Languages and formations generated by D_4 and Q_8*

Jean-Éric Pin[1], Xaro Soler-Escrivà[2]

Abstract

We describe the two classes of languages recognized by the groups D_4 and Q_8, respectively. Then we show that the formations of languages generated by these two classes are the same. We also prove that these two formations are closed under inverses of morphisms, which yields a language theoretic proof of the fact that the group formations generated by D_4 and Q_8, respectively, are two equal varieties.

Most monoids and groups considered in this paper are finite. In particular, we use the term *variety of groups* for *variety of finite groups*. Similarly, all languages considered in this paper are regular languages and hence their syntactic monoid is finite.

1 Introduction

A nontrivial question is to describe the regular languages corresponding to well-studied families of finite groups. Only a few cases have been investigated in the literature: abelian groups [6], p-groups [6, 20, 21, 22], nilpotent groups [6, 19], soluble groups [17, 21] and supersoluble groups [4]. More recently [2], the authors addressed the following question: is it possible to obtain a reasonable description of the languages corresponding to a given formation of groups? Recall that a *formation of groups* is a class of finite groups closed under taking quotients and subdirect products.

This question was motivated by the importance of formations in finite group theory, notably in the development of a generalised Sylow's theory. The theory of formations was born with the seminal paper [7] of Gaschütz in 1963, where a broad extension of Sylow's and Hall's theories was presented. The new theory was not arithmetic, that is, based on the orders of subgroups.

[1]IRIF, CNRS and Université Paris-Diderot, Case 7014, 75205 Paris Cedex 13, France.

[2]Dpt. de Matemàtiques, Universitat d'Alacant, Sant Vicent del Raspeig, Ap. Correus 99, E – 03080 Alacant.

*The first author is supported by Proyecto MTM2014-54707-C3-1-P from MINECO (Spain) and FEDER (European Union) and partially funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 670624) and by the DeLTA project (ANR-16-CE40-0007).

It was concerned instead with group classes sharing certain properties, the so-called formations of groups, which have played a fundamental role in the study of groups since then [5, 1].

In [2], the authors extended Eilenberg's correspondence theorem between varieties of monoids and varieties of languages [6] to the setting of formations. More precisely, they spotted a bijective correspondence between formations of finite monoids and the so-called *formations of languages*. Using this "formation theorem" the authors not only recovered the previously mentioned results on nilpotent groups, soluble groups and supersoluble groups, but, relying on the local definition of a saturated formation [5], they exhibited new examples, like the class of groups having a Sylow tower [3].

The present paper focuses on the language interpretation of two results dealing with the dihedral group D_4 and the quaternion group Q_8. The first result asserts that D_4 and Q_8 generate the same formation [5, Exercise 9, p. 344]. The second one states that this formation is a variety of groups, that is, is closed under taking subgroups. This latter result is actually an instance of a more general result, due to Neumann [10], which states that any formation generated by a single nilpotent group is a variety (see [5, IV.1.16, p. 342] for an alternative proof).

The main result of this paper provides a purely language theoretic proof of these two results on D_4 and Q_8. To do so, we first translate them in terms of languages: the formations of languages $\mathcal{F}_1$ and $\mathcal{F}_2$ associated to D_4 and Q_8, respectively, are the same (first result) and they form a variety of languages (second result). The main difficulty in proving these results by pure language theoretic means is to establish the inclusion $\mathcal{F}_1 \subseteq \mathcal{F}_2$. The lengthy proof of Theorem 5.1 should convince the reader that it is a nontrivial property.

Our proofs rely on a systematic use of the binomial coefficients of two words. This is not really a surprise, since binomial coefficients modulo p are the main tool for describing languages recognized by p-groups, and D_4 and Q_8 are 2-groups. In this paper, we present two explicit formulas with an algorithmic flavour. First, we discuss the behaviour of binomial coefficients under morphisms (Formula 3.4). Next, we show that a language of A^* is recognized by a p-group if and only if it is a finite union of languages defined by linear algebraic constraints involving the binomial coefficients. Finally, we give an algorithm to obtain such a decomposition when the p-group is a group of unitriangular matrices over $\mathbb{F}_p$.

Our paper is organised as follows. In order to keep the paper self-contained, prerequisites (Section 2) include formations and varieties, syntactic monoids and the Formation Theorem. Section 3 is devoted to binomial coefficients on words. We present in Section 4 various descriptions of the languages recognized by p-groups and the corresponding algorithms. Section 5 contains the proof of our main theorem.

2 Prerequisites

2.1 Formations and varieties

A *formation of groups* is a class of groups **F** satisfying the two conditions:

(1) any quotient of a group of **F** also belongs to **F**,

(2) the subdirect product of any finite family of groups of **F** is also in **F**.

Formations of finite algebras can be defined in the same way [14, 16, 15]. In particular, a *formation of monoids* is a class of finite monoids closed under taking quotients and subdirect products. If S is a set of finite monoids, the formation *generated by* S is the smallest formation containing S. It is also the set of quotients of subdirect products of members of S (see [5, II.2.2, p. 272] for group formations, [15, Chapter I, Theorem 2.2] and [8, Lemma 3.2] for general algebraic systems and [2, Proposition 1.4] for a self-contained proof for monoid formations).

A *variety of groups* is a class of groups **V** satisfying the three conditions:

(1) any subgroup of a group of **V** also belongs to **V**,

(2) any quotient of a group of **V** also belongs to **V**,

(3) the direct product of any finite family of groups of **V** is also in **V**.

Varieties of monoids are defined in the same way. It follows from the definition that a formation of groups [monoids] is a variety if and only if it is closed under taking subgroups [submonoids]. Note that a formation is not necessarily a variety. For instance, the formation of groups generated by the alternating group A_5 is known to be the class of all direct products of copies of A_5, which is not a variety [1, Lemma 2.2.3, p. 91], [5, II.2.13].

2.2 Regular languages

A language is *regular* if it is representable by a regular expression. According to Kleene's theorem, a language is regular if and only if it is *recognizable*, that is, recognized by some finite automaton.

There is an equivalent definition in terms of monoids. A language L of A^* is *recognized by a monoid morphism* $\varphi : A^* \to M$ if there exists a subset P of the monoid M such that $L = \varphi^{-1}(P)$. By extension, L is said to be *recognized by a monoid* M if there exists a monoid morphism $\varphi : A^* \to M$ that recognizes L. The equivalence mentioned above can now be stated as follows: a language is recognizable if and only if it is recognized by a finite monoid (see for instance [11, p. 15]).

Let L be a language of A^* and let u be a word of A^*. Then the language

$$u^{-1}L = \{v \mid uv \in L\}$$

is the *left quotient of* L *by* u.

The *Nerode automaton* of L is the deterministic automaton $\mathcal{A}(L) = (Q, A, \cdot, L, F)$ where $Q = \{u^{-1}L \mid u \in A^*\}$, $F = \{u^{-1}L \mid u \in L\}$ and the transition function is defined, for each $a \in A$, by the formula

$$(u^{-1}L)\cdot a = a^{-1}(u^{-1}L) = (ua)^{-1}L.$$

Each state of $\mathcal{A}(L)$ is a left quotient of L by a word, and hence is a language of A^*. The initial state is the language L, and the set of final states is the set of all left quotients of L by a word of L.

Proposition 2.1. *A language L is recognizable if and only if the set $\{u^{-1}L \mid u \in A^*\}$ is finite. In this case, L is recognized by its Nerode automaton.*

2.3 Syntactic monoids

Let L be a language and let x and y be words. The *quotient* $x^{-1}Ly^{-1}$ of L by x and y is defined by the formula

$$x^{-1}Ly^{-1} = \{u \in A^* \mid xuy \in L\}$$

The *syntactic monoid* of a language L of A^* is the monoid obtained as the quotient of A^* by the *syntactic congruence* of L, defined on A^* as follows: $u \sim_L v$ if and only if, for every $x, y \in A^*$,

$$xvy \in L \iff xuy \in L$$

The natural morphism $\eta : A^* \to A^*/{\sim_L}$ is the *syntactic morphism* of L. The syntactic monoid is the smallest monoid recognizing a language. In particular, a language is regular if and only if its syntactic monoid is finite.

A *class* of regular languages $\mathcal{C}$ associates with each finite alphabet A a set $\mathcal{C}(A^*)$ of regular languages of A^*. It is *closed under quotients* if for each language $L \in \mathcal{C}(A^*)$ and for each pair of words (x, y) of A^*, the language $x^{-1}Ly^{-1}$ belongs to $\mathcal{C}$.

2.4 The Formation Theorem

Just as formations of finite monoids extend the notion of a variety of finite monoids, formations of languages are more general than varieties of languages. Like varieties, formations are classes of regular languages closed under Boolean operations and quotients. But while varieties are closed under inverse of morphisms, formations of languages only enjoy a weak version of this property — Property (F_2) below — and thus comprise more general classes of languages than varieties.

The following definition was first given in [2]. A *formation of languages* is a class of regular languages $\mathcal{F}$ satisfying the following conditions:

(F_1) for each alphabet A, $\mathcal{F}(A^*)$ is closed under Boolean operations and quotients,

(F_2) if L is a language of $\mathcal{F}(B^*)$ and $\eta : B^* \to M$ denotes its syntactic morphism, then for each monoid morphism $\alpha : A^* \to B^*$ such that $\eta \circ \alpha$ is surjective, the language $\alpha^{-1}(L)$ belongs to $\mathcal{F}(A^*)$.

Observe that a formation of languages is closed under inverse of surjective morphisms, but this condition is not equivalent to (F_2).

To each formation of monoids $\mathbf{F}$, let us associate the class of languages $\mathcal{F}(\mathbf{F})$ defined as follows: for each alphabet A, $\mathcal{F}(\mathbf{F})(A^*)$ is the set of languages of A^* whose syntactic monoid belongs to $\mathbf{F}$.

