JOURNAL OF ALGEBRA 11, 469–482 (1969)

Baer Invariants and the Birkhoff-Witt Theorem

P. J. HIGGINS

Department of Mathematics, King's College, Strand, London W.C. 2

Received July 4, 1968

1. INTRODUCTION

The Birkhoff-Witt theorem asserts that, under certain conditions on a Lie algebra M (over a commutative ring R), the graded algebra associated with the enveloping algebra of M is canonically isomorphic with the symmetric algebra of the underlying module of M. It then follows that M is embedded in its enveloping algebra. It is known that not all Lie algebras are so embedded, but the theorem has been proved in the following cases:

(i) when M is a free R-module, for any R (Birkhoff [1], Witt [7]);
(ii) when R is a Dedekind domain, for any M (Lazard [5], Cartier [2]);
(iii) when R is an algebra over the rationals, for any M (Cohn [3]).

These conditions refer only to the module structure of M, and we aim to give a unified treatment in which the Lie product on M and the enveloping algebra of M are eliminated from the discussion at an early stage. They are replaced by a homological invariant $B(M)$, defined for an arbitrary module M, which depends only on the tensor algebra of M and is one of the Baer invariants defined for algebras by Fröhlich [4]. The vanishing of $B(M)$ implies the Birkhoff-Witt theorem for all Lie algebras with underlying module M, and we shall show that $B(M) = 0$ in the three cases mentioned above. In fact the condition $B(M) = 0$ is necessary and sufficient for the validity of a closely related embedding theorem for "Lie structures over M" in "associative structures over M." The definition of these concepts is our starting-point.

2. LIE STRUCTURES AND ASSOCIATIVE STRUCTURES

Let R be a commutative ring (with identity) and let M be an R-module. We denote by $T(M) = \bigoplus_{n \geqslant 0} (\bigotimes^n M)$ the tensor algebra of M and identify M with the homogeneous part of $T(M)$ of degree 1. (All tensor products are taken over R unless otherwise stated). The canonical map from $T(M)$ to the

481/11/4-1

symmetric algebra $S(M)$ has as kernel the ideal of $T(M)$ generated by all commutators $xy - yx$ $(x, y \in M)$. We denote this kernel by $K(M)$; it is a homogeneous ideal of $T(M)$.

If we consider $K(M)$ as a $T(M)$-bimodule, it is generated by the commutators $xy - yx$, and these satisfy some obvious relations which we use as axioms in the following definition. A *Lie structure over the R-module M* is a $T(M)$-bimodule A together with a bilinear function $M \otimes M \to A$ (denoted by $x \otimes y \mapsto \langle x, y\rangle$), satisfying the axioms:

(L1) $\langle x, x\rangle = 0 \quad (x \in M)$;

(L2) $\langle x, y\rangle\, t(uv - vu) = (xy - yx)\, t\langle u, v\rangle \quad (x, y, u, v \in M, t \in T(M))$;

(L3) $$(\langle x, y\rangle z - z\langle x, y\rangle) + (\langle y, z\rangle x - x\langle y, z\rangle) + (\langle z, x\rangle y - y\langle z, x\rangle) = 0 \quad (x, y, z \in M).$$

It is easy to check that $K(M)$ is a Lie structure with $\langle x, y\rangle = xy - yx$. So is $T(M)$ with the same definition of $\langle x, y\rangle$, and other examples will occur below.

As with Lie algebras, one can easily obtain Lie structures from similar associative structures. We define an *associative structure over* M to be a $T(M)$-bimodule B together with a bilinear function $M \otimes M \to B$ (denoted by $x \otimes y \mapsto (x, y)$) satisfying the associative law:

(A) $(x, y)z = x(y, z) \quad (x, y, z \in M)$.

Then B becomes a Lie structure over M if we define $\langle x, y\rangle = (x, y) - (y, x)$. Axioms (L1) and (L3) are obviously satisfied. Axiom (L2) need only be checked when $t = z_1 z_2 \cdots z_n$ $(z_i \in M)$, and in this case axiom (A) implies $(x, y)\, z_1 z_2 \cdots z_n uv = xyz_1 z_2 \cdots z_n(u, v)$, from which (L2) follows easily.

For any given module M there is a universal Lie structure $L(M)$ over M which can be described as follows. $L(M)$ is generated as $T(M)$-bimodule by symbols $\langle x, y\rangle$, one for each pair of elements x, y of M, and has defining relations (L1), (L2), (L3) together with the relations that assert the bilinearity of the function $\langle , \rangle$. This $L(M)$ is characterised up to isomorphism by the universal property: if $(A, \langle\!\langle , \rangle\!\rangle)$ is any Lie structure over M then there is a unique morphism $L(M) \to A$ of $T(M)$-bimodules such that $\langle x, y\rangle \mapsto \langle\!\langle x, y\rangle\!\rangle$ for all $x, y \in M$.

There is also, of course, a universal associative structure over M, but we do not need a special notation for it since it can easily be identified. In $T(M)$, the ideal $M^2T(M) = \bigoplus_{n \geqslant 2} (\otimes^n M)$ is an associative structure over M with $(x, y) = xy$, and we have

THEOREM 1. *$M^2T(M)$ is the universal associative structure over M.*

Proof. Let $(B, (,))$ be any associative structure over M. For $j \geqslant 2$ we can map $\otimes^j M$ to B by the rule $x_1 \otimes x_2 \otimes \cdots \otimes x_j \mapsto (x_1, x_2)\, x_3 \cdots x_j$ since the latter is an R-multilinear function. This gives a map $\theta : M^2T(M) \to B$ which is clearly a homomorphism of right $T(M)$-modules and maps xy to (x, y). To show that it is also a left $T(M)$-homomorphism it is enough to show that $y_1 y_2 \cdots y_n(x_1, x_2)\, x_3 \cdots x_j = (y_1, y_2)\, y_3 \cdots y_n x_1 \cdots x_j$, $(y_r, x_s \in M)$, and this is a consequence of the associative law for B. The map θ is unique because the elements xy $(x, y \in M)$ generate $M^2T(M)$ as $T(M)$-bimodule.

