HOOK INEQUALITIES

IGOR PAK*, FEDOR PETROV†, AND VIACHESLAV SOKOLOV‡

Abstract. We give an elementary proof of the recent *hook inequality* given in [MPP3]:

$$\prod_{u\in\lambda} h(u) \le \prod_{u\in\lambda} h^*(u),$$

where $h(u)$ is the usual hook in a Young diagram λ, and $h^*(i,j) = i + j - 1$. We then obtain a large variety of similar inequalities and their high-dimensional generalizations.

Introduction

In Enumerative Combinatorics, the results are usually easy to state. Essentially, you are counting the number of certain combinatorial objects: exactly, asymptotically, bijectively or otherwise. Judging the importance of the results is also relatively easy: the more natural or interesting the objects are, and the stronger or more elegant is the final formula, the better. In fact, the story or the context behind the results is usually superfluous since they speak for themselves. In the words of Gian-Carlo Rota, one of the founding fathers of modern enumerative combinatorics:

> *"Combinatorics is an honest subject. [...] You count balls in a box, and you either have the right number or you haven't."* [Rota]

It is only occasionally the context makes the difference; this paper is an exception.

The subject of this paper are certain new combinatorial inequalities for hook numbers of Young diagrams, rooted trees and their generalizations. In two special cases these inequalities are known, have technical proofs, and arise in somewhat different areas. Not only do we give elementary proofs of these results, we also setup a new framework which allows us to obtain their advanced generalizations. Viewed by themselves, our most general inequalities overwhelm the senses – they are just too far removed from anything the reader would know and recognize as interesting (see Section 6).

We structure the paper in a way so as to postpone stating the most general results until the end. First, we tell the story of two combinatorial inequalities which would ring a lot of bells for people in the area (sections 1 and 2). We then introduce combinatorial tools to prove the inequalities (*majorization* and *shuffling*), see sections 3 and 4. We then gradually proceed with our generalizations, hoping not to lose the reader in the process.

1. First story: The number of standard Young tableaux

1.1. **Hook-length formula.** Our story starts with the classical *hook-length formula* [FRT] for the number of *standard Young tableaux* of a given shape. This formula is "*a beautiful result in enumerative combinatorics that is both mysterious and extremely well studied*" [MPP1]. We recall it first starting from basic definitions.

*Department of Mathematics, UCLA, Los Angeles, CA, 90095. Email: pak@math.ucla.edu.

†Steklov Mathematical Institute, St. Petersburg, Russia. Email: fedyapetrov@gmail.com.

‡Steklov Mathematical Institute, St. Petersburg, Russia. Email: vi.soksok@gmail.com.

September 6, 2019.

Let $\lambda = (\lambda_1, \ldots, \lambda_\ell)$, $\lambda_1 \geq \ldots \geq \lambda_\ell > 0$ be an *integer partition*. We say that λ is a *partition of* n, write $\lambda \vdash n$ or $|\lambda| = n$, if $\lambda_1 + \lambda_2 + \ldots + \lambda_\ell = n$. Here n is also called the *size* of λ. Let λ' denotes the *conjugate partition* of λ. An example of $\lambda = (4,3,1)$ and $\lambda' = (3,2,2,1)$ is given in Figure 1.

The *Young diagram* of a partition λ is a subset of $\mathbb{N}^2$ (visualized as the set of unit squares), with λ_1 squares in the first row, λ_2 squares in the second row, etc. By tradition, a diagram is drawn as in Figure 1, with the top left corner $(1,1)$. For a square $u = (i,j)$ of the Young diagram, we write $(i,j) \in \lambda$, with the first coordinate increasing downward and the second from left to right.

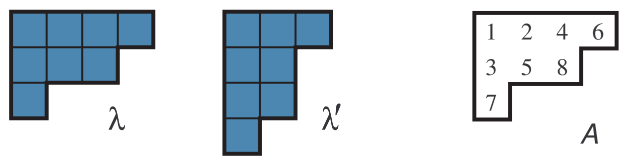

Figure 1. Young diagram $\lambda = (4,3,1)$, conjugate diagram $\lambda' = (3,2,2,1)$, and standard Young tableau $A \in \mathrm{SYT}(\lambda)$.

A *standard Young tableau* (SYT) of shape λ is an array A of shape λ with the numbers $1, \ldots, n$, where $n = |\lambda|$, each appearing once, increasing in rows and columns (see Figure 1). Denote by $\mathrm{SYT}(\lambda)$ the set of standard Young tableaux of shape λ. The number $\left|\mathrm{SYT}(\lambda)\right|$ is fundamental in Algebraic Combinatorics; it is equal to the dimension of the corresponding irreducible representation of S_n and generalizes binomial coefficients, Catalan numbers and ballot numbers. We refer to [Sta, Ch. 7] for background and further references.

The *hook length* $h_\lambda(i,j) := \lambda_i - i + \lambda'_j - j + 1$ of a square $u = (i,j) \in [\lambda]$ is the number of squares directly to the right or directly below u, including u. For a partition $\lambda = (6,6,5,4,1)$ and $u = (2,3)$, we have $h(2,3) = 6$, see Figure 2. The celebrated *hook-length formula* by Frame, Robinson and Thrall [FRT] states:

$$\left|\mathrm{SYT}(\lambda)\right| \, = \, n! \prod_{u \in \lambda} \frac{1}{h(u)}\,, \tag{1.1}$$

(see §7.2). For example, for $\lambda = (4,3,1)$ as in the figure, we have:

$$\left|\mathrm{SYT}(4,3,1)\right| \, = \, \frac{8!}{6 \cdot 4 \cdot 4 \cdot 3 \cdot 2} \, = \, 70.$$

1.2. **Hook inequalities.** We are now ready to state the first result which motivated much of our study. Denote by $h^*(i,j) = i + j - 1$ the *anti-hook length* of the square $u = (i,j) \in \lambda$. It is the number of squares directly to the left or directly above u, including u. For a partition $\lambda = (6,6,5,4,1)$ and $u = (2,3)$, we have $h^*(2,3) = 4$, see Figure 2.

