On Pro-2 Identities of $2 \times 2$ Linear Groups

David El-Chai Ben-Ezra, Efim Zelmanov

Abstract
Let $\hat{F}$ be a free pro-$p$ non-abelian group, and let $\Delta$ be a local commutative complete ring with a maximal ideal $I$ such that $\text{char}(\Delta/I) = p$. In [Zu], Zubkov showed that when $p \neq 2$, the pro-$p$ congruence subgroup $GL_2^p(\Delta) = \ker(GL_2(\Delta) \xrightarrow{\Delta/I} GL_2(\Delta/I))$ admits a pro-$p$ identity. I.e. there exists an element $1 \neq w \in \hat{F}$ that vanishes under any continuous homomorphism $\hat{F} \rightarrow GL_2^p(\Delta)$.

In this paper we investigate the case $p = 2$. The main result is that when $\text{char}(\Delta) = 2$, the pro-$2$ group $GL_2^p(\Delta)$ admits a pro-$2$ identity. This result was obtained by the use of trace identities that are originated in PI-theory.

Mathematics Subject Classification (2010): primary: 20E18, 16R30, secondary: 20E05, 20H25.

Key words and phrases: pro-$p$ identities, linear pro-$p$ group, trace identities, PI-theory.

Contents

1 Introduction 2

2 The case $\text{char}(\Delta) = 2$ 3

2.1 The Universal Representation 3

2.2 Reduction of Theorem 2.2 4

2.3 Some useful lemmas 6

2.4 The ring of the Pseudo-Generic Matrices 10

2.5 Some properties of $\hat{G}'$ 13

2.6 Minimal $t(x)$-related components 15

2.7 Interim conclusions 23

2.8 Proof of Theorem 2.2 27

3 Some remarks regarding $p = 2$ versus $p \neq 2$ 29

3.1 The Universal Representation 29

3.2 The Pseudo Generic Matrices 31

3.3 The structure of a minimal component 32

3.4 Proving Proposition 3.3 38
1 Introduction

It is well known that discrete free non-abelian groups are linear, and can easily be embedded even in the group $GL_2(\mathbb{Z})$. Surprisingly, turning to the category of pro-$p$ groups, the problem of linearity is still open.

We say that $\Delta$ is a pro-$p$ ring, if $\Delta$ is a local commutative complete ring with a maximal ideal $I$, such that $\Delta/I$ is a finite field of characteristic $p$. In this case

$$\Delta = \lim_{\leftarrow n} \Delta/I^n$$

is a profinite ring, and for any $d$, the congruence subgroup

$$GL_d^1(\Delta) = \ker(GL_d(\Delta) \xrightarrow{\Delta \rightarrow \Delta/I} GL_d(\Delta/I))$$

is a pro-$p$ group.

Problem 1.1. Let $\hat{F}$ be a non-abelian free pro-$p$ group. Can $\hat{F}$ be continuously embedded in $GL_d^1(\Delta)$ for some $d$ and a pro-$p$ ring $\Delta$?

Infact, the known partial results give the impression that the answer should be negative. For example, it is known that a free pro-$p$ non-abelian group cannot be embedded as a closed subgroup in the pro-$p$ groups (see [DMSS], [BL])

$$GL_d^1(\mathbb{Z}_p) = \ker(GL_d(\mathbb{Z}_p) \xrightarrow{\mathbb{Z}_p \rightarrow \mathbb{F}_p} GL_d(\mathbb{F}_p))$$

$$GL_d^1(\mathbb{F}_p \langle \langle t \rangle \rangle) = \ker(GL_d(\mathbb{F}_p \langle \langle t \rangle \rangle) \xrightarrow{\langle \langle t \rangle \rangle \rightarrow 0} GL_d(\mathbb{F}_p)).$$

Let $\hat{F}$ be a free pro-$p$ group, and $\hat{H}$ a pro-$p$ group. We say that an element $1 \neq w \in \hat{F}$ is a pro-$p$ identity of $\hat{H}$ if $w$ vanishes under every (continuous) homomorphism $\hat{F} \rightarrow \hat{H}$. In [Zu], using the idea of generic matrices, Zubkov showed that given a fixed $d$, the following conditions are equivalent [Zu]:

- $\hat{F}$ cannot be embedded in $GL_d^1(\Delta)$ for some pro-$p$ ring $\Delta$.
- There exists an element $1 \neq w \in \hat{F}$ that serves as a pro-$p$ identity of every pro-$p$ group of the form $GL_d^1(\Delta)$, where $\Delta$ is a pro-$p$ ring.

Then, Zubkov showed that these conditions are satisfied for $d = 2$ whenever $p \neq 2$ [Zu]. In particular, for every $p \neq 2$, a free pro-$p$ group cannot be embedded, as a closed subgroup, in $GL_2^1(\Delta)$, where $\Delta$ is a pro-$p$ ring. Later, using ideas from the solution of the Specht problem, the second author announced that given a fixed $d$, the aforementioned conditions are satisfied for every large enough prime $d \ll p$ (see [Ze1], [Ze2]).

Given these results, the following natural question is what happens when $p$ is not large enough? Or let’s be even more specific, what happens in the case where $d = p = 2$? Investigating this case is the main purpose of this paper. Here is the main result (see [2]).
Theorem 1.2. Let $\Delta$ be a pro-$2$ ring of $\text{char}(\Delta) = 2$. Then, $GL_1^2(\Delta)$ admits a pro-$2$ identity that is independent in $\Delta$.

From Theorem 1.2, we get that a free pro-$2$ group cannot be embedded in $GL_1^2(\Delta)$ when $\Delta$ is a pro-$2$ ring of $\text{char}(\Delta) = 2$. Actually, one can derive from here that a free pro-$2$ group cannot be embedded in $GL_1^2(\Delta)$ whenever $\text{char}(\Delta) = 2^m$ for some $m$. The main idea of Theorem 1.2’s proof is the use of trace identities that are originated in PI-theory (see [R, P, K, DF, BKR]). We note that the problem whether a free pro-$2$ group can be embedded in $GL_1^2(\Delta)$ when $\Delta$ is a pro-$2$ ring of $\text{char}(\Delta) = 0$ is still open.

To the end of the paper (see §3) we give a review of Zubkov’s approach, and we describe where exactly Zubkov’s argument fails when $d = p = 2$. This description allows us to show that when $d = 2$, there is a dichotomy between $p = 2$ and $p \neq 2$, and in some sense, $2 \times 2$ linear pro-$2$ groups have less pro-$2$ identities (if any).

Acknowledgments: During the period of the research, the first author was supported by NSF research training grant (RTG) # 1502651.

2 The case $\text{char}(\Delta) = 2$

2.1 The Universal Representation

We denote the free pro-$2$ group generated by $X, Y$, by $\hat{F} = \langle\langle X, Y \rangle\rangle$. Let $x_{i,j}$ and $y_{i,j}$ for $1 \leq i, j \leq 2$ be free commuting variables, and let

$$\Lambda_* = (\mathbb{Z}/2\mathbb{Z}) \langle\langle x_{i,j}, y_{i,j} \mid 1 \leq i, j \leq 2 \rangle\rangle$$

be the associative ring (with identity) of formal power series on $x_{i,j}$ and $y_{i,j}$ over $\mathbb{Z}/2\mathbb{Z}$. Every element in $\Lambda_*$ can be written as $f = \sum_{i=0}^{\infty} f_i$ when $f_i$ is homogeneous of degree $i$.$^{1}$ The finite index ideals

$$\Lambda_* \triangleright P_{*n} = \left\{ f = \sum_{i=0}^{\infty} f_i \in \Lambda_* \mid f_0, ..., f_{n-1} = 0 \right\}.$$

serve as a basis of neighborhoods of zero to a profinite topology on $\Lambda_*$, making $\Lambda_*$ a pro-$2$ ring, with $P_{*1}$ as its maximal ideal. Notice that $P_{*n} = P_{*1}^n$ for every $n$.

$^{1}$By degree we mean that $\deg(\prod_{i=1}^{2} x_{i,j}^{\alpha_{i,j}} y_{i,j}^{\beta_{i,j}}) = \sum_{i,j=1}^{2}(\alpha_{i,j} + \beta_{i,j})$. 

3
Endowed with the topology that comes from the congruence ideals
\[ M_2(\Lambda_*, P_n) = \ker(M_2(\Lambda_*) \rightarrow M_2(\Lambda_*/P_n)) \]
as a basis of neighborhoods of zero, \( M_2(\Lambda_*) \) is a profinite ring. It is easy to check that this topology makes the group \( 1 + M_2(\Lambda_*, P_1) \) a pro-2 group. Denoting the generic matrices
\[
\begin{align*}
x_* &= \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}, &
y_* &= \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \in M_2(\Lambda_*, P_1).
\end{align*}
\]
we have then a natural (continuous) homomorphism \( \pi_* : \hat{F} \rightarrow 1 + M_2(\Lambda_*, P_1) \) defined by
\[
X \mapsto 1 + x_*, \quad Y \mapsto 1 + y_*.
\]
We denote \( G_* = \langle 1 + x_*, 1 + y_* \rangle \subseteq 1 + M_2(\Lambda_*, P_1) \) for the (discrete) subgroup generated by \( 1 + x_* \) and \( 1 + y_* \), and \( \hat{G}_* \subseteq 1 + M_2(\Lambda_*, P_1) \) to be its closure in \( 1 + M_2(\Lambda_*, P_1) \). Adopting Zubkov’s terminology, \( \pi_* : \hat{F} \rightarrow \hat{G}_* \) is called the universal representation.

The following proposition, which is actually based on the proof of of Theorem 2.1 in [Zu], justifies the name of \( \pi_* \):

**Proposition 2.1.** Let \( \Delta \) be a pro-2 ring with \( \text{char}(\Delta) = 2 \). Then, every \( 1 \neq w(X, Y) \in \ker \pi_* \) is a pro-2 identity of \( GL_2(\Delta) \).

Proposition 2.1 shows that Theorem 1.2 boils down to the following theorem:

**Theorem 2.2.** The universal representation \( \pi_* : \hat{F} \rightarrow \hat{G}_* \) is not injective.

### 2.2 Reduction of Theorem 2.2

We want to replace the generic matrices \( x_*, y_* \) by matrices \( x, y \) that satisfy the condition \( \det(x) = \det(y) = 0 \).

Let \( h \) be a rational function on \( x_{i,j}, y_{i,j} \) over \( \mathbb{Z}/2\mathbb{Z} \), and denote the discrete ring
\[
\Lambda_# = (\mathbb{Z}/2\mathbb{Z}) \langle x_{i,j}, y_{i,j} \mid 1 \leq i, j \leq 2 \rangle \leq \Lambda_*.
\]
We say that \( h \) is homogeneous if there exist homogeneous polynomials \( 0 \neq f, g \in \Lambda_# \) such that \( h = \frac{f}{g} \). In this case we define
\[
\deg(h) = \deg(f) - \deg(g).
\]
It is easy to see that \( \deg(h) \) is well defined, i.e. if \( \frac{f_1}{g_1} = \frac{f_2}{g_2} \) then \( \deg \left( \frac{f_1}{g_1} \right) = \deg \left( \frac{f_2}{g_2} \right) \). Now, consider the set of all power series of the form
\[
\Gamma = \left\{ \sum_{i=0}^{\infty} h_i \mid h_i = 0 \text{ or } \deg(h_i) = i, \ m \in \mathbb{Z} \right\}.
\]
It is easy to see that \( \Gamma \) has a natural ring structure. In addition, let \( 0 \neq h \in \Gamma \). Then, we can write \( h = \sum_{i=m}^{\infty} h_i \) where \( h_m \neq 0 \). Hence
\[
h = h_m(1 + a) \quad \text{where} \quad a = h_m^{-1} \sum_{i=m+1}^{\infty} h_i.
\]
Then as \( h_m \) and \( 1 + a \) are invertible in \( \Gamma \), we obtain that \( h \) is invertible as well.

It follows that \( \Gamma \) is actually a field that contains \( \Lambda_\ast \). Consider now the following quadratic polynomials on the variables \( \mu \) and \( \nu \) over the field \( \Gamma \):
\[
p(\mu) = \det((1 + x_\ast)(1 + \mu \cdot 1) - 1) \\
= (1 + t(x_\ast) + \det(x_\ast)) \cdot \mu^2 + (2 \det(x_\ast) + t(x_\ast)) \cdot \mu + \det(x_\ast)
\]
\[
q(\nu) = \det((1 + y_\ast)(1 + \nu \cdot 1) - 1) \\
= (1 + t(y_\ast) + \det(y_\ast)) \cdot \nu^2 + (2 \det(y_\ast) + t(y_\ast)) \cdot \nu + \det(y_\ast)
\]

where \( t(x_\ast) \) and \( t(y_\ast) \) are the traces of \( x_\ast \) and \( y_\ast \) respectively.

**Lemma 2.3.** Let \( \bar{\mu} \) and \( \bar{\nu} \) be roots of \( p(\mu) \) and \( q(\nu) \) respectively. Then, for the field extension \( \Gamma(\bar{\mu}, \bar{\nu}) \) we have
\[
\Gamma(\bar{\mu}, \bar{\nu}) = \Gamma \oplus \bar{\mu} \cdot \Gamma \oplus \bar{\nu} \cdot \Gamma \oplus \bar{\mu} \cdot \bar{\nu} \cdot \Gamma
\]
as vector spaces.

**Proof.** As \( p(\mu) \) and \( q(\nu) \) are of degree 2, it is enough to show that \( \bar{\mu} \notin \Gamma \), hence \( \Gamma(\bar{\mu}) = \Gamma \oplus \bar{\mu} \cdot \Gamma \), and then, that \( \bar{\nu} \notin \Gamma(\bar{\mu}) \). We will assume that \( \Gamma(\bar{\mu}) = \Gamma \oplus \bar{\mu} \cdot \Gamma \) and show that \( \bar{\nu} \notin \Gamma(\bar{\mu}) \). The other part is easier. Clearly, for showing that \( \bar{\nu} \notin \Gamma(\bar{\mu}) \), it is enough to show that \( q(\nu) \) does not have a root in \( \Gamma(\bar{\mu}) = \Gamma \oplus \bar{\mu} \cdot \Gamma \). Assume negatively that
\[
h = k + \bar{\mu} \cdot l
\]
for some \( k, l \in \Gamma \) is a root of \( q(\nu) \). Substituting \( h \) in \( q(\nu) \), and using the identity (recall that \( \text{char}(\Gamma) = 2 \))
\[
\bar{\mu}^2 = \frac{\bar{\mu} \cdot t(x_\ast) + \det(x_\ast)}{1 + t(x_\ast) + \det(x_\ast)}
\]
we obtain that \( l, k \) satisfy the following system of equations:
\[
\begin{align*}
1. \quad & (1 + t(y_\ast) + \det(y_\ast)) \cdot (k^2 + \frac{\det(x_\ast)}{1 + t(x_\ast) + \det(x_\ast)} \cdot l^2) + t(y_\ast) \cdot k + \det(y_\ast) = 0 \\
2. \quad & \frac{t(x_\ast)(1 + t(y_\ast) + \det(y_\ast))}{1 + t(x_\ast) + \det(x_\ast)} \cdot l^2 + t(y_\ast) \cdot l = 0
\end{align*}
\]
The solutions for Equation 2. are:

\[ l = 0, \]
\[ l = \frac{t(y_*) \cdot (1 + t(x_*) + \det(x_*))}{t(x_*) \cdot (1 + t(y_*) + \det(y_*))} \]

So we have two cases to negate:

**Case 1.** \( l = 0 \). In this case, Equation 1. becomes

\[ (1 + t(y_*) + \det(y_*)) \cdot k_2^2 + t(y_*) \cdot k + \det(y_*) = 0 \quad (2.1) \]

Write \( k = \sum_{i=m}^{\infty} k_i \) where \( k_m \neq 0 \). It is easy to see that if \( \text{deg}(k_m) = m \leq 0 \), then Equation (2.1) implies \( k_m^2 = 0 \), so \( k_m = 0 \), and we get a contradiction. On the other hand, if \( \text{deg}(k_m) = m \geq 2 \) then \( \det(y_*) = 0 \) which also gives a contradiction. So \( m = 1 \), and we get that \( k_1 \) satisfies the equation

\[ k_1^2 + t(y_*) \cdot k_1 + \det(y_*) = 0. \]

Write \( k_1 = 4 \over g \) for \( f, g \in \Lambda_\# \) such that \( f, g \) do not have a common divisor in \( \Lambda_\# \). Then

\[ f^2 + f \cdot t(y_*) + g^2 \cdot \det(y_*) = 0. \]

We obtain that \( g \) divides \( f^2 \), so as \( \Lambda_\# \) is a unique factorization domain, by the assertion that \( f, g \) do not have a common divisor, we get that \( g \) is invertible in \( \Lambda_\# \). It follows that we can assume that \( g = 1 \), and thus

\[ 0 = f^2 + f \cdot t(y_*) + \det(y_*) \]
\[ = f^2 + f \cdot (y_{11} + y_{22}) + y_{11}y_{22} + y_{12}y_{21} \]

where \( f \in \Lambda_\# \) with \( \text{deg}(f) = 1 \). It is easy to check that such \( f \) does not exist, what gives a contradiction to our negative assumption.

