

Sharp Bounds on Davenport-Schinzel Sequences of Every Order*

Seth Pettie
University of Michigan
Ann Arbor, MI

ABSTRACT

One of the oldest unresolved problems in extremal combinatorics is to determine the maximum length of *Davenport-Schinzel sequences*, where an *order- s* DS sequence is defined to be one over an n -letter alphabet that avoids alternating subsequences of the form $a \cdots b \cdots a \cdots b \cdots$ with length $s + 2$. These sequences were introduced by Davenport and Schinzel in 1965 to model a certain problem in differential equations and have since become an indispensable tool in computational geometry and the analysis of discrete geometric structures.

Let $\lambda_s(n)$ be the extremal function for such sequences. What is λ_s asymptotically? This question has been answered satisfactorily (by Hart and Sharir, Agarwal, Sharir, and Shor, and Nivasch) when s is even or $s \leq 3$. However, since the work of Agarwal, Sharir, and Shor in the 1980s there has been a persistent gap in our understanding of the odd orders, a gap that is just as much qualitative as quantitative.

In this paper we establish the following bounds on $\lambda_s(n)$ for every order s .

$$\lambda_s(n) = \begin{cases} n & s = 1 \\ 2n - 1 & s = 2 \\ 2n\alpha(n) + O(n) & s = 3 \\ \Theta(n2^{\alpha(n)}) & s = 4 \\ \Theta(n\alpha(n)2^{\alpha(n)}) & s = 5 \\ n2^{(1+o(1))\alpha^t(n)/t!} & s \geq 6, t = \lfloor \frac{s-2}{2} \rfloor \end{cases}$$

These results refute a conjecture of Alon, Kaplan, Nivasch, Sharir, and Smorodinsky and run counter to common sense. When s is odd, λ_s behaves essentially like λ_{s-1} .

*This work is supported by NSF CAREER grant no. CCF-0746673, NSF grant no. CCF-1217338, and a grant from the US-Israel Binational Science Foundation.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SoCG'13, June 17–20, 2013, Rio de Janeiro, Brazil.

Copyright 2013 ACM 978-1-4503-2031-3/13/06 ...\$15.00.

Categories and Subject Descriptors

G.2.1 [Discrete Mathematics]: Combinatorics

General Terms

Theory

Keywords

Davenport-Schinzel sequences, lower envelopes

1. INTRODUCTION

Consider the classic problem of bounding the complexity of the lower envelope of n univariate functions f_1, \dots, f_n , each pair of which crosses at most s times. In other words, how many maximal connected intervals of the $\{f_i\}$ are contained in the graph of $f_{\min}(x) = \min\{f_1(x), \dots, f_n(x)\}$? In the absence of further information about the functions $\{f_i\}$ this problem can be completely stripped of its geometry by *transcribing* the lower envelope f_{\min} as a *Davenport-Schinzel (DS) sequence* of order s , namely, a repetition-free sequence over the alphabet $\{1, \dots, n\}$ that does not contain any alternating subsequences of the form $a \cdots b \cdots a \cdots b \cdots$ with length $s + 2$, for any $a, b \in \{1, \dots, n\}$.¹ Although Davenport and Schinzel [9] introduced this problem nearly 50 years ago, DS sequences only became well known in the computational geometry community in the 1980s [5, 24]. Since then DS sequences and lower envelopes have found a truly startling number of geometric applications, with a growing number [20, 16, 4, 6, 11, 18] that are not overtly geometric. See Sharir and Agarwal [23] for a survey of DS sequences up to 1995 and Klazar's 2002 survey [14] for a more detailed history of the problem.

Davenport and Schinzel [9] established $n^{1+o(1)}$ upper bounds on the length of DS sequences of every order s . In order to properly survey the improvements that followed [8, 25, 10, 21, 22, 15, 2, 13, 17] we must define some notation for forbidden subsequences and their extremal functions.

Sequence Notation and Terminology.

Let $|\sigma|$ be the length of a sequence $\sigma = (\sigma_i)_{1 \leq i \leq |\sigma|}$ and let $\|\sigma\|$ be the size of its alphabet $\Sigma(\sigma) = \{\sigma_i\}$. Two equal length sequences are *isomorphic* if they are the same up to a renaming of their alphabets. We say σ is a *subsequence* of

¹If the sequence corresponding to the lower envelope contained an alternating subsequence $abab \cdots$ with length $s + 2$ then the functions f_a and f_b must have crossed at least $s + 1$ times, a contradiction.

σ' , written $\sigma \succ \sigma'$, if σ can be obtained by deleting symbols from σ' . The predicate $\sigma \prec \sigma'$ asserts that σ is isomorphic to a subsequence of σ' . If $\sigma \not\prec \sigma'$ we say σ' is σ -free. If P is a set of sequences, $P \not\prec \sigma'$ holds if $\sigma \not\prec \sigma'$ for every $\sigma \in P$. The assertion that σ appears in or occurs in or is contained in σ' means either $\sigma \prec \sigma'$ or $\sigma \succ \sigma'$, which one being clear from context. The projection of a sequence σ onto $G \subseteq \Sigma(\sigma)$ is obtained by deleting all non- G symbols from σ . A sequence σ is k -sparse if whenever $\sigma_i = \sigma_j$ and $i \neq j$, then $|i - j| \geq k$. A block is a sequence of distinct symbols. If σ is understood to be partitioned into a sequence of blocks, $\llbracket \sigma \rrbracket$ is the number of blocks. The predicate $\llbracket \sigma \rrbracket = m$ asserts that σ can be partitioned into m blocks. The extremal functions for generalized Davenport-Schinzel sequences are defined as

$$\begin{aligned} \text{Ex}(\sigma, n, m) &= \max\{|S| : \sigma \not\prec S, \|S\| = n, \text{ and } \llbracket S \rrbracket \leq m\} \\ \text{Ex}(\sigma, n) &= \max\{|S| : \sigma \not\prec S, \|S\| = n, \text{ and } S \text{ is } \|\sigma\|\text{-sparse}\} \end{aligned}$$

where σ may be a single sequence or a set of sequences. The conditions “ $\llbracket S \rrbracket \leq m$ ” and “ S is $\|\sigma\|$ -sparse” guarantee that the extremal functions are finite. The extremal functions for (standard) Davenport-Schinzel sequences are defined to be $\lambda_s(n, m) = \text{Ex}(\sigma_{s+2}, n, m)$ and $\lambda_s(n) = \text{Ex}(\sigma_{s+2}, n)$, where $\sigma_{s+2} = abab \dots$ is the alternating sequence with length $s+2$. Note that the sparseness condition in the definition of $\lambda_s(n)$ only forbids immediate repetitions since $\|\sigma_{s+2}\| = 2$.

Bounds on generalized Davenport-Schinzel sequences are expressed as a function of “the” inverse-Ackermann function, though there is no universally agreed-upon definition. All definitions in the literature differ by at most a constant, which usually obviates the need for more precision. Our upper bounds refer to the following variant of Ackermann’s function.

$$\begin{aligned} a_{1,j} &= 2^j & j &\geq 1 \\ a_{i,1} &= 2 & i &\geq 2 \\ a_{i,j} &= a_{i,j-1} \cdot a_{i-1,a_{i,j-1}} & i, j &\geq 2 \end{aligned}$$

Note that in the table of $\{a_{i,j}\}$ values, the first column is constant ($a_{i,1} = 2$) and the second merely exponential ($a_{i,2} = 2^i$), so we have to look to the third column to find Ackermann-type growth. The double and single argument versions of the inverse-Ackermann function are defined to be

$$\begin{aligned} \alpha(n, m) &= \min\{i \mid a_{i,j} \geq m, \text{ where } j = \max\{\lceil n/m \rceil, 3\}\} \\ \alpha(n) &= \alpha(n, n) \end{aligned}$$

We could have defined $\alpha(n, m)$ without direct reference to Ackermann’s function. Note that $j = \log(a_{1,j})$. One may convince oneself that $j = \log^*(a_{2,j}) - O(1)$, $j = \log^{**}(a_{3,j}) - O(1)$, and in general, that $j = \log^{[i-1]}(a_{i,j}) - O(1)$, where $[i-1]$ is short for $i-1$ stars.² Up to $O(1)$ differences $\alpha(n, m)$ is $\min\{i \mid \log^{[i-1]}(m) \leq \max\{\lceil n/m \rceil, 3\}\}$. We state previous results in terms of the single argument version of α . However, they all generalize to the two-argument version by replacing $\lambda_s(n)$ with $\lambda_s(n, m)$ and $\alpha(n)$ with $\alpha(n, m)$.

A Brief History of λ_s .

After introducing the problem in 1965, Davenport and Schinzel [9] proved that $\lambda_1(n) = n$, $\lambda_2(n) = 2n - 1$, $\lambda_3(n) =$

²In general, if $f : \mathbb{N} \setminus \{0\} \rightarrow \mathbb{N}$ is a decreasing function, $f^*(m) = \min\{\ell \mid f^{(\ell)}(m) \leq 1\}$, where $f^{(0)}(m) = m$ and $f^{(\ell)}(m) = f(f^{(\ell-1)}(m))$.

$O(n \log n)$, and for all $s \geq 4$, that $\lambda_s(n) = n \cdot 2^{O(\sqrt{\log n})}$, where the leading constant in the exponent depends on s . Shortly thereafter Davenport [8] improved the bound on $\lambda_3(n)$ to $O(n \log n / \log \log n)$. In 1973 Szemerédi [25] dramatically improved the upper bounds for all $s \geq 3$, showing that $\lambda_s(n) = O(n \log^s n)$, where the leading constant depends on s .

