4

Algebraic groups and rings of invariants

In general it is hard to construct rings of invariants – that is, to determine explicitly a set of generators and relations. However, this is not actually necessary in order to say that a moduli space exists as an algebraic variety. For this one would like to understand – in the precise manner of a Galois theory, so to speak – the relationship between the invariant ideals in a ring and the ideals in its subring of invariants. What we need here is that the group that is acting is *linearly reductive*: this is the central notion in this chapter.

We begin by giving a careful definition of a representation of an algebraic group. Various important properties can be deduced only by following closely to this definition; for example, it allows us to deduce that all representations are locally finite-dimensional. The set of local distributions supported at the identity in an algebraic group G admits a convolution product, making it into a (noncommutative) algebra $\mathcal{H}(G)$, called the *distribution algebra*. The tangent space of G at the identity element, called the *Lie space* $\mathfrak{g} = \text{Lie } G$, is a vector subspace of $\mathcal{H}(G)$. As is well known, it inherits a Lie algebra structure, although we will not use this in this book. As well as the Lie space, $\mathcal{H}(G)$ also contains a distinguished element Ω, called the Casimir element, constructed using an invariant inner product on the Lie space (Section 4.2). In Section 4.3 we use the Casimir element to prove the linear reductivity of $SL(n)$. We then prove Hilbert's Theorem 4.53 that if a linearly reductive algebraic group acts on a finitely generated algebra, then the invariant subalgebra is finitely generated. The key ingredient in the proof of this is Hilbert's Basis Theorem.

In Section 4.4 we determine the Hilbert series of the rings of classical binary invariants. Using the relation $e \star f - f \star e = h$ in the distribution algebra, we prove the dimension formula for (invariants of) $SL(2)$. As an application we derive the Cayley-Sylvester formula for the Hilbert series of the classical invariants for binary forms.

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

In the final section of this chapter we give an alternative proof of linear reductivity for $SL(2)$. This yields, in addition, a proof of geometric reductivity of $SL(2)$ over a ground field of positive characteristic.

4.1 Representations of algebraic groups

Let $G = \operatorname{Spm} A$ be an affine algebraic group over the field k.

Definition 4.1. An *(algebraic) representation* of the group G (or of the algebra A) is a pair consisting of a vector space V over k and a linear map $\mu_V : V \to V \otimes_k A$ satisfying the following conditions.

(i) The composition

$$V \xrightarrow{\mu_V} V \otimes_k A \xrightarrow{1 \otimes \epsilon} V$$

is the identity, where $\epsilon : A \to k$ is the coidentity.

(ii) The following diagram commutes, where $\mu_A : A \to A \otimes_k A$ is the coproduct.

$$\begin{tikzcd} V \arrow[r, "\mu_V"] \arrow[d, "\mu_V"'] & V \otimes_k A \arrow[d, "\mu_V \otimes 1_A"] \\ V \otimes_k A \arrow[r, "1_V \otimes \mu_A"] & V \otimes_k A \otimes_k A \end{tikzcd}$$

□

Example 4.2. The coordinate ring A of the group G, together with the coproduct μ_A, is itself an algebraic representation. □

Remark 4.3. Let us check that this is equivalent to the usual definition of a representation as a linear action of G on V. If $\rho : G \to GL(n)$ is such a representation, then write $k[X_{ij}, (\det X)^{-1}]$ for the coordinate ring of $GL(n)$ and let $f_{ij} = \rho^* X_{ij} \in A$ be the pull-back of X_{ij} under ρ. In other words, the f_{ij} are the entries of the $n \times n$ matrix representation, viewed as functions on the group G. We then obtain an algebraic representation of G, in the sense of Definition 4.1, by taking an n-dimensional vector space V, a basis $\{e_i\}$, and the linear map given by

$$\mu_V : e_i \mapsto \sum_j e_j \otimes f_{ji}.$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Conversely, suppose that $\mu : V \to A \otimes V$ is a finite-dimensional algebraic representation, and let $x_1, \dots, x_n \in V$ be a basis. Then μ extends naturally to a homomorphism of polynomial rings

$$k[x_1, \dots, x_n] \to A[x_1, \dots, x_n].$$

This is just the same thing as a linear action of G on the dual space of V, viewed as an affine space, and this construction is inverse to the first. □

All of the usual notions concerning group representations can be defined in the spirit of Definition 4.1. In what follows we shall often drop the subscript and write $\mu = \mu_V$ when there is no risk of confusion.

Definition 4.4. Given a representation $\mu : V \to V \otimes A$ of a group $G = \operatorname{Spm} A$:

(i) a vector $x \in V$ is said to be G-invariant if $\mu(x) = x \otimes 1$;
(ii) a subspace $U \subset V$ is called a subrepresentation if $\mu(U) \subset U \otimes A$. □

Remark 4.5. In characteristic zero the coordinate ring A of a connected algebraic group is an integral domain. It follows from this that the above definitions are in this case equivalent to the usual notions for a rational representation $\rho : G \to GL(V)$. (See Exercise 4.8.)

The above definitions have some immediate consequences. The first says, in the language of Remark 4.3, that in any infinite dimensional representation only finitely many of the matrix entries f_{ij} are nonzero:

Proposition 4.6. *Every representation V of G is locally finite-dimensional. In other words, every $x \in V$ is contained in a finite-dimensional subrepresentation of the group.*

Proof. We can write $\mu(x)$ as a finite sum $\sum_i x_i \otimes f_i$ for some $x_i \in V$ and linearly independent elements $f_i \in A$. The linear span $U \subset V$ of the vectors x_i is then exactly what we require. First, it follows from Definition 4.1(i) that

$$x = \sum_i \epsilon(f_i) x_i,$$

so that $x \in U$. Second, the commutative diagram in Definition 4.1(ii) says that

$$\sum_i \mu(x_i) f_i = \sum_i x_i \mu_A(f_i) \in U \otimes A \otimes A.$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Since the f_i are linearly independent, this implies that $\mu(x_i) \in U \otimes A$ for each i, so $U \subset V$ is indeed a (finite-dimensional) subrepresentation. □

For the multiplicative group $\mathbb{G}_\mathrm{m}$ of Example 3.51, the representations are particularly simple to describe. Given a vector space V and an integer $m \in \mathbb{Z}$, consider the map

$$V \to V \otimes k[t, t^{-1}], \qquad v \mapsto v \otimes t^m.$$

This defines an algebraic representation of $\mathbb{G}_\mathrm{m}$, called its representation of *weight m*. By taking direct sums of these representations we get all representations of $\mathbb{G}_\mathrm{m}$.

Proposition 4.7. *Every representation V of $\mathbb{G}_\mathrm{m}$ is a direct sum $V = \sum_{m \in \mathbb{Z}} V_{(m)}$, where each $V_{(m)} \subset V$ is a subrepresentation of weight m.*

Given such a representation $V = \bigoplus V_{(m)}$ of $\mathbb{G}_\mathrm{m}$, a vector $v \in V_{(m)}$ is said to be *homogeneous of weight m*.

Proof. For each integer $m \in \mathbb{Z}$ define

$$V_{(m)} = \{v \in V \mid \mu(v) = v \otimes t^m\}.$$

It is easy to verify that this is a subrepresentation of V (see Exercise 4.4), and by construction it has weight m. The proof that $V = \bigoplus_{m \in \mathbb{Z}} V_{(m)}$ is very similar to the proof of Proposition 4.6: begin by writing, for an arbitrary $v \in V$,

$$\mu(v) = \sum_{m \in \mathbb{Z}} v_m \otimes t^m \in V \otimes k[t, t^{-1}].$$

It follows from Definition 4.1(i) that $v = \sum v_m$, so it just remains to check that each $v_m \in V_{(m)}$. This will prove the direct sum decomposition since obviously $V_{(m)} \cap V_{(n)} = 0$ whenever $m \neq n$. However, Definition 4.1(ii) tells us that

$$\sum_{m \in \mathbb{Z}} \mu(v_m) t^m = \sum_{m \in \mathbb{Z}} v_m \otimes t^m \otimes t^m \in V \otimes A \otimes A,$$

and so by linear independence of the $t^m \in A$ it follows that $\mu(v_m) = v_m \otimes t^m$ for each $m \in \mathbb{Z}$; hence $v_m \in V_{(m)}$. □

It is also easy to classify the representations of the additive group $\mathbb{G}_a$ (Example 3.50). Note, incidentally, that our assumption that the field k has characteristic zero is essential in the next proposition, as well as in the two examples which follow.

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Proposition 4.8. *Every representation V of $\mathbb{G}_{\mathrm{a}} = \mathrm{Spm}\, k[s]$ is given by*

$$\mu(v) = \sum_{n=0}^{\infty} f^n(v) \otimes \frac{s^n}{n!}$$

for some endomorphism $f \in \mathrm{End} V$ which is locally nilpotent (that is, every vector is eventually killed by iterates of f).

Proof. We have a sequence of linear maps $\delta_n : V \to V$ defined by

$$\mu(v) = \sum_{n=0}^{\infty} \delta_n(v) \otimes s^n \in V \otimes k[s].$$

By Definition 4.1 we see that $\delta_0(v) = v$ and

$$\sum_{n=0}^{\infty} \mu(\delta_n(v)) \otimes s^n = \sum_{n=0}^{\infty} \delta_n(v) \otimes (s \otimes 1 + 1 \otimes s)^n,$$

from which it follows that

$$\delta_m \circ \delta_n = \binom{m+n}{m} \delta_{m+n}.$$

The map $f = \delta_1$ therefore has the properties stated in the proposition. □

In the previous chapter (see Definition 3.54) we defined an action of a group $G = \mathrm{Spm}\, A$ on an affine variety $X = \mathrm{Spm}\, R$. We can now interpret this as simply a representation

$$\mu_R : R \to R \otimes_k A$$

which is also a ring homomorphism. The subset of G-invariants

$$R^G = \{f \in R \mid \mu_R(f) = f \otimes 1\}$$

is a subring of R.

Example 4.9. An action of the multiplicative group $\mathbb{G}_{\mathrm{m}} = \mathrm{Spm}\, k[t, t^{-1}]$ on $X = \mathrm{Spm}\, R$ is equivalent to specifying a grading

$$R = \bigoplus_{m \in \mathbb{Z}} R_{(m)}, \qquad R_{(m)} R_{(n)} \subset R_{(m+n)}.$$

The invariants of $\mathbb{G}_{\mathrm{m}}$ are then the homogeneous elements of weight 0 under this grading. Moreover, the linear endomorphism of R which rescales each

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

summand $R_{(m)}$ by m,

$$E : R \to R, \qquad \sum f_m \mapsto \sum m f_m,$$

is a derivation of R. E is called the *Euler operator*. Trivially, the $\mathbb{G}_\mathrm{m}$-invariants in R are the elements killed by E, that is, $R^{\mathbb{G}_\mathrm{m}} = \ker E$. □

Example 4.10. An action of the additive group $\mathbb{G}_\mathrm{a} = \operatorname{Spm} k[s]$ on $X = \operatorname{Spm} R$ is equivalent to specifying a locally nilpotent derivation $D \in \operatorname{End} R$ of the function ring R (see Definition 4.15 below). The action $\mu_R : R \to R \otimes k[s]$ is then given by

$$\mu_R(f) = \sum_{n=0}^{\infty} D^n(f) \otimes \frac{s^n}{n!}.$$

The $\mathbb{G}_\mathrm{a}$-invariants in R are the elements killed by D, that is, $R^{\mathbb{G}_\mathrm{a}} = \ker D$. □

We will later need to consider semiinvariants of group representations as well as invariants (see Chapter 6), and for these we make the next two definitions.

