

Steinhaus Piercing with Slack: The Exact Asymptotic Constant

Anton Joha

Independent Researcher, Norwich, United Kingdom
antonjoha@gmail.com

14 October 2026

Abstract

Let $s(d)$ be the largest order N for which the first $n + d$ points of a sequence in $[0, 1)$ meet every uniform n -cell for all $n \leq N$. We prove $s(d) = d/(1 - \log 2) + O(\sqrt{d})$, determining the limit asked for by Anholcer et al. and disproving their conjectured coefficient. For strong piercing, which requires meeting every interval of length $1/n$, we obtain an exact finite harmonic capacity and rational constructions. A Dirichlet covering lemma and a packing estimate for disjoint gaps transfer the capacity bound to ordinary grids. We also solve the finite largest-gap minimax problem, establish rigidity for infinite heap refinements, and extend the asymptotic estimate to affine arrival budgets. Ordinary near-extremizers satisfy universal laws for root sizes and first-split times, with a corresponding entropy correction. An explicit finite obstruction shows why root constraints alone do not characterize ordinary piercing.

Keywords: piercing sequences; irregularities of distribution; interval refinement; harmonic capacity; entropy.

2020 Mathematics Subject Classification: Primary 11K38; Secondary 05D99.

1 Introduction

Steinhaus asked how long points can be revealed in the unit interval so that the first n points meet all n equal cells at every stage [7, Problems 6–7]. The classical maximum is seventeen [2, 8]. Allowing an initial supply of extra points leads to a finite-horizon question: how far can a fixed amount of slack extend such a sequence?

For integers $d \geq 0$ and $N \geq 1$, call $x_1, \dots, x_{N+d} \in [0, 1)$ an **ordinary piercing sequence of order N with slack d** if

$$\{x_1, \dots, x_{n+d}\} \cap [k/n, (k+1)/n) \neq \emptyset \quad (1 \leq n \leq N, 0 \leq k < n).$$

Write $s(d)$ for the largest such order. Repetitions are permitted in the definition; our constructions have distinct coordinates. All unsubscripted logarithms are natural.

Anholcer et al. [1] gave linear lower and upper bounds and asked whether $s(d)/d$ has a limit. Their Conjecture 10 proposes $\log 2/(1 - \log 2)$. We determine the limit with a square-root error.

Theorem 1.1. As $d \rightarrow \infty$,

$$s(d) = \frac{d}{1 - \log 2} + O(\sqrt{d}), \quad \lim_{d \rightarrow \infty} \left(\frac{s(d)}{d} \right) = \frac{1}{1 - \log 2} = 3.258891353 \dots$$

The coefficient differs from the conjectured value by exactly one. The proof occupies Sections 2–3 and is independent of finite enumeration or computational certificates.

Call a sequence **strongly piercing** if its first $n + d$ points meet every half-open interval $[y, y + 1/n) \subseteq [0, 1)$ for all $n \leq N$. Let $\sigma(d)$ be its maximal order, so $\sigma(d) \leq s(d)$. We include the boundary gaps at 0 and 1 when measuring gaps. Strong piercing forces every gap to have length at most $1/n$, and a strict inequality for every gap is sufficient. At equality, the endpoint positions matter under the half-open convention.

For $a = \lceil N/2 \rceil$, define

$$\Phi_N(d) = \sum_{j=a}^{N-1} \frac{1}{j} + \frac{2a + d + 1 - N}{N}.$$

Theorem 2.1 proves that strong piercing of order N is possible exactly when $\Phi_N(d) > 1$. A rational binary refinement tree attains the capacity, giving $\sigma(d) = d/(1 - \log 2) + O(1)$. It also solves the finite minimax problem for the normalized largest gap (Proposition 2.2).

The logarithmic splitting pattern of de Bruijn and Erdős [3] supplies the dynamic part of the construction. For a finite horizon, however, the initial slack can populate a complementary static region. This allocation yields the improved lower bound. Proposition 2.3 quantifies the distinction between separate finite-horizon designs and one infinite realization within the heap splitting schedule. Earlier work on finite irregularities includes Graham [4], Levy [6] and Konyagin [5]; the generalized piercing problem is treated in [1].

For the ordinary upper bound, we consider the gaps present at stage $\lceil N/2 \rceil$ and assign each its first later split time. An empty gap can cross cell boundaries without containing a complete cell. A Dirichlet covering lemma controls the locations of gaps exceeding their strong capacities; disjointness bounds their total positive excess by $176/\sqrt{N}$. Thus every ordinary sequence of order N satisfies

$$1 \leq \Phi_N(d) + \frac{176}{\sqrt{N}},$$

which proves Theorem 1.1. This packing scale is sharp for arbitrary disjoint avoiding intervals, although sharpness of the error in $s(d)$ remains open.

Section 4 treats general arrival budgets, and Section 5 derives universal size, split-time and entropy laws for near-extremizers. Appendix A describes the arithmetic fluctuations of the strong endpoint. Appendix B gives explicit obstructions showing why symmetric gap statistics and first-split root constraints do not determine ordinary coverage.

2 Exact capacity for strong piercing

2.1 Harmonic capacity

For $N \geq 1$ and $d \geq 0$, put $a = \lceil N/2 \rceil$ and define

$$\Phi_N(d) = \sum_{j=a}^{N-1} \frac{1}{j} + \frac{2a + d + 1 - N}{N}. \quad (2.1)$$

Theorem 2.1. A strongly $(n + d)$ -piercing sequence of order N exists if and only if $\Phi_N(d) > 1$. Consequently,

$$\sigma(d) = \max\{N \geq 1 : \Phi_N(d) > 1\}. \quad (2.2)$$

For $d = 0, 1, \dots, 10$, the values are 4, 8, 11, 14, 17, 21, 24, 27, 30, 34, 37.

Proof. At stage a , the $a + d$ points give $a + d + 1$ gaps, with endpoints 0 and 1 included. Each gap has either a first later insertion or no later insertion. A gap first split at stage $j + 1$ has its original length at stage j , so that length is at most $1/j$. A gap never split has length at most $1/N$. At most one old gap has its first split at each insertion. Assigning the baseline $1/N$ to every old gap and allowing the increment $1/j - 1/N$ at every subsequent insertion gives

$$1 \leq \frac{a + d + 1}{N} + \sum_{j=a}^{N-1} \left(\frac{1}{j} - \frac{1}{N} \right) = \Phi_N(d). \quad (2.3)$$

This accounting permits repeated insertions in an already split gap and zero-length gaps.

