Descent algebras

related topics: noncommutative symmetric functions / quasisymmetric functions
hyperplane arrangements
random walks on a group
reflection / Coxeter groups

Work in the symmetric group algebra $\mathbb{R}S_n$
(the permutations are a basis, add pointwise, multiply by linear extension of compositions of permutations.)

To interpret the multiplication in $\mathbb{R}S_n$: suppose $\alpha = \sum_{\sigma \in S_n} a_\sigma \sigma$; $\beta = \sum_{\sigma \in S_n} b_\sigma \sigma$.
then the coefficient of σ in $\alpha\beta$ is the probability of going from the identity to σ by multiplying by a random permutation with probability a_σ, then multiplying by another random permutation with probability b_σ.

e.g. if $\pi_\sigma^{(k)}$ is the chance of moving from the identity to σ in k repeats of the same process, and $\rho = \sum_{\sigma \in S_n} \pi_\sigma^{(1)} \sigma$, then $\rho^k = \sum_{\sigma \in S_n} \pi_\sigma^{(k)} \sigma$

Suppose ρ models a card-shuffle. Then a question of interest is, how many shuffles are necessary to mix the deck?

The deck is well-mixed if it is equally likely to be in any of the $n!$ orders
$\therefore$ define U to be $\sum_{\sigma \in S_n} \frac{1}{n!} \sigma$

Then the rigorous mathematical question is, given $\varepsilon > 0$, what value of k ensures $\| \rho^k - U \| < \varepsilon$, for some metric $\| \ \|$ on $\mathbb{R}S_n$?

The study of riffle-shuffles uses the following "magic" formula in $\mathbb{R}(\mathbb{S}_n)[t]$:

$$\sum_{i=1}^{n} e_i t^i = \sum_{\sigma \in \mathbb{S}_n} \frac{(t-d(\sigma))^{(n)}}{n!}\sigma$$

here, e_i are particular elements of $\mathbb{R}(\mathbb{S}_n)$

$d(\sigma)$ is the number of descents in σ — i.e. $|\{i \in \{1,\dots,n\} \mid \sigma_{i+1} > \sigma_i\}|$

$x^{(n)}$ denotes the increasing factorial $x(x+1)\cdots(x+n-1)$.

e.g. in $\mathbb{S}_3$, the right hand side is

$$\frac{t(t+1)(t+2)}{6}[\![123]\!] + \frac{(t-1)t(t+1)}{6}\left([\![132]\!]+[\![213]\!]+[\![231]\!]+[\![312]\!]\right) + \frac{(t-2)(t-1)t}{6}[\![321]\!]$$

so, collecting the powers of t, we get $e_1 = \frac{1}{3}[\![123]\!] + \frac{1}{6}$

$$e_2 = \frac{1}{2}[\![123]\!] - \frac{1}{2}[\![321]\!]$$

$$e_3 = \frac{1}{6}\left([\![123]\!]+[\![132]\!]+[\![213]\!]+[\![231]\!]+[\![312]\!]+[\![321]\!]\right)$$

It is always the case that $e_n = \frac{1}{n!}\sum_{\sigma \in \mathbb{S}_n}\sigma$.

The e_i turn out to be orthogonal idempotents — i.e. $e_i^2 = e_i$, and $e_i e_j = 0$ if $i \neq j$.
Taking $t=1$ in the magic formula: $e_1 + e_2 + \cdots + e_n = \mathrm{id}$

Taking $t=2$ in the magic formula:
because $x^{(n)} = 0$ if $-n+1 \leq x \leq 0$, the right hand side is $(n+1)\,\mathrm{id} + \sum_{\sigma \in \mathbb{S}_n, d(\sigma)=1}\sigma$.
and, dividing this by 2^n gives the probabilities of the riffle-shuffle.
(see the paper by Bayer-Diaconis.)

So we are looking for the minimal k such that $\left(\frac{1}{2^n}\sum_{i=1}^{n} e_i 2^i\right)^k$ is close to uniform.

But, because the e_i are orthogonal idempotents, this is simply $\frac{1}{2^{nk}}\sum_{i=1}^{n} e_i 2^{ik}$.
Using the right-hand-side of the magic formula, we take the coefficients of each permutation, and discover that the k-step transition probabilities are

$$\pi_\sigma^{(k)} = \frac{(2^k - d(\sigma))^{(n)}}{n!}$$

Note that this is non-zero only when $2^k > d(\sigma)$ $\therefore$ this is non-zero for all σ when $k \sim \log_2 n$. And indeed, Bayer-Diaconis shows that $\frac{3}{2}\log_2 n$ riffle shuffles are necessary to mix a deck.
(The idempotent e_1 is originally due to Michael Barr, from Hochschild cohomology)

The descent algebra is originally due to Solomon:

let $D(\sigma) := \{ i \mid 1 \le i \le n-1,\ \sigma_i > \sigma_{i+1} \} \subseteq \{1, 2, \dots, n-1\}$

in $\mathbb{Q}S_n$, let $y_T := \sum_{D(\sigma)=T} \sigma$, where T is any of the 2^{n-1} subsets of $\{1, 2, \cdots, n-1\}$

e.g for $n=3$: $y_\emptyset = [\![123]\!]$ ($y_\emptyset$ = identity, always)

$$y_{\{1\}} = [\![213]\!] + [\![312]\!]$$
$$y_{\{2\}} = [\![132]\!] + [\![231]\!]$$
$$y_{\{1,2\}} = [\![321]\!]$$

Theorem of Solomon (which holds for all finite Coxeter groups)

the y_T span a subalgebra of $\mathbb{Q}S_n$: this is $\Sigma[S_n]$, the descent algebra

Equivalently, for two subsets T, R of $\{1, 2, \cdots, n-1\}$, then $y_T y_R = \sum a^K_{TR} y_K$.

Indeed, the a^K_{TR} are integral, and Solomon gives an expression for them explicitly.

Since $y_\emptyset$ is the identity permutation, the descent algebra is an algebra with identity.

e.g. for $n=3$: $y_\emptyset$ acts as the identity.

$$y_{\{1,2\}}\, y_{\{1,2\}} = y_\emptyset$$
$$y_{\{1,2\}}\, y_{\{1\}} = y_{\{2\}}$$
$$y_{\{1,2\}}\, y_{\{2\}} = y_{\{1\}}$$
$$y_{\{1\}}\, y_{\{1\}} = y_\emptyset + y_{\{2\}} + y_{\{1,2\}}.$$

<u>Coxeter groups</u>: these are pairs (W, S) such that W is a finite group, and S generates W

$s, t \in S$ satisfy the relations $s^2 = id$, $(st)^{m_{st}} = id$ for some $m_{st} \in \mathbb{N}$, with $m_{st} \ge 2$.

if $m_{s,t} = 2$, then $(st)(st) = id \Rightarrow \cancel{s}\cancel{t}s\cancel{t}\cancel{t} = st \Rightarrow ts = st$. i.e. s, t commute.