Definition 4.11. Let $G = \operatorname{Spm} A$ be an affine algebraic group. A *(1-dimensional) character* of G is a function $\chi \in A$ satisfying

$$\mu_A(\chi) = \chi \otimes \chi, \qquad \iota(\chi)\chi = 1.$$

□

Note that the characters of G are invertible elements of the function ring A, and in fact they form a multiplicative subgroup of these.

Lemma 4.12. *The characters of the general linear group* $GL(n) = \operatorname{Spm} k[X_{ij}, (\det X)^{-1}]$ *are precisely the integer powers of the determinant* $(\det X)^n$*,* $n \in \mathbb{Z}$*.*

Proof. This is trivial: since $\det(X_{ij})$ is an irreducible polynomial, the only invertible elements of $k[X_{ij}, (\det X)^{-1}]$ are, up to multiplication by a scalar, the powers $(\det X)^n$. For every $n \in \mathbb{Z}$ and scalar $\lambda \in k$, moreover, $\lambda(\det X)^n$ is a character precisely when $\lambda = 1$. □

Definition 4.13. Let χ be a character of an affine algebraic group G, and let V be a representation of G. A vector $x \in V$ satisfying

$$\mu_V(x) = x \otimes \chi$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

is called a semiinvariant of G with weight χ. The semiinvariants of V belonging to a given character χ of G form a subrepresentation (see Exercise 4.4), which we shall denote by $V_\chi \subset V$. □

An algebraic group $T = \operatorname{Spm} A$ which is isomorphic to a direct product of copies of $\mathbb{G}_{\mathrm{m}}$ is called an *algebraic torus*. In this case the set $X(T)$ of characters of T is a basis over k of the algebra A. The following fact follows from Proposition 4.7.

Proposition 4.14. *Let T be an algebraic torus and let $X(T)$ be its set of characters. Then every representation V of T is the direct sum of all its semiinvariant subrepresentations:*

$$V = \bigoplus_{\chi \in X(T)} V_{(\chi)}.$$

4.2 Algebraic groups and their Lie spaces

In this section we will define the Casimir operator associated to a representation of an algebraic group.

(a) Local distributions

Definition 4.15. Let R be a commutative ring over k and M an R-module (see Chapter 8). An *M-valued derivation* is a k-linear map

$$D : R \to M$$

satisfying the Leibniz rule $D(xy) = xD(y) + yD(x)$ for $x, y \in R$. □

An R-valued derivation $D : R \to R$ will simply be referred to as a derivation of R.

Example 4.16. Let $R = k[t_1, \ldots, t_n]$. The first examples of derivations are the partial derivatives

$$\frac{\partial}{\partial t_i} : R \to R.$$

Fixing $a_1, \ldots, a_n \in k$, a second example is the evaluation

$$\alpha : R \to k, \qquad f \mapsto \frac{\partial}{\partial t_i}(a_1, \ldots, a_n).$$

This is a derivation with values in the R-module $k = R/(t_1 - a_1, \ldots, t_n - a_n)$. □

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

One can generalise this example to any affine variety. Let p be a point of $X = \operatorname{Spm} R$, with maximal ideal $\mathfrak{m}_p \subset R$. Then a $k = R/\mathfrak{m}_p$-valued derivation is a linear map

$$\alpha : R \to k$$

with the property that

$$\alpha(fg) = f(p)\alpha(g) + g(p)\alpha(f)$$

for all $f, g \in R$. We shall sometimes refer to α as a *derivation of X at the point $p \in X$*. Note, in particular, that such a derivation vanishes on $\mathfrak{m}_p^2 \subset R$. It is this idea that we want to generalise next.

Definition 4.17. Let p be a point of $X = \operatorname{Spm} R$ with maximal ideal $\mathfrak{m}_p \subset R$. A *local distribution with support at $p \in X$* is a k-linear map $\alpha : R \to k$ with the property that $\alpha(\mathfrak{m}_p^N) = 0$ for sufficiently large $N \in \mathbb{N}$. □

The degree $\deg \alpha$ of a local distribution supported at p is the minimum $d \in \mathbb{N}$ such that $\alpha(\mathfrak{m}_p^{d+1}) = 0$. Every local distribution of degree 0 is a scalar multiple of the evaluation map

$$ev_p : R \to k, \qquad f \mapsto f(p).$$

Lemma 4.18. *For a k-linear map $\alpha : R \to k$, the following are equivalent:*

(i) α is a derivation of $X = \operatorname{Spm} R$ at the point $p \in X$;
(ii) α is a local distribution, supported at $p \in X$, of degree 1 and satisfying $\alpha(1) = 0$.

Proof. (i) $\Longrightarrow$ (ii) It has already been observed that α is a local distribution of degree 1, while $\alpha(1) = \alpha(1 \cdot 1) = \alpha(1) + \alpha(1)$. Hence $\alpha(1) = 0$.

(ii) $\Longrightarrow$ (i) If $f, g \in R$, then $f - f(p), g - g(p) \in \mathfrak{m}_p$, so that $\deg \alpha = 1$ implies

$$\alpha((f - f(p))(g - g(p))) = 0.$$

Expanding and using the fact that $\alpha(f(p)g(p)) = f(p)g(p)\alpha(1) = 0$ gives

$$\alpha(fg) = \alpha(f)g(p) + \alpha(g)f(p),$$

as required. □

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

It follows from the lemma that the vector space $\mathrm{Der}_k(R, R/\mathfrak{m}_p)$ of derivations of $X = \mathrm{Spm}\, R$ at $p \in X$ is isomorphic to the dual of $\mathfrak{m}_p/\mathfrak{m}_p^2$.

Definition 4.19. The k-vector space $(\mathfrak{m}_p/\mathfrak{m}_p^2)^\vee \cong \mathrm{Der}_k(R, R/\mathfrak{m}_p)$ is called the *Zariski tangent space* of X at the point p. □

Remark 4.20. The dimension of the Zariski tangent space at $p \in X$ is greater than or equal to the dimension of the variety X at p; X can be defined to be nonsingular at p if and only if the two dimensions are equal. (This is equivalent to Definition 9.44 in Section 9.3(b).) Over a field k of characteristic zero an algebraic group is always nonsingular. □

We now have a vector space isomorphism

$$\{\text{local distributions with degree} \le d\} \cong (R/\mathfrak{m}_p^{d+1})^\vee,$$

and when $d = 1$ this decomposes into

$$k \oplus (\mathfrak{m}_p/\mathfrak{m}_p^2)^\vee,$$

where, by Lemma 4.18, the two summands are spanned by the evaluation map ev_p and by derivations at p, respectively.

More generally, for each $d \le e$ the natural projection $R/\mathfrak{m}_p^{e+1} \to R/\mathfrak{m}_p^{d+1}$ induces an injection

$$(R/\mathfrak{m}_p^{d+1})^\vee \hookrightarrow (R/\mathfrak{m}_p^{e+1})^\vee.$$

There is therefore an ascending sequence:

$$k \subset (R/\mathfrak{m}_p^2)^\vee \subset (R/\mathfrak{m}_p^3)^\vee \subset \cdots \subset (R/\mathfrak{m}_p^{d+1})^\vee \subset \cdots.$$

The space of local distributions supported at p can thus be identified with the limit (that is, the union) of this sequence.

(b) The distribution algebra

If $G = \mathrm{Spm}\, A$ is an affine algebraic group with coordinate ring A, we will denote by $\mathcal{H}(G)$ the vector space of distributions $\alpha : A \to k$ supported at the identity element $e \in G$. The Zariski tangent space of G at this point is called the *Lie space* of G and is denoted by $\mathfrak{g} = \mathrm{Lie}(G) \subset \mathcal{H}(G)$.

Remark 4.21. The vector space $\mathfrak{g}$ acquires from $\mathcal{H}(G)$ (together with its convolution product, which we are about to define) the structure of a *Lie algebra*.

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

This is outlined in Exercise 4.3. However, we are not going to use the Lie algebra structure in this book. □

Let $\mu = \mu_A : A \to A \otimes A$ be the coproduct on G.

Definition 4.22. If $\alpha, \beta \in \mathcal{H}(G)$ are distributions supported at the identity, the *convolution product* $\alpha \star \beta$ of α and β is the composition

$$A \xrightarrow{\mu} A \otimes A \xrightarrow{\alpha \otimes \beta} k \otimes k \xrightarrow{\sim} k.$$

□

Lemma 4.23. *The convolution product of $\alpha, \beta \in \mathcal{H}(G)$ is again a distribution supported at the identity $\alpha \star \beta \in \mathcal{H}(G)$, and*

$$\deg \alpha \star \beta \le \deg \alpha + \deg \beta.$$

Proof. Since $(e, e) \mapsto e$ under the group operation $G \times G \to G$, we have

$$\mu(\mathfrak{m}) \subset \mathfrak{m} \otimes A + A \otimes \mathfrak{m},$$

where $\mathfrak{m} = \mathfrak{m}_e$. Since μ is a ring homomorphism, this implies, for $a, b \in \mathbb{N}$,

$$\mu(\mathfrak{m}^{a+b+1}) \subset \sum_{i+j=a+b+1} \mathfrak{m}^i \otimes \mathfrak{m}^j.$$

Taking $a = \deg \alpha$ and $b = \deg \beta$, it follows from $\alpha(\mathfrak{m}^{a+1}) = 0$ and $\beta(\mathfrak{m}^{b+1}) = 0$ that $\alpha \star \beta(\mathfrak{m}^{a+b+1}) = 0$, which proves the lemma. □

Evaluation at the identity $\epsilon = ev_e \in \mathcal{H}(G)$ is an identity element for the convolution product. Moreover, it follows from the associative law for the coproduct μ (Definition 3.48(i)) that $\star$ is associative and thus makes $\mathcal{H}(G)$ into an associative algebra, called the *distribution algebra* of the algebraic group G.

Remark 4.24. In general, the distribution algebra is infinite-dimensional and noncommutative. If k has characteristic zero, it is a theorem of Cartier that $\mathcal{H}(G)$ is the universal enveloping algebra of a Lie algebra. □

Example 4.25. Consider the multiplicative group $G = \mathbb{G}_\mathrm{m} = \operatorname{Spm} k[t, t^{-1}]$. The vector space of distributions supported at the identity is, by definition,

$$\mathcal{H}(\mathbb{G}_\mathrm{m}) = \lim_{n \to \infty} \left(k[t]/(t-1)^n \right)^\vee.$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

As an algebra this is isomorphic to the polynomial ring $k[E]$ where

$$E = \frac{d}{dt}\bigg|_{t=1} : k[t, t^{-1}] \to k,$$

and $\mathrm{Lie}(\mathbb{G}_{\mathrm{m}})$ is the 1-dimensional linear span of E. In fact, by definition $E \star E$ takes $f(t) \in k[t, t^{-1}]$ to

$$\frac{\partial^2.}{\partial t \partial t'} f(tt')\bigg|_{t=t'=1} = \left(\frac{d}{dt}\right)^2 f(t)\bigg|_{t=1}.$$

Similarly, the n-th power $E^{\star n} = E \star \cdots \star E$ is equal to $(\partial^2/\partial t^n)|_{t=1}$. □

Example 4.26. Similarly to the previous example, the additive group $\mathbb{G}_{\mathrm{a}} = \mathrm{Spm}\, k[s]$ has distribution algebra

$$\mathcal{H}(\mathbb{G}_{\mathrm{a}}) = \lim_{n\to\infty} \big(k[s]/(s^n)\big)^{\vee} = k[D].$$

This is the polynomial ring generated by

$$D = \frac{d}{ds}\bigg|_{s=0} : k[s] \to k.$$

□

A homomorphism of algebraic groups $G \to G'$ induces a ring homomorphism $\mathcal{H}(G) \to \mathcal{H}(G')$ and a linear map of Lie spaces $\mathrm{Lie}(G) \to \mathrm{Lie}(G')$. (This is also a homomorphism of Lie algebras in the sense of Exercise 4.3.) Note that the induced homomorphism of function rings is in the reverse direction, $A' \to A$, but our constructions dualise this once more so that both functors $\mathcal{H}$ and Lie are covariant.