Equality cannot occur. For $N \geq 4$, Bertrand's postulate gives a prime p with $N/2 < p < N$. In the harmonic sum in (2.1), exactly one denominator is divisible by p , namely p . Since p does not divide N , the reduced denominator of $\Phi_N(d)$ is divisible by p , and $\Phi_N(d)$ cannot equal 1. Direct substitution excludes equality for $N = 1, 2, 3$ as well.

For sufficiency, take the binary tree with internal nodes $1, \dots, N - 1$ and terminal nodes $N, \dots, 2N - 1$, the children of j being $2j$ and $2j + 1$. Give each terminal node capacity $1/N$, and recursively set

$$w_j = \min \left\{ \frac{1}{j}, w_{2j} + w_{2j+1} \right\}, \quad j < N. \quad (2.4)$$

If $a \leq j < N$, both children are terminal and $w_j = 1/j$. If $j < a$ and both children are internal, their capacities sum to at most $1/(2j) + 1/(2j + 1) < 1/j$. The only remaining case has $N = 2j + 1$; its capacities sum to $1/(N - 1) + 1/N < 1/j$. Thus the minimum is inactive for node indices $j < a$, and

$$w_1 = \sum_{j=a}^{N-1} \frac{1}{j} + \frac{2a - N}{N}. \quad (2.5)$$

For $N = 1$, the root is the sole terminal node, and (2.5) still holds. Add $d + 1$ static gaps of capacity $1/N$. Together with the dynamic root, their total capacity is $F = \Phi_N(d) > 1$. Make the root interval have length w_1/F , followed by $d + 1$ static intervals of length $1/(NF)$. The left endpoints of the static intervals are the initial $d + 1$ points.

Split a node interval in the ratio $w_{2j} : w_{2j+1}$, inserting its dividing point in node order $j = 1, 2, \dots, N - 1$. Every node interval has length at most w_j/F . After $n - 1$ splits the active dynamic nodes are exactly $n, \dots, 2n - 1$. Each internal active node has length at most $1/(nF)$, as do the terminal active nodes and the static gaps. Every gap is therefore strictly shorter than $1/n$, and the prefix has exactly $n + d$ points. All capacities and coordinates are rational. This proves sufficiency. $\square$

Equation (2.3) is a capacity bound on old gaps and their available first-split times; the harmonic tree attains it. Expansion of (2.1) gives

$$\Phi_N(d) = \log 2 + \frac{d + 3/2}{N} + O(N^{-2}), \quad \sigma(d) = \frac{d}{1 - \log 2} + O(1). \quad (2.6)$$

2.2 A logarithmic construction

An explicit logarithmic tree explains the limiting mass allocation. Put $q = d + 1$ and $b = \log 2$. Initially place the q points

$$b + \frac{r(1-b)}{q}, \quad r = 0, 1, \dots, q-1. \quad (2.7)$$

They divide $[b, 1]$ into q static gaps. For $j \geq 1$, set $k = \lfloor \log_2 j \rfloor$ and split the interval

$$I_j = \left[\log \frac{j}{2^k}, \log \frac{j+1}{2^k} \right] \quad \text{at} \quad z_j = \log \frac{2j+1}{2^{k+1}}. \quad (2.8)$$

The children are I_{2j} and I_{2j+1} . Inserting $z_1, z_2, \dots$ leaves nodes $n, \dots, 2n-1$ active at stage n , each of length at most $\log(1 + 1/n) < 1/n$. The static gaps have length $(1 - \log 2)/q$. Hence

$$s(d) \geq \sigma(d) \geq \lfloor \frac{d+1}{1 - \log 2} \rfloor. \quad (2.9)$$

At equality in a static gap bound its left endpoint is present; the left boundary gap is always strict, so the half-open convention causes no exception. This is the de Bruijn–Erdős logarithmic splitting pattern on a shorter interval, with the complementary interval populated at the outset. That initial allocation is available because the horizon is finite.

2.3 The exact finite minimax problem

Let $\Delta_m(x)$ be the largest gap after the first m points, including boundary gaps, and define

$$D_{N,d} = \inf_{x_1, \dots, x_{N+d}} \max_{1 \leq n \leq N} n \Delta_{n+d}(x). \quad (2.10)$$

For the capacity argument, take the gaps after the first $a + d$ points, where $a = \lceil N/2 \rceil$, as roots.

Proposition 2.2. For every $N \geq 1$ and $d \geq 0$, $D_{N,d} = 1/\Phi_N(d)$, and distinct rational coordinates attain the infimum.

Proof. If the objective is D , a root first split at stage $j+1$ has length at most D/j , and a never-split root has length at most D/N . Therefore $1 \leq D\Phi_N(d)$. For attainment, use the harmonic tree and static capacities in Theorem 2.1, now normalized by $\Phi_N(d)$ whether or not it exceeds one. The same induction gives every stage- n gap at most $1/(n\Phi_N(d))$. $\square$

For an arbitrary fixed refinement tree with node capacities b_v , its maximum realizable root mass is obtained recursively as the minimum of the parent capacity and the sum of its child capacities. Equivalently it is the minimum of $\sum_{v \in \mathcal{A}} b_v$ over cutsets meeting every root-to-leaf path exactly once. Induction proves both statements. This is tree flow–cut duality; complete prefix-free cutsets with $b_v = 2^{-|v|}$ give the equality case of the binary Kraft identity. Here the harmonic formula explicitly evaluates the relevant cutset capacity.

2.4 One infinite heap refinement

The minimax problem permits the point sequence to depend on the horizon. To compare it with a fixed infinite construction, start with the interval $[0, 1]$ labelled 1, of mass $w_1 = 1$. At insertion j , split node j into children $2j$ and $2j+1$, with positive masses adding to w_j . Either child may lie to the left. This is the **heap schedule**; after n insertions the $n+1$ active gaps have labels $n+1, \dots, 2n+1$.

Write $R_N = \max_{1 \leq n \leq N} n \Delta_n$ and $C_\infty = 1/\log 2$. The logarithmic masses below are those of de Bruijn and Erdős [3]. The rigidity statement concerns indexed masses; the child orientations remain free.