**Case 2.** \( l = \frac{t(y_*) \cdot (1 + t(x_*) + \det(x_*))}{t(x_*) \cdot (1 + t(y_*) + \det(y_*))} \). In this case, Equation 1. becomes

\[ (1 + t(y_*) + \det(y_*)) \cdot k^2 + t(y_*) \cdot k + \det(y_*) = 0. \]

It follows that \( k' = \frac{(1 + t(y_*) + \det(y_*))}{t(y_*)} \cdot k \in \Gamma \) satisfies the equation

\[ k'^2 + k' + \frac{(1 + t(y_*) + \det(y_*)) \cdot \det(y_*)}{t(y_*)^2} \quad (2.2) \]
\[ + \frac{(1 + t(x_*) + \det(x_*)) \cdot \det(x_*)}{t(x_*)^2} = 0. \]
As a corollary from Lemma 2.3 we have:

and thus the sum

Now, notice that as

we have

This finishes the proof.

where

If we can assume that

have a common divisor, we get that

As \( \Lambda_{\#} \) is a unique factorization domain, it follows that \( g \) is divisible by \( t(x_\ast) \) and in a similar way, \( g \) is also divisible by \( t(y_\ast) \). So writing \( g = g' t(x_\ast) t(y_\ast) \) we get that

We obtain that \( g' \) divides \( f^2 \), so by the assertion that \( f, g \) do not have a common divisor, we get that \( g' \) is invertible in \( \Lambda_{\#} \). It follows that we can assume that \( g' = 1 \), and thus

where \( f \in \Lambda_{\#} \) with \( \deg(f) = 2 \). It is easy to check that such \( f \) does not exist, what gives a contradiction to our negative assumption. This finishes the proof.

Now, notice that as \( 1 + t(x_\ast) + \det(x_\ast) \) and \( 1 + t(y_\ast) + \det(y_\ast) \) are invertible over \( \Lambda_\ast \), we have

and thus the sum \( \Lambda = \Lambda_\ast + \bar{\mu} \cdot \Lambda_\ast + \bar{\nu} \cdot \Lambda_\ast + \bar{\mu} \bar{\nu} \cdot \Lambda_\ast \) is actually a subring of \( \Gamma \).

As a corollary from Lemma 2.3 we have:
Corollary 2.4. The ring $\Lambda$ satisfies
\[ \Lambda = \Lambda_1 \oplus \bar{\mu} \cdot \Lambda_2 \oplus \bar{\nu} \cdot \Lambda_3 \oplus \bar{\mu} \bar{\nu} \cdot \Lambda_4. \]  
(2.5)

Recall the ideals $\Lambda_\ast \triangleright P_\ast n = \{ f = \sum_{i=0}^{\infty} f_i \in \Lambda_\ast \mid f_0, \ldots, f_{n-1} = 0 \}$. For $n \geq 0$ define the subsets
\[ P_n = P_\ast n + \bar{\mu} \cdot P_{\ast n-1} + \bar{\nu} \cdot P_{\ast n-1} + \bar{\mu} \bar{\nu} \cdot P_{\ast n-2} \subseteq \Lambda \]  
(2.6)
when we denote $P_{-1} = P_0 = \Lambda_\ast$, so $P_1 = P_\ast 1 + \bar{\mu} \cdot \Lambda_2 + \bar{\nu} \cdot \Lambda_3 + \bar{\mu} \bar{\nu} \cdot \Lambda_4$. Notice that as the sum in (2.5) is direct, so is the sum in (2.6). Hence, for every $n$ we have $P_n \cap \Lambda_\ast = P_\ast n$. The following properties of $\Lambda$ are easy to verify by using the identities in (2.4), and induction on $n$:

Proposition 2.5. The ring $\Lambda$ is a pro-2 ring, with a maximal ideal $P_\ast 1$. Moreover, $P_n = P_\ast n$ for every $n$.

As $P_n \cap \Lambda_\ast = P_\ast n$ it follows that the topology of $\Lambda_\ast \subseteq \Lambda$ that is induced by the topology of $\Lambda$ coincides with the topology of $\Lambda_\ast$ defined by the basis $P_\ast n$.

Similarly, endowed with the topology that comes from the congruence ideals $M_2(\Lambda, P_n) = \ker(M_2(\Lambda) \to M_2(\Lambda/P_n))$ as a basis of neighborhoods of zero, $M_2(\Lambda)$ is also a profinite ring, that contains the profinite ring $M_2(\Lambda_\ast)$. Also here, it is easy to show that this topology makes the group $1 + M_2(\Lambda, P_\ast 1)$ a pro-2 group that contains the pro-2 subgroup $1 + M_2(\Lambda_\ast, P_\ast 1)$. Notice that as $P_n \cap \Lambda_\ast = P_\ast n$, the profinite topology that $1 + M_2(\Lambda, P_\ast 1)$ induces on $1 + M_2(\Lambda_\ast, P_\ast 1)$ coincides with the profinite topology of $1 + M_2(\Lambda_\ast, P_\ast 1)$ defined by the basis $1 + M_2(\Lambda_\ast, P_\ast n)$.

Now, we are ready to define the pseudo-generic matrices
\[ x = (1 + x_\ast)(1 + \bar{\mu} \cdot 1) - 1 \in M_2(\Lambda, P_\ast 1) \]
\[ y = (1 + y_\ast)(1 + \bar{\nu} \cdot 1) - 1 \in M_2(\Lambda, P_\ast 1). \]

By the construction of $x$ and $y$ we have $\det(x) = \det(y) = 0$.

Let now $G = (1 + x, 1 + y) \subseteq 1 + M_2(\Lambda, P_\ast 1)$ be the (discrete) group generated by $1 + x$ and $1 + y$, and let $\tilde{G} \subseteq 1 + M_2(\Lambda, P_\ast 1)$ be its closure. Notice that by the discussion above, both $G_\ast$ and $\tilde{G}$ are embedded in $1 + M_2(\Lambda, P_\ast 1)$, and their profinite topology is induced by the one of $1 + M_2(\Lambda, P_\ast 1)$.

Now, by the definition of the pseudo-generic matrices
\[ 1 + x = (1 + x_\ast)(1 + \bar{\mu} \cdot 1) \]
\[ 1 + y = (1 + y_\ast)(1 + \bar{\nu} \cdot 1) \]
where $1 + \bar{\mu} \cdot 1$ and $1 + \bar{\nu} \cdot 1$ are central in $M_2(\Lambda)$. Hence, for every (discrete) commutator element $w(X, Y) \in F'$ one has

\[ w(1 + x, 1 + y) = w(1 + x_\ast, 1 + y_\ast) \in G_\ast' \subseteq 1 + M_2(\Lambda_\ast, P_\ast 1) \subseteq 1 + M_2(\Lambda, P_\ast 1). \]

As we saw that both the profinite topology of $G_\ast$ and $\tilde{G}$ are induced by the profinite topology of $1 + M_2(\Lambda, P_\ast 1)$, it follows that actually

\[ G' = G_\ast' \subseteq 1 + M_2(\Lambda_\ast, P_\ast 1) \subseteq 1 + M_2(\Lambda, P_\ast 1) \]
for the commutator subgroups $\hat{G}_*$ and $\hat{G}$. Similarly, the lower central series of $\hat{G}$ and $G_*$ are equal (apart from the first term). Hence, in order to prove Theorem 2.6, it is enough to prove that:

**Theorem 2.6.** The homomorphism $\pi = \pi_*|_{\hat{F}'} : \hat{F}' \to \hat{G}' = \hat{G}$ is not injective.

Notice that obviously, $\pi$ can also be seen as the restriction of the map $\hat{F} \to \hat{G}$ defined by $X \mapsto 1 + x, Y \mapsto 1 + y$, to $\hat{F}'$.

From now on, we assume negatively that Theorem 2.6 is false, and that $\pi$ is injective. In particular, we assume that $\hat{G}'$ is isomorphic to $\hat{F}'$ through $\pi$.

### 2.3 Some useful lemmas

We state some general properties of $2 \times 2$ matrices. By the Cayley–Hamilton theorem, for every $a \in M_2(\Lambda)$, we have $a^2 + t(a)a + (\det(a)\cdot 1 = 0$ where $t(a)$ is the trace of $a$. As a corollary of that, we have the following lemmas:

**Lemma 2.7.** Let $a, b \in M_2(\Lambda)$. Then:

1. $[a, b] = t(a)b + t(b)a + (t(ab) + t(a)t(b)) \cdot 1$.
2. $[a, b, a] = t(a)[a, b]$.
3. If $t(a) = 0$, then $a^2 = \det(a)\cdot 1 \in M_2(\Lambda)$ is central.
4. If $a, b \in M_2(\Lambda)$ then the trace $t([a, b]b^n) = 0$ for every $n \geq 0$.

**Proof.** Applying the Cayley–Hamilton theorem for $a + b$ one has

\[
0 = (a + b)^2 + t(a + b)(a + b) + \det(a + b) \cdot 1
\]

\[
= a^2 + t(a)a + b^2 + t(b)b + ab + ba + t(a)b + t(b)a + \det(a + b) \cdot 1.
\]

Subtracting the equations $a^2 + t(a)a + \det(a) \cdot 1 = 0$ and $b^2 + t(b)b + \det(b) \cdot 1 = 0$ it follows that

\[
[a, b] = ab + ba = t(a)b + t(b)a + (\det(a) + \det(b) + \det(a + b)) \cdot 1
\]

\[
= t(a)b + t(b)a + (t(ab) + t(a)t(b)) \cdot 1
\]

so we get Part 1. Part 2 is an immediate consequence of Part 1. Part 3 is an immediate consequence of the Cayley–Hamilton theorem.

For Part 4 we use induction on $n$. So for $n = 0$ the claim is easy, and for $n = 1$ it follows from the observation $[a, b]b = [ab, b]$. For $n \geq 2$: by the Cayley–Hamilton theorem we have $b^2 = t(b)b + \det(b)\cdot 1$. Hence, by the induction hypothesis

\[
t([a, b]b^n) = t(b) \cdot t([a, b]b^{n-1}) + \det(b) \cdot t([a, b]b^{n-2}) = 0
\]

as required.

**Lemma 2.8.** For the pseudo-generic matrices $x, y$ we have
1. \( x^2 = t(x)x, \quad y^2 = t(y)y, \quad (xy)^2 = t(xy)xy, \quad (yx)^2 = t(xy)y. \)

2. \( xyx = t(xy)x, \quad yxy = t(xy)y. \)

3. \([x, y]^2 = (t(xy)^2 + t(x)t(y)t(xy)) \cdot 1.\)

**Proof.** Part 1 follows from the Cayley–Hamilton theorem and the property \( \det(x) = \det(y) = 0. \) The identity \( xyx = t(xy)x \) follows from the previous part and Part 1 of Lemma 2.7 by the following computation

\[
xyx = (yx + t(x)y + t(y)x + t(xy) + t(x)t(y)) \cdot 1 \cdot x
= yx^2 + t(x)yx + t(y)x^2 + t(xy)x + t(x)t(y)x = t(xy)x.
\]

The identity \( yxy = t(xy)y \) follows similarly. Part 3 follows from the previous properties by the following computation

\[
[x, y]^2 = (xy + yx)(xy + yx)
= t(xy)(xy + yx) + t(xy)t(x)y + t(xy)t(y)x
= t(xy)(t(xy) + t(x)t(y)) \cdot 1.
\]

\[\Box\]

**Remark 2.9.** As the expression \([x, y]^2\) is central in \( M_2(\Lambda) \), sometime we will consider it as an element of \( \Lambda \) and just write \([x, y]^2 = (t(xy)^2 + t(x)t(y)t(xy)).\]

### 2.4 The ring of the Pseudo-Generic Matrices

**Definition 2.10.** We define the (discrete) subrings (with identity) of \( \Lambda\) and \( M_2(\Lambda)\)

\[
S = \langle t(x), t(y), [x, y]^2 \rangle \subseteq \Lambda
T = \langle t(x), t(y), t(xy) \rangle \subseteq \Lambda
R = \langle x, y, T \cdot 1 \rangle \subseteq M_2(\Lambda).
\]

The ring \( R \) will be called the ring of the pseudo generic matrices.

**Proposition 2.11.** The ring \( T \) is freely generated by \( t(x), t(y), t(xy) \) as a commutative ring over \( \mathbb{Z}/2\mathbb{Z}, \)

**Proof.** Consider the free commutative variables \( \lambda, \theta, \vartheta \) over \( \mathbb{Z}/2\mathbb{Z}, \) and define

\[
\bar{x} = \begin{pmatrix} \lambda & \vartheta - \lambda \theta \\ 0 & 0 \end{pmatrix}, \quad \bar{y} = \begin{pmatrix} \theta & 0 \\ 1 & 0 \end{pmatrix}.
\]

Denote \( \bar{T} = \langle t(\bar{x}), t(\bar{y}), t(\bar{x}\bar{y}) \rangle. \) As \( \det(\bar{x}) = \det(\bar{y}) = 0 \) it is easy to verify that we have a natural homomorphism of discrete rings

\[
\Lambda \geq (\mathbb{Z}/2\mathbb{Z}) \langle x_{ij}, y_{ij}, \mu, \nu | 1 \leq i, j \leq 2 \rangle \rightarrow (\mathbb{Z}/2\mathbb{Z}) \langle \lambda, \theta, \vartheta \rangle
\]

10
defined by
\[
\begin{align*}
  x_{11} &\mapsto \lambda & y_{11} &\mapsto \theta \\
  x_{12} &\mapsto \lambda - \lambda \theta & y_{12} &\mapsto 0 \\
  x_{21} &\mapsto 0 & y_{21} &\mapsto 1 \\
  x_{22} &\mapsto 0 & y_{22} &\mapsto 0 \\
  \bar{\mu} &\mapsto 0 & \bar{\nu} &\mapsto 0
\end{align*}
\]
that induces a natural ring homomorphism from the discrete ring generated by
the pseudo-generic matrices \(x, y\) to the discrete ring generated by \(\bar{x}, \bar{y}\) sending
\(x \mapsto \bar{x}\) and \(y \mapsto \bar{y}\). Hence, we have a natural ring homomorphism
\(T \to \bar{T}\) by
\[
\begin{align*}
t(x) &\mapsto t(\bar{x}) = \lambda \\
t(y) &\mapsto t(\bar{y}) = \theta \\
t(xy) &\mapsto t(\bar{x}\bar{y}) = \vartheta.
\end{align*}
\]
So as the ring generated by \(\lambda, \theta, \vartheta\) is freely generated by them as a commutative
ring over \(\mathbb{Z}/2\mathbb{Z}\), the same is valid for \(T\).

As a corollary of Part 3 in Lemma 2.8 and Proposition 2.11 one can easily prove:

**Proposition 2.12.** The ring \(S \subseteq T\) and it is free on \(t(x), t(y), [x, y]^2\). In
addition, \(T\) is freely generated by \(t(x), t(y), t(xy)\) as an \(S\)-module.

We also have:

**Proposition 2.13.** The ring \(R\) is freely generated as a module over \(T\) by \(1, x, y\) and \(xy\).

**Proof.** It follows from Lemma 2.8 that indeed \(R\) is generated as a module over \(T\) by \(1, x, y\) and \(xy\). We have to show that the way to write \(a = \alpha + \beta \cdot [x, y] + \gamma \cdot y + \delta \cdot xy\)
for \(\alpha, \beta, \gamma, \delta \in T\) is unique. Observe that as the traces \(t([x, y]) = t([x, y]x) = t([x, y]y) = 0\) and \(x[x, y]x = y[x, y]y = 0\), given such \(a\), we have
\[
\begin{align*}
t(x[x, y]a[x, y]y) &\mapsto \alpha \cdot t(xy)[x, y]^2 \\
t([x, y]ya) &\mapsto \beta \cdot [x, y]^2 \\
t([x, y]ax) &\mapsto \gamma \cdot [x, y]^2 \\
t([x, y]a) &\mapsto \delta \cdot [x, y]^2.
\end{align*}
\]
Hence, as \(T\) is a domain (as a free commutative ring by Proposition 2.11), given
such \(a\), we can uniquely restore its coefficients in \(T\), as required.

Denote now the (discrete) ring
\[
R_* = \langle x_*, y_*, t(x_*) \cdot 1, t(y_*) \cdot 1, t(x_*y_*) \cdot 1 \rangle \subseteq M_2(\Lambda_*).
\]
Then \(R_*\), inheriting the degree of \(\Lambda_*\), can be seen as \(R_* = \oplus_{n=0}^\infty R_*^{(n)}\) where
\(R_*^{(n)}\) is the additive group of homogeneous elements of degree \(n\). This direct
sum makes \(R_*\) a graded ring. As \(x_*, y_*\) are generic matrices, we can define a
map \(\tau : R_* \to R\) such that \(x_* \mapsto x\) and \(y_* \mapsto y\). Denote \(R^{(n)} = \tau(R_*^{(n)})\).
Corollary 2.14. One has $R = \oplus_{n=0}^{\infty} R^{(n)}$, what makes $R$ a graded ring.