Given a formation of languages $\mathcal{F}$, let $\mathbf{F}(\mathcal{F})$ denote the formation of monoids generated by the syntactic monoids of the languages of $\mathcal{F}$. The following statement is the main result of [2].

Theorem 2.2 (Formation Theorem)**.** *The correspondences* $\mathbf{F} \to \mathcal{F}(\mathbf{F})$ *and* $\mathcal{F} \to \mathbf{F}(\mathcal{F})$ *are two mutually inverse, order preserving, bijections between formations of monoids and formations of languages.*

3 Binomial coefficients on words

Binomial coefficients on words were first defined in [6, p. 238]. Useful references include [9, Chapter 6] and [12].

3.1 Definition of binomial coefficients on words

A word $u = a_1 a_2 \cdots a_n$ (where $a_1, \ldots, a_n$ are letters) is a *subword* of a word v if v can be factored as $v = v_0 a_1 v_1 \cdots a_n v_n$. For instance, ab is a subword of $cacbc$. Given two words u and v, we denote by $\binom{v}{u}$ the number of distinct ways to write u as a subword of v.

More formally, if $u = a_1 a_2 \cdots a_n$, then

$$\binom{v}{u} = \mathrm{Card}\{(v_0, v_1, \ldots, v_n) \mid v_0 a_1 v_1 \cdots a_n v_n = v\}$$

Observe that if u is a letter a, then $\binom{v}{a}$ is simply the number of occurrences of the letter a in v, also denoted by $|v|_a$. These binomial coefficients satisfy the following recursive formula, where $u, v \in A^*$ and $a, b \in A$:

$$\begin{cases} \binom{u}{1} = 1 \\ \binom{1}{u} = 0 \text{ if } u \neq 1 \\ \binom{va}{ub} = \begin{cases} \binom{v}{ub} & \text{if } a \neq b \\ \binom{v}{ub} + \binom{v}{u} & \text{if } a = b \end{cases} \end{cases} \tag{3.1}$$

An alternative definition of the binomial coefficients is given below in Formula (3.3). We shall later use the following elementary result.

Proposition 3.1. *Let $u \in \{a,b\}^*$. Then the following formula holds*

$$\binom{u}{a}\binom{u}{b} + \binom{u}{ab} + \binom{u}{ba} \equiv 0 \bmod 2 \tag{3.2}$$

Proof. Let us prove (3.2) by induction on $|u|$. The result is trivial if $|u| = 0$. For the induction step, it suffices to prove the result for ua, the case ub being symmetrical.

$$\begin{aligned}\binom{ua}{a}\binom{ua}{b} + \binom{ua}{ab} + \binom{ua}{ba} &= \left(\binom{u}{a} + 1\right)\binom{u}{b} + \binom{u}{ab} + \binom{u}{ba} + \binom{u}{b} \\ &\equiv \binom{u}{a}\binom{u}{b} + \binom{u}{ab} + \binom{u}{ba} \equiv 0 \bmod 2.\end{aligned}$$ □

3.2 Binomial coefficients and morphisms

Let $\mathbb{Z}\langle A\rangle$ be the ring of noncommutative polynomials with coefficients in $\mathbb{Z}$ and variables in A (see [9, Chapter 6] or [12]). Given a polynomial $P \in \mathbb{Z}\langle A\rangle$ and a word x, we let $\langle P, x\rangle$ denote the coefficient of x in P. Thus all but a finite number of these coefficients are null and $P = \sum_{x\in A^*} \langle P, x\rangle x$.

In this section, we study the behaviour of binomial coefficients under monoid morphisms. More precisely, given a monoid morphism $\varphi : A^* \to B^*$ and words $u \in A^*$ and $x \in B^*$, we give a formula to compute $\binom{\varphi(u)}{x}$.

The proof of this result relies on properties of the *Magnus automorphism* of the ring $\mathbb{Z}\langle A\rangle$. This automorphism μ_A is defined, for each letter $a \in A$, by $\mu_A(a) = 1 + a$. Its inverse is defined by $\mu_A^{-1}(a) = a - 1$. The following *binomial identity* [9, Formula 6.3.4]

$$\text{for all } u \in A^*, \quad \mu_A(u) = \sum_{x\in A^*} \binom{u}{x} x \tag{3.3}$$

can be used to give an alternative definition of the binomial coefficients.

If $\varphi : A^* \to B^*$ is a monoid morphism, then φ can be extended by linearity to a ring morphism from $\mathbb{Z}\langle A\rangle$ to $\mathbb{Z}\langle B\rangle$. Let $\gamma : \mathbb{Z}\langle A\rangle \to \mathbb{Z}\langle B\rangle$ be the ring morphism defined by $\gamma = \mu_B \circ \varphi \circ \mu_A^{-1}$.

We are now ready to present the announced formula:

Proposition 3.2. *Let $\varphi : A^* \to B^*$ be a morphism and let $u \in A^*$ and $x \in B^*$. Then*

$$\binom{\varphi(u)}{x} = \sum_{|s|\leqslant|x|} \binom{u}{s} \langle \gamma(s), x\rangle \tag{3.4}$$

Proof. Observing that $\mu_A^{-1}(a) = a - 1$ for each letter $a \in A$, one gets

$$\gamma(a) = \mu_B(\varphi(a) - 1) = \mu_B(\varphi(a)) - 1 = \Big(\sum_{x\in B^*} \binom{\varphi(a)}{x} x\Big) - 1 = \sum_{x\in B^+} \binom{\varphi(a)}{x} x$$

and thus $\langle \gamma(a), 1\rangle = 0$. It follows that $\langle \gamma(s), x\rangle = 0$ if $|x| < |s|$. Furthermore, for each $u \in A^*$, one gets on the one hand from (3.3)

$$\mu_B(\varphi(u)) = \sum_{x \in B^*} \binom{\varphi(u)}{x} x$$

and on the other hand, using (3.3),

$$\gamma(\mu_A(u)) = \gamma\Big(\sum_{s \in A^*} \binom{u}{s} s\Big) = \sum_{s \in A^*} \binom{u}{s} \gamma(s) = \sum_{s \in A^*} \sum_{x \in B^*} \binom{u}{s} \langle \gamma(s), x\rangle x$$

Now since $\gamma \circ \mu_A = \mu_B \circ \varphi$, the polynomials $\mu_B(\varphi(u))$ and $\gamma(\mu_A(u))$ have the same coefficients, which gives (3.4). $\square$

Example 3.1. To illustrate the use of (3.4), let us show how to compute $\binom{\varphi(u)}{ab}$. Let $A = \{a, b, c\}$, $B = \{a, b\}$ and let $\varphi : A^* \to B^*$ be the morphism defined by $\varphi(a) = a$, $\varphi(b) = ab$ and $\varphi(c) = a^2b$. First, $\gamma = \mu_B \circ \varphi \circ \mu_A^{-1}$ is defined as follows:

$$\begin{aligned}
\gamma(a) &= \mu_B(\varphi(a-1)) = \mu_B(a-1) = \mu_B(a) - \mu_B(1) = (a+1) - 1 = a \\
\gamma(b) &= \mu_B(\varphi(b-1)) = \mu_B(ab-1) = (1+a)(1+b) - 1 = a + b + ab \\
\gamma(c) &= \mu_B(\varphi(c-1)) = \mu_B(a^2b-1) = \mu_B(a^2b) - 1 \\
&= (1+a)(1+a)(1+b) - 1 = 2a + aa + b + 2ab + aab
\end{aligned}$$

Thus we get by (3.4)

$$\binom{\varphi(u)}{ab} = \sum_{s \in A^*} \binom{u}{s} \langle \gamma(s), ab\rangle = \sum_{|s| \leqslant 2} \binom{u}{s} \langle \gamma(s), ab\rangle$$

We now need to compute the coefficients $\langle \gamma(s), ab\rangle$ for $|s| \leqslant 2$. The non-zero coefficients are the following:

$$\begin{array}{cccc}
\langle \gamma(b), ab\rangle = 1 & \langle \gamma(c), ab\rangle = 2 & \langle \gamma(ab), ab\rangle = 1 & \langle \gamma(ac), ab\rangle = 1 \\
\langle \gamma(bb), ab\rangle = 1 & \langle \gamma(bc), ab\rangle = 1 & \langle \gamma(cb), ab\rangle = 2 & \langle \gamma(cc), ab\rangle = 2
\end{array}$$

and finally

$$\binom{\varphi(u)}{ab} = \binom{u}{b} + 2\binom{u}{c} + \binom{u}{ab} + \binom{u}{ac} + \binom{u}{bb} + \binom{u}{bc} + 2\binom{u}{cb} + 2\binom{u}{cc}.$$

4 Languages recognized by p-groups

Let p be a prime number. A *p-group* is a group whose order is a power of p. A *p-group language* is a language whose syntactic monoid is a p-group.

4.1 Two descriptions of the p-group languages

The following result is credited to Eilenberg and Schützenberger in [6].

Proposition 4.1. *A language of A^* is a p-group language if and only if it is a Boolean combination of languages of the form*

$$L(x, r, p) = \{u \in A^* \mid \binom{u}{x} \equiv r \bmod p\}, \tag{4.5}$$

where $0 \leqslant r < p$ and $x \in A^*$.

We now give another characterization. A function $f : A^* \to \mathbb{Z}$ is said to be a *linear combination of binomial coefficients* if there exist $c_1, \ldots, c_n \in \mathbb{Z}$ and $x_1, \ldots, x_n \in A^*$ such that, for all $u \in A^*$,

$$f(u) = c_1\binom{u}{x_1} + \cdots + c_n\binom{u}{x_n} \tag{4.6}$$

Since the function $f(u) = c\binom{u}{1}$ maps every word to the constant c, every constant function is a linear combination of binomial coefficients.

