MAPPING CLASSES AND CHARACTER VARIETIES

MAX WEINREICH

These are notes on the paper *Bers and Hénon, Painlevé and Schrödinger* by Serge Cantat [Can09]. The paper builds a bridge from mapping class dynamics on the 2-manifolds $S_{g,n} = S_{1,1}$ and $S_{0,4}$ to algebraic dynamics on certain Markoff-like cubic surfaces M_D, and more generally $M_{A,B,C,D}$ in $\mathbb{C}^3$ and $\mathbb{R}^3$; then this bridge is used to motivate the proofs of some surprising statements phrased purely in terms of complex and real dynamics. Here $S_{g,n}$ is an oriented surface (2-manifold) with n punctures, and for any $A, B, C, D \in \mathbb{C}$, the surface $M_{A,B,C,D}$ is the cubic cut out by

$$x^2 + y^2 + z^2 - xyz = Ax + By + Cz + D.$$

The surface M_D is the special case $M_{0,0,0,D}$.

Contents:

- Section 1 introduces Markoff surfaces and Markoff dynamics.
- Section 2 introduces the action of the modular group on $\mathbb{H}$.
- Section 3 introduces the extended mapping class group, its action on character varieties, and explain how this ties the modular group to Markoff dynamics. This material expands on Section 2.1 of Cantat's paper.
- Section 4 describes complex Markoff dynamics (Section 3 of Cantat).
- Section 5 briefly describes real Markoff dynamics (Section 5 of Cantat).
- Section 6 considers how different kinds of representations determine invariant loci in Markoff surfaces, and studies some interesting orbits related to Teichmüller space (Section 4 of Cantat).
- Section 7 applies real Markoff dynamics to the spectral theory of discrete Schrödinger operators (Section 6 of Cantat).
- Section 8 desceribes the link from Markoff dynamics to Painlevé VI, a differential equation (Section 7 of Cantat).

1. The Markoff Surface

The *Markoff surface* M is defined as the variety in $\mathbb{C}^3$ defined by

$$M := \{(x, y, z) \in \mathbb{C}^3 : x^2 + y^2 + z^2 - 3xyz = 0\}.$$

The surface M first studied by Markoff in 1880 as a part of a study on Diophantine approximation. The Markoff surface is interesting because it admits many nontrivial self-maps that can all be built out of just three generators. Define

$$\begin{aligned} s_x : M &\to M, \\ (x, y, z) &\mapsto (x', y, z), \end{aligned}$$

where (x', y, z) is the unique other point on M with these particular y and z coordinates. There may be points where this point is a double intersection, and these are fixed points of s_x.

In other words, we have intersected the line parallel to the x-axis through (x, y, z) with M. We compute x' by factoring the polynomial $X^2 + y^2 + z^2 - 3Xyz \in \mathbb{C}[X]$ into its two factors and taking the factor besides $X - x$. Since $x + x'$ is the trace $3yz$ by Vieta's formulas, this gives the explicit formula

$$x' = 3yz - x.$$

We similarly define s_y and s_z, and we call these three involutions the *Vieta switches* or *Vieta involutions* on M.

The Vieta switches map $M(\mathbb{R})$ to itself, and in fact map the part of M in the positive orthant $(\mathbb{R}_{>0})^3$ to itself. They generate a group $\mathcal{A}$ of algebraic automorphisms of M that is a free product:

$$\mathrm{Aut}(M) \supset \mathcal{A} := \langle s_x, s_y, s_z \rangle \cong \frac{\mathbb{Z}}{2\mathbb{Z}} * \frac{\mathbb{Z}}{2\mathbb{Z}} * \frac{\mathbb{Z}}{2\mathbb{Z}}.$$

Remark 1.1. It is hardly clear from the definitions that these maps generate the claimed free product. We should worry that some complicated word in s_x, s_y, s_z is somehow the identity. In fact, one can find a fundamental domain in the part M^+ of M in the positive orthant and prove directly that s_x, s_y, s_z freely generate a tessellation of M^+ by ideal triangles; this is explained visually in a companion paper by Cantat-Loray [CL07, Figure 4], or see Section 6.1. Incidentally, this shows that the trivalent graph representing the orbit of $(1, 1, 1)$ is really a tree. (I said this was unknown in my talk, but I was mistaken.)

We also give names to composed pairs:

$$g_x = s_z \circ s_y,$$

$$g_y = s_x \circ s_z,$$

$$g_z = s_y \circ s_x.$$

What are the dynamics of the elements of $\mathcal{A}$?

- The dynamics of one involution, e.g. s_x, are trivial.
- Any word that can be expressed in terms of just two of the three involutions s_x, s_y, s_z preserves a conic fibration, since e.g. g_x has the invariant function z and the fibers of z are conics. On each conic, we have "chess billiards" (see [NT20] for more on this). But this makes it sound more interesting than it is; really, on each conic g_x must be a Mobius transformation.
- After this, things get difficult. Describing the dynamics of more complicated words, like $s_x \circ s_y \circ s_z$, is the subject of Cantat's paper and this talk.

This group $\mathcal{A}$ can be defined in the same way for any of the surfaces $M_{A,B,C,D}$. The only important difference is that the Vieta switch formula may be more complicated, and in particular, the positivity-preserving property of the Vieta switches over $\mathbb{R}$ may be lost.

Remark 1.2. There are automorphisms of M not in $\mathcal{A}$, for instance

$$(x, y, z) \mapsto (x, -y, -z)$$

and

$$(x, y, z) \mapsto (y, x, z).$$

But a theorem of El-Huti shows that $\mathcal{A}$ is finite index in $\mathrm{Aut}(M)$, so any automorphism of M has an iterate in $\mathcal{A}$ [Èl'H74]. So, the dynamics for general automorphisms of M have no essential differences from those in $\mathcal{A}$. Further, the same holds for all surfaces $M_{A,B,C,D}$ we consider, and for a generic surface in this family, we have $\mathrm{Aut}(M_{A,B,C,D}) = \mathcal{A}$.

Remark 1.3. Almost no algebraic varieties admit interesting self-maps. This may sound strange, because obviously $\mathbb{A}^n$ admits lots of self-maps. But a generic variety is of general type, and a theorem of Matsumura says that projective varieties of general type can admit only finitely many dominant rational self-maps [Mat63]. So Markoff surfaces are very special.

2. A tale of two trees

There is another source of trivalent trees in our world. The group $\mathrm{PSL}_2(\mathbb{Z})$ acts on the upper half-plane $\mathbb{H}$ or the open disk $\mathbb{D}$ by Möbius transformations. A fundamental domain for this action on $\mathbb{H}$ is the hyperbolic "ideal triangle" with vertices 0, 1, and ∞. The group $\mathrm{PSL}_2(\mathbb{Z})$ contains elements that move points in the fundamental domain across the three edges of the triangle in well-understood ways, to neighboring triangles; this process tessellates $\mathbb{H}$ with distinct triangles. The dual graph of this tessellation reflects that there are elements of order 2 and 3 in $\mathrm{PSL}_2(\mathbb{Z})$. In fact, $\mathrm{PSL}_2(\mathbb{Z})$ can be expressed as a free product

$$\mathrm{PSL}_2(\mathbb{Z}) \cong \frac{\mathbb{Z}}{2\mathbb{Z}} * \frac{\mathbb{Z}}{3\mathbb{Z}}.$$

We extend this to an action of $\mathrm{PGL}_2(\mathbb{Z})$ on $\mathbb{H}$ by letting elements with determinant -1 act by their complex conjugates (to ensure the image of $\mathbb{H}$ is $\mathbb{H}$). We define a group $\Gamma_2^\pm \subset \mathrm{PGL}_2(\mathbb{Z})$ to be the subgroup of matrices congruent to the identity modulo 2, in the entries. The subgroup $\Gamma_2^\pm$ has a larger fundamental domain – an ideal triangle with vertices $0, 2$, and ∞ – and a tesselation of $\mathbb{H}$ by ideal triangles again, where the dual graph of the triangular tessellation is a trivalent tree. Thus

$$\Gamma_2^\pm \cong \frac{\mathbb{Z}}{2\mathbb{Z}} * \frac{\mathbb{Z}}{2\mathbb{Z}} * \frac{\mathbb{Z}}{2\mathbb{Z}}.$$

Obviously $\mathcal{A} \cong \Gamma_2^\pm$ as groups. The amazing thing is that this isomorphism is not a coincidence.

