Efficient computations in Schubert calculus

Robert Zeier
zeier@eecs.harvard.edu

Applicable Algebra and Computer Science:
A Meeting in Honor of Jacques Calmet
(Logroño)

February 28, 2008

Outline

Fundamentals

Lie groups and Lie algebras

- We consider a Lie group $\mathfrak{G}$ as a (closed) linear matrix group.
- a Lie algebra $\mathfrak{g}$ (of $\mathfrak{G}$) is
  - the tangent space to the Lie group $\mathfrak{G}$ at the identity
  - a vector space with a bilinear and skew-symmetric multiplication $[\cdot,\cdot]$:
    1. $[g_1, g_2] \in \mathfrak{g} \quad (g_i \in \mathfrak{g})$
    2. $[[g_1, g_2], g_3] + [[g_3, g_1], g_2] + [[g_2, g_3], g_1] = 0$ (Jacobi identity)

our model example: the special unitary group $\mathfrak{G} = \mathrm{SU}(d)$

- $\mathfrak{G} = \mathrm{SU}(d) = \{\, U \in \mathrm{GL}(d, \mathbb{C}) |\, UU^\dagger = 1 \text{ and } \det(U) = 1\}$
- $\mathfrak{G}$ is real, semisimple, and compact
- $\mathfrak{g} = \mathfrak{su}(d) = \{ g \in \mathfrak{gl}(d, \mathbb{C}) |\, g + g^\dagger = 0 \text{ and } \mathrm{Tr}(g) = 0\}$

Schubert calculus: overview

purpose [see, e.g., Hiller (1981,1982)]

- module basis (with relations) of the integral cohomology group
- explicit description of the multiplicative structure

origin

- named after Hermann Schubert (1848-1911)
  who solved counting problems in projective geometry
- relies in its current form on ideas of Chevalley ($\sim$1958, publ. 1994)
- worked out by Demazure (1973,1974) and Bernstein et al. (1973)
- Lascoux and Schützenberger (1982) introduced a special form
  (Schubert polynomials)

Schubert calculus: important special case

cohomology of Grassmannians $G(m, n)$

- $G(m, n) =$ complex m-dimensional linear subspaces of $\mathbb{C}^{m+n}$
- module basis of the cohomology ring: X_λ
- multiplicative structure: $\mathrm{X}_\lambda \cdot \mathrm{X}_\mu = \sum_\nu c^\nu_{\lambda\mu} \mathrm{X}_\nu$

Littlewood-Richardson rule

- Lesieur (1947): irreducible (poly.) representations V^λ of $\mathrm{GL}(n, \mathbb{C})$ satisfy $V^\lambda \otimes V^\mu = \sum_\nu c^\nu_{\lambda\mu} V^\nu$
  ($\lambda, \mu, \nu =$ integer partitions with $\leq n$ parts)
- the Littlewood-Richardson rule (1934) for the $c^\nu_{\lambda\mu}$ is exceptional

Restricting Schubert calculus to a subgroup

basis elements in the Schubert module basis of the cohomology ring

vanishing or nonvanishing?

- if you restrict a basis element to a subgroup
  it may or may not vanish
- plausible that this can be decided in polynomial time
  for each basis element

our problem [see, e.g., Purbhoo (2006)]

find efficient algorithms to determine all nonvanishing basis elements

main example

restrict Schubert calculus from $\mathrm{SU}(2^n)$ to $\mathrm{SU}(2)^{\otimes n} = \mathrm{SU}(2) \otimes \cdots \otimes \mathrm{SU}(2)$

Motivation

analyze and understand the cosets $\mathrm{SU}(2^n)/\mathrm{SU}(2)^{\otimes n}$

- nonvanishing cohomology elements are important for computing the cohomology of the coset
- compare with de Rham cohomology computations using the Koszul complex
- (potential) applications in quantum computing

restricting representations (and corresponding convexity theorems)

- Berenstein and Sjamaar (2000): explicit algorithm to compute restricted representations (asymptotic version)
- restricted Schubert calculus plays a significant role in this algorithm

Outline

The adjoint representation and the Weyl group

adjoint representation $\mathrm{Ad}_{\mathfrak{g}}$ of the Lie group $\mathfrak{G}$ on its Lie algebra $\mathfrak{g}$

