a_i, b_i, c_i, d_i are integer sequences and $N \in \mathbb{Z}_{>0} \cup \{\infty\}$ the *standard form* of S. We call the sequences a_i, b_i, c_i, d_i the *coefficients* of S.

To prove that there is a unique Ulam sequence starting with $1, X$, we start by showing that the beginning of its standard form is uniquely determined.

Lemma 12. *Let U be an Ulam sequence starting with $1, X$. Then $U \cap [1, 2X+1] = \{1\} \cup [X, 2X]$.*

Proof. Obviously, U contains both 1 and X, and hence it contains $X+1$. By induction, it must also contain $X+k$ for all $k \in \mathbb{N}$. It follows that it must contain elements $2X-k$ with k arbitrarily large, since otherwise $U \cap [1, 2X-k]$ would fail to have a maximum for some $k \in \mathbb{N}$. But $2X-k$ must be the sum of two prior elements $u < v$, and if $u, v \geq X$, then it would have to be at least $2X+1 = X+(X+1)$. Therefore, $u = 1$ and $v = 2X-k-1$, which is to say that $[X, 2X] \subset U$. Finally, since $2X+1$ can be written as the sum of distinct elements in U in two different ways, $2X+1 \notin U$. □

Theorem 13. *There exists a unique Ulam sequence $U(1, X)$ starting with $1, X$. Furthermore, $U(1, X)$ is a union of intervals with finite intersections.*

Proof. We prove by induction that for all $k \in \mathbb{N}$, there exists $l_k \in \mathbb{Z}$ and a unique set $U \subset [1, kX + l_k]$ such that

1. $U \cap [1, X] = \{1, X\}$,
2. U is a finite union of intervals, and
3. for every $u \in [1, kX + l_k]$, $u \in U$ if and only if there exists a unique pair $p < q \in U$ such that $u = p + q$.

We already know that this is true for $k = 0, 1$. So, assume that it is true for $k-1$; we prove it for k. Start with a set $U \subset [1, (k-1)X + l_{k-1}]$ which has the desired properties. By induction, there is a unique extension $U' \subset [1, (k-1)X + l]$ for all $l \in \mathbb{N}$ such that

1. $U' \cap [1, X] = \{1, X\}$ and
2. for every $u \in [1, (k-1)X + l]$, $u \in U'$ if and only if there exists a unique pair $p < q \in U'$ such that $u = p + q$.

Moreover, our claim is that if l is sufficiently large, then either $U' \cap [(k-1)X + l, (k-1)X + l'] = [(k-1)X + l, (k-1)X + l']$ for all larger l', or $U' \cap [(k-1)X + l, (k-1)X + l'] = \emptyset$. This is because if $u \in U'$, then $u + 1 \in U'$ unless there exist $X \leq p < q$ such that $u + 1 = p + q$. But it must be that $p < q < (k-2)X + l$ for some $l \in \mathbb{Z}$ since otherwise $p + q$ would be too big. Therefore, p, q both belong to one of the intervals of U—there are only finitely many of these, hence

$$[1, kX] \setminus \{p + q | 1 < p < q \in U\}$$

is a finite union of intervals. Therefore, if l is sufficiently large, the intersection of this set with $[(k-1)X + l, kX - l]$ is either all of $[(k-1)X + l, kX - l]$ or empty. In the first case, $U' \cap [(k-1)X + l, (k-1)X + l'] = \emptyset$; in the second case, either $U' \cap [(k-1)X + l, (k-1)X + l'] = \emptyset$ or $U' \cap [(k-1)X + l, (k-1)X + l'] = [(k-1)X + l, (k-1)X + l']$, depending on whether $(k-1)X + l \in U'$.

If there exists $l \in \mathbb{N}$ such that $U' \cap [(k-1)X+l, (k-1)X+l'] = [(k-1)X+l, (k-1)X+l']$ for all larger l', then there exists some $l'' \in \mathbb{Z}$ such that

$$([1, kX] \setminus \{p+q | 1 < p < q \in U\}) \cap [(k-1)X + l, kX - l''] = \emptyset.$$

Therefore, if we define $l_k = l''$, $U'' = U \cup [(k-1)X + l, kX + l'']$, then

1. $U'' \cap [1, X] = \{1, X\}$,
2. U'' is a finite union of intervals, and
3. for every $u \in [1, kX + l_k]$, $u \in U''$ if and only if there exists a unique pair $p < q \in U''$ such that $u = p + q$.

Moreover, this is the only possible extension of U' to $[1, kX + l_k]$, since the requirement that $U'' \cap [1, kX + l_k]$ have a maximum requires that there must be some element $kX + l''' \in U''$—but this forces $kX + l''' - 1, kX + l''' - 2, \ldots \in U''$ and $kX + l''' + 1, kX + l''' + 2, \ldots, kX + l_k \in U$.

Otherwise, $U' \cap [(k-1)X + l, kX - l] = \emptyset$ for sufficiently large l. We can choose l large enough such that for all $p \in [(k-1)X + l, kX - l]$, either there exists more than one pair $X \leq u, v < kX - l$ such that $p = u + v$, or there are no such pairs. Then if we define $l_k = -l$, then U' is a subset of $[1, kX + l_k]$ such that

1. $U' \cap [1, X] = \{1, X\}$,
2. U' is a finite union of intervals, and
3. for every $u \in [1, kX + l_k]$, $u \in U'$ if and only if there exists a unique pair $p < q \in U'$ such that $u = p + q$.

Moreover, this is the only possible way to extend U' to $[1, kX + l_k]$—if we were to add any element $kX + l$, it would require that $kX + l - 1, kX + l - 2, \ldots \in U'$, which would contradict the requirement that $U' \cap [(k-1)X + l, kX + l_k]$ has a minimum.

