SOLUTION OF THE RESTRICTED BURNSIDE PROBLEM FOR GROUPS OF ODD EXPONENT

To cite this article: E I Zelmanov 1991 Math. USSR Izv. 36 41

View the article online for updates and enhancements.

Related content
- Efim Isaakovich Zelmanov is 60 years old

Recent citations
- Non-abelian tensor product of residually finite groups
  R. Bastos and N. R. Rocco
- Length-type parameters of finite groups with almost unipotent automorphisms
  E. I. Khukhro and P. Shumyatsky
- Algorithmically complex residually finite groups
  Olga Kharkampovich et al
SOLUTION OF THE RESTRICTED BURNSIDE PROBLEM
FOR GROUPS OF ODD EXPONENT

E. I. ZEL'MANOV

Abstract. Let $n$ be an odd number and $m$ a positive integer. It is shown that there are only finitely many $m$-generator groups of exponent $n$.

Bibliography: 23 titles.

Introduction

This paper is dedicated to V. P. Platonov on his fiftieth birthday.

Burnside posed his famous problem in 1902: is the free $m$-generator group $B(m, n)$ satisfying the identical relation $x^n = 1$ finite? At the close of the thirties, the following weakened form of the Burnside conjecture having to do with finite groups was discussed in group-theoretical circles (see [1]-[4]): is it true that $B(m, n)$ has only finitely many finite homomorphic images?

In other words, the problem is about the existence of a universal finite $m$-generator group $B_0(m, n)$ of exponent $n$ such that all other finite $m$-generator groups of exponent $n$ are homomorphic images of it. Later on, the light hand of Magnus [5] was instrumental in naming the second of these questions the restricted Burnside problem (RBP).

In 1968, Adian and Novikov [6] found the negative solution to the Burnside problem for groups of odd exponent $n \geq 4381$ (this has since been done [7] for odd $n \geq 665$).

It is thus clear that RBP represents the portion of the circle of Burnside problems that is most likely to have a positive solution. Two circumstances have encouraged the hope for such a solution:

1) Reduction to groups of prime-power exponent. Hall and Higman showed in [8] that the positive solution of RBP for groups of exponent $n = p_1^{k_1} \cdots p_r^{k_r}$, where the $p_i$ are different primes, follows from that of RBP for the exponents $p_i^{k_i}$, $1 \leq i \leq r$, on the assumption that there are only finitely many finite simple groups of exponent $n$.

2) The connection between RBP for prime-power exponent $p^k$ and problems of Burnside type in Lie algebras (see [2], [3], [5], [9], and [10]): this was noted in the thirties and early forties. Let us explain this connection. Suppose that $G$ is a finite group of exponent $p^k$, and consider the lower central series

$$G = G_0 > G_1 > \cdots > G_{s+1} = \{1\}, \quad G_{i+1} = [G_i, G], \quad 0 \leq i \leq s,$$

1980 Mathematics Subject Classification (1985 Revision). Primary 20F50; Secondary 20F45.

©1991 American Mathematical Society 0025-5726/91 $1.00 + $.25 per page
together with the direct sum

\[ L = L(G) = \bigoplus_{i=0}^{s} G_i/G_{i+1} \]

of abelian groups. The operation \([a_iG_{i+1}, a_jG_{j+1}] = (a_i, b_j)G_{i+j+1}\), where \(a_i \in G_i\), \(b_j \in G_j\), and \((, )\) denotes the group commutator, gives \(L\) a Lie ring structure. It is clear that the Lie ring \(L\) has the same nilpotency class as the group \(G\).

If \(n = p\), a prime, then \(L(G)\) is a Lie algebra over the field \(Z_p\) of order \(p\) satisfying the Engel identity

\[ \text{ad}(x)^{p-1} = 0 \]  
(see [2], [5], and [9]). In this way, RBP for groups of prime exponent reduces to the following problem of Burnside type in Lie algebras: is every Lie algebra over \(Z_p\) satisfying the identity \((E_{p-1})\) locally nilpotent?

This problem was solved in the positive by Kostrikin [11], [12], and RBP for prime exponent was thereby resolved (see also the book [13], where RBP is discussed in detail).

**Theorem (A. I. Kostrikin).** Every Lie algebra over \(Z_p\) satisfying \((E_{p-1})\) is locally nilpotent.

The ideas and methods of [11] play a fundamental part in this paper, and we shall therefore make frequent reference to them.

As regards finite groups \(G\) of prime-power exponent \(p^k\), the associated Lie rings do not in general satisfy the Engel identity \((E_{p^k-1})\)—this was proved in [14]. However, they do have two important properties, neither of which on its own is enough to ensure local nilpotency:

1) \(L(G)/pL(G)\) satisfies the linearized identity \((E_{p^k-1})\) [15]; that is, for arbitrary elements \(a_1, \ldots, a_{p^k-1}\) in \(L(G)\), the following inclusion is satisfied:

\[ L(G) \sum_{\sigma \in S_{p^k-1}} \text{ad}(a_{\sigma(1)}) \cdots \text{ad}(a_{\sigma(p^k-1)}) \subseteq pL(G) \]

2) For every commutator \(\rho\) in the elements \(gG_1, g \in G, L(G)\) satisfies the equality

\[ \text{ad}(\rho)^{kp^k-1} = 0 \]

(see [9]).

The following theorem is the main result of this paper. It was stated as a problem by the author in [16] and included in [13].

**Theorem 1.** Let \(L\) be a Lie algebra over \(Z_p\) generated by elements \(x_1, \ldots, x_m\). Assume further that

1) \(L\) satisfies the linearized Engel identity \((E_n)\), and

2) there exists a natural number \(s\) such that \(\text{ad}(\rho)^s = 0\) for every commutator \(\rho\) in the generators \(x_1, \ldots, x_m\).

Then \(L\) is nilpotent.

**Corollary.** Every Lie ring satisfying an Engel identity is locally nilpotent.

From Theorem 1 and the results of Higman and Sanov quoted above, it follows that there exists a function of \(m, p, \) and \(k\) which is an upper bound for the nilpotency
class of the quotient algebra \( L(G)/pL(G) \), where \( G \) is any finite \( m \)-generator group of exponent \( p^k \). It is not hard to deduce from this (see [17], Lemma 2) that the nilpotency class of \( L(G) \) is also bounded above in terms of \( m, p, \) and \( k \). This gives the positive solution of RBP for prime-power exponents \( p^k \) with \( p \neq 2 \). Since every finite group of odd exponent is solvable [18], our second main theorem follows from this and the Hall-Higman factorization theorem:

**Theorem 2.** The restricted Burnside problem for groups of odd exponent has a positive solution.

§1. Idea and plan of the proof

Everywhere below, we shall consider algebras over fields of characteristic \( p \neq 2 \). As usual, \( \text{ad}(x) \) denotes the operator of commutation with \( x \), and

\[
[x_0, x_1, \ldots, x_m] = \cdots [[x_0, x_1], x_2], \ldots, x_m] = x_0 \text{ad}(x_1) \cdots \text{ad}(x_m)
\]

is the left normed commutator of the elements \( x_0, x_1, \ldots, x_m \). When \( x_1 = \cdots = x_m = x \), we shall write this simply as \([x_0 x^m]\).

Consider the \( \mathbb{Z}_p \)-algebra \( K \) on generators \( e_i, i = 1, 2, \ldots \), and relations \( e_i^2 = 0 \) and \( e_i e_j = e_j e_i \) for \( i, j = 1, 2, \ldots \). If \( L \) is an algebra satisfying conditions 1) and 2) of Theorem 1, then \( \tilde{L} = L \otimes \mathbb{Z}_p K \) satisfies the Engel identity \((E_n)\).

In essence, the proof of Theorem 1 was begun in [17], and this paper is a straight continuation. It was proved in [17] that the nilpotency of \( L \) is established once one proves that every algebra in the variety \( \text{var}(\tilde{L}) \) generated by \( \tilde{L} \) is locally nilpotent.

Let \( F \) be an infinite field of characteristic \( p \). Since \( \tilde{L} \) satisfies \((E_n)\) and all its partical linearizations, it follows that the scalar extension \( \tilde{L} \otimes \mathbb{Z}_p F \) also satisfies \((E_n)\) (see [23]). Thus, without loss of generality, we shall assume that \( L \) is an algebra over an infinite field and satisfies \((E_n)\); we denote by \( K \) the \( F \)-algebra on generators \( e_i, i \geq 1 \), with relations \( e_i^2 = 0 \) and \( e_i e_j = e_j e_i \) for \( i, j \geq 1 \); \( \tilde{L} \) denotes \( L \otimes F K \).

Following Kostrikin [13], an element \( a \) of an algebra \( \mathcal{L} \) is called a sandwich if \([\mathcal{L} a^2] = 0\). A Lie algebra is said to be a sandwich algebra if it is generated by finitely many sandwiches.

A fundamental part in our paper is played by a theorem about sandwich algebras, proved recently by Kostrikin and the author.

