EXTREMAL ELEMENTS IN LIE ALGEBRAS

ARJEH M. COHEN

Abstract. This is the content of a talk at the seminario de Álgebra de la Facultad de Ciencias of the Universidad de Málaga on October 6, 2014.

The main result discussed in this lecture is an elementary proof of the following theorem: If \( L \) is a simple Lie algebra over \( F \) of characteristic distinct from 2 and 3 having an extremal element that is not a sandwich, then either \( F \) has characteristic 5 and \( L \) is isomorphic to the 5-dimensional Witt algebra \( W_{1,1}(5) \), or \( L \) is generated by extremal elements.

We will also pay attention to the following theorem: If \( L \) is a simple Lie algebra generated by extremal elements that are not sandwiches, then it is classical, i.e., essentially a Lie algebra of Chevalley type. This result, of which various geometric proofs are emerging (mainly thanks to Cuypers, Fleischmann, Roberts, and Shpectorov), gives a new proof of the classification of classical simple Lie algebras of characteristic distinct from 2 and 3.

This is joint work with Gábor Ivanyos and Dan Roozemond. For the full paper, see [7]

1. Introduction

1. Definitions

Throughout the talk, let \( L \) be a finite-dimensional Lie algebra over a field \( F \) of characteristic \( p \) distinct from 2 and 3.

- An element \( x \in L \) is said to be extremal if \([x,[x,L]] \subseteq Fx\).
- If \([x,[x,L]] = 0\) we say \( x \) is a sandwich.

Let \( E = E(L) \) be the set of nonzero extremal elements of \( L \).

2. Four Examples

(a) A nilpotent element in \( \mathfrak{sl}_2(F) \) (non-sandwich) or a non-central element in the Heisenberg algebra (sandwich).

(b) A long root element \( X_\alpha \) in a ‘classical’ Lie algebra.

\[
[X_\alpha,[X_\alpha,X_\beta]] \subseteq FX_{2\alpha+\beta} \subseteq FX_\alpha
\]

\[
[X_\alpha,[X_\alpha,H]] \subseteq FX_\alpha = \{0\}
\]

(c) The Witt algebra \( W_{1,1}(5) \) over \( F \) of char 5 of dim 5 can be defined as follows. As a vector space \( W_{1,1}(5) \) has basis \( z^i \partial_z \), for \( i = 0, \ldots, 4 \). The Lie bracket is defined on two of these elements by

\[
[z^i \partial_z,z^j \partial_z] := (j-i)z^{i+j-1} \partial_z,
\]

with the convention that

\[
z^k := 0 \text{ whenever } k \notin \{0,\ldots,4\}.
\]

Now \( v := 2^4 \partial_z \) is a sandwich and \( x = z^2 \partial_z \) is extremal but not a sandwich in \( W \). Together with \( y = \partial_z \), the extremal element \( x \) forms an \( \mathfrak{sl}_2 \)-pair in \( W \). Moreover, \([v,y] = 2z^5 \partial_z \), so \( W \) is generated by \( x,y,v \). But \([y,[y,v]] = x \), so \( y \) is not extremal in \( W \).
We construct an extension \( \widetilde{W}_{1,1}(5) \) of \( W_{1,1}(5) \) by adding a basis element, namely \( z^6 \partial_z \), and adapting (c):

\[ z^k := 0 \text{ whenever } k \notin \{0, 1, 2, 3, 4, 6\}. \]

The only entry of the multiplication table that differs between \( \widetilde{W}_{1,1}(5) \) and \( W_{1,1}(5) \) is \([z^3 \partial_z, z^4 \partial_z] \): This is 0 in \( W_{1,1}(5) \) and \( z^6 \partial_z \) in \( \widetilde{W}_{1,1}(5) \). Furthermore, \( \widetilde{W}_{1,1}(5) \) is an extension of \( W_{1,1}(5) \) by the 1-dimensional center \( F z^6 \partial_z \). This extension was constructed by Block in [3]. A similar construction over the complex numbers of the infinite-dimensional Weyl algebra is also known as the Virasoro algebra. The elements \( x, y \), and \( v = 2z^4 \partial_z \) generate \( \widetilde{W}_{1,1}(5) \) and the element \( z^2 \partial_z \) is again extremal and no sandwich, but \( v \) is no longer a sandwich as \([v, [v, y]] = z^0 \partial_z \).

