

(February 19, 2005)

# Poincaré-Birkhoff-Witt Theorem

Paul Garrett [garrett@math.umn.edu](mailto:garrett@math.umn.edu) <http://www.math.umn.edu/~garrett/>

We prove the Poincaré-Birkhoff-Witt Theorem on the structure of enveloping algebras of Lie algebras. The argument here is basically that given in Jacobson, and applies to not-necessarily finite-dimensional Lie algebras over arbitrary fields. The same argument is also that reproduced later in Varadarajan. By contrast, there is a somewhat different argument given in Bourbaki, which is what Humphreys gives.

N. Bourbaki, *Groupes et algèbres de Lie, Chap. 1*, Paris: Hermann, 1960.

N. Jacobson, *Lie Algebras*, Dover, 1962.

J. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag, 1972.

V.S. Varadarajan, *Lie Groups, Lie Algebras, and their Representations*, Springer-Verlag, 1974, 1984.

It may not be clear *a priori* that the *Jacobi identity* necessarily plays an essential role in the argument. At the same time, apart from Jacobson's device of use of the endomorphism 'L' (see below), the argument is mostly very natural. And we note that the Jacobi identity is the assertion that the adjoint action of a Lie algebra on itself is a Lie algebra representation.

Let  $\mathcal{G}$  be a Lie algebra over a field  $k$  of characteristic 0. Let  $\mathcal{T}_n$  be the space

$$\mathcal{T}_n = \underbrace{\mathcal{G} \otimes \dots \otimes \mathcal{G}}_n$$

of **homogeneous tensors of degree  $M$** . Let  $\mathcal{T}$  be the **tensor algebra**

$$\mathcal{T} = k \oplus \mathcal{T}_1 \oplus \mathcal{T}_2 \oplus \dots = \bigotimes \bullet \mathcal{G} = k \oplus \mathcal{G} \oplus (\mathcal{G} \otimes \mathcal{G}) \oplus (\mathcal{G} \otimes \mathcal{G} \otimes \mathcal{G}) \dots$$

of  $\mathcal{G}$ . For  $x, y \in \mathcal{G}$  let

$$u_{x,y} = (x \otimes y - y \otimes x) - [x, y] \in \mathcal{T}_2 + \mathcal{T}_1 \subset \mathcal{T}$$

Let  $\mathcal{I}$  be the two-sided ideal in  $\mathcal{T}$  generated by the set of all elements  $u_{x,y}$ . Since  $u_{x,y} \in \mathcal{T}_1 + \mathcal{T}_2$ , the ideal  $\mathcal{I}$  contains no elements of  $\mathcal{T}_0 \approx k$ , so  $\mathcal{I}$  is a *proper* ideal in  $\mathcal{T}$ .

Let  $\mathcal{A} = \mathcal{T}/\mathcal{I}$  be the quotient, the **universal enveloping algebra** of  $\mathcal{G}$ . Let

$$q : \mathcal{T} \rightarrow \mathcal{A}$$

be the quotient map.

For any basis  $\{x_i : i \in I\}$  of  $\mathcal{G}$  the images  $q(x_{i_1} \otimes \dots \otimes x_{i_n})$  in  $\mathcal{A}$  of tensor monomials  $x_{i_1} \otimes \dots \otimes x_{i_n}$  *span* the enveloping algebra over  $k$ , since they span the tensor algebra.

Suppose now that the indexing set  $I$  is *ordered*. Using the Lie bracket  $[,]$ , we can do a certain amount of rearranging of the  $x_{i_j}$  in a monomial. We *anticipate* that everything in  $\mathcal{A}$  can be rewritten to be as sum of monomials  $x_{i_1} \dots x_{i_n}$  where

$$i_1 \leq i_2 \leq \dots \leq i_n$$

A monomial in which the indices possess this ordering is a **standard monomial**.

**Theorem:** (*Poincaré-Birkhoff-Witt*) The standard monomials form a *basis* for the enveloping algebra  $\mathcal{A}$  of a Lie algebra  $\mathcal{G}$ . In particular, the Lie algebra  $\mathcal{G}$  itself *injects* into the enveloping algebra.