**Theorem [19].** Every sandwich Lie algebra is locally nilpotent.

This theorem suggests the following plan for the proof of Theorem 1 as set out in [17]. Suppose that we have managed to find, for every nonzero Lie \( F \)-algebra \( L \) satisfying \((E_n)\), a polynomial \( f(x_1, \ldots, x_r) \) that is not identically zero on \( L \) and such that every element of \( f(L) \) (the set of values of \( f \) on \( L \)) is a sandwich in \( L \). We denote by \( V_f(L) \) the \( F \)-linear span of \( f(L) \). Since \( F \) is infinite, it follows (see [20], the proof of Theorem 4.2.9, for instance) that \( V_f(L) \) is an ideal of \( L \). By the theorem on sandwich algebras, \( V_f(L) \) is a nonzero locally nilpotent ideal. It is then enough to use the fact that Engel Lie algebras have locally nilpotent radicals [21], [22].

Therefore, the proof-plan in [17] comes to constructing a polynomial \( f \) with the required properties. However, the author has not succeeded in constructing such a polynomial: its existence follows a posteriori from Theorem 1. Instead, we consider
polynomials with extended signature (with the addition of "separated powers" of ad-operators); they are called *U-polynomials* below. For every nonzero $F$-algebra $L$ satisfying $(E_n)$ we construct a $U$-polynomial $f$ which is not identically zero on $\tilde{L} = L \otimes_F K$, and is such that every element of $f(\tilde{L})$ is a sandwich of $\tilde{L}$. We prove further that the complete linearization $\hat{f}$ of the $U$-polynomial $f$ is an ordinary Lie polynomial which is also not identically zero on $\tilde{L}$. Every value of $\hat{f}$ on $\tilde{L}$ is a linear combination of a fixed number of elements of $f(\tilde{L})$.

Thus, instead of a polynomial each of whose values on $L$ is a sandwich, we construct a polynomial each of whose values on $\tilde{L}$ is a linear combination of a fixed number of sandwiches. If $\mathcal{L}^F$ is a free algebra of the variety $\text{var}(\tilde{L})$, the theorem on sandwich algebras shows that the ideal $V_f(\mathcal{L}^F)$ of $\mathcal{L}^F$ is locally nilpotent. All that remains is to prove that the quotient algebra $\mathcal{L}_1 = \mathcal{L}^F / V_f(\mathcal{L}^F)$ is locally nilpotent. To do that, we construct a polynomial $\hat{f}_1$ for the algebra $\mathcal{L}_1 = \mathcal{L}_1 \otimes_F K$, etc. An analysis of the construction of our polynomials shows that this process is finite. We give the definition of $U$-polynomials in §2, and show that the complete linearizations of $U$-polynomials are ordinary polynomials. In §§3 and 4, we construct a $U$-polynomial for a nonzero Engel Lie algebra $L$ that is not identically zero on $\tilde{L}$, such that each of its values on $\tilde{L}$ is a sandwich in $\tilde{L}$.

§2. *U*-polynomials

As basis for the commutative algebra $K = F \langle e_i | e_i^2 = 0, e_i e_j = e_j e_i \rangle$ we can take the elements $e_\pi = e_{i_1} \cdots e_{i_r}$, where $\pi = \{i_1 < \cdots < i_r\}$ is an ordered multi-index. Thus, every element $a$ of the algebra $\tilde{L} = L \otimes_F K$ has a unique representation as a finite sum

$$a = \sum_{\pi} \alpha_\pi \otimes e_\pi = \sum_{\pi} \alpha_\pi,$$

where $\alpha_\pi \in L$ and $a_\pi = \alpha_\pi \otimes e_\pi$. We call (1) the canonical decomposition of $a$.

We denote by $V_i$ the ideal of $\tilde{L}$ consisting of the elements $a$ such that $a_\pi = 0$ whenever $i$ does not occur in $\pi$. It is clear that $V_i^2 = 0$ and $\tilde{L} = \sum_1^\infty V_i$.

Let $\mathfrak{A} = \{a_i, i \in I\}$ be a finite family of elements of $\tilde{L}$ such that (i) every element $a_i$ in $\mathfrak{A}$ lies in one of the ideals $V_j$, and (ii) $[a_i, a_j] = 0$ for every pair $a_i, a_j$ of elements of $\mathfrak{A}$.

We consider the following linear operator on $L$:

$$U_k(\mathfrak{A}) = \sum \text{ad}(a_{i_1}) \cdots \text{ad}(a_{i_k}),$$

where the summation extends over all $k$-element choices $i_1, \ldots, i_k$ from $I$. If $i_1, \ldots, i_k$ contains a repeat, the corresponding summand is zero, because of property (i). It is clear that

$$U_1(\mathfrak{A}) = \text{ad} \left( \sum_{i \in I} a_i \right).$$

We further set $U_0(\mathfrak{A}) = \text{Id}$, the identity operator. If the characteristic of the ground field exceeds $k$, we have

$$U_k(\mathfrak{A}) = \frac{1}{k!} \text{ad} \left( \sum_{i \in I} a_i \right)^k,$$
so that the operators $U_k$ play the role of separated powers of the operators $\text{ad}$.

In what follows, we shall write linear operators to the right of the elements they act on.

We shall define a set of words for a Lie algebra $L$ in the alphabet $x_i$, $U_k$, $(,)$, $\text{ad}$, $[,]$, which we call $U$-words relative to $L$. If $L$ is fixed, we shall speak simply of $U$-words. The fact that no letter $x_i$ other than $x_1, \ldots, x_r$ occurs in the expression for a $U$-word $W$ will be expressed by writing $W = W(x_1, \ldots, x_r)$. If the subject variables $x_i$ are associated with values $a_i$ of them in $L$, the $U$-word $W$ is associated with the linear operator $W(a_1, \ldots, a_r)$ acting on $L$.

By definition, we make the following four assumptions.

I) If $p, \ldots, s$ are commutators in $x_1, \ldots, x_r$ with arbitrary disposition of brackets, the word $W = \text{ad}(p) \cdot \ldots \cdot \text{ad}(s)$ is a $U$-word. If the letters $x_1, \ldots, x_r$ are associated with elements $a_1, \ldots, a_r$, the corresponding word $W$ is associated with the linear operator $W(a_1, \ldots, a_r) = \text{ad}(p)(a_1, \ldots, a_r) \cdot \ldots \cdot \text{ad}(s)(a_1, \ldots, a_r)$.

II) If $W = W(a_1, \ldots, a_r)$ is a U-word in which the letters $x_i$, $\ldots$, $x_i$ do not occur, then $W' = \text{ad}(x_i) \cdot \ldots \cdot \text{ad}(x_i)$ is a $U$-word. If the letters $x_1, \ldots, x_r$ and $x_i, \ldots, x_i$ are associated with elements $a_1, \ldots, a_r$ and $a_i, \ldots, a_i$ respectively, then $W'$ is associated with the linear operator $\text{ad}(a_i) \cdot \ldots \cdot \text{ad}(a_i)$.

III) Assume that $W = W(a_1, \ldots, a_r)$ is a $U$-word such that, for arbitrary elements $a$, $b$, and $a_1, \ldots, a_r$ of $\tilde{L}$,

$$[aW(a_1, \ldots, a_r), bW(a_1, \ldots, a_r)] = 0,$$

and $x_i$ does not occur in the expression for $W$. Then $W' = U_k(x_iW)$ is a $U$-word. If the letters $x_1, \ldots, x_r$ and $x_i$ are associated with elements $a_1, \ldots, a_r$ and $a$ respectively, and $a = \sum a_n$ is the canonical decomposition of $a$, then the set $\mathfrak{A} = \{a_nW(a_1, \ldots, a_r)\}$ satisfies conditions (i) and (ii). By definition, the linear operator $U_k(\mathfrak{A})$ is a value of $W'$.

IV) A word is a $U$-word relative to $L$ if it is a $U$-word because of conditions I)-III).

By a $U$-polynomial relative to $L$ we understand a system consisting of three objects: a homogeneous Lie polynomial $f(t_i; y_j; z_k)$ whose variables are divided into three disjoint subsets $T = \{t_i\}$, $Y = \{y_j\}$, $Z = \{z_k\}$; and two $U$-words $W = W(x_1, \ldots, x_r)$ and $W' = W'(x_{r+1}, \ldots, x_{r+s})$ without common variables. If the variables $t_i, y_j, z_k$ are given values $a_i, b_j, c_k$ respectively in $\tilde{L}$, and $x_1, \ldots, x_{r+s}$ are given values $d_1, \ldots, d_{r+s}$ in $\tilde{L}$, then the following element is called a value of the $U$-polynomial $(f(t_i, y_j, z_k); W; W')$:

$$f(a_iW(d_1, \ldots, d_r), b_jW'(d_{r+1}, \ldots, d_{r+s}), c_k).$$

In what follows, we occasionally use the notation $f(t_iW, y_jW', z_k)$ for a $U$-polynomial. If the value of the $U$-polynomial is zero for all values of its arguments in $\tilde{L}$, we write $f(t_iW, y_jW', z_k) \equiv 0$. 
Assume that the $U$-word $W = W(x_1, \ldots, x_r)$ satisfies condition (4). It can be verified immediately that the following equality holds for arbitrary elements $b', b''$, and $a_1, \ldots, a_r$ of $\tilde{L}$:

$$U_k((b' + b'')W(a_1, \ldots, a_r)) = \sum_{i=0}^{k} U_i(b' W(a_1, \ldots, a_r))U_{k-i}(b'' W(a_1, \ldots, a_r)). \tag{5}$$

For $k < p$, (5) becomes Newton's binomial formula.

