Advances in Group Theory and Applications
© 2018 AGTA - www.advgrouptheory.com/journal
6 (2018), pp. 3–21 ISSN: 2499-1287
DOI: 10.32037/agta-2018-002

Fox Derivatives: a Unique Connection Between Group Presentations and the Integral Group Ring *

Aryan Ghobadi

(*Received Sep. 9, 2017; Accepted Aug. 20, 2018 — Communicated by E. Jespers*)

Dedicated to Iraj and Mehri

Abstract

We give a criterion to determine whether generators can be removed from a finite presentation via Tietze transformations. We prove that for a generator in a presentation $\langle X|R\rangle$ to be removable, there must exist a word in the normal closure of relators, $\overline{R}$, whose Fox derivative is an invertible element in $\mathbb{Z}G$. Furthermore, in this case all elements of $\mathbb{Z}G$ can be written as the derivative of words in $\overline{R}$, with respect to the removable generator. We further discuss the application of this result on the theory units of group rings.

Mathematics Subject Classification (2010): 20F05, 20F10
Keywords: group presentation; Fox derivative; unit of the integral group ring

1 Introduction

Motivated by Knot Theory, Ralph H. Fox defined what we now call the Fox Derivative for free groups and extended the definition to all groups via group presentations. As a result, the Jacobian matrix and the elementary

* This work is partially supported by the UCL Mathematics Department

ideals, as defined by Fox, reveal critical information about both the group's isomorphism type and its presentation. Most of these results are described in [1] and [2]. In particular, in [2] it is demonstrated that Tietze Transformations, when applied to the Jacobian, give an equivalent matrix. The work here, as described in the Preliminaries, is inspired by the cases where the inverse of this correspondence fails. Specifically, if the presentation has a superfluous generator, then the Jacobian matrix has a column and row of certain description. Our question is whether a presentation with a Jacobian satisfying that description, must have a generator that is removable in the following sense.

Definition 1.1 *Given a presentation* $\langle X|R\rangle$ *for group* G *and* $y \in X$*, we say that generator* y *is removable, if there exists a word* $w \in F(X \setminus \{y\})$ *such that* $y^{-1}w \in \overline{R}$*.*

Although the correspondence of presentations and equivalent Jacobians fails, our main result shows that if we find a column and row of the mentioned description, a generator is indeed removable, and superfluous in the presentation. Specifically, this happens if and only if the Fox derivative of a word in the normal closure of R is invertible.

Theorem 1.2 *Let* $\langle X|R\rangle$ *be a presentation for group* G *and* $y \in X$*. Then, there exists a word* $f \in \overline{R}$ *such that*

$$\Psi\frac{\partial f}{\partial y}$$

is an invertible element in $\mathbb{Z}G$ *if and only if the generator* y *is removable in this presentation.*

Here Ψ is the map evaluating words in $F(X)$ as group elements in G, when linearly extended to $\mathbb{Z}F(X)$, and

$$\frac{\partial}{\partial y} : F(X) \to \mathbb{Z}F(X)$$

is the Fox derivative with respect to y.

In Section 4, we discuss the potential applications of the main theorem to the theory of group rings. In particular, Corollary 4.2 states if y is removable, all non-trivial units in $\mathbb{Z}G$ can be written as the derivatives of words in the normal closure of R. Furthermore, in the proof of the main result, we further describe such words. Hence, we discuss the potential use of the Theorem for finding non-trivial units in the integral group ring.

2 Preliminaries

First we recall the theory of group presentations: given a set X and a subset R of the free group $F(X)$, we say $\langle X|R\rangle$ is a *presentation* for group G, if there exists a surjective homomorphism

$$\psi : F(X) \to G$$

with $\mathrm{Ker}(\psi) = \overline{R}$, where $\overline{R}$ is the normal closure of the set R in $F(X)$. Elements of X and R are called *generators* and *relators*, respectively. Furthemore, we say $\langle X|R\rangle$ is a *finite presentation* if both X and R are finite sets. As a standard reference on presentations of groups we refer to [3].

We recall Tietze's four transformations on the presentation of a group from [3]:

(T1) Addition of relator: $X' = X, R' = R \cup \{r\}$ where $r \in \overline{R} \setminus R$.

(T2) Removal of a relator: $X' = X, R' = R \setminus \{r\}$ where $r \in R \cap \overline{R \setminus \{r\}}$.

(T3) Addition of a generator: $X' = X \cup \{y\}, R' = R \cup \{y^{-1}w\}$ where $y \notin X$ and $w \in F(X)$.

(T4) Removal of generator: $X' = X \setminus \{y\}, R' = R \setminus \{y^{-1}w\}$ where $y \in X$, $w \in F(X \setminus \{y\})$ and $y^{-1}w$ is the only word in R involving y.

where each transformation takes us from a presentation $\langle X|R\rangle$ to presentation $\langle X'|R'\rangle$ for the same group. Notice that given a finite presentation for group G, if generator y is removable as in Definition 1.1, one can first add the relator $y^{-1}w$ to the presentation by a (T1) transformation, and then substitute y by w in all other relators by a series of (T1) and (T2) transformations. This will not change the group isomorphism type. Consequently, $y^{-1}w$ will be the only remaining relator containing y and generator y can then be removed by (T4).

We now recall Fox's construction of derivatives on the free group $F(X)$ from [1]: for any $x \in X$, there exists a map

$$\frac{\partial}{\partial x} : F(X) \to \mathbb{Z}F(X)$$

with defining properties

$$\text{(a)} \quad \frac{\partial y}{\partial x} = \begin{cases} 0 & y \neq x \\ 1 & y = x \end{cases} \qquad\qquad \text{(b)} \quad \frac{\partial wv}{\partial x} = \frac{\partial w}{\partial x} v + \frac{\partial v}{\partial x}$$

for $y \in X$ and $w, v \in F(X)$. From the above axioms one can easily conclude that for any word

$$w = u_m x^{p_m} u_{m-1} \ldots x^{p_1} u_0 \in F(X),$$

where $u_i \in F(X \setminus \{x\})$ and p_i are non-zero integers,

$$\frac{\partial w}{\partial x} = \sum_{i=1}^{m} \left(x^{p_i - 1} + \ldots + x + 1 \right) u_{i-1} x^{p_{i-1}} \ldots x^{p_1} u_0 \tag{2.1}$$

Another important corollary of (a) and (b) is the identity

$$\frac{\partial w^{-1}}{\partial x} = -\frac{\partial w}{\partial x} w^{-1} \tag{2.2}$$

holding for any word $w \in F(X)$.