Theorem 1.1 ([MPP3, Prop. 12.1])**.** *For every Young diagram λ, we have:*

$$\prod_{u \in \lambda} h(u) \, \leq \, \prod_{u \in \lambda} h^*(u).$$

Moreover, the equality holds if and only if λ has a rectangular shape.

For example, for $\lambda = (4,3,1)$ as in the figure, we have:

$$6 \cdot 4 \cdot 4 \cdot 3 \cdot 2 \, = \, 576 \, \leq \, 4 \cdot 4 \cdot 3 \cdot 3 \cdot 3 \cdot 2 \cdot 2 \, = \, 1728.$$

Figure 2. Hook length $h(2,3) = 6$, anti-hook length $h(2,3) = 4$ for a partition $\lambda = (6,6,5,4,1)$. Table of hook numbers $\mathrm{H}_\lambda = \{h(i,j)\}$ and table of anti-hook numbers $\mathrm{H}^*_\lambda = \{h^*(i,j)\}$ for $\lambda = (4,3,1) \vdash 8$.

Remark 1.2. The theorem is a corollary of a much more general inequality we prove in [MPP3, Thm 3.3], which in turn is a corollary of the *Naruse hook-length formula* for the number of standard Young tableaux of *skew shape*. This formula was discovered by Naruse in the intersection of Representation Theory and Schubert Calculus; see [MPP1, MPP2] for both algebraic and combinatorial proof and further references. We refer to [Swa] for a purely combinatorial proof of Theorem 1.1, from a very different point of view.

To really understand the above theorem, start with the following:

Observation 1.3. *For every Young diagram λ, we have:*

$$\sum_{u\in\lambda} h(u) \,=\, \sum_{u\in\lambda} h^*(u).$$

Proof. Note that both sides count the number of pairs of squares $x, y \in \lambda$, such that x is either directly above y or directly to the left of y, or $x = y$. □

There are two implications of this observation. First, the theorem cannot have a "trivial" proof by using monotonicity of anti-hooks over hooks. Second, this suggests the distribution of hooks and anti-hooks have the same mean, but the variance of hooks is larger than the variance of anti-hooks. That is, in fact, true and follows from our later results:

Corollary 1.4. *For every Young diagram λ, we have:*

$$\sum_{u\in\lambda} h(u)^2 \,\geq\, \sum_{u\in\lambda} \big(h^*(u)\big)^2.$$

For example, for $\lambda = (4,3,1)$ as above, we have:

$$6^2+4^2+4^2+3^2+2^2+1^2+1^2+1^2 \,=\, 84 \,\geq\, 4^2+4^2+3^2+3^2+3^2+2^2+2^2+1^2 \,=\, 68.$$

1.3. **More hook inequalities.** Here is yet another corollary from our general results. For $(i,j) \in \lambda$, let

$$a(i,j) \,:=\, \big|\{(p,q)\in\lambda \,:\, i\leq p,\, j\leq q\}\big| \quad \text{and} \quad a^*(i,j) \,:=\, i\cdot j\,.$$

Here $a(i,j)$ is the area of λ below and to the right of (i,j). Similarly, $a^*(i,j)$ is the area of λ above and to the left of (i,j). We refer to these as *area* and *anti-area numbers*. For $\lambda = (6,6,5,4,1)$ and $u = (2,3)$ as in Figure 2, we have $a(2,3) = 9$ and $a^*(2,3) = 6$. We have the equality analogous to Observation 1.3 and the following analogue of Theorem 1.1.

Corollary 1.5. *For every Young diagram λ, we have:*

$$\prod_{u\in\lambda} a(u) \,\leq\, \prod_{u\in\lambda} a^*(u).$$

For example, for $\lambda = (4,3,1)$ as in the Figure 1, we have:

$$8 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \,=\, 960 \;\leq\; 6 \cdot 4 \cdot 4 \cdot 3 \cdot 3 \cdot 2 \cdot 2 \,=\, 3456.$$

The reader will have to wait until Section 3 to see how these results fit together.

2. Second story: The number of increasing trees and linear extensions

2.1. **Increasing trees.** Let τ be a *rooted tree* on n vertices, one of which is called the *root* $R \in V$. We draw the trees as in Figure 3. An *increasing tree of shape* τ is a labeling of τ with the numbers $1,\ldots,n$, each appearing once, increasing downwards, away from R. Denote by $\mathrm{IT}(\tau)$ the set of increasing trees of shape τ.

The number $|\mathrm{IT}(\tau)|$ famously has its own analogue of the "hook-length formula" (see §7.2). For a vertex $v \in \tau$, a *branch* in τ below v consists of vertices $w \in \tau$ such that the shortest path from w to R goes through v; denote by $b(v)$ the size of this branch. We have:

$$|\mathrm{IT}(\tau)| \,=\, n! \prod_{v\in\tau} \frac{1}{b(v)}\,. \tag{2.1}$$

For example, for τ as in the figure, we have:

$$|\mathrm{IT}(\tau)| \,=\, \frac{8!}{8\cdot 4\cdot 3\cdot 2} \,=\, 210.$$

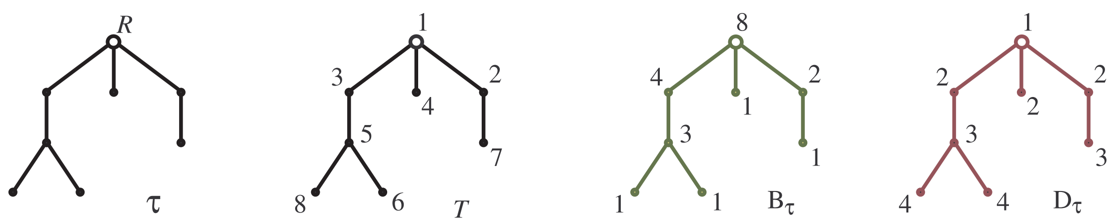

Figure 3. Rooted tree τ with $n = 8$ vertices, increasing tree T of shape τ, table of branch sizes $\mathrm{B}_\tau = \{b(v)\}$ and table of distances $\mathrm{D}_\tau = \{d(v)\}$.

