Appendix A

A.1 Universal enveloping algebra

In this part, we will discuss the link between Lie algebras and associative algebras.

Definition A.1.1. *An* **associative algebra** *over $\mathbb{R}$ is a $\mathbb{R}$-vector space A equipped with a bilinear multiplication $(X, Y) \to XY$ that is associative, i.e. $(XY)Z = X(YZ)$. If A contains a multiplicative identity element 1, it is said to be* **unital**.

An **ideal** *of A is a vector subspace I such that $XY \in I$, $YX \in I$ for all $X \in I$, $Y \in A$.*

Example A.1.2. For a vector space V, the infinite dimensional vector space $\bigoplus_{n=0}^{\infty} V^{\otimes n}$ is an associative algebra, with its multiplication being the natural extension of the tensor product. Here $V^{\otimes 0} = \mathbb{R}$.

It is easy to verify that when I is an ideal of A, the quotient A/I inherits a natural assoicative algebra structure from A.

For another example, the space $M_{d\times d}(\mathbb{R})$ of $d \times d$ real matrices is an associative algebra. Just as $M_{d\times d}(\mathbb{R})$ can be realized as the lie algebra $\mathfrak{gl}(d, \mathbb{R})$, every associative algebra can be equpped with a Lie algebra structure by letting $[X, Y] = XY - YX$.

On the other hand, all Lie algebras can be embedded into an associative algebra though such an embedding is in general not unique. However there is a largest one among such embeddings, called the universal enveloping algebra.

Definition A.1.3. *The universal enveloping algebra $U_{\mathfrak{g}}$ of the Lie algebra is the quotient of the associative algebra $\bigoplus_{n=0}^{\infty} \mathfrak{g}^{\otimes n}$ by the ideal I generated by $\{X \otimes Y - Y \otimes X - [X, Y] : X, Y \in \mathfrak{g}\}$.*

We say a linear transform ψ from a Lie algebra $\mathfrak{g}$ to an associative algebra A is a **morphism** if $\psi X \cdot \psi Y - \psi Y \cdot \psi X = \psi[X, Y]$. An **embedding** is an injective morphism. It is easy to see that the natural map $\iota(X) = X$ from

$\mathfrak{g}$ to $U_{\mathfrak{g}}$ is a morphism. In $U_{\mathfrak{g}}$, the element $X \otimes Y - Y \otimes X$ is identified with $[X, Y]$.

Moreover, $U_{\mathfrak{g}}$ has the following universal property:

Proposition A.1.4. If $\psi : \mathfrak{g} \to A$ is a morphism into a unital associative algebra A, then there is a unique associative algebra morphism $\phi : U_{\mathfrak{g}} \to A$ such that $\phi \circ \iota = \psi$.

By abusing notation, for $Y \in \mathfrak{g}$, write Y for $\iota(Y)$ even though we have not shown yet ι is injective. The proof of the proposition is not hard: for $Y_j \in \mathfrak{g}$, the linear transform ϕ should map $Y_1 Y_2 \cdots Y_n \in U_{\mathfrak{g}}$ to $\psi(Y_1)\psi(Y_2)\cdots\psi(Y_n)$. It remains to show that this map is well defined because ψ is a morphism and after the projection π, $XY - YX$ is identified with $[X, Y]$. For details, see e.g. [Jac79, Ch. V].

Theorem A.1.5 (Poincaré-Birkhoff-Witt Theorem)**.** *Suppose* $\{X_1, \cdots, X_m\}$ *is a basis of* $\mathfrak{g}$*, then the monomials*

$$\{X_1^{r_1} \cdots X_m^{r_m} : r_1, \cdots, r_m \geq 0\} \tag{A.1}$$

form a basis of $U_{\mathfrak{g}}$*.*

Proof. Write $\mathbf{X}^{\mathbf{r}}$ for the element $X_1^{r_1} \cdots X_m^{r_m}$ of $U_{\mathfrak{g}}$, where $\mathbf{r} = (r_1, \cdots, r_m)$. From now on, denote $\sum_{i=1}^m r_i$ by $|\mathbf{r}|$.