We assume next that the Lie polynomial $f(t_1, \ldots, t_m, \ast)$ is multilinear in $t_1, \ldots, t_m$, and fix values of the remaining variables. Suppose that $W(x_1, \ldots, x_r)$ is a $U$-word satisfying (4). Then, on repeated application of (5), the expression

$$f(c_\alpha, b_\beta, a_\gamma, \ast) = f\left(c_1 U_k \left(\sum_{i=1}^{m} b_i W(a_1, \ldots, a_r)\right), \ldots, c_m U_k \left(\sum_{i=1}^{m} b_i W(a_1, \ldots, a_r), \ast\right)\right),$$

can be represented as a sum of expressions of the form

$$f(c_1 U_{a_{i_1}} (b_{\beta_{i_1}} W(a_1, \ldots, a_r)) \ldots U_{a_{i_k}} (b_{\beta_{i_k}} W(a_1, \ldots, a_r)), \ldots, c_m U_{a_{m_1}} (b_{\beta_{m_1}} W(a_1, \ldots, a_r)) \ldots U_{a_{m_k}} (b_{\beta_{m_k}} W(a_1, \ldots, a_r), \ast), \tag{6}$$

where $0 \leq \alpha_{ij} \leq k$, $1 \leq \beta_{ij} \leq mk$, and $\sum_{j=1}^{k} \alpha_{ij} = k$, $1 \leq i \leq m$.

We group summands in (6) according to the degrees that $b_1, \ldots, b_{mk}$ occur in them; set

$$f\left(c_\alpha U_k \left(\sum_{i=1}^{mk} b_i W(a_1, \ldots, a_r)\right), \ast\right) = f' + f'' \tag{7}$$

where $f''$ is a sum of expressions like (6), multilinear in $b_1, \ldots, b_{mk}$, and $f'$ is the sum of those expressions of the form (6) in which at least one of the elements $b_1, \ldots, b_{mk}$ occurs with degree 2 or greater. In the expressions of the form (6) occurring in $f''$, all the $\alpha_{ij}$ are 0 or 1; by (3), they are of the form

$$f([c_1, b_{i_1} W(a_1, \ldots, a_r)], \ldots, b_{i_k} W(a_1, \ldots, a_r)]), \ldots, [c_m, b_{(m-1)k+1}} W(a_1, \ldots, a_r), \ldots, b_{i_{mk}} W(a_1, \ldots, a_r)]). \tag{8}$$

Standard arguments using Vandermonde determinants ($F$ is infinite!) show that $f''$ is a linear combination of not more than $(mk)^{mk}$ elements

$$f(c_\alpha, d_1 b_1, \ldots, d_{mk} b_{mk}, a_\gamma, \ast),$$

where the $d_\beta$ are different scalars in $F$.

If $b = b_1 + \cdots + b_{mk}$ is the canonical decomposition of an element $b$ of $\tilde{L}$, then for arbitrary $i$, $1 \leq i \leq mk$, it is clear that $U_j(b_i W(a_1, \ldots, a_r)) = 0$ for $j \geq 2$, so that

$$f'(c_\alpha, b_1, \ldots, b_{mk}, a_\gamma, \ast) = 0$$

and

$$f(c_\alpha U_k (b W(a_1, \ldots, a_r), \ast) = f''(c_\alpha, b_1, \ldots, b_{mk}, a_\gamma, \ast).$$
We show next that if the canonical decomposition of an element \( b \) contains more than \( mk \) nonzero summands, \( b = \sum_{i} b_i, s > mk \), then
\[
f(c_{\alpha} U_k(b W(a_1, \ldots, a_r)), *) = \sum f''(c_{\alpha}, b_{i_1}, \ldots, b_{i_{mk}}, a_{\gamma}, *),
\]
where summation is over all \( mk \)-element choices from \( 1, \ldots, s \).

Let \( b_1, \ldots, b_s \) be any elements of \( L \). As above, we decompose the expression
\[
f \left( c_{\alpha} U_k \left( \left( \sum_{i=1}^{s} b_i \right) W(a_1, \ldots, a_r) \right), * \right)
\]
as a linear combination of elements of the form (6), and group the latter according to the degrees that \( b_i, \ldots, b_r \) occur in them. We have
\[
f \left( c_{\alpha} U_k \left( \left( \sum_{i=1}^{s} b_i \right) W(a_1, \ldots, a_r) \right), * \right) = F' + \sum F_{i_1, \ldots, i_{mk}}(c_{\alpha}, b_{i_1}, \ldots, b_{i_{mk}}, a_{\gamma}, *)
\]
where \( F_{i_1, \ldots, i_{mk}}(c_{\alpha}, b_{i_1}, \ldots, b_{i_{mk}}, a_{\gamma}, *) \) is the component having degree 1 in \( b_{i_1}, \ldots, b_{i_{mk}} \), the summation is over all \( mk \)-element choices from \( 1, \ldots, s \), and the terms that are nonlinear in one of the \( b_{i}, 1 \leq i \leq s \), are gathered together in \( F' \). Setting \( b_j = 0 \) for \( j \not\in \{i_1, \ldots, i_{mk}\} \), we get
\[
f \left( c_{\alpha} U_k \left( \left( \sum_{\mu=1}^{mk} b_{i_{\mu}} \right) W(a_1, \ldots, a_r) \right), * \right)
\]
\[= F'_{i_1, \ldots, i_{mk}} + F_{i_1, \ldots, i_{mk}}(c_{\alpha}, b_{i_1}, \ldots, b_{i_{mk}}, a_{\gamma}, *)
\]
where \( F'_{i_1, \ldots, i_{mk}} \) is a linear combination of summands of the form (6), in each of which one of the elements \( b_{i_1}, \ldots, b_{i_{mk}} \) occurs nonlinearly. It is clear from this that, if \( b = \sum_{i} b_i \) is the canonical decomposition, then
\[
f \left( c_{\alpha} U_k \left( \left( \sum_{i=1}^{s} b_i \right) W(a_1, \ldots, a_r) \right), * \right) = \sum F_{i_1, \ldots, i_{mk}}(c_{\alpha}, b_{i_1}, \ldots, b_{i_{mk}}, a_{\gamma}, *)
\]
\[
f \left( c_{\alpha} U_k \left( \left( \sum_{\mu=1}^{mk} b_{i_{\mu}} \right) W(a_1, \ldots, a_r) \right), * \right) = F'_{i_1, \ldots, i_{mk}}(c_{\alpha}, b_{i_1}, \ldots, b_{i_{mk}}, a_{\gamma},*) = f''(c_{\alpha}, b_{i_1}, \ldots, b_{i_{mk}}, a_{\gamma}, *)
\]
This proves equality (9).

We are now ready to prove

**Proposition 1.** Let \( f(t, W, y, W', z_k) \) be a \( U \)-polynomial that is not identically zero on \( \tilde{L} \). Then there exists a multilinear Lie polynomial \( \tilde{f} \) (the linearization of \( f \)) and a natural number \( N \) such that
\[(1) \tilde{f} \) is not identically zero on \( \tilde{L} \),
\[(2) \) every value of \( \tilde{f} \) on \( \tilde{L} \) can be written as a linear combination of at most \( N \)}}
values of \( f \), and

(3) every value of the \( U \)-polynomial \( f \) can be written as a linear combination of values of \( \tilde{f} \).

**Proof.** We consider the complete linearization \( \tilde{f} \) of the polynomial \( f(t_i, y_j, z_k) \) (see [23]). Assume that the variables \( t_{i\mu}, \mu \geq 1 \), appear on linearization with respect to \( t_i \), and the groups of variables \( y_{j\mu}, \mu \geq 1 \), and \( z_{k\mu}, \mu \geq 1 \), appear on linearization with respect to \( y_j \) and \( z_k \) respectively: \( \tilde{f} = \tilde{f}(t_{i\mu}, y_{j\mu}, z_{k\mu}) \). If \( a_i = \sum \mu a_{i\mu}, b_j = \sum \mu b_{j\mu}, \) and \( c_k = \sum \mu c_{k\mu} \) are the canonical decompositions of \( a_i, b_j, \) and \( c_k \) respectively, then

\[
\tilde{f}(a_i W(d_1, \ldots, d_r), b_j W'(d_{r+1}, \ldots, d_s), c_k)
= \sum \tilde{f}(a_{i\mu} W(d_1, \ldots, d_r), b_{j\mu} W'(d_{r+1}, \ldots, d_s), c_{k\mu}),
\]

whence it follows that the \( U \)-polynomial \( \tilde{f}(t_{i\mu} W, y_{j\mu} W', z_{k\mu}) \) is not identically zero on \( \tilde{L} \). Moreover, every value of this \( U \)-polynomial is a linear combination of a fixed number of values of the \( U \)-polynomial \( f(t_i W, y_j W', z_k) \). We shall assume that the polynomial \( f(t_1, \ldots, t_m, y_j, z_k) \) is multilinear.

We establish the proposition by induction on the sum of the lengths of the \( U \)-words \( W \) and \( W' \). If they are both empty or of the form \( \text{ad}(\rho_1) \cdots \text{ad}(\rho_s) \), where the \( \rho_i \) are commutators, we are dealing with an ordinary Lie polynomial. Its complete linearization is the required polynomial \( \tilde{f} \); as we saw above, assertion 3) holds for \( \tilde{f} \), and thus \( \tilde{f} \) is not identically zero on \( \tilde{L} \).