3. **Theorem** (Premet, [10, 11]) If \( F = \mathbb{F} \) with \( p \neq 2, 3, \) and \( L \) is simple, then \( E \neq \emptyset \).

A self-contained proof for \( p \neq 2, 3, 5 \) is in [13].

4. **Theorem** (AMC, Ivanyos, Roozemond, [7]) If \( p \neq 2, 3 \) and \( L \) is simple and has an extremal element that is not a sandwich, then
   - either \( p = 5 \) and \( L \) is isomorphic to \( W_{1,1}(5) \),
   - or \( L \) is generated by \( E \).

New on the following counts:
   - the field need not be algebraically closed;
   - characteristic 5 is included;
   - the proof is elementary.

We will discuss the proof of this theorem, but first some definitions and a discussion of its significance.

5. **Basic properties of extremal elements**

Let \( x \in E \).

(i) There is a map \( g_x : L \rightarrow \mathbb{F} \) such that

\[ [x, [x, y]] = 2g_x(y)x, \]

\[ [[x, y], [x, z]] = g_x([y, z])x + g_x(z)[x, y] - g_x(y)[x, z], \]

and

\[ [x, [y, [x, z]]] = g_x([y, z])x - g_x(z)[x, y] - g_x(y)[x, z] \]

hold for every \( y, z \in L \).

These are the **Premet identities** and are used for a definition of extremity of \( p = 2 \).

(ii) For \( t \in \mathbb{F} \) define the map \( \exp(x, t) : L \rightarrow L \) by

\[ \exp(x, t)y = y + t[x, y] + t^2 g_x(y)x. \]

Then \( \exp(x, t) \in \text{Aut}(L) \).

6. **Target Theorem** If \( L = \langle E \rangle \) is simple, then it is of classical type.

Three proofs.

a. There is an algebraic proof of this fact for \( p > 5 \). It uses the classification of finite-dimensional simple Lie algebras as described in Premet, Strade, Benkart, Block, Kostrikin, et al. See [1, 12].
b. Using the automorphisms above, it can be shown that \( \text{Aut}(L) \) contains a normal subgroup \( G \) generated by abstract root subgroups in the sense of Timmesfeld. (We do not need his classification.) \( G \) is an algebraic group with Lie algebra \( L(G) = L \). So the classification of these algebras is brought back to the classification of algebraic groups, which depends on the work of Killing! (NOT a circular argument!) More details are in [4].

c. The classification can also be derived from geometric arguments using the theory of spherical buildings up to small rank cases. It uses the result by Cuypers and Fleischmann [8] that the building determines a unique Lie algebra generated by extremal elements up to isomorphism.

7. Some more details on the geometric proof Let \( \mathcal{E} \) be the set of projective points corresponding to members of \( E \). For \( i \in \{-2, -1, 0, 1, 2\} \), define the relation \( \mathcal{E}_i \) on \( \mathcal{E} \) as follows, where \( x, y \in E \).

\(-2\) \( \mathcal{E}_{-2} \) is equality.
\(-1\) \( \mathcal{E}_{-1} \) is defined by \((F \cdot x, F \cdot y) \in \mathcal{E}_{-1}\) if and only if \( \dim((x, y)) = 2 \) and \([x, y] = 0\), and, for every \( z \in L \),
\[
[x, [y, z]] = g_y(z)x + g_x(z)y
\]
(in other words: each projective point of \((x, y)\) is in \( \mathcal{E} \)). We write \( \mathcal{F} \) for the collection of projective lines of the form \((x, y)\) with \((F \cdot x, F \cdot y) \in \mathcal{E}_{-1}\).

\(0\) \( \mathcal{E}_0 \) stands for commuting, but not in \( \mathcal{E}_{-2} \cup \mathcal{E}_{-1} \).
\(1\) \( \mathcal{E}_1 \) is defined by \((F \cdot x, F \cdot y) \in \mathcal{E}_1\) if and only if \( g_x(y) = 0 \) and \([x, y] \neq 0\).
\(2\) \( \mathcal{E}_2 \) consists of all pairs \((F \cdot x, F \cdot y) \in \mathcal{E} \times \mathcal{E}\) with \([x, y] \neq 0\).