Proposition 2.3. In the heap schedule, if $R_N \leq C_\infty$ for every N , all masses are forced:

$$w_j = C_\infty \log \frac{j+1}{j}. \quad (2.11)$$

In this case $R_N = (N/\log 2) \log(1 + 1/(N+1))$. For every other mass assignment some finite $R_M > C_\infty$. Consequently every infinite heap realization satisfies

$$\liminf_{N \rightarrow \infty} N(R_N - D_{N,0}) \geq \frac{3(1 - \log 2)}{2(\log 2)^2}, \quad (2.12)$$

with equality for (2.11) and infinite left side otherwise.

Proof. Node $k \geq 2$ survives through $k-1$ insertions, so $w_k \leq C_\infty/(k-1)$. The depth- h descendants of node j have labels $2^h j, \dots, 2^h(j+1)-1$. Sum their bounds and let $h \rightarrow \infty$ to obtain $w_j \leq C_\infty \log((j+1)/j)$. On every active frontier $n+1, \dots, 2n+1$, these bounds sum to $C_\infty \log 2 = 1$, forcing equality throughout. The formula for R_N follows because $n \log(1 + 1/(n+1))$ increases. A different assignment must violate the bound at a finite stage; its cumulative excess persists. Finally, Proposition 2.2 and (2.6) give $D_{N,0} = 1/\log 2 - 3/(2(\log 2)^2 N) + O(N^{-2})$. $\square$

3 Transfer to ordinary grids

A Dirichlet covering lemma bounds the excess lengths of disjoint gaps. The resulting estimate transfers the limiting constant of the strong problem to uniform grids. Lemma 3.1 is a Dirichlet-approximation counterpart of the Farey covering lemma of Anholcer et al. [1, Lemma 8], with different constants and exceptional neighborhoods of small-denominator fractions.

Lemma 3.1 (Dirichlet covering). Let $W \geq 2$ and $k \geq 8W^2$ be integers; set $\alpha = 1 + 4/W$, $\beta = 1 + 2/W$, $L = \lceil \beta k \rceil$ and $U = \lfloor \alpha k \rfloor$. Every interval $[y, y + 1/k) \subset [0, 1)$ either contains a full n -grid cell strictly away from both endpoints for some $L \leq n \leq U$, or satisfies

$$|y - a/q| \leq \frac{W}{qk} \quad \text{for a reduced } a/q \in [0, 1], \quad q < 4W. \quad (3.1)$$

Proof. Put $Q = U - L + 1 \geq 2k/W - 1 \geq 16W - 1$. Dirichlet's pigeonhole argument supplies a reduced a/q with $q \leq Q$ and $|y - a/q| \leq 1/(qQ)$. If $q < 4W$, use $Q \geq k/W$. Otherwise, the q points $(L+t)y \bmod 1$, $0 \leq t < q$, lie within circular distance $1/Q$ of the translated uniform q -grid. Its perturbed gaps are at most

$$\frac{1}{q} + \frac{2}{Q} \leq \frac{1}{4W} + \frac{2}{16W-1} < \frac{1}{W}.$$

Some $n = L + t$ therefore satisfies $\{ny\} \in (1 - 1/W, 1)$. For $c = \lceil ny \rceil$, $c/n > y$ and $(c+1)/n - y = (2 - \{ny\})/n < (1 + 1/W)/n < 1/k$. The cell $[c/n, (c+1)/n)$ is strictly inside the interval. $\square$

Lemma 3.2 (packing the positive excess). Let I_r be pairwise interior-disjoint intervals in $[0, 1]$, of lengths ℓ_r , with integer deadlines $\lceil N/2 \rceil \leq h_r \leq N$. Suppose the interior of I_r contains no full grid cell of any order at most h_r . Then

$$\sum_r (\ell_r - 1/h_r)_+ \leq \frac{176}{\sqrt{N}}. \quad (3.2)$$

Proof. Set $e_r = N(\ell_r - 1/h_r)_+$. Avoidance of the h_r -grid gives $\ell_r \leq 2/h_r$, hence $e_r \leq 2$. Every positive-excess interval is longer than $1/N$, so their left endpoints are more than $1/N$ apart and there are at most N of them.

For $N \geq 64W^2$, put $k_r = \lfloor h_r/(1 + 4/W) \rfloor$. Directly from the floor inequality,

$$k_r \geq 8W^2, \quad k_r \geq N/7, \quad N(1/k_r - 1/h_r) \leq \frac{8/W + 12/N}{1 - 6/N} \leq \frac{9}{W}. \quad (3.3)$$

If $e_r > 9/W$, then $\ell_r > 1/k_r$. Lemma 3.1 places the left endpoint within $7W/(qN)$ of a reduced fraction of denominator $q < 4W$; its grid-cell alternative is excluded. At most $14W/q + 1$ separated endpoints fit in such a neighborhood. There are at most q fractions of denominator q , apart from the extra endpoint at denominator one. Thus

$$\#\{r : e_r > 9/W\} \leq \sum_{q=1}^{4W-1} q(14W/q + 1) + (14W + 1) = 64W^2 - 2W + 1 < 64W^2.$$

For $0 < t \leq 2$, take $W = \lceil 9/t \rceil \leq 11/t$. If the size condition on N fails, the trivial bound by N gives the same conclusion. Hence

$$\begin{aligned} \#\{r : e_r > t\} &\leq \min\{N, 7744/t^2\}, \\ \sum_r e_r &\leq \int_0^\infty \min\{N, 7744/t^2\} dt = 176\sqrt{N}. \end{aligned} \quad (3.4)$$

Division by N proves (3.2). $\square$

Theorem 3.3. Put $\eta = 1 - \log 2$. Every ordinary sequence of order N satisfies

$$1 \leq \Phi_N(d) + \frac{176}{\sqrt{N}}, \quad \eta N \leq d + 3 + 176\sqrt{N}. \quad (3.5)$$

Consequently, as $d \rightarrow \infty$,

$$s(d) = \frac{d}{1 - \log 2} + O(\sqrt{d}), \quad 0 \leq s(d) - \sigma(d) = O(\sqrt{d}). \quad (3.6)$$

Proof. Put $A = \lceil N/2 \rceil$ and call the gaps between the first $A + d$ points the roots. Assign deadline j to a root first split by the arrival at stage $j + 1$, and deadline N to a root never split. Through its deadline a root contains no selected point in its interior, and therefore no full grid cell. Lemma 3.2 applies. The nominal capacities sum to at most $\Phi_N(d)$, by the same first-split accounting as (2.3). The lengths sum to one, proving the first inequality. The harmonic estimate $\Phi_N(d) \leq \log 2 + (d + 3)/N$ gives the second.

