Simplicity of Lyapunov spectra: proof of the Zorich–Kontsevich conjecture

by

ARTUR AVILA

Université Pierre et Marie Curie
Paris, France

MARCELO VIANA

Instituto National de Matemática Pura e Aplicada
Rio de Janeiro, Brazil

1. Introduction

A *translation surface* is a compact orientable surface endowed with a flat metric with finitely many conical singularities and a unit parallel vector field. Another equivalent description is through complex analysis: a translation surface is a pair (M, ω), where M is a Riemann surface and ω is an Abelian differential, that is, a complex holomorphic 1-form: the metric is then given by $|\omega|$ and a unit parallel vector field is specified by $\omega \cdot v = i$. We call the flow ϕ_t along the vector field v *vertical flow*. It is defined for all times except for finitely many orbits that meet the singularities in finite time.

The space of all translation surfaces of genus g, modulo isometries preserving the parallel vector fields, is thus identified with the moduli space of Abelian differentials $\mathcal{M}_g$. If we also specify the orders $\varkappa$ of the zeroes of ω, the space of translation surfaces, modulo isometry, is identified with a *stratum* $\mathcal{M}_{g,\varkappa} \subset \mathcal{M}_g$. Each such stratum is a complex affine variety, and so is endowed with a Lebesgue measure class.

1.1. The Zorich phenomenon

Pick a typical (with respect to the Lebesgue measure class) translation surface (M, ω) in $\mathcal{M}_{g,\varkappa}$. Consider segments of orbit $\Gamma_x^T = \{\phi_t(x) : 0 \leqslant t \leqslant e^T\}$, where x is chosen arbitrarily such that the vertical flow is defined for all times, and let γ_x^T be a closed loop obtained by concatenating Γ_x^T with a segment of smallest possible length joining $\phi_{e^T}(x)$ to x. Then $[\gamma_x^T] \in H_1(M)$ is asymptotic to $e^T c$, where $c \in H_1(M) \setminus \{0\}$ is the Schwartzman [Sc] *asymptotic cycle*.

If $g=1$ then this approximation is quite good: the deviation from the line F_1 spanned by the asymptotic cycle is bounded. When $g=2$, one gets a richer picture: $[\gamma_x^T]$ oscillates around F_1 with amplitude roughly $e^{\lambda_2 T}$, where $0<\lambda_2<1$. Moreover, there is an asymptotic isotropic 2-plane F_2: deviations from F_2 are bounded. More generally, in genus $g>1$, there exists an asymptotic Lagrangian flag, that is, a sequence of nested isotropic spaces F_i of dimension i, $1\leqslant i\leqslant g$, and numbers $1>\lambda_2>...>\lambda_g>0$ such that $[\gamma_x^T]$ oscillates around F_i with amplitude roughly $e^{\lambda_{i+1} T}$, $1\leqslant i\leqslant g-1$, and the deviation from F_g is bounded:

$$\limsup_{T\to\infty} \frac{1}{T} \log \operatorname{dist}([\gamma_x^T], F_i) = \lambda_{i+1} \quad \text{for every } 1\leqslant i\leqslant g-1,$$

$$\sup \operatorname{dist}([\gamma_x^T], F_g) < \infty.$$

Moreover, the deviation spectrum $\lambda_2>...>\lambda_g$ is universal: it depends only on the connected component of the stratum to which (M,ω) belongs.

The picture we just described is called the Zorich phenomenon, and was discovered empirically by Zorich [Zo1]. It was shown by Kontsevich and Zorich [Zo1], [Zo2], [KZ] that this picture would follow from a statement about the Lyapunov exponents of the Kontsevich–Zorich cocycle, that we now discuss.

1.2. The Kontsevich–Zorich cocycle

The Kontsevich–Zorich cocycle can be described roughly as follows. It is better to work on (the cotangent bundle of) the Teichmüller space, that is, to consider translation structures on a surface up to isometry isotopic to the identity. The moduli space is obtained by taking the quotient by the modular group. Let (M_T, ω_T) be obtained from (M, ω) by applying the Teichmüller flow for time T, that is, we change the flat metric by contracting (by e^{-T}) the vertical direction, and expanding (by e^T) the orthogonal horizontal direction. The Riemann surface M_T now looks very distorted but it can be brought back to a fundamental domain in the Teichmüller space by applying an element $\Phi_T(M, \omega)$ of the mapping class group. The action of Φ_T on the cohomology $H^1(M)$ is, essentially, the Kontsevich–Zorich cocycle.

The Kontsevich–Zorich cocycle is, thus, a linear cocycle over the Teichmüller flow. It was shown by Masur [Ma] and Veech [Ve1] that the Teichmüller flow restricted to each connected component of strata is ergodic, provided we normalize the area. The cocycle is measurable (integrable), and so it has Lyapunov exponents $\lambda_1\geqslant...\geqslant\lambda_{2g}$. The statement about the cocycle that implies the Zorich phenomenon is that the cocycle has simple Lyapunov spectrum, that is, its Lyapunov exponents are all distinct:

$$\lambda_1 > \lambda_2 > ... > \lambda_{2g-1} > \lambda_{2g} \tag{1.1}$$

(see [Zo4, Theorem 2 and Conditional theorem 4]). The λ_i's, $1 < i \leqslant g$, are the same that appear in the description of the Zorich phenomenon, and the asymptotic flag is related to the Oseledets decomposition. One has $\lambda_i = -\lambda_{2g-i+1}$ for all i, because the cocycle is symplectic. It is easy to see that $\lambda_1 = 1 = -\lambda_{2g}$. The λ_i's are also related to the Lyapunov exponents of the Teichmüller flow on the corresponding connected component of strata: the latter are, exactly,

$$2 \geqslant 1 + \lambda_2 \geqslant \ldots \geqslant 1 + \lambda_g \geqslant 1 = \ldots = 1 \geqslant 1 + \lambda_{g+1} \geqslant \ldots \geqslant 1 + \lambda_{2g-1} \geqslant 0, \qquad (1.2)$$

together with their symmetric values. Zero is a simple exponent, corresponding to the flow direction. There are $\sigma - 1$ exponents equal to 1 (and to -1), arising from the action on relative cycles joining the σ singularities. It is clear that (1.1) is equivalent to saying that all the inequalities in (1.2) are strict.

It follows from the work of Veech [Ve2] that $\lambda_2 < 1$, and so the first and the last inequalities in (1.2) are strict. Thus, the Teichmüller flow is non-uniformly hyperbolic. Together with Kontsevich [Ko], Zorich conjectured that the λ_i's are, indeed, all distinct. They also established formulas for the sums of Lyapunov exponents, but it has not been possible to use them for proving this conjecture. The fundamental work of Forni [Fo] established that $\lambda_g > \lambda_{g+1}$, which is the same as saying that no exponent vanishes. This means that 0 does not belong to the Lyapunov spectrum, that is, the Kontsevich–Zorich cocycle is non-uniformly hyperbolic. Besides giving substantial information on the general case, this result settles the conjecture in the particular case $g=2$, and has also been used to obtain other dynamical properties of translation flows [AF].

1.3. Main result

Previously to the introduction of the Kontsevich–Zorich cocycle, Zorich had already identified and studied a discrete time version, called the Zorich cocycle. Its precise definition will be recalled in §3. While the base dynamics of the Kontsevich–Zorich cocycle is the Teichmüller flow on the space of translation surfaces, the basis of the Zorich cocycle is a renormalization dynamics in the space of interval exchange transformations. The link between interval exchange transformations and translation flows is well known: the first return map to a horizontal cross-section to the vertical flow is an interval exchange transformation, and the translation flow can be reconstructed as a special suspension over this 1-dimensional transformation. Typical translation surfaces in the same connected component of strata are special suspensions over interval exchange transformations whose combinatorics belong to the same *Rauzy class*. The Zorich cocycle is a measurable (integrable) linear cocycle (over the renormalization dynamics) acting on a vector space H

that can be naturally identified with $H_1(M)$. So, it has Lyapunov exponents $\theta_1 \geqslant \ldots \geqslant \theta_{2g}$. The θ_i's are linked to the λ_i's by

$$\lambda_i = \frac{\theta_i}{\theta_1} \quad \text{for all } 1 \leqslant i \leqslant 2g$$

(see [AF, §4.5] and references therein), so that properties of the θ_i's can be deduced from those of the λ_i's, and conversely. The Zorich conjecture (see [Zo3, Conjecture 1] or [Zo4, Conjecture 2]) states that the Lyapunov spectrum of the Zorich cocycle is simple, that is, all the θ_i's are distinct. By the previous discussion, it implies the full picture of the Zorich phenomenon. Here we prove this conjecture.

THEOREM 1.1. *The Zorich cocycle has simple Lyapunov spectrum on every Rauzy class.*

The most important progress in this direction so far, the work of Forni [Fo], was via the Kontsevich–Zorich cocycle. Here we address the Zorich cocycle directly. Though many ideas can be formulated in terms of the Kontsevich–Zorich cocycle, our approach is mostly dynamical and does not involve the extra geometrical and complex analytic structures present in the Teichmüller flow (particularly the $\mathrm{SL}(2, \mathbf{R})$ action on the moduli space) that were crucial in [Fo]. Our arguments contain, in particular, a new proof of Forni's main result.

A somewhat more extended discussion of the Zorich–Kontsevich conjecture, including an announcement of our present results, can be found in [AV1].

Remark 1.2. Besides the Zorich phenomenon, the Lyapunov exponents of the Zorich cocycle are also linked to the behavior of ergodic averages of interval exchange transformations (a first result in this direction is given in [Zo2]) and translation flows or area-preserving flows on surfaces [Fo]. Let us point out that it is now possible to treat the case of interval exchange transformations in a very elegant way, using the results of Marmi–Moussa–Yoccoz [MMY]. A result for translation flows can then be recovered by suspension, which also implies the result for area-preserving flows.

1.4. A geometric motivation for the proof

Our proof of the Zorich conjecture has two distinct parts:

(1) a general criterion for the simplicity of the Lyapunov spectrum of locally constant cocycles;

(2) a combinatorial analysis of Rauzy diagrams to show that the criterion can be applied to the Zorich cocycle on any Rauzy class.

The basic idea of the criterion is that it suffices to find orbits of the base dynamics over which the cocycle exhibits certain forms of behavior, that we call "twisting" and "pinching". Roughly speaking, twisting means that certain families of subspaces are put in general position, and pinching means that a large part of the Grassmannian (consisting of subspaces in general position) is concentrated in a small region, by the action of the cocycle. We will be a bit more precise in a while. We call the cocycle *simple* if it meets both requirements (notice that "simple" really means that the cocycle's behavior is quite rich). Then, according to our criterion, the Lyapunov spectrum is simple.

It should be noted that the orbits on which these types of behavior are observed are very particular and, a priori, correspond to zero measure subsets of the orbits. Nevertheless, they are able to "persuade" almost every orbit to have a simple Lyapunov spectrum. For this, one assumes that the base dynamics is rather chaotic (which is the case for the Teichmüller flow). In a purely random situation, this persuasion mechanism has been understood for quite some time, through the works of Furstenberg [Fu], Guivarc'h and Raugi [GR], and Goldsheid and Margulis [GM]. That this happens also in chaotic, but not random, situations was unveiled by the works of Ledrappier [Le] and Bonatti, Gomez-Mont, Viana [BGV], [BV], [Vi]. The particular situation needed for our arguments is, however, not covered in those works, and was dealt with in our twin paper [AV].

The present paper is devoted to part (2) of the proof. However, to make the paper self-contained, in Appendix 7.1 we also prove an instance of the criterion (1) that covers the present situation.

The basic idea to prove that the Zorich cocycle is "rich", in the sense described above, is to use induction on the complexity. The geometric motivation is more transparent when one thinks in terms of the Kontsevich–Zorich cocycle. As explained, we want to find orbits of the Teichmüller flow inside any connected component of a stratum $\mathcal{C}$ with some given behavior. To this end, we look at orbits that spend a long time near the boundary of $\mathcal{C}$. While there, these orbits pick up the behavior of the boundary dynamics of the Teichmüller flow, which contains the dynamics of the Teichmüller flow restricted to connected components of strata $\mathcal{C}'$ with simpler combinatorics (corresponding to certain ways to degenerate $\mathcal{C}$).

This is easy to make sense of when the stratum $\mathcal{C}$ is not closed in the moduli space, since the whole Teichmüller flow provides a broader ambient dynamics where everything takes place. It is less clear how to formalize the idea when $\mathcal{C}$ is closed (this is the case when there is only one singularity). In this case, the boundary dynamics corresponds to the Teichmüller flow acting on surfaces of smaller genera, and here we will not attempt to describe geometrically what happens. Let us point out that Kontsevich–Zorich [KZ]

considered the inverse of such a degeneration process, that they called "bubbling a handle". It is worth noting that, though they used many different techniques, and as much geometric reasoning as possible, this was one of the few steps that needed to be done using combinatorics of interval exchange transformations (encoded in Rauzy diagrams).

On the other hand, our degeneration process is very simple when viewed in terms of interval exchange transformations: we just make one interval very small. This small interval remains untouched by the renormalization process for a very long time, while the other intervals are acted upon by a degenerate renormalization process. This is what allows us to put in place an inductive argument. To really control the effect, we must choose our small interval very carefully. It is also sometimes useful to choose particular permutations in the Rauzy class we are analyzing. A particularly sophisticated choice is needed when we must change the genus of the underlying translation surface: at this point we use Lemma 20 of Kontsevich–Zorich [KZ], which allowed them to obtain the inverse process of "bubbling a handle".

Let us say a few more words on how the Zorich cocycle will be shown to be simple. The fact that the cocycle is symplectic is important for the arguments. By induction, we show that it acts minimally on the space of Lagrangian flags. This is used to derive that the cocycle is "twisting": certain orbits can be used to put families of subspaces in general position. In the induction, there must be some gain of information at each step, when we must change genus: in this case, this gain regards the action of the cocycle on lines, and it comes from the rather easy fact that this action is minimal.

Also by induction, we show that certain orbits of the cocycle are "pinching": they take a large amount of the Grassmannian and concentrate it into a small region. Here the gain of information when we must change genus has to come from the action on Lagrangian spaces, and it is far from obvious. One can use Forni's theorem [Fo] and, indeed, we did so in a previous version of the arguments. However, the proof we will give is independent of his result, so that our work gives a new proof of Forni's main theorem. Indeed, in the (combinatorial) argument that we will give, the pinching of Lagrangian subspaces comes from orbits that have a pair of zero Lyapunov exponents, but present some parabolic behavior in the central subspace.

1.5. Outline of the paper

In §2 general background on linear and symplectic actions of monoids is given. In §3 we collect well-known material on interval exchange transformations. We introduce the Zorich cocycle and discuss the combinatorics of the Rauzy diagram. The presentation follows [MMY] closely. The matrices appearing in the Zorich cocycle can be studied

in terms of the natural symplectic action of a combinatorial object that we call the Rauzy monoid. We give properties of the projective action of the Rauzy monoids, and discuss some special elements in Rauzy classes. In §4, we introduce the twisting and pinching properties of symplectic actions of monoids, and use them to define the notion of simplicity for monoid actions, which is our basic sufficient condition for simplicity of the Lyapunov spectrum. In §5 and §6 we prove the twisting and pinching properties, and thus simplicity, for the action of the Rauzy monoid. The proof involves a combinatorial analysis of relations between different Rauzy classes, as explained before. In §7, we state the basic sufficient criterion for simplicity of the Lyapunov spectrum (Theorem 7.1) and show that it is satisfied by the Zorich cocycle on any Rauzy class. Theorem 1.1 follows. For completeness, in Appendix 7.1 we give a proof of that sufficient criterion.

2. General background

2.1. Grassmannian structures

Throughout this paper, all vector spaces are finite-dimensional vector spaces over $\mathbf{R}$. The notation $\mathbf{P}H$ always represents the projective space, that is, the space of lines (1-dimensional subspaces) of a vector space H. More generally, $\mathrm{Grass}(k, H)$ will represent the *Grassmannian* of k-planes, $1 \leqslant k \leqslant \dim H - 1$, in the space H. For $0 \leqslant k \leqslant \dim H$, we denote by $\Lambda^k(H)$ the kth exterior product of H. If $F \in \mathrm{Grass}(k, H)$ is spanned by linearly independent vectors $v_1, ..., v_k$, then the exterior product of the v_i's is defined up to multiplication by a scalar. This defines an embedding $\mathrm{Grass}(k, H) \to \mathbf{P}\Lambda^k(H)$.

A *geometric line* in $\Lambda^k H$ is a line that is contained in $\mathrm{Grass}(k, H)$. The duality between $\Lambda^k H$ and $\Lambda^{\dim H - k} H$ allows one to define a *geometric hyperplane* of $\Lambda^k H$ as the dual of a geometric line of $\Lambda^{\dim H - k} H$. A *hyperplane section* is the intersection with $\mathrm{Grass}(k, H)$ of the projectivization of a geometric hyperplane in $\Lambda^k H$. In other words, it is the set of all $F \in \mathrm{Grass}(k, H)$ having non-trivial intersection with a given $E \in \mathrm{Grass}(\dim H - k, H)$. Thus, hyperplane sections are closed subsets of $\mathrm{Grass}(k, H)$ with empty interior. In particular, $\mathrm{Grass}(k, H)$ cannot be written as a finite, or even countable, union of hyperplane sections.

A *linear arrangement* in $\Lambda^k H$ is any finite union of finite intersections of geometric hyperplanes. The intersection of the projectivization of a linear arrangement with $\mathrm{Grass}(k, H)$ will be called a linear arrangement in $\mathrm{Grass}(k, H)$. It will be called non-trivial if it is neither empty nor the whole $\mathrm{Grass}(k, H)$.

The *flag space* $\mathcal{F}(H)$ is the set of all $(F_i)_{i=1}^{\dim H - 1}$, where F_i is a subspace of H of dimension i, and $F_i \subset F_{i+1}$ for all i. It is useful to see $\mathcal{F}(H)$ as a fiber bundle $\mathcal{F}(H) \to \mathbf{P}H$, through the projection $(F_i)_{i=1}^{\dim H - 1} \mapsto F_1$. The fiber over $\lambda \in \mathbf{P}H$ is naturally isomorphic to $\mathcal{F}(H/\lambda)$, via the isomorphism

$$(F_i)_{i=1}^{\dim H - 1} \longmapsto (F_{i+1}/\lambda)_{i=1}^{\dim H - 2}.$$

We are going to state a few simple facts about the family of finite non-empty unions of linear subspaces of any given vector space, and deduce corresponding statements for the family of linear arrangements. The proofs are elementary, and we leave them to the reader.

LEMMA 2.1. *Any finite union W of linear subspaces of some vector space admits a canonical expression $W = \bigcup_{V \in \mathcal{X}} V$ which is minimal in the following sense: if $W = \bigcup_{V \in \mathcal{Y}} V$ is another way to express W as a union of linear subspaces then $\mathcal{X} \subset \mathcal{Y}$.*

It is clear that the family of finite unions of linear subspaces is closed under finite unions and finite intersections. The next statement implies that it is even closed under arbitrary intersections.

LEMMA 2.2. *If $\{V^\alpha : \alpha \in A\}$ is an arbitrary family of finite unions of linear subspaces of some vector space, then $\bigcap_{\alpha \in A} V^\alpha$ coincides with the intersection of the V^α's over a finite subfamily.*

COROLLARY 2.3. *Any totally ordered (under inclusion) family $\{V^\alpha : \alpha \in A\}$ of finite unions of linear subspaces of some vector space is well ordered.*

COROLLARY 2.4. *If V is a finite union of linear subspaces of some vector space and x is an isomorphism of the vector space such that $x \cdot V \subset V$, then $x \cdot V = V$.*

Intersections of geometric hyperplanes are (special) linear subspaces of $\Lambda^k H$, and linear arrangements in $\Lambda^k H$ are finite unions of those linear subspaces. So, the previous results apply, in particular, to the family of linear arrangements in the exterior product. Let us call a linear arrangement S in $\Lambda^k H$ *economical* if it is contained in any other linear arrangement S' such that $\mathbf{P}S \cap \mathrm{Grass}(k, H) = \mathbf{P}S' \cap \mathrm{Grass}(k, H)$. There is a natural bijection between economical linear arrangements of the exterior product and linear arrangements of the Grassmannian that preserves the inclusion order. Hence, the previous results translate immediately to linear arrangements in the Grassmannian. We summarize this in the next corollary.

COROLLARY 2.5. *Both in $\Lambda^k H$ and in $\mathrm{Grass}(k, H)$,*

(1) *the set of all linear arrangements is closed under finite unions and arbitrary intersections;*

(2) *any totally ordered (under inclusion) set of linear arrangements is well ordered;*

(3) *if S is a linear arrangement and x is a linear isomorphism such that $x \cdot S \subset S$, then $x \cdot S = S$.*

2.2. Symplectic spaces

A *symplectic form* on a vector space H is a bilinear form which is antisymmetric, that is, $\omega(u, v) = -\omega(v, u)$ for all u and v, and non-degenerate, that is, for every $u \in H \setminus \{0\}$ there exists $v \in H$ such that $\omega(u, v) \neq 0$. We say that (H, ω) is a *symplectic space*. Notice that $\dim H$ is necessarily an even number $2g$. A *symplectic isomorphism* $A \colon (H, \omega) \to (H', \omega')$ is an isomorphism satisfying $\omega(u, v) = \omega'(A \cdot u, A \cdot v)$. By Darboux's theorem, all symplectic spaces with the same dimension are symplectically isomorphic.

Given a subspace $F \subset H$, the *symplectic orthogonal* of F is the set H_F of all $v \in H$ such that for every $u \in F$ we have $\omega(u, v) = 0$. Its dimension is complementary to the dimension of F, that is, $\dim H_F = 2g - \dim F$. Notice that if two subspaces of H intersect non-trivially and have complementary dimensions, then their symplectic orthogonals also intersect non-trivially.

A subspace $F \subset H$ is called *isotropic* if for every $u, v \in F$ we have $\omega(u, v) = 0$ or, in other words, if F is contained in its symplectic orthogonal. This implies that $\dim F \leqslant g$. A *Lagrangian subspace* is an isotropic subspace of maximal possible dimension g. We represent by $\mathrm{Iso}(k, H) \subset \mathrm{Grass}(k, H)$ the space of isotropic k-planes, for $1 \leqslant k \leqslant g$. This is a closed, hence compact, subset of the Grassmannian. Notice that

$$\mathrm{Iso}(1, H) = \mathrm{Grass}(1, H) = \mathbf{P}H.$$

Given $F \in \mathrm{Iso}(k, H)$, the *symplectic reduction* of H by F is the quotient space

$$H^F = H_F / F.$$

Notice that H^F admits a canonical symplectic form ω^F defined by $\omega^F([u], [v]) = \omega(u, v)$.