- $\mathrm{Ad}_{\mathfrak{g}}(G) : \mathfrak{g} \to \mathfrak{g}, \quad g \mapsto G^{-1} g G \quad (G \in \mathfrak{G}, g \in \mathfrak{g})$
- $\mathrm{Ad}_{\mathfrak{g}}(\mathfrak{H}) := \{\mathrm{Ad}_{\mathfrak{g}}(H) \colon H \in \mathfrak{H}\} \quad (\mathfrak{H} \subset \mathfrak{G})$

Weyl group $\mathcal{W}_{\mathfrak{H}}(\mathfrak{a}) := N_{\mathfrak{H}}(\mathfrak{a}) / C_{\mathfrak{H}}(\mathfrak{a}) \ (\mathfrak{H} \subset \mathfrak{G})$

- $\mathfrak{a}$ is an Abelian subalgebra of $\mathfrak{g}$
- centralizer $C_{\mathfrak{H}}(\mathfrak{a}) = \{H \in \mathfrak{H} | \forall a \in \mathfrak{a} \colon (\mathrm{Ad}_{\mathfrak{g}}(H))(a) = a\}$
- normalizer $N_{\mathfrak{H}}(\mathfrak{a}) = \{H \in \mathfrak{H} | (\mathrm{Ad}_{\mathfrak{g}}(H))(\mathfrak{a}) \subset \mathfrak{a}\}$
- $\mathcal{W}_{\mathrm{SU}(3)}(\mathfrak{a}) =$ symmetric group $\mathrm{S}_3, \quad |\mathrm{S}_3| = 6$

Roots $\Delta = \Delta_{\mathfrak{g}}$

roots $\Delta_{\mathfrak{g}}$ of $\mathfrak{g}$ w.r.t. an Abelian subalgebra $\mathfrak{a}$

$$\Delta_{\mathfrak{g}} = \{\alpha \in \mathfrak{a}^* \,|\, \mathfrak{g}^\alpha \neq \{0\} \text{ and } \alpha \neq 0\}$$

where $\mathfrak{g}^\alpha = \{g \in \mathfrak{g} + \mathrm{i}\mathfrak{g} \,|\, \forall a \in \mathfrak{a}\colon [a, g] = \mathrm{i}\alpha(a) g\}$ and $\mathfrak{a}^* =$ dual space to $\mathfrak{a}$

- choose $\mathfrak{a}$ as a maximal Abelian subalgebra
- positive and negative roots: $\Delta = \Delta^+ \uplus \Delta^-$ where $|\Delta^+| = |\Delta^-|$
- simple roots $\Sigma \subset \Delta^+$ cannot written as sum of two positive roots

- Killing form $B(g, h) := \mathrm{Tr}_{\mathfrak{g}}(\mathrm{ad}(g) \circ \mathrm{ad}(h))$, $\mathrm{ad}(g)\colon h \mapsto [g, h]$
- $\mathfrak{g}$ semisimple $\Rightarrow \exists$ unique $a_\lambda \in \mathfrak{a}\colon \lambda(\mathfrak{a}) = B(a_\lambda, \mathfrak{a})$ for $\lambda \in \mathfrak{a}^*$

Cartan matrix, weights, and the Weyl group $\mathcal{W} = \mathcal{W}_{\mathfrak{G}}(\mathfrak{a})$

Cartan matrix $n(\alpha, \beta) = 2\,(\alpha, \beta)\,/\,(\alpha, \alpha) = (\alpha, \check{\beta})$ where $\alpha, \beta \in \Delta^+$

- coroot $\check{\alpha} := 2\alpha/\,(\alpha, \alpha)$ where $(\lambda, \mu) := B(a_\lambda, a_\mu)$

- Killing form $B(g, h) := \mathrm{Tr}_{\mathfrak{g}}(\mathrm{ad}(g) \circ \mathrm{ad}(h))$, $\mathrm{ad}(g)\colon h \mapsto [g, h]$
- $\mathfrak{g}$ semisimple $\Rightarrow \exists$ unique $a_\lambda \in \mathfrak{a}\colon \lambda(\mathfrak{a}) = B(a_\lambda, \mathfrak{a})$ for $\lambda \in \mathfrak{a}^*$

weight lattice $\Theta = \{\lambda \in \mathfrak{a}^*\colon \lambda(\check{\alpha}) \in \mathbb{Z} \quad \forall \alpha \in \Delta\}$