Now recall Fox's extension of the derivative to the group ring $\mathbb{Z}G$. Let

$$\Psi : \mathbb{Z}F(X) \to \mathbb{Z}G$$

be the natural linear extension of the map ψ. In [2], Fox explores the results which arise when applying the composite map

$$\Psi \frac{\partial}{\partial x} : F(X) \to \mathbb{Z}G$$

on the relators of the presentation.

Notation

- From this point we will use notation $\simeq$ for equality in $\mathbb{Z}G$, and $=$ for equality in $\mathbb{Z}F$, to avoid confusion between two words being equal and two words representing the same group element. Moreover, when we say a word w has *derivative* equal to $\alpha \in \mathbb{Z}G$, we mean $\Psi \frac{\partial w}{\partial x} \simeq \alpha$.
- Given a group presentation $\langle X | R \rangle$ with $y \in X$, we say y *appears* in a word $w \in F(X)$ if w belongs to $F(X) \setminus F(X \setminus \{y\})$.

The identities in the following propositions follow from (2.2) and the definition of the Fox derivative. We will regularly use them without further notice.

Proposition 2.1 *For $s, t \in \overline{R}$, the following identities hold:*

- $\Psi \frac{\partial s^{-1}}{\partial x} \simeq -\Psi \frac{\partial s}{\partial x}$
- $\Psi \frac{\partial st}{\partial x} \simeq \Psi \frac{\partial s}{\partial x} + \Psi \frac{\partial t}{\partial x}$
- $\Psi \frac{\partial w^{-1} s w}{\partial x} \simeq \left(\Psi \frac{\partial s}{\partial x}\right) \psi(w), \quad w \in F(X)$

We now recall the definition of the Jacobian from [2]. For a finite presentation

$$\langle x_1, x_2, \ldots x_g | f_1, f_2, \ldots, f_m \rangle$$

of group G, we call the matrix

$$J = \begin{pmatrix} \Psi \frac{\partial f_1}{\partial x_1} & \Psi \frac{\partial f_1}{\partial x_2} & \ldots & \Psi \frac{\partial f_1}{\partial x_g} \\ \Psi \frac{\partial f_2}{\partial x_1} & \Psi \frac{\partial f_2}{\partial x_2} & \ldots & \Psi \frac{\partial f_2}{\partial x_g} \\ \vdots & \vdots & & \vdots \\ \Psi \frac{\partial f_m}{\partial x_1} & \Psi \frac{\partial f_m}{\partial x_2} & \ldots & \Psi \frac{\partial f_m}{\partial x_g} \end{pmatrix}$$

the Jacobian of the presentation. In [2], Fox showed that Jacobians generated by two distinct presentations of the same group are *equivalent* matrices, where equivalence is as described below in Definition 2.2 (b). This is done by showing that Tietze transformations keep the matrices in the same equivalence class. Recall Tietze's criterion for a generator y to be removable from a presentation, (T4): the only relator the generator appears in, must have the form $y^{-1}w$ where $w \in F(X \setminus \{y\})$. Hence,

$$\Psi \frac{\partial y^{-1} w}{\partial y} = -1$$

and the derivative of the other relators are zero with respect to y.

The motivation behind our work is to construct a pathway in the opposite direction and possibly a one-to-one correspondence between presentations of a group and an equivalence class of matrices via Jacobians.

Definition 2.2

(a) *For positive integers* k *and* l, *by right linear combinations of*

$$a_1, a_2, \ldots, a_l \in \mathbb{Z}G^k,$$

we refer to terms

$$a_1 b_1 + a_2 b_2 + \ldots + a_l b_l \in \mathbb{Z}G^k$$

where $b_i \in \mathbb{Z}G$.

(b) *Given a Jacobian* J, *a matrix* J' *is said to be equivalent to* J *if a sequence of the following elementary operations can be applied to* J *to obtain* J'.

(I) *Permute rows or columns.*

(II) *Adjoin a new row to the matrix and the new row can be written as a right linear combination of other rows.*

(III) *Remove a row, a row that can be written as a right linear combination of the other rows of the matrix.*

(IV) *Adjoin a column and row to the matrix, where the intersection of the new row and column is* -1 *and the other entries of the column are zeros:*

$$M \longrightarrow \begin{pmatrix} M & 0 \\ * & -1 \end{pmatrix}.$$

(V) *Remove a column, a column that has a unique non-zero entry of* -1 *is removed along with the respective row*

$$\begin{pmatrix} M & 0 \\ * & -1 \end{pmatrix} \longrightarrow M.$$

Notice that operation (I) corresponds to permuting the generators or relators and does not change the group isomorphism type and thereby does not affect our discussion. Moreover, operations (II) and (III) correspond to the addition and removal of a relator by (T1) and (T2). However, the two remaining operations (IV) and (V) obstruct the desired correspondence. Operation (IV) does not necessarily correspond to adjoining a generator directly since the entries of the new row are a random set of elements in $\mathbb{Z}G$ and need not be derivatives of a word in $F(X)$. However, if they are derivatives of a word $w \in F(X)$, then the operation can be thought of adjoining a new generator y and its corresponding relator $y^{-1}w$.