Thus, we have shown that there is a unique $U \subset [1, kX + l_k]$ for all k. But this gives a unique $U \subset \langle 1, X \rangle$ which is just a union of all these sets—it is easy to see that this U is the unique Ulam sequence starting with $1, X$, and that it is a union of intervals with finite intersections. □

3 Numerical results

In light of Theorem 13, we know that $U(1, X)$ is uniquely defined; we now discuss some of the numerical data gathered about this set. In Section 5, we produce an algorithm $\mathcal{A}_{\text{Ulam}}$ such that on an input of $k \in \mathbb{N}$, it returns the first k coefficients of $U(1, X)$ and an integer N such that for all $n \geq N$,

$$\text{eval}_{X \to n} (U(1, X) \cap [1, c_k X + d_k]) = U(1, n) \cap [1, c_k n + d_k],$$

where $\text{eval}_{X\to n} : \mathbb{Z}[X] \to \mathbb{Z}$ is the evaluation map sending X to n. This algorithm has in fact been implemented in Python by the author; unfortunately, in practice $\mathcal{A}_{\text{Ulam}}$ is not very efficient, as we shall discuss.

An easier approach toward computing the coefficients a_i, b_i, c_i, d_i is to assume that Conjecture 3 is true and that Ulam sequences grow linearly, and then to compute $U(1,4)$ and $U(1,5)$ up to a suitably large number of terms, from which one can compute the coefficients a_i, b_i, c_i, d_i. The results of this method can be proven correct by using Theorem 3.1 from a paper of Hinman, Kuca, Schlesinger, and the present author [7]; one need only verify that there exists a $B \approx 0.13901$ such that

1. $|b_i - Ba_i|, |d_i - Bc_i| < 2.5$ for all i, and
2. for $n = 4, 5, \ldots 14$,

$$U(1,n) \cap [1, c_k n + d_k] = \bigcup_{i=0}^{k} [a_i n + b_i, c_i n + d_i].$$

It is immediate from these calculations that the computed coefficients a_i, b_i, c_i, d_i are actually the coefficients of $U(1,X)$ since we know that

$$\text{eval}_{X\to n}(U(1,X) \cap [1, c_k X + d_k + 1]) = U(1,n) \cap [1, c_k n + d_k + 1],$$

for sufficiently large n and that the standard form of $U(1,X)$ is uniquely determined.

In this way, using a basic $\Theta(k^2)$ algorithm for computing Ulam sequences $U(1,n)$, the author was able to compute coefficients a_i, b_i, c_i, d_i such that for all $n \geq 4$,

$$U(1,n) \cap [1, c_{217529} n + d_{217529}] = \bigcup_{i=0}^{217529} [a_i n + b_i, c_i n + d_i].$$

As $(a_{217529}, b_{217529}, c_{217529}, d_{217529}) = (966409, 134342, 966410, 134340)$, this is the full list of coefficients a_i, b_i, c_i, d_i such that $c_i n + d_i \leq 966410n + 134340$. This data gives further numerical evidence for Conjecture 2; specifically, defining $\lambda_n = 3n + 0.417031$, we consider the sets $U(1,n) \bmod \lambda_n$. We find that 99.9% of the terms smaller than $10^7 n$ lie in the interval $[\lambda_n/3, 2\lambda_n/3]$—this is shown in Figures 3 and 4.

With this motivation, it is natural to investigate whether there might be some "magic polynomial" $\lambda(X)$ and a way to define $U(1,X) \bmod \lambda(X)$ such that the resulting distribution has interesting properties. Not only is this possible, but the results are startling. First, we give a couple of definitions.

Definition 14. Given elements $p, q \in \mathbb{R}[X]$, define their *remainder set* to be

$$\mathcal{R}_{p,q} := \{p - sq | s \in \mathbb{Z}[X],\ p - sq \geq 0\}.$$

If $\mathcal{R}_{p,q}$ has a smallest element, we define $p \bmod q = \min \mathcal{R}_{p,q}$.

n	# of exceptions	# elements of $U(1,n)$ computed	Percent of exceptions
4	411	635045	0.0647198%
5	416	814686	0.0510626%
6	423	994573	0.0425308%
7	426	1174402	0.0362738%
8	427	1354386	0.0315272%
9	430	1534323	0.0280254%
10	430	1714404	0.0250816%
11	432	1894463	0.0228033%
12	433	2074581	0.0208717%
13	436	2254856	0.019336%
14	436	2435164	0.0179043%

Figure 3: Computations of the number of elements $u \in U(1,n)$ such that $u \bmod \lambda_n$ is not between $\lambda_n/3$ and $2\lambda_n/3$.

Notice in particular that if we define $\lambda(X) = 3X + .417031$, then $U(1,X) \bmod \lambda(X)$ is a well-defined subset of $\{aX+b | a \in \{0,1,2\},\ b \in \mathbb{Q}\} \subset \mathbb{Q}[X]$.

Definition 15. We say that an interval $[p,q] \subset \mathbb{Q}[X]$ is *long* if $\deg(q-p) > 0$. Otherwise, we say that the interval is *short*.

Remarkably, long intervals in $U(1,X)$ seem to occur almost exactly four times as often as short ones—for the first 217530 intervals, we find that 79.98% are long, and 20.02% are short. These two interval types seem to exhibit slightly different statistical behaviors modulo $\lambda(X)$, so we split into two cases accordingly.

Long Intervals: Let $\mathcal{U}_L$ be the subset of $U(1,X)$ consisting of all long intervals, and let a_i', b_i', c_i', d_i' be the coefficients of $\mathcal{U}_L$. For all i such that $c_i'X + d_i' \leq 966410X + 134340$, we find that $a_i' \equiv 1 \pmod 3$ and $c_i' \equiv 2 \pmod 3$, hence $a_i'X + b_i' \bmod \lambda(X) = X + \sigma_{i,1}$ and $c_i'X + d_i' \bmod \lambda(X) = 2X + \sigma_{i,2}$. Furthermore, we find that $|\sigma_{i,1}|, |\sigma_{i,2}| < 3$ for all i. The statistical distributions of $\sigma_{i,1}$ and $\sigma_{i,2}$ are given in Figure 5.