- fundamental weights $\bar{\omega}_\alpha$: $(\bar{\omega}_\alpha, \check{\beta}) = \delta_{\alpha\beta}$
- root lattice $\Xi = \langle \Delta \rangle \subset \Theta$

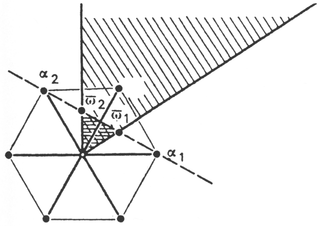
$\mathcal{W} = \langle s_\alpha\colon \alpha \in \Delta^+ \rangle$ where $s_\alpha(\beta) = \alpha - n(\alpha, \beta)\beta$ is a reflection

- length function: $l(w) = \min\{k\colon w = s_{\alpha_1} \cdots s_{\alpha_k}\}$
- $\exists$ unique $w_0\colon l(w_0) = \max l(w) = |\Delta^+|$

Example: SU(3)

roots and weights

- simple roots $\Sigma = \{\alpha = \alpha_1, \beta = \alpha_2\}$
- positive roots $\Delta^+ = \Sigma \cup \{\alpha + \beta\}$
- Cartan matrix $n(\alpha, \beta) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$
- fundamental weights: $\bar{\omega}_1 = \bar{\omega}_\alpha = (2\alpha + \beta)/3$, $\bar{\omega}_2 = \bar{\omega}_\beta(\alpha + 2\beta)/3$



the Weyl group $\mathcal{W} = \langle s_\alpha : \alpha \in \Delta^+ \rangle$ $[s_\alpha(\beta) = \alpha - n(\alpha, \beta)\beta]$

- $\mathcal{W}_{\mathrm{SU}(3)}(\mathfrak{a})$ = symmetric group S_3, $|\mathrm{S}_3| = 6$
- elements: id, s_α, s_β, $s_\alpha s_\beta$, $s_\beta s_\alpha$, and $w_0 = s_\alpha s_\beta s_\alpha$

Schubert calculus: theory (1)

symmetric algebra $\mathrm{S}(\mathfrak{a}^*)$ ($\mathfrak{a}$ = maximal Abelian subalgebra of $\mathfrak{g}$)

- $\mathfrak{a}^* = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$ is the dual space to $\mathfrak{a}$
- $\mathrm{S}(\mathfrak{a}^*)$ is isomorphic to the polynomial algebra $\mathbb{R}[\lambda_1, \lambda_2, \dots, \lambda_m]$

cohomology ring $H_{\mathcal{W}}$

$\{\mathrm{X}_w : w \in \mathcal{W}\}$ = module basis index by the Weyl group $\mathcal{W}$

the operators $\partial_\alpha := (1 - s_\alpha)/\alpha$ where $\partial_\alpha : \mathrm{S}(\mathfrak{a}^*) \to \mathrm{S}(\mathfrak{a}^*)$

- $\partial_\alpha^2 = 0$, $\partial_\beta(\bar{\omega}_\alpha) = \delta_{\alpha\beta}$, and $\partial_\alpha(uv) = \partial_\alpha(u)v + s_\alpha(u)\partial_\alpha(v)$
- define $\partial_w = \partial_{\alpha_1} \cdots \partial_{\alpha_k}$ if $w = s_{\alpha_1} \cdots s_{\alpha_k}$ and $l(w) = k$
- $\partial_w \cdot \partial_{w'} = \begin{cases} \partial_{ww'} & \text{if } l(ww') = l(w) + l(w'), \\ 0 & \text{otherwise} \end{cases}$

Schubert calculus: theory (2)

the operators $\partial_\alpha := (1 - s_\alpha)/\alpha$ where $\partial_\alpha \colon \mathrm{S}(\mathfrak{a}^*) \to \mathrm{S}(\mathfrak{a}^*)$

- $\partial_\alpha^2 = 0$, $\partial_\beta(\bar{\omega}_\alpha) = \delta_{\alpha\beta}$, and $\partial_\alpha(uv) = \partial_\alpha(u)v + s_\alpha(u)\partial_\alpha(v)$
- define $\partial_w = \partial_{\alpha_1} \cdots \partial_{\alpha_k}$ if $w = s_{\alpha_1} \cdots s_{\alpha_k}$ and $l(w) = k$

characteristic homomorphism $c \colon \mathrm{S}(\mathfrak{a}^*) \to H_{\mathcal{W}}$ = cohomology ring