Theorem 2.1. *There is an explicit isomorphism $\Gamma_2^\pm \to \mathcal{A}$ that arises for topological reasons, that is, via the natural action of the extended mapping class group of $S_{1,1}$ or $S_{0,4}$ on its 2-dimensional character variety.*

The natural action of Theorem 2.1 is defined in the next section.

3. Natural actions of the mapping class group

Expanded versions of many arguments in this section can be found in the Primer [FM12, Chapter 2 and 8]. The only difference is that, in Chapter 2, the Primer uses the standard mapping class group.

The *extended mapping class group* $\mathrm{MCG}^\pm$ of a surface S is defined as the group of self-homeomorphisms of S, modulo isotopy. It differs from the standard mapping class group in that we allow for orientation-reversing maps.

Proposition 3.1. *The fundamental group of the once-punctured torus $S_{1,1}$ is a free group F_2 on two generators. Thus $H_1(S_{1,1}) \cong \mathbb{Z}^2$. The extended mapping class group acts on homology, inducing a map $\mathrm{MCG}^\pm(S_{1,1}) \to \mathrm{GL}_2(\mathbb{Z})$. In fact, this map is an isomorphism*

$$\mathrm{MCG}^\pm(S_{1,1}) \cong \mathrm{GL}_2(\mathbb{Z}).$$

The fact that the map $\mathrm{MCG}^{\pm}(S_{1,1}) \to \mathrm{GL}_2(\mathbb{Z})$ is surjective follows from finding explicit examples, which is not hard: we start with linear maps of $\mathbb{R}^2$ with $\mathbb{Z}$-coefficients, and descend to the torus.

For our application, we actually want more refined information about how this map interacts with the fundamental group. Let $p \in S_{1,1}$ be any base point.

We can write

$$\pi_1(S_{1,1}, p) = \langle \alpha, \beta \rangle \cong F_2,$$

where α, β are paths along the boundary of the diagram; the commutator $[\alpha, \beta]$ goes once around the puncture. To see that α, β freely generate π_1, use the homotopy equivalence between $S_{1,1}$ and $S^1 \vee S^1$.

Given a base point $p \in S_{1,1}$, one would like to have a natural action of $\mathrm{MCG}^{\pm}$ on $\pi_1(S_{1,1}, p)$, hence a map $\mathrm{MCG}^{\pm} \to \mathrm{Aut}(\pi_1(S_{1,1}, p))$. But this doesn't quite work! The issue is that a self-homeomorphism f does not necessarily preserve the base point, at least not in a natural way. We have to choose a path from p to $f(p)$ to get an action on π_1. Choosing different base points or different paths from p to $f(p)$ just conjugates the result, so there is still a well-defined action on $\pi_1(S)/\sim$, and a map

$$r : \mathrm{MCG}^{\pm}(S) \to \mathrm{Out}(\pi_1(S)).$$

Remark 3.2. In fact, the Dehn-Nielsen-Baer theorem says that this map is an isomorphism

$$\mathrm{MCG}^{\pm}(S) \cong \mathrm{Out}(\pi_1(S)).$$

But we don't need this.

The action of $\mathrm{MCG}^{\pm}$ on homology factors through the action on $\pi_1(S)/\sim$, so we can upgrade Proposition 3.1 to

$$\mathrm{MCG}^{\pm}(S_{1,1}) \cong \mathrm{Out}(\pi_1(S_{1,1})) \cong \mathrm{Out}(F_2) \cong \mathrm{GL}_2(\mathbb{Z}). \tag{1}$$

By refining our view of the action of $\mathrm{MCG}^{\pm}$ in this way, we can say exactly what happens to the commutator.

Proposition 3.3. *The natural action of* $\mathrm{MCG}^{\pm}$ *on* $\pi_1(S_{1,1})/\sim$ *either fixes or inverts the class of the commutator* $[\alpha, \beta]$.

Proof. Draw a sufficiently small loop around the puncture. The puncture needed to map to a puncture, and there is only one; so the result is another small loop around the puncture, although perhaps the orientation has reversed. □

Remark 3.4. It may be helpful to write out explicitly what r does on words. Let $m : S_{1,1} \to S_{1,1}$ be a representative of a mapping class. Up to homotopy, it sends the generators α and β of F_2 to words

$$m_*(\alpha) = \alpha^{e_1}\beta^{e_1'}\cdots\alpha^{e_j}\beta^{e_j'},$$
$$m_*(\beta) = \alpha^{f_1}\beta^{f_1'}\cdots\alpha^{f_j}\beta^{f_j'}.$$

The induced matrix of m in $\mathrm{GL}_2(\mathbb{Z})$ is

$$\begin{bmatrix} \sum e_i & \sum f_i \\ \sum e_i' & \sum f_i'. \end{bmatrix}$$

Notice that conjugating m by an element of F_2 won't affect the resulting matrix, hence this operation factors through $\mathrm{Out}(F_2)$.

3.1. **Character varieties.** In this section, we show there is a bijection between (1) a set containing almost all representations of the fundamental group of $S_{1,1}$ in SL_2 up to conjugacy, and (2) the affine space $\mathbb{A}^3$. A representation of $S_{1,1}$ is just a pair (A, B) of matrices in SL_2, and we will classify (most of) them up to simultaneous conjugation in SL_2.

Given a surface S, let $\mathrm{Rep}(S)$ denote the set of representations of the fundamental group of S into $\mathrm{SL}_2(\mathbb{C})$. Typically, one is interested in classifying representations up to conjugacy $\sim$. For finite groups, there are finitely many, classified by character theory; for infinite groups, there may be infinitely many, and there is not necessarily a nice geometric space that parametrizes all these representations exactly. Instead, we can build a variety that parametrizes *most* representations well.

In the case of free groups, representations can be described by specifying the image of each generator. This motivates the next definition.

Definition 3.5. Given a free group F_n of rank at least 2, the *character variety* $\chi(F_n)$ for $\mathrm{SL}_2(\mathbb{C})$ is the quotient variety

$$\chi(F_n) := (\mathrm{SL}_2(\mathbb{C}))^n /\!/ \mathrm{SL}_2(\mathbb{C}),$$

where the action is simultaneous conjugation.

If a surface S has free fundamental group of rank at least 2, we define its *character variety* $\chi(S)$ to be $\chi(\pi_1(S))$.

Proposition 3.6 (Fricke)**.** *We have*

$$\chi(S_{1,1}) = (\mathrm{SL}_2(\mathbb{C}) \times \mathrm{SL}_2(\mathbb{C})) /\!/ \mathrm{SL}_2(\mathbb{C}) \cong \mathbb{A}^3,$$
$$(A, B) \mapsto (x, y, z) := (\operatorname{tr} A, \operatorname{tr} B, \operatorname{tr} AB).$$

We do not quite prove this, but see Section 3.2 for some ideas about how to think about it.

The symbol // means *categorical quotient*, and it is an algebraic concept. What it literally means is that x, y, z generate the ring of conjugation-invariant functions on $\mathrm{SL}_2 \times \mathrm{SL}_2$. Thus every conjugation-invariant algebraic function on $\mathrm{SL}_2 \times \mathrm{SL}_2$, e.g. $\operatorname{tr}(AB^2A)$ or $\operatorname{tr}(ABA^{-1}B^1)$, can be expressed as a polynomial in terms of these three functions $\operatorname{tr} A, \operatorname{tr} B, \operatorname{tr} AB$.

For any surface, we saw that $\mathrm{MCG}^{\pm}$ acts on $\pi_1(S)/\sim$, so $\mathrm{MCG}^{\pm}$ also acts on $\mathrm{Rep}(S)/\sim$. For $S_{1,1}$, after identifying with $\mathbb{A}^3$, we get an action of $\mathrm{MCG}^{\pm}(S_{1,1}) = \mathrm{GL}_2(\mathbb{Z})$ on $\mathbb{A}^3$ by algebraic (polynomial) automorphisms.

