WILF-EQUIVALENCE OF CLASSICAL PERMUTATION PATTERNS: THE CODIMENSION-TWO MAXIMUM

KOREA SUPERINTELLIGENCE LABS

ABSTRACT. Write $\mathrm{Av}_n(\beta)$ for the set of permutations of length n that avoid the classical pattern β, and call two patterns Wilf-equivalent when these sets have the same size at every n. Ray and West proved that for β of length m the number $g_2(\beta)$ of permutations of length $m+2$ containing β is $(m^4+2m^3+m^2+4m+4-2j)/2$ for an integer $j=j(\beta)$ with $0 \le j \le m-1$; expressing j as a statistic of β is a problem of Vatter's, open since 2007. We determine the extreme case. For every pattern of length 3, 4 and 5 we prove, by exhaustive machine-checked computation, that $|\mathrm{Av}_{m+2}(\beta)|$ is largest — equivalently $j(\beta)=m-1$ — exactly when β is layered or reverse-layered, that is, avoids both 231 and 312, or both 132 and 213; an independent computation confirms this through length 8. There are 2^m-2 such patterns, and neither half of the disjunction can be dropped. The same machinery reproduces the published codimension-one and codimension-two data, the lower bounds 3 and 16 on the number of Wilf classes at lengths 4 and 5, the three shape-Wilf classes of length-3 patterns, and Stankova's Wilf-equivalence $1342 \sim 2413$ that is not shape-Wilf; at length 8 it realises all eight avoidance counts that Ray–West's formula permits at $n=10$. Separately, Section 7 reports a computation carried out outside the formal development and *not* machine-checked: there are exactly 4755 Wilf-equivalence classes of permutations of length 8, and the published theorems account for all of them. Except where explicitly marked as external, every statement below is machine-checked in Lean 4.

1. Introduction

Let $\mathfrak{S}_n$ be the set of permutations of $\{1,\dots,n\}$, written in one-line notation. A permutation $\pi \in \mathfrak{S}_n$ *contains* the *pattern* $\beta \in \mathfrak{S}_m$ if some subsequence of π of length m is order-isomorphic to β, and *avoids* β otherwise; $\mathrm{Av}_n(\beta) \subseteq \mathfrak{S}_n$ is the set of avoiders. Two patterns β, γ are *Wilf-equivalent*, written $\beta \sim \gamma$, when $|\mathrm{Av}_n(\beta)| = |\mathrm{Av}_n(\gamma)|$ for every n.

The eight symmetries of the square act on plots of permutations — they are generated by reverse, complement and inverse — and patterns in the same symmetry orbit are trivially Wilf-equivalent. The converse fails, and fails immediately: 231 and 321 lie in different symmetry orbits and are both counted by the Catalan numbers. Classifying patterns up to Wilf-equivalence is therefore a genuine problem, and it is a hard one. As Vatter puts it [9], *"the only general method for proving that $\pi \not\sim \sigma$ is to enumerate the π- and σ-avoiding permutations until the counts disagree"*. The classification is complete for $m \le 7$, with $1,1,1,3,16,91,595$ classes for $m=1,\dots,7$ (OEIS `A099952`), and has been stuck there since Stankova and West [8]. Section 3 of [9] collects the resulting questions; two of them drive this paper.

Problem 3.7 of [9]**.** Ray and West [5, Theorem 8.7] showed that, writing $g_k(\beta)$ for the number of permutations of length $m+k$ that *contain* a pattern β of length m,

$$g_2(\beta) \;=\; \frac{m^4+2m^3+m^2+4m+4-2j}{2}, \qquad 0 \le j \le m-1, \tag{1}$$

where $j=j(\beta)$ is defined operationally, as a count of "active lefthand triples" [5, Corollary 8.6]. *Express j in terms of statistics of β.* The problem has been open since Permutation Patterns 2007.

Question 3.1 of [9]**.** *How many Wilf-equivalence classes of permutations of length 8 are there?* Vatter states the fork explicitly: perhaps the existing theorems suffice to explain every equivalence at length 8, or perhaps there are new Wilf-equivalences waiting to be discovered.

