18.745 Introduction to Lie Algebras	November 30, 2010

Lecture 22 — The Universal Enveloping Algebra

Prof. Victor Kac | *Scribe: Aaron Potechin*

Definition 22.1. Let g be a Lie algebra over a field F. An enveloping algebra of g is a pair (φ, U), where U is a unital associative algebra and $\varphi : g \to U_-$ is a Lie algebra homomorphism, where U_- stands for U with the bracket $[a, b] = ab - ba$.

Example 22.1. Let $\varphi : g \to \text{End } V$ be a representation of g in a vector space V. Then the pair $(\text{End } V, \varphi)$ is an enveloping algebra of g.

Definition 22.2. The universal enveloping algebra of g is an enveloping algebra $(\Phi, U(g))$ which has the following universal mapping property: for any enveloping algebra (φ, U) of g there exists a unique associative algebra homomorphism $f : U(g) \to U$ such that $\varphi = f \circ \Phi$.

Exercise 22.1. Prove that the universal enveloping algebra is unique (if it exists).

Solution: Assume $(\Phi_1, U_1(g))$ and $(\Phi_2, U_2(g))$ are both universal enveloping algebras of g. Then by the universal mapping property of the universal enveloping algebra, we have unique maps $f_{11}, f_{12}, f_{21}, f_{22}$ such that for $i, j \in \{1, 2\}$, $\Phi_i = f_{ij} \circ \Phi_j$. Now $\Phi_i = id \circ \Phi_i$, so by uniqueness, $f_{ii} = id$. $\Phi_i = f_{ij} \circ f_{ji} \circ \Phi_i$, so by uniqueness $f_{ij} \circ f_{ji} = f_{ii} = id$. $f_{12} = f_{21}^{-1}$, so $(\Phi_1, U_1(g))$ and $(\Phi_2, U_2(g))$ are isomorphic, as needed. □

Existence of universal enveloping algebras:
Let $T(g)$ be the free unital associative algebra on a basis $a_1, a_2, \cdots$ of g and let $J(g)$ be the two-sided ideal of $T(g)$ generated by the elements $a_i a_j - a_j a_i - [a_i, a_j]$. Then $U(g) = T(g)/J(g)$.
Define $\Phi : g \to U(g)_-$ by letting $\Phi(a_i) =$ the image of a_i in $U(g)$ and extending linearly.

Remark: $T(g)$ is called the tensor algebra over the vector space g.
$T(g) = F \oplus g \oplus (g \otimes g) \oplus (g \otimes g \otimes g) \oplus \cdots$ with the concatenation product.
Good linear algebra textbooks: Artin, Vinberg.

Exercise 22.2. Prove that $(\Phi, U(g))$ is the universal enveloping algebra (i.e the universality property holds).

Solution: First, we will show that $(\Phi, U(g))$ is an enveloping algebra. $T(g)$ is a unital associative algbera, so $U(g) = T(g)/J(g)$ is a unital associative algebra. To check that Φ is a Lie algebra homomorphism, by linearity, it suffices to check that for all i, j, $\Phi([a_i, a_j]) = [\Phi(a_i), \Phi(a_j)]$. But this is clear, as $\Phi([a_i, a_j]) - [\Phi(a_i), \Phi(a_j)] = [a_i, a_j] - a_i a_j - a_j a_i$ is in $J(g)$, so it is zero in $U(g)$.

Now let (U, φ) be another enveloping algebra. Define the map $f : T(g) \to U$ by taking
$f(\prod_{k=1}^{l} a_{i_k}) = \prod_{k=1}^{l} \varphi(a_{i_k})$ and extending linearly. This is clearly a well-defined unital associative algebra homomorphism.
Now for all i, j, $f([a_i, a_j] - a_i a_j - a_j a_i) = \varphi([a_i, a_j]) - \varphi(a_i)\varphi(a_j) - \varphi(a_j)\varphi(a_i) = 0$ because $\varphi : g \to U_-$ is a Lie algebra homomorphism. Thus, $J(g) \subseteq \ker(f)$, so $f : U(g) \to U$ is a well-defined unital associative algebra homomorphism.

For all i, $(f \circ \Phi)(a_i) = f(a_i) = \varphi(a_i)$, so by linearity $\varphi = f \circ \Phi$.