Some terminology:
- A space is a pair consisting of a set \( X \) of points and a collection of subsets of \( X \) of size at least two, the lines of the space.
- A space is said to be partial linear if each pair of points lies on at most one line.
- A subspace of a space is a subset of the space such that every line having two distinct points from the subset lies entirely in the subset.
- A subspace is called a hyperplane if every line meets it.
- The space is called connected if its collinearity graph is connected.

Proposition The pair \((\mathcal{E}, \mathcal{F})\) is a partial linear space and \(\{\mathcal{E}_i\}_{-2 \leq i \leq 2}\) is a quintuple of disjoint symmetric relations partitioning \( \mathcal{E} \times \mathcal{E} \) such that

(A) The relation \( \mathcal{E}_{-2} \) is equality on \( \mathcal{E} \).
(B) The relation \( \mathcal{E}_{-1} \) is collinearity of distinct points of \( \mathcal{E} \).
(C) There is a map \( \mathcal{E}_1 \rightarrow \mathcal{E}_1 \), denoted by \((u, v) \mapsto [u, v]\) such that, if \((u, v) \in \mathcal{E}_1\) and \( x \in \mathcal{E}_i(u) \cap \mathcal{E}_j(v) \), then \([u, v] \in \mathcal{E}_{i+j}(x)\).
(D) For each \((x, y) \in \mathcal{E}_2\), we have \( \mathcal{E}_{x0}(x) \cap \mathcal{E}_{y0}(y) = \emptyset \).
(E) For each \( x \in \mathcal{E} \), the subsets \( \mathcal{E}_{x-1}(x) \) and \( \mathcal{E}_{x0}(x) \) are subspaces of \((\mathcal{E}, \mathcal{F})\).
(F) For each \( x \in \mathcal{E} \), the subset \( \mathcal{E}_{x1}(x) \) is a hyperplane of \((\mathcal{E}, \mathcal{F})\).

We call a space \((\mathcal{E}, \mathcal{F})\) with the above properties a root filtration space. Moreover, \((\mathcal{E}, \mathcal{F})\) is called non-degenerate if the following two conditions are satisfied.

(G) For each \( x \in \mathcal{E} \) the set \( \mathcal{E}_2(x) \) is not empty.
(H) The graph \((\mathcal{E}, \mathcal{E}_{-1})\) is connected.

Theorem (see [5, 6]) Let \((\mathcal{E}, \mathcal{F})\) be a non-degenerate root filtration space. If the singular rank of \((\mathcal{E}, \mathcal{F})\) is finite, then \((\mathcal{E}, \mathcal{F})\) is isomorphic to a shadow space of type \( A_{n,1,n} \) \((n \geq 2)\), \( B(C)_{n,2} \) \((n \geq 3)\), \( D_{n,2} \) \((n \geq 4)\), \( E_{6,2}, E_{7,1}, E_{8,8}, F_{4,1} \), or
$G_{2,2}$. Moreover, $(E,F)$ has a universal embedding in a projective space, which is polarized.

Here,

- the Bourbaki labeling used for the Dynkin diagrams;
- an embedding of a root filtration space $S$ in a projective space is an injective map from points of the geometry to points of the projective space such that lines of the space $S$ are mapped onto lines of the projective space;
- it is universal if every other embedding can be obtained by factorization through it;
- it is polarized if, for each point $x$ of $S$, the image of the subspace $E_{\leq 1}(x)$ is contained in a hyperplane of the projective space.

We swept the hyperbolic polar spaces (case without lines) under the rug; these can be handled by theorems of Cuypers.

We obtained a map

\[
\{\text{simple finite-dimensional Lie algebras without sandwiches}\} \downarrow \{\text{shadow spaces of spherical buildings on neighbors of highest root, with a universal embedding in a Lie algebra}\}
\]

The fact that the embedding of $E(L)$ is polarized, makes it possible to prove

**Theorem** (Fleischmann’s PhD thesis [8]) *The Lie algebra structure of $L$ is uniquely determined by the geometry $E(L)$ embedded in $\mathbb{P}(L)$.*

Conclusion: This completes the classification of simple Lie algebras with Lie rank at least three, $p \neq 2, 3$ with $E$ having strongly commuting elements.