This inequality first gives $N = O(d + 1)$ and then $N \leq d/\eta + O(\sqrt{d + 1})$. The exact strong construction in Theorem 2.1 gives the matching lower bound with error $O(1)$. The form $O(\sqrt{d + 1})$ also covers $d = 0$. $\square$

The limiting coefficient $1/(1 - \log 2) = 3.258891353\dots$ differs from $\log 2/(1 - \log 2)$ proposed in [1, Conjecture 10]. The tail estimate (3.4) is a weak-square-summability bound; integrating it charges each exceptional root according to its actual excess.

Proposition 3.4 (sharpness of the packing scale). For every $N \geq 16$ there are disjoint intervals avoiding all complete grid cells through order N for which

$$\sum_I (|I| - 1/N)_+ \geq \frac{1}{16\sqrt{N}}. \quad (3.7)$$

Proof. Set $t = 1/(4\sqrt{N})$ and $\ell = (1+t)/N$. Among starts in $[0, 1-\ell]$, the measure forbidden by all cells of orders at most N is at most $\sum_{1/\ell < n \leq N} n(\ell - 1/n) \leq (Nt+1)t \leq 1/8$. Close the forbidden intervals to exclude endpoint equalities. The allowed set has measure greater than $3/4$. A maximal family of starts separated by at least ℓ has at least $3/(8\ell)$ members, since their radius- ℓ neighborhoods cover the allowed set. Its total excess is at least $3t/[8(1+t)] > 1/(16\sqrt{N})$. $\square$

Thus Lemma 3.2 is sharp for disjoint avoiding intervals. Improving the sequence bound (3.6) requires compatibility of later subdivisions or cancellation between signed defects.

4 General arrival budgets

For a nonnegative nondecreasing integer budget $M(n)$, put $A = \lceil N/2 \rceil$ and

$$\mathcal{P}_N(M) = \frac{M(A)+1}{N} + \sum_{j=A}^{N-1} (M(j+1) - M(j)) \left(\frac{1}{j} - \frac{1}{N} \right). \quad (4.1)$$

Proposition 4.1. If the first $M(n)$ points cover each n -grid through order N , then $1 \leq \mathcal{P}_N(M) + 176/\sqrt{N}$. For strong piercing the error term is unnecessary.

Proof. There are $M(A)+1$ roots. At transition j at most $M(j+1) - M(j)$ roots receive their first split, each increasing the baseline capacity by $1/j - 1/N$. Lemma 3.2 bounds their total positive excess. $\square$

The midpoint minimizes this nominal capacity for every arrival schedule. If the starting stage is allowed to vary, the difference between consecutive choices is

$$\mathcal{P}_N(M; A+1) - \mathcal{P}_N(M; A) = (M(A+1) - M(A)) (2/N - 1/A). \quad (4.2)$$

The sign changes at the midpoint. The stated packing error is asserted at that midpoint.

For fixed $\gamma > 0$, let $s_\gamma(d)$ be the maximal order with budget $\lfloor \gamma n \rfloor + d$. If no positive order is feasible, its value is defined to be zero. This includes $\gamma < 1$, subject to the necessary cardinality condition. The original endpoint is $s_1(d) = s(d)$.

Corollary 4.2. For every fixed $0 < \gamma < 1/\log 2$,

$$s_\gamma(d) = \frac{d}{1 - \gamma \log 2} + O_\gamma(\sqrt{d+1}). \quad (4.3)$$

For $\gamma \geq 1/\log 2$, an infinite strongly piercing sequence exists with budget $\lfloor \gamma n \rfloor$ and zero additive slack.

Proof. Summation by parts in (4.1), using the bounded discrepancy of $\lfloor \gamma n \rfloor - \gamma n$, gives $\mathcal{P}_N(M) = \gamma \log 2 + d/N + O_\gamma(1/N)$. Proposition 4.1 supplies the upper estimate. For the lower bound, let $b = \gamma \log 2 < 1$. Divide $(b, 1)$ into d static gaps, with their left endpoints present initially, and scale the logarithmic tree of Section 2.2 to $(0, b)$. After $\lfloor \gamma n \rfloor$ dynamic splits every dynamic gap is shorter than $\gamma/(\lfloor \gamma n \rfloor + 1) < 1/n$. The static gaps satisfy the target through every order with $N(1-b) \leq d$. This gives the lower estimate as $d \rightarrow \infty$. At or above the critical rate, normalize the logarithmic tree to $(0, 1)$. After $m = \lfloor \gamma n \rfloor$ splits its largest gap is below $1/((m+1)\log 2) < 1/n$. $\square$