Suppose that \( W = U_k(x_i W_1) \), where the \( U \)-word \( W_1 = W_1(x_1, \ldots, x_r) \) satisfies condition (4). We consider the \( U \)-polynomial

\[
f''(q_1 W_1(d_1, \ldots, d_r), \ldots, q_{mk} W_1(d_1, \ldots, d_r), b_j W'(d_{r+1}, \ldots, d_s), a_i, c_k),
\]

defined by (8). If \( q = \sum q_i \) is the canonical decomposition of \( q \), by (9) we have

\[
f(a_i U_k(q W_1(d_1, \ldots, d_r)), \ldots, a_m U_k(q W_1(d_1, \ldots, d_r)), b_j W'(d_{r+1}, \ldots, d_s), c_k)
= \sum f''(a_1, \ldots, a_m, q_{i_1} W_1(d_1, \ldots, d_r), \ldots, q_{i_k} W_1(d_1, \ldots, d_r), b_j W', c_k).
\]

As we remarked above, every value of the \( U \)-polynomial \( f'' \) can be written as a linear combination of not more than \( (mk)^m \) values of the \( U \)-polynomial \( f(t_i W, y_j W', z_k) \). The induction hypothesis applies to \( f'' \).

If \( W = \text{ad}(x_i W_1) \cdots \text{ad}(x_i W_1) \), the \( U \)-polynomial under consideration is obtained by the action of the operators \( W_1 \) and \( W' \) on the arguments of the Lie polynomial

\[
h(u_1, \ldots, u_k; y_j; z_k; t_i) = f([t_i, u_1, \ldots, u_k], y_j, z_k),
\]

and we can again use the induction hypothesis. This completes the proof of the proposition.

§3. Construction of a \( U \)-polynomial-sandwich

In this section we fix a nonzero Lie algebra \( L \) over an infinite field \( F \) satisfying (\( E_n \)), and construct a \( U \)-polynomial that is not identically zero on \( \tilde{L} \) such that each of its values is a sandwich of \( \tilde{L} \).
If \( \mathfrak{A} \) is a finite family of elements of \( \widetilde{L} \) satisfying conditions (i) and (ii), then
\[
U_i(\mathfrak{A}) U_j(\mathfrak{A}) = C_{i+j} U_{i+j}(\mathfrak{A}), \quad i, j \geq 0.
\]
Moreover, for an arbitrary element \( a \) of \( \widetilde{L} \) we have
\[
\text{ad}(a U_m(\mathfrak{A})) = \sum_{i=0}^{m} (-1)^i U_i(\mathfrak{A}) \text{ad}(a) U_{m-i}(\mathfrak{A}).
\]
Assume that \( U_i(\mathfrak{A}) = 0 \) for \( i \geq m \).

Applying (11) to the operators \( \text{ad}(a U_{2(m-1)}(\mathfrak{A})) \), \( \text{ad}(a U_{2m-3}(\mathfrak{A})) \), \( \text{ad}(a U_{m-1+i}(\mathfrak{A})) \), \( i \geq 1 \), and \( \text{ad}(a U_{2m-4}(\mathfrak{A})) \), respectively, we get
\[
U_{m-1}(\mathfrak{A}) \text{ad}(a) U_{m-1}(\mathfrak{A}) = 0,
\]
\[
U_{m-1}(\mathfrak{A}) \text{ad}(a) U_{m-2}(\mathfrak{A}) = U_{m-2}(\mathfrak{A}) \text{ad}(a) U_{m-1}(\mathfrak{A}),
\]
\[
U_i(\mathfrak{A}) \text{ad}(a) U_{m-1}(\mathfrak{A}) = \sum_{j=1}^{m-1} (-1)^j U_{i+j}(\mathfrak{A}) \text{ad}(a) U_{m-1-j}(\mathfrak{A}),
\]
\[
U_{m-2}(\mathfrak{A}) \text{ad}(a) U_{m-2}(\mathfrak{A}) = U_{m-1}(\mathfrak{A}) \text{ad}(a) U_{m-3}(\mathfrak{A}) + U_{m-3}(\mathfrak{A}) \text{ad}(a) U_{m-1}(\mathfrak{A}).
\]

**Lemma 1.** Suppose that \( U_i(\mathfrak{A}) = 0 \) for \( i \geq m \). Then, for arbitrary elements \( a_1, \ldots, a_r \) of \( \widetilde{L} \) and numbers \( k_0, k_1, \ldots, k_r \geq 0 \) such that \( \sum_0^r k_i > (m-2)(r+1) + 1 \), we have
\[
U_{k_0}(\mathfrak{A}) \text{ad}(a_1) U_{k_1}(\mathfrak{A}) \cdots \text{ad}(a_r) U_{k_r}(\mathfrak{A}) = 0.
\]

**Proof.** We proceed by induction on \( r \). When \( r = 0 \), we have \( k_1 > m - 1 \) by assumption, which implies that \( U_{k_1}(\mathfrak{A}) = 0 \). Suppose that \( r \geq 1 \). If \( k_0 \leq m - 2 \), the induction hypothesis gives that \( U_{k_1}(\mathfrak{A}) \text{ad}(a_2) \cdots \text{ad}(a_r) U_{k_r}(\mathfrak{A}) = 0 \). Suppose now that \( k_0 = m - 1 \). If \( k_1 \leq m - 3 \), then again by induction we have \( U_{k_2}(\mathfrak{A}) \text{ad}(a_3) \cdots \text{ad}(a_r) U_{k_r}(\mathfrak{A}) = 0 \). If \( k_1 = m - 1 \), then \( U_{k_0}(\mathfrak{A}) \text{ad}(a_1) U_{k_1}(\mathfrak{A}) = 0 \) by (12). We still have to consider the case \( k_1 = m - 2 \). By (13) we have
\[
U_{m-1}(\mathfrak{A}) \text{ad}(a_1) U_{m-2}(\mathfrak{A}) \cdots U_{k_r}(\mathfrak{A})
\]
\[
= U_{m-2}(\mathfrak{A}) \text{ad}(a_1) U_{m-1}(\mathfrak{A}) \cdots U_{k_r}(\mathfrak{A}),
\]
and this case was considered above. This proves the lemma.

**Lemma 2.** Suppose that \( U_i(\mathfrak{A}) = 0 \) for \( i \geq m, m \geq 4 \). Then, for arbitrary elements \( a_1, \ldots, a_r \) of \( \widetilde{L} \) and numbers \( k_0, \ldots, k_r \geq 0 \) such that \( \sum_0^r k_i > (m-2)(r+1) + 1 \), we have
\[
U_{k_0}(\mathfrak{A}) \text{ad}(a_1) U_{k_1}(\mathfrak{A}) \cdots \text{ad}(a_r) U_{k_r}(\mathfrak{A})
\]
\[
= \alpha U_{m-1}(\mathfrak{A}) \text{ad}(a_1) U_{m-2}(\mathfrak{A}) \cdots \text{ad}(a_r) U_{m-2}(\mathfrak{A}),
\]
where \( \alpha \) is an integer.

**Proof.** Let \( (k_0, \ldots, k_r) \) be a lexicographically maximal set of numbers for which the lemma is false. If \( k_i = 0 \) for some \( i \), then at least one of the operators
\[
U_{k_0}(\mathfrak{A}) \text{ad}(a_1) \cdots U_{k_{i-1}}(\mathfrak{A}) , \quad U_{k_{i+1}}(\mathfrak{A}) \text{ad}(a_{i+2}) \cdots U_{k_r}(\mathfrak{A})
\]

\(^1\) In general, the equation \( U_m(\mathfrak{A}) = 0 \) does not imply that \( U_i(\mathfrak{A}) = 0 \) for \( i > m \).
satisfies the conditions of Lemma 1. In fact, if \( k_0 + \cdots + k_{i-1} \leq (m-2)i + 1 \) and \( k_{i+1} + \cdots + k_r \leq (m-2)(r-i) + 1 \), then \( k_0 + \cdots + k_r \leq (m-2)r + 2 < (m-2)(r+1) + 1 \) for \( m \geq 4 \).

Suppose that \( k_0, \ldots, k_r \geq 1 \). Then, by (14) and the lexicographic maximality of the vector \((k_0, \ldots, k_r)\), all the numbers \( k_1, \ldots, k_r \) are at most \( m-2 \). This means that \( k_0 \geq (m-2)(r+1) + 1 - (m-2)r = m - 1 \), whence it follows that \( k_0 = m - 1 \). And now, since \( k_1 + \cdots + k_r = (m-2)r \), it follows that \( k_1 = \cdots = k_r = m - 2 \). This contradiction proves the lemma.