It is unfortunately not always true that $K(M)$ is the universal Lie structure over M. Indeed, it is precisely when $K(M) \simeq L(M)$ that we can prove the Birkhoff–Witt theorem for Lie algebras on the module M. The proof of this theorem given in the next section is essentially that of Lazard [5], with some simplifications made possible by the axiomatic approach.

3. The Birkhoff–Witt Theorem

Suppose that we are given a multiplication $M \otimes M \to M$ $(x \otimes y \mapsto [x, y])$ which makes M a Lie algebra over R. The elements $\langle x, y\rangle = xy - yx - [x, y]$ of $T = T(M)$ $(x, y \in M)$ generate a (non-homogeneous) ideal J, and the quotient algebra $E = T/J$ is the *enveloping algebra* of the Lie algebra M. The filtration $T_0 \subset T_1 \subset T_2 \subset \cdots$ of T given by $T_n = \bigoplus_{i \leqslant n} (\otimes^i M)$ induces filtrations of $K = K(M)$, $S = S(M)$, J and E as follows: $K_n = K \cap T_n$, $S_n = (T_n + K)/K$, $J_n = J \cap T_n$, $E_n = (T_n + J)/J$. E is then a filtered algebra and we denote by $G = \bigoplus_{n \geqslant 1} (E_n/E_{n-1})$ the associated graded algebra. The homogeneous components of G are

$$E_n/E_{n-1} \cong (T_n + J)/(T_{n-1} + J) \cong T_n/(T_n \cap (T_{n-1} + J)) \cong T_n/(T_{n-1} + J_n).$$

Now S is itself graded with homogeneous components

$$S_n/S_{n-1} \cong T_n/(T_{n-1} + K_n),$$

and it is clear from the definition of J that $T_{n-1} + J_n \supset T_{n-1} + K_n$. We therefore have canonical surjections $\sigma_n : S_n/S_{n-1} \to E_n/E_{n-1}$ with kernels $(T_{n-1} + J_n)/(T_{n-1} + K_n)$. These give the canonical surjection $\sigma : S \to G$ which is in fact the algebra homomorphism induced by the canonical map from M to the (commutative) algebra G. If σ is an isomorphism we say that the Lie algebra M has the Birkhoff–Witt property, and this is clearly equivalent to the condition

$$J_n \subset T_{n-1} + K_n \qquad (n \geqslant 1). \tag{1}$$

Theorem 2. *If $(K(M), xy - yx)$ is the universal Lie structure over M then every Lie algebra on M has the Birkhoff–Witt property.*

Proof. Following Lazard, we introduce modules $J_{(n)}$ $(n \geqslant 1)$ which consist of those elements of J that are most obviously in T_n, namely $J_{(n)} = \sum (T_r\langle x, y\rangle T_s \mid x, y \in M, r + s = n - 2)$. Then $J_{(1)} = 0$, $J_{(2)}$ is the R-module spanned by all $\langle x, y\rangle = xy - yx - [x, y]$, and $J_{(1)} \subset J_{(2)} \subset \cdots \subset J_{(n)} \subset \cdots \subset J$. Also $J_{(n)} \subset J_n$, and $\bigcup_n J_{(n)} = J$. We write $A_n = J_{(n)}/J_{(n-1)}$ $(n \geqslant 2)$, and form the graded R-module $A = \bigoplus_{n\geqslant 2} A_n$ associated with this new filtration of J. Since $J_{(n)}M + MJ_{(n)} \subset J_{(n+1)}$, A has the structure of a T-bimodule with $A_nM + MA_n \subset A_{n+1}$. Also, $\langle x, y\rangle \in J_{(2)} = A_2$ for all $x, y \in M$.

LEMMA. $(A, \langle,\rangle)$ *is a Lie structure over* M.

Proof. (L1): this is trivially true since $\langle x, x\rangle = [x, x] = 0$ in M.

(L2): let $w = \langle x, y\rangle\, t(uv - vu) - (xy - yx)\, t\langle u, v\rangle$, where $x, y, u, v \in M$, and where $t \in T$ is homogeneous of degree n. Then, calculating in A, we have $w \in A_{n+4}$ since $\langle x, y\rangle \in A_2$ and $\langle u, v\rangle \in A_2$. But if we calculate in $J_{(n+4)}$ we find that

$$\begin{aligned} w &= \langle x, y\rangle\, t\{\langle u, v\rangle + [u, v]\} - \{\langle x, y\rangle + [x, y]\}\, t\langle u, v\rangle \\ &= \langle x, y\rangle\, t[u, v] - [x, y]\, t\langle u, v\rangle, \end{aligned}$$

and this lies in $J_{(n+3)}$ since $[u, v]$ and $[x, y]$ lie in M. Hence $w = 0$ in A, and (L2) follows for all t by linearity.