The case which interests us is that of operation (V). The issue here is that simple examples can be found where the removal of a column of qualifying form, with a unique non-zero entry -1, and its corresponding row in the Jacobian can not be tracked back to the removal of a generator. The main problem being that the matrix is evaluated at $\mathbb{Z}G$ and if $\Psi\frac{\partial w}{\partial y} \simeq 0$, one can not conclude that the word w does not include powers of y. Furthermore, not every word f with $\Psi\frac{\partial f}{\partial y} \simeq -1$ has the form $y^{-1}w$.

Example 2.3 Consider the presentation

$$\left\langle a, b, y \mid a^3, b^2, abab, ay^{-1}ay^{-1}ay^2, a^3y^{-1}a^2bay^{-1}ay^{-1}ay^2b \right\rangle$$

for the dihedral group D_6. By consecutive use of the fourth, second and first relators, it's easy to derive from the fifth relator that

$$y^{-1}a^2 \simeq a^3(y^{-1}a^2(b(ay^{-1}ay^{-1}ay^2)b)) \in \overline{R}$$

Hence, using $y \simeq a^2$, we calculate the derivative of the relators accordingly;

$$\begin{aligned}
\frac{\partial a^3}{\partial y} = \frac{\partial b^2}{\partial y} = \frac{\partial abab}{\partial y} = 0 \quad &\Rightarrow \quad \Psi\frac{\partial a^3}{\partial y} \simeq \Psi\frac{\partial b^2}{\partial y} \simeq \Psi\frac{\partial abab}{\partial y} \simeq 0 \\
\frac{\partial ay^{-1}ay^{-1}ay^2}{\partial y} &= -y^{-1}ay^{-1}ay^2 - y^{-1}ay^2 + y + 1 \\
\Psi\frac{\partial ay^{-1}ay^{-1}ay^2}{\partial y} &\simeq -a^2 - 1 + a^2 + 1 \simeq 0 \\
\frac{\partial a^3y^{-1}a^2bay^{-1}ay^{-1}ay^2b}{\partial y} &= -y^{-1}a^2bay^{-1}ay^{-1}ay^2b \\
&\quad - y^{-1}ay^{-1}ay^2b - y^{-1}ay^2b + yb + b \\
\Psi\frac{\partial a^3y^{-1}a^2bay^{-1}ay^{-1}ay^2b}{\partial y} &\simeq -1 - a^2b - b + a^2b + b \simeq -1
\end{aligned}$$

First observe that $ay^{-1}ay^{-1}ay^2$ has four appearances of y but its derivative is 0. Additionally, $a^3y^{-1}a^2bay^{-1}ay^{-1}ay^2b$ does not have the form $y^{-1}w$

but has derivative -1.

However, we will show that if one can find a word in $\overline{R}$ with derivative -1, then it is possible to construct a word in $\overline{R}$ of the form $y^{-1}w$, such that w does not contain powers of y. Consequently, Theorem 1.2 implies that given a Jacobian with a column of qualifying form, although one can not directly remove the column, a row and the generator straight away, the generator is indeed removable. Hence, given a Jacobian of the following form

$$\begin{pmatrix} J & 0 \\ * & -1 \end{pmatrix}$$

Although reducing the matrix to J might not correspond to a presentation of the same group, the generator corresponding to the last column can indeed be removed. For instance, in Example 2.3, we can not remove generator y directly, along with the fifth relator. However, the following transformation is possible: we add relator $y^{-1}w$, replace y in the other relators and remove generator y by (T4) i.e.

$$\begin{pmatrix} J & 0 \\ * & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} J & 0 \\ * & -1 \\ \star & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} J' & 0 \\ *' & 0 \\ \star & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} J' \\ *' \end{pmatrix}$$

where J' is of the same size of J and the entries of J' and rows $\star$ and $*'$ depend on the word w with which we have replaced generator y with.

3 Main result

We recall the result that given a group G, the integral group ring $\mathbb{Z}G$ is *Dedekind-finite* i.e. left-invertible elements in $\mathbb{Z}G$ are right-invertible and vice-versa. For a proof of this result refer to [4]. Hence, we can refer to left (or right)- invertible elements in $\mathbb{Z}G$ as simply *invertible/units* and denote the group of units in $\mathbb{Z}G$ by $\mathcal{U}(\mathbb{Z}G)$.

Lemma 3.1 *The following statements are equivalent:*

(a) *There exists a word* $f \in \overline{R}$ *such that* $\Psi\frac{\partial f}{\partial y}$ *is an invertible element in* $\mathbb{Z}G$.

(b) *There exists a word* $f \in \overline{R}$ *such that* $\Psi\frac{\partial f}{\partial y} \simeq 1$.

(c) *There exists a word* $f \in \overline{R}$ *such that* $\Psi\frac{\partial f}{\partial y} \simeq -1$.

(d) *The map* $\Psi\frac{\partial}{\partial y} : \overline{R} \to \mathbb{Z}G$ *is surjective.*

Proof — The equivalence of (b) and (c) follows directly from Proposition 2.1. The rest of the Lemma follows from the below observation.

Consider $f \in \overline{R}$ and $\beta \in \mathbb{Z}G$, then there exist

$$\lambda_i = \pm 1 \quad \text{and} \quad g_i \in G$$

for $1 \leqslant i \leqslant M$, for some finite number M such that

$$\beta \simeq \sum_{i=1}^{M} \lambda_i g_i$$

Notice that here we are allowing repetitions of the same group element in g_i i.e. we are writing $\beta \simeq 3g \in \mathbb{Z}\langle g \rangle$ as

$$\beta \simeq \sum_{i=1}^{3} \lambda_i g_i,$$

with $g_i = g$ and $\lambda_i = 1$ for all i. Since the map ψ is surjective, then there exist words $w_i \in F(X)$ such that $\psi(w_i) \simeq g_i$, and thereby

$$\beta \simeq \Psi \left(\sum_{i=1}^{M} \lambda_i w_i \right)$$

We now use the previous identities from Proposition 2.1:

$$\left(\Psi \frac{\partial f}{\partial y} \right) \beta \simeq \Psi \left(\sum_{i=1}^{M} \lambda_i \frac{\partial f}{\partial y} w_i \right)$$

$$\simeq \Psi \frac{\partial \left((w_1^{-1} f w_1)^{\lambda_1} (w_2^{-1} f w_2)^{\lambda_2} \dots (w_M^{-1} f w_M)^{\lambda_M} \right)}{\partial y}.$$