Assume that $f' : U(g) \to U$ is another unital associative algebra homomorphism with $\varphi = f' \circ \Phi$. Then $f'(1) = f(1) = 1$ and for all i, $f(a_i) = \varphi(a_i) = (f' \circ \Phi)(a_i) = f'(a_i)$. $1, a_1, a_2, \cdots$ generate $U(g)$ as a unital associative algebra, so $f = f'$. Thus, f is unique, as needed. $\square$

Corollary 22.1. *Any representation $\pi : g \to End\ V$ extends uniquely to a homomorphism of associative algebras $U(g) \to End\ V$ (so that $a_i \mapsto \pi(a_i)$).*

Theorem 22.2. *Poincaré-Birkhoff-Witt (PBW) theorem: Let $a_1, a_2, \cdots$ be a basis of g. Then the monomials* (*) $a_{i_1} a_{i_2} \cdots a_{i_s}$ with $i_1 \leq i_2 \leq \cdots \leq i_s$ form a basis of $U(g)$.

Proof. Easy part: the monomials(*) span $U(g)$.
Proof is by induction on the pair (s, N), where s is the degree of the monomial and N is the number of inversions, i.e. number of pairs i_m, i_n for which $m < n$ but $i_m > i_n$, lexicographically ordered, i.e. $((s, N) > (s', N')$ if $s > s'$ or $s = s'$ and $N > N')$
For $N = 0$ there is nothing to prove.
If $N \geq 1$, then in the monomial we have $a_{i_t} a_{i_{t+1}}$ where $i_t > i_{t+1}$, as otherwise the monomial is already in our set of monomials(*).
But we have the relation $a_{i_t} a_{i_{t+1}} = a_{i_{t+1}} a_{i_t} + [a_{i_t}, a_{i_{t+1}}]$, so that in $U(g)$:
$a_{i_1} a_{i_2} \cdots a_{i_s} = a_{i_1} \cdots a_{i_{t+1}} a_{i_t} \cdots a_{i_s} + a_{i_1} \cdots a_{i_{t-1}} [a_{i_t}, a_{i_{t+1}}] a_{i_{t+2}} \cdots a_{i_s}$. The first term is a monomial of degree s and $N - 1$ inversions, and the second term is the sum of monomials with degree $s - 1$. Thus, by the inductive hypothesis, each of these monomials is generated by the monomials(*), so $a_{i_1} a_{i_2} \cdots a_{i_s}$ is also generated by these monomials, as needed.

The hard part: why are the monomials(*) linearly independent?
Let B_s be the vector space over F with basis $b_{i_1} \cdots b_{i_s}$ with $i_1 \leq i_2 \leq \cdots i_s$.
Take $B_0 = F$ and let $B = \oplus_{s \geq 0} B_s$. We shall construct a linear map $f : T(g) \to B$ such that $J(g) \subseteq \ker(f)$ and $f(a_{i_1} \cdots a_{i_s}) = (b_{i_1} \cdots b_{i_s})$ if $i_1 \leq i_2 \leq \cdots \leq i_s$.
This will induce a linear map $f : U(g) = T(g)/J(g) \to B$. Hence the monomials(*) are linearly independent because $b_{i_1} \cdots b_{i_s}$, $i_1 \leq i_2 \leq \cdots i_s$ are linearly independent.

Construction: $f(1) = 1$, $f(a_{i_1} \cdots a_{i_s}) = (b_{i_1} \cdots b_{i_s})$ if $i_1 \leq i_2 \leq \cdots \leq i_s$, and
(1) $f(a_{i_1} \cdots a_{i_t} a_{i_{t+1}} \cdots a_{i_s}) = f(a_{i_1} \cdots a_{i_{t+1}} a_{i_t} \cdots a_{i_s}) + f(a_{i_1} \cdots a_{i_{t-1}} [a_{i_t}, a_{i_{t+1}}] a_{i_{t+2}} \cdots a_{i_s})$
if $i_t > i_{t+1}$

By induction on (s, N), we can use the inversion (1) to reduce $f(a_{i_1} \cdots a_{i_s})$ to a sum of terms of the form $f(a_{j_1} \cdots a_{j_{s'}})$, where $j_1 \leq j_2 \leq \cdots j_{s'}$. We just need to check that the final expression is independent of which sequence of inversions we choose. We do this by induction on (s, N).