Let $\mathcal{L}(H)$ be the space of *Lagrangian flags*, that is, the set of $(F_i)_{i=1}^g$, where F_i is an isotropic subspace of H of dimension i and $F_i \subset F_{i+1}$ for all i. There exists a canonical embedding

$$\mathcal{L}(H) \longrightarrow \mathcal{F}(H),$$
$$(F_i)_{i=1}^g \longmapsto (F_i)_{i=1}^{2g-1},$$

where F_i is the symplectic orthogonal of F_{2g-i}. One can see $\mathcal{L}(H)$ as a fiber bundle over each $\mathrm{Iso}(k,H)$, through the projection

$$\Upsilon_k\colon \mathcal{L}(H) \longrightarrow \mathrm{Iso}(k,H),$$
$$(F_i)_{i=1}^g \longmapsto F_k. \tag{2.1}$$

We will use mostly the particular case $k=1$:

$$\Upsilon = \Upsilon_1\colon \mathcal{L}(H) \longrightarrow \mathbf{P}H,$$
$$(F_i)_{i=1}^g \longmapsto F_1. \tag{2.2}$$

The fiber over a given $\lambda\in\mathbf{P}H$ is naturally isomorphic to $\mathcal{L}(H^\lambda)$, via the isomorphism

$$\Upsilon^{-1}(\lambda) \longrightarrow \mathcal{L}(H^\lambda),$$
$$(F_i)_{i=1}^g \longmapsto (F_{i+1}/\lambda)_{i=1}^{g-1}. \tag{2.3}$$

LEMMA 2.6. *Let* $1\leqslant k\leqslant g$. *Every hyperplane section in* $\mathrm{Grass}(k,H)$ *intersects* $\mathrm{Iso}(k,H)$ *in a non-empty compact subset with empty interior.*

Proof. Let S be a geometric hyperplane dual to some $E\in\mathrm{Grass}(2g-k,H)$. It is easy to see that S meets $\mathrm{Iso}(k,H)$: just pick any $\lambda\in\mathbf{P}H$ contained in E and consider any element F of $\mathrm{Iso}(k,H)$ containing λ. By construction, F is in $S\cap\mathrm{Iso}(k,H)$. Clearly, the intersection is closed in $\mathrm{Iso}(k,H)$, and so it is compact. We are left to show that the complement of $S\cap\mathrm{Iso}(k,H)$ is dense. If $k=1$ the result is clear; in particular, this takes care of the case $g=1$. Next, assuming that the result is true for $(k-1,g-1)$, we deduce that it is also true for (k,g). Indeed, for a dense subset D of lines $\lambda\in\mathbf{P}H$, we have that $\lambda\not\subset E$ and $E\not\subset H_\lambda$. Consequently, $E_\lambda=(E\cap H_\lambda)/\lambda$ is a subspace of H^λ of dimension $2g-k-1$. By the induction hypothesis, a dense subset $\mathcal{D}_\lambda$ of $\mathrm{Iso}(k-1,H^\lambda)$ is not in the hyperplane section S_λ dual to E_λ. Then, in view of (2.1), the set $\mathcal{L}_\lambda=\Upsilon_{k-1}^{-1}(\mathcal{D}_\lambda)$ of Lagrangian flags whose $(k-1)$-dimensional subspace is contained in $\mathcal{D}_\lambda$ is dense in $\mathcal{L}(H^\lambda)$. By (2.3), we may think of $\mathcal{L}_\lambda$ as a subset of $\mathcal{L}(H)$ contained in the fiber over λ. So, taking the union of the $\mathcal{L}_\lambda$'s over all $\lambda\in D$, we obtain a dense subset $\mathcal{L}$ of $\mathcal{L}(H)$. Using (2.1) once more, we conclude that $\Upsilon_k(\mathcal{L})$ is a dense subset of $\mathrm{Iso}(k,H)$. It suffices to prove that no element of this set belongs to S. Let $F\in\Upsilon_k(\mathcal{L})$. By definition, there exists $\lambda\subset F$ with $\lambda\not\subset E$ and $(F/\lambda)\cap((E\cap H_\lambda)/\lambda)=\{0\}$. Since F is contained in H_λ, as it is isotropic, this implies that $F\cap E=\{0\}$, which is precisely what we wanted to prove. $\square$

2.3. Linear actions of monoids

A *monoid* $\mathcal{B}$ is a set endowed with a binary operation that is associative and admits a neutral element (the same axioms as in the definition of group, except for the existence of

inverse). The monoids that interest us most in this context correspond to spaces of loops through a given vertex in a Rauzy diagram, relative to the concatenation operation. A *linear action* of a monoid is an action by isomorphisms of a finite-dimensional vector space H. It induces actions on the projective space $\mathbf{P}H$, the Grassmannian spaces $\mathrm{Grass}(k, H)$, and the flag space $\mathcal{F}(H)$. Given a subspace $F \subset H$, we denote by $\mathcal{B}_F$ the stabilizer of F, that is, the subset of $x \in \mathcal{B}$ such that $x \cdot F = F$.

A *symplectic action* of a monoid is an action by symplectic isomorphisms on a symplectic space (H, ω). It induces actions on $\mathrm{Iso}(k, H)$ and the space of Lagrangian flags $\mathcal{L}(H)$. These actions are compatible with the fiber bundles (2.1), in the sense that they are conjugated to each other by Υ_k. Given $\lambda \in \mathbf{P}H$, the stabilizer $\mathcal{B}_\lambda \subset \mathcal{B}$ of λ acts symplectically on H^λ.

Let $\mathcal{B}$ be a monoid acting by homeomorphisms of a compact space X. By *minimal set* for the action of $\mathcal{B}$ we mean a non-empty closed set $C \subset X$ which is invariant, that is $x \cdot C = C$ for all $x \in \mathcal{B}$, and which has no proper subset with these properties.[1] We say that the action is *minimal* if the whole space is the only minimal set. Detecting minimality of actions of a monoid in fiber bundles can be reduced to detecting minimality for the action in the basis and in the fiber. We will need the following particular case of this idea.

LEMMA 2.7. *Let $\mathcal{B}$ be a monoid acting symplectically on (H, ω). Assume that the action of $\mathcal{B}$ on $\mathbf{P}H$ is minimal and that there exists $\lambda \in \mathbf{P}H$ such that the stabilizer $\mathcal{B}_\lambda \subset \mathcal{B}$ of λ acts minimally on $\mathcal{L}(H^\lambda)$. Then $\mathcal{B}$ acts minimally on $\mathcal{L}(H)$.*

Proof. Let C be a closed invariant subset of $\mathcal{L}(H)$. For $\lambda' \in \mathbf{P}H$, let $C_{\lambda'}$ be the intersection of C with the fiber over λ'. Notice that C_λ may be seen as a closed set invariant for the action of $\mathcal{B}_\lambda$ on $\mathcal{L}(H^\lambda)$. So, C_λ is either empty or the whole fiber. In the first case, let Λ be the set of λ' such that $C_{\lambda'}$ is non-empty. Then Λ is a closed invariant set which is not the whole $\mathbf{P}H$, and so it is empty. This means that C itself is empty. In the second case, let Λ be the set of λ' such that $C_{\lambda'}$ coincides with the whole fiber. This time, Λ is a non-empty invariant set and so it is the whole $\mathbf{P}H$. In other words, in this case $C = \mathcal{L}(H)$. $\square$

2.4. Singular values, Lyapunov exponents

If we supply a vector space H with an inner product, we identify H with the dual H^*, and we also introduce a metric on the Grassmannians: given $F, F' \in \mathrm{Grass}(k, H)$, the

---

[1] If $\mathcal{B}$ is actually a group, a minimal set can be equivalently defined as a non-empty closed set $C \subset X$ such that $\mathcal{B} \cdot x$ is dense in C for every $x \in C$.

distance between F and F' is defined as $\max_{\lambda \subset F} \min_{\lambda' \subset F'} \text{angle}(\lambda, \lambda')$, taken over all lines $\lambda \subset F$ and $\lambda' \subset F'$. We will often consider balls with respect to this metric. All balls will be assumed to have radius less than $\pi/2$.

The inner product also allow us to speak of the *singular values* of a linear isomorphism x acting on H: those are the square roots of the eigenvalues (counted with multiplicity) of the positive self-adjoint operator $x^* x$. We always order them

$$\sigma_1(x) \geqslant \ldots \geqslant \sigma_{\dim H}(x) > 0.$$

A different inner product gives singular values differing from the σ_i's by bounded factors, where the bound is independent of x.

The *Lyapunov exponents* of a linear isomorphism x acting on H are the logarithms of the absolute values of its eigenvalues, counted with multiplicity. We denote and order the Lyapunov exponents of x as

$$\theta_1(x) \geqslant \ldots \geqslant \theta_{\dim H}(x).$$

Alternatively, they can be defined by

$$\theta_i(x) = \lim_{n \to \infty} \frac{1}{n} \log \sigma_i(x^n).$$

Given a linear isomorphism x such that $\sigma_k(x) > \sigma_{k+1}(x)$, we let $E_k^+(x)$ and $E_k^-(x)$ be the orthogonal spaces of dimension k and $\dim H - k$, respectively, such that $E_k^+(x)$ is spanned by the eigenvectors of $x^* x$ with eigenvalue at least $\sigma_k(x)^2$, and $E_k^-(x)$ is spanned by those with eigenvalue at most $\sigma_{k+1}(x)^2$. For any $h = h_k^+ + h_k^- \in E_k^+(x) \oplus E_k^-(x)$,

$$\|x \cdot h\| \geqslant \sigma_k(x) \|h_k^+\|. \tag{2.4}$$

This is useful for the following reason. If $F \in \text{Grass}(k, H)$ is transverse to $E_k^-(x)$, then

$$\|x \cdot h\| \geqslant c \sigma_k(x) \|h\| \quad \text{for all } h \in F, \tag{2.5}$$

where $c > 0$ does not depend on x, but only on the distance between F and the hyperplane section dual to $E_k^-(x)$ in $\text{Grass}(k, H)$.

We collect here several elementary facts from linear algebra that will be useful in the sequel. Related ideas appear in [AV, §7.2].

LEMMA 2.8. *Let x_n be a sequence of linear isomorphisms of H such that*

$$\log \sigma_k(x_n) - \log \sigma_{k+1}(x_n) \to \infty, \quad x_n \cdot E_k^+(x_n) \to E_k^u \quad \text{and} \quad E_k^-(x_n) \to E_k^s,$$

as $n \to \infty$. If K is a compact subset of $\text{Grass}(k, H)$ which does not intersect the hyperplane dual to E_k^s, then $x_n \cdot K \to E_k^u$ as $n \to \infty$.

Proof. Considering a subsequence if necessary, let $F_n \in K$ be such that $x_n \cdot F_n$ converges to some $F' \neq E_k^u$. Take $h_n \in F_n$ with $\|h_n\| = 1$ such that

$$\frac{x_n \cdot h_n}{\|x_n \cdot h_n\|} \to h \notin E_k^u,$$

as $n \to \infty$. By (2.5), there exists $c_1 > 0$ depending only on the distance from K to the hyperplane dual to E_k^s such that

$$\|x_n \cdot h_n\| \geqslant c_1 \sigma_k(x_n).$$

Moreover, since $x_n \cdot E_k^+(x_n) = E_{\dim H - k}^-(x_n^{-1})$, there exists $c_2 > 0$ depending only on the distance from h to E_k^u such that, for n large,

$$1 = \|h_n\| = \|x_n^{-1} x_n \cdot h_n\| \geqslant c_2 \sigma_{\dim H - k}(x_n^{-1}) \|x_n \cdot h_n\|.$$

Consequently, since $\sigma_{\dim H - k}(x_n^{-1}) = \sigma_{k+1}(x_n)^{-1}$,

$$1 \geqslant c_1 c_2 \frac{\sigma_k(x_n)}{\sigma_{k+1}(x_n)}$$

for all n, and this contradicts the hypothesis. $\square$

LEMMA 2.9. *Let x_n be a sequence of linear isomorphisms of H and let $1 \leqslant r \leqslant \dim H - 1$ be such that $\log \sigma_k(x_n) - \log \sigma_{k+1}(x_n) \to \infty$, as $n \to \infty$, for all $1 \leqslant k \leqslant r$. Assume that $x_n \cdot E_k^+(x_n) \to E_k^u$ and $E_k^-(x_n) \to E_k^s$, and that x is a linear isomorphism of H such that $(x \cdot E_k^u) \cap E_k^s = \{0\}$ for all $1 \leqslant k \leqslant r$. Then there exists $C > 0$ such that*

$$|\theta_k(xx_n) - \log \sigma_k(x_n)| < C$$

for all $1 \leqslant k \leqslant r$ and n large.

Proof. Let U_k be an open ball around $x \cdot E_k^u$ such that $F \cap E_k^s = \{0\}$ for every $F \in \overline{U}_k$. By the previous lemma, for n large we have $xx_n \cdot \overline{U}_k \subset U_k$. In particular, there exists $E_{k,n}^u \in U_k$ such that $xx_n \cdot E_{k,n}^u = E_{k,n}^u$ and

$$\|xx_n \cdot h\| \geqslant c_1 \sigma_k(x_n) \|h\| \quad \text{for all } h \in E_{k,n}^u,$$

where c_1 depends only on x and the distance from $\overline{U}_k$ to E_k^s. Consequently,

$$e^{\theta_k(xx_n)} \geqslant c_1 \sigma_k(x_n) \quad \text{for all } 1 \leqslant k \leqslant r. \tag{2.6}$$

Clearly, we also have

$$\prod_{k=1}^{j} e^{\theta_k(xx_n)} \leqslant \prod_{k=1}^{j} \sigma_k(xx_n) \leqslant c_2^{-1} \prod_{k=1}^{j} \sigma_k(x_n) \quad \text{for all } 1 \leqslant j \leqslant r,$$

where $c_2 = c_2(x)$. Using (2.6), we conclude that

$$e^{\theta_j(xx_n)} \leqslant c^{-j} \sigma_j(x_n) \quad \text{for all } 1 \leqslant j \leqslant r,$$

with $c = \min\{c_1, c_2\}$, and the result follows. $\square$

LEMMA 2.10. *Let x_n be a sequence of linear isomorphisms of H. Suppose that there is $F \in \mathrm{Grass}(k, H)$ such that the set $\{F' \in \mathrm{Grass}(k, H) : x_n \cdot F' \to F\}$ is not contained in a hyperplane section. Then $\log \sigma_k(x_n) - \log \sigma_{k+1}(x_n) \to \infty$, and*

$$F = \lim_{n \to \infty} x_n \cdot E_k^+(x_n).$$

Proof. Assume that $\log \sigma_k(x_n) - \log \sigma_{k+1}(x_n)$ is bounded (along some subsequence). Passing to a subsequence, and replacing x_n by $y_n x_n z_n$, with y_n, y_n^{-1}, z_n and z_n^{-1} bounded, we may assume that there exists $l \leqslant k < r$ such that $\sigma_j(x_n) = \sigma_k(x_n)$ if $l \leqslant j \leqslant r$ and $|\log \sigma_j(x_n) - \log \sigma_k(x_n)| \to \infty$ otherwise, and there exists an orthonormal basis $\{e_i\}_{i=1}^{\dim H}$, independent of n, such that $x_n \cdot e_i = \sigma_i(x_n) e_i$ for every i and n. Let E^u, E^c and E^s be the spans of $\{e_i\}_{i=1}^{l-1}$, $\{e_i\}_{i=l}^{r}$ and $\{e_i\}_{i=r+1}^{\dim H}$, respectively. Notice that, for any $F' \in \mathrm{Grass}(k, H)$,

    (1) $(E^c \oplus E^s) + F' = H$ if and only if any limit of $x_n \cdot F'$ contains E^u;

    (2) $F' \cap E^s = \{0\}$ if and only if any limit of $x_n \cdot F'$ is contained in $E^u \oplus E^c$;

    (3) if both (1) and (2) hold, then $x_n \cdot F' \to F$, where $F = E^u \oplus F^c$ and $F^c \subset E^c$ is such that $F^c \oplus E^s = F' \cap (E^c \oplus E^s)$.

Thus, if $x_n \cdot F' \to F$ then F' is contained in the hyperplane section dual to the subspace $G \in \mathrm{Grass}(\dim H - k, H)$ chosen as follows. If $E^u \not\subset F$, then G can be any subspace contained in $E^c \oplus E^s$. If $F \not\subset E^u \oplus E^c$, then G can be any subspace containing E^s. If $E^u \subset F \subset E^c \oplus E^u$, then G can be any subspace containing E^s and such that

$$G \cap F \cap (E^c \oplus E^s) \neq \{0\}.$$

This shows that if $\{F' \in \mathrm{Grass}(k, H) : x_n \cdot F' \to F\}$ is not contained in a hyperplane section, then the difference $\log \sigma_k(x_n) - \log \sigma_{k+1}(x_n) \to \infty$. To conclude, we may assume that $E_k^-(x_n)$ converges to some E_k^s. By Lemma 2.8, if $\{F' \in \mathrm{Grass}(k, H) : x_n \cdot F' \to F\}$ is not contained in the hyperplane section dual to E_k^s, then $F = \lim_{n \to \infty} x_n \cdot E^+(x_n)$. $\qquad\square$

LEMMA 2.11. *Let x_n be a sequence of linear isomorphisms of H, ϱ be a probability measure on $\mathrm{Grass}(k, H)$ and $\varrho^{(n)}$ be the push-forwards of ϱ under x_n.*

    (1) *Assume that ϱ gives zero weight to any hyperplane section. If*

$$\log \sigma_k(x_n) - \log \sigma_{k+1}(x_n) \to \infty \quad \text{and} \quad x_n \cdot E_k^+(x_n) \to E_k^u,$$

then $\varrho^{(n)}$ converges in the weak* topology to a Dirac mass on E_k^u.

    (2) *Assume that ϱ is not supported in a hyperplane section. If $\varrho^{(n)}$ converges in the weak* topology to a Dirac mass on E_k^u, then*

$$\log \sigma_k(x_n) - \log \sigma_{k+1}(x_n) \to \infty \quad \text{and} \quad x_n \cdot E_k^+(x_n) \to E_k^u.$$

Proof. Let us first prove (1). We may assume that $E_k^-(x_n)$ also converges to some E_k^s. Take a compact set K disjoint from the hyperplane section dual to E_k^s, and such that $\varrho(K) > 1 - \varepsilon$. Then $\varrho^{(n)}(x_n \cdot K) > 1 - \varepsilon$ and $x_n \cdot K$ is close to E_k^u for all large n, by Lemma 2.8. This shows that $\varrho^{(n)}$ converges to the Dirac measure on E_k^u.

Now, let us prove (2). The hypothesis implies that, passing to a subsequence of an arbitrary subsequence, $x_n \cdot F \to E_k^u$ for a full measure set, and this set is not contained in a hyperplane section. Using Lemma 2.10, we conclude that

$$\log \sigma_k(x_n) - \log \sigma_{k+1}(x_n) \to \infty \quad \text{and} \quad x_n \cdot E_k^+(x_n) \to E_k^u. \qquad \square$$

LEMMA 2.12. *Let $F^u \in \mathrm{Grass}(k, H)$ and $F^s \in \mathrm{Grass}(\dim H - k, H)$ be orthogonal subspaces. Assume that x is a linear isomorphism of H such that $x \cdot F^u = F^u$ and $x \cdot F^s = F^s$, and there is an open ball U around F^u such that $x \cdot \overline{U} \subset U$. Then $\sigma_k(x) > \sigma_{k+1}(x)$,*

$$E_k^+(x) = F^u \quad and \quad E_k^-(x) = F^s.$$

Proof. Up to composition with orthogonal operators preserving F^u and F^s, we may assume that x is diagonal with respect to some orthonormal basis $e_1, ..., e_{\dim H}$, and that F^u and F^s are spanned by elements of the basis. We may order the elements of the basis so that $x \cdot e_i = \sigma_i(x) e_i$. Let F^u be spanned by $\{e_{i_j}\}_{j=1}^k$. If $\sigma_l(x) \leqslant \sigma_r(x)$ for some $l \in \{i_j\}_{j=1}^k$ and $r \notin \{i_j\}_{j=1}^k$, let F_θ be spanned by the e_s's which are either of the form e_{i_j}, with $i_j \neq l$, or of the form $\cos 2\pi\theta e_l + \sin 2\pi\theta e_r$. If $\sigma_l(x) = \sigma_r(x)$ then $x \cdot F_\theta = F_\theta$ for every θ, and if $\sigma_l(x) < \sigma_r(x)$ then $x^n \cdot F_\theta \to F_{\pi/2} \notin U$ for every θ such that $F_\theta \neq F_0$. In both cases, this gives a contradiction (by considering some $F_\theta \in \partial U$). $\qquad \square$

If (H, ω) is a symplectic space (with $\dim H = 2g$), the choice of an inner product defines an antisymmetric linear isomorphism Ω on H satisfying

$$\omega(\Omega \cdot u, \Omega \cdot v) = \langle u, \Omega \cdot v \rangle.$$

One can always choose the inner product so that Ω is also orthogonal (we call such an inner product *adapted* to ω). If a linear isomorphism x is symplectic, we have, for this particular choice of inner product,

$$\sigma_i(x)\sigma_{2g-i+1}(x) = 1 \quad \text{for all } i = 1, ..., g,$$

and hence

$$\theta_i(x) = -\theta_{2g-i+1}(x) \quad \text{for all } i = 1, ..., g.$$

LEMMA 2.13. *Let x_n be a sequence of symplectic isomorphisms of H such that $\sigma_k(x_n) \to \infty$, $\sigma_k(x_n) > \sigma_{k+1}(x_n)$ and such that $x_n \cdot E_k^+(x_n)$ converges to some space E_k^u. Then E_k^u is isotropic.*

Proof. By (2.4), if $h_n \in x_n \cdot E_k^+(x_n)$ then

$$\|x_n^{-1} \cdot h_n\| \leqslant \sigma_k(x_n)^{-1} \|h_n\|.$$

Thus, if $u_n, v_n \in x_n \cdot E_k^+(x_n)$ are such that $\|u_n\| = \|v_n\| = 1$, then

$$|\omega(u_n, v_n)| = |\omega(x_n^{-1} \cdot u_n, x_n^{-1} \cdot v_n)| \leqslant c^{-1} \sigma_k(x_n)^{-2}.$$

Passing to the limit as $n \to \infty$, we get $\omega(u, v) = 0$ for every $u, v \in E_k^u$. $\qquad\square$

3. Rauzy classes and the Zorich cocycle

3.1. Interval exchange transformations

We follow the presentation of [MMY]. An interval exchange transformation is defined as follows. Let $\mathcal{A}$ be some fixed alphabet on $d \geqslant 2$ symbols. All intervals will be assumed to be closed on the left and open on the right.

- Take an interval $I \subset \mathbf{R}$ and break it into subintervals $\{I_x\}_{x \in \mathcal{A}}$.
- Rearrange the intervals in a new order, via translations, inside I.

Up to translations, we may always assume that the left endpoint of I is 0. Thus the interval exchange transformation is entirely defined by the following data:

(1) the lengths of the intervals $\{I_x\}_{x \in \mathcal{A}}$;

(2) their orders before and after rearranging.

The first is called *length data*, and is given by a vector $\lambda \in \mathbf{R}_+^{\mathcal{A}}$. The second is called *combinatorial data*, and is given by a pair of bijections $\pi = (\pi_0, \pi_1)$ from $\mathcal{A}$ to $\{1, ..., d\}$ (we will sometimes call such a pair of bijections a permutation). We denote the set of all such pairs of bijections by $\mathfrak{S}(\mathcal{A})$. We view a bijection $\mathcal{A} \to \{1, ..., d\}$ as a row where the elements of $\mathcal{A}$ are displayed in the right order. Thus we can see an element of $\mathfrak{S}(\mathcal{A})$ as a pair of rows, the top (corresponding to π_0) and the bottom (corresponding to π_1) of π. The interval exchange transformation associated with this data will be denoted by $f = f(\lambda, \pi)$.

