A conjecture for Ulam Sequences

Philip Gibbs

Abstract: A conjecture on the quasi-periodic behaviour of Ulam sequences.

An Ulam sequence is an increasing sequence a_n $n \geq 1$ of positive integers such that each element after the second is the smallest positive integer greater than its predecessor which is the sum of two previous distinct elements of the sequence in exactly one way [1]. Such a sequence is determined by its first two elements. For example if $a_1 = 1$ and $a_2 = 2$ (the main Ulam sequence) then the sequence begins:

1,2,3,4,6,8,11,13,16,18,26,…

Numerical studies of Ulam sequences suggests that surprisingly they have a positive density rather than becoming rarer for increasing n. Although the numbers are often chaotically distributed they have been found by Steinerberger to have a natural wavelength $\lambda = \frac{2\pi}{\alpha}$ where in the case above $\alpha = 2.5714474995\ldots$ and $\cos(\alpha a_n) < 0$ for the first 10 million numbers on the sequence except 2,3,47 and 69 [2].

The main point of this paper is to make a more general conjecture as follows

For any Ulam sequence a_n there is a natural wavelength $\lambda \geq 2 \in \mathbb{R}$ such that if r_n is the residual of a_n mod λ in the interval [0,λ) then for any $\varepsilon > 0$ there are only a finite number of elements in the Ulam sequence such that $r_n < \frac{\lambda}{3} - \varepsilon$ or $r_n > \frac{2\lambda}{3} + \varepsilon$.

This means that the residuals almost always lie in the middle third of their possible range. When a number does not fall in this range we call it an **outlier**. For the above sequence the outliers are 2, 3, 8, 13, 36, 47, 53, 57, 69, 97, …

Steinerberger's observation that $\cos(\alpha a_n) < 0$ in all but a few cases follows from the conjecture with $\varepsilon = \frac{\lambda}{12}$ since this requires that the Ulam numbers fall in the middle half of the range where the cosine is negative in all but a finite number of exceptions.

It is easy to see that the sum of two numbers which are not outliers must itself be an outlier. This means that most numbers in an Ulam sequence include at least one outlier in their sum despite and because of the rarity of outliers. If the conjecture could be proven it would therefore go some way towards explaining the behaviour of the sequences.

In the case of Ulam sequences where $a_1 = 2$ and $a_2 \geq 5$ is odd it is known that the sequence has only a finite number of even elements and is eventually periodic [3,4]. This confirms the conjecture for these cases with $\lambda = 2$

We can model the behaviour of an Ulam sequence heuristically in the light of this conjecture using a stochastic approximation. Write p_a for the probability that a positive integer a is in the Ulam sequence and then treat these probabilities as independent so that

$$p_a = \left(\sum_{0<b<a} \frac{q_{a,b}}{(1-q_{a,b})}\right)\left(\prod_{0<b<a} (1-q_{a,b})\right)$$

$$q_{a,b} = p_b p_{a-b}(1-\delta_{a,b-a})$$

If the density of the sequence is low we can approximate this using a Poisson distribution

$$p_a = Q_q e^{-Q_a}, \quad Q_a = \sum_{0<b<a} q_{a,b}$$

Assume that for large a we can treat the probability as a function $p(r)$ of the residual in the middle third and as one for a in the set S of outliers

$$p_a = p(a \, mod \, \lambda) + \sum_{b \in S} \delta_{a,b}$$

$$p(r) = 0 \,, r < \frac{\lambda}{3} or \; r > \frac{2\lambda}{3}$$

When a is large with a residual $r = a \, mod \, \lambda$ outside the middle third

$$Q_a \gtrsim a \int_{r+\frac{\lambda}{3}}^{\frac{2\lambda}{3}} p(s)p(r+\lambda-s)ds \,, \qquad 0 < r < \frac{\lambda}{3}$$

$$Q_a \gtrsim a \int_{\frac{\lambda}{3}}^{r-\frac{\lambda}{3}} p(s)p(r-s)ds \,, \qquad \frac{2\lambda}{3} < r < \lambda$$