Proposition 4.2. *A language of A^* is a p-group language if and only if it is a finite union of languages of the form*

$$L(f_1, \ldots, f_r, p) = \{u \in A^* \mid f_1(u) \equiv \cdots \equiv f_r(u) \equiv 0 \bmod p\} \tag{4.7}$$

where $f_1, \ldots, f_r$ are linear combinations of binomial coefficients.

Proof. Let $\mathcal{G}_p$ be the Boolean algebra generated by the languages of the form $L(x, r, p)$ and let $\mathcal{S}_p$ be the set of languages that are finite unions of languages of the form $L(f_1, \ldots, f_r, p)$.

Step 1. $\mathcal{S}_p$ is a Boolean algebra. First, $\mathcal{S}_p$ is closed under union by definition. It is also closed under intersection since

$$L(f_1, \ldots, f_r, p) \cap L(g_1, \ldots, g_s, p) = L(f_1, \ldots, f_r, g_1, \ldots, g_s, p). \tag{4.8}$$

In particular,

$$L(f_1, \ldots, f_r, p) = L(f_1, p) \cap \cdots \cap L(f_r, p). \tag{4.9}$$

It remains to show that $\mathcal{S}_p$ is closed under complementation. Since $\mathcal{S}_p$ is closed under union and intersection, it suffices to prove that the complement of each language of the form $L(f, p)$, where f is a linear combination of binomial coefficients, belongs $\mathcal{S}_p$. Now

$$\begin{aligned}
L(f, p)^c &= \{u \in A^* \mid f(u) \not\equiv 0 \bmod p\} \\
&= \bigcup_{c \in \mathbb{F}_p \setminus \{0\}} \{u \in A^* \mid f(u) \equiv c \bmod p\} \\
&= \bigcup_{c \in \mathbb{F}_p \setminus \{0\}} \{u \in A^* \mid (f - c)(u) \equiv 0 \bmod p\}
\end{aligned}$$

It remains to observe that $f-c$ is a linear combination of binomial coefficients to conclude.

Step 2: $\mathcal{S}_p \subseteq \mathcal{G}_p$. It suffices to show that every language of the form $L(f,p)$ belongs to $\mathcal{G}_p$. Now if f is given by (4.6), one gets

$$L(f,p) = \bigcup_{\{(r_1,\ldots,r_n) \mid c_1r_1+\cdots+c_nr_n \equiv 0 \bmod p\}} \big(L(x_1,r_1,p) \cap \cdots \cap L(x_n,r_n,p)\big) \tag{4.10}$$

and thus $L(f,p) \in \mathcal{G}_p$ as required. Thus $\mathcal{S}_p \subseteq \mathcal{G}_p$.

Step 3: $\mathcal{G}_p \subseteq \mathcal{S}_p$. This immediately follows from the formula

$$L(x,r,p) = L(f,p) \text{ where } f(u) = -r\binom{u}{1} + \binom{u}{x}.$$

Thus $\mathcal{G}_p = \mathcal{S}_p$ and it now suffices to apply Proposition 4.1 to conclude the proof. □

As explained in Section 2.2, one can compute the minimal automaton of a language of the form $L(f_1,\ldots,f_r,p)$ by computing its derivatives as follows:

$$u^{-1}L = \{x \in A^* \mid f_1(ux) = f_2(ux) = \cdots = f_n(ux) \equiv 0 \bmod p\}\,.$$

4.2 An algorithm for p-group languages

Let p be a prime number and let $U_n(\mathbb{F}_p)$ be the group of unitriangular[1] $n\times n$-matrices with coefficients in $\mathbb{F}_p$, the finite field of order p. Then $U_n(\mathbb{F}_p)$ is a p-group and it is a well-known fact that every p-group is isomorphic to a subgroup of some $U_n(\mathbb{F}_p)$, for a suitable choice of n. See for instance [13, p. 276, Corollary 5.48].

Let $\pi : A \to U_{n+1}(\mathbb{F}_p)$ be a map[2] and let G be the subgroup of $U_{n+1}(\mathbb{F}_p)$ generated by $\pi(A)$. Then π extends to a surjective monoid morphism $\pi :\ A^* \to G$ which maps every word $a_1 \cdots a_k \in A^*$ to the matrix $\pi(a_1)\cdots\pi(a_k)$. For $1 \leqslant i < j \leqslant n+1$, we let $\pi_{i,j} : A^* \to \mathbb{F}_p$ be the map defined, for all $u \in A^*$, by

$$\pi_{i,j}(u) = (\pi(u))_{i,j} \tag{4.11}$$

By definition, a language K is recognized by π if there exists a subset S of G such that $K = \pi^{-1}(S)$. According to Proposition 4.2, K is a finite union of languages of the form $L(f_1,\ldots,f_r,p)$. We now give an algorithm to obtain this representation explicitly.

[1] An $n \times n$-matrix is *unitriangular* if its diagonal coefficients are all equal to 1 and all its coefficients below the diagonal are equal to 0.

[2] The switch from n to $n+1$ will be justified later on.

Setting, for each $s \in S$, $K_s = \pi^{-1}(s)$, one gets

$$K = \bigcup_{s \in S} K_s \quad \text{and}$$
$$K_s = \{\, u \in A^* \mid \text{for } 1 \leqslant \ i < j \leqslant n+1,\ \pi_{i,j}(u) = s_{i,j} \,\}$$

It just remains to verify that the languages K_s are of the form $L(f_1, \ldots, f_r, p)$. But this follows immediately from the following result:

Proposition 4.3. *Each function $\pi_{i,j}$ is a linear combination of binomial coefficients.*

Proof. Let $\theta : A \to U_{n+1}(\mathbb{F}_p)$ be the map defined by $\theta(a) = \pi(a) - 1$ for all $a \in A$. Then θ extends to a ring morphism $\theta : \mathbb{Z}\langle A\rangle \to U_{n+1}(\mathbb{F}_p)$ and for $1 \leqslant i < j \leqslant n+1$, the maps $\theta_{i,j} : A^* \to \mathbb{F}_p$ are defined as in (4.11). Since $\theta(a)$ is a strictly triangular matrix for all $a \in A$, it follows that $\theta(x) = 0$ for all words x of length $> n$. Note however that $\theta(x)$ is not in general equal to $\pi(x) - 1$.

Let also $\mu : A^* \to \mathbb{Z}\langle A\rangle$ be the monoid morphism defined by $\mu(a) = 1+a$ for all $a \in A$. Thus μ is the restriction to A^* of the Magnus automorphism introduced in Section 3.2. Since the formula $\theta(\mu(a)) = \theta(1+a) = 1+\theta(a) = \pi(a)$ holds for all $a \in A$, one has $\pi = \theta \circ \mu$.

$$\begin{tikzcd} A^* \arrow[r, "\mu"] \arrow[rr, bend left, "\pi"] & \mathbb{Z}\langle A\rangle \arrow[r, "\theta"] & U_{n+1}(\mathbb{F}_p) \end{tikzcd}$$

It follows by (3.3) that

$$\pi(u) = \theta(\mu(u)) = \theta\Big(\sum_{x \in A^*} \binom{u}{x} x\Big) = \sum_{x \in A^*} \binom{u}{x}\theta(x) = \sum_{|x| \leqslant n} \binom{u}{x}\theta(x)$$

and hence

$$\pi_{i,j}(u) = \sum_{|x| \leqslant n} \theta_{i,j}(x)\binom{u}{x} \tag{4.12}$$

which shows that $\pi_{i,j}$ is a linear combination of binomial coefficients. $\square$

An interesting special case occurs if the language is defined by constraints on the first row of the matrix, for instance for a language of the form

$$L = \{\, u \in A^* \mid \pi_{1,2}(u) = \ \cdots \ = \pi_{1,n}(u) = 0 \,\}$$

Observing that L can also be written as

$$L = \{\, u \in A^* \mid (1, 0, \ldots, 0)\pi(u) = (1, 0, \ldots, 0) \,\}$$

one can directly obtain a deterministic automaton for L by taking $\mathbb{F}_p^n$ as set of states, the state $(0,\dots,0)$ as initial and unique final state and by defining the transitions, for each $(z_1,\dots,z_n) \in \mathbb{F}_p^n$ and each letter a, by setting

$$\begin{aligned}(z_1,\dots,z_n)\cdot a &= (z'_1,\dots,z'_n),\\ &\qquad \text{where } (1,z_1,\dots,z_n)\pi(a) = (1,z'_1,\dots,z'_n),\end{aligned} \tag{4.13}$$

that is,

$$\begin{aligned}z'_1 &= \pi_{1,2}(a) + z_1,\\ z'_2 &= \pi_{1,3}(a) + \pi_{2,3}(a)z_1 + z_2,\\ z'_3 &= \pi_{1,4}(a) + \pi_{2,4}(a)z_1 + \pi_{3,4}(a)z_2 + z_3, \text{ etc.}\end{aligned}$$

This algorithm is illustrated by the examples presented in Section 4.3.

4.3 Three examples

These examples will be used in Section 5. The languages of the first two examples were also considered by Thérien [19].

