THE KOSTANT–PLÜCKER RELATIONS

FRAN BURSTALL

1. Motivation

1.1. **Grassmannian of lines.** Let $V = \mathbb{C}^{n+1}$ and contemplate the Grassmannian of lines in $\mathbb{P}^n$:

$$G_2(V) = \{W \leq V \colon \dim W = 2\} = \{X \in \mathbb{P}^n \colon \dim X = 1\}.$$

This is a projective variety thanks to the Plücker embedding $k : G_2(V) \to \mathbb{P}(\wedge^2 V) : W \mapsto \wedge^2 W$ with image $\{[\omega] \colon \omega \text{ is decomposable}\}$. To see that the Plücker image really is a projective variety, that is, cut out by homogeneous polynomial equations, we readily prove that it is given by

$$\{[\omega] \colon \omega \wedge \omega = 0\}$$

and so is the intersection of a $\dim \wedge^4 V$-dimensional family of quadrics.

1.2. **Twisted cubic.** Now let $n = 1$ so that $V = \mathbb{C}^2$ and contemplate the twisted cubic $c : \mathbb{P}(V) \to \mathbb{P}(S^3 V) = \mathbb{P}^3 : [v] \mapsto [v^3]$. In suitable homogeneous coordinates, this reads $c : [s, t] \mapsto [s^3, s^2 t, st^2, t^3]$ and again the image is an intersection of the three quadrics:

$$\begin{aligned} X_0 X_3 &= X_1 X_2 \\ X_0 X_2 &= X_1^2 \\ X_1 X_3 &= X_2^2. \end{aligned}$$

1.3. **Synthesis.** One can understand both examples from a common viewpoint: in each case, our projective variety is an orbit of $G = \mathrm{SL}(V)$, the unique closed orbit, in fact, and the quadratic equations cutting out the orbit have a representation theoretic origin.

In the first case, G acts on $\wedge^2 V$ and so on the symmetric square $S^2(\wedge^2 V)$. We have a G-morphism $P : S^2(\wedge^2 V) \to \wedge^4 V : \omega_1 \omega_2 \mapsto \omega_1 \wedge \omega_2$ with irreducible kernel so that

$$S^2(\wedge^2 V) \cong \ker P \oplus \wedge^4 V$$

is a decomposition into irreducible G-modules and the equations defining the Plücker image read $P(\omega^2) = 0$.

Again, in the twisted cubic case, G acts on $S^3 V$ and so on $S^2(S^3 V)$. The Clebsch–Gordan formulae tell us that the latter decomposes into irreducibles as follows

$$S^2(S^3 V) \cong S^6 V \oplus S^2 V$$

and one can show that, with $P : S^2(S^3 V) \to S^2 V$ the corresponding projection, the equations of the twisted cubic are again $P(u^2) = 0$, $u \in S^3 V$.

We are going to show that this is a very general phenomenon: for any complex, connected semisimple Lie group G and irreducible G-module V, there is a unique

Date: November 16, 2016.

Zariski closed G-orbit in $\mathbb{P}(V)$ (the *projective highest weight orbit*) and this is cut out by quadratic equations coming from the decomposition of S^2V into irreducible submodules. This will need preparation to which we now turn.

2. Ingredients

We work with a complex, connected, semisimple Lie group G with Lie algebra $\mathfrak{g}$.

2.1. **Roots and weights.** Fix a Cartan subalgebra $\mathfrak{h} \leq \mathfrak{g}$ with roots $\Delta \subset \mathfrak{h}^*$ and associated root space decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha \in \Delta} \mathfrak{g}_\alpha.$$

The Killing form of $\mathfrak{g}$ induces an inner product on the real span $\mathfrak{h}^*_{\mathbb{R}}$ of the roots which we denote by $(\, , \,)$ and a pairing $\mathfrak{h}^*_{\mathbb{R}} \times \mathfrak{h}^*_{\mathbb{R}} \to \mathbb{R}$ given by

$$\langle \lambda, \alpha \rangle = 2\frac{(\lambda, \alpha)}{(\alpha, \alpha)}.$$

The *weight lattice* Λ is now given by

$$\Lambda = \{\lambda \in \mathfrak{h}^*_{\mathbb{R}} \colon \langle \lambda, \alpha \rangle \in \mathbb{Z} \text{ for all } \alpha \in \Delta \}.$$

Let V be a finite-dimensional $\mathfrak{g}$-module and, for $\mu \in \mathfrak{h}^*$, let $V_\mu \leq V$ be given by

$$V_\mu = \{v \in V \colon Hv = \lambda(H)v \text{ for all } H \in \mathfrak{h}\}.$$

If $V_\mu \neq 0$ say that V_μ is a *weight space* of V with *weight* μ. We further say that μ has *multiplicity* $\dim V_\mu$ in V.