*Proof:* It is relatively easy to form the induction that proves that the (images of) standard monomials *span*  $\mathcal{A}$ . Indeed, given a monomial  $x_{i_1} \dots x_{i_n}$  with some indices not correctly ordered, it must be that there is at least one index  $j$  so that

$$i_j > i_{j+1}$$

Since

$$x_{i_j}x_{i_{j+1}} - x_{i_{j+1}}x_{i_j} - [x_{i_j}, x_{i_{j+1}}] \in \mathcal{I}$$

we have

$$\begin{aligned} x_{i_1} \cdots x_{i_n} &= x_{i_1} \cdots x_{i_{j-1}} \cdot (x_{i_j}x_{i_{j+1}} - x_{i_{j+1}}x_{i_j} - [x_{i_j}, x_{i_{j+1}}]) \cdot x_{i_{j+2}} \cdots x_{i_n} \\ &\quad + x_{i_1} \cdots x_{i_{j-1}}x_{i_{j+1}}x_{i_j}x_{i_{j+2}} \cdots x_{i_n} + x_{i_1} \cdots x_{i_{j-1}}[x_{i_j}, x_{i_{j+1}}]x_{i_{j+2}} \cdots x_{i_n} \end{aligned}$$

The first summand lies inside the ideal  $\mathcal{I}$ , while the third is a tensor of smaller degree. Thus, we do induction on degree of tensors, and for each fixed degree we do induction on the number of pairs of indices out of order.

The far more serious assertion is *linear independence*. For brevity, given a tensor monomial  $x_{i_1} \otimes \cdots \otimes x_{i_n}$ , the **defect** of this monomial is the number of pairs of indices  $i_j, i_{j'}$  so that  $j < j'$  but  $i_j > i_{j'}$ . Suppose that we can define a linear map

$$L : \mathcal{T} \rightarrow \mathcal{T}$$

with the properties that, first,  $L$  is the identity map on a standard monomial, and, second, that whenever  $i_j > i_{j+1}$  we have

$$\begin{aligned} L(x_{i_1} \otimes \cdots \otimes x_{i_n}) &= L(x_{i_1} \otimes \cdots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \cdots \otimes x_{i_n}) \\ &\quad + L(x_{i_1} \otimes \cdots \otimes [x_{i_j}, x_{i_{j+1}}] \otimes \cdots \otimes x_{i_n}) \end{aligned}$$

If so, then  $L(\mathcal{I}) = 0$ , while  $L$  acts as the identity on any linear combination of standard monomials. This would prove that the subspace of  $\mathcal{T}$  consisting of linear combinations of standard monomials meets the ideal  $\mathcal{I}$  just at 0, so maps injectively to the enveloping algebra.

Note that, incidentally,  $L$  would have the property that

$$\begin{aligned} L(y_{i_1} \otimes \cdots \otimes y_{i_n}) &= L(y_{i_1} \otimes \cdots \otimes y_{i_{j+1}} \otimes y_{i_j} \otimes \cdots \otimes y_{i_n}) \\ &\quad + L(y_{i_1} \otimes \cdots \otimes [y_{i_j}, y_{i_{j+1}}] \otimes \cdots \otimes y_{i_n}) \end{aligned}$$

for *any* vectors  $y_{i_j}$  in  $\mathcal{G}$ .

Thus, the problem reduces to defining such a map  $L$ . We do an induction to define  $L$ . First, define  $L$  to be the identity on  $\mathcal{T}_0 + \mathcal{T}_1$ . Note that the second condition on  $L$  is vacuous here, and the first condition is met since every monomial tensor of degree 1 or 0 is standard.

Now fix  $n \geq 2$ , and attempt to define  $L$  on monomials in  $\mathcal{T}_{\leq n}$  inductively by using the second required property: define  $L(x_{i_1} \otimes \cdots \otimes x_{i_n})$  by

$$\begin{aligned} L(x_{i_1} \otimes \cdots \otimes x_{i_n}) &= L(x_{i_1} \otimes \cdots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \cdots \otimes x_{i_n}) \\ &\quad + L(x_{i_1} \otimes \cdots \otimes [x_{i_j}, x_{i_{j+1}}] \otimes \cdots \otimes x_{i_n}) \end{aligned}$$

where  $i_j > i_{j+1}$ . One term on the right-hand side is of lower degree, and the other is of smaller defect. Thus, we do induction on degree of tensor monomials, and for each fixed degree do induction on defect.

The potential problem is the well-definedness of this definition. Monomials of degree  $n$  and of defect 0 are already standard. For monomials of degree  $n$  and of defect 1 the definition is unambiguous, since there is just one pair of indices that are out of order.