Short Intervals: Let $\mathcal{U}_S$ be the subset of $U(1,X)$ consisting of all short intervals, and let a_i', b_i', c_i', d_i' be the coefficients of $\mathcal{U}_S$. For all i such that $c_i'X + d_i' \leq 966410X + 134340$, we find that $b_i' = d_i'$ with just a single exception, namely the interval $[25X+4, 25X+5]$. Furthermore, we find that $a_i' \equiv 0 \pmod 3$ if and only if $(a_i', b_i', c_i', d_i') = (0,1,0,1)$, $a_i' \equiv 1 \pmod 3$ is true for $\approx 43.7\%$ of indices i, and $a_i' \equiv 2 \pmod 3$ is true for $\approx 56.3\%$ of indices i. Therefore, $a_i'X + b_i' \bmod \lambda(X) = X + \sigma_{i,3}$ or $a_i'X + b_i' \bmod \lambda(X) = 2X + \sigma_{i,4}$, and we find that $|\sigma_{i,3}|, |\sigma_{i,4}| \leq 2$ for all i. The statistical distributions of $\sigma_{i,3}$ and $\sigma_{i,4}$ are given in Figure 6.

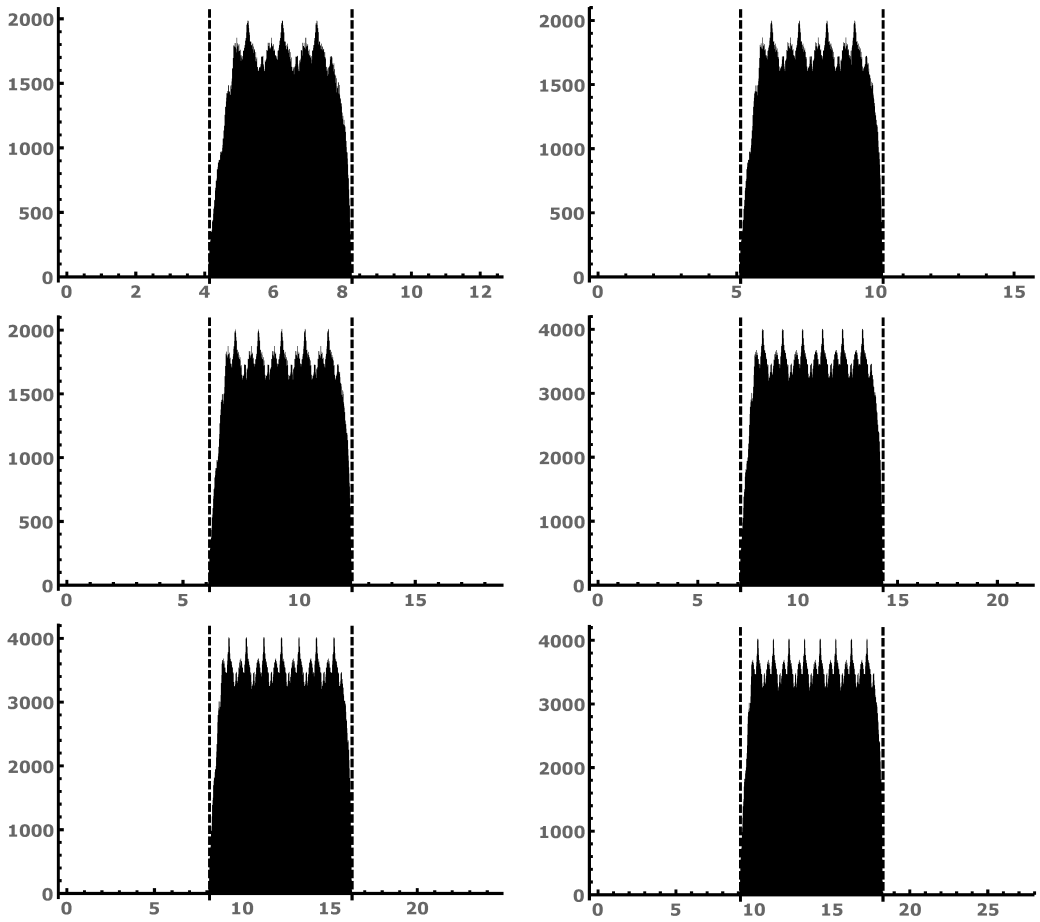

Figure 4: From left to right, and top to bottom, histograms of $U(1,n) \bmod \lambda_n$ for $n = 4,5,6,7,8,9$. The dashed lines split the interval into thirds.

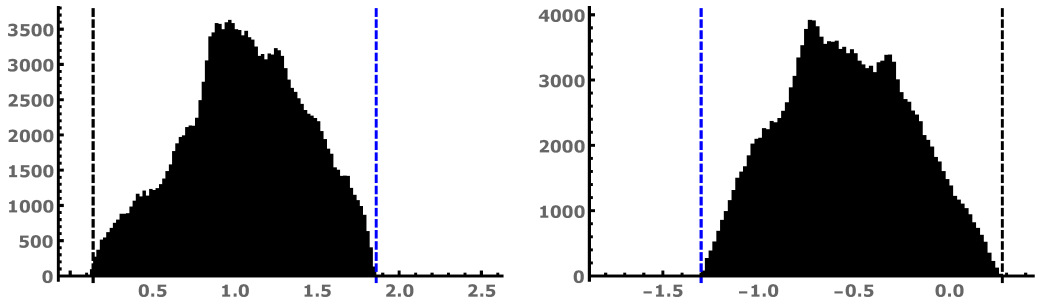

Figure 5: The histograms of $\sigma_{i,1}$ (left) and $\sigma_{i,2}$ (right), together with black lines at $x = \lambda'/3$ (left) and $x = 2\lambda'/3$ (right), and blue lines at $x = 1.86$ (left) and $x = -1.3$ (right).