Example 4.1. The subgroup of $U_3(\mathbb{F}_2)$ generated by the two matrices

$$a = \begin{pmatrix}1&1&0\\0&1&0\\0&0&1\end{pmatrix} \quad\text{and}\quad b = \begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}$$

is isomorphic to D_4. A confluent rewriting system for this group is $a^2 \to 1$, $b^2 \to 1$ and $baba \to abab$. The group consists of the matrices

$$1 = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} \quad a = \begin{pmatrix}1&1&0\\0&1&0\\0&0&1\end{pmatrix} \quad b = \begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix} \quad ab = \begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}$$
$$ba = \begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix} \quad aba = \begin{pmatrix}1&0&1\\0&1&1\\0&0&1\end{pmatrix} \quad bab = \begin{pmatrix}1&1&1\\0&1&0\\0&0&1\end{pmatrix} \quad abab = \begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}$$

Let $\pi : A^* \to D_4$ be the natural morphism and let

$$L_1 = \{u \in A^* \mid \pi_{1,2}(u) = \pi_{1,3}(u) = 0\}.$$

To obtain a deterministic automaton for L_1, we take $\mathbb{F}_2^2$ as the set of states and define the transitions, for all $(z_1,z_2) \in \mathbb{F}_2^2$, by setting

$$\begin{cases}(z_1,z_2)\cdot a = (1+z_1,z_2) & (4.14)\\ (z_1,z_2)\cdot b = (z_1,z_1+z_2) & (4.15)\end{cases}$$

The resulting automaton, which turns out to be minimal, is the following:

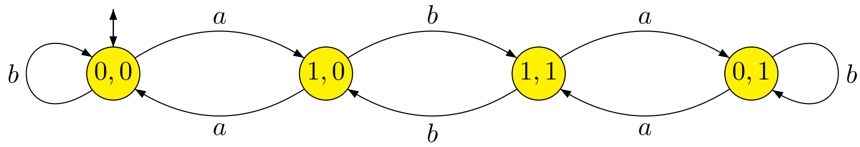

Figure 4.1: The minimal automaton of L_1.

The syntactic monoid of L_1 is the group D_4 presented by the relations $a^2 = 1$, $b^2 = 1$ and $(ba)^2 = (ab)^2$. Its syntactic image is $\{1, b\}$.

	1	2	3	4
$*\ 1$	1	2	3	4
a	2	1	4	3
b	1	3	2	4
ab	3	1	4	2

	1	2	3	4
ba	2	4	1	3
aba	4	2	3	1
bab	3	4	1	2
$abab$	4	3	2	1

Applying (4.12) with $n = 2$ one gets

$$\begin{aligned}
\pi_{1,2}(u) &= \sum_{|x| \leqslant 2} \binom{u}{x} \theta_{1,2}(x) = \binom{u}{1}\theta_{1,2}(1) + \binom{u}{a}\theta_{1,2}(a) + \binom{u}{b}\theta_{1,2}(b) \\
&\quad + \binom{u}{aa}\theta_{1,2}(aa) + \binom{u}{ab}\theta_{1,2}(ab) + \binom{u}{ba}\theta_{1,2}(ba) + \binom{u}{bb}\theta_{1,2}(bb) \\
&= \binom{u}{a} \\
\pi_{1,3}(u) &= \sum_{|x| \leqslant 2} \binom{u}{x} x_{1,3} = \binom{u}{1}\theta_{1,3}(1) + \binom{u}{a}\theta_{1,3}(a) + \binom{u}{b}\theta_{1,3}(b) \\
&\quad + \binom{u}{aa}\theta_{1,3}(aa) + \binom{u}{ab}\theta_{1,3}(ab) + \binom{u}{ba}\theta_{1,3}(ba) + \binom{u}{bb}\theta_{1,3}(bb) \\
&= \binom{u}{ab}
\end{aligned}$$

It follows that

$$L_1 = \left\{ u \in \{a, b\}^* \mid \binom{u}{a} \equiv \binom{u}{ab} \equiv 0 \bmod 2 \right\} \tag{4.16}$$

Moreover, for all $u \in \{a, b\}^*$,

$$(0, 0)\cdot u = \left(\binom{u}{a}, \binom{u}{ab} \right)$$

where the binomial coefficients are computed modulo 2. Thus the states of the minimal automaton of L_1 encode the possible values modulo 2 of these two binomial coefficients. Now, one can recover (4.14) and (4.15) by observing that, if

$$(0, 0)\cdot u = (z_1, z_2) = \left(\binom{u}{a}, \binom{u}{ab} \right)$$

then

$$(0,0)\cdot ua = (z_1,z_2)\cdot a = \left(\binom{ua}{a},\binom{ua}{ab}\right) = \left(\binom{u}{a}+1,\binom{u}{ab}\right)$$
$$= (z_1+1, z_2)$$

and

$$(0,0)\cdot ub = (z_1,z_2)\cdot b = \left(\binom{ub}{a},\binom{ub}{ab}\right) = \left(\binom{u}{a},\binom{u}{ab}+\binom{u}{a}\right)$$
$$= (z_1, z_1+z_2)\,.$$

Example 4.2. The group D_4 is also generated by the two matrices

$$a = \begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix} \quad\text{and}\quad b = \begin{pmatrix}1&1&1\\0&1&0\\0&0&1\end{pmatrix}$$

A confluent rewriting system for this group is $b^2 \to 1$, $aba \to b$, $ba^2 \to a^2b$, $bab \to a^3$, $a^4 \to 1$ and $a^3b \to ba$. The group consists of the matrices

$$1 = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} \quad a = \begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix} \quad b = \begin{pmatrix}1&1&1\\0&1&0\\0&0&1\end{pmatrix} \quad a^2 = \begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}$$
$$ab = \begin{pmatrix}1&0&1\\0&1&1\\0&0&1\end{pmatrix} \quad ba = \begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix} \quad a^3 = \begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix} \quad a^2b = \begin{pmatrix}1&1&0\\0&1&0\\0&0&1\end{pmatrix}$$

Let $\pi : A^* \to D_4$ be the natural morphism and let

$$L_2 = \{u \in A^* \mid \pi_{1,2}(u) = \pi_{1,3}(u) = 0\}.$$

To obtain a deterministic automaton for L_2, we take $\mathbb{F}_2^2$ as the set of states and define the transitions, for all $(z_1, z_2) \in \mathbb{F}_2^2$, by setting

$$\begin{cases}(z_1,z_2)\cdot a = (1+z_1, z_1+z_2) & (4.17)\\ (z_1,z_2)\cdot b = (1+z_1, 1+z_2) & (4.18)\end{cases}$$

The resulting automaton, which turns out to be minimal, is the following:

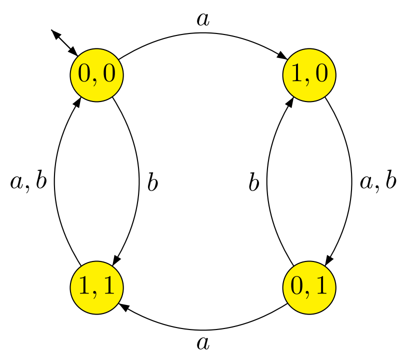

Figure 4.2: The minimal automaton of L_2.

Applying (4.12) with $n = 2$ one gets[3]

$$\begin{aligned}
\pi_{1,2}(u) &= \sum_{|x|\leqslant 2} \binom{u}{x}\theta_{1,2}(x) = \binom{u}{1}\theta_{1,2}(1) + \binom{u}{a}\theta_{1,2}(a) + \binom{u}{b}\theta_{1,2}(b) \\
&\quad + \binom{u}{aa}\theta_{1,2}(aa) + \binom{u}{ab}\theta_{1,2}(ab) + \binom{u}{ba}\theta_{1,2}(ba) + \binom{u}{bb}\theta_{1,2}(bb) \\
&= \binom{u}{a} + \binom{u}{b} \\
\pi_{1,3}(u) &= \sum_{|x|\leqslant 2} \binom{u}{x}\theta_{1,3}(x) = \binom{u}{1}\theta_{1,3}(1) + \binom{u}{a}\theta_{1,3}(a) + \binom{u}{b}\theta_{1,3}(b) \\
&\quad + \binom{u}{aa}\theta_{1,3}(aa) + \binom{u}{ab}\theta_{1,3}(ab) + \binom{u}{ba}\theta_{1,3}(ba) + \binom{u}{bb}\theta_{1,3}(bb) \\
&= \binom{u}{b} + \binom{u}{aa} + \binom{u}{ba}
\end{aligned}$$

It follows that

$$L_2 = \Big\{ u \in \{a,b\}^* \mid \binom{u}{a} + \binom{u}{b} \equiv \binom{u}{aa} + \binom{u}{ab} \equiv 0 \bmod 2 \Big\} \tag{4.19}$$

Moreover, for all $u \in \{a,b\}^*$,

$$(0,0)\cdot u = \left(\binom{u}{a} + \binom{u}{b}, \binom{u}{b} + \binom{u}{aa} + \binom{u}{ba} \right)$$

where the binomial coefficients are computed modulo 2. Thus the states of the minimal automaton of L_1 encode the possible values modulo 2 of these two linear combinations of binomial coefficients. Now, one can recover (4.17) and (4.18) by observing that, if

$$(0,0)\cdot u = (z_1, z_2) = \left(\binom{u}{a} + \binom{u}{b}, \binom{u}{b} + \binom{u}{aa} + \binom{u}{ba} \right)$$

then

$$\begin{aligned}
(0,0)\cdot ua = (z_1, z_2)\cdot a &= \left(\binom{ua}{a} + \binom{ua}{b}, \binom{ua}{b} + \binom{ua}{aa} + \binom{ua}{ba} \right) \\
&= \left(\binom{u}{a} + 1 + \binom{u}{b}, \binom{u}{b} + \binom{u}{aa} + \binom{u}{a} + \binom{u}{ba} + \binom{u}{b} \right) \\
&= (z_1 + 1, z_1 + z_2)
\end{aligned}$$

and

$$\begin{aligned}
(0,0)\cdot ub = (z_1, z_2)\cdot b &= \left(\binom{ub}{b} + \binom{ub}{a}, \binom{ub}{b} + \binom{ub}{aa} + \binom{ub}{ba} \right) \\
&= \left(\binom{u}{a} + \binom{u}{b} + 1, \binom{u}{b} + 1 + \binom{u}{aa} + \binom{u}{ba} \right) \\
&= (z_1 + 1, z_2 + 1)\,.
\end{aligned}$$

[3]It is easy to make mistakes in this computation. Recall that in general $\theta(x) \neq \pi(x) - 1$. Thus for instance $\theta(ba) = \theta(b)\theta(a) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $\pi(ba) - 1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, whence $\theta_{1,3}(ba) = 1$.