Proof. Denote $T^{(n)} = R^{(n)} \cap T \cdot 1$ for $n \geq 0$, and $T^{(-1)} = T^{(-2)} = 0$. Using Lemma 2.8 it is easy to check that
\[ R^{(n)} = \tau(R^{(n)}_\ast) = T^{(n)} \cdot x + T^{(n-1)} \cdot y + T^{(n-2)} \cdot xy. \]
Hence, by Proposition 2.11 we have
\[ R^{(n)} = T^{(n)} \oplus T^{(n-1)} \cdot x \oplus T^{(n-1)} \cdot y \oplus T^{(n-2)} \cdot xy. \]

By Proposition 2.13 it is clear that $T \cdot 1 = \oplus_{n=0}^{\infty} T^{(n)}$. Hence
\begin{align*}
R &= T \oplus T \cdot x \oplus T \cdot y \oplus T \cdot xy \\
&= \oplus_{n=0}^{\infty} T^{(n)} \oplus \oplus_{n=0}^{\infty} T \cdot x \oplus \oplus_{n=0}^{\infty} T \cdot y \oplus \oplus_{n=0}^{\infty} T \cdot xy = \oplus_{n=0}^{\infty} R^{(n)}
\end{align*}
as required. \( \square \)

Now, as the ring $R = \langle x, y, 1 \cdot T \rangle = \oplus_{n=0}^{\infty} R^{(n)}$ is a graded ring, we can define the ring $U$ of power series on $x, y$ and $1 \cdot T$ with relation to this grading on $R$. Then, the ideals
\[ U_n = \left\{ \sum_{i=0}^{\infty} f_i \mid f_i \in R_i, f_0, ..., f_{n-1} = 0 \right\} \]
serve as a basis of neighborhoods of zero to a topology on $U$, making $U$ a profinite ring. Now, as for every $n \leq k$ we have $R^{(k)} \subseteq M_2(\Lambda, P_n)$ the profinite topology of $R$ induced by the grading of $R$ is apriori stronger than the profinite topology of $R$ induced by the profinite topology of $M_2(\Lambda)$. Hence, we have a (continuous) ring homomorphism $\sigma : U \to M_2(\Lambda)$ that sends the copy of $R$ in $U$ to its copy in $M_2(\Lambda)$. Hence, $\sigma$ induces also a group homomorphism from the profinite completion of $G = \langle 1 + x, 1 + y \rangle$ in $U$ to its profinite completion in $M_2(\Lambda)$, by sending the generators to their copy in $M_2(\Lambda)$. Hence, we get the diagram
\[ \hat{F} = \langle X, Y \rangle \to \langle \langle 1 + x, 1 + y \rangle \rangle \subseteq U \]
\[ \downarrow \]
\[ \langle \langle 1 + x, 1 + y \rangle \rangle = \hat{G} \subseteq M_2(\Lambda). \]
Recall now that by our negative assumption, we have an isomorphism $\pi : \hat{F}' \to \hat{G}'$ induced by the map
\[ X \mapsto 1 + x \in M_2(\Lambda) \quad Y \mapsto 1 + y \in M_2(\Lambda). \]
Hence, by the diagram we obtain that also the map
\[ X \mapsto 1 + x \in U \quad Y \mapsto 1 + y \in U \]
induces an isomorphism $\hat{F}'$. Hence, we can replace $\hat{G} = \langle \langle 1 + x, 1 + y \rangle \rangle \subseteq M_2(\Lambda)$ by the completion of $G$ in $U$. So from now on the notation $G = \langle \langle 1 + x, 1 + y \rangle \rangle$ refers to the completion of $G$ in $U$, and the degree of terms of elements in $\hat{G}$ will be determined by the grading of $R$. In other words, we regard the elements in $R^{(n)}$ as the elements of $R$ of degree $n$. 

12
2.5 Some properties of $\hat{G}'$

**Proposition 2.15.** If $g \in \hat{G}'$, then $g = 1 + [x, y] \sum_{i=0}^{\infty} a_i$ for some $a_i \in R$.

**Proof.** Having the observations

$$
(1 + [x, y]b_1)(1 + [x, y]b_2) = 1 + [x, y](b_1 + b_2 + b_1[x, y]b_2)
$$

$$
(1 + [x, y]b)^{-1} = 1 + [x, y](b \sum_{i=0}^{\infty} ([x, y]b)^i)
$$

it is enough to show that $\hat{G}'$ has a set of (topological) generators of the form $1 + [x, y] \sum_{i=0}^{\infty} a_i$. One can see (cf. [MKS], Section 2.4, Problem 13) that the set of commutators

$$
\{X^nY^m X^{-n}Y^{-m} | n, m \in \mathbb{Z}\}
$$

generates $F'$, the commutator subgroup of $F = \langle X, Y \rangle$. Hence, it is enough to show that the elements of the set

$$
\{(1 + x)^n (1 + y)^m (1 + x)^{-n} (1 + y)^{-m} | n, m \in \mathbb{Z}\}
$$

are of the form $1 + [x, y] \sum_{i=0}^{\infty} a_i$. Let $n, m \in \mathbb{Z}$. Then

$$
(1 + x)^n = 1 + x \cdot p(x) = 1 + x \cdot p(t(x))
$$

$$
(1 + y)^m = 1 + y \cdot q(y) = 1 + y \cdot q(t(y))
$$

where $p(x)$ and $q(y)$ are power series on $x$ and $y$ respectively. Now

$$
(1 + x)^n (1 + y)^m (1 + x)^{-n} (1 + y)^{-m}
$$

$$
= 1 + [x \cdot p(t(x)), y \cdot q(t(y))] \sum_{i,j=0}^{\infty} (x \cdot p(t(x))^i (y \cdot q(t(y)))^j
$$

$$
= 1 + [x, y] \cdot p(t(x)) \cdot q(t(y)) \sum_{i,j=0}^{\infty} (x \cdot p(t(x))^i (y \cdot q(t(y)))^j
$$

which has the needed form. \qed

Denote now the $T$-module generated by $[x, y]x$, $[x, y]y$, $[x, y]^2$, $[x, y]xy$, by $J$. Then:

**Proposition 2.16.** The $T$-module $J$ is a two sided ideal of $R$. Moreover, $J$ is freely generated as a $T$-module by

$$
[x, y]x, [x, y]y, [x, y]^2, [x, y]xy.
$$

**Proof.** So for showing that $J$ is a two sided ideal of $R$, it is enough to show that $J$ is closed under multiplication by $x$, $y$. Indeed, for the generators of $J$ as a
Corollary 2.17. Let $g = 1 + \sum_{i=1}^{\infty} a_i \in \hat{G}''$ where $\hat{G}''$ is the second derived subgroup of $G$. Then, we have $a_i \in J$ for every $i$.

Proof. By definition, $\hat{G}''$ is generated (topologically) by commutotors of elements of $\hat{G}'$. So as $J$ is an ideal of $R$, it is enough to show that for every two elements in $h, k \in \hat{G}'$ the terms of $g = hkh^{-1}k^{-1}$ lie in $J$. By Proposition 2.16 we can write $h = 1 + [x, y] \sum_{i=1}^{\infty} b_i$ and $k = 1 + [x, y] \sum_{j=1}^{\infty} c_j$ for some $b_i, c_j \in R$. Hence, for $g$ we have

$$g = 1 + ([x, y] \sum_{i=1}^{\infty} b_i, [x, y] \sum_{j=1}^{\infty} c_j) \cdot \sum_{i,j=0}^{\infty} ([x, y] \sum_{i=1}^{\infty} b_i)([x, y] \sum_{j=1}^{\infty} c_j)^2.$$ 

Now, as by Part 2 of Lemma 2.7 we have

$$x[x, y] = t(x)[x, y] + [x, y]x, \quad y[x, y] = t(y)[x, y] + [x, y]y$$

we deduce that $R[x, y] = [x, y]R$. Hence, one can write $g = 1 + [x, y]^2 \cdot \sum_{i=1}^{\infty} d_i$ for some $d_i \in R$. As $[x, y]^2 \in J$, and $J$ is an ideal of $R$, the claim follows. \hfill \Box
We continue with the following lemma:

**Lemma 2.18.** Let \( \hat{F} \) be a free pro-p group, and \( \hat{F} = \hat{F}_1, \hat{F}_2, \hat{F}_3, \ldots \) its lower central series. If \( 1 \neq g, h \in \hat{F} \) commute, then, there exists \( n \) such that \( g, h \in \hat{F}_n - \hat{F}_{n+1} \). In other words, \( g, h \) lie in the same term of the lower central series of \( \hat{F} \).

**Proof.** Let \( \hat{H} = \langle \langle g, h \rangle \rangle \) be the closure of the group generated by \( g, h \) in \( \hat{F} \). Obviously, \( \hat{H} \) is abelian. In addition, by the well known result of Serre, \( \hat{H} \) is a free pro-p group. It follows that \( \hat{H} \cong \mathbb{Z}_p \) and that there exists an element \( 1 \neq k \in \hat{H} \) such that \( \hat{H} = \langle \langle k \rangle \rangle \). As \( \hat{F} \) is pro-nilpotent, there exists \( n \) such that \( k \in \hat{F}_n - \hat{F}_{n+1} \). So \( \hat{H} \subseteq \hat{F}_n \). Consider the homomorphic image of \( \hat{H} \) into the quotient \( \mathbb{Y}_n(\hat{F}) = \hat{F}_n/\hat{F}_{n+1} \). As \( \mathbb{Y}_n(\hat{F}) \) is known to be isomorphic to \( \mathbb{Z}_p \) for some \( l \) (See [L], Proposition 2.7), the image of \( \hat{H} \) in \( \mathbb{Y}_n(\hat{F}) \) is a procyclic pro-p group of infinite order. Hence, this image is isomorphic to \( \mathbb{Z}_p \cong \hat{H} \). As \( \hat{H} \) is Hopfian, it follows that the homomorphism of \( \hat{H} \) into \( \mathbb{Y}_n(\hat{F}) \) is an embedding. In particular, the image of \( h, g \) in \( \mathbb{Y}_n(\hat{F}) \) is non-trivial, so the claim follows. \( \square \)

**Corollary 2.19.** As we assume that \( \hat{G}' = \langle \langle 1 + x, 1 + y \rangle \rangle \) is isomorphic to \( \hat{F}' = \langle \langle X, Y \rangle \rangle \) through the map \( X \mapsto 1 + x \) and \( Y \mapsto 1 + y \), for every \( 1 \neq g \in \hat{G}' \) we have \( [(1 + x)^m, [y]] \neq 1 \) for every \( m \geq 1 \).

### 2.6 Minimal \( t(x) \)-related components

For every \( n \geq 0 \) denote the following \( T \)-submodules of \( J \):

\[
\begin{align*}
J_n &= t(x)^nJ = t(x)^n[x, y](xT + yT + [x, y]T + xyT) \subseteq J \\
\tilde{J}_n &= t(x)^n[x, y](xT + yT) \subseteq J_n \subseteq J \\
C_n &= t(x)^n[x, y]^2T \subseteq J_n \subseteq J.
\end{align*}
\]

One can see that as a corollary from Proposition 2.10, the \( T \)-module \( J_n \) is a two sided ideal of \( R \). The following proposition follows from Proposition 2.10 and Part 3 of Lemma 2.8 and due to that as a free commutative ring, \( T \) is a unique factorization domain.

**Proposition 2.20.** The \( T \)-modules \( J_n, \tilde{J}_n, C_n \) are free, and for every \( a \in J \) we have

\[
\begin{align*}
a \in J_n &\iff t(x)a \in J_{n+1} \iff t(y)a, [x, y]^2a \in J_n \\
a \in \tilde{J}_n &\iff t(x)a \in \tilde{J}_{n+1} \iff t(y)a, [x, y]^2a \in \tilde{J}_n \\
a \in C_n &\iff t(x)a \in C_{n+1} \iff t(y)a, [x, y]^2a \in C_n.
\end{align*}
\]

By the above, it is easy to see that \( J_n = \tilde{J}_n \oplus C_n \oplus t(x)^nT \cdot [x, y]xy \).

**Proposition 2.21.** Let \( a \in J_n \). If \( t(a) = 0 \), then \( a \in \tilde{J}_n + C_n \). I.e.

\[
a = \alpha \cdot [x, y]x + \beta \cdot [x, y]y + \gamma \cdot [x, y]^2
\]

for some \( \alpha, \beta, \gamma \in t(x)^nT \).
Proof. As \( a \in J_n \) it has the form
\[
a = \alpha \cdot [x, y]x + \beta \cdot [x, y]y + \gamma \cdot [x, y]^2 + \delta \cdot [x, y]xy
\]
for some \( \alpha, \beta, \gamma, \delta \in t(x)^n T \). Now, as \( t([x, y]^2) = t([x, y]x) = t([x, y]y) = 0 \) we have
\[
0 = t(a) = \delta \cdot t([x, y]xy) = \delta \cdot [x, y]^2.
\]
As \( T \) is a domain, \( \delta = 0 \), as required.

We define now the following objects. Denote the subring of \( S \) generated by \( t(y) \) and \( [x, y]^2 \) by \( \tilde{S} = \langle t(y), [x, y]^2 \rangle \). Then, for every \( n \), we denote the \( \tilde{S} \)-module
\[
\tilde{J}_n = t(x)^n (\tilde{S} + t(xy)\tilde{S})[x, y]x + (\tilde{S} + t(xy)\tilde{S})[x, y]y [x, y]y \subseteq \tilde{J}_n.
\]
From Proposition 2.12 one deduces that \( S = \oplus_{i=0}^{\infty} t(x)^i \cdot \tilde{S} \), and therefore \( T = S + t(xy)S = \oplus_{i=0}^{\infty} t(x)^i (\tilde{S} + t(xy)\tilde{S}) \). Thus:

**Proposition 2.22.** For every \( n \geq 0 \) one has \( \tilde{J}_n = \oplus_{i=n}^{\infty} \tilde{J}_i = \tilde{J}_n \oplus \tilde{J}_{n+1} \) as a direct sum.

We continue with the following definition:

**Definition 2.23.** We define \( \tilde{G} \) to be the subset of \( \hat{G}' \) of all elements that all their terms are in \( J_3 \):
\[
\tilde{G} = \left\{ g = 1 + \sum_{i=1}^{\infty} a_i \in \hat{G}' | \forall i \ a_i \in J_3 \right\}.
\]

Using the identity \((1 + a)(1 + b)(1 + a)^{-1} = 1 + b + [a, b] \sum_{i=0}^{\infty} a^i \) and the fact that \( J_3 \) is a two sided ideal of \( R \), it is easy to see that:

**Proposition 2.24.** The set \( \tilde{G} \) is a normal closed subgroup of \( \hat{G}' \).

The reason for taking \( J_3 \) in the definition of \( \tilde{G} \) will be clear later. In the following definition, given an element \( a \in J_n = \tilde{J}_n \oplus \hat{G}_n \oplus t(x)^n T \cdot [x, y]xy \) we denote its projection to \( J_n \) by \( \bar{a} \).

**Definition 2.25.** Let \( 1 \neq g \in \tilde{G} \) and write \( g = 1 + \sum_{i=0}^{\infty} a_i \), where \( a_i \) is the term of degree \( i \). We define
\[
\begin{align*}
n(g) &= \max\{n \mid \forall i \ a_i \in J_n\} \\
\bar{n}(g) &= \max\{n \mid \forall i \ \bar{a}_i \in \tilde{J}_n\} \\
\bar{i}(g) &= \min\{i \mid \bar{a}_i \notin \tilde{J}_{\bar{n}(g)+1}\}.
\end{align*}
\]
Notice that as \( g \neq 1 \), \( n(g) \) is well defined. If \( g \) does not have terms in \( \tilde{J}_n \), for any \( n \) we denote \( \bar{n}(g) = \bar{i}(g) = \infty \). Notice that by definition \( n(g) \leq \bar{n}(g) \). If \( \bar{n}(g) < \infty \), by Proposition 2.22, for \( \bar{a}_{\bar{i}(g)} \) we can decompose
\[
\bar{a}_{\bar{i}(g)} = a + \text{elements from } \tilde{J}_{\bar{n}(g)+1}
\]
for a unique element \( a \in \tilde{J}_{\bar{n}(g)} \). In this case we define \( \min_x(g) = a \).
**Definition 2.26.** We say that an element $1 \neq g \in \hat{G}$ is **good** if $\bar{n}(g) < \infty$ and it has the form

$$g = 1 + \min_x(g) + \text{terms in } \tilde{J}_{\bar{n}(g)} \text{ of degree } > \deg(\min_x(g)) + \text{terms in } \tilde{J}_{\bar{n}(g)+1} + \text{terms in } C_{\bar{n}(g)-1} + \text{terms in } J_{\bar{n}(g)+2}.$$ 

**Proposition 2.27.** Let $g, h$ be good elements such that $n_0 = \bar{n}(g) = \bar{n}(h)$, $i_0 = \bar{i}(g) = \bar{i}(h)$ and $\min_x(g) \neq \min_x(h)$. Then $g \cdot h$ is also good, $\bar{n}(gh) = n_0$, $\bar{i}(gh) = i_0$, and

$$\min_x(g \cdot h) = \min_x(g) + \min_x(h).$$

**Proof.** We remind that by the definition of $\hat{G}$, the terms of $g, h$ are in $J_3 \subseteq t(x)^3 R$. Hence, $g \cdot h$ has the form

$$g \cdot h = 1 + \min_x(g) + \min_x(h) + \text{terms in } \tilde{J}_{n_0} \text{ of degree } > i_0 + \text{terms in } \tilde{J}_{n_0+1} + \text{terms in } C_{n_0-1} + \text{terms in } J_{n_0+2}.$$ 

By assumption $0 \neq \min_x(g) + \min_x(h) \in \tilde{J}_{n_0}$ and of degree $i_0$. By the above description of $g \cdot h$ it follows that $\bar{n}(g \cdot h) = n_0$, $\bar{i}(g \cdot h) = i_0$, and that $\min_x(g \cdot h) = \min_x(g) + \min_x(h)$. Hence, $g \cdot h$ is good, as required. \qed

**Definition 2.28.** Let $1 \neq g \in \hat{G}$. We define the operator

$$\varphi_x(g) = [(1 + x)^4, g]_{\hat{G}} = [1 + t(x)^3x, g]_{\hat{G}}.$$ 

When it will be convenient, we will use also the notation $g_x = \varphi_x(g)$.