Since $|\mathrm{Av}_{m+2}(\beta)| = (m+2)! - g_2(\beta)$, the extreme case $j(\beta) = m-1$ of (1) is exactly the case in which β is *hardest to contain two steps up*: $|\mathrm{Av}_{m+2}(\beta)|$ attains its maximum

$$\max_{\beta\in\mathfrak{S}_m} |\mathrm{Av}_{m+2}(\beta)| \;=\; (m+2)! - \frac{m^4+2m^3+m^2+2m+6}{2}, \tag{2}$$

which is 42, 513, 4582 and 3626197 for $m = 3, 4, 5, 8$. Our main result identifies the patterns that attain it. Call β *layered* if it is a direct sum of decreasing blocks, equivalently if it avoids both 231 and 312, and *reverse-layered* if its complement is layered, equivalently if it avoids both 132 and 213.

Theorem 1.1 (Theorem 4.3)**.** *For every pattern β of length $m \in \{3, 4, 5\}$,*

$$|\mathrm{Av}_{m+2}(\beta)| \text{ is maximal} \iff \beta \text{ is layered or reverse-layered.}$$

The proof is an exhaustive check over all 6, 24 and 120 patterns of those lengths, run inside the kernel of a proof assistant; a separate, independently written program confirms the same equivalence for every pattern of length 6, 7 and 8. The patterns attaining the maximum are therefore $2^m - 2$ in number (Lemma 4.2), and both halves of the disjunction are needed: 12354 is layered and not reverse-layered, 23451 is reverse-layered and not layered, and both attain 4582. This settles the extreme value of j and nothing more, so Problem 3.7 itself remains open; we state the general case as Conjecture 4.5.

Question 3.1 is a different kind of object, and it is not what is proved here. Separating the 5282 symmetry classes of length-8 patterns requires computing $|\mathrm{Av}_n(\beta)|$ up to $n = 14$, which is three orders of magnitude past what the proof kernel can carry. Section 6 records what *is* machine-checked at length 8: all eight values permitted by (1) occur at $n = 10$, the first length at which any two length-8 patterns can differ at all, so there are at least eight Wilf classes. Beyond that, Section 7 reports — separately, and marked throughout as an external computation that is *not* machine-checked — a value for #Wilf(8) together with the conclusion that the published theorems account for every Wilf-equivalence at length 8, which is the first branch of Vatter's fork. The argument there is a two-sided squeeze, a separation computation from below meeting the closure of the published equivalence theorems from above, and it is described in enough detail to be repeated.

The remaining sections reproduce the published record with the same machinery (Section 5) and record the negative controls that keep the whole development honest (Section 8). Section 9 states what has been formally verified and how.

2. Preliminaries

Throughout, m denotes the length of a pattern β and n the length at which avoidance counts are compared. Patterns are written in one-line notation over $\{1, \dots, m\}$; in the formal development they are indexed from 0 instead, so that 13425 appears there as 02314. We write $\iota_m = 12\cdots m$ for the increasing pattern.

Containment is defined through the rank sequence: for a list s of distinct integers, $\mathrm{pat}(s)$ replaces each entry by the number of entries of s below it, and π contains β exactly when some subsequence s of π has $\mathrm{pat}(s) = \beta$. Then $\mathrm{Av}_n(\beta) = \{\pi \in \mathfrak{S}_n : \pi \text{ does not contain } \beta\}$ and

$$g_k(\beta) \;=\; |\mathfrak{S}_{m+k}| - |\mathrm{Av}_{m+k}(\beta)| \;=\; (m+k)! - |\mathrm{Av}_{m+k}(\beta)|.$$

Two elementary facts are used constantly. First, if $n < m$ then every permutation of length n avoids β, so $|\mathrm{Av}_n(\beta)| = n!$. Second — and this is the whole engine of the subject — if $|\mathrm{Av}_n(\beta)| \neq |\mathrm{Av}_n(\gamma)|$ for a single n, then $\beta \not\sim \gamma$; consequently a list of patterns whose avoidance counts at one common n are pairwise distinct is a list of pairwise non-Wilf-equivalent patterns.

At codimension 1 the count does not see the pattern at all: $g_1(\beta) = m^2 + 1$ for every $\beta \in \mathfrak{S}_m$, so $|\mathrm{Av}_{m+1}(\beta)| = (m+1)! - m^2 - 1$. At codimension 2 it begins to, through (1): $|\mathrm{Av}_{m+2}(\beta)|$ takes at most m values, namely $(m+2)! - (m^4 + 2m^3 + m^2 + 4m + 4)/2 + j$ for $0 \le j \le m-1$, and Ray and West add that every value of j occurs once $m \ge 5$. Specialising, $|\mathrm{Av}_5(\beta)| = 40 + j$ with $j \le 2$,

$|\mathrm{Av}_6(\beta)| = 510 + j$ with $j \leq 3$, $|\mathrm{Av}_7(\beta)| = 4578 + j$ with $j \leq 4$, and $|\mathrm{Av}_{10}(\beta)| = 3626190 + j$ with $j \leq 7$ for β of length 8. Formula (2) is (1) at $j = m - 1$.