From a purely numerical perspective Szemerédi’s bound settled the problem for all values of n one might encounter in nature (the log-star function being at most 5 for n less than $10^{19,000}$). However, the problem of quantitatively estimating $\lambda_s(n)$ has, in our view, always been a proxy for several qualitative questions: is $\lambda_s(n)$ linear or nonlinear? what is the structure of extremal sequences realizing $\lambda_s(n)$? and does it even matter what s is? In 1984 Hart and Sharir [10] answered the first question. They gave a bijection between order-3 (ababa-free) DS sequences and so-called generalized postorder path compression schemes. Although these schemes resembled the path compressions found in set-union data structures, Tarjan’s analysis [26] did not imply any non-trivial upper or lower bounds on their length. Hart and Sharir proved that such path compression schemes have length $\Theta(n\alpha(n))$, thereby settling the asymptotics of $\lambda_3(n)$. This result implied that $\lambda_s(n)$ is nonlinear for all $s \geq 3$ but it left open the possibility that $\lambda_s(n)$ could be $O(n\alpha(n))$, where s only influences the leading constant.

Improving on results of Sharir [21, 22], Agarwal, Sharir, and Shor [2] gave asymptotically tight bounds on order-4 DS sequences and reasonably tight bounds on all higher order sequences.

$$\begin{aligned} \lambda_4(n) &= \Theta(n \cdot 2^{\alpha(n)}) \\ \lambda_s(n) &\begin{cases} > n \cdot 2^{(1+o(1))\alpha^t(n)/t!} \\ < n \cdot 2^{(1+o(1))\alpha^t(n)} \end{cases} & \text{even } s \geq 6 \\ &\begin{cases} < n \cdot (\alpha(n))^{(1+o(1))\alpha^t(n)} \end{cases} & \text{odd } s \geq 5 \end{aligned}$$

where $t = \lfloor \frac{s-2}{2} \rfloor$

For even s the lower bounds are tight up to the constant in the exponent: 1 for the upper bound and $1/t!$ for the lower bound. Moreover, their lower bounds gave a qualitatively satisfying answer to the question of how extremal sequences are structured at the even orders. For odd s the gap between upper and lower bounds was wider, the base of the exponent being 2 at the lower bound and $\alpha(n)$ at the upper bound.

REMARK 1.1. *The results of Agarwal, Sharir, and Shor [2] force us to confront another question, namely, when is it safe to declare victory and call the problem closed? As Nivasch [17] observed, the “+o(1)” in the exponent necessarily hides a $\pm\Omega(\alpha^{t-1}(n))$ term if we express the bound in an “Ackermann-invariant” fashion, that is, in terms of the generic $\alpha(n)$, without specifying the precise variant of Ackermann’s function for which it is the inverse. Furthermore, under any of the definitions in the literature $\alpha(n)$ is an integer-valued function whereas $\lambda_s(n)/n$ must increase fairly smoothly with n , that is, an estimate of $\lambda_s(n)/n$ that is expressed as a primitive recursive function of any integer-valued $\alpha(n)$ must be off by at least a $2^{\Omega(\alpha^{t-1}(n))}$ factor. A reasonable definition of sharp bound (when dealing with generalized Davenport-Schinzel sequences) is an expression that cannot be improved, given $\pm\Theta(1)$ uncertainty in the defini-*

tion of $\alpha(n)$. For example, $\lambda_4(n) = \Theta(n2^{\alpha(n)})$ is sharp in this sense since the constant hidden by Θ reflects this uncertainty. In contrast, $\lambda_3(n) = \Theta(n\alpha(n))$ is not sharp in an Ackermann-invariant sense. See the tighter bounds on $\lambda_3(n)$ cited below and in Theorem 1.2.

In 2009 Nivasch [17] presented a simplified construction of even-order sequences and a better analysis of $\lambda_s(n)$ for both even and odd s . In addition, he provided a tight lower bound on the leading constant of $\lambda_3(n)$, matching an earlier upper bound of Klazar [13].

$$\lambda_s(n) = \begin{cases} 2n\alpha(n) + O(n\sqrt{\alpha(n)}) & s = 3; \text{ see [13]} \\ \Theta(n \cdot 2^{\alpha(n)}) & s = 4 \\ n \cdot 2^{(1+o(1))\alpha^t(n)/t!} & \text{even } s \geq 6 \\ \lambda_s(n) < n \cdot (\alpha(n))^{(1+o(1))\alpha^t(n)/t!} & \text{odd } s \geq 5 \end{cases}$$

$$\text{where } t = \lfloor \frac{s-2}{2} \rfloor \quad (\text{Niv})$$

This essentially closed the problem for even $s \geq 6$ (the leading constant in the exponent being precisely $1/t!$) but left the odd case open. The (Niv) bounds are actually corollaries of a more general theorem in [17] concerning the length of sequences avoiding *catenated permutations*,³ which were introduced by Klazar [12]. Define $\text{Perm}(r, s+1)$ to be the set of sequences obtained by concatenating $s+1$ permutations over an r -letter alphabet. For example, $abcd \ cbad \ badc \ abcd \ dcba \in \text{Perm}(4, 5)$. Let $\Lambda_{r,s}(n) = \text{Ex}(\text{Perm}(r, s+1), n)$ be the extremal function for $\text{Perm}(r, s+1)$ -free sequences. The “ $s+1$ ” here is chosen to highlight the parallels with order- s DS sequences. Every member of $\text{Perm}(2, s+1)$ contains an alternating sequence $abab \dots$ with length $s+2$,⁴ so order- s DS sequences are also $\text{Perm}(2, s+1)$ -free, implying that $\lambda_s(n) \leq \Lambda_{2,s}(n)$.

Alon, Kaplan, Nivasch, Sharir, and Smorodinsky [3, 17] conjectured that the upper bounds (Niv) for odd orders are tight, that is, the base of the exponent is, in fact, $\alpha(n)$. This conjecture was spurred by their discovery of similar functions that arose in an apparently unrelated combinatorial problem.

New Results.

We give new bounds on the length of Davenport-Schinzel sequences and in the process refute conjectures due to Alon et al. [3, 17] and Pettie [19].

THEOREM 1.2. *Let $\lambda_s(n)$ be the maximum length of a repetition-free sequence over an n -letter alphabet avoiding subsequences isomorphic to $abab \dots$ (length $s+2$). Then λ_s satisfies:*

$$\lambda_s(n) = \begin{cases} n & s = 1 \\ 2n - 1 & s = 2 \\ 2n\alpha(n) + O(n) & s = 3 \\ \Theta(n2^{\alpha(n)}) & s = 4 \\ \Theta(n\alpha(n)2^{\alpha(n)}) & s = 5 \\ n2^{(1+o(1))\alpha^t(n)/t!} & s \geq 6, t = \lfloor \frac{s-2}{2} \rfloor \end{cases}$$

³Nivasch called these formation-free sequences.

⁴The first permutation contributes 2 symbols and every subsequent permutation contributes at least one.

Theorem 1.2 is optimal in that it provides the sharpest bounds on $\lambda_s(n)$ that can be expressed in an Ackermann-invariant fashion (see Remark 1.1), and in this sense closes the Davenport-Schinzel problem. (For $s \geq 6$, the exponent is actually the Ackermann-invariant expression $\alpha^t(n)/t! + O(\alpha^{t-1}(n))$.) However, we believe our primary contributions are not the tight asymptotic bounds *per se* but the structural differences they reveal between even and odd s . We can now give a cogent explanation for *why* odd orders $s \geq 5$ behave essentially like the preceding even orders and yet why they are intrinsically more difficult to understand.

To what extent can Theorem 1.2 be extended to larger classes of generalized Davenport-Schinzel sequences? Surprisingly little, it turns out. We are able to extend the techniques to *double* Davenport-Schinzel sequences. Define $\lambda_s^{\text{dbl}}(n)$ to be the extremal function of $\text{dbl}(\sigma_{s+2})$ -free sequences, where $\sigma_{s+2} = abab \dots$ is the alternating sequence with length $s+2$ and $\text{dbl}(\sigma)$ is obtained by doubling every symbol in σ save the first and last. For example, $\text{dbl}(abab) = abbaab$.⁵ We can show that $\lambda_s^{\text{dbl}}(n)$ obeys all the upper and lower bounds of Theorem 1.2, except at $s=5$, where the upper bound is weaker by an $\alpha^{1+o(1)}(n)$ factor.

The situation becomes stranger when we attempt to extend the bounds of Theorem 1.2 to $\text{Perm}(r, s+1)$ -free sequences. When $s=3$ or s is even it was known that $\Lambda_{r,s}(n)$ behaves like $\lambda_s(n)$ for any $r \geq 2$. (Nivasch’s upper bounds [17] match the lower bounds of [10] and [2] for these parameters.) We prove that $\Lambda_{r,s}(n)$ behaves like $\lambda_s(n)$ *only if* $s=3$, or s is even, or $r=2$. For all odd $s \geq 5$ and $r \geq 3$ we give a new lower bound construction showing that Nivasch’s upper bound is essentially tight.