We return now to the representations of $G = \mathrm{Spm}\, A$. Let $\mu_V : V \to V \otimes A$ be an algebraic representation, with an associated linear representation $\rho : G \to GL(V)$. For each distribution $\alpha \in \mathcal{H}(G)$ consider the k-linear map

$$V \xrightarrow{\mu} V \otimes A \xrightarrow{1\otimes\alpha} V \otimes k \xrightarrow{\sim} V.$$

We will denote this composition by $\widetilde{\rho}(\alpha) \in \mathrm{End}\, V$. From the associativity of μ (Definition 4.1(ii)) we obtain:

Lemma 4.27. *The map* $\widetilde{\rho} : \mathcal{H}(G) \to \mathrm{End}\, V$ *is a ring homomorphism.* □

In other words, the representation V is a (noncommutative) $\mathcal{H}(G)$-module. A vector $v \in V$ is G-invariant if and only if

$$\widetilde{\rho}(\alpha)v = \alpha(1)v$$

for all $\alpha \in \mathcal{H}(G)$. In particular (by Lemma 4.18(i)), $\widetilde{\rho}(\alpha)v = 0$ if $\alpha \in \mathfrak{g} = \mathrm{Lie}(G)$. That is, the Lie space $\mathfrak{g}$ kills all of the G-invariants in V.

Example 4.28. Let V be a representation of $\mathbb{G}_\mathrm{m}$, and decompose a vector $v \in V$ into its homogeneous components $v = \sum_{m\in\mathbb{Z}} v_m$ under the action of $\mathbb{G}_\mathrm{m}$. Then

$$\mu(v) = \sum v_m \otimes t^m,$$

and so the generating distribution

$$E = \left.\frac{d}{dt}\right|_{t=1} \in \mathcal{H}(\mathbb{G}_\mathrm{m}) = k[E]$$

acts by $\widetilde{\rho}(E) : v \mapsto \sum m v_m$. In other words, it coincides with the Euler operator of Example 4.9. □

Consider the action of G on itself by conjugation:

$$G \times G \to G, \qquad (x, g) \mapsto gxg^{-1}.$$

This induces an action of G on its coordinate ring A. Since the identity element $e \in G$ is fixed under conjugation, the action preserves the maximal ideal $\mathfrak{m} = \mathfrak{m}_e$, and in particular it induces a k-linear action on each quotient $A/\mathfrak{m}^n$ and on its dual space. It follows that $\mathcal{H}(G)$ becomes a linear representation of G. On the other hand, given a representation $\rho : G \to GL(V)$, the space End V also becomes a linear representation, by conjugation $T \mapsto \rho(g)T\rho(g)^{-1}$ for $T \in$ End V and $g \in G$. With respect to this action we have:

Lemma 4.29. *The map $\widetilde{\rho} : \mathcal{H}(G) \to$ End V is a homomorphism of G-representations.* □

In particular, the Lie space $\mathfrak{g}$ is a subrepresentation of $\mathcal{H}(G)$; this is called the *adjoint representation* and is denoted by

$$\mathrm{Ad} : G \to GL(\mathfrak{g}).$$

Example 4.30. Consider the general linear group $GL(n)$ with coordinate ring $A = k[X_{ij}, (\det X)^{-1}]$. The Lie space $\mathfrak{gl}(n)$ of $GL(n)$ is the k-vector space

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

with basis

$$E_{ij} = \left.\frac{\partial}{\partial X_{ij}}\right|_{X=I_n} : k[X_{ij}, (\det X)^{-1}] \to k$$

for $1 \leq i, j \leq n$. (Note that this generalises Example 4.25.) This space is isomorphic to the vector space of $n \times n$ matrices over k, by mapping E_{ij} to the matrix with a 1 in the (i, j)-th entry and 0s elsewhere. With this identification the adjoint representation is

$$\mathrm{Ad}(g) : \mathfrak{gl}(n) \to \mathfrak{gl}(n) \qquad M \mapsto gMg^{-1}.$$

Example 4.31. The special linear group $SL(n)$ has coordinate ring $A = k[X_{ij}]/(\det X - 1)$. We can identify its Lie space $\mathfrak{s}l(n)$ as follows. It is the Zariski tangent space $(\mathfrak{m}/\mathfrak{m}^2)^\vee$, where $\mathfrak{m} \subset A$ is the ideal at the identity matrix, and a tangent vector is therefore a ring homomorphism $f : A \to k[t]/(t^2)$ for which the composition $A \xrightarrow{f} k[t]/(t^2) \to k[t]/(t) = k$ coincides with the map $A \to A/\mathfrak{m}$. In other words, if we write ϵ for the residue class t mod t^2 and $k[\epsilon] = k[t]/(t^2)$, then a tangent vector is a matrix $I + \epsilon M$ which satisfies

$$1 = \det(I + \epsilon M) = 1 + \epsilon \operatorname{tr} M,$$

since $\epsilon^2 = 0$. Hence $\mathfrak{s}l(n) \subset \mathfrak{gl}(n)$ is the vector space of $n \times n$ matrices over k with trace zero. □

(c) The Casimir operator

Consider now an inner product on the Lie space, that is, a symmetric and nondegenerate bilinear form

$$\kappa : \mathfrak{g} \times \mathfrak{g} \to k.$$

We will assume that κ is invariant under the adjoint representation $\mathrm{Ad} : G \to GL(\mathfrak{g})$.

Definition 4.32. Let κ be a G-invariant inner product on $\mathfrak{g}$, as above. Let $X_1, \dots, X_N \in \mathfrak{g}$ be a basis of $\mathfrak{g}$ and let $X'_1, \dots, X'_N \in \mathfrak{g}$ be its dual basis with respect to κ. The distribution

$$\Omega = X_1 \star X'_1 + \cdots + X_N \star X'_N \in \mathcal{H}(G)$$

is called the *Casimir element* over G with respect to κ. □

Proposition 4.33. *The Casimir element* Ω *is independent of the choice of basis* $\{X_1, \dots, X_N\}$.

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Proof. A second basis $\{Y_1, \dots, Y_N\}$, with dual basis $\{Y_1', \dots, Y_N'\}$ is related to the first by

$$Y_i = \sum_{j=1}^{N} a_{ij} X_j, \quad Y_i' = \sum_{j=1}^{N} a_{ij}' X_j', \qquad i = 1, \dots, N,$$

where $A = (a_{ij})$, $A' = (a_{ij}')$ are some matrices satisfying $A^t A' = I_N$. So we compute

$$\begin{aligned}\sum_{i=1}^{N} Y_i \star Y_i' &= \sum_{i=1}^{N} \left(\sum_{j=1}^{N} a_{ij} X_j\right) \star \left(\sum_{k=1}^{N} a_{ik}' X_k'\right) \\ &= \sum_{j,k} \left(\sum_{i=1}^{N} a_{ij} a_{ik}'\right) X_j \star X_k' \\ &= \sum_{j,k} \delta_{jk} X_j \star X_k' = \Omega,\end{aligned}$$

as required. □

For the Casimir element for $SL(n)$, see Example 4.48 below.

Since the inner product κ is assumed to be G-invariant, for each $g \in G$ the sets

$$\{\mathrm{Ad}(g)(X_1), \dots, \mathrm{Ad}(g)(X_N)\}, \qquad \{\mathrm{Ad}(g)(X_1'), \dots, \mathrm{Ad}(g)(X_N')\},$$

are again dual bases. We therefore deduce:

Corollary 4.34. *The Casimir element $\Omega \in \mathcal{H}(G)$ is invariant under the action of G on the distribution algebra.* □

Let $\rho : G \to GL(V)$ be a representation of G. This is an $\mathcal{H}(G)$-module via the homomorphism $\widetilde{\rho} : \mathcal{H}(G) \to \mathrm{End}\, V$ of Lemma 4.27. In particular, the Casimir element Ω determines a linear endomorphism of V,

$$\widetilde{\rho}(\Omega) : V \to V,$$

called the *Casimir operator* (with respect to the inner product κ). By Lemma 4.29 and Corollary 4.34 this is invariant under the conjugation action

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

of G on End V: that is, it commutes with the action of G. Moreover, since $\mathfrak{g}$ kills the G-invariants $V^G \subset V$, so does the Casimir operator. In other words:

Corollary 4.35. *The Casimir operator is an endomorphism of each representation V of G, and*

$$V^G \subset \ker\,(\widetilde{\rho}(\Omega))\,.$$

4.3 Hilbert's Theorem

(a) Linear reductivity

Definition 4.36. An algebraic group G is said to be *linearly reductive* if, for every epimorphism $\phi : V \to W$ of G representations, the induced map on G-invariants $\phi^G : V^G \to W^G$ is surjective. □

There are various equivalent definitions (see also Lemma 4.74 below):

Proposition 4.37. *The following conditions are all equivalent.*

(i) G is linearly reductive.

(ii) For every epimorphism $\phi : V \to W$ of finite-dimensional representations the induced map on G-invariants $\phi^G : V^G \to W^G$ is surjective.

(iii) If V is any finite-dimensional representation and $v \in V$ is G-invariant modulo a proper subrepresentation $U \subset V$, then the coset $v + U$ contains a nontrivial G-invariant vector.

Proof. (i) implies (ii) trivially. Applying (ii) to the quotient map $V \to V/U$ gives (iii), so we just have to show that (iii) implies (i). We suppose that $\phi : V \to W$ is an epimorphism of G representations and that $\phi(v) = w \in W^G$ for some $v \in V$. By local finite dimensionality (Proposition 4.6) there exists a finite-dimensional subrepresentation $V_0 \subset V$ containing v. Now $v \in V_0$ is G-invariant modulo the subrepresentation $U_0 = V_0 \cap \ker \phi$, so by property (iii) there exists a G-invariant vector $v' \in V_0$ such that $v' - v \in U_0$. Since $\phi(v') = w$, we have shown that $\phi^G : V^G \to W^G$ is surjective. □

Proposition 4.38. *Every finite group is linearly reductive.*

Proof. Suppose that V is a finite-dimensional representation and $v \in V$ is a vector invariant modulo a subrepresentation $U \subset V$, and set

$$v' = \frac{1}{|G|} \sum_{g \in G} g \cdot v.$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Clearly v' is G-invariant, while

$$v' - v = \frac{1}{|G|}\sum_{g \in G}(g \cdot v - v)$$

is contained in U. So we have verified condition (iii) of Proposition 4.37. □

Remark 4.39. The homomorphism $R : V \to V^G$ used in this proof, $R = \frac{1}{|G|}\sum_{g \in G} g$, is called a *Reynolds operator* and corresponds to *Cayley's* Ω-*process* in the work of Hilbert. (See Sturmfels [28].) One could, alternatively, prove the proposition by using R to verify Definition 4.36 directly, but we have used the criterion of Proposition 4.37 because this is the approach that we will take to prove the linear reductivity of $SL(n)$ (Theorem 4.43). □

Direct products of linearly reductive algebraic groups are linearly reductive; moreover, if $H \subset G$ is a normal subgroup and G is linearly reductive, then so is the quotient G/H. Conversely, if both H and G/H are linearly reductive, then G is linearly reductive.