If S is the disjoint union $S_1 \amalg S_2$, and $m_{st} = 2$ for all $s \in S_1$, $t \in S_2$

then the group W generated by S is a cartesian product $W_1 \times W_2$, where S_i generates W_i

$\therefore$ it is enough to understand irreducible Coxeter groups – these have been classified by Coxeter.

If we draw a graph whose vertices are S, and draw $s \quad t$ if $m_{st} = 2$

s —— t if $m_{st} = 3$

s ═══ t if $m_{st} = 4$,

then the irreducible Coxeter groups correspond to connected graphs.

Example: the symmetric group S_n corresponds to the graph •—•—•—• ··· —• with $n-1$ vertices

the generator s_i is the transposition $(i, i+1)$
the relations are $(s_i s_j)^2 = id$ if $|j-i| \geq 2$
$(s_i s_{i+1})^3 = id$.

For any $w \in W$, let $l(w)$ be the length of a minimal-length expression of w as a product of generators in S.
Then, for $w \in W$, its descent set is $D(w) = \{i \mid l(ws_i) < l(w)\} \subseteq S$
In the case of the symmetric group: right-multiplication by s_i exchanges the images of i and $i+1$ — i.e. it sends $\sigma_1\sigma_2 \cdots \sigma_n$ to $\sigma_1\sigma_2 \cdots \sigma_{i-1}\sigma_{i+1}\sigma_i\sigma_{i+2}\cdots\sigma_n$.
The length of a permutation is its number of inversions: $|\{(i,j) \mid i<j, \sigma_i > \sigma_j\}|$
And exchanging the images of i and $i+1$ reduces the number of inversions precisely when $\sigma_{i+1} > \sigma_i$ ∴ this more general definition of descents agree with our previous definition.

The general version of Solomon's theorem:
for any subset T of S, set $y_T := \sum_{D(w)=T} w$,
then the y_T span a subalgebra of $\mathbb{Q}W$ of dimension $2^{|S|}$, and there is an explicit description of the coefficients in the products $y_T y_R$.

In the case of the symmetric group, the basis elements y_T can also be indexed by compositions c of n (written $c \vDash n$), since there is a classical bijection between subsets of $n-1$ and compositions of n: $\{i_1, \ldots, i_\ell\} \to (i_1, i_2-i_1, i_3-i_2, \ldots)$

There is a second basis for $\Sigma[S_n]$: let $B_T := \sum_{D(\sigma) \subseteq T} \sigma$ — i.e. the permutations who might have a descent at T. The change from B_T to y_T is upper-unitriangular, so B_T is a basis. Its multiplication table is much easier, and proves that $\Sigma[S_n]$ is an algebra: $B_C B_D = \sum_M B_{C(M)}$ where the sum is over all matrices M with non-negative integral entries whose j^{th} column sums to c_j (j^{th} part of the composition C), and whose i^{th} row sums to d_i.
Then $C(M)$ is the composition formed by reading the matrix entries along the rows from left to right, from the top row to the bottom row, and deleting zeroes.

In $\mathbb{Q}[S_n]$, there are two anti-isomorphic algebras: the descent algebra $\Sigma[S_n]$, and l'algèbre de battage B_n.

the anti-isomorphism is given by $\sigma \to \sigma^{-1}$ (so $\alpha\beta \to \beta^{-1}\alpha^{-1}$, hence an anti-isomorphism).

The image of B_T under this anti-isomorphism (in B_n) is ($T = \{t_1, t_2, \dots, t_k\}$)

$12\cdots t_1 \sqcup\!\sqcup\, t_1+1\, t_1+2 \cdots t_1+t_2 \sqcup\!\sqcup \cdots \sqcup\!\sqcup\, t_k+1,\, t_k+2 \cdots n$.

(e.g. the image of $B_{\{3,5\}}$ is $123 \sqcup\!\sqcup 45 \sqcup\!\sqcup 678$)

(Here, $\sqcup\!\sqcup$ denotes the shuffle product: for α, β who between them have distinct letters, $\alpha \sqcup\!\sqcup \beta := \sum_\gamma$ over all γ which contain α, β as complimentary substrings — i.e. $\alpha = \gamma_{i_1}\gamma_{i_2}\cdots\gamma_{i_k}$, $\beta = \gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_l}$, with $i_1 < i_2 < \cdots < i_k$; $j_1 < j_2 < \cdots < j_l$, $\{i_1, i_2, \dots, i_k\} \sqcup\!\sqcup \{j_1, j_2, \dots, j_l\} = \{1, 2, \dots\}$. This is an associative product.)

These images of B_T multiply by an analogous matrix rule, except the matrix composition is read down the columns, from the left column to the right column. (This is easy to prove, and the multiplication of B_T may be deduced from this.)

The same multiplication rule describes the Kronecker product (also called internal product) of complete symmetric functions of the same degree. (One definition of this product is that the power sums satisfy $p_\lambda \cdot p_\mu = 0$ if $\lambda \neq \mu$, and $p_\mu \cdot p_\mu = z_\mu p_\mu$, where z_μ is such that $\frac{n!}{z_\mu}$ is the number of permutations of cycle type μ. Equivalently, p_λ / z_λ are orthogonal idempotents.)

As a result, we can define a surjective algebra morphism $\Psi : \Sigma[S_n] \to \Lambda_n$, $\Psi(B_c) = h_c$. This surjection is in fact split — i.e. $\Sigma[S_n]$ is a direct sum of the kernel of S_n and a subspace which is isomorphic to Λ_n via Ψ. So we can take the preimages of p_λ / z_λ in this subspace — these are the idempotents E_λ of Garsia and Reutenauer, and summing over all λ of the same number of parts give the e_i from before. More details to follow.

It's easy to verify that the sum of orthogonal idempotents is also an idempotent. We are interested in primitive idempotents, which cannot be written as the sum of orthogonal idempotents.

In a finite-dimensional algebra, orthogonal idempotents are necessarily linearly independent. So the number of mutually orthogonal idempotents is at most the dimension of the algebra. Any set of mutually orthogonal idempotents $\{e_1, e_2, \dots, e_r\}$ can be extended so that $\sum e_i$ = unit of algebra: this is because, if e is an idempotent, then unit $- e$ is an idempotent orthogonal

to e. If $\sum e_i = 1$, then $\{e_1, \dots, e_r\}$ is a complete family of idempotents.
Note that complete families of primitive orthogonal idempotents
for example, for the algebra of 2×2 upper-triangular matrices, $\{\begin{pmatrix}1&0\\0&0\end{pmatrix}, \begin{pmatrix}0&0\\0&1\end{pmatrix}\}$
and $\{\begin{pmatrix}1&1\\0&0\end{pmatrix}, \begin{pmatrix}0&-1\\0&1\end{pmatrix}\}$ are both such families.
Idempotents are important for studying the representations of an algebra.
Expanding an algebra element in terms of idempotents is akin to a Fourier transform.

