10. Polynomial dynamical systems associated with substitutions

It is natural, in many situations related to physics or geometry, to study representations of the free group with values in $SL(2,\mathbb{C})$. Conjugate representations just differ by a change of coordinates, so one should be interested in quantities invariant by conjugacy; the trace is one such quantity, so that, for a representation ϕ, one is interested in studying $\operatorname{tr}\phi(W)$, for an element W of the free group.

For instance, consider the following problem. Given two complex 2×2-matrices $\mathbf{A}_0$ and $\mathbf{B}_0$ with determinant 1, define, for $n \geq 0$, $\mathbf{A}_{n+1} = \mathbf{A}_n\mathbf{B}_n$ and $\mathbf{B}_{n+1} = \mathbf{B}_n\mathbf{A}_n$ (in other words, $\mathbf{A}_n$ is a product of 2^{n+1} matrices, the factors being chosen according to the beginning of the Thue-Morse sequence). How to compute the traces of $\mathbf{A}_n$ and $\mathbf{B}_n$?

One can obtain $\big(\operatorname{tr}\mathbf{A}_n, \operatorname{tr}\mathbf{B}_n, \operatorname{tr}\mathbf{A}_n\mathbf{B}_n\big)$ by iterating the function Φ : $(x,y,z) \longmapsto (z,z,xyz-x^2-y^2+2)$. To be more precise, one has

$$\big(\operatorname{tr}\mathbf{A}_{n+1}, \operatorname{tr}\mathbf{B}_{n+1}, \operatorname{tr}\mathbf{A}_{n+1}\mathbf{B}_{n+1}\big) = \Phi\big(\operatorname{tr}\mathbf{A}_n, \operatorname{tr}\mathbf{B}_n, \operatorname{tr}\mathbf{A}_n\mathbf{B}_n\big).$$

Had we defined $\mathbf{A}_{n+1} = \mathbf{A}_n\mathbf{B}_n$ and $\mathbf{B}_n = \mathbf{A}_n$ (this time using the Fibonacci substitution), we would have obtained

$$\big(\operatorname{tr}\mathbf{A}_{n+1}, \operatorname{tr}\mathbf{B}_{n+1}, \operatorname{tr}\mathbf{A}_{n+1}\mathbf{B}_{n+1}\big) = \big(\operatorname{tr}\mathbf{A}_n\mathbf{B}_n, \operatorname{tr}\mathbf{A}_n, \operatorname{tr}\mathbf{A}_n \operatorname{tr}\mathbf{A}_n\mathbf{B}_n - \operatorname{tr}\mathbf{B}_n\big).$$

This behavior is general: given a substitution σ on the two-letter alphabet $\{\mathbf{A}_0, \mathbf{B}_0\}$, there exists a polynomial map Φ from $\mathbb{C}^3$ into itself such that, if $\mathbf{A}_n = \sigma^n(\mathbf{A}_0)$ and $\mathbf{B}_n = \sigma^n(\mathbf{B}_0)$, one has $\big(\operatorname{tr}\mathbf{A}_{n+1}, \operatorname{tr}\mathbf{B}_{n+1}, \operatorname{tr}\mathbf{A}_{n+1}\mathbf{B}_{n+1}\big) = \Phi\big(\operatorname{tr}\mathbf{A}_n, \operatorname{tr}\mathbf{B}_n, \operatorname{tr}\mathbf{A}_n\mathbf{B}_n\big)$.

Of course, to find such a recursion relation, one could think of expressing the eight entries of $\mathbf{A}_{j+1}$ and $\mathbf{B}_{j+1}$ in terms of those of $\mathbf{A}_j$ and $\mathbf{B}_j$, and then getting, by elimination, a recursion relation linking nine successive values of $\operatorname{tr}\mathbf{A}_j$. As a matter of fact, on the one hand, this method is not so bad: had we considered a recursion involving n matrices, we should have obtained a recursion relation, the length of which grows linearly in n, for the traces. On the other hand, eliminating variables could be an untractable operation, even when using computer algebra software. Besides, this method gives no idea of the algebraic properties of these recurrence formulae.

[1] This chapter has been written by J. Peyrière

Another way of operating, which is developed here, is to exploit polynomial identities in rings of matrices. This will provide an effective algorithm for constructing such recursion relations for traces, the so-called trace maps. Besides, these trace maps exhibit very interesting algebraic and geometric properties.

More generally, given a representaion ϕ in the case of the free group Γ_2 on two elements with generators a, b, for any element $W \in \Gamma_2$, there exists a unique polynomial $P_W(x, y, z)$ such that

$$\mathrm{tr}\, \phi(W) = P_W(\mathrm{tr}\, \phi(a)\, \mathrm{tr}\, \phi(b), \mathrm{tr}\, \phi(ab));$$

in other words, the trace of any product of 2 matrices $\mathbf{A}, \mathbf{B}$ can be computed by using only $\mathrm{tr}\, \mathbf{A}, \mathrm{tr}\, \mathbf{B}, \mathrm{tr}\, \mathbf{AB}$. Hence, the traces of the representation is completely determined by $[T]_\phi = (\mathrm{tr}\, \phi(a), \mathrm{tr}\, \phi(b), \mathrm{tr}\, \phi(ab)) \in \mathbb{C}^3$.

An object of particular interest are free subgroups of $SL(2, \mathbb{C})$ whose generators $\mathbf{A}, \mathbf{B}$ have a parabolic commutator (that is, $\mathbf{ABA}^{-1}\mathbf{B}^{-1}$ has trace 2, or $\mathbf{A}$ and $\mathbf{B}$ have a common eigenvector). A computation shows that the polynomial P_W associated with the word $W = aba^{-1}b^{-1}$, as defined above, is $P_W(x, y, z) = x^2 + y^2 + z^2 - xyz - 2 = \lambda(x, y, z) + 2$, where $\lambda(x, y, z) = x^2 + y^2 + z^2 - xyz - 4$. Hence, these free subgroups are given by representations ϕ such that $\lambda([T]_\phi) = 0$.

In Chap. 9, we studied endomorphisms σ of the free group; for any representation ϕ whose image is a free group, such an endomorphism gives rise to a new representation $\phi \circ \sigma$, and, by the above, one can find a polynomial map $\Phi_\sigma : \mathbb{C}^3 \to \mathbb{C}^3$ such that $[T]_{\phi \circ \sigma} = \Phi_\sigma([T_\phi])$. We will show that $\lambda \circ \Phi_\sigma$ is always divisible by λ, and equal to λ if σ is an automorphism. Thus, we recover a dynamical system on the surface $\lambda(x, y, z) = 0$ associated with the automorphisms of the free group Γ_2.

More is true: in Chap. 9, we saw that the inner automorphisms are exactly those whose abelianization is the identity. We will prove here that they are also exactly the automorphisms σ such that $\Phi_\sigma = Id$.

It is of course tempting to try to generalize, by increasing the number of letters, or the dimension. Indeed, when dealing with more than two matrices, the situation is more complex. This time, one gets for Φ a polynomial map from a certain affine algebraic variety into itself. We will show some results in this direction, but everything here becomes more difficult, as we saw already in Chap. 9.

The case of representations ϕ with value in $SL(2, \mathbb{R})$ such that $\lambda([T]_\phi) = 0$ is of particular geometric interest. Indeed, consider a once-punctured torus with an hyperbolic metric; such a torus is obtained as a quotient of the hyperbolic plane under the action of a free subgroup of $SL(2, \mathbb{R})$ of rank 2 whose generators have a parabolic commutator, and it is classical that the conjugacy classes of such groups, with given generators, parametrize the Teichmüller space of the once-punctured torus. In fact, the elements $[T]_\phi$ offer an explicitation of this parametrization, as will be proved below.

One is also interested in the modular space: one would like to forget the particular set of generators of the group. A change of generators is nothing but an automorphism σ of the free group, and this is done, at the level of the parametrization, by the polynomial map Φ_σ. Hence the modular space is the quotient of the surface $\lambda(x, y, z) = 0$ by the action of the automorphism group, via Φ_σ. Since the action of the inner automorphism group is trivial, one obtains an action of the outer automorphism group, which is isomorphic to $SL(2\mathbb{Z})$.

We recover, in a completely different way, something very similar to the modular surface discussed at the end of Chap. 6. This is just the beginning of a long story: it may be proved that the modular surface of the compact torus and of the once-punctured torus are isomorphic in a canonical way; the once-punctured torus (whose fundamental group is Γ_2) can be seen as a non-commutative version of the compact torus (whose fundamental group is $\mathbb{Z}^2$); the group of automorphisms of the free group in the first case plays the role of $SL(2, \mathbb{Z})$ in the second, and the linear representations we consider enter naturally by considering geometric structures, instead of considering lattices as we did in Chap. 6. It turns out that continued fractions and Sturmian sequences also enter naturally in the hyperbolic version of the theory (for example, by way of the parametrization of geodesics without self-intersection on the hyperbolic once-punctured torus).

The trace maps are also useful in studying certain physical problems, namely the heat or electric conduction in one-dimensional quasicrystals, modeled as chains of atoms disposed according to a substitutive sequence.

These trace maps have been widely used and studied from the point of view of iteration. But applications as well as the dynamical properties of trace maps are not within the scope of this chapter, which only aims at defining these dynamical systems.

Part of the material of this chapter comes from [323]. In the same volume, which is the proceedings of a school on quasicrystals and deterministic disorder, one can find, besides mathematical developments, many courses showing the importance of finite automata and substitutive sequences for modeling and describing certain situations in condensed matter physics.

The main additions to the lecture given at Les Houches School in Condensed Matter Physics [323] are the following: the proof that $Q_\sigma = 1$ characterizes automorphisms, and the trace maps for 3×3-matrices.

