Appendix A
Linear Algebra

A.1 The Cartan Decomposition

We briefly recall the so-called polar or Cartan decomposition in $\mathrm{GL}_d(\mathbb{R})$: Every element $g \in \mathrm{GL}_d(\mathbb{R})$ can be written in the form

$$g = kak'$$

with $k, k' \in \mathrm{O}_d(\mathbb{R})$ and a positive diagonal matrix a whose eigenvalues are arranged in decreasing size. Moreover, if $g \in \mathrm{SL}_d(\mathbb{R})$ then it is possible to achieve $k, k' \in \mathrm{SO}_d(\mathbb{R})$.

To obtain this decomposition, we only have to apply basic linear algebra. Indeed, notice that gg^{t} is positive and symmetric. Hence gg^{t} has an orthonormal basis of eigenvectors. Ordering them by size of the respective eigenvalues, and switching the sign of the last eigenvector if necessary, these eigenvectors together form the columns of a matrix $k \in \mathrm{SO}_d(\mathbb{R})$ with the property that $D = k^{\mathrm{t}} gg^{\mathrm{t}} k$ is diagonal, with positive eigenvalues arranged in decreasing size. We define the diagonal matrix a to be the positive square root of $D = a^2$, and define k' to be $a^{-1}k^{\mathrm{t}}g$. Then $g = kak'$ and

$$k'(k')^{\mathrm{t}} = a^{-1} \underbrace{k^{\mathrm{t}} gg^{\mathrm{t}} k}_{=D} a^{-1} = a^{-1} D a^{-1} = I$$

shows that $k' \in \mathrm{O}_d(\mathbb{R})$ as claimed.

If $g \in \mathrm{SL}_d(\mathbb{R})$ then $\det k = 1$ and $\det a > 0$ forces $\det k' = 1$, and so we also have $k' \in \mathrm{SO}_d(\mathbb{R})$ as claimed.

A.2 The Amitsur–Levitzki Identity for Matrix Algebras

The *standard polynomial* s_d in d non-commuting variables $T_1, \dots, T_d$ is defined by

$$s_d(T_1, \dots, T_d) = \sum_{\sigma \in \mathrm{S}_d} \operatorname{sign}(\sigma) T_{\sigma(1)} T_{\sigma(2)} \cdots T_{\sigma(d)}. \tag{A.1}$$

With this notation, we can already state the main result of this appendix, a result of Amitsur and Levitzki [1].

Theorem A.1 (Amitsur–Levitzki). *For $n \in \mathbb{N}$ and matrices $T_1, \dots, T_{2n}$ in $\mathrm{Mat}_{n,n}(\mathbb{C})$ we have the* standard polynomial identity

$$s_{2n}(T_1, \dots, T_{2n}) = 0. \tag{A.2}$$

Moreover, there are elements $T_1, \dots, T_{2n} \in \mathrm{Mat}_{n+1,n+1}(\mathbb{C})$ that fail this identity. In fact, if $E_{i,j}$ denotes the elementary matrix with a single 1 in row i and column j and 0 elsewhere and we choose

$$T_1{=}E_{1,1},\ T_2{=}E_{1,2},\ T_3{=}E_{2,2},\ T_4{=}E_{2,3}, \dots,\ T_{2n}{=}E_{n,n+1},$$

then

$$s_{2n}(T_1, \dots, T_{2n}) = E_{1,n+1}. \tag{A.3}$$

The second part of the theorem is quite easy to prove. In fact we have $E_{i,j}E_{k,\ell} = \delta_{j,k}E_{i,\ell}$ for all $i, j, k, \ell \in \{1, \dots, n+1\}$ and so

$$T_1 \cdots T_{2n} = E_{1,1}E_{1,2}E_{2,2}E_{2,3} \cdots E_{n,n}E_{n,n+1} = E_{1,n+1}.$$

However, if the matrices are permuted in any non-trivial fashion, then somewhere in the product a term of the form $E_{i,j}E_{k,\ell}$ will appear for which $j \neq k$ and $i, j, k, \ell \in \{1, \dots, n+1\}$, so that the product vanishes. Summing over all $\sigma \in \mathrm{S}_{2n}$ as in (A.1) is then quite straightforward, and gives (A.3).