Define $U_{\mathfrak{g},n} \subset U_{\mathfrak{g}}$ be the by

$$\text{span}\{Y_1 Y_2 \cdots Y_n : Y_1, \cdots, Y_n \in \mathfrak{g}\}. \tag{A.2}$$

Then $U_{\mathfrak{g}} = \bigcup_{n=0}^{\infty} U_{\mathfrak{g},n}$.

It is not hard to show by induction that:

$$ZW - WZ \in U_{\mathfrak{g},n}, \forall Z \in U_{\mathfrak{g},1}, \forall W \in U_{\mathfrak{g},n},$$

and then by a second layer of induction that:

$$ZW - WZ \in U_{\mathfrak{g},l+n-1}, \forall Z \in U_{\mathfrak{g},l}, \forall W \in U_{\mathfrak{g},n}.$$

Using these facts, it is straight forward to show that $U_{\mathfrak{g},n}$ is spanned by $\{\mathbf{X}^{\mathbf{r}} : \mathbf{r} \in (\mathbb{N} \cup \{0\})^m, |\mathbf{r}| \leq n\}$. This would show that the monomials from (A.1) span $U_{\mathfrak{g}}$. We will focus on proving that these monomials are linearly independent.

Let V be the vector space freely spanned by $\{\mathbf{x}^{\mathbf{r}} : \mathbf{r} \in (\mathbb{N} \cup \{0\})^m\}$ where $\mathbf{x}^{\mathbf{r}}$ stands for symbolic monomial $x_1^{r_1} \cdots x_m^{r_m}$. Below, let $V_n \subset V$ be the span of $\{\mathbf{x}^{\mathbf{r}} : |\mathbf{r}| \leq n\}$.

We first build a Lie algebra representation of $\mathfrak{g}$ on V, i.e. a Lie algebra morphism $\rho : \mathfrak{g} \to \mathrm{End}(V)$. For this purpose, one will inductively define $\rho(X_i)\mathbf{x}^{\mathbf{r}}$ for all i and $\mathbf{r}$.

When $|\mathbf{r}| = 0$, then $\mathbf{r} = \mathbf{0}$, let $\rho(X_i)\mathbf{x}^{\mathbf{0}} = \mathbf{x}^{\mathbf{e}_i} = x_i \in V_1$.

Suppose $|\mathbf{r}| \geq 1$, $1 \leq i \leq m$ and that $\rho(X_{i'})\mathbf{x}^{\mathbf{r}'}$ has been defined and belongs to $V_{|\mathbf{r}'|+1}$, for all i' and $\mathbf{r}'$ such that $|\mathbf{r}'| < |\mathbf{r}|$, as well as for all $j < i$ when $\mathbf{r}'$ satisfies $|\mathbf{r}'| = |\mathbf{r}|$.

Write $\mathbf{x}^{\mathbf{r}}$ as $x_k\mathbf{x}^{\mathbf{r}'}$ where $k = k(\mathbf{r})$ is the smallest k such that $r_k > 0$. If $i \leq k$, then simply set

$$\rho(X_i)\mathbf{x}^{\mathbf{r}} = x_i\mathbf{x}^{\mathbf{r}} = x_ix_k\mathbf{x}^{\mathbf{r}'}; \tag{A.3}$$

otherwise, define

$$\rho(X_i)\mathbf{x}^{\mathbf{r}} = \rho(X_k)\rho(X_i)\mathbf{x}^{\mathbf{r}'} + \rho([X_i, X_k])\mathbf{x}^{\mathbf{r}'}. \tag{A.4}$$

Note that by inductive hypothesis, $\rho(X_i)\mathbf{x}^{\mathbf{r}'}$ and $\rho([X_i, X_k])\mathbf{x}^{\mathbf{r}'}$ are already defined inside $V_{|\mathbf{r}|}$. Thus as $k < i$, again by inductive hypothesis, $\rho(X_k)\rho(X_i)\mathbf{x}^{\mathbf{r}'}$ is defined and belongs to $V_{|\mathbf{r}|+1}$.

By iteration, this defines a linear map $\rho : \mathfrak{g} \to \mathrm{End}(V)$. One needs to show that this is a Lie algebra morphism, i.e. for all i, j and $\mathbf{r}$,

$$\rho(X_i)\rho(X_j)\mathbf{x}^{\mathbf{r}} - \rho(X_j)\rho(X_i)\mathbf{x}^{\mathbf{r}} = \rho([X_i, X_j])\mathbf{x}^{\mathbf{r}}. \tag{A.5}$$

Let $k = k(\mathbf{r})$, where we set $k(\mathbf{r}) = m$ when $|\mathbf{r}| = 0$. After multiplying by -1 if necessary, we may assume $i \leq j$ in (A.5). When $i = j$, (A.5) holds trivially. So we may assume $i < j$.