- $c(u) = \sum_{w \in \mathcal{W}} \varepsilon[\partial_w(u)]\, \mathrm{X}_w$ (projection $\varepsilon \colon \mathrm{S}(\mathfrak{a}^*) \to \mathrm{S}^0(\mathfrak{a}^*) \approx \mathbb{R}$)
- $c(\frac{1}{|\mathcal{W}|} \prod_{\alpha \in \Delta^+} \alpha) = \mathrm{X}_{w_0}$, $c(\bar{\omega}_\alpha) = \mathrm{X}_{s_\alpha}$, and $c(\alpha) = \sum_{\beta \in \Sigma} (\alpha, \check{\beta})\, \mathrm{X}_{s_\beta}$

- $\mathrm{Ker}(\partial_\alpha) = \mathrm{S}(\mathfrak{a}^*)^{\langle s_\alpha \rangle}$ and $\mathrm{Ker}(c) = (\mathrm{S}(\mathfrak{a}^*)^{\mathcal{W}})_+$
- $\mathcal{W}$ = complex reflection group $\Rightarrow \mathrm{S}(\mathfrak{a}^*)^{\mathcal{W}}$ = polynomial algebra [Shephard-Todd, Chevalley]

Schubert calculus: theory (3) and example SU(3)

characteristic homomorphism $c\colon \mathrm{S}(\mathfrak{a}^*) \to H_{\mathcal{W}}$ = cohomology ring

- $c(\frac{1}{|\mathcal{W}|}\prod_{\alpha\in\Delta^+}\alpha) = \mathrm{X}_{w_0}$, $c(\bar\omega_\alpha) = \mathrm{X}_{s_\alpha}$, and $c(\alpha) = \sum_{\beta\in\Sigma}(\alpha,\check\beta)\,\mathrm{X}_{s_\beta}$
- $\mathrm{X}_{w_0} = c(d)$; $\mathrm{X}_{s_\alpha} = c(\bar\omega_\alpha)$; $\mathrm{X}_{s_\beta} = c(\bar\omega_\beta)$ $[d = \alpha\beta(\alpha+\beta)/6]$
- $\partial_\alpha(d) = \beta(\alpha+\beta)/3$

Giambelli formula: $c(\partial_{w^{-1}w_0}\frac{1}{|\mathcal{W}|}\prod_{\alpha\in\Delta^+}\alpha) = \mathrm{X}_w$

- $\mathrm{X}_{s_\alpha s_\beta} = (-\mathrm{X}_{s_\alpha}^2 + \mathrm{X}_{s_\beta}\mathrm{X}_{s_\alpha} + 2\mathrm{X}_{s_\beta}^2)/3$
- computes the preimage of c (modulo $(\mathrm{S}(\mathfrak{a}^*)^{\mathcal{W}})_+$)

Pieri formula: $\mathrm{X}_{s_\alpha}\cdot\mathrm{X}_w = \sum_{\beta\in\Delta^+,\, l(ws_\beta)=l(w)+1}\left(\check\beta,\bar\omega_\alpha\right)\mathrm{X}_{ws_\beta}$

- $\mathrm{X}_{s_\alpha}^2 = \mathrm{X}_{s_\beta s_\alpha}$; $\mathrm{X}_{s_\beta}\mathrm{X}_{s_\alpha} = \mathrm{X}_{s_\beta s_\alpha} + \mathrm{X}_{s_\alpha s_\beta}$; $\mathrm{X}_{s_\beta}^2 = \mathrm{X}_{s_\alpha s_\beta}$
- describes the multiplicative structure of $H_{\mathcal{W}}$

Outline

Induced action on the cohomology ϕ^*

induced actions given by the embedding $f\colon \tilde{\mathfrak{G}} \hookrightarrow \mathfrak{G}$