10.1 Polynomial identities in the algebra of 2×2-matrices

10.1.1 Some identities for 2×2-matrices

In this section upper case letters will stand for 2×2-matrices the entries of which are complex numbers. The basic identity is given by the Cayley-

Hamilton theorem:

$$\mathbf{A}^2 - (\operatorname{tr}\mathbf{A})\mathbf{A} + \mathbf{I}\det\mathbf{A} = 0, \tag{10.1}$$

where $\mathbf{I}$ is the identity matrix of order 2, and $\operatorname{tr}\mathbf{A}$ stands for the trace of $\mathbf{A}$.

As a consequence, one has

$$\mathbf{A}^n = p_n(\operatorname{tr}\mathbf{A}, \det\mathbf{A})\mathbf{A} - p_{n-1}(\operatorname{tr}\mathbf{A}, \det\mathbf{A})\,(\det\mathbf{A})\,\mathbf{I},$$

where p_n's are polynomials in two variables, independent of $\mathbf{A}$, with integer coefficients. If A is invertible, such a formula is also valid for negative n. The polynomials p_n's are closely related to the Chebyschev polynomials of the second kind. Indeed, if variables are denoted by x and u, we have the following recursion formula: $p_{n+1} = xp_n - up_{n-1}$, from which it results that $p_n(2\cos\varphi, 1) = \sin n\varphi / \sin\varphi$.

One has $\det\mathbf{A} = \lambda\mu = \left[(\lambda+\mu)^2 - (\lambda^2+\mu^2)\right]/2$, if λ and μ are the eigenvalues of $\mathbf{A}$. Therefore the Cayley-Hamilton relation can be rewritten as

$$\mathbf{A}^2 - \mathbf{A}\operatorname{tr}\mathbf{A} + \frac{1}{2}\left[(\operatorname{tr}\mathbf{A})^2 - \operatorname{tr}\mathbf{A}^2\right] I = 0\,.$$

This form allows bilinearization: writing this formula for $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{A}+\mathbf{B}$, one gets

$$\mathbf{AB} + \mathbf{BA} = \operatorname{tr}\mathbf{AB} - (\operatorname{tr}\mathbf{A})(\operatorname{tr}\mathbf{B}) + \mathbf{A}\operatorname{tr}\mathbf{B} + \mathbf{B}\operatorname{tr}\mathbf{A} \tag{10.2}$$

(we dropped the identity matrix $\mathbf{I}$ as, from now on, we identify scalars and scalar matrices). As this identity is a polynomial identity with integral coefficients linking the entries of matrices $\mathbf{A}$ and $\mathbf{B}$, it is valid for matrices with entries in any commutative ring.

By using (10.2), one gets $\mathbf{A}(\mathbf{AB}+\mathbf{BA}) = (\operatorname{tr}\mathbf{AB} - tr\mathbf{A}\operatorname{tr}\mathbf{B})\,\mathbf{A} + \mathbf{A}^2\operatorname{tr}\mathbf{B} + \mathbf{AB}\operatorname{tr}\mathbf{A}$. Then $\mathbf{ABA} = (\operatorname{tr}\mathbf{AB} - tr\mathbf{A}\operatorname{tr}\mathbf{B})\,\mathbf{A} + \mathbf{A}^2\operatorname{tr}\mathbf{B} + \mathbf{A}B\operatorname{tr}\mathbf{A} - \mathbf{A}^2\mathbf{B} = \mathbf{A}\operatorname{tr}\mathbf{AB} + (\mathbf{A}^2 - \mathbf{A}\operatorname{tr}\mathbf{A})\operatorname{tr}\mathbf{B} - (\mathbf{A}^2 - \mathbf{A}\operatorname{tr}\mathbf{A})\mathbf{B}$. Finally, we get the formula

$$\mathbf{ABA} = \mathbf{A}\operatorname{tr}\mathbf{AB} + \mathbf{B}\det\mathbf{A} - \det\mathbf{A}\operatorname{tr}\mathbf{B} \tag{10.3}$$

which will be useful later.

For the sake of simplicity, we shall mostly deal with complex matrices having determinant 1, i.e., elements of $SL(2,\mathbb{C})$).

Proposition 10.1.1. *If $m_1, n_1, m_2, n_2, \cdots, m_k, n_k$ is a sequence of integers, there exist four polynomials p, q, r, and s in three variables with integer coefficients such that, for any pair of matrices $\mathbf{A}$ and $\mathbf{B}$ in $SL(2,\mathbb{C})$, one has*

$$\begin{aligned}\mathbf{A}^{m_1}\mathbf{B}^{n_1}\mathbf{A}^{m_2}\mathbf{B}^{n_2}\cdots\mathbf{A}^{m_k}\mathbf{B}^{n_k} = \; & p(\operatorname{tr}\mathbf{A}, \operatorname{tr}\mathbf{B}, \operatorname{tr}\mathbf{AB}) + \\ & q(\operatorname{tr}\mathbf{A}, \operatorname{tr}\mathbf{B}, \operatorname{tr}\mathbf{AB})\,\mathbf{A} + \\ & r(\operatorname{tr}\mathbf{A}, \operatorname{tr}\mathbf{B}, \operatorname{tr}\mathbf{AB})\,\mathbf{B} + \\ & s(\operatorname{tr}\mathbf{A}, \operatorname{tr}\mathbf{B}, \operatorname{tr}\mathbf{AB})\,\mathbf{AB}.\end{aligned}$$

Proof. By using repeatedly the Cayley-Hamilton theorem for $\mathbf{A}$ and $\mathbf{B}$, one is left with a linear combination with polynomial coefficients of $\mathbf{I}$, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{AB}$, $\mathbf{BA}$, and of products of the form $\mathbf{ABA}\cdots$ or $\mathbf{BAB}\cdots$. Then by using the Cayley-Hamilton theorem for $\mathbf{AB}$ and $\mathbf{BA}$, one is left with a linear combination of I, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{AB}$, $\mathbf{BA}$, $\mathbf{ABA}$, and $\mathbf{BAB}$. We conclude by using (10.2) and (10.3). ■

Corollary 10.1.2. *Given m's and n's as in Proposition 10.1.1, there exists a unique polynomial P in three variables with integer coefficients such that, for any pair* $(\mathbf{A}, \mathbf{B})$ *of unimodular* 2×2*-matrices, one has*

$$\operatorname{tr} \mathbf{A}^{m_1} \mathbf{B}^{n_1} \cdots \mathbf{A}^{m_k} \mathbf{B}^{m_k} = P(\operatorname{tr} \mathbf{A}, \operatorname{tr} \mathbf{B}, \operatorname{tr} \mathbf{AB}) .$$

Proof. The existence results from the above proposition. The uniqueness follows from the fact that $(\operatorname{tr} \mathbf{A}, \operatorname{tr} \mathbf{B}, \operatorname{tr} \mathbf{AB})$ can assume any value (x, y, z). To see this, just take

$$\mathbf{A} = \begin{pmatrix} x & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 0 & t \\ z+t & y \end{pmatrix}$$

with $t(z+t)+1=0$. ■

Remark. Had we not restricted the determinants of $\mathbf{A}$ and $\mathbf{B}$ to be 1, the trace of the product above would have been expressed as a polynomial in the five variables $\operatorname{tr} \mathbf{A}$, $\operatorname{tr} \mathbf{B}$, $\operatorname{tr} \mathbf{AB}$, $\det \mathbf{A}$, and $\det \mathbf{B}$.

Examples. Let $\mathbf{A}$ and $\mathbf{B}$ be two unimodular 2×2-matrices, $x = \operatorname{tr} \mathbf{A}$, $y = \operatorname{tr} \mathbf{B}$, $z = \operatorname{tr} \mathbf{AB}$.

1.

$$\begin{aligned}(\mathbf{AB} - \mathbf{BA})^2 &= (\mathbf{AB})^2 + (\mathbf{BA})^2 - \mathbf{AB}^2\mathbf{A} - \mathbf{BA}^2\mathbf{B} \\ &= z(\mathbf{AB} + \mathbf{BA}) - 2 - y\mathbf{ABA} - x\mathbf{BAB} + \mathbf{A}^2 + \mathbf{B}^2 \\ &= z(z - xy + y\mathbf{A} + x\mathbf{B}) - 2 - y(z\mathbf{A} + \mathbf{B} - y) \\ &\quad - x(z\mathbf{B} + \mathbf{A} - x) + x\mathbf{A} + y\mathbf{B} - 2 \\ &= x^2 + y^2 + z^2 - xyz - 4.\end{aligned}$$

This result is not surprising because $\operatorname{tr}(\mathbf{AB} - \mathbf{BA}) = 0$. The polynomial

$$\lambda(x, y, z) = x^2 + y^2 + z^2 - xyz - 4 \tag{10.4}$$

will play an important role in the sequel. The above formula says that the determinant of $\mathbf{AB} - \mathbf{BA}$ is $-\lambda(x, y, z)$. This is an easy exercise in linear algebra to show that $\det(\mathbf{AB} - \mathbf{BA}) = 0$ if and only if the matrices $\mathbf{A}$ and $\mathbf{B}$ have a common eigendirection. Indeed, let $\mathbf{e}$ be a non-zero element in $\ker(\mathbf{AB} - \mathbf{BA})$. If $\mathbf{e}$ and $\mathbf{Ae}$ are independent, then $(\mathbf{AB} - \mathbf{BA})\mathbf{Ae} = \mathbf{AABe} - \mathbf{BAAe} = x(\mathbf{AB} - \mathbf{BA})\mathbf{e} = 0$, which means that $\mathbf{AB} = \mathbf{BA}$ (recall that we are here in dimension 2). If $\mathbf{Ae} = \rho\mathbf{e}$, then $\rho\mathbf{Be} = \mathbf{ABe}$, which means that, if $\mathbf{A} \neq \rho\mathbf{I}$, there exists ρ' such that $\mathbf{Be} = \rho'\mathbf{e}$.