If $k \geq i$ (note this includes the $|\mathbf{r}| = 0$ case), then $\rho(X_i)\mathbf{x}^{\mathbf{r}} = x_i\mathbf{x}^{\mathbf{r}}$ and (A.5) directly follows from (A.4). Hence when proving (A.5), we may assume $k < i < j$.

Suppose for induction that (A.5) is known for all configurations i', j',$\mathbf{r}'$ satisfying: (1) either $\mathbf{r}'$ with $|\mathbf{r}'| < |\mathbf{r}|$, (2) or $|\mathbf{r}'| = |\mathbf{r}|$ and $\min(i', j') < \min(i, j)$. The starting step is the $|\mathbf{r}| = 0$ case settled above.

Like before, $\mathbf{x}^{\mathbf{r}} = x_k\mathbf{x}^{\mathbf{r}'}$ where $|\mathbf{r}'| = |\mathbf{r}| - 1$. We can also write $\mathbf{x}^{\mathbf{r}} = \rho(X_k)\mathbf{x}^{\mathbf{r}}$ in this case.

Then by induction,

$$\begin{aligned}
&\rho(X_i)\rho(X_j)\mathbf{x}^{r} \\
=&\rho(X_i)\rho(X_j)\rho(X_k)\mathbf{x}^{\mathbf{r}'} \\
=&\rho(X_i)\rho(X_k)\rho(X_j)\mathbf{x}^{\mathbf{r}'} + \rho(X_i)\rho([X_j, X_k])\mathbf{x}^{\mathbf{r}'} \\
=&\rho(X_k)\rho(X_i)\rho(X_j)\mathbf{x}^{\mathbf{r}'} + \rho([X_i, X_k])\rho(X_j)\mathbf{x}^{\mathbf{r}'} + \rho(X_i)\rho([X_j, X_k])\mathbf{x}^{\mathbf{r}'}.
\end{aligned}$$

Here the induction hypothesis applies in the third line because $|\mathbf{r}'| < |\mathbf{r}|$, and in the last line because $\rho(X_j)\mathbf{x}^{\mathbf{r}'} \in V_{|\mathbf{r}|}$ and $\min(i,k) = k < i = \min(i,j)$. Similarly,

$$\begin{aligned}
&\rho(X_j)\rho(X_i)\mathbf{x}^r \\
=&\rho(X_j)\rho(X_i)\rho(X_k)\mathbf{x}^{\mathbf{r}'} \\
=&\rho(X_j)\rho(X_k)\rho(X_i)\mathbf{x}^{\mathbf{r}'} + \rho(X_j)\rho([X_i,X_k])\mathbf{x}^{\mathbf{r}'} \\
=&\rho(X_k)\rho(X_j)\rho(X_i)\mathbf{x}^{\mathbf{r}'} + \rho([X_j,X_k])\rho(X_i)\mathbf{x}^{\mathbf{r}'} + \rho(X_j)\rho([X_i,X_k])\mathbf{x}^{\mathbf{r}'}.
\end{aligned}$$

So again by repeatedly applying the inductive hypothesis,

$$\begin{aligned}
&\rho(X_i)\rho(X_j)\mathbf{x}^{\mathbf{r}} - \rho(X_j)\rho(X_i)\mathbf{x}^{\mathbf{r}} \\
=&\rho(X_k)\Big(\rho(X_i)\rho(X_j) - \rho(X_j)\rho(X_i)\Big)\mathbf{x}^{\mathbf{r}'} + \rho([X_i,X_k])\rho(X_j)\mathbf{x}^{\mathbf{r}'} \\
&\quad + \rho(X_i)\rho([X_j,X_k])\mathbf{x}^{\mathbf{r}'} - \rho([X_j,X_k])\rho(X_i)\mathbf{x}^{\mathbf{r}'} - \rho(X_j)\rho([X_i,X_k])\mathbf{x}^{\mathbf{r}'} \\
=&\rho(X_k)\rho([X_i,X_j])\mathbf{x}^{\mathbf{r}'} + \rho([X_j,[X_k,X_i]])\mathbf{x}^{\mathbf{r}'} + \rho([X_i,[X_j,X_k]])\mathbf{x}^{\mathbf{r}'} \\
=&\rho(X_k)\rho([X_i,X_j])\mathbf{x}^{\mathbf{r}'} - \rho([X_k,[X_i,X_j]])\mathbf{x}^{\mathbf{r}'} \\
=&\rho([X_i,X_j])\rho(X_k)\mathbf{x}^{\mathbf{r}'} \\
=&\rho([X_i,X_j])\mathbf{x}^{\mathbf{r}},
\end{aligned}$$