Given $m \in \mathrm{MCG}^{\pm}(S_{1,1})$, let $f_m : \mathbb{A}^3 \to \mathbb{A}^3$ be the corresponding map. On the fundamental group, this m preserved $[\alpha, \beta]$ up to conjugacy and inversion. Thus the function $\operatorname{tr} \rho([\alpha, \beta])$ is f_m-invariant. Proposition 3.6 that we can expand the conjugation-invariant function $\operatorname{tr} \rho([\alpha, \beta])$ as a polynomial in x, y, z; in fact we have

$$\operatorname{tr} \rho([\alpha, \beta]) = x^2 + y^2 + z^2 - xyz - 2.$$

Thus the whole mapping class group action admits an invariant fibration by level sets of this function. So for any $D \in \mathbb{C}$, there is an algebraic action of the mapping class group $\mathrm{GL}_2(\mathbb{Z})$ on the surface

$$M_D: \quad x^2 + y^2 + z^2 - xyz = D.$$

We summarize the main properties of this correspondence in the following elaboration of Theorem 2.1.

Theorem 3.7. *The map* $\mathrm{MCG}^{\pm}(S_{1,1}) \to \mathrm{GL}_2(\mathbb{Z}) \to \mathrm{Aut}(M_D)$ *conjugates* Γ_2 *to* $\mathcal{A}$, *modulo the action of*

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

which reduces to the identity on $\mathbb{A}^3$. *Thus there is a* $\mathrm{PGL}_2(\mathbb{Z})$*-action on* $\mathcal{A}$.

The Vieta switches s_x, s_y, s_z *correspond to*

$$\begin{bmatrix} -1 & 0 \\ -2 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & -2 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Remark 3.8. The Markoff surface M corresponds to $D = 0$ after making a minor change of variables.

Remark 3.9. Categorical quotients come up in algebraic geometry even when you can't get a bona fide topological quotient space. You can think of it as meaning $\mathbb{A}^3$ is almost a topological quotient space, but some orbits of the action get mapped to the same point of $\mathbb{A}^3$ (so we can't tell everything apart). It is easy to see that this is happening for the trace map above, because if A and B are both upper-triangular, we can't distinguish them from their diagonal parts with this map. There is a precise description of the orbits that collapse this way in terms of the underlying representations, but we don't need it.

3.2. **Intuition for $\chi(S_{1,1})$.** We are going to talk about representations of infinite discrete groups, but to motivate the idea, we should recall some representation theory of finite groups.

For our purposes, a *representation* of a group G is a group homomorphism $\rho : G \to \mathrm{SL}_N(\mathbb{C})$, where $N \in \mathbb{N}$. If S is a surface, a *representation* of that surface is simply a representation of $\pi_1(S)$. Two representations ρ, ρ' are considered equivalent or *conjugate* if a simultaneous change of basis takes ρ to ρ'.

In the representation theory of finite groups, classifying representations up to conjugacy is done by classifying associated conjugacy-invariant functions called characters. The *character* of a representation ρ of G is the function

$$\begin{aligned} \chi(\rho) : G &\to \mathbb{C}, \\ g &\mapsto \operatorname{tr} \rho(g). \end{aligned}$$

Writing $n := \#G$, It is helpful to think of the trace as being a map from the set $\mathrm{Rep}(G)$ of all G-representations to $\mathbb{C}^n$, that is,

$$\chi : \frac{\mathrm{Rep}(G)}{\sim} \to \mathbb{C}^n.$$

The $\sim$ here refers to conjugacy. Since trace is a conjugacy invariant, so is χ. It turns out that only finitely many χ arise in this way and that they are in 1-1 correspondence with irreducible representations of G up to conjugacy.

To execute this method for countably infinite groups, our first try would be to define χ as a $\mathbb{C}^{\mathbb{N}}$-valued function. However, for the particular case of representations of free groups F_k in $\mathrm{SL}_2(\mathbb{C})$, this χ turns out to have lots of redundacy. We can do much better and only keep track of finitely many traces.

To illustrate, we will work with $F_2 = \langle \alpha, \beta \rangle$. Let $A = \rho(\alpha), B = \rho(\beta)$. The naive trace we defined kept track of

$$(1, \operatorname{tr} A, \operatorname{tr} B, \operatorname{tr} A^2, \operatorname{tr} AB, \operatorname{tr} BA, \operatorname{tr} B^2, \ldots).$$

Visibly, there is a lot of redundancy in this list because e.g. $\operatorname{tr} AB = \operatorname{tr} BA$. We can use *trace identities* to show constructively that in fact all the traces except $1, \operatorname{tr} A, \operatorname{tr} B, \operatorname{tr} AB$ are redundant. These trace identities are:

(1) For any matrices,
$$\operatorname{tr} MN = \operatorname{tr} NM$$
(2) For any invertible matrices,
$$\operatorname{tr} M = \operatorname{tr} N^{-1}MN$$
(3) In SL_2,
$$\operatorname{tr} MN^{-1} + \operatorname{tr} MN = \operatorname{tr} M \operatorname{tr} N$$
(4) In SL_2,
$$\operatorname{tr} M = \operatorname{tr} M^{-1}.$$

Note that if one specializes to SO_2, then these recover cosine identities – pretty neat.

So far, we have indicated why all traces can be reduced to just three. The result of Proposition 3.6, that $\chi(S_{1,1}) \cong \mathbb{A}^3$ via this map, is even stronger; it says that *all* conjugacy-invariant functions can be similarly reduced to these three traces, and that there are no relations between $\operatorname{tr} A, \operatorname{tr} B, \operatorname{tr} AB$.

Even if we couldn't come up with such a thing on our own, there are reasons to expect that variety is a three-dimensional moduli space:

- Since $\dim \mathrm{SL}_2 = 3$, we expect to have a 3-dimensional moduli space
$$(\mathrm{SL}_2 \times \mathrm{SL}_2) /\!/ \mathrm{SL}_2 .$$
(But verifying this takes more work; the analogous count for $\chi(F_1)$ fails because of nontrivial stabilizers in the action.)
- Single matrices $A \in \mathrm{SL}_2$ are classified up to conjugation by their traces, except for non-diagonalizable matrices. This is because the characteristic polynomial of A is $X^2 - (\operatorname{tr} A)X + 1$. Thus there are at least two conjugation-invariant functions on $\mathrm{SL}_2 \times \mathrm{SL}_2$, namely $\operatorname{tr} A$ and $\operatorname{tr} B$. A third is provided by the cross-ratio of the eigenvector directions in $\mathbb{P}^1$ of A and B.

Proposition 3.6 has a cute corollary:

Corollary 3.10. *Given any two matrices $A, B \in \mathrm{SL}_2(\mathbb{C})$, there is a simultaneous change of basis in $\mathrm{SL}_2(\mathbb{C})$ that conjugates A to A^{-1} and B to B^{-1}.*

Proof. The values of x, y, z agree, by the trace identities. This concludes the proof.

A second proof is constructive. Let $\lambda_1, 1/\lambda_1$ be the eigenvalues of A, and let $\lambda_2, 1/\lambda_2$ be the eigenvalues of B; let the corresponding eigenvector directions be v_1, v_1', v_2, v_2'. The change of basis C, acting on $\mathbb{P}^1(\mathbb{C})$, needs to take (v_1, v_1', v_2, v_2') to (v_1', v_1, v_2', v_2). These two 4-tuples have the same cross-ratio, so it's possible.

A third proof is topological, using the hyperelliptic involution on the once-punctured torus (but I didn't quite get this). $\square$

3.3. **The 4-punctured sphere.** All the ideas in the argument for Theorem 3.7 have direct analogues for the 4-punctured sphere $S_{0,4}$. Since the ideas are so similar, we just give a brief overview of what happens for $S_{0,4}$.

Theorem 3.11.