Proposition 2.1. *V is a direct sum of its weight spaces and each weight lies in Λ.*

Now choose a Weyl chamber so that we have simple roots $\alpha_1, \dots, \alpha_l$ or, equivalently, a choice of positive roots $\Delta^+ \subset \Delta$, or, equivalently, a Borel subalgebra $\mathfrak{h} \leq \mathfrak{b} \leq \mathfrak{g}$ given by

$$\mathfrak{b} = \mathfrak{h} \oplus \sum_{\alpha \in \Delta^+} \mathfrak{g}_\alpha.$$

We distinguish *fundamental weights* $\lambda_1, \dots, \lambda_l \in \Lambda$ by requiring $\langle \lambda_i, \alpha_j \rangle = \delta_{ij}$. Then $\Lambda = \oplus_i \mathbb{Z}\lambda_i$ and we define the *dominant weights* Λ^+ by

$$\Lambda^+ = \{\lambda \in \Lambda \colon \langle \lambda, \alpha_i \rangle \geq 0 \text{ for all } i\}.$$

The final data we get from our Weyl chamber is a partial ordering on Λ: say that $\mu \prec \lambda$ if $\lambda - \mu = \sum_i n_i \alpha_i$ with the n_i non-negative integers.

2.2. **Irreducible $\mathfrak{g}$-modules.** The Theorem of the Highest Weight says that Λ^+ parametrises the finite-dimensional irreducible $\mathfrak{g}$-modules. Here is how.

Definition 2.2. Let V be a $\mathfrak{g}$-module. A *highest weight vector* of V is a non-zero $v \in V$ whose span is stable under the Borel $\mathfrak{b}$: $\mathfrak{b}[v] = [v]$.

Clearly such a v lies in a weight space V_λ and we say that λ is a *highest weight* of V.

We now have:

Theorem 2.3.

(1) *A finite-dimensional irreducible $\mathfrak{g}$-module V has a unique, up to scale, highest weight vector v_λ and the corresponding highest weight is dominant.*

(2) *Given a dominant weight* $\lambda \in \Lambda^+$*, there is a unique, up to isomorphism, finite-dimensional irreducible* $\mathfrak{g}$*-module* V^λ *with highest weight* λ*.*
(3) *For any other weight* μ *of* V^λ*, we have* $\mu \prec \lambda$*.*

2.3. **Casimir operator.** We now introduce the main tool for what follows: the quadratic *Casimir operator*.

The Killing form of $\mathfrak{g}$ is a non-degenerate $\mathfrak{g}$-invariant element of $\mathfrak{g}^* \otimes \mathfrak{g}^*$ and so provides a $\mathfrak{g}$-invariant element of $\mathfrak{g} \otimes \mathfrak{g}$ after using it to raise indices. The multiplication in the universal enveloping algebra $U(\mathfrak{g})$ now gives a $\mathfrak{g}$-morphism $\mathfrak{g} \otimes \mathfrak{g} \to U(\mathfrak{g})$ and so a $\mathfrak{g}$-invariant, hence central, element of $U(\mathfrak{g})$: this is the Casimir operator C. Thus, from a practical point of view, $C = \sum_h e_h e^h$ where $e_1, \dots, e_n$ and $e^1, \dots, e^n$ are Killing dual bases of $\mathfrak{g}$.