Notice that if the combinatorial data is such that the set of the first k elements in the top and bottom of π coincide for some $1 \leqslant k < d$ then, irrespective of the length data, the interval exchange transformation splits into two simpler transformations. Thus we will consider only combinatorial data for which this does not happen, which we will call *irreducible*. Let $\mathfrak{S}^0(\mathcal{A}) \subset \mathfrak{S}(\mathcal{A})$ be the set of irreducible combinatorial data.

3.2. Translation vector

The positions of the intervals I_x before and after applying the interval exchange transformation differ by translations by

$$\delta_x = \sum_{\pi_1(y) < \pi_1(x)} \lambda_y - \sum_{\pi_0(y) < \pi_0(x)} \lambda_y.$$

We let $\delta = \delta(\lambda, \pi) \in \mathbf{R}^{\mathcal{A}}$ be the *translation vector*, whose coordinates are given by the δ_x's. Notice that the "average translation" $\langle \lambda, \delta \rangle = \sum_{x \in \mathcal{A}} \lambda_x \delta_x$ is zero. We can write the relation between δ and (λ, π) as

$$\delta(\lambda, \pi) = \Omega(\pi) \cdot \lambda,$$

where $\Omega(\pi)$ is a linear operator on $\mathbf{R}^{\mathcal{A}}$ given by

$$\langle \Omega(\pi) \cdot e_x, e_y \rangle = \begin{cases} 1, & \text{if } \pi_0(x) > \pi_0(y) \text{ and } \pi_1(x) < \pi_1(y), \\ -1, & \text{if } \pi_0(x) < \pi_0(y) \text{ and } \pi_1(x) > \pi_1(y), \\ 0, & \text{otherwise} \end{cases} \tag{3.1}$$

(here e_x is the canonical basis of $\mathbf{R}^{\mathcal{A}}$, and the inner product $\langle \cdot, \cdot \rangle$ is the natural one which makes this canonical basis orthonormal). Notice that $\Omega(\pi)$ can be viewed as an antisymmetric matrix with integer entries. Notice also that $\Omega(\pi)$ may not be invertible.

We let $H(\pi) = \Omega(\pi) \cdot \mathbf{R}^{\mathcal{A}}$, which is the space spanned by all possible translation vectors $\delta(\lambda, \pi)$. We define a symplectic form $\omega = \omega_\pi$ on $H(\pi)$ by putting

$$\omega_\pi(\Omega(\pi) \cdot u, \Omega(\pi) \cdot v) = \langle u, \Omega(\pi) \cdot v \rangle.$$

We let $2g(\pi)$ be the dimension of $H(\pi)$, where $g(\pi)$ is called the *genus*. This terminology is justified by the construction described in §3.10.

3.3. Rauzy diagrams and monoids

A *diagram* (or directed graph) consists of two kinds of objects, vertices and (oriented) arrows joining two vertices. Thus, an arrow has a start and an end. A *path* of length $m \geqslant 0$ is a sequence $v_0, ..., v_m$ of vertices and a sequence of arrows $a_1, ..., a_m$ such that a_i starts at v_{i-1} and ends in v_i. If γ_1 and γ_2 are paths such that the end of γ_1 is the start of γ_2, their concatenation is also a path, denoted by $\gamma_1 \gamma_2$. The set of all paths starting and ending at a given vertex v is a monoid with the operation of concatenation. We can identify paths of length zero with vertices and paths of length one with arrows.

Given $\pi \in \mathfrak{S}^0(\mathcal{A})$, we consider two operations. Let x and y be the last elements of the top and bottom rows. The *top* operation keeps the top row; on the other hand, it takes y

and inserts it back in the bottom immediately to the right of the position occupied by x. When applying this operation to π, we will say that x *wins* and y *loses*. The *bottom* operation is defined in a dual way, just interchanging the words top and bottom, and the roles of x and y. In this case we say that y *wins* and x *loses*. Notice that both operations preserve the first elements of both the top and the bottom row.

It is easy to see that those operations give bijections of $\mathfrak{S}^0(\mathcal{A})$. The *Rauzy diagram* associated with $\mathcal{A}$ has the elements of $\mathfrak{S}^0(\mathcal{A})$ as its vertices, and its arrows join each vertex to the ones obtained from it through either of the operations we just described. So, every vertex is the start and end of two arrows, one top and one bottom. Thus, every arrow has a start, an end, a type (top/bottom), a winner and a loser.

The set of all paths is denoted by $\Pi(\mathcal{A})$. The orbit of π under the monoid generated by the actions of the top and bottom operations on $\mathfrak{S}^0(\mathcal{A})$ will be called the *Rauzy class* of π, and denoted by $\mathfrak{R}(\pi)$. The set of all paths inside a given Rauzy class will be denoted by $\Pi(\mathfrak{R})$. The set of all paths that begin and end at $\pi \in \mathfrak{S}^0(\mathcal{A})$ will be denoted by $\Pi(\pi)$. We call $\Pi(\pi)$ a *Rauzy monoid*.

In Figure 1, we present the restriction of the Rauzy diagram to the Rauzy class of the permutation

$$\pi = \begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix}.$$

The Rauzy class has 7 elements, and there are 14 arrows in the diagram. One element of the monoid $\Pi(\pi)$ is, for instance, the closed path

$$\gamma \colon \begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix} \longrightarrow \begin{pmatrix} A & B & C & D \\ D & A & C & B \end{pmatrix} \longrightarrow \begin{pmatrix} A & B & C & D \\ D & B & A & C \end{pmatrix} \longrightarrow \begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix}.$$

3.4. Rauzy induction

Let $\mathfrak{R} \subset \mathfrak{S}^0(\mathcal{A})$ be a Rauzy class, and define $\Delta_\mathfrak{R}^0 = \mathbf{R}_+^\mathcal{A} \times \mathfrak{R}$. Given (λ, π) in $\Delta_\mathfrak{R}^0$, we say that we can apply Rauzy induction to (λ, π) if $\lambda_x \neq \lambda_y$, where $x, y \in \mathcal{A}$ are the last elements of the top and bottom rows of π, respectively. Then we define (λ', π') as follows:

(1) let $\gamma = \gamma(\lambda, \pi)$ be a top or bottom arrow starting at π in the Rauzy diagram, according to whether $\lambda_x > \lambda_y$ or $\lambda_y > \lambda_x$;

(2) let

$$\lambda_z' = \begin{cases} \lambda_z, & \text{if } z \text{ is not the winner of } \gamma, \\ |\lambda_x - \lambda_y|, & \text{if } z \text{ is the winner of } \gamma; \end{cases}$$

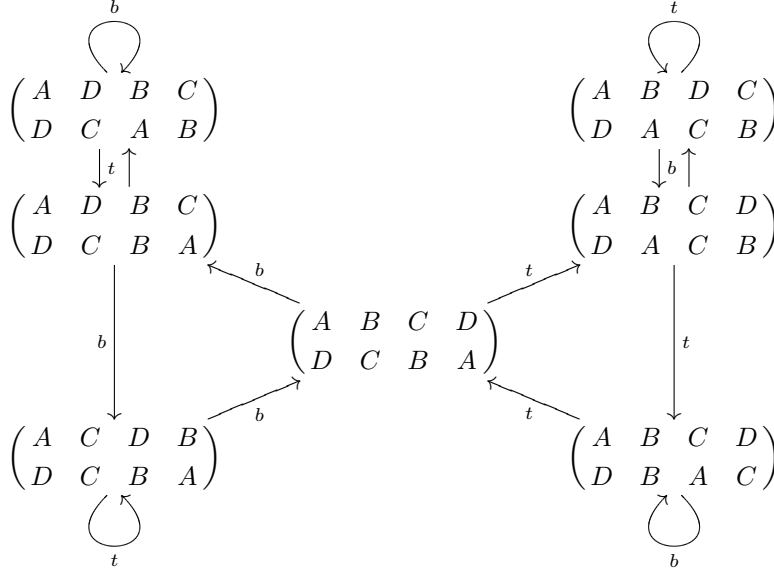(3) let π' be the end of γ.

Figure 1. Rauzy class and Rauzy diagram.

We say that (λ', π') is obtained from (λ, π) by applying *Rauzy induction*, of type *top* or *bottom* depending on whether the type of γ is top or bottom. We have that $\pi' \in \mathfrak{S}^0(\mathcal{A})$ and $\lambda' \in \mathbf{R}_+^{\mathcal{A}}$. The interval exchange transformations $f: I \to I$ and $f': I' \to I'$ specified by the data (λ, π) and (λ', π') are related as follows. The map f' is the first return map of f to a subinterval of I, obtained by cutting from I a subinterval with the same right endpoint and of length λ_z, where z is the loser of γ. The map $Q_R: (\lambda, \pi) \to (\lambda', \pi')$ is called the *Rauzy induction map*. Its domain of definition, the set of all $(\lambda, \pi) \in \Delta_{\mathfrak{R}}^0$ such that $\lambda_x \neq \lambda_y$, will be denoted by $\Delta_{\mathfrak{R}}^1$.

3.5. Relation between translation vectors

Let us calculate the relation between the translation vectors $\delta = \delta(\lambda, \pi)$ and $\delta' = \delta(\lambda', \pi')$. Let γ be the arrow that starts at π and ends at π'. We have $\delta'_z = \delta_z$ if z is not the loser of γ, and $\delta'_z = \delta_x + \delta_y$ if z is the loser of γ (we continue to denote by x and y the last symbols in the top and bottom rows of π, respectively). In other words, we can write

$$\delta(\lambda', \pi') = \Theta(\gamma) \cdot \delta(\lambda, \pi),$$

where $\Theta(\gamma)$ is the linear operator of $\mathbf{R}^{\mathcal{A}}$ defined by $\Theta(\gamma){\cdot}e_z=e_x+e_y$ if z is the winner of γ, and $\Theta(\gamma){\cdot}e_z=e_z$ otherwise. Notice that we also have

$$\lambda = \Theta(\gamma)^*{\cdot}\lambda',$$

where the adjoint $\Theta(\gamma)^*$ is taken with respect to the natural inner product on $\mathbf{R}^{\mathcal{A}}$ that renders the canonical basis orthonormal. Consequently, if $h\in\mathbf{R}^{\mathcal{A}}$ and $h'=\Theta(\gamma){\cdot}h$ then $\langle\lambda,h\rangle=\langle\lambda',h'\rangle$.

Since $H(\pi)$ and $H(\pi')$ are spanned by possible translation vectors, $\Theta(\gamma){\cdot}H(\pi)= H(\pi')$, and so the dimension of $H(\pi)$ only depends on the Rauzy class of π. One can check that

$$\Theta(\gamma)\Omega(\pi)\Theta(\gamma)^* = \Omega(\pi'), \tag{3.2}$$

which implies that $\Theta(\gamma)\colon(H(\pi),\omega_\pi){\to}(H(\pi'),\omega_{\pi'})$ is a symplectic isomorphism. Notice that $\Theta(\gamma)$ can be viewed as a matrix with non-negative integer entries and determinant 1. We extend the definition of Θ from arrows to paths in $\Pi(\mathcal{A})$ in the natural way, setting $\Theta(\gamma_1\gamma_2)=\Theta(\gamma_2)\Theta(\gamma_1)$. In this way Θ induces a representation on $\mathrm{SL}(\mathcal{A},\mathbf{Z})$ of the Rauzy monoids $\Pi(\pi)\subset\Pi(\mathcal{A})$, $\pi\in\mathfrak{S}^0(\mathcal{A})$.

3.6. Iterates of Rauzy induction

The connected components of $\Delta_{\mathfrak{R}}^0=\mathbf{R}_+^{\mathcal{A}}\times\mathfrak{R}$ are naturally labelled by the elements of $\mathfrak{R}$ or, in other words, by length-0 paths in $\Pi(\mathfrak{R})$. The connected components of the domain $\Delta_{\mathfrak{R}}^1$ of the induction map Q_R are naturally labelled by arrows, that is, length-1 paths in $\Pi(\mathfrak{R})$. One easily checks that each connected component of $\Delta_{\mathfrak{R}}^1$ is mapped bijectively to some connected component of $\Delta_{\mathfrak{R}}^0$. Now let $\Delta_{\mathfrak{R}}^n$ be the domain of Q_R^n, for each $n{\geqslant}2$. The connected components of $\Delta_{\mathfrak{R}}^n$ are naturally labelled by length-n paths in $\Pi(\mathfrak{R})$: if γ is obtained by concatenation of arrows $\gamma_1,...,\gamma_n$, then

$$\Delta_\gamma = \{x\in\Delta^0(\lambda,\pi):Q_R^{k-1}(x)\in\Delta_{\gamma_k},1\leqslant k\leqslant n\}.$$

If γ is a length-n path in $\Pi(\mathfrak{R})$ ending at $\pi\in\mathfrak{R}$, then

$$Q_R^\gamma = Q_R^n\colon\Delta_\gamma \longrightarrow \Delta_\pi$$

is a bijection.

The set $\Delta_{\mathfrak{R}}=\bigcap_{n\geqslant 0}\Delta_{\mathfrak{R}}^n$ of (λ,π), to which we can apply Rauzy induction infinitely many times,[2] contains all (λ,π) such that the coordinates of λ are rationally independent, and so it is a full Lebesgue measure subset of $\Delta_{\mathfrak{R}}^0$. Let $(\lambda,\pi)\in\Delta_{\mathfrak{R}}$. The

---

[2] There is a nice explicit characterization of this set, called the Keane property. See [MMY] for a statement.

connected components of $\Delta_{\mathfrak{R}}^m$, $m \geqslant 0$, form a nested sequence of convex cones containing the half-line $\{(t\lambda, \pi) : t \in \mathbf{R}_+\}$, and their intersection is the connected component of (λ, π) in $\Delta_{\mathfrak{R}}$. In general, it is not true that this connected component reduces to the half-line $\{(t\lambda, \pi) : t \in \mathbf{R}_+\}$ ("combinatorial rigidity"): this happens precisely when the interval exchange transformation defined by (λ, π) is uniquely ergodic, and one can find counterexamples as soon as the genus $g = \dim H(\pi)/2$ is at least 2.

3.7. The Rauzy renormalization map

Since Q_R commutes with dilations, it projectivizes to a map $R_R \colon \mathbf{P}\Delta_{\mathfrak{R}}^1 \to \mathbf{P}\Delta_{\mathfrak{R}}^0$, that we call the *Rauzy renormalization map*. Let R_R^γ be the projectivization of Q_R^γ, for each path γ.

THEOREM 3.1. (Veech [Ve1]) *The map R_R has an ergodic conservative absolutely continuous invariant infinite measure μ_R. The density is analytic and positive in $\mathbf{P}\Delta_{\mathfrak{R}}^0$.*

Conservativeness means that if a measurable set contains its pre-image then the difference between the two has zero measure. It ensures that Poincaré recurrence holds in this context, despite the fact that the measure is infinite. This theorem is the key step in the proof by Masur and Veech that almost every interval exchange transformation is uniquely ergodic. Indeed, they show that if x is recurrent under R_R, then the connected component of x in $\mathbf{P}\Delta_{\mathfrak{R}}$ reduces to a point. The latter implies unique ergodicity of the interval exchange transformation given by the (λ, π)'s that projectivize to x.

3.8. The Zorich map

The Rauzy renormalization map does not admit an absolutely continuous invariant probability measure, because it is too slow. For instance, in the case of two intervals, the Rauzy renormalization map is just the Farey map, which exhibits a parabolic fixed point. Zorich introduced a way to "accelerate" Rauzy induction, that produces a new renormalization map, which is more expanding, and always admits an absolutely continuous invariant probability measure. In the case of two intervals, the Zorich renormalization map is, essentially, the Gauss map.

We say that we can apply Zorich induction to some $(\lambda, \pi) \in \Delta_{\mathfrak{R}}^0$ if there exists a smallest $m \geqslant 1$ such that we can apply Rauzy induction $m+1$ times to (λ, π), and in doing so we use both kinds of operations, top and bottom. Then we define $Q_Z(\lambda, \pi) = Q_R^m(\lambda, \pi)$. The domain of this *Zorich induction map* is the union of Δ_γ over all paths γ of length $m+1 \geqslant 2$ which are obtained by concatenating m arrows of one type (top or bottom) followed by an arrow of the other type. If we can apply Rauzy induction infinitely many

times to (λ,π) then we can also apply Zorich induction infinitely many times. The projectivization of the Zorich induction map Q_Z is called the *Zorich renormalization map* R_Z or, simply, the *Zorich map*.

THEOREM 3.2. (Zorich [Zo2]) *The Zorich map R_Z admits an ergodic absolutely continuous invariant probability measure. The density is analytic and positive in* $\mathbf{P}\Delta_{\mathfrak{R}}^1$.

We call the measure given by this theorem the *Zorich measure* μ_Z.

3.9. The Zorich cocycle

We define a linear cocycle $(x,h)\mapsto(R_Z(x),B^Z(x)\cdot h)$ over the Zorich map $x\mapsto R_Z(x)$ as follows. Let x belong to a connected component $\mathbf{P}\Delta_\gamma$ of the domain of R_Z, where γ is a path of length $m+1\geqslant 2$, and let $\tilde\gamma$ be the length-m path obtained by dropping the very last arrow in γ (the one that has type distinct from all the others). Then $B^Z(x)=\Theta(\tilde\gamma)$. To specify the linear cocycle completely we also have to specify where h is allowed to vary. There are two natural possibilities: either $H(\pi)$ for $x\in\mathbf{P}\Delta_\pi$, or the whole $\mathbf{R}^{\mathcal{A}}$. In the first case we speak of the *Zorich cocycle*, whereas in the second one we call this the *extended Zorich cocycle*.

THEOREM 3.3. (Zorich [Zo2]) *The (extended) Zorich cocycle is measurable*:

$$\int \log \|B^Z(x)\|\, d\mu_Z(x)<\infty.$$

The relation (3.2) gives that the Zorich cocycle is symplectic. Consequently, its Lyapunov exponents $\theta_1\geqslant...\geqslant\theta_{2g}$ satisfy $\theta_i=-\theta_{2g-i+1}$ for all i, where $g=g(\mathfrak{R})$ is the genus. We say that the Lyapunov spectrum is symmetric. The Lyapunov spectrum of the extended Zorich cocycle consists of the Lyapunov spectrum of the Zorich cocycle together with additional zeros.

3.10. Zippered rectangles

Let $\mathfrak{R}\subset\mathfrak{S}^0(\mathcal{A})$ be a Rauzy class. Let $\pi=(\pi_0,\pi_1)\in\mathfrak{R}$. Let $\Gamma_\pi\subset\mathbf{R}^{\mathcal{A}}$ be the set of all τ such that

$$\sum_{\pi_0(\xi)\leqslant k}\tau_\xi>0 \quad\text{and}\quad \sum_{\pi_1(\xi)\leqslant k}\tau_\xi<0 \quad\text{for all } 1\leqslant k\leqslant d-1. \tag{3.3}$$

Notice that Γ_π is an open convex polyhedral cone. It is non-empty, since the vector τ with coordinates $\tau_\xi=\pi_1(\xi)-\pi_0(\xi)$ belongs to Γ_π.

From the data (λ,π,τ), it is possible to define a translation surface $S=S(\lambda,\pi,\tau)$ in some $\mathcal{M}_{g,\varkappa}$, where g and $\varkappa$ depend only on π (see [MMY, §3.2]). This surface is obtained,

in the *zippered rectangles* construction, by gluing rectangles of horizontal sides λ_α and vertical sides h_α, where the height vector $h \in \mathbf{R}^\mathcal{A}$ is given by $h = -\Omega(\pi) \cdot \tau$. From the relations (3.3) and the definition (3.1) of $\Omega(\pi)$ one easily sees that all the coordinates of h are positive. The area of the translation surface S is $A(\lambda, \pi, \tau) = -\langle \lambda, \Omega \cdot \tau \rangle$. The space $H(\pi)$ may be identified with the first homology group of the surface, and so its dimension $2g(\pi)$ coincides with twice the genus of S. The numbers λ_ξ and τ_ξ also admit a geometric interpretation in terms of the surface: the zippered rectangles construction distinguishes a basis of the homology of S relative to the singularities, whose periods are the complex numbers $\lambda_\xi + i\tau_\xi$, $\xi \in \mathcal{A}$.

3.11. Inverse limit: renormalization for zippered rectangles

If γ is the top arrow ending at some permutation π', let Γ_γ be the set of all $\tau \in \Gamma_{\pi'}$ such that $\sum_{x \in \mathcal{A}} \tau_x < 0$, and if γ is a bottom arrow ending at π', let Γ_γ be the set of all $\tau \in \Gamma_{\pi'}$ such that $\sum_{x \in \mathcal{A}} \tau_x > 0$. If γ is an arrow starting at π and ending at π' then

$$\Theta(\gamma)^* \cdot \Gamma_\gamma = \Gamma_\pi.$$

Thus, the map

$$\widehat{Q}_R^\gamma \colon \Delta_\gamma \times \Gamma_\pi \longrightarrow \Delta_{\pi'} \times \Gamma_\gamma,$$
$$(x, \tau) \longmapsto (Q_R(x), (\Theta(\gamma)^{-1})^* \cdot \tau)$$

is invertible. Now we can define an invertible skew-product $\widehat{Q}_R$ over Q_R by putting together the $\widehat{Q}_R^\gamma$'s for every arrow γ. This is a map from $\bigcup \Delta_\gamma \times \Gamma_\pi$ (where the union is taken over all $\pi \in \mathfrak{R}$ and all arrows γ starting at π) to $\bigcup \Delta_{\pi'} \times \Gamma_\gamma$ (where the union is taken over all $\pi' \in \mathfrak{R}$ and all arrows ending at π'). Thus, $\widehat{Q}_R$ is an extension of the renormalization operator Q_R to zippered rectangles. The translation surfaces S and S' corresponding to (λ, π, τ) and $\widehat{Q}_R(\lambda, \pi, \tau)$, respectively, are obtained by appropriate cutting and pasting, so they correspond to the same element in the moduli space $\mathcal{M}_{g,\varkappa}$. See [MMY, §4.1].

Let $\hat{\Delta}_{\mathfrak{R}}^0 = \bigcup_{\pi \in \mathfrak{R}} \Delta_\pi \times \Gamma_\pi$, and let $\hat{\Delta}_{\mathfrak{R}} \subset \hat{\Delta}_{\mathfrak{R}}^0$ be the set of all points that can be iterated infinitely many times forward and backwards by $\widehat{Q}_R$. Note that $((\lambda, \pi), \tau) \in \hat{\Delta}_{\mathfrak{R}}^0$ can be iterated infinitely many times forward/backwards by $\widehat{Q}_R$ provided the coordinates of λ/τ are rationally independent. Projectivization in λ and τ gives a map $\widehat{R}_R \colon \widehat{\mathbf{P}\Delta}_{\mathfrak{R}} \to \widehat{\mathbf{P}\Delta}_{\mathfrak{R}}$. This is an invertible map that can be seen as the inverse limit of R_R: the connected components of $\widehat{\mathbf{P}\Delta}_{\mathfrak{R}}$ reduce to points Lebesgue almost everywhere. There is a natural infinite ergodic conservative invariant measure $\hat{\mu}_R$ for $\widehat{R}_R$ that is equivalent to Lebesgue

measure. The projection μ_R of $\hat{\mu}_R$ is the measure appearing in Theorem 3.1: this is how μ_R is actually constructed in [Ve1].