**Proposition 2.29.** Let $1 \neq g = 1 + \sum_{i=0}^{\infty} a_i \in \hat{G}$. Then $n(g_x) = n_0 + 3$ where $n_0 = \max\{n \mid \forall i \ n_0 \in J_n\}$, and

$$g_x = 1 + t(x)^3 \sum_{i=0}^{\infty} [x, a_i] + \text{terms in } J_{n(g_x)+3}.$$ 

**Proof.** Notice first that as $g \neq 1$ and $\hat{G} \subseteq \hat{G}'$, then by Corollary $g$ does not commute with $1 + x$, and hence there exists $i$ such that $[x, a_i] \neq 0$. Thus, $n_0$ is well defined. We claim now that $n(g_x) = n_0 + 3$. Indeed, denote $g_x = 1 + \sum_{i=0}^{\infty} b_i$, and write

$$g_x = 1 + \sum_{i=0}^{\infty} b_i = 1 + t(x)^3[x, \sum_{i=0}^{\infty} a_i] \sum_{k,l=0}^{\infty} (t(x)^3)^k (\sum_{i=0}^{\infty} a_i)^l.$$ 

17
Now, by assumption, for every \( i \) we have \( a_i \in J_3 \subseteq t(x)^3 R \), and \([x, a_i] \in J_{n_0}\). Hence, for every \( i \) we have
\[
b_{i+4} = t(x)^3[x, a_i] \mod J_{n_0+3+3}.
\]
Hence \( b_{i+4} \in J_{n_0+3} \) for every \( i \), and thus \( n(g_x) \geq n_0 + 3 \). On the other hand, by assumption, there exists \( i_0 \) such that \([x, a_{i_0}] \notin J_{n_0+1} \), so by Proposition 2.20 \( t(x)^3[x, a_{i_0}] \notin J_{n_0+4} \). Therefore
\[
b_{i_0+4} = t(x)^3[x, a_{i_0}] + \text{terms in } J_{n_0+6}
\]
is not in \( J_{n_0+4} \). Thus \( n(g_x) = n_0 + 3 \). Notice now that as \( n_0 + 6 = n(g_x) + 3 \), in the course of the proof we got also that \( b_{i+4} = t(x)^3[x, a_i] \mod J_{n(g_x)+3} \) for every \( i \), as required.

**Corollary 2.30.** Let \( 1 \neq g \in \mathcal{G} \). Then \( g_x \) is good, and \( n(g_x) \leq \bar{n}(g_x) \leq n(g_x) + 1 \).

**Proof.** Write \( g = 1 + \sum_{i=0}^\infty a_i \), and denote
\[
a_i = a_i [x, y] x \beta_i + \gamma_i [x, y]^2 + \delta_i [x, y] xy
\]
for \( \alpha_i, \beta_i, \gamma_i, \delta_i \in T \). By Lemmas 2.7 and 2.8 one has
\[
[x, a_i] = (t(x) \alpha_i + t(xy) \delta_i)[x, y] x + t(x) \beta_i [x, y] y + \beta_i [x, y]^2.
\]
Set \( \lambda_i = t(x) \alpha_i + t(xy) \delta_i \). Now, for \( g_x = 1 + \sum_{i=0}^\infty b_i \), let \( i \) be an index such that \( b_{i+4} \notin J_{n(g_x)+3} \) (which exists by definition). Then, by the previous proposition
\[
b_{i+4} = t(x)^3[x, a_i] + \text{terms in } J_{n(g_x)+3}
\]
\[
= t(x)^3(\lambda_i [x, y] x + t(x) \beta_i [x, y] y + \beta_i [x, y]^2) + \text{terms in } J_{n(g_x)+3}
\]
Hence, by definition, it means that we have two options:

1. \( t(x)^3 \lambda_i \notin t(x)^{n(g_x)+1} T \). In this case we also have \( \bar{b}_{i+4} \notin \bar{J}_{n(g_x)+1} \). In particular \( \bar{n}(g_x) \leq n(g_x) \). Otherwise:

2. \( t(x)^3 \beta_i \notin t(x)^{n(g_x)+1} T \). Hence \( t(x)^4 \beta_i \notin t(x)^{n(g_x)+2} T \). So in this case we have \( \bar{b}_{i+4} \notin \bar{J}_{n(g_x)+2} \). In particular \( \bar{n}(g_x) \leq n(g_x) + 1 \).

In both cases we get \( \bar{n}(g_x) \leq n(g_x) + 1 \) so in total \( n(g_x) \leq \bar{n}(g_x) \leq n(g_x) + 1 \). In particular, \( \bar{n}(g_x) < \infty \). Hence \( \min_x(g_x) \) exists and we have
\[
g_x = 1 + t(x)^3 \sum_{i=0}^\infty \left( \lambda_i [x, y] x + t(x) \beta_i [x, y] y + \beta_i [x, y]^2 \right) + \text{terms in } J_{n(g_x)+3}
\]
\[
= 1 + \min_x(g_x)
\]
+ terms in \( \bar{J}_{n(g_x)} \) of degree \( > \deg(\min_x(g_x)) \)
+ terms in \( \bar{J}_{n(g_x)+1} \)
+ terms in \( \bar{C}_{n(g_x)} \)
+ terms in \( J_{n(g_x)+3} \).

(2.7)
Combining the latter estimation for $\bar{n}(g_x)$ with Equation (2.7) we get that $g_x$ is good, as required.

**Proposition 2.31.** Let $g \in \tilde{G}$ be a good element. Then $\varphi_x(g)$ is also good, and

$$\min_x(\varphi_x(g)) = t(x)^4 \min_x(g)$$

In particular $\bar{n}(\varphi_x(g)) = \bar{n}(g) + 4$.

**Proof.** So the fact that $\varphi_x(g) = g_x$ is good was already proven previously in a more general case. Now, write $g = 1 + a$ for $a$ of the form

$$a = \min_x(g)$$

+ terms in $\bar{J}_{\bar{n}(g)}$ of degree $> \deg(\min_x(g))$
+ terms in $\bar{J}_{\bar{n}(g)+1}$
+ terms in $\bar{C}_{\bar{n}(g)-1}$
+ terms in $\bar{J}_{\bar{n}(g)+2}$.

We remind that the terms of $g$ lie in $J_3 \subseteq t(x)^3R$. Hence

$$\varphi_x(g) = 1 + t(x)^3 [x,a] \sum_{k,l=0}^\infty (t(x)^3 x)^k a^l = 1 + t(x)^3 [x,a] + \text{terms in } J_{\bar{n}(g)+6}.$$

Noticing the identities

$$[x, [x,y]x] = t(x)[x,y]x \quad \text{and} \quad [x, [x,y]y] = t(y)[x,y]x + [x,y]^2$$

we have:

- $[x, \min_x(g)] = t(x)\min_x(g)$ + terms in $\bar{C}_{\bar{n}(g)}$.
- $[x, \text{terms in } \bar{J}_{\bar{n}(g)} \text{ of degree } > \deg(\min_x(g))]$
  = terms in $\bar{J}_{\bar{n}(g)+1} \text{ of degree } > \deg(t(x)\min_x(g))$
  + terms in $\bar{C}_{\bar{n}(g)}$.
- $[x, \text{terms in } \bar{J}_{\bar{n}(g)+1}]$ = terms in $\bar{J}_{\bar{n}(g)+2}$ + terms in $\bar{C}_{\bar{n}(g)+1}$.
- $[x, \text{terms in } \bar{C}_{\bar{n}(g)-1}] = 0$.
- $[x, \text{terms in } \bar{J}_{\bar{n}(g)+2}]$ = terms in $\bar{J}_{\bar{n}(g)+2}$.

Hence

$$\varphi_x(g) = 1 + t(x)^3 [x,a] + \text{terms in } J_{\bar{n}(g)+6}$$

$$= t(x)^4 \min_x(g)$$

+ terms in $\bar{J}_{\bar{n}(g)+4}$ of degree $> \deg(t(x)^4 \min_x(g))$
+ terms in $\bar{J}_{\bar{n}(g)+5}$
+ terms in $\bar{C}_{\bar{n}(g)+3}$
+ terms in $\bar{J}_{\bar{n}(g)+5}$.
Recalling Proposition 2.20, it follows that \( \bar{n}(\varphi_x(g)) = \bar{n}(g) + 4 \), and that we have \( \min_x(\varphi_x(g)) = t(x)^3 \min_x(g) \) as required.

**Proposition 2.32.** Let \( g \in \tilde{G} \) be a good element, and let \( \varphi_y \) be the operator \( \varphi_y(g) = [1 + y, g]_{\tilde{G}} \). Then, \( \varphi_y(g) \) is also good, and
\[
\min_x(\varphi_y(g)) = t(y) \min_x(g).
\]
In particular \( \bar{n}(\varphi_y(g)) = \bar{n}(g) \).

**Proof.** Write \( g = 1 + a \) where \( a \) is built up from
\[
a = \min_x(g)
\]
\[
\quad + \text{terms in } J_{\bar{n}(g)} \text{ of degree } > \deg(\min_x(g))
\]
\[
\quad + \text{terms in } J_{\bar{n}(g)+1}
\]
\[
\quad + \text{terms in } C_{\bar{n}(g)-1}
\]
\[
\quad + \text{terms in } J_{\bar{n}(g)+2}.
\]
In addition, we can think on \( a \) as built up from terms in \( t(x)^3 R \). Hence
\[
\varphi_y(g) = 1 + [y, a] \sum_{k=0}^{\infty} y^k a^l = 1 + [y, a] \sum_{k=0}^{\infty} y^k + \text{terms in } J_{\bar{n}(g)+3}.
\]

Now, using Part 4 of Lemma 2.7, Proposition 2.21 and the identities
\[
[y, [x, y]x] = t(y)[x, y]x + [x, y]^2, \quad [y, [x, y]y] = t(y)[x, y]y
\]
we have:

- \([y, \min_x(g)] \sum_{k=0}^{\infty} y^k\)
  \[= t(y) \min_x(g) + \text{terms in } C_{\bar{n}(g)} + \text{terms in } J_{\bar{n}(g)} \text{ of degree } > \deg(t(y) \min_x(g))\]
- \([y, \text{terms in } J_{\bar{n}(g)} \text{ of degree } > \deg(\min_x(g))] \sum_{k=0}^{\infty} y^k\)
  \[= \text{terms in } J_{\bar{n}(g)} \text{ of degree } > \deg(t(y) \min_x(g)) + \text{terms in } C_{\bar{n}(g)}\]
- \([y, \text{terms in } J_{\bar{n}(g)+1}] \sum_{k=0}^{\infty} y^k = \text{terms in } J_{\bar{n}(g)+1} + \text{terms in } C_{\bar{n}(g)+1}\]
- \([y, \text{terms in } C_{\bar{n}(g)-1}] \sum_{k=0}^{\infty} y^k = 0\]
- \([y, \text{terms in } J_{\bar{n}(g)+2}] \sum_{k=0}^{\infty} y^k = \text{terms in } J_{\bar{n}(g)+2}\].
Hence

\[ \varphi_y(g) = 1 + [y, a] \sum_{k=0}^{\infty} y^k + \text{terms in } J_{\bar{n}(g)+3} \]

\[ = 1 + t(y) \min_x(g) \]

+ terms in \( J_{\bar{n}(g)+1} \)

+ terms in \( C_{\bar{n}(g)} \)

Recalling Proposition 2.20, we obtain the assertions in the proposition. □

Proposition 2.33. Let \( g \in \bar{G} \) be a good element, and let \( \psi \) be the operator \( \psi(g) = [(1 + x, 1 + y)_{\bar{G}}^2, g]_{\bar{G}} \). Then \( \psi(g) \) is also good, and

\[ \min_x(\psi(g)) = [x, y]^4 \min_x(g) \]

In particular \( \bar{n}(\psi(g)) = \bar{n}(g) \).

Proof. As in the previous proposition, we write \( g = 1 + a \) where \( a \) can be written as

\[ a = \min_x(g) \]

+ terms in \( J_{\bar{n}(g)+1} \)

+ terms in \( C_{\bar{n}(g)} \)

In addition, we think on \( a \) as built up from terms in \( t(x)^3 R \).

Denote \( ([1 + x, 1 + y]_{\bar{G}})^2 = 1 + b. \) then

\[ \psi(g) = 1 + [b, a] \sum_{k,l=0}^{\infty} b^k a^l = 1 + [b, a] \sum_{k=0}^{\infty} b^k + \text{terms in } J_{\bar{n}(g)+3}. \]

Notice that for every \( p, q \in M_2(A) \) we have \( t(pq + qp) = t([p, q]) = 0. \) Also, notice that \( ([x, y]xy)^2 = [x, y]^3 xy \) and that \( ([x, y]x)^2 = ([x, y]y)^2 = 0. \) Hence, a direct computation shows that

\[ 1 + b = 1 + ([x, y] \sum_{i,j=0}^{\infty} x^iy^j)^2 \]

\[ = 1 + [x, y]^2 + [x, y]^3 xy \]

+ terms of zero trace of degree \( \geq 5 \)

+ terms of degree \( \geq 9. \)
Observe that by Part 1 of Lemma 2.7 and Corollary 2.21 for $p,q \in J$ such that $t(p) = t(q) = 0$ we have $[p,q] \in [x,y]^2T$. Hence, using Part 4 of Lemma 2.7, Corollary 2.21, and the identities

$$ [[x,y]xy, [x,y]x] = [x,y]^2 \cdot [x,y]x, \quad [[x,y]xy, [x,y]y] = [x,y]^2 \cdot [x,y]y. $$

We have:

- $[b, \min_x(g)] \sum_{k=0}^{\infty} b^k$
  
  $$ = [b, \min_x(g)] + \text{terms in } \tilde{J}_{\bar{n}(g)} \text{ of degree } > \deg([x,y]^{4}\min_x(g)) $$
  $$ + \text{terms in } \tilde{C}_{\bar{n}(g)} $$
  $$ = [[x,y]^3xy, \min_x(g)] + \text{terms in } \tilde{J}_{\bar{n}(g)} \text{ of degree } > \deg([x,y]^{4}\min_x(g)) $$
  $$ + \text{terms in } \tilde{C}_{\bar{n}(g)} $$
  $$ = [x,y]^{4}\min_x(g) + \text{terms in } \tilde{J}_{\bar{n}(g)} \text{ of degree } > \deg([x,y]^{4}\min_x(g)) $$
  $$ + \text{terms in } \tilde{C}_{\bar{n}(g)} $

- $[b, \text{terms in } \tilde{J}_{\bar{n}(g)} \text{ of degree } > \deg(\min_x(g))] \sum_{k=0}^{\infty} b^k$
  
  $$ = [b, \text{terms in } \tilde{J}_{\bar{n}(g)} \text{ of degree } > \deg(\min_x(g))] $$
  $$ + \text{terms in } \tilde{J}_{\bar{n}(g)} \text{ of degree } > \deg([x,y]^{4}\min_x(g)) $$
  $$ + \text{terms in } \tilde{C}_{\bar{n}(g)} $$
  $$ = \text{terms in } \tilde{J}_{\bar{n}(g)} \text{ of degree } > \deg([x,y]^{4}\min_x(g)) $$
  $$ + \text{terms in } \tilde{C}_{\bar{n}(g)} $

- $[b, \text{terms in } \tilde{J}_{\bar{n}(g)+1}] \sum_{k=0}^{\infty} b^k = \text{terms in } \tilde{J}_{\bar{n}(g)+1} + \text{terms in } \tilde{C}_{\bar{n}(g)+1}$.

- $[b, \text{terms in } \tilde{C}_{\bar{n}(g)}] \sum_{k=0}^{\infty} b^k = 0.$

- $[b, \text{terms in } \tilde{J}_{\bar{n}(g)+2}] \sum_{k=0}^{\infty} b^k = \text{terms in } \tilde{J}_{\bar{n}(g)+2}$.

Hence

$$ \psi(g) = 1 + [b, a] \sum_{k=0}^{\infty} b^k + \text{terms in } \tilde{J}_{\bar{n}(g)+3} $$

$$ = 1 + [x,y]^{4}\min_x(g) $$

$$ + \text{terms in } \tilde{J}_{\bar{n}(g)} \text{ of degree } > \deg([x,y]^{4}\min_x(g)) $$

$$ + \text{terms in } \tilde{J}_{\bar{n}(g)+1} $$

$$ + \text{terms in } \tilde{C}_{\bar{n}(g)} $$

$$ + \text{terms in } \tilde{J}_{\bar{n}(g)+2}. $$

Recalling Proposition 2.20 we obtain the assertions in the proposition. □
2.7 Interim conclusions

Denote the ring (with identity) \( V = \langle t(x)^4, t(y), [x, y]^4 \rangle \subseteq S \), and the ideal \( I = t(x)^4 V \triangleleft V \). It is easy to see that \( S \) is a finitely generated \( V \)-module, and therefore \( J \) is a finitely generated \( V \)-module (see Propositions 2.12 and 2.16).

Define now the set
\[
M = \left\{ \min_x(g) \mid g \in \hat{G} \text{ is good} \right\}
\]
and define \( N \) to be the \( V \)-submodule of \( J \) generated by the elements of \( M \). So as \( V \) is Noetherian and \( J \) is a finitely generated \( V \)-module, by the Artin-Rees lemma, there exists a number \( \rho \) such that for every \( k \geq 0 \)
\[
M \cap (J_{4(\rho+k)}) \subseteq N \cap (J \cdot I^{\rho+k}) \subseteq N \cdot I^k.
\]
Fix this \( \rho \). Recall the lower central series of \( \hat{G} = \hat{G}_1, \hat{G}_2, \ldots \). We have:

**Proposition 2.34.** Let \( 1 \neq g \in \hat{G} \) such that \( \bar{n}(g_x) \geq 4\rho + 4k + 4 \) for some \( k \geq 1 \). Then, there exists an element \( h \in \hat{G} \cap \hat{G}_k \) that satisfies one of the following conditions

- \( \bar{n}((gh)_x) > \bar{n}(g_x) \) or
- \( \bar{n}((gh)_x) = \bar{n}(g_x) \) and \( \bar{i}((gh)_x) > \bar{i}(g_x) \).