The syntactic monoid of L_2 is the group D_4, but this time presented by the group relations $b^2 = 1$, $a^4 = 1$ and $a^3b = ba$. Its syntactic image is $\{1, ba\}$.

	1	2	3	4
$*\ 1$	1	2	3	4
a	2	3	4	1
b	4	3	2	1
a^2	3	4	1	2

	1	2	3	4
ab	3	2	1	4
ba	1	4	3	2
a^3	4	1	2	3
a^2b	2	1	4	3

Example 4.3. The subgroup of $U_4(\mathbb{F}_2)$ generated by the two matrices

$$a = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is isomorphic to Q_8. A confluent rewriting system for this group is $b^2 \to a^2$, $aba \to b$, $ba^2 \to a^2b$, $bab \to a$, $a^4 \to 1$ and $a^3b \to ba$. The group consists of the matrices of the following form, where $\varepsilon_1, \varepsilon_2, \varepsilon_3 \in \mathbb{F}_2$.

$$\begin{pmatrix} 1 & \varepsilon_1 & \varepsilon_2 & \varepsilon_3 \\ 0 & 1 & 0 & \varepsilon_1 + \varepsilon_2 \\ 0 & 0 & 1 & \varepsilon_2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Let $\pi : A^* \to Q_8$ be the natural morphism and let

$$L_3 = \{u \in A^* \mid \pi_{1,2}(u) = \pi_{1,3}(u) = \pi_{1,4}(u) = 0\}.$$

To obtain a deterministic automaton for L_2, we take $\mathbb{F}_2^3$ as the set of states and define the transitions, for all $(z_1, z_2, z_3) \in \mathbb{F}_2^3$, by setting

$$\begin{cases} (z_1, z_2, z_3) \cdot a = (z_1 + 1, z_2, z_1 + z_3) & (4.20) \\ (z_1, z_2, z_3) \cdot b = (z_1, z_2 + 1, z_1 + z_2 + z_3) & (4.21) \end{cases}$$

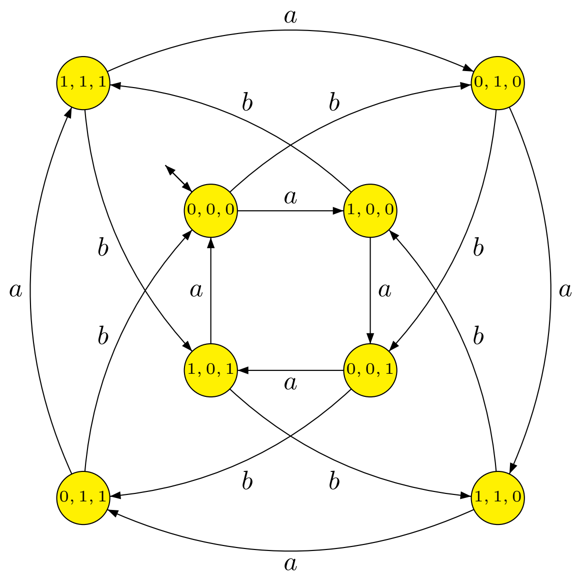

Figure 4.3: The minimal automaton of L_3.

Applying (4.12) with $n = 3$ one gets

$$\begin{aligned}
\pi_{1,2}(u) &= \sum_{|x| \leqslant 3} \binom{u}{x}\theta_{1,2}(x) = \binom{u}{1}\theta_{1,2}(1) + \binom{u}{a}\theta_{1,2}(a) + \binom{u}{b}\theta_{1,2}(b) \\
&\quad + \binom{u}{aa}\theta_{1,2}(aa) + \binom{u}{ab}\theta_{1,2}(ab) + \binom{u}{\theta}(ba)_{1,2}ba + \binom{u}{bb}\theta_{1,2}(bb) \\
&= \binom{u}{a} \\
\pi_{1,3}(u) &= \sum_{|x| \leqslant 3} \binom{u}{x}\theta_{1,3}(x) = \binom{u}{1}\theta_{1,3}(1) + \binom{u}{a}\theta_{1,3}(a) + \binom{u}{b}\theta_{1,3}(b) \\
&\quad + \binom{u}{aa}\theta_{1,3}(aa) + \binom{u}{ab}\theta_{1,3}(ab) + \binom{u}{ba}\theta_{1,3}(ba) + \binom{u}{bb}\theta_{1,3}(bb) \\
&= \binom{u}{b} \\
\pi_{1,4}(u) &= \sum_{|x| \leqslant 3} \binom{u}{x}\theta_{1,4}(x) = \binom{u}{1}\theta_{1,4}(1) + \binom{u}{a}\theta_{1,4}(a) + \binom{u}{b}\theta_{1,4}(b) \\
&\quad + \binom{u}{aa}\theta_{1,4}(aa) + \binom{u}{ab}\theta_{1,4}(ab) + \binom{u}{ba}\theta_{1,4}(ba) + \binom{u}{bb}\theta_{1,4}(bb) \\
&= \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb}
\end{aligned}$$

It follows that

$$L_3 = \left\{ u \in \{a, b\}^* \mid \binom{u}{a} \equiv \binom{u}{b} \equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb} \equiv 0 \bmod 2 \right\} \quad (4.22)$$

Moreover, for all $u \in \{a,b\}^*$,

$$(0,0,0)\cdot u = \left(\binom{u}{a}, \binom{u}{b}, \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb}\right)$$

where the binomial coefficients are computed modulo 2. Thus the states of the minimal automaton of L_1 encode the possible values modulo 2 of these two linear combinations of binomial coefficients. Now, one can recover (4.20) and (4.21) by observing that, if

$$(0,0,0)\cdot u = (z_1, z_2, z_3) = \left(\binom{u}{a}, \binom{u}{b}, \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb}\right)$$

then

$$\begin{aligned}(0,0,0)\cdot ua = (z_1, z_2, z_3)\cdot a &= \left(\binom{ua}{a}, \binom{ua}{b}, \binom{ua}{aa} + \binom{ua}{ab} + \binom{ua}{bb}\right)\\ &= \left(\binom{u}{a} + 1, \binom{u}{b}, \binom{u}{aa} + \binom{u}{a} + \binom{u}{ab} + \binom{u}{bb}\right)\\ &= (z_1 + 1, z_2, z_1 + z_3)\end{aligned}$$

and

$$\begin{aligned}(0,0,0)\cdot ub = (z_1, z_2, z_3)\cdot b &= \left(\binom{ub}{a}, \binom{ub}{b}, \binom{ub}{aa} + \binom{ub}{ab} + \binom{ub}{bb}\right)\\ &= \left(\binom{u}{a}, \binom{u}{b} + 1, \binom{u}{aa} + \binom{u}{ab} + \binom{u}{a} + \binom{u}{bb} + \binom{u}{b}\right)\\ &= (z_1, z_2 + 1, z_1 + z_2 + z_3)\end{aligned}$$

The syntactic monoid of L_3 is the group Q_8 presented by the group relations $a^4 = 1$, $b^2 = a^2$ and $a^3b = ba$. Its syntactic image is $\{1\}$.

	1	2	3	4	5	6	7	8
$*\,1$	1	2	3	4	5	6	7	8
a	2	3	4	1	6	7	8	5
b	6	5	8	7	4	3	2	1
a^2	3	4	1	2	7	8	5	6

	1	2	3	4	5	6	7	8
ab	5	8	7	6	3	2	1	4
ba	7	6	5	8	1	4	3	2
a^3	4	1	2	3	8	5	6	7
a^2b	8	7	6	5	2	1	4	3

The Cayley graph of this group is represented in Figure 4.4. As one can see, this is exactly the same automaton as in Figure 4.3, up to the following renaming of the states:

$$\begin{array}{llll}(0,0,0) \leftrightarrow 1 & (1,0,0) \leftrightarrow a & (0,0,1) \leftrightarrow a^2 & (1,0,1) \leftrightarrow a^3\\ (0,1,0) \leftrightarrow b & (1,1,0) \leftrightarrow ba & (0,1,1) \leftrightarrow a^2b & (1,1,1) \leftrightarrow ab\end{array} \tag{4.23}$$

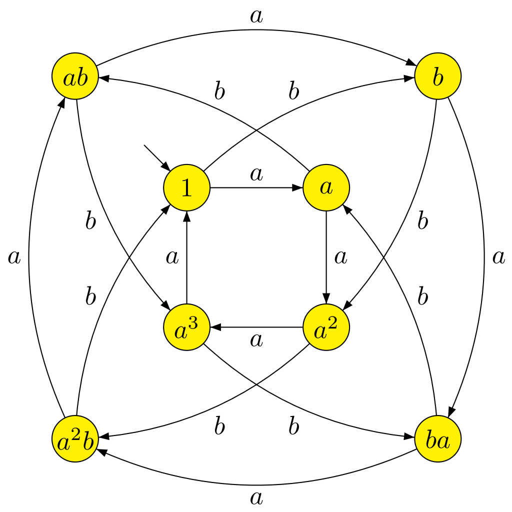

Figure 4.4: The Cayley graph of Q_8.

4.4 The varieties of languages $\mathcal{V}_{c,p}$

In this section, we revisit the congruences first introduced in [6, p. 240] and also studied in [19]. Let c be a nonnegative integer. For each alphabet A, let $\sim_{p,c}$ be the congruence on A^* defined by $u \sim_{p,c} v$ if and only if, for all words x such that $0 \leqslant |x| \leqslant c$,

$$\binom{u}{x} \equiv \binom{v}{x} \bmod p$$

This congruence has finite index and the languages which are saturated for this congruence form a Boolean algebra $\mathcal{V}_{c,p}(A^*)$, which is also the Boolean algebra generated by the languages $L(x, r, p)$ for $0 \leqslant r < p$ and $|x| \leqslant c$.