where the third equality is an application of the Jacobi identity. This shows (A.5) for i, j, $\mathbf{r}$. Therefore we conclude that ρ is a Lie algebra representation. By the universal property of U_g, ρ extends to a morphism $U_{\mathfrak{g}} \to \mathrm{End}(V)$ of associative algebras.

Now suppose $\sum_{\mathbf{r}\in R} c_{\mathbf{r}}\mathbf{X}^{\mathbf{r}} = 0$ for a finite set R of multi-indices and coefficients $c_{\mathbf{r}} \in \mathbb{R}$. Then $\rho(\sum_{\mathbf{r}\in R} c_{\mathbf{r}}\mathbf{X}^{\mathbf{r}})\mathbf{x}^{\mathbf{0}} = \sum_{\mathbf{r}\in R} c_{\mathbf{r}}\rho(\mathbf{X}^{\mathbf{r}})\mathbf{x}^{\mathbf{0}} = \sum_{\mathbf{r}\in R} c_{\mathbf{r}}\mathbf{x}^{\mathbf{r}} = 0$. Because $\{\mathbf{x}^{\mathbf{r}}\}$ is a basis of V, all the $c_{\mathbf{r}}$'s vanish. This shows that the $\mathbf{X}^{\mathbf{r}}$'s are linearly independent. □

Remark that the proof in fact shows that $a \to \rho(a)\mathbf{x}^{\mathbf{0}}$ is a isomorphism of vector spaces between $U_{\mathfrak{g}}$ and V.

An immediate consequence to Theorem A.1.5 is:

Corollary A.1.6. *The morphism $\iota : \mathfrak{g} \to U_{\mathfrak{g}}$ is an embedding.*

A.2 Nilpotent case of Ado's Theorem

We are now going to prove Ado's Theorem, or in fact the following strengthening of it, for the special case of nilpotent Lie algebras.

Theorem A.2.1. *If $\mathfrak{g}$ is a finite dimensional nilpotent Lie algebra, then there exists an injective Lie algebra morphism $\psi : \mathfrak{g} \to \mathfrak{gl}(F)$ for some finite-dimensional vector space F, and $M \in \mathbb{N}$ such that $\psi(Y_1) \cdots \psi(Y_M) = 0$ for all $Y_1, \cdots, Y_M \in \mathfrak{g}$.*

Proof. The proof is by induction in the step of nilpotency of $\mathfrak{g}$. Suppose $\mathfrak{g}$ is 1-step nilpotent, i.e. abelian, then $\mathfrak{g} \cong \mathbb{R}^d$. Then the morphism $\psi(\mathbf{t}) = \begin{pmatrix} 1 & \mathbf{t} \\ & \mathrm{Id} \end{pmatrix}$ from $\mathbb{R}^d$ to $\mathfrak{gl}(d+1, \mathbb{R})$ and $n = 2$ verify the conclusion.