(1) *The fundamental group* $\pi_1(S_{0,4})$ *is generated, with redundancy, by four loops* $\alpha, \beta, \gamma, \delta$ *around the punctures, and*
$$\pi_1(S_{0,4}) = \langle \alpha, \beta, \gamma, \delta : \alpha\beta\gamma\delta = 1 \rangle \cong F_3.$$
(2) *The extended mapping class group is x*
$$\mathrm{MCG}^{\pm}(S_{0,4}) \cong \mathrm{PGL}_2(\mathbb{Z}) \ltimes \left(\frac{\mathbb{Z}}{2\mathbb{Z}} \times \frac{\mathbb{Z}}{2\mathbb{Z}} \right).$$
(3) *The subgroup* $\mathrm{PGL}_2(\mathbb{Z}) \times 1$ *in the semidirect product contains* $\Gamma_2^{\pm}$ *as a subgroup, and the action of* $\Gamma_2^{\pm}$ *on* π_1 *fixes each of* $\alpha, \beta, \gamma, \delta$ *up to conjugacy and inversion.*
(4) *The character variety* $\chi(S_{0,4})$ *is 6-dimensional. There is a map*
$$\begin{aligned} R : \chi(S_{0,4}) &\to \mathbb{C}^7, \\ \rho &\mapsto (a, b, c, d, x, y, z) \\ &:= (\operatorname{tr}\rho(\alpha), \operatorname{tr}\rho(\beta), \operatorname{tr}\rho(\gamma), \operatorname{tr}\rho(\delta), \\ &\qquad \operatorname{tr}\rho(\alpha\beta), \operatorname{tr}\rho(\beta\gamma), \operatorname{tr}\rho(\gamma\alpha)). \end{aligned}$$
There exist $A, B, C, D \in \mathbb{Z}[a, b, c, d]$ *such that the image of* R *is the hypersurface cut out by*
$$x^2 + y^2 + z^2 - xyz = Ax + By + Cz + D.$$
(5) *The functions a, b, c, d are invariant for the* $\Gamma_2^{\pm}$*-action on* $\chi(S_{0,4})$*, hence so are A, B, C, D. The map* $(a, b, c, d) \to (A, B, C, D)$ *surjects to* $\mathbb{C}^4$*. Thus* $\Gamma_2^{\pm}$ *acts algebraically on each surface* $M_{A,B,C,D} \subset \mathbb{C}^3$*, for any choices of* A, B, C, D*.*
(6) *The map* $\Gamma_2^{\pm} \to \mathrm{Aut}(M_{A,B,C,D})$ *identifies* $\Gamma_2^{\pm}$ *with* $\mathcal{A}$*, with the generators' images just as in Theorem 3.7.*

Notes on interesting points of difference in the proof:

(1) This is standard topology. We could have also written a redundant presentation for $\pi_1(S_{1,1})$.
(2) Mapping classes on $S_{0,4}$ can be found by taking mapping classes on the 4-punctured torus $S_{1,4}$ and descending via the hyperelliptic involution. Mapping classes on $S_{1,4}$ can be found by combining translations that stabilize the set of punctures, and the usual torus mapping classes. We have to do this in a way that respects the hyperelliptic involution.
(3) The point here is that this copy of $\Gamma_2^{\pm}$ fixes the punctures rather than stabilizing them as a set.
(4) On one hand, it is a theorem that any trace of an element of F_n can be reduced via trace identities to a polynomial in a certain list of $2^n - 1$ traces; this suggests the map to $\mathbb{C}^7$. (Although this list of traces is not quite standard and I'm not sure this is a useful perspective to have here.) On the other hand, the dimension count comes from standard moduli space theory and shows that these traces can't all be dependent.
(5) The map $(a, b, c, d) \to (A, B, C, D)$ is finite-to-1 but not 1-1.

(6) This is the same as for $S_{1,1}$. Put all the previous parts together and see where the generators of $\Gamma_2^{\pm}$ go.

4. Complex dynamics on Markoff surfaces

All that we have discussed so far is fundamental material in the theory of Markoff-like surfaces. The main theorems of Cantat describe complex and real Markoff dynamics in terms of the corresponding matrices in $\Gamma_2^{\pm}$. In this section, we describe Cantat's results for $\mathbb{C}$.

There is a holomorphic nonvanishing 2-form ω on M that is preserved by all the maps in $\mathcal{A}$, in the sense that if $f \in \mathcal{A}$, then $f^*\omega = \pm\omega$. This means that each $f \in \mathcal{A}$ is area-preserving or area-inverting. A similar statement holds for other dynamical systems on character varieties, e.g. monomial maps.

Instead of working with the affine surface $M \coloneqq M_{A,B,C,D}$, we compactify. Let $\bar{M} = M \cup \Delta$ be the projective compactification of M in $\mathbb{P}^3$, denoting the boundary at infinity by Δ. The set Δ has the structure of an algebraic variety. Working in homogeneous coordinates $[W : X : Y : Z]$ on $\mathbb{P}^3$, the dominant term of the equation for M at infinity is xyz, so the boundary Δ is defined by $W = 0$ and $XYZ = 0$. So Δ is a union of three (complex) lines forming a triangle. The three pairwise intersections will be called the *vertices* of Δ.

The map $f : M \to M$ extends to a birational map $f : \bar{M} \dashrightarrow \bar{M}$. The indeterminacy locus of f is contained in Δ. Its structure depends on what kind of element of $\Gamma_2^{\pm}$ produced f.

From now on, we identify $\Gamma_2^{\pm} \cong \mathcal{A}$ via the $S_{0,4}$-induced isomorphism of Theorem 3.11. (If you like, little is lost in thinking in terms of $S_{1,1}$.)

Each element $T \in \mathrm{PGL}_2(\mathbb{Z})$ is either *elliptic*, *parabolic*, or *hyperbolic*. Elliptic elements are those of finite order. Parabolic elements are those of infinite order that have a nontrivial Jordan block. Hyperbolic elements are those with eigenvalues λ and $1/\lambda$ not equal to 1. We will always assume that λ is the eigenvalue satisfying $|\lambda| > 1$.

Definition 4.1. An element $f \in \mathcal{A}$ is *hyperbolic* if it is induced by a hyperbolic element T of $\Gamma_2^{\pm}$.

There are many kinds of hyperbolicity in dynamics. These hyperbolic Markoff maps f turn out to be "hyperbolic" in many other ways. However, for now, the motivation for the definition is the dictionary from Theorem 3.11:

$\mathrm{MCG}^{\pm}(S_{0,4})$	$\supset$	$\Gamma_2^{\pm}$	$\cong$	$\mathcal{A}$
Finite order	$\rightsquigarrow$	Elliptic	$\leftrightsquigarrow$	words in one Vieta switch
Dehn twist	$\rightsquigarrow$	Parabolic	$\leftrightsquigarrow$	words in two Vieta switches
Pseudo-Anosov	$\rightsquigarrow$	Hyperbolic	$\leftrightsquigarrow$	words in three Vieta switches

Proposition 4.2. *If $f \in \mathcal{A}$ is hyperbolic, then after a birational change of coordinates, the indeterminacy loci* $\operatorname{Ind} f$ *and* $\operatorname{Ind} f^{-1}$ *consist of two distinct vertices v_- and v_+ of Δ, and*

$$f(\Delta \setminus \{v_-\}) = \{v_+\}.$$

In fact, the vertices v_- and v_+ correspond in a rigorous way to the repelling and attracting fixed points of the inducing matrix T on $\partial\mathbb{H}$.

We can speak of "bounded orbits" in terms of the standard Euclidean norm on $\mathbb{C}^3$, restricted to M.

From now on, assume f is already conjugated so that the result of Proposition 4.2 holds. Proposition 4.2 implies that any unbounded f-orbit starting in M is in the attracting basin of v_+. Backwards iterates go towards v_-, by symmetry.

Definition 4.3. The *filled Julia sets* of f are

$$\begin{aligned} K_f^+ &:= \{p \in M : f^n(p) \not\to v_+\}, \\ K_f^- &:= \{p \in M : f^{-n}(p) \not\to v_-\}, \\ K_f &:= K_f^+ \cap K_f^-. \end{aligned}$$

Points in K_f are bounded both forwards and backwards. In particular, all periodic points of f are contained in K_f.