Let $\Upsilon^0_{\mathfrak{R}}=\bigcup \mathbf{P}\Delta_\gamma \times \mathbf{P}\Gamma_{\gamma'}$, where the union is taken over all pairs of arrows γ' and γ of distinct types (one is top and the other is bottom) such that γ' ends at the start of γ. Let $\Upsilon_{\mathfrak{R}}=\Upsilon^0_{\mathfrak{R}}\cap\widehat{\mathbf{P}\Delta}_{\mathfrak{R}}$ and $\widehat{R}_Z\colon \Upsilon_{\mathfrak{R}}\to\Upsilon_{\mathfrak{R}}$ be the first return map. The choice of $\Upsilon_{\mathfrak{R}}$ is so that $\hat{\mu}_R(\Upsilon_{\mathfrak{R}})$ is finite. So, we may normalize $\hat{\mu}_R|_{\Upsilon_{\mathfrak{R}}}$ to get a probability measure $\hat{\mu}_Z$ on $\Upsilon_{\mathfrak{R}}$ which is invariant under $\widehat{R}_Z$. One checks that $\hat{\mu}_Z$ is ergodic. Moreover, the map $\widehat{R}_Z$ is a skew-product over R_Z that can be seen as the inverse limit of R_R. The projection μ_Z of $\hat{\mu}_Z$ is the probability measure described in Theorem 3.2: this is how μ_Z is constructed in [Zo2].

The *invertible Zorich cocycle* is the lift $(\widehat{R}_Z, \widehat{B}^Z)$ of the Zorich cocycle to a cocycle over the invertible Zorich map, defined by $\widehat{B}^Z((\lambda,\pi),\tau)=B^Z(\lambda,\pi)$.

3.12. Minimality in the projective space

The aim of this section is to prove Corollary 3.6, regarding the projective behavior of the matrices involved in the Zorich cocycle. This result easily follows from known constructions, but will be quite useful in the sequel .

Given $x\in\widehat{\mathbf{P}\Delta}_{\mathfrak{R}}$, obtained by projectivizing $((\lambda,\pi),\tau)$, consider the subspaces E^{uu}, E^c and E^{ss} of $H(\pi)$ defined as follows: firstly, E^{uu} is the line spanned by $\Omega(\pi)\cdot\tau$; secondly, E^{ss} is the line spanned by $\Omega(\pi)\cdot\lambda$; and, finally, E^c is the symplectic orthogonal to the plane $E^{uu}\oplus E^{ss}$. Notice that E^{uu} is not symplectically orthogonal to E^{ss}: $\langle\lambda,\Omega(\pi)\cdot\tau\rangle<0$, since both λ and $h=-\Omega(\pi)\cdot\tau$ have only positive coordinates. Thus, E^c has codimension 2 in $H(\pi)$ and $H(\pi)=E^{uu}\oplus E^c\oplus E^{ss}$.

LEMMA 3.4. *The splitting $E^{uu}\oplus E^c\oplus E^{ss}$ is invariant under the invertible Zorich cocycle. The spaces E^{uu} and E^{ss} correspond to the largest and the smallest Lyapunov exponents, respectively, and E^c corresponds to the remaining exponents.*

Proof. The invariance of E^{uu} and E^{ss} follows directly from the definitions and (3.2). The invariance of E^c is a consequence, since the Zorich cocycle is symplectic. Since E^{uu} is the projectivization of a direction in the positive cone and the matrices of the invertible Zorich cocycle have non-negative entries, E^{uu} must be contained in the subspace corresponding to the largest Lyapunov exponent. Since the largest Lyapunov exponent is simple (see [Zo2]), E^{uu} must be the Oseledets direction associated with it. This implies that E^{ss} corresponds to the smallest Lyapunov exponent, since E^{ss} is an invariant direction which is not contained in the symplectic orthogonal to E^{uu}. It follows that E^c must correspond to the other Lyapunov exponents. $\square$

It follows from the definitions that $-\Omega(\pi)\cdot\Gamma_\pi \subset H(\pi)\cap\mathbf{R}_+^\mathcal{A}$ and hence $-\Omega(\pi)\cdot\Gamma_\pi$ is an open cone in $H(\pi)$. In fact, it is the interior of the cone $H^+(\pi)$ of Veech [Ve1].

LEMMA 3.5. *For every* $[h]\in\mathbf{P}H(\pi)$, *the set* $\Theta(\Pi(\pi))\cdot[h]$ *contains a dense subset of* $\mathbf{P}(-\Omega(\pi)\cdot\Gamma_\pi)$.

Proof. For almost every $x\in\Upsilon_\mathfrak{R}$, $[h]$ is not symplectically orthogonal to the span $E^{ss}(x)$ of $\Omega(\pi)\cdot\lambda$. Consequently, $[h]\not\subset E^c(x)\oplus E^{ss}(x)$. By the Oseledets theorem,

$$\widehat{B}^Z(R_Z^{m-1}(x))\dots B^Z(x)\cdot[h] \quad \text{and} \quad \Omega(\pi^{(m)})\cdot[\tau^{(m)}]=E^{uu}(x^{(m)})$$

are asymptotic. On the other hand, by ergodicity, the sequence

$$\widehat{R}_Z^m(x)=x^{(m)}=(([\lambda^{(m)}],\pi^{(m)}),[\tau^{(m)}])$$

is dense in $\Upsilon_\mathfrak{R}$ for almost every x. This implies that $E^{uu}(x^{(m)})$ is dense in $\mathbf{P}(-\Omega(\pi)\cdot\Gamma_\pi)$. The result follows from these two observations. $\square$

COROLLARY 3.6. *The action, via* Θ, *of the Rauzy monoid* $\Pi(\pi)$ *on* $\mathbf{P}H(\pi)$ *is minimal.*

Proof. By Lemma 3.5, the closure of any orbit must intersect every other orbit. This implies minimality, obviously. $\square$

3.13. Choices of permutations in Rauzy classes

Let us say that π is *standard* if the first letter in the top/bottom is the last letter in the bottom/top. Note that a standard permutation is always irreducible.

LEMMA 3.7. (Rauzy [Ra]) *In every Rauzy class there exists a standard permutation.*

Let $\pi\in\mathfrak{S}^0(\mathcal{A})$ be standard. Assuming $d\geqslant3$, we call π *degenerate* if there exists $B\in\mathcal{A}$ which is either second of the top and bottom rows or last but one of the top and bottom rows. Assuming $d\geqslant4$, we call π *good* if forgetting the first (and last) letters of the top and the bottom rows give an irreducible permutation. Notice that a standard permutation cannot be both degenerate and good.

LEMMA 3.8. ([KZ, Lemma 20]) *In every Rauzy class with* $\#\mathcal{A}\geqslant3$ *there exists either a good permutation or a degenerate permutation.*

Proof. We give a proof here for the convenience of the reader, and to avoid confusion with the slightly different language of [KZ]. For $d=\#\mathcal{A}=3$ the result is immediate. In what follows we suppose $d\geqslant4$.

Let A be the first letter in the top and E be the first letter in the bottom. Suppose, by contradiction, that no standard permutation in $\mathfrak{R}$ is either degenerate or good. Let π be a standard permutation and π' be obtained by forgetting A and E. Then π' is reducible, and so there exists a maximal $1 \leqslant k \leqslant d-3$ such that the set of the first k symbols in the top and in the bottom of π' coincide. Since π is non-degenerate, we must have $2 \leqslant k \leqslant d-4$. Assume that π has been chosen so that the resulting k is maximal. Let $x_1, ..., x_k$ be the first k letters in the top and $y_1, ..., y_k$ be the first k letters in the bottom of π' (so $\{x_1, ..., x_k\} = \{y_1, ..., y_k\}$). Let C be the letter in position $d-1$ in the top of π, and l be its position in the bottom of π. Then $k+2 \leqslant l \leqslant d-2$. Let C' be the letter preceding C in the bottom of π, and r be its position in the top of π. Then $2 \leqslant r \leqslant d-2$. Let us consider the following Rauzy path starting from π:

(1) apply $d-r$ bottom iterations to π, so that C' becomes last in the top;

(2) apply $d-l$ top iterations, so that C becomes last in the bottom;

(3) apply $r-1$ bottom iterations, so that E becomes last in the top;

(4) finally, apply 1 top iteration, so that A becomes last in the bottom.

Notice that step (1) sends C to the position $d-r$ in the top, preceding E, and step (2) sends A to the position $d-1$ in the bottom. In the end, we get a new standard permutation $\tilde{\pi}$: A is last in the bottom and E is last in the top. Let $\tilde{\pi}'$ be obtained from $\tilde{\pi}$ by forgetting the letters A and E. There are two cases to consider.

If $l > k+2$ then $r > k+1$. The calculation for this case is detailed in the following formula:

$$\begin{pmatrix} A & x_1 & \cdot & x_k & \cdot & \cdot & \cdot & \cdot & C' & \cdot & C & E \\ E & y_1 & \cdot & y_k & \cdot & C' & C & \cdot & \cdot & \cdot & \cdot & A \end{pmatrix}$$

$$\longmapsto \begin{pmatrix} A & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & C & E & x_1 & \cdot & x_k & \cdot & C' \\ E & y_1 & \cdot & y_k & \cdot & C' & C & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & A \end{pmatrix}$$

$$\longmapsto \begin{pmatrix} A & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & C & E & x_1 & \cdot & x_k & \cdot & C' \\ E & y_1 & \cdot & y_k & \cdot & C' & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & A & C \end{pmatrix}$$

$$\longmapsto \begin{pmatrix} A & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & C & x_1 & \cdot & x_k & \cdot & C' & E \\ E & y_1 & \cdot & y_k & \cdot & C' & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & A & C \end{pmatrix}$$

$$\longmapsto \begin{pmatrix} A & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & C & x_1 & \cdot & x_k & \cdot & C' & E \\ E & C & y_1 & \cdot & y_k & \cdot & C' & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & A \end{pmatrix}.$$

Notice that C precedes $x_1, ..., x_k$ in the top and precedes $y_1, ..., y_k$ in the bottom of $\tilde{\pi}$. By assumption, $\tilde{\pi}$ is neither degenerate nor good. The first of these facts implies that C is not the second letter in the top of $\tilde{\pi}$. The second one means that $\tilde{\pi}'$ is reducible: there exists $1 \leqslant \tilde{k} \leqslant d-3$ such that the first $\tilde{k}$ elements in the top and the bottom of $\tilde{\pi}'$ coincide.

In view of the previous observations, this implies that the first $\tilde{k}$ elements in the bottom of $\tilde{\pi}'$ include C and $y_1, ..., y_k$. Thus $\tilde{k} > k$, contradicting the choice of k.

If $l = k+1$, then $C' = y_k = x_{r-1}$. The calculation for this case is detailed in the following formula:

$$\begin{pmatrix} A & x_1 & \cdot & x_{r-1} & \cdot & x_k & \cdot & \cdot & C & E \\ E & y_1 & \cdot & \cdot & \cdot & y_k & C & \cdot & D & A \end{pmatrix}$$

$$\longmapsto \begin{pmatrix} A & \cdot & \cdot & D & \cdot & \cdot & \cdot & \cdot & C & E & x_1 & \cdot & \cdot & x_{r-1} \\ E & y_1 & \cdot & \cdot & \cdot & y_k & C & \cdot & \cdot & \cdot & \cdot & \cdot & D & A \end{pmatrix}$$

$$\longmapsto \begin{pmatrix} A & \cdot & \cdot & D & \cdot & \cdot & \cdot & \cdot & C & E & x_1 & \cdot & \cdot & x_{r-1} \\ E & y_1 & \cdot & \cdot & \cdot & y_k & \cdot & \cdot & \cdot & \cdot & \cdot & D & A & C \end{pmatrix}$$

$$\longmapsto \begin{pmatrix} A & \cdot & \cdot & D & \cdot & \cdot & \cdot & \cdot & C & x_1 & \cdot & \cdot & x_{r-1} & E \\ E & y_1 & \cdot & \cdot & \cdot & y_k & \cdot & \cdot & \cdot & \cdot & \cdot & D & A & C \end{pmatrix}$$

$$\longmapsto \begin{pmatrix} A & \cdot & \cdot & D & \cdot & \cdot & \cdot & \cdot & C & x_1 & \cdot & \cdot & x_{r-1} & E \\ E & C & y_1 & \cdot & \cdot & \cdot & y_k & \cdot & \cdot & \cdot & \cdot & \cdot & D & A \end{pmatrix}.$$

Let D be the letter in position $d-1$ in the bottom of π. Notice that $C \neq D$, since π is non-degenerate. After the first step, C precedes E which precedes $x_1, ..., x_{r-1} = y_k$. In particular, D appears before C in the top. It follows that D appears before C in the top of $\tilde{\pi}$. On the other hand, after the second step, D precedes A in the bottom. It follows that D is in position $d-1$ in the bottom of $\tilde{\pi}$. So, C is the first letter and D is the last letter in the bottom of $\tilde{\pi}'$, and D appears before C in the top. This implies that $\tilde{\pi}'$ is irreducible, which contradicts the hypothesis. So, this case cannot really occur. $\qquad\square$

4. Twisting and pinching

In this section we consider actions of a monoid $\mathcal{B}$ by linear isomorphisms of a vector space H. We introduce certain properties of monoids that we call twisting, twisting of isotropic spaces, pinching, strong pinching, and simplicity, and describe some logical relations between them.

4.1. Twisting

We say that a monoid twists a subspace F if it contains enough elements to send F outside any finite union of hyperplane sections. More precisely, we give the following definition.

Definition 4.1. We say that $\mathcal{B}$ *twists* some $F \in \mathrm{Grass}(k, H)$ if, for every finite subset $\{F_i\}_{i=1}^m$ of $\mathrm{Grass}(\dim H - k, H)$, there exists $x \in \mathcal{B}$ such that

$$(x \cdot F) \cap F_i = \{0\}, \quad 1 \leqslant i \leqslant m.$$

In connection with the next lemma, observe that a linear arrangement S is $\mathcal{B}$-invariant (that is, $x \cdot S = S$ for every $x \in \mathcal{B}$) if and only if it is $\mathcal{B}^{-1}$-invariant, where $\mathcal{B}^{-1} = \{x^{-1} : x \in \mathcal{B}\}$. In particular, the lemma implies that $\mathcal{B}$ *twists* F if and only if $\mathcal{B}^{-1}$ *twists* F.

LEMMA 4.2. *A monoid $\mathcal{B}$ twists $F \in \mathrm{Grass}(k, H)$ if and only if F does not belong to any non-trivial invariant linear arrangement in $\mathrm{Grass}(k, H)$.*

Proof. Let $S \subset \mathrm{Grass}(k, H)$ be a non-trivial linear arrangement containing F. Then S is contained in a finite union $\widetilde{S}$ of hyperplane sections $S_i = \{F' \in \mathrm{Grass}(k, H) : F' \cap F_i \neq \{0\}\}$ for all $1 \leqslant i \leqslant m$. If $\mathcal{B}$ twists F then there exists $x \in \mathcal{B}$ such that $(x \cdot F) \cap F_i = \{0\}$, $1 \leqslant i \leqslant m$, that is $x \cdot F \notin \widetilde{S}$. Since $S \subset \widetilde{S}$, it follows that S is not invariant. On the other hand, if $\mathcal{B}$ does not twist F then there exists a finite union of hyperplane sections $\widetilde{S} \subset \mathrm{Grass}(k, H)$ such that $x \cdot F \in \widetilde{S}$ for every $x \in \mathcal{B}$. From Corollary 2.5 (1) we get that $S = \bigcap_{x \in \mathcal{B}} x^{-1} \cdot \widetilde{S}$ is a non-trivial linear arrangement containing F. It is clear that $x^{-1} \cdot S \supset S$ for every $x \in \mathcal{B}$. From Corollary 2.5 (3) it follows that $x \cdot S = S$ for every $x \in \mathcal{B}$. $\square$

Next, we prove that given any finite family of hyperplane sections and subspaces, possibly with variable dimensions, there exists some isomorphism in $\mathcal{B}$ that sends every subspace outside the corresponding hyperplane section (simultaneously), provided $\mathcal{B}$ twists each one of the subspaces individually.

LEMMA 4.3. *For $1 \leqslant j \leqslant m$, let k_j satisfy $1 \leqslant k_j \leqslant 2g-1$, let $F_j \in \mathrm{Grass}(k_j, H)$, and let $F_j' \in \mathrm{Grass}(\dim H - k_j, H)$. Assume that F_j is twisted by $\mathcal{B}$ for every j. Then there exists $x \in \mathcal{B}$ such that $x \cdot F_j \cap F_j' = \{0\}$ for all $1 \leqslant j \leqslant m$.*

Proof. Let $S_j \subset \Lambda^{k_j} H$ be the hyperplane dual to F_j'. Consider the vector space $X = \prod_{j=1}^m \Lambda^{k_j} H$ and let $W = \bigcap_{x \in \mathcal{B}} x^{-1} \cdot Y$, where

$$Y = \bigcup_{j=1}^m \left[\prod_{i < j} \Lambda^{k_i} H \times S_j \times \prod_{i > j} \Lambda^{k_i} H \right] \subset X.$$

Let $u_j \in \Lambda^{k_j} H$ projectivize to F_j. If the conclusion does not hold, then $u = \{u_j\}_{j=1}^m$ belongs to W. Using Lemmas 2.1 and 2.2, we may write W as a finite union (uniquely defined, up to order) of subspaces W_l, $l = 1, ..., L$, where L is minimal and the W_l's are of the form

$$W_l = W_{l,1} \times ... \times W_{l,m},$$

where each $W_{l,j}$ is an intersection of geometric hyperplanes of $\Lambda^{k_j} H$, or else coincides with the whole exterior product. Given any $x \in \mathcal{B}$, we have $x^{-1} \cdot W \supset W$ and so $x^{-1} \cdot W = W$, by Corollary 2.5 (3). This means that each $x \in \mathcal{B}$ permutes the W_l's and so, for every j, it permutes the $W_{l,j}$'s (counted with multiplicity). We have $u \in W_{l_0}$ for some l_0, that is, $u_j \in W_{l_0,j}$ for all j. Since $W \neq X$, there exists j_0 such that $W_{l_0,j_0} \neq \Lambda^{k_{j_0}} H$, and so $\mathrm{Grass}(k_{j_0}, H)$ is not contained in the projectivization of W_{l_0,j_0}. Note that, by construction, the intersection of $\mathrm{Grass}(k_{j_0}, H)$ with the projectivization of W_{l_0,j_0} contains F_{j_0}. Let W^0 be the union of all W_{l,j_0}'s whose projectivization intersects but does not contain $\mathrm{Grass}(k_{j_0}, H)$. Then the projectivization of W^0 intersected with $\mathrm{Grass}(k_{j_0}, H)$ is a non-trivial linear arrangement in the Grassmannian, invariant for $\mathcal{B}$ and containing F_{j_0}. This contradicts the assumption that $\mathcal{B}$ twists F_{j_0}. $\qquad \square$

LEMMA 4.4. *Let $\mathcal{B}$ be a monoid acting symplectically on (H, ω), and assume that the action of $\mathcal{B}$ on the space of Lagrangian flags $\mathcal{L}(H)$ is minimal. Then $\mathcal{B}$ twists isotropic subspaces.*

Proof. If $\mathcal{B}$ acts minimally on $\mathcal{L}(H)$, then it also acts minimally on each $\mathrm{Iso}(k, H)$, $1 \leqslant k \leqslant g$ (where $\dim H = 2g$). We will only use this latter property. Let $F \in \mathrm{Iso}(k, H)$ and S be a non-trivial invariant linear arrangement in $\mathrm{Grass}(k, H)$. If $F \in S$ then $S \cap \mathrm{Iso}(k, H)$ is a non-empty closed invariant set under the action of $\mathcal{B}$. By minimality, it must be the whole of $\mathrm{Iso}(k, H)$. This contradicts Lemma 2.6. Therefore, $F \notin S$. In view of Lemma 4.2, this proves the claim. $\qquad \square$

4.2. Pinching

Assume that $\mathcal{B}$ acts symplectically on (H, ω), $\dim H = 2g$.

Definition 4.5. We say that $\mathcal{B}$ is *strongly pinching* if for every $C > 0$, there exists $x \in \mathcal{B}$ such that

$$\log \sigma_g(x) > C \quad \text{and} \quad \log \sigma_i(x) > C \log \sigma_{i+1}(x) \text{ for all } 1 \leqslant i \leqslant g-1. \qquad (4.1)$$

This is independent of the choice of the inner product used to define the singular values.

The converse statement in the next lemma says that strong pinching together with the twisting of all isotropic spaces provide good separation of the Lyapunov exponents.

LEMMA 4.6. *If for every $C > 0$ there exists $x \in \mathcal{B}$ such that*

$$\theta_g(x) > 0 \quad and \quad \theta_k(x) > C\theta_{k+1}(x) \text{ for all } 1 \leqslant k \leqslant g-1,$$

then $\mathcal{B}$ is strongly pinched. Most important, the converse holds if $\mathcal{B}$ twists isotropic spaces.

Proof. Clearly, if x has simple Lyapunov spectrum, we have

$$\sigma_k(x^n) \asymp e^{n\theta_k(x)}.$$

This gives the first assertion. For the second one, let $x_n \in \mathcal{B}$ be such that

$$\sigma_g(x_n) > n \quad \text{and} \quad \log \sigma_k(x_n) > n \log \sigma_{k+1}(x_n) \ \text{ for all } \ 1 \leqslant k \leqslant g-1.$$

We may assume that $x_n \cdot E_k^+(x_n)$ converges to some $E_k^u \in \mathrm{Grass}(k, H)$ and $E_k^-(B_n)$ converges to some $E_k^s \in \mathrm{Grass}(2g-k, H)$, $1 \leqslant k \leqslant g$. By Lemma 2.13, the subspaces E_k^u are isotropic. It follows from Lemma 4.3 that if $\mathcal{B}$ twists isotropic subspaces, there exists $x \in \mathcal{B}$ such that

$$x \cdot E_k^u \cap E_k^s = \{0\} \quad \text{for all } \ 1 \leqslant k \leqslant g.$$

By Lemma 2.9, there exists $C > 0$ such that

$$|\theta_k(xx_n) - \log \sigma_k(x_n)| < C \quad \text{for all } \ 1 \leqslant k \leqslant g,$$

which implies the result. $\qquad\qquad\square$

The next lemma indicates a useful special situation where one has the strong pinching property.

LEMMA 4.7. *If for every $C > 0$ there exists $x \in \mathcal{B}$ for which 1 is an eigenvalue of geometric multiplicity 1 (the dimension of the eigenspace is equal to 1), and we have*

$$\theta_{g-1}(x) > 0 \quad \text{and} \quad \theta_k(x) > C\theta_{k+1}(x) \ \text{ for all } \ 1 \leqslant k \leqslant g-2,$$

then $\mathcal{B}$ is strongly pinched.