- embedding $f_*\colon \tilde{\mathfrak{g}} \hookrightarrow \mathfrak{g}$, restriction $f^*\colon \mathfrak{g}^* \to \tilde{\mathfrak{g}}^*$
- restriction $\phi^*\colon H_{\mathcal{W}} \to H_{\tilde{\mathcal{W}}}$ ($\mathcal{W}, \tilde{\mathcal{W}}$ are the Weyl groups)

computation of ϕ^* ($\mathfrak{a}, \tilde{\mathfrak{a}}$ are max. commutative subalgebras)

$$\begin{array}{ccc} \mathrm{S}(\mathfrak{a}^*) & \xrightarrow{\mathrm{S}(f^*)} & \mathrm{S}(\tilde{\mathfrak{a}}^*) \\ \downarrow c & & \downarrow \tilde{c} \\ H_{\mathcal{W}} & \xrightarrow[\phi^*]{} & H_{\tilde{\mathcal{W}}} \end{array}$$

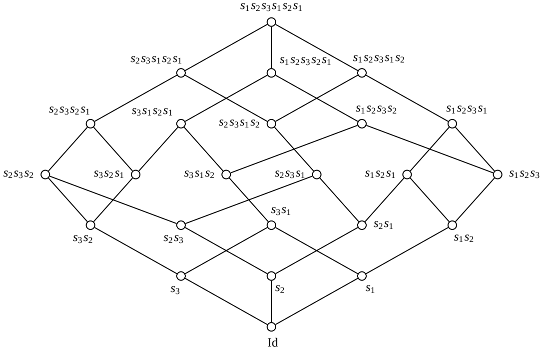
$\mathrm{S}(f^*)$ uniquely given by f^* (and surjective if f^* is);
Demazure (1973) $\Rightarrow$ c (or $\tilde{c}$) is surjective if $\mathfrak{G}$ (or $\tilde{\mathfrak{G}}$) contains only SU-components and the corresponding Cartan matrix is nondegenerate (no torsion effects in this case);
compute $\phi^* = \tilde{c} \circ \mathrm{S}(f^*) \circ c^{-1}$
by inverting c using Giambelli

$\Rightarrow$ we are in a good situation if $\mathfrak{G} = \mathrm{SU}(2^n)$ and $\tilde{\mathfrak{G}} = \mathrm{SU}(2)^{\otimes n}$

Inverting c using the Giambelli formula (1/3): [Ex. $\mathrm{SU}(4)$]

weak order

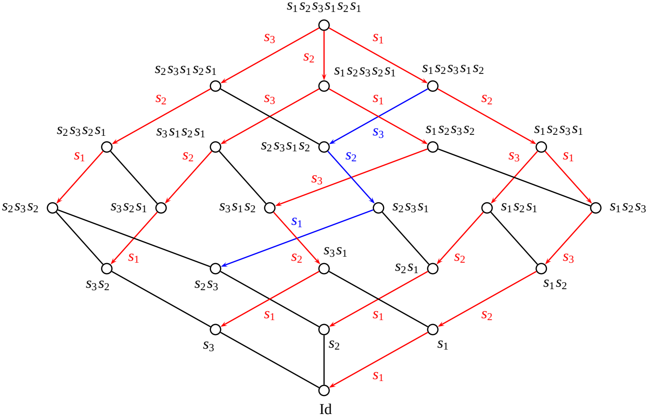
$w_1 \leq w_2 \Leftrightarrow w_2 = w_1 s_{j_1} \ldots s_{j_k}$
s.t. $l(w_1 s_{j_1} \ldots s_{j_i}) = l(w_1) + i$
for all $i \in \{1, \ldots, k\}$



Giambelli formula: $c(\partial_{w^{-1}w_0} \frac{1}{|\mathcal{W}|} \prod_{\alpha \in \Delta^+} \alpha) = \mathrm{X}_w$

- $\partial_w = \partial_{\alpha_1} \cdots \partial_{\alpha_k}$ if $w = s_{\alpha_1} \cdots s_{\alpha_k}$ and $l(w) = k$
- $\partial_\alpha := (1 - s_\alpha)/\alpha$

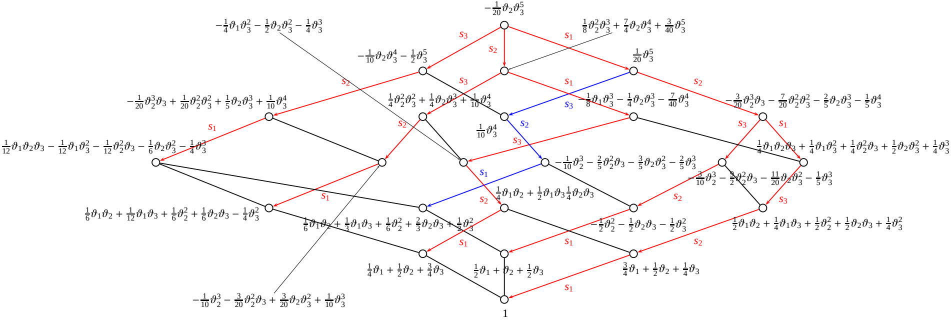
Inverting c using the Giambelli formula (2/3): [Ex. SU(4)]