Here is what we need to know about C:

(1) By Schur's Lemma, C acts as a scalar $C(\lambda) \in \mathbb{R}$ on the irreducible V^λ.
(2) By choosing dual bases adapted to the root space decomposition and evaluating C on v_λ, it is straightforward to compute that
$$C(\lambda) = (\lambda + \rho, \lambda + \rho) - (\rho, \rho),$$
where $\rho = \lambda_1 + \dots + \lambda_l = \frac{1}{2} \sum_{\alpha \in \Delta^+} \alpha$.
(3) If $\mu \neq \lambda \in \Lambda^+$ with $\mu \prec \lambda$, then $C(\mu) < C(\lambda)$: indeed,
$$C(\lambda) - C(\mu) = (\lambda + \mu + 2\rho, \lambda - \mu) = \sum_i n_i (\lambda + \mu + 2\rho, \alpha_i),$$
with each $n_i \geq 0$, and $(\lambda + \mu + 2\rho, \alpha_i) \geq (\rho, \alpha_i) > 0$.

We use the Casimir operator to gain insight into the tensor product of irreducible representations: contemplate such a product $U = V^{\mu_1} \otimes \dots \otimes V^{\mu_k}$. Since $\mathfrak{g}$ acts as a derivation over tensor product, the weight spaces of U are (sums of) tensor products of those of the V^{μ_j} and the weights are sums of the corresponding weights. In particular, $v_{\mu_1} \otimes \dots \otimes v_{\mu_n}$ is a highest weight vector, the corresponding highest weight $\mu_1 + \dots + \mu_n$ appears with multiplicity one and all other weights ν of U have $\nu \prec \mu_1 + \dots + \mu_n$. Write U as a sum of irreducibles:
$$U = V^{\mu_1 + \dots + \mu_n} \oplus V^{\nu_1} \oplus \dots \oplus V^{\nu_k}.$$
Applying our observations to the ν_j, we see that $C(\nu_j) < C(\mu_1 + \dots + \mu_n)$ so that
$$V^{\mu_1 + \dots + \mu_n} = \{u \in U : C(u) = C(\mu_1 + \dots + \mu_n) u\}.$$
We shall use this characterisation of $V^{\mu_1 + \dots + \mu_n}$ repeatedly below.

3. Quadratic relations

Let V be a finite-dimensional irreducible G-module. Then V is also an irreducible $\mathfrak{g}$-module (differentiate!) and the preceding theory applies. Thus $V = V^\lambda$ with highest weight vector v_λ, for some $\lambda \in \Lambda^+$.

We are interested in the projective highest weight orbit $\mathcal{O} = G[v_\lambda] \subset \mathbb{P}(V^\lambda)$. We are going to show that $\mathcal{O}$ is a projective variety and, in fact, the intersection of an explicit family of quadrics. Along the way, we will identify the homogeneous coordinate ring of $\mathcal{O}$ as a G-module.

We start with some preliminary observations: $v_\lambda \otimes v_\lambda \in V^{2\lambda} \subset S^2 V^\lambda$ so that, for any $[u] \in \mathcal{O}$, $u \otimes u = g v_\lambda \otimes g v_\lambda = g(v_\lambda \otimes v_\lambda) \in V^{2\lambda}$ also. Said another way, $\mathcal{O}$

satisfies quadratic equations: write

$$S^2V^\lambda = V^{2\lambda} \oplus W,$$

a decomposition of G-modules, and let $P : S^2V^\lambda \to W$ be the corresponding projection. Then we have just seen that

$$\mathcal{O} \subset \{[u] \in \mathbb{P}(V^\lambda) \colon P(u \otimes u) = 0\} \tag{3.1}$$

which last is the intersection of a $\dim W$-dimensional family of quadrics. We will show equality in (3.1).

For this, let $V^\mu = (V^\lambda)^*$ and identify $S^kV^\mu \cong (S^kV^\lambda)^*$ with the homogeneous polynomials of degree k on V^λ via $f \mapsto (v \mapsto f(v^k))$. In this way, we have an isomorphism of $\mathfrak{g}$-modules between the coordinate ring $\mathbb{C}[V^\lambda]$ of the affine space V^λ and the symmetric algebra $S^\bullet V^\mu$.

Set $Q = W^*$ so that

$$S^2V^\mu = V^{2\mu} \oplus Q$$

and view Q as a space of quadratic functions that vanish on $\mathcal{O}$ thanks to (3.1).

We can now state and prove the following theorem of Kostant:

Theorem 3.1.