Thus $\mathbf{A}$ *and* $\mathbf{B}$ *have a common eigendirection if and only if* $\lambda(x, y, z) = 0$.

2.

$$\begin{aligned}\mathbf{ABA}^{-1}\mathbf{B}^{-1} &= \mathbf{AB}(x - \mathbf{A})(y - \mathbf{B}) \\ &= (\mathbf{AB})^2 + xy\mathbf{AB} - y\mathbf{ABA} - x\mathbf{AB}^2 \\ &= (\mathbf{AB})^2 + x\mathbf{A} - y\mathbf{ABA} \\ &= z\mathbf{AB} - 1 + (x - yz)\mathbf{A} - y(\mathbf{B} - y) \\ &= z\mathbf{AB} + (x - yz)\mathbf{A} - y\mathbf{B} + y^2 - 1.\end{aligned}$$

Therefore $\operatorname{tr} \mathbf{ABA}^{-1}\mathbf{B}^{-1} = \lambda(x, y, z) + 2$. ∎

We now turn our attention to formulae involving more than two elements of $SL(2, \mathbb{C})$. As a consequence of (10.2), we have the following proposition.

Proposition 10.1.3. *If* $\{\mathbf{A}_j\}_{1 \leq j \leq n}$ *are elements of* $SL(2, \mathbb{C})$*, then*

1. *any product constructed from these matrices or their inverses, can be written as a linear combination of the* 2^n *matrices* $\mathbf{A}_{i_1}\mathbf{A}_{i_2}\cdots\mathbf{A}_{i_k}$ $(0 \leq k \leq n,\ i_1 < i_2 < \cdots < i_k)$ [1] *the coefficients of which are polynomials in the* $2^n - 1$ *variables* $\operatorname{tr} \mathbf{A}_{i_1}\mathbf{A}_{i_2}\cdots\mathbf{A}_{i_k}$ $(1 \leq k \leq n,\ i_1 < i_2 < \cdots < i_k)$,
2. *the trace of such a product can be expressed as a polynomial with integer coefficients in the* $2^n - 1$ *traces defined above.*

We now turn to formulae which involve three matrices $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ in $SL(2, \mathbb{C})$. Let x_1, x_2, x_3, y_1, y_2, and y_3 denote the traces of $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$, $\mathbf{A}_2\mathbf{A}_3$, $\mathbf{A}_3\mathbf{A}_1$, and $\mathbf{A}_1\mathbf{A}_2$.

Define the following polynomials:

$$p(X, Y) = x_1y_1 + x_2y_2 + x_3y_3 - x_1x_2x_3 \tag{10.5}$$

$$\begin{aligned}q(X, Y) = x_1^2 + x_2^2 + x_3^2 + y_1^2 + y_2^2 + y_3^2 \\ - x_1x_2y_3 - x_2x_3y_1 - x_3x_1y_2 + y_1y_2y_3 - 4\end{aligned} \tag{10.6}$$

where X stands for the collection of x's and similarly for Y.

Proposition 10.1.4 (Fricke lemma). *One has*

1. $\operatorname{tr}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3) + \operatorname{tr}(\mathbf{A}_1\mathbf{A}_3\mathbf{A}_2) = p(X, Y)$
2. $\operatorname{tr}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3)\operatorname{tr}(\mathbf{A}_1\mathbf{A}_3\mathbf{A}_2) = q(X, Y)$.

[1] Of course, for $k = 0$, this product should be interpreted as the identity matrix.

Proof. To prove assertion 1, write $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3 + \mathbf{A}_1\mathbf{A}_3\mathbf{A}_2 = \mathbf{A}_1(\mathbf{A}_2\mathbf{A}_3 + \mathbf{A}_3\mathbf{A}_2)$ and use (10.2).

To prove assertion 2, write $[\mathbf{A}_1\,(\mathbf{A}_2\mathbf{A}_3)\,\mathbf{A}_1]\,\mathbf{A}_3\mathbf{A}_2 = \mathbf{A}_1\mathbf{A}_2\,(\mathbf{A}_3\mathbf{A}_1\mathbf{A}_3)\,\mathbf{A}_2$ and use (10.3) twice. We obtain

$$\mathbf{A}_1\mathbf{A}_3\mathbf{A}_2\,\mathrm{tr}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3) + \mathbf{A}_2\mathbf{A}_3^2\mathbf{A}_2 - y_1\mathbf{A}_3\mathbf{A}_2 = \\ y_2\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3\mathbf{A}_2 + (\mathbf{A}_1\mathbf{A}_2)^2 - x_1\mathbf{A}_1\mathbf{A}_2^2.$$

By reducing further, we get

$$\mathbf{A}_1\mathbf{A}_3\mathbf{A}_2\,\mathrm{tr}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3) = -2 + x_3^2 + y_1^2 - y_1x_2x_3 + (x_1 - y_2x_3)\mathbf{A}_1 + x_2\mathbf{A}_2 \\ -(x_3 - y_1x_2)\mathbf{A}_3 + (y_1y_2 + y_3 - x_1x_2)\mathbf{A}_1\mathbf{A}_2 + y_2\mathbf{A}_1\mathbf{A}_3 - y_1\mathbf{A}_2\mathbf{A}_3\,,$$

from which assertion 2 follows by taking the trace. ∎

Corollary 10.1.5. $\mathrm{tr}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3)$ *and* $\mathrm{tr}(\mathbf{A}_1\mathbf{A}_3\mathbf{A}_2)$ *are the roots of the equation*

$$z^2 - p(X,Y)z + q(X,Y) = 0.$$

This leads to define a polynomial in seven variables

$$\Lambda(X,Y,z) = z^2 - p(X,Y)z + q(X,Y) \tag{10.7}$$

This corollary means that variables x's, y's, and $z = \mathrm{tr}(\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3)$ are not independent. Indeed the set of polynomials P, in seven variables with integer coefficients such that, for any triple $\{\mathbf{A}_j\}_{1\le j\le 3}$ of elements of $SL(2,\mathbb{C})$, one has

$$P(\mathrm{tr}\,\mathbf{A}_1, \mathrm{tr}\,\mathbf{A}_2, \mathrm{tr}\,\mathbf{A}_3, \mathrm{tr}\,\mathbf{A}_2\mathbf{A}_3, \mathrm{tr}\,\mathbf{A}_3\mathbf{A}_1, \mathrm{tr}\,\mathbf{A}_1\mathbf{A}_2, \mathrm{tr}\,\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3) = 0,$$

is an ideal containing Λ. It can be shown that this ideal is generated by Λ. Therefore, the polynomial the existence of which is asserted in Proposition 10.1.3-2 is not unique when $n > 2$. In the case $n = 3$ it is defined up to a multiple of Λ.

Proposition 10.1.6. *Let* z *stand for* $\mathrm{tr}\,\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3$. *One has*

$$2\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3 = z - x_1y_1 - x_3y_3 + x_1x_2x_3 + (y_1 - x_3x_3)\mathbf{A}_1 - y_2\mathbf{A}_2 \\ +(y_3 - x_1x_2)\mathbf{A}_3 + x_3\mathbf{A}_1\mathbf{A}_2 + x_2\mathbf{A}_1\mathbf{A}_3 + x_1\mathbf{A}_2\mathbf{A}_3\,.$$

Proof. In $\mathbf{A}_1(\mathbf{A}_2\mathbf{A}_3)$, commute $\mathbf{A}_1$ and $\mathbf{A}_2\mathbf{A}_3$ by using (10.2). We get, among other terms, $-\mathbf{A}_2\mathbf{A}_3\mathbf{A}_1$. By using (10.2) twice, on can make $\mathbf{A}_1$ to jump over $\mathbf{A}_3$ and $\mathbf{A}_2$. So $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3$ can be written as $-\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3$ plus a linear combination of I, $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$, and products of two such matrices. ∎

Corollary 10.1.7. *If* n *is larger than 3 and* $\{\mathbf{A}_j\}_{1\le j\le n}$ *are elements of* $SL(2,\mathbb{C})$, *then*

1. *any product constructed from these matrices or their inverses, can be written as a linear combination of the matrices* $\mathbf{I}$, $\mathbf{A}_i$ $(1 \leq i \leq n)$, *and* $\mathbf{A}_{i_1}\mathbf{A}_{i_2}$ $(1 \leq i_1 < i_2 \leq n)$, *the coefficients of which are polynomials in the variables* $\operatorname{tr}\mathbf{A}_i$ $(1 \leq i \leq n)$, $\operatorname{tr}\mathbf{A}_{i_1}\mathbf{A}_{i_2}$ $(1 \leq i_1 < i_2 \leq n)$, *and* $\operatorname{tr}\mathbf{A}_{i_1}\mathbf{A}_{i_2}\mathbf{A}_{i_3}$ $(1 \leq i-1 < i_2 < i_3 \leq n)$.
2. *the trace of such a product can be expressed as a polynomial with rational coefficients in the* $n(n^2+5)/6$ *traces defined above.*

This last corollary is a significant improvement on Proposition 10.1.3 when n is larger than 3.

As a matter of fact, one can go further reducing the number of traces needed. It results from [271] that the trace of a product of the kind considered above can be expressed as a rational fraction in the variables $\operatorname{tr}\mathbf{A}_i$ $(1 \leq i \leq n)$, $\operatorname{tr}\mathbf{A}_i\mathbf{A}_j$ $(i < j,\ 1 \leq i \leq 3,\ 1 < j \leq n)$, and $\operatorname{tr}\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3$ (see also [324]).

10.1.2 Free groups and monoids

Let $\mathcal{A} = \{A_1, A_2, \cdots, A_n\}$ be a finite set called alphabet.