Assume below that $s \geq 2$, $\mathfrak{g}$ is s-step nilpotent, and that the theorem is true for all nilpotent Lie algebras of lower steps. Let $\mathfrak{z}$ be the center of $\mathfrak{g}$. Since $\mathfrak{g}$ is not abelian, $\mathfrak{z} \neq \mathfrak{g}$. Thus by Lemma 1.4.12, $\mathfrak{z} + \mathfrak{g}_{(2)} \neq \mathfrak{g}$. We first choose an arbitrary one dimensional line $\mathfrak{l} \subset \mathfrak{g}$ that is not in $\mathfrak{g}_{(2)}$. Then there is an vector subspace $\mathfrak{h} \subset \mathfrak{g}$ of codimension 1 such that $\mathfrak{z} + \mathfrak{g}_{(2)} \subseteq \mathfrak{h}$ and $\mathfrak{h} \cap \mathfrak{g}_{(2)} = \{0\}$. Both $\mathfrak{l}$ and $\mathfrak{h}$ are Lie subalgebras. For $\mathfrak{l}$, this is because the Lie bracket vanishes within any one dimensional subspace. For $\mathfrak{h}$, this is because $[\mathfrak{h}, \mathfrak{h}] \subseteq \mathfrak{g}_{(2)} \subseteq \mathfrak{h}$. Moreover, $\mathfrak{g} = \mathfrak{l} \oplus \mathfrak{h}$ as vector spaces (but not as Lie algebras).

By inductive assumption, there is an injective representation $\phi : \mathfrak{h} \to \mathfrak{gl}(E)$ for some E and such that $\phi(Y_1) \cdots \phi(Y_N) = 0$ for some N and all $Y_1, \cdots, Y_N \in \mathfrak{h}$. By universality of $U_\mathfrak{h}$, ϕ extends to an associative algebra morphism $U_\mathfrak{h} \to \mathfrak{gl}(E)$. Let J be the ideal generated by $\{Y_1 \cdots Y_N : Y_1, \cdots Y_N \in \mathfrak{h}\}$. Then ϕ vanishes on J, and thus descends to a morphism $U_\mathfrak{h}/J \to \mathfrak{gl}(E)$. Because $\phi : \mathfrak{h} \to U_\mathfrak{h} \to U_\mathfrak{h}/J \to \mathfrak{gl}(E)$ is injective, the composition $\mathfrak{h} \to U_\mathfrak{h}/J$ is injective. Furthermore, $U_\mathfrak{h}/J$ is spanned by $\{Y_1 \cdots Y_k : Y_1, \cdots, Y_n \in \mathfrak{h}, n < N\}$, and therefore finite dimensional.

Remark that as $U_\mathfrak{h}$ is unital, so is the quotient $U_\mathfrak{h}/J$. The assignment to each $Z \in U_\mathfrak{h}/J$ the left multiplication $L_Z : W \to ZW$ is an injective linear transform from $U_\mathfrak{h}/J$ to $\mathfrak{gl}(U_\mathfrak{h}/J)$, since $L_Z(1) = Z$. So we conclude that the composition

$$\phi' : \mathfrak{h} \to U_\mathfrak{h} \to U_\mathfrak{h}/J \to \mathfrak{gl}(U_\mathfrak{h}/J)$$

is injective.

Furthermore, we can extend the map ϕ' to $\mathfrak{l}$. For this, define first a morphism $\widetilde{\mathrm{ad}} : \mathfrak{l} \to \mathfrak{gl}(U_\mathfrak{h})$ by

$$\widetilde{\mathrm{ad}}_X(Y_1 Y_2 \cdots Y_n) = \sum_{i=1}^{n} Y_1 \cdots Y_{i-1} (\mathrm{ad}_X Y_i) Y_{i+1} \cdots Y_n$$

for all $X \in \mathfrak{l}$ and $Y_1, \cdots, Y_n \in \mathfrak{h}$. This is well-defined on $\bigoplus_{n=1}^{\infty} \mathfrak{h}^{\otimes n}$ because $\mathrm{ad}_X Y_i \in \mathfrak{g}_{(2)} \subseteq \mathfrak{h}$, and descends to $U_\mathfrak{h}$ because ad_X is a Lie algebra mor-

phism. By definition, $\widetilde{\mathrm{ad}}_X$ preserves J. Hence $\widetilde{\mathrm{ad}}$ descends to a Lie algebra morphism $\mathfrak{l} \to \mathfrak{gl}(U_\mathfrak{h}/J)$, which we adopt as the definition of ϕ' on $\mathfrak{l}$.