So far for a kind of geometric revisionism of the classification of Lie algebras.

2. **The proof of Theorem 4.**

Recall that $p$ is distinct from 2 and 3 and let $x \in E$. We show that $L$ is generated by $E$. The proof is in six steps. Some of these were found by experiments with the GAP computer system package GBNP.

1. **Jacobson-Morozov** If $w$ is an element for which $g_x(w) = -1$, then, with $h = [x,w]$, there is $y \in L$ for which

\[
[x,y] = h, \quad [h,x] = 2x, \quad \text{and} \quad [h,y] = -2y.
\]

The elements $x, y$ are the usual nilpotent generators of the Lie algebra $\mathfrak{sl}_2(\mathbb{F})$ of $2 \times 2$ matrices of trace 0 over $\mathbb{F}$; such a pair is called an $\mathfrak{sl}_2$-pair.

The proof is much like the original.

In the remainder of this section we suppose that $x \in E, y \in L$ is an $\mathfrak{sl}_2$-pair and write $h = [x,y]$ and $S = \langle x, y, [x,y] \rangle$.

2. **Quadratic action** $\text{ad}_y$ acts quadratically on $L/S$, i.e., $\text{ad}_y^2(L/S) = 0$.

**Proof** Consider $L$ as a module on which $S$ acts. Obviously $S$ is an invariant subspace, so $L/S$ is an $S$-module. Write $X, Y$ for the action of $\text{ad}_x, \text{ad}_y$, respectively,
on \( L/S \). As \( \text{ad}_Y^2(L) \subseteq \mathbb{F}x \subseteq S \), we have \( X^2 = 0 \). We list the known relations for \( X \) and \( Y \), and the quadraticity of \( X \) that we just found.

\[
\begin{align*}
(R1) & \quad X^2Y - 2XYX + YX^2 + 2X = 0 \\
(R2) & \quad -XY^2 + 2YXY - Y^2X - 2Y = 0 \\
(R3) & \quad X^2 = 0
\end{align*}
\]

The relations \( (R1) \) and \( (R3) \) immediately imply

\[
XYX = X.
\]

Multiplying \( (R2) \) from the left by \( X \) gives

\[
-X^2Y^2 + 2XYXY - XY^2X - 2XY = 0,
\]

which, after application of \( (R3) \) and \( (R4) \), gives

\[
XY^2X = 0.
\]

Denote by \( R_2 \) the left hand side of \( (R2) \). Then, by \( (R3) \),

\[
0 = YR_2YX - YXYR_2 + 2Y^2XR_2 - R_2YXY + XYR_2Y - 3YR_2 \\
-2YXR_2Y + 3R_2Y - 2YXR_2Y - 6R_2Y + 2XR_2Y^2 \\
= 12Y^2 - 3XY^3 + 7YXY^2 + 5Y^2XY - 2XY^3X + 3XYXY^3 \\
-7YXYXY^2 + 5Y^2XYXY - Y^3XYX.
\]

Replacing \( XYX \) by \( X \) and \( X^2 \) by 0, using \( (R4) \) and \( (R3) \), we find

\[
0 = 12Y^2 - 3XY^3 + 7YXY^2 - 5Y^2XY + Y^3X + 3XY^3 \\
-7YXY^2 + 5Y^2XY - Y^3X \\
= 12Y^2.
\]

As \( p \neq 2, 3 \), we conclude that \( Y^2 = 0 \). 

\[ \square \]

3. A grading Denote by \( L_\lambda \) the \( \lambda \)-eigenspace of \( -\text{ad}_h \) in \( L \).

Suppose that \( x \) is not a sandwich. Then \( \text{ad}_h \) is diagonalizable with eigenvalues 0, ±1, ±2 and satisfies \( L_{-2} = \mathbb{F}x \) and \( L_2 = \mathbb{F}y \).

**Proof** On \( L/S \), the linear map \( \text{ad}_h \) has minimal polynomial \( \lambda^2 - \lambda \). (Indeed \( (XY - YX)^3 = XYXYXY - YXYXYX = XYXY - YXYX \).