We can now state the inequality analogous to Theorem 1.1. For a vertex $v \in \tau$, denote by $d(v)$ the *distance* from v to the root R, defined as the number of vertices on the shortest path from v to R (see Figure 3).

Theorem 2.1. *For every rooted tree* τ*, we have:*

$$\prod_{v\in\tau} b(v) \,\leq\, \prod_{v\in\tau} d(v).$$

Moreover, the equality holds if and only if τ *is a path with an endpoint at* R*.*

For example, for τ as in Figure 3, we have:

$$8 \cdot 4 \cdot 3 \cdot 2 \,=\, 192 \;\leq\; 4 \cdot 4 \cdot 3 \cdot 3 \cdot 2 \cdot 2 \cdot 2 \,=\, 1152.$$

The theorem is easier than Theorem 1.1 and has a straightforward proof by induction which we leave to the reader. However, the following direct analogue of Observation 1.3 is a warning that it is still nontrivial:

Observation 2.2. *For every rooted tree* τ*, we have:*

$$\sum_{v\in\tau} b(v) \,=\, \sum_{v\in\tau} d(v).$$

We omit the easy double counting proof of the observation. Before we derive Theorem 2.1, let us first state the analogue of Corollary 1.4, which is of independent interest.

Corollary 2.3. *For every rooted tree τ, we have:*

$$\sum_{v\in\tau} b(v)^2 \,\geq\, \sum_{v\in\tau} d(v)^2.$$

Again, we derive the corollary from our general results.

2.2. **Linear extensions.** Let $\mathcal{P}$ be a ranked poset on a finite set X, $|X| = n$, with linear ordering denoted by $\prec$. The *upper order ideal* $\mathcal{P}_{\succeq x} := \{y \in X \ : \ y \succeq x\} \subseteq \mathcal{P}$ is the set of elements in $\mathcal{P}$ greater or equal to x. Similarly, the *lower order ideal* $\mathcal{P}_{\preceq x} := \{y \in X \ : \ y \preceq x\} \subseteq \mathcal{P}$ is the set of elements in $\mathcal{P}$ smaller or equal to x. A *linear extension* of $\mathcal{P}$ is a bijection $f : X \to \{1, \ldots, n\}$ such that $f(x) < f(y)$ for all $x \prec y$. Let $\mathrm{LE}(\mathcal{P})$ be the set of *linear extensions* of $\mathcal{P}$.

Note that when $\mathcal{P}$ is a Young diagram λ or a rooted tree τ with natural ordering, the number $\bigl|\mathrm{LE}(\mathcal{P})\bigr|$ of linear extensions coincides with $\bigl|\mathrm{SYT}(\lambda)\bigr|$ and $\bigl|\mathrm{IT}(\tau)\bigr|$, respectively. However, for general posets the number $\bigl|\mathrm{LE}(\mathcal{P})\bigr|$ is hard to find both theoretically and computationally. In fact, $\bigl|\mathrm{LE}(\mathcal{P})\bigr|$ is #P-complete to compute [BW] (see also [DP]), so no simple product formula is expected to exist in full generality (cf. [Pro]).

Denote by $r_{\prec}(x) := \bigl|\mathcal{P}_{\succeq x}\bigr|$ the number of elements in $\mathcal{P}$ greater or equal to x. We have the following general inequality:

Theorem 2.4 ([HP, Cor. 2])**.** *For every poset $\mathcal{P}$, in the notation above, we have:*

$$\bigl|\mathrm{LE}(\mathcal{P})\bigr| \,\geq\, n! \prod_{x\in\mathcal{P}} \frac{1}{r_{\prec}(x)}\,.$$

This lower bound was proposed by Stanley [Sta, Exc. 3.57] and proved by Hammett and Pittel in [HP] by an involved probabilistic argument. Heuristically, the theorem says that events $A_x :=$ "random bijection $f : X \to \{1, \ldots, n\}$ has $f(x) < f(z)$ for all $z \succ x$" have positive correlation:

$$P(A_x \cap A_y) \,\geq\, P(A_x) \cdot P(A_y),$$

cf. [She].

When $\mathcal{P}$ is a rooted tree τ on n vertices, with the order "$\prec$" increasing down, the inequality in the theorem is an equality, and the sizes of upper order ideal are exactly the branch size numbers: $r_{\prec}(v) = b(v)$. When the order is reversed, we obtain the distance numbers: $r_{\succ}(v) = d(v)$. We thus have:

$$n! \prod_{v\in\tau} \frac{1}{b(v)} \,=\, \bigl|\mathrm{IT}(\tau)\bigr| \,\geq\, n! \prod_{v\in\tau} \frac{1}{d(v)}\,,$$

which implies the inequality in Theorem 2.1.

3. Majorization approach

3.1. **Definitions and background.** Let $\mathbf{a} = (a_1, \ldots, a_n), \mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$ be two non-increasing sequences $a_1 \geq \ldots \geq a_n > 0$ and $b_1 \geq \ldots \geq b_n > 0$. We say that $\mathbf{a}$ *majorizes* $\mathbf{b}$, write $\mathbf{a} \trianglerighteq \mathbf{b}$, if

$$\begin{aligned} & a_1 + \ldots + a_k \,\geq\, b_1 + \ldots + b_k \quad \text{for all } \ 1 \leq k < n, \quad \text{and} \\ & a_1 + \ldots + a_n \,=\, b_1 + \ldots + b_n\,. \end{aligned} \tag{3.1}$$

See [MOA] for a thorough treatment of majorization. We need the following important result:

Theorem 3.1 (Karamata's inequality)**.** *Fix a strictly convex function* $\varphi : \mathbb{R} \to \mathbb{R}$*. Let* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ *be two sequences, s.t.* $\mathbf{a} \trianglerighteq \mathbf{b}$*. Then*

$$\sum_{i=1}^{n} \varphi(a_i) \, \geq \, \sum_{i=1}^{n} \varphi(b_i).$$

Moreover, the equality holds if and only if $\mathbf{a} = \mathbf{b}$*.*

This result is classical and the converse also holds, but we will not need it. For streamlined proofs and further references see e.g. [HLP, §3.17] and [BB, §28, §30].