Let us first show that the class $\mathcal{V}_{c,p}$ is closed under inverses of morphisms. This result is due to Thérien [18], but for the convenience of the reader, we give a self-contain proof. This relies on the following result.

Proposition 4.4. *Let $\varphi : A^* \to B^*$ be a morphism. Let u and v be two words of A^* such that $u \sim_{p,c} v$. Then $\varphi(u) \sim_{p,c} \varphi(v)$.*

Proof. If $u \sim_{p,c} v$, one has, for $0 \leqslant |s| \leqslant c$, $\binom{u}{s} \equiv \binom{v}{s} \bmod p$. Therefore by (3.4) we obtain for $|x| \leqslant c$,

$$\binom{\varphi(u)}{x} - \binom{\varphi(v)}{x} = \sum_{|s| \leqslant |x|} \left(\binom{u}{s} - \binom{v}{s} \right) \langle \gamma(s), x \rangle \equiv 0 \bmod p$$

Thus $\varphi(u) \sim_{p,c} \varphi(v)$. $\square$

We can now state:

Proposition 4.5. *Let $\varphi : A^* \to B^*$ be a morphism and L a language of $\mathcal{V}_{c,p}(B^*)$. Then $\varphi^{-1}(L)$ belongs to $\mathcal{V}_{c,p}(A^*)$.*

Proof. Let L be a language of $\mathcal{V}_{c,p}(B^*)$. Let $u \in \varphi^{-1}(L)$ and let v be a word such that $u \sim_{p,c} v$. Then $\varphi(u) \sim_{p,c} \varphi(v)$ by Proposition 4.4, and since $u \in L$ and L is saturated by $\sim_{p,c}$, we get $\varphi(v) \in L$, that is, $v \in \varphi^{-1}(L)$. This proves that $\varphi^{-1}(L)$ is saturated by $\sim_{p,c}$ and therefore $\varphi^{-1}(L)$ belongs to $\mathcal{V}_{c,p}(A^*)$. $\square$

Proposition 4.6 ([18]). *For each c, the class $\mathcal{V}_{c,p}$ is a variety of languages.*

Proof. Proposition 4.5 shows that the class $\mathcal{V}_{c,p}$ is closed under inverses of morphisms. Furthermore, $\mathcal{V}_{c,p}(A^*)$ is by definition a Boolean algebra, generated by the languages of the form $L(x, r, p)$. We claim that it is closed under left quotient by a word u. Arguing on induction on the length of u, it suffices to consider the case where u is a letter a. Now, since left quotients commute with Boolean operations, it suffices to prove that any left quotient of the form $a^{-1}\big(L(x, r, p)\big)$ belongs to $\mathcal{V}_{c,p}(A^*)$. If x is the empty word, then $L(x, r, p)$ is either empty or equal to A^* and the result is trivial. Suppose that x is nonempty. Then, we get by (3.1):

$$\begin{aligned}
a^{-1}\big(L(x,r,p)\big) &= \left\{u \in A^* \mid \binom{au}{x} \equiv r \bmod p\right\} \\
&= \begin{cases} \{u \in A^* \mid \binom{u}{x} + \binom{u}{s} \equiv r \bmod p\} & \text{if } x = as \text{ for some } s \\ \{u \in A^* \mid \binom{u}{x} \equiv r \bmod p\} & \text{otherwise} \end{cases} \\
&= \begin{cases} \cup_{r_1+r_2 \equiv r \bmod p} \big(L(x, r_1, p) \cap L(s, r_2, p)\big) & \text{if } x = as \\ L(x, r, p) & \text{otherwise} \end{cases}
\end{aligned}$$

which proves the claim. A dual argument proves that $\mathcal{V}_{c,p}(A^*)$ is closed under right quotient. Thus $\mathcal{V}_{c,p}$ is a variety of languages. $\square$

5 The formation generated by D_4 and by Q_8

We are now ready to prove our main result.

Theorem 5.1. *The groups D_4 and Q_8 generate the same formation and the associated formation of languages is the variety $\mathcal{V}_{2,2}$.*

Proof. Let $\mathbf{F}_1$ [$\mathbf{F}_2$] be the formation generated by D_4 [Q_8] and let $\mathcal{F}_1$ [$\mathcal{F}_2$] be the associated formation of languages. Let $\mathcal{V} = \mathcal{V}_{2,2}$ and let $\mathbf{V}$ be the associated group formation, which is actually a variety. For each alphabet A, $\mathcal{V}(A^*)$ is by definition the Boolean algebra generated by the languages $L(x, r, 2)$ for $0 \leqslant r < 2$ and $|x| \leqslant 2$. Proposition 4.6 shows that $\mathcal{V}$ is a variety. We shall prove successively the following properties:

(1) D_4 and Q_8 belong to $\mathbf{V}$, and hence $\mathcal{F}_1$ and $\mathcal{F}_2$ are contained in $\mathcal{V}$,
(2) for each alphabet A, for $0 \leqslant r < 2$ and $|x| \leqslant 1$, the language $L(x, r, 2)$ belongs to $\mathcal{F}_1(A^*)$ and to $\mathcal{F}_2(A^*)$,
(3) $\mathcal{V}$ is contained in $\mathcal{F}_1$ and hence $\mathcal{V} = \mathcal{F}_1$,
(4) $\mathcal{F}_1$ is contained in $\mathcal{F}_2$.

In the sequel, the languages L_1, L_2 and L_3 refer to the examples discussed in Section 4.3.

Step 1. The syntactic monoid of L_1 is equal to D_4 and that of L_3 is equal to Q_8. Formula (4.16) shows that L_1 belongs to $\mathcal{V}(\{a, b\}^*)$ and thus D_4 belongs to $\mathbf{V}$. Moreover, Formula (4.22) shows that L_3 can be written as

$$L(a, 0, 2) \cap L(b, 0, 2) \cap \Big(\bigcup_{i+j+k \equiv 0 \bmod 2} (L(ab, i, 2) \cap L(aa, j, 2) \cap L(bb, k, 2)) \Big)$$

and thus L_3 belongs to $\mathcal{V}(\{a, b\}^*)$. It follows that Q_8 belongs to $\mathbf{V}$.

Step 2. If $x = 1$, the result is trivial. If $x = a$, where a is a letter, the syntactic monoid of $L(a, r, 2)$ is the cyclic group C_2. Since C_2 is a quotient of both D_4 and Q_8, it belongs to $\mathbf{F}_1$ and to $\mathbf{F}_2$ and thus $L(a, r, 2)$ belongs to $\mathcal{F}_1(A^*)$ and to $\mathcal{F}_2(A^*)$.

Step 3. Let A be an alphabet. It suffices to prove that, for $|x| \leqslant 2$ and $r = 0$ or $r = 1$, the language $L(x, r, 2)$ belongs to $\mathcal{F}_1(A^*)$. Let $c(x)$ be the set of all letters occurring in x. In the minimal automaton of $L(x, r, 2)$, every letter of $A \setminus c(x)$ acts as the identity on the set of states. It follows that the languages $L(x, r, 2)$ and the language

$$\Big\{ u \in c(x)^* \mid \binom{u}{x} \equiv r \bmod 2 \Big\}$$

have the same syntactic monoid. Therefore, we may assume without loss of generality that $A = \{a, b\}$.

Suppose first that $x = ab$ with $a \neq b$. It already follows from (2) that for $|x| \leqslant 1$, $L(x, r, 2)$ belongs to $\mathcal{F}_1(A^*)$. Then the minimal automaton of $L(ab, 0, 2)$ is obtained from the automaton of Figure 4.1 by taking $(0, 0)$ and $(1, 0)$ as final states. Indeed in this way the parameter $z_2 = \binom{u}{ab}$ will be equal to zero modulo 2. Thus the syntactic monoid of $L(ab, 0, 2)$ is D_4 and since D_4 belongs to $\mathbf{F}_1$, the language $L(ab, 0, 2)$ belongs to $\mathcal{F}_1(A^*)$ and so does its complement $L(ab, 1, 2)$.

Consider now the case $x = aa$. The automaton obtained from the automaton of Figure 4.2 by taking $(0, 0)$ and $(1, 0)$ as final states recognizes the language

$$K = \Big\{ u \in \{a, b\}^* \mid \binom{u}{b} + \binom{u}{ba} + \binom{u}{aa} \equiv 0 \bmod 2 \Big\}$$

The syntactic monoid of K is also D_4 and thus $K \in \mathcal{F}_1(A^*)$. Now since

$$\begin{aligned} L(aa,0,2) = {} & (K \cap L(b,0,2) \cap L(ba,0,2)) \cup (K \cap L(b,1,2) \cap L(ba,1,2)) \\ & \cup (K^c \cap L(b,0,2) \cap L(ba,1,2)) \cup (K^c \cap L(b,1,2) \cap L(ba,0,2)) \end{aligned}$$

the language $L(aa,0,2)$ and its complement $L(aa,1,2)$ belong to $\mathcal{F}_1(\{a,b\}^*)$. Since the languages $L(bb,r,2)$ and $L(aa,r,2)$ have the same syntactic monoid, we also have $L(bb,r,2) \in \mathcal{F}_1(\{a,b\}^*)$ for $r = 0$ and $r = 1$.

Step 4. We will show that some language L having D_4 as syntactic monoid belongs to $\mathcal{F}_2$. By the Formation Theorem, this will show that D_4 belongs to F_2 and hence that $\mathcal{F}_1$ is contained in $\mathcal{F}_2$ as required. We choose for L the language of Example 4.2:

$$L = \varphi^{-1}(1) = \left\{ u \in \{a,b\}^* \mid \binom{u}{a} \equiv \binom{u}{b} \equiv \binom{u}{aa} + \binom{u}{ab} \equiv 0 \bmod 2 \right\}$$

Let us now view D_4 as the group $\{1, a, b, a^2, ab, ba, a^3, a^2b\}$ presented by the group relations $b^2 = 1$, $a^4 = 1$ and $a^3b = ba$ and Q_8 as the group $\{1, a, b, a^2, ab, ba, a^3, a^2b\}$ presented by the group relations $a^4 = 1$, $b^2 = a^2$ and $a^3b = ba$.