If the integrals are greater than zero then $Q_a = O(a)$ and $p_a = O(a)e^{-O(a)} \longrightarrow 0$ so for large a there will be very few outliers except close to $\frac{\lambda}{3}$ or $\frac{2\lambda}{3}$ where the convolution integral goes to zero. This is consistent with the conjecture. The requirement that the integral is everywhere greater than zero imposes limits on the size of gaps where the probability function is greater than zero. In particular if there is a gap such that $p(r) = 0$ for all r in a range $\frac{\lambda}{3} < r_0 < r < r_1 < \frac{2\lambda}{3}$ then there is a constraint

$$r_1 < \min\left(2r_0 - \frac{\lambda}{3}, \frac{r_0}{2} + \frac{\lambda}{3}\right)$$

The set S of outliers can be split into two parts: the right movers R with residual between 0 and $\frac{\lambda}{3}$ and the left movers L with residuals between $\frac{2\lambda}{3}$ and λ . The conjecture about the outliers is equivalent to the condition that the residuals of the outliers in R can be ordered into an increasing sequence $r^R{}_i$, $i = 1,2,\ldots$ which is either finite or tending to $\frac{\lambda}{3}$, and the residuals of the outliers in L can be ordered into a decreasing sequence $r^L{}_i, i = 1,2,\ldots$ which is either finite or tending to $\frac{2\lambda}{3}$

When the recurrence relation is applied to residuals in the middle third a different and more exact expression can be formed combining the outliers with the distribution function

$$p(r) = (\Sigma_L + \Sigma_R)\Pi_L\Pi_R$$

$$\Sigma_L = \sum_i \frac{p(r - r^R{}_i)}{1 - p(r - r^R{}_i)}$$

$$\Sigma_R = \sum_j \frac{p\left(r + (\lambda - r^L{}_j)\right)}{1 - p\left(r + (\lambda - r^L{}_j)\right)}$$

$$\Pi_L = \prod_i \left(1 - p(r - r^R{}_i)\right)$$

$$\Pi_R = \prod_j \left(1 - p\left(r + (\lambda - r^L{}_j)\right)\right)$$

These equations can be regarded as a non-linear recursion relation for $p(r)$ which under suitable conditions should converge to the distributions found computationally by Steinerberger. The number of terms in each sum or factors in each product is finite because of the conjecture.

If $r^L{}_1 - r^R{}_1 < \frac{2\lambda}{3}$ then there are no terms or factors for a central range so that

$$p(r) = 0, \qquad r^L{}_1 - \frac{\lambda}{3} < r < r^R{}_1 + \frac{\lambda}{3}$$

Furthermore there will be only one term and factor in a neigbouring range which gives

$$p(r) = p(r - r^R{}_1), \qquad r^R{}_1 + \frac{\lambda}{3} < r < r^R{}_2 + \frac{\lambda}{3}$$

$$p(r) = p\left(r + (\lambda - r^L{}_1)\right), \qquad r^L{}_2 - \frac{\lambda}{3} < r < r^L{}_1 - \frac{\lambda}{3}$$

As r approaches the ends of the middle third range the number of terms increases so that the products tend to zero much faster than the sums making the value of $p(r)$ drop to zero if the number of left and right movers is infinite.

To make the analysis more concrete the left and right movers for the main Ulam sequence can be computed as follows

i	outlier	$\frac{r^R{}_1}{\lambda}$	outlier	$\frac{r^L{}_1}{\lambda}$
1	3	0.22777	2	0.81851
2	47	0.23515	102	0.74437
3	69	0.23884	339	0.73865
4	8	0.27406	36	0.73330
5	2581	0.29639	273	0.72759
6	983	0.30118	400	0.70343

This first left and right movers imply a central gap for $3.0485 < \frac{2\pi r}{\lambda} < 3.5255$ as observed in Steinerberger's computed distribution (Figure 3 of [2]) The first left mover accounts for the repetition of the peak separated by $2\pi\left(1 - \frac{r^L{}_1}{\lambda}\right) = 1.1403$

References

[1] S.M. Ulam, Problems in modern mathematics. Science Editions John Wiley & Sons, Inc., New York 1964

[2] S Steinerberger, A Hidden signal in the Ulam sequence. arXiv:1507.00267 [math.CO]

[3] S. Finch, On the regularity of certain 1-additive sequences. J. Combin. Theory Ser. A 60 (1992), no. 1, 123-130.

[4] J. Schmerl and E. Spiegel, The regularity of some 1-additive sequences. J. Combin. Theory Ser. A 66 (1994), no. 1, 172-175.