Giambelli formula: $c(\partial_{w^{-1}w_0}\frac{1}{|\mathcal{W}|}\prod_{\alpha\in\Delta^+}\alpha) = X_w$

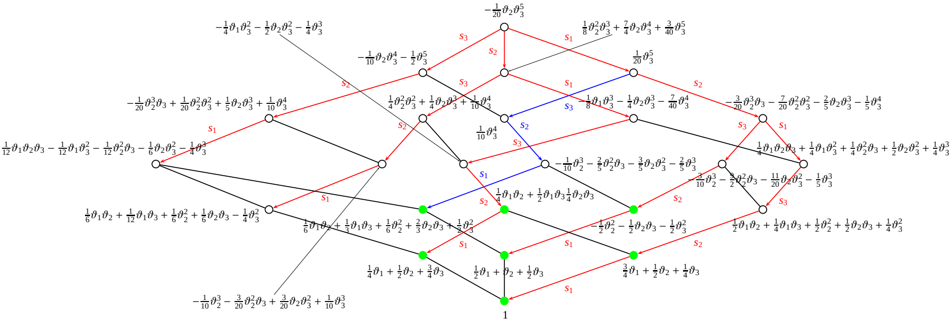
- $\partial_w = \partial_{\alpha_1}\cdots\partial_{\alpha_k}$ if $w = s_{\alpha_1}\cdots s_{\alpha_k}$ and $l(w) = k$
- $\partial_\alpha := (1 - s_\alpha)/\alpha$

Inverting c using the Giambelli formula (3/3): [Ex. $\mathrm{SU}(4)$]



- $\partial_{w^{-1}w_0}\frac{1}{|\mathcal{W}|}\prod_{\alpha\in\Delta^+}\alpha$ is normalized w.r.t. $(\mathrm{S}(\mathfrak{a}^*)^{\mathcal{W}})_+$
- $\frac{1}{|\mathcal{W}|}\prod_{\alpha\in\Delta^+}\alpha$ is normalized to $-\frac{1}{20}\vartheta_2\vartheta_3^5$

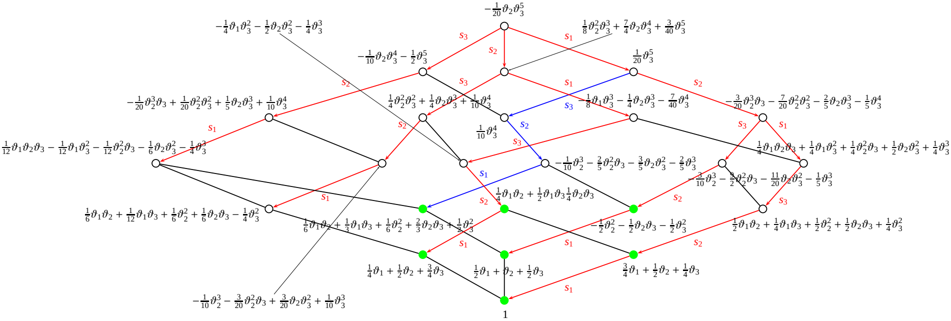
Restricting from $\mathfrak{G} = \mathrm{SU}(4)$ to $\tilde{\mathfrak{G}} = \mathrm{SU}(2) \otimes \mathrm{SU}(2)$ (1/2)



1. use $\mathrm{S}(f^*)$ (uniquely given by f^* and surjective if f^* is)
2. use $\tilde{c}(u) = \sum_{w \in \tilde{\mathcal{W}}} \tilde{\varepsilon}[\partial_w(u)]\, \mathrm{X}_w$ (projection $\tilde{\varepsilon}\colon \mathrm{S}(\tilde{\mathfrak{a}}^*) \to \mathrm{S}^0(\tilde{\mathfrak{a}}^*) \approx \mathbb{R}$)