Extending ϕ' linearly from $\mathfrak{h}$ and $\mathfrak{l}$ to $\mathfrak{g} = \mathfrak{l} \oplus \mathfrak{h}$, we obtain a linear transform $\phi' : \mathfrak{g} \to \mathfrak{gl}(U_\mathfrak{h}/J)$. Indeed, ϕ' is also a Lie algebra morphism. Since this is already true within each of $\mathfrak{h}$ and $\mathfrak{l}$, it suffices to show that

$$\phi'([X,Y]) = [\phi'(X), \phi'(Y)], \forall X \in \mathfrak{l}, Y \in \mathfrak{h}. \tag{A.6}$$

Given $Y_1, \cdots, Y_n \in \mathfrak{h}$,

$$\phi'([X,Y])(Y_1 \cdots Y_n) = L_{[X,Y]}(Y_1 \cdots Y_n) = (\mathrm{ad}_X Y) \cdot Y_1 \cdots Y_n,$$

and

$$\begin{aligned}
&[\phi'(X), \phi'(Y)](Y_1 \cdots Y_n) \\
=&\big(\phi'(X)\phi'(Y) - \phi'(Y)\phi'(X)\big)(Y_1 \cdots Y_n) \\
=&\widetilde{\mathrm{ad}}_X(YY_1 \cdots Y_n) - Y\widetilde{\mathrm{ad}}_X(Y_1 \cdots Y_n) \\
=&(\mathrm{ad}_X Y) \cdot Y_1 \cdots Y_n.
\end{aligned}$$

Equation (A.6) follows. So ϕ' is a Lie algebra morphism.

We now show that for sufficiently large M,

$$\phi'(Z_1)\phi'(Z_2)\cdots\phi'(Z_M) = 0, \forall Z_1, \cdots, Z_M \in \mathfrak{g}. \tag{A.7}$$

Without loss of generality, one can assume each Z_i is either in $\mathfrak{l}$ or $\mathfrak{h}$. For an n-fold product $Y_1 \cdots Y_n$ in $U_\mathfrak{h}$, define its order as $\sum_{i=1}^n \min_{Y_i \in \mathfrak{g}_{(j)}} j$. Recall that $\phi'(Z_k)$ sends $Y_1 \cdots Y_n$ either to an $(n+1)$-fold product in $U_\mathfrak{h}/J$ when $Z_i \in \mathfrak{h}$, or to a linear combination of n-fold products when $Z_i \in \mathfrak{l}$. But in the later case, the order of each n-fold component at least increased by 1 compared to that of the input n-fold product. On the one hand, all n-fold products vanish in $U_\mathfrak{h}/J$ unless $n < N$. On the other hand, the order of a non-trivial n-fold product is at most ns. Hence $\phi'(Z_1)\phi'(Z_2)\cdots\phi'(Z_M)Y_1 \cdots Y_n$ vanishes in $U_\mathfrak{h}/J$ if $M > N(N-1)s + N$.

To summarize, we have so far proved, under the inductive hypothesis, that there is a finite dimensional linear representaion $\phi' : \mathfrak{g} \to \mathfrak{gl}(U_\mathfrak{h}/J)$, such that: (1) $\phi'(Z_1)\cdots\phi'(Z_M) = 0$ for $M \geq N^2 s$ and all $Z_i \in \mathfrak{g}$, (2) ϕ' is injective on $\mathfrak{h}$.

Let us also consider the linear representation $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. It satisfies: (1) Because $\mathfrak{g}$ is s-step nilpotent, $\mathrm{ad}_{Z_1} \cdots \mathrm{ad}_{Z_M} = 0$ for $M \geq s+1$ and all $Z_i \in \mathfrak{g}$, (2) $\mathrm{ad}_Z \neq 0$ unless Z is in the center $\mathfrak{z}$ of $\mathfrak{g}$.

Finally, take the finite dimensional representation $\psi = \phi' \oplus \mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}\Big(\big((U_\mathfrak{h})/J\big) \oplus \mathfrak{g}\Big)$, then $\psi(Z_1)\cdots\psi(Z_M) = 0$ for all $Z_i \in \mathfrak{g}$ when $M \geq \max(N^2 s, s+1)$.

In addition, if $\psi(Z) = 0$, then $Z \in \mathfrak{z}$ as $\mathrm{ad}_Z = 0$. Since $\mathfrak{z} \subseteq \mathfrak{h}$ by construction, $Z \in \mathfrak{h}$. So $Z = 0$ as $\phi'(Z) = 0$. This shows ψ is injective. The proof is complete. $\square$