For the case of (a)$\Rightarrow$(b), consider $f \in \overline{R}$, where $\Psi \frac{\partial f}{\partial y}$ is invertible in $\mathbb{Z}G$ and $\beta \in \mathbb{Z}G$ is its inverse, such that

$$\left(\Psi \frac{\partial f}{\partial y} \right) \beta \simeq 1.$$

Then the word

$$v = (w_1^{-1} f w_1)^{\lambda_1} (w_2^{-1} f w_2)^{\lambda_2} \dots (w_M^{-1} f w_M)^{\lambda_M}$$

satisfies $\Psi \frac{\partial v}{\partial y} \simeq 1$ and belongs to $\overline{R}$.

For the case of (b)⇒(d), consider $f \in \overline{R}$, where

$$\Psi \frac{\partial f}{\partial y} = 1.$$

Then for any $\beta \in \mathbb{Z}G$, with notation as before, the word

$$v = (w_1^{-1} f w_1)^{\lambda_1} (w_2^{-1} f w_2)^{\lambda_2} \dots (w_M^{-1} f w_M)^{\lambda_M}$$

satisfies $\Psi \frac{\partial v}{\partial y} \simeq \beta$ and belongs to $\overline{R}$. □

For $w, v \in F(X)$, we will say v is a *subword* of w, if there exist words

$$u, p \in F(X)$$

such that $w = uvp$ in its reduced form. For example, we say yx is a subword of $w = xyx$, however, $x^{-1}yx$ is not a subword of w, although

$$x^2 x^{-1} yx = xyx.$$

Furthermore, we say a subword v of $w \in F(X)$ is a *right-subword* of w, if for some $u \in F(X)$, $w = uv$ in its reduced form.

Lemma 3.2 *There exists a word* $f \in \overline{R}$ *such that*

$$\Psi \frac{\partial f}{\partial y} \simeq -1,$$

if and only if there exists a word $f' \in \overline{R}$ *and* $w \in F(X)$ *such that* f' *has the form*

$$f' = y^{-1} w$$

with $\Psi \frac{\partial f'}{\partial y} \simeq -1$.

Proof — If we look at the derivative evaluated at $\mathbb{Z}F$

$$\frac{\partial f}{\partial y} = \sum_{i=1}^{n} (-1)^{\mu_n} w_i \tag{3.1}$$

where $\mu_n = 0, 1$ and, by the definition of the Fox derivative, $w_i \in F(X)$ are all right subwords of f. Then by applying Ψ, we have

$$\Psi \frac{\partial f}{\partial y} \simeq \Psi \left(\sum_{i=1}^{n} (-1)^{\mu_n} w_i \right) \simeq \sum_{i=1}^{n} (-1)^{\mu_n} \psi(w_i) \simeq -1$$

where $\psi(w_i) \in G$. Therefore, there exists a w_i such that $\psi(w_i) \simeq 1$ and $\mu_i = 1$ and furthermore, the sum of the other terms is zero. Hence by re-naming the terms of the original sum we get the following identity

$$\frac{\partial f}{\partial y} = -\omega + \sum_{i=1}^{m} (v_i - v_i') \tag{3.2}$$

where $\omega, v_i, v_i' \in \{w_1, \ldots, w_n\}$ and $\psi(\omega) \simeq 1$, $\psi(v_i) \simeq \psi(v_i')$. It is also clear that $m = (n-1)/2$. Notice by the definition of the derivative, each word in expansion (3.1) is equal to the derivative of a subword of f with respect to y. Since the derivative only admits negative signs for negative powers of y, subword ω of f must have form $y^{-1}w'$ for some $w' \in F(X)$:

$$\frac{\partial y^{-1}w'}{\partial y} = -y^{-1}w' + \frac{\partial w'}{\partial y} = -\omega + \frac{\partial w'}{\partial y}$$

Since $y^{-1}w'$ is a right-subword of f, there must also exist a word $w'' \in F(X)$ so that

$$f = w''y^{-1}w'.$$

Thereby, since $w''y^{-1}w' \in \overline{R}$, the word conjugated by w'' will also remain in the normal closure, and

$$f' := y^{-1}w'w'' \in \overline{R}.$$

Because $\psi(\omega) \simeq 1$, one can conclude that $y^{-1}w' \in \overline{R}$ and thereby $w'' \in \overline{R}$.

$$\Psi\frac{\partial f}{\partial y} \simeq \Psi\frac{\partial w''y^{-1}w'}{\partial y} \simeq \Psi\left(\frac{\partial w''}{\partial y}y^{-1}w' - y^{-1}w' + \frac{\partial w'}{\partial y}\right)$$

$$\simeq \Psi\frac{\partial w''}{\partial y} - 1 + \Psi\frac{\partial w'}{\partial y}$$

$$\Psi\frac{\partial f'}{\partial y} \simeq \Psi\frac{\partial y^{-1}w'w''}{\partial y} \simeq \Psi\left(-y^{-1}w'w'' + \frac{\partial w'}{\partial y}w'' + \frac{\partial w''}{\partial y}\right)$$

$$\simeq -1 + \Psi\frac{\partial w'}{\partial y} + \Psi\frac{\partial w''}{\partial y}$$

Hence, $f' = y^{-1}w'w''$ belongs to $\overline{R}$ and satisfies $\Psi\frac{\partial f'}{\partial y} \simeq \Psi\frac{\partial f}{\partial y} \simeq -1$. □

