# FindStat – a database and search engine for combinatorial statistics and maps

Martin Rubey and Christian Stump

## Definitions

- *combinatorial collection:* a collection $\mathcal{C} = \bigcup_n \mathcal{C}_n$ of finite sets (e.g. the set of permutations)
- *combinatorial map:* a map $\phi : \mathcal{C} \to \mathcal{C}'$ between collections (e.g. the inverse of a permutation)
- *combinatorial statistic:* a map $\beta : \mathcal{C} \to \mathbb{Z}$ (e.g. the order of a permutation)

## Feature

*Given (a few) values of a combinatorial statistic $\alpha : \mathcal{C} \to \mathbb{Z}$, find maps $\phi_1$, $\phi_2$, $\phi_3$ and a statistic $\beta$ in the database such that*

$$\alpha = \beta \circ \phi_3 \circ \phi_2 \circ \phi_1 \qquad \text{(search for values)}$$

$$\text{or} \quad \sum_{c \in \mathcal{C}_n} x^{\alpha(c)} = \sum_{c \in \mathcal{C}_n} x^{\beta \circ \phi_3 \circ \phi_2 \circ \phi_1(c)} \qquad \text{(search for distribution)}$$

# identify your statistic

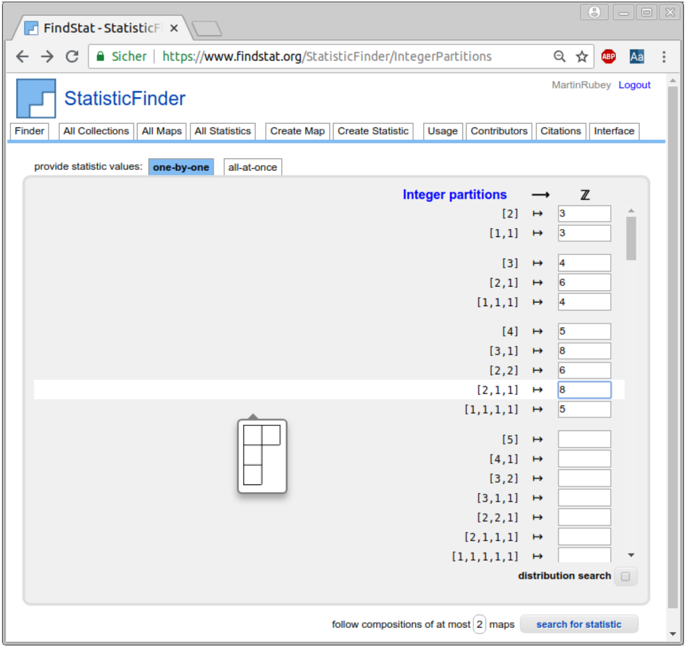
`www.mathoverflow.net/q/132338`:

> *I've come across a function from the set of integer partitions to the natural numbers which I don't recognise but which probably ought to be familiar.*

- $f(\emptyset) = 1$
- $f(\lambda) = \binom{i+j}{i} f(\mu) f(\nu)$
  - $(i, j)$: coordinates of a box with $i + j$ is maximal
  - $\mu$: remove first $i$ rows from $\lambda$
  - $\nu$: remove first $j$ columns from $\lambda$

---

```
[1] => 2,
[2] => 3, [1,1] => 3,
[3] => 4, [2,1] => 6, [1,1,1] => 4,
[4] => 5, [3,1] => 8, [2,2]   => 6, [2,1,1] => 8, [1,1,1,1] => 5
```


FindStat - StatisticF
Sicher
https://www.findstat.org/StatisticFinder/IntegerPartitions
StatisticFinder
MartinRubey
Logout
Finder
All Collections
All Maps
All Statistics
Create Map
Create Statistic
Usage
Contributors
Citations
Interface
provide statistic values:
one-by-one
all-at-once
Integer partitions ⟶ ℤ
[2] ↦ 3
[1,1] ↦ 3
[3] ↦ 4
[2,1] ↦ 6
[1,1,1] ↦ 4
[4] ↦ 5
[3,1] ↦ 8
[2,2] ↦ 6
[2,1,1] ↦ 8
[1,1,1,1] ↦ 5
[5] ↦
[4,1] ↦
[3,2] ↦
[3,1,1] ↦
[2,2,1] ↦
[2,1,1,1] ↦
[1,1,1,1,1] ↦
distribution search
follow compositions of at most 2 maps
search for statistic

over all $\mu$, as noticed by Philippe Nadeau in a comment.
This statistic arises in the homogeneous Garnir relations for the universal graded Specht modules for cyclotomic quiver Hecke algebras.