THEOREM 1.3. *Let $\Lambda_{r,s}(n)$ be the maximum length of an r -sparse, $\text{Perm}(r, s+1)$ -free sequence over an n -letter alphabet. For all $s \geq 1$, $t = \lfloor \frac{s-2}{2} \rfloor$, and all $r \geq 2$, we have*

$$\Lambda_{r,s}(n) = \begin{cases} \Theta(n) & s \in \{1, 2\}, r \geq 2 \\ \Theta(n\alpha(n)) & s = 3, r \geq 2 \\ \Theta(n2^{\alpha(n)}) & s = 4, r \geq 2 \\ \Omega(n\alpha(n)2^{\alpha(n)}) \text{ and } O(n\alpha^{2+o(1)}(n)2^{\alpha(n)}) & s = 5, r = 2 \\ n \cdot (\alpha(n))^{(1+o(1))\alpha(n)} & s = 5, r \geq 3 \\ n \cdot 2^{(1+o(1))\alpha^t(n)/t!} & \text{even } s \geq 6, r \geq 2 \\ n \cdot 2^{(1+o(1))\alpha^t(n)/t!} & \text{odd } s \geq 7, r = 2 \\ n \cdot (\alpha(n))^{(1+o(1))\alpha^t(n)/t!} & \text{odd } s \geq 7, r \geq 3 \end{cases}$$

Theorem 1.3 is rather surprising, even in retrospect and even given Theorem 1.2. Refer to [7] for implications of Theorem 1.3 on the size of sets of permutations with fixed VC-dimension.

Organization.

In Section 2 we introduce notation used throughout the paper and review Nivasch’s recurrence [17], which captures even-order DS sequences well. In Section 3 we discuss the

⁵Why not consider higher multiplicities? It is fairly easy to show that repeating symbols more than twice, or repeating the first and last at all, affects the extremal function by at most a constant factor. See Adamec, Klazar, and Valtr [1].

difficulties of analyzing the odd orders. Our approach is to view DS sequences not just as 1-dimensional objects but in terms of a hierarchical decomposition called a *derivation tree*. Section 4 defines the derivation tree and identifies useful structural properties. In Section 5 we present a new recurrence for odd orders. The recurrences of Sections 2 and 5 ultimately lead to the upper bounds of Theorem 1.2, with the exception of order $s = 5$. The matching lower and upper bound on λ_s are omitted from this extended abstract.

2. BASIC UPPER BOUNDS

2.1 Sequence Decomposition

Let S be a sequence over an $n = \|S\|$ letter alphabet consisting of $m = \llbracket S \rrbracket$ blocks. Suppose we partition S into \hat{m} intervals of consecutive blocks $S_1 S_2 \cdots S_{\hat{m}}$, where $m_q = \llbracket S_q \rrbracket$ is the number of blocks in interval q . Let \check{S}_q be the alphabet of symbols *local* to S_q (that do not appear in any S_p , $p \neq q$) and let $\hat{S} = \Sigma(S) \cup \bigcup_q \check{S}_q$ be the alphabet of all other *global* symbols. The cardinalities of \check{S}_q and \hat{S} are \tilde{n}_q and \hat{n} , thus $n = \hat{n} + \sum_{q=1}^{\hat{m}} \tilde{n}_q$. A global symbol in S_q is called *first*, *last*, or *middle* if it appears in no earlier interval, no later interval, or appears in both earlier and later intervals, respectively. Let $\check{S}_q, \hat{S}_q, \check{S}_q, \check{S}_q$ be the subset of $\Sigma(S_q)$ consisting of, respectively, first, last, middle, and all global symbols, and let $\hat{n}_q, \hat{n}_q, \tilde{n}_q$, and \hat{n}_q be their cardinalities. Let $\check{S}_q, \hat{S}_q, \check{S}_q, \check{S}_q, \bar{S}_q$ be the projection of S_q onto $\check{S}_q, \hat{S}_q, \check{S}_q, \check{S}_q$, and \bar{S}_q . Note that \hat{S}_1 consists solely of first occurrences; if the last occurrence of a symbol appeared in \hat{S}_1 the symbol would be classified as local to S_1 , not global. The same argument shows that $\check{S}_{\hat{m}}$ consists solely of last occurrences. Let $\check{S}, \hat{S}, \check{S}, \check{S}$, and \bar{S} be the subsequences of local, global, first, last, and middle occurrences, respectively, that is, $\check{S} = \check{S}_1 \cdots \check{S}_{\hat{m}}$, $\hat{S} = \hat{S}_1 \cdots \hat{S}_{\hat{m}}$, $\check{S} = \check{S}_1 \cdots \check{S}_{\hat{m}-1}$, $\check{S} = \check{S}_2 \cdots \check{S}_{\hat{m}}$, and $\bar{S} = \bar{S}_2 \cdots \bar{S}_{\hat{m}-1}$, the last of which would be empty if $\hat{m} = 2$. Let $\hat{S}' = \hat{S}_1 \cdots \hat{S}_{\hat{m}}$ be an \hat{m} -block sequence obtained from \hat{S} by replacing each \hat{S}_q with a single block containing its alphabet \hat{S}_q , listed in order of *first* appearance in \hat{S}_q .

2.2 Nivasch's Recurrence

We can reduce any m -block sequence to a 2-sparse one by removing $m-1$ duplicate symbols at block boundaries, hence $\lambda_s(n, m) \leq \lambda_s(n) + m - 1$ for any order s . Lemma 2.1 follows from this observation and the known bounds on $\lambda_1(n)$ and $\lambda_2(n)$.

LEMMA 2.1. $\lambda_1(n, m) = n + m - 1$ and, for $m \geq 2$, $\lambda_2(n, m) = 2n + m - 2$.

Nivasch's [17] upper bounds (Niv) are a consequence of a recurrence for λ_s that builds on that of Agarwal, Sharir, and Shor [2]. Here we present a streamlined version of Nivasch's recurrence.

RECURRENCE 2.2. *Let m, n , and $s \geq 3$ be the block count, alphabet size, and order parameters. For any block partition $\{m_q\}_{1 \leq q \leq \hat{m}}$ and alphabet partition $\{\hat{n}\} \cup \{\tilde{n}_q\}_{1 \leq q \leq \hat{m}}$, where $m = \sum_q m_q$ and $n = \hat{n} + \sum_q \tilde{n}_q$, we have $\lambda_s(n, m) \leq \sum_{q=1}^{\hat{m}} \lambda_s(\tilde{n}_q, m_q) + 2 \cdot \lambda_{s-1}(\hat{n}, m) + \lambda_{s-2}(\lambda_s(\hat{n}, \hat{m}) - 2\hat{n}, m)$.*

PROOF. We adopt the notation and definitions from Section 2.1, where S is an extremal order- s DS sequence with

$\|S\| = n$ and $\llbracket S \rrbracket = m$. We shall bound $|S|$ by considering its four constituent subsequences $\check{S}, \hat{S}, \check{S}$, and \bar{S} .

Each \check{S}_q is an order- s DS sequence, therefore the contribution of local symbols is $|\check{S}| \leq \sum_{q=1}^{\hat{m}} \lambda_s(\tilde{n}_q, m_q)$. We claim each \hat{S}_q is an order- $(s-1)$ DS sequence. By virtue of being categorized as *first* in \hat{S}_q , every symbol in \hat{S}_q appears at least once after \hat{S}_q . Therefore an occurrence of an alternating sequence $\sigma_{s+1} = abab \cdots$ (length $s+1$), in \hat{S}_q would imply an occurrence of σ_{s+2} in S , a contradiction. By symmetry it also follows that \check{S}_q is an order- $(s-1)$ DS sequence, hence $|\hat{S}'| + |\check{S}'| = \sum_{q=1}^{\hat{m}-1} \lambda_{s-1}(\hat{n}_q, m_q) + \sum_{q=2}^{\hat{m}} \lambda_{s-1}(\hat{n}_q, m_q) \leq \lambda_{s-1}(\hat{n}, m - m_{\hat{m}}) + \lambda_{s-1}(\hat{n}, m - m_1) < 2 \cdot \lambda_{s-1}(\hat{n}, m)$. Note that $\sum_q \hat{n}_q = \hat{n}$ and $\sum_q \tilde{n}_q = \hat{n}$ as each sum counts each global symbol exactly once. Furthermore, λ_s is clearly superadditive.⁶

The same argument shows that \bar{S}_q is an order- $(s-2)$ DS sequence. Symbols in \bar{S}_q were categorized as *middle*, so an alternating subsequence $\sigma_s = baba \cdots$ (length s) in \bar{S}_q , together with an a preceding \bar{S}_q and either an a or b following \bar{S}_q (depending on whether s is even or odd), yields an instance of σ_{s+2} in S , a contradiction. Thus the contribution of middle symbols is

$$\begin{aligned} |\bar{S}| &\leq \sum_{q=2}^{\hat{m}-1} \lambda_{s-2}(\tilde{n}_q, m_q) \\ &\leq \lambda_{s-2} \left(\sum_{q=2}^{\hat{m}-1} \tilde{n}_q, m - m_1 - m_{\hat{m}} \right) \\ &\leq \lambda_{s-2}(|\hat{S}'| - 2\hat{n}, m - m_1 - m_{\hat{m}}) \quad (1) \\ &\leq \lambda_{s-2}(\lambda_s(\hat{n}, \hat{m}) - 2\hat{n}, m) \quad (2) \end{aligned}$$