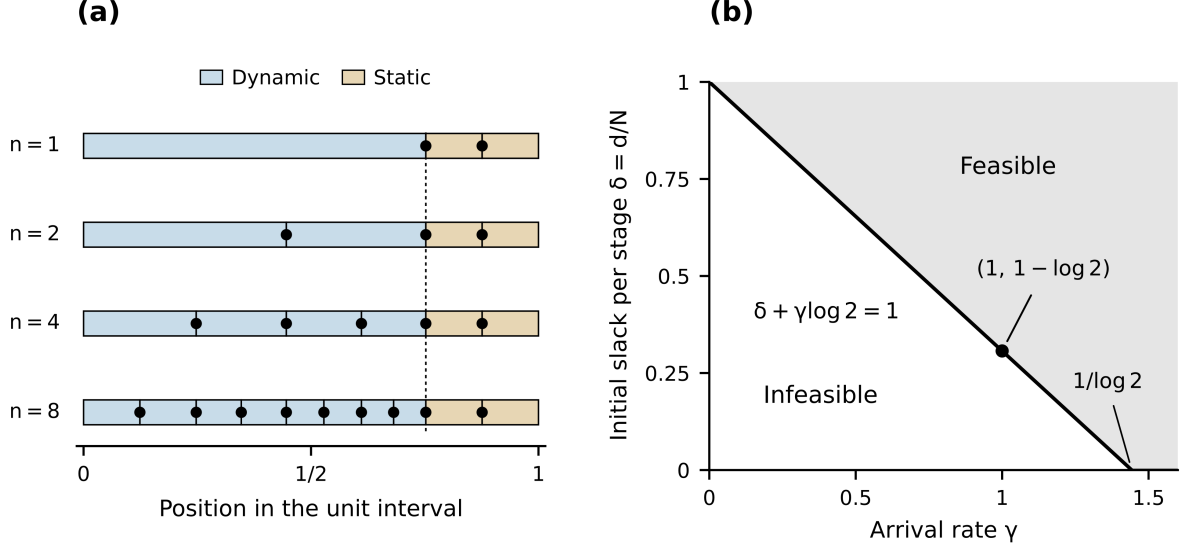


Figure 1: The finite harmonic construction and the affine budget boundary. Left: the order-eight, slack-one construction at four stages; two static gaps persist while the dynamic interval is split. Right: the limiting boundary (4.4), with unit arrival rate and normalized slack $1 - \log 2$ marked.

In normalized coordinates $\delta = d/N$, the limiting boundary is

$$\delta + \gamma \log 2 = 1. \quad (4.4)$$

Strict excess above the boundary is achievable at sufficiently large horizons, while a strict deficit is impossible. For $\gamma < 1$, the cardinality upper bound $N \leq d/(1 - \gamma)$ is weaker than (4.3). Figure 1 pairs the harmonic construction with this affine-budget boundary.

5 Geometry near the maximal horizon

The asymptotic endpoint determines the limiting geometry of ordinary piercing sequences. Throughout this section $N \geq 2$, $A = \lceil N/2 \rceil$, $\eta = 1 - \log 2$, and the roots are the gaps between the first $A + d$ points, including zero gaps. Give a root R deadline $h_R = j$ if its first split occurs at transition j to $j + 1$, and deadline N if it remains unsplit. Put

$$\begin{aligned} \epsilon &= \Phi_N(d) - 1, \\ P &= \sum_R (|R| - 1/h_R)_+, \\ Q &= \sum_R (1/h_R - |R|)_+. \end{aligned} \quad (5.1)$$

Both sums include the never-split roots. Lemma 3.2 gives $P \leq 176N^{-1/2}$ for ordinary sequences; for strongly piercing sequences $P = 0$.

5.1 First-split times and unused slots

Let $J \subset \{A, \dots, N - 1\}$ be the slots that do not first split a root, and let $u = |J|$. Write ℓ_j for the length first split in slot j , or zero for an unused slot, and U for the total never-split length. Define

$$\nu_N = \sum_{j=A}^{N-1} \ell_j \delta_{j/N} + U \delta_\infty, \quad \mu = \frac{dt}{t}|_{[1/2,1]} + \eta \delta_\infty. \quad (5.2)$$

Thus ν_N samples a location uniformly in space and records the first split of its root. Use the usual distance between finite times and distance one from the isolated point ∞ to every finite time; write W_1 for the corresponding Wasserstein distance.

Theorem 5.1. With $e = \epsilon + P$, one has

$$\epsilon = Q - P + \sum_{j \in J} (1/j - 1/N), \quad e \geq \frac{u(u+1)}{2N^2}, \quad (5.3)$$

and

$$W_1(\nu_N, \mu) \leq \max\{P, e + \sqrt{2e}\} + \frac{3}{N}. \quad (5.4)$$

Consequently, whenever $d/N \rightarrow \eta$ through ordinary feasible sequences, $\nu_N \rightarrow \mu$. At the maximal ordinary horizon $N = s(d)$, $W_1(\nu_N, \mu) = O(N^{-1/4})$ and $u = O(N^{3/4})$.

Proof. Exactly $N - A - u$ roots are first split. Sum their capacities $1/j$ and the other capacities $1/N$, then subtract Q and add P , to obtain the identity. Among u unused slots the last u minimize the displayed sum; their cost is at least $\sum_{r=1}^u r/N^2$. This proves the second assertion.

Compare ν_N with

$$\lambda_N = \sum_{j=A}^{N-1} \frac{1}{j} \delta_{j/N} + \left(1 - \sum_{j=A}^{N-1} \frac{1}{j}\right) \delta_\infty.$$

The finite-atom positive and negative discrepancies satisfy $F_+ = \sum_j (\ell_j - 1/j)_+ \leq P$ and $F_- = \sum_j (1/j - \ell_j)_+ \leq e + u/N$. At infinity the discrepancy is $F_- - F_+$, so the total positive mass to transport is $\max(F_+, F_-)$. The metric has diameter one.

For the reference comparison, move dt/t on each interval $[j/N, (j+1)/N]$ to its left endpoint, at cost at most $1/N$. The additional mass needed to reach $1/j$ sums to at most $1/A - 1/N \leq 1/N$. The omitted interval $[1/2, A/N]$ has mass less than $1/N$. Routing these discrepancies through infinity gives $W_1(\lambda_N, \mu) < 3/N$. This proves (5.4). Finally, $\epsilon = d/N - \eta + O(1/N)$, and Theorem 3.3 gives $|d/N - \eta| = O(N^{-1/2})$ at $N = s(d)$. $\square$

In particular, the never-split mass tends to η . Conditional on a split, the limiting density is $1/(t \log 2)$ on $[1/2, 1]$, or equivalently $\log_2(2T)$ is uniform. This law holds for every ordinary near-extremizer.

For maximal strong sequences, $P = 0$ and $\epsilon = O(1/N)$, so the rates improve to $O(N^{-1/2})$ and $O(\sqrt{N})$ respectively. There is also a sharper ordinary rate if the statistic identifies “never split” with time 1. Writing bars for these censored laws, transfer discrepancies via 1 to obtain

$$W_1(\bar{\nu}_N, \bar{\mu}) \leq \epsilon + 2P + 3/N. \quad (5.5)$$

Indeed the finite cost is at most $\sum_{j \notin J} |\ell_j - 1/j| (1 - j/N) + \sum_{j \in J} (1/j - 1/N) \leq e + P$. Thus the censored law has error $O(N^{-1/2})$ at maximal ordinary horizons. This metric no longer distinguishes an untouched root from one first split immediately before the horizon.