Since our hook numbers do not have a natural ordering, we will need a multiset analogue of majorization. For two multisets $\mathcal{A}$ and $\mathcal{B}$ with n real numbers, let $\mathbf{a}$ and $\mathbf{b}$ be non-increasing ordering of $\mathcal{A}$ and $\mathcal{B}$, respectively. We say that $\mathcal{A}$ *majorizes* $\mathcal{B}$, write $\mathcal{A} \trianglerighteq \mathcal{B}$, whenever $\mathbf{a} \trianglerighteq \mathbf{b}$. The following is a sufficient condition for majorization.

Lemma 3.2. *Let* $\mathcal{A}, \mathcal{B}$ *two multisets with* n *real numbers. Suppose for all* $\mathcal{B}' \subseteq \mathcal{B}$*, there exist* $\mathcal{A}' \subseteq \mathcal{A}$*, such that* $|\mathcal{A}'| = |\mathcal{B}'|$ *and*

$$\sum_{x \in \mathcal{A}'} x \, \geq \, \sum_{y \in \mathcal{B}'} y \, .$$

Then $\mathcal{A} \trianglerighteq \mathcal{B}$*.*

The lemma is also standard and straightforward. Since we use it repeatedly, we include a simple proof for completeness.

Proof. Let $\mathbf{a}$ and $\mathbf{b}$ be non-increasing ordering of $\mathcal{A}$ and $\mathcal{B}$, respectively. Take $\mathcal{B}' = \{b_1, \ldots, b_k\}$ and the corresponding $\mathcal{A}' = \big\{a_{i_1}, \ldots, a_{i_k}\big\}$, where $1 \leq i_1 < \cdots < i_k \leq n$. We have:

$$b_1 + \ldots + b_k \, \leq \, a_{i_1} + \ldots + a_{i_k} \, \leq \, a_1 + \ldots + a_k \, ,$$

as desired. □

3.2. **Applications to hook inequalities.** Using the notation of Section 1, we can now state the following new result.

Theorem 3.3. *Let* λ *be a Young diagram. Denote by* $\mathcal{H}_\lambda = \big\{h(i,j), (i,j) \in \lambda\big\}$ *the multiset of hook numbers, and by* $\mathcal{H}^*_\lambda = \big\{h^*(i,j), (i,j) \in \lambda\big\}$ *be the multiset of anti-hook numbers. Then* $\mathcal{H}_\lambda \trianglerighteq \mathcal{H}^*_\lambda$*.*

Here the equality part of majorization (3.1) is exactly Observation 1.3. The theorem will be deduced in the next section from Lemma 3.2.

Now, since $\varphi(z) = z^2$ is strictly convex, Theorem 3.3 and Theorem 3.1 imply Corollary 1.4. Similarly, since $\varphi(z) = -\log(z)$ is strictly convex on $\mathbb{R}_{>0}$, we obtain Theorem 1.1 from Theorem 3.3 and Theorem 3.1 by taking logs on both sides. The details are straightforward.

Theorem 3.4. *Let* λ *be a Young diagram. Denote by* $\mathcal{A}_\lambda = \big\{a(i,j), (i,j) \in \lambda\big\}$ *the multiset of area numbers, and by* $\mathcal{A}^*_\lambda = \big\{a^*(i,j), (i,j) \in \lambda\big\}$ *the multiset of anti-area numbers. Then* $\mathcal{A}_\lambda \trianglerighteq \mathcal{A}^*_\lambda$*.*

By analogy with the argument above, we obtain Corollary 1.5 from Theorem 3.4 and Theorem 3.1 by taking logs on both sides.

Theorem 3.5. *Let* τ *be a rooted tree. Denote by* $\mathcal{R}_\tau = \big\{r(v), v \in \tau\big\}$ *the multiset of branch numbers, and by* $\mathcal{D}_\tau = \big\{d(v), v \in \tau\big\}$ *the multiset of distance numbers. Then* $\mathcal{R}_\tau \trianglerighteq \mathcal{D}_\tau$*.*

Now Corollary 2.3 follows from Theorem 3.5 and Theorem 3.1 by taking $\varphi(z) = z^2$, see the argument above.

4. Shuffling in the plane

4.1. **Hook majorization.** For a square $(i, j) \in \lambda$, denote by $\alpha(i, j) = \lambda_i - j$ the *arm* at (i, j), and by $\beta(i, j) = \lambda'_j - i$ the *leg* at (i, j). Similarly, denote by $\alpha^*(i, j) = j$ and $\beta^*(i, j) = i$ the *anti-arm* and *anti-leg*. For a partition $\lambda = (6, 6, 5, 4, 1)$ and $u = (2, 3)$ as in Figure 2, we have $\alpha(2, 3) = 3$, $\alpha^*(2, 3) = 2$, $\beta(2, 3) = 2$ and $\beta^*(2, 3) = 1$.

Shuffling in the plane. Let $X \subseteq \lambda$ be a subset of squares. Perform two steps:
(1) Push all squares in X all the way up inside each column,
(2) Push all squares in X all the way left inside each row.

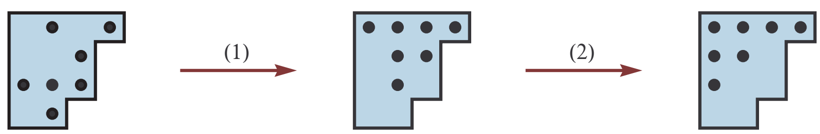

Figure 4. Shuffling of 7 squares in $\lambda = (4, 3, 3, 2)$.

Proof of Theorem 3.3. Take $X = \{x_1, \ldots, x_k\} \subseteq \lambda$ and shuffle it as above. After step (1), we obtain set $X' = \{x'_1, \ldots, x'_k\} \subseteq \lambda$, s.t.

$$\begin{aligned} &\alpha(x'_i) \ge \alpha(x_i), \quad \alpha^*(x'_i) = \alpha^*(x_i) \quad \text{for all} \quad 1 \le i \le k, \ \text{ and} \\ &\beta(x'_1) + \ldots + \beta(x'_k) \ \ge\ \beta^*(x_1) + \ldots + \beta^*(x_k). \end{aligned} \tag{4.1}$$

The first two inequalities hold by definition. The last inequality holds since all squares are pushed up, where they maximize the legs (cf. the proof of Lemma 3.2).