Case 1:
$a_{i_1} \cdots a_{i_s} = a_{i_1} \cdots a_{i_t} a_{i_{t+1}} \cdots a_{i_r} a_{i_{r+1}} \cdots a_{i_s}$, where $i_t > i_{t+1}$ and $i_r > i_{r+1}$
Using the left inversion first gives us
$f(a_{i_1} \cdots a_{i_{t+1}} a_{i_t} \cdots a_{i_r} a_{i_{r+1}} \cdots a_{i_s}) + f(a_{i_1} \cdots [a_{i_t}, a_{i_{t+1}}] \cdots a_{i_r} a_{i_{r+1}} \cdots a_{i_s})$
By the inductive hypothesis, we may use any sequence of inversions (1) to evaluate this, so using the right inversion on each term, we get
$f(a_{i_1} \cdots a_{i_{t+1}} a_{i_t} \cdots a_{i_{r+1}} a_{i_r} \cdots a_{i_s}) + f(a_{i_1} \cdots a_{i_{t+1}} a_{i_t} \cdots [a_{i_r}, a_{i_{r+1}}] \cdots a_{i_s}) +$
$f(a_{i_1} \cdots [a_{i_t}, a_{i_{t+1}}] \cdots a_{i_{r+1}} a_{i_r} \cdots a_{i_s}) + f(a_{i_1} \cdots [a_{i_t}, a_{i_{t+1}}] \cdots [a_{i_r}, a_{i_{r+1}}] \cdots a_{i_s})$

Using the right inversion first gives us
$f(a_{i_1}\cdots a_{i_t}a_{i_{t+1}}\cdots a_{i_{r+1}}a_{i_r}\cdots a_{i_s})+f(a_{i_1}\cdots a_{i_t}a_{i_{t+1}}\cdots[a_{i_r},a_{i_{r+1}}]\cdots a_{i_s})$.
By the inductive hypothesis, we may use any sequence of inversions (1) to evaluate this, so using the left inversion on each term, we get
$f(a_{i_1}\cdots a_{i_{t+1}}a_{i_t}\cdots a_{i_{r+1}}a_{i_r}\cdots a_{i_s})+f(a_{i_1}\cdots[a_{i_t},a_{i_{t+1}}]\cdots a_{i_{r+1}}a_{i_r}\cdots a_{i_s})+$
$f(a_{i_1}\cdots a_{i_{t+1}}a_{i_t}\cdots[a_{i_r},a_{i_{r+1}}]\cdots a_{i_s})+f(a_{i_1}\cdots[a_{i_t},a_{i_{t+1}}]\cdots[a_{i_r},a_{i_{r+1}}]\cdots a_{i_s})$
These expressions are the same, so we get the same result whether we begin with the left inversion or the right inversion.

Case 2: Inversions overlap
$a_{i_1}\cdots a_{i_s}=a_{i_1}\cdots a_{i_t}a_{i_{t+1}}a_{i_{t+2}}\cdots a_{i_s}$ with $i_t>i_{t+1}>i_{t+2}$.

Exercise 22.3. Show that we get the same result whether we start with the inversion on $a_{i_t}a_{i_{t+1}}$ or the inversion on $a_{i_{t+1}}a_{i_{t+2}}$.

Solution: First using the left inversion gives us
$f(a_{i_1}\cdots a_{i_{t+1}}a_{i_t}a_{i_{t+2}}\cdots a_{i_s})+f(a_{i_1}\cdots[a_{i_t},a_{i_{t+1}}]a_{i_{t+2}}\cdots a_{i_s})$
Using the inversion on $a_{i_t}a_{i_{t+2}}$ in the first term, we get
$f(a_{i_1}\cdots a_{i_{t+1}}a_{i_{t+2}}a_{i_t}\cdots a_{i_s})+f(a_{i_1}\cdots a_{i_{t+1}}[a_{i_t},a_{i_{t+2}}]\cdots a_{i_s})+f(a_{i_1}\cdots[a_{i_t},a_{i_{t+1}}]a_{i_{t+2}}\cdots a_{i_s})$
Using the inversion on $a_{i_{t+1}}a_{i_{t+2}}$ in the first term, we get
$f(a_{i_1}\cdots a_{i_{t+2}}a_{i_{t+1}}a_{i_t}\cdots a_{i_s})+f(a_{i_1}\cdots[a_{i_{t+1}},a_{i_{t+2}}]a_{i_t}\cdots a_{i_s})+f(a_{i_1}\cdots a_{i_{t+1}}[a_{i_t},a_{i_{t+2}}]\cdots a_{i_s})+$
$f(a_{i_1}\cdots[a_{i_t},a_{i_{t+1}}]a_{i_{t+2}}\cdots a_{i_s})$