Shape-Wilf-equivalence, written $\alpha \overset{\mathrm{s}}{\sim} \beta$, is the refinement of Wilf-equivalence obtained by counting avoiders inside Ferrers boards rather than inside the full square; it implies Wilf-equivalence, and it is the notion in which the strongest known equivalence theorems are proved. Two of those theorems are used below. Backelin, West and Xin [2] proved $12 \cdots k \overset{\mathrm{s}}{\sim} k \cdots 21$ for every k, and Stankova and West [8] proved $231 \overset{\mathrm{s}}{\sim} 312$; these are Theorems 3.3 and 3.4 of [9]. Moreover, by a closure property first observed by Babson and West [1] and made explicit by Backelin, West and Xin [2] — Proposition 3.2 of [9] — shape-Wilf-equivalence is preserved by direct sums:

$$\alpha \overset{\mathrm{s}}{\sim} \beta \implies \alpha \oplus \gamma \overset{\mathrm{s}}{\sim} \beta \oplus \gamma \quad \text{for every } \gamma, \tag{3}$$

where $\alpha \oplus \gamma$ places γ above and to the right of α. Since $\overset{\mathrm{s}}{\sim}$ implies $\sim$, a proof that $\alpha \oplus 1 \not\sim \beta \oplus 1$ refutes $\alpha \overset{\mathrm{s}}{\sim} \beta$; this is the route used in Section 5.

All question, problem and result numbers quoted from [9] follow version 2 of that preprint (25 February 2026), the latest version at the time of writing; Question 3.1 stands open there.

3. Evaluating $|\mathrm{Av}_n(\beta)|$ inside a proof kernel

Every count below is produced by literal enumeration — there is no formula to appeal to — so the cost of the enumeration decides what is provable. The definition of containment quoted in Section 2 evaluates badly: deciding it for a single π builds all $\binom{n}{m}$ subsequences of π and rank-normalises each in $O(m^2)$ steps. In the regime that matters here the *codimension* $k = n - m$ is small (2, 3 or 4), and it is far cheaper to enumerate the k deleted *values*.

Proposition 3.1. *Let π be a permutation of $\{0, \ldots, n-1\}$ and let β be a pattern with $|\beta| \leq n$, and put $k = n - |\beta|$. Then π contains β if and only if there is a k-element subset $S \subseteq \{0, \ldots, n-1\}$ such that deleting from π every entry whose value lies in S and replacing each surviving entry v by $v - |\{u \in S : u < v\}|$ produces β. Consequently the count obtained by this route agrees with $|\mathrm{Av}_n(\beta)|$ for all n and β with $|\beta| \leq n$.*

The proof is elementary but not vacuous: the arithmetic re-ranking $v \mapsto v - |\{u \in S : u < v\}|$ agrees with the rank sequence of the surviving subsequence precisely because the entries of π are exactly $0, \ldots, n-1$, and this is where the hypothesis that π is a permutation of a full range enters. Deletion, re-ranking and comparison against β are then fused into a single left-to-right walk with early exit. The identification is proved, not assumed, so every statement in this paper is about the transparent definition of $|\mathrm{Av}_n(\beta)|$ while the machine runs the fast route. Measured at $n = 8$ the speedup is a factor of about 7 (3.6 seconds against 26.7), and it is what brings $n = 10$ for length-8 patterns inside budget.

Proposition 3.2. *If $n < |\beta|$ then $|\mathrm{Av}_n(\beta)| = n!$. If L is a list of patterns whose avoidance counts $|\mathrm{Av}_n(\beta)|$, $\beta \in L$, are pairwise distinct for some single n, then the members of L are pairwise non-Wilf-equivalent.*

Both are proved for all n, with no computation; the second is the formal counterpart of the sentence of Vatter's quoted in Section 1, and every separation result below is an instance of it.