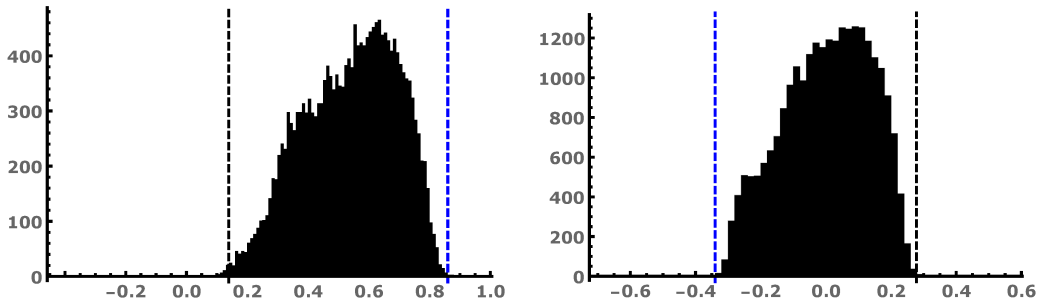

Figure 6: The histograms of $\sigma_{i,3}$ (left) and $\sigma_{i,4}$ (right), together with black lines at $x = \lambda'/3$ (left) and $x = 2\lambda'/3$ (right), and blue lines at $x = 0.86$ (left) and $x = -0.34$ (right).

One final, but important, numerical observation is that the bulk of of the $\sigma_{i,j}$ appear to be bounded—specifically, $\sigma_{i,1}, \sigma_{i,3} \geq \lambda'/3$ and $\sigma_{i,2}, \sigma_{i,4} \leq 2\lambda'/3$ for almost all i. Additionally, $\sigma_{i,1} \leq 1.86$, $\sigma_{i,3} \leq 0.86$, $\sigma_{i,2} \geq -1.3$, and $\sigma_{i,4} \geq -0.34$ for almost all i. Taking all of this information together, we precisely come up with Conjecture 6. Note additionally that $1.86 \approx 2 - \lambda'/3$ and $-1.3 \approx -1 - 2\lambda'/3$; unfortunately, the numerical evidence is insufficient to conjecture this with any strong degree of certainty.

4 Relations between conjectures

The importance of Conjecture 6 is that, if true, then it sheds light on other open questions about Ulam sequences. We give a few examples.

Theorem 16. *If Conjecture 6 is true, then taking $B = \lambda'/3$, for all $\epsilon > 0$, if i is sufficiently large,*

$$|b_i - Ba_i|, |d_i - Bc_i| < \min\left(\sigma_1 - \frac{\lambda'}{3} + \epsilon, \frac{2\lambda'}{3} - \sigma_2 + \epsilon\right).$$

Proof. Note that if $i > 0$, then

$$\begin{aligned} a_iX + b_i \bmod 3X + \lambda' &= a_iX + b_i - (3X + \lambda')\lfloor a_i/3 \rfloor \\ &= (a_i - 3\lfloor a_i/3 \rfloor)\, X + b_i + \frac{\lambda'}{3}\left(a_i - \lfloor a_i/3 \rfloor\right) - \frac{\lambda' a_i}{3} \\ &= \begin{cases} X + b_i + \frac{\lambda'}{3} - \frac{\lambda' a_i}{3}, & \text{if } a_i \equiv 1 \pmod 3; \\ 2X + b_i + \frac{2\lambda'}{3} - \frac{\lambda' a_i}{3}, & \text{if } a_i \equiv 2 \pmod 3. \end{cases} \end{aligned}$$

Therefore, for sufficiently large i, either

$$\begin{aligned} b_i + \frac{\lambda'}{3} - \frac{\lambda' a_i}{3} &\in \left(\frac{\lambda'}{3} - \epsilon, \sigma_1 + \epsilon \right) \\ b_i - \frac{\lambda' a_i}{3} &\in \left(-\epsilon, \sigma_1 - \frac{\lambda'}{3} + \epsilon \right) \\ \left| b_i - \frac{\lambda' a_i}{3} \right| &< \sigma_1 - \frac{\lambda'}{3} + \epsilon \end{aligned}$$

or

$$\begin{aligned} b_i + \frac{2\lambda'}{3} - \frac{\lambda' a_i}{3} &\in \left(\sigma_2 - \epsilon, \frac{2\lambda'}{3} + \epsilon \right) \\ b_i - \frac{\lambda' a_i}{3} &\in \left(\sigma_2 - \frac{2\lambda'}{3} - \epsilon, \epsilon \right) \\ \left| b_i - \frac{\lambda' a_i}{3} \right| &< \frac{2\lambda'}{3} - \sigma_2 + \epsilon. \end{aligned}$$

The argument for $c_i X + d_i$ is identical. □

Theorem 17. *If Conjecture 3 and Conjecture 6 are true, then Conjecture 2 is true for $n \geq 4$.*

Proof. Fix an $n \geq 4$ and define $\lambda_n = 3n + \lambda'$. Choose an $\epsilon > 0$. By Conjecture 6, we get that if i is sufficiently large, then

$$\begin{aligned} a_i n + b_i - \lambda_n \left\lfloor \frac{a_i}{3} \right\rfloor &= \left(a_i - 3 \left\lfloor \frac{a_i}{3} \right\rfloor \right) n + \left(b_i - \lambda' \left\lfloor \frac{a_i}{3} \right\rfloor \right) \\ &\in \left(n + \frac{\lambda'}{3} - \epsilon, n + \sigma_1 + \epsilon \right) \cup \left(2n + \sigma_2 - \epsilon, 2n + \frac{2\lambda'}{3} + \epsilon \right) \\ &\subset \left(\frac{\lambda_n}{3} - \epsilon, \frac{2\lambda_n}{3} + \epsilon \right). \end{aligned}$$

Since this gives a real number between 0 and λ_n, we conclude that for i sufficiently large,

$$a_i n + b_i \bmod \lambda_n \in \left(\frac{\lambda_n}{3} - \epsilon, \frac{2\lambda_n}{3} + \epsilon \right),$$

and similarly for $c_i n + d_i$—thus, by Conjecture 3, it suffices to prove that for all $a_i n + b_i < u < c_i n + d_i$,

$$u \bmod \lambda_n \in \left(\frac{\lambda_n}{3} - \epsilon, \frac{2\lambda_n}{3} + \epsilon \right).$$