For \( k \geq n \), \( L \) satisfies the identity \( \text{ad}(a) = 0 \). Thus, for any elements \( a, b \in L \) and any \( \alpha \in F \), we have \( \text{ad}(a + \alpha b)^k = \sum_0^k \alpha^i F_i(a, b) = 0 \), where \( F_i(a, b) \) is homogeneous of degree \( i \) in \( b \) and \( k - i \) in \( a \). Since the field \( F \) is infinite, every homogeneous component \( F_i(a, b) \) is zero, and, in particular

\[
F_i(a, b) = \sum_{i=0}^{k-1} \text{ad}(a)^i \text{ad}(b) \text{ad}(a)^{k-1-i} = 0. \tag{16}
\]

The following identities come from (16) with \( k = 2n-2, 2n-3, n-1+i \) (for \( i \geq 1 \)), and \( 2n-4 \):

\[
\text{ad}(a)^{n-1} \text{ad}(b) \text{ad}(a)^{n-1} = 0; \tag{12'}
\]

\[
\text{ad}(a)^{n-2} \text{ad}(b) \text{ad}(a)^{n-2} = - \text{ad}(a)^{n-2} \text{ad}(b) \text{ad}(a)^{n-1}; \tag{13'}
\]

\[
\text{ad}(a)^i \text{ad}(b) \text{ad}(a)^{n-1} = - \sum_{j=1}^{n-1} \text{ad}(a)^{i+j} \text{ad}(b) \text{ad}(a)^{n-1-j}; \tag{14'}
\]

\[
\text{ad}(a)^{n-2} \text{ad}(b) \text{ad}(a)^{n-2} = - \text{ad}(a)^{n-1} \text{ad}(b) \text{ad}(a)^{n-3}
- \text{ad}(a)^{n-3} \text{ad}(b) \text{ad}(a)^{n-1}. \tag{15'}
\]

Lemmas 1 and 2 also have analogues for ordinary powers of the operators \( \text{ad}(a) \). We shall simply state the corresponding results, and leave to the reader the verification that the proofs are word-for-word repeats of those of Lemmas 1 and 2 (the only difference is that reference is made to (12')-(15'), not (12)-(15)).

**Lemma 1'**. For arbitrary elements \( a, a_1, \ldots, a_r \) of \( \tilde{L} \) and numbers \( k_0, k_1, \ldots, k_r \geq 0 \) such that \( \sum_0^r k_i > (n-2)(r+1) + 1 \), we have

\[
\text{ad}(a)^{k_0} \text{ad}(a_1)^{k_1} \cdots \text{ad}(a_r)^{k_r} = 0.
\]

**Lemma 2'**. Suppose that \( n \geq 4 \). Then, for arbitrary elements \( a, a_1, \ldots, a_r \) of \( \tilde{L} \) and numbers \( k_0, \ldots, k_r \geq 0 \) such that \( \sum_0^r k_i = (n-2)(r+1) + 1 \), we have

\[
\text{ad}(a)^{k_0} \text{ad}(a_1)^{k_1} \cdots \text{ad}(a_r)^{k_r} = \alpha \text{ad}(a)^{n-1} \text{ad}(a_1)^{n-2} \cdots \text{ad}(a_r)^{n-2},
\]

where \( \alpha \) is an integer.

We return now to the operators \( U_i(\mathfrak{A}) \).

**Lemma 3**. Suppose that \( U_i(\mathfrak{A}) = 0 \) for \( i \geq m, \ m \geq 4 \). The expression

\[
U_{m-1}(\mathfrak{A}) \text{ad}(a_1) U_{m-2}(\mathfrak{A}) \text{ad}(a_2) \cdots \text{ad}(a_r) U_{m-2}(\mathfrak{A})
\]


is skew-symmetric in $a_1, \ldots, a_r$ if $m$ is even, and symmetric in $a_1, \ldots, a_r$ if $m$ is odd.

**Proof.** By (13), it is enough to show that the expression $U_{m-1}(\mathfrak{a})\text{ad}(a_1)U_{m-2}(\mathfrak{a})\times \text{ad}(a_2)U_{m-2}(\mathfrak{a})$ is (skew) symmetric in $a_1$ and $a_2$. By (15), we have

$$U_{m-1}(\mathfrak{a})\text{ad}(a_1)U_{m-2}(\mathfrak{a})\text{ad}(a_2)U_{m-2}(\mathfrak{a}) = U_{m-1}(\mathfrak{a})\text{ad}(a_1)U_{m-3}(\mathfrak{a})\text{ad}(a_2)U_{m-1}(\mathfrak{a}).$$

Further, it follows from (11) that

$$\text{ad}([a_1U_{m-3}(\mathfrak{a}), a_2]) = \text{ad}(a_1)U_{m-3}(\mathfrak{a})\text{ad}(a_2) - (-1)^{m-3}\text{ad}(a_2)U_{m-3}(\mathfrak{a})\text{ad}(a_1) + \sum_{\alpha+\beta>0} \pm U_{\alpha}(\mathfrak{a})(\cdots)U_{\beta}(\mathfrak{a}).$$

From this and (12) and (10) it follows that

$$U_{m-1}(\mathfrak{a})\text{ad}(a_1)U_{m-3}(\mathfrak{a})\text{ad}(a_2)U_{m-1}(\mathfrak{a}) = (-1)^{m-3}U_{m-1}(\mathfrak{a})\text{ad}(a_2)U_{m-3}(\mathfrak{a})\text{ad}(a_1)U_{m-1}(\mathfrak{a}).$$

This completes the proof of the lemma.

Our immediate aim is to construct a $U$-word $W(x_1, \ldots, x_r)$ such that $W(a_1, \ldots, a_r) \neq 0$ for some elements $a_i \in \tilde{L}$, but $U_3(x_{r+1}W)_{\mathfrak{a}}$ is identically zero on $\tilde{L}$.

Assume that $W_1(x_i) = \text{ad}(x_i)^{n-1}$. We may assume that $W_1$ is not identically zero on $\tilde{L}$, since otherwise $\tilde{L}$ satisfies $(E_{n-1})$. Moreover, it follows from (12) that, for every $a_1$ in $\tilde{L}$,

$$[\tilde{L}W_1(a_1), \tilde{L}W_1(a_1)] = 0,$$

and by Lemma 1', the following equality holds for arbitrary $a$ and $a_1, \ldots, a_n$:

$$\text{ad}([a_1a_i\cdots a_n])\cdots \text{ad}([a_na_i\cdots a_n]) = 0.$$

Thus, the words $U_i(x_2W_1(x_1)), i \geq n$ are identically zero on $\tilde{L}$.

Suppose that we have already constructed a $U$-word $W_k(x_1, \ldots, x_r)$ that is not identically zero on $L$, satisfies condition (4) and is such that $U_i(x_{r+1}W_k) = 0$ for $i \geq m$.

If $U_3(x_{r+1}W_k) \equiv 0$, then $W_k$ is the required $U$-word. We suppose therefore that $m \geq 4$, and $U_{m-1}(x_{r+1}W_k)$ is not identically zero on $\tilde{L}$. By (12), the $U$-word $U_{m-1}(x_{r+1}W_k)$ satisfies condition (4), and Lemma 1 gives that

$$U_i(x_{r+2}U_{m-1}(x_{r+1}W_k)) \equiv 0$$

for $i \geq m$. If $U_{m-1}(x_{r+2}U_{m-1}(x_{r+1}W_k)) \equiv 0$, we set $W_{k+1} = U_{m-1}(x_{r+1}W_k)$.

Suppose that the $U$-word $U_{m-1}(x_{r+2}U_{m-1}(x_{r+1}W_k))$ is not identically zero. We set

$$W_{k,0} = \text{ad}(x_{r+2}U_{m-1}(x_{r+1}W_k)) \cdots \text{ad}(x_{r+m}U_{m-1}(x_{r+1}W_k)).$$

In view of our assumption, the $U$-word $W_{k,0}$ is not identically zero on $\tilde{L}$. By Lemma 2, for any elements $a_1, \ldots, a_{r+m}$, the operator $W_{k,0}(a_1, \ldots, a_{r+m})$ is a multiple of

$$U_{m-2}(a_{r+1}W_k(a_1, \ldots, a_r))\text{ad}(a_{r+2})\cdots U_{m-2}(a_{r+1}W_k(a_1, \ldots, a_r)) \times \text{ad}(a_{r+m})U_{m-1}(a_{r+1}W_k(a_1, \ldots, a_r)).$$
In particular,
\[
\tilde{L}W_{k,0}(a_1, \ldots, a_{r+m}) \subseteq \tilde{L}U_{m-1}(a_{r+1}W_k(a_1, \ldots, a_r)).
\] (17)

It follows from (17) and Lemma 1 that
\[
[\tilde{L}W_{k,0}(a_1, \ldots, a_{r+m}), \tilde{L}W_{k,0}(a_1, \ldots, a_{r+m})] = 0
\]
and \(U_j(x_{r+m+1}W_k,0) \equiv 0\) for \(j \geq m\).

Suppose we have already constructed a \(U\)-word \(W_{k,i}(x_1, \ldots, x_{r+m+i})\) satisfying (4) and such that \(U_j(x_{r+m+i+1}W_k, i) \equiv 0\) for \(j \geq m\). We set \(W_{k,i+1} = U_{m-1}(x_{r+m+i+1}W_k, i)\). In view of (12) once more, the \(U\)-word \(W_{k,i+1}\) satisfies condition (4), and, by Lemma 1, \(U_j(x_{r+m+i+2}W_{k, i+1}) \equiv 0\) for \(j \geq m\).