First using the right inversion gives us
$f(a_{i_1}\cdots a_{i_t}a_{i_{t+2}}a_{i_{t+1}}\cdots a_{i_s})+f(a_{i_1}\cdots a_{i_t}[a_{i_{t+1}},a_{i_{t+2}}]\cdots a_{i_s})$
Using the inversion on $a_{i_t}a_{i_{t+2}}$ in the first term, we get
$f(a_{i_1}\cdots a_{i_{t+2}}a_{i_t}a_{i_{t+1}}\cdots a_{i_s})+f(a_{i_1}\cdots[a_{i_t},a_{i_{t+2}}]a_{i_{t+1}}\cdots a_{i_s})+f(a_{i_1}\cdots a_{i_t}[a_{i_{t+1}},a_{i_{t+2}}]\cdots a_{i_s})$
Using the inversion on $a_{i_t}a_{i_{t+1}}$ in the first term, we get
$f(a_{i_1}\cdots a_{i_{t+2}}a_{i_{t+1}}a_{i_t}\cdots a_{i_s})+f(a_{i_1}\cdots a_{i_{t+2}}[a_{i_t},a_{i_{t+1}}]\cdots a_{i_s})+f(a_{i_1}\cdots[a_{i_t},a_{i_{t+2}}]a_{i_{t+1}}\cdots a_{i_s})+$
$f(a_{i_1}\cdots a_{i_t}[a_{i_{t+1}},a_{i_{t+2}}]\cdots a_{i_s})$

Now look at equation (1).
$f(a_{i_1}\cdots a_{i_t}a_{i_{t+1}}\cdots a_{i_s})=f(a_{i_1}\cdots a_{i_{t+1}}a_{i_t}\cdots a_{i_s})+f(a_{i_1}\cdots[a_{i_t},a_{i_{t+1}}]\cdots a_{i_s})$ if $i_t>i_{t+1}$
By skew-symmetry, this gives

$$\begin{aligned}
f(a_{i_1}\cdots a_{i_{t+1}}a_{i_t}\cdots a_{i_s}) &= f(a_{i_1}\cdots a_{i_t}a_{i_{t+1}}\cdots a_{i_s})-f(a_{i_1}\cdots[a_{i_t},a_{i_{t+1}}]\cdots a_{i_s})\\
&= f(a_{i_1}\cdots a_{i_t}a_{i_{t+1}}\cdots a_{i_s})+f(a_{i_1}\cdots[a_{i_{t+1}},a_{i_t}]\cdots a_{i_s})
\end{aligned}$$

Thus, (1) holds regardless of whether $i_t>i_{t+1}$ or $i_{t+1}>i_t$. By the inductive hypothesis, we may use this relation freely for monomials of dimension $s-1$ and it will not change the result. By linearity, we have that for any $b,c\in g$
$f(a_{i_1}\cdots bc\cdots a_{i_{s-1}})-f(a_{i_1}\cdots cb\cdots a_{i_{s-1}})=f(a_{i_1}\cdots[b,c]\cdots a_{i_{s-1}})$.
Applying this, we find that the difference between the first expression and the second expression above is
$f(a_{i_1}\cdots[[a_{i_{t+1}},a_{i_{t+2}}],a_{i_t}]\cdots a_{i_s})+f(a_{i_1}\cdots[a_{i_{t+1}},[a_{i_t},a_{i_{t+2}}]]\cdots a_{i_s})+f(a_{i_1}\cdots[[a_{i_t},a_{i_{t+1}}],a_{i_{t+2}}]\cdots a_{i_s})$
Using skew-symmetry, this is
$f(a_{i_1}\cdots[[a_{i_{t+1}},a_{i_{t+2}}],a_{i_t}]\cdots a_{i_s})+f(a_{i_1}\cdots[[a_{i_{t+2}},a_{i_t}],a_{i_{t+1}}]\cdots a_{i_s})+f(a_{i_1}\cdots[[a_{i_t},a_{i_{t+1}}],a_{i_{t+2}}]\cdots a_{i_s})$

which is 0 by the Jacobi identity.
Thus, we get the same result regardless of which inversion we start with. □