We exploit the \( \text{ad}_h \)-grading with five components.

4. The dichotomy Either \( p = 5 \) and \( [y, [y, v]] = x \) for some \( v \in L_{-1} \), or \( y \in E \), the \( L_i \) give a \( \mathbb{Z} \)-grading of \( L \), \( [x, L_1] = L_{-1} \), and \( [y, L_{-1}] = L_1 \).

**Proof** By assumption, \( S \cong \mathfrak{sl}_2(F) \). Suppose \( y \notin E \). As \( \text{ad}_h^2L_i \subseteq \mathbb{F}y \) for \( i \neq \pm 1 \) and \( \text{ad}_hL_{-1} \subseteq L_1 \), this can only happen if \( [y, L_1] \neq 0 \). Then, by the grading properties, \( [y, L_1] \subseteq L_3 \) and so 3 is equal to a member \( i \) of \( \{-2, -1, 0, 1, 2\} \) modulo \( p \). As \( p \geq 5 \), this implies \( p = 5 \) and \( i = -2 \).

Thus \( [y, L_1] = \mathbb{F}x \). It follows that, for every \( u \in L_1 \), \( \text{ad}_u \text{ad}_h u = 0 \), whence \( \text{ad}_u \text{ad}_h u = (\text{ad}_h \text{ad}_u - \text{ad}_h)u = -u \). Therefore \( [y, [y, L_{-1}]] \supseteq [y, [y, [x, L_1]]] = [y, L_1] = \mathbb{F}x \), and, by homogeneity, \( [y, [y, L_{-1}]] \subseteq L_{-2} = \mathbb{F}x \), so the first case holds.

Otherwise, \( y \) is extremal as well and the other statements follow routinely. 

\[ \square \]

5. The \( p = 5 \) case Suppose \( L \) is simple and \( p = 5 \) and \( \text{ad}_Y^2(L_{-1}) \neq \{0\} \). Then \( L \) is isomorphic to the Witt algebra \( W_{1,1}(5) \).
it suffices to show that for each \( z \)

\[
\text{words, }
\]

\[
\text{the subalgebra }
\]

\[
\text{is isomorphic to the Witt algebra }
\]

\[
W_y,
\]

\[
\text{such that }
\]

\[
\text{It remains to prove that }
\]

\[
L
\]

\[
\text{is simple. To this end, suppose that }
\]

\[
L
\]

\[
\text{generates is simple of dimension } 5.
\]

\[
\text{Computations yield that if } v, y, L \text{ generates is simple of dimension } 5.
\]

\[
\text{As }
\]

\[
\text{Proof 6 ARJEH M. COHEN}
\]

\[
\text{subalgebra generated by }
\]

\[
\text{L}
\]

\[
\text{diagonizable and the components }
\]

\[
\text{h}
\]

\[
\text{by }
\]

\[
\text{Assume that }
\]

\[
\text{Note that }
\]

\[
\text{Proof}
\]

\[
\text{two force ad}
\]

\[
\text{and hence on } L/S
\]

\[
\text{generated by ad}
\]

\[
\text{W}
\]

\[
\text{a module on which}
\]

\[
\text{End}(L/W)
\]

\[
\text{Applying Step 2, we find that ad}_x \text{ and ad}_y \text{ act quadratically on } L/S \text{ and hence on } L/W.
\]

\[
\text{The non-commutative polynomial relations between the}
\]

\[
\text{two force ad}_z = 0.
\]

\[
\text{6. The generic case Assume that } L \text{ is a simple Lie algebra not isomorphic to}
\]

\[
W_{1,1}(5).
\]

\[
\text{Then } L \text{ is generated by } E.
\]

\[
\text{Proof Note that } [y, [y, L_{-1}]] = 0 \text{ as } y \in E \text{ and so Step 3 gives that } h = [x, y]
\]

\[
\text{is diagonalizable and the components }
\]

\[
L_i = L_i(-\text{ad}_h) \text{ (i = } -2, -1, 0, 1, 2 \text{) of the grading by } h \text{ satisfy } L_{-2} = \mathbb{F}x, L_{-1} = [x, L_1], L_2 = \mathbb{F}y, \text{ and } L_1 = [y, L_{-1}].
\]