Similarly, after step (2) applied to X', we obtain set $Y = \{y_1, \ldots, y_k\} \subseteq \lambda$, s.t.

$$\begin{aligned} &\alpha(y_1) + \ldots + \alpha(y_k) \ \ge\ \alpha^*(x'_1) + \ldots + \alpha^*(x'_k), \ \text{ and} \\ &\beta(y_i) \ge \beta(x_i), \quad \beta^*(y_i) = \beta^*(x'_i) \quad \text{for all} \quad 1 \le i \le k. \end{aligned} \tag{4.2}$$

Combining equations (4.1) and (4.2) we obtain:

$$\begin{aligned} &\alpha(y_1) + \ldots + \alpha(y_k) \ \ge\ \alpha^*(x'_1) + \ldots + \alpha^*(x'_k) \ =\ \alpha^*(x_1) + \ldots + \alpha^*(x_k), \\ &\beta(y_1) + \ldots + \beta(y_k) \ \ge\ \beta(x'_1) + \ldots + \beta(x'_k) \ \ge\ \beta^*(x_1) + \ldots + \beta^*(x_k). \end{aligned} \tag{4.3}$$

Using

$$h(i, j) \ =\ \alpha(i, j) + \beta(i, j) + 1, \qquad h^*(i, j) \ =\ \alpha^*(i, j) + \beta^*(i, j) + 1$$

we can sum two equations in (4.3) to obtain

$$h(y_1) + \ldots + h(y_k) \ \ge\ h^*(x_1) + \ldots + h^*(x_k). \tag{4.4}$$

In summary, for every $X \subseteq \lambda$ there exists $Y \subseteq \lambda$, $|Y| = |X|$, such that (4.4) holds. Now by Lemma 3.2 we obtain the result. □

4.2. **Proof deconstruction.** Before we proceed to several generalizations let us reformulate the proof of Theorem 3.3 using a different language.

Let $h^\circ(i, j) = j + \lambda'_j - i + 1 = \alpha^*(i, j) + \beta(i, j) + 1$ be the *semi-hook length*. For a given subset X of squares in λ, denote by $h(X)$, $h^\circ(X)$ and $h^*(X)$ the sums of hooks, semi-hooks and anti-hook numbers, respectively.

In the proof above, the sum of inequalities in (4.1) and in (4.2), gives

$$h^*(X) \ \le\ h^\circ(X') \quad \text{and} \quad h^\circ(X') \ \le\ h(Y),$$

respectively. Combined, these give $h^*(X) \le h(Y)$, a consequence of inequalities in (4.3). This implies majorization by Lemma 3.2. $\square$

Note that in the proof we split hooks and anti-hooks into sums of arms/legs and anti-arms/anti-legs, respectively. This is out of necessity as individually these parameters are easier to control. The approach we take below is to split the generalized hooks into even more parameters.

4.3. **Area majorization.** We need new notation. For a square $(i,j) \in \lambda$ and integers $p,q \in \mathbb{Z}$, denote

$$a_{pq}(i,j) \, := \begin{cases} 1 & \text{if } \ (i+p, j+q) \in \lambda, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$a(i,j) \, := \sum_{(p,q)\in\mathbb{N}^2} a_{pq}(i,j)\,, \qquad a^*(i,j) \, := \sum_{(p,q)\in\mathbb{N}^2} a_{-p,-q}(i,j)\,.$$

Indeed, think of subscript (p,q) as the vector shift. Then the first sum is the number of squares in λ to the right and below (i,j), while the second sum is the number of squares in λ to the left and above (i,j), as desired.

Proof of Theorem 3.4. We proceed as in the proof above. Let

$$a^\circ(i,j) \, := \sum_{(p,q)\in\mathbb{N}^2} a_{p,-q}(i,j)\,.$$

For a subset of squares $X \subseteq \lambda$ denote by

$$a_{pq}(X) \, := \sum_{(i,j)\in X} a_{p,q}(i,j), \quad a^\circ_{pq}(X) := \sum_{(i,j)\in X} a^\circ_{p,-q}(i,j), \quad a^*_{pq}(X) := \sum_{(i,j)\in X} a^*_{-p,-q}(i,j).$$

Now consider the shuffling of squares $X \to X' \to Y$ as in the proof above. For every $(p,q) \in \mathbb{N}$, we have:

$$a^*_{pq}(X) \, \le \, a^\circ_{pq}(X') \, \le \, a_{pq}(Y)\,.$$

This follows simply by the geometry of partitions: rows above are larger or equal than rows below, and the columns to the left are larger or equal than columns to the right. Summing this over all $(p,q) \in \mathbb{N}$, we conclude:

$$a^*(X) \, \le \, a^\circ(X') \, \le \, a(Y)\,.$$

This implies majorization by Lemma 3.2. $\square$

4.4. **Generalization to weighted squares.** We need new notation. Fix a function $g : \mathbb{N}^2 \to \mathbb{R}_+$ non-increasing in both coordinates. For a square $(i,j) \in \lambda$, denote

$$\psi_{pq}(i,j) \, := \begin{cases} g(p,q) & \text{if } \ (p+i, q+j) \in \lambda, \\ 0 & \text{otherwise.} \end{cases}$$

and

$$\psi(i,j) \, := \sum_{(p,q)\in\mathbb{N}^2} \psi_{pq}(i,j)\,, \qquad \psi^*(i,j) \, := \sum_{(p,q)\in\mathbb{N}^2} \psi_{i-p,j-q}(p,q)\,.$$

Think of subscript (p,q) as the vector shift. Then $\psi(i,j)$ is a sum of function g over the fourth quadrant starting at (i,j), while $\psi^*(i,j)$ is a sum of function g over the second quadrant starting at (i,j).