By induction, for any monomial $a_{i_1} \cdots a_{i_s}$, the evaluation of $f(a_{i_1} \cdots a_{i_s})$ is independent of the sequence of inversions(1) that we use, so $f(a_{i_1} \cdots a_{i_s})$ is well-defined. Thus, $f : T(g) \to B$ is well-defined.

It remains to show that $J(g) \subseteq \ker(f)$. By linearity, it suffices to show that for all i, j, for all $A, B \in T(g)$, $f(A(a_i a_j - a_j a_i - [a_i, a_j])B) = 0$. If $i > j$, this is just the equation (1). If $i < j$, the eqation (1) gives
$f(Aa_j a_i B) = f(Aa_i a_j B) + f(A[a_j, a_i]B)$. Using skew-symmetry,
$f(Aa_i a_j B) = f(Aa_j a_i B) - f(A[a_j, a_i]B) = f(Aa_j a_i B) + f(A[a_i, a_j]B)$.
$f(A(a_i a_j - a_j a_i - [a_i, a_j])B) = 0$, as needed. This completes the proof. □

The Casimir Element of $U(g)$

We assume that $\dim g < \infty$ and g carries a non-degenerate symmetric invariant bilinear form $(\cdot, \cdot)$ (e.g. g is semi-simple and $(a, b) = K(a, b)$)

Choose a basis $\{a_i\}$ of g and let b_i be the dual basis i.e. $(a_i, b_j) = \delta_{ij}$. The Casimir element is the following element of $U(g)$: $\Omega = \sum_{i=1}^{\dim g} a_i b_i$.

Exercise 22.4. Show that Ω is independent of the choice of the basis $\{a_i\}$.

Solution: From linear algebra, to go from one basis to another, it is sufficient to use the following three operations:
1. $a'_i = a_i$ if $i \neq j$, $a'_j = ca_j$ where $c \in F$, $c \neq 0$
2. $a'_i = a_i$ if $i \neq j$, $i \neq k$. $a'_j = a_k$, $a'_k = a_j$
3. $a'_i = a_i$ if $i \neq j$. $a'_j = a_j + ca_k$, where $c \in F$
To show that Ω is independent of the choice of the basis $\{a_i\}$, it is sufficient to show that Ω is invariant under these three operations. For these three operations, the dual basis changes as follows:
1. $b'_i = b_i$ if $i \neq j$, $b'_j = \frac{1}{c} a_j$
2. $b'_i = b_i$ if $i \neq j$, $i \neq k$. $b'_j = b_k$, $b'_k = b_j$
3. $b'_i = b_i$ if $i \neq k$. $b'_k = b_k - cb_j$
In all three cases, it is easily verified that $\Omega' = \sum_{i=1}^{\dim g} a'_i b'_i = \sum_{i=1}^{\dim g} a_i b_i = \Omega$. □

Lemma on dual basis:

Lemma 22.3. *For any $a \in g$ write*
(2) $[a, a_i] = \sum_j \alpha_{ij} a_j$, (3) $[a, b_i] = \sum_j \beta_{ij} b_j$, $\alpha_{ij}, \beta_{ij} \in F$. Then $\alpha_{ij} = -\beta_{ji}$.