Let $B = \{a,b\}$ and $C = \{a,b,c\}$. Consider the following diagram,

$$\begin{tikzcd} B^* \arrow[rr, "\alpha"] \arrow[rrrr, bend left, "\nu_i"] \arrow[dr, "\varphi"'] & & C^* \arrow[rr, two heads, "\alpha_i"] \arrow[dl, "\kappa"] \arrow[dr, "\kappa_i"'] & & B^* \arrow[dl, "\varphi_i"] \\ & D_4 & & Q_8 & \end{tikzcd}$$

in which the morphisms are defined by

$$\begin{array}{llll} \varphi(a) = a & \varphi(b) = b & \alpha(a) = c & \alpha(b) = a \\ \varphi_1(a) = a & \varphi_1(b) = b & \varphi_2(a) = a & \varphi_2(b) = b \\ \nu_1(a) = a^2b & \nu_1(b) = a & \nu_2(a) = 1 & \nu_2(b) = a \end{array}$$

and

$$\begin{array}{lll} \alpha_1(a) = a & \alpha_1(b) = b & \alpha_1(c) = a^2b \\ \alpha_2(a) = a & \alpha_2(b) = b & \alpha_2(c) = 1 \\ \kappa_1(a) = a & \kappa_1(b) = b & \kappa_1(c) = a^2b \\ \kappa_2(a) = a & \kappa_2(b) = b & \kappa_2(c) = 1 \\ \kappa(a) = b & \kappa(b) = 1 & \kappa(c) = a \end{array}$$

Note that $\varphi_1 = \varphi_2$, but we keep two distinct names to preserve homogeneity of the notation. All these morphisms make the diagram commutative. Let

$$R_1 = \varphi_1^{-1}(1) = \varphi_2^{-1}(1) \qquad R_b = \varphi_1^{-1}(b) = \varphi_2^{-1}(b)$$
$$R_{a^2} = \varphi_1^{-1}(a^2) = \varphi_2^{-1}(a^2) \qquad R_{a^2b} = \varphi_1^{-1}(a^2b) = \varphi_2^{-1}(a^2b)$$

By construction, the languages R_1, R_b, R_{a^2} and R_{a^2b} are all recognized by Q_8 and hence belong to $\mathcal{F}_2(B^*)$.

Using the state renaming described in (4.23), one sees that R_1, R_b, R_{a^2} and R_{a^2b} are also accepted by the automaton represented in Figure 4.3 by taking as final state $(0,0,0)$, $(0,1,0)$, $(0,0,1)$ and $(0,1,1)$ respectively. Coming back to the interpretation of these states as linear combinations of binomial coefficients, as described in Example 4.3, one gets the following explicit descriptions:

$$\begin{aligned}
R_1 &= \Big\{u \in B^* \mid \tbinom{u}{a} \equiv \tbinom{u}{b} \equiv \tbinom{u}{aa} + \tbinom{u}{ab} + \tbinom{u}{bb} \equiv 0 \bmod 2\Big\}\\
R_b &= \Big\{u \in B^* \mid \tbinom{u}{a} \equiv \tbinom{u}{b} + 1 \equiv \tbinom{u}{aa} + \tbinom{u}{ab} + \tbinom{u}{bb} \equiv 0 \bmod 2\Big\}\\
R_{a^2} &= \Big\{u \in B^* \mid \tbinom{u}{a} \equiv \tbinom{u}{b} \equiv \tbinom{u}{aa} + \tbinom{u}{ab} + \tbinom{u}{bb} + 1 \equiv 0 \bmod 2\Big\}\\
R_{a^2b} &= \Big\{u \in B^* \mid \tbinom{u}{a} \equiv \tbinom{u}{b} + 1 \equiv \tbinom{u}{aa} + \tbinom{u}{ab} + \tbinom{u}{bb} + 1 \equiv 0 \bmod 2\Big\}
\end{aligned}$$

Let

$$\begin{aligned}
R = \big(\alpha_1^{-1}(R_1) \cap \alpha_2^{-1}(R_1)\big) \cup \big(\alpha_1^{-1}(R_b) \cap \alpha_2^{-1}(R_b)\big) \cup \\
\big(\alpha_1^{-1}(R_{a^2}) \cap \alpha_2^{-1}(R_{a^2})\big) \cup \big(\alpha_1^{-1}(R_{a^2b}) \cap \alpha_2^{-1}(R_{a^2b})\big)
\end{aligned}$$

We claim that

$$R = \Big\{u \in C^* \mid \tbinom{u}{a} \equiv \tbinom{u}{c} \equiv \tbinom{u}{ac} + \tbinom{u}{bc} + \tbinom{u}{cb} + \tbinom{u}{cc} \equiv 0 \bmod 2\Big\}$$

Indeed, Formula (3.4) leads to the following computations:

$$\begin{aligned}
\alpha_1^{-1}(R_1) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} + \binom{u}{c} \\
&\equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{ac} + \binom{u}{bb} + \binom{u}{bc} + \binom{u}{cb} + \binom{u}{cc} \equiv 0 \bmod 2\Big\} \\
\alpha_2^{-1}(R_1) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} \equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb} \equiv 0 \bmod 2\Big\} \\
\alpha_1^{-1}(R_b) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} + \binom{u}{c} + 1 \\
&\equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{ac} + \binom{u}{bb} + \binom{u}{bc} + \binom{u}{cb} + \binom{u}{cc} \equiv 0 \bmod 2\Big\} \\
\alpha_2^{-1}(R_b) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} + 1 \equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb} \equiv 0 \bmod 2\Big\} \\
\alpha_1^{-1}(R_{a^2}) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} + \binom{u}{c} \\
&\equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{ac} + \binom{u}{bb} + \binom{u}{bc} + \binom{u}{cb} + \binom{u}{cc} + 1 \equiv 0 \bmod 2\Big\} \\
\alpha_2^{-1}(R_{a^2}) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} \equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb} + 1 \equiv 0 \bmod 2\Big\} \\
\alpha_1^{-1}(R_{a^2b}) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} + \binom{u}{c} + 1 \\
&\equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{ac} + \binom{u}{bb} + \binom{u}{bc} + \binom{u}{cb} + \binom{u}{cc} + 1 \equiv 0 \bmod 2\Big\} \\
\alpha_2^{-1}(R_{a^2b}) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} + 1 \equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb} + 1 \equiv 0 \bmod 2\Big\}
\end{aligned}$$

It follows that

$$\begin{aligned}
\alpha_1^{-1}(R_1) \cap \alpha_2^{-1}(R_1) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} \equiv \binom{u}{c} \equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb} \\
&\equiv \binom{u}{ac} + \binom{u}{bc} + \binom{u}{cb} + \binom{u}{cc} \equiv 0 \bmod 2\Big\} \\
\alpha_1^{-1}(R_b) \cap \alpha_2^{-1}(R_b) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} + 1 \equiv \binom{u}{c} \equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb} \\
&\equiv \binom{u}{ac} + \binom{u}{bc} + \binom{u}{cb} + \binom{u}{cc} \equiv 0 \bmod 2\Big\} \\
\alpha_1^{-1}(R_{a^2}) \cap \alpha_2^{-1}(R_{a^2}) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} \equiv \binom{u}{c} \equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb} + 1 \\
&\equiv \binom{u}{ac} + \binom{u}{bc} + \binom{u}{cb} + \binom{u}{cc} \equiv 0 \bmod 2\Big\} \\
\alpha_1^{-1}(R_{a^2b}) \cap \alpha_2^{-1}(R_{a^2b}) &= \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{b} + 1 \equiv \binom{u}{c} \equiv \binom{u}{aa} + \binom{u}{ab} + \binom{u}{bb} + 1 \\
&\equiv \binom{u}{ac} + \binom{u}{bc} + \binom{u}{cb} + \binom{u}{cc} \equiv 0 \bmod 2\Big\}
\end{aligned}$$

Finally R is the union of these four languages and hence

$$R = \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{c} \equiv \binom{u}{ac} + \binom{u}{bc} + \binom{u}{cb} + \binom{u}{cc} \equiv 0 \bmod 2\Big\}$$

Now, by (3.2), one gets $\binom{u}{bc} + \binom{u}{cb} = \binom{u}{b}\binom{u}{c}$ and $\binom{u}{ac} + \binom{u}{ca} = \binom{u}{a}\binom{u}{c}$. It follows that

$$R = \Big\{u \in C^* \mid \binom{u}{a} \equiv \binom{u}{c} \equiv \binom{u}{ca} + \binom{u}{cc} \equiv 0 \bmod 2\Big\}$$

The syntactic monoid of R is D_4 and is syntactic morphism is κ.

Lemma 5.2. *The language R belongs to $\mathcal{F}_2(C^*)$.*

Proof. For $i = 1, 2$, the morphism $\varphi_i \circ \alpha_i$ is equal to κ_i and thus is surjective. By definition of a formation of languages, the languages $\alpha_i^{-1}(R_1)$, $\alpha_i^{-1}(R_b)$, $\alpha_i^{-1}(R_{a^2})$ and $\alpha_i^{-1}(R_{a^2b})$ belong to $\mathcal{F}_2(C^*)$. It follows that R belongs to $\mathcal{F}_2(C^*)$. □

Lemma 5.3. *The language $\alpha^{-1}(R)$ belongs to $\mathcal{F}_2(B^*)$.*

Proof. The syntactic morphism of R is κ. Then since $\kappa \circ \alpha = \varphi$, $\kappa \circ \alpha$ is surjective and by definition of a formation of languages, $\alpha^{-1}(R)$ belongs to $\mathcal{F}_2(B^*)$. □

The last step consists in computing $\alpha^{-1}(R)$.