**Proof.** We will do it in a few steps.

**Step 1 - Claim:** There exist good elements \( h_1, \ldots, h_m \) such that
\[
\min_x(g_x) = t(x)^{4k+4} \sum_{j=1}^m \min_x(h_j).
\]
Indeed: By Corollary 2.30, \( g_x \) is good. Hence, by assumption \( \min_x(g_x) \in M \cap J_{4(\rho+k+1)} \). From the Artin-Rees argument above it follows that \( \min_x(g_x) \in N \cdot I^{k+1} \). Hence, there exist good elements \( \hat{h}_1, \ldots, \hat{h}_m \in \hat{G} \) and elements \( \lambda_1, \ldots, \lambda_m \in I^{k+1} \) such that
\[
\min_x(g_x) = \lambda_1 \cdot \min_x(\hat{h}_1) + \cdots + \lambda_m \cdot \min_x(\hat{h}_m).
\]
Without loss of generality we can assume that each \( \lambda_j \) has the form
\[
\lambda_j = t(x)^{4u_j + 4k+4}t(y)^{v_j}[x, y]^{4w_j}
\]
for some numbers \( u_j, v_j, w_j \geq 0 \). Now, by Propositions 2.31, 2.32 and 2.33 we have
\[
\min_x(h_{j}) = t(x)^{4u_j}t(y)^{v_j}[x, y]^{4w_j}\min_x(\hat{h}_{j})
\]
where \( h_{j} = \psi^{u_j} \circ \varphi^{v_j} \circ \varphi^{w_j}(\hat{h}_{j}) \) are good elements, as required.
**Step 2 - Claim:** There exists a good element $h \in \tilde{G} \cap G_k$ such that
$$\min_x(g_x) = t(x)^4 \min_x(h).$$

We continue from the previous stage. By definition we have
$$\min_x(g_x) \in \tilde{J}_{\tilde{n}(g_x)}, \quad \deg(\min_x(g_x)) = \tilde{i}(g_x)$$
$$\min_x(h_j) \in \tilde{J}_{\tilde{n}(h_j)}, \quad \deg(\min_x(h_j)) = \tilde{i}(h_j) \ \forall j.$$

Denote
$$\Omega = \{j \mid 4k + 4 + \tilde{n}(h_j) = \tilde{n}(g_x), \ 4k + 4 + \tilde{i}(h_j) = \tilde{i}(g_x)\}.$$

By Proposition 2.22 it follows that
$$\min_x(g_x) = t(x)^{4k+4} \sum_{j \in \Omega} \min_x(h_j) \tag{2.8}$$

Hence, we can assume that for every $j$ we have $4k + 4 + \tilde{n}(h_j) = \tilde{n}(g_x), \ 4k + 4 + \tilde{i}(h_j) = \tilde{i}(g_x)$ and in particular $\tilde{n}(h_j) = \tilde{n}(h_j')$ and $\tilde{i}(h_j) = \tilde{i}(h_j')$ for every $j, j'$.

In addition, without loss of generality, we can assume that for every $1 \leq l \leq m - 1$ one has
$$\sum_{j=1}^{l} \min_x(h_j) \neq \min_x(h_{l+1}).$$

Otherwise, if $\sum_{j=1}^{l} \min_x(h_j) = \min_x(h_{l+1})$ for some $l$, we can omit $h_1, ..., h_{l+1}$ from the sum in Equation (2.8). Define
$$h = \varphi_x^k(\prod_{j=1}^{m} h_j) \in \tilde{G} \cap G_k.$$ 

By Proposition 2.27 $\prod_{j=1}^{m} h_j$ is good and
$$\min_x(\prod_{j=1}^{m} h_j) = \sum_{j=1}^{m} \min_x(h_j).$$

Hence, by Proposition 2.31 we get Step 2.

**Step 3:** We want now to show that $h$ satisfies the conditions in the proposition. By the previous stage and Proposition 2.31 we have $\tilde{n}(g_x) = \tilde{n}(h_x) = \tilde{n}(h) + 4$. Write $g = 1 + a = 1 + \sum_{i=0}^{\infty} a_i$ and $h = 1 + b = 1 + \sum_{i=0}^{\infty} b_i$, so
$$gh = 1 + a + b + ab.$$ 

Now, as $k \geq 1$, by the construction, $h$ is of the form $h = \varphi_x(h')$ for some good element $h' \in \tilde{G}$. Hence $\tilde{n}(h) - 1 \leq n(h)$ by Corollary 2.30. Thus, as the terms of $a$ lie in $t(x)^3 R$ we have
$$ab = \text{terms in } J_{n(h)+3} \subseteq J_{\tilde{n}(h)+2}.$$ 

24
Denote
\[ n_a = \max \{ n \mid \forall i \ [x, a_i] \in J_n \} \]
\[ n_b = \max \{ n \mid \forall i \ [x, b_i] \in J_n \}. \]

By Proposition 2.29 we have \( n(g_x) = n_a + 3 \), and \( n(h_x) = n_b + 3 \). Hence
\[ n_a = n(g_x) - 3 \geq \tilde{n}(g_x) - 4 = \tilde{n}(h) \]
\[ n_b = n(h_x) - 3 \geq \tilde{n}(h_x) - 4 = \tilde{n}(h). \]

Therefore
\[
(gh)_x = 1 + t(x)^3[x, a + b + ab] \sum_{i,j=0}^{\infty} (t(x)^3)^i(a + b + ab)^j
\]
\[
= 1 + t(x)^3[x, a] + t(x)^3[x, b] + \text{terms in } J_{\tilde{n}(h)+5}.
\]

Now, as in the proof of Corollary 2.30 we have
\[
t(x)^3[x, a] = \min_x(g_x) + \text{terms in } J_{\tilde{n}(g_x)} \text{ of degree } > \deg(\min_x(g_x)) + \text{terms in } \bar{J}_{\tilde{n}(g_x)+1} + \text{terms in } \bar{C}_{\tilde{n}(g_x)-1}
\]
\[
t(x)^3[x, b] = \min_x(h_x) + \text{terms in } J_{\tilde{n}(h_x)} \text{ of degree } > \deg(\min_x(h_x)) + \text{terms in } \bar{J}_{\tilde{n}(h_x)+1} + \text{terms in } \bar{C}_{\tilde{n}(h_x)-1}.
\]

Now, as \( h \) is good, we have
\[
\min_x(h_x) = t(x)^4 \min_x(h) = \min_x(g_x)
\]
by the construction of \( h \) and Proposition 2.31. In addition, \( \tilde{n}(g_x) = \tilde{n}(h) + 4 = \tilde{n}(h_x) \). Hence, we get that
\[
(gh)_x = 1 + t(x)^3[x, a] + t(x)^3[x, b] + \text{terms in } J_{\tilde{n}(g_x)+1}
\]
\[
= 1 + \text{terms in } J_{\tilde{n}(g_x)} \text{ of degree } > \deg(\min_x(g_x)) + \text{terms in } \bar{C}_{\tilde{n}(g_x)-1} + \text{terms in } \bar{J}_{\tilde{n}(g_x)+1}.
\]

Which is equivalent to our assertion. \( \square \)

Denote now the ring (with identity) \( \hat{V} = \langle t(y), [x, y]^4 \rangle \subseteq \hat{S} = \langle t(y), [x, y]^2 \rangle \),
and let \( \hat{I} < \hat{V} \) be the ideal generated by \( t(y), [x, y]^4 \). In other words \( \hat{I} = \hat{V} - \{1\} \).
It is easy to see that \( \hat{S} \) is a finitely generated \( \hat{V} \)-module, and hence, for every \( n \)
\[
\hat{J}_n = t(x)^n((\hat{S} + t(xy)\hat{S})[x, y]x + (\hat{S} + t(xy)\hat{S})[x, y]y)
\]

25
is a finitely generated $\tilde{V}$-module.

Define now the set $M_n = \{ \min_x(g) \in \tilde{J}_n \mid g \in \tilde{G} \text{ is good} \}$, and define $N_n$ to be the $\tilde{V}$-submodule of $\tilde{J}_n$ generated by the elements of $M_n$. So as $\tilde{V}$ is Noetherian and $\tilde{J}_n$ is a finitely generated $\tilde{V}$-module, by the Artin-Rees lemma, for every $n$, there exists a number $\rho_n$ such that for every $k \geq 0$

$$M_n \cap (\tilde{J}_n \cdot \tilde{I}^{\rho_n+k}) \subseteq N_n \cap (\tilde{J}_n \cdot \tilde{I}^{\rho_n+k}) \subseteq N_n \cdot \tilde{I}^k$$

Fix these numbers $\rho_n$, and recall $\rho$ from Proposition 2.34. We have:

**Proposition 2.35.** Let $1 \neq g \in \tilde{G}$, and assume that $n = \tilde{n}(g_x) \geq 4\rho + 8$. Set $k \geq 1$. If

$$\tilde{i}(g_x) \geq n + 5 + 8(k + \rho_{n} - s)$$

then, there exists an element $h \in \tilde{G} \cap \tilde{G}_k$ that satisfies one of the following conditions

- $\tilde{n}((gh)_x) > \tilde{n}(g_x)$ or
- $\tilde{n}((gh)_x) = \tilde{n}(g_x)$ and $\tilde{i}((gh)_x) > \tilde{i}(g_x)$.

**Proof.** Under the assumption $n = \tilde{n}(g_x) \geq 4\rho + 8$, we saw in the proof of the previous proposition that there exists $\tilde{h} \in \tilde{G}$ of the form $\phi_x(\tilde{h}') = \tilde{h}$ where $\tilde{h}'$ is good, such that $\min_x(g_x) = t(x)^4\min_x(\tilde{h})$. In other words, there exists a good element $\tilde{h}' \in \tilde{G}$ such that $\min_x(g_x) = t(x)^8\min_x(\tilde{h}')$. In particular

$$\tilde{i}(g_x) = \tilde{i}(\tilde{h}') + 8$$

$$n = \tilde{n}(g_x) = \tilde{n}(\tilde{h}') + 8.$$ 

Hence

$$\deg(\min_x(\tilde{h}')) = \tilde{i}(\tilde{h}') = \tilde{i}(g_x) - 8 \geq n - 8 + 8(k + \rho_{n} - s) + 5.$$ 

It follows that $\min_x(\tilde{h}')$ has the form

$$\min_x(\tilde{h}') = t(x)^n - s \cdot ((\alpha + t(xy)\beta)[x,y]x + (\gamma + t(xy)\delta)[x,y]y$$

for some $\alpha, \beta, \gamma, \delta \in \tilde{S}$ of degree $\geq 8(k + \rho_{n} - s)$. In particular $\alpha, \beta, \gamma, \delta \in \tilde{S}^n + k$. It follows that

$$\min_x(\tilde{h}') \in M_{n-s} \cap (\tilde{J}_{n-s} \cdot \tilde{I}^{\rho_{n-s}+k}) \subseteq N_{n-s} \cdot \tilde{I}^k$$

and hence there exist good elements $\tilde{h}_1, ..., \tilde{h}_m \in \tilde{G}$ with $\min_x(\tilde{h}_j) \in \tilde{J}_{n-s}$ and elements $\lambda_1, ..., \lambda_m \in \tilde{I}^k$ such that

$$\min_x(\tilde{h}') = \lambda_1 \cdot \min_x(\tilde{h}_1) + ... + \lambda_m \cdot \min_x(\tilde{h}_m).$$

Without loss of generality we can assume that each $\lambda_j$ has the form

$$\lambda_j = t(y)^{v_j}[x,y]^{4w_j}.$$
for some numbers \( v_j, w_j \geq 0 \) such that \( v_j + w_j \geq k \). Now, by Propositions 2.32 and 2.33 we have
\[
\min_x(h_j) = t(y)^{v_j}[x,y]^{4w_j}\min_x(\tilde{h}_j).
\]
for \( h_j = \psi^{v_j} \circ \varphi^{w_j}(\tilde{h}_j) \in \tilde{G} \cap \hat{G}_k \). Clearly \( \bar{n}(h_j) = n - 8 \) and without loss of generality we can assume that \( \bar{i}(h_j) = \bar{i}(\tilde{h'}) \) for every \( j \). In addition, without loss of generality, like in the proof of Proposition 2.34 we can assume that for every \( 1 \leq l \leq m - 1 \) one has
\[
\sum_{j=1}^{l} \min_x(h_j) \neq \min_x(h_{l+1}).
\]
We define now
\[
h = \varphi_x(\prod_{j=1}^{m} h_j) \in \tilde{G} \cap \hat{G}_k.
\]
Then, by Propositions 2.27 and 2.31 we have
\[
\min_x(h) = t(x)^4 \sum_{j=1}^{m} \min_x(h_j) = t(x)^4 \sum_{j=1}^{m} t(y)^{v_j}[x,y]^{4w_j}\min_x(\tilde{h}_j)
= t(x)^4 \min_x(\tilde{h'}) = \min_x(\tilde{h}).
\]
Therefore, we got \( h \in \tilde{G} \cap \hat{G}_k \) that has the form \( h = \varphi_x(h') \) for some \( h' \in \tilde{G} \) and \( \min_x(g_x) = t(x)^4\min_x(h) \). Hence, like in Step 3 in the proof of Proposition 2.34 we get that
\[
(gh)_x = 1 + \text{terms in } \bar{J}_{\bar{n}(g_x)} \text{ of degree } > \deg(\min_x(g_x))
+ \text{terms in } \bar{C}_{\bar{n}(g_x)-1}
+ \text{terms in } J_{\bar{n}(g_x)+1},
\]
as required. \( \square \)

2.8 Proof of Theorem 2.2

Let \( v \in \tilde{G}'' \). By Corollary 2.17 all the terms of \( v \) lie in \( J \). By Corollary 2.17 as \( v \in \tilde{G}'' \subseteq \hat{G}' \) it does not commute with \((1 + x)^2\). Therefore, \([v,(1+x)^2]_{\tilde{G}} \neq 1\). As \( \tilde{G} \) is pro nilpotent, there exists a unique \( s \) such that
\[
[v,(1+x)^2]_{\tilde{G}} \in \tilde{G}_s - \tilde{G}_{s+1}.
\]
where \( \tilde{G} = \tilde{G}_1, \tilde{G}_2, \tilde{G}_3, \ldots \) is the lower central series of \( \tilde{G} \). As by the negative assumption, \( \hat{G}' \cong \hat{F}' = \langle [X,Y] \rangle' \) through the map \( X \mapsto 1 + x \) and \( Y \mapsto 1 + y \), the quotient \( Y_s(\hat{G}) = \hat{G}_s/\hat{G}_{s+1} \) is a free abelian pro-2 group [E]. Hence
Let \( ([v, (1 + x)^2]_\mathcal{G})^{2^r} \in \mathcal{G}_s - \mathcal{G}_{s+1} \) for every \( r \). Recall \( \rho \) from Proposition 2.34 and choose \( r \) large enough such that \( 2^r \geq 4\rho + 4(s + 1) + 4 \). Define

\[
g = ([v, (1 + x)^2]_\mathcal{G})^{2^r} = ([v, 1 + t(x)x]_\mathcal{G})^{2^r}.
\]

As obviously, all the terms of \([v, (1 + x)^2]_\mathcal{G} = [v, 1 + t(x)x]_\mathcal{G}\) lie in \( t(x)J \), it follows that all the terms of \( g \) lie in \( t(x)^2 J = J_{2^r} \). In particular \( g \in \mathcal{G} \) and \( \tilde{n}(g_x) \geq n(g_x) \geq 4\rho + 4(s + 1) + 4 \). Now, we have two options:

- \( \tilde{i}(g_x) \leq \tilde{n}(g_x) + 5 + 8(s + 1 + \rho \tilde{n}(g_x) - s) \). In this case, by Proposition 2.34 we construct an element \( h_1 \) such that \( h_1 \in \mathcal{G} \cap \mathcal{G}_{s+1} \) and

  - \( \tilde{n}((gh_1)_x) > \tilde{n}(g_x) \) or
  - \( \tilde{n}((gh_1)_x) = \tilde{n}(g_x) \) and \( \tilde{i}((gh_1)_x) > \tilde{i}(g_x) \).

- \( \tilde{i}(g_x) > \tilde{n}(g_x) + 5 + 8(s + 1 + \rho \tilde{n}(g_x) - s) \) and \( k_1 \geq 1 \) is the largest with this property. In this case, by Proposition 2.35 we construct an element \( h_1 \) such that \( h_1 \in \mathcal{G} \cap \mathcal{G}_{s+k_1} \) and

  - \( \tilde{n}((gh_1)_x) > \tilde{n}(g_x) \) or
  - \( \tilde{n}((gh_1)_x) = \tilde{n}(g_x) \) and \( \tilde{i}((gh_1)_x) > \tilde{i}(g_x) \).

We proceed in this way. There are two options:

- For some \( i \) we will have: \( \tilde{n}((gh_1 \cdot \ldots \cdot h_i)_x) > \tilde{n}(g_x) \). In the first time it happens we define \( w_1 = h_1 \cdot \ldots \cdot h_i \).