4. The codimension-two maximum

Definition 4.1. A pattern β is *layered* if it avoids both 231 and 312, and *reverse-layered* if it avoids both 132 and 213.

Layered permutations are exactly the direct sums of decreasing blocks, equivalently $\mathrm{Av}(231, 312)$ [9, Section 2], and reverse-layered permutations are exactly the skew sums of increasing blocks, their complements. Neither identification is used in what follows: Definition 4.1 is the avoidance condition, and the avoidance condition is what the formal development decides.

Lemma 4.2. *For $m \geq 2$ there are exactly $2^m - 2$ patterns of length m that are layered or reverse-layered.*

Proof. A layered permutation of length m is determined by the composition of m recording its block sizes, so there are 2^{m-1} of them, and as many reverse-layered ones. If β is both, and its layered block decomposition has r blocks of sizes $c_1, \dots, c_r$ with $r \geq 2$ and $c_i \geq 2$ for some i, then taking the two entries of a descent inside block i together with a larger entry from a later block gives an occurrence of 213 when $i < r$, and taking a smaller entry from an earlier block together with the descent gives an occurrence of 132 when $i = r$; either contradicts reverse-layeredness. Hence β is ι_m or $m \cdots 21$, and inclusion–exclusion gives $2^{m-1} + 2^{m-1} - 2$. $\square$

For $m = 4$ and $m = 5$ the counts 14 and 30 predicted by Lemma 4.2 are also verified by direct enumeration in the formal development.

Theorem 4.3. *Let β be a pattern of length m.*

(i) *If $m = 3$, then $|\mathrm{Av}_5(\beta)| = 42$ for all six patterns, and all six are layered or reverse-layered.*
(ii) *If $m = 4$, then $|\mathrm{Av}_6(\beta)| = 513$ — the maximum — if and only if β is layered or reverse-layered, and there are 14 such β.*
(iii) *If $m = 5$, then $|\mathrm{Av}_7(\beta)| = 4582$ — the maximum — if and only if β is layered or reverse-layered, and there are 30 such β.*

Equivalently, in the notation of (1), $j(\beta) = m - 1$ if and only if β is layered or reverse-layered, for every pattern of length $m \leq 5$.

Proof sketch. Each part is a single exhaustive computation, carried out in the kernel of the proof assistant and resting on no other input. For (iii), for instance, the finite search runs over all 120 patterns β of length 5; for each one it enumerates all $7! = 5040$ permutations of length 7, decides avoidance of β by Proposition 3.1, compares the resulting count with 4582, and compares that Boolean with the Boolean "β avoids $\{231, 312\}$ or β avoids $\{132, 213\}$", itself decided by enumerating the $\binom{5}{3} = 10$ subsequences of length 3 of β. The theorem is the assertion that the two Booleans agree in all 120 cases. Parts (i) and (ii) are the same search over 6 patterns at $n = 5$ and 24 patterns at $n = 6$. That the maximum in (ii) and (iii) is 513 and 4582 is (2), and is also re-derived by the same searches, which certify that no pattern of the relevant length exceeds those values (Section 5). $\square$

Theorem 4.4. *Neither half of the disjunction in Theorem 4.3 may be dropped:* 12354 *is layered, is not reverse-layered, and has* $|\mathrm{Av}_7(12354)| = 4582$*, while* 23451 *is reverse-layered, is not layered, and also has* $|\mathrm{Av}_7(23451)| = 4582$.

Two independent computations extend Theorem 4.3 beyond the reach of the kernel. A separate program, written in a different language and evaluating $g_2(\beta)$ by enumerating occurrences of β rather than by filtering permutations, checked the same equivalence for every pattern of length 6, 7 and 8, and the counts $2, 6, 14, 30, 62, 126, 254$ of maximisers for $m = 2, \dots, 8$ agree with Lemma 4.2. These checks are outside the formal development. On that evidence we state:

Conjecture 4.5. For every $m \geq 2$ and every $\beta \in \mathfrak{S}_m$: $j(\beta) = m - 1$ — equivalently $|\mathrm{Av}_{m+2}(\beta)|$ is maximal — if and only if β is layered or reverse-layered.

We emphasise what this does *not* do. It pins one value of j, the largest. Problem 3.7 asks for j as a statistic of β for all values, and that remains open; the distribution of j over $\mathfrak{S}_5$, for instance, is $(2, 12, 36, 40, 30)$ for $j = 0, \dots, 4$, and we know of no description of the classes with $0 < j < m - 1$.