If $a_i = b_i$, then $u = a_i n + r$, where $b_i < r < d_i$. Therefore,

$$\begin{aligned} u - \lambda_n \left\lfloor \frac{u}{3} \right\rfloor &= \left(a_i - 3 \left\lfloor \frac{a_i}{3} \right\rfloor \right) n + \left(r - \lambda' \left\lfloor \frac{a_i}{3} \right\rfloor \right) \\ &\in \left(\frac{\lambda_n}{3} - \epsilon, \frac{2\lambda_n}{3} + \epsilon \right), \end{aligned}$$

hence we have the desired conclusion. If $a_i \neq b_i$, then $b_i = a_i + 1$, and therefore $(a_i, b_i) \bmod 3 = (1, 2)$. In that case, we know that

$$a_i n + b_i \bmod \lambda_n \in \left(n + \frac{\lambda'}{3} - \epsilon, n + \sigma_1 + \epsilon \right).$$

Since $u - (a_i n + b_i) < n$ and $2n + \sigma_1 + \epsilon < \lambda_n$ if ϵ is small enough, we conclude that

$$u \bmod \lambda_n = (a_i n + b_i \bmod \lambda_n) + u - a_i n + b_i,$$

whence

$$u \bmod \lambda_n \in \left(\frac{\lambda_n}{3} - \epsilon, \frac{2\lambda_n}{3} + \epsilon \right).$$

□

5 Algorithms

We close by giving the technical details behind Theorem 5 and the promised algorithm $\mathcal{A}_{\text{Ulam}}$. Before we present this, we will need four intermediate subroutines.

Algorithm 5.1. On an input of two intervals $I_1 = [p_1, q_1]$, $I_2 = [p_2, q_2]$ such that $p_2 + 1 < q_1$, this algorithm returns a pair of unions of intervals with finite intersections S_1, S_2 such that S_1 consists of all elements that can be written as the sum of an element of I_1 and an element of I_2 in exactly one way, and S_2 consists of all elements that can be written as the sum of an element of I_1 and an element of I_2 in more than one way.

```
1:  procedure SumDistinct(I1,I2)
2:      start ← p1 + p2
3:      end ← q1 + q2
4:      if #I1 = 1 or #I2 = 1 then
5:          S1 ← [start, end]
6:          S2 ← {}
7:      else
8:          S1 ← {start} ∪ {end}
9:          S2 ← [start + 1, end − 1]
10:     return S1, S2
```

Proof of Correctness. Clearly, $I_1 + I_2 = [\texttt{start}, \texttt{end}]$, so it is solely a question of how this set is partitioned between S_1 and S_2. If $\#I_1 = 1$ or $\#I_2 = 1$, then it is easy to see that S_2 is empty and $S_1 = [\texttt{start}, \texttt{end}]$. Otherwise, $\texttt{start}, \texttt{end} \in S_1$, but for every $x \in [\texttt{start} + 1, \texttt{end} - 1]$ we can write $x = u + v = (u - 1) + (v + 1)$ for some $u, (u - 1) \in I_1$, $v, (v + 1) \in I_2$. □

Algorithm 5.2. On an input of an interval $I = [p, q]$, this algorithm returns a pair of unions of intervals with finite intersections S_1, S_2 such that S_1 consists of all elements that can be written as the sum of two distinct elements of I in exactly one way, and S_2 consists of all elements that can be written as the sum of two distinct elements of I in more than one way.

```
1: procedure SumWithItself(I)
2:     if #I = 1 then
3:         S_1 ← {}
4:         S_2 ← {}
5:     else if #I = 2 then
6:         S_1 ← {p + q}
7:         S_2 ← {}
8:     else if #I = 3 then
9:         S_1 ← [2p + 1, 2p + 3]
10:        S_2 ← {}
11:    else
12:        S_1 ← [2p + 1, 2p + 2] ∪ [2q − 2, 2q − 1]
13:        S_2 ← [2p + 3, 2q − 3]
14:    return S_1, S_2
```

Proof of Correctness. It is clear that the subset of $\langle 1, X \rangle$ representable by pairwise sums of distinct elements of I is $[2p+1, 2q-1]$. It is easy to see that $2p+1, 2p+2, 2q-2, 2q-1 \in S_1$ if they are in this subset, while the remainder must be in S_2. □

Algorithm 5.3. On an input of four unions of intervals with finite intersections S_1, T_1, S_2, T_2 such that $S_1 \cap T_1 = S_2 \cap T_2 = \emptyset$, this algorithm returns the unique pair (S, T) of unions of intervals with finite intersections such that $S \cap T = \emptyset$, $S \cup T = S_1 \cup S_2 \cup T_1 \cup T_2$, and S consists of all elements $x \in (S_1 \backslash T_2) \cup (S_2 \backslash T_1)$ such that $x \notin (S_1 \backslash T_2) \cap (S_2 \backslash T_1)$.

```
1: procedure Amalgam(S_1, T_1, S_2, T_2)
2:     S_3 ← S_1\T_2
3:     S_4 ← S_2\T_1
4:     S ← S_3△S_4                ▷ By △ we mean the symmetric difference
5:     T ← S_1 ∪ S_2 ∪ T_1 ∪ T_2
6:     T ← T\S
7:     return S, T
```

Proof of Correctness. It is clear that the defined pair of sets is unique and it is simple to verify that the given operations produce it. Although we have not shown explicitly how to do it here, both intersections and unions are easy to implement on the standard form of unions of intersections with finite intersection, using either a loop or recursion. Since the set of unions of intervals with finite intersection is closed under finite unions and intersections, the result of these operations is guaranteed to be a union of intervals with finite intersection. □

Remark 18. While it is clear what this algorithm does, it might not be clear why we need it. The idea is that S_1, S_2 will represent sets of elements x for which only one representation $x = u + v$ with $u \neq v \in U(1, X)$ has been found, and T_1, T_2 will represent sets of elements for which multiple representations have been found. The new sets S, T are what is needed to combine this information: S contains everything in S_1 and S_2 except those elements that are now confirmed to have multiple representations, which get moved to T along with everything else. An element is confirmed to have multiple representations if: 1) it is in either T_1 or T_2, or 2) it is not in T_1 or T_2 but is in both S_1 and S_2.