Example 4.40. If G is an algebraic group whose connected component at the identity is linearly reductive, then G is linearly reductive. □

Proposition 4.41. *Every algebraic torus* $(\mathbb{G}_\mathrm{m})^N$ *is linearly reductive.*

Proof. It is enough to prove this for $T = \mathbb{G}_\mathrm{m}$; again, we shall check condition (iii) of Proposition 4.37. By Proposition 4.7, a representation V and a subrepresentation U have weight decompositions

$$V = \bigoplus_{m \in \mathbb{Z}} V_{(m)}, \qquad U = \bigoplus_{m \in \mathbb{Z}} U_{(m)},$$

with $U_{(m)} \subset V_{(m)}$. T-invariance of an element $v = \sum v_{(m)}$ modulo U means that $v_{(m)} \in U$ for all $m \neq 0$. It follows that $v_{(0)}$ is a T-invariant element of the coset $v + U$, as required. □

Example 4.42. An example of a group which is not linearly reductive is the additive group $\mathbb{G}_\mathrm{a} \cong k$. To see this, consider the 2-dimensional representation given by

$$\mathbb{G}_\mathrm{a} \to GL(2), \qquad t \mapsto \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Then restriction to the x-axis,

$$V := k[x, y] \to W := k[x, y]/(y) = k[x],$$

is a surjective homomorphism of $\mathbb{G}_a$-representations. But $V^{\mathbb{G}_a} = k[y]$, $W^{\mathbb{G}_a} = k[x]$, so the induced homomorphism on invariants is not surjective. □

Our aim in the remainder of this section is to prove:

Theorem 4.43. *The special linear group $SL(n)$ is linearly reductive.*

The general linear group $GL(n)$ is generated by its centre, which is isomorphic to $\mathbb{G}_m$, and the subgroup $SL(n)$. It can therefore be expressed as a quotient $GL(n) = (\mathbb{G}_m \times SL(n))/\mathbb{Z}_n$, and so we obtain:

Corollary 4.44. *$GL(n)$ is linearly reductive.* □

Remark 4.45. The proof of Theorem 4.43 will be modelled on that of Proposition 4.38, using a Reynold's operator $R : V \to V^{SL(n)}$ (Remark 4.39). In this case R will be constructed using the Casimir operator for the representation V. For the case $n = 2$, on the other hand, we will give an alternative and more direct proof in Section 4.5. □

Let U be any finite-dimensional vector space. The Lie space of $GL(U)$ is canonically isomorphic to End U, and the adjoint representation is the conjugation action of $GL(U)$ on this space (Example 4.30). Associating to a pair of endomorphisms of U the trace of their composite

$$\kappa : \text{End } U \times \text{End } U \to k, \qquad (f, g) \mapsto \text{tr } fg, \tag{4.1}$$

defines a nondegenerate inner product (symmetric bilinear form) on End U. We will write $\kappa(f) = \kappa(f, f)$. The following is clear.

Lemma 4.46. *κ is invariant under the adjoint action of $GL(U)$. In other words,*

$$\kappa(\alpha f \alpha^{-1}) = \kappa(f)$$

is satisfied for all $f \in$ End U and $\alpha \in GL(U)$. □

The Lie space $\mathfrak{sl}(U)$ of the special linear group $SL(U)$, as a subgroup of $GL(U)$, is the subalgebra of End U consisting of trace zero endomorphisms

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

(Example 4.31). We shall denote the restriction of the inner product κ to $\mathfrak{sl}(U)$ by the same symbol.

Lemma 4.47. *κ is a nondegenerate inner product on $\mathfrak{sl}(U)$ invariant under the adjoint action of $SL(U)$.*

Proof. κ is nondegenerate on $\operatorname{End} U$, and with respect to κ the subspace $\mathfrak{sl}(U)$ is the orthogonal complement of the identity element I_U. Since $\kappa(I_U) = \dim U \neq 0$, it follows that κ is nondegenerate on $\mathfrak{sl}(U)$. □

Example 4.48. Let us calculate the Casimir element for $SL(n)$ using this inner product. Let e_{ij} denote the sparse $n \times n$ matrix with a single 1 in the i-th row and the j-th column. Let $f_{ij} = e_{ji}$ and, for $i = 1, \ldots, n-1$, let

$$h_i = e_{ii} - e_{i+1,i+1}, \qquad m_i = e_{11} + \cdots + e_{ii} - \frac{i}{n}(e_{11} + \cdots + e_{nn}).$$

Then the Lie space $\mathfrak{sl}(n)$ has dual bases

$$\{e_{ij}\}_{i<j} \cup \{f_{ij}\}_{i<j} \cup \{h_i\}_i, \qquad \{f_{ij}\}_{i<j} \cup \{e_{ij}\}_{i<j} \cup \{m_i\}_i.$$

By Definition 4.32, we now compute

$$\Omega = \sum_{i<j}(e_{ij} \star f_{ij} + f_{ij} \star e_{ij}) + e_{11}^2 + \cdots + e_{nn}^2 - \frac{1}{n}(e_{11} + \cdots + e_{nn})^2.$$

For example, in the case $n = 2$, the Casimir element is

$$\Omega = e \star f + f \star e + \tfrac{1}{2} h \star h,$$

where

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(See also Section 4.4(a).) □

Proposition 4.49. *For a representation $\rho : SL(n) \to GL(V)$ the following are equivalent.*

(i) *The representation is trivial.*
(ii) $\widetilde{\rho}(\Omega) = 0$.
(iii) $\operatorname{tr} \widetilde{\rho}(\Omega) = 0$.

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Proof. (i) implies (ii) by Corollary 4.35, and this implies (iii) trivially; so we just have to show that (iii) implies (i).

Let $T \subset SL(n)$ be the torus subgroup of diagonal matrices and $\mathfrak{h} \subset \mathfrak{sl}(n)$ its Lie space. Under the action of T the representation V has, by Proposition 4.14, a character space decomposition

$$V = \bigoplus_{\chi \in X(T)} V_\chi .$$

Each $\chi : T \to \mathbb{G}_{\mathrm{m}}$ corresponds to a linear form with integer coefficients $\overline{\chi} : \mathfrak{h} \to k$, and for each $h \in \mathfrak{h}$ we have

$$\operatorname{tr} \widetilde{\rho}(h)^2 = \sum_{\chi \neq 1} (\dim V_\chi) \overline{\chi}(h)^2.$$

It follows that if $\operatorname{tr} \widetilde{\rho}(h) = 0$ for all $h \in \mathfrak{h}$, then $\dim V_\chi = 0$ for all nontrivial characters $\chi \in X(T)$ – or, in other words, V is a trivial representation of the torus T. Since diagonalisable elements are dense in $SL(n)$, this implies that V is also trivial as a representation of $SL(n)$.

The proposition is proved, therefore, if we can show that $\operatorname{tr} \widetilde{\rho}(\Omega) = 0$ implies $\operatorname{tr} \widetilde{\rho}(h) = 0$ for all $h \in \mathfrak{h}$. We will do this just for $SL(2)$; the general case is similar. We have seen that the Casimir element is

$$\Omega = e \star f + f \star e + \tfrac{1}{2} h \star h \in \mathcal{H}(SL(2))$$

$$= \tfrac{1}{2}(e+f)^2 + \tfrac{1}{2}(\sqrt{-1}e - \sqrt{-1}f)^2 + \tfrac{1}{2}h^2,$$

where the matrices

$$e + f = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sqrt{-1}(e-f) = \begin{pmatrix} 0 & \sqrt{-1} \\ -\sqrt{-1} & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are conjugate in $\mathfrak{sl}(2)$. Hence

$$\operatorname{tr} \widetilde{\rho}(\Omega) = \frac{3}{2} \operatorname{tr} \widetilde{\rho}(h)^2,$$

and we are done. □

In particular, if the Casimir operator $\rho(\Omega)$ of a representation V is nilpotent, then the representation is trivial. Applying this to the subrepresentation $\ker \rho(\Omega)^n$ shows that $\ker\, (\widetilde{\rho}(\Omega))^m \subset V^{SL(n)}$ for any integer $m \geq 0$. Combining this with Corollary 4.35, we conclude:

Corollary 4.50. $V^{SL(n)} = \bigcup_{m \geq 0} \ker\, (\widetilde{\rho}(\Omega))^m.$ □

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Proof of Theorem 4.43. Given a finite-dimensional representation V of $SL(n)$ we shall construct a Reynolds operator $R : V \to V^{SL(n)}$ analogous to the averaging operator $R = 1/|G| \sum_{g \in G} g$ used in the proof of Proposition 4.38 for the case of a finite group. This is constructed from the Casimir operator $C_V := \widetilde{\rho}(\Omega) \in \operatorname{End} V$ as follows. Let $N = \dim V$, and let

$$\chi_V(t) = t^N + c_1 t^{N-1} + \cdots + c_{N-m} t^m$$

be the characteristic polynomial of C_V, where $c_{N-m} \neq 0$. The Reynolds operator is constructed by substituting the Casimir into the polynomial

$$P(t) := \frac{1}{c_{N-m} t^m} \chi_V(t) = \frac{1}{c_{N-m}} (t^{N-m} + c_1 t^{N-m-1} + \cdots + c_{N-m}).$$

That is, $R := P(C_V)$. This is a homomorphism of $SL(n)$-representations, and by the Cayley-Hamilton theorem it satisfies

$$C_V^m P(C_V) = 0.$$

It follows from Corollary 4.50 that the image of $P(C_V)$ is contained in $V^{SL(n)}$.

We can now follow the proof of Proposition 4.38 and apply the criterion in Proposition 4.37(iii) for linear reductivity. Suppose that $U \subset V$ is a subrepresentation and that $v \in V$ is $SL(n)$-invariant modulo U. This means that $\widetilde{\rho}(\mathfrak{sl}(n))v \subset U$, and hence $C_V v \in U$. It follows from the way we have defined $P(t)$ that $P(C_V)v - v \in U$, and so the vector $v' = P(C_V)v$ satisfies the requirements of Proposition 4.37(iii). □

(b) Finite generation

Let G be an algebraic group acting on a polynomial ring S, preserving the grading.

Theorem 4.51 (Hilbert [19]). *If G is linearly reductive, then the ring of invariant polynomials S^G is finitely generated.*

Proof. We will essentially follow the original reasoning of Hilbert. The invariant ring is graded by

$$S^G = \bigoplus_{e \geq 0} S^G \cap S_e.$$

Let $S_+^G \subset S^G$ be the span of the invariants of positive degree and denote by $J \subset S$ the ideal generated in S by S_+^G. By Theorem 2.2, in fact, J is generated

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

by finitely many polynomials $f_1, \ldots, f_N \in S^G_+$. In other words, the S-module homomorphism

$$\phi : S \oplus \cdots \oplus S \to J \qquad (h_1, \ldots, h_N) \mapsto \sum_{i=1}^{N} h_i f_i$$

is surjective.

Claim: S^G is generated by $f_1, \ldots, f_N$.

We pick an arbitrary homogeneous invariant $h \in S^G$. To show that h belongs to $k[f_1, \ldots, f_N]$ we shall use induction on $\deg h$. If $\deg h = 0$, then h is a constant, so this is clear. If $\deg h > 0$, then h belongs to the homogeneous ideal J, and therefore to the invariant subspace J^G, where we can view J as a representation of G. The map ϕ above is a surjective homomorphism of G-representations, so by linear reductivity the induced map of invariants $S^G \oplus \cdots \oplus S^G \to J^G$ is surjective. There therefore exist invariant polynomials $h'_1, \ldots, h'_N \in S^G$ such that

$$h = \sum_{i=1}^{N} h'_i f_i.$$

The f_i all have positive degree, so $\deg h'_i < \deg h$. By the inductive hypothesis we may therefore assume that each $h'_i \in k[f_1, \ldots, f_N]$ and hence $h \in k[f_1, \ldots, f_N]$ also. □

Note that the last part of this proof is really just Proposition 2.41.