The "complex hyperbolicity" of f is described in the following theorem. The key insight is that many dynamical invariants of f can be read from the spectral radius $|\lambda(T)|$ of the inducing matrix T. We denote this quantity by $|\lambda(f)|$.

Theorem 4.4 (Cantat)**.** *Let $f \in \mathcal{A}$ be hyperbolic, induced by $T \in \Gamma_2^{\pm}$, and in the form of Proposition 4.2.*

(1) *The topological entropy of f on $M(\mathbb{C})$ is*

$$h_{\text{top}}(f) = h_{\text{top}}(T) = |\lambda(f)|.$$

(2) *For all $p \in M \setminus K_f^+$, the attraction rate to v_+ is*

$$\log \|f^n(p)\| \sim |\lambda(f)|^n.$$

(3) *As $n \to \infty$, the set $\mathrm{Per}_n(f)$ of n-periodic points equidistributes to a measure μ_f of maximal entropy with support contained in K_f.*

(4) *Almost all points in $\mathrm{Per}_n(f)$ are saddle-type, and*

$$\#\,\mathrm{Per}_n(f) \approx |\lambda(f)|^n.$$

(5) *There exist plurisubharmonic Green's functions G_f^+ and G_f^-, and*

$$\mu_f = dd^c G_f^+ \wedge dd^c G_f^-.$$

This theorem is mostly proved in Iwasaki-Uehara [IU07].

Remark 4.5. All these facts follow from methods in several complex variable dynamics that were first applied to understand Hénon maps. A Hénon map is a particular kind of polynomial automorphism of $\mathbb{A}^2$ of the form

$$(x, y) \mapsto (y, x + P(y)).$$

These are one of the only types of map on $\mathbb{C}^2$ where we understand the dynamics reasonably well. The key ingredient of the proofs of each statement of Theorem 4.4 is the existence of boundary points v_+ and v_- behaving as in Proposition 4.2. In the Hénon case, Such boundary points also exist on the line at infinity. This is the sole link between Markoff dynamics and Hénon maps.

If one is not familiar with Hénon maps, a looser analogue is "hyperbolic" monomial maps on $\mathbb{C}^2$, such as $\mu(x, y) = (xy, x)$. This map is induced by the hyperbolic matrix

$$T = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \in \mathrm{GL}_2(\mathbb{Z}).$$

Hyperbolic monomial maps preserve the 2-form $dx/x \wedge dy/y$ up to ± 1, and have an invariant set $S^1 \times S^1$ on which the map is a pseudo-Anosov real toral endomorphism. Off that invariant set, and off the coordinate axes, all points approach a unique superattracting fixed point at the boundary at a rate equal to $|\lambda(T)|$. The periodic points are all contained in $S^1 \times S^1$, and they equidistribute relative to Haar measure, and they are all saddles. (One can think of this whole example as dynamics on the character variety of $S_{1,0}$.)

Yet there is a critical difference: the inverse of μ is not a polynomial self-map of $\mathbb{C}^2$. It is more natural to look at this map on $(\mathbb{C}^*)^2$, where it has an algebraic inverse. So monomial maps do not have the nice forwards-backwards time symmetry of Markoff dynamics.

Remark 4.6. In fact, this μ commutes with $(x, y) \mapsto (x^{-1}, y^{-1})$, and taking an appropriate quotient allows one to conjugate the dynamics of μ to Markoff dynamics on a particular Markoff-like surface called the Cayley cubic.

5. Real dynamics on Markoff-type surfaces

Let $M := M_{A,B,C,D}$ be a Markoff-type surface. If all the parameters $A, B, C, D \in \mathbb{R}$, then all the maps in $\mathcal{A}$ restrict to self-maps of the real part $M(\mathbb{R})$. The topology of $M(\mathbb{R})$ depends on the parameters. For convenience we will assume $A, B, C = 0$ and $D > 4$; this guarantees that $M(\mathbb{R})$ is connected, and that is a hypothesis in Cantat's analysis of the real dynamics.

Theorem 5.1 (Cantat)**.** *Assume $M := M_{A,B,C,D}$ has real parameters and that its real part is connected.*

(1) The complex filled Julia set K_f is contained in the real locus $M(R)$. In particular, all periodic points of f are real.
(2) On K_f, the map f is uniformly hyperbolic.

6. Cool orbits (Bers)

Let's summarize the picture so far. There are natural actions of the extended mapping class group $\mathrm{MCG}^{\pm}$ on $\pi_1(S)/\sim$ and on $\chi(S)$. In the case of $S = S_{1,1}$ and $S = S_{0,4}$, the action on $\pi_1(S)/\sim$ stabilizes the set of loops around punctures, up to inversion of loops. This provides extra structure, in the following way. Working with a subgroup $\Gamma_2^{\pm}$ of $\mathrm{MCG}^{\pm}$, the conjugacy classes of these loops are all fixed up to inversion, so there is an invariant fibration of $\chi(S)$ for the action. The invariant fibers are denoted M_D in the case of $S_{1,1}$, and $M_{A,B,C,D}$ in the case of $S_{0,4}$, and the latter family specializes to the former. The automorphisms of $M_{A,B,C,D}$ obtained from this recipe form a set denoted $\mathcal{A}$. Hyperbolic elements of $\mathrm{MCG}^{\pm} \cong \mathrm{PGL}_2(\mathbb{Z})$ give rise to the most interesting elements of $\mathcal{A}$.

We will mostly discuss M_D in this section. The fundamental group $\pi_1(S_{1,1})$ is generated by two loops α and β. The definition of D is

$$D := \operatorname{tr} \rho([\alpha, \beta]) + 2.$$

Since the surfaces $M_{A,B,C,D}$ are moduli spaces of $\mathrm{SL}_2(\mathbb{C})$-representations, we should expect their geometry to reflect the representations they parameterize. For instance, what are the real points of $M_{A,B,C,D}$?

Let M be a surface in this family. Each $\mathbb{R}$-point of M represents a $\mathrm{SL}_2(\mathbb{R})$-representation or a SU_2-representation. Since SU_2 is compact and connected, the corresponding locus in M is compact and connected.

Remark 6.1. The fact that classes of SU_2-representations are defined over $\mathbb{R}$ can be seen in two ways. First, there is the general fact that matrices in SU_2 have real trace. Second, we can consider "field of definition vs. field of moduli". The key observation here is that a pair $(A,B) \in \mathrm{SL}_2(\mathbb{C}) \times \mathrm{SL}_2(\mathbb{C})$ generating an irreducible representation can be simultaneously conjugated to (A^{-1}, B^{-1}); we showed this in Section 3.2. For special unitary representations we thus have

$$(\bar{A}, \bar{B}) = (A^{-1}, B^{-1}) \sim (A, B),$$

so the pair (A,B) is invariant for complex conjugation up to matrix conjugation. It thus descends to a real point of the quotient.

Example 6.2. Choosing $D = 0$, the surface M_0 has five real components. Four of them are unbounded, simply-connected smooth sheets. The fifth is the singleton $\{(0,0,0)\}$, a singularity. Viewing M_0 as a character variety, the point $(0,0,0)$ corresponds to the class of the representation with image equal to the quaternion group. Indeed, the quaternion representation is in SU_2.

Example 6.3. Choosing $D = 2$, the surface M_2 has five real components. They are all smooth. Four are unbounded, and one of these is inside the positive orthant. The fifth is homeomorphic to S^2, consists of special unitary representations' classes, and is denoted M_{SU}.

Example 6.4. Choosing $D = 4$ produces the surface M_4, called the *Cayley cubic.* It appears all over the place; one way it arises is as the quotient of $\mathbb{C}^2$ by $(x,y) \mapsto (x^{-1}, y^{-1})$. The Cayley cubic has four singularities and one connected real component.

Example 6.5. When $D > 4$, the surface M_D is connected, smooth, and homeomorphic to a 4-punctured sphere.

Remark 6.6. The dependence of the topology of the real locus on the parameters A,B,C,D is more complicated and is largely worked out in [BG99].