(L3): let $x, y, z \in M$, and put

$$u = \{\langle x, y\rangle z - z\langle x, y\rangle\} + \{\langle y, z\rangle x - x\langle y, z\rangle\} + \{\langle z, x\rangle y - y\langle z, x\rangle\}.$$

Then, calculating in A, we have $u \in A_3$. On the other hand, calculating in $J_{(3)}$, if we replace $\langle x, y\rangle$ by $xy - yx - [x, y]$, we obtain

$$u = \{[x, y]z - z[x, y]\} + \{[y, z]x - x[y, z]\} + \{[z, x]y - y[z, x]\},$$

the other terms cancelling. Now $[x, y] \in M$, so $\langle [x, y], z\rangle \in J_{(2)}$, that is,

$$[x, y]z - z[x, y] \equiv [[x, y], z] \qquad (\mathrm{mod}\ J_{(2)}).$$

Permuting x, y, z cyclically and adding, we therefore have $u \in J_{(2)}$ by the Jacobi law in M, and this means that $u = 0$ in A. The lemma is now proved.

To prove the theorem we use the hypothesis that $K(M)$ is the *universal* Lie structure over M to obtain a morphism $\theta : K \to A$ of T-bimodules sending $xy - yx$ to $\langle x, y\rangle$ for all $x, y \in M$. We claim that θ is an isomorphism. For let δ_n denote the R-linear map which sends any element of T to its homogeneous part of degree n. It is clear from the definition of $J_{(n)}$ that δ_n maps $J_{(n)}$ into K_n and therefore induces a map

$$\delta_n^* : A_n = J_{(n)}/J_{(n-1)} \to K_n/K_{n-1}.$$

The maps δ_n^* combine to give a map $\delta : A \to \bigoplus (K_n/K_{n-1}) = K$ which sends $\langle x, y\rangle$ to $xy - yx$ $(x, y \in M)$, and it is easy to check that δ is a morphism of T-bimodules. Since A and K are generated as T-bimodules by all $\langle x, y\rangle$ and all $xy - yx$, respectively, we see that θ and δ are inverse isomorphisms. In particular, $\delta_n^* : J_{(n)}/J_{(n-1)} \to K_n/K_{n-1}$ is an injection and it follows that every element of $J_{(n)}$ not in $J_{(n-1)}$ has leading term of degree exactly n. Since $\bigcup_n J_{(n)} = J$, this implies that $J_{(n)} = J \cap T_n = J_n$ for all n. But $J_{(n)} \subset K_n + T_{n-1}$, so we have established the condition (1) which is equivalent to the Birkhoff–Witt property.

4. Lie Structures over Free Modules

Before investigating Lie structures over arbitrary modules M we need to know the situation for the special case when M is free. Our next theorem, combined with Theorem 2, gives a new proof of the Birkhoff–Witt theorem in this case.

THEOREM 3. *Let M be a free R-module. Then every Lie structure $(A, \langle,\rangle)$ over M can be embedded in an associative structure $(B, (,))$ over M so that $\langle x, y\rangle = (x, y) - (y, x)$.*

Proof. Let A be a given Lie structure and put $B = A \oplus S(M)$. We shall show how to make B an associative structure with the required property. Let X be a basis for M over R and take a fixed total ordering $\leqslant$ of X. If $x_i \in X$ we denote by ξ_i its image in $S(M) = S$. Then S has a basis consisting of all products $\xi_1\xi_2 \cdots \xi_n$ $(n \geqslant 0)$, where the x_i are in X and $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n$.

We make B a T-bimodule as follows. The action of T on A is to be the given action. To define the action of T on S we need only define maps $M \otimes S \to B$ $(m \otimes \sigma \to m\sigma)$ and $S \otimes M \to B$ $(\sigma \otimes m \to \sigma m)$ such that $(m\sigma)n = m(\sigma n)$ for all $m, n \in M$, $\sigma \in S$. Since M is free we can define $x\sigma$ and σy arbitrarily in B for $x, y \in X$ and σ a basis element $\xi_1\xi_2 \cdots \xi_n$ of S, and we need only check that

$$(x\sigma)y = x(\sigma y) \tag{2}$$

in this case. So let $\sigma = \xi_1\xi_2 \cdots \xi_n$ $(x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n$ in $X)$ and let ξ, η be the images of x, y in S. We define

$$x\sigma = x \circ \sigma + \xi\sigma, \qquad \sigma y = \sigma \circ y + \sigma\eta,$$

where $\xi\sigma$, $\sigma\eta$ are products in S and $x \circ \sigma$, $\sigma \circ y$ are the elements of A defined by

$$x \circ \sigma = \sum_{x_i < x} x_1 x_2 \cdots x_{i-1} \langle x, x_i \rangle x_{i+1} \cdots x_n ,$$
$$\sigma \circ y = \sum_{x_i > y} x_1 x_2 \cdots x_{i-1} \langle x_i , y \rangle x_{i+1} \cdots x_n ,$$

with the convention that $x \circ 1 = 1 \circ y = 0$. (These definitions are just an imitation of the operations in the associative structure $T(M)$ in terms of the splitting $T(M) = K(M) + S(M)$). Since $(\xi\sigma)\eta = \xi(\sigma\eta)$ in S, equation (2) is equivalent to

$$(\xi\sigma) \circ y - x(\sigma \circ y) = x \circ (\sigma\eta) - (x \circ \sigma) y. \tag{3}$$