Let $p(x)$ be a polynomial $(x-a_1)(x-a_2)\cdots(x-a_n)$, where a_i are pairwise distinct.
Then, in the algebra, $\mathbb{C}[x]/\langle p(x)\rangle$, a complete family of primitive orthogonal
idempotents are the Lagrange interpolation formulae: $e_k(x) = \dfrac{(x-a_1)\cdots(x-a_{k-1})(x-a_{k+1})\cdots(x-a_n)}{(a_k-a_1)\cdots(a_k-a_{k-1})(a_k-a_{k+1})\cdots(a_k-a_n)}$
This is because every element of $\mathbb{C}[x]/\langle p(x)\rangle$ is determined by its value
on $a_1, a_2, \dots, a_n$, and e_k is the indicator function on a_k.
So idempotents of algebras are often interesting.

Because $\{p_\lambda/z_\lambda\}$ gives a basis of Λ_n, they must be a complete family of
primitive orthogonal idempotents. The unit of the Kronecker product of Λ_n is
$\sum_{\lambda\vdash n} p_\lambda/z_\lambda = h_n$.
It happens that their preimages E_λ is a complete family of primitive orthogonal
idempotents for $\mathbb{C}[S_n]$.
Analogous families exist for other Coxeter groups - but what plays the role of
partitions? What is the indexing set?

Let $H(z)$, $E(z)$, $P(z)$ be generating functions for the complete, elementary and
power sum symmetric functions: e.g. $H(z) = \sum_{n=0}^{\infty} h_n z^n$.
Then the following relations hold: $H(z) = e^{P(z)}$, $E(z) = e^{-P(-z)}$, $H(z)E(-z) = 1$

$$P(z) = \log(H(z)) = -\log(E(-z))$$
$$P'(z) = H'(z)/H(z)$$

These allow us to express one basis in terms of another.

To define a noncommutative analogue of the symmetric functions, we take S_i to
be generators analogous to h_i, and define other bases via noncommutative
interpretations of a relation from a basis to h_i. For instance, one analogue Ψ_k of

p_k is defined via $\sum \Psi_k z^k := (\sum_{n\geq 1} n S_n z^n)(\sum_{k\geq 1}(-1)^k(\sum_{n\geq 1} S_n z^n)^k)$, where multiplication of S_i on the right hand side is noncommutative. Using a different relation between p_λ and h_λ may give a different basis.
The S_i are non-commutative generators of this algebra SYM, so a basis of SYM is indexed by compositions. The basis elements indexed by compositions of n span the subspace of degree n.
There is a graded projection $\pi: SYM \to \Lambda$ sending S_i to h_i, which preserves degree.
The map $\Psi: \sum[\mathbb{S}_n] \to \Lambda_n$ lifts to a map $\Theta: \sum[\mathbb{S}_n] \to SYM_n$, by sending B_T to S_T.
Θ is a (vector space) isomorphism. Hence Θ allows the definition of an internal ($SYM_n \times SYM_n \to SYM_n$) product on SYM_n (although it is more useful to define it to be the opposite of the product on $\sum[\mathbb{S}_n]$). Similarly, Θ equips $\bigoplus_n \sum[\mathbb{S}_n]$ with an external ($\sum[\mathbb{S}_i] \times \sum[\mathbb{S}_j] \to \sum[\mathbb{S}_{i+j}]$) product. This writing of $\sum[\mathbb{S}_n]$ for all n is useful for studying the idempotents of $\sum[\mathbb{S}_n]$. (Indeed, the study of descent algebras was much of the motivation for the definition of SYM.) Expressing the idempotents of $\sum[\mathbb{S}_n]$ in terms of the basis B_T is analogous to changing from noncommutative power sums to the basis S_T -- and this involves the external product.

Now we look more closely at the idempotents E_λ. Recall that they do not form a basis.
However it is possible to get a basis I_α such that $I_\alpha I_\beta = \begin{cases} w(\mu)\, I_\beta & \text{if } \lambda(\alpha)=\lambda(\beta)=\mu \\ 0 & \text{otherwise} \end{cases}$

Here, $\lambda(\alpha)$ is the partition obtained by putting the parts of α in decreasing order, and $w(\mu)$ is the number of compositions whose partition is μ
So $I_\alpha / w(\lambda(\alpha))$ is an idempotent, but these are not orthogonal.
Indeed, $I_\alpha := \sum_{\alpha \leq \beta} \frac{(-1)^{\ell(\beta)-\ell(\alpha)}}{f(\alpha,\beta)} B_\beta$ (so the sum is over all refinements β of α – i.e β is the concatenation $\beta^{(1)}\beta^{(2)}\cdots\beta^{(\ell(\alpha))}$ with $\beta^{(i)} \models \alpha_i$.)
where $f(\alpha,\beta)$ is the product of the lengths of the $\beta^{(i)}$
Inversing this relation shows $B_\alpha = \sum_{\alpha\leq\beta} \frac{1}{g(\alpha,\beta)} I_\beta$,

where $g(\alpha,\beta)$ is the product of the factorials of the lengths of the $\beta^{(i)}$.
It is a theorem of Garsia/Reutenauer that the I_α multiply as described above, and that the orthogonal idempotents E_λ are sums of the I_α over rearrangements of λ:
$E_\mu = \frac{1}{\ell(\mu)!} \sum_{\lambda(\alpha)=\mu} I_\alpha$

To understand this, work in SYM. Define elements F_m via the generating function equality

$$\sum_{n=1}^{\infty} F_n x^n := \log\left(1+\sum_{k=1}^{\infty} S_k x^k\right) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\left(\sum_{k=1}^{\infty} S_k x^k\right)^n \quad \left(\text{i.e. use Taylor series for } \log(1+x)\right)$$

For example: $F_1 = S_1$

$$F_2 = S_2 - \tfrac{1}{2} S_{11}$$

$$F_3 = S_3 - \tfrac{1}{2}(S_{12} + S_{21}) + \tfrac{1}{3} S_{111}.$$

Also, define $\mathcal{E}_\lambda$ by

$$\sum_\lambda \mathcal{E}_\lambda x_\lambda := \exp\left(\sum_{n\geq 1} F_n x_n\right)$$

Note that this involves infinitely-many variables x_n, and x_λ denotes $x_{\lambda_1} x_{\lambda_2} \cdots x_{\lambda_{\ell(\lambda)}}$.