Inequality (1) follows from the fact that $\sum_q \tilde{n}_q$ counts the length of \hat{S}' , save the first and last occurrence of each global symbol, that is, $2\hat{n}$ occurrences in total. Since \hat{S}' is a subsequence of S , it too is an order- s DS sequence, so $|\hat{S}'| \leq \lambda_s(\hat{n}, \hat{m})$. Inequality (2) follows. \square

Recurrence 2.2 only offers us the freedom to choose the block partition $\{m_q\}_{1 \leq q \leq \hat{m}}$ and this is where Ackermann's function comes into play. When we invoke Recurrence 2.2 it is with respect to a constant $c \geq s-2$ and an index $i \geq 1$. Let j be minimum such that $m \leq a_{i,j}^c$ and let $w = a_{i,j-1}$. When $i, j > 1$ we choose $\hat{m} = \lceil m/w^c \rceil$ and partition the given sequence into intervals of w^c blocks, so $m_q = w^c$ for $q \leq \lceil m/w^c \rceil - 1$ and the leftover $m_{\hat{m}}$ may be smaller. We call this a *uniform block partition* with width w^c . The intention is that $\lambda_s(\tilde{n}_q, m_q)$ will be bounded by invoking Recurrence 2.2 with parameters c, i , and $j-1$, $\lambda_{s-1}(\hat{n}, m)$ with parameters c, i , and j , $\lambda_s(\hat{n}, \hat{m})$ with parameters $c, i-1$, and w ,⁷ and $\lambda_{s-2}(\lambda_s(\hat{n}, \hat{m}) - 2\hat{n}, m)$ with parameters c, i , and j . In our analysis of the odd orders it is important that the sequence decomposition (determined by the choice of \hat{m}) be independent of s , hence our introduction of the parameter c . In the base cases we apply Lemma 2.1 when $s \leq 2$ and Lemma 2.3 when $i = 1$ or $j = 1$. Lemma 2.3 implies a weak bound of $\lambda_s(n, m) = O(n + m \log^{s-2} m)$.

⁶That is, $\lambda_s(n', m') + \lambda_s(n'', m'') \leq \lambda_s(n' + n'', m' + m'')$ for all n', n'', m', m'' .

⁷Note that $\hat{m} = \lceil m/w^c \rceil = \lceil m/a_{i,j-1}^c \rceil \leq (a_{i,j}/a_{i,j-1})^c = a_{i-1,w}^c$, so the "j" parameter for index $i-1$ is w .

LEMMA 2.3. Fix $s \geq 2$ and $c \geq \max\{s-2, 1\}$. Whether $i = 1$ and $j \geq 1$ or $j = 1$ and $i > 1$, if $m \leq a_{i,j}^c$ then

$$\lambda_s(n, m) \leq 2^{s-1}n + (cj)^{s-2}(m-1).$$

PROOF. Fix $c = 1$. We will prove that for $m \leq a_{1,j}$, $\lambda_s(n, m) \leq 2^{s-1}n + j^{s-2}(m-1)$. This will establish the lemma since for $c > 1$, $a_{1,j}^c = (2^j)^c = a_{1,cj}$ and $a_{i,1}^c = 2^c = a_{1,c}$.

At $s = 2$ the claim follows directly from Lemma 2.1. At $s \geq 3, j = 1$, the claim is trivial since there are only $2 = a_{1,1}$ blocks and $\lambda_s(n, 2) = 2n$. In the general case we have $s \geq 3$ and $j > 1$. Let S be an order- s , m -block sequence over an n -letter alphabet, where $m \leq a_{1,j}$. Let $\hat{m} = 2$ and $S = S_1S_2$ be the uniform block partition with width $a_{1,j-1} = 2^{j-1}$, so $\llbracket S_1 \rrbracket = a_{1,j-1}$ and $\llbracket S_2 \rrbracket = m - a_{1,j-1}$. Note that $\hat{S}' = \beta_1\beta_2$ consists of two blocks, where each β_q is some permutation of the global alphabet $\hat{\Sigma}$. Since there are no middle occurrences in S we can apply a simplified version of Recurrence 2.2.

$$\begin{aligned} \lambda_s(n, m) &\leq \sum_{q=1,2} \lambda_s(\hat{n}_q, \llbracket S_q \rrbracket) + \lambda_{s-1}(\hat{n}, \llbracket S_1 \rrbracket) + \lambda_{s-1}(\hat{n}, \llbracket S_2 \rrbracket) \\ &\hspace{15em} \{\text{local, first, and last}\} \\ &\leq 2^{s-1}(n - \hat{n}) + (j-1)^{s-2}(m-2) + 2(2^{s-2}\hat{n}) \\ &\quad + (j-1)^{s-3}(m-2) \quad \{\text{Inductive hypothesis}\} \\ &< 2^{s-1}n + j^{s-2}(m-1) \end{aligned}$$

The last inequality follows from the fact that $(j-1)^{s-2} + (j-1)^{s-3} \leq j^{s-2}$. This concludes the induction. \square

3. DIFFICULTIES AT THE ODD ORDERS

The proof of Recurrence 2.2 is indifferent to the parity of s . However, when analyzing the contribution of middle symbols we hinted at some structural differences between even and odd s . It was claimed that \bar{S}_q is an order- $(s-2)$ DS sequence. If \bar{S}_q contained an alternating subsequence $\sigma_s = baba \dots$ with length s then there must be an occurrence of σ_{s+2} in S because both a and b appear both before and after S_q in S . The longest permissible alternating subsequence $\sigma_{s-1} = baba \dots$ in \bar{S}_q therefore has length $s-1$. When s is even, $s-1$ is odd, meaning σ_{s-1} begins and ends with b , whereas when s is odd, σ_{s-1} begins with b and ends with a . Since we can only afford to introduce *one* alternation between a and b at each boundary of S_q , the pattern of as and bs on either side of S_q in S (and therefore on either side of the q th block of \hat{S}') must be of the form

$$\begin{array}{ll} a^* b^* S_q b^* a^* & \text{when } s \text{ is even and} \\ a^* b^* S_q a^* b^* & \text{when } s \text{ is odd.} \end{array}$$

In the first case a and b are *nested* and in the second case they are *interleaved* (w.r.t. q). If the (Niv) bounds prove to be tight there must be two systems for generating sequences, one where nesting is the norm, and one where interleaving is the norm. By *norm* we mean nesting/interleaving is sufficiently plentiful such that \bar{S}_q can be close to an extremal order- $(s-2)$ DS sequence.

Entertain the idea that interleaving is forbidden, that is, every pair of middle symbols in a block of \hat{S}' are nested. When s is odd, \bar{S}_q must be an order- $(s-3)$ DS sequence. If \bar{S}_q were to contain an even-length alternating subsequence

$\sigma_{s-1} = ba \dots ba$ with length $s-1$, then S would contain either

$$ab \mid ba \dots ba \mid ba \quad \text{or} \quad ba \mid ba \dots ba \mid ab$$

where the portion between the bars comes from \bar{S}_q , that is, S would contain σ_{s+2} , a contradiction.

We cannot simply assume interleaving does not exist, but we can attempt to bound its prevalence. Measuring *prevalence* is tricky since nesting/interleaving is a non-transitive pairwise relation but to argue that \bar{S}_q has order- $(s-3)$ we need *every* pair of its symbols to be nested. Our approach is to fully unroll the induction implicit in Recurrence 2.2. The result is what we call a *derivation tree* \mathcal{T} , the nodes of which correspond to blocks encountered in the decomposition of S . Whereas S 's blocks occupy the leaves of \mathcal{T} , derived sequences like \hat{S}' occupy levels higher in \mathcal{T} . Fix a symbol $a \in \Sigma(S)$. The *projection* of \mathcal{T} onto a , call it $\mathcal{T}_{|a}$, is obtained by deleting from \mathcal{T} all nodes whose blocks do not contain a .

Suppose we are interested in the nestedness of middle symbols a, b in some block β in S (a leaf in \mathcal{T}). The projection trees $\mathcal{T}_{|a}$ and $\mathcal{T}_{|b}$ both contain node β . Some nodes in projection trees are intrinsically bad; these are called *feathers* in Section 4. Whether a node in $\mathcal{T}_{|a}$ is a feather depends solely on its position in $\mathcal{T}_{|a}$, not how $\mathcal{T}_{|a}$ is embedded in \mathcal{T} . Our main structural lemma states that if β is not a feather in $\mathcal{T}_{|a}$ and not a feather in $\mathcal{T}_{|b}$, then a and b are nested with respect to β . In Section 5 we use this characterization to state two new recurrences: one on $\Phi_s(n, m)$, the total number of feathers in an order- s DS sequence, and one on $\lambda_s(n, m)$ for odd $s \geq 5$.

4. THE DERIVATION TREE

The *derivation tree* $\mathcal{T} = \mathcal{T}(S, c, i)$ of a sequence S is an ordered, rooted binary tree whose nodes are identified with the blocks encountered in recursively decomposing S , as in Lemma 2.3 and the remarks following Recurrence 2.2. The sequence alone does not determine \mathcal{T} , hence the need to define it with respect to i and c . Define j to be minimum such that $\llbracket S \rrbracket \leq a_{i,j}^c$.

Let $\mathcal{B}(v)$ denote the block associated with a node $v \in \mathcal{T}$. The p th block of S is always identified with the p th leaf of \mathcal{T} in left-to-right order. As we are sometimes indifferent to the order of symbols within a block, $\mathcal{B}(v)$ is often treated as a set. We adopt the notation from Section 2.1 and the proof of Recurrence 2.2.

Base Cases.