Proof. Since the action is symplectic, the eigenvalue 1 has even algebraic multiplicity. Considering the Jordan form of x, the hypotheses imply that

$$\sigma_g(x^n) \asymp n, \tag{4.2}$$

$$\sigma_k(x^n) \asymp e^{\theta_k(x)n} \quad \text{for all } \ 1 \leqslant k \leqslant g-1. \tag{4.3}$$

The claim follows immediately. $\qquad\qquad\square$

LEMMA 4.8. *Assume that $\mathcal{B}_0 \subset \mathcal{B}$ twists isotropic subspaces and is strongly pinching. Then for every $x \in \mathcal{B}$ and any $C > 0$ there exists $x_0 \in \mathcal{B}_0$ such that*

$$\theta_g(x_0) > 0 \quad \text{and} \quad \theta_i(x_0) > C\theta_{i+1}(x_0) \ \text{ for all } \ 1 \leqslant i \leqslant g-1, \tag{4.4}$$

$$\theta_g(xx_0) > 0 \quad \text{and} \quad \theta_i(xx_0) > C\theta_{i+1}(xx_0) \ \text{ for all } \ 1 \leqslant i \leqslant g-1. \tag{4.5}$$

Proof. Let $x_2 \in \mathcal{B}_0$ satisfy the conditions in (4.4): $\theta_g(x_2) > 0$ and $\theta_i(x_2) > C\theta_{i+1}(x_2)$ for all $1 \leqslant i \leqslant g-1$. Let $E_k^u(x_2) = \lim_{n \to \infty} x_2^n \cdot E_k^+(x_2^n)$, which is the space spanned by the k eigenvectors with largest eigenvalues, and let $E_k^s(x_2) = \lim_{n \to \infty} E_k^-(x_2^n)$, which is the space spanned by the $2g-k$ remaining eigenvectors. Using Lemma 4.3, select $x_1 \in \mathcal{B}_0$ such that $x_1 \cdot E_k^u \cap E_k^s = \{0\}$ and $xx_1 \cdot E_k^u \cap E_k^s = \{0\}$ for all $1 \leqslant k \leqslant g$. Then $x_0 = x_1 x_2^n$ satisfies all the conditions for n large, by the same argument as in Lemma 4.6. $\square$

4.3. Simple actions

Definition 4.9. We say that $\mathcal{B}$ is *twisting* if it twists any $F \in \mathrm{Grass}(k, H)$ for any $1 \leqslant k \leqslant \dim H - 1$.

From the observation preceding Lemma 4.2 we have that $\mathcal{B}$ is twisting if and only if $\mathcal{B}^{-1}$ is twisting.

Definition 4.10. We say that $\mathcal{B}$ is *pinching* if for every $C > 0$ there exists $x \in \mathcal{B}$ such that

$$\sigma_i(x) > C\sigma_{i+1}(x) \quad \text{for all } 1 \leqslant i \leqslant \dim H - 1.$$

This is independent of the choice of the inner product used to define the singular values.

Definition 4.11. We say that $\mathcal{B}$ is *simple* if it is both twisting and pinching.

LEMMA 4.12. *Let $\mathcal{B}$ be a simple monoid. Then the inverse monoid $\mathcal{B}^{-1}$ is also simple.*

Proof. Twisting follows directly from the observation preceding Lemma 4.2. Pinching follows from the fact that $\log \sigma_i(x) = -\log \sigma_{\dim H - i + 1}(x^{-1})$ for all $1 \leqslant i \leqslant \dim H - 1$. $\square$

We suspect that there could be sufficient conditions for twisting along the following lines.

Problem 4.13. If $\mathcal{B}$ acts minimally on $\mathbf{P}H$, is it necessarily twisting? For symplectic actions, one can ask a weaker question: Does minimal action on $\mathbf{P}H$ imply twisting of isotropic subspaces?

LEMMA 4.14. *Let $\mathcal{B}$ be a monoid acting symplectically on (H, ω). If $\mathcal{B}$ twists isotropic subspaces and is strongly pinching then it is simple.*

Proof. To see that strong pinching implies pinching, it is enough to consider an adapted inner product, for which (4.1) implies also

$$-\log \sigma_{g+1}(x) > C, \tag{4.6}$$

$$-\log \sigma_{i+1}(x) > -C\log \sigma_i(x) \quad \text{for all } g+1 \leqslant i \leqslant 2g-1, \tag{4.7}$$

where $\dim H = 2g$. We are left to show that, under the hypotheses, $\mathcal{B}$ twists any $F \in \mathrm{Grass}(k, H)$. It is easy to see that $\mathcal{B}$ twists F if and only if it twists its symplectic orthogonal. So it is enough to consider the case $1 \leqslant k \leqslant g$. Let $S \subset \mathrm{Grass}(k, H)$ be a non-trivial invariant linear arrangement. By Lemma 4.6, there exists $x_1 \in \mathcal{B}$ with simple Lyapunov spectrum. Let $E^s = \lim_{n \to \infty} E_g^-(x_1^n)$ and $E^u = \lim_{n \to \infty} x_1^n \cdot E_g^+(x_1^n)$ be the stable and unstable eigenspaces of x_1. Thus both E^s and E^u are isotropic subspaces of H. If $F \in S$ then there exists $x_0 \in \mathcal{B}$ such that $(x_0 \cdot F) \cap E^s = \{0\}$. It follows that $x_1^n x_0 \cdot F$ converges to a subspace of E^u. Since S is closed, we conclude that $S \cap \mathrm{Iso}(k, H) \neq \varnothing$. This contradicts the assumption that $\mathcal{B}$ twists all elements of $\mathrm{Iso}(k, H)$. $\qquad \square$

Remark 4.15. An argument similar to the proof of the previous lemma shows that if $\mathcal{B}$ is simple and acts symplectically, then any closed invariant set in the flag space $\mathcal{F}(H)$ intersects the embedding of the space of Lagrangian flags $\mathcal{L}(H)$, that is, it contains some $(F_i)_{i=1}^{2g-1}$ such that $(F_i)_{i=1}^{g}$ is a Lagrangian flag and F_{2g-i} is the symplectic orthogonal of F_i.

LEMMA 4.16. *Let $\mathcal{B}_0 \subset \mathcal{B}$ be a large submonoid in the sense that there exists a finite subset $Y \subset \mathcal{B}$ and $z \in \mathcal{B}$ such that for every $x \in \mathcal{B}$ there is some $y \in Y$ such that $yxz \in \mathcal{B}_0$. If $\mathcal{B}$ is twisting or pinching then $\mathcal{B}_0$ also is. Assuming that the action of $\mathcal{B}$ is symplectic, if $\mathcal{B}$ twists isotropic subspaces or is strongly pinching then the same holds for $\mathcal{B}_0$.*

Proof. Notice that $|\log \sigma_i(x) - \log \sigma_i(yxz)| < C$, where C only depends on y, z and the choice of the inner product used to define the singular values. Thus if $\mathcal{B}$ is (strongly) pinching then $\mathcal{B}_0$ is also (strongly) pinching. Let $S \subset \mathrm{Grass}(k, H)$ be a non-trivial linear arrangement invariant for $\mathcal{B}_0$. Then

$$\bigcap_{x \in \mathcal{B}} x^{-1} \cdot \bigcup_{y \in Y} y^{-1} \cdot S \supset \bigcap_{y \in Y} \bigcap_{x_0 \in \mathcal{B}_0} (y^{-1} x_0 z^{-1})^{-1} \cdot \bigcup_{y \in Y} y^{-1} \cdot S \supset z \cdot \bigcap_{y \in Y} \bigcap_{x_0 \in \mathcal{B}_0} x_0^{-1} \cdot S = z \cdot S.$$

So, $z \cdot S$ is contained in a non-trivial linear arrangement invariant for $\mathcal{B}$. This shows that if $\mathcal{B}$ is twisting then so is $\mathcal{B}_0$. If $\mathcal{B}$ acts symplectically and S intersects $\mathrm{Iso}(k, H)$ then $z \cdot S$ also does, so if $\mathcal{B}$ twists isotropic subspaces then $\mathcal{B}_0$ also does. $\qquad \square$

LEMMA 4.17. *Let $\mathcal{B}$ be a simple monoid. Then there exists $x \in \mathcal{B}$ with simple Lyapunov spectrum $\theta_i(x) > \theta_{i+1}(x)$, $1 \leqslant i \leqslant \dim H - 1$.*

Proof. Let $x_n \in \mathcal{B}$ be such that $\log \sigma_i(x) - \log \sigma_{i+1}(x) > n$, $1 \leqslant i \leqslant \dim H - 1$. We may assume that $E_i^-(x_n)$ and $x_n \cdot E_i^+(x_n)$ converge to spaces E_i^s and E_i^u, $1 \leqslant i \leqslant \dim H - 1$. Let $x \in \mathcal{B}$ be such that $(x \cdot E_i^u) \cap E_i^s = \{0\}$, $1 \leqslant i \leqslant \dim H - 1$. By Lemma 2.9, there exists $C > 0$ such that $|\theta_i(xx_n) - \log \sigma_i(x_n)| < C$, $1 \leqslant i \leqslant \dim H - 1$. This gives the result. $\qquad \square$

5. Twisting of Rauzy monoids

Recall that a Rauzy monoid $\Pi(\pi)$ acts symplectically on $H(\pi)$ (by $\gamma \cdot h = \Theta(\gamma) \cdot h$). In particular, it acts on the space of Lagrangian flags $\mathcal{L}(H(\pi))$. Our aim in this section is to prove the following result.

THEOREM 5.1. *Let π be irreducible. The action of the Rauzy monoid $\Pi(\pi)$ on the space of Lagrangian flags $\mathcal{L}(H(\pi))$ is minimal.*

COROLLARY 5.2. *Let π be irreducible. The action of the Rauzy monoid $\Pi(\pi)$ on $H(\pi)$ twists isotropic subspaces.*

Proof. This follows directly from Theorem 5.1 and Lemma 4.4. □

5.1. Simple reduction

Let $\mathcal{A}$ be an alphabet on $d \geqslant 3$ symbols, $B \in \mathcal{A}$, and $\mathcal{A}' = \mathcal{A} \setminus \{B\}$. Given $\pi \in \mathfrak{S}^0(\mathcal{A})$, let π' be obtained from π by removing B from the top and bottom rows. If $\pi' \in \mathfrak{S}^0(\mathcal{A}')$, we say that π' is a *simple reduction* of π. Let $2g(\pi) = \dim H(\pi) = \operatorname{rank} \Omega(\pi)$ and analogously for π'. Let $P \colon \mathbf{R}^{\mathcal{A}} \to \mathbf{R}^{\mathcal{A}'}$ be the natural projection, and $P^* \colon \mathbf{R}^{\mathcal{A}'} \to \mathbf{R}^{\mathcal{A}}$ be its adjoint (the natural inclusion). Notice that

$$P\Omega(\pi)P^* = \Omega(\pi'). \tag{5.1}$$

LEMMA 5.3. *Let π' be a simple reduction of π. Then either $g(\pi) = g(\pi')$ or $g(\pi) = g(\pi') + 1$. Moreover, the following are equivalent:*
(1) $g(\pi) = g(\pi')$;
(2) $H(\pi)$ *is spanned by* $\{\Omega(\pi) \cdot e_x, x \in \mathcal{A}'\}$;
(3) $e_B \notin H(\pi)$;
(4) e_B *does not belong to the span of* $\{\Omega(\pi) \cdot e_x, x \in \mathcal{A}'\}$;
(5) P *restricts to a symplectic isomorphism* $H(\pi) \to H(\pi')$.

Proof. There are two possibilities for the value of $\operatorname{rank} P\Omega(\pi)P^*$:

(a) We may have $\operatorname{rank} P\Omega(\pi)P^* = \operatorname{rank} \Omega(\pi)$. Since the rank cannot increase by composition, this implies that

$$\operatorname{rank} P\Omega(\pi)P^* = \operatorname{rank} P\Omega(\pi) = \operatorname{rank} \Omega(\pi)P^* = \operatorname{rank} \Omega(\pi).$$

(b) We may have $\operatorname{rank} P\Omega(\pi)P^* < \operatorname{rank} \Omega(\pi)$. Notice that

$$\operatorname{rank} P\Omega(\pi)P^* \geqslant \operatorname{rank} \Omega(\pi) - 2. \tag{5.2}$$

Since $\operatorname{rank} \Omega(\pi)$ and $\operatorname{rank} P\Omega(\pi)P^*$ are even, we must have $\operatorname{rank} P\Omega(\pi)P^* = \operatorname{rank} \Omega(\pi) - 2$. This implies that $\operatorname{rank} P\Omega(\pi) = \operatorname{rank} \Omega(\pi)P^* = \operatorname{rank} \Omega(\pi) - 1$.

By (5.1), in case (a) we have $g(\pi)=g(\pi')$ and in case (b) we have $g(\pi)=g(\pi')+1$. Notice that

(1) is equivalent to rank $P\Omega(\pi)P^*=\operatorname{rank}\Omega(\pi)$;

(2) is equivalent to $\Omega(\pi)\cdot\mathbf{R}^{\mathcal{A}}=\Omega(\pi)(P^*\cdot\mathbf{R}^{\mathcal{A}'})$, that is, rank $\Omega(\pi)=\operatorname{rank}\Omega(\pi)P^*$;

(3) is equivalent to $(\Omega(\pi)\cdot\mathbf{R}^{\mathcal{A}})\cap\operatorname{Ker}P=\{0\}$, that is, rank $P\Omega(\pi)=\operatorname{rank}\Omega(\pi)$;

(4) is equivalent to $(\Omega(\pi)(P^*\cdot\mathbf{R}^{\mathcal{A}'}))\cap\operatorname{Ker}P=\{0\}$, that is,

$$\operatorname{rank}P\Omega(\pi)P^* = \operatorname{rank}\Omega(\pi)P^*.$$

Thus, (1), (2), (3) and (4) are all equivalent. It is clear that (5) implies (1). Notice that, by (5.1), we have $P\cdot H(\pi)\supset H(\pi')$. So, (1) implies that P restricts to an isomorphism $H(\pi)\to H(\pi')$. To see that this isomorphism is symplectic, let ω and ω' be the symplectic forms on $H(\pi)$ and $H(\pi')$. Then, using (5.1) and the definition of ω and ω',

$$\omega'(P\Omega(\pi)\cdot e_x, P\Omega(\pi)\cdot e_y) = \omega'(\Omega(\pi')\cdot e_x, P\Omega(\pi)\cdot e_y) = \langle e_x, P\Omega(\pi)\cdot e_y\rangle$$
$$= \langle P^*\cdot e_x, \Omega(\pi)\cdot e_y\rangle = \omega(\Omega(\pi)\cdot e_x, \Omega(\pi)\cdot e_y)$$

for all $x,y\in\mathcal{A}'$. In view of (2), this implies that $P\colon H(\pi)\to H(\pi')$ is symplectic. $\square$

5.2. Simple extension

Let $\mathcal{A}'$ be an alphabet on $d\geqslant 2$ symbols and $\pi'\in\mathfrak{S}^0(\mathcal{A}')$. Let A be the first letter in the top and E be the first letter in the bottom. If $B\notin\mathcal{A}'$, let $\mathcal{A}=\mathcal{A}'\cup\{B\}$. Let $C,D\in\mathcal{A}'$ be such that $(A,E)\neq(C,D)$ (we allow $C=D$). Let $\pi=\mathcal{E}(\pi')$ be obtained by inserting B in π' just before C in the top and before D in the bottom. This operation is described by

$$\pi' = \begin{pmatrix} A & \cdot & C & \cdot & \cdot & \cdot \\ E & \cdot & \cdot & \cdot & D & \cdot \end{pmatrix} \longmapsto \pi = \begin{pmatrix} A & \cdot & B & C & \cdot & \cdot & \cdot & \cdot \\ E & \cdot & \cdot & \cdot & \cdot & B & D & \cdot \end{pmatrix}.$$

LEMMA 5.4. $\pi=\mathcal{E}(\pi')$ is irreducible.

Proof. Let $k<\#\mathcal{A}$. Suppose that the first k symbols in the top and the bottom of π coincide. Note that $k\geqslant 2$. If these symbols do not include B, then they are also the first k symbols in the top and the bottom of π'. Otherwise, removing B, we obtain the first $k-1$ symbols in the top and the bottom of π'. In either case, this contradicts the assumption that π' is irreducible. $\square$

This immediately extends to a map $\mathcal{E}$ defined (by the same rule) on the whole Rauzy class $\mathfrak{R}(\pi')$. We are going to see that $\mathcal{E}$ takes values in the Rauzy class of π. Given an

arrow γ' in $\mathfrak{R}(\pi')$, let $\gamma = \mathcal{E}_*(\gamma')$ be the path defined as follows.

(1) If C is last in the top of π' and γ' is a bottom arrow, then γ is the sequence of the two bottom arrows starting at the $\mathcal{E}$-image of the start of γ'. This is described by

$$\gamma': \begin{pmatrix} \cdot & * & \cdot & \cdot & \cdot & C \\ \cdot & \cdot & \cdot & D & \cdot & * \end{pmatrix} \longmapsto \begin{pmatrix} \cdot & * & C & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & D & \cdot & * \end{pmatrix},$$

$$\mathcal{E}_*(\gamma'): \begin{pmatrix} \cdot & * & \cdot & \cdot & \cdot & \cdot & B & C \\ \cdot & \cdot & \cdot & B & D & \cdot & \cdot & * \end{pmatrix} \longmapsto \begin{pmatrix} \cdot & * & C & \cdot & \cdot & \cdot & \cdot & B \\ \cdot & \cdot & \cdot & \cdot & B & D & \cdot & * \end{pmatrix}$$

$$\longmapsto \begin{pmatrix} \cdot & * & B & C & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & B & D & \cdot & * \end{pmatrix}.$$

(2) If D is last in the bottom of π' and γ' is a top arrow, then γ is the sequence of two bottom arrows starting at the $\mathcal{E}$-image of the start of γ'. This is analogous to the previous case.

(3) Otherwise, γ is the arrow of the same type as γ' starting at the $\mathcal{E}$-image of the start of γ'. For example,

$$\gamma': \begin{pmatrix} \cdot & * & \cdot & \cdot & \cdot & C & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & D & \cdot & \cdot & \cdot & \cdot & * \end{pmatrix} \longmapsto \begin{pmatrix} \cdot & * & \cdot & \cdot & \cdot & \cdot & C & \cdot & \cdot \\ \cdot & \cdot & \cdot & D & \cdot & \cdot & \cdot & \cdot & * \end{pmatrix},$$

$$\mathcal{E}_*(\gamma'): \begin{pmatrix} \cdot & * & \cdot & \cdot & \cdot & \cdot & B & C & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & B & D & \cdot & \cdot & \cdot & \cdot & \cdot & * \end{pmatrix}$$

$$\longmapsto \begin{pmatrix} \cdot & * & \cdot & \cdot & \cdot & \cdot & \cdot & B & C & \cdot & \cdot \\ \cdot & \cdot & \cdot & B & D & \cdot & \cdot & \cdot & \cdot & \cdot & * \end{pmatrix}.$$

In all cases, γ starts at the image by $\mathcal{E}$ of the start of γ' and ends at the image by $\mathcal{E}$ of the end of γ'. This shows that γ is a path in the Rauzy class of π, and that $\mathcal{E}$ takes values in $\mathfrak{R}(\pi)$. We extend $\mathcal{E}_*$ to a map $\Pi(\mathfrak{R}(\pi')) \to \Pi(\mathfrak{R}(\pi))$ in the whole space of paths, compatible with concatenation, and call $\mathcal{E}: \mathfrak{R}(\pi') \to \mathfrak{R}(\pi)$ and $\mathcal{E}_*: \Pi(\mathfrak{R}(\pi')) \to \Pi(\mathfrak{R}(\pi))$ *extension maps*.

Remark 5.5. If π is a simple extension of π', then π' is a simple reduction of π. The converse is true if and only if the omitted letter is not the last on the top nor on the bottom of π.

Observe that if $\gamma' = \mathcal{E}_*(\gamma)$ then

$$P\Theta(\gamma) = \Theta(\gamma')P. \tag{5.3}$$

Indeed, this may be rewritten as follows (recall that $P(e_B)=0$ and $P(e_x)=e_x$ for every $x \neq B$):

$$\langle \Theta(\gamma) \cdot e_x, e_y \rangle = \langle \Theta(\gamma') \cdot P(e_x), e_y \rangle \quad \text{for all } x \in \mathcal{A} \text{ and } y \in \mathcal{A}'. \tag{5.4}$$

It is enough to check the case when γ' is an arrow, because $\mathcal{E}_*$ is compatible with concatenation. In case (1) of the definition above, $\Theta(\gamma')\cdot e_*=e_C+e_*$ and $\Theta(\gamma')\cdot e_x=e_x$ for any $x\in\mathcal{A}'\backslash\{*\}$. On the other hand, $\Theta(\gamma)\cdot e_*=e_B+e_C+e_*$ and $\Theta(\gamma')\cdot e_x=e_x$ for any $x\in\mathcal{A}\backslash\{*\}$. In particular, $\Theta(\gamma)\cdot e_B=e_B$. The claim (5.3) follows in this case, and the other two are analogous.

LEMMA 5.6. *Let π be a simple extension of π', and assume that $g(\pi)=g(\pi')$. There exists a symplectic isomorphism $H(\pi)\rightarrow H(\pi')$ which conjugates the actions of $\mathcal{E}_*(\gamma')$ on $H(\pi)$ and of γ' on $H(\pi')$, for every $\gamma'\in\Pi(\pi')$.*

Proof. By Lemma 5.3, the natural projection $P\colon\mathbf{R}^{\mathcal{A}}\rightarrow\mathbf{R}^{\mathcal{A}'}$ restricts to a symplectic isomorphism $H(\pi)\rightarrow H(\pi')$. Then (5.3) shows that this isomorphism conjugates $\Theta(\gamma)|_{H(\pi)}$ and $\Theta(\gamma')|_{H(\pi')}$. $\qquad\square$

LEMMA 5.7. *Let $\pi\in\mathfrak{S}^0(\mathcal{A})$. If $\#\mathcal{A}\geqslant3$ then there exist $B\in\mathcal{A}$ and $\pi'\in\mathfrak{S}^0(\mathcal{A}\backslash\{B\})$ such that π is a simple extension of π'.*

Proof. Let A be the first in the top and E be the first in the bottom of π. Let

$$\begin{cases} B=E, & \text{if } \pi_0(E)<\pi_1(A), \\ B=A, & \text{if } \pi_0(E)>\pi_1(A), \\ B\in\{A,E\} \text{ be arbitrary}, & \text{if } \pi_0(E)=\pi_1(A)<\#\mathcal{A}, \\ B\in\mathcal{A}\backslash\{A,E\} \text{ be arbitrary}, & \text{if } \pi_0(E)=\pi_1(A)=\#\mathcal{A}. \end{cases}$$

Let π' be obtained by forgetting B on π. Notice that π' is irreducible. Indeed, suppose that the first k letters in the top coincide with the first k letters in the bottom, $1\leqslant k\leqslant d-2$. In the first case, one must have $k\geqslant\pi_1'(A)$ and then, adding the letter E to the list, one gets that, for π, the first $k+1$ letters in the top coincide with the first $k+1$ letters in the bottom. This contradicts the assumption that π is irreducible. A symmetric argument applies in the second case, and the same reasoning applies also in the third case. Finally, admissibility is obvious in the fourth case, because the permutation is standard. In all cases, π is a simple extension of π'. $\qquad\square$

5.3. Proof of Theorem 5.1

We prove the theorem by induction on the size $\#\mathcal{A}$ of the alphabet. By Corollary 3.6, the conclusion holds when $\#\mathcal{A}=2$, since in this case $\mathcal{L}(H(\pi))$ is just $\mathbf{P}H(\pi)$. Assume that it holds for $\#\mathcal{A}=d-1\geqslant2$, and let us show that it also does for $\#\mathcal{A}=d$.

By Lemma 5.7, there exists $\pi'\in\mathfrak{S}^0(\mathcal{A}')$, $\mathcal{A}'=\mathcal{A}\backslash\{B\}$, such that π is a simple extension of π'. If $g(\pi)=g(\pi')$, the result follows immediately from the induction hypothesis and Lemma 5.6. Thus, we may assume $g(\pi)=g(\pi')+1$. Let $H=H(\pi)$, $H'=H(\pi')$,

$\Omega=\Omega(\pi)$, $\Omega'=\Omega'(\pi)$, $\omega=\omega_\pi$ and $\omega'=\omega_{\pi'}$. If $\gamma'\in\Pi(\pi')$ then $\gamma=\mathcal{E}_*(\gamma')$ is contained in the stabilizer of e_B, that is $\Theta(\gamma)\cdot e_B=e_B$ (because B never wins).