Now we can consider a word $f = y^{-1}w \in \overline{R}$ where

$$\Psi\frac{\partial f}{\partial y} \simeq -1.$$

In order for y to be removable, w must not have any powers of y. Hence, the next step is to show we can replace w with a word with fewer powers of y. Similar to (3.2)

$$\frac{\partial f}{\partial y} = -y^{-1}w + \sum_{i=1}^{m}(v_i - v_i') \tag{3.3}$$

where $\psi(v_i) \simeq \psi(v_i')$. Since the Fox derivative admits a term for each appearance of y in the word, then v_i and v_i' are subwords of w, with their corresponding powers of y appearing in w as well. In other words

$$\frac{\partial w}{\partial y} = \sum_{i=1}^{m}(v_i - v_i') \quad \Rightarrow \quad \Psi\left(\frac{\partial w}{\partial y}\right) = \sum_{i=1}^{m}(\psi(v_i) - \psi(v_i')) = 0$$

Since for any given i, both v_i and v_i' are right-subwords of w, one must be a right-subword of the other i.e. either

$$v_i = r_i v_i' \quad \text{or} \quad v_i' = r_i v_i$$

for a word $r_i \in F(X)$. Furthermore,

$$\psi(v_i) \simeq \psi(v_i'),$$

hence $r_i \in \overline{R}$. By the definition of the derivative, we also know that each v_i' must have the form $v_i' = y^{-1}v_i''$ to admit a negative coefficient. Similarly, w must have right-subwords of the form yv_i so that v_i appear in its derivative with positive coefficient. Therefore, w has right-subwords of the form

$$yv_i = yr_iy^{-1}v_i'' \quad \text{or} \quad v_i' = y^{-1}v_i'' = r_iv_i.$$

In the latter case, by the previous argument, yv_i must be a right-subword of w and v_i', and thereby $r_i = y^{-1}s_iy$ for a word s_i. Since $r_i \in \overline{R}$, then in these cases $s_i \in \overline{R}$. Hence, we rename r_i's in the second cases so that subwords of the form $y^{\epsilon_i}r_iy^{-\epsilon_i}$ appear in w, where $\epsilon_i = \pm 1$. Moreover, we define η_i to be the words such that $y^{\epsilon}r_iy^{-\epsilon}\eta_i$ is a right-subword of w. Observe that η_i is either v_i'' or v_i. Therefore, each power of y in w has a corresponding opposite power with a word from $\overline{R}$ in between them. Under the derivative the subwords $y^{\epsilon}r_iy^{-\epsilon}\eta_i$ appear as one of the following forms:

$$\epsilon_i = 1: \quad r_iy^{-1}\eta_i - y^{-1}\eta_i \qquad\qquad \epsilon_i = -1: \quad -y^{-1}r_iy\eta_i + \eta_i$$

which agrees with expansion (3.3) and $\psi(v_i) \simeq \psi(v_i')$ holds, since

$$\psi(r_i y^{-1}\eta_i) \simeq \psi(y^{-1}\eta_i) \qquad\qquad \psi(y^{-1}r_i y\eta_i) \simeq \psi(\eta_i).$$

Hence, we have proved the following lemma.

Lemma 3.3 *Given $w \in F(X)$, $\Psi\frac{\partial w}{\partial y} \simeq 0$ if and only if for each power of y, there exists an opposite power of y in the word w, such that the subword between the two powers of y, belongs to $\overline{R}$.*

Now we describe how given a word w satisfying

$$\Psi\frac{\partial w}{\partial y} \simeq 0,$$

one can construct a word $w' \simeq w$, such that

$$\Psi\frac{\partial w'}{\partial y} \simeq 0,$$

and w' has a reduced number of appearances of y, compared to w.

Definition 3.4 *By a* matching π *on finite set A, we mean a bijective map*

$$\pi : A \to A,$$

such that $\pi(x) \neq x$ for any $x \in A$ and $\pi^2 = \mathrm{id}$.

In its most general form we know

$$w = w_0 y^{\epsilon_1} w_1 y^{\epsilon_2} w_2 y^{\epsilon_3} w_3 \ldots y^{\epsilon_{2m}} w_{2m}$$

where $w_i \in F(X \setminus \{y\})$ and $\epsilon_i = \pm 1$. By Lemma 3.3, there exists a matching

$$\pi : \{1, 2, \ldots, 2m\} \to \{1, 2, \ldots, 2m\}$$

where $\epsilon_{\pi(i)} = -\epsilon_i$, for any given $i \in \{1, 2, \ldots, 2m\}$. We also denote the word between y^{ϵ_i} and $y^{\epsilon_{\pi(i)}}$ by $r_i \in \overline{R}$, with $r_i = r_{\pi(i)}$.

Steps to reduce m

We take the first power of y i.e. y^{ϵ_1} and consider the matching π. The following cases are possible.

(I) $\pi(1) = 2$.

(II) $\pi(k) \leqslant \pi(1)$ for $1 \leqslant k \leqslant \pi(1)$.

(III) There exists $k < \pi(1)$ such that $\pi(k) \geqslant \pi(1)$.

In case (I) one can remove

$$y^{\epsilon_1} w_1 y^{\epsilon_2}$$

from w in its entirety, since $w_1 = r_1 \in \overline{R}$ and $y^{\epsilon_1} w_1 y^{\epsilon_2} = y^{\epsilon_1} w_1 y^{-\epsilon_1} \in \overline{R}$. Let $\upsilon := w_2 y^{\epsilon_3} w_3 \dots y^{\epsilon_{2m}} w_{2m}$, then

$$\begin{aligned}
0 \simeq \Psi \frac{\partial w}{\partial y} &\simeq \Psi \frac{\partial w_0 y^{\epsilon_1} w_1 y^{-\epsilon_1} \upsilon}{\partial y} \\
&\simeq \Psi \left(\frac{\partial w_0}{\partial y} y^{\epsilon_1} r_1 y^{-\epsilon_1} \upsilon + \frac{\partial y^{\epsilon_1}}{\partial y} r_1 y^{-\epsilon_1} \upsilon + \frac{\partial y^{-\epsilon_1}}{\partial y} \upsilon + \frac{\partial \upsilon}{\partial y} \right) \\
&\simeq \Psi \frac{\partial w_0}{\partial y} \upsilon + \Psi \frac{\partial \upsilon}{\partial y} \\
&\simeq \Psi \frac{\partial w_0 \upsilon}{\partial y}
\end{aligned}$$