We choose arbitrary elements \(a_j \in \tilde{L}, 1 \leq j \leq r + m + i\); let \(a_{r+1} = \sum_{\pi} a_{r+1, \pi}\) be the canonical decomposition, and set \(\mathfrak{A} = \{a_{r+1, \pi}W(a_1, \ldots, a_r) | \pi\}\). By Lemma 2 (and a simple induction on \(i\)), it follows that \(W_{k,i}(a_1, \ldots, a_{r+m+1})\) is a linear combination of operators of the form
\[
U_{m-2}(\mathfrak{A})\text{ad}(c_1)U_{m-2}(\mathfrak{A})\cdots \text{ad}(c_d)U_{m-1}(\mathfrak{A}),
\] (18)
where the elements \(a_{r+2}, \ldots, a_{r+m}\) occur \((m - 1)^i\) times each among \(c_1, \ldots, c_d\), and the remaining \(c_a\) are the canonical components of the elements \(a_j\) with \(r + m < j \leq r + m + i\). If \((m - 1)^i \geq n\), by Lemma 3 the letters \(a_{r+2}\) can be brought together, with the operator (18) written in the form
\[
\cdots \text{ad}(a_{r+2}U_{m-2}(\mathfrak{A}))U_{m-1}(\mathfrak{A}) = 0.
\]

Let \(i\) be the smallest number such that the operator \(W_{k,i}\) is identically zero on \(\tilde{L}\), and set \(W_{k+1} = W_{k,i}(x_1, \ldots, x_{r+m+i})\). The word \(W_{k+1}\) satisfies condition (4), and \(U_j(x_{r+m+i+1}W_{k+1}) \equiv 0\) for \(j \geq m - 1\). At some stage we will get a \(U\)-word \(W = W_k(x_1, \ldots, x_q), k \leq n - 2\), having the required properties.

We take arbitrary elements \(a\) and \(a_1, \ldots, a_q\) in \(\tilde{L}\) and set \(b = aW_k(a_1, \ldots, a_q)\). It is clear that \(U_1(aW_k(a_1, \ldots, a_q)) = \text{ad}(b)\). Moreover, since the characteristic of the ground field is not 2, \(U_2(aW_k(a_1, \ldots, a_q)) = \text{ad}(b)^2/2\). By (11), we have for each \(u\) in \(\tilde{L}\) that
\[
\text{ad}(b)^2\text{ad}(u)\text{ad}(b) = \text{ad}(b)\text{ad}(u)\text{ad}(b)^2,
\] (19)
\[
\text{ad}(b)^3 = 6U_3(aW_k(a_1, \ldots, a_q)) = 0.
\] (20)

We could now conclude the construction of our \(U\)-polynomial-sandwich by referring to [13], §2.2. However, in the case \(\text{char} F = 3\) we would have to refer to the proof, that is, to assertions that can be extracted from [13], §2.2, but which are not stated explicitly there. Thus, for the convenience of the reader we prefer to carry through the construction right to the end, tidying up that in §2.2 of [13] in passing.

We choose two sets of elements \(a^\sigma, a_1^\sigma, \ldots, a_q^\sigma\) in \(\tilde{L}\), with \(\sigma = \pm\), and set \(b^\sigma = a^\sigma W(a_1^\sigma, \ldots, a_q^\sigma)\). Consider the subspace \(Q^+ = [\tilde{L}b^2]\) and \(Q^- = [\tilde{L}b^{-2}]\). It follows from (19) and (20) that
\[
[\tilde{L}, Q^\sigma, Q^\sigma] \subseteq Q^\sigma, \quad [Q^\sigma, Q^\sigma] = 0, \quad \sigma = \pm.
\] (21)
Lemma 4. If the relation \([x[y, x]^m y^2] = 0\) is satisfied for \(m \geq 1\) by all elements \(x \in Q^\sigma\) and \(y, y' \in Q^{-\sigma}\) (\(\sigma = \pm\)), then \([x[y, x]^{m-1} y^2]\) is a sandwich of \(\tilde{L}\) whenever \(x \in Q^\sigma\) and \(y, y' \in Q^{-\sigma},\ \sigma = \pm\).

We shall devote a separate section to the proof of this lemma.

Let \(m\) be the smallest number such that \([x[y, x]^m y^2] = 0\) holds for arbitrary elements \(a^\sigma, a_1^\sigma, \ldots, a_q^\sigma\) in \(\tilde{L}\), \(\sigma = \pm\), and all \(x \in Q^\sigma\) and \(y, y' \in Q^{-\sigma}\). If \(m \geq 1\), by Lemma 4 the \(U\)-polynomial

\([x[y, x]^{m-1} y^2]\),

is the one we want, where

\[x = [u, (a^+ W(a_1^+, \ldots, a_q^+))^2], \quad y = [v, (a^- W(a_1^-, \ldots, a_q^-))^2],\]

\[y' = [v', (a^- W(a_1^-, \ldots, a_q^-))^2].\]

Assume now that \(m = 0\); that is, \([Q^{-\sigma}, Q^\sigma, Q^\sigma] = 0\) for all \(a^\pm, a_1^\pm, \ldots, a_q^\pm\) in \(\tilde{L}\). Consider the \(U\)-polynomial \(f(u, a, a_1, \ldots, a_q) = [u(a W(a_1, \ldots, a_q))^2]\) and its linearization \(\tilde{f}\). By Proposition 1, the \(F\)-span of the set of values \(f\) is just the ideal \(V_f(\tilde{L})\) of \(\tilde{L}\), and, by the assumption just made,

\([V_f(\tilde{L}) f(u, a, a_1, \ldots, a_q^2)] = 0.\]

Lemma 5. Let \(L\) be any Lie algebra over a field \(F\) of characteristic \(\neq 2\), \(I\) an ideal of \(L\), and \(a\) and \(b\) sandwiches of the algebra \(I\). Then \([a, b]\) is a sandwich of \(L\).

Proof. We have

\([L[a, b]^2] \subseteq [L, a, b, a, b] + [L, b, a, b, a] + [L, a b^2, a] + [L, b a^2, b].\)

Clearly,

\([L, a b^2] \subseteq [I b^2] = 0, \quad [L, b a^2] \subseteq [I a^2] = 0.\)

Moreover, since \([b a^2] = 0\), it follows that

\([L, a, b, a, b] \subseteq [L a^2 b^2] + [L, b a^2, b] = 0,\]

\([L, b, a, b, a] \subseteq [L b^2 a^2] + [L, b a^2, b] = 0.\]

This completes the proof of the lemma.

By Lemma 5, every value of the \(U\)-polynomial

\[h(u, a, a_1, \ldots, a_q, u', a', a'_1, \ldots, a'_q)\]

\[= [f(u, a, a_1, \ldots, a_q), f(u', a', a'_1, \ldots, a'_q)]\]

is a sandwich of \(\tilde{L}\). If

\([f(u, a, a_1, \ldots, a_q), f(u', a', a'_1, \ldots, a'_q)] \equiv 0,\]

then \(V_f(\tilde{L})\) is an abelian ideal of \(\tilde{L}\), and so every value of \(\tilde{f}\) is a sandwich of \(\tilde{L}\).

Thus, we have constructed a \(U\)-polynomial \(h\) different from zero on \(\tilde{L}\) such that every value of \(h\) is a sandwich of \(\tilde{L}\). More than that, an analysis of the construction
shows that we have proved:

**Proposition 2.** There exist a finite sequence of multilinear polynomials \( \hat{h}_1 = 0, \ldots, \hat{h}_s \) and a natural number \( N \geq 1 \) such that, if \( L \) is a Lie algebra over a field \( F \) of characteristic \( \neq 2 \) satisfying the Engel identity \( E_n \) and \( \hat{h}_1, \ldots, \hat{h}_{i-1} \) are identically zero on \( \tilde{L} \) for \( 2 \leq i \leq s \), then every value of \( \hat{h}_i \) on \( \tilde{L} \) is a linear combination of not more than \( N \) sandwiches of \( \tilde{L} \).

Suppose that every Lie \( F \)-algebra satisfying the identities \( E_n \) and \( \hat{h}_1, \ldots, \hat{h}_i \) is locally nilpotent; we shall prove that every Lie algebra satisfying \( E_n \) and \( \hat{h}_1, \ldots, \hat{h}_{i-1} \) is likewise locally nilpotent.

As we observed above, it is proved in [17] that the local nilpotency of an algebra satisfying the identities \( (E_n) \) and \( \hat{h}_1, \ldots, \hat{h}_i \) follows from that of the variety generated by the algebra \( \tilde{L} = L \otimes_F K \). We denote a free algebra in this variety by \( L^F \), and show that \( L^F \) is locally nilpotent.

By the theorem on sandwich algebras [19], there exists a function \( S(m) \) of the natural argument \( m \geq 1 \) such that every Lie ring generated by \( m \) sandwiches is nilpotent of class at most \( S(m) \). By assumption, every value of the polynomial \( \hat{h}_i \) on \( \tilde{L} \) is a linear combination of not more than \( N \) sandwiches of \( \tilde{L} \). Therefore, any \( m \) elements in the set \( \hat{h}_i(\tilde{L}) \) of values of \( \hat{h}_i \) together generate a subalgebra of nilpotency class at most \( S(mN) \). This last assertion is equivalent to the statement that a certain system of identities holds in \( \tilde{L} \). Hence, every set of \( m \) elements in \( \hat{h}_i(L^F) \) generates a subalgebra of class at most \( S(mN) \), and the ideal \( V_{\hat{h}_i}(L^F) \) is locally nilpotent. The quotient algebra \( L^F/V_{\hat{h}_i}(L^F) \) satisfies \( (E_n) \) and \( \hat{h}_1, \ldots, \hat{h}_i \), and, by our assumption, it is also locally nilpotent. All we need do now is to refer to the existence of a locally nilpotent radical in varieties of Engel algebras [21], [22], and the local nilpotency of \( L^F \) is established.