When $i \in \{0, 1\}$ the construction of $\mathcal{T}(S, c, i)$ is independent of c .

At $i = 0$ we only consider a 2-block sequence $S = \beta_1\beta_2$, where each block contains the whole alphabet $\Sigma(S)$. The tree $\mathcal{T}(S, c, 0)$ consists of three nodes u, u_1, u_2 , where u is the parent of u_1 and u_2 , $\mathcal{B}(u)$ is empty, $\mathcal{B}(u_1) = \beta_1$, and $\mathcal{B}(u_2) = \beta_2$. For every $a \in \Sigma(S)$ call u its *crown* and u_1 and u_2 its left and right *heads*, respectively. These nodes are denoted $\text{cr}_{|a}$, $\text{lhe}_{|a}$, and $\text{rhe}_{|a}$.

At $i = 1$ we let j' be minimum such that $\llbracket S \rrbracket \leq a_{1,j'}$. If $j' = 1$ ($\llbracket S \rrbracket = 2$) then $\mathcal{T}(S, c, 1) = \mathcal{T}(S, c, 0)$, otherwise take $S = S_1S_2$ to be the uniform block partition with width $a_{1,j'-1}$. Let $\hat{\mathcal{T}} = \mathcal{T}(\hat{S}', c, 0)$ be the three-node derivation tree for \hat{S}' and let $\tilde{\mathcal{T}}_1, \tilde{\mathcal{T}}_2$ be the local trees, where $\tilde{\mathcal{T}}_q =$

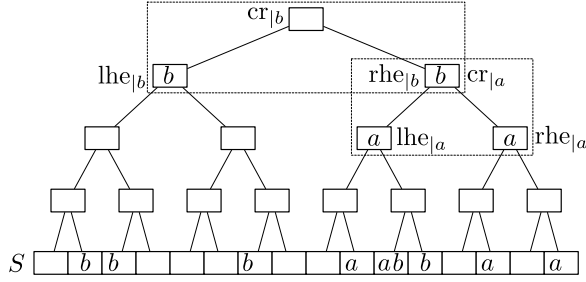


Figure 1: A derivation tree $\mathcal{T}(S, c, 1)$ for a 16-block sequence S with $i = 1$. The dashed boxes isolate the derivation trees with $i = 0$ that assign the crown and heads for symbols $a, b \in \Sigma(S)$.

$\mathcal{T}(\tilde{S}_q, c, 1)$. The tree $\mathcal{T} = \mathcal{T}(S, c, 1)$ is formed by identifying the root of $\tilde{\mathcal{T}}_q$ with the q th leaf of $\tilde{\mathcal{T}}$, then placing the blocks of S at the leaves of \mathcal{T} . This last step is necessary since only local symbols appear in $\tilde{\mathcal{T}}_1$ and $\tilde{\mathcal{T}}_2$ whereas the leaves of \mathcal{T} must be identified with the blocks of S , which include both local and global symbols. See Figure 1 for an illustration.

The case $i > 1, j = 1$ is handled just like the $i = 1$ case, where the derivation tree $\mathcal{T}(S, c, i)$ will have height $\log(a_{i,1}^c) = \log(a_{1,c}) = c$.

Inductive Case.

The construction when $i, j > 1$ is similar to the $i = 1$ case except we take the c parameter into account. Recall that j is minimum such that $\llbracket S \rrbracket \leq a_{i,j}^c$ and $w = a_{i,j-1}$. The tree $\mathcal{T}(S, c, i)$ is formed by composing $\hat{\mathcal{T}} = \mathcal{T}(\hat{S}', c, i-1)$ and $\tilde{\mathcal{T}}_1, \dots, \tilde{\mathcal{T}}_{\hat{m}}$, where $\tilde{\mathcal{T}}_q = \mathcal{T}(\tilde{S}_q, c, i)$, $\hat{m} = \lceil \llbracket S \rrbracket / w^c \rceil$, and $\llbracket \tilde{S}_q \rrbracket = w^c$ if $q < \hat{m}$. To be more specific, we identify the root of $\tilde{\mathcal{T}}_q$ (whose block is empty) with the q th leaf of $\hat{\mathcal{T}}$, then assign the blocks of S to the leaves of \mathcal{T} . The crown and heads of each symbol $a \in \Sigma(S)$ are inherited from $\hat{\mathcal{T}}$, if a is global, or some $\tilde{\mathcal{T}}_q$ if a is local to S_q .

4.1 Anatomy of the Tree

The *projection* of \mathcal{T} onto $a \in \Sigma(S)$, denoted $\mathcal{T}|_a$, is the tree on the node set $\{\text{cr}|_a\} \cup \{v \in \mathcal{T} | a \in \mathcal{B}(v)\}$ that inherits the ancestor/descendant relation from \mathcal{T} , that is, the parent of v in $\mathcal{T}|_a$, where $v \notin \{\text{cr}|_a, \text{lhc}|_a, \text{rhc}|_a\}$, is v 's nearest strict ancestor u for which $a \in \mathcal{B}(u)$.

DEFINITION 4.1. (*Anatomy*)

- The leftmost and rightmost leaves of $\mathcal{T}|_a$ are wingtips, denoted $\text{lwt}|_a$ and $\text{rwt}|_a$.
- The left and right wings are those paths in $\mathcal{T}|_a$ extending from $\text{lhc}|_a$ to $\text{lwt}|_a$ and from $\text{rhc}|_a$ to $\text{rwt}|_a$.
- Descendants of $\text{lhc}|_a$ and $\text{rhc}|_a$ in $\mathcal{T}|_a$ are called doves and hawks, respectively.
- A child of a wing node that is not itself on the wing is called a quill.
- A leaf is called a feather if it is the rightmost descendant of a dove quill or leftmost descendant of a hawk quill.

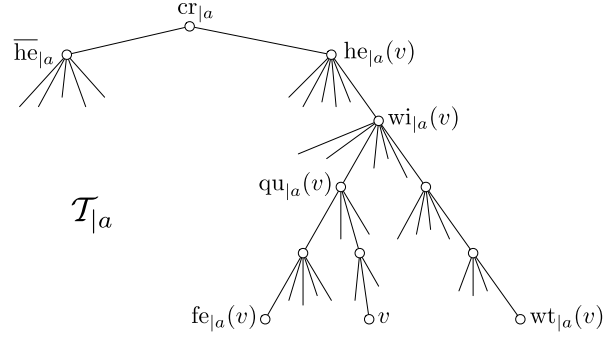


Figure 2: In this example v is a hawk leaf in $\mathcal{T}|_a$ since its head $\text{hc}|_a(v) = \text{rhc}|_a$ is the right child of $\text{cr}|_a$. Its wing node $\text{wi}|_a(v)$, wingtip $\text{wt}|_a(v)$, quill $\text{qu}|_a(v)$, and feather $\text{fe}|_a(v)$ are indicated.

- Suppose v is a node in $\mathcal{T}|_a$. Let $\text{hc}|_a(v)$ be the head ancestral to v and $\overline{\text{hc}}|_a(v)$ be the other head; let $\text{wt}|_a(v)$ and $\overline{\text{wt}}|_a(v)$ be the wingtips descending from $\text{hc}|_a(v)$ and $\overline{\text{hc}}|_a(v)$. Let $\text{wi}|_a(v)$ be the nearest wing node ancestor of v , $\text{qu}|_a(v)$ the quill ancestral to v , and $\text{fe}|_a(v)$ the feather descending from $\text{qu}|_a(v)$. See Figure 2 for an illustration.

Once $a \in \Sigma(S)$ is known or specified, we will use these terms (feather, wingtip, etc.) to refer to nodes in $\mathcal{T}|_a$, to their associated blocks, or to the occurrences of a within those blocks. For example, an occurrence of a in S would be a feather if it appears in a block $\mathcal{B}(v)$ in S , where v is a feather in $\mathcal{T}|_a$.

Note that the nodes $\text{hc}|_a(v)$, $\text{wi}|_a(v)$, $\text{qu}|_a(v)$, $\text{wt}|_a(v)$, and $\text{fe}|_a(v)$ are not necessarily distinct. It may be that $\text{hc}|_a(v) = \text{wi}|_a$, and it may be that $v = \text{qu}|_a(v) = \text{fe}|_a(v)$ if v 's parent in $\mathcal{T}|_a$ is $\text{wi}|_a(v)$.

Lemma 4.2 identifies one property of \mathcal{T} used in the proof of Lemma 4.3.

LEMMA 4.2. Suppose that on a leaf-to-root path in \mathcal{T} we encounter nodes u, v, x , and y (the last two possibly identical), where $u, x \in \mathcal{T}|_a$ and $v, y \in \mathcal{T}|_b$. It must be that $a \in \mathcal{B}(v)$ and therefore $v \in \mathcal{T}|_a$.