$|\mathcal{W}| = 24$ and $|\tilde{\mathcal{W}}| = 4$

Restricting from $\mathfrak{G} = \mathrm{SU}(4)$ to $\tilde{\mathfrak{G}} = \mathrm{SU}(2) \otimes \mathrm{SU}(2)$ (2/2)



seven nonvanishing basis elements ($|\mathcal{W}| = 24$ and $|\tilde{\mathcal{W}}| = 4$)

- $\phi^*(\mathrm{X}_{\mathrm{id}}) = \mathrm{X}_{\mathrm{id}}$
- $\phi^*(\mathrm{X}_{s_1}) = \phi^*(\mathrm{X}_{s_3}) = \mathrm{X}_{s_1} + \mathrm{X}_{s_2}$ and $\phi^*(\mathrm{X}_{s_2}) = 2\mathrm{X}_{s_1}$
- $\phi^*(\mathrm{X}_{s_2 s_1}) = \phi^*(\mathrm{X}_{s_3 s_1}) = \phi^*(\mathrm{X}_{s_2 s_3}) = 2\mathrm{X}_{s_2 s_1}$

General case: $\tilde{\mathfrak{G}} = \mathrm{SU}(2)^{\otimes n} \hookrightarrow \mathfrak{G} = \mathrm{SU}(2^n)$ $(n > 2)$

Challenge: combinatorial explosion $(|\tilde{\mathcal{W}}| = 2^n$ and $|\mathcal{W}| = (2^n)!)$

- $n = 2 \Rightarrow |\tilde{\mathcal{W}}| = 4$ and $|\mathcal{W}| = 24$
- $n = 3 \Rightarrow |\tilde{\mathcal{W}}| = 8$ and $|\mathcal{W}| = 40320$
- $n = 4 \Rightarrow |\tilde{\mathcal{W}}| = 16$ and $|\mathcal{W}| = 20922789888000$ (This is a problem!)

Example: $\tilde{\mathfrak{G}} = \mathrm{SU}(2) \otimes \mathrm{SU}(2) \otimes \mathrm{SU}(2) \hookrightarrow \mathfrak{G} = \mathrm{SU}(8)$

83 nonvanishing basis elements $(|\tilde{\mathcal{W}}| = 8$ and $|\mathcal{W}| = 8! = 40320)$

- $\phi^*(\mathrm{X}_{\mathrm{id}}) = \mathrm{X}_{\mathrm{id}}$, $\phi^*(\mathrm{X}_{s_1}) = \phi^*(\mathrm{X}_{s_7}) = \mathrm{X}_{s_1} + \mathrm{X}_{s_2} + \mathrm{X}_{s_2}$,
  $\phi^*(\mathrm{X}_{s_2}) = \phi^*(\mathrm{X}_{s_6}) = 2\mathrm{X}_{s_1} + 2\mathrm{X}_{s_2}$,
  $\phi^*(\mathrm{X}_{s_3}) = \phi^*(\mathrm{X}_{s_5}) = 3\mathrm{X}_{s_1} + \mathrm{X}_{s_2} + \mathrm{X}_{s_3}$, and $\phi^*(\mathrm{X}_{s_4}) = 4\mathrm{X}_{s_1}$,
- 25 nonvanshing elements X_w where $w \in \mathcal{W}$ is of length $l(w) = 2$
- 50 nonvanshing elements X_w where $w \in \mathcal{W}$ is of length $l(w) = 3$
- no nonvanshing elements X_w where $w \in \mathcal{W}$ is of length $l(w) > 3$

Further ideas

tree-free computations (work in progress)

- local decision rule for finding descendants (ideas of Stembridge)
- potential speed-up in the case of $\tilde{\mathfrak{G}} = \mathrm{SU}(2) \otimes \mathrm{SU}(2) \otimes \mathrm{SU}(2) \hookrightarrow \mathfrak{G} = \mathrm{SU}(8)$

algorithms which do not touch all basis elements in $H_{\mathcal{W}}$ (future)

- Can $\mathrm{S}(f^*)$ be inverted? (respecting $(\mathrm{S}(\mathfrak{a}^*)^{\mathcal{W}})_+$ and $(\mathrm{S}(\tilde{\mathfrak{a}}^*)^{\tilde{\mathcal{W}}})_+$)
- This may give us a significantly smaller candidate set in $\mathcal{W}$!

Thank you for your attention!