By lowering the index \( i \) in this way, we arrive at the conclusion that every Lie algebra satisfying \( (E_n) \) is locally nilpotent. Theorems 1 and 2 now follow.

§4. Proof of Lemma 4

In this section we shall avoid the words “Jordan algebra” with a persistence worthy of better application, and restrict ourselves to elementary combinatorics. Reference to the theory of Jordan algebras would no doubt explain the meaning of the calculations to be performed below, but this would require very specialist preparation of the reader.

We establish lemma 4 in several steps:

\( 4.1 \) It follows from (21) that

\[
[\tilde{L}, Q^1, Q^2, Q^3] = 0.
\] (22)

For any element \( x = \{tb^{a^2}\} \in Q^1, t \in \tilde{L} \), we consider the canonical decomposition \( t = \sum t_i \) and the system \( \mathcal{A} = \{ \{t_i b^{a^2}\} \} \) of elements. By (22), \( U_3(\mathcal{A}) = 0 \). It follows from this and (11) that for every \( u \) in \( \tilde{L} \)

\[
\text{ad}(x)^2 \text{ad}(u) \text{ad}(x) = \text{ad}(x) \text{ad}(u) \text{ad}(x)^2.
\] (23)

\( 4.2 \) Choose arbitrary \( x \in Q^1 \) and \( y, y' \in Q^{-1} \). Our problem is to show that \( [x[y, x]^{m-1}y^2] \) is a sandwich in \( \tilde{L} \). We prove that it is enough to establish the
equality
\[ Q^{-\sigma}[x[y, x]]^m = Q^{-\sigma}, Q^{-\sigma} = 0. \] (24)

Set \( z = [x[y, x]]^m \in Q^\sigma \). Using the equality \( \text{ad}(y')^3 = \text{ad}(y')^2 \text{ad}(z) \text{ad}(y')^2 = 0 \), we have
\[
\text{ad}([zy'2]) \text{ad}([zy'2]) = \text{ad}(y')^2 \text{ad}(z)^2 \text{ad}(y')^2 - 2\text{ad}(y')^2 \text{ad}(z) \text{ad}(y') \text{ad}(z) \text{ad}(y')^2
- 2\text{ad}(y') \text{ad}(z) \text{ad}(y') \text{ad}(z) \text{ad}(y')^2
+ 4\text{ad}(y') \text{ad}(z) \text{ad}(y')^2 \text{ad}(z) \text{ad}(y')
= \text{ad}(y')^2 \text{ad}(z)^2 \text{ad}(y')^2,
\] (25)
since \( \text{ad}(y')^2 \text{ad}(z) \text{ad}(y') = \text{ad}(y') \text{ad}(z) \text{ad}(y')^2 \) by (4.1). Thus,
\[ [\mathcal{L}[zy'2]] \subseteq [Q^{-\sigma} z^2 y'^2]. \]

(4.3) We set \( x^{(i)} = [x[y, x]]^{-1} \) and \( y^{(i)} = [y[x, y]]^{-1} \) for \( i \geq 1 \), and show that
\[
x^{(2i)} = -[y x^2 (y^2 x^2) (y^2 x^2) \cdots (y^2 x^2)]^{i-1},
\]
\[
x^{(2i-1)} = [x (y^2 x^2) (y^2 x^2) \cdots (y^2 x^2)]^{i-1}.
\] (26)

For \( i = 1 \), there is nothing to prove. Let us assume that the assertion has been proved for \( x^{(k-1)} \), and prove it for \( x^{(k)} \). If \( k = 2i \), we have
\[
x^{(2i)} = [x^{(2i-1)}, [y, x]] = [xy^2 x^2 y^2 x^2 \cdots y^2 x^2, y, x]
= [x, y, xy^2 x^2 \cdots y^2 x^2] = -[yx^2 y^2 x^2 \cdots y^2 x^2]
\]
by (23). If \( k = 2i + 1 \), then
\[
x^{(2i+1)} = [x^{(2i)}, [y, x]] = -[yx^2 y^2 x^2 \cdots y^2 x^2, y, x]
= -[y, x, xy^2 x^2 \cdots y^2 x^2] = [xy^2 x^2 y^2 \cdots y^2 x^2].
\]

Similar formulas are valid for the \( y^{(k)} \).

(4.4) It follows from (25) and (4.3) that
\[ \text{ad}(x^{(i)})^2 = \text{ad}(x)^2 (\text{ad}(y)^2 \text{ad}(x)^2)^{i-1}. \] (27)

In particular,
\[ [yx^{(i)^2}] = -x^{(2i)}. \] (28)

(4.5) We claim that the following equality holds for all \( u \) in \( Q^{-\sigma} \):
\[ [u, x^{(i)}, y x^2] = [ux^2, y, x^{(i)}]. \] (29)
For \( i = 1 \), (29) follows from (23). Suppose that (29) has been proved for all \( 1 \leq i < k \); we establish it for \( k \). Suppose that \( k = 2i \). By (28), we have
\[
[u, x^{(k)}, y x^2] = -[u, [yx^{(i)^2}], y x^2] = -[u, y x^{(i)^2}, y x^2] + 2[u, x^{(i)}, y, x^{(i)}, y x^2]
- [ux^{(i)^2}, y^2 x^2] = 2[u, x^{(i)}, y, x^{(i)}, y x^2] - [ux^{(i)^2}, y^2 x^2].
\]
Similarly,
\[ [ux^2, y, x^{(k)}] = -[ux^2, y, [yx^{(i)}]] = 2[ux^2, y, x^{(i)}] - yx^2x^{(i)}]. \]

By the induction hypothesis,
\[ [u, x^{(i)}, y, x^{(i)}] = 2[ux^2, y, x^{(i)}]. \]

Further, it follows from (27) that
\[ [ux^{(i)^2}y^2x^2] = [ux^2(y^2x^2)^i] = [ux^2x^{(i)^2}]. \]

Therefore,
\[ [u, x^{(k)^2}, yx^2] = [ux^2, y, x^{(k)}]. \]

Suppose now that \( k = 2i + 1 \). We have \( x^{(i+1)} = [x^{(i)}, [y, x]] \). This means that, for all \( u \) in \( Q^{-a} \),
\[ [u, x^{(i+1)}, x^{(i)}] = [u, x^{(i)}, [y, x], x^{(i)}] = [u, [y, x]x^{(i)^2}]. \] (30)

On the other hand,
\[ [u, x^{(i+1)}, x^{(i)}] = [u, x^{(i)}, x^{(i+1)}] = [ux^{(i)^2}, [y, x]] - [u, x^{(i)}, [y, x], x^{(i)}]. \] (31)

By (27) and (23), we have
\[ [u, [y, x]x^{(i)^2}] = -[u, x, yx^2(y^2x^2)^{i-1}] = -[ux^2(y^2x^2)^{i-1}, y, x] \]
\[ = -[ux^{(i)^2}, [y, x]]. \]

It follows from this and (30) and (31) that
\[ [u, x^{(i)}, [y, x], x^{(i)}] = 0, \]
\[ [u, x^{(i+1)}, x^{(i)}] = [ux^{(i)^2}, [y, x]] = [u, [y, x]x^{(i)^2}]. \] (32)

In particular, for \( u = y \) we get \( [y, x^{(i+1)}, x^{(i)}] = -x^{(2i+1)} \). Now
\[ [u, x^{(2i+1)}, yx^2] = -[u[y, x^{(i+1)}, x^{(i)}], yx^2] \]
\[ = [u, x^{(i+1)}, y, x^{(i)}], yx^2] + [u, x^{(i)}, y, x^{(i+1)}, yx^2] \]
\[ - [u, x^{(i+1)}, x^{(i)}], yx^2] \]
\[ = -[ux^2, y, [y, x^{(i+1)}, x^{(i)}]] = [ux^2, y, x^{(i+1)}, y, x^{(i)}] \]
\[ + [ux^2, y, x^{(i)}, y, x^{(i+1)}] - [ux^2y^2, x^{(i+1)}, x^{(i)}]. \]

By the induction hypothesis,
\[ [u, x^{(i+1)}, y, x^{(i)}], yx^2] = [ux^2, y, x^{(i+1)}, y, x^{(i)}] + [ux^2, y, x^{(i)}, y, x^{(i+1)}]. \]

By (32) and (27),
\[ [u, x^{(i+1)}, x^{(i)}], yx^2] = [u, x, yx^2(y^2x^2)^i] = [ux^2(y^2x^2)^i, y, x] \]
\[ = [ux^2y^2, x^{(i+1)}, x^{(i)}]. \]