Finally, let E_λ and I_α be the images of $\mathcal{E}_\lambda$ and of $F_{\alpha_1} \cdots F_{\alpha_\ell}$ under the identification Θ^{-1} of SYM with $\Sigma_1[\mathbb{S}_n]$.

(These calculations are in the paper of Garsia/Reutenauer, though not in the language of SYM, as the paper predates the definition of SYM)

For example, $F_3 F_1 F_2 = S_{312} - \frac{1}{2} S_{1212} - \frac{1}{2} S_{2112} + \frac{1}{3} S_{11112} - \frac{1}{2} S_{3111} + \frac{1}{4} S_{12111} + \frac{1}{4} S_{21111} - \frac{1}{6} S_{111111}$

so $I_{312} = B_{312} - \frac{1}{2} B_{1212} - \frac{1}{2} B_{2112} + \frac{1}{3} B_{11112} - \frac{1}{2} B_{3111} + \frac{1}{4} B_{12111} + \frac{1}{4} B_{21111} - \frac{1}{6} B_{111111}$

which agrees with the formula at the start of this section, in terms of summing over refinements. (This explains why those formulae resemble the series for log and exp).

We have analogous idempotents E_λ for the descent algebras of all finite Coxeter groups. Details are in the paper by Bergeron/Bergeron/Howlett/Taylor.

For example, take the dihedral group with presentation $\langle s, r \mid s^2 = r^2 = 1, (sr)^p = 1 \rangle$ for some $p \geq 3$. This has $2p$ elements: $1, s, r, sr, rs, srs, rsr, \ldots$

In the notation of Solomon, for a subset A of the generating reflections, $x_A := \sum_{\mathcal{D}(w)\cap A = \emptyset} w$, the sum of all elements which lengthen when multiplied by each element of A.

So we find $x_{r,s} = 1$

$$x_r = 1 + s + rs + \cdots$$

$$x_s = 1 + r + sr + \cdots$$

$$x_\emptyset = \sum_{w\in W} w$$

When p is even, the multiplication table of these elements read:

	1	x_r	x_s	x_ϕ
1	1	x_r	x_s	x_ϕ
x_r	x_r	$2x_r + \frac{p-2}{2} x_\phi$	$\frac{p}{2} x_\phi$	$p x_\phi$
x_s	x_s	$\frac{p}{2} x_\phi$	$2x_s + \frac{p-2}{2} x_\phi$	$p x_\phi$
x_ϕ	x_ϕ	$p x_\phi$	$p x_\phi$	$2p x_\phi$

And from this we find a complete family of primitive orthogonal idempotents:

$$e_{r,s} = 1 - \tfrac{1}{2} x_r - \tfrac{1}{2} x_s + \tfrac{p-1}{2p} x_\phi$$
$$e_r = \tfrac{1}{2}(x_r - \tfrac{1}{2} x_\phi)$$
$$e_s = \tfrac{1}{2}(x_s - \tfrac{1}{2} x_\phi)$$
$$e_\phi = \tfrac{1}{2p} x_\phi$$

The case for p odd is similar, but in that case the descent algebra is not commutative, so the analogous formulae result in e_r, e_s which are not orthogonal. Hence there are only three primitive orthogonal idempotents: $e_{r,s}$, $e_r + e_s$, e_ϕ.

The analysis for type B, the hyperoctahedral group, was less immediate.
(This is the group of $n \times n$ matrices with a single non-zero entry in each row and column, and each such entry is 1 or -1 e.g. $\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. Its cardinality is $2^n n!$)
It took a while to find an analogue of the symmetric group formula $I_\alpha = \sum_{\alpha \le \beta} \frac{(-1)^{\ell(\beta)-\ell(\alpha)}}{f(\alpha,\beta)} B_\beta$,

because the coefficients here are in fact simplified from $\prod_{i=1}^{\ell(\alpha)} \frac{(-1)(-2)\cdots(-\ell(\beta^{(i)})+1)}{(2)(3)\cdots(\ell(\beta^{(i)}))}$,

and $1, 2, \cdots, n-1$ are the exponents of the group S_n, being one less than the degrees $2, 3, \ldots, n$.
These two sequences are defined for any Coxeter group, and for B_n they are $1, 3, 5, \ldots, 2n-1$ and $2, 4, 6, \ldots, 2n$.
Thus one advantage of generalising away from S_n to other Coxeter groups is that it "explains" the results in S_n.

It is sometimes useful to view the symmetric functions as operators on the variables $x_1, x_2, \cdots$. Denote this set of variables by $X := x_1 + x_2 + \cdots$
Now impose these formal calculation rules, for $f, g \in \Lambda$, and a constant c:

$$(f \cdot g)[X] = f[X]\, g[X]; \quad (f+g)[X] = f[X] + g[X]; \quad c[X] = c$$

Consequently, to calculate the result of any symmetric function acting on any set of variables, it suffices to calculate the actions of the power sums p_k. (explicitly, expand $f \in \Lambda$ as $f = \sum_{\lambda \vdash n} a_\lambda p_\lambda$; then $f[X] = \sum a_\lambda \prod_{i=1}^{\ell(\lambda)} (p_{\lambda_i}[X])$)
These actions are defined by, for "expressions" A and B:

$p_k[A+B] = p_k[A] + p_k[B]$ $\qquad$ $p_k[A-B] = p_k[A] - p_k[B]$

$p_k[AB] = p_k[A]\, p_k[B]$ $\qquad$ $p_k\left[\frac{A}{B}\right] = p_k[A] / p_k[B]$

$p_k[x] = x^k$ if x is a variable, $\quad p_k[c] = c$ if c is a constant

(so one must specify in an expression A what is a variable and what is a constant.)

Since the symmetric functions are operators, they need not commute with other operators e.g. if ev_{-1} denotes evaluation at $x=-1$, then $p_k\, ev_{-1}[x] = p_k[-1] = -1$, but $ev_{-1}\, p_k[x] = ev_{-1}[x^k] = (-1)^k$.

One application of this operator viewpoint is to define the coproduct on Λ:
$\Delta(f) = \sum_i g_i \otimes h_i$ if $f[X+Y] = \sum_i g_i[X]\, h_i[Y]$.

Theory of species.

Start with two categories: $\mathcal{B}$ = category of finite sets, with bijections
$\mathcal{V}$ = category of finite dimensional vector spaces over some field K, with invertible linear transformations
Consider the functor: $\mathcal{B} \to \mathcal{V}$, sending a set to a vector space with this set as basis.