We turn now to the general case of G acting on a finitely generated k-algebra R. We shall show that, again, the invariant subring R^G is finitely generated, by reducing to the case of a polynomial ring as above.

Lemma 4.52. *Suppose that an algebraic group acts on a finitely generated k-algebra R. Then there exists a set of generators $r_1, \ldots, r_N$ of R whose k-linear span $\langle r_1, \ldots, r_N \rangle \subset R$ is a G-invariant vector subspace.*

Proof. Let $s_1, \ldots, s_M \in R$ be any set of generators. By local finiteness (Proposition 4.6) each s_i is contained in a finite-dimensional subrepresentation $V_i \subset R$ of G. It therefore suffices to extend $s_1, \ldots, s_M$ to a basis $r_1, \ldots, r_N$ of the finite-dimensional subspace $\sum_i V_i \subset R$. □

Geometrically, this lemma says that an affine algebraic variety acted on by an algebraic group can always be equivariantly embedded in an affine space $\mathbb{A}^N$ on which G acts linearly.

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Theorem 4.53. *If a linearly reductive algebraic group G acts on a finitely generated k-algebra R, then the invariant ring R^G is finitely generated.*

Proof. Pick generators $r_1, \dots, r_N$ of R as in Lemma 4.52. Then there is a surjective k-algebra homomorphism

$$S = k[x_1, \dots, x_N] \to R$$

given by mapping $x_i \mapsto r_i$ for each $i = 1, \dots, N$. Via this map G acts on the ring S, and by Theorem 4.51 the invariant ring S^G is finitely generated, while by linear reductivity the induced map $S^G \to R^G$ is surjective. It follows that R^G is finitely generated. □

4.4 The Cayley-Sylvester Counting Theorem

In order to gain a concrete understanding of any invariant ring it is essential to be able to compute its Hilbert series. In this section we shall describe some methods for determining the Hilbert series for the case of classical binary invariants.

(a) SL(2)

Let us write a general unimodular 2×2 matrix as

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2).$$

The Lie space $\mathfrak{sl}(2)$ of $SL(2)$ has a basis consisting of three derivations:

$$e = \left.\frac{\partial}{\partial b}\right|_{X=I}, \quad f = \left.\frac{\partial}{\partial c}\right|_{X=I}, \quad h = \left.\left(\frac{\partial}{\partial a} - \frac{\partial}{\partial d}\right)\right|_{X=I}.$$

These correspond to the three subgroups of $SL(2)$

$$N^+ = \left\{ \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \mid s \in k \right\} \quad \cong \mathbb{G}_{\mathrm{a}},$$

$$N^- = \left\{ \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix} \mid s \in k \right\} \quad \cong \mathbb{G}_{\mathrm{a}},$$

$$T = \left\{ \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \mid t \in k^\times \right\} \cong \mathbb{G}_{\mathrm{m}}.$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

We can represent these basis elements, in the manner of Example 4.30, as matrices

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The adjoint action of $SL(2)$ on $\mathfrak{sl}(2)$ is by conjugation, and its restriction to $T \subset SL(2)$ is therefore given by

$$\mathrm{Ad}\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} : \begin{cases} e \mapsto t^2 e \\ f \mapsto t^{-2} f \\ h \mapsto 0. \end{cases} \tag{4.2}$$

The invariant inner product (4.1) from Section 4.3(a) is given by the symmetric matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

with respect to the basis $e, f, h \in \mathfrak{sl}(2)$. We may therefore construct an orthonormal basis

$$\frac{e+f}{\sqrt{2}}, \quad \frac{e-f}{\sqrt{-2}}, \quad \frac{h}{\sqrt{2}}$$

and, as we have seen in Example 4.48, Casimir element

$$\Omega = e \star f + f \star e + \tfrac{1}{2} h \star h \in \mathcal{H}(SL(2)). \tag{4.3}$$

We now consider the basic representation $S = k[x, y]$ of $SL(2)$, and the action on S of the Lie space $\mathfrak{s}l(2) \subset \mathcal{H}(SL(2))$ and of the Casimir element on S, via the homomorphism $\widetilde{\rho} : \mathcal{H}(SL(2)) \to \mathrm{End}\, S$. For simplicity we will usually drop the tilde and write just $\rho : \mathcal{H}(SL(2)) \to \mathrm{End}\, S$. This should not cause any confusion.

Example 4.54. $SL(2)$ acts on the right on the polynomial algebra $S = k[x, y]$ by

$$f(x, y) \cdot \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = f(\alpha x + \beta y, \gamma x + \delta y).$$

Let us compute the derivative at the identity of the restriction of this action to the subgroup $N^+ \subset SL(2)$:

$$f(x, y) \cdot \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} = f(x + sy, y),$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

so we obtain

$$\frac{d}{ds}\bigg|_{s=0}\left\{f(x,y)\cdot\begin{pmatrix}1 & s\\ 0 & 1\end{pmatrix}\right\}=y\frac{\partial}{\partial x}f(x,y).$$

Similarly, for $N^{-}\subset SL(2)$ we find

$$\frac{d}{ds}\bigg|_{s=0}\left\{f(x,y)\cdot\begin{pmatrix}1 & 0\\ s & 1\end{pmatrix}\right\}=x\frac{\partial}{\partial y}f(x,y).$$

For the subgroup $T\subset SL(2)$ the restriction of the adjoint action is given by

$$f(x,y)\cdot\begin{pmatrix}t & 0\\ 0 & t^{-1}\end{pmatrix}=f(tx,t^{-1}y),$$

and we must differentiate at $t=1$:

$$\frac{d}{dt}\bigg|_{t=1}\left\{f(x,y)\cdot\begin{pmatrix}t & 0\\ 0 & t^{-1}\end{pmatrix}\right\}=x\frac{\partial f}{\partial x}-y\frac{\partial f}{\partial y}.$$

We conclude that the representation of the Lie space $\rho:\mathfrak{sl}(2)\to\operatorname{End}S$ is given by

$$\rho(e)=y\frac{\partial}{\partial x},\quad \rho(f)=x\frac{\partial}{\partial y},\quad \rho(h)=x\frac{\partial}{\partial x}-y\frac{\partial}{\partial y}.$$

From (4.3) the Casimir operator is

$$\begin{aligned}\rho(\Omega)&=\rho(e)\rho(f)+\rho(f)\rho(e)+\tfrac{1}{2}\rho(h)^2\\ &=E+\tfrac{1}{2}E^2,\end{aligned}$$

where $E=x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}$. □

In this example, for each $d\in\mathbb{N}$ the binary forms of degree d give a subrepresentation $V_d\subset S$. Note that by Euler's Theorem the Casimir operator on V_d is the scalar $d+d^2/2$.

Remark 4.55. The following inhomogeneous description of the representation V_d will reappear at the end of this chapter. Namely, V_d can be viewed as the $d+1$-dimensional vector subspace of $k[x]$ consisting of polynomials of degree at most d. Then $SL(2)$ acts on the right on V_d by:

$$f(x)\cdot\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}=(\gamma x+\delta)^d f(\frac{\alpha x+\beta}{\gamma x+\delta}).$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

(b) The dimension formula for SL(2)

If V is a finite-dimensional representation of $SL(2)$, then under the action of the torus $T \subset SL(2)$ it has a weight-space decomposition (see Proposition 4.7)

$$V = \bigoplus_{m\in\mathbb{Z}} V_{(m)}. \tag{4.4}$$

The following follows from (4.2):

Proposition 4.56 (Weight shift). *Under $\rho : \mathfrak{sl}(2) \to \operatorname{End} V$ we have*

$$\rho(e) : V_{(m)} \to V_{(m+2)}, \qquad \rho(f) : V_{(m)} \to V_{(m-2)}.$$

Lemma 4.57. *In the distribution algebra $\mathcal{H}(SL(2))$ the following relation holds:*

$$e \star f - f \star e = h.$$

Proof. We fix independent variables a, b, c, d and a', b', c', d' as the entries of matrices

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad X' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}.$$

By definition of the convolution product, the element $e \star f$ evaluated on a polynomial $F(X)$ in a, b, c, d gives

$$\left.\frac{\partial^2 F(XX')}{\partial b \partial c'}\right|_{X=X'=I}.$$

One easily checks that this is equal to

$$\left.\frac{\partial^2 F(X)}{\partial b \partial c}\right|_{X=I} + \left.\frac{\partial F(X)}{\partial a}\right|_{X=I}.$$

Similarly, the value of $f \star e$ at the polynomial $F(X)$ is

$$\left.\frac{\partial^2 F(X)}{\partial b \partial c}\right|_{X=I} + \left.\frac{\partial F(X)}{\partial d}\right|_{X=I}.$$

Subtracting these two expressions yields the identity in the lemma. □

Corollary 4.58. *The Casimir element of $SL(2)$ is*

$$\Omega = e \star f + f \star e + \tfrac{1}{2} h \star h = 2e \star f - h + \tfrac{1}{2} h \star h = 2f \star e + h + \tfrac{1}{2} h \star h.$$

□

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

This allows us to locate explicitly the invariants of $SL(2)$ in the representation V:

Corollary 4.59. $V^{SL(2)} = \ker\{e : V_{(0)} \to V_{(2)}\}$.

Proof. Clearly $V^{SL(2)} \subset V_{(0)}$, and indeed it is contained in ker e since it is killed by the Lie space $\mathfrak{sl}(2)$. But conversely, every $v \in \ker e$ is killed by the Casimir operator, by Corollary 4.58. By Proposition 4.50, this implies that $v \in V^{SL(2)}$. □

The dimension formula for $V^{SL(2)}$ will follow once we show that the map in Corollary 4.59 is surjective. In fact:

Lemma 4.60. *The composition* $e^2 : V_{(-2)} \xrightarrow{e} V_{(0)} \xrightarrow{e} V_{(2)}$ *is an isomorphism.*

Proof. First let us show that e^2 is injective. If $v \in \ker e^2$, then the vector $e(v) \in V_{(0)}$ is $SL(2)$-invariant by Corollary 4.59. In particular, $fe(v) = 0$. It follows that

$$\rho(\Omega)v = \rho(2f \star e + h + \tfrac{1}{2}h \star h)v = 0,$$

so that, by Proposition 4.50, $v \in V^{SL(2)}$. But v lies in the -2-weight space, which contains no nonzero invariants; hence $v = 0$.

By a similar argument the homomorphism $f^2 : V_{(2)} \to V_{(-2)}$ is injective. It follows that $\dim V_{(-2)} = \dim V_{(2)}$, and hence that e^2 is an isomorphism. □

It follows from this that $e : V_{(0)} \to V_{(2)}$ is surjective, and we deduce:

Dimension Formula 4.61. *If V is any finite-dimensional representation of $SL(2)$, with weight-space decomposition (4.4) with respect to the torus $T \subset SL(2)$, then*

$$\dim V^{SL(2)} = \dim V_{(0)} - \dim V_{(2)}.$$

□

The generating function

$$\mathrm{ch}_V(q) = \sum_{m\in\mathbb{Z}} \dim V_{(m)} q^m \in \mathbb{Z}[q, q^{-1}]$$

of the weight-space decomposition (4.4) is called the *(formal) character* of the representation V. For example, the space V_d of binary forms of degree d

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

has character

$$\mathrm{ch}_{V_d}(q) = q^{-d} + q^{-d+2} + \cdots + q^{d-2} + q^d = \frac{q^{d+1} - q^{-d-1}}{q - q^{-1}}.$$

Corollary 4.62. $\dim V^{SL(2)} = -\operatorname*{Res}_{q=0}(q - q^{-1})\mathrm{ch}_V(q)$. □

(c) A digression: Weyl measure

By Cauchy's integral formula we can re-express Corollary 4.62:

$$\dim V^{SL(2)} = -\frac{1}{2\pi i}\oint (q - q^{-1})\mathrm{ch}_V(q)dq,$$

where the integral is taken with winding number 1 around the origin. Taking the unit circle with parametrisation $q = e^{i\theta}$ transforms the integral to

$$\dim V^{SL(2)} = \frac{1}{\pi}\int_0^{\pi}(1 - \cos 2\theta)\mathrm{ch}_V(e^{i\theta})d\theta. \tag{4.5}$$

(We have used here the Weyl symmetry $\mathrm{ch}_V(q) = \mathrm{ch}_V(q^{-1})$. This can be seen from the definition of the character, using conjugation by the element $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in SL(2)$.)