- For every \( i \) we have \( \tilde{n}((gh_1 \cdot \ldots \cdot h_i)_x) = \tilde{n}(g_x) \). In this case

\[
\tilde{i}((gh_1 \cdot \ldots \cdot h_i)_x) \xrightarrow{i \to \infty} \infty
\]

and therefore \( k_i \xrightarrow{i \to \infty} \infty \). Hence, the sequence \( \mathcal{G} \cap \mathcal{G}_{s+k_i} \ni h_i \xrightarrow{i \to \infty} 1 \), and therefore, the sequence \( h_1 \cdot \ldots \cdot h_i \) converges to an element \( w_i \in \mathcal{G} \cap \mathcal{G}_{s+1} \) with the property

\[
\tilde{n}((gw_i)_x) > \tilde{n}(g_x).
\]

We proceed in this way until

\[
\tilde{n}((gw_1 \cdot \ldots \cdot w_i)_x) \geq 4\rho + 4(s + 2) + 4
\]

and in a similar way we construct the next \( w_i \)-s so that they will lie also in \( w_i \in \mathcal{G} \cap \mathcal{G}_{s+2} \). We proceed in this pattern.

At the end of the process we get a sequence \( w_i \xrightarrow{i \to \infty} 1 \) and therefore the product converges to an element

\[
h = \lim_{i \to \infty} w_1 \cdot \ldots \cdot w_i \in \mathcal{G} \cap \mathcal{G}_{s+1}
\]

with the property \( (gh)_x = 1 \). In particular, the element \( gh \in \mathcal{G}_s - \mathcal{G}_{s+1} \) commutes with \( (1 + x)^2 \), and this is a contradiction to Corollary 2.19. This finishes the proof of Theorem 2.2.
3 Some remarks regarding $p = 2$ versus $p \neq 2$

3.1 The Universal Representation

In this section we describe where exactly Zubkov’s approach fails when $p = 2$, and that in some sense, $2 \times 2$ pro-$2$ linear groups indeed have less pro-$2$ identities (if any). We note that in this section we use similar notations as in Section 2 for some objects that are slightly different, but play a similar role in this section.

Let $x_{i,j}$ and $y_{i,j}$ for $1 \leq i, j \leq d$ be free commuting variables, and let

$$\Pi_* = \mathbb{Z}_p \langle \langle x_{i,j}, y_{i,j} | 1 \leq i, j \leq d \rangle \rangle$$

be the associative ring (with identity) of formal power series on $x_{i,j}$ and $y_{i,j}$ over the $p$-adic numbers $\mathbb{Z}_p$. Every element in $\Pi_*$ can be written as $f = \sum_{i=0}^{\infty} f_i$ when $f_i$ is homogeneous of degree $i$. Denote

$$\Pi_* \triangleright Q_{*n} = \left\{ f = \sum_{i=0}^{\infty} f_i \in \Pi_* | f_0, ..., f_{n-1} = 0 \right\}.$$ 

The finite index ideals $B_{n,m} = Q_{*n} + p^m \Pi_*$ serve as a basis of neighborhoods of zero to a topology on $\Pi_*$, making $\Pi_*$ a pro-$p$ ring, with a maximal ideal $B_{1,1} = Q_{*1} + p \Pi_*$.  

Endowed with the topology that comes from the congruence ideals $M_d(\Pi_*, B_{n,m}) = \ker(M_d(\Pi_*) \to M_d(\Pi_*/B_{n,m}))$ as a basis of neighborhoods of zero, $M_d(\Pi_*)$ is a profinite ring. One can see that this topology makes the group $1 + M_d(\Pi_*, Q_{*1})$ a pro-$p$ group.

Denote the generic matrices $x_*, y_* \in M_d(\Pi_*, Q_{*1})$ by

$$x_* = (x_{i,j})_{i,j=1}^d, \quad y_* = (y_{i,j})_{i,j=1}^d.$$ 

Let $\hat{F} = \langle \langle X, Y \rangle \rangle$ be the free pro-$p$ group on $X, Y$. By the above, there is a natural (continuous) homomorphism $\pi_* : \hat{F} \to 1 + M_d(\Pi_*, Q_{*n})$ defined by

$$X \mapsto 1 + x_*, \quad Y \mapsto 1 + y_*.$$ 

We denote $G_* = \langle 1 + x_*, 1 + y_* \rangle \subseteq 1 + M_d(\Pi_*, Q_{*1})$ for the (discrete) subgroup generated by $1 + x_*$ and $1 + y_*$, and $\hat{G}_* = \langle \langle 1 + x_*, 1 + y_* \rangle \rangle \subseteq 1 + M_d(\Pi_*, Q_{*1})$ to be its closure in $1 + M_d(\Pi_*, Q_{*1})$. Then, $\hat{G}_*$ is a pro-$p$ group. The map

$$\pi_* : \hat{F} \to \hat{G}_* = \langle \langle 1 + x_*, 1 + y_* \rangle \rangle \subseteq 1 + M_d(\Lambda_*, Q_{*1})$$

is called the universal representation. We have [Zu] (see the proof of Theorem 2.1 in [Zu]):

**Theorem 3.1.** Let $\Delta$ be a pro-$p$ ring. Then, every $1 \neq w(X,Y) \in \ker \pi_*$ is a pro-$p$ identity of $GL_d^1(\Delta)$. 

29
The conclusion of Theorem 3.1 is that in some sense, the gap between \( \hat{G}^* \) from being a free pro-\( p \) group measures the amount of pro-\( p \) identities of \( d \times d \) pro-\( p \) linear groups.

Now, let \( \hat{F} = \hat{F}_1, \hat{F}_2, \hat{F}_3, \ldots \) be the lower central series of \( \hat{F} \), and let \( \Upsilon_n(\hat{F}) = \hat{F}_n/\hat{F}_{n+1} \). Then, the abelian groups \( \Upsilon_n(\hat{F}) \cong \mathbb{Z}_p^{l_2(n)} \) are free \( \mathbb{Z}_p \)-modules of the rank given by the Witt formula (See [L], Proposition 2.7)

\[
l_2(n) = \text{rank}_{\mathbb{Z}_p}(L^{(n)}) = \frac{1}{n} \sum_{m|n} \mu(m) : 2^{\frac{n}{m}}
\]

when \( \mu \) is the Mobius function.

Let \( \hat{H} \) be a pro-\( p \) group, generated by two elements. let \( \hat{H} = \hat{H}_1, \hat{H}_2, \hat{H}_3, \ldots \) be the lower central series of \( \hat{H} \), and denote \( \Upsilon_n(\hat{H}) = \hat{H}_n/\hat{H}_{n+1} \). Notice that from the definition of the lower central series, for every \( n \) we have an epimorphism \( \hat{F}_n \rightarrow \hat{H}_n \) that induces an epimorphism \( \Upsilon_n(\hat{F}) \rightarrow \Upsilon_n(\hat{H}) \).

The following proposition suggests that one way to measure the gap between \( \hat{H} \) from being a free pro-\( p \) group is to evaluate \( \Upsilon_n(\hat{H}) \).

**Proposition 3.2.** The map \( \hat{F} \rightarrow \hat{H} \) is an isomorphism if and only if for every \( n \) the abelian pro-\( p \) group \( \Upsilon_n(\hat{H}) \) is isomorphic to \( \Upsilon_n(\hat{F}) \cong \mathbb{Z}_p^{l_2(n)} \).

**Proof.** If \( \hat{F} \rightarrow \hat{H} \) is an isomorphism then \( \Upsilon_n(\hat{H}) \cong \Upsilon_n(\hat{F}) \cong \mathbb{Z}_p^{l_2(n)} \). On the other hand, if

\[
\Upsilon_n(\hat{F}) \cong \mathbb{Z}_p^{l_2(n)}, \quad \Upsilon_n(\hat{H}) \cong \mathbb{Z}_p^{l_2(n)}
\]

for every \( n \), then as \( \mathbb{Z}_p^{l_2(n)} \) is Hopfian (as a finitely generated profinite group), we get that the surjective map

\[
\Upsilon_n(\hat{F}) \rightarrow \Upsilon_n(\hat{H})
\]

is an isomorphism for every \( n \). Now, assume negatively that \( \hat{F} \rightarrow \hat{H} \) is not an isomorphism, and let \( 1 \neq g \in \ker(\hat{F} \rightarrow \hat{H}) \). As \( \hat{F} \) is pro nilpotent, there exists a unique \( n \) such that \( g \in \hat{F}_n - \hat{F}_{n+1} \). For this \( n \) the map \( \Upsilon_n(\hat{F}) \rightarrow \Upsilon_n(\hat{H}) \) is not injective, what leads to a contradiction. \( \square \)

Fix \( d = 2 \). We are going to prove the following proposition:

**Proposition 3.3.** For every \( p \) and every \( n \leq 5 \) one has \( \Upsilon_n(\hat{G}_*) \cong \mathbb{Z}_p^{l_2(n)} \). Continuing to \( n = 6 \) we have:

- For \( p \neq 2 \) (Zubkov): \( \Upsilon_6(\hat{G}_*) \cong \mathbb{Z}_p^6 \), and in general \( \Upsilon_n(\hat{G}_*) \cong \mathbb{Z}_p^{m(n)} \) where

\[
m(n) = \begin{cases} 
  n(n+2)/8 & \text{if } n \text{ is even} \\
  (n-1)(n+1)/4 & \text{if } n \text{ is odd}.
\end{cases}
\]
For $p = 2$: $\Upsilon_6(\hat{G}_*)$ is an abelian group that is generated by at least $l_2(6) = 9$ generators.

Considering Proposition 3.2 and Theorem 3.1, Proposition 3.3 shows that in some sense $\hat{G}_*$ is closer to be a free pro-$p$ group when $p = 2$, and that there is a real difference between the case $p = 2$ and the cases $p \neq 2$. We notice that we do not have a useful tool to say substantially more informative details regarding this difference. We cannot even say that $\Upsilon_6(\hat{G}_*) \cong \mathbb{Z}_p^9$ when $p = 2$, i.e. that $\Upsilon_6(\hat{G}_*)$ is torsion free. However, on the way of proving Proposition 3.3, we will show how this dichotomy arises in more details than are actually needed in order to prove the proposition (see Proposition 3.13).

### 3.2 The Pseudo Generic Matrices

We strat with presenting appropriate “pseudo generic matrices”, that will help us to present Zubkov’s approach in a bit simpler way than in [Zu]. We notice that we can define these “pseudo generic matrices” in a similar way as we did in Section 2. However, it turns out that over $\mathbb{Z}_p$, in order to move from the original generic matrices to the pseudo generic matrices, one can use a simpler argument, that allows us to use a much simpler definition for the pseudo generic matrices.

Let $x$ and $y$ denote the pseudo generic matrices

\[
x = \begin{pmatrix} x_{11} & x_{12} \\ 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} y_{11} & 0 \\ y_{21} & 0 \end{pmatrix} \in M_2(\Pi_*) .
\]

Notice that $\det(x) = \det(y) = 0$. Being careful to distinguish between $+$ and $-$, the following lemmas are obtained in a similar way as Lemmas 2.7 and 2.8:

**Lemma 3.4.** Let $a, b \in M_2(\Pi_*)$. Then:
1. \( ab + ba = t(a)b + t(b)a + (t(ab) - t(a)t(b)) \cdot 1 . \)
2. \( [a, b, a] = -t(a)[a, b] + 2[a, b]a . \)
3. If $t(a) = 0$, then $a^2 = -\det(a) \cdot 1 \in M_2(\Pi_*)$ is central.

**Lemma 3.5.** For the pseudo-generic matrices $x, y$ we have
1. \( x^2 = t(x)x, \quad y^2 = t(y)y, \quad (xy)^2 = t(xy)xy, \quad (yx)^2 = t(xy)yx . \)
2. \( xyx = t(xy)x, \quad yxy = t(xy)y . \)
3. \( [x, y]^2 = (t(xy)^2 - t(x)t(y)t(xy)) \cdot 1 . \)

We define the subrings (with identity) of $\Pi_*$

\[
T = \mathbb{Z}_p \langle t(x), t(y), t(xy) \rangle \subseteq \Pi_*
\]

\[
S = \mathbb{Z}_p \langle t(x)^2, t(y)^2, [x, y]^2 \rangle \subseteq \Pi_*
\]

The following proposition is proved by similar (but easier) arguments as Proposition 2.11.
Proposition 3.6. The ring $T$ is freely generated by $t(x), t(y), t(xy)$ as a commutative ring over $\mathbb{Z}_p$.

As a corollary of Part 3 in Lemma 3.5 and the above proposition one can easily prove that:

Corollary 3.7. The ring $S \subseteq T$, and it is free on $t(x)^2, t(y)^2, [x, y]^2$.

3.3 The structure of a minimal component

Recall the generic matrices

$$x_\ast = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}, \quad y_\ast = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \in M_2(\Pi_\ast).$$

Let’s have also the following notations:

- $\hat{\Pi}_\ast = \mathbb{Q}_p \langle \langle x_{i,j}, y_{i,j} | 1 \leq i, j \leq n \rangle \rangle$ is the ring of power series on $x_{i,j}$ and $y_{i,j}$ for $1 \leq i, j \leq n$ over $\mathbb{Q}_p$.
- $\hat{Q}_{s,n} = \{ f = \sum_{i=0}^{\infty} f_i \in \hat{\Pi}_\ast | f_0, \ldots, f_{n-1} = 0 \}$.
- $\hat{A}_\ast = \mathbb{Q}_p \langle \langle x_\ast, y_\ast \rangle \rangle \subseteq M_2(\hat{\Pi}_\ast)$ is the ring of power series on $x_\ast, y_\ast$ over $\mathbb{Q}_p$.
- $A_\ast = \mathbb{Z}_p \langle \langle x_\ast, y_\ast \rangle \rangle \subseteq M_2(\Pi_\ast)$ is the ring of power series on $x_\ast, y_\ast$ over $\mathbb{Z}_p$.
- $\hat{L}_\ast$ is the Lie ring on $x_\ast, y_\ast$ over $\mathbb{Q}_p$.
- $\hat{L}_\ast^{(n)}$ is the subspace of $\hat{L}_\ast$ of homogeneous elements of degree $n$.
- $L_\ast \subseteq \hat{L}_\ast$ is the Lie ring on $x_\ast, y_\ast$ over $\mathbb{Z}_p$.
- $L_\ast^{(n)} \subseteq \hat{L}_\ast^{(n)}$ is the additive subgroup of $L_\ast$ of homogeneous elements of degree $n$.

For an element of the form $g = 1 + a_n + a_{n+1} + \ldots \in 1 + M_2(\hat{\Pi}_\ast, \hat{Q}_{s,1})$ where $a_i$ is the term of $g$ of degree $i$, and $a_n \neq 0$, we denote

$$\min(g) = a_n.$$

The following proposition was proved in [Zu]:

Proposition 3.8. Let $\hat{G}_\ast = \langle \langle 1 + x_\ast, 1 + y_\ast \rangle \rangle \subseteq 1 + M_2(\Pi_\ast, Q_{s,1})$, and let $1 \neq g \in \hat{G}_\ast$. Then, for some $n$ we have

$$\min(g) \in \hat{L}_\ast^{(n)} \cap A_\ast.$$
Here is an outline of the proof. It is obvious that $\min(g) \in A_\ast$. To show the additional inclusion, write

$$e^{x_\ast} = 1 + x_\ast + \left( \begin{array}{cc} x'_{11} & x'_{12} \\ x'_{21} & x'_{22} \end{array} \right)$$

$$e^{y_\ast} = 1 + y_\ast + \left( \begin{array}{cc} y'_{11} & y'_{12} \\ y'_{21} & y'_{22} \end{array} \right)$$

where $x'_{ij}$ and $y'_{ij}$ are built up from terms of degree $\geq 1$. Then, the ring homomorphism $\phi_\ast : \tilde{\Pi}_\ast \to \tilde{\Pi}_\ast$ defined by sending $x_{ij} \to x_{ij} + x'_{ij}$ and $y_{ij} \to y_{ij} + y'_{ij}$ gives rise to a ring homomorphism $M_2(\tilde{\Pi}_\ast) \to M_2(\tilde{\Pi}_\ast)$ that gives rise to a group homomorphism $\Psi_\ast : G_\ast \to 1 + M_2(\tilde{\Pi}_\ast, Q_{s1})$ defined by

$$1 + x_\ast \mapsto e^{x_\ast}$$

$$1 + y_\ast \mapsto e^{y_\ast}.$$ 

By the Baker–Campbell–Hausdorff formula, one has

$$\min(\Psi_\ast(g)) \in \tilde{L}_\ast^{(n)}$$

for some $n$. But as $\Psi_\ast$ is originally induced by the ring homomorphism $\phi_\ast$, one obtains that

$$\min(g) = \min(\Psi_\ast(g)) \in \tilde{L}_\ast^{(n)}$$

as claimed in Proposition 6.8

In his paper, Zubkov gave the following formulas regarding the generic matrices:

$$[x_\ast, y_\ast, x_\ast, x_\ast] = \alpha_\ast [x_\ast, y_\ast] \quad \text{for} \quad \alpha_\ast = t(x_\ast)^2 - 4 \cdot \det(x_\ast)$$

$$[x_\ast, y_\ast, y_\ast, x_\ast] = \beta_\ast [x_\ast, y_\ast] \quad \text{for} \quad \beta_\ast = 2t(x_\ast y_\ast) - t(x_\ast) t(y_\ast)$$

$$[x_\ast, y_\ast, y_\ast, y_\ast] = \gamma_\ast [x_\ast, y_\ast] \quad \text{for} \quad \gamma_\ast = t(y_\ast)^2 - 4 \cdot \det(y_\ast)$$

It follows that

$$\tilde{L}_\ast^{(n)} = \begin{cases} \sum_{r+s+t=(n-2)/2} \left( Q_p \alpha_\ast^r \beta_\ast^s \gamma_\ast^t [x_\ast, y_\ast] \right) & n \text{ even} \\ \sum_{r+s+t=(n-3)/2} \left( Q_p \alpha_\ast^r \beta_\ast^s \gamma_\ast^t [x_\ast, y_\ast, x_\ast] + Q_p \alpha_\ast^r \beta_\ast^s \gamma_\ast^t [x_\ast, y_\ast, y_\ast] \right) & n \text{ odd} \end{cases}$$

and we have a similar description of $L_\ast^{(n)}$. We remark that in [Zu], Zubkov shows that the sums in (3.1) are actually direct.