In this setup, a species is a (covariant) functor $F: \mathcal{B} \to \mathcal{B}$ — so, given a finite set A, F constructs another set $F[A]$, and $F[A]$ and $F[B]$ are in bijection if A, B are in bijection.
Associate to a species F a generating function $F(x) := \sum f_n \frac{x^n}{n!}$ where $f_n := |F[\{1, \cdots, n\}]|$

Here are some operations on species, which are all functorial:
The sum $(F+G)[A]$ is the disjoint union $F[A] + G[A]$
The product $(F \cdot G)[A]$ is $\sum_{B+C=A} F[B] \times G[C]$ — that is, split A into a disjoint union of B and C, and take the cartesian product of $F[B]$ and $G[C]$.
The substitution $(F \circ G)[A]$, defined when $G[\phi] = \phi$, is $\sum_{\pi \in \mathrm{Part}(A)} F(\pi) \times \prod_{B \in \pi} G[B]$ — that is, partition the set A, take the cartesian product of F[set of blocks] with G[restriction of A to each block].
It is not hard to show that the corresponding operations on the generating functions are $(F+G)(x) = F(x) + G(x)$
$(F \cdot G)(x) = F(x) \cdot G(x)$ (the $n!$ in the denominator is required for this)
$(F \circ G)(x) = F(G(x))$

Because F is a functor, there is an action of S_A, the permutation group on A, on $F[A]$.
So, for each $\sigma \in S_A$, we can define $\mathrm{Fix}_F(\sigma) = \{x \in F[A] \mid \sigma \cdot x = x\}$
And we have a representation of S_A on the vector space on $F[A]$ (i.e. the image of $F[A]$ in $\mathcal{V}$), whose character is the cardinalities of $\mathrm{Fix}_F(\sigma)$.
A refinement of the generating function $F(x)$ is the cycle index series of F (as from Polya theory): $Z_F := \sum_{n \geq 0} \frac{1}{n!} \sum_{\sigma \in S_n} |\mathrm{Fix}_F(\sigma)| \, p_{\text{cycle type of } \sigma}$ (a symmetric function of unbounded degree)
This is the Frobenius characteristic of the representations of S_n on $F[\{1, \cdots, n\}]$.
To recover $F(t)$ from Z_F, we take the coefficients of $p_{1,1,\cdots,1}$ (which would count the fixed points for the identity permutation) — equivalently, we take the inner product with $\frac{1}{1-p_1 t}$.

The species operations also transfer well to the cycle index series:

$Z_{F+G} = Z_F + Z_G$

$Z_{F\cdot G} = Z_F Z_G$

$Z_{F\circ G} = Z_F[Z_G]$ (plethysm)

Note that the cardinality of $\mathrm{Fix}_F(\sigma)$ depends only on the cycle type of σ — write this as $\mathrm{fix}_F(\lambda)$, where λ is the cycle type of σ.

Then Z_F simplifies as $Z_F = \sum_{n\geq 0}\sum_{\lambda\vdash n} \mathrm{fix}_F(\lambda)\frac{p_\lambda}{z_\lambda}$, where $\frac{n!}{z_\lambda}$ is the number of permutations of cycle type λ.

In this expression, the plethysm $Z_F[Z_G]$ is $\sum_\lambda \mathrm{fix}_F(\lambda)\frac{p_\lambda[Z_G]}{z_\lambda}$.

Example: F is the species of simple graphs — i.e. $F[A]$ is the set of all simple graphs whose vertex set is A (equivalently, $F[A]$ is the set of subsets of $\mathbb{P}_2[A] := \{(x,y) \mid x,y\in A,\ x\neq y\}$.

Example: Let $\mathbb{S}[A]$ be the set of permutations on A (i.e. bijections $: A\to A$)

let $\mathbb{L}[A]$ be the set of lists on A (i.e. bijections $\{1,2,\cdots,|A|\}\to A$)

$\mathbb{S}(x) = \mathbb{L}(x) = \sum_{n\geq 0} n!\frac{x^n}{n!} = \frac{1}{1-x}$

However, these are not isomorphic: $\mathbb{S}_A$ acts on $\mathbb{S}[A]$ by conjugation, but on $\mathbb{L}[A]$ by multiplication (the difference is that $\{1,2,\cdots,|A|\}$ is ordered, but A is not)

Example: let E be the set species — that is, $E[A] := \{A\}$

let $\mathcal{C}$ be the species of cycles — $\mathcal{C}[A] := \{\sigma : A\to A \text{ a cyclic bijection}\}$.

Then the fact that "a permutation is a disjoint union of cycles" translates to $\mathbb{S} = E\circ\mathcal{C}$ (where "=" means isomorphic via a natural transformation).

One can also view a permutation as the disjoint union of a fixed point set and a derangement: $\mathbb{S} = E\cdot D$, where D is the species of derangements.

Since $E(x) = \sum_{n=0}^{\infty}\frac{x^n}{n!} = e^x$, we find that $\frac{1}{1-x} = \mathbb{S}(x) = E(x)D(x) = e^x D(x)$, so $D(x) = \frac{e^{-x}}{1-x}$.

For any species F, define $F^+[A] := \begin{cases} F[A] & \text{if } A \neq \emptyset \\ \emptyset & \text{if } A = \emptyset \end{cases}$; $F_n[A] := \begin{cases} F[A] & \text{if } |A| = n \\ \emptyset & \text{otherwise} \end{cases}$

Example: The species PART of partitions is $E \circ E^+$ – a partition is a set of non-empty sets. Its series is $PART(x) = E(E^+(x)) = e^{e^x - 1}$.

Example: The endofunctions species is $ENDO[A] := \{f \mid f: A \to A\}$. We can visualize such functions as a graph on A, connecting each element to its image. Then each connected component is a cycle, with rooted trees (edges directed towards the roots) attached to each node in the cycle. Hence an endofunction is a permutation of rooted trees: $ENDO = \mathbb{S} \circ \mathcal{A}$, where $\mathcal{A}$ is the species of rooted trees. Hence $\frac{1}{1-\mathcal{A}(x)} = \mathbb{S}(\mathcal{A}(x)) = ENDO(x) = \sum_{n=0}^{\infty} n^n \frac{x^n}{n!}$.

$\mathcal{A}$ has the recursive definition $\mathcal{A} = X \cdot (E \circ \mathcal{A})$. Here X is the singleton indicator:

$X[A] := \begin{cases} \{A\} & \text{if } |A| = 1 \\ \emptyset & \text{otherwise} \end{cases}$

(so $X \cdot E[A] = A$ – it picks out the elements of a set.) So the recursive definition says that a rooted tree is a root to which is attached a set of rooted trees.

Example: A vertebrate is a list of rooted trees (think of connecting the roots in a backbone): $VERT = \mathbb{L} \circ \mathcal{A}$.