\[
\text{Consider the subalgebra } I \text{ of } L \text{ generated by } x, y, \text{ and } L_1.
\]

\[
\text{As } L_{-1} = [x, L_1], \text{ the subalgebra } I \text{ contains the linear subspace } J = L_{-2} + L_{-1} + L_1 + L_2 \text{ of } L.
\]

\[
\text{As } [J, L_0] \subseteq J \text{ and } J \text{ generates } I, \text{ we have } [I, L_0] \subseteq I. \text{ This implies } [I, L] = I.
\]

\[
\text{In other words, } I \text{ is an ideal of } L, \text{ and so, by simplicity of } L, \text{ it coincides with } L.
\]

\[
\text{Therefore, it suffices to show that for each } z \in L_1 \text{ there exists } u \in E \text{ such that } z \text{ is in the}
\]

\[
\text{subalgebra generated by } x, y, \text{ and } u.
\]
To this end, let \( z \in L_1 \). Put \( h = [x, y] \). The following relations hold in \( L \), for some \( \alpha \in F \).

(1) \[ [h, x] = 2x, \]
(2) \[ [h, y] = -2y, \]
(3) \[ [z, h] = z, \]
(4) \[ [y, z] = 0, \]
(5) \[ [x, [x, z]] = 0, \]
(6) \[ [y, [x, z]] = z, \]
(7) \[ [y, [z, [z, x]]] = 0, \]
(8) \[ [x, [z, [z, x]]] = 0, \]
(9) \[ [y, [z, [z, x]]] = 0, \]
(10) \[ [x, [x, [z, [z, x]]]] = 0, \]
(11) \[ [y, [x, [z, [z, x]]]] = [z, [z, [z, x]]], \]
(12) \[ [z, [z, [z, [z, x]]]] = \alpha y. \]

By computation, we establish that the Lie subalgebra \( L' \) of \( L \) generated by \( x, y \), and \( z \) is linearly spanned by the following set \( B \) of eight elements, where \( h_1 = [[x, z], z] \).

\[ x \in L_{-2}; \quad [x, z], [[h_1, z], x] \in L_{-1}; \quad h, h_1 \in L_0; \quad z, [h_1, z] \in L_1; \quad y \in L_2 \]

We then exhibit an element \( u \in L' \) as specified. Because of the grading induced by \( \text{ad}_h \) on \( L \), the endomorphism \( \text{ad}_z \) on \( L \) is nilpotent of order at most 5 and \( \exp(\text{ad}_z) \) is a linear transformation of \( L \) (it is well defined as \( p \neq 2, 3 \)). Put

\[ u = \exp(\text{ad}_z)x = x + \text{ad}_z(x) + \frac{1}{2}\text{ad}_z^2(x) + \frac{1}{6}\text{ad}_z^3(x) + \frac{1}{24}\text{ad}_z^4(x). \]

A straightforward computation in \( L' \) using the above equations shows that \( y \) and \( u \) are an \( \mathfrak{sl}_2 \)-pair in \( L \). By the assumption that \( L \) is not isomorphic to \( W_{1,1}(5) \), this implies \( u \in E \).

We verify that \( z \) lies in the subalgebra \( L'' \) of \( L \) generated by the three extremal elements \( x, y \), and \( u \). Observe that

\[ \text{ad}_z(x) + \frac{1}{2}\text{ad}_z^2(x) + \frac{1}{6}\text{ad}_z^3(x) = u - x - \frac{\alpha}{24}y \in L''. \]

Acting by \( \text{ad}_y \) and using (6), (7), (9), we find

\[ z = -\text{ad}_y\text{ad}_z(x) - \frac{1}{2}\text{ad}_y\text{ad}_z^2(x) - \frac{1}{6}\text{ad}_y\text{ad}_z^3(x) \in \text{ad}_yL'' \subset L''. \]

This proves that \( z \) belongs to \( L'' \) and so we are done. \( \square \)
REFERENCES


(Arjeh M. Cohen) DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE TECHNISCHE UNIVERSITET EINDHOVEN P.O. BOX 513, 5600 MB EINDHOVEN, NETHERLANDS
E-mail address: arjehmcohen@gmail.com