(1) *$\mathcal{O}$ is the zero set of Q so that we have equality in* (3.1).
(2) *The ideal of $\mathcal{O}$ is generated by Q.*
(3) *The homogeneous coordinate ring of $\mathcal{O}$ is $\bigoplus_{k\geq 0} V^{k\mu}$.*

Proof. First we need to see that $\mathcal{O}$ is a projective variety at all (so that it is determined by the homogeneous polynomials that vanish on it). For this, first note that $\mathbb{P}(V^\lambda)$ does contain at least one closed orbit[1]. Now contemplate the subgroup $B \leq G$ with Lie algebra $\mathfrak{b}$: this is (maximal) solvable and so by the Borel Fixed Point Theorem[2] has a fixed point on such an orbit. However, a fixed point is the span of a highest weight vector and so must be $[v_\lambda]$. Our closed orbit is therefore $\mathcal{O}$. In short, $\mathcal{O}$ is the unique closed G-orbit in $\mathbb{P}(V^\lambda)$.

Let $X \subset V^\lambda$ be the cone over $\mathcal{O}$: this is an affine variety. Let $R = \mathbb{C}[V^\lambda]$, $A = \mathbb{C}[X]$ and I the ideal of X so that $A = R/I$. Let $J = RQ$, the ideal generated by Q and set $B = R/J$. We know from (3.1) that $J \leq I$ so that the restriction $B \to A$ surjects.

The heart of the matter will be to show that $B = \bigoplus_{k\geq 0} V^{k\mu}$, or, more precisely that the degree k part of B is $V^{k\mu}$. For this, start by observing that $B_1 = R_1 = V^\mu$, since J is generated in degree 2, while $B_2 = V^{2\mu}$ by construction. Now let $x, y \in B_1$: we are going to compute $C(xy)$ in two different ways. Firstly $xy \in V^{2\mu}$ so that $C(xy) = C(2\mu)xy$. On the other hand, $C = \sum_h a_hb_h$ with $a_h, b_h \in \mathfrak{g}$ so that both a_h and b_h act by derivations on B. Thus

$$C(xy) = \sum_h (a_hb_h)xy = C(x)y + C(y)x + \sum_h \big((a_hx)(b_hy) + (b_hx)(a_hy)\big)$$

and, using the simple identity $C(2\mu) - 2C(\mu) = 2(\mu, \mu)$, we conclude that

$$\sum_h \big((a_hx)(b_hy) + (b_hx)(a_hy)\big) = 2(\mu, \mu)xy. \tag{3.2}$$

[1] Any orbit is (Zariski) open in its closure so that an orbit of minimal dimension is closed [2, §8.3].

[2] A connected solvable group acting on a complete variety has a fixed point [2, Theorem 21.2].

Now let $u = x_1 \dots x_k \in B_k$ and compute Cu via the Leibniz Rule:

$$C(u) = \sum_i \sum_h x_1 \dots a_h b_h x_i \dots x_k + \sum_{i<j} \sum_h (x_1 \dots a_h x_i \dots b_h x_j \dots x_k + x_1 \dots b_h x_i \dots a_h x_j \dots x_k).$$

In view of (3.2), this simplifies to

$$C(u) = \big(kC(\mu) + 2\binom{k}{2}(\mu, \mu)\big)u = C(k\mu)u$$

and the discussion in section 2.3 assures us that $u \in V^{k\mu}$.

Thus $B_k = V^{k\mu}$ is irreducible so that if some B_k lay in the kernel of $B \to A$, we would have R_k vanishing on X forcing $X = \{0\}$. We conclude that $A = B = \bigoplus_{k \geq 0} V^{k\mu}$ and $I = J$ as required. □

Remark 3.2. Theorem 3.1 is due to Kostant but appeared, with appropriate attribution, in Lancaster–Towber [3]. Item (1) of the theorem was proved independently by Lichtenstein [4]. The present exposition follows that of Procesi [5] closely.