One reason that the moduli space perspective is very powerful for understanding the dynamics of $\Gamma_2^{\pm}$ on $M_{A,B,C,D}$ is that it furnishes invariant subsets. We should think of elements of $\Gamma_2^{\pm}$ as being automorphism-like operations on representations (speaking loosely): they don't really change the representation, but rather the choice of generators. Thus any generator-independent property of the representation gives rise to an invariant subset of $\Gamma_2^{\pm}$. (This is just informal language, since an automorphism of a representation is something different, literally speaking – it is a self-conjugacy.)

In particular, let $\mathrm{DF} \subset M_D(\mathbb{C})$ be the locus of classes of discrete, faithful representations, and let $M_D^{\mathrm{SU}} \subset M_D(\mathbb{R})$ denote the locus of classes of special unitary representations. Then DF and M_{SU} are invariant subsets. These two subsets contain very different kinds of representations! A typical SU_2-representation is not discrete.

Using a refinement of Theorem 4.4, which characterized complex dynamics of Markoff maps, Cantat constructs interesting orbits of the $\Gamma_2^{\pm}$-action on M. These orbits have the counterintuitive property that they have limit points in both the special unitary set and the discrete faithful set.

Theorem 6.7 (Cantat)**.**

(1) *In M_0, there is a hyperbolic element $f \in \mathcal{A}$ and an f-orbit with both $(0,0,0)$ and a discrete, faithful representation class in its closure.*
(2) *In M_2, there is a $\Gamma_2^{\pm}$-orbit with all of the special unitary locus M_2^{SU} and a discrete faithful representation class in its closure.*

We sketch the proof of (2). The idea is that, for a particular $f \in \mathcal{A}$, namely

$$f = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

there exist saddle fixed points in both M^2_{SU} and in DF. There is a heteroclinic intersection h for this pair of saddle fixed points (this means that the stable and unstable manifolds intersect). Thus following h forward and backward, one gets a point of DF and a point ρ_{SU} of M_2^{SU} in the closure. Then one shows that ρ_{SU} has dense $\mathrm{MCG}^{\pm}$-orbit in M_2^{SU}.

6.1. **Saddle fixed points from Teichmüller space.** Warning: my understanding of this part is hazy!

The proof of Theorem 6.7 (2) is easy to verify but hard to discover. One can verify that the claimed saddle points exist once you know where to look by computing the differential of f as a map on $\mathbb{C}^3$, but one must know where these saddle points are and what kinds of representations they correspond to.

Remarkably, the action of $\mathrm{PGL}_2(\mathbb{Z})$ on the open complex unit disk $\mathbb{D}$ is conjugate to the action of $\mathrm{PGL}_2(\mathbb{Z})$ on M_D. The complex disk $\mathbb{D}$ is the Teichmüller space of $S_{1,1}$ consisting of hyperbolic metrics of a fixed finite volume. Points in $\mathbb{D}$ also correspond to complex structures on $S_{1,1}$ up to an appropriate equivalence relation, in a topologically nice way. Complex structures (and hyperbolic metrics) can be pulled back along mapping classes, so $\Gamma_2^{\pm}$ acts on $\mathbb{D}$, and it turns out that this action is the standard one by Möbius transformations (and their complex conjugates). The conjugating map

$$\mathbb{D} \to M_0$$

is constructed as follows. A point in Teichmüller space endows $S_{1,1}$ with a hyperbolic metric of finite volume. The Uniformization Theorem says that the universal cover of $S_{1,1}$ is $\mathbb{D}$ and that $S_{1,1}$ with the chosen metric is the quotient of $\mathbb{D}$ by a group of isometries. Each path of $\pi_1(S_{1,1})$ gives rise to a $\mathrm{PSL}_2(\mathbb{Z})$-element that moves us correctly along that path in the model $\mathbb{D}$ of the universal cover, and this correspondence is a group homomorphism. That representation on $\pi_1(S_{1,1})$ in $\mathrm{PSL}_2(\mathbb{Z})$ can be lifted in four ways to $\mathrm{SL}_2(\mathbb{R})$, landing us in $\mathbb{A}^3$ in four different ways. Note that the $\mathrm{PSL}_2(\mathbb{Z})$ element that moves you along path $[\alpha, \beta]$ in the universal cover $\mathbb{D}$ must fix exactly one point on the boundary $\partial\mathbb{D}$, a cusp. So $[\alpha, \beta]$ has a parabolic image up to some sign change, so you land in M_0. The four lifts correspond to the four unbounded components in $M_0(\mathbb{R})$.

This correspondence gives us a map

$$\mathbb{D} \to M_0(\mathbb{R})^+$$

that conjugates the mapping class group dynamics on each side. In fact, this map is bijective and real-analytic!

So, one idea for coming up with fixed points is to import them from the action of $\Gamma_2^{\pm}$ on $\mathbb{D}$, as follows. Background knowledge on hyperbolic isometries tells us that every hyperbolic isometry of $\mathbb{D}$ has two fixed points on the boundary $\partial\mathbb{D}$: one attracting, one repelling. The conjugacy from $\mathbb{D}$ to $M_0(\mathbb{R})^+$ would transfer the fixed points on $\partial\mathbb{D}$ in a useful way, if the conjugacy extended to that boundary. Unfortunately, it does not; the fixed points on $\partial\mathbb{D}$ turn out to correspond to indeterminacy on the boundary Δ of $\bar{M}_0$.

Nevertheless, we can do something similar using Bers' parametrization of the quasi-Fuchsian locus, which is a kind of "simultaneous uniformization theorem." Let QF denote the subset of $M_0(\mathbb{C})$ consisting of quasi-Fuchsian representations; these are deformations of $M_0(\mathbb{R})^+$ in the discrete faithful locus, and QF is exactly the interior of the discrete faithful locus. Bers gives a map

$$\mathbb{H} \times -\mathbb{H} \to \mathrm{QF}$$

that, in a certain sense, is natural for mapping class dynamics. This map extends to the boundary by work of Minsky [Min03], and McMullen proved that the corresponding fixed points in $\partial\,\mathrm{QF} \subseteq \mathrm{DF}$ are saddle points [McM96].

7. Discrete Schrödinger Operators

There is a surprising application of Markoff dynamics to Schrödinger operators; see e.g. [Cas86]. We give a brief overview of this connection and Cantat's contribution. Cantat proves a result describing how the spectrum of discrete Schrödinger operators changes along 1-parameter families.

Definition 7.1. Let v be a bounded function $\mathbb{Z} \to \mathbb{C}$. The *discrete Schrödinger operator with potential* v is the bounded linear operator

$$\begin{aligned} &H_v : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z}), \\ &\quad [H_v\psi](i) = \psi(i-1) + \psi(i+1) - v(i)\psi(i). \end{aligned}$$

If $v \equiv 0$, then we have the *discrete Laplacian operator*

$$\begin{aligned} &\Delta : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z}), \\ &\quad [\Delta\psi](i) = \psi(i-1) + \psi(i+1). \end{aligned}$$

Remark 7.2. The normalization of the discrete Laplacian is unfortunate here, but we choose it for consistency with our main reference [DF22]. In our opinion, the discrete Laplacian should be $\Delta - 2$. But this is a minor quibble, since in terms of spectral theory, one can go back and forth between Δ and $\Delta - 2$ easily.

The discrete Laplacian is one of the most fundamental operators in mathematics. Schrödinger operators can be thought of as deformations of the Laplacian. In particular, one can ask about the behavior of Schrodinger operators approaching the Laplacian.

For finite-dimensional vector spaces, the spectrum is the set of eigenvalues, which are roots of the characteristic polynomial, so they vary algebraically in 1-parameter families of maps, determining a *spectral curve*. In the infinite-dimensional setting, the spectrum is some subset of $\mathbb{C}$, and it is much more delicate to say how it varies in 1-parameter families.

However, there is at least one class of discrete Schrödinger operators that we understand well: the real periodic ones, that is, those for which the potential v is n-periodic for some $n \in \mathbb{N}$. In this case, the spectrum H_v is a union of at most n compact real intervals. As H_v moves in 1-parameter families, the endpoints of these intervals move algebraically.