The proof of (A.2) is not easy, and we will follow the argument presented in a short paper of Rosset [57]. For the convenience of the reader, and as a warm-up for what follows, we first review some classical material needed for the proof. We start by recalling the Newton–Girard formulæ for symmetric polynomials. Let $x_1, \dots, x_n$ be commuting variables, and define the *power sums* p_k for $k \geqslant 1$ by

$$p_k = p_k(x_1, \dots, x_n) = \sum_{i=1}^{n} x_i^k,$$

and the *elementary symmetric polynomials* by

$$\begin{aligned}
e_0 &= e_0(x_1,\dots,x_n) = 1,\\
e_1 &= e_1(x_1,\dots,x_n) = \sum_{i=1}^n x_i,\\
e_2 &= e_2(x_1,\dots,x_n) = \sum_{i_1<i_2} x_{i_1}x_{i_2},\\
&\vdots\\
e_k &= e_k(x_1,\dots,x_n) = \sum_{i_1<\dots<i_k} x_{i_1}\cdots x_{i_k}
\end{aligned}$$

for $1 \leqslant k \leqslant n$,

$$\begin{aligned}
&\vdots\\
e_n &= e_n(x_1,\dots,x_n) = x_1x_2\cdots x_n,
\end{aligned}$$

and

$$e_k = e_k(x_1,\dots,x_n) = 0$$

for $k > n$. Then the Newton–Girard identity is given by

$$ke_k = \sum_{i=1}^k (-1)^{i+1} e_{k-i} p_i \tag{A.4}$$

for all $k \geqslant 1$ and any number $n \geqslant 1$ of variables.

The reason why the elementary symmetric polynomials are of interest (both in general and for the application at hand) is the simple relation

$$\prod_{i=1}^n (t - x_i) = \sum_{k=0}^n (-1)^k e_k(x_1,\dots,x_n) t^{n-k}.$$

To prove (A.4) divide the above by t^n and replace t by t^{-1} to obtain

$$\sum_{k=0}^n (-1)^k e_k t^k = \prod_{i=1}^n (1 - x_i t),$$

differentiate with respect to t and multiply by t to get

$$\sum_{k=0}^n (-1)^k k e_k t^k = t \sum_{i=1}^n (-x_i) \prod_{j\neq i} (1 - x_j t).$$

We now work in the ring $\mathbb{Z}[x_1,\dots,x_n][\![t]\!]$ of formal power series in t and use the relation

$$(1 - x_i t)^{-1} = \sum_{\ell=0}^{\infty} (x_i t)^\ell$$

to obtain

$$\begin{aligned}
\sum_{k=0}^{n} (-1)^k k e_k t^k &= -\left(\sum_{i=1}^{n} \frac{x_i t}{1 - x_i t} \right) \prod_{j=1}^{n} (1 - x_j t) \\
&= -\left(\sum_{i=1}^{n} \sum_{m=1}^{\infty} x_i^m t^m \right) \left(\sum_{\ell=0}^{n} (-1)^\ell e_\ell t^\ell \right) \\
&= -\left(\sum_{m=1}^{\infty} p_m t^m \right) \left(\sum_{\ell=0}^{n} (-1)^\ell e_\ell t^\ell \right) \\
&= \sum_{k=0}^{\infty} \sum_{m=1}^{k} p_m (-1)^{k-m+1} e_{k-m} t^k.
\end{aligned}$$

This gives (A.4) for all $k \geqslant 1$ and over any commutative ring with a unit (see Exercise A.2).