Algorithm 5.4. Let `one_rep_set` and `mult_rep_set` be unions of intervals with finite intersection such that there exists $u \in \langle 1, X\rangle \backslash U(1, X)$ so that

1. `one_rep_set` consists of all elements greater than u that can be written as the sum of two distinct elements in $U(1, X) \cap [1, u]$ in exactly one way, and

2. `mult_rep_set` consists of all elements greater than u that can be written as the sum of two distinct elements in $U(1, X) \cap [1, u]$ in more than one way.

On an input of two such sets `one_rep_set`, `mult_rep_set`, this algorithm returns the unique interval $[a, b]$ such that if we define $U' = (U(1, X) \cap [1, u]) \cup [a, b]$, then $U' = U(1, X) \cap [1, b+1]$.

1: **procedure** NEXTINTERVAL(`one_rep_set`, `mult_rep_set`)
2: $[p, q] \leftarrow$ smallest interval in `one_rep_set`
3: **if** $p = q$ **then**
4: $b \leftarrow p + X$
5: $b' \leftarrow \min(\texttt{one_rep_set} \cap [p + 1, \infty))$
6: $b'' \leftarrow \min(\texttt{mult_rep_set} \cap [p + 1, \infty))$
7: $b \leftarrow \min(b, b', b'')$
8: **return** $[p, b - 1]$
9: **else**
10: **return** $[p, p]$

Proof of Correctness. Let p be the smallest element in $U(1, X) \cap [u, \infty)$. It must be the sum of two unique elements $u, v \in U(1, X)$, and it is not hard to see that actually $u, v \in U(1, X) \cap [1, u]$, hence $p \in$ `one_rep_set`—specifically, p is the starting point of the smallest interval $[p, q]$ in `one_rep_set`. If $p \neq q$, we are done: the next interval of $U(1, X)$ is $[p, p]$.

This is because $p+1$ can already be written as the sum of two prior Ulam elements in two different ways: on the one hand, there is the representation coming from `one_rep_set`; on the other hand, $1, p \in U(1,X)$.

If $p = q$, then things are more complicated, and we need to find the largest b such that $[p, b-1] \subset U(1,X)$. Note that if $u \in U(1,X)$ then so is $u+1$ unless there exist some other $v, w \in U(1,X)$ such that $u+1 = v+w$. This condition can occur in one of three ways: first, if $u+1 = p+X$; second, if $u+1 \in$ `one_rep_set`; third, if $u+1 \in$ `mult_rep_set`. Therefore, b is the minimum of $p+X$, `one_rep_set` $\cap\, [p+1, \infty)$, and `mult_rep_set` $\cap\, [p+1, \infty)$. □

We are ready to give an algorithm for computing the coefficients of $U(1,X)$.

Algorithm 5.5. On an input of a natural number k, this algorithm returns the first $k+1$ coefficients a_i, b_i, c_i, d_i of $U(1,X)$.

1: **procedure** UlamCoefficients(k)
2: `ulam_set` $\leftarrow \{1\} \cup [X, 2X]$
3: `one_rep_set` $\leftarrow \{\}$
4: `mult_rep_set` $\leftarrow \{2X+1\}$
5: `largest_ulam` $\leftarrow \max($`ulam_set`$)$
6: **for** $1 < i \leq k$ **do**
7: `last` $\leftarrow$ last interval in `ulam_set`
8: **for** `interval` $\in$ `ulam_set` **do**
9: **if** `interval` $=$ `last` **then**
10: $(S,T) \leftarrow$ SumWithItself(`last`)
11: **else**
12: $(S,T) \leftarrow$ SumDistinct(`interval`, `last`)
13: $S \leftarrow S \cap [$`largest_ulam` $+2, \infty)$
14: $T \leftarrow T \cap [$`largest_ulam` $+2, \infty)$
15: (`one_rep_set`, `mult_rep_set`) $\leftarrow$ Amalgam(`one_rep_set`, `mult_rep_set`, S, T)
16: $I \leftarrow$ NewInterval(`one_rep_set`, `mult_rep_set`)
17: `ulam_set` $\leftarrow$ `ulam_set` $\cup\, I$
18: `largest_ulam` $\leftarrow \max($`ulam_set`$)$
19: `one_rep_set` $\leftarrow$ `one_rep_set` $\cap\, [$`largest_ulam` $+2, \infty)$
20: `mult_rep_set` $\leftarrow$ `mult_rep_set` $\cap\, [$`largest_ulam` $+2, \infty)$
21: **return** coefficients of `ulam_set`

Remark 19. The basic idea behind this algorithm is that it keeps track of three unions of intervals with finite intersections.

1. `ulam_set` maintains the subset of $U(1,X)$ computed thus far.

2. `one_rep_set` maintains a subset of $\langle 1, X\rangle$ such that every element of `one_rep_set` is larger than every element of `ulam_set`, and each element of `one_rep_set` can be written as a sum of distinct elements in `ulam_set` in exactly one way.

3. `mult_rep_set` maintains a subset of $\langle 1, X\rangle$ such that every element of `mult_rep_set` is larger than every element of `ulam_set`, and each element of `mult_rep_set` can be written as a sum of distinct elements in `ulam_set` in more than one way.

Proof of Correctness. For $l \in \mathbb{N}$, let $\mathcal{U}_l$ consist of the first $l+1$ intervals of $U(1, X)$. We show that at the end of each cycle of the outer for-loop indexed over i, `ulam_set` $= \mathcal{U}_i$; `largest_ulam` $= \max \mathcal{U}_i$; `one_rep_set` consists of all elements of $\mathcal{U}_{i-1} + \mathcal{U}_{i-1}$ larger than `largest_ulam`$+1$ that can be written as a sum of two distinct elements in $\mathcal{U}_{i-1}$ in exactly one way; and `mult_rep_set` consists of all elements of $\mathcal{U}_{i-1} + \mathcal{U}_{i-1}$ larger than `largest_ulam` $+ 1$ that can be written as a sum of two distinct elements in $\mathcal{U}_{i-1}$ in more than one way. We prove this claim by induction on i.