Free semi-group generated by $\mathcal{A}$. Let $\mathcal{A}^*$ be the set of words over the alphabet $\mathcal{A}$. Recall that the product W_1W_2 of two of its elements is just the word obtained by putting the word W_2 after the word W_1. This operation is called concatenation. It is associative and has a unit element, the empty word, denoted by ε. The set $\mathcal{A}^*$ endowed with this structure is called the free semi-group or free monoid generated by $\mathcal{A}$.

Free group generated by $\mathcal{A}$. We perform the same construction as above with the alphabet $\{A_1, A_2, \cdots, A_n,\ \ A_1^{-1}, \cdots, A_n^{-1}\}$, but we introduce the following simplification rules:

$$A_jA_j^{-1} = A_j^{-1}A_j = \varepsilon \quad (\text{for } 1 \leq j \leq n).$$

We obtain a group[2] which we will denote by $\Gamma_{\mathcal{A}}$ and call the free group generated by $\mathcal{A}$.

In the case of a two-letter alphabet, $\Gamma_{\{a,b\}}$ will be simply denoted by Γ.

Abelianization map. The notations defined for the free monoid extend to the free group: if W is an element of $\Gamma_{\mathcal{A}}$ and $a \in \mathcal{A}$, $|W|_a$ stands for the sums of exponents of a in W (one can easily be convinced that $|W|_a$ only depends on W and not of the particular word used to represent it). Moreover, $|W|_a$ is independent of the order of the factors in W. Recall that the mapping $W \longmapsto \mathbf{l}(W) = (|W|_{A_1}, \cdots, |W|_{A_n})$ (where $A_1, A_2, \cdots, A_n$ are the elements of $\mathcal{A}$) is a group homomorphism of $\Gamma_{\mathcal{A}}$ onto $\mathbb{Z}^n$ known as the abelianization map.

[2] Strictly speaking, elements of this group are not words, but equivalence classes of words.

For instance,

$$|aba^{-1}|_a = 0 \qquad |aba^{-1}|_b = 1$$
$$|ab^{-1}a^{-2}|_a = -1 \quad |ab^{-1}a^{-2}|_b = -1.$$

Representations in $SL(2,\mathbb{C})$. A representation of $\mathcal{A}^*$ or $\Gamma_{\mathcal{A}}$ in $SL(2,\mathbb{C})$ is a mapping φ from $\mathcal{A}^*$ or $\Gamma_{\mathcal{A}}$ into $SL(2,\mathbb{C})$ such that

$$\varphi(W_1W_2) = \varphi(W_1)\varphi(W_2)$$

for any W_1 and W_2 in $\mathcal{A}^*$ or $\Gamma_{\mathcal{A}}$.

Such a representation is determined by the values $\mathbf{A}_j$ of $\varphi(A_j)$ for $j = 1, 2, \cdots, n$. Computing $\varphi(W)$ simply consists in replacing each letter in W by the corresponding matrix.

10.1.3 Reformulation in terms of polynomial identities algebras

The following proposition is mainly a reformulation of the corollary to Proposition 10.1.1.

Proposition 10.1.8. *For any $W \in \Gamma$, there exists a unique polynomial P_W with integer coefficients such that, for any representation φ of Γ in $SL(2,\mathbb{C})$, one has*

$$\operatorname{tr}\varphi(W) = P_W(\operatorname{tr}\varphi(a), \operatorname{tr}\varphi(b), \operatorname{tr}\varphi(ab)).$$

Moreover, if $\mathbf{l}(W_1) = \mathbf{l}(W_2)$, the polynomial $P_{W_1} - P_{W_2}$ is divisible by λ.

Proof. We only have to prove the second assertion. Consider a representation φ such that the matrices $\mathbf{A} = \varphi(a)$ and $\mathbf{B} = \varphi(b)$ share an eigenvector. As these matrices are simultaneously trigonalizable, the trace of a product of $\mathbf{A}$'s and $\mathbf{B}$'s does not depend on the order of factors. This means that $\operatorname{tr}\varphi(W_1) = \operatorname{tr}\varphi(W_2)$. In other terms, we have $P_{W_1}(x,y,z) = P_{W_2}(x,y,z)$ as soon as $\lambda(x,y,z) = 0$. The conclusion then follows from the irreducibility of λ. ■

According to Horowitz [209], polynomials P_W are called Fricke characters of Γ.

The following notation will prove to be convenient: if φ is a representation of Γ in $SL(2,\mathbb{C})$, set

$$[T]\varphi = \big(\operatorname{tr}\varphi(a), \operatorname{tr}\varphi(b), \operatorname{tr}\varphi(ab)\big). \tag{10.8}$$

With this notation, the equation of definition of P_W is $\operatorname{tr}\varphi(W) = P_W([T]\varphi)$.

The reader may wonder whether it was necessary to replace lower case letters by upper case ones (i.e., to replace a letter by its image under a representation) in the previous calculations. Indeed, this is not compulsory. If we analyze what we have done, we have just considered a and b to be generators of an algebra on the ring $\mathbb{Z}[x, y, z]$, subject to the following relations:

$$a^2 - xa + 1 = 0,\ b^2 - yb + 1 = 0, \text{and } ab + ba = z - xy + ya + xb\,.$$

It is an algebra with polynomial identities (a PI algebra) which we call the Procesi-Razmyslov algebra on a two-letter alphabet. Then Proposition 10.1.1 can be interpreted as giving a homomorphism of the group algebra[3] of Γ to the Procesi-Razmyslov algebra.

We denote again by tr the $\mathbb{Z}[x, y, z]$-linear form on this algebra which maps 1, a, b, and ab respectively on 2, x, y, and z. Then, for any $W \in \Gamma$, one has $\operatorname{tr} W = P_W$. Moreover, it is easy to show that, if u and v are two elements of this algebra, one has $\operatorname{tr} uv = \operatorname{tr} vu$.

10.2 Trace maps

10.2.1 Endomorphisms of free groups

A map σ from $\Gamma_{\mathcal{A}}$ to $\Gamma_{\mathcal{A}}$ is an endomorphism of $\Gamma_{\mathcal{A}}$ if, for any W_1 and W_2 in $\Gamma_{\mathcal{A}}$, one has

$$\sigma(W_1 W_2) = \sigma(W_1)\sigma(W_2)\ .$$

In the case where none of the words $\sigma(A_j)$ contains negative powers, σ is an endomorphism of $\mathcal{A}^*$ and we recover the notion of substitution on the alphabet $\mathcal{A}$.

Obviously, an endomorphism σ is determined by $\sigma(A_j)$ $(j = 1, 2, \cdots, n)$. Hereafter we shall identify an endomorphism σ and the collection of words $\big(\sigma(A_1), \sigma(A_2), \cdots, \sigma(A_n)\big)$.

For instance[4], $\sigma = (ab, a)$ means that σ is the endomorphism, so called the Fibonacci substitution, such that $\sigma(a) = ab$ and $\sigma(b) = a$. In this case, as an example, let us compute $\sigma(aba^{-1})$:

$$\begin{aligned}\sigma(aba^{-1}) &= \sigma(a)\sigma(b)\sigma(a)^{-1}\\ &= aba(ab)^{-1} = abab^{-1}a^{-1}\ .\end{aligned}$$

The composition of endomorphisms is simply the composition of maps. This is illustrated by the following examples:

- $(ab, ba) \circ (ab, a) = (abba, ab)$,
- $(ab, a) \circ (b, b^{-1}a) = (b, b^{-1}a) \circ (ab, a) = (a, b)$ (this means that, as an endomorphism of $\Gamma_{\mathcal{A}}$, the Fibonacci substitution is invertible).

[3] This algebra is the set of finite formal linear combinations of elements of Γ endowed with the bilinear multiplication which extends the product in Γ.

[4] In the case of a two-letter alphabet, we prefer to denote by a and b the generators instead of A_1 and A_2.

Recall that the $n \times n$-matrix $\mathbf{M}_\sigma$ whose entry of indices (i,j) is $|\sigma(A_j)|_{A_i}$ is, by definition, the matrix of the endomorphism σ.

For instance, the Fibonacci substitution matrix is $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, and the one of the Morse substitution, (ab, ba), is $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

One has, for any σ and W,

$$\mathbf{l}(\sigma(W)) = \mathbf{M}_\sigma \mathbf{l}(W)$$

and, for any pair of endomorphisms,

$$\mathbf{M}_{\sigma_1 \circ \sigma_2} = \mathbf{M}_{\sigma_1} \times \mathbf{M}_{\sigma_2} .$$

10.2.2 Trace maps (two-letter alphabet)

Recall that Γ stands for the free group on the two generators a and b.

Definition of trace maps. Let σ be an endomorphism of Γ. We define the trace map associated with σ to be

$$\Phi_\sigma = \left(P_{\sigma(a)}, P_{\sigma(b)}, P_{\sigma(ab)}\right) . \tag{10.9}$$

It can be considered as well as a map from $\mathbb{C}^3$ to $\mathbb{C}^3$.

Let us compute the trace map for the Morse substitution $\sigma = (ab, ba)$. We operate in the Procesi-Razmyslov algebra. We wish to compute $(\operatorname{tr} ab, \operatorname{tr} ba, \operatorname{tr} ab^2a)$. We have $\operatorname{tr} ab = \operatorname{tr} ba = z$ and $ab^2a = a(yb-1)a = y(za+b-y) - (xa-1)$, so $\operatorname{tr} ab^2a = xyz - y^2 - x^2 + 2$. At last, $\Phi_\sigma = \left(z, z, xyz - x^2 - y^2 + 2\right)$.

Here are a few examples of trace maps.