To simplify the notation, let $\tau = \xi\sigma\eta$ (product in S), and rename x, y and $x_1, x_2, \ldots, x_n$ so that $\tau = \eta_1\eta_2 \cdots \eta_k$ with $y_1 \leqslant y_2 \leqslant \cdots \leqslant y_k$ $(k = n + 2)$ and $x = y_r$, $y = y_s$ $(r \neq s)$. Then $\sigma = (\eta_1\eta_2 \cdots \eta_k)_{\hat{r}\hat{s}}$, where the subscripts $\hat{r}$, $\hat{s}$ denote that the factors with subscripts r, s are to be omitted. Also $\xi\sigma = (\eta_1\eta_2 \cdots \eta_k)_{\hat{s}}$ and $\sigma\eta = (\eta_1\eta_2 \cdots \eta_k)_{\hat{r}}$. Writing $\epsilon_{rs} = 0$ if $r \leqslant s$ and $\epsilon_{rs} = 1$ if $r > s$, the left hand side of (3) becomes

$$(\xi\sigma) \circ y - x(\sigma \circ y) = \sum_j \epsilon_{js}(y_1 \cdots y_{j-1} \langle y_j , y_s \rangle y_{j+1} \cdots y_k)_{\hat{s}}$$
$$- y_r \sum_{j \neq r} \epsilon_{js}(y_1 \cdots y_{j-1} \langle y_j , y_s \rangle y_{j+1} \cdots y_k)_{\hat{r}\hat{s}} .$$

The terms $j \neq r$ in the first sum appear in the second sum with y_r moved to the left hand end. This move can be accomplished by adding terms containing commutators $y_i y_r - y_r y_i$ and possibly $\langle y_j , y_s \rangle y_r - y_r \langle y_j , y_s \rangle$. We therefore have $(\xi\sigma) \circ y - x(\sigma \circ y) = U + V + W$, where

$$U = \epsilon_{rs}(y_1 \cdots y_{r-1} \langle y_r , y_s \rangle y_{r+1} \cdots y_k)_{\hat{s}} ,$$
$$V = -\sum_{j \neq r} \sum_{\substack{i \neq s \\ i \neq j}} \epsilon_{js}\epsilon_{ri}(y_1 \cdots y_{i-1}(y_r y_i - y_i y_r) y_{i+1} \cdots y_{j-1} \langle y_j , y_s \rangle y_{j+1} \cdots y_k)_{\hat{r}\hat{s}} ,$$
$$W = \sum_j \epsilon_{js}\epsilon_{rj}(y_1 \cdots y_{j-1}(\langle y_j , y_s \rangle y_r - y_r \langle y_j , y_s \rangle) y_{j+1} \cdots y_k)_{\hat{r}\hat{s}} .$$

(The notation in V is not meant to imply that $i < j$). Similarly, the right hand side of (3) is $U' + V' + W'$, where

$$U' = \epsilon_{rs}(y_1 \cdots y_{s-1} \langle y_r , y_s \rangle y_{s+1} \cdots y_k)_{\hat{r}} ,$$
$$V' = -\sum_{i \neq s} \sum_{\substack{j \neq r \\ j \neq i}} \epsilon_{ri}\epsilon_{js}(y_1 \cdots y_{i-1} \langle y_r , y_i \rangle y_{i+1} \cdots y_{j-1}(y_j y_s - y_s y_j) y_{j+1} \cdots y_k)_{\hat{r}\hat{s}} ,$$
$$W' = \sum_j \epsilon_{rj}\epsilon_{js}(y_1 \cdots y_{j-1}(y_s \langle y_r , y_j \rangle - \langle y_r , y_j \rangle y_s) y_{j+1} \cdots y_k)_{\hat{r}\hat{s}} .$$

Now $V = V'$ by axiom (L2) in A. Also, by axiom (L3),

$$\begin{aligned} W - W' &= \sum_{s<j<r} (y_1 \cdots y_{j-1}(y_j\langle y_s, y_r\rangle - \langle y_s, y_r\rangle y_j)y_{j+1} \cdots y_k)_{\hat{r}\hat{s}} \\ &= \epsilon_{rs}\{(y_1 \cdots y_{r-1}\langle y_s, y_r\rangle y_{r+1} \cdots y_k)_{\hat{s}} \\ &\quad - (y_1 \cdots y_{s-1}\langle y_s, y_r\rangle y_{s+1} \cdots y_k)_{\hat{r}}\} \\ &= U' - U \end{aligned}$$

since (L1) implies that $\langle y_s, y_r\rangle = -\langle y_r, y_s\rangle$. Thus equation (3) is satisfied and B is a T-bimodule.

Now for $x, y \in X$ we have

$$x\eta = \xi y = \begin{cases} \xi\eta & \text{if } x \leqslant y \\ \xi\eta + \langle x, y\rangle & \text{if } x \geqslant y. \end{cases} \tag{4}$$

We may therefore define $(x, y) = x\eta = \xi y$ for $x, y \in X$, and extend linearly to the whole of M. Then B is an associative structure over M since the equation $(x, y)z = x(y, z)$ is R-multilinear and is a special case of (2) when $x, y, z \in X$. Finally, $(x, y) - (y, x) = x\eta - y\xi = \langle x, y\rangle$ by (4) when $x, y \in X$, and the equality holds for $x, y \in M$ by linearity.

COROLLARY 1. *If M is a free R-module then $K(M)$ is the universal Lie structure over M.*

Proof. Let $(A, \langle,\rangle)$ be any Lie structure over M and embed A in the associative structure B as above. By Theorem 1 there is a map $\theta : M^2T(M) \to B$ of $T(M)$-bimodules sending xy to (x, y) for $x, y \in M$. Since $(x, y) - (y, x) = \langle x, y\rangle \in A$, θ induces a map of $T(M)$-bimodules $K(M) \to A$ sending $xy - yx$ to $\langle x, y\rangle$.