LEMMA 5.8. *There is a symplectic isomorphism $H^{\lambda_B}\to H'$ that conjugates the action of $\mathcal{E}_*(\gamma')$ on H^{λ_B} and the action of γ' on H', for every $\gamma'\in\Pi(\pi')$.*

Proof. Write $h\in H$ as

$$h=\sum_{x\in\mathcal{A}}u_x(\Omega\cdot e_x)$$

with $u_x\in\mathbf{R}$. Then the condition $\omega(h,e_B)=0$ corresponds to

$$\omega(h,e_B)=\sum_{x\in\mathcal{A}}u_x\langle e_x,e_B\rangle=u_B=0.$$

In other words, $h\in H_{\lambda_B}$ if and only h belongs to the span of $\Omega\cdot e_x$, $x\in\mathcal{A}'$. Using (5.1) we obtain that $P\colon\mathbf{R}^\mathcal{A}\to\mathbf{R}^{\mathcal{A}'}$ takes H_{λ_B} to H'. The quotient $P\colon H^{\lambda_B}\to H'$ is symplectic: using again (5.1), the definition of ω and ω', and (5.4),

$$\omega'(P\Omega\cdot e_x,P\Omega\cdot e_y)=\omega'(\Omega'\cdot e_x,\Omega'\cdot e_y)=\langle e_x,\Omega'\cdot e_y\rangle=\langle e_x,\Omega\cdot e_y\rangle=\omega(\Omega\cdot e_x,\Omega\cdot e_y)$$

for all $x,y\in\mathcal{A}'$. Moreover, by (5.3), this map conjugates the action of γ on H^{λ_B} with the action of γ' on H'. $\square$

By the induction hypothesis, we conclude that $\mathcal{E}_*(\Pi(\pi'))\subset\Pi(\pi)$ acts minimally on $\mathcal{L}(H^{\lambda_B})$. By Lemma 2.7 and Corollary 3.6, this implies that the action of $\Pi(\pi)$ on $\mathcal{L}(H)$ is minimal. $\square$.

6. Pinching of Rauzy monoids

Our aim in this section is to prove the following result.

THEOREM 6.1. *Let π be irreducible. The action of the Rauzy monoid $\Pi(\pi)$ on $H(\pi)$ is strongly pinching.*

COROLLARY 6.2. *Let π be irreducible. The action of the Rauzy monoid $\Pi(\pi)$ on $H(\pi)$ is simple.*

Proof. This follows directly from Corollary 5.2, Theorem 6.1 and Lemma 4.14. $\square$

Remark 6.3. In view of Corollary 6.2, we may apply Remark 4.15 to the action of the Rauzy monoid on the space of flags: any closed invariant set in $\mathcal{F}(H(\pi))$ intersects the embedding of the space $\mathcal{L}(H(\pi))$ of Lagrangian flags. Since the action on the latter is minimal, by Theorem 5.1, we get that the only minimal set of the action of $\Pi(\pi)$ on the space of flags is the embedding of Lagrangian flags.

We start the proof of Theorem 6.1 by observing that the strong pinching property only depends on the Rauzy class.

LEMMA 6.4. *If π is such that the action of $\Pi(\pi)$ on $H(\pi)$ is strongly pinching and $\tilde{\pi} \in \mathfrak{R}(\pi)$ then the action of $\Pi(\tilde{\pi})$ on $H(\tilde{\pi})$ is strongly pinching.*

Proof. Let $\gamma_0 \in \Pi(\mathfrak{R}(\pi))$ be a path starting at $\tilde{\pi}$ and ending at π, and let $\gamma_1 \in \Pi(\mathfrak{R}(\pi))$ be a path starting at π and ending at $\tilde{\pi}$. Then $\Pi(\tilde{\pi}) \supset \gamma_0 \Pi(\pi) \gamma_1$, and we conclude, as in the first part of the proof of Lemma 4.16, that $\Pi(\tilde{\pi})$ is strongly pinching. $\square$

6.1. Minimal Rauzy classes

Let us call a Rauzy class $\mathfrak{R} \subset \mathfrak{S}^0(\mathcal{A})$ *minimal* if $\#\mathcal{A} = 2g(\mathfrak{R})$. Recall the definitions of degenerate permutation and good permutation in §3.13. In particular, these permutations are standard, by definition.

LEMMA 6.5. *Let $\pi \in \mathfrak{S}^0(\mathcal{A})$ be a degenerate permutation. Then $e_B \notin H(\pi)$, where B is the letter appearing in the second (or last but one) position of both top and bottom.*

Proof. Let $\Omega = \Omega(\pi)$. Let A be first in the top (last in the bottom) and E be the first in the bottom (last in the top). If $e_B \in H(\pi)$ then, by the equivalence of (3) and (4) in Lemma 5.3, we can write

$$e_B = \sum_{x \neq B} u_x (\Omega \cdot e_x) \quad \text{with } u_x \in \mathbf{R}.$$

Since π is standard, with first/last letters A and E, the definition (3.1) gives $\langle \Omega \cdot e_x, e_A \rangle = 1$ for all $x \neq A$, and $\langle \Omega \cdot e_x, e_E \rangle = -1$ for all $x \neq E$. Thus, the previous relation implies that

$$0 = \langle e_B, e_A \rangle = \sum_{x \neq A,B} u_x \quad \text{and} \quad 0 = \langle e_B, e_E \rangle = -\sum_{x \neq E,B} u_x.$$

This implies that $u_A - u_E = 0$. On the other hand,

$$1 = \langle e_B, e_B \rangle = \sum_{x \neq B} \langle \Omega, e_B \rangle = u_E - u_A,$$

because B is the second letter in both top and bottom (use (3.1) once more). This contradicts the previous conclusion, and this contradiction proves that $e_B \notin H(\pi)$. $\square$

LEMMA 6.6. *Any minimal Rauzy class $\mathfrak{R} \subset \mathfrak{S}^0(\mathcal{A})$ with $g(\mathfrak{R}) \geqslant 2$ contains a good permutation.*

Proof. By Lemma 3.8, if $\mathfrak{R}$ does not contain a good permutation, then it contains some degenerate permutation π. Let $B \in \mathcal{A}$ be as in Lemma 6.5, and let π' be obtained by forgetting B in π. Then π' is a simple reduction of π, and by the equivalence of (1) and (3) in Lemma 5.3, $g(\pi')=g(\pi)$. Thus $\#\mathcal{A} \geqslant 2g(\pi')+1=2g(\pi)+1$, and so $\mathfrak{R}$ is not minimal. $\qquad\square$

LEMMA 6.7. *If π is good and π'' is obtained by forgetting the first (and last) letters of the top and bottom rows, then $g(\pi) \leqslant g(\pi'')+1$. In particular, if $\mathfrak{R}(\pi)$ is minimal, then $\mathfrak{R}(\pi'')$ is also minimal.*

Proof. To prove the first claim, let A be first in the top (last in the bottom) and E be first in the bottom (last in the top) of π. Let π' be obtained by forgetting A in π. Then π'' is a simple reduction of π', which is itself a simple reduction of π. Let $\Omega=\Omega(\pi)$, $\Omega'=\Omega(\pi')$, $\Omega''=\Omega(\pi'')$, $g=g(\pi)$, $g'=g(\pi')$ and $g''=g(\pi'')$. If $g=g'$ then $g=g' \leqslant g''+1$, by Lemma 5.3, and the conclusion follows. Otherwise, $g=g'+1$ and, by the equivalence of (1) and (4) in Lemma 5.3, we can write

$$e_A = \sum_{x \neq A} u_x (\Omega \cdot e_x) \quad \text{with } u_x \in \mathbf{R}.$$

This implies that (recall the definition (3.1) of Ω)

$$0 = \langle e_A, e_E \rangle = - \sum_{x \neq A, E} u_x \quad \text{and} \quad 1 = \langle e_A, e_A \rangle = \sum_{x \neq A} u_x,$$

so that $u_E=1$. From (3.1) we have that $\Omega' \cdot e_x = \Omega \cdot e_x - e_A$ for all $x \neq A$. Consequently, using the previous equalities,

$$\Omega' \cdot e_E = \Omega \cdot e_E - \sum_{x \neq A} u_x (\Omega \cdot e_x) = - \sum_{x \neq A, E} u_x (\Omega \cdot e_x) = - \sum_{x \neq A, E} u_x (\Omega' \cdot e_x).$$

By the equivalence of (1) and (2) in Lemma 5.3, this gives $g'=g''$ and so $g=g''+1$.

The last claim in the lemma is an immediate consequence of the first one, because $\#\mathcal{A}=2g(\pi)$ and

$$\#\mathcal{A}-2 = \#(\mathcal{A} \setminus \{A, E\}) \geqslant 2g(\pi''),$$

and so the equality must hold if $g(\pi) \leqslant g(\pi'')+1$. $\qquad\square$

6.2. Reduction to the case of minimal Rauzy classes

LEMMA 6.8. *Let $\mathfrak{R}\subset\mathfrak{S}^0(\mathcal{A})$ be a Rauzy class such that $\#\mathcal{A}>2g(\mathfrak{R})$. Then there exists $\pi\in\mathfrak{R}$ which is a simple extension of some π' with $g(\pi)=g(\pi')$.*

Proof. By Lemma 3.7, we may consider some standard permutation $\tilde{\pi}\in\mathfrak{R}$. Since $\#\mathcal{A}>2g$, there exists $B\in\mathcal{A}$ such that $e_B\notin H(\tilde{\pi})$. Let $\mathcal{A}'=\mathcal{A}\backslash\{B\}$. Let A be the first letter in the top (last in the bottom) and E be the first letter in the bottom (last in the top) of $\tilde{\pi}$. If $B=A$, let π be obtained from $\tilde{\pi}$ by applying a top arrow. If $B=E$, let π be obtained from $\tilde{\pi}$ by applying a bottom arrow. Otherwise, let $\pi=\tilde{\pi}$. The first two cases of this definition are condensed in the following formula:

$$\tilde{\pi}=\begin{pmatrix} A & \cdot & \cdot & E \\ E & \cdot & \cdot & A \end{pmatrix} \longmapsto \pi=\begin{pmatrix} A & \cdot & \cdot & E \\ E & A & \cdot & \cdot \end{pmatrix} \quad \text{or} \quad \pi=\begin{pmatrix} A & E & \cdot & \cdot \\ E & \cdot & \cdot & A \end{pmatrix}.$$

In all cases, the fact that $e_B\notin H(\tilde{\pi})$ easily implies that $e_B\notin H(\pi)$. Let π' be obtained from π by forgetting B. Then π is a simple extension of π' and, from the equivalence of (1) and (3) in Lemma 5.3, we obtain that $g(\pi)=g(\pi')$. $\square$

Lemma 6.4 says that replacing any permutation by another one in the same Rauzy class does not affect the strong pinching property. By Lemma 6.8, if a class is non-minimal then it contains some permutation π which is a simple extension of some π' with $g(\pi)=g(\pi')$. Then, Lemma 5.6 says that the action of $\Pi(\pi')$ is conjugate to the action of a submonoid of $\Pi(\pi)$. It follows that strong pinching for $\Pi(\pi')$ implies strong pinching for $\Pi(\pi)$. Repeating this procedure, one eventually reaches a permutation in some minimal class. This means that these results reduce the proof of Theorem 6.1 to the case of minimal Rauzy classes.

We will also need a special formulation of this reduction which has a stronger conclusion, since it provides properties of a specific submonoid.

LEMMA 6.9. *Let $\pi'\in\mathfrak{S}^0(\mathcal{A}')$ be such that $E\in\mathcal{A}'$ is the last letter in the top and the first in the bottom. Let $\mathcal{A}''=\mathcal{A}'\backslash\{E\}$ and π'' be obtained from π' by forgetting E. Assume that π'' is irreducible and that $g(\pi')=g(\pi'')$. Assume also that the action of $\Pi(\pi'')$ on $H(\pi'')$ twists isotropic subspaces and is strongly pinching. Let $\widetilde{\Pi}(\pi')\subset\Pi(\pi')$ be the submonoid of all γ such that $\Theta(\gamma)\cdot e_E=e_E$. Then the action of $\widetilde{\Pi}(\pi')$ on $H(\pi')$ twists isotropic subspaces and is strongly pinching.*

Proof. Notice that π' is not a simple extension of π'', since E is the last letter in the top (recall Remark 5.5), and so we cannot apply Lemma 5.6 directly. Let $d=\#\mathcal{A}$. Let D be the last letter in the bottom for π'. Then D is in the kth position in the top row for some $1\leqslant k\leqslant d-1$. In fact, $k\leqslant d-2$, because we assume that π'' is irreducible. Let γ'_0

be the bottom arrow starting at π', and let $\tilde{\pi}'$ be the end of γ_0'. Let γ_1' be a sequence of $d-k-1$ bottoms starting at $\tilde{\pi}'$. Notice that γ_1' ends at π'. This is illustrated in the following formula:

$$\pi' = \begin{pmatrix} \cdot & \cdot & D & \cdot & \cdot & E \\ E & \cdot & \cdot & \cdot & \cdot & D \end{pmatrix} \overset{\gamma_0'}{\longmapsto} \tilde{\pi}' = \begin{pmatrix} \cdot & \cdot & D & E & \cdot & \cdot \\ E & \cdot & \cdot & \cdot & \cdot & D \end{pmatrix} \overset{\gamma_1'}{\longmapsto} \pi'.$$

Then $\tilde{\pi}'$ is a simple extension of π''. By Lemma 5.6, the action of $\mathcal{E}_*(\Pi(\pi''))\subset\Pi(\tilde{\pi}')$ on $H(\tilde{\pi}')$ is conjugate to the action of $\Pi(\pi'')$ on $H(\pi'')$, and so it also twists isotropic subspaces and is strongly pinching. Since $D\neq E$ is the winner of all arrows of γ_0' and γ_1', $\Theta(\gamma_0')$ and $\Theta(\gamma_1')$ preserve e_E. It follows that

$$\gamma_0'\mathcal{E}_*(\Pi(\pi''))\gamma_1'$$

is contained in the submonoid $\widetilde{\Pi}(\pi')$ that stabilizes e_E. We conclude, as in the proof of Lemma 4.16, that $\widetilde{\Pi}(\pi')$ twists isotropic subspaces and is strongly pinching. $\qquad\square$

6.3. Proof of Theorem 6.1

The proof is by induction on $\#\mathcal{A}$. The case $\#\mathcal{A}=2$ is easy because any arrow, top or bottom, gives a parabolic element for the action. Let us show that if strong pinching holds for $\#\mathcal{A}=d-1\geqslant 2$ then it holds for $\#\mathcal{A}=d$.

Let $\pi\in\mathfrak{S}^0(\mathcal{A})$. As explained before, if $d>2g(\pi)$ then the result follows from the induction hypothesis using Lemmas 5.6, 6.4 and 6.8. So, we may assume that $d=2g(\pi)$. Let $g=g(\pi)$. Notice that $g\geqslant 2$ because $d>2$. By Lemma 6.4, the conclusion does not change if we replace π by any other permutation in the same Rauzy class. By Lemma 6.6, the Rauzy class contains some good permutation $\tilde{\pi}\in\mathfrak{R}(\pi)$. Hence, we may suppose from the start that π is the permutation obtained by applying a top arrow to $\tilde{\pi}$. Lemma 4.7 reduces the proof that the action of $\Pi(\pi)$ on $H(\pi)=\mathbf{R}^{\mathcal{A}}$ is strongly pinching to proving the following result.

LEMMA 6.10. *For every $C>0$ there exists $B\in\Theta(\Pi(\pi))$ for which 1 is an eigenvalue of geometric multiplicity 1 (the eigenspace has dimension 1) and*

$$\theta_{g-1}(B)>0 \quad and \quad \theta_k(B)>C\theta_{k+1}(B) \ \ for \ all \ \ 1\leqslant k\leqslant g-2. \tag{6.1}$$

Proof. Let A be first in the top (last in the bottom) and E be first in the bottom (last in the top) of π. Let π' be obtained by forgetting A from π and $\mathcal{A}'=\mathcal{A}\setminus\{A\}$. Then E is last in the top and first in the bottom of π', and this ensures that π' is irreducible.

Moreover, π is a simple extension of π', and π' is a simple reduction of both π and $\tilde{\pi}$. This is illustrated in the formula that follows:

$$\tilde{\pi} = \begin{pmatrix} A & \cdot & \cdot & E \\ E & \cdot & \cdot & A \end{pmatrix}, \quad \pi = \begin{pmatrix} A & \cdot & \cdot & E \\ E & A & \cdot & \cdot \end{pmatrix}, \quad \pi' = \begin{pmatrix} \cdot & \cdot & E \\ E & \cdot & \cdot \end{pmatrix}.$$

Let $\gamma_0' \in \Pi(\pi')$ be the sequence of $2g-2$ top arrows starting (and ending) at π', and let $\gamma_0 = \mathcal{E}_*(\gamma_0')$. Then $\gamma_0 \in \Pi(\pi)$ is a sequence of $2g-1$ top arrows starting and ending at π. Notice that E is the winner in all these arrows, and so

$$\Theta(\gamma_0) = \begin{pmatrix} 1 & 0 & \dots & 0 & 1 \\ 0 & 1 & \dots & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

where the column made of 1's is the one indexed by E. More precisely,

$$\Theta(\gamma_0) \cdot e_x = e_x \qquad \text{for all } x \in \mathcal{A} \setminus \{E\}, \tag{6.2}$$

$$\Theta(\gamma_0) \cdot e_E = \sum_{x \in \mathcal{A}} e_x. \tag{6.3}$$

Now let $\widetilde{\Pi}(\pi') \subset \Pi(\pi')$ be the set of all γ' such that $\Theta(\gamma') \cdot e_E = e_E$. We claim that the action of $\widetilde{\Pi}(\pi')$ on $H(\pi')$ twists isotropic subspaces and is strongly pinching. This is easily obtained from Lemma 6.9, as follows. Let π'' be obtained from π' by forgetting the letter E. Equivalently, π'' is obtained from the good permutation $\tilde{\pi}$ by forgetting A and E. According to Lemma 6.7, $g(\tilde{\pi}) = g(\pi'') + 1$. We also have

$$2g(\pi'') \leqslant 2g(\pi') \leqslant \#\mathcal{A} - 1 = 2g(\tilde{\pi}) - 1.$$

Consequently, $g(\pi'') = g(\pi')$. This means that we are in a position to apply Lemma 6.9: since the action of $\Pi(\pi'')$ on $H(\pi'')$ is strongly pinching and twists isotropic subspaces, by the induction hypothesis and Corollary 5.2, the same is true for the action of $\widetilde{\Pi}(\pi')$ on $H(\pi')$, as claimed. Then we may apply Lemma 4.8 to find $\gamma' \in \widetilde{\Pi}(\pi')$ such that (take $x = \gamma_0'$ and $x_0 = \gamma'$ in the lemma) 1 is not an eigenvalue of γ' acting on $H(\pi')$, and the Lyapunov exponents of $\gamma' \gamma_0'$ acting on $H(\pi')$ satisfy

$$\theta_{g-1}(\gamma' \gamma_0') > 0 \quad \text{and} \quad \theta_i(\gamma' \gamma_0') > C \theta_{i+1}(\gamma' \gamma_0') \text{ for all } 1 \leqslant i \leqslant g-2. \tag{6.4}$$

Let us show that $B = \Theta(\mathcal{E}_*(\gamma' \gamma_0'))$ has the required properties.

Write $\gamma = \mathcal{E}_*(\gamma')$, so that $B = \Theta(\gamma_0)\Theta(\gamma)$. Since E is never a winner for any of the arrows that form γ' (because γ' is chosen in the stabilizer of e_E), it is never a winner for γ either. Moreover, noting that E is first and A is second in the bottom of π, we see that A is never a winner nor a loser for γ. Together with (6.2) and (6.3), this gives

$$\Theta(\gamma) \cdot e_A = e_A \quad \text{and} \quad \Theta(\gamma) \cdot e_E = e_E, \tag{6.5}$$

$$\langle \Theta(\gamma) \cdot e_x, e_A \rangle = 0 \quad \text{for all } x \in \mathcal{A} \setminus \{A\}. \tag{6.6}$$

In other words, the matrix of $\Theta(\gamma)$ has the form

$$\Theta(\gamma) = \begin{pmatrix} 1 & 0 & \cdot & 0 & 0 \\ 0 & * & \cdot & * & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & * & \cdot & * & 0 \\ 0 & * & \cdot & * & 1 \end{pmatrix}. \tag{6.7}$$

This implies that $B = \Theta(\gamma_0)\Theta(\gamma)$ satisfies

$$B \cdot e_A = e_A, \tag{6.8}$$

$$\langle B \cdot e_x, e_E \rangle = \langle B \cdot e_x, e_A \rangle \quad \text{for all } x \in \mathcal{A} \setminus \{A\}. \tag{6.9}$$

By (6.8), we have that 1 is an eigenvalue of B. Let us check that its geometric multiplicity is 1. Indeed, otherwise there would exist $h \in \mathbf{R}^{\mathcal{A}} \setminus \{0\}$ such that $B \cdot h = h$ and $\langle h, e_A \rangle = 0$. By (6.9),

$$\langle h, e_E \rangle = \langle B \cdot h, e_E \rangle = \langle B \cdot h, e_A \rangle = \langle h, e_A \rangle = 0.$$

Let $P \colon \mathbf{R}^{\mathcal{A}} \to \mathbf{R}^{\mathcal{A}'}$ be the natural projection and $h' = P \cdot h$. Notice that $P^* \cdot h' = h$, because h is orthogonal to e_A. In particular, h' is non-zero. We have

$$\Theta(\gamma_0') = P\Theta(\gamma_0)P^*, \quad \Theta(\gamma') = P\Theta(\gamma)P^* \quad \text{and} \quad \Theta(\gamma_0')\Theta(\gamma') = PBP^*. \tag{6.10}$$

This implies that

$$\Theta(\gamma_0')\Theta(\gamma') \cdot h' = PBP^* \cdot h' = PB \cdot h = P \cdot h = h'. \tag{6.11}$$

We also have $\langle h', e_E \rangle = \langle h, e_E \rangle = 0$. Now, (6.2) and (6.10) imply that $\Theta(\gamma_0') \cdot e_x = e_x$ for all $x \in \mathcal{A}' \setminus \{E\}$. Consequently,

$$\Theta(\gamma_0') \cdot h' = h' \quad \text{or, equivalently,} \quad h' = \Theta(\gamma_0')^{-1} \cdot h'. \tag{6.12}$$

From (6.11) and (6.12) we immediately get that

$$\Theta(\gamma') \cdot h' = h'. \tag{6.13}$$

But (6.5) implies that $\Theta(\gamma') \cdot e_E = P\Theta(\gamma)P^* \cdot e_E = e_E$, and so (6.13) implies that 1 is an eigenvalue of $\Theta(\gamma')$ with geometric multiplicity at least 2. Since $H(\pi')$ has codimension 1 in $\mathbf{R}^{\mathcal{A}'}$, this implies that 1 is an eigenvalue of $\Theta(\gamma')$ acting on $H(\pi')$, which contradicts the choice of γ'. This proves that 1 is an eigenvalue of B with geometric multiplicity 1, as claimed. To obtain the properties in (6.1), notice that (6.8) and (6.10) imply that the matrix of B has a block triangular form, with the matrix of $\Theta(\gamma'_0)\Theta(\gamma')$ as a diagonal block. It follows that the eigenvalues of B are, precisely, the eigenvalues of $\Theta(\gamma'_0)\Theta(\gamma')$, together with an additional eigenvalue 1 associated with the eigenvector e_A. Observe that this eigenvalue must have *algebraic* multiplicity 2, since the action is symplectic. Therefore, the Lyapunov spectrum of B consists of the Lyapunov spectrum of $\Theta(\gamma'_0)\Theta(\gamma')$ on $H(\pi')$ together with two additional zero Lyapunov exponents. Thus, (6.4) implies (6.1). This finishes the proof of Lemma 6.10. □

At this point the proof of Theorem 6.1 is complete.