For example, when $g(p,q) = 1$ we get the area and anti-area numbers: $\psi(i,j) = a(i,j)$ and $\psi^*(i,j) = a^*(i,j)$. Similarly, let

$$g(p,q) \, := \begin{cases} 1 & \text{if } \ p = 0 \ \text{ or } \ q = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.5}$$

Then we get the hook and anti-hook numbers in this case: $\psi(i,j) = h(i,j)$ and $\psi^*(i,j) = h^*(i,j)$.

Theorem 4.1. *Let λ be a Young diagram and fix $g : \mathbb{N}^2 \to \mathbb{R}_+$. Denote by*

$$\Psi\langle\lambda, g\rangle := \big\{\psi(i,j),\,(i,j) \in \lambda\big\} \quad \text{and} \quad \Psi^*\langle\lambda, g\rangle := \big\{\psi^*(i,j),\,(i,j) \in \lambda\big\}$$

the multisets of numbers defined above. Then $\Psi\langle\lambda, g\rangle \trianglerighteq \Psi^*\langle\lambda, g\rangle$.

This result is a common generalization of both Theorem 3.4 and Theorem 3.3. In the examples above, when $g(p,q) = 1$ we have $\Psi\langle\lambda, g\rangle = \mathcal{A}_\lambda$ and $\Psi^*\langle\lambda, g\rangle = \mathcal{A}^*_\lambda$. Similarly, for g as in (4.5), we have $\Psi\langle\lambda, g\rangle = \mathcal{H}_\lambda$ and $\Psi^*\langle\lambda, g\rangle = \mathcal{H}^*_\lambda$. Since the proof follows verbatim the proof of Theorem 3.4, we omit the details.

5. Generalized shuffling

5.1. **Hooks in space.** In $\mathbb{R}^3$, *solid partitions* Λ are defined as lower order ideals of the poset $\big(\mathbb{N}^3, \prec\big)$, with the natural ordering making $(0,0,0)$ the smallest (see §2.2). Here 'down along the x-axis' means decreasing the x-coordinate, and similarly with 'up' and other axes.

Note that solid partitions generalize Young diagrams and can be viewed as arrangements of unit cubes, such that whenever $(i,j,k) \in \Lambda$ we also have $(p,q,r) \in \Lambda$ for all $p \le i$, $q \le j$, and $r \le k$.

Now, there are two natural notions of the hooks: 1- and 2-dimensional. More precisely, let

$$R(i,j,k) := \Big|\big\{(p,j,k) \in \Lambda \,:\, i \le p\big\} \cup \big\{(i,q,k) \in \Lambda \,:\, j \le q\big\} \cup \big\{(i,j,r) \in \Lambda \,:\, k \le r\big\}\Big|$$

$$Q(i,j,k) := \Big|\big\{(p,q,k) \in \Lambda \,:\, i \le p, j \le q\big\} \cup \big\{(p,j,r) \in \Lambda \,:\, i \le p, k \le r\big\} \cup \big\{(i,q,r) \in \Lambda \,:\, j \le q, k \le r\big\}\Big|$$

These are subsets of squares of Λ along the 1-dimensional rays or 2-dimensional quadrants emanating from a cube $(i,j,k) \in \Lambda$. The anti-hooks $R^*(i,j,k)$ and $Q^*(i,j,k)$ are defined analogously, with the direction of rays/quadrants reversed.

Theorem 5.1. *Let Λ be a solid partition. Denote by $\mathcal{R}_\Lambda$, $\mathcal{Q}_\Lambda$, $\mathcal{R}^*_\Lambda$, and $\mathcal{Q}^*_\Lambda$ the multisets of 1- and 2-dimensional hook numbers and anti-hook numbers defined above. Then $\mathcal{R}_\Lambda \trianglerighteq \mathcal{R}^*_\Lambda$ and $\mathcal{Q}_\Lambda \trianglerighteq \mathcal{Q}^*_\Lambda$.*

We prove the theorems later in this section. But first, by analogy with the arguments in Section 3, we obtain the following generalizations of Theorem 1.1.

Corollary 5.2. *Let Λ be a solid partition. Then:*

$$\prod_{(i,j,k)\in\Lambda} R(i,j,k) \le \prod_{(i,j,k)\in\Lambda} R^*(i,j,k), \quad \text{and} \quad \prod_{(i,j,k)\in\Lambda} Q(i,j,k) \le \prod_{(i,j,k)\in\Lambda} Q^*(i,j,k).$$

There is also volume and anti-volume numbers, generalizing area and anti-area numbers:

$$V(i,j,k) := \Big|\big\{(p,q,r) \in \Lambda \,:\, i \le p, j \le q, k \le r\big\}\Big|,$$

$$V^*(i,j,k) := \Big|\big\{(p,q,r) \in \Lambda \,:\, p \le i, q \le j, r \le k\big\}\Big|.$$

Theorem 5.3. *Let Λ be a solid partition. Denote by $\mathcal{V}_\Lambda$ and $\mathcal{V}^*_\Lambda$ the multisets of volume and anti-volume numbers defined above. Then $\mathcal{V}_\Lambda \trianglerighteq \mathcal{V}^*_\Lambda$.*

This is a generalization of Theorem 3.4. By the same argument as above, we have the following natural result generalizing Corollary 1.5.

Corollary 5.4. *Let Λ be a solid partition. Then:*

$$\prod_{(i,j,k)\in\Lambda} V(i,j,k) \le \prod_{(i,j,k)\in\Lambda} V^*(i,j,k)\,.$$

This corollary has a curious implication. By Theorem 2.4, we have two lower bounds

$$\big|\mathrm{LE}(\mathcal{P}_\Lambda)\big| \,\ge\, n! \prod_{(i,j,k)\in\Lambda} \frac{1}{V(i,j,k)}\,, \quad \text{and} \quad \big|\mathrm{LE}(\mathcal{P}_\Lambda)\big| \,\ge\, n! \prod_{(i,j,k)\in\Lambda} \frac{1}{V^*(i,j,k)}\,, \tag{5.1}$$

where $\mathcal{P}_\Lambda$ is the subposet of $\mathbb{N}^3$ restricted to Λ. Corollary 5.4 implies that the first of the bounds in (5.1) is always sharper than the second. Of course, when Λ is a box these lower bounds coincide.