Using Theorem 2 we now obtain

COROLLARY 2. *Every Lie algebra over R whose underlying module is free has the Birkhoff–Witt property.*

5. Baer Invariants of Tensor Algebras

We now introduce two invariants of a module M analogous to the invariants $[F, F]/[F, R]$ and $(R \cap [F, F])/[F, R]$ for a group presented as a quotient F/R of a free group. In Fröhlich's notation [4] they are $D_0V(T(M))$

and $D_1U(T(M))$, where U and V are the functors on algebras associated with the variety of commutative algebras: $U(T(M)) = S(M)$, $V(T(M)) = K(M)$. We recall briefly their definition and main properties.

We start with a short exact sequence of R-modules

$$0 \to Q \to P \to M \to 0 \tag{5}$$

with P projective, and we make identifications so that $Q \subset P \subset T(P)$. Then we have a commutative diagram with exact rows and columns

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & K(P) \cap \underset{\sim}{Q} & \to & \underset{\sim}{Q} & \to & \underset{\sim}{Q}^* & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & K(P) & \to & T(P) & \to & S(P) & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \to & K(M) & \to & T(M) & \to & S(M) & \to & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

where $\underset{\sim}{Q}$ is the ideal $T(P)QT(P)$ of $T(P)$ generated by Q, and $\underset{\sim}{Q}^*$ is its image in $S(P)$. Let W be the set of words

$$w_i(\mathbf{x}) = w_i(x_1, x_2, \ldots, x_n) = x_1x_2 \cdots x_{i-1}[x_i, x_{i+1}]\, x_{i+2} \cdots x_n$$

($n \geqslant 2$, $1 \leqslant i \leqslant n - 1$), where $[a, b]$ denotes, as always from now on, the additive commutator $ab-ba$. If the x_j take values running through P then W generates $K(P)$ as additive group. Let W' be the set of derived words $w_i'(\mathbf{x}, \mathbf{y}) = w_i(\mathbf{x} + \mathbf{y}) - w_i(\mathbf{x})$. If the x_j run through P and the y_j run through Q, W' generates (as additive group) an R-module Z. Clearly the image of any such element $w_i'(\mathbf{p}, \mathbf{q})$ in $T(M)$ or in $S(P)$ is 0, so $Z \subset K(P) \cap \underset{\sim}{Q}$, and it is not difficult to check that in fact $Z = [T(P), \underset{\sim}{Q}]$, the ideal of $T(P)$ generated by all $[t, q]$ ($t \in T(P)$, $q \in Q$). We write $B(M) = (K(P) \cap \underset{\sim}{Q})/Z$ and $C(M) = K(P)/Z$. (These are respectively $D_1UT(M)$ and $D_0VT(M)$).

Now $K(P)$, $\underset{\sim}{Q}$ are $T(P)$-bimodules and $K(P)\underset{\sim}{Q} + \underset{\sim}{Q}K(P) \subset Z$ (because e.g. $p_1p_2 \cdots [p_i, p_{i+1}] \cdots q \cdots p_n$ is of the form $w_i'(\mathbf{p}, \mathbf{q})$). Hence $B(M)$ and $C(M)$ are $T(M)$-bimodules, and we have an exact sequence of $T(M)$-bimodules:

$$0 \to B(M) \to C(M) \to K(M) \to 0. \tag{6}$$

As the notation suggests, $B(M)$ and $C(M)$ depend only on M and not on its presentation (5). To see this directly, let $0 \to Q' \to P' \to M \to 0$ be another presentation of M with P' projective. Then there are R-linear maps σ, σ_* making the following diagram commute:

$$\begin{tikzcd} 0 \arrow[r] & \underset{\sim}{Q} \arrow[r] \arrow[d, "\sigma_*"] & P \arrow[r] \arrow[d, "\sigma"] & M \arrow[r] \arrow[d, "1_M"] & 0 \\ 0 \arrow[r] & \underset{\sim}{Q}' \arrow[r] & P' \arrow[r] & M \arrow[r] & 0. \end{tikzcd}$$

The map $T(\sigma) : T(P) \to T(P')$ is an algebra homomorphism and sends $\bar{\underset{\sim}{Q}}$ into $\bar{\underset{\sim}{Q}}'$. Also, since $K(P)$ and Z are defined by algebra words, $T(\sigma)$ sends $K(P)$ into $K(P')$ and Z into Z'. Hence σ induces maps $B(M) \to B'(M)$ and $C(M) \to C'(M)$ (in the obvious notation). Similarly, there is a map $\tau : P' \to P$ inducing maps in the opposite direction. It is enough, therefore, to show that if $\rho : P \to P$ induces the identity map on M then it induces the identity map on $B(M)$ and on $C(M)$. But this is clear since if $u \in K(P)$ then u is a sum of elements $w_i(\mathbf{p})$ $(p_j \in P)$, and so $u\rho - u = \sum (w_i(\mathbf{p}\rho) - w_i(\mathbf{p})) \in Z$ (because $p_j\rho - p_j \in \underset{\sim}{Q}$). A similar argument with 1_M replaced by an arbitrary map of R-modules shows that B and C are functors from R-modules to R-modules.

We now show that $C(M) = K(P)/Z$ is a Lie structure over M in a natural way. We have already shown that $C(M)$ is a $T(M)$-bimodule. We also know that $K(P)$ is a Lie structure over P with respect to the operation $[x, y] = xy - yx$ $(x, y \in P)$. Suppose that $x \equiv x' \pmod{\underset{\sim}{Q}}$ and $y \equiv y' \pmod{\underset{\sim}{Q}}$. Then $[x, y] \equiv [x', y'] \pmod{Z}$, so for $\xi, \eta \in M$ we may define $[\![\xi, \eta]\!] \in C(M) = K(P)/Z$ to be the image in $C(M)$ of $[x, y]$, where $x, y \in P$ have images ξ, η in M. The axioms for a Lie structure hold in $C(M)$ over M because they hold in $K(P)$ over P.