5.2 Gap sizes under two sampling conventions

Put $M = A + d + 1$, $B_N = 2A + d + 1 - N$, and $\rho = 1/2 + \eta$. Define the root-count law and its harmonic reference by

$$\Pi_N = \frac{1}{M} \sum_R \delta_{N|R|}, \quad \Pi_N^{\text{ref}} = \frac{1}{M} \left(\sum_{j=A}^{N-1} \delta_{N/j} + B_N \delta_1 \right). \quad (5.6)$$

Theorem 5.2. With the usual metric on the real line,

$$W_1(\Pi_N, \Pi_N^{\text{ref}}) \leq \frac{N}{M}(\epsilon + 2P). \quad (5.7)$$

If $d/N \rightarrow \eta$ through ordinary feasible sequences, then

$$\begin{aligned} \Pi_N &\rightarrow \Pi_* = \frac{1}{\rho} \left(\frac{dz}{z^2} \Big|_{[1,2]} + \eta \delta_1 \right), \\ \Xi_N &:= \sum_R |R| \delta_{N|R|} \rightarrow \Xi_* = \frac{dz}{z} \Big|_{[1,2]} + \eta \delta_1. \end{aligned} \quad (5.8)$$

At $N = s(d)$ both distances are $O(N^{-1/2})$; at $N = \sigma(d)$ for strong sequences both are $O(1/N)$.

Proof. Match each split root to its reference atom N/j . Move the reference atom of every unused slot from N/j to 1, then pair the remaining atoms with the never-split roots. Before division by M , the total cost is

$$N \left(P + Q + \sum_{j \in J} (1/j - 1/N) \right) = N(\epsilon + 2P).$$

This proves (5.7) without a coverage assumption. Riemann sums for $t \mapsto 1/t$ give $W_1(\Pi_N^{\text{ref}}, \Pi_*) = O(|d/N - \eta| + 1/N)$. Combine this with (5.7) and Lemma 3.2. For spatial sampling, $\Xi_N[f] = (M/N) \Pi_N[zf(z)]$. Ordinary roots have $N|R| \leq 2N/A \leq 4$, so multiplication by z preserves a uniform Lipschitz bound on the common support. The same rates follow. $\square$

The two measures record different experiments: choose a root uniformly, or choose a point uniformly and inspect its root. Figure 2 compares their cumulative distributions. The limiting atom at 1 represents roots whose lengths concentrate near $1/N$. Corollary 5.4 below identifies their common joint law with censored first-split time.

5.3 Entropy and collisions

For the root masses let $H_A = -\sum_R |R| \log |R|$, with $0 \log 0 = 0$.

Corollary 5.3. Every ordinary near-extremizing family with $d/N \rightarrow \eta$ satisfies $H_A = \log N - (\log 2)^2/2 + o(1)$. At $N = s(d)$,

$$\begin{aligned} H_A &= \log N - \frac{(\log 2)^2}{2} + O(N^{-1/2}), \\ \log M - H_A &\rightarrow \kappa = \log(3/2 - \log 2) + \frac{(\log 2)^2}{2}. \end{aligned} \quad (5.9)$$

Here $\kappa = 0.02561249973766 \dots > 0$ is the limiting relative entropy of the root-mass vector with respect to uniform mass on its M roots.

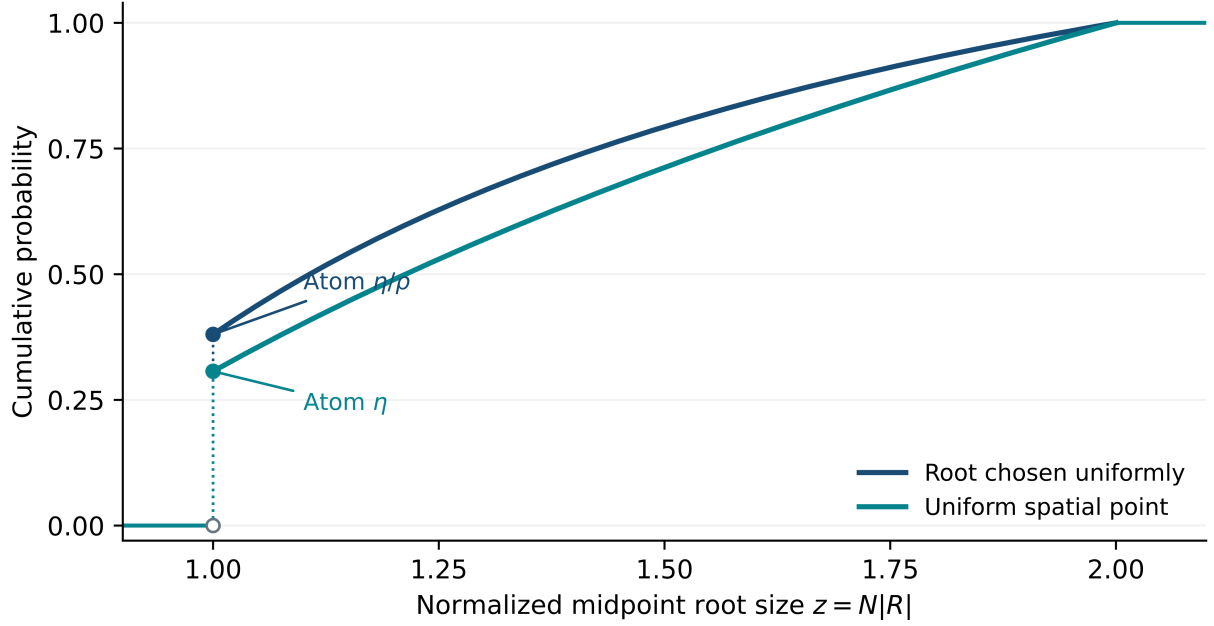


Figure 2: Limiting cumulative distributions of the normalized midpoint root size $z = N|R|$. Uniform sampling of roots gives Π_* , whereas a uniform spatial location gives the length-biased law Ξ_* . The jumps at 1 have masses η/ρ and η , respectively. Both laws are forced for ordinary near-extremizers by Theorem 5.2.