σ	Φ_σ
inner automorphism	(x, y, z)
(a^{-1}, b^{-1})	(x, y, z)
(b, a)	(y, x, z)
(ab, b^{-1})	(z, y, x)
(b, a^{-1})	$(y, x, xy - z)$
(ab, a)	$(z, x, xz - y)$
$(b, b^{-1}a)$	$(y, xy - z, x)$
(ab, ba)	$(z, z, xyz - x^2 - y^2 + 2)$
(aba, b)	$(xz - y, y, z^2 - 2)$
(a^2b, ba)	$(xz - y, z, x^2yz - x^3 - xy^2 - yz + 3x)$
(aab, bab)	$(xz - y, yz - x, xyz^2 - (x^2 + y^2 - 1)z)$

First properties of trace maps.

Proposition 10.2.1. *For any endomorphisms σ and τ of Γ, we have $\Phi_{\sigma \circ \tau} = \Phi_\tau \circ \Phi_\sigma$.*

Proof. We have the following characterization of Φ_σ: for any representation φ,

$$[T](\varphi \circ \sigma) = \Phi_\sigma([T]\varphi).$$

The proposition then results from

$$\Phi_{\sigma\circ\tau}([T]\varphi) = [T](\varphi \circ \sigma \circ \tau) = \Phi_\tau([T](\varphi \circ \sigma)) = \Phi_\tau \circ \Phi_\sigma([T]\varphi)\,.$$

■

Corollary 10.2.2. *For any endomorphism σ of Γ, and for any $W \in \Gamma$, one has*

$$P_{\sigma(W)} = P_W \circ \Phi_\sigma.$$

Proof. Let τ be the endomorphism (W, b). Then $P_{\sigma(W)}$ is the first component of $\Phi_{\sigma\circ\tau}$, i.e., the first component of Φ_τ composed with Φ_σ.

As a consequence, if σ has a fixed point W, the corresponding trace map Φ_σ leaves the surfaces of $P_W(x,y,z) = \text{constant}$ globally invariant. ■

If σ is invertible, then $\Phi_\sigma \circ \Phi_{\sigma^{-1}} = \text{id}$. Taking the Jacobian, we get

$$\det(\Phi'_\sigma \circ \Phi_{\sigma^{-1}}) \det(\Phi'_{\sigma^{-1}}) = 1\,.$$

As these determinants are polynomials with integer coefficients, we must have $\det \Phi'_\sigma \equiv 1$ or $\det \Phi'_\sigma \equiv -1$.

As an example, consider the Morse substitution for which we have

$$\Phi'_\sigma = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ yz - 2x & xz - 2y & xy \end{pmatrix}.$$

The corresponding determinant is 0, so the Morse substitution is not invertible.

Let us consider another example: $\sigma = (aba, b)$. Then

$$\Phi'_\sigma = \begin{pmatrix} z & -1 & x \\ 0 & 1 & 0 \\ 0 & 0 & 2z \end{pmatrix}.$$

The determinant equals $2z^2$, therefore σ is not invertible.

Proposition 10.2.3. *For any endomorphism σ of Γ, there exists a polynomial Q_σ with integer coefficients such that*

$$\lambda \circ \Phi_\sigma = \lambda \cdot Q_\sigma.$$

Proof. Let x, y, and z be such that $\lambda(x,y,z) = 0$ and $z \neq 0$. Choose unimodular matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\operatorname{tr}\mathbf{A} = x$, $\operatorname{tr}\mathbf{B} = y$, and $\operatorname{tr}\mathbf{AB} = z$, and consider the representation φ defined by $\varphi(a) = \mathbf{A}$ and $\varphi(b) = \mathbf{B}$. As $\mathbf{A}$ and $\mathbf{B}$ share an eigenvector, so do $\varphi(\sigma(a))$ and $\varphi(\sigma(b))$. So we have $\lambda(\mathrm{T}(\varphi \circ \sigma)) = 0$. Therefore $\lambda(x,y,z) = 0$ implies $\lambda(\Phi_\sigma(x,y,z)) = 0$. Since λ is irreducible, it divides $\lambda \circ \Phi_\sigma$. ∎

As a consequence, any Φ_σ leaves globally invariant the surface Ω the equation of which is $\lambda(x,y,z) = 0$. Moreover, the restriction of Φ_σ to Ω only depends on $\mathbf{M}_\sigma$.

Lemma 10.2.4. *For any σ, we have $Q_\sigma(0,0,0) = 0$ or 1.*

Proof. This is checked by testing on matrices $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$. ∎

Proposition 10.2.5. *If σ and τ are endomorphisms, we have*

$$Q_{\sigma\circ\tau} = Q_\sigma \cdot Q_\tau \circ \Phi_\sigma \,.$$

Proof. We have

$$\lambda \cdot Q_{\sigma\circ\tau} = \lambda \circ \Phi_{\sigma\circ\tau} = (\lambda \circ \Phi_\tau) \circ \Phi_\sigma = \lambda \circ \Phi_\sigma \cdot Q_\tau \circ \Phi_\sigma = \lambda \cdot Q_\sigma \cdot Q_\tau \circ \Phi_\sigma \,.$$

∎

Corollary 10.2.6. *If σ is invertible, then $Q_\sigma \equiv 1$ and Φ_σ leaves globally invariant each surface $\lambda(x,y,z) = \text{constant}$.*

Characterization of automorphisms of Γ in terms of $\mathbf{Q}_\sigma$. In this section, if $W \in \Gamma$, we shall also denote the polynomial P_W by $\operatorname{tr} W$. Let us set $\operatorname{tr} a = x$, $\operatorname{tr} b = y$, and $\operatorname{tr} ab = z$. As we have seen in Sec. 10.1.1, we have $\operatorname{tr} a^n = u_n(x)a - u_{n-1}(x)$ for $n \in \mathbb{Z}$, where the polynomials u_n satisfy $u_0 = 0$, $u_1(x) = 1$, and $u_{n+1}(x) + u_{n-1}(x) = x\, u_n(x)$.

Two elements W and W' of Γ are conjugate if there exists $V \in \Gamma$ such that $W' = VWV^{-1}$. In this case, we have $\operatorname{tr} W = \operatorname{tr} W'$.

Any $W \in \Gamma$ is conjugate to ε, a^m, b^n, or to $a^{m_1}b^{n_1}a^{m_2}b^{n_2}\cdots a^{m_k}b^{m_k}$ with $\prod_{j=1}^{k} m_j n_j \neq 0$ (such a form will be called a cyclic reduction of W). In the latter case, we shall say that k is the width of W, otherwise the width of W is 0.

Lemma 10.2.7. *If $W \in \Gamma$, the degree, $\mathrm{d}^\circ_z P_W$, of P_W with respect to the variable z equals the width of W.*

Proof. It goes by induction on the width k of W. This is true for $k = 0$.

Suppose that $k \geq 1$ and the lemma is true for any word of width less than k. Consider $W = a^{m_1} b^{n_1} a^{m_2} b^{n_2} \cdots a^{m_k} b^{m_k}$, and write $W = a^{m_1} b^{n_1} W'$. One has $W = a^{m_1} b^{n_1} a^{m_1} a^{-m_1} W'$. Equality (10.3) shows that

$$\begin{aligned} W &= \left[a^{m_1} \operatorname{tr} a^{m_1} b^{n_1} - b^{-n_1}\right] a^{-m_1} W' \\ &= W' \operatorname{tr} a^{m_1} b^{n_1} - b^{-n_1} a^{-m_1} W'. \end{aligned}$$

But, the second term of the last equality has a width less than k. One has $a^{m_1} b^{n_1} = \left(u_{m_1}(x) a - u_{m_1-1}(x)\right)\left(u_{n_1}(y) b - u_{n_1-1}(y)\right)$, from which it results that $\mathrm{d}_z^\circ \operatorname{tr} a^{m_1} b^{n_1} = 1$. Therefore, $\mathrm{d}_z^\circ \operatorname{tr} W = \mathrm{d}_z^\circ \operatorname{tr} W' + 1$. This proves the lemma. ■

Lemma 10.2.8. *If $W \in \Gamma$ is such that $\operatorname{tr} W = \alpha z$, with $\alpha \in \mathbb{Z}$, then $\alpha = 1$ and the cyclic reduction of W is either ab or $a^{-1}b^{-1}$.*

Proof. If we had $\alpha = 0$, the cyclic reduction of W would be ε, a^m, or b^n. But, the trace is nonzero in any of these cases. Therefore $\alpha \neq 0$.

So, by the preceding lemma, a cyclic reduction of W is of the form $a^m b^n$ with $mn \neq 0$. Therefore we have

$$\alpha z = u_m(x) u_n(y)\, z - y\, u_{m-1}(x) u_n(y) - x\, u_m(x) u_{n-1}(y) + 2 u_{m-1}(x) u_{n-1}(y).$$

By looking at the coefficients of z in both sides, we get $|m| = |n| = 1$. It is then easy to show that we must have $mn = 1$. ■

Lemma 10.2.9. *For $W \in \Gamma$, if $\operatorname{tr} W = \alpha x$, then $\alpha = 1$ and W is conjugate either to a or to a^{-1}.*

Proof. Consider the following automorphism of Γ: $\sigma = (ab, b^{-1})$. As one has $\Phi_\sigma(x, y, z) = (z, y, x)$, it results from the corollary to Proposition 10.2.1 that $\operatorname{tr} \sigma(W) = \alpha z$. Then, by virtue of the preceding lemma, $\alpha = 1$ and $\sigma(W)$ writes $V(ab)^{\pm 1} V^{-1}$. Then, $W = \sigma^{-1}(V) a^{\pm 1} \sigma^{-1}(V^{-1})$. ■

Of course, a similar result holds if $\operatorname{tr} W = \alpha y$.