THEOREM 4. *For any R-module M, $(C(M), [\![,]\!])$ is the universal Lie structure over M.*

Proof. Let $(A, \langle,\rangle)$ be any Lie structure over M. In constructing $C(M)$ we may choose P to be a free R-module, and in this case we know (Theorem 3, Corollary 1) that $(K(P), [,])$ is the universal Lie structure over P. Now A can be viewed as a Lie structure over P *via* the map $\theta : P \to M$ of the presentation, so there is a unique map $\alpha : K(P) \to A$ of $T(P)$-bimodules sending $[x, y]$ to $\langle x\theta, y\theta\rangle$ for all $x, y \in P$. If $x_1, x_2, \ldots, x_n \in P$ then

$$\begin{aligned} w_i(x_1, x_2, \ldots, x_n)\alpha &= (x_1 \cdots x_{i-1}[x_i, x_{i+1}]\, x_{i+2} \cdots x_n)\alpha \\ &= (x_1\theta) \cdots (x_{i-1}\theta)\langle x_i\theta, x_{i+1}\theta\rangle(x_{i+2}\theta) \cdots (x_n\theta). \end{aligned}$$

Hence, for $q_1, q_2, \ldots, q_n \in \underset{\sim}{Q}$, $(w_i(\mathbf{x} + \mathbf{q}) - w_i(\mathbf{x}))\alpha = 0$, i.e. $Z \subset \operatorname{Ker} \alpha$. Thus α induces a map $\beta : C(M) = K(M)/Z \to A$ sending $[\![\xi, \eta]\!]$ to $\langle \xi, \eta\rangle$ for all

$\xi, \eta \in M$. It is easy to see that β is a morphism of $T(M)$-bimodules, and it is unique since the $[x, y]$ generate $K(P)$ as $T(P)$-bimodule and therefore the $[[\xi, \eta]]$ generate $C(M)$ as $T(M)$-bimodule.

The canonical maps $C(M) \to L(M)$ given by this theorem form a natural equivalence of functors $C \simeq L$. We may therefore write $L(M)$ for $C(M)$ from now on, and we have the exact sequence

$$0 \to B(M) \to L(M) \to K(M) \to 0 \tag{6'}$$

for all R-modules M. Clearly this gives an exact sequence of functors $0 \to B \to L \to K \to 0$, and combining it with $0 \to K \to T \to S \to 0$ we obtain the exact sequence of functors

$$0 \to B \to L \to T \to S \to 0. \tag{7}$$

THEOREM 5. *For any R-module M the following are equivalent:*

(i) $B(M) = 0$;

(ii) $(K(M), [,])$ *is the universal Lie structure over* M;

(iii) *every Lie structure over M is embeddable in an associative structure over M.*

Proof. The equivalence of (i) and (ii) follows from the exact sequence (6′) (which is a consequence of Theorem 4).

(ii) $\Rightarrow$ (iii). Let $(A, \langle,\rangle)$ be any Lie structure over M and let A_0 be the $T(M)$-bimodule generated in A by the elements $\langle x, y\rangle$ $(x, y \in M)$. If (ii) holds then there is a unique morphism $\alpha : K(M) \to A$ of $T(M)$-bimodules sending $[x, y]$ to $\langle x, y\rangle$. The kernel D of α is a $T(M)$-bimodule, i.e. an ideal of $T(M)$. The algebra $A^* = T(M)/D$ is an associative structure over M and the Lie structure $A_0 \cong K(M)/D$ is embedded in it. To extend this to an embedding of A itself is a trivial matter. Any $T(M)$-bimodule containing A^* is also an associative structure over M, so we need only form the fibre coproduct of A and A^* with respect to the embeddings $A_0 \to A$ and $A_0 \to A^*$. It is clear that $\langle x, y\rangle$ goes to xy–yx in the resulting embedding of A.

(iii) $\Rightarrow$ (ii). If (iii) holds then the universal Lie structure $L = (L(M), \langle,\rangle)$ is embeddable in an associative structure $(L^*, (,))$ so that $(x, y) - (y, x) = \langle x, y\rangle$. By Theorem 1, there is a morphism of $T(M)$-bimodules $\alpha : M^2T(M) \to L^*$ sending xy to (x, y), and this induces a morphism $\beta : K(M) \to L$ sending $xy - yx$ to $\langle x, y\rangle$. Clearly β is inverse to the canonical map $L(M) \to K(M)$, so $K(M) \cong L(M)$.

6. Modules with $B(M) = 0$

Our main result is an immediate consequence of Theorems 2 and 5:

Theorem 6. *If $B(M) = 0$ for the R-module M then the Birkhoff–Witt theorem holds for all Lie algebras over R with underlying module M.*

To show that this theorem contains the known results quoted in the introduction we now look for conditions on the R-module M which ensure that $B(M) = 0$.

Theorem 7. *Let R be a fixed commutative ring and let M be any R-module.*

(i) *If M is R-projective then $B(M) = 0$.*

(ii) *If M is uniquely divisible as Abelian group* (i.e. *M is a rational vector space*) *then* $B(M) = 0$.

(iii) *If M is a direct sum of cyclic* (i.e. *one-generator*) *modules then* $B(M) = 0$.