Recall the pseudo generic matrices

$$x = \left( \begin{array}{cc} x_{11} & x_{12} \\ 0 & 0 \end{array} \right), \quad y = \left( \begin{array}{cc} y_{11} & 0 \\ y_{21} & 0 \end{array} \right) \in M_2(\Pi)$$

where $\Pi = \mathbb{Z}_p \langle (x_{12}, x_{11}, y_{11}, y_{21}) \rangle \subseteq \Pi_\ast$. Notice that $\Pi$ can also be viewed as the image of $\Pi_\ast$ under the projection $\varphi : \Pi_\ast \to \Pi$ defined by

$$x_{21}, x_{22}, y_{12}, y_{22} \overset{\varphi}{\to} 0.$$ 

We use the following notations regarding the pseudo generic matrices:
• \( \Pi = \varphi(\hat{\Lambda}) = \mathbb{Q}_p \langle \langle x_{12}, x_{11}, y_{11}, y_{21} \rangle \rangle \).
• \( \bar{\Pi} = \varphi(\hat{\Lambda}) = \mathbb{Z}_p \langle \langle x_{12}, x_{11}, y_{11}, y_{21} \rangle \rangle \subseteq \bar{\Pi} \).
• \( Q_n = \varphi(Q_n) \) and \( \bar{Q}_n = \varphi(\bar{Q}_n) \).
• \( \hat{A} = \mathbb{Q}_p \langle \langle x, y \rangle \rangle \subseteq M_2(\bar{\Pi}) \).
• \( A = \mathbb{Z}_p \langle \langle x, y \rangle \rangle \subseteq M_2(\Pi) \).
• \( \hat{L} = \) the Lie ring on \( x, y \) over \( \mathbb{Q}_p \).
• \( L^{(n)} = \) the subspace of \( L \) of homogeneous elements of degree \( n \).
• \( L \subseteq \hat{L} = \) the Lie ring on \( x, y \) over \( \mathbb{Z}_p \).
• \( L^{(n)} \subseteq \hat{L}^{(n)} = \) the additive subgroup of \( L \) of homogeneous elements of degree \( n \).

From Zubkov’s formulas one can see that regarding the pseudo generic matrices, one has

\[
[x, y, x, x] = \alpha [x, y] \quad \text{for} \quad \alpha = t(x)^2
\]
\[
[x, y, x, y] = [x, y, x, y] = \beta [x, y] \quad \text{for} \quad \beta = 2t(xy) - t(x)t(y)
\]
\[
[x, y, y, y] = \gamma [x, y] \quad \text{for} \quad \gamma = t(y)^2
\]

and hence

\[
\hat{L}^{(n)} = \begin{cases} 
\sum_{r+s+t=(n-2)/2} (Q_p \alpha^r \beta^s \gamma^t [x, y]) & n = \text{even} \\
\sum_{r+s+t=(n-3)/2} (Q_p \alpha^r \beta^s \gamma^t [x, y, x] + Q_p \alpha^r \beta^s \gamma^t [x, y, y]) & n = \text{odd}
\end{cases}
\]

(3.2)

and we have a similar description of \( L^{(n)} \). Let \( \mathbb{Q}_p \cdot T = \mathbb{Q}_p \langle t(x), t(y), t(xy) \rangle \subseteq \hat{\Pi} \). Like in Propositions 3.6 and 2.11 \( \mathbb{Q}_p \cdot T \) is freely generated by \( \langle t(x), t(y), t(xy) \rangle \) over \( \mathbb{Q}_p \), and hence a unique factorization domain. The following proposition is an easy corollary of this fact:

**Proposition 3.9.** The ring \( \mathbb{Q}_p \langle \alpha, \beta, \gamma \rangle \subseteq \mathbb{Q}_p \cdot T \) is freely generated by \( \alpha, \beta, \gamma \) over \( \mathbb{Q}_p \).

**Corollary 3.10.** The sums in (3.2) are direct. Moreover, the natural map \( \hat{L}^{(n)} \to L^{(n)} \) is an isomorphism of vector spaces for every \( n \).

**Proof.** We start with showing that the sums in (3.2) are direct when \( n \) is even. By Proposition 3.9 it is enough to show that given an element in \( \hat{L}^{(n)} \) of the form \( a = \epsilon \cdot [x, y] \) for \( \epsilon \in \mathbb{Q}_p \langle \alpha, \beta, \gamma \rangle \), \( \epsilon \) is determined uniquely. Indeed, consider

\[
\mathbb{Q}_p \cdot T \ni t(axy) = t(\epsilon \cdot [x, y] xy) = \epsilon \cdot [x, y]^2.
\]

Hence, as \( \mathbb{Q}_p \cdot T \) is a domain, given such \( a \) we can restore \( \epsilon \), as required.
For the odd case, assume that 
\[ a = \varepsilon_1 \cdot [x, y, x] + \varepsilon_2 \cdot [x, y, y] \text{ for } \varepsilon_1, \varepsilon_2 \in \mathbb{Q}_p \langle \alpha, \beta, \gamma \rangle. \]
Consider
\[ \mathbb{Q}_p \cdot T \ni t(ax) = t(\varepsilon_1 \cdot [x, y, x]x + \varepsilon_2 \cdot [x, y, y]x) = -2\varepsilon_2 \cdot [x, y]^2 \]
\[ \mathbb{Q}_p \cdot T \ni t(ay) = t(\varepsilon_1 \cdot [x, y, x]y + \varepsilon_2 \cdot [x, y, y]y) = 2\varepsilon_1 \cdot [x, y]^2. \]
So similarly, we can restore \(\varepsilon_1, \varepsilon_2\). Thus, the sums in (3.2) are direct. Considering Proposition 3.9 it is clear now that the map \(\tilde{L}^{(n)} \to \hat{L}^{(n)}\) is an isomorphism.

**Corollary 3.11.** ([Zu]) We have
\[
\text{rank}_{\mathbb{Z}_p}(L^{(n)}) = \dim_{\mathbb{Q}_p}(\tilde{L}^{(n)}) = \begin{cases} 
\frac{n(n+2)}{8} & n \text{ is even} \\
\frac{(n-1)(n+1)}{4} & n \text{ is odd.} 
\end{cases}
\]

Considering Proposition 3.8 we obtain:

**Corollary 3.12.** The ring homomorphism \(\varphi\) induces a natural isomorphism
\[ \hat{G} = \langle (1 + x_*, 1 + y_*) \rangle \cong \hat{G} = \langle (1 + x, 1 + y) \rangle \]
and for every \(g \in \hat{G}\) one has \(\min(g) \in \tilde{L}^{(n)} \cap A\) for some \(n\).

Following the outline of Proposition 3.8 proof, we can write
\[ e^x = 1 + x + \begin{pmatrix} x'_{11} & x'_{12} \\ 0 & 0 \end{pmatrix} \]
\[ e^y = 1 + y + \begin{pmatrix} y'_{11} \\ y'_{21} \end{pmatrix} \]
where \(x'_{ij}\) and \(y'_{ij}\) are built up from terms of degree > 1. Hence, the ring homomorphism \(\phi : \hat{\Pi} \to \hat{\Pi}\) defined by sending \(x_{ij} \to x_{ij} + x'_{ij}\) and \(y_{ij} \to y_{ij} + y'_{ij}\) gives rise to a ring homomorphism \(M_2(\hat{\Pi}) \to M_2(\hat{\Pi})\) that gives rise to a group homomorphism \(\Psi : \hat{G} \to 1 + M_2(\hat{\Pi}, \mathbb{Q}_1)\) defined by \(1 + x \mapsto e^x\) and \(1 + y \mapsto e^y\). Therefore, also here we have
\[ \min(g) = \min(\Psi(g)) \in \hat{L}^{(n)} \tag{3.3} \]
for some \(n\). We will use this property later.

It is obvious that in general \(L^{(n)} \subseteq \tilde{L}^{(n)} \cap A\). In [Zu], Zubkov shows that when \(p \neq 2\), we actually have \(L^{(n)} = \tilde{L}^{(n)} \cap A\). However, this is not the case when \(p = 2\). Here is the full description of \(\tilde{L}^{(n)} \cap A\) when \(p = 2\):

**Proposition 3.13.** Denote
\[ \delta = \frac{\beta^2 - \alpha \gamma}{4} = t(xy)^2 - t(xy)t(x)t(y) = [x, y]^2 \]
and \(S = \mathbb{Z}_2 \langle \alpha, \gamma, \delta \rangle\). Then:
• If $n$ is even, then $\tilde{L}^{(n)} \cap A = S \cdot [x, y] + \beta \cdot S \cdot [x, y]$.

• If $n$ is odd, then $\tilde{L}^{(n)} \cap A$ is equal to

$$S \cdot [x, y, x] + S \cdot [x, y, y] + S \cdot \beta [x, y, x] + \alpha [x, y, y] + S \cdot \gamma [x, y, x] + \beta [x, y, y] \over 2.$$  

(3.4)

Proof. One direction of the inclusion, namely $\supseteq$ is easy to verify, using Part 2 of Lemma 3.4 for (3.4). We turn to the opposite inclusion, starting with the even case. Let $p(\alpha, \beta, \gamma) \cdot [x, y] \in \tilde{L}^{(n)} \cap A$ when $p(\alpha, \beta, \gamma)$ is a homogeneous polynomial on $\alpha, \beta, \gamma$ over $\mathbb{Q}_2$, say of degree $m$. In particular $p(\alpha, \beta, \gamma) \in \mathbb{Q}_2 \cdot T$. Consider

$$t(p(\alpha, \beta, \gamma) \cdot [x, y]xy) = p(\alpha, \beta, \gamma) \cdot t([x, y]xy) = p(\alpha, \beta, \gamma) \cdot [x, y]^2.$$  

On the other hand, $p(\alpha, \beta, \gamma) \cdot [x, y]xy \in A$, so

$$p(\alpha, \beta, \gamma) \cdot [x, y]^2 = p(\alpha, \beta, \gamma) \cdot (t(xy)^2 - t(xy)t(x)y)) \in T.$$  

One can see that it follows that $p(\alpha, \beta, \gamma) \in T$. So write

$$p(\alpha, \beta, \gamma) = \sum_{2i+j+k=2m} \varepsilon_{i,j,k} t(xy)^it(xy)^j t(y)^k$$

for some $\varepsilon_{i,j,k} \in \mathbb{Z}_2$. We want to show that $p(\alpha, \beta, \gamma) \in S + \beta \cdot S$. Order the triples $(i, j, k)$ in the lexicographical order and let $(i_0, j_0, k_0)$ be the maximal for which $i_0$ is even, and $\varepsilon_{i_0,j_0,k_0} \neq 0$. We have two options:

• If $j_0, k_0$ are even, reduce $\varepsilon_{i_0,j_0,k_0} \cdot \delta^{i_0} \cdot \alpha^{j_0 \over 2} \cdot \gamma^{k_0 \over 2} \in S$ from $p(\alpha, \beta, \gamma)$. One can see that by doing this, the maximal triple $(i_0, j_0, k_0)$ for which $i_0$ is even, and $\varepsilon_{i_0,j_0,k_0} \neq 0$ is reduced.

• If $j_0, k_0$ are odd, noticing that $(i_0 + 1)^{-1} \in \mathbb{Z}_2$, consider

$$(i_0 + 1)^{-1} \cdot \varepsilon_{i_0,j_0,k_0} \cdot \beta \cdot \delta^{i_0} \cdot \alpha^{i_0 - 1 \over 2} \cdot \gamma^{k_0 - 1 \over 2}$$

$$= (i_0 + 1)^{-1} \cdot \varepsilon_{i_0,j_0,k_0} \cdot \alpha^{i_0 - 1 \over 2} \cdot \gamma^{k_0 - 1 \over 2} \cdot 2t(xy)^i_0 + 2 \cdot \frac{t}{2} \cdot t(x)y(t(xy)^i_0 - t(x)t(y)y^i_0)$$

$$+ \text{terms with lower powers of } t(xy)$$

$$= (i_0 + 1)^{-1} \cdot \varepsilon_{i_0,j_0,k_0} t(xy)^i_0 \cdot (t(x) \alpha^{i_0 - 1 \over 2} \cdot t(y) \gamma^{k_0 - 1 \over 2})$$

$$+ \text{terms with lower powers of } t(xy).$$

Hence, by adding $(i_0 + 1)^{-1} \cdot \varepsilon_{i_0,j_0,k_0} \cdot \beta \cdot \delta^{i_0} \cdot \alpha^{i_0 - 1 \over 2} \cdot \gamma^{k_0 - 1 \over 2} \in \beta \cdot S$ to $p(\alpha, \beta, \gamma)$, the maximal triple $(i_0, j_0, k_0)$ for which $i_0$ is even, and $\varepsilon_{i_0,j_0,k_0} \neq 0$ is reduced.
We continue in this way until all the terms for which \( i_0 \) is even are reduced. We claim that in this case \( p(\alpha, \beta, \gamma) = 0 \). Indeed, write

\[
p(\alpha, \beta, \gamma) = \sum_{i=0}^{r} \beta^i \cdot q_i(\alpha, \gamma)
\]

Clearly, the highest \( i_0 \) for which \( q_{i_0}(\alpha, \gamma) \neq 0 \) is the highest degree of \( t(xy) \). So by assumption, \( i_0 \) is odd. Hence, the term for which the degree of \( t(xy) \) is \( i_0 - 1 \) is given by

\[
2^{i_0-1} \cdot t(xy)^{i_0-1}(-i_0 \cdot t(x)t(y)q_{i_0}(\alpha, \gamma) + q_{i_0-1}(\alpha, \gamma)).
\]

As \( q_{i_0}(\alpha, \gamma) \neq 0 \), the degrees of the terms of \( t(x)t(y)q_{i_0}(\alpha, \gamma) \) in \( t(x), t(y) \) are odd, and the degrees of the terms of \( q_{i_0-1}(\alpha, \gamma) \) in \( t(x), t(y) \) are even, one has

\[
-i_0 \cdot t(x)t(y)q_{i_0}(\alpha, \gamma) + q_{i_0-1}(\alpha, \gamma) \neq 0
\]

and this is a contradiction to the assumption that \( p(\alpha, \beta, \gamma) \) does not have terms with even degree in \( t(xy) \).

For the odd case, write

\[
v = p(\alpha, \beta, \gamma) \cdot [x, y, x] + q(\alpha, \beta, \gamma) \cdot [x, y, y] \in \tilde{L}^{(n)} \cap A
\]

when \( p(\alpha, \beta, \gamma), q(\alpha, \beta, \gamma) \) are homogeneous on \( \alpha, \beta, \gamma \) over \( \mathbb{Q}_2 \), say of degree \( m \). Now, using the identities \( [x, y, x] = -t(x)[x, y] + 2[x, y]x \) and \( [x, y, y] = -t(y)[x, y] + 2[x, y]y \) we have

\[
v = (-p(\alpha, \beta, \gamma)t(x) - q(\alpha, \beta, \gamma)t(y))[x, y] + 2p(\alpha, \beta, \gamma)[x, y]x + 2q(\alpha, \beta, \gamma)[x, y]y.
\]

As \( v \in A \) we have

\[
-t(xy)[x, y]^2(p(\alpha, \beta, \gamma)t(x) + q(\alpha, \beta, \gamma)t(y)) = t(x \cdot v \cdot [x, y]y) \in T.
\]

Also here, one can see that it follows that \( p(\alpha, \beta, \gamma)t(x) + q(\alpha, \beta, \gamma)t(y) \in T \). So write

\[
p(\alpha, \beta, \gamma)t(x) + q(\alpha, \beta, \gamma)t(y) = \sum_{2i+j+k=2m+1} \varepsilon_{i,j,k} t(xy)^i t(x)^j t(y)^k
\]

for some \( \varepsilon_{i,j,k} \in \mathbb{Z}_2 \). Order the triples \((i, j, k)\) in the lexicographical order and let \((i_0, j_0, k_0)\) be the maximal for which \( \varepsilon_{i_0,j_0,k_0} \neq 0 \). We have a few options:

- If \( i_0 \) is even, \( j_0 \) is odd, \( k_0 \) is even, subtract the following from \( v \):

  \[
  \varepsilon_{i_0,j_0,k_0} \cdot \delta^{i_0} \cdot \alpha^{\frac{j_0}{2}} \cdot \gamma^{\frac{k_0-1}{2}} [x, y, x] \in S \cdot [x, y, x].
  \]

- If \( i_0 \) is even, \( j_0 \) is even, \( k_0 \) is odd, subtract the following from \( v \):

  \[
  \varepsilon_{i_0,j_0,k_0} \cdot \delta^{i_0} \cdot \alpha^{\frac{j_0}{2}} \cdot \gamma^{\frac{k_0-1}{2}} [x, y, y] \in S \cdot [x, y, y].
  \]
• If \( i_0 \) is odd, \( j_0 \) is odd, \( k_0 \) is even, subtract the following from \( v \):

\[
\varepsilon_{i_0, j_0, k_0} \delta_{\frac{j_0 - 1}{2}} \cdot \alpha \frac{m}{2} \cdot \gamma^{k_0} \left( \frac{\beta[x, y, x] + \alpha[x, y, y]}{2} \right) \in S \frac{\beta[x, y, x] + \alpha[x, y, y]}{2}.
\]

• If \( i_0 \) is odd, \( j_0 \) is even, \( k_0 \) is odd, subtract the following from \( v \):

\[
\varepsilon_{i_0, j_0, k_0} \delta_{\frac{j_0 - 1}{2}} \cdot \alpha \frac{m}{2} \cdot \gamma^{k_0} \left( \frac{\gamma[x, y, x] + \beta[x, y, y]}{2} \right) \in S \frac{\gamma[x, y, x] + \beta[x, y, y]}{2}.
\]

One can see that in the new element, the maximal \((i_0, j_0, k_0)\) for which \( \varepsilon_{i_0, j_0, k_0} \neq 0 \) is lower. We continue the process until the expression

\[
p(\alpha, \beta, \gamma)t(x) + q(\alpha, \beta, \gamma)t(y)
\]

is vanished. We claim that in this case, \( p(\alpha, \beta, \gamma) = q(\alpha, \beta, \gamma) = 0 \), i.e. \( v = 0 \).