Remark 6.11. In the above argument we concatenate two Rauzy paths to generate another one exhibiting some parabolic behavior. It would be interesting to find a geometric interpretation, in terms of the Teichmüller flow, of how that behavior arises.

Remark 6.12. Some of the richness properties we have been discussing, like twisting, depend only on the group generated by the monoid (because invariance under the monoid is equivalent to invariance under the group generated), while others, such as strong pinching, depend on the monoid itself. It is interesting to investigate how large the generated group is. The only obvious restriction is that this group preserves the integer lattice $H(\pi) \cap \mathbf{Z}^{\mathcal{A}}$. What is its Zariski closure (which plays a role in [GM])? Such questions had already been raised by Zorich in [Zo4, Appendix A.3], where he made some specific conjectures. We believe it is possible to proceed further in this direction with the arguments of the present paper, particularly our induction procedure in terms of relations between Rauzy classes.

7. Simplicity of the spectrum

7.1. A criterion for simplicity of the spectrum

Let (Δ, μ_0) be a probability space, and let $T : \Delta \to \Delta$ be a transformation such that there exists a finite or countable partition of Δ into sets $\Delta^{(l)}$, $l \in \Lambda$, of positive measure such that $T : \Delta^{(l)} \to \Delta$ is an invertible measurable transformation and $T_*(\mu_0|_{\Delta^{(l)}})$ is equivalent to μ_0. Let Ω be the set of finite sequences of elements of Λ, including the empty sequence. If $\underline{l} = (l_1, ..., l_m) \in \Omega$, let

$$\Delta^{\underline{l}} = \{x \in \Delta : T^k(x) \in \Delta^{(l_{k+1})} \text{ for } 0 \leqslant k < m\},$$

and $T^{\underline{l}}=T^m\colon\Delta^{\underline{l}}\to\Delta$. The $\Delta^{\underline{l}}$'s have positive measure and $T^{\underline{l}}$ is an invertible measurable transformation. We say that (T,μ_0) has *approximate product structure* if there exists $C>0$ such that

$$\frac{1}{C}\leqslant\frac{1}{\mu_0(\Delta^{\underline{l}})}\frac{dT^{\underline{l}}_*(\mu_0|\Delta^{\underline{l}})}{d\mu_0}\leqslant C\quad\text{for all }\underline{l}\in\Omega.\tag{7.1}$$

Notice that T is measurable with respect to the *cylinder σ-algebra* of T, that is, the σ-algebra generated by the $\Delta^{\underline{l}}$'s, $\underline{l}\in\Omega$. It is not difficult to check, using (7.1), that T is ergodic with respect to the cylinder σ-algebra, and there is a unique probability measure μ on the cylinder σ-algebra which is invariant under T and is absolutely continuous with respect to μ_0. Moreover,

$$\frac{1}{C}\leqslant\frac{d\mu}{d\mu_0}\leqslant C$$

and (T,μ) has approximate product structure as well.

Let (T,μ) have approximate product structure and H be some finite-dimensional vector space. Let $A^{(l)}\in\mathrm{SL}(H)$, $l\in\Lambda$, and define $A\colon\Delta\to\mathrm{SL}(H)$ by $A(x)=A^{(l)}$ if $x\in\Delta^{(l)}$. We will say that (T,A) is a *locally constant cocycle*. The *supporting monoid* of (T,A) is the monoid generated by the $A^{(l)}$'s, $l\in\Lambda$. By ergodicity and the Oseledets theorem [Os], if the cocycle is measurable, that is, if $\int\log\|A(x)\|\,d\mu(x)<\infty$, then it has a well-defined Lyapunov spectrum.

THEOREM 7.1. *Let (T,A) be a locally constant measurable cocycle. If the supporting monoid is simple then the Lyapunov spectrum is simple.*

This is an adaptation of the main result in [AV] to the present situation. For completeness, a proof is provided in Appendix A.

7.2. Locally constant projective cocycles

Fix any $p\geqslant 2$. We call $\mathbf{PR}^p_+$ the *standard simplex*. A *projective contraction* is a projective transformation taking the standard simplex into itself or, in other words, it is the projectivization of some matrix $B\in\mathrm{GL}(p,\mathbf{R})$ with non-negative entries. The image of the standard simplex by a projective contraction is called a *simplex*.

A *projective expanding map* T is a map $\bigcup\Delta^{(l)}\to\Delta$, where Δ is a simplex compactly contained in the standard simplex, the $\Delta^{(l)}$'s form a finite or countable family of pairwise disjoint simplexes contained in Δ and covering almost all of Δ, and $T^{(l)}=T\colon\Delta^{(l)}\to\Delta$ is a bijection such that $(T^{(l)})^{-1}$ is the restriction of a projective contraction.

LEMMA 7.2. *If $T\colon\bigcup\Delta^{(l)}\to\Delta$ is a projective expanding map then it has approximate product structure with respect to Lebesgue measure.*

Proof. This follows from the well-known fact that the logarithm of the Jacobian of a projective contraction relative to Lebesgue measure is (uniformly) Lipschitz with respect to the projective metric. See the proof of [AF, Lemma 2.1] or [AV, Appendix] for details. $\square$

7.3. Proof of Theorem 1.1

Let $\gamma \in \Pi(\mathfrak{R})$ be any path such that $\mathbf{P}\Delta_\gamma$ is compactly contained in $\mathbf{P}\Delta_{\mathfrak{R}}^1$. This is easy to satisfy: for every $x \in \mathbf{P}\Delta_{\mathfrak{R}}$, the connected component of x in $\mathbf{P}\Delta_{\mathfrak{R}}^n$ is compactly contained in $\mathbf{P}\Delta_{\mathfrak{R}}^1$ for every n sufficiently large. Let T be the first return map to $\mathbf{P}\Delta_\gamma$ under the Zorich map. We define a linear skew-product $(x, h) \mapsto (T(x), A(x) \cdot h)$ over $T : \mathbf{P}\Delta_\gamma \to \mathbf{P}\Delta_\gamma$, by setting $A(x) = B^Z(R_Z^{m-1}(x)) \dots B^Z(x)$, where $m = m(x)$ is the first return time of x to $\mathbf{P}\Delta_\gamma$. The map T preserves the probability measure

$$\mu = \frac{1}{\mu_Z(\mathbf{P}\Delta_\gamma)} \mu_Z|_{\mathbf{P}\Delta_\gamma},$$

and the Lyapunov exponents of (T, A) are obtained by multiplying the Lyapunov exponents of the cocycle (R_Z, B^Z) by $1/\mu_Z(\mathbf{P}\Delta_\gamma)$. Notice that T is a projective expanding map. So, by Lemma 7.2, (T, A) is a locally constant cocycle.

Let π' be the start of γ. Let γ_0 be an arrow ending at π' and which is either top or bottom according to whether γ starts by a bottom or a top arrow. Let π be the start of γ_0. Let γ_1 be a path starting by γ and ending at π. Then the monoid $\mathcal{B}$ generated by the $A(x)$ contains $\Theta(\gamma_1 \Pi(\pi) \gamma_0)$. Hence $\mathcal{B}$ is a submonoid of $\Theta(\Pi(\pi'))$ containing $\Theta(\gamma_1 \gamma_0 \Pi(\pi') \gamma_1 \gamma_0)$. Since, the action of $\Theta(\Pi(\pi'))$ on $H(\pi')$ is simple, Lemma 4.16 gives that the monoid $\mathcal{B}$ generated by the $A(x)$ is also simple. Now Theorem 1.1 is a consequence of Theorem 7.1.

Appendix A. Proof of Theorem 7.1

In this appendix we prove our sufficient criterion for simplicity of the Lyapunov spectrum. See also [AV].

A.1. Symbolic dynamics

Let us say that two locally constant cocycles (T, A) and (T', A') are *equivalent* if

$$\mu(\Delta^{\underline{l}}) = \mu'(\Delta'^{\underline{l}}) \text{ for all } \underline{l} \in \Omega \quad \text{and} \quad A^{(l)} = A'^{(l)} \text{ for all } l \in \Lambda.$$

It is clear that two measurable locally constant cocycles have the same Lyapunov spectrum if they are equivalent. Any locally constant cocycle is equivalent to a *symbolic* locally constant cocycle, that is, one for which the base transformation is the shift

$$\{x_i\}_{i=0}^{\infty} \longmapsto \{x_{i+1}\}_{i=0}^{\infty}$$

over $\Lambda^{\mathbf{N}}$, where $\mathbf{N} = \{n \in \mathbf{Z} : n \geqslant 0\}$. Thus, it is enough to prove Theorem 7.1 in the symbolic case. Let us introduce some convenient notation to deal with this case.

Let $\Sigma = \Lambda^{\mathbf{N}}$. Given $\underline{l} = (l_1, ..., l_m) \in \Omega$, let $\Sigma^{\underline{l}}$ be the set of $\{x_i\}_{i=0}^{\infty}$ such that $x_i = l_{i+1}$ for $0 \leqslant i < m$. We also write $|\underline{l}| = m$. Let $f : \Sigma \to \Sigma$ be the shift map. A cylinder on Σ is a set of the form $f^{-n}(\Sigma^{\underline{l}})$ with $n \in \mathbf{N}$. Denote by $f^{\underline{l}}$ the restriction of f^m to $\Sigma^{\underline{l}}$. A probability measure on Σ is defined once specified on cylinders. Thus, an f-invariant probability measure is completely determined by its value on the $\Sigma^{\underline{l}}$'s. Given an f-invariant probability measure μ, let $\phi_\mu : \Omega \to \mathbf{R}^+$ be defined by $\phi_\mu(\underline{l}) = \mu(\Sigma^{\underline{l}})$. For $\underline{l}_1 = (l_1, ..., l_m)$ and $\underline{l}_2 = (l_{m+1}, ..., l_{m+n})$ we set $\underline{l}_1 \underline{l}_2 = (l_1, ..., l_{m+n})$.

We will say that μ has *bounded distortion* for f if it gives positive measure to all cylinders and there exists $C(\mu) > 0$ such that

$$\frac{1}{C(\mu)} \leqslant \frac{\phi_\mu(\underline{l}_1 \underline{l}_2)}{\phi_\mu(\underline{l}_1)\phi_\mu(\underline{l}_2)} \leqslant C(\mu) \quad \text{for all } \underline{l}_1, \underline{l}_2 \in \Omega.$$

Notice that this last condition is equivalent to

$$\frac{1}{C(\mu)} \leqslant \frac{1}{\mu(\Sigma^{\underline{l}})} \frac{df^{\underline{l}}_*(\mu|_{\Sigma^{\underline{l}}})}{d\mu} \leqslant C(\mu) \quad \text{for all } \underline{l} \in \Omega.$$

Hence, μ has bounded distortion for f if and only if (f, μ) has approximate product structure.

A.2. The inverse limit

Let $\Sigma_- = \Lambda^{\mathbf{Z} \backslash \mathbf{N}}$ and $\widehat{\Sigma} = \Lambda^{\mathbf{Z}}$. Then, $\widehat{\Sigma} = \Sigma_- \times \Sigma$. Let π_- and π be the coordinate projections. Let $\hat{f} : \widehat{\Sigma} \to \widehat{\Sigma}$ be the shift map. Then $\pi \circ \hat{f} = f \circ \pi$. Let $\widehat{\Sigma}^{\underline{l}} = \Sigma_- \times \Sigma^{\underline{l}}$ and $\widehat{\Sigma}^{\underline{l}}_- = \hat{f}^m(\Sigma^{\underline{l}})$, with $m = |\underline{l}|$, and $\Sigma^{\underline{l}}_- \subset \Sigma_-$ be such that $\widehat{\Sigma}^{\underline{l}}_- = \Sigma^{\underline{l}}_- \times \Sigma$. A cylinder on $\widehat{\Sigma}$ is a set of the form $\hat{f}^n(\widehat{\Sigma}^{\underline{l}}_-)$, $n \in \mathbf{Z}$.

A probability measure on $\widehat{\Sigma}$ is defined once specified on cylinders. An $\hat{f}$-invariant probability measure is, thus, completely determined by its values on the $\widehat{\Sigma}^{\underline{l}}$'s. Thus, there is a bijection $\hat{\mu} \mapsto \mu$ between $\hat{f}$- and f-invariant probability measures given by $\pi_* \hat{\mu} = \mu$. We call $\hat{\mu}$ the *lift* of μ. We say that $\hat{\mu}$ has bounded distortion for $\hat{f}$ if μ has bounded distortion for f. This implies that $\hat{f}^n$ is ergodic with respect to $\hat{\mu}$ for every $n \geqslant 1$. We also denote by μ_- the projection of $\hat{\mu}$ to Σ_-.

A.3. Invariant section

Fix some probability measure μ with bounded distortion and let $\hat{\mu}$ be its lift. Given a locally constant measurable cocycle (f, A), we may consider its lift $(\hat{f}, \hat{A})$, defined by $\hat{A}(x) = A(\pi(x))$. It naturally acts on $\widehat{\Sigma} \times \mathrm{Grass}(k, H)$, for any $1 \leqslant k \leqslant \dim H - 1$. Let k be fixed. Given $\underline{l} \in \Omega$, let $\xi_{\underline{l}} = A^{\underline{l}} \cdot E_k^+(A^{\underline{l}})$ if defined, otherwise choose $\xi_{\underline{l}}$ arbitrarily. For $x \in \Sigma_-$ and $n \in \mathbf{N}$, let $\underline{l}(x, n) \in \Omega$ be such that $|\underline{l}(x, n)| = n$ and $x \in \Sigma_-^{\underline{l}(x,n)}$.

THEOREM A.1. *Let (f, A) be a symbolic locally constant measurable cocycle, and assume that the supporting monoid is simple. Then,*
    (1) *there exists a measurable function*

$$\xi \colon \Sigma_- \longrightarrow \mathrm{Grass}(k, H)$$

such that $\hat{\xi} = \xi \circ \pi_-$ satisfies $\hat{\xi}(\hat{f}(x)) = \hat{A}(x) \cdot \hat{\xi}(x)$;
    (2) *for μ_--almost all $x \in \Sigma_-$ we have*

$$\log \sigma_k(A^{\underline{l}(x,n)}) - \log \sigma_{k+1}(A^{\underline{l}(x,n)}) \to \infty \quad and \quad \xi_{\underline{l}(x,n)} \to \xi(x);$$

    (3) *for every hyperplane section $S \subset \mathrm{Grass}(k, H)$ there exists a positive μ_--measure set of $x \in \Sigma_-$ such that $\xi(x) \notin S$.*

The proof of this theorem will take several steps.

A.4. Measures on Grassmannians

Let $\widehat{m}$ be a probability measure on $\widehat{\Sigma} \times \mathrm{Grass}(k, H)$. We say that $\widehat{m}$ is a *u-state* if its projection on $\widehat{\Sigma}$ is $\hat{\mu}$ and there exists $C(\widehat{m}) > 0$ such that for any Borel set $X \subset \mathrm{Grass}(k, H)$ and any $\underline{l}_0, \underline{l}, \underline{l}' \in \Omega$, we have

$$\frac{1}{\mu(\Sigma^{\underline{l}'})} \widehat{m}(\Sigma_-^{\underline{l}_0} \times \Sigma^{\underline{l}'} \times X) \leqslant C(\widehat{m}) \frac{1}{\mu(\Sigma^{\underline{l}})} \widehat{m}(\Sigma_-^{\underline{l}_0} \times \Sigma^{\underline{l}} \times X).$$

It is easy to give examples of *u*-states: take any probability measure ν in $\mathrm{Grass}(k, H)$ and let $\widehat{m} = \hat{\mu} \times \nu$ (in this case $C(\widehat{m}) = C(\mu)^2$ holds). From this we can get examples of *u*-states invariant under $(\hat{f}, \hat{A})$, as follows.

Let $\mathcal{N}$ be the space of probability measures on $\widehat{\Sigma} \times \mathrm{Grass}(k, H)$ that project to $\hat{\mu}$ on $\widehat{\Sigma}$. We introduce on $\mathcal{N}$ the smallest topology such that the map $\eta \mapsto \int \psi \, d\eta$ is continuous, for every bounded continuous function $\psi \colon \widehat{\Sigma} \times \mathrm{Grass}(k, H) \to \mathbf{R}$. We will call this the weak* topology. Notice that $\mathcal{N}$ is a compact separable space. This is easy to see in the following alternative description of the topology. Let $K_n \subset \widehat{\Sigma}$, $n \geqslant 1$, be disjoint compact

sets such that $\hat{\mu}(K_n)>0$ and $\sum_{n=1}^{\infty}\hat{\mu}(K_n)=1$. Let $\mathcal{N}_n$ be the space of measures on $K_n\times\mathrm{Grass}(k,H)$ that project to $\hat{\mu}|_{K_n}$. The usual weak* topology makes $\mathcal{N}_n$ into a compact separable space. Given $\eta\in\mathcal{N}$, let $\eta_n\in\mathcal{N}_n$ be obtained by restriction of η. This identifies $\mathcal{N}$ with $\prod_{n=1}^{\infty}\mathcal{N}_n$, and the weak* topology on $\mathcal{N}$ corresponds to the product topology on $\prod_{n=1}^{\infty}\mathcal{N}_n$.

LEMMA A.2. *Let $\widehat{m}_0$ be a u-state. For every $n\geqslant 0$,*

$$\widehat{m}_0^{(n)}=(\hat{f},\hat{A})_*^n(\widehat{m}_0)$$

is a u-state and $C(\widehat{m}_0^{(n)})\leqslant C(\widehat{m}_0)C(\mu)^2$.

Proof. We must show, for $\underline{l}_0,\underline{l},\underline{l}'\in\Omega$ and $X\subset\mathrm{Grass}(k,H)$ measurable, that

$$\frac{\mu(\Sigma^{\underline{l}})}{\mu(\Sigma^{\underline{l}'})}\widehat{m}_0^{(n)}(\Sigma_-^{\underline{l}_0}\times\Sigma^{\underline{l}'}\times X)\leqslant C(\widehat{m}_0)C(\mu)^2\widehat{m}_0^{(n)}(\Sigma_-^{\underline{l}_0}\times\Sigma^{\underline{l}}\times X).$$

It is enough to consider the case where $|\underline{l}_0|\geqslant n$. In this case write $\underline{l}_0=\underline{l}_1\underline{l}_2$ with $|\underline{l}_2|=n$. Then,

$$\begin{aligned}\frac{\mu(\Sigma^{\underline{l}})}{\mu(\Sigma^{\underline{l}'})}\widehat{m}_0^{(n)}(\Sigma_-^{\underline{l}_0}\times\Sigma^{\underline{l}'}\times X)&=\frac{\mu(\Sigma^{\underline{l}})}{\mu(\Sigma^{\underline{l}'})}\widehat{m}_0(\Sigma_-^{\underline{l}_1}\times\Sigma^{\underline{l}_2\underline{l}'}\times(A^{\underline{l}_2})^{-1}\cdot X)\\
&\leqslant C(\widehat{m}_0)\frac{\mu(\Sigma^{\underline{l}})}{\mu(\Sigma^{\underline{l}'})}\frac{\mu(\Sigma^{\underline{l}_2\underline{l}'})}{\mu(\Sigma^{\underline{l}_2\underline{l}})}\widehat{m}_0(\Sigma_-^{\underline{l}_1}\times\Sigma^{\underline{l}_2\underline{l}}\times(A^{\underline{l}_2})^{-1}\cdot X)\\
&\leqslant C(\widehat{m}_0)C(\mu)^2\widehat{m}_0(\Sigma_-^{\underline{l}_1}\times\Sigma^{\underline{l}_2\underline{l}}\times(A^{\underline{l}_2})^{-1}\cdot X)\\
&=C(\widehat{m}_0)C(\mu)^2\widehat{m}_0^{(n)}(\Sigma_-^{\underline{l}_0}\times\Sigma^{\underline{l}}\times X),\end{aligned}$$

as claimed. $\square$

For any fixed $C>0$, the space of u-states $\widehat{m}$ with $C(\widehat{m})\leqslant C$ is a convex compact subset in the weak* topology. So, the next statement implies the existence of invariant u-states.

COROLLARY A.3. *Let $\widehat{m}_0$ be a u-state. Let $\widehat{m}$ be a Cesàro weak* limit of the sequence*

$$(\hat{f},\hat{A})_*^n(\widehat{m}_0).$$

Then $\widehat{m}$ is an invariant u-state with $C(\widehat{m})\leqslant C(\widehat{m}_0)C(\mu)^2$.

LEMMA A.4. *Let $\widehat{m}$ be a probability measure on $\widehat{\Sigma}\times\mathrm{Grass}(k,H)$. For $x\in\Sigma_-$, let $\widehat{m}^{(n)}(x)$ be the probability measure on $\mathrm{Grass}(k,H)$ obtained by normalized projection of $\widehat{m}$ restricted to $\Sigma_-^{\underline{l}(x,n)}\times\Sigma\times\mathrm{Grass}(k,H)$. Then, for μ_--almost every x, the sequence $\widehat{m}^{(n)}(x)$ converges in the weak* topology to some probability measure $\widehat{m}(x)$.*

Proof. Since $\mathrm{Grass}(k, H)$ is a compact metric space, and keeping in mind the definition of the weak* topology, it is enough to show that if $\phi\colon \mathrm{Grass}(k, H) \to \mathbf{R}$ is continuous, then for almost every $x \in \Sigma_-$ the integral $\phi^{(n)}(x)$ of ϕ with respect to $\widehat{m}^{(n)}(x)$ converges as $n \to \infty$. This is a simple application of the martingale convergence theorem. A direct proof goes as follows.