Proof. Let $0 \to \underset{\sim}{Q} \to P \to M \to 0$ be a presentation of M with P projective. Then, in the notation of Section 5, $B(M) = (K(P) \cap \bar{\underset{\sim}{Q}})/Z$. Item (i) is clear since if M is projective we may take $P = M$ and $\underset{\sim}{Q} = 0$. To prove (ii) and (iii) we first observe that $K(P)$, $\bar{\underset{\sim}{Q}}$ and Z are homogeneous ideals of $T(P)$, so it is enough to take $u \in K(P) \cap \bar{\underset{\sim}{Q}}$ homogeneous of degree $n (n \geqslant 2)$ and show that $u \in Z$. Now the symmetric group $\mathscr{S}_n$ acts on the homogeneous part $\otimes^n P$ of $T(P)$, and if $\pi \in \mathscr{S}_n$, $u \in \otimes^n P$, then $u - u\pi \in K(P)$. Moreover, if $u \in \bar{\underset{\sim}{Q}}$, then $u - u\pi$ is a sum of elements of type $p_1 p_2 \cdots p_{i-1}[p_i, q]\, p_{i+2} \cdots p_n$ or $p_1 \cdots p_{i-1}[p_i, p_{i+1}]\, p_{i+2} \cdots q \cdots p_n$, where $p_j \in P$ and $q \in \underset{\sim}{Q}$. All such elements are in Z, by definition, so $u - u\pi \in Z$ whenever $u \in \bar{\underset{\sim}{Q}} \cap \otimes^n P$. On the other hand, if $u \in K(P)$ and is homogeneous of degree n then u is a sum of elements of the form $v - v\tau$, where $\tau \in \mathscr{S}_n$ is a transposition. Hence $\sum_{\pi \in \mathscr{S}_n} u\pi = \sum_{\pi \in \mathscr{S}_n} (v\pi - v\tau\pi) = 0$ in this case. Thus, for any $u \in \bar{\underset{\sim}{Q}} \cap K(P) \cap \otimes^n P$, we have $n!u = \sum_{\pi \in \mathscr{S}_n} (u - u\pi) \in Z$. This shows that $B(M)$ is always a torsion group. It is graded by degree: $B(M) = \bigoplus B^n(M)$, and $n!B^n(M) = 0$.

Suppose now that M is uniquely divisible. Then for each integer $k > 0$ we have an isomorphism $k : M \to M$ $(x \mapsto kx)$. Since B is a functor this induces an isomorphism $B(k) : B(M) \to B(M)$ which in dimension n is multiplication by k^n. Taking $k = n!$ we see that $B^n(M) = (n!)^n B^n(m) = 0$, which proves (ii).

To prove (iii) we suppose that $M = \bigoplus_{x \in X} M_x$, where $M_x = R/I_x$ is cyclic, and we take P to be the free R-module on X with the obvious map $P \to M$. Then $\underset{\sim}{Q}$ is spanned by certain elements of the form λx, where $\lambda \in R$ and $x \in X$. We take a fixed total ordering $\leqslant$ of X and denote by S^* the

R-submodule of $T(P)$ spanned by all products $x_1 x_2 \cdots x_n$ with $x_i \in X$ and $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n$ $(n \geqslant 0)$. Then there is an R-linear map $\theta : T(P) \to S^*$ which sends any product of x's to be the product of the same x's in correct order. The kernel of θ is exactly $K(P)$. Suppose now that $u \in \bar{Q} \cap \otimes^n P$. Because of the special form of our presentation we have $u = u_1 + u_2 + \cdots + u_k$ where each u_i is still in $\bar{Q}$ and is of the form $\lambda x_1 x_2 \cdots x_n$ with $\lambda \in R$ and $x_1, x_2, \ldots, x_n \in X$. Then $u_i \theta = u_i \pi_i$ for some $\pi_i \in \mathcal{S}_n$, so $u_i - u_i \theta = u_i - u_i \pi_i \in Z$, as we have already shown. Hence $u - u\theta \in Z$. If now $u \in \bar{Q} \cap K(P) \cap \otimes^n P$ then $u\theta = 0$, and we have $u \in Z$ as required.

COROLLARY. *If R is the direct sum of a finite number of fields or is an algebra over the rationals then $B(M) = 0$ for all R-modules M. If R is a principal ideal domain then $B(M) = 0$ for all finitely generated R-modules M.*

We can extend this last result by general arguments as follows.

THEOREM 8. *If $\{M_\alpha\}$ is a directed system of R-modules, and $M = \varinjlim M_\alpha$, then $B(M) = \varinjlim B(M_\alpha)$.*

Proof. The exact sequence of functors (7) gives rise to a directed system of exact sequences

$$0 \to B(M_\alpha) \to L(M_\alpha) \to T(M_\alpha) \to S(M_\alpha) \to 0.$$

Since $\varinjlim$ is an exact functor for R-modules, we obtain a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \varinjlim B(M_\alpha) & \longrightarrow & \varinjlim L(M_\alpha) & \longrightarrow & \varinjlim T(M_\alpha) & \longrightarrow & \varinjlim S(M_\alpha) & \longrightarrow & 0 \\ & & \downarrow & & \downarrow{\scriptstyle\lambda} & & \downarrow{\scriptstyle\tau} & & \downarrow{\scriptstyle\sigma} & & \\ 0 & \longrightarrow & B(M) & \longrightarrow & L(M) & \longrightarrow & T(M) & \longrightarrow & S(M) & \longrightarrow & 0 \end{array}$$

with exact rows, the limits in the upper row being taken in the category of R-modules. It is enough, therefore, to show that λ, τ, σ are isomorphisms. We give the proof for λ; the other cases are proved by similar "general nonsense" and are in any case well known.