Matches **St000085** ○ **Mp00026** ○ **Mp00043** in **10 values** Result quality: 100%

**Mapping** Mp00043: Integer partitions —to Dyck path⟶ Dyck paths
Mp00026: Dyck paths —to ordered tree⟶ Ordered trees
St000085: Ordered trees ⟶ $\mathbb{Z}$

**Values**

```
[2] => [1,1,0,0,1,0] => [[[]],[]] => 3
[1,1] => [1,0,1,1,0,0] => [[],[[]]] => 3
[3] => [1,1,1,0,0,0,1,0] => [[[[]]],[]] => 4
[2,1] => [1,0,1,0,1,0] => [[],[],[]] => 6
[1,1,1] => [1,0,1,1,1,0,0,0] => [[],[[[]]]] => 4
[4] => [1,1,1,1,0,0,0,0,1,0] => [[[[[]]]],[]] => 5
[3,1] => [1,1,0,1,0,0,1,0] => [[[],[]],[]] => 8
[2,2] => [1,1,0,0,1,1,0,0] => [[[]],[[]]] => 6
[2,1,1] => [1,0,1,1,0,1,0,0] => [[],[[],[]]] => 8
[1,1,1,1] => [1,0,1,1,1,1,0,0,0,0] => [[],[[[[]]]]] => 5
```

**Description** The number of linear extensions of the tree.
We use Knuth's hook length formula for trees [pg.70, 1]. For an ordered tree $T$ on $n$ vertices, the number of linear extensions is

$$\frac{n!}{\prod_{v\in T} |T_v|},$$

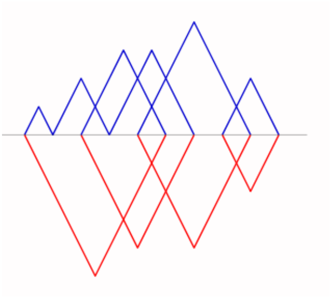
where $T_v$ is the number of vertices of the subtree rooted at $v$.

Matches **St000110** ○ **Mp00025** ○ **Mp00043** in **10 values** Result quality: 100%

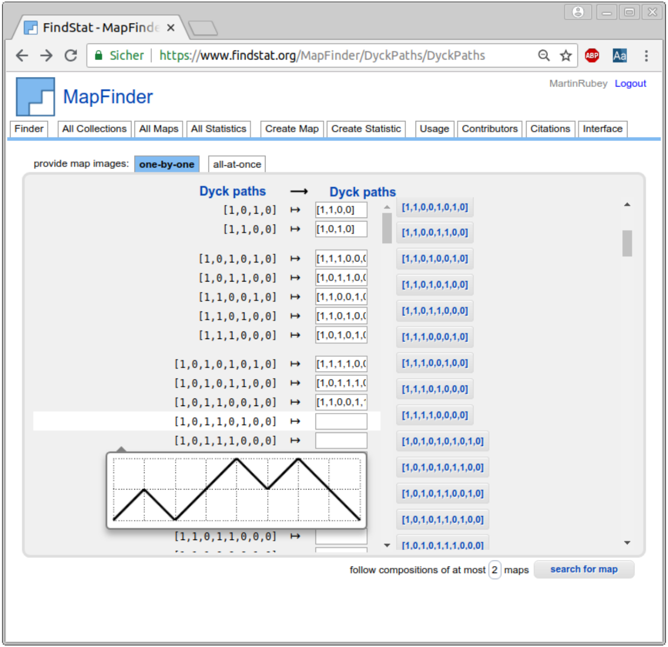
**Mapping** Mp00043: Integer partitions —to Dyck path⟶ Dyck paths
Mp00025: Dyck paths —to 132-avoiding permutation⟶ Permutations