Given a rooted tree, picking any node creates a vertebrate – let the root be the head in the backbone, and the picked node the tail. This is an example of the pointing operation on species: $F^\bullet[A] := A \times F[A]$. It is obvious that pointing multiplies the coefficient of x^n in a generating series by n.

Hence $\mathcal{A}^\bullet = VERT$, and $VERT(x) = \mathbb{L} \circ \mathcal{A}(x) = \mathbb{S} \circ \mathcal{A}(x) = ENDO(x) = \sum_{n=0}^{\infty} n^n \frac{x^n}{n!}$

Hence the number of rooted trees on n vertices is n^{n-1}. (Proof due to Andre Joyal)

Example: The cycle index series of the set species is $Z_E = \sum_\lambda \frac{p_\lambda}{z_\lambda} = \sum_{n=0}^{\infty} h_n = e^{\sum_{k \geq 1} p_k/k}$

Example: To find the cycle index series of the list species $\mathbb{L}$, first note that $\mathbb{L}_n = X^{\cdot n}$

Hence $Z_{\mathbb{L}} = \sum_{n=0}^{\infty} (Z_X)^n = \frac{1}{1-Z_X} = \frac{1}{1-p_1}$.

Example: The cycle index series for partitions is $Z_{PART} = Z_E[Z_{E^+}]$

$$= e^{\sum_{k\geq 1} \frac{1}{k} p_k [e^{\sum_{j\geq 1} p_j/j} - 1]} = \prod_{j,k\geq 1} \frac{1}{1-p_j^k}, \text{ after some work.}$$

Polya theory is concerned with counting objects up to a group action – i.e. counting the number of orbits. The generating series for these counts can be obtained as $\langle Z_F, \sum_{n=0}^{\infty} h_n \rangle$ – this uses Burnside's lemma that the number of orbits of a group action is equal to the average number of fixed points.

Example: To find the decomposition into irreducibles of the S_n-representation on pairs, view the set of pairs as the species $E_2 \cdot E_{n-2}$. This converts the problem into finding the Schur-expansion of $Z_{E_2 \cdot E_{n-2}} = Z_{E_2} Z_{E_{n-2}}$ $= h_2 h_{n-2} = h_2 s_{n-2}$. Now use the Pieri-rule: $h_2 s_{n-2} = s_n + s_{n-1,1} + s_{n-2,2}$ for $n \geq 4$.

Similarly, the decomposition into irreducibles of the permutation action of S_n is given by $h_1 s_{n-1} = s_n + s_{n-1,1}$, for $n \geq 2$.

In both cases, the copy of the trivial representation is spanned by the sum of the objects/pairs, and a complement is spanned by linear combinations of objects/pairs whose coefficients sum to zero.

In general, expanding a cycle-index series in Schur functions is a hard problem.

One way to construct the irreducible representation of S_n corresponding to the partition λ is in the polynomial ring $\mathbb{R}[x_1, \cdots, x_n]$: fill the partition of shape λ in all possible ways with $x_1, \cdots, x_n$ each appearing once, and take the space spanned by $(\prod_{i,j:\, x_i \text{ appears above } x_j} x_i - x_j)$ over all such fillings.

For example, the irreducible representation for the partition $(n-1,1)$ is spanned by $\{x_i - x_j\}$, since there is one column of length two and all other columns have length 1.

Other irreducible representations come from multiplying this space by a symmetric function.

A similar construction exists for other Coxeter groups.

Recall that a finite Coxeter group W is generated by reflections in hyperplanes. Thus W acts on $\mathbb{R}^n$ for some n, and also on polynomials in n variables. We are interested in the polynomials which are invariant under this action.
For example, the hyperoctahedral group B_n acts as signed permutation matrices, and its invariants are symmetric functions is $x_1^2, x_2^2, \cdots, x_n^2$.

It is a theorem that, for every group generated by reflections, its ring of invariants is a polynomial ring in certain homogeneous polynomials $f_1, f_2, \cdots, f_n$. The f_i are not unique, but their degrees are. In fact, a group is generated by reflections if and only if its ring of invariants is polynomial.

If d_i are the degrees of the generating polynomials f_i, then the dimensions of the ring of invariants in each degree is encoded by $\prod_{i=1}^{n} \frac{1}{1-q^{d_i}}$
In the case of the symmetric group, we can choose the generating polynomials f_i to be the power sums p_i (in n variables), or the elementary symmetric functions e_i, or the complete symmetric functions h_i. For all these, $d_i = i$, and the dimension identity is the generating function for partitions with part size at most n: $\prod_{i=1}^{n} \frac{1}{1-q^i}$.

For the hyperoctahedral group, $d_i = 2i$, and the dimension identity is the generating function for partitions with even parts, and each part size at most $2n$.

Let I_W be the ideal generated by the f_i – the ideal of W-invariants of positive degree.
Let R_W be the quotient of the polynomial ring by this ideal, the coinvariants.
It is a theorem that $\dim(R_W) = |W|$ (and this is true only for reflection groups)
Furthermore, the coinvariant ring is graded (take the degree of the smallest representative.) and its graded dimension is given by $(\prod \frac{1}{1-q}) / (\prod \frac{1}{1-q^{d_i}}) = \prod \frac{1-q^{d_i}}{1-q}$.
Taking $q=1$, this shows $|W| = d_1 d_2 \cdots d_n$.
It is another theorem that the polynomial ring can be decomposed as the tensor product of the invariants and the coinvariants – this is an isomorphism of graded W-modules.
This decomposition holds if and only if W is generated by reflections.
Using an inner product, it is possible to embed the coinvariants (as a graded W-module) into the polynomial ring, as the space H_W of W-harmonic polynomials. This space is spanned by all multivariate derivatives of the Jacobian $\Delta_W := \det \frac{\partial f_i}{\partial x_j}$.

For the symmetric group, taking f_i to be the power sums gives $\Delta_W = \det(x_j^{i-1})$, the Vandermonde determinant. For the hyperoctahedral group $\Delta_W = \det(x_j^{2i-1})$. In general, Δ_W is independent of the choice of f_i – indeed, it is the product of the positive roots.