Hence, the word

$$w' = w_0 \upsilon = w_0 w_2 y^{\epsilon_3} w_3 \dots y^{\epsilon_{2m}} w_{2m}$$

has $2m - 2$ powers of y and $y^{-1} w'$ belongs to $\overline{R}$ with

$$\Psi \frac{\partial y^{-1} w'}{\partial y} \simeq -1.$$

In case (II), y^{ϵ_1} is in a block which can be removed from the word since its derivative is zero. Consider the subword

$$v := y^{\epsilon_1} w_1 y^{\epsilon_2} w_2 y^{\epsilon_3} w_3 \dots w_{\pi(1)-1} y^{\epsilon_{\pi(1)}}$$

Since $\pi(k) \leqslant \pi(1)$ for $1 \leqslant k \leqslant \pi(1)$, the map π is a matching when restricted to the domain

$$\{1, 2, \dots, \pi(1)\}.$$

Since v is a subword of w, then the subwords contained between matched powers of y in v also belong to $\overline{R}$ and by Lemma 3.3, $\psi \frac{\partial v}{\partial y} \simeq 0$. Furthermore, since

$$v = y^{\epsilon_1} r_1 y^{-\epsilon_1}$$

belongs to $\overline{R}$, then we can remove it from f. Hence

$$w' = w_0 w_{\pi(1)} y^{\epsilon_{\pi(1)+1}} w_{\pi(1)+1} \dots y^{\epsilon_{2m}} w_{2m}$$

has derivative zero and $y^{-1} w' \in \overline{R}$. Thereby, we have reduced the number of appearances of y by $\pi(1) \geqslant 2$. Notice that case (I) is a specific case of (II) where $\pi(1) = 2$.

In case (III) we look at the subword which y^{ϵ_1} is interacting with. Let $k \leqslant \pi(1)$ such that $\pi(1) < \pi(k)$. Then there exist subwords v_0, v_1 and v_2 such that $r_1 = v_0 y^{\epsilon_k} v_1$ and $r_k = v_1 y^{-\epsilon_1} v_2$. Hence

$$y^{\epsilon_1} w_1 y^{\epsilon_2} w_2 y^{\epsilon_3} w_3 \dots w_{\pi(k)-1} y^{\epsilon_{\pi(k)}} = y^{\epsilon_1} v_0 y^{\epsilon_k} v_1 y^{-\epsilon_1} v_2 y^{-\epsilon_k}.$$

Since $r_1, r_k \in \overline{R}$, we consider the following relations

$$\text{Under} \quad \psi \begin{cases} r_1 = v_0 y^{\epsilon_k} v_1 \simeq e \\ r_k = v_1 y^{-\epsilon_1} v_2 \simeq e \end{cases} \Rightarrow \quad v_0 y^{\epsilon_k} \simeq v_1^{-1} \simeq y^{-\epsilon_1} v_2$$

$$v_0 y^{\epsilon_k} \simeq y^{-\epsilon_1} v_2 \quad \Rightarrow \quad \begin{cases} v_0 \simeq y^{-\epsilon_1} v_2 y^{-\epsilon_k} \\ v_2 \simeq y^{\epsilon_1} v_0 y^{\epsilon_k} \end{cases}.$$

Hence, we can replace v_0 and v_2 with the above identities:

$$\text{Under} \quad \psi : \quad y^{\epsilon_1} v_0 y^{\epsilon_k} v_1 y^{-\epsilon_1} v_2 y^{-\epsilon_k} \simeq v_2 v_1 v_0.$$

Recognize that this change reduces 4 appearances of powers of y in f. Now we show that this change does not affect the Fox derivative:

$$\psi \frac{\partial v_2 v_1 v_0}{\partial y} \simeq \psi \frac{\partial v_2}{\partial y} v_1 v_0 + \psi \frac{\partial v_1}{\partial y} v_0 + \psi \frac{\partial v_0}{\partial y}.$$

Now using the identities $v_0 \simeq y^{-\epsilon_1} v_2 y^{-\epsilon_k}$ and $v_0 y^{\epsilon_k} v_1 \simeq y^{\epsilon_k} v_1 v_0 \simeq e$:

$$\begin{aligned} &\psi \frac{\partial y^{\epsilon_1} v_0 y^{\epsilon_k} v_1 y^{-\epsilon_1} v_2 y^{-\epsilon_k}}{\partial y} \\ &\simeq \psi \frac{\partial v_0}{\partial y} y^{\epsilon_k} v_1 y^{-\epsilon_1} v_2 y^{-\epsilon_k} + \psi \frac{\partial v_1}{\partial y} y^{-\epsilon_1} v_2 y^{-\epsilon_k} + \psi \frac{\partial v_2}{\partial y} y^{-\epsilon_k} \\ &\simeq \psi \frac{\partial v_0}{\partial y} y^{\epsilon_k} v_1 v_0 + \psi \frac{\partial v_1}{\partial y} v_0 + \psi \frac{\partial v_2}{\partial y} y^{-\epsilon_k} (y^{\epsilon_k} v_1 v_0) \\ &\simeq \psi \frac{\partial v_0}{\partial y} + \psi \frac{\partial v_1}{\partial y} v_0 + \psi \frac{\partial v_2}{\partial y} v_1 v_0 \simeq \psi \frac{\partial v_2 v_1 v_0}{\partial y}. \end{aligned}$$

Thereby, replacing $y^{\epsilon_1} v_0 y^{\epsilon_k} v_1 y^{-\epsilon_1} v_2 y^{-\epsilon_k}$ by $v_2 v_1 v_0$ in the word f does not change the derivative and by the identities above this replacement is possible. Let

$$w' = w_0 v_2 v_1 v_0 w_{\pi(k)} y^{\epsilon_{\pi(k)+1}} \dots y^{\epsilon_{2m}} w_{2m}$$

Then w' has $2m - 4$ powers of y, has derivative zero and $f' = y^{-1} w'$ belongs to $\overline{R}$. Hence, we can remove all powers of y from w by repeating the above process. Therefore, one can construct a word $y^{-1}w$ in $\overline{R}$ with $w \in F(X \setminus \{x\})$, starting with any word $f = y^{-1}w$ such that

$$\Psi \frac{\partial y^{-1} w}{\partial y} \simeq -1 + \Psi \frac{\partial w}{\partial y} \simeq -1,$$

proving Theorem 1.2.