$L_\alpha = L(M_\alpha)$ is the universal Lie structure over M_α. Its structure is given by canonical maps $M_\alpha \otimes M_\alpha \to L_\alpha$, $L_\alpha \otimes M_\alpha \to L_\alpha$ and $M_\alpha \otimes L_\alpha \to L_\alpha$ satisfying axioms (L1), (L2), (L3). If $\theta : M_\alpha \to M_\beta$ is R-linear then $\theta^* = L(\theta) : L_\alpha \to L_\beta$ is obtained by viewing L_β as a Lie structure over M_α *via* the map θ. It is therefore not only R-linear but is compatible with the structure maps, that is, for $x, y \in M$ and $a \in L$, we have $\langle x, y\rangle \theta^* = \langle x\theta, y\theta\rangle$,

$(ax)\,\theta^* = (a\theta^*)(x\theta)$ and $(xa)\,\theta^* = (x\theta)(a\theta^*)$. Hence, writing $\Lambda = \varinjlim L_\alpha$, the structure maps of the various L_α induce maps

$$M \otimes M = \varinjlim (M_\alpha \otimes M_\alpha) \to \Lambda, \qquad \Lambda \otimes M = \varinjlim (L_\alpha \otimes M_\alpha) \to \Lambda$$

and $M \otimes \Lambda \to \Lambda$. The axioms (L1), (L2), (L3) carry over to the limits since each axiom involves only a finite number of symbols and therefore each instance of it is implied by the corresponding axiom for some pair M_α, L_α. Thus Λ is in a canonical way a Lie structure over M. If now A is any Lie structure over M then A can be viewed as a Lie structure over M_α. Hence there is a unique morphism of Lie structures $L_\alpha \to A$ for each α. These induce a unique morphism $\Lambda \to A$ of Lie structures over M which, in the particular case $A = L(M)$, is the map λ. The standard argument for universal objects now shows that λ is an isomorphism.

COROLLARY. *If R is a principal ideal domain then $B(M) = 0$ for all R-modules M.*

Finally, we consider change-of-ring arguments. If R' is a commutative R-algebra and M is an R-module then $M' = M \otimes_R R'$ is an R'-module and we may form its Baer invariant as such. We write $B_{R'}(M')$ to indicate that we are calculating with R'-modules.

THEOREM 9. *If the R-algebra R' is flat over R then $B_{R'}(M \otimes_R R') = B_R(M) \otimes_R R'$.*

Proof. The argument is similar to the one given for direct limits. Since R' is flat over R we have, for any R-module M an exact sequence of R'-modules

$$0 \to B_R(M) \underset{R}{\otimes} R' \to L_R(M) \underset{R}{\otimes} R' \to T_R(M) \underset{R}{\otimes} R' \to S_R(M) \underset{R}{\otimes} R' \to 0.$$

We write $M' = M \otimes_R R'$, $L = L_R(M)$, $L' = L_R(M) \otimes_R R'$. The structure maps for $L : M \otimes_R M \to L$, $L \otimes_R M \to L$, $M \otimes_R L \to L$ induce R'-linear maps $M' \otimes_{R'} M' \to L'$ etc. which clearly make L' a Lie structure over M', and it is easy to check that L' is then the universal Lie structure $L_{R'}(M')$. Similarly, we may identify $T_R(M) \otimes_R R'$ with $T_{R'}(M')$ and $S_R(M) \otimes_R R'$ with $S_{R'}(M')$, and the theorem follows.

In particular, the local $R_\mathfrak{p}$ at a prime ideal $\mathfrak{p}$ of R is flat over R (see, for example, Nagata [6], p. 19). Writing $M_\mathfrak{p}$ for $M \otimes_R R_\mathfrak{p}$ we therefore have the following.

COROLLARY 1. *For any R-module M and any prime ideal $\mathfrak{p}$ of R, $B_{R_\mathfrak{p}}(M_\mathfrak{p}) = (B_R(M))_\mathfrak{p}$. Hence $B_R(M) = 0$ if and only if $B_{R_\mathfrak{p}}(M_\mathfrak{p}) = 0$ for all prime ideals (or all maximal ideals) $\mathfrak{p}$.*

Since the local rings of a Dedekind domain are principal ideal domains, the corollary to Theorem 8 now gives

COROLLARY 2. *If R is a Dedekind domain then $B(M) = 0$ for all R-modules M.*

REFERENCES

1. BIRKHOFF, G. Representability of Lie algebras and Lie groups by matrices. *Annals Math.* **38** (1937), 526–532.
2. CARTIER, P. Remarques sur la théorème de Birkhoff–Witt. *Ann. Scuola norm. sup. Pisa, Sci. fis. mat.* 3 Ser. 12 (1958), 1–4.
3. COHN, P. M. A remark on the Birkhoff–Witt theorem. *J. London Math. Soc.* **38** (1963), 197–203.
4. FRÖHLICH, A. Baer-invariants of algebras. *Trans. Am. Math. Soc.* **109** (1963), 221–244.
5. LAZARD, M. Sur les algèbres enveloppantes universelles de certaines algèbres de Lie. *Publ. Sci. Univ. Alger*, Sér. A. 1 (1954), 281–294.
6. NAGATA, M. "Local rings," (Interscience, New York, 1962).
7. WITT, E. Treuer Darstellung Liescher Ringe. *J. reine angew. Math.* **177** (1937), 152–160.