Note that every conjugacy class of the maximal compact subgroup $SU(2) \subset SL(2, \mathbb{C})$ has a unique representative of the form

$$A(\theta) = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \qquad \text{for } 0 \leq \theta \leq \pi.$$

Thus, if we define a *Weyl measure*

$$\mu(\theta) = 1 - \cos 2\theta = 2\sin^2\theta,$$

then, noting that $\int_0^{\pi}\mu(\theta)d\theta = \pi$, we see that the dimension formula (4.5) has the form of an average of the class function $\mathrm{ch}_V(\theta) = \mathrm{ch}_V(e_{i\theta})$ with respect to the measure μ.

Let us compare this situation with the Dimension Formula 1.11 for representations of a finite group G. If we denote by $\mathfrak{c}_1, \ldots, \mathfrak{c}_k \subset G$ its conjugacy classes, then, noting that the character $\chi : G \to \mathbb{C}$ of the representation is a class function, the Dimension Formula 1.11 can be written:

$$\dim V^G = \frac{1}{|G|}\sum_{i=1}^{k}|\mathfrak{c}_i|\chi(\mathfrak{c}_i).$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

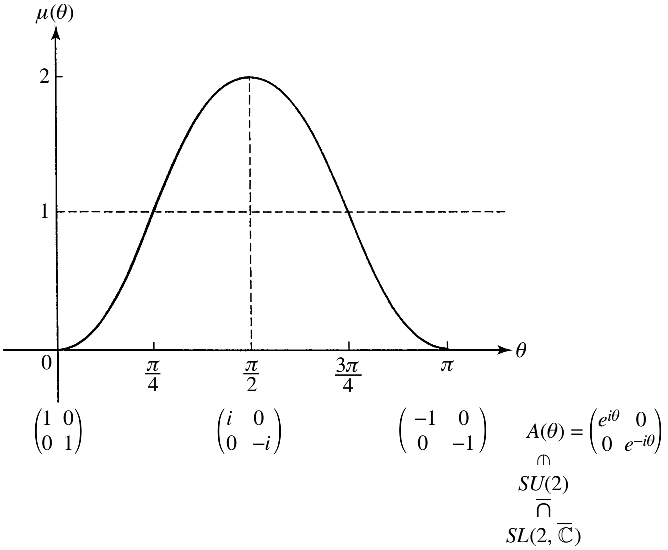

Figure 4.1: *Weyl measure* $\mu(\theta)=2\sin^2\theta$

In other words (again noting that $\sum_{i=1}^{k}\mu(\mathfrak{c}_i)=|G|$), the dimension is given as the average of the character with respect to the measure $\mu(\mathfrak{c})=|\mathfrak{c}|$.

In conclusion, then, in both cases of G a finite group or the compact Lie group $SU(2)$ we see that the dimension formula can be interpreted as the average of the character over the conjugacy classes of the group, with respect to cardinality or the Weyl measure, respectively.

(d) The Cayley-Sylvester Formula

As an application of the dimension formula we are now going to compute the Hilbert series of the classical binary invariant ring. If V is any n-dimensional representation of $SL(2)$, we consider the induced action of $SL(2)$ on the polynomial ring $S(V)=k[x_1,\ldots,x_n]$ of functions on V. Let $a_1,\ldots,a_n\in\mathbb{Z}$ be the weights (not necessarily distinct) of the torus $T\subset SL(2)$ (see Section 4.4(a)) occuring in the weight-space decomposition of the representation V (Proposition 4.14). The function

$$P(q;t)=\frac{1}{(1-q^{a_1}t)(1-q^{a_2}t)\cdots(1-q^{a_n}t)}=\det\left(I_V-t\begin{pmatrix}q & 0\\0 & q^{-1}\end{pmatrix}_V\right)^{-1}$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

is called the *q-Hilbert series* of the representation. Then Molien's Theorem 1.10 for finite groups has the following analogue for $SL(2)$.

Proposition 4.63. *The invariant ring $S(V)^{SL(2)}$ has Hilbert series*

$$P(t) = -\operatorname*{Res}_{q=0}\,(q - q^{-1})P(q;t).$$

Equivalently, if $P(q;t) = \sum_{m\in\mathbb{Z}} c_m(t)q^m$, then $P(t) = c_0(t) - c_2(t)$.

Proof. This is similar to the proof of Molien's Theorem. First, by making a linear change of coordinates we can assume that $x_1, \dots, x_n$ diagonalise the action of $T \cong \mathbb{G}_\mathrm{m}$. Then we note that the power series expansion of

$$R(x_1, \dots, x_n) = \frac{1}{(1-x_1)(1-x_2)\cdots(1-x_n)}$$

lists once each all the monomials of the ring S, and that the action on this expression of

$$\begin{pmatrix} q & 0 \\ 0 & q^{-1} \end{pmatrix} \in T$$

yields

$$R(q^{a_1}x_1, \dots, q^{a_n}x_n) = \frac{1}{(1-q^{a_1}x_1)(1-q^{a_2}x_2)\cdots(1-q^{a_n}x_n)}.$$

When this is expanded, the sum of the coefficients in degree e is precisely the formal character $\mathrm{ch}_{V_e}(q)$ of the representation V_e. It follows from this that

$$\sum_{e=0}^{\infty} \mathrm{ch}_{V_e}(q)t^e = R(q^{a_1}t, \dots, q^{a_n}t) = P(q;t).$$

From Corollary 4.62

$$\dim S^{SL(2)} \cap V_e = -\operatorname*{Res}_{q=0}\,(q - q^{-1})\mathrm{ch}_{V_e}(q),$$

and hence the Hilbert series is

$$P(t) = -\operatorname*{Res}_{q=0} \sum_{e=0}^{\infty}(q - q^{-1})\mathrm{ch}_{V_e}(q)t^e = -\operatorname*{Res}_{q=0}\,(q - q^{-1})P(q;t).$$

□

In particular, consider the representation $V = V_d$ of binary forms of degree d (see Example 4.54). The $d+1$ monomials x^d, $x^{d-1}y, \dots, y^d$ are already a basis

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

diagonalising the action of T and are transformed to

$$q^d x^d, q^{d-2}x^{d-1}y, \dots, q^{-d}y^d$$

by the element $\begin{pmatrix} q & 0 \\ 0 & q^{-1} \end{pmatrix}$. The q-Hilbert series is therefore

$$P(q;t) = \prod_{i=0}^{d} \frac{1}{1-q^{d-2i}t}.$$

Definition 4.64. The q-analogue of an integer d, of its factorial and of the binomial coefficients are, respectively:

(i) $[d]_q = q^{d-1} + q^{d-3} + \cdots + q^{-d+3} + q^{-d+1} = \frac{q^d - q^{-d}}{q-q^{-1}},$

(ii) $$[d]_q! = \prod_{i=1}^{d} [i]_q,$$

(iii) $$\begin{bmatrix} d+e \\ e \end{bmatrix}_q = \frac{[d+e]_q!}{[d]_q![e]_q!}.$$

□

The corresponding classical notions are obtained by letting $q \to 1$. In this sense the following proposition is the q-analogue of the binomial theorem

$$\frac{1}{(1-t)^{d+1}} = \sum_{e \geq 0} \binom{d+e}{e} t^e.$$

Proposition 4.65.

$$\prod_{i=0}^{d} \frac{1}{1-q^{d-2i}t} = \sum_{e \geq 0} \begin{bmatrix} d+e \\ e \end{bmatrix}_q t^e.$$

Proof. Denote the left-hand side by

$$\phi(q,t) = \prod_{i=0}^{d} \frac{1}{1-q^{d-2i}t},$$

and its power series expansion in t by

$$\phi(q,t) = \sum_{e \geq 0} a_e(q) t^e, \qquad a_0(q) = 1.$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Note that $\phi(q,t)$ satisfies the functional equation

$$\phi(q, q^2 t) = \frac{1 - q^{-d} t}{1 - q^{d+2} t} \phi(q, t).$$

Comparing terms on both sides, this gives

$$a_e(q) q^{2e} - a_{e-1} q^{2e+d} = a_e(q) - a_{e-1} q^{-d}.$$

Rearranging, we obtain the recurrence relation

$$a_e(q) = \frac{q^{e+d} - q^{-e-d}}{q^e - q^{-e}} a_{e-1}(q)$$

for the coefficients of $\phi(q,t)$, from which the proposition follows. □

Returning to the representation V_d of $SL(2)$, with basis $x^d, x^{d-1}y, \ldots, y^d$, let $\xi_0, \ldots, \xi_d$ be the dual basis of $V_d^\vee$. From Propositions 4.63 and 4.65 we deduce:

Theorem 4.66. *The Hilbert series of the classical invariant ring $k[\xi_0, \ldots, \xi_d]^{SL(2)}$ for binary forms of degree d is given by:*

$$P^{(d)}(t) = -\sum_{e \geq 0} \left\{ \operatorname*{Res}_{q=0} (q - q^{-1}) \begin{bmatrix} d+e \\ e \end{bmatrix}_q \right\} t^e.$$

For the purpose of computing it is convenient to make a change of variable $u = q^2$. Then

$$\begin{aligned} \begin{bmatrix} d+e \\ e \end{bmatrix}_q &= \frac{[e+1]_q [e+2]_q \cdots [e+d]_q}{[1]_q [2]_q \cdots [d]_q} \\ &= q^{-de} \frac{(1-u^{e+1})(1-u^{e+2}) \cdots (1-u^{e+d})}{(1-u)(1-u^2) \cdots (1-u^d)} \end{aligned}$$

and (note that the denominator begins with the quadratic factor!)

$$-(q - q^{-1}) \begin{bmatrix} d+e \\ e \end{bmatrix}_q = q^{-de-1} \frac{(1-u^{e+1})(1-u^{e+2}) \cdots (1-u^{e+d})}{(1-u^2) \cdots (1-u^d)}. \qquad (4.6)$$

For a formal power series $f(u) \in \mathbb{Z}[[u]]$ we shall denote the coefficient of u^n by $[f(u)]_n \in \mathbb{Z}$.