5. Reproducing the published record

The results of this section are known. They are included because the machinery of Section 3 is only worth trusting if it reproduces what is already in print, and because two of the computations below are asserted in [9] without the numbers being printed.

Proposition 5.1. (i) $|\mathrm{Av}_n(231)| = |\mathrm{Av}_n(321)| = C_n$ *for* $1 \le n \le 8$*, the values being* 1*,* 2*,* 5*,* 14*,* 42*,* 132*,* 429*,* 1430*.*

(ii) $|\mathrm{Av}_7(\beta)| = 5003$ *for every one of the* 720 *patterns* β *of length* 6*; that is,* $g_1(\beta) = 37 = 6^2 + 1$ *throughout. The same statement is verified for all patterns of length* 3*,* 4 *and* 5*, with* $|\mathrm{Av}_4| = 14$*,* $|\mathrm{Av}_5| = 103$ *and* $|\mathrm{Av}_6| = 694$*.*

(iii) $|\mathrm{Av}_7(\beta)| \in \{4578, 4579, 4580, 4581, 4582\}$ *for every one of the* 120 *patterns of length* 5*, which is Ray–West's range* $0 \le j \le 4$*; and all five values occur, realised by* 25314*,* 23514*,* 13425*,* 12453 *and* ι_5 *respectively. Correspondingly* $|\mathrm{Av}_5(\beta)| = 42$ *for all six patterns of length* 3 *(*$j = 2$ *throughout) and* $|\mathrm{Av}_6(\beta)| \in \{512, 513\}$ *for all* 24 *patterns of length* 4 *(only* $j = 2, 3$ *occur).*

Part (iii) reproduces Ray and West's assertion that every value of j is realised as soon as $m \ge 5$ [5, Theorem 8.7], together with their remark that at $m = 4$ only $j \in \{2, 3\}$ occur and at $m = 3$ only $j = 2$. Two of the five witnesses above, ι_5 and 25314, are Ray and West's own; for $j = 1, 2, 3$ they print 41253, 14253 and 14523 instead, of which only 41253 lies in the symmetry orbit of the witness used here. Each clause is one exhaustive search over all patterns of the stated length, with a full enumeration of $\mathfrak{S}_n$ inside it.

Remark 5.2. The attribution of $g_1(\beta) = m^2 + 1$ is unsettled. Vatter [9] credits an observation made by Pratt in 1973 [4]; Ray and West [5] write instead that the formula "was originally obtained in 1990 by Bloomberg, but so far as we are aware remains unpublished". The statement is verified here for every pattern of length at most 6; the attribution we leave to others.

Proposition 5.3. *There are at least* 3 *Wilf-equivalence classes of length-*4 *patterns and at least* 16 *Wilf-equivalence classes of length-*5 *patterns.*

Proof sketch. For length 4: $|\mathrm{Av}_7(1234)| = 2761$, $|\mathrm{Av}_7(1324)| = 2762$ and $|\mathrm{Av}_7(1342)| = 2740$ are pairwise distinct, so Proposition 3.2 applies. For length 5: sixteen explicit patterns have pairwise distinct values of $|\mathrm{Av}_9|$, ranging over $258757, \ldots, 261863$; each of the sixteen values is one enumeration of the $9! = 362880$ permutations of length 9. □

Both bounds are tight: the true counts are 3 and 16. The matching upper bounds are the *equalities*, and they are not proved here. At length 4 they are West's $1234 \sim 1243 \sim 2143$ [10], Stankova's $1342 \sim 2413$ [6], and Stankova's $1234 \sim 4123$ [7], the last of which completed the classification; at length 5 the classification was completed by Babson and West [1]. We follow the account of this history given in the introduction of [8]. Only the separating half is proved here. The lengths at which the separations happen are themselves informative. At length 4, 1234 and 1324 agree at every $n \le 6$ (both give 513 at $n = 6$) and first split at $n = 7$. At length 5, the 23 symmetry classes take only 14 distinct values of $|\mathrm{Av}_8|$, so $n = 8$ is not enough and $n = 9$ is genuinely required.