Proof. For $0 \leq z \leq 4$ and $1 \leq y \leq 2$, $|z \log z - y \log y| \leq (1 + \log 4) |z - y|$. For $z \geq 1$ use the mean-value theorem; for $z < 1$ first use $-z \log z \leq 1 - z$. Applying this estimate to the coupling in (5.7) gives the finite bound

$$\left| H_A - \left(\log N - \sum_{j=A}^{N-1} \frac{\log(N/j)}{j} \right) \right| \leq (1 + \log 4) (\epsilon + 2P). \quad (5.10)$$

The Riemann sum tends to $(\log 2)^2/2$, with error $O(1/N)$. Theorem 3.3 and $M/N \rightarrow \rho$ finish the proof. $\square$

For maximal strong sequences the error in (5.9) is $O(1/N)$. More generally, each fixed $p > 1$ gives, at a maximal ordinary horizon,

$$N^{p-1} \sum_R |R|^p = \eta + \frac{2^{p-1} - 1}{p - 1} + O_p(N^{-1/2}). \quad (5.11)$$

This follows from (5.8) applied to z^p on $[0, 4]$. In particular, $\sum_R |R|^2 = (2 - \log 2)/N + O(N^{-3/2})$, the probability that two independent uniform locations lie in the same root.

5.4 The joint law of size and censored split time

The marginal laws above arise from one common geometric relation. For each root put $T_R = h_R/N$ and $Z_R = N|R|$, identifying an untouched root with the censored time $T_R = 1$ as in (5.5).

Corollary 5.4 (joint rigidity). *If $d/N \rightarrow \eta$ through ordinary feasible sequences, the joint law obtained by choosing a root uniformly satisfies*

$$\Omega_N := \frac{1}{M} \sum_R \delta_{(T_R, Z_R)} \rightarrow \Omega_* := \frac{1}{\rho} \left(\int_{1/2}^1 \delta_{(t, 1/t)} dt + \eta \delta_{(1, 1)} \right). \quad (5.12)$$

The convergence has Wasserstein rate $O(N^{-1/2})$ at $N = s(d)$, for the sum metric on the two coordinates. Sampling by spatial mass instead gives the limit $\int_{1/2}^1 \delta_{(t, 1/t)} dt/t + \eta \delta_{(1, 1)}$, with the same rate.

$$\frac{1}{M} \sum_R |T_R Z_R - 1| \leq \frac{N}{M} (\epsilon + 2P). \quad (5.13)$$

Proof. Couple roots to the reference atoms $(j/N, N/j)$ and $(1, 1)$ exactly as in Theorem 5.2. A used slot requires no movement in time. Moving an unused slot to time 1 costs at most $1 - j/N \leq N(1/j - 1/N)$. The total coupling cost is therefore at most $2N(\epsilon + 2P)/M$, by (5.3). The reference measures converge by Riemann sums, with error $O(|d/N - \eta| + 1/N)$. Theorem 3.3 gives the stated rate. Spatial sampling weights each atom by $Z_R M/N$; the support is bounded by $0 \leq Z_R \leq 4$, so the same estimate applies to Lipschitz test functions.

$|T_R Z_R - 1| = h_R ||R| - 1/h_R| \leq N ||R| - 1/h_R|$. Summation and $P + Q \leq \epsilon + 2P$ prove (5.13). $\square$

Thus the limiting size and censored split time lie on the curve $TZ = 1$. Among split roots, counting roots gives a uniform time law, whereas sampling a spatial location gives the logarithmic law of Theorem 5.1. These statements describe near-extremizers without prescribing a refinement tree.

6 Further questions

The exact strong endpoint has bounded error, while the transfer to ordinary grids gives $s(d) - \sigma(d) = O(\sqrt{d})$. Proposition 3.4 shows that the auxiliary packing estimate alone cannot improve this scale. A sharper sequence bound must use further compatibility between subdivisions, or cancellation between positive and negative capacity defects. Appendix B exhibits a finite failure of the root constraints to encode this compatibility.

The laws in Section 5 also leave open optimal convergence rates and the fluctuations of individual root gaps. Proposition 2.3 gives finite-versus-infinite separation within the heap schedule; corresponding rigidity for more general splitting schedules is a separate question.

A Arithmetic of the strong endpoint

Put $C = 1/(1 - \log 2)$. The bounded error in $\sigma(d) = Cd + O(1)$ has a more precise arithmetic description.

Proposition A.1. For every $N \geq 1$,

$$\begin{aligned} \Phi_N(d) &= \log 2 + \frac{d + 3/2}{N} + (-1)^N h_N, \\ h_N &= \frac{1}{2} \int_0^\infty e^{-Nt} \tanh(t/2) dt, \end{aligned} \quad (\text{A.1})$$

where $1/(4N^2) - 1/(8N^4) < h_N < 1/(4N^2)$. Every increment $\sigma(d+1) - \sigma(d)$ is 3 or 4, with respective limiting frequencies $4 - C$ and $C - 3$. Outside a set of natural density zero, $\sigma(d) = \lfloor C(d + 3/2) \rfloor$. The centered endpoint $\sigma(d) - Cd$, sampled uniformly over $0 \leq d < D$, converges in distribution as $D \rightarrow \infty$ to the uniform law on $[3C/2 - 1, 3C/2]$.

Proof. Set $J_N = \int_0^1 x^{N-1}/(1+x) dx$. The recurrence $J_N + J_{N+1} = 1/N$, starting with $J_1 = \log 2$, gives (A.1) after separating the even and odd harmonic sums; $h_N = J_N - 1/(2N)$. Integrating $u - u^3/3 < \tanh u < u$ proves the two bounds.

Write $a_d = C(d + 3/2)$ and $B = \lfloor a_d \rfloor$. If B is even, the exact threshold gives $\sigma(d) = B$. If B is odd, the only alternatives are $B - 1, B, B + 1$: the first occurs when $\eta(a_d - B) < Bh_B$, the last when $\eta(B + 1 - a_d) < (B + 1)h_{B+1}$, and otherwise the endpoint is B . The two exceptional inequalities cannot both hold since $B \geq 4$ and $C/(4B) + C/(4(B + 1)) < 1$. Equality would contradict Theorem 2.1.

For the increments, set $T_N = \eta N - 3/2 - (-1)^N N h_N$, so feasibility is $d > T_N$. The bounds give $T_{N+3} - T_N \leq 3\eta + 1/(4N) + 1/(4(N + 3)) < 1$ for $N \geq 8$, and $T_{N+5} - T_{N+1} \geq 4\eta - 1/(4(N + 1)) - 1/(4(N + 5)) > 1$ for $N \geq 4$. Apply these at $N = \sigma(d)$; the case $d = 0$ is checked from $\sigma(0) = 4$, $\sigma(1) = 8$. Telescoping proves the frequencies.