Proposition 10.2.10. *Let σ be an endomorphism of Γ. Then $\Phi_\sigma = \mathrm{Id}$ if and only if σ is either an inner automorphism of Γ or an inner automorphism composed with the involution (a^{-1}, b^{-1}).*

Proof. Suppose $\Phi_\sigma = \mathrm{Id}$. It results from the preceding lemma that $\sigma(a) = U a^\varepsilon U^{-1}$ and $\sigma(b) = V b^\eta V^{-1}$, with $|\varepsilon| = |\eta| = 1$. By Proposition 10.1.8, λ divides $\operatorname{tr} \sigma(ab) - \operatorname{tr} a^\varepsilon b^\eta$. This implies $\varepsilon = \eta$. By composing, if necessary, with (a^{-1}, b^{-1}), we may suppose that $\varepsilon = \eta = 1$.

We assume that the words UaU^{-1} and VbV^{-1} are reduced (i.e., there are no cancellations). If $U = V = \varepsilon$, there is nothing to be proved. Suppose that $|U| > 1$ and write $U = W b^n$, with either $W = \varepsilon$ or W ending with an a. Then, $\sigma(ab) = W b^n a b^{-n} W^{-1} V b V^{-1}$; so, $z = \operatorname{tr} \sigma(ab) =$

$\operatorname{tr}(ab^{-n}W^{-1}VbV^{-1}Wb^n)$. This means that, once reduced, $W^{-1}V$ does not contain a. Therefore $W^{-1}V = b^k$. This shows that σ is an inner automorphism.

Now we turn to the study of polynomial maps from $\mathbb{C}^3$ to $\mathbb{C}^3$ which leave λ invariant. Let us set

$$\mathcal{G} = \left\{ \psi \in \mathbb{C}[x,y,z]^3 \;\middle|\; \lambda \circ \psi = \lambda \right\}.$$

Of course $\mathcal{G}$ contains $\left\{ \Phi_\sigma \;\middle|\; \sigma \in \operatorname{Aut}\Gamma \right\}$ (by the corollary to Proposition 10.2.5).

It will be convenient to name some elements of $\operatorname{Aut}\Gamma$:

$$\alpha = (b,a), \quad \beta = (ab, b^{-1}), \quad \gamma = (a, b^{-1}).$$

The corresponding trace maps are (y,x,z), (z,y,x), and $(x,y,xy-z)$.

We also consider the following elements of $\mathcal{G}$: $\rho(x,y,z) = (-x,-y,z)$ and $\theta(x,y,z) = (-x,y,-z)$.

We shall use the following notations: d° stands for the total degree of a polynomial in three variables, and, if $\psi = (\psi_1,\psi_2,\psi_3) \in \mathbb{C}[x,y,z]^3$, $\deg\psi = \sum_{j=1}^{3} \mathrm{d}^\circ \psi_j$. ■

Lemma 10.2.11. *If $\psi = (\psi_1,\psi_2,\psi_3) \in \mathcal{G}$, then, for $j = 1, 2, 3$, we have $\mathrm{d}^\circ \psi_j \geq 1$*

Proof. If, for instance, we had $\psi_3 = c \in \mathbb{C}$, then we would have $\psi_1^2 + \psi_2^2 - c\psi_1\psi_2 = x^2 + y^2 + z^2 - xyz - c^2$, which is impossible, for the left handside is reducible whereas the right handside is not. ■

Lemma 10.2.12. *The set $\mathcal{L} = \left\{ \psi \in \mathcal{G} \;\middle|\; \deg\psi = 3 \right\}$ is the group generated by Φ_α, Φ_β, and ρ.*

Proof. Let us call the variables x_1, x_2, and x_3 instead of x, y, and z. We have $\psi_j = \ell_j + h_j$, where ℓ_j is linear and $h_j \in \mathbb{C}$. We have

$$\sum_{j=1}^{3} (\ell_j + h_j)^2 - \prod_{j=1}^{3} (\ell_j + h_j) = x_1^2 + x_2^2 + x_3^2 - x_1x_2x_3.$$

Looking at terms of degree 3 gives $\ell_j = k_j x_{\tau(j)}$, where τ is a permutation, $k_j \in \mathbb{C}$, and $k_1k_2k_3 = 1$. Looking at quadratic terms gives $h_1 = h_2 = h_3 = 0$ and $k_1^2 = k_2^2 = k_3^2 = 1$. The result follows easily. ■

Lemma 10.2.13. *If $\psi \in \mathcal{G}$ is such that $\deg\psi > 3$, there exits σ in $\langle \alpha,\beta,\gamma \rangle$, the group generated by α, β, and γ, such that $\deg \Phi_\sigma \circ \psi < \deg\psi$.*

Proof. By replacing ψ by $\Phi_\sigma \circ \psi$, where σ is a suitable element of $\langle \alpha, \beta \rangle$, we may suppose $\mathrm{d}^\circ \psi_3 \geq \mathrm{d}^\circ \psi_2 \geq \mathrm{d}^\circ \psi_1$. Moreover, since $\deg \psi > 3$, we have $\mathrm{d}^\circ \psi_3 \geq 2$.

Since $\psi \in \mathcal{G}$, we have

$$\psi_3(\psi_3 - \psi_1\psi_2) + \psi_2^2 + \psi_1^2 = x^2 + y^2 + z^2 - xyz. \tag{10.10}$$

If we had $\mathrm{d}^\circ \psi_3 \neq \mathrm{d}^\circ \psi_1\psi_2$, the degree of the left handside of (10.10) would be

$$\max(2\,\mathrm{d}^\circ \psi_3, \mathrm{d}^\circ \psi_1 + \mathrm{d}^\circ \psi_2 + \mathrm{d}^\circ \psi_3) \geq 4,$$

which is impossible.

We have $\mathrm{d}^\circ \psi_3 = \mathrm{d}^\circ \psi_1\psi_2 > \mathrm{d}^\circ \psi_2$. If we had $\mathrm{d}^\circ(\psi_3 - \psi_1\psi_2) \geq \mathrm{d}^\circ \psi_3$, then we would have $2\,\mathrm{d}^\circ \psi_3 = 3$, which is absurd. Therefore $\mathrm{d}^\circ(\psi_3 - \psi_1\psi_2) < \mathrm{d}^\circ \psi_3$, and $\deg \Phi_\gamma \circ \psi < \deg \psi$. ■

Proposition 10.2.14. *$\mathcal{G}$ is the group generated by Φ_α, Φ_β, Φ_γ, and ρ.*

Proof. Apply Lemma 10.2.13 repeatedly and conclude by using Lemma 10.2.12. ■

Proposition 10.2.15. *For an endomorphism σ of Γ, $Q_\sigma = 1$ if and only if σ is an automorphism.*

Proof. $Q_\sigma = 1$ is equivalent to $\Phi_\sigma \in \mathcal{G}$. Due to Proposition 10.2.14 and to commutation relations

$$\Phi_\alpha \rho \Phi_\alpha = \rho, \quad \Phi_\beta \theta \Phi_\beta = \theta, \quad \Phi_\alpha \theta \Phi_\alpha = \Phi_\beta \rho \Phi_\beta = \rho\theta,$$

there exists $\tau \in \langle \alpha, \beta, \gamma \rangle$ such that $\Phi_\tau \circ \Phi_\sigma \in \langle \rho, \theta \rangle$. Then Lemma 10.2.8 and 10.2.9 and Proposition 10.2.10 show that $\sigma \circ \tau$ is an automorphism. ■

Corollary 10.2.16. $\mathrm{Aut}\, F = \langle \alpha, \beta, \gamma \rangle$.

Proof. If $\sigma \in \mathrm{Aut}\, F$, then $\Phi_\sigma \in \mathcal{G}$. So, there exists $\tau \in \langle \alpha, \beta, \gamma \rangle$ such that $\tau\sigma$ is an inner automorphism or an inner automorphism composed with (a^{-1}, b^{-1}). But, as $(a^{-1}, b^{-1}) = (\alpha\gamma)^2$, the corollary will be proved once we have shown that an inner automorphism is in $\langle \alpha, \beta, \gamma \rangle$. It is easily checked that, if i_W stands for the inner automorphism $V \mapsto WVW^{-1}$, we have $i_a = \alpha\gamma\beta\alpha\gamma\alpha\gamma\beta\alpha\gamma$ and $i_b = \alpha i_a \alpha$. ■

For further properties of trace maps, see [325].

10.2.3 Trace maps (n-letter alphabet)

Three-letter alphabet. If φ is a representation of $\Gamma_\mathcal{A}$ in $SL(2, \mathbb{C})$, we define $[T]\varphi$ to be the following collection of traces:

$$\big(\operatorname{tr}\varphi(A_1), \operatorname{tr}\varphi(A_2), \operatorname{tr}\varphi(A_3), \operatorname{tr}\varphi(A_2A_3), \operatorname{tr}\varphi(A_3A_1),$$
$$\operatorname{tr}\varphi(A_1A_2), \operatorname{tr}\varphi(A_1A_2A_3)\big)$$

and recall the definitions of several polynomials

$$\Lambda(X, Y, z) = z^2 - p(X, Y)z + q(X, Y)$$

where

$$p(X, Y) = x_1y_1 + x_2y_2 + x_3y_3 - x_1x_2x_3$$

and

$$q(X, Y) = x_1^2 + x_2^2 + x_3^2 + y_1^2 + y_2^2 + y_3^2 - x_1x_2y_3 - x_2x_3y_1 - x_3x_1y_2 + y_1y_2y_3 - 4$$

(as previously, X stands for the collection of x's, and similarly for Y).

Let $\mathcal{V}$ be the hyper-surface in $\mathbb{C}^7$ the equation of which is $\Lambda(X, Y, z) = 0$. It can be seen that any point of $\mathcal{V}$ is of the form $[T]\varphi$ (see [325, 324]).