Proposition 5.4. (i) $123 \oplus 1 = 1234$*,* $132 \oplus 1 = 1324$ *and* $231 \oplus 1 = 2314$ *are pairwise non-Wilf-equivalent, separated at* $n = 7$ *by the values* 2761*,* 2762*,* 2740*.*

(ii) $1342 \oplus 1 = 13425$ *and* $2413 \oplus 1 = 24135$ *agree at* $n = 7$*, both giving* 4580*, and separate at* $n = 8$*:* $|\mathrm{Av}_8(13425)| = 33256$ *while* $|\mathrm{Av}_8(24135)| = 33254$*.*

By (3), (i) says that 123, 132 and 231 are pairwise not shape-Wilf-equivalent; combined with $123 \overset{\mathrm{s}}{\sim} 321$ (Backelin–West–Xin), with $213 = 21 \oplus 1 \overset{\mathrm{s}}{\sim} 12 \oplus 1 = 123$ (the same theorem at $k = 2$, pushed through (3)), and with $231 \overset{\mathrm{s}}{\sim} 312$ (Stankova–West), this gives the three shape-Wilf classes $\{123, 321, 213\}$, $\{231, 312\}$, $\{132\}$ of length-3 patterns. Statement (ii) says that $1342 \overset{\mathrm{s}}{\not\sim} 2413$, although $1342 \sim 2413$ by Stankova [6]: a Wilf-equivalence that does not come from shape-Wilf-equivalence. Both computations are asserted in Section 3 of [9] without the numbers; the numbers are $2761, 2762, 2740$ and $33256, 33254$. The last two are the OEIS entries `A256202` and `A256201`, at the symmetries 53241 and 35241, which both carry data to $n = 26$ and agree.

6. Length-eight patterns

For a pattern of length 8 the first length at which *any* separation is possible is $n = 10$. Indeed every permutation of length $n \leq 7$ avoids β; at $n = 8$ exactly $8! - 1$ do; at $n = 9$ exactly $9! - 65$ do, since $g_1(\beta) = 8^2 + 1 = 65$ does not depend on β; and only at $n = 10$ does (1) leave room, giving $|\mathrm{Av}_{10}(\beta)| = 3626190 + j$ with $0 \leq j \leq 7$.

Proposition 6.1. *All eight values permitted at $n = 10$ occur. The patterns*

13526847, 12468375, 12358647, 12356847, 12357468, 12346758, 12345786, 12345678

have $|\mathrm{Av}_{10}| = 3626190, \ldots, 3626197$ respectively, that is $j = 0, \ldots, 7$. Consequently there are at least eight Wilf-equivalence classes of length-8 patterns.

Proof sketch. Eight independent exhaustive computations, one per pattern: each enumerates all $10! = 3628800$ permutations of length 10 and decides avoidance by Proposition 3.1. Each costs about 212 seconds of CPU time and 4.17 gigabytes of memory — the memory is the enumeration of $\mathfrak{S}_{10}$ itself, so it does not shrink with a better matcher, and $n = 11$ is out of reach on this encoding. The separation is then Proposition 3.2. □

Eight is the most that $n = 10$ can give, and it is a long way from the truth. The maximum 3626197 is attained by $\iota_8 = 12345678$, which is both layered and reverse-layered, as Conjecture 4.5 requires.

7. An external computation: $\#\mathrm{Wilf}(8) = 4755$

The results of this section were obtained by computations carried out outside the formal development described in Section 9, and are not machine-checked. They are reported as such.

Length-8 patterns fall into 5282 symmetry classes (OEIS `A000903`), and every Wilf class is a union of symmetry classes. Two partitions of these 5282 classes squeeze the answer from opposite sides.

From below, by separation. Wilf-equivalent patterns have equal counts, so partitioning the symmetry classes by the vector $(|\mathrm{Av}_{10}|, |\mathrm{Av}_{11}|, \ldots)$ gives a partition coarser than the Wilf partition, and its class count is a lower bound. The counts $g_k(\beta)$ were computed by enumerating occurrences: for each choice of m positions and m values one lays β across them, fills the remaining $n - m$ cells in all $(n - m)!$ ways, and deduplicates the resulting permutations by their Lehmer rank in a bitset, at a cost of $\binom{n}{m}^2 (n - m)!$ per pattern, independent of $n!$. Run adaptively — level k only on the classes still tied after level $k - 1$ — this gives

level	n	tied going in	classes out
g_2	10	5282	8
g_3	11	5282	256
g_4	12	5269	4210
g_5	13	1145	4751
g_6	14	8	**4755**