4. Bilinear relations

4.1. **Incidence of parabolic subalgebras.** Recall that a subalgebra $\mathfrak{p} \leq \mathfrak{g}$ is *parabolic* if it contains a Borel subalgebra. Thus the infinitesimal stabiliser of any point $[v] \in \mathcal{O} \subset \mathbb{P}(V^\lambda)$ in a projective highest weight orbit is parabolic. Moreover, any parabolic subalgebra arises this way many times: the conjugacy class of parabolics stabilising points of $G \cdot [v_\lambda]$, $\lambda = \sum_i m_i \lambda_i$ depends only on $\{i \colon m_i \neq 0\}$.

Parabolic subalgebras admit elaborate combinatorial structures of which perhaps the simplest is an incidence geometry: say that parabolics $\mathfrak{p}_1, \mathfrak{p}_2$ are incident if $\mathfrak{p}_1 \cap \mathfrak{p}_2$ is also parabolic, that is, if the $\mathfrak{p}_i$ contain a common Borel.

We now show that this incidence, viewed as a relation on projective highest weight vectors, is simply a bilinear version of Kostant's quadratic equations.

4.2. **Incidence as a bilinear relation.** For $\lambda, \mu \in \Lambda^+$ and $[v] \in \mathcal{O}_\lambda$, $[w] \in \mathcal{O}_\mu$ points in the corresponding projective highest weight orbits with infinitesimal stabilisers $\mathfrak{p}_\lambda$ and $\mathfrak{p}_\mu$, note that the infinitesimal stabiliser of $[v \otimes w] \in V^\lambda \otimes V^\mu$ is just $\mathfrak{p}_\lambda \cap \mathfrak{p}_\mu$. Thus this intersection is parabolic if and only if $v \otimes w$ is a highest weight vector, necessarily of weight $\lambda + \mu$, and so $[v \otimes w] \in \mathcal{O}_{\lambda+\mu} \subset \mathbb{P}(V^{\lambda+\mu}) \subset \mathbb{P}(V^\lambda \otimes V^\mu)$.

We use theorem 3.1 to show that this obtains precisely when $v \otimes w \in V^{\lambda+\mu}$.

Theorem 4.1. *For $[v] \in \mathcal{O}_\lambda$ and $[w] \in \mathcal{O}_\mu$, $[v \otimes w] \in \mathcal{O}_{\lambda+\mu}$ if and only if $v \otimes w \in V^{\lambda+\mu}$.*

Proof. The forward implication is immediate so let us assume that $u = v \otimes w \in V^{\lambda+\mu}$. By theorem 3.1, we need to show that $u^2 \in V^{2(\lambda+\mu)}$, or, equivalently, $C(u \otimes u) = C(2\lambda + 2\mu)u \otimes u$.

Note that we already have $Cv = C(\lambda)v$, $Cw = C(\mu)w$ and, by hypothesis, $C(v \otimes w) = C(\lambda + \mu)v \otimes w$ from which we conclude

$$\sum_h (a_h v \otimes b_h w + b_h v \otimes a_h w) = \big(C(\lambda + \mu) - C(\lambda) - C(\mu)\big)v \otimes w. \tag{4.1}$$

To compute $C(u\otimes u)$, we reorder the tensor product to make the derivations easier to think about and compute

$$\begin{aligned}C(v\otimes v\otimes w\otimes w) &= C(v\otimes v)\otimes w\otimes w + v\otimes v\otimes C(w\otimes w)\\ &\quad+\sum_h\big(a_h(v\otimes v)\otimes b_h(w\otimes w)+b_h(v\otimes v)\otimes a_h(w\otimes w)\big).\end{aligned}$$

Further applications of the Leibniz rule show that the last summand is a sum of four terms with the following flavour:

$$\sum_h\big((a_hv)\otimes v\otimes(b_hw)\otimes w+(b_hv)\otimes v\otimes(a_hw)\otimes w\big),$$

each of which is $(C(\lambda+\mu)-C(\lambda)-C(\mu))v\otimes v\otimes w\otimes w$ by (4.1). On the other hand, theorem 3.1 tells us that $C(v\otimes v)=C(2\lambda)v\otimes v$ and $C(w\otimes w)=C(2\mu)w\otimes w$ so that

$$C(u\otimes u)=\big(C(2\lambda)+C(2\mu)+4(C(\lambda+\mu)-C(\lambda)-C(\mu))\big)u\otimes u=C(2\lambda+2\mu)u\otimes u,$$

as a short calculation verifies. □

5. Fundamental relations

In an effort to squeeze the last drop of utility from this circle of ideas, we exhibit an affine variety cut out by quadratic relations of the kind considered here whose coordinate ring contains every finite-dimensional, irreducible $\mathfrak{g}$-module exactly once.