5.2. **Shuffling in space.** We give a brief outline of the proof of theorems above. First, define the shuffling in the space as follows.

Shuffling in $\mathbb{N}^3$. Let $X \subseteq \Lambda$ be a subset of cubes. Perform three steps:
(1) Push all cubes in X all the way down along the x axis,
(2) Push all cubes in X all the way down along the y axis,
(3) Push all cubes in X all the way down along the z axis.

Let us concentrate on the 1-dimensional hook numbers. We introduce two intermediate hook numbers: $R^\circ(i,j,k)$ and $R^{\circ\circ}(i,j,k)$ with two rays pointing up along the y,z axes, and one ray pointing up along the z axis, respectively. We then prove that

$$\sum_{(i,j,k)\in X} \mathcal{R}^*(i,j,k) \,\le \sum_{(i,j,k)\in X'} \mathcal{R}^\circ(i,j,k) \,\le \sum_{(i,j,k)\in X''} \mathcal{R}^{\circ\circ}(i,j,k) \,\le \sum_{(i,j,k)\in Y} \mathcal{R}(i,j,k),$$

where $X \to_{(1)} X' \to_{(2)} X'' \to_{(3)} Y$ is the result of making shuffling steps. Finally, by Lemma 3.2 we obtain the first part of Theorem 5.1. We omit the details.

5.3. **Back to trees.** Now that the reader is exhausted by 3-dimensional generalizations of the hooks, the reader can restart anew with the following:

Shuffling in trees. Let $X \subseteq \tau$ be a subset of vertices and let $w \in X$ be a vertex with the smallest distance $d(w)$. Perform the following steps:
• Move w into R. Now do the shuffling in each branch in $\tau \smallsetminus R$ by induction.

Sketch of proof of Theorem 3.5. For a set of vertices $X \subseteq \tau$, denote by Y the result of the shuffling. We need to show that

$$\sum_{v\in Y} d(v) \,\le\, \sum_{v\in X} r(v)\,.$$

Observe that $r(R) = |\tau| \ge d(w)$, so moving w to the root is always advantageous for the sum of branch numbers. We continue by induction leaving the details to the reader. □

6. Greatest generalization

Let us now state the most general version of the hook inequalities that we promised in the introduction. Denote by $T_1,\ldots,T_d$ infinite locally finite rooted trees. We view each T_i as a poset with ordering towards the root R_i, and denote by d_i the distance function to R_i in the tree T_i. Consider a poset $\mathcal{P} := T_1 \times \ldots \times T_d$ with the natural ordering towards the root, and let Ω be a finite lower ideal in $\mathcal{P}$.

Let $g : \mathbb{N}^d \to \mathbb{R}_+$ be a fixed weight function non-increasing in all coordinates. We say that elements $\boldsymbol{v} = (v_1,\ldots,v_d)$ and $\boldsymbol{w}(w_1,\ldots,w_d)$ in Ω have a $\boldsymbol{m} = (m_1,\ldots,m_d)$ shift, if $v_i \preccurlyeq w_i$ and

$d(w_i) - d(v_i) = m_i$ for all i. In this case we write $\boldsymbol{v} \to_{\boldsymbol{m}} \boldsymbol{w}$. We can now define the hook and anti-hook numbers

$$H(\boldsymbol{v}) := \sum_{\boldsymbol{m}\in\mathbb{N}^d} g(\boldsymbol{m}) \sum_{\boldsymbol{w}\in\Omega} 1_{\{\boldsymbol{v}\to_{\boldsymbol{m}}\boldsymbol{w}\}}\,, \quad \text{and}$$
$$H^*(\boldsymbol{v}) := \sum_{\boldsymbol{m}\in\mathbb{N}^d} g(\boldsymbol{m}) \sum_{\boldsymbol{w}\in\Omega} 1_{\{\boldsymbol{w}\to_{\boldsymbol{m}}\boldsymbol{v}\}}\,.$$

Theorem 6.1. *Let Ω be a lower order ideal in the product of trees $T_1 \times \ldots \times T_d$ poset. Fix $g : \mathbb{N}^d \to \mathbb{R}_+$. Denote by $\mathcal{H}_{\Omega,g} := \big\{H(\boldsymbol{v}),\, \boldsymbol{v} \in \Omega\big\}$ and $\mathcal{H}^*_{\Omega,g} := \big\{H^*(\boldsymbol{v}),\, \boldsymbol{v} \in \Omega\big\}$ the multisets of hook and anti-hook numbers defined above. Then $\mathcal{H}_\Omega \trianglerighteq \mathcal{H}^*_\Omega$.*

We leave to the reader to see how this result generalizes all previous results, and what inequalities follow by Theorem 3.1. The shuffling in the product of trees is also straightforward: first shuffle in T_1, then in T_2, etc. We omit the proof.

7. Final remarks

7.1. There is no universal agreement on the term "anti-hook". While we follow e.g. [Jon], the term "cohook" is also frequently used, see e.g. [BLRS, Sul]. Since both "cohook" and "anti-hook" are also used in other contexts, we opted for the latter due to personal preferences.

7.2. The hook-length formula (1.1) has a number of different proofs (bijective, probabilistic, analytic, etc.) as well as generalizations different from Naruse's. The literature is too numerous to include; see [CKP, §6.2] for a brief survey, and [Sta, Ch. 7] for further references.

The hook-length formula for trees (2.1) is due to Knuth [Knu, §5.1.4, Exc. 20], and can be easily proved by induction via removing the root. Probabilistic and bijective proofs are given in [SY] and [Beá], respectively.

7.3. Both hook-length formulas (1.1) and (2.1) have well understood q-analogues in terms of the numbers of certain descents of the standard Young tableaux and increasing trees, respectively (see e.g. [CKP, Sta, MPP1]). Since $\varphi(n) := (n)_q = 1+q+\ldots+q^{n-1} = (1-q^n)/(1-q)$ is logarithmically concave in n, this gives a q-deformation of Theorem 1.1. It would be interesting to see if our hook inequalities have natural q-analogues as polynomials in q (i.e. for each coefficient), cf. [Bre, §2].