From above, by the published theorems. A proved equivalence is a real one, so the partition generated by the known equivalence theorems is finer than the Wilf partition, and its class count is an upper bound. The generators used are: the eight plot symmetries; $12 \cdots k \overset{\mathrm{s}}{\sim} k \cdots 21$ [2]; $231 \overset{\mathrm{s}}{\sim} 312$ [8]; closure of $\overset{\mathrm{s}}{\sim}$ under $\oplus$, (3); the implication $\overset{\mathrm{s}}{\sim} \Rightarrow \sim$; and $1342 \sim 2413$ [6], which is a Wilf-equivalence only and therefore is *not* propagated through $\oplus$. A union-find closure over $\mathfrak{S}_m$ under these rules gives shape-Wilf class counts $1, 1, 3, 19, 108, 680, 4862, 39326$ and Wilf class counts $1, 1, 1, 3, 16, 91, 595, 4755$ for $m = 1, \ldots, 8$. The first seven Wilf counts are the true ones.

The two bounds meet at 4755. Since a coarser and a finer partition of a finite set with equal numbers of classes coincide, the Wilf partition of $\mathfrak{S}_8$ is thereby determined exactly, and not merely counted; in particular every Wilf-equivalence at length 8 is accounted for by the published theorems, which is the first branch of Vatter's fork in Question 3.1. The sequence `A099952` currently holds seven terms and carries the OEIS keyword `more`; on this computation its eighth term is 4755.

It was close. At $n = 12$ the separation left 545 merges unexplained by the upper bound; at $n = 13$, four; $n = 14$ removed those four:

pair	g_6 at $n = 14$	margin
13426758 vs. 13427568	1705417708 vs. 1705417704	4
13678254 vs. 32718564	1716923933 vs. 1716921913	2020
14567823 vs. 14567832	1692573439 vs. 1692573824	385
23815764 vs. 25876314	1717958606 vs. 1717959654	1048

A margin of 4 in 1.7×10^9 decides the first pair. A run stopping at $n = 12$ or $n = 13$ would have reported spurious new Wilf-equivalences, and no amount of agreement between counts would have revealed it. What makes stopping safe is the upper bound, because it says exactly how many merges still need explaining.

The computation was checked five ways: against a brute-force containment count for every pattern of length 3 and 4; against the formally verified values of Proposition 6.1, which are produced by a completely different route (filtering all of $\mathfrak{S}_{10}$, with no occurrence enumeration and no Lehmer ranking) and agree on all eight; against the literature, by reproducing `A000903` to $n = 8$ and `A099952` at lengths $4, 5, 6, 7$; against the union-find closure, in that each of the 4755 known-theorem classes lies inside a single count-vector class, with zero violations; and, by accident, in duplicate, a shell redirection error having caused every one of the 1145 patterns at the $n = 13$ level to be computed twice independently, with all duplicates agreeing. The whole run cost about five core-hours, roughly half of it lost to that duplication; the binding constraint at $n = 14$ is memory, the deduplication bitset there occupying 14! bits, or 10.9 gigabytes.

The same pipeline applied to length 9 would face 46066 symmetry classes and a deduplication bitset of about 87 gigabytes at $n = 15$, so the Lehmer-rank scheme would have to be replaced before that level is reachable at all. The upper bound, by contrast, is cheap at any length, and it is the useful half to compute first: it says how far the separation has to be pushed before stopping is safe.

8. Negative controls

Every assertion in this paper is a finite count, so the ways to be wrong are to claim something too strong — a separation that is not there, a characterisation that over-reaches — or too weak — a count off by one, a bound not attained. Nine controls, each machine-checked, refute such variants.

Proposition 8.1. (i) Too strong: "Wilf-equivalent implies symmetric". 231 *and* 321 *have equal avoidance counts at every* $n \leq 8$*, yet* 321 *does not lie in the symmetry orbit of* 231*; the orbit is certified closed under reverse, complement and inverse, so this is not a search that stopped early.*

(ii) Too strong: "no two patterns of the same length ever separate". *Refuted at length* 4*:* $1234 \not\sim 1324$.

(iii) Too strong: "1234 and 1324 separate at $n = 6$". *They do not: both give* 513*, which is why the separation of Proposition 5.4(i) needs* $n = 7$.

(iv) Too strong: "1342 and 2413 separate below $n = 9$". *They do not, at* $n = 6$*,* 7 *or* 8*; the separation that does exist is one level up, between* 13425 *and* 24135.