Outside the world of periodic potentials, it is difficult to describe the spectra of Schrödinger operators. A natural place to start is to come up with almost-periodic potentials, whatever that means. One specific kind of almost-periodicity is *quasi-periodicity*, which comes from dynamics.

Theorem 7.3 (Cantat, slogan form)**.** *In certain 1-parameter families of (quasi-periodic) dynamically-defined potentials valued in a two-element set $\{v_1, v_2\} \subset \mathbb{R}$ (obtained from symbolic dynamics on 2 symbols conjugate to an irrational rotation), the Hausdorff dimension* $\operatorname{HDim} \sigma(H_v)$ *varies real-analytically.*

One can ask about the spectrum of a random Schrödinger operator, meaning that the potential v is sampled from some probability space. This is very hard, but we can get more structure by asking for the samples to come from symbolic dynamics with an ergodic measure; these are called "dynamically defined" potentials. In this situation, the spectrum $\sigma(H_v)$ is defined up to a measure 0 subset of the probability space.

7.1. **Details.** A *Banach space* is a vector space equipped with a norm; we will only be interested in the Banach space $\ell^2(\mathbb{Z})$ of square-summable sequences $\mathbb{Z} \to \mathbb{C}$, which is also a Hilbert space. A linear map of infinite-dimensional Banach spaces is called a *linear operator*. A linear operator $H : V \to W$ is *bounded* if there is a constant C such that, for all vectors $v \in V$ of norm 1, the image Hv has norm at most C. The set of bounded operators forms a vector space denoted $\mathcal{B}(V, W)$.

Definition 7.4. Let V be a Banach space, and let I denote the identity operator $V \to V$. The *spectrum* $\sigma(H)$ of a linear operator $H : V \to V$ is the set

$$\{E \in \mathbb{C} : H - EI \text{ has an inverse in } \mathcal{B}(V, V)\}.$$

An *eigenvalue* of H is a value $E \in \mathbb{C}$ such that the *eigenvalue equation* can be solved for some $\psi \in V$:

$$H\psi = E\psi.$$

Try to forget everything you know from finite-dimensional linear algebra. For linear operators, the eigenvalue set is contained in the spectrum, but there can be points in the spectrum that are not eigenvalues, because of the issues of boundedness and square-summability.

Example 7.5. Let $z : \mathbb{Z} \to \mathbb{C}$ be a bounded sequence of complex numbers, and consider the linear operator

$$\begin{aligned} Z : \ell^2(\mathbb{Z}) &\to \ell^2(\mathbb{Z}), \\ [Z\psi](i) &= z(i)\psi(i). \end{aligned}$$

Then the set of eigenvalues is the set $z(\mathbb{Z}) = \{z(i) : i \in \mathbb{Z}\}$. To see this, consider vectors ψ with all 0 except a 1 in position i; this is a $z(i)$-eigenvector. There are no other eigenvalues, as we can see by considering the effect on Z in each position $i \in \mathbb{Z}$. Yet the spectrum $\sigma(Z)$ is the closure $\overline{z(\mathbb{Z})}$. Indeed, for each E outside $z(\mathbb{Z})$, there is an inverse linear operator $(Z - E)^{-1} : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ defined by

$$[Z^{-1}\psi](i) = \frac{1}{z(i) - E}\psi(i).$$

This inverse is a bounded operator if and only if $1/(z(i) - E)$ is bounded as a sequence in $\mathbb{C}$.

The spectrum of a bounded operator H is a nonempty, compact subset of $\mathbb{C}$. In fact $\sigma(H)$ is contained in the closed ball centered at 0 of radius equal to the operator norm of H, which is the infimum of all constants C for H in the definition of bounded operator. If the operator is also self-adjoint, then the spectrum is contained in $\mathbb{R}$, by a Spectral Theorem.

Example 7.6. The discrete Laplacian is bounded, with operator norm 2, is self-adjoint, and has spectrum $\sigma(\Delta) = [-2, 2]$. (The remarks above show $\sigma(\Delta) \subseteq [-2, 2]$.)

We show that Δ has no eigenvalues. Suppose $E \in \mathbb{C}$; let us attempt to solve

$$(\Delta - E)\psi = 0 \tag{2}$$

A *choice of initial conditions* for this difference equation (2) is a pair $\psi_1, \psi_2 \in \mathbb{C}$. Any choice of initial conditions determines a unique sequence $\psi \in \mathbb{C}^{\mathbb{Z}}$ satisfying the difference equation, as follows. The *transfer matrix* T_i depending on E, for each $i \in \mathbb{Z}$, is the matrix

$$T_i = \begin{bmatrix} E & -1 \\ 1 & 0 \end{bmatrix}.$$

A sequence $\psi \in \mathbb{C}^{\mathbb{Z}}$ satisfies (2) if and only if, for each $i \in \mathbb{Z}$, the matrix T_i satisfies

$$T_i \begin{bmatrix} \psi_i \\ \psi_{i-1} \end{bmatrix} = \begin{bmatrix} \psi_{i+1} \\ \psi_i \end{bmatrix}.$$

Thus there is a 1-1 correspondence between the space $\mathbb{C}^2$ of choices of initial conditions and the space of solutions in $\mathbb{C}^{\mathbb{Z}}$. Further, no nontrivial solution in $\mathbb{C}^{\mathbb{Z}}$ decays bi-infinitely, that is, as n goes to ∞ and $-\infty$, because T_i always has an eigenvalue at least 1.

Example 7.6 generalizes readily to any discrete Schrödinger operator, but the transfer matrices T_i do actually depend on i in general.

Example 7.7. If v is an n-periodic potential, then the transfer matrices T_i are n-periodic in i. Choices of initial conditions and transfer matrices are defined as in Example 7.6. Since $(H_v - E)\psi = 0$ is a linear difference equation, acting on the choice of initial conditions by a linear transformation in $\mathrm{GL}_2(\mathbb{C})$ has the effect of transforming all solutions in ψ by the same amount. The operator that shifts ψ by n units to the right commutes with $H_v - E$ (by periodicity), so the growth of the solutions is controlled by the *monodromy matrices* $T_n \dots T_1$. This is essentially the reason why the spectrum varies algebraically in 1-parameter families: the monodromy is on a finite-dimensional space.

Let R be a substitution rule on two letters, that is, a pair of words w_a and w_b in two letters a, b, viewed as a self-map of the set of all finite words in $\{a, b\}$ by the rule $a \mapsto w_a$, $b \mapsto w_b$, with concatenation. If R is induced by a hyperbolic automorphism of F_2, then it is primitive (we will not define this carefully). By basic symbolic dynamics, any primitive rule R has a unique infinite fixed word $w_R : \mathbb{N} \to \{a, b\}$. Let Ω be the ω-limit set of the shift operator applied to any left completion of w_R in the space of bi-infinite words. Then Ω is an invariant set for w_R that admits an ergodic measure.

Remark 7.8. here is a way of framing all this as the dynamics of a quadratic irrational rotation. The most famous example is the Fibonacci substitution, which gives rise to the "Fibonacci Hamiltonian."

Definition 7.9. A *quasi-periodic potential* for rule R is a potential $v : \mathbb{Z} \to \mathbb{R}$ given by any element w of Ω and $v(i) = 0$ if $w(i) = a$, $v(i) = 1$ if $w(i) = b$.

The spectrum of H_v is the same for almost any v obtained from Ω; it is called the *almost sure spectrum* of the rule R. This also applies to any scaling κv of v, where $\kappa \in \mathbb{R}$, so we can define the *almost sure spectrum* of the pair (R, κ). The formal statement of Theorem 7.3 is:

Theorem 7.10. *Fix a primitive substitution rule R. The Hausdorff dimension of the almost sure spectrum of (R, κ) varies analytically in κ.*

The idea of the proof is that, for fixed E and κ, there are only two distinct transfer matrices, because the potential is only two-valued. This provides a $\mathrm{SL}_2(\mathbb{R}(E,k))$-representation of the free group on two generators. This provides a map of a rational curve, the *Schrödinger curve*, parametrized by E to the character variety $\mathbb{A}^3(\mathbb{R}(k))$. One can show that a point is in the spectrum for f if and only if the f-orbit of the corresponding representation is bounded. As κ varies, so does the intersection of the curve with the forward Julia set. (Note that the Markoff surface that one lands in varies with κ.) Then hyperbolic dynamics and thermodynamical formalism tell you everything.