The base case $i = 1$ is obvious—this is just the initialization of `ulam_set`, `largest_ulam`, `one_rep_set`, and `mult_rep_set` prior to the for-loop. For all subsequent i, note that to find all elements $\mathcal{U}_{i-1} + \mathcal{U}_{i-1}$ that need to be added to `one_rep_set` and `mult_rep_set`, it suffices to consider the sums of the intervals in $\mathcal{U}_{i-2}$ with the last interval of $\mathcal{U}_{i-1}$, since all other sums have already been handled in prior steps. Thus, for each interval I of `ulam_set`, we add it to the last interval, producing a pair of sets S, T, where S consists of all elements that can be written down as a sum in just one way, and T consists of all elements that can be written down as a sum in multiple ways—this is established in the proofs of correctness of Algorithms 5.1 and 5.2. We remove anything smaller than `largest_ulam`$+2$ from both of these sets. We then use Algorithm 5.3 to merge S and T into `one_rep_set` and `mult_rep_set`, as described in Comment 18. Thus, by the end of the inner for-loop, we have constructed the desired sets `one_rep_set` and `mult_rep_set`. Algorithm 5.4 then uses these two sets to compute the next interval I of $U(1, X)$ such that $\mathcal{U}_i = \mathcal{U}_{i-1} \cup I$. We use this to then update `ulam_set`, `largest_ulam`, `one_rep_set`, and `mult_rep_set` as needed. □

What relation does this algorithm have to Theorem 5? To explain, it is helpful to first reformulate the result we wish to prove slightly. For each $n \in \mathbb{Z}$, let $\mathrm{eval}_{X \to n} : \mathbb{Z}[X] \to \mathbb{Z}$ be the evaluation homomorphism defined by $X \mapsto n$. Conjecture 3 can be stated as

$$\mathrm{eval}_{X \to n}\left(U(1, X)\right) = U(1, n)$$

for all $n \geq 4$. Similarly, Theorem 5 states that there exists an algorithm $\mathcal{A}_{\mathrm{Ulam}}$ such that on an input of $k > \mathbb{N}$, it computes the coefficients a_i, b_i, c_i, d_i of $U(1, X)$ and an $N \in \mathbb{N}$ such that for all $n \geq N$,

$$\mathrm{eval}_{X \to n}\left(U(1, X) \cap [1, c_k n + d_k + 1]\right) = U(1, n) \cap [1, c_k n + d_k + 1].$$

We can now show how this follows from Algorithm 5.5.

Proof of Theorem 5. Modify Algorithm 5.5 as follows: each time a comparison $p(X) < q(X)$, $p(X) \leq q(X)$, $p(X) > q(X)$, $p(X) \geq q(X)$, or $p(X) = q(X)$ is made, compute the smallest natural number M such that for all $m \geq M$, replacing X with m does not change the truth value of the comparison. It is not hard to see that such an M must exist by considering the behavior of these polynomials as $M \to \infty$. At the end of the computation, along with the coefficients a_i, b_i, c_i, d_i, return the maximum N of all of the aforementioned M. Call this new algorithm $\mathcal{A}_{\text{Ulam}}$.

Now, take any $n \geq N$ and consider modifying $\mathcal{A}_{\text{Ulam}}$ by replacing every set $S \subset \mathbb{Z}[X]$ used with $\text{eval}_{X \to n}(S)$. It is easy enough to see that this will be an algorithm that computes the coefficients of $U(1, n)$. But since N is chosen such that all comparisons of elements yield the same result as before, the output of this algorithm is the same as that of $\mathcal{A}_{\text{Ulam}}$, which proves that

$$\text{eval}_{X \to n}\left(U(1, X) \cap [1, c_k n + d_k + 1]\right) = U(1, n) \cap [1, c_k n + d_k + 1].$$

Thus, $\mathcal{A}_{\text{Ulam}}$ has the desired properties. □

Unfortunately, the particular algorithm $\mathcal{A}_{\text{Ulam}}$ defined in the proof of this theorem is not an efficient method of computing the coefficients a_i, b_i, c_i, d_i of $U(1, X)$—a naive implementation puts it in the $\Theta(k^3 \log_2(k))$ complexity class, due to the two nested for-loops and the need to perform binary search for the set operations. However, the ideas of the construction of the algorithm $\mathcal{A}_{\text{Ulam}}$ can be generalized—there are in fact families of algorithms with its essential properties. The importance of this is that it might be possible to construct a more efficient algorithm using, for example, ideas of Gibbs and Judson [5, 6]. Before we show what this family of algorithms is, we need some definitions.

Definition 20. Let S be a computable ordered ring. We say that a formula is *S-expressible* if it is

1. a variable x,
2. a constant $c \in S \cup \{\top, \bot\}$,
3. an empty list $[]$,

or if it can be constructed recursively from other S-expressible formulas according to the following rules:

1. If $\mathcal{E}_1, \mathcal{E}_2$ are S-expressible, then $\mathcal{E}_1 = \mathcal{E}_2$ is S-expressible.
2. If $\mathcal{E}_1, \mathcal{E}_2$ are S-expressible, then $\mathcal{E}_1 + \mathcal{E}_2, \mathcal{E}_1 - \mathcal{E}_2, \mathcal{E}_1 \cdot \mathcal{E}_2, \mathcal{E}_1 / \mathcal{E}_2$ are S-expressible if defined.
3. If $\mathcal{E}_1, \mathcal{E}_2$ are S-expressible, then $\mathcal{E}_1 < \mathcal{E}_2, \mathcal{E}_1 \leq \mathcal{E}_2, \mathcal{E}_1 > \mathcal{E}_2, \mathcal{E}_1 \geq \mathcal{E}_2$ are S-expressible if defined. (We call such formulas *comparisons*.)
4. If $\mathcal{E}_1, \mathcal{E}_2$ are S-expressible, then $\neg \mathcal{E}_1, \mathcal{E}_1 \vee \mathcal{E}_2, \mathcal{E}_1 \wedge \mathcal{E}_2, \mathcal{E}_1 \Rightarrow \mathcal{E}_2$ are S-expressible if defined.

5. If $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ are S-expressible, then $[\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n]$ is S-expressible.