This proves (29).
(4.6) We show that, for all $u$ in $Q^{-\sigma}$,
\[ [u, x^{(i)}, yx^2] = [u, x^{(i+1)}, x]. \]
Suppose that $\alpha \in f$. It follows from (23) that
\[ [u, (\alpha x^{(i)} + x), y(\alpha x^{(i)} + x)^2] = [u(\alpha x^{(i)} + x)^2, y, \alpha x^{(i)} + x]. \]
Equating the components that are linear in $\alpha$, we get
\[ [u, x^{(i)}, yx^2] + 2[u, x, y, x^{(i)}, x] = [ux^2, y, x^{(i)}] + 2[u, x^{(i)}, x, y, x]. \]
Using (4.5) and the fact that $\text{char } F \neq 2$, we get
\[ [u, x^{(i)}, yx^2] + 2[u, x, y, x^{(i)}, x] = [ux^2, y, x^{(i)}] + 2[u, x^{(i)}, x, y, x]. \]
Using (4.5) and the similar formulas for $y^{(2i)}$ and $y^{(2i-1)}$ that
\[ x^{(j)} = -[y^{(j-1)}x^2], \quad j \geq 2. \]
(4.7) It follows form (26) and the similar formulas for $y^{(2i)}$ and $y^{(2i-1)}$ that
\[ x^{(j)} = -[y^{(j-1)}x^2], \quad j \geq 2. \]
(4.8) The following holds:
\[ [x^{(i)}, y^{(j)}] = j[x^{(i+j-1)}, y] + (j - 1)[y^{(i+j-1)}, x], \quad j \geq 1. \]
In fact, for $j = 1$ there is nothing to prove. For $j \geq 2$ we have $y^{(j)} = [y^{(j-1)}, [x^{(j-1)}, y]]$. Hence
\[ [x^{(i)}, y^{(j)}] = [x^{(i)}, [y^{(j-1)}, [x, y]]] = [x^{(i)}, y^{(j-1)}, [x, y]] - [x^{(i)}, [x, y], y^{(j-1)}] \]
\[ = (j - 1)[x^{(i+j-2)}, y, x, y] - (j - 1)[x^{(i+j-2)}, y, x, y] \]
\[ + (j - 2)[y^{(i+j-2)}x^2, y] - (j - 2)[y^{(i+j-2)}, x, y, x] + [x^{(i)}, y^{(j-1)}] \]
\[ = (j - 1)[x^{(i+j-1)}, y] + (j - 1)[y^{(i+j-1)}, x] - (j - 2)[x^{(i+j-1)}, y] \]
\[ - (j - 2)[y^{(i+j-1)}, x] + (j - 1)[x^{(i+j-1)}, y] + (j - 2)[y^{(i+j-1)}, x] \]
\[ = j[x^{(i+j-1)}, y] + (j - 1)[y^{(i+j-1)}, x]. \]
(4.9) By assumption, $[x^{(m+1)}, Q^{-\sigma}, Q^{-\sigma}] = 0$. We denote by $I$ the ideal of the algebra $Q = Q^{-\sigma} + [Q^{-\sigma}, Q^{-\sigma}] + Q^{\sigma}$ generated by $x^{(m+1)}$, and show that $[I, Q^{-\sigma}, Q^{-\sigma}] = 0$. Let $r$ be the smallest number such that there exist elements $x_1, \ldots, x_r$ in $Q^{-\sigma} \cup [Q^{-\sigma}, Q^{\sigma}] \cup Q^{\sigma}$ and $z \in Q^{-\sigma}$ such that
\[ w = [x^{(m+1)}, x_1, \ldots, x_r, z^2] \neq 0. \]
Since $r$ is minimal, for every permutation $\pi$ in $S_r$ we have
\[ [x^{(m+1)}, x_{\pi(1)}, \ldots, x_{\pi(r)}, z^2] = [x^{(m+1)}, x_1, \ldots, x_r, z^2]. \]
If at least one of the $x_i, 1 \leq i \leq r$, lies in $Q^{\sigma}$, we may assume that $x_1 \in Q^{\sigma}$, in which case $w = [x^{(m+1)}, x_1, \ldots] = 0$. If some $x_i, 1 \leq i \leq r$, lies in $Q^{-\sigma}$, we may assume that $x_r \in Q^{-\sigma}$, and then $w = [\cdots, x_r, z^2] = 0$. This means that all the $x_i$ lie
in \([Q^{-\sigma}, Q^{-\sigma}]\). Then
\[
w = [x^{(m+1)}_1, x_1, \ldots, x_{r-1} z^2, x_r] - 2[x^{(m+1)}], x_1, \ldots, x_{r-1}, [z, x_r], z] = 0
\]
by the minimality of \(r\).

(4.10) We prove that, for all \(u \in Q^{-\sigma}\),
\[
[u, x, y^{(i)} x^2] = [u, x^{(i+1)}, x].
\]
In fact,
\[
[u, x^{(i+1)}, x] = -[u, [y^{(i)} x^2], x] = -[ux^2, y^{(i)} x] + 2[u, x, y^{(i)} x^2] = [u, x, y^{(i)} x^2].
\]

(4.11) We can now complete the proof of the lemma. Our problem is to show that, for all \(u \in Q^{-\sigma}\),
\[
[u(x^{(m)}), Q^{-\sigma}, Q^{-\sigma}] = 0.
\]
Suppose first that \(m = 1\). Then
\[
[u x^2, Q^{-\sigma}, Q^{-\sigma}] = -[x, [u, x], Q^{-\sigma}, Q^{-\sigma}] = 0
\]
by assumption. Suppose now that \(m \geq 2\). By (27), we have \([ux^{(m)}] = [ux^2(y^2 x^2)^{m-1}]\). Further, by (26) and our assumption about \(m\), for every \(\alpha\) in \(F\) we have
\[
[(\alpha u + y)x^2((\alpha u + y)^2 x^2)^{m-1}, Q^{-\sigma}, Q^{-\sigma}] = -[x[\alpha u + y, x]^{2m}, Q^{-\sigma}, Q^{-\sigma}] = 0.
\]
Since \(F\) is infinite, every component of this expression homogeneous in \(\alpha\) is zero. In particular,
\[
[u x^2(y^2 x^2)^{m-1}, u, y x^2(y^2 x^2)^{m-2}] + \cdots + 2[y x^2(y^2 x^2)^{m-2}, u, y x^2], Q^{-\sigma}, Q^{-\sigma}] = 0.
\]
We claim that
\[
[y x^2(y^2 x^2)^{i-1}, u, y x^2(y^2 x^2)^{m-1-i}] \in I, \quad 1 \leq i \leq m - 1. \tag{35}
\]
Lemma 4 will then follow from (34) and (4.9). Denote congruence modulo the ideal \(I\) by \(\equiv\). Applying in turn (26), (27), (26), (25), and (4.6) we get
\[
- [yx^2(y^2 x^2)^{i-1}, u, y x^2(y^2 x^2)^{m-1-i}] = -[x^{(2i)}], u, y(x^{(m-i)})^2]
= [u, x^{(2i)}, y[y^{(m-i-1)} x^2]] = [u, x^{(2i)}, y x^2, y^{(m-i-1)}, y^{(m-i-1)} x^2]
= [u, x, x^{(2i+1)} y^{(m-i-1)}, x^{(m-i-1)} x^2] = [u, x, y^{(m-i-1)}, y^{(m-i-1)}, x^{(2i+1)} x^2]
+ [u, x, [x^{(2i+1)}, y^{(m-i-1)}], y^{(m-i-1)} x^2]
+ [u, x, y^{(m-i-1)} [x^{(2i+1)}, y^{(m-i-1)}] x^2]
= [u, x, y^{(m-i-1)} [x^{(2i+1)}, y^{(m-i-1)}] x^2]
+ [u, x, [x^{(2i+1)}, y^{(m-i-1)}], y^{(m-i-1)}] x^2
\]
since \([u, x, y^{(m-i-1)}, y^{(m-i-1)}] = 0\). By (4.8),
\[
[u, x, y^{(m-i-1)}, [x^{(2i+1)}, y^{(m-i-1)}] x^2]
= (m - i - 1)[u, x, y^{(m-i-1)}, [x^{(m+i-1)}, y] x^2]
+ (m - i - 2)[u, x, y^{(m-i-1)}, [y^{(m-i-1)}, x] x^2].
\]
It now follows from (4.6) that
\[
[u, x, y^{(m-i-1)}, [x^{(m+i+1)}, y]x^2] = [u, x, y^{(m-i-1)}, x^{(m+i-1)}, yx^2] \\
= [u, x, y^{(m-i-1)}, x^{(m+i)}, x] \in I,
\]
and from (4.10) that
\[
[u, x, y^{(m-i-1)}, [y^{(m+i-1)}, x]x^2] = -[u, x, y^{(m-i-1)}, x, y^{(m+i-1)}x^2] \\
= -[u, x, y^{(m-i-1)}, x^{(m+i)}, x] \in I.
\]
Finally,
\[
[u, x, [x^{(2i+1)}, y^{(m-i-1)}, y^{(m-i-1)}]x^2] = -[u, x, y^{(2m-2)}x^2] = -[u, x^{(2m-1)}, x]
\]
again by (4.10). Since \( m \geq 2 \), we have \( 2m - 1 \geq m + 1 \) and \( [u, x^{(2m-1)}, x] \in I \), and this establishes the inclusion (35).

**BIBLIOGRAPHY**


Translated by J. WIEGOLD

---