Next we recall that for a matrix $A = (a_{i,j})_{i,j} \in \mathrm{Mat}_{n,n}(\mathbb{C})$, the characteristic polynomial is defined by

$$\chi_A(t) = \det(tI_n - A),$$

and the *Cayley–Hamilton theorem* states that

$$\chi_A(A) = 0. \tag{A.5}$$

One way to prove (A.5) is by perturbation. First recall that any matrix A in $\mathrm{Mat}_{n,n}(\mathbb{C})$ can be conjugated to become triangular, and that the eigenvalues of A then appear precisely along the diagonal. Next notice that the set $\mathcal{D}$ of matrices with distinct eigenvalues is open and dense in the space of all matrices. In fact, after triangularisation it is easy to perturb the eigenvalues along the diagonal to ensure that they are distinct. Now notice that for any matrix $A \in \mathcal{D}$ with distinct eigenvalues the claim holds, since in fact any such A can be diagonalised and (A.5) holds for diagonal matrices. Finally, (A.5) follows since $\chi_A(A)$ depends continuously on $A \in \mathrm{Mat}_{n,n}(\mathbb{C})$ and vanishes on a dense subset.

As our last preparation we recall that (for example, after triangularisation) if $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ are the eigenvalues of A (counted with multiplicities), then

$$\chi_A(t) = \prod_{j=1}^{n} (t - \lambda_j) = \sum_{k=0}^{n} (-1)^k e_k(\lambda_1, \dots, \lambda_n) t^{n-k}$$

and

$$\operatorname{tr}(A^k) = \sum_{j=1}^{n} \lambda_j^k = p_k(\lambda_1, \dots, \lambda_n)$$

for all $k \geqslant 1$. Combining the Cayley–Hamilton theorem (A.5) and the Newton–Girard formula (A.4) shows that there exists a formula of the form

$$A^n = \sum_{k=1}^{n} q_k(\operatorname{tr}(A), \dots, \operatorname{tr}(A^n)) A^{n-k} \tag{A.6}$$

for some polynomials $q_1, \dots, q_n$ with rational coefficients without constant terms.

We note that the claims above may be expressed as equalities of polynomial expressions with integer coefficients in the entries of the matrices. In particular, (A.4) holds over any commutative ring R with a unit, (A.5) and (A.6) hold for all matrices in R (see Exercise A.2). As we will see, the proof of the Amitsur–Levitzki theorem due to Rosset uses all of the above.

Exercise A.2. Justify the lifting of the identities (A.3)–(A.5) to arbitrary commutative ring R by the following steps.
(a) Show that for any elements $a_1, \dots, a_n \in R$ the evaluation map $\mathbb{Z}[X_1, \dots, X_n] \to R$ defined by sending $f \in \mathbb{Z}[X_1, \dots, X_n]$ to $f(a_1, \dots, a_n)$ is a ring homomorphism.
(b) Show that if $f \in \mathbb{C}[X_1, \dots, X_n]$ vanishes on a non-empty open subset of $\mathbb{C}^n$ then $f = 0$ as elements of $\mathbb{C}[X_1, \dots, X_n]$.
(c) Show that the set of matrices in $\operatorname{Mat}_n(\mathbb{C})$ with distinct eigenvalues contains a non-empty open set.

Proof of Theorem A.1. Given $n \in \mathbb{N}$ we define the *exterior algebra* E in $2n$ dimensions over $\mathbb{C}$ by

$$E = \bigoplus_{k=0}^{2n} \bigwedge^k (\mathbb{C}^{2n}),$$

where we identify $\bigwedge^0(\mathbb{C}^{2n})$ with $\mathbb{C}$ and note that

$$R = \bigoplus_{k=0}^{n} \bigwedge^{2k} (\mathbb{C}^{2n})$$

is a central sub-algebra. That is, if $t \in E$ and $r_1, r_2 \in R$ then $r_1 \wedge r_2 \in R$ and $t \wedge r_1 = r_1 \wedge t$. We also define the algebra $M_E = \operatorname{Mat}_{n,n}(\mathbb{C}) \otimes_{\mathbb{C}} E$, its sub-algebra $M_R = \operatorname{Mat}_{n,n}(\mathbb{C}) \otimes_{\mathbb{C}} R \subseteq M_E$, and note that