Lemma 5.4. *One has $\alpha^{-1}(R) = L$ and thus L belongs to $\mathcal{F}_2(\{a, b\}^*)$.*

Proof. Since $\nu_i = \alpha_i \circ \alpha$, one gets

$$\alpha^{-1}(R) = \big(\nu_1^{-1}(R_1) \cap \nu_2^{-1}(R_1)\big) \cup \big(\nu_1^{-1}(R_b) \cap \nu_2^{-1}(R_b)\big) \cup \\ \big(\nu_1^{-1}(R_{a^2}) \cap \nu_2^{-1}(R_{a^2})\big) \cup \big(\nu_1^{-1}(R_{a^2b}) \cap \nu_2^{-1}(R_{a^2b})\big)$$

We claim that

$$\begin{aligned}\alpha^{-1}(R) &= \Big(\nu_1^{-1}(R_1) \cap \nu_2^{-1}(R_1)\Big) \cup \Big(\nu_1^{-1}(R_{a^2}) \cap \nu_2^{-1}(R_{a^2})\Big) \\ &= \Big\{u \in B^* \mid \binom{u}{a} \equiv \binom{u}{b} \equiv \binom{u}{aa} + \binom{u}{ba} \equiv 0 \bmod 2\Big\}\end{aligned}$$

Indeed, Formula (3.4) leads to the following computations:

$$\nu_1^{-1}(R_1) = \Big\{u \in B^* \mid \binom{u}{b} \equiv \binom{u}{a} \equiv \binom{u}{bb} + \binom{u}{ba} + \binom{u}{aa} \equiv 0 \bmod 2\Big\}$$
$$\nu_2^{-1}(R_1) = \Big\{u \in B^* \mid \binom{u}{b} \equiv \binom{u}{bb} \equiv 0 \bmod 2\Big\}$$
$$\nu_1^{-1}(R_b) = \Big\{u \in B^* \mid \binom{u}{b} \equiv \binom{u}{a} + 1 \equiv \binom{u}{bb} + \binom{u}{ba} + \binom{u}{aa} \equiv 0 \bmod 2\Big\}$$
$$\nu_2^{-1}(R_b) = \Big\{u \in B^* \mid \binom{u}{b} + 1 \equiv \binom{u}{bb} \equiv 0 \bmod 2\Big\}$$
$$\nu_1^{-1}(R_{a^2}) = \Big\{u \in B^* \mid \binom{u}{b} \equiv \binom{u}{a} \equiv \binom{u}{bb} + \binom{u}{ba} + \binom{u}{aa} + 1 \equiv 0 \bmod 2\Big\}$$
$$\nu_2^{-1}(R_{a^2}) = \Big\{u \in B^* \mid \binom{u}{b} \equiv \binom{u}{bb} + 1 \equiv 0 \bmod 2\Big\}$$
$$\nu_1^{-1}(R_{a^2b}) = \Big\{u \in B^* \mid \binom{u}{b} \equiv \binom{u}{a} + 1 \equiv \binom{u}{bb} + \binom{u}{ba} + \binom{u}{aa} + 1 \equiv 0 \bmod 2\Big\}$$
$$\nu_2^{-1}(R_{a^2b}) = \Big\{u \in B^* \mid \binom{u}{b} + 1 \equiv \binom{u}{bb} + 1 \equiv 0 \bmod 2\Big\}$$

It follows that

$$\nu_1^{-1}(R_1) \cap \nu_2^{-1}(R_1) = \Big\{u \in B^* \mid \binom{u}{a} \equiv \binom{u}{b} \equiv \binom{u}{bb} \equiv \binom{u}{aa} + \binom{u}{ba} \equiv 0 \bmod 2\Big\}$$
$$\nu_1^{-1}(R_b) \cap \nu_2^{-1}(R_b) = \emptyset$$
$$\nu_1^{-1}(R_{a^2}) \cap \nu_2^{-1}(R_{a^2}) = \Big\{u \in B^* \mid \binom{u}{a} \equiv \binom{u}{b} \equiv \binom{u}{bb} + 1 \equiv \binom{u}{aa} + \binom{u}{ba} \equiv 0 \bmod 2\Big\}$$
$$\nu_1^{-1}(R_{a^2b}) \cap \nu_2^{-1}(R_{a^2b}) = \emptyset$$

and thus

$$\begin{aligned}\alpha^{-1}(R) &= \Big(\nu_1^{-1}(R_1) \cap \nu_2^{-1}(R_1)\Big) \cup \Big(\nu_1^{-1}(R_{a^2}) \cap \nu_2^{-1}(R_{a^2})\Big) \\ &= \Big\{u \in B^* \mid \binom{u}{a} \equiv \binom{u}{b} \equiv \binom{u}{aa} + \binom{u}{ba} \equiv 0 \bmod 2\Big\}\end{aligned}$$

Finally, Proposition 3.1 shows that when $\binom{u}{a} \equiv \binom{u}{b} \equiv 0 \bmod 2$, then $\binom{u}{ab} \equiv \binom{u}{ba} \equiv 0 \bmod 2$. It follows that $\alpha^{-1}(R) = L$. □

This concludes the proof of Theorem 5.1. □

Important remark. It is tempting to prove directly that the languages $\nu_1^{-1}(R_1)$, $\nu_2^{-1}(R_1)$, etc. belong to $\mathcal{F}_2(\{a,b\}^*)$. However, the morphism $\varphi_2 \circ \nu_2$ is not surjective and one cannot conclude directly.

6 Conclusion

We used language theory to prove that D_4 and Q_8 generate the same formation and that this formation is a variety of groups. Our project for the future would be to show, also by language theoretic means, that any formation generated by a single nilpotent group is a variety.

Acknowledgements

We would like to thank Ramón Esteban-Romero, Adolfo Ballester-Bolinches and the anonymous referees for their useful comments and suggestions.

References

[1] A. Ballester-Bolinches and L. M. Ezquerro, *Classes of finite groups*, *Mathematics and Its Applications (Springer)* vol. 584, Springer, Dordrecht, 2006.

[2] A. Ballester-Bolinches, J.-É. Pin and X. Soler-Escrivà, Formations of finite monoids and formal languages: Eilenberg's variety theorem revisited, *Forum Math.* **26** (2014), 1737–1761.

[3] A. Ballester-Bolinches, J.-É. Pin and X. Soler-Escrivà, Languages associated with saturated formations of groups, *Forum Math.* **27** (2015), 1471–1505.

[4] O. Carton, J.-E. Pin and X. Soler-Escrivà, Languages Recognized by Finite Supersoluble Groups, *Journal of Automata, Languages and Combinatorics* **14**,2 (2009), 149–161.

[5] K. Doerk and T. Hawkes, *Finite soluble groups*, *de Gruyter Expositions in Mathematics* vol. 4, Walter De Gruyter & Co., Berlin, 1992.

[6] S. Eilenberg, *Automata, languages, and machines. Vol. B*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976. Pure and Applied Mathematics, Vol. 59.

[7] W. Gaschütz and U. Lubeseder, Kennzeichnung gesättigter Formationen, *Math. Z.* **82** (1963), 198–199.

[8] W. Guo and K. P. Shum, Formation operators on classes of algebras, *Comm. Algebra* **30**,7 (2002), 3457–3472.

[9] M. Lothaire, *Combinatorics on words*, Cambridge University Press, Cambridge, 1997. With a foreword by Roger Lyndon and a preface by Dominique Perrin, corrected reprint of the 1983 original, with a new preface by Perrin.

[10] P. M. Neumann, A note on formations of finite nilpotent groups, *Bull. London Math. Soc.* **2** (1970), 91.

[11] J.-É. Pin, *Varieties of formal languages*, North Oxford, London et Plenum, New-York, 1986. (Translation of Variétés de langages formels).

[12] C. Reutenauer, *Free Lie algebras*, *London Mathematical Society Monographs. New Series* vol. 7, The Clarendon Press, Oxford University Press, New York, 1993. Oxford Science Publications.

[13] J. J. Rotman, *Advanced modern algebra*, Prentice Hall, Inc., Upper Saddle River, NJ, 2002.

[14] L. A. Shemetkov, Product of formations of algebraic systems, *Algebra and Logic* **23**,6 (1984), 484–490.

[15] L. A. Shemetkov and A. N. Skiba, *Formations of algebraic systems. (Formatsii algebraicheskikh sistem.)*, Современная Алгебра. *[Modern Algebra]*, Sovremennaya Algebra. Moskva: Nauka. 256 P. R. 3.00, Moscow, 1989. With an English summary.

[16] A. N. Skiba, Finite subformations of varieties of algebraic systems, in *Problems in algebra, No. 2 (Russian)*, pp. 7–20, 126, "Universitet·Skoe", Minsk, 1986.

[17] H. Straubing, Families of recognizable sets corresponding to certain varieties of finite monoids, *J. Pure Appl. Algebra* **15**,3 (1979), 305–318.

[18] D. Thérien, Classification of finite monoids: the language approach, *Theoret. Comput. Sci.* **14**,2 ang. (1981), 195–208.

[19] D. Thérien, Subword counting and nilpotent groups, in *Combinatorics on words (Waterloo, Ont., 1982)*, pp. 297–305, Academic Press, Toronto, ON, 1983.

[20] P. Weil, An extension of the Schützenberger product, in *Lattices, semigroups, and universal algebra (Lisbon, 1988)*, pp. 315–321, Plenum, New York, 1990.

[21] P. Weil, Products of languages with counter, *Theoret. Comput. Sci.* **76** (1990), 251–260.

[22] P. Weil, Closure of varieties of languages under products with counter, *J. Comput. System Sci.* **45** (1992), 316–339.

Contents