7.4. For solid partitions, there is no known closed formula for the number of linear extensions. However, in this case both lower bounds (5.1) are not very sharp and weaker than other, more elementary bounds (see [MPP3]).

7.5. Despite their origin, in some ways our hook inequalities have a different nature than the inequalities usually studied in Enumerative and Algebraic Combinatorics, whose proofs are non-robust and rely on the global algebraic structure (see e.g. [Brä, Pak]). Instead, they resemble many isoperimetric and other geometric inequalities, where the proofs are obtained by incremental improvements, sometimes disguised by the variational principle (see e.g. [BZ]). In fact, it is not hard to give an analytic generalization of Theorem 4.1 to the case when λ is replaced by a monotone curve, and g is replaced by a Lebesgue measurable non-negative function. It would be interesting to see if there are any new applications of this generalization.

7.6. Swanson writes: "It would be interesting to find a bijective explanation of [Thm 1.1]" [Swa]. The methods of this paper are elementary, combinatorial in nature, but fundamentally non-bijective. They are likely to be the most "bijective explanation" possible in this case. In fact, we believe that our shuffling approach is truly a "proof from the book".

Acknowledgements. We are grateful to Joshua Swanson for telling us about [Swa] and for the careful reading of the manuscript. The first author is thankful to Alejandro Morales and Greta Panova for numerous interesting conversations about the Naruse hook-length formula. We are also grateful to BIRS in Banff, Canada, for hosting the first two authors at the *Asymptotic Algebraic Combinatorics* workshop in March 2019, when this paper was finalized. This collaboration started when the first author asked a question on `MathOverflow`.[1] The first author was partially supported by the NSF. The second and the third authors were partially supported by the RSCF (grant 17-71-20153).

References

[BLRS] K. Barrese, N. Loehr, J. Remmel and B. E. Sagan, m-level rook placements, *J. Combin. Theory, Ser. A* **124** (2014), 130–165.

[Beá] B. Beáta, Bijective proofs of the hook formula for rooted trees, *Ars Combin.* **106** (2012), 483–494.

[BB] E. Beckenbach and R. Bellman, *Inequalities*, Springer, Berlin, 1961

[Brä] P. Brändén, Unimodality, log-concavity, real-rootedness and beyond, in *Handbook of Enumerative Combinatorics*, CRC Press, Boca Raton, FL, 2015, 437–483; `arXiv:1410.6601`.

[Bre] F. Brenti, Log-concave and unimodal sequences in algebra, combinatorics, and geometry: an update, in *Jerusalem Combinatorics*, AMS, Providence, RI, 1994, 71–89

[BW] G. Brightwell and P. Winkler, Counting linear extensions, *Order* **8** (1991), 225–247.

[BZ] Yu. D. Burago and V. A. Zalgaller, *Geometric inequalities*, Springer, Berlin, 1988.

[CKP] I. Ciocan-Fontanine, M. Konvalinka and I. Pak, The weighted hook length formula, *J. Combin. Theory, Ser. A* **118** (2011), 1703–1717.

[DP] S. Dittmer and I. Pak, Counting linear extensions of restricted posets, preprint (2018), 33 pp.; `arXiv:`1802.06312.

[FRT] J. S. Frame, G. de B. Robinson and R. M. Thrall, The hook graphs of the symmetric group, *Canad. J. Math.* **6** (1954), 316–324.

[HP] A. Hammett and B. Pittel, How often are two permutations comparable? *Trans. AMS* **360** (2008), 4541–4568.

[HLP] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities* (Second ed.), Cambridge Univ. Press, 1952.

[Jon] B. F. Jones, Singular Chern classes of Schubert varieties via small resolution, *Int. Math. Res. Not.* **2010** (2010), 1371–1416.

[Knu] D. Knuth, *The Art of Computer Programming*, Vol. III, Addison-Wesley, Reading, MA, 1973.

[MOA] A. Marshall, I. Olkin and B. C. Arnold, *Inequalities: theory of majorization and its applications* (Second ed.), Springer, New York, 2011.

[MPP1] A. H. Morales, I. Pak and G. Panova, Hook formulas for skew shapes I. q-analogues and bijections, *J. Combin. Theory, Ser. A* **154** (2018), 350–405.

[MPP2] A. H. Morales, I. Pak and G. Panova, Hook formulas for skew shapes II. Combinatorial proofs and enumerative applications, *SIAM J. Discrete Math.* **31** (2017), 1953–1989.

[MPP3] A. H. Morales, I. Pak and G. Panova, Asymptotics for the number of standard Young tableaux of skew shape, *European J. Combin.* **70** (2018), 26–49.

[Pak] I. Pak, Combinatorial inequalities, *Notices of the AMS*, to appear; available at `https://tinyurl.com/y26efyoj`.

[Pro] R. A. Proctor, d-Complete posets generalize Young diagrams for the hook product formula, in *RIMS Kôkyûroku* **1913** (2014), 120–140.

[Rota] G.-C. Rota and D. Sharp, Mathematics, Philosophy and Artificial Intelligence, dialogue in *Los Alamos Science*, No. 12 (Spring/Summer 1985), 94–104.

[SY] B. E. Sagan and Y. N. Yeh, Probabilistic algorithms for trees, *Fibonacci Quart.* **27** (1989), 201–208.

[She] L. A. Shepp, The XYZ conjecture and the FKG inequality, *Ann. Probab.* **10** (1982), 824–827.

[Sta] R. P. Stanley, *Enumerative Combinatorics*, vol. 1 (Second ed.) and 2, Cambridge Univ. Press, 2012 and 1999.

[Sul] R. Sulzgruber, Symmetry properties of the Novelli-Pak-Stoyanovskii algorithm, in *Proc. 26th FPSAC (Chicago, USA)*, DMTCS, Nancy, 2014, 205–216.

[Swa] J. P. Swanson, On the existence of tableaux with given modular major index, *Algebraic Combinatorics* **1** (2018), 3–21.

[1] See http://mathoverflow.net/q/243846.