Proposition 10.1.3 (or the corollary to Proposition 10.1.6), shows that, for any $W \in \Gamma_{\mathcal{A}}$, there exits a polynomial P_W such that $\operatorname{tr}\varphi(W) = P_W([T]\varphi)$ for any representation φ. As we have already observed, this polynomial is no longer unique. It is indeed defined up to the addition of a multiple of polynomial Λ (i.e., modulo the ideal $\mathcal{I}$ generated by Λ).

Now, if we have an endomorphism σ of $\Gamma_{\mathcal{A}}$, we choose a collection of polynomials

$$\Phi_\sigma = \big(P_{\sigma(A_1)}, P_{\sigma(A_2)}, P_{\sigma(A_3)}, P_{\sigma(A_2A_3)}, P_{\sigma(A_3A_1)}, P_{\sigma(A_1A_2)}, P_{\sigma(A_1A_2A_3)}\big) .$$

This Φ_σ defines a map from $\mathbb{C}^7$ to $\mathcal{V}$, the restriction of which to $\mathcal{V}$ does not depend on the different choices. Indeed, this is this map from $\mathcal{V}$ to $\mathcal{V}$ which is the trace map and which we call Φ_σ. As previously, $[T](\varphi \circ \sigma) = \Phi_\sigma([T]\varphi)$ for any φ, and $\Phi_{\sigma\circ\tau} = \Phi_\tau \circ \Phi_\sigma$.

In order to show that, as previously, there exists an algebraic sub-manifold Ω of $\mathcal{V}$ which is globally invariant under any Φ_σ, we need the following lemma of which we omit the proof.

Lemma 10.2.17. *Three matrices $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{A}_3$ in $SL(2,\mathbb{C})$ have a common eigendirection if and only if $\Lambda(X, Y, z) = \lambda(x_1, x_2, y_3) = \lambda(x_2, x_3, y_1) = \lambda(x_3, x_1, y_2) = 0$ and $p(X, Y)^2 - 4q(X, Y) = 0$, where*

$$z = \operatorname{tr}\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3,\ X = (x_1, x_2, x_3) = (\operatorname{tr}\mathbf{A}_1, \operatorname{tr}\mathbf{A}_2, \operatorname{tr}\mathbf{A}_3),$$
$$\text{and}\ \ Y = (y_1, y_2, y_3) = (\operatorname{tr}\mathbf{A}_2\mathbf{A}_3, \operatorname{tr}\mathbf{A}_3\mathbf{A}_1, \operatorname{tr}\mathbf{A}_1\mathbf{A}_2).$$

Let Ω be the manifold associated with the ideal $\mathcal{J}$ generated by the polynomials Λ, $\lambda(x_1, x_2, y_3)$, $\lambda(x_2, x_3, y_1)$, $\lambda(x_3, x_1, y_2)$, and $p^2 - 4q$.

Then an argument similar to the one used in the proof of Proposition 10.2.3 shows that Ω is invariant under any Φ_σ.

***n*-letter alphabet.** When n is larger than 3, some complications occur and the situation is less easy to describe. In view of the corollary to Proposition 10.1.6, we need $n(n^2+5)/6$ variables. The ideal $\mathcal{I}$ of relations between these variables is no longer principal. The trace maps take the variety $\mathcal{V}$ of $\mathcal{I}$ to itself. They still leave globally invariant a sub-variety Ω of $\mathcal{V}$ defined by an ideal $\mathcal{J}$ the definition of which comes from expressing that n elements of $SL(2,\mathbb{C})$ have a common eigendirection.

10.3 The case of 3 × 3-matrices

If $\mathbf{M}$ is a 3×3-matrix, the Cayley-Hamilton identity can be written as

$$\mathbf{N}^3-(\operatorname{tr}\mathbf{N})\mathbf{N}^2+\frac{1}{2}\big((\operatorname{tr}\mathbf{N})^2-\operatorname{tr}\mathbf{N}^2\big)\mathbf{N}-\frac{1}{6}(\operatorname{tr}\mathbf{N})^3 \\ +\frac{1}{2}(\operatorname{tr}\mathbf{N})(\operatorname{tr}\mathbf{N}^2)-\frac{1}{3}\operatorname{tr}\mathbf{N}^3=0. \tag{10.11}$$

So, by trilinearization, ones get the formula

$$\begin{aligned}\sum_\eta \mathbf{N}_{\eta(1)}\mathbf{N}_{\eta(2)}\mathbf{N}_{\eta(3)}&-(\operatorname{tr}\mathbf{N}_{\eta(1)})\mathbf{N}_{\eta(2)}\mathbf{N}_{\eta(3)}\\ &+\frac{1}{2}(\operatorname{tr}\mathbf{N}_{\eta(1)})(\operatorname{tr}\mathbf{N}_{\eta(2)})\mathbf{N}_{\eta(3)}\\ &-\frac{1}{2}(\operatorname{tr}\mathbf{N}_{\eta(1)}\mathbf{N}_{\eta(2)})\mathbf{N}_{\eta(3)}-\frac{1}{6}\operatorname{tr}\mathbf{N}_{\eta(1)}\mathbf{N}_{\eta(2)}\mathbf{N}_{\eta(3)}\\ &+\frac{1}{2}(\operatorname{tr}\mathbf{N}_{\eta(1)})\operatorname{tr}\mathbf{N}_{\eta(2)}\mathbf{N}_{\eta(3)}\\ &-\frac{1}{3}(\operatorname{tr}\mathbf{N}_{\eta(1)})(\operatorname{tr}\mathbf{N}_{\eta(2)})(\operatorname{tr}\mathbf{N}_{\eta(3)})=0,\end{aligned}$$

where the summation runs over the permutations η of $\{1,2,3\}$ and where $\mathbf{N}_1$, $\mathbf{N}_2$, and $\mathbf{N}_3$ are arbitrary 3×3-matrices. If in this formula one takes $\mathbf{N}_1=\mathbf{N}_2=\mathbf{M}$ and $\mathbf{N}_3=\mathbf{N}$, one gets

$$\begin{aligned}\mathbf{MNM}+\mathbf{M}^2\mathbf{N}+\mathbf{NM}^2=(\operatorname{tr}\mathbf{N})\mathbf{M}^2&+(\operatorname{tr}\mathbf{M})(\mathbf{NM}+\mathbf{MN})\\ &-(\operatorname{tr}\mathbf{M}\operatorname{tr}\mathbf{N}-\operatorname{tr}\mathbf{MN})\mathbf{M}\\ &-\frac{1}{2}\big((\operatorname{tr}\mathbf{M})^2-\operatorname{tr}\mathbf{M}^2\big)(\mathbf{N}-\operatorname{tr}\mathbf{N})\\ &+\operatorname{tr}\mathbf{M}^2\mathbf{N}-\operatorname{tr}\mathbf{M}\operatorname{tr}\mathbf{MN}.\end{aligned}\tag{10.12}$$

Let $\mathcal{A}=\{a,b\}$ be a two-letter alphabet. Consider the subset $\mathcal{S}_0=\{\varepsilon\}$ of the free monoid $\mathcal{A}^*$, of which the unit ε is the only element. We are going to construct by induction a sequence $\mathcal{S}_n$ of subsets of $\mathcal{A}^n$. Suppose that we know $\mathcal{S}_n$. Then, $\mathcal{S}_{n+1}$ will be the set $\mathcal{S}_n\,a\cup\mathcal{S}_n\,b$ from which the elements ending exactly by a^3, aba, bab, b^3, ab^2a, ba^2b, a^2ba^2, or b^2ab^2 have been removed (by "ending exactly", we mean, for instance, that a^2ba is *not* removed). We get

$$
\begin{aligned}
\mathcal{S}_1 &= \{a, b\} \\
\mathcal{S}_2 &= \{a^2, ab, ba, b^2\} \\
\mathcal{S}_3 &= \{ba^2, b^2a, a^2b, ab^2\} \\
\mathcal{S}_4 &= \{b^2a^2, a^2ba, b^2ab, a^2b^2\} \\
\mathcal{S}_5 &= \{a^2b^2a, b^2a^2b\} \\
\mathcal{S}_6 &= \{b^2a^2ba, a^2b^2ab\} \\
\mathcal{S}_7 &= \emptyset
\end{aligned}
\tag{10.13}
$$

Define

$$\mathcal{S} = \bigcup_{n=0}^{6} \mathcal{S}_n. \tag{10.14}$$

The property of the $\mathcal{S}_n$'s which matters to us is the following. Suppose we are given a representation φ from $\mathcal{A}^*$ in $\mathcal{M}_3(\mathbb{C})$, the ring of 3×3-matrices with complex entries. Then, for any $W \in \mathcal{S}_n$, the matrices $\varphi(Wa)$ and $\varphi(Wb)$ can be expressed as a linear combination of the matrices $\{\varphi(V)\}_{V \in \bigcup_{i=0}^{n+1} \mathcal{S}_i}$, of which the coefficients are polynomials, which can be chosen independent of φ, in the variables $\operatorname{tr}\varphi(a^3)$, $\operatorname{tr}\varphi(b^3)$, and $\{\operatorname{tr}\varphi(V)\}_{V \in \bigcup_{i=0}^{n+1} \mathcal{S}_i}$. The verification of this property is left to the reader. It involves repeated use of (10.11) and (10.12). Also, it is important to notice that the traces of $\varphi(V)$, for $V \in \mathcal{S}_6$, are not involved.

This can be summarized in the following proposition.