For this, let $V=V^{\lambda_1}\oplus\dots\oplus V^{\lambda_l}$ be the sum of the fundamental $\mathfrak{g}$-modules. The coordinate ring $R=\mathbb{C}[V]$ is graded by multi-degree with $R_{m_1\dots m_l}=(S^{m_1}V^{\mu_1})\otimes\dots\otimes(S^{m_l}V^{\mu_l})$, where the μ_i are the permutation of the λ_i for which V^{μ_i} and V^{λ_i} are dual.

Define $\mathfrak{g}$-submodules Q_i, Q_{ij} of total degree 2 by

$$\begin{aligned}S^2V^{\mu_i}&=V^{2\mu_i}\oplus Q_i\\ V^{\mu_i}\otimes V^{\mu_j}&=V^{\mu_i+\mu_j}\oplus Q_{ij}\end{aligned}$$

and let I be the ideal in R generated by the Q_i, $1\le i\le l$, and the Q_{ij}, $1\le i<j\le l$. We prove:

Theorem 5.1. *$R/I\cong\sum_{\mu\in\Lambda^+}V^\mu$. In fact, $(R/I)_{m_1\dots m_l}\cong V^{m_1\mu_1+\dots+m_l\mu_l}$.*

Proof. We will prove the second identity and start by observing that it holds by construction when $\mathbf{m}=(m_1,\dots,m_l)$ has $|\mathbf{m}|\le 2$. In particular, with $x_i\in V^{\mu_i}$ and $y_j\in V^{\mu_j}$, a now familiar argument starting from $x_iy_j\in V^{\mu_i+\mu_j}$ yields

$$\sum_h\big((a_hx_i)(b_hy_j)+(b_hx_i)(a_hy_j)\big)=2(\mu_i,\mu_j)x_iy_j, \tag{5.1}$$

in agreement with (3.2) when $i=j$.

We prove the result by induction on total degree and contemplate $C(xq)$ where $x\in V^{\mu_i}$ and $q\in V^\mu$ has multi-degree $\mathbf{m}$. Then, the induction hypothesis tells us that

$$C(xq)=C(\mu_i)xq+C(\mu)xq+\sum_h\big((a_hx)(b_hq)+(b_hx)(a_hq)\big).$$

Further expansion of the last summand along with $|\mathbf{m}|$ applications of (5.1) yield

$$C(xq)=\big(C(\mu_i)+C(\mu)+2(\mu_i,\mu)\big)=C(\mu_i+\mu)xq$$

so that $xq\in V^{\mu_i+\mu}$ and we are done. □

Remarks.

(1) Such descriptions of all representations by "generators and relations" are discussed in Fulton–Harris [1, p. 235–237] and Lancaster–Towber [3].

(2) The affine space $X \subset V$ cut out by the ideal J of Theorem 5.1 is the affine closure of the so-called "base affine space" of Bernstein–Gelfand–Gelfand.

References

[1] W. Fulton and J. Harris, *Representation theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991. A first course, Readings in Mathematics. MR1153249 (93a:20069)

[2] J. E. Humphreys, *Linear algebraic groups*, Springer-Verlag, New York-Heidelberg, 1975. Graduate Texts in Mathematics, No. 21. MR0396773 (53 #633)

[3] G. Lancaster and J. Towber, *Representation-functors and flag-algebras for the classical groups. I*, J. Algebra **59** (1979), no. 1, 16–38, DOI 10.1016/0021-8693(79)90149-2. MR541667 (80i:14020)

[4] W. Lichtenstein, *A system of quadrics describing the orbit of the highest weight vector*, Proc. Amer. Math. Soc. **84** (1982), no. 4, 605–608. MR643758 (83h:14046)

[5] C. Procesi, *Lie groups: An approach through invariants and representations*, Universitext, Springer, New York, 2007. MR2265844 (2007j:22016)

Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK

E-mail address: feb@maths.bath.ac.uk