8. Painlevé with rich monodromy

Warning: this is not my area of expertise! I recommend [RR21] and [Wik22] for further information. There is also a nice series of video lectures on foliations by Loray that one can find on YouTube.

The final application of Markoff dynamics is to the monodromy of the Painlevé VI differential equation,

$$\begin{aligned} q'' = \quad & \frac{1}{2}\left(\frac{1}{q}+\frac{1}{q-1}+\frac{1}{q-t}\right)(q')^2+\left(\frac{1}{t}+\frac{1}{t-1}+\frac{1}{q-t}\right)q' \\ & + \frac{q(q-1)(q-t)}{t^2(t-1)^2}\left(\alpha+\beta\frac{q}{t}+\gamma\frac{t-1}{(q-1)^2}+\delta\frac{t(t-1)}{(q-t)^2}\right). \end{aligned}$$

The exact form of the equation does not concern us. What matters are the following aspects of the equation:

- We work over $\mathbb{C}$. The independent (time) variable is t, and the dependent variable is q. It is a second-order nonlinear differential equation that happens to be linear in q''.
- There are four parameters $\alpha, \beta, \gamma, \delta \in \mathbb{C}$. Each choice of parameters is a special case of Painlevé VI.
- The *Painlevé property*: except on a finite subset X of $\mathbb{P}^1_{\mathbb{C}}$ (with variable t), every solution $q(t)$ to Painlevé VI can be meromorphically continued. We have $X = \{0, 1, \infty\}$. There are also moving singularities (i.e. singularities that move with the choice of initial conditions), but these are just poles, so they do not prevent meromorphic continuation.

Remark 8.1. An example of an ODE that does not satisfy the Painlevé property is $dq/dt = \frac{1}{2q}$. This can be solved with separation of variables, and the solutions $q = \pm\sqrt{t+c}$ cannot be meromorphically continued through $t = c$, where c depends on t. Such singularities are called *movable*.

The Painlevé I-VI differential equations are six families of second-order differential equations with the above properties.

One motivating interest in the Painlevé equations is to define new special functions, constructed as solutions of ODEs, that cannot be expressed in terms of classical special functions such as e^x. The Painlevé property is a very restrictive one and provides special functions with a lot of structure.

In the case of first-order ODEs, the only ODEs with the Painlevé property that give new special functions are the Riccati and Weierstrass equations. Painlevé I-VI are a complete set of second-order ODEs for the solutions that are new in second order.

Note also that, for certain members of the Painlevé families (that is, certain choices of parameters), the solutions behave in a very special way; for instance, they might be algebraic. The specialness of the solutions is only a generic feature of the family of differential equations.

8.1. **Connection to Markoff dynamics.** The *naive phase space* of Painlevé VI is $\mathbb{C}^3$, with coordinates t, q, q', where $q' = \partial q/\partial t$. This is because knowledge of q, q', t determines knowledge of q', q'', t, hence gives us a complex vector field on $\mathbb{C}^3$. The vector field defines a meromorphic foliation of $\mathbb{C}^3$ of dimension 1. Given $t \in \mathbb{P}^1 \smallsetminus X$, the fiber consisting of all points (q, q', t) above t is isomorphic to $\mathbb{C}^2$. Any path in $\mathbb{P}^1 \smallsetminus X$ gives rise to a diffeomorphism $\mathbb{C}^2 \to \mathbb{C}^2$ obtained by following the foliation as t goes along that curve. The resulting diffeomorphism is well-defined up to homotopy of the path, so we get a *monodromy representation*

$$\pi_1(\mathbb{P}^1 \smallsetminus X) \to \mathrm{Diff}(\mathbb{C}^2). \tag{3}$$

Amazingly, this action is essentially the same as Markoff dynamics. There is a transcendental map, the Riemann-Hilbert map, that conjugates $\mathbb{C}^2$ to an appropriate $M_{A,B,C,D}$, taking these diffeomorphisms to Markoff dynamics.

The link between Painlevé VI and Markoff dynamics is via the notion of *isomonodromic deformation*.

A *Fuchsian system of ODEs* is a linear complex ODE with a finite set of singularities in $\mathbb{P}^1_{\mathbb{C}}$ with variable t, all of which are simple poles. A second-order Fuchsian system with singularities at $T = \{t_1, \dots, t_n\}$ has phase space $\mathbb{C}^2 \times (\mathbb{P}^1 \smallsetminus T)$. A loop in $\mathbb{P}^1 \smallsetminus T$ gives rise to a self-map of $\mathbb{C}^2$, via the monodromy representation. (Same notion as in (3), but different setting). The linearity hypothesis implies that this self-map is linear, so the monodromy representation lands in $\mathrm{GL}_2(\mathbb{C})$. The *Riemann-Hilbert map* takes a Fuchsian system to its monodromy representation.

If T consists of just four elements, then one can normalize by assuming that some three of the elements are 0,1, and ∞, and letting the last element be called t. As t varies, the monodromy changes, but we can vary the other parameters of the Fuchsian system to force the monodromy to stay the same. Then we obtain an *isomonodromic family of Fuchsian systems*. The Painlevé VI equation arises in nature as the ODE that the parameters of the Fuchsian system must satisfy in an isomonodromic family. Thus each solution of Painlevé VI parameterizes an isomonodromic family of Fuchsian systems.

Mapping classes act on $\mathbb{P}^1 \smallsetminus T$, hence also on each solution of Painlevé VI.

Remark 8.2. There is a nice partial compactification of the Painlevé phase space using Hirzebruch surfaces that resolves some singularities of the system.

References

[BG99] Robert L. Benedetto and William M. Goldman. The topology of the relative character varieties of a quadruply-punctured sphere. *Experiment. Math.*, 8(1):85–103, 1999.

[Can09] Serge Cantat. Bers and Hénon, Painlevé and Schrödinger. *Duke Math. J.*, 149(3):411–460, 2009.

[Cas86] Martin Casdagli. Symbolic dynamics for the renormalization map of a quasiperiodic Schrödinger equation. *Comm. Math. Phys.*, 107(2):295–318, 1986.

[CL07] Serge Cantat and Frank Loray. Holomorphic dynamics, Painlevé VI equation and character varieties, 2007.

[DF22] David Damanik and Jake Fillman. *One-dimensional Schrödinger Operators I. General Theory.* Graduate Studies in Mathematics. 2022.

[Èl'H74] M. H. Èl'-Huti. Cubic surfaces of Markov type. *Mat. Sb. (N.S.)*, 93(135):331–346, 487, 1974.

[FM12] Benson Farb and Dan Margalit. *A primer on mapping class groups*, volume 49 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 2012.

[IU07] Katsunori Iwasaki and Takato Uehara. An ergodic study of Painlevé VI. *Math. Ann.*, 338(2):295–345, 2007.

[Mat63] Hideyuki Matsumura. On algebraic groups of birational transformations. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat. (8)*, 34:151–155, 1963.

[McM96] Curtis T. McMullen. *Renormalization and 3-manifolds which fiber over the circle*, volume 142 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1996.

[Min03] Yair N. Minsky. End invariants and the classification of hyperbolic 3-manifolds. In *Current developments in mathematics, 2002*, pages 181–217. Int. Press, Somerville, MA, 2003.

[NT20] Arnaldo Nogueira and Serge Troubetzkoy. Chess billiards, 2020. `https://arxiv.org/abs/2007.14773`.

[RR21] Julio Rebelo and Roland Roeder. Dynamics of groups of automorphisms of character varieties and Fatou/Julia decomposition for Painlevé 6, 2021. `https://arxiv.org/abs/2104.09256`.

[Wik22] Wikipedia. Isomonodromic deformation, 2022. `https://en.wikipedia.org/wiki/Isomonodromic_deformation`.