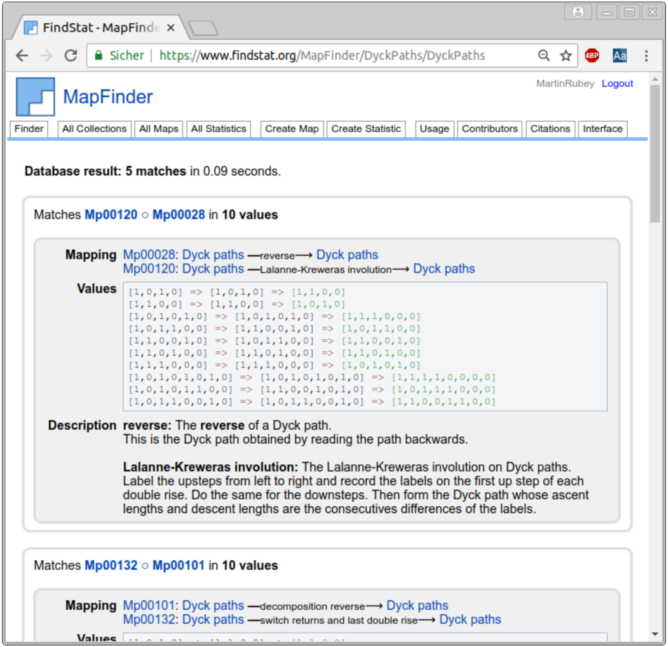
# identify your map

https://irma.math.unistra.fr/~chapoton/dycatalan.html:



*On a représenté un chemin de Dyck en rouge et son dual en bleu. Cette dualité est une involution qui échange les chemins à k pics avec les chemins à n-k pics. Cette involution commute à la symétrie droite-gauche. Le principe est le suivant : on décompose le chemin de Dyck en ses montagnes comme sur la figure. Puis on impose que chaque point de la ligne horizontale soit le lieu d'une et une seule "réflexion ou réfraction". On obtient ainsi un unique chemin de Dyck dual.*


FindStat - MapFinde
Sicher
https://www.findstat.org/MapFinder/DyckPaths/DyckPaths
MartinRubey Logout
MapFinder
Finder
All Collections
All Maps
All Statistics
Create Map
Create Statistic
Usage
Contributors
Citations
Interface
provide map images: one-by-one all-at-once
Dyck paths ⟶ Dyck paths
[1,0,1,0] ↦ [1,1,0,0]
[1,1,0,0] ↦ [1,0,1,0]
[1,0,1,0,1,0] ↦ [1,1,1,0,0,(
[1,0,1,1,0,0] ↦ [1,0,1,1,0,(
[1,1,0,0,1,0] ↦ [1,1,0,0,1,(
[1,1,0,1,0,0] ↦ [1,1,0,1,0,(
[1,1,1,0,0,0] ↦ [1,0,1,0,1,(
[1,0,1,0,1,0,1,0] ↦ [1,1,1,1,0,(
[1,0,1,0,1,1,0,0] ↦ [1,0,1,1,1,(
[1,0,1,1,0,0,1,0] ↦ [1,1,0,0,1,
[1,0,1,1,0,1,0,0] ↦
[1,0,1,1,1,0,0,0] ↦
[1,1,0,1,1,0,0,0] ↦
[1,1,0,0,1,0,1,0]
[1,1,0,0,1,1,0,0]
[1,1,0,1,0,0,1,0]
[1,1,0,1,0,1,0,0]
[1,1,0,1,1,0,0,0]
[1,1,1,0,0,0,1,0]
[1,1,1,0,0,1,0,0]
[1,1,1,0,1,0,0,0]
[1,1,1,1,0,0,0,0]
[1,0,1,0,1,0,1,0,1,0]
[1,0,1,0,1,0,1,1,0,0]
[1,0,1,0,1,1,0,0,1,0]
[1,0,1,0,1,1,0,1,0,0]
[1,0,1,0,1,1,1,0,0,0]
follow compositions of at most 2 maps
search for map