So suppose that the defect is at least two. Let  $j < j'$  be two indices so that both  $i_j > i_{j+1}$  and  $i_{j'} > i_{j'+1}$ . To prove well-definedness it suffices to show that the two right-hand sides of the defining relation for  $L(x_{i_1} \otimes \cdots \otimes x_{i_n})$  are actually the same element of  $\mathcal{T}$ .

Consider the case that  $j + 1 < j'$ . Necessarily  $n \geq 4$ . (In this case the two rearrangements do not interact with each other.) Doing the rearrangement specified by the index  $j$ , we have

$$L(x_{i_1} \otimes \dots \otimes x_{i_n}) = L(x_{i_1} \otimes \dots \otimes x_{i_{j+1}} \otimes x_{i_j} \otimes \dots \otimes x_{i_n}) \\ + L(x_{i_1} \otimes \dots \otimes [x_{i_j}, x_{i_{j+1}}] \otimes \dots \otimes x_{i_n})$$

The first summand on the right-hand side has smaller defect, and the second has smaller degree, so we can use the inductive definition to evaluate them both. And still has  $i_{j'} > i_{j'+1}$ . Nothing is lost if we simplify notation by taking  $j = 1, j' = 3$ , and  $n = 4$ , since all the other factors in the monomials are inert. Further, to lighten the notation write  $x$  for  $x_{i_1}$ ,  $y$  for  $x_{i_2}$ ,  $z$  for  $x_{i_3}$ , and  $w$  for  $x_{i_4}$ . We use the inductive definition to obtain

$$L(x \otimes y \otimes z \otimes w) = L(y \otimes x \otimes z \otimes w) + L([x, y] \otimes z \otimes w) \\ = L(y \otimes x \otimes w \otimes z) + L(y \otimes x \otimes [z, w]) \\ + L([x, y] \otimes w \otimes z) + L([x, y] \otimes [z, w])$$

But then it is clear (or can be computed analogously) that the same expression is obtained when the roles of  $j$  and  $j'$  are reversed. Thus, the induction step is completed in case  $j + 1 < j'$ .

Now consider the case that  $j + 1 = j'$ , that is, the case in which the interchanges *do* interact. Here nothing is lost if we just take  $j = 1, j' = 2$ , and  $n = 3$ . And write  $x$  for  $x_{i_1}$ ,  $y$  for  $x_{i_2}$ ,  $z$  for  $x_{i_3}$ . Thus,

$$i_1 > i_2 > i_3$$

Then, on one hand, applying the inductive definition by first interchanging  $x$  and  $y$ , and then doing further reshufflings, we have

$$L(x \otimes y \otimes z) = L(y \otimes x \otimes z) + L([x, y] \otimes z) = L(y \otimes z \otimes x) + L(y \otimes [z, x]) + L([x, y] \otimes z) \\ = L(z \otimes y \otimes x) + L([y, z] \otimes x) + L(y \otimes [z, x]) + L([x, y] \otimes z)$$

On the other hand, starting by doing the interchange of  $y$  and  $z$  gives

$$L(x \otimes y \otimes z) = L(x \otimes z \otimes y) + L(x \otimes [y, z]) = L(z \otimes x \otimes y) + L([x, z] \otimes y) + L(x \otimes [y, z]) \\ = L(z \otimes y \otimes x) + L(z \otimes [x, y]) + L([x, z] \otimes y) + L(x \otimes [y, z])$$

It remains to see that the two right-hand sides are the same.

Since  $L$  is already well-defined, by induction, for tensors of degree  $n - 1$  (here in effect  $n - 1 = 2$ ), we can invoke the property

$$L(v \otimes w) = L(w \otimes v) + L([v, w])$$

for all  $v, w \in \mathcal{G}$ . Apply this to the second, third, and fourth terms in the first of the two previous computations, to obtain

$$L(x \otimes y \otimes z) \\ = L(z \otimes y \otimes x) + \left( L(x \otimes [y, z]) + L([[y, z], x]) \right) + \left( L([x, z] \otimes y) + L([y, [x, z]]) \right) + \left( L(z \otimes [x, y]) + L([[x, y], z]) \right)$$

The latter differs from the right-hand side of the *second* computation just by the expressions involved doubled brackets, namely

$$L([[y, z], x]) + L([y, [x, z]]) + L([[x, y], z])$$

Thus, we wish to prove that the latter is 0. Reversing *both* (so no net sign change) the brackets in the middle summand gives the equivalent requirement

$$L([[y, z], x]) + L([[z, x], y]) + L([[x, y], z])$$

which is the **Jacobi identity**. ///