PROOF. Consider the decomposition of \mathcal{T} into a global derivation tree $\hat{\mathcal{T}}$ and local derivation trees $\{\tilde{\mathcal{T}}_q\}$. If v were an internal node in some $\tilde{\mathcal{T}}_q$ then b would be classified as local. This implies $y \in \tilde{\mathcal{T}}_q$ as well and the claim follows by induction on the construction of $\tilde{\mathcal{T}}_q$. If v were an internal node in $\hat{\mathcal{T}}$ then let u' be the leaf of $\hat{\mathcal{T}}$ ancestral to u . The nodes $u', v, x, y \in \hat{\mathcal{T}}$ also satisfy the criteria of the lemma; the claim follows by induction on the construction of $\hat{\mathcal{T}}$. Thus, we can assume u is a leaf of \mathcal{T} and v is a leaf of $\tilde{\mathcal{T}}$. By construction all global symbols in $\mathcal{B}(u)$ also appear in $\mathcal{B}(v)$. Since $x \in \hat{\mathcal{T}}$, the symbol a is classified as global and must appear in $\mathcal{B}(v)$. \square

4.2 Habitual Nesting

Suppose a block β in S contains two symbols a, b which make neither their first nor last appearance in β . We call a and b nested in β if S contains either $ab\beta ba$ or $ba\beta ab$

and call them *interleaved in β* otherwise, that is, if the occurrences of a and b in S take the form $a^*b^*\beta a^*b^*$ or $b^*a^*\beta b^*a^*$. Lemma 4.3 is the critical structural lemma used in our analysis. It provides us with simple criteria for nestedness.

LEMMA 4.3. *Suppose that $v \in \mathcal{T}$ is a leaf and a, b are symbols in $\mathcal{B}(v)$. If the following two criteria are satisfied then a and b are nested in $\mathcal{B}(v)$.*

- i. v is not a wingtip in either \mathcal{T}_a or \mathcal{T}_b .
- ii. v is not a feather in either \mathcal{T}_a or \mathcal{T}_b .

PROOF. Without loss of generality we can assert two additional criteria.

- iii. $cr_{|b}$ is equal to or strictly ancestral to $cr_{|a}$.
- iv. v is a dove in \mathcal{T}_a .

By Criterion (iv) the leftmost leaf descendant of $w_{|a}(v)$ is $wt_{|a}(v)$. Let u be its rightmost leaf descendant. According to Criteria (i,ii) v is distinct from both $wt_{|a}(v)$ and u since u must be a feather. We partition the sequence outside of $\mathcal{B}(v)$ into the following four intervals.

- I_1 : everything preceding the a in $\mathcal{B}(wt_{|a}(v))$,
- I_2 : everything from the end of I_1 to $\mathcal{B}(v)$,
- I_3 : everything from $\mathcal{B}(v)$ to the a in $\mathcal{B}(u)$, and
- I_4 : everything following I_3 .

By Criterion (i) there must be occurrences of b both before and after $\mathcal{B}(v)$. If, contrary to the claim, a and b are not nested in $\mathcal{B}(v)$, all other occurrences of b must appear exclusively in I_1 and I_3 or exclusively in I_2 and I_4 . We show that both possibilities lead to contradictions. Figures 3 and 4 may help illuminate the proof.

Case 1: b does not appear in I_1 or I_3 .

According to Criterion (i) the left wingtip $lwt_{|b}$ of \mathcal{T}_b is distinct from v , and therefore appears in interval I_2 . Since $lwt_{|b}$ and v are descendants of $w_{|a}(v)$, which is a *strict* descendant of $cr_{|a}$, which, by Criterion (iii), is a descendant of $cr_{|b}$, it must also be that $lwt_{|b}$ and v descend from the same child of $cr_{|b}$, that is,

- iv. v is a dove in \mathcal{T}_b and therefore $wt_{|b}(v) = lwt_{|b}$.

We shall argue below that

- v. In \mathcal{T} , $qu_{|b}(v)$ is a strict descendant of $w_{|a}(v)$ and a strict ancestor of u , and $fe_{|b}(v)$ lies in interval I_4 .

The least common ancestor of v and $wt_{|b}(v)$ in \mathcal{T}_b is by definition $w_{|b}(v)$. The quill $qu_{|b}(v)$ is a child of $w_{|b}(v)$ not on a wing, hence $qu_{|b}(v)$ cannot be ancestral to $wt_{|b}(v)$, and hence $qu_{|b}(v)$ must be a strict descendant of $w_{|a}(v)$. By Criterion (ii) and Inference (iv), $fe_{|b}(v)$ is the rightmost leaf descendant of $qu_{|b}(v)$ and distinct from v . However, by supposition I_3 contains no occurrences of b , so $fe_{|b}(v)$ must lie in interval I_4 . For $qu_{|b}(v)$ to have descendants in both I_2 and I_4 it must be a strict ancestor of u in \mathcal{T} . As we explain below, a consequence of Inference (v) is that

- vi. $\overline{wt}_{|a}(v)$ lies to the right of $fe_{|b}(v)$.

According to Inference (v) $qu_{|b}(v)$ is a descendant of $w_{|a}(v)$, which is a descendant of $he_{|a}(v)$. According to Criterion (iv) $he_{|a}(v)$ is the *left* head of \mathcal{T}_a . Since $\overline{wt}_{|a}(v)$ is a descendant of $he_{|a}(v)$, the right sibling of $he_{|a}(v)$, $\overline{wt}_{|a}(v)$ must lie to the right of $fe_{|b}(v)$.

Let us review the situation. Scanning the leaves from left to right we see the blocks $wt_{|a}(v)$, $wt_{|b}(v)$, v , u , $fe_{|b}(v)$, and $\overline{wt}_{|a}(v)$. It may be that $wt_{|a}(v)$ and $wt_{|b}(v)$ are equal and it may be that u and $fe_{|b}(v)$ are equal. If either of these cases hold then the a precedes the b in the given block. The blocks $wt_{|a}(v)$, $wt_{|b}(v)$, v , $fe_{|b}(v)$, $\overline{wt}_{|a}(v)$ certify that a and b are nested in $\mathcal{B}(v)$.

Case 2: b does not appear in I_2 or I_4 .

By Criterion (i) the right wingtip $rwt_{|b}$ is distinct from v and must therefore lie in I_3 . Following the same reasoning from Case 1 we can deduce that

- vii. v is a hawk in \mathcal{T}_b .
- viii. In \mathcal{T} , $qu_{|b}(v)$ is strict descendant of $w_{|a}(v)$ and a strict ancestor of $wt_{|a}(v)$.

Inference (vii) follows since v and $rwt_{|b}$ must be descendants of the same head in \mathcal{T}_b . This implies that $fe_{|b}(v)$ is the *leftmost* leaf descendant of $qu_{|b}(v)$. Since $fe_{|b}(v)$ is distinct from v and interval I_2 is free of bs , it must be that $fe_{|b}(v)$ lies in I_1 and that $qu_{|b}(v)$ is a strict descendant of $w_{|a}(v)$ and a strict ancestor of $wt_{|a}(v)$. Inference (viii) follows. See Figure 4.

It follows from Criterion (iii) and Inference (viii) that on a leaf-to-root path one encounters the nodes $wt_{|a}(v)$, $qu_{|b}(v)$, $w_{|a}(v)$, and $cr_{|b}$, in that order. Lemma 4.2 implies that $a \in \mathcal{B}(qu_{|b}(v))$. We have deduced that $qu_{|b}(v)$ is in \mathcal{T}_a , is a strict descendant of $w_{|a}(v)$, and is an ancestor of both $wt_{|a}(v)$ and v . This contradicts the fact that $w_{|a}(v)$ is the least common ancestor of v and $wt_{|a}(v)$ in \mathcal{T}_a . \square

5. A RECURRENCE FOR ODD ORDERS

Lemma 4.3 may be rephrased as follows. Every blocked sequence S is the union of four sequences: two comprising wingtips (first occurrences and last occurrences, each of length n), one comprising all feathers, and one comprising non-wingtip non-feathers. The last sequence is distinguished by the property that each pair of symbols in any block is nested with respect to S , which is a “good” thing if we are intent on giving strong upper bounds on odd-order sequences. The sequence comprising feathers is “bad” in this sense, therefore we must obtain better-than-trivial upper bounds on its length if this strategy is to bear fruit.

Recall that *feather* is a term that can be applied to nodes in some \mathcal{T}_a or the corresponding *occurrences* of a in the given sequence S . This definition is with respect to *one* derivation tree \mathcal{T} for S , which is not necessarily the best one. Let $\mathcal{T}^*(S)$ be the derivation tree for S that minimizes the total number of feathers.⁸

RECURRENCE 5.1. *Define $\Phi_s(n, m)$ to be the maximum number of feathers in any order- s , m -block DS sequence S over an n -letter alphabet, with respect to the optimal derivation tree $\mathcal{T}^*(S)$. When $m \leq 2$ we have $\Phi_s(n, m) = 0$.*

⁸In other words, when $i > 1$ the construction can choose any $\hat{m} \in \llbracket S \rrbracket$ and any block partition $\{m_q\}_{1 \leq q \leq \hat{m}}$. When $i = 1$ the construction is the same as $\mathcal{T}(S, \cdot, 1)$.