Next, we work towards a basis for $\mathcal{H}_{S_n}$.
In lexicographic order, the highest term in Δ_{S_n} is $x_1^{n-1} x_2^{n-2} \cdots x_{n-1}^{1} x_n^{0}$.
Given any n-tuple $\alpha := (a_1, \cdots, a_n)$, the polynomial $\partial x^\alpha \Delta_{S_n} := \frac{\partial^{a_1}}{\partial x_1^{a_1}} \cdots \frac{\partial^{a_n}}{\partial x_n^{a_n}} \Delta_{S_n}$ is either 0 (if $a_i > n-i$ for some i), or has highest term equal to a multiple of $x_1^{n-1-a_1} x_2^{n-2-a_2} \cdots x_{n-1}^{1-a_{n-1}} x_n^{0-a_n}$. By triangularity, the $n!$ non-zero such are linearly independent; to conclude they give a basis, it suffices to show $\dim \mathcal{H}_{S_n} = n!$

Recall that the polynomial ring is the tensor product of the invariants and the coinvariants, and hence also the tensor product of the invariants and $\mathcal{H}_W$. (This means that any polynomial has a unique expression $\sum_\alpha g_\alpha(x)\, \partial x^\alpha \Delta_{S_n}(x)$ over $\alpha = (a_1, \cdots, a_n)$ with $0 \le a_i \le n-i$, for symmetric functions g_α).
The space of polynomials in n variables is spanned by monomials, which, in species language, is the substitution of the species of polynomials in a single variable into the species E_n of sets of size n. Hence its graded Frobenius characteristic is the plethysm $h_n\left[\frac{X}{1-q}\right] = \sum_{\lambda \vdash n} \frac{p_\lambda}{z_\lambda}\left[\frac{X}{1-q}\right] = \sum_{\lambda \vdash n} \frac{p_\lambda}{z_\lambda (1-q^{\lambda_1})(1-q^{\lambda_2})\cdots(1-q^{\lambda_\ell})}$, since

$p_k\left[\frac{X}{1-q}\right] = \frac{p_k[X]}{1-q^k}$.

Now this is equal to the product of the graded Frobenius characteristics of the invariants and of $\mathcal{H}_{S_n}$ – this is $\prod_{k=1}^{n} \frac{1}{1-q^k} \mathrm{Frob}_q(\mathcal{H}_{S_n})$.

Hence
$$\mathrm{Frob}_q(\mathcal{H}_{S_n}) = \left(\prod_{k=1}^{n} (1-q^k)\right) \sum_{\lambda \vdash n} \frac{p_\lambda}{z_\lambda (1-q^{\lambda_1}) \cdots (1-q^{\lambda_\ell})}$$
$$= \prod_{k=1}^{n} (1+q+\cdots+q^{k-1}) \sum_{\lambda \vdash n} \frac{p_\lambda (1-q^{\ell(\lambda)})}{z_\lambda (1+\cdots+q^{\lambda_1-1}) \cdots (1+\cdots+q^{\lambda_\ell-1})}$$

When $q=1$, only the summand $\lambda = (1, \cdots, 1)$ remains, and the character is that of the regular representation. Hence $\dim(\mathcal{H}_{S_n}) = n!$.
For other Coxeter groups, it is possible to make an analogue of this argument with a graded character (though there is in general no interpretation in terms of an algebra like Λ)

To see that $\mathcal{H}_W$ is isomorphic to the coinvariants, define the following scalar product on the space of polynomials: $\langle f, g \rangle := \partial f(g)$ evaluated at $x_1 = \cdots = x_n = 0$, where the

differential operator associated to f is $\partial f := f(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \cdots, \frac{\partial}{\partial x_n})$. Note that $\langle x^\alpha, x^\alpha \rangle = \alpha!$ (where $(a_1, \cdots, a_n)! = a_1! \cdots a_n!$), and $\langle x^\alpha, x^\beta \rangle = 0$ for $\alpha \neq \beta$, because $\partial x^\alpha (x^\beta) = 0$ if all entries of α are greater than the corresponding entries of β, and otherwise some x_i remains.

So the orthogonal complement of the ideal I_W is the set of all polynomials g with $\partial f(g)$ evaluated at $x_1 = x_2 = \cdots = 0$, for each $f \in I_W$. This condition is in fact equivalent to $\partial f(g) \equiv 0$ for each $f \in I_W$: if $f \in I_W$, then $x^\gamma f \in I_W$, so $\partial x^\gamma \partial f(g)$ must also be 0 when evaluated at $x_1 = x_2 = \cdots = 0$. Taking γ to be the highest term in $\partial f(g)$ shows that $\partial f(g)$ must be 0.

Hence $I_W^\perp = \{ g \mid \partial f(g) = 0 \text{ for all } f \in I_W \} = \{ g \mid \partial f_i(g) = 0 \text{ for each generating } f_i \}$.

Note that $\Delta_W \in I_W^\perp$, since Δ_W is the W-antisymmetric polynomial of lowest degree, and, for $f \in I_W$, $(\partial f)\Delta_W$ is W-antisymmetric but of lower degree, so necessarily $(\partial f)\Delta_W = 0$.

Note also that, for any monomial x^α, $\partial f(\partial x^\alpha \Delta_W) = \partial x^\alpha (\partial f \Delta_W) = 0$, so $\partial x^\alpha \Delta_W \in I_W^\perp$.

Hence H_W, defined as the space of all derivatives of the Jacobian, is a subspace of $I_W^\perp$.

The graded character argument then shows $\dim H_W = |W|$, which is also the dimension of the coinvariants.

The proof that the coinvariants have dimension $|W|$ uses some algebraic geometry.

Given any ideal I of a ring R, let $g_1, \cdots, g_m$ denote its Gröbner-basis generators, and $D(g_i)$ the leading term in g_i. Then, a basis for the quotient R/I is the set of monomials of lower degree than all the $D(g_i)$. In particular, the dimension of R/I depends only on $D(g_i)$, not on g_i. So $\dim(R/I) = \dim(R/\mathrm{gr}(I))$, where the ideal $\mathrm{gr}\, I$ is generated by the highest degree pieces of each g_i.

Let x be a point not on any hyperplane of W, and let I_x denote the ideal of polynomials vanishing on the orbit of x. It can be shown that $\mathrm{gr}\, I_x = I_W$, and $\dim(R/I_x) = |W|$ since the orbit of x has $|W|$ points, and R/I_x are the polynomials on this orbit.

Multizeta values, quasisymmetric functions and shuffling (presented by christophe)

To a composition $(i_1, \dots, i_k)$ of n, associate the series

$$\zeta(i_1, \dots, i_k) = \sum_{n_1 > n_2 > \dots n_k \geq 1} \frac{1}{n_1^{i_1} n_2^{i_2} \cdots n_k^{i_k}},$$ which has weight $i_1 + \dots + i_k = n$.

If this series converges, then its limit is a multizeta value.

Example: $\zeta(1) = \sum_{n \geq 1} \frac{1}{n}$, the harmonic series, which diverges.

However, $\lim_{n \to \infty} (1 + \frac{1}{2} + \dots + \frac{1}{n} - \log n) = \gamma$, Euler's constant.