4 Remarks and applications

First observe that Lemma 3.1 helps us extend Theorem 1.2 as follows:

Corollary 4.1 *Let $\langle X|R\rangle$ be a presentation for group G and $y \in X$. Then the following statements are equivalent:*

(a) *The generator y is removable in this presentation.*

(b) *There exists a word $f \in \overline{R}$ such that $\Psi \frac{\partial f}{\partial y} \in \mathcal{U}(\mathbb{Z}G)$.*

(c) *There exists a word $f \in \overline{R}$ such that $\Psi \frac{\partial f}{\partial y} \simeq 1$.*

(d) *There exists a word $f \in \overline{R}$ such that $\Psi \frac{\partial f}{\partial y} \simeq -1$.*

(e) *The map $\Psi \frac{\partial}{\partial y} : \overline{R} \to \mathbb{Z}G$ is surjective.*

4.1 Application to the theory of group rings

In particular statement (e) implies that for the subset

$$\mathcal{U}_y = \left(\Psi \frac{\partial}{\partial y}\right)^{-1} (\mathcal{U}(\mathbb{Z}G)) \cap \overline{R}$$

of words in $\overline{R}$, the map $\Psi \frac{\partial}{\partial y} : \mathcal{U}_y \to \mathcal{U}(\mathbb{Z}G)$ is surjective.

Corollary 4.2 *Let $\langle X|R\rangle$ be a presentation for group G and $y \in X$. Then the generator y is removable if and only if for all $\beta \in \mathcal{U}(\mathbb{Z}G)$ there exists a word $f \in \overline{R}$ such that $\Psi \frac{\partial f}{\partial y} \simeq \beta$.*

A prominent question in the theory of group rings, is to understand the units of group ring $\mathbb{Z}G$ for a given group G. In particular, other than the *trivial units* $\pm g$, where $g \in G$, there are several known recipes for constructing units, such as Bass and bicyclic units. However, a full description of the unit group $\mathcal{U}(\mathbb{Z}G)$ is available for a small number of groups G. For standard

references on units of group rings, we refer to Chapter 8 of [5] and the survey [6].
If $\langle X \mid R \rangle$ is a presentation for group G, then for any arbitrary word $w \in F(X)$, we have an alternate presentation

$$\left\langle X \cup \{y\} | R \cup \{y^{-1}w\} \right\rangle$$

of the same group with a removable generator y. Hence, Corollary 4.2 implies that any non-trivial unit in $\mathbb{Z}G$ is the derivative of a word in $\overline{R}$. In fact, statement (e) of Corollary 4.1 implies that all elements of $\mathbb{Z}G$ are derivatives of words in $\overline{R}$, but while proving Theorem 1.2 we have further characterised those words whose derivative belongs to $\mathcal{U}(\mathbb{Z}G)$.

Definition 4.3 *Given the presentation $\langle X | R \rangle$ for a group G with generator $y \in X$,*

(I) *we denote* $\left(\Psi \frac{\partial}{\partial y}\right)^{-1}(0)$ *by* Z_y*;*

(II) *if y is removable, we denote* $\left[\left(\Psi \frac{\partial}{\partial y}\right)^{-1}(\mathcal{U}(\mathbb{Z}G))\right] \cap \overline{R}$ *by* $\mathcal{U}_y$*;*

(III) *if y is removable, we denote* $\left[\left(\Psi \frac{\partial}{\partial y}\right)^{-1}(1)\right] \cap \overline{R}$ *by* O_y*.*

In the proofs of Lemmas 3.1, 3.2 and 3.3 the following relations between $\mathcal{U}_y, O_y$ and Z_y were established.
If $f \in \overline{R}$ and $y \in X$ is removable, then

(A) $f \in \mathcal{U}_y$ if and only if there exists $\beta \in \mathbb{Z}G$, the inverse of $\Psi \frac{\partial f}{\partial y}$, such that

$$\Psi \frac{\partial \left((w_1^{-1} f w_1)^{\lambda_1} (w_2^{-1} f w_2)^{\lambda_2} \ldots (w_M^{-1} f w_M)^{\lambda_M}\right)}{\partial y} \simeq \Psi \frac{\partial f}{\partial y} \beta \simeq 1$$

where

$$\beta \simeq \Psi \left(\sum_{i=1}^{M} \lambda_i w_i \right).$$

In other words, $f \in \mathcal{U}_y$ if and only if there exists positive integer M,w_i and $\lambda_i = \pm 1$ for $1 \leqslant i \leqslant M$, such that

$$(w_1^{-1} f w_1)^{\lambda_1} (w_2^{-1} f w_2)^{\lambda_2} \ldots (w_M^{-1} f w_M)^{\lambda_M} \in O_y$$

(B) (Symmetric argument to Lemma 3.2) $f \in O_y$ if and only if $f = wyr$ for words $r, w \in F(X)$, $r \in \overline{R}$ (or equivalently $wy \in \overline{R}$) and $rw \in Z_y$.

(C) (Lemma 3.3) $rw \in Z_y$ if and only if when

$$rw = w_0 y^{\epsilon_1} w_1 y^{\epsilon_2} w_2 y^{\epsilon_3} w_3 \dots y^{\epsilon_{2m}} w_{2m}$$

with $w_i \in F(X \setminus \{y\})$ and $\epsilon_i = \pm 1$, then there exists a matching

$$\pi : \{1, 2, \dots, 2m\} \to \{1, 2, \dots, 2m\}$$

where $\epsilon_{\pi(i)} = -\epsilon_i$ and the word between ϵ_i and $\epsilon_{\pi(i)}$ belongs to $\overline{R}$ for any given $1 \leqslant i \leqslant 2m$.