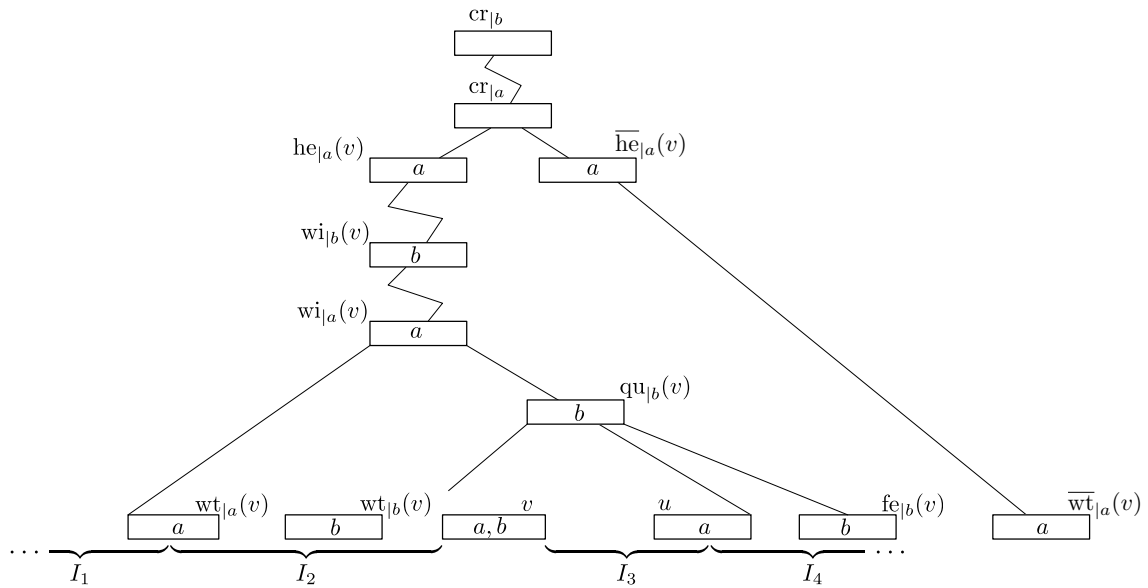


Figure 3: Boxes represent nodes in \mathcal{T} and their associated blocks. The blocks at the leaf-level correspond to those in S . In Case 1 all occurrences of b outside of $\mathcal{B}(v)$ appear in intervals I_2 and I_4 . Contrary to the depiction, it may be that cr_a and cr_b are identical, that $wt_a(v)$ and $wt_b(v)$ are identical, that u and $fe_b(v)$ are identical, and that $wi_b(v)$ is not a strict ancestor of $wi_a(v)$.

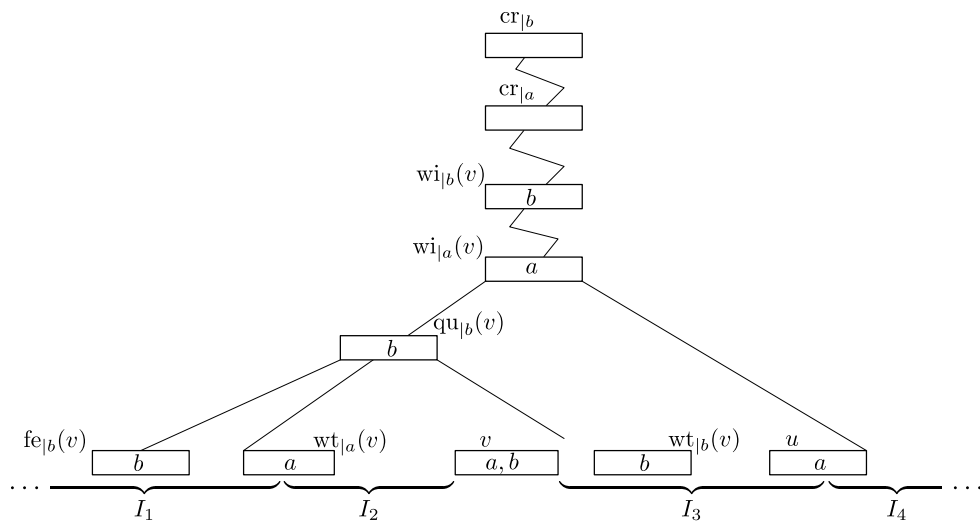


Figure 4: In Case 2 all occurrences of b outside of $\mathcal{B}(v)$ appear in intervals I_1 and I_3 .

For any block partition $\{m_q\}_{1 \leq q \leq \hat{m}}$ and alphabet partition $\{\hat{n}\} \cup \{\tilde{n}_q\}_{1 \leq q \leq \hat{m}}$, we have

$$\begin{aligned} & \Phi_s(n, m) \\ & \leq \sum_{q=1}^{\hat{m}} \Phi_s(\tilde{n}_q, m_q) + \Phi_s(\hat{n}, \hat{m}) + 2 \cdot (\lambda_{s-1}(\hat{n}, m) - \hat{n}) \end{aligned}$$

PROOF. When $m \leq 2$, $\Phi_s(n, m)$ is trivially 0 since every occurrence in S is a wingtip, and feathers are not wingtips. When $m > 2$ we shall decompose S as in Section 2.1. The choice of \hat{m} and the block partition $\{m_q\}_{1 \leq q \leq \hat{m}}$ are not necessarily those of the optimal derivation tree, but we do not need them to be. We are only interested in an upper bound on $\Phi_s(n, m)$. Let $\hat{\mathcal{T}}^*$ and $\{\check{\mathcal{T}}_q^*\}_{1 \leq q \leq \hat{m}}$ be the optimal derivation trees for \hat{S}' and $\{\check{S}_q\}_{1 \leq q \leq \hat{m}}$, and let \mathcal{T} be their composition, with the blocks of S placed at \mathcal{T} 's leaves.

The number of occurrences of local feathers w.r.t. $\{\check{\mathcal{T}}_q^*\}$ is at most $\sum_q \Phi_s(\tilde{n}_q, m_q)$. An occurrence of $a \in \mathcal{B}(v)$ in \hat{S} will be a dove feather in \mathcal{T} if either (i) v is the rightmost child of a dove feather in $\hat{\mathcal{T}}_{|a}^*$ or (ii) v is a non-wingtip child of the left wingtip in $\hat{\mathcal{T}}_{|a}^*$, which corresponds to an occurrence of a in \check{S} . The same statement is true of hawk feathers, swapping the roles of left and right and substituting \check{S} for \hat{S} . There are at most $\Phi_s(\hat{n}, \hat{m})$ feathers of type (i) and, since \check{S} and \hat{S} are order- $(s-1)$ DS sequences, at most $2 \cdot (\lambda_{s-1}(\hat{n}, m) - \hat{n})$ of type (ii). \square

We now have all the elements in place to provide a recurrence for odd-order Davenport-Schinzel sequences.

RECURRENCE 5.2. *Let m, n , and $s \geq 3$ be the block count, alphabet size, and order parameters, where $s \geq 5$ is odd. For any block partition $\{m_q\}_{1 \leq q \leq \hat{m}}$ and alphabet partition $\{\hat{n}\} \cup \{\tilde{n}_q\}_{1 \leq q \leq \hat{m}}$, we have*

$$\begin{aligned} \lambda_s(n, m) & \leq \sum_{q=1}^{\hat{m}} \lambda_s(\tilde{n}_q, m_q) + 2 \cdot \lambda_{s-1}(\hat{n}, m) \\ & \quad + \lambda_{s-2}(\Phi_s(\hat{n}, \hat{m}), m) + \lambda_{s-3}(\lambda_s(\hat{n}, \hat{m}), m) \end{aligned}$$

PROOF. As always, we adopt the notation from Section 2.1. Define $\hat{\mathcal{T}}^*$, $\{\check{\mathcal{T}}_q^*\}$, and \mathcal{T} as in the proof of Recurrence 5.1. In Recurrence 2.2 we partitioned S into local and global symbols and partitioned the occurrences of global symbols into first, middle, and last. We now partition the middle occurrences one step further. Define \check{S}' and \check{S}'' to be the subsequences of \check{S}' consisting of feathers (according to $\hat{\mathcal{T}}^*$) and non-feather, non-wingtips, respectively. That is, $|\check{S}'| = |\check{S}'| + |\check{S}''| + 2\hat{n}$. In an analogous fashion define \check{S} and \check{S}'' to be the subsequences of \check{S} consisting of children of occurrences in \check{S}' and \check{S}'' . The sequences \check{S} and \check{S}'' represent the children of dove and hawk wingtips in $\hat{\mathcal{T}}^*$. Thus, $|\check{S}| = \sum_q |\check{S}_q| + |\check{S}'| + |\check{S}''| + |\check{S}''| + |\check{S}''|$.

The local sequences $\{\check{S}_q\}$ are order- s DS sequences. According to the standard argument \check{S} and \check{S}'' are order- $(s-1)$ DS sequences and $\check{S} = \check{S}_1 \cdots \check{S}_{\hat{m}}$ is obtained from \check{S}' by substituting for its q th block an order- $(s-2)$ DS sequence \check{S}_q over the same alphabet. Thus $|\check{S}| \leq \lambda_{s-2}(|\check{S}'|, m) \leq \lambda_{s-2}(\Phi_s(\hat{n}, \hat{m}), m)$.