(v) Too weak: an off-by-one in the split of Proposition 5.4(ii). $|\mathrm{Av}_8(13425)| \neq 33255$ *and* $|\mathrm{Av}_8(24135)| \neq 33256$*: the split is by exactly two.*

(vi) Too strong: "$j = m - 1$ always", i.e. $|\mathrm{Av}_{m+2}|$ is constant. $|\mathrm{Av}_7(25314)| = 4578$*, four below the maximum, so codimension* 2 *— unlike codimension* 1 *— does see the pattern.*

(vii) Too weak: at length 5, "$|\mathrm{Av}_7(\beta)|$ can reach 4577, or 4583". *It cannot, so the range* $\{4578, \ldots, 4582\}$ *of Proposition 5.1(iii) is exactly right on both sides.*

(viii) Too strong: "layered alone characterises the maximum". 23451 *is reverse-layered, is not layered, and attains* 4582*; symmetrically* 12354 *is layered, is not reverse-layered, and attains* 4582.

(ix) Too weak: "some pattern that is neither layered nor reverse-layered attains the maximum". *At length* 5 *none does — the contrapositive form of Theorem 4.3(iii), which is what a counterexample search would have to refute.*

9. Verification

Every statement in this paper other than those of Section 7, of Lemma 4.2, and of Conjecture 4.5 has been formally verified in Lean 4 (toolchain `v4.32.0`) against *Mathlib*, in a development that is free of `sorry`; the theorems as stated above correspond one-to-one to theorems there. Propositions 3.1 and 3.2 are ordinary proofs and depend on no axioms beyond `propext`, `Classical.choice` and `Quot.sound`. What is delegated to compiled evaluation, through Lean's `native_decide`, is exactly the finite searches: one search over all 6, 24, respectively 120 patterns for the three parts of Theorem 4.3, each enumerating $\mathfrak{S}_5$, $\mathfrak{S}_6$, respectively $\mathfrak{S}_7$ for every pattern; one over all 720 patterns of length 6 for Proposition 5.1(ii) and over all 120 of length 5 for (iii); one avoidance count per value in Propositions 5.3 and 5.4; and one enumeration of $\mathfrak{S}_{10}$ for each of the eight counts of Proposition 6.1. The heavy files together consume about 49 minutes of CPU time. Numerical values were produced beforehand by independent implementations in C or Python, and the formal statements were written against them; the eight length-8 counts of Proposition 6.1 agree exactly between the two routes. The computations of Section 7 are outside this development and are not machine-checked.

References

[1] E. Babson and J. West, *The permutations* $123p_4\cdots p_m$ *and* $321p_4\cdots p_m$ *are Wilf-equivalent*, Graphs and Combinatorics **16** (2000) 373–380.

[2] J. Backelin, J. West and G. Xin, *Wilf-equivalence for singleton classes*, Advances in Applied Mathematics **38** (2007) 133–148.

[3] OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, `https://oeis.org`; entries `A000903` (inequivalent placements of n non-attacking rooks, i.e. symmetry classes of permutations), `A099952` (Wilf classes in $\mathfrak{S}_n$), `A256201` and `A256202` (permutations avoiding 35241, respectively 53241). All four consulted in August 2026.

[4] V. R. Pratt, *Computing permutations with double-ended queues, parallel stacks and parallel queues*, in: STOC '73: Proceedings of the Fifth Annual ACM Symposium on Theory of Computing, ACM, 1973, pp. 268–277.

[5] N. Ray and J. West, *Posets of matrices and permutations with forbidden subsequences*, Annals of Combinatorics **7** (2003) 55–88.

[6] Z. Stankova, *Forbidden subsequences*, Discrete Mathematics **132** (1994) 291–316.

[7] Z. Stankova, *Classification of forbidden subsequences of length* 4, European Journal of Combinatorics **17** (1996) 501–517.

[8] Z. Stankova and J. West, *A new class of Wilf-equivalent permutations*, Journal of Algebraic Combinatorics **15** (2002) 271–290.

[9] V. Vatter, *An assortment of problems in permutation patterns: unimodality, equivalence, derangements, and sorting*, arXiv:2602.16355v2 (2026), Section 3.

[10] J. West, *Generating trees and the Catalan and Schröder numbers*, Discrete Mathematics **146** (1995) 247–262.