Proposition 10.3.1. *Given a word W in $\{a, b\}^*$, there exists polynomials $\{p_V\}_{V \in \mathcal{S}}$ in ten variables with rational coefficients such that, for any representation φ of $\{a, b\}^*$ in $\mathcal{M}_3(\mathbb{C})$, one has*

$$\varphi(W) = \sum_{V \in \mathcal{S}} p_V(T_\varphi)\varphi(V),$$

where

$$
\begin{aligned}
T_\varphi = \Big(&\operatorname{tr}\varphi(a), \operatorname{tr}\varphi(a^2), \operatorname{tr}\varphi(a^3), \\
&\operatorname{tr}\varphi(b), \operatorname{tr}\varphi(b^2), \operatorname{tr}\varphi(b^3), \operatorname{tr}\varphi(ab), \operatorname{tr}\varphi(ab^2), \operatorname{tr}\varphi(a^2b), \operatorname{tr}\varphi(a^2b^2)\Big).
\end{aligned}
$$

Lemma 10.3.2. *There exists a polynomial p in ten variables with rational coefficients such that, for any representation φ of $\mathcal{A}^*$ in $\mathcal{M}_3(\mathbb{C})$, one has*

$$\operatorname{tr}\varphi(b^2a^2ba) + \operatorname{tr}\varphi(a^2b^2ab) = p(T_\varphi).$$

Proof. Multiply on the left identity (10.12) by $\mathbf{N}^2\mathbf{M}$, then use (10.11) and (10.12). ■

Lemma 10.3.3. *There exists a polynomial q such that, for any representation of $\mathcal{A}^*$ in $\mathcal{M}_3(\mathbb{C})$, we have*

$$\operatorname{tr}\varphi(a^2b^2ab)\ \operatorname{tr}\varphi(b^2a^2ba) = q(T_\varphi).$$

Proof. By multiplying (10.12) on the right by $\mathbf{L}$, one gets

$$\begin{aligned}\operatorname{tr}\mathbf{MN}\operatorname{tr}\mathbf{ML} = {} & \operatorname{tr}\mathbf{MNML} + \operatorname{tr}\mathbf{M}^2\mathbf{NL} + \operatorname{tr}\mathbf{M}^2\mathbf{NL} - \operatorname{tr}\mathbf{N}\operatorname{tr}\mathbf{M}^2\mathbf{L}\\ & - \operatorname{tr}\mathbf{L}\operatorname{tr}\mathbf{M}^2\mathbf{N} - (\operatorname{tr}\mathbf{M})(\operatorname{tr}\mathbf{MLN} + \operatorname{tr}\mathbf{MNL})\\ & + \operatorname{tr}\mathbf{M}\operatorname{tr}\mathbf{N}\operatorname{tr}\mathbf{ML} + \operatorname{tr}\mathbf{M}\operatorname{tr}\mathbf{L}\operatorname{tr}\mathbf{MN}\\ & + \frac{1}{2}\Big((\operatorname{tr}\mathbf{M})^2 - \operatorname{tr}\mathbf{M}^2\Big)(\operatorname{tr}\mathbf{NL} - \operatorname{tr}\mathbf{N}\operatorname{tr}\mathbf{L}).\end{aligned}$$

By putting $\mathbf{M} = \varphi(ab)$, $\mathbf{N} = \varphi(a^2b^2)$, and $\mathbf{L} = \varphi(ba^2b)$ in the preceding identity, one gets

$$\begin{aligned}&\operatorname{tr}\varphi(b^2a^2ba) \times \operatorname{tr}\varphi(a^2b^2ab) =\\ &p_1(T_\varphi)\ \operatorname{tr}\varphi(b^2a^2ba) + p_2(T_\varphi)\ \operatorname{tr}\varphi(a^2b^2ab) + p_3(T_\varphi),\end{aligned} \tag{10.15}$$

where p_1, p_2, and p_3 are polynomials in ten variables.

Then, by replacing in (10.15) the matrices $\varphi(a)$ and $\varphi(b)$ by their transpose, one gets

$$\begin{aligned}&\operatorname{tr}\varphi(b^2a^2ba) \times \operatorname{tr}\varphi(a^2b^2ab) =\\ &p_2(T_\varphi)\ \operatorname{tr}\varphi(b^2a^2ba) + p_1(T_\varphi)\ \operatorname{tr}\varphi(a^2b^2ab) + p_3(T_\varphi).\end{aligned} \tag{10.16}$$

By adding (10.15) and (10.16) and taking Lemma 10.3.2 into account, one gets

$$\operatorname{tr}\varphi(b^2a^2ba) \times \operatorname{tr}\varphi(a^2b^2ab) = \frac{1}{2}\Big(p_1(T_\varphi) + p_2(T_\varphi)\Big) + p_3(T_\varphi).$$

Let us consider the following polynomial in eleven variables with rational coefficients

$$\Lambda = \tau^2 - p\tau + q,$$

where p and q are defined in Lemma 10.3.2 and 10.3.3. It results from these lemmas that, for any homomorphism φ of $\Gamma_{\langle a,b\rangle}$ in $\mathcal{M}_3(\mathbb{C})$, the roots of $\Lambda(T_\varphi, \tau)$ are $\operatorname{tr} a^2b^2ab$ and $\operatorname{tr} b^2a^2ba$. ■

Proposition 10.3.4. *The polynomial Λ is irreducible on $\mathbb{C}$.*

Proof. Consider the homomorphism φ of $\Gamma_{\langle a,b\rangle}$ in $\mathcal{M}_3(\mathbb{C})$ so defined

$$\varphi(a) = \begin{pmatrix} 0 & t & 0 \\ 0 & 0 & t^{-1} \\ 1 & 0 & 0 \end{pmatrix} \text{ and } \varphi(b) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

It is straightforward to check that $T\varphi = (0,0,3,0,0,3,0,x,x,x,0)$, $p(T_\varphi) = x^2 - 3$, and $q(T_\varphi) = 2x^3 - 6x^2 + 9$, where $x = 1 + t + t^{-1}$. This gives $(p^2 - 4q)(0,0,3,0,0,3,0,x,x,0) = (x-3)^3(x+1)$.

If Λ were not irreducible, the polynomial $p^2 - 4q$ would be a square and so would be the polynomial $(p^2-4q)(0,0,3,0,0,3,0,x,x,0)$, which it is not.■

Proposition 10.3.5. *For any $W \in \Gamma_{\langle a,b\rangle}$ there exists a polynomial P_W with rational coefficients in eleven variables such that, for any homomorphism φ from $\Gamma_{\langle a,b\rangle}$ to $\mathcal{M}_3(\mathbb{C})$, one has*

$$\operatorname{tr}\varphi(W) = P_W(T_\varphi, \operatorname{tr}\varphi(a^2b^2ab)).$$

This polynomial is unique modulo the principal ideal generated by Λ.

Proof. The existence of P_W results from Proposition 10.3.1 and Lemma 10.3.2. Its uniqueness modulo λ comes from the fact that the derivative of the mapping $\varphi \longmapsto (T_\varphi, \operatorname{tr}(\varphi(a^2b^2ab)))$ is of rank 10 at some φ, for instance for φ such that

$$\varphi(a) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \varphi(b) = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix}.$$

As in the case of 2×2-matrices, one can define a trace map associated with an endomorphism of $\Gamma_{\langle a,b\rangle}$: it is a polynomial map of the variety of $\langle\Lambda\rangle$ into itself. ■

10.4 Comments

Fricke formula and the corollary to Proposition 10.1.4 (Fricke lemma) appear in [175], but were also stated by Vogt in 1889.

Proposition 10.1.3-1 has been stated by Fricke [175] and proved by Horowitz [208]. Since then, it has been rediscovered several times: Allouche and Peyrière [20] for $n = 2$, for general n by Kolář and Nori [246] (although they gave a formula involving a number of traces much larger than $2^n - 1$), and Peyrière et al. [325].

Traina [427, 428] gave an efficient algorithm for computing P_W in the case of a two-letter alphabet; also in this case Wen Z.-X. and Wen Z.-Y. [441] determined the leading term of P_W. Procesi [332] and Razmyslov [348], instead of considering relations between traces only, used more general polynomial identities. This gives simple algorithms for computing polynomials P_W with an arbitrary alphabet. This is this method which is exposed here.

Proposition 10.1.6 and its corollary appear in Avishai, Berend and Glaubman [48].

The trace map appears in Horowitz [209]. It has also been rediscovered a number of times: by Kohmoto et al. [312] in the case of Fibonacci, by Allouche and Peyrière [20] and Peyrière [322] for $n = 2$, by Peyrière et al. [325] for $n > 2$.

Proposition 10.2.3 essentially appears in Horowitz [209]. It has also be rediscovered. Kolář and Ali [245] conjectured it after having used a computer algebra software. The proof given here appears in Peyrière [322]. For a generalization, see [326]

Results in Sec. 10.2.2 can be found in Horowitz [209] and Peyrière et al. [325]. For recent developments, see Wen &Wen [444, 443]: they prove that $Q_\sigma(2, 2, z) \equiv 1$ implies that σ is an automorphism; they show that, on a two-letter alphabet, invertible substitutions (i.e., morphisms of the free monoid which extend as automorphisms of the free group) are generated by three substitutions.

The structure of the ideal $\mathcal{I}$, for a four-letter alphabet, is studied by Whittemore [450] and completely elucidated by Magnus [271] for an arbitrary alphabet. It also results from Magnus [271] that, for a n-letter alphabet, one can use $3n - 3$ variables only in trace maps with the counterpart that Φ_σ is a rational map instead of being a polynomial one. See also [324]

For a study of the quotient ring modulo $\mathcal{I}$ (the ring of Fricke characters) see Magnus [271].

Polynomial identities for $p \times p$-matrices are studied by Procesi [332], Razmyslov [348], and Leron [258]. Wen Z.-X. [440, 441] gives some algorithms for getting such identities. He also constructs a trace map for 3×3-matrices and a two-letter alphabet. This is his derivation which is given in this course.

For basic references on free groups, see [272, 307, 308, 309].