We claim that $\check{S} = \check{S}_1 \cdots \check{S}_{\hat{m}}$ is obtained from \check{S}' by substituting for its q th block an order- $(s-3)$ DS sequence \check{S}_q , which, if true, would imply that $|\check{S}| \leq \lambda_{s-3}(|\check{S}'|, m) \leq$

$\lambda_{s-3}(\lambda_s(\hat{n}, \hat{m}), m)$. Suppose for the purpose of obtaining a contradiction that the q th block β in \check{S}' contains $a, b \in \check{\Sigma}$, and that \check{S}_q is not an order- $(s-3)$ DS sequence, that is, it contains an alternating subsequence $ab \cdots ab$ of length $s-1$. Note that $s-1$ is even. By definition β is a non-feather, non-wingtip in both $\mathcal{T}_{|a}$ and $\mathcal{T}_{|b}$. According to Lemma 4.3, a and b must be nested in β , which implies that S contains a subsequence of the form

$$\begin{array}{c} a \cdots b \cdots \left| \overbrace{\cdots a \cdots b \cdots a \cdots b \cdots}^{s-1} \right| \cdots b \cdots a \\ \text{or} \quad b \cdots a \cdots \left| \overbrace{\cdots a \cdots b \cdots a \cdots b \cdots}^{s-1} \right| \cdots a \cdots b \end{array}$$

where the portion between bars is in S_q . In either case S contains an alternating subsequence with length $s+2$, contradicting the fact that S is an order- s DS sequence. \square

The dependencies between λ and Φ established by Recurrences 2.2, 5.1, and 5.2 are rather intricate. For even s , λ_s depends on λ_{s-1} and λ_{s-2} and for odd s , λ_s depends on λ_{s-1} , λ_{s-2} , λ_{s-3} , and Φ_s while Φ_s depends on λ_{s-1} . Lemma 5.3 states bounds on $\lambda_s(n, m)$ that imply parts of Theorem 1.2. Its proof is omitted from this extended abstract.

LEMMA 5.3. *Let $s \geq 1, i \geq 1$, and $c \geq s-2$. The following upper bounds on $\lambda_s(n, m)$ and $\Phi_s(n, m)$ hold for all $s \geq 1$ and all odd $s \geq 5$, respectively, where j is defined to be minimum such that $m \leq a_{i,j}^c$.*

$$\begin{aligned} \lambda_1(n, m) & \leq n + m - 1 \\ \lambda_2(n, m) & \leq 2n + m - 2 \\ \lambda_3(n, m) & \leq (2i + 2)n + (3i - 2)cj(m - 1) \\ \lambda_s(n, m) & \leq \mu_{s,i} (n + (cj)^{s-2}(m - 1)) \quad \text{all } s \geq 4 \\ \Phi_s(n, m) & \leq \nu_{s,i} (n + (cj)^{s-2}(m - 1)) \quad \text{odd } s \geq 5 \end{aligned}$$

The values $\{\mu_{s,i}, \nu_{s,i}\}$ are defined as follows, where $t = \lfloor \frac{s-2}{2} \rfloor$.

$$\begin{aligned} \mu_{s,i} & = \begin{cases} 2^{\binom{i+t+3}{t}} - 3(2(i+t+1))^t & \text{even } s \geq 4 \\ \frac{3}{2}(2(i+t+1))^{t+1} 2^{\binom{i+t+3}{t}} & \text{odd } s \geq 5 \end{cases} \\ \nu_{s,i} & = 4 \cdot 2^{\binom{i+t+3}{t}} \quad \text{odd } s \geq 5 \end{aligned}$$

It is not difficult to show that the optimal choice of i is $\alpha(n, m) + O(1)$. Since $\binom{i+t+O(1)}{t} = i^t/t! + O(i^{t-1})$, Lemma 5.3 immediately yields the following bounds on $\lambda_s(n, m)$.

$$\lambda_s(n, m) = \begin{cases} O((n+m)\alpha(n, m)) & s = 3 \\ O((n+m)2^{\alpha(n, m)}) & s = 4 \\ O((n+m)\alpha^2(n, m)2^{\alpha(n, m)}) & s = 5 \\ O((n+m)2^{\alpha^t(n, m)/t! + O(\alpha^{t-1}(n, m))}) & s \geq 6 \end{cases}$$

Using existing machinery it is possible to obtain analogous bounds on $\lambda_s(n)$ (in terms of $\alpha(n)$ rather than $\alpha(n, m)$)

and even to show that $\lambda_3(n) = 2n\alpha(n) + O(n)$. This still leaves an $\alpha(n)$ gap at order-5 between the $\Omega(n\alpha(n)2^{\alpha(n)})$ construction (which is omitted from this extended abstract) and the $O(n\alpha^2(n)2^{\alpha(n)})$ upper bound. To shave off the extra $\alpha(n)$ factor we need to consider a *two layer* derivation tree for order-5 DS sequences that incorporates the derivation trees for all order-4 DS sequences (\acute{S} , \acute{S}) encountered during the sequence decomposition.

6. CONCLUSION

We have provided the strongest bounds on $\lambda_s(n)$ that can be expressed in an *Ackermann-invariant* fashion, that is, in a way that is tolerant to $O(1)$ uncertainty in the definition of $\alpha(n)$. However, this is not necessarily the last word on the problem. It has been pointed out to us (M. Sharir and G. Nivasch, personal communication) that although $\lambda_{s-1}(n)$ and $\lambda_s(n)$ cannot be distinguished by Ackermann-invariant expressions, for odd $s \geq 7$, one *could* conceivably bound the ratio $\lambda_s(n)/\lambda_{s-1}(n)$ in some non-constructive fashion. Our analyses provide compelling evidence that $\lambda_s(n)/\lambda_{s-1}(n) = \Theta(\alpha(n))$ for all odd $s \geq 3$.

7. REFERENCES

- [1] R. Adamec, M. Klazar, and P. Valtr. Generalized Davenport-Schinzel sequences with linear upper bound. *Discrete Math.*, 108(1-3):219–229, 1992.
- [2] P. Agarwal, M. Sharir, and P. Shor. Sharp upper and lower bounds on the length of general Davenport-Schinzel sequences. *J. Combin. Theory Ser. A*, 52:228–274, 1989.
- [3] N. Alon, H. Kaplan, G. Nivasch, M. Sharir, and S. Smorodinsky. Weak ϵ -nets and interval chains. *J. ACM*, 55(6), 2008.
- [4] S. Alstrup, P. W. Lauridsen, P. Sommerlund, and M. Thorup. Finding cores of limited length. In *Proceedings 5th International Workshop on Algorithms and Data Structures (WADS)*, pages 45–54, 1997.
- [5] M. J. Atallah. Some dynamic computational geometry problems. *Computers & Mathematics with Applications*, 11(12):1171–1181, 1985.
- [6] R. E. Burkard and H. Dollani. Robust location problems with pos/neg weights on a tree. *Networks*, 38(2):102–113, 2001.
- [7] J. Cibulka and J. Kynčl. Tight bounds on the maximum size of a set of permutations with bounded VC-dimension. *J. Combin. Theory Ser. A*, 119(7):1461–1478, 2012.
- [8] H. Davenport. A combinatorial problem connected with differential equations. II. *Acta Arith.*, 17:363–372, 1970/1971.
- [9] H. Davenport and A. Schinzel. A combinatorial problem connected with differential equations. *American J. Mathematics*, 87:684–694, 1965.
- [10] S. Hart and M. Sharir. Nonlinearity of Davenport-Schinzel sequences and of generalized path compression schemes. *Combinatorica*, 6(2):151–177, 1986.
- [11] M. M. Klawe. Superlinear bounds for matrix searching problems. *J. Algor.*, 13(1):55–78, 1992.
- [12] M. Klazar. A general upper bound in extremal theory of sequences. *Comment. Math. Univ. Carolin.*, 33(4):737–746, 1992.
- [13] M. Klazar. On the maximum lengths of Davenport-Schinzel sequences. In *Contemporary Trends in Discrete Mathematics, Štířín Castle 1997 (Czech Republic)*, American Mathematical Society, Providence RI, pages 169–178, 1999.
- [14] M. Klazar. Generalized Davenport-Schinzel sequences: results, problems, and applications. *Integers*, 2:A11, 2002.
- [15] P. Komjáth. A simplified construction of nonlinear Davenport-Schinzel sequences. *J. Combin. Theory Ser. A*, 49(2):262–267, 1988.
- [16] M. C. López-de-los Mozos, J. Puerto, and A. M. Rodríguez-Chía. Robust mean absolute deviation problems on networks with linear vertex weights. *Networks*, 2012.
- [17] G. Nivasch. Improved bounds and new techniques for Davenport-Schinzel sequences and their generalizations. *J. ACM*, 57(3), 2010.
- [18] S. Pettie. Splay trees, Davenport-Schinzel sequences, and the deque conjecture. In *Proceedings 19th ACM-SIAM Symposium on Discrete Algorithms*, pages 1115–1124, 2008.
- [19] S. Pettie. Generalized Davenport-Schinzel sequences and their 0-1 matrix counterparts. *J. Comb. Theory Ser. A*, 118(6):1863–1895, 2011.
- [20] A. Di Salvo and G. Proietti. Swapping a failing edge of a shortest paths tree by minimizing the average stretch factor. *Theoretical Computer Science*, 383(1):23–33, 2007.
- [21] M. Sharir. Almost linear upper bounds on the length of general Davenport-Schinzel sequences. *Combinatorica*, 7(1):131–143, 1987.
- [22] M. Sharir. Improved lower bounds on the length of Davenport-Schinzel sequences. *Combinatorica*, 8(1):117–124, 1988.
- [23] M. Sharir and P. Agarwal. *Davenport-Schinzel Sequences and their Geometric Applications*. Cambridge University Press, 1995.
- [24] M. Sharir, R. Cole, K. Kedem, D. Leven, R. Pollack, and S. Sifrony. Geometric applications of Davenport-Schinzel sequences. In *Proceedings 27th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 77–86, 1986.
- [25] E. Szemerédi. On a problem of Davenport and Schinzel. *Acta Arith.*, 25:213–224, 1973/74.
- [26] R. E. Tarjan. Efficiency of a good but not linear set merging algorithm. *J. ACM*, 22(2):215–225, 1975.