So there is another constant γ' for which $1 + \frac{1}{2} + \dots + \frac{1}{n} \leq \gamma' + \log n \leq \alpha \log n$ for another constant α, for sufficiently large n. (choose α such that $\gamma' \leq (\alpha - 1) \log 2$, so $\gamma' \leq (\alpha - 1) \log n$ for all $n \geq 2$).

Example: $\zeta(2) = \sum_{n \geq 1} \frac{1}{n^2} = \frac{\pi^2}{6}$ (an elementary proof is in Aigler-Ziegler, "Proofs from the Book")

$$\zeta(2h) = \pi^{2h} \frac{(-1)^h 2^{2h-1}}{2h-1} \left(\frac{-B_{2h}}{2h} \right),$$ where B_{2h} are the Bernoulli numbers

Proposition: $\zeta(i_1, \dots, i_k)$ converges if and only if $i_1 \geq 2$

Proof: This is straightforward if $k=1$. So suppose $k=2$.

First, suppose $i_1 = 1$.

$$\zeta(i_1, \dots, i_k) = \sum_{n_1 \geq k} \frac{1}{n_1} \left(\sum_{n_1 > n_2 > \dots > n_k \geq 1} \frac{1}{n_2^{i_2} \cdots n_k^{i_k}} \right)$$

Now $\sum_{n_1 > \dots > n_k \geq 1} \frac{1}{n_2^{i_2} \cdots n_k^{i_k}} > \frac{1}{(k-1)^{i_2} (k-2)^{i_2} \cdots 1^{i_k}}$, which is independent of n_1.

So $\zeta(i_1, \dots, i_k)$ diverges.

If, instead $i_1 \geq 2$, then

$$\sum_{n_1 > \dots > n_k \geq 1} \frac{1}{n_2^{i_2} \cdots n_k^{i_k}} \leq \sum_{n_1 > \dots > n_k \geq 1} \frac{1}{n_2 \cdots n_k} \leq \sum_{n_1 \geq n_2 \geq 1} \frac{1}{n_2 \cdots n_k} = \left(\sum_{n_1 \geq j \geq 1} \frac{1}{j} \right)^{k-1}$$
$$\leq (\alpha \log n_1)^{k-1}$$

For n_1 sufficiently large, $\alpha \log n_1 \leq n_1^{1/k}$, so

$$\frac{1}{n_1^{i_1}} \sum_{n_1 > \dots > n_k \geq 1} \frac{1}{n_2^{i_2} \cdots n_k^{i_k}} \leq \frac{1}{n_1^{i_1 - k-1/k}},$$ and $\sum_{n_1 \geq k} \frac{1}{n_1^{i_1 - k-1/k}}$ converges as

$i_1 - \frac{k-1}{k} > 1$ when $i_1 \geq 2$.

Proposition: $\zeta(3) = \zeta(2,1)$ – in other words, $\sum_{n\geq 1} \frac{1}{n^3} = \sum_{n>m\geq 1} \frac{1}{n^2 m}$

Proof: First observe that the limit of the telescoping sum $\sum_{n\geq 1}\left(\frac{1}{n} - \frac{1}{n+m}\right)$ is $1 + \frac{1}{2} + \cdots + \frac{1}{m}$.

So $\frac{1}{m^2}\sum_{n\geq 1} \frac{1}{n} - \frac{1}{n+m} = \frac{1}{m^2}\left(1 + \frac{1}{2} + \cdots + \frac{1}{m}\right)$. $\quad *$

Now $\frac{1}{m^2}\left(\frac{1}{n} - \frac{1}{n+m}\right) = \frac{1}{m^2}\frac{m}{n(n+m)} = \frac{1}{mn(n+m)} = \frac{1}{m(m+n)^2} + \frac{1}{n(m+n)^2}$

So summing the left hand side of $*$ gives $\sum_{m,n\geq 1} \frac{1}{m(m+n)^2} + \frac{1}{n(m+n)^2} = 2\zeta(2,1)$.

The right hand side of $*$ sums to $\sum_{m\geq 1} \frac{1}{m^2}\left(1 + \frac{1}{2} + \cdots + \frac{1}{m-1}\right) + \frac{1}{m^3} = \zeta(2,1) + \zeta(3)$.

Let Z_n be the $\mathbb{Q}$-vector space generated by $\zeta(C)$ over compositions C of n.

So $Z_1 = 0$, $Z_2 = \mathbb{Q}\zeta(2) = \mathbb{Q}\pi^2$, $Z_3 = \mathbb{Q}\zeta(3) + \mathbb{Q}\zeta(2,1) = \mathbb{Q}\zeta(3)$ since $\zeta(2,1) = \zeta(3)$, and $\zeta(1,2)$, $\zeta(1,1,1)$ don't converge.

Goncharov's conjecture: $Z := \sum Z_n = \bigoplus Z_n$ (i.e. the Z_n are linearly independent)

Zagier's conjecture: the dimensions d_n of the spaces Z_n are characterized by

$$d_0 = 1,\ d_1 = 0,\ d_2 = 1, \text{ and } d_n = d_{n-2} + d_{n-3}.$$

Theorem: Z is generated as a $\mathbb{Q}$ vector space by $\zeta(C)$ where each part of C has size 2 or 3.

e.g. $\zeta(4) = \zeta(2,1,1) = \frac{4}{3}\zeta(2,2)$; $\zeta(3,1) = \frac{1}{3}\zeta(2,2)$.

So Zagier's conjecture is equivalent to such $\zeta(C)$ being linearly independent.

Z is in fact closed under multiplication – that is, the product of two multizeta values is a $\mathbb{Q}$-linear (in fact $\mathbb{N}$-linear) combination of multizeta values.

e.g. $(\zeta(2))^2 = \left(\sum_{n\geq 1}\frac{1}{n^2}\right)^2 = \sum_{n,m\geq 1}\frac{1}{n^2}\frac{1}{m^2} = \sum_{n\geq 1}\frac{1}{n^4} + \sum_{n>m\geq 1}\frac{1}{n^2}\frac{1}{m^2} + \sum_{m>n\geq 1}\frac{1}{n^2}\frac{1}{m^2}$.

$= \zeta(4) + 2\zeta(2,2)$.

Similarly, $\zeta(2,1)\zeta(2) = \sum_{n\geq m\geq 1, p\geq 1}\frac{1}{n^2}\frac{1}{m^2}\frac{1}{p} = 2\zeta(2,2,1) + \zeta(4,1) + \zeta(2,3) + \zeta(2,1,2)$.

This product on compositions is the quasishuffle.

Indeed, let $QSym_0$ be the subalgebra of $QSym$ generated by monomial quasi-symmetric functions for compositions with first part at least 2. Then the map: $QSym_0 \to \mathbb{R}$, $M_C \to \zeta(C)$ is a $\mathbb{Q}$-algebra morphism.