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Cayley-Sylvester Formula 4.67. *The vector space $k[\xi_0, \ldots, \xi_d]_e^{SL(2)}$ of classical invariants of degree e for binary d-ics has dimension*

$$m(d,e) = \begin{cases} \left[\dfrac{(1-u^{e+1})(1-u^{e+2})\cdots(1-u^{e+d})}{(1-u^2)\cdots(1-u^d)}\right]_{de/2} & \textit{if } de \textit{ is even,} \\ 0 & \textit{if } de \textit{ is odd.} \end{cases}$$

Proof. The dimension $m(d,e)$ is equal to the residue appearing in Theorem 4.66, and we may note that, if de is odd, then this residue vanishes since the expansion of (4.6) contains only even powers of q. We shall therefore assume that de is even. Writing

$$R(u) = \frac{(1-u^{e+1})(1-u^{e+2})\cdots(1-u^{e+d})}{(1-u^2)\cdots(1-u^d)},$$

we have

$$m(d,e) = \frac{1}{2\pi i}\oint q^{-de-1}R(q^2)dq,$$

where the path of integration is a small circle around $q = 0$. Under the change of variable $u = q^2$, $du = 2qdq$, this is equal to

$$\frac{1}{2\pi i}\oint u^{-de/2}\frac{u^{-1}}{2}R(u)du,$$

where the contour now has winding number 2 about $u = 0$ and is therefore

$$\begin{aligned} m(d,e) &= \frac{1}{2\pi i}\oint u^{-de/2-1}R(u)du \\ &= \operatorname*{Res}_{u=0} u^{-de/2-1}R(u) \\ &= [R(u)]_{de/2}. \end{aligned}$$

□

Corollary 4.68 (Hermite reciprocity). $m(d,e) = m(e,d)$. □

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

(e) Some computational examples

Proposition 4.69. *For $2 \leq d \leq 6$, the Hilbert series $P^{(d)}(t)$ of the classical invariant ring for binary d-ics is given by the following table:*

d	$P^{(d)}(t)$
2	$\dfrac{1}{1-t^2}$
3	$\dfrac{1}{1-t^4}$
4	$\dfrac{1}{(1-t^2)(1-t^3)}$
5	$\dfrac{1+t^{18}}{(1-t^4)(1-t^8)(1-t^{12})}$
6	$\dfrac{1+t^{15}}{(1-t^2)(1-t^4)(1-t^6)(1-t^{10})}$

Proof. We shall just do the cases $d = 4, 5$, leaving the others to the reader. From the Cayley-Sylvester formula,

$$m(4,e) = \left[\frac{(1-u^{e+1})(1-u^{e+2})(1-u^{e+3})(1-u^{e+4})}{(1-u^2)(1-u^3)(1-u^4)}\right]_{2e}.$$

In this expression we can expand the numerator, ignoring terms of degree greater than $2e$:

$$\begin{aligned} m(4,e) &= \left[\frac{1-u^{e+1}-u^{e+2}-u^{e+3}-u^{e+4}}{(1-u^2)(1-u^3)(1-u^4)}\right]_{2e} \\ &= \left[\frac{1}{(1-u)(1-u^{3/2})(1-u^2)}\right]_e - \left[\frac{u+u^2+u^3+u^4}{(1-u^2)(1-u^3)(1-u^4)}\right]_e \\ &= \left[\frac{1+u^{3/2}}{(1-u)(1-u^2)(1-u^3)}\right]_e - \left[\frac{u}{(1-u)(1-u^2)(1-u^3)}\right]_e \end{aligned}$$

(where for the last term we have simply factorised $1-u^4$). Exchanging $u^{3/2}$ and u between the two numerators, and noting that the second then has no integer powers in its expansion, we see that

$$m(4,e) = \left[\frac{1}{(1-u^2)(1-u^3)}\right]_e,$$

and hence that $P^{(4)}(t) = 1/(1-t^2)(1-t^3)$.

We now turn to the case $d = 5$. Since $m(5,e) = 0$ whenever e is odd, it is enough to consider even values $e = 2a$. Then

$$m(5,2a) = \left[\frac{(1-u^{2a+1})(1-u^{2a+2})(1-u^{2a+3})(1-u^{2a+4})(1-u^{2a+5})}{(1-u^2)(1-u^3)(1-u^4)(1-u^5)}\right]_{5a}.$$

Expanding the numerator and ignoring terms of degree greater than $5a$, we obtain:

$$\begin{aligned} m(5,2a) = &\left[\frac{1}{(1-u^2)(1-u^3)(1-u^4)(1-u^5)}\right]_{5a} \\ &-\left[\frac{u+u^2+u^3+u^4+u^5}{(1-u^2)(1-u^3)(1-u^4)(1-u^5)}\right]_{3a} \\ &+\left[\frac{u^3+u^4+2u^5+2u^6+2u^7+u^8+u^9}{(1-u^2)(1-u^3)(1-u^4)(1-u^5)}\right]_{a}. \end{aligned} \tag{4.7}$$

We deal with each of these three brackets. The first can be rewritten:

$$\begin{aligned} &\left[\frac{1}{(1-u^2)(1-u^3)(1-u^4)(1-u^5)}\right]_{5a} \\ &= \left[\frac{(1+u^2+u^4+u^6+u^8)(1+u^3+u^6+u^9+u^{12})(1+u^4+u^8+u^{12}+u^{16})}{(1-u^{10})(1-u^{15})(1-u^{20})(1-u^5)}\right]_{5a} \\ &= \left[\frac{1+u^5+4u^{10}+5u^{15}+7u^{20}+4u^{25}+3u^{30}}{(1-u^5)(1-u^{10})(1-u^{15})(1-u^{20})}\right]_{5a}, \end{aligned}$$

where, in the last step, the numerator has been expanded ignoring terms that are not divisible by 5.

The second bracket in (4.7) can be rearranged similarly:

$$\left[\frac{u+u^2+u^3+u^4+u^5}{(1-u^2)(1-u^3)(1-u^4)(1-u^5)}\right]_{3a}$$

$$=\left[\frac{u(1+u^2+u^4)(1+u^4+u^8)(1+u+u^2)}{(1-u^6)(1-u^3)(1-u^{12})(1-u^3)}\right]_{3a}$$

$$=\left[\frac{2u^3+2u^6+3u^9+u^{12}+u^{15}}{(1-u^3)^2(1-u^6)(1-u^{12})}\right]_{3a}.$$

The third bracket is:

$$\left[\frac{u^3+u^4+2u^5+2u^6+2u^7+u^8+u^9}{(1-u^2)(1-u^3)(1-u^4)(1-u^5)}\right]_a$$

$$=\left[\frac{u^3(1+u+u^2+u^3+u^4)(1+u^2)}{(1-u^2)(1-u^3)(1-u^4)(1-u^5)}\right]_a$$

$$=\left[\frac{u^3}{(1-u)(1-u^2)^2(1-u^3)}\right]_a.$$

It follows from these computations that the Hilbert series is given by

$$P^{(5)}(\sqrt{t})=\frac{1+t+4t^2+5t^3+7t^4+4t^5+3t^5}{(1-t)(1-t^2)(1-t^3)(1-t^4)}-\frac{2t+2t^2+3t^3+t^4+t^5}{(1-t)^2(1-t^2)(1-t^4)}$$

$$+\frac{t^3}{(1-t)(1-t^2)^2(1-t^3)}$$

$$=\frac{1+t^9}{(1-t^2)(1-t^4)(1-t^6)}.$$

□

In the cases $d=2,3$, the discriminant $D(\xi)\in k[\xi_0,\dots,\xi_d]^{SL(2)}$ has degree 2, 4, respectively (Examples 1.21 and 1.22).

For the case $d=4$, we have constructed in Chapter 1 (see Section 1.3(b)) invariants $g_2(\xi), g_3(\xi)\in k[\xi_0,\dots,\xi_4]^{SL(2)}$ of degrees 2, 3, respectively. In fact, g_2 and g_3 are algebraically independent: this can be seen by restricting to the subspace $\xi_0=\xi_4$, $\xi_1=\xi_3=0$, on which the invariants reduce to

$$g_2(\xi)=\xi_0^2+3\xi_2^2, \qquad g_3(\xi)=(\xi_0^2-\xi_2^2)\xi_2.$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

These determine a map $k^2 \to k^2$, $(\xi_0, \xi_2) \mapsto (g_2, g_3)$ which is clearly surjective (since one can separate variables), so g_2, g_3 cannot satisfy any polynomial identity. From Propositions 4.69 and 1.9 we can therefore conclude:

Corollary 4.70. *The ring of invariants $k[\xi_0, \dots, \xi_d]^{SL(2)}$ is generated by the discriminant $D(\xi)$ when $d = 2, 3$, and by $g_2(\xi), g_3(\xi)$ when $d = 4$.* □

The higher degree cases are less simple, but by constructing the invariant rings explicitly the following results are known. (See also Schur [26].)

Example 4.71. For $d = 5$, the Hilbert series can be written as

$$P^{(5)}(t) = \frac{1 - t^{36}}{(1 - t^4)(1 - t^8)(1 - t^{12})(1 - t^{18})},$$

and, indeed, the invariant ring $k[\xi_0, \dots, \xi_5]^{SL(2)}$ has four generators of degrees 4, 8, 12, 18 satisfying a single relation of degree 36. Similar for $d = 6$,

$$P^{(6)}(t) = \frac{1 - t^{30}}{(1 - t^2)(1 - t^4)(1 - t^6)(1 - t^{10})(1 - t^{15})},$$

and the invariant ring has five generators of degrees 2, 4, 6, 10, 15 and a single relation of degree 30. □

Example 4.72 (Shioda [27]). For $d = 8$, the Hilbert series is

$$\begin{aligned} P^{(8)}(t) &= \frac{1 + t^8 + t^9 + t^{10} + t^{18}}{(1 - t^2)(1 - t^3)(1 - t^4)(1 - t^5)(1 - t^6)(1 - t^7)} \\ &= \frac{1 + \sum_{a=16}^{20} t^a + \sum_{b=25}^{29} t^b - t^{45}}{(1 - t^2)(1 - t^3)(1 - t^4)(1 - t^5)(1 - t^6)(1 - t^7)(1 - t^8)(1 - t^9)(1 - t^{10})}, \end{aligned}$$

where the first expression is obtained from the Cayley-Sylvester formula, and the second is a convenient rearrangement. In this case the ring $k[\xi_0, \dots, \xi_8]^{SL(2)}$ is generated by nine invariants $J_2(\xi), \dots, J_{10}(\xi)$. These satisfy five relations of degrees $16, \dots, 20$, which in turn satisfy five syzygies. More precisely, the relations can be expressed as the Pfaffians of the five principal 4 × 4 minors of a skew-symmetric matrix

$$\begin{pmatrix} 0 & f_6(J) & f_7(J) & f_8(J) & f_9(J) \\ & 0 & f_8(J) & f_9(J) & f_{10}(J) \\ & & 0 & f_{10}(J) & f_{11}(J) \\ & - & & 0 & f_{12}(J) \\ & & & & 0 \end{pmatrix},$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

where each $f_i(J)$ is a weighted homogeneous polynomial of degree i, with $\deg J_m = m$. □

Remark 4.73. In fact, a Gorenstein ring of codimension 3 is always defined by an odd number $2k + 1$ of relations in its generators, and, by a theorem of Buchsbaum and Eisenbud [22], these relations can always be expressed as the Pfaffians of the principal $2k \times 2k$ minors of a skew-symmetric matrix, as above. □

The case $d = 7$ we prefer quietly to omit, though the interested reader may like to compute $P^{(7)}(t)$ for him or herself, or consult Dixmier and Lazard [24].

4.5 Geometric reductivity of $SL(2)$

We shall give in this section an alternative proof of linear reductivity in the special case of $SL(2)$. We begin by extending Proposition 4.37.

Lemma 4.74. *For an algebraic group G the following conditions are equivalent.*

(i) G is linearly reductive.

(ii) Given a finite-dimensional representation V of G and a surjective G-invariant linear form $f : V \to k$ there exists an invariant vector $w \in V^G$ such that $f(w) \neq 0$.

(iii) Given a finite-dimensional representation V of G and an invariant vector $w \in V^G$ there exists an G-invariant linear form $f : V \to k$ such that $f(w) \neq 0$.

Proof. (i) implies (ii) immediately from Definition 4.36, with G acting trivially on $W = k$. Conversely, (ii) implies (i) using the formulation of Proposition 4.37(ii). If $v \in W^G$, then we can decompose W as a representation of G, as $W = k\{v\} \oplus W'$ (Exercise 4.6). Then by condition (ii) the composition $V^G \subset V \to W \to k\{v\}$ is surjective.