We say that an algorithm $\mathcal{A}$ *has basic steps over* S if it is composed of the following basic steps:

1. Initializing $x \leftarrow \mathcal{E}$, where x is a variable and $\mathcal{E}$ is S-expressible.
2. Retrieving the i-th index of a list l.
3. Checking an if-statement $\textbf{if}(\mathcal{E})$, where $\mathcal{E}$ is S-expressible, and branching accordingly.
4. Running a while-loop $\textbf{while}(\mathcal{E})$, where $\mathcal{E}$ is S-expressible.
5. Returning an S-expressible formula $\mathcal{E}$.

For $\mathbb{Z}[X]$ and $\mathbb{Q}[X]$, specifically, we also need a way to extend the evaluation homomorphism $\text{eval}_{X\to n}$ to expressible formulas.

Definition 21. Let $\mathcal{E}$ be an $\mathbb{Q}[X]$-expressible formula. We define $\text{eval}_{X\to n}(\mathcal{E})$ to be the formula produced by replacing each instance of a constant c in $\mathcal{E}$ with $\text{eval}_{X\to n}(c)$.

With these definitions, we can state the following result.

Theorem 22. *Define $\mathcal{C}_{Ulam}$ to be the set of all algorithms $\mathcal{A}$ with basic steps over $\mathbb{Q}[X]$ such that on an input of $k \in \mathbb{N}$, $\mathcal{A}$ returns the first k coefficients of $U(1, X)$ and a proof of correctness. There exists an algorithm $\mathcal{P}$ such that, on an input of an algorithm $\mathcal{A} \in \mathcal{C}_{Ulam}$, it returns an algorithm $\mathcal{P}(\mathcal{A})$ which on an input of $k \in \mathbb{N}$ returns an $N \in \mathbb{N}$ such that for all $n \geq N$,*

$$\text{eval}_{X\to n}\left(U(1, X) \cap [1, c_k n + d_k + 1]\right) = U(1, n) \cap [1, c_k n + d_k + 1].$$

Proof. The algorithm runs as follows: on an input of $k \in \mathbb{N}$, run $\mathcal{A}(k)$, but each time a comparison $\mathcal{E}$ comes up in the computation, compute the smallest integer M such that for all $m \geq M$, $\mathcal{E} = \text{eval}_{X\to m}(\mathcal{E})$. Then return the maximum N of all of these M—there will only be finitely many comparisons for a given input k.

Why does N have the claimed properties? For each $n \geq N$, define an algorithm $\mathcal{A}_n$ produced by replacing each $\mathbb{Q}[X]$-expressible formula $\mathcal{E}$ in $\mathcal{A}$ with $\text{eval}_{X\to n}(\mathcal{E})$. Since the truth value of comparisons is preserved and $\mathcal{A}$ has basic steps over $\mathbb{Q}[X]$, $\mathcal{A}_n$ produces the same coefficients a_i, b_i, c_i, d_i as $\mathcal{A}$, and verifies that if we define a subset

$$\mathcal{U}_n = \bigcup_{i=0}^{k} \left[a_i n + b_i, c_i n + d_i\right],$$

then

1. $a_0 n + b_0 = c_0 n + d_0 = 1$ and $a_1 n + b_1 = n$,

2. $a_{i+1}n + b_{i+1} > c_i n + d_i + 1$ for all $i \in \mathbb{N}$, and

3. for every $x \in \mathcal{U}_n$, x is the smallest element of $\mathbb{Z}$ with the property that there exists exactly one way such that x can be written as the sum of two distinct elements $y, z \in \mathcal{U}_n$.

In other words, $\mathcal{U}_n = U(1,n) \cap [1, c_k n + d_k + 1] = \mathrm{eval}_{X \to n}(U(1,X) \cap [1, c_k X + d_k + 1])$. Ergo, we have the desired result. □

6 Acknowledgments

The author is exceptionally grateful to Oleg Sheydvasser, for lending his expertise in implementing efficient versions of the algorithms discussed in this paper, and for lending his processor time to run them. The author would also like to thank Stefan Steinerberger for providing helpful discussion and feedback, and Camp Minnetonka for the atmosphere in which this paper was finished.

References

[1] J. Cassaigne and S. R. Finch, A class of 1-additive sequences and quadratic recurrences, *Exp. Math.* **4** (1995), 49–60.

[2] S. R. Finch, Conjectures about s-additive sequences, *Fibonacci Quart.* **29** (1991), 209–214.

[3] S. R. Finch, On the regularity of certain 1-additive sequences, *J. Combin. Theory Ser. A* **60** (1992), 123–130.

[4] S. R. Finch, Patterns in 1-additive sequences, *Exp. Math.* **1** (1992), 57–63.

[5] P. Gibbs, An efficient method for computing Ulam numbers. Available at http://vixra.org/abs/1711.0134, 2015.

[6] P. Gibbs and J. McCranie, The Ulam numbers up to one trillion. Available at http://vixra.org/abs/1508.0085, 2017.

[7] J. Hinman, B. Kuca, A. Schlesinger, and A. Sheydvasser, Rigidity of Ulam sets and sequences, *Involve* **12** (2019), 521–539.

[8] J. Hinman, B. Kuca, A. Schlesinger, and A. Sheydvasser, The unreasonable rigidity of Ulam sequences, *J. Number Theory* **194** (2019), 409–425.

[9] N. Kravitz and S. Steinerberger, Ulam sequences and Ulam sets, *Integers* **18** (2018), A80.

[10] R. Queneau, Sur les suites s-additives, *J. Combin. Theory Ser. A* **12** (1972), 31–71.

[11] J. Schmerl and E. Spiegel, The regularity of some 1-additive sequences, *J. Combin. Theory Ser. A* **66** (1994), 172–175.

[12] S. Steinerberger, A hidden signal in the Ulam sequence, *Exp. Math.* **26** (2017), 460–467.

[13] S. Ulam, Combinatorial analysis in infinite sets and some physical theories, *SIAM Rev.* **6** (1964), 343–355.

2010 *Mathematics Subject Classification*: Primary 11Y55; Secondary 11B83, 11U10. *Keywords:* additive number theory, Ulam sequence, computational number theory.

(Concerned with sequence A002858.)

Received July 11 2021; revised version received November 30 2021. Published in *Journal of Integer Sequences*, December 27 2021.

Return to Journal of Integer Sequences home page.