FindStat - MapFinde
Sicher | https://www.findstat.org/MapFinder/DyckPaths/DyckPaths
MapFinder
MartinRubey Logout
Finder
All Collections
All Maps
All Statistics
Create Map
Create Statistic
Usage
Contributors
Citations
Interface
Database result: 5 matches in 0.09 seconds.
Matches Mp00120 ○ Mp00028 in 10 values
Mapping
Mp00028: Dyck paths —reverse⟶ Dyck paths
Mp00120: Dyck paths —Lalanne-Kreweras involution⟶ Dyck paths
Values
[1,0,1,0] => [1,0,1,0] => [1,1,0,0]
[1,1,0,0] => [1,1,0,0] => [1,0,1,0]
[1,0,1,0,1,0] => [1,0,1,0,1,0] => [1,1,1,0,0,0]
[1,0,1,1,0,0] => [1,1,0,0,1,0] => [1,0,1,1,0,0]
[1,1,0,0,1,0] => [1,0,1,1,0,0] => [1,1,0,0,1,0]
[1,1,0,1,0,0] => [1,1,0,1,0,0] => [1,1,0,1,0,0]
[1,1,1,0,0,0] => [1,1,1,0,0,0] => [1,0,1,0,1,0]
[1,0,1,0,1,0,1,0] => [1,0,1,0,1,0,1,0] => [1,1,1,1,0,0,0,0]
[1,0,1,0,1,1,0,0] => [1,1,0,0,1,0,1,0] => [1,0,1,1,1,0,0,0]
[1,0,1,1,0,0,1,0] => [1,0,1,1,0,0,1,0] => [1,1,0,0,1,1,0,0]
Description
reverse: The reverse of a Dyck path.
This is the Dyck path obtained by reading the path backwards.
Lalanne-Kreweras involution: The Lalanne-Kreweras involution on Dyck paths.
Label the upsteps from left to right and record the labels on the first up step of each double rise. Do the same for the downsteps. Then form the Dyck path whose ascent lengths and descent lengths are the consecutives differences of the labels.
Matches Mp00132 ○ Mp00101 in 10 values
Mapping
Mp00101: Dyck paths —decomposition reverse⟶ Dyck paths
Mp00132: Dyck paths —switch returns and last double rise⟶ Dyck paths
Values

# refine your equidistribution

Arbeitsgemeinschaft Diskrete Mathematik Wien, November 2018:

*Are there any equidistribution results for triples of permutation statistics* $(\varepsilon, \sigma, \mu)$*, such that*

$\varepsilon$ *is Eulerian*
*(equidistributed with the number of descents),*
$\sigma$ *is a Stirling statistic*
*(equidistributed with the number of cycles),*
$\mu$ *is a Mahonian statistic*
*(equidistributed with the Major index)?*

Idea:

- Let $E$ be all Eulerian statistics,
- let $S$ be all Stirling statistics,
- let $M$ be all Mahonian statistics `www.findstat.org/St000004`

for each pair in $E$ select all pairs in $S$, and then all pairs in $M$, such that the result is jointly equidistributed...

# refine your equidistribution

| | | |
|---|---|---|
| cyc: | number of cycles | www.findstat.org/St000031 |
| sal: | number of saliances | www.findstat.org/St000007 |
| srt: | sorting index | www.findstat.org/St000224 |
| ninv: | number of non-inversions | www.findstat.org/St000246 |
| dfc: | number of deficiencies | www.findstat.org/St000703 |
| leh: | number of repeated entries in the Lehmer code | www.findstat.org/St001298 |

## Conjecture

$$\sum_{\pi \in \mathcal{S}_n} x^{\mathsf{cyc}(\pi)} y^{\mathsf{srt}(\pi)} z^{\mathsf{dfc}(\pi)} = \sum_{\sigma \in \mathcal{S}_n} x^{\mathsf{sal}(\pi)} y^{\mathsf{ninv}(\pi)} z^{\mathsf{leh}(\pi)}$$

# further ideas

- compute mean and variance for statistics in the database, check whether they are given by polynomials
- use full text search to quickly find definitions of maps or statistics
- use generalized distribution search if you know which values occur, but not for which elements
- browse identities for maps
- ...

# metastatistics

- www.findstat.org/Contributors
- www.findstat.org/Citations