(ii) is equivalent to (iii) by replacing V by its dual $V^\vee$ and noting that the space of G-invariant forms is $\mathrm{Hom}_G(V^\vee, k) = \mathrm{Hom}_G(k, V) = V^G$, where G acts trivially on k. □

It is the formulation of linear reductivity given by part (iii) of the lemma that we shall verify for $SL(2)$. First, by linear reductivity of $\mathbb{G}_\mathrm{m} \cong T \subset SL(2)$, we can find a T-invariant linear form $e : V \to k$ such that $e(w) = 1$, and using

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

this form we can define a map

$$\phi : V \to k[SL(2)]$$

by $\phi(x)(g) = e(g \cdot x)$ for $x \in V$ and $g \in SL(2)$. Equivalently, ϕ is the composition of the group action (see Definition 4.1) with $e \otimes 1$:

$$V \xrightarrow{\mu_V} V \otimes k[SL(2)] \xrightarrow{e\otimes 1} k \otimes k[SL(2)] = k[SL(2)].$$

The following properties of ϕ are easy to check.

Lemma 4.75.

(i) $\phi(w)$ is the constant function 1.
(ii) For all $x \in V$ the function $\phi(x)$ is invariant under the right-action of T, that is, $\phi(V) \subset k[SL(2)]^T$.
(iii) ϕ is a homomorphism of $SL(2)$ representations.

We will give an explicit description of the invariant ring $k[SL(2)]^T$. The coordinate ring of $SL(2)$ is

$$k[SL(2)] = k[x, y, z, t]/(xt - yz - 1),$$

and this carries left- and right-actions of the group by left- and right-translation:

$$\begin{pmatrix} x & y \\ z & t \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix}\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}.$$

We consider the right-action restricted to the torus $T \subset SL(2)$ and the left-action of $SL(2)$ on the invariant subalgebra $k[SL(2)]^T$. Since

$$\begin{pmatrix} x & y \\ z & t \end{pmatrix}\begin{pmatrix} q & 0 \\ 0 & q^{-1} \end{pmatrix} = \begin{pmatrix} qx & q^{-1}y \\ qz & q^{-1}t \end{pmatrix},$$

we have

$$k[x, y, z, t]^T = k[xy, xt, zy, zt].$$

The polynomial $xt - yz - 1$ is itself T-invariant, and so

$$k[SL(2)]^T = k[xy, xt, zy, zt]/(xt - yz - 1).$$

We can give this ring another description. For each natural number $n \in \mathbb{N}$ let R_n be the following set of rational functions in variables u, v:

$$R_n = \left\{ \frac{f(u, v)}{(u - v)^n} \mid \deg_u f(u, v) \le n,\ \deg_v f(u, v) \le n \right\}.$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

This is a vector space of dimension $(n+1)^2$, and $R_n \subset R_{n+1}$. The union

$$R = \bigcup_{n \geq 0} R_n = \lim_{n \to \infty} R_n$$

is a subalgebra of $k[u, v, 1/(u-v)] \subset k(u, v)$, while the function field $k(u, v)$ is a representation of $SL(2)$ via

$$u \mapsto \frac{au+b}{cu+d}, \quad v \mapsto \frac{av+b}{cv+d}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2).$$

Note that under this action

$$u - v \mapsto \frac{au+b}{cu+d} - \frac{av+b}{cv+d} = \frac{u-v}{(cu+d)(cv+d)},$$

from which it follows that $R_n \subset k(u, v)$ is a subrepresentation. More precisely, R_n is isomorphic to the tensor product $V_n \otimes V_n$, where V_n is the $(n+1)$-dimensional irreducible representation of $SL(2)$ as described in Remark 4.55.

Lemma 4.76. *There exists an isomorphism*

$$R \xrightarrow{\sim} k[SL(2)]^T = k[xy, xt, zy, zt]/(xt - yz - 1)$$

induced by mapping $u \mapsto x/z$, $v \mapsto y/t$, $1/(u-v) \mapsto zt$. □

The proof of this is easy and is left to the reader.

Second proof of Theorem 4.43 for $n = 2$. We have to construct an invariant linear form $f \in \mathrm{Hom}_{SL(2)}(V, k)$ such that $f(w) \neq 0$, and we begin with $\phi : V \to k[SL(2)]$ constructed above. By Lemmas 4.75 and 4.76, the image of $\phi : V \to k[SL(2)]^T \cong R$ is contained in the finite-dimensional vector space R_n for some n. By definition, a general element is of the form

$$\frac{f(u,v)}{(u-v)^n} \qquad \text{where } f(u,v) = \sum_{0 \leq i,j \leq n} a_{ij} u^i v^j.$$

Taking the determinant of the coefficient matrix $(a_{ij})_{0 \leq i,j \leq n}$ defines a function $\det : R_n \to k$ which is homogeneous of degree $n+1$ and is $SL(2)$-invariant. Note that at the constant function

$$1 = \frac{(u-v)^n}{(u-v)^n} = \frac{u^n - nu^{n-1}v + \binom{n}{2}u^{n-2}v^2 - \cdots + (-v)^n}{(u-v)^n} \in R_n \subset k[SL(2)]^T$$

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

det takes the value

$$\det \begin{vmatrix} & & & & 1 \\ & & & -n & \\ & & \binom{n}{2} & & \\ & \cdots & & & \\ (-1)^{n-1}n & & & & \\ (-1)^n & & & & \end{vmatrix} = \prod_{i=0}^{n} \binom{n}{i}.$$

Let $h : R_n \to k$ be the formal differential of det at $1 \in R_n$, that is, the linear coefficient in the expansion of

$$\det \left(1 + \epsilon \frac{f(u, v)}{(u - v)^n} \right).$$

Then h is an $SL(2)$-invariant linear map given by

$$\frac{f(u, v)}{(u - v)^n} \mapsto \sum_{i=0}^{n} (-1)^{n-i} a_{i,n-i} \prod_{j \neq i} \binom{n}{j}.$$

In particular,

$$h(1) = (-1)^n (n + 1) \prod_{j=0}^{n} \binom{n}{j} \neq 0. \tag{4.8}$$

Hence (using Lemma 4.75(i)) the composition

$$f = h \circ \phi : V \to R_n \to k$$

is an invariant linear form $f \in \mathrm{Hom}_{SL(2)}(V, k)$ with the required property. □

Notice that in this proof the only place where we have used the assumption that the field k has characteristic zero is in the final step (4.8). In positive characteristic, even if we cannot use the differential h, we can nevertheless move the goalposts and use the function det to modify the condition of Lemma 4.74(iii):

Definition 4.77. An algebraic group G is *geometrically reductive* if, given a finite-dimensional representation V of G and an invariant vector $w \in V^G$, there exists a G-invariant homogeneous polynomial function $f : V \to k$ satisfying $f(w) = 1$. □

Theorem 4.78 (Seshadri [76]). *If char $k = p > 0$, then $SL(2)$ is geometrically reductive.*

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

Proof. In the proof above we may take $n = p^\nu - 1$ for sufficiently large ν. Then

$$(u-v)^n = \frac{(u-v)^{p^\nu}}{u-v} = \frac{u^{p^\nu} - v^{p^\nu}}{u-v} = \sum_{i=0}^{n} u^{n-i} v^i .$$

In particular, the form $\det : R_n \to k$ takes the value 1 at $1 \in R_n$. So in this case the composition

$$f = \det \circ \phi : V \to R_n \to k$$

has exactly the properties asserted in the theorem. □

Remarks 4.79.

(i) Obviously linear reductivity implies geometric reductivity, and over a field k of characteristic zero the converse is also true and the two conditions are equivalent. In positive characteristic, however, geometric reductivity is a strictly weaker condition, and in fact the only connected linearly reductive groups are tori. In particular, $SL(n)$ for $n \geq 2$ is geometrically reductive but not linearly reductive.

(ii) It is actually the property of geometric reductivity that the construction of quotient varieties depends upon. It turns out that both the finite-generatedness of the invariant ring and the separation of orbits by the invariants follow from the geometric reductivity of the group.

Exercises

1. If D and D' are derivations of a ring R, show that their commutator $[D, D'] = DD' - D'D$ is also a derivation.
2. Let $\alpha \in \mathcal{H}(G)$ be a distribution supported at the identity of an algebraic group G, let $R = k[G]$, and denote by $D_\alpha \in \operatorname{End} R$ the k-linear endomorphism

$$R \xrightarrow{\mu} R \otimes_k R \xrightarrow{\alpha \otimes 1} k \otimes_k R \xrightarrow{\sim} R.$$

Show that the following two conditions are equivalent:

(i) $\alpha : R \to k$ is a $k = R/\mathfrak{m}$-valued derivation;

(ii) $D_\alpha : R \to R$ is a derivation of R.

3. Show that the set $\mathrm{Lie}(G)$ of derivations at the identity is closed in distribution algebra $\mathcal{H}(G)$ under the commutator $[\alpha, \beta] = \alpha \star \beta - \beta \star \alpha$. (The Lie space $\mathrm{Lie}(G)$ equipped with this commutator product is called the *Lie algebra* of the algebraic group G.)

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press

4. Let $\chi \in k[G]$ be a character of an algebraic group G, and let V be a linear representation of G. Show that

$$V_\chi := \{v \in V \mid \mu_V(v) = v \otimes \chi\}$$

is a subrepresentation.

5. Let $\mu_V : V \to V \otimes k[G]$ and $\mu_W : W \to W \otimes k[G]$ be two representations of an algebraic group G.
   (i) Show that the tensor product $U \otimes V$ is a representation of G via the composition

$$\begin{aligned} U \otimes V &\xrightarrow{\mu_U \otimes \mu_V} U \otimes k[G] \otimes V \otimes k[G] \xrightarrow{\sim} U \otimes V \otimes k[G] \otimes k[G] \\ &\xrightarrow{1 \otimes m} U \otimes V \otimes k[G], \end{aligned}$$

   where $m : k[G] \otimes k[G] \to k[G]$ is multiplication in the ring.
   (ii) Show that the space $\mathrm{Hom}_k(U, V)$ of k-linear maps from U to V is also a representation of G.
   (iii) Show that $f \in \mathrm{Hom}_k(U, V)$ is G-invariant if and only if f is a G-module homomorphism.

6. Suppose that G is linearly reductive, and let W be a subrepresentation of a finite-dimensional representation V of G. Prove that V decomposes as a direct sum of representations $V = W \oplus W'$. *Hint:* Apply linear reductivity to the surjective map of representations $\mathrm{Hom}_k(V, W) \to \mathrm{Hom}_k(W, W)$.

7. Let

$$0 \to U \to V \to W \to 0$$

be an exact sequence of representations of G. Show that the induced sequence of spaces of invariants

$$0 \to U^G \to V^G \to W^G$$

is exact. (That is, the functor which takes invariants is left-exact.)

8. Let $\mu : V \to V \otimes k[G]$ be a representation of an algebraic group G. Let $g \in G(k)$ be a k-valued point and $\mathfrak{m}_g \subset k[G]$ the corresponding maximal ideal. Denote by $\rho(g) \in \mathrm{End}\, V$ the composition

$$V \xrightarrow{\mu} V \otimes k[G] \xrightarrow{\text{mod } \mathfrak{m}_g} V \otimes k \xrightarrow{\sim} V.$$

Show that, if the coordinate ring $k[G]$ is an integral domain, then a vector $v \in V$ such that $\rho(g)(v) = v$ for every $g \in G(k)$ is a G-invariant.

https://doi.org/10.1017/CBO9781316257074.006 Published online by Cambridge University Press