Proof. Taking the inner product of (2) with b_j and of (3) with a_j, we get
$([a, a_i], b_j) = \alpha_{ij}$ and $([a, b_i], [a_j]) = \beta_{ij}$. We also have
$([a, a_i], b_j) = (a, [a_i, b_j])$ and $([a, b_i], [a_j]) = (a, [b_i, a_j]) = -(a, [a_j, b_i])$, hence $\alpha_{ij} = -\beta_{ji}$. □

Definition 22.3. Let g be a Lie algbera, and V be a g-module.
(We used the language of a representation π of g in V, notation $\pi(g)V$, $g \in g$, $v \in V$. A little more

convenient is the equivalent language of a g-module V, notation: gV)
A 1-cocycle of g with coefficients in a g-module V is a linear map $f : g \to V$ such that
(4) $f([a,b]) = af(b) - bf(a)$

Example: Trivial 1-cocycle: for $v \in V$, let $f_v(a) = av$.

Exercise 22.5. Show that $f_v : g \to V$ is a 1-cocycle.

Solution: f_v is clearly linear, and $f([a,b]) = [a,b](v) = a(b(v)) - b(a(v)) = af(b) - bf(a)$, as needed. Thus, f_v is a 1-cocycle of g. □

Denote by $Z^1(g,V)$ the space of all 1-cocycles of g with coefficients in V. Then by exercise 22.5, trivial cocycles form a subspace denoted by $B^1(g,V)$.

Definition 22.4. $H^1(g,V) = Z^1(g,v)/B^1(g,V)$ is called the first coboundary.

Note that $H^1(g,V) = 0$ just means that any 1-cocycle of g, i.e. any linear map $f : g \to V$ satisfying (4) is trivial, i.e. of the form $f = f_v$ for some V.

Theorem 22.4. *If g is a semi-simple Lie algebra over a field F of characteristic 0 and V is a finite-dimensional g-module, then $H^1(g,V) = 0$*

Lemma 22.5. *If $\{a_i\}$ and $\{b_j\}$ are dual bases of g and f is a 1-cocycle of g with values in V then for any $a \in g$ we have $a(\sum_i a_i f(b_i)) = \Omega f(a)$*

Exercise 22.6. Prove this using Lemma 1.

Solution: The equation (4) gives that for all $a, b \in g$, $af(b) = bf(a) + f([a,b])$.

$$\begin{aligned}
a(\sum_i a_i f(b_i)) &= \sum_i a(a_i f(b_i)) \\
&= \sum_i a(f([a_i,b_i])) + \sum_i a(b_i f(a_i)) \\
&= \sum_i f([a,[a_i,b_i]]) + \sum_i [a_i,b_i](f(a)) + \sum_i a(b_i f(a_i))
\end{aligned}$$

By the Jacobi identity, for all $a, b, c \in g$, $[a,[b,c]] + [b,[c,a]] + [c,[a,b]] = 0$, so
$[a,[b,c]] = [b,[a,c]] - [c,[a,b]]$.

$$\begin{aligned}
\sum_i f([a,[a_i,b_i]]) &= \sum_i f([a_i,[a,b_i]]) - \sum_i f([b_i,[a,a_i]]) \\
&= \sum_i \sum_j f([a_i, \beta_{ij} b_j]) - \sum_i \sum_j f([b_i, \alpha_{ij} a_j]) \\
&= \sum_i \sum_j \beta_{ij} f([a_i,b_j]) + \sum_j \sum_i \alpha_{ji} f([a_i,b_j]) = 0 \text{ by Lemma 1}
\end{aligned}$$

$$
\begin{aligned}
a(\sum_i a_i f(b_i)) &= \sum_i [a_i, b_i](f(a)) + \sum_i a(b_i f(a_i)) \\
&= \sum_i a_i b_i(f(a)) - \sum_i b_i a_i(f(a)) + \sum_i a(b_i f(a_i)) \\
&= \Omega f(a) - \sum_i b_i a(f(a_i)) - \sum_i b_i f([a_i, a]) + \sum_i a(b_i f(a_i)) \\
&= \Omega f(a) + \sum_i b_i f([a, a_i]) + \sum_i a(b_i f(a_i)) - \sum_i b_i a(f(a_i)) \\
&= \Omega f(a) + \sum_i \sum_j b_i f(\alpha_{ij} a_j) + \sum_i [a, b_i] f(a_i) \\
&= \Omega f(a) + \sum_i \sum_j \alpha_{ij} b_i f(a_j) + \sum_i \sum_j \beta_{ij} b_j f(a_i) = \Omega f(a) \text{ by Lemma 1}
\end{aligned}
$$

□