Indeed, if we write \( p(\alpha, \beta, \gamma) \) as an expression in \( t(x, y), t(x), t(y) \) (with coefficients in \( \mathbb{Q}_2 \)) and order the monomials with the above lexicographical order, the highest monomial of \( p(\alpha, \beta, \gamma)t(x) \) will have the form \( \varepsilon \cdot t(xy)^j t(x)^i t(y)^k \) where \( j \) is odd and \( k \) is even, and the highest monomial of \( q(\alpha, \beta, \gamma)t(y) \) will have the form \( \varepsilon \cdot t(xy)^j t(x)^i t(y)^k \) where \( j \) is even and \( k \) is odd. So they cannot cancel each other, and the only way for \( (3.5) \) to be 0 is that \( p(\alpha, \beta, \gamma) = q(\alpha, \beta, \gamma) = 0 \), as required.

### 3.4 Proving Proposition 3.3

We want now to come up with some conclusions regarding the universal representation when \( p \neq 2 \) and when \( p = 2 \). Recall the ideals of the form

\[
Q_n = \{ f = \sum_{r=0}^{\infty} f_r \in \Pi \mid f_0, \ldots, f_{n-1} = 0 \}
\]

and denote

\[
\omega_n(\tilde{G}) = \ker(\tilde{G} \to GL_2(\Pi/Q_n))
\]

\[
\Omega_n(\tilde{G}) = \omega_n(\tilde{G})/\omega_{n+1}(\tilde{G}).
\]

It is easy to verify that in general, \( \tilde{G}_n \subseteq \omega_n(\tilde{G}) \). Hence, for every \( n \) we have a natural map

\[
\Upsilon_n(\tilde{G}) \to \Omega_n(\tilde{G}).
\]

Now, notice that in general, by Corollary 3.12 one can view \( \Omega_n(\tilde{G}) \) as an abelian subgroup of \( L^{(n)} \cap A \) such that

\[
\Upsilon_n(\tilde{G}) \to L^{(n)} \leq \Omega_n(\tilde{G}) \leq L^{(n)} \cap A.
\]

Let’s start with proving the first part of Proposition 3.13, namely, that for every \( p \) and every \( n \leq 5 \) one has \( \Upsilon_n(\tilde{G}) \cong \mathbb{Z}_{p}^{l_2(n)} \). Actually:

**Proposition 3.14.** For every \( p \) we have:

- \( \tilde{G}_n = \omega_n(\tilde{G}) \) for every \( n \leq 6 \).
• $\Upsilon_n(\hat{F}) \cong \Upsilon_n(\hat{G}) \cong \Omega_n(\hat{G}) \cong L^{(n)} \cong \mathbb{Z}^{l_2(n)}$ for every $n \leq 5$.

Proof. By definition, we have $\hat{G} = \hat{G}_1 = \omega_1(\hat{G})$. We want to show, by induction on $n$, that the same is valid for every $n \leq 6$. Let $n \leq 6$, and assume that $\hat{G}_n = \omega_n(\hat{G})$. Then, under this assumption, $\Upsilon_n(\hat{G}) \to \Omega_n(\hat{G})$ and hence, Equation (3.6) gives

$$\Upsilon_n(\hat{F}) \to \Upsilon_n(\hat{G}) \to L^{(n)} = \Omega_n(\hat{G}).$$

Hence, as for every $n \leq 5$ the formula in Corollary 3.11 coincides with the Witt formula, we have

$$\Upsilon_n(\hat{F}) \cong \mathbb{Z}^{l_2(n)}, \quad \Omega_n(\hat{G}) \cong L^{(n)} \cong \mathbb{Z}^{l_2(n)}.$$  

As $\mathbb{Z}^{l_2(n)}$ is Hopfian (as a finitely generated profinite group), we get that the composition map $\Upsilon_n(\hat{F}) \to \Omega_n(\hat{G})$ is an isomorphism. It follows that

$$\Upsilon_n(\hat{F}) \to \Upsilon_n(\hat{G}) = \hat{G}_n/\hat{G}_{n+1} \to \Omega_n(\hat{G}) = \omega_n(\hat{G})/\omega_{n+1}(\hat{G})$$

is an isomorphism. In particular $\hat{G}_{n+1} = \omega_{n+1}(\hat{G})$ as required. Notice that in the course of the proof we also proved the second statement. \hfill $\square$

Now, notice that as when $p \neq 2$ we have $L^{(n)} = L^{(n)} \cap A$, Equation (3.6) actually gives

$$\Upsilon_n(\hat{G}) \to \Omega_n(\hat{G}) \cong L^{(n)} \cong \mathbb{Z}^{m(n)}$$

where $m(n) = \begin{cases} n(n+2)/8 & \text{n is even} \\ (n-1)(n+1)/4 & \text{n is odd} \end{cases}$.

Thus, in this case, one can easily deduce the following proposition which yields the second part of Proposition 3.13 (see [Zu], Theorem 4.1):

**Proposition 3.15.** When $p \neq 2$ we have $\hat{G}_n = \omega_n(\hat{G})$ and hence $\Upsilon_n(\hat{G}) = \Omega_n(\hat{G}) \cong L^{(n)}$ for every $n$.

The meaning of Proposition 3.15 is that when $p \neq 2$, given an element $g \in \hat{G}$, there is a correspondence between the degree of $\min(g)$ and the location of $g$ in lower central series of $\hat{G}$. When $p = 2$, this correspondence is broken, as we demonstrate below.

Assume now that $p = 2$. Recall that $\hat{G}_{7} \subseteq \omega_7(\hat{G})$ and by Proposition 3.14 we have $\hat{G}_6 = \omega_6(\hat{G})$. It follows that we have the exact sequence

$$1 \to (\omega_7(\hat{G}) \cap \hat{G}_6)/\hat{G}_7 \to \Upsilon_6(\hat{G}) = \hat{G}_6/\hat{G}_7 \to \Omega_6(\hat{G}) = \omega_6(\hat{G})/\omega_7(\hat{G}) \to 1.$$  

Notice that by Equation (3.6), the surjective map $\Upsilon_6(\hat{G}) \to \Omega_6(\hat{G})$ yields that $\Omega_6(\hat{G}) \cong L^{(6)} \cong \mathbb{Z}^2_2$.

Given a pro-$p$ group $\hat{H}$, we denote the minimum number of (topological) generators for $\hat{H}$ by $d(\hat{H})$. We have the following lemma:
Lemma 3.16. Let $1 \to \hat{H}_1 \to \hat{H}_2 \to \hat{H}_3 \to 1$ be an exact sequence of finitely generated abelian pro-$p$ groups, such that $\hat{H}_3$ is free (as an abelian pro-$p$ group). Then: $d(H_2) = d(H_1) + d(H_3)$.

Proof. As $\hat{H}_3$ is free, we have a section map $\hat{H}_3 \to \hat{H}_2$ such that the composition map $\hat{H}_3 \to \hat{H}_2 \to \hat{H}_3$ is a natural isomorphism. It follows that: $\hat{H}_2 \cong \hat{H}_1 \times \hat{H}_3$. Hence $d(H_2) \leq d(H_1) + d(H_3)$. On the other hand, $\hat{H}_2$ has the vector space

$$(\hat{H}_1/\Phi(\hat{H}_1)) \times (\hat{H}_3/\Phi(\hat{H}_3)) \cong (\mathbb{Z}/p\mathbb{Z})^{d(\hat{H}_1)} \times (\mathbb{Z}/p\mathbb{Z})^{d(\hat{H}_3)} \cong (\mathbb{Z}/p\mathbb{Z})^{d(\hat{H}_1)+d(\hat{H}_3)}$$

as a homomorphic image, where $\Phi(\hat{H}_1), \Phi(\hat{H}_3)$ are the Frattini subgroups of $\hat{H}_1, \hat{H}_3$. Hence, $d(H_2) \geq d(H_1) + d(H_3)$, as required.

Recall that $\Omega_6(\hat{G}) \cong \mathbb{Z}_2^6$. By the lemma, it follows that in order to prove the last part of Proposition 3.3 it is enough to show that for

$$\hat{H} = (\omega_7(\hat{G}) \cap \hat{G}_6)/\hat{G}_7$$

we have $d(\hat{H}) \geq 3$. By Corollary 3.12 every element of $\omega_7(\hat{G})$ can be written as $g = 1 + \min(g) + \text{terms of degree } \geq 8$, when $\min(g) \in \bar{L}(7)$. Therefore, by mapping each $g \in \omega_7(\hat{G})$ to its corresponding minimal term\(^\dagger\) we have a natural map $\varrho : \omega_7(\hat{G}) \to \bar{L}(7)$. Clearly, the image of $G_7$ under $\varrho$ is $L(7)$. Hence

$$\hat{H} = \varrho(\omega_7(\hat{G}) \cap \hat{G}_6)/\varrho(\hat{G}_7) = \varrho(\omega_7(\hat{G}) \cap \hat{G}_6)/L(7)$$

is a homomorphic image of $\hat{H}$. We are going to show that the abelian group $\hat{H}$ contains a copy of $(\mathbb{Z}/2\mathbb{Z})^3$. It will follow that $d(\hat{H}) \geq 3$, and thus $d(\hat{H}) \geq 3$ as well, as required. We prove the following technical proposition:

Lemma 3.17. The following elements belong to $\omega_7(\hat{G}) \cap \hat{G}_6$:

$$\begin{align*}
g_1 &= [1 + x, 1 + y, 1 + x, 1 + x, [1 + x, 1 + y]_G]_G \\
g_2 &= [1 + x, 1 + y, 1 + x, 1 + y, [1 + x, 1 + y]_G]_G \\
g_3 &= [1 + x, 1 + y, 1 + y, 1 + y, [1 + x, 1 + y]_G]_G
\end{align*}$$

In addition, the image of $g_i$ under $\varrho$ is

$$\begin{align*}
\varrho(g_1) &= \min(g_1) = \frac{\beta^2 - \alpha \gamma}{2} [x, y, x] = 2\delta \cdot [x, y, x] \mod L(7) \\
\varrho(g_2) &= \min(g_2) = \beta \cdot \left( \frac{\beta [x, y, x] + \alpha [x, y]}{2} \right) \\
&\quad + \beta \cdot \left( \frac{\beta [x, y, y] + \gamma [x, y, x]}{2} \right) \mod L(7) \\
\varrho(g_3) &= \min(g_3) = \frac{\beta^2 - \alpha \gamma}{2} [x, y, y] = 2\delta \cdot [x, y, y] \mod L(7)
\end{align*}$$

\(^\dagger\)Elements $g \in \omega_7(\hat{G})$ with $\deg(\min(g)) > 7$ are mapped to $0 \in \bar{L}(7)$.
Before we prove this lemma, we want to show that it yields that \( \bar{H} \) contains a copy of \((\mathbb{Z}/2\mathbb{Z})^3\). Indeed, we saw that the elements
\[
\alpha^r \beta^s \gamma^t \{x, y, x\}, \ r + s + t = 2
\]
\[
\alpha^r \beta^s \gamma^t \{x, y, y\}, \ r + s + t = 2
\]
generate the \( \mathbb{Z}_2 \)-module \( L(7) \). In addition, as these elements give a basis to the vector space \( \tilde{L}^{(7)} \), they generate \( L^{(7)} \) freely as a \( \mathbb{Z}_2 \)-module. It follows that as \( 1 \in \mathbb{Z}_2 \), the elements \( \bar{\rho}(g_i) \) are not in \( L(7) \). Moreover, obviously, the order of \( \bar{\rho}(g_i) \mod L^{(7)} \) is 2. Eventually, one can easily see that \( \bar{\rho}(g_i) \) are different mod \( L^{(7)} \), and no one can be expressed by the others. It follows that the subgroup of \( \bar{H} \) generated by \( \bar{\rho}(g_i) \mod L^{(7)} \) is isomorphic to \((\mathbb{Z}/2\mathbb{Z})^3 \leq \bar{H} \), as required. So it remains to prove Lemma 3.17.

**Proof.** (of Lemma 3.17) By the formula given in Equation (3.3), in order to evaluate \( \min(g_1) \) it is enough to evaluate
\[
\min([e^x, e^y, e^x, e^y, [e^x, e^y]_{\hat{G}}, [e^x, e^y]_{\hat{G}}])
\]
So in general, if \( z, w \in M_2(\Pi/\mathbb{Q}_1) \) then a direct computation throuth the Baker–Campbell–Hausdorff formula gives
\[
\ln([e^x, e^y]_{\hat{G}}) = [z, w] - \frac{1}{2} [z, w, z] - \frac{1}{2} [z, w, w] + \text{higher commutators on } z, w.
\]
Using this formula three times one has
\[
\ln([e^x, e^y, e^x, e^x]_{\hat{G}}) = \alpha [x, y] - \frac{3\alpha}{2} [x, y, x] - \frac{\beta}{2} [x, y, x] + \text{terms of degree } \geq 6.
\]
Therefore, using the identities
\[
[x, y, [x, y, x]] = [x, y, x] - [x, y, x, x, y] = \beta [x, y, x] - \alpha [x, y, y]
\]
\[
[x, y, [x, y, y]] = [x, y, y, x] - [x, y, y, x, y] = \gamma [x, y, x] - \beta [x, y, y]
\]
one has
\[
\ln([e^x, e^y, e^x, e^x]_{\hat{G}}) = [\alpha [x, y] - \frac{3\alpha}{2} [x, y, x] - \frac{\beta}{2} [x, y, x], [x, y] - \frac{1}{2} [x, y, x] - \frac{1}{2} [x, y, y]] + \text{terms of degree } \geq 8
\]
\[
= \frac{\beta^2 - \alpha \gamma}{2} \cdot [x, y, x] + \alpha \beta [x, y, x] - \alpha^2 [x, y, y] + \text{terms of degree } \geq 8.
\]
It follows that
\[
\min(g_1) = \min([e^x, e^y, e^x, e^y, [e^x, e^y]_{\hat{G}}, [e^x, e^y]_{\hat{G}}]) = \frac{\beta^2 - \alpha \gamma}{2} \cdot [x, y, x] \mod L^{(7)}.
\]

A similar computation shows that
\[
\min(g_3) = \min((e^x, e^y, e^y, [e^x, e^y]_{\Phi(\hat{G})}) = \frac{\beta^2 - \alpha \gamma}{2} \cdot [x, y, y] \mod L(7).
\]
Regarding \(g_2\) we have
\[
\ln((e^x, e^y, e^x, e^y)_{\Phi(\hat{G})}) = \beta [x, y] - \alpha [x, y, y] - \beta [x, y, y] + \text{terms of degree } \geq 6.
\]

Hence
\[
\min(g_2) = \min((e^x, e^y, e^x, e^y, [e^x, e^y]_{\Phi(\hat{G})}) = \beta \cdot (\frac{\beta [x, y, x] + \alpha [x, y, y]}{2}) + \beta \cdot (\frac{\beta [x, y, y] + \gamma [x, y, x]}{2}) \mod L(7)
\]
as required.
[MKS] W. Magnus, A. Karrass, D. Solitar, Combinatorial group theory: Presentations of groups in terms of generators and relations, Dover Publications, Inc., New York, 1976.

[N] M. F. Newman, The soluble length of soluble linear groups, Math. Z. 126, 1972, 59–70.

[P] C. Procesi, The invariant theory of $n \times n$ matrices, Advances in Math. 19, 1976, 306–381.

[R] Ju. P. Razmyslov, Identities with trace in full matrix algebras over a field of characteristic zero (Russian), Izv. Akad. Nauk SSSR Ser. Mat. 38, 1974, 723–756.

[Ze1] E. Zelmanov, Infinite algebras and pro-$p$ groups, Infinite groups: geometric, combinatorial and dynamical aspects, Progr. Math. 248, Birkhäuser, Basel, 2005, 403–413.

[Ze2] E. Zelmanov, Groups with identities, Note. Mat. 36, 2016, 101–113.

[Zu] A. N. Zubkov, Non-abelian free pro-$p$-groups cannot be represented by 2-by-2 matrices, Siberian Mathematical Journal 28, 1987, 742-747.

Ben-Ezra, David El-Chai
Department of Mathematics
University of California in San-Diego
La Jolla, CA 92093
USA
davidel-chai.ben-ezra@mail.huji.ac.il

Zelmanov, Efim
Department of Mathematics
University of California in San-Diego
La Jolla, CA 92093
USA
eelmanov@ucsd.edu