The author believes that for a *finite group* G, it should be possible to write an algorithm in GAP which detects all words satisfying conditions (B) and (C). The next step would be to take a word in $f \in \overline{R}$ and check for positive integers $1 \leqslant M$, upto a bound $\overline{M}$, if for arbitrary group elements $\psi(w_i)$ and $\lambda_i = \pm 1$, the word

$$(w_1^{-1} f w_1)^{\lambda_1} (w_2^{-1} f w_2)^{\lambda_2} \dots (w_M^{-1} f w_M)^{\lambda_M}$$

satisfies (B) and (C). Such an algorithm would both give the unit

$$\alpha \simeq \Psi \frac{\partial f}{\partial y} \quad \text{and its inverse} \quad \alpha^{-1} \simeq \sum_{i=1}^{M} \lambda_i \psi(w_i).$$

The issue with this method however, is the length of the words in question. Observe that if

$$\Psi \frac{\partial f}{\partial y} \simeq \alpha \simeq \sum_{i=1}^{n} a_{g_i} g_i \in \mathbb{Z}G,$$

then the value

$$\sum_{i=1}^{n} |a_{g_i}|$$

is less or equal to the number of appearances of y in f, since the derivative admits a word for every appearance of y in f. Additionally, if

$$\alpha^{-1} \simeq \sum_{i=1}^{n} b_{g_i} g_i,$$

then the number

$$\sum_{i=1}^{n} |b_{g_i}|$$

is less or equal to M. This observation implies that very “large” units might be hard to detect via such an algorithm, where *largeness* of the unit refers to

the number
$$\sum_{i=1}^{n} |a_{g_i}|.$$

For instance, lets look at the example of S. K. Sehgal in [6]:

$$\begin{aligned} u = -372099 + 114985a + 301035a^2 - 301035a^3 - 114985a^4 + 372100a^5 \\ - 114985^6 - 301035a^7 + 301035a^8 + 114985a^9 \end{aligned}$$

is a unit in $\mathbb{Z}G$ where G is the cyclic group of order 10 with presentation $\langle a | a^{10} \rangle$. Such a unit, would be the derivative of a word of approximate length 2×10^6. Further the inverse of u, as presented in [6], is also large of the same order, hence when checking conditions (B) and (C) for u, the program would need to deal with a word of approximate length 4×10^{12}.

Remark 4.4 Observe that at no point during the proof of Theorem 1.2 was the group presentation $\langle X|R \rangle$ assumed to be finite. However if R is not finite, and a relator $y^{-1}w$, with w not containing powers of y exists, then infinitely many Tietze transformations might be required to remove the generator from the presentation. This is the case when y appears in infinitely many relators. In fact our emphasis on G being a finitely presented group during the preliminaries is because the Jacobian is only well-defined in this case, and our original question concerned the Jacobian. Moreover, the algorithm mentioned above would only be possible for groups with decidable word problem. In particular for automatic groups, packages such as kbmag in GAP exist, which can evaluate whether a word belongs to $\overline{R}$.

4.2 Implication for Kaplansky's Unit Conjecture

A prominent question in the theory of group rings is *Kaplansky's Unit Conjecture* which states: If $\mathbb{K}$ is a field and G is a torsion-free group, then $\mathbb{K}G$ has only trivial units. In the case where $\mathbb{K}$ has positive characteristic, the conjecture is trivially true, since for all but a few groups G, the group $\mathbb{K}G$ only has trivial units (see Proposition 8.1.3 of [5]). On the other hand, for any field $\mathbb{K}$ of characteristic 0, we have

$$\mathbb{Z}G \subset \mathbb{K}G.$$

Hence, if Kaplansky's conjecture holds for a group G and a field of characteristic 0, then $\mathcal{U}(\mathbb{Z}G) = \pm G$. In this case, $M = 1$ in condition (A) and thereby
$$\mathcal{U}_y = \{w^{-1} f^{\lambda} w \mid w \in F(X \setminus \{y\}) \quad , \lambda = \pm 1 \quad , f \in O_y\}$$

Hence, a possible strategy for finding a counterexample to Kaplansky's conjecture would be finding a presentation of a Torsion-free group with removable generator y and a word in $\mathcal{U}_y$ not of the above form.

Acknowledgements

Much thanks must go out to Isidoros Strouthos for both introducing the author to the subject and many helpful discussions during his supervision of the original research project in July 2016, funded by the UCL Mathematics Department, although the results presented here were discovered later. The author is also extremely grateful for the referee's many suggestions and corrections, which motivated the addition of Section 4. F.E.A. Johnson and W.H. Mannan must also be thanked for looking at earlier drafts of this work.

REFERENCES

[1] R.H. Fox: "Free differential calculus. I: Derivation in the free group ring", *Ann. Math.* 57 (1953), 547–560.

[2] R.H. Fox: "Free differential calculus. II: The isomorphism problem of groups", *Ann. Math.* 59 (1954), 196–210.

[3] D. L. Johnson: "Presentations of Groups", *Cambridge University Press*, Cambridge (1997).

[4] D.S. Passman: "Idempotents in group rings", *Proc. Amer. Math. Soc.* 28 (1971), 371–374.

[5] C. Polcino Milies – S. K. Sehgal: "An Introduction to Group Rings", *Springer*, Berlin (2002).

[6] S.K. Sehgal: "Units of integral group rings - A survey", in Algebraic Structures and Number Theory, *World Scientific*, Teaneck (1990), 255–268.

Aryan Ghobadi
School of Mathematical Sciences
Queen Mary University of London
Mile End Road
London (UK)
e-mail: aryan.gh75@yahoo.com