$$M_R \cong \operatorname{Mat}_{n,n}(R) \tag{A.7}$$

by identifying $(aE_{i,j}) \otimes r$ with $(ar)E_{i,j}$ for $a \in \mathbb{C}$, $i, j \in \{1, \dots, n\}$ and $r \in R$. Let $T_1, \dots, T_{2n} \in \operatorname{Mat}_{n,n}(\mathbb{C})$ and define

$$A = \sum_{j=1}^{2n} T_j \otimes e_j \in M_E,$$

which satisfies

$$A^2 = \sum_{j_1,j_2=1}^{2n} T_{j_1}T_{j_2} \otimes (e_{j_1} \wedge e_{j_2}) = \sum_{j_1<j_2} s_2(T_{j_1}, T_{j_2}) \otimes (e_{j_1} \wedge e_{j_2})$$

and, more generally,

$$A^k = \sum_{j_1<j_2<\cdots<j_k} s_k(T_{j_1}, \ldots, T_{j_k}) \otimes (e_{j_1} \wedge \cdots \wedge e_{j_k})$$

for $k = 2, \ldots, 2n$. In particular,

$$A^{2n} = s_{2n}(T_1, \ldots, T_{2n}) \otimes (e_1 \wedge \cdots \wedge e_{2n}), \tag{A.8}$$

and the theorem is precisely the statement that $A^{2n} = 0$.

We will now focus on $A^2 \in M_R$ and its powers. For this, it is useful to use the isomorphism in (A.7) and consider A^2 as a matrix with entries in $\mathrm{Mat}_{n,n}(R)$. In particular, the trace of A^2 over R is given by

$$\mathrm{tr}_R(A^2) = \sum_{j_1<j_2} \mathrm{tr}_{\mathbb{C}}\big(s_2(T_{j_1}, T_{j_2})\big) e_{j_1} \wedge e_{j_2} = 0,$$

since the trace over $\mathbb{C}$ of

$$s_2(T_{j_1}, T_{j_2}) = [T_{j_1}, T_{j_2}] = T_{j_1}T_{j_2} - T_{j_2}T_{j_1} \in \mathrm{Mat}_{n,n}(\mathbb{C})$$

is zero for any $j_1, j_2 \in \{1, \ldots, 2n\}$. We also claim that

$$\mathrm{tr}_R(A^{2k}) = 0. \tag{A.9}$$

Indeed, if $T_1, \ldots, T_{2k} \in \mathrm{Mat}_{n,n}(\mathbb{C})$ and $\sigma \in \mathrm{S}_{2k}$ is any permutation, then the products

$$T_{\sigma(1)}T_{\sigma(2)}T_{\sigma(3)} \cdots T_{\sigma(2k-1)}T_{\sigma(2k)}$$

and

$$T_{\sigma(2k)}T_{\sigma(1)}T_{\sigma(2)} \cdots T_{\sigma(2k-2)}T_{\sigma(2k-1)}$$

appear with different signs in the definition of $s_{2k}(T_1, \ldots, T_{2k})$ and have the same trace. Pairing the terms in the sum up in this way, we see that

$$\mathrm{tr}_{\mathbb{C}}(s_{2k}(T_1, \ldots, T_{2k})) = 0$$

for any $T_1, \ldots, T_{2k} \in \mathrm{Mat}_{n,n}(\mathbb{C})$. Using this for $T_{j_1}, \ldots, T_{j_{2k}}$ for any finite list $j_1 < \cdots < j_{2k}$ in $\{1, \ldots, 2n\}$, we obtain (A.9).

Applying (A.6) to $A^2 \in \mathrm{Mat}_{n,n}(R)$ and using (A.9), we see that

$$(A^2)^n = A^{2n} = 0,$$

which gives the theorem by (A.8). □