An exception to the floor formula requires $\text{dist}(a_d, \mathbb{Z}) = O(1/d)$. Irrationality of C and equidistribution of its rotation therefore imply density zero of exceptions. Outside them, $\sigma(d) - Cd = 3C/2 - \{a_d\}$, proving the limiting law. $\square$

The first four exactly checked exceptions occur at $d = 12, 97, 603, 2515$: the uncorrected floors are 43, 321, 1969, 8201, while the true endpoints are 44, 320, 1970, 8200. An exception at order N requires a one-sided approximation to $\log 2$ by $(2N - 2d - 3)/(2N)$ with error below $1/(2N)^2$; reduction preserves odd numerator and even denominator.

B Limits of gap and first-split information

The capacity argument uses necessary conditions on root lengths and first-split times. The following examples identify information these conditions omit.

First, the point sets

$$\{1/10, 2/5, 51/100, 3/5, 9/10\}, \quad \{1/10, 21/100, 51/100, 3/5, 9/10\}$$

have the same complete gap multiset $\{9/100, 1/10, 1/10, 11/100, 3/10, 3/10\}$, including boundary gaps. The first meets all four uniform cells; the second misses $[1/4, 1/2)$. Thus symmetric gap statistics, including entropy and every power sum, do not determine grid coverage.

Anholcer et al. [1, Section 5] report Oliveira e Silva's computer-search result $s(1) = 31$, which already rules out any slack-one extension to order 32. Proposition B.1 shows that the separate root conditions can nevertheless be satisfied, and gives a direct nonextendability certificate for the displayed prefix.

Proposition B.1. There exists a slack-one piercing prefix of order 16 whose roots satisfy all grid-avoidance and distinct first-split-slot conditions for an order-32 extension, but which admits no such extension.

Proof. In arrival order take

$$(x_1, \dots, x_6) = (41/840, 169/714, 337/624, 61/62, 233/754, 755/1044),$$

$$(x_7, \dots, x_{12}) = (321/736, 755/1218, 433/496, 109/594, 125/1302, 419/532),$$

$$(x_{13}, \dots, x_{17}) = (251/700, 551/1150, 749/1100, 1103/1334, 929/992).$$

Direct rational comparison gives

$$\{\lfloor nx_i \rfloor : 1 \leq i \leq n+1\} = \{0, \dots, n-1\} \quad (1 \leq n \leq 16).$$

Hence this is a feasible prefix. Write $0 = y_0 < y_1 < \dots < y_{17} < y_{18} = 1$ for its sorted coordinates with boundary points. Assign the eighteen roots (y_{i-1}, y_i) , in spatial order, the deadlines

$$(20, 28, 16, 25, 19, 26, 18, 29, 24, 17, 22, 30, 21, 32, 27, 23, 31, 32).$$

A deadline $h < 32$ schedules the root's first split at the transition from stage h to stage $h+1$; deadline 32 leaves it untouched. For each root with deadline h_i , exact comparison gives

$$\lceil ny_{i-1} \rceil + 1 > ny_i \quad (1 \leq n \leq h_i).$$

Consequently no root contains a whole grid cell through its deadline. The deadline multiset is $\{16, \dots, 31, 32, 32\}$, so the sixteen required first splits occupy distinct available slots.

Nevertheless, let $G_{n,j} = [(j-1)/n, j/n)$. The following seventeen cells are pairwise disjoint and contain none of the x_i :

Grid n	Cell indices j
30	4, 5, 12, 13, 23, 26
32	3, 7, 9, 11, 15, 17, 19, 21, 23, 29, 31

An order-32 slack-one extension has only $33 - 17 = 16$ further points. Every listed cell must contain one of them, and their disjointness makes this impossible. $\square$

The obstruction uses cells from two later grids. It detects compatibility requirements absent from the separate root-avoidance and first-split-slot conditions.

Data availability

All proofs and the exact rational data needed to verify Appendix B are contained in this article. An earlier, extended version of this work is available at <https://doi.org/10.6084/m9.figshare.34383261.v1>.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests

The author declares no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Declaration of generative AI assistance

During the research and preparation of this work, the author used ChatGPT (OpenAI) and Claude (Anthropic) to assist with organization, clarity, presentation, references and formatting. ChatGPT also assisted with mathematical exploration and the development of enumeration, SAT-encoding, certificate-checking and plotting code for the broader project. Figures were generated from the formulas and data stated in the work. The author reviewed and edited the content and takes responsibility for the manuscript and its mathematical and computational claims.

References

- [1] M. Anholcer, B. Bosek, J. Grytczuk, G. Gutowski, J. Przybyło, R. Pyzik, M. Zając, On a problem of Steinhaus, *Bull. Lond. Math. Soc.* 55 (6) (2023) 2635–2652. <https://doi.org/10.1112/blms.12885>
- [2] E.R. Berlekamp, R.L. Graham, Irregularities in the distributions of finite sequences, *J. Number Theory* 2 (1970) 152–161. [https://doi.org/10.1016/0022-314X\(70\)90015-6](https://doi.org/10.1016/0022-314X(70)90015-6)
- [3] N.G. de Bruijn, P. Erdős, Sequences of points on a circle, *Proc. Sect. Sci. K. Ned. Akad. Wet. Amsterdam* 52 (1949) 14–17 = *Indag. Math.* 11 (1949) 46–49.
- [4] R.L. Graham, A note on irregularities of distribution, *Integers* 13 (2013) A53.
- [5] S.V. Konyagin, On irregularity of finite sequences, *Proc. Steklov Inst. Math.* 314 (2021) 90–95. <https://doi.org/10.1134/S0081543821040052>
- [6] K. Levy, Lower and upper bounds on irregularities of distribution, *Integers* 20 (2020) A26.
- [7] H. Steinhaus, *Sto zadań*, PWN, Warszawa, 1958; English translation: *One Hundred Problems in Elementary Mathematics*, Basic Books, New York, 1964; Dover reprint, 1979, Problems 6–7.
- [8] M. Warmus, A supplementary note on the irregularities of distributions, *J. Number Theory* 8 (3) (1976) 260–263. [https://doi.org/10.1016/0022-314X\(76\)90002-0](https://doi.org/10.1016/0022-314X(76)90002-0)