We may assume that $0 \leqslant \phi(z) \leqslant 1$, $z \in \mathrm{Grass}(k, H)$. Consider the measure $\nu \leqslant \mu_-$ defined by

$$\nu(X) = \int_{X \times \Sigma \times \mathrm{Grass}(k, H)} \phi(z)\, d\widehat{m}(x, y, z).$$

Then, $\nu(\Sigma_-^{\underline{l}(x,n)}) = \phi^{(n)}(x)\mu_-(\Sigma_-^{\underline{l}(x,n)})$. Let

$$\phi^+(x) = \limsup_{n \to \infty} \phi^{(n)}(x) \quad \text{and} \quad \phi^-(x) = \liminf_{n \to \infty} \phi^{(n)}(x).$$

We have to show that for every $a < b$ the set

$$U_{a,b} = \{x \in \Sigma_- : \phi^-(x) < a < b < \phi^+(x)\}$$

has zero μ_--measure. To do this it is enough to show that $\nu(U_{a,b}) \geqslant b\mu_-(U_{a,b}) - \varepsilon$ and $\nu(U_{a,b}) \leqslant a\mu_-(U_{a,b}) + \varepsilon$ for every $\varepsilon > 0$. We will show only the first inequality, the second one being analogous. Fix an open set $U \supset U_{a,b}$ with $\mu_-(U \setminus U_{a,b}) < \varepsilon$. Let

$$N(x) = \min\{n : \Sigma_-^{\underline{l}(x,n)} \subset U\} \quad \text{and} \quad n(x) = \min\{n \geqslant N(x) : \phi^{(n)}(x) > b\}.$$

Let $U^N = \bigcup_{x \in U_{a,b}} \Sigma_-^{\underline{l}(x, n(x))}$. Then $U_{a,b} \subset U^N \subset U$ and there exists a (finite or countable) sequence $x_j \in U_{a,b}$ such that $U^N = \bigcup_j \Sigma_-^{\underline{l}(x_j, n(x_j))}$. In particular, $\nu(U^N) - \nu(U_{a,b}) < \varepsilon$ and

$$\nu(U^N) = \sum_j \nu\big(\Sigma_-^{\underline{l}(x_j, n(x_j))}\big) \geqslant \sum_j b\mu_-\big(\Sigma_-^{\underline{l}(x_j, n(x_j))}\big) \geqslant b\mu_-(U_{a,b}).$$

This implies that $\nu(U_{a,b}) \geqslant b\mu_-(U_{a,b}) - \varepsilon$, as required. $\square$

LEMMA A.5. *There exist $N > 0$, $\delta > 0$, $m > 0$, $\underline{l}_p$, $\underline{l}_q$ and $\underline{l}_{q_i}$, $1 \leqslant i \leqslant m$, with the following properties*:

(1) $|\underline{l}_p| = |\underline{l}_q| = |\underline{l}_{q_i}| = N$;

(2) *the matrix $A^{\underline{l}_p}$ has simple Lyapunov spectrum*;

(3) $A^{\underline{l}_q} \cdot F \cap G = \{0\}$ *for every $F \in \mathrm{Grass}(k, H)$ and $G \in \mathrm{Grass}(\dim H - k, H)$ which are sums of eigenspaces of $A^{\underline{l}_p}$*;

(4) *for every $F \in \mathrm{Grass}(k, H)$ and $G \in \mathrm{Grass}(\dim H - k, H)$ such that F is a sum of eigenspaces of $A^{\underline{l}_p}$, there exists i such that the $A^{\underline{l}_{q_i}} \cdot F$ and G form angle at least δ.*

Proof. Since $\mathcal{B}$ is simple, there exists $\underline{l}_{\tilde{p}}$ such that the matrix $A^{l_{\tilde{p}}}$ has simple Lyapunov spectrum. Let $\mathcal{B}' \subset \mathcal{B}$ be the monoid consisting of the A^l's where $|\underline{l}|$ is a multiple of $|\underline{l}_{\tilde{p}}|$. Then $\mathcal{B}'$ is a large submonoid of $\mathcal{B}$, in the sense of Lemma 4.16, and so $\mathcal{B}'$ is simple. By the definition of twisting and Lemma 4.3, there exists $\underline{l}_{\tilde{q}}$ such that $|\underline{l}_{\tilde{q}}|$ is a multiple of $|\underline{l}_{\tilde{p}}|$ and $A^{l_{\tilde{q}}} \cdot F \cap G = \{0\}$ for every $F \in \mathrm{Grass}(k, H)$ and every $G \in \mathrm{Grass}(\dim H - k, H)$ which are sums of eigenspaces of $A^{l_{\tilde{p}}}$. For the same reasons, for every $G \in \mathrm{Grass}(\dim H - k, H)$ there exists $\underline{l}_{\tilde{q}(G)}$ such that $|\underline{l}_{\tilde{q}(G)}|$ is a multiple of $|\underline{l}_{\tilde{p}}|$ and $A^{l_{\tilde{q}(G)}} \cdot F \cap G = \{0\}$ for every $F \in \mathrm{Grass}(k, H)$ which is a sum of eigenspaces of $A^{l_{\tilde{p}}}$. By compactness of $\mathrm{Grass}(\dim H - k, H)$, one can choose finitely many $\underline{l}_{\tilde{q}_i}$ among all the $\underline{l}_{\tilde{q}(G)}$, so that for every $G \in \mathrm{Grass}(k, H)$,

$$\max_i \min_F \mathrm{angle}(A^{l_{\tilde{q}_i}} \cdot F, G) > 0,$$

where the minimum is over all $F \in \mathrm{Grass}(k, H)$ that are sums of eigenspaces of $A^{l_{\tilde{p}}}$. Using compactness once more, the expression on the left is even uniformly bounded below by some $\delta > 0$. Let N be the maximum of $|\underline{l}_{\tilde{p}}|$, $|\underline{l}_{\tilde{q}}|$ and $\underline{l}_{\tilde{q}_i}$. Let $\underline{l}_p$, $\underline{l}_q$ and $\underline{l}_{q_i}$ be obtained from $\underline{l}_{\tilde{p}}$, $\underline{l}_{\tilde{q}}$ and $\underline{l}_{\tilde{q}_i}$ by adding at the beginning as many copies of $\underline{l}_{\tilde{p}}$ as necessary to get to length N. One easily checks the required properties. $\square$

For $\underline{l} \in \Omega$, let $\underline{l}^n$ be obtained by repeating $\underline{l}$ exactly n times. Given a probability measure ϱ on $\mathrm{Grass}(k, H)$, let ϱ^l be the push-forward of ϱ under

$$A^l \colon \mathrm{Grass}(k, H) \longrightarrow \mathrm{Grass}(k, H).$$

LEMMA A.6. *Given $\varepsilon > 0$ and any probability measure ϱ in $\mathrm{Grass}(k, H)$, there is $n_0 = n_0(\varepsilon, \varrho)$ and, given any $\underline{l}_0 \in \Omega$, there is $i = i(\underline{l}_0)$ such that $\varrho^l(B) > 1 - \varepsilon$ for every $n > n_0$, where $\underline{l} = \underline{l}_p^n \underline{l}_q \underline{l}_p^n \underline{l}_{q_i} \underline{l}_0$ and B is the ε-neighborhood of $\xi_{\underline{l}}$.*

Proof. By (2) in Lemma A.5, when n is large, most of the mass of $\varrho^{l_p^n}$ is concentrated near sums of eigenspaces of A^{l_p} of dimension k. Then most of the mass of $\varrho^{l_p^n l_q}$ is concentrated near the A^{l_q}-image of those sums. Using (3) in Lemma A.5, it follows that, when n is large, most of the mass of $\varrho^{l_p^n l_q l_p^n}$ is concentrated near the subspace $F \in \mathrm{Grass}(k, H)$ given by the sum of the eigenspaces associated with the k eigenvalues of A^{l_p} with largest norms. Now let $G(\underline{l}_0) \in \mathrm{Grass}(\dim H - k, H)$ be spanned by eigenvectors of $(A^{l_0})^* A^{l_0}$ corresponding to $\dim H - k$ smallest singular values; if $\sigma_k(A^{l_0}) > \sigma_{k+1}(A^{l_0})$ then the only choice is $G(\underline{l}_0) = E_k^-(A^{l_0})$. Then, in general, the family $\{A^{l_0}\}_{\underline{l}_0}$ is uniformly continuous restricted to the subset of k-dimensional subspaces whose angle to $G(\underline{l}_0)$ is larger than any fixed $\delta/2$. Choose $\delta > 0$ as in Lemma A.5, and let $i = i(\underline{l}_0)$ be as in (4) of that lemma, for this choice of F and $G = G(\underline{l}_0)$. Then most of the mass of $\varrho^{l_p^n l_q l_p^n l_{q_i}}$ is

concentrated near $A^{l_{q_i}} \cdot F$, and so most of the mass of $\varrho^{\underline{l}}$ is concentrated near $A^{l_{q_i} \underline{l}_0} \cdot F$, as long as n is large; moreover, the equicontinuity property allows us to take the condition on n uniform on $\underline{l}_0$. In particular, we get that $\varrho^{\underline{l}}$ converges to the Dirac measure at $A^{l_{q_i} \underline{l}_0} \cdot F$ when n goes to infinity.

Now it suffices to show that $\xi_{\underline{l}}$ is close to $A^{l_{q_i} \underline{l}_0} \cdot F$ when n is large, uniformly on $\underline{l}_0$. This can be done applying the previous argument to the case when ϱ is the Lebesgue measure on $\mathrm{Grass}(k, H)$, and using (2) in Lemma 2.11 for $x_n = A^{l_p^n l_q l_p^n}$: we conclude that $\log \sigma_k(A^{\underline{l}}) - \log \sigma_{k+1}(A^{\underline{l}}) \to \infty$ and that $\xi_{\underline{l}} = A^{\underline{l}} \cdot E_k^+(A^{\underline{l}})$ converges to $A^{l_{q_i} \underline{l}_0} \cdot F$ when $n \to \infty$, as claimed. $\qquad \square$

LEMMA A.7. *There exist $C_0 = C_0(f, A, \mu) > 0$ and a sequence of sets $X_n \subset \Omega$ such that $|\underline{l}| \geqslant n$ for all $\underline{l} \in X_n$, the $\widehat{\Sigma}_-^{\underline{l}}$'s, $\underline{l} \in X_n$, are pairwise disjoint with $\sum_{\underline{l} \in X_n} \mu_-(\Sigma_-^{\underline{l}}) > C_0^{-1}$ and, given any probability measure ϱ on $\mathrm{Grass}(k, H)$ and any $n > n_0(\varepsilon, \varrho)$, we have*

$$\varrho^{\underline{l}}(B) > (1 - \varepsilon) \quad \text{for all } \underline{l} \in X_n,$$

where B is an ε-neighborhood of $\xi_{\underline{l}}$.

Proof. Let us order the elements of $\underline{l} \in \Omega$ by inclusion of the $\widehat{\Sigma}_-^{\underline{l}}$'s. Let X_n^0 be the collection of elements of Ω of the form $\underline{l}_p^n \underline{l}_q \underline{l}_p^n \underline{l}_{q_i} \underline{l}_0$, where $|\underline{l}_0|$ is a multiple of $(2n+2)N$. Let $X_n^1 \subset X_n^0$ be the maximal elements. Then the $\widehat{\Sigma}_-^{\underline{l}}$'s, $\underline{l} \in X_n^1$, are disjoint and (by ergodicity of $\hat{f}^{(2n+2)N}$)

$$\sum_{\underline{l} \in X_n^1} \mu_-(\Sigma_-^{\underline{l}}) = 1.$$

Notice that if $l_p^n \underline{l}_q l_p^n \underline{l}_{q_i} \underline{l}_0 \in X_n^1$ for some $1 \leqslant i \leqslant m$ then $l_p^n \underline{l}_q l_p^n \underline{l}_{q_i} \underline{l}_0 \in X_n^1$ for every $1 \leqslant i \leqslant m$. Let X_n be the set of $\underline{l} \in X_n^1$ such that $\underline{l} = l_p^n \underline{l}_q l_p^n \underline{l}_{q_i} \underline{l}_0$ with $i = i(\underline{l}_0)$ as in the previous lemma. Then

$$\sum_{\underline{l} \in X_n} \mu_-(\Sigma_-^{\underline{l}}) \geqslant C(\mu)^{-2} \min_{1 \leqslant i \leqslant m} \mu_-(\Sigma_-^{l_{q_i}}).$$

To conclude, apply the previous lemma. $\qquad \square$

Proof of Theorem A.1. Let $\widehat{m}$ be an invariant u-state, given by Corollary A.3, and let ν be its projection on $\mathrm{Grass}(k, H)$. For almost every $x \in \Sigma_-$, let $\widehat{m}^{(n)}(x)$ and $\widehat{m}(x)$ be as in Lemma A.4. Notice that we have, for any Borel set $Y \subset \mathrm{Grass}(k, H)$,

$$\frac{1}{C(\widehat{m})} \leqslant \frac{\widehat{m}^{(n)}(x)(Y)}{\nu^{\underline{l}(x,n)}(Y)} \leqslant C(\widehat{m}).$$

In particular, $\widehat{m}(x)$ is equivalent to any limit of $\nu^{\underline{l}(x,n)}$. By Lemma A.7, there is $Z \subset \Sigma_-$ with $\mu_-(Z) \geqslant C_0^{-1}$ such that $\nu^{\underline{l}(x,n)}$ accumulates at a Dirac mass for all $x \in Z$. Using

Lemma A.4, it follows that $\widehat{m}(x)$ is a Dirac mass for almost every $x \in Z$. By ergodicity and equivariance, $\widehat{m}(x)$ is a Dirac mass for almost every $x \in \Sigma_-$. We will denote the support of this Dirac mass by $\xi(x)$. Hence, $\xi(x)$ is the support of $\lim \nu^{\underline{l}(x,n)}$. Let $\hat{\xi} = \xi \circ \pi_-$. Then $\hat{\xi}(f(x,y)) = \hat{A}(x,y) \cdot \hat{\xi}(x,y)$. Notice that the push-forward of μ_- by ξ is equal to ν. In particular, the support of ν is forward invariant under the action of the supporting monoid $\mathcal{B}$. Observe that the support of ν cannot be contained in any hyperplane section S: otherwise, $\bigcap_{x \in \mathcal{B}} x^{-1} \cdot S$ would be a non-trivial invariant linear arrangement, and this cannot exist because $\mathcal{B}$ is simple. It follows from part (2) of Lemma 2.11 that $\log \sigma_k(A^{\underline{l}(x,n)}) - \log \sigma_{k+1}(A^{\underline{l}(x,n)}) \to \infty$ and $\xi(x) = \lim_{n \to \infty} \xi_{\underline{l}(x,n)}$. $\qquad\square$

Remark A.8. It follows that the push-forward of $\hat{\mu}$ by $(x,y,z) \mapsto (x,y,\xi(x))$ is an invariant u-state. It is not difficult to see that it must coincide with $\widehat{m}$. Thus, the invariant u-state is unique. See also [BV, Remark 5.4].

A.5. The inverse cocycle

Given $\underline{l} \in \Omega$, let $\xi_{\underline{l}}^- = E_k^-(A^{\underline{l}})$ if defined, otherwise choose $\xi_{\underline{l}}^-$ arbitrarily. For $x \in \Sigma$, let $\underline{l}^-(x,n) \in \Omega$ be such that $|\underline{l}^-(x,n)| = n$ and $x \in \Sigma^{\underline{l}^-(x,n)}$.

THEOREM A.9. *Let (f,A) be a symbolic locally constant measurable cocycle, with simple supporting monoid.*

(1) *There exists a measurable function $\xi^- \colon \Sigma \to \mathrm{Grass}(\dim H - k, H)$ such that $\hat{\xi}^- = \xi^- \circ \pi$ satisfies $\hat{\xi}^-(\hat{f}(x)) = \hat{A}(x) \cdot \hat{\xi}^-(x)$.*

(2) *For μ-almost every $x \in \Sigma$,*

$$\log \sigma_k(A^{\underline{l}^-(x,n)}) - \log \sigma_{k+1}(A^{\underline{l}^-(x,n)}) \to \infty \quad and \quad \xi_{\underline{l}^-(x,n)}^- \to \xi^-(x).$$

(3) *For every hyperplane section $S \subset \mathrm{Grass}(\dim H - k, H)$ there exists a positive μ-measure set of $x \in \Sigma$ such that $\xi^-(x) \notin S$.*

Proof. The cocycle $(\hat{f}, \hat{A})^{-1}$ is also measurable. Let $J \colon \widehat{\Sigma} \to \widehat{\Sigma}$ be given by

$$J(x_i)_{i \in \mathbf{Z}} = (x_{-i-1})_{i \in \mathbf{Z}}.$$

Then J conjugates $(\hat{f}, \hat{A})^{-1}$ to a locally constant symbolic cocycle $(\hat{f}, \widehat{B})$, where $\widehat{B}$ is defined by $\widehat{B}(x) = \hat{A}^{-1}(J(x)) = \hat{A}(\hat{f}^{-1}(J(x)))^{-1}$. The corresponding supporting monoid is also simple, by Lemma 4.12. The result then follows by application of Theorem A.1 to $(\hat{f}, \widehat{B})$, with $\dim H - k$ in the place of k. The invariant sections are related by

$$\hat{\xi}_{\hat{A}}^- = \hat{\xi}_{\widehat{B}}^- \circ J. \qquad\qquad\square$$

A.6. Proof of Theorem 7.1

Let $1 \leqslant k \leqslant \dim H - 1$. We must show that $\theta_k(f, A) > \theta_{k+1}(f, A)$. Let $\xi \colon \Sigma_- \to \operatorname{Grass}(k, H)$ be as in Theorem A.1, and let $\xi^- \colon \Sigma \to \operatorname{Grass}(\dim H - k, H)$ be as in Theorem A.9. We claim that $\xi(x) \cap \xi^-(y) = \{0\}$ for $\hat{\mu}$-almost every $(x, y) \in \Sigma_- \times \Sigma$. Indeed, if this was not the case then, by ergodicity, we would have $\xi(x) \cap \xi^-(y) \neq \{0\}$ for $\hat{\mu}$-almost every (x, y). By Fubini's theorem and bounded distortion, it would follow that $\xi(x) \cap \xi^-(y) \neq \{0\}$ for μ_--almost every $x \in \Sigma^-$ and μ-almost every $y \in \Sigma$. This would imply that $\xi(x)$ is contained in the hyperplane section dual to $\xi^-(y)$ for μ_--almost every $x \in \Sigma^-$, contradicting Theorem A.1.

Let $F^u \in \operatorname{Grass}(k, H)$ and $F^s \in \operatorname{Grass}(\dim H - k, H)$ be subspaces in the support of $\xi_* \mu_-$ and $\xi_*^- \mu$, respectively, such that $F^u \cap F^s = \{0\}$. Choose an inner product so that F^u and F^s are orthogonal. Let $\varepsilon_0 > 0$ be such that if F and F' belong to the balls of radius ε_0 around F^u and F^s, respectively, then $F \cap F' = \{0\}$. Fix $0 < \varepsilon_3 < \varepsilon_2 < \varepsilon_1 < \varepsilon_0$. Take $0 < \varepsilon_4 < \varepsilon_3$ small and let $\widehat{X} = X_- \times X$ be the set of (x, y) such that $\xi(x)$ belongs to the ball of radius ε_4 around F^u and $\xi^-(y)$ belongs to the ball of radius ε_4 around F^s. Let $n(x, y) \geqslant 0$ be minimal such that $A^{l(x,n)}$ takes the ball of radius ε_1 around F^u into the ball of radius ε_3 around F^u whenever $n > n(x, y)$ is such that $f^{-n}(x, y) \in \widehat{X}$. Then $n(x, y) < \infty$ for almost every $(x, y) \in X_- \times X$, by Theorems A.1, A.9 and Lemma 2.8.

Let $\widehat{Z} \subset \widehat{X}$ be a positive measure set such that the minimum of the first return time $r(x, y)$ of f to $\widehat{Z}$ is bigger than the maximum of $n(x, y)$ over $(x, y) \in \widehat{Z}$. Let R denote the return map to $\widehat{Z}$. For every $(x, y) \in \widehat{Z}$, choose $C(x, y) \in \operatorname{SL}(H)$ in a measurable fashion in the $c^{-1} \varepsilon_4$-neighborhood of id, such that $C(x, y) \cdot \xi(x) = F^u$ and $C(x, y) \cdot \xi^-(y) = F^s$. Let $B(x, y) = C(R(x, y)) A^l C(x, y)^{-1}$, where $\underline{l} = \underline{l}^-(y, r(x, y))$. Then the Lyapunov exponents of (R, B) relative to the normalized restriction of $\hat{\mu}$ to $\widehat{Z}$ are obtained by multiplying the Lyapunov exponents of (f, A) by $1/\hat{\mu}(\widehat{Z})$. Taking ε_4 sufficiently small, we guarantee that $B(x, y)$ takes the closure of the ball of radius ε_2 around F^u into its interior. It follows from Lemma 2.12 that $\sigma_k(B(x, y)) > \sigma_{k+1}(B(x, y))$ for almost every $(x, y) \in \widehat{Z}$, and $E_k^+(B(x, y)) = F^u$ and $E_k^-(B(x, y)) = F^s$. We conclude that

$$\theta_k(R, B) - \theta_{k+1}(R, B) \geqslant \frac{1}{\hat{\mu}(Z)} \int_Z [\log \sigma_k(B(x, y)) - \log \sigma_{k+1}(B(x, y))] \, d\hat{\mu}(x, y) > 0,$$

which implies the conclusion of the theorem.

A.7. Generalizations

Theorem 7.1 may be seen as a criterion for simplicity of the Lyapunov spectrum for products of matrices generated by a stationary stochastic process. Although we have

stated it in the case where there is a finite or countable number of states, our method can actually be used to prove the following more general result. Let (T, A), with

$$T \colon (\Delta, \mu) \longrightarrow (\Delta, \mu) \quad \text{and} \quad A \colon \Delta \longrightarrow \mathrm{SL}(H),$$

be a measurable cocycle. Given $k \geqslant 1$ and any set $Z \in \mathrm{SL}(H)^k$, we set

$$Z^\wedge = \{x \in \Delta : (A(x), ..., A(T^{k-1}(x))) \in Z\}.$$

We say that the cocycle has *approximate local product structure* if there exists $C \geqslant 1$ such that, for any $n, m \geqslant 1$ and any measurable sets $X \in \mathrm{SL}(H)^n$, $Y \in \mathrm{SL}(H)^m$,

$$C^{-1}\mu(X^\wedge)\mu(Y^\wedge) \leqslant \mu \left((X \times Y)^\wedge\right) \leqslant C\mu(X^\wedge)\mu(Y^\wedge).$$

THEOREM A.10. *Let (T, A), with $T \colon (\Delta, \mu) \to (\Delta, \mu)$ and $A \colon \Delta \to \mathrm{SL}(H)$, be a measurable cocycle with approximate local product structure, and assume that the submonoid of $\mathrm{SL}(H)$ generated by the support of the measure $A_*\mu$ is simple. Then the Lyapunov spectrum of (T, A) is simple.*

References

[AF]   AVILA, A. & FORNI, G., Weak mixing for interval exchange transformations and translation flows. *Ann. of Math.*, 165 (2007).

[AV1]   AVILA, A. & VIANA, M., Dynamics in the moduli space of abelian differentials. *Port. Math.*, 62 (2005), 531–547.

[AV]   — Simplicity of Lyapunov spectra: a sufficient criterion. To appear in *Port. Math.*

[BGV]   BONATTI, C., GÓMEZ-MONT, X. & VIANA, M., Généricité d'exposants de Lyapunov non-nuls pour des produits déterministes de matrices. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 20 (2003), 579–624.

[BV]   BONATTI, C. & VIANA, M., Lyapunov exponents with multiplicity 1 for deterministic products of matrices. *Ergodic Theory Dynam. Systems*, 24 (2004), 1295–1330.

[Fo]   FORNI, G., Deviation of ergodic averages for area-preserving flows on surfaces of higher genus. *Ann. of Math.*, 155 (2002), 1–103.

[Fu]   FURSTENBERG, H., Noncommuting random products. *Trans. Amer. Math. Soc.*, 108 (1963), 377–428.

[GM]   GOLDSHEID, I. Y. & MARGULIS, G. A., Lyapunov exponents of a product of random matrices. *Uspekhi Mat. Nauk*, 44:5 (1989), 13–60 (Russian); English translation in *Russian Math. Surveys*, 44:5 (1989), 11–71.

[GR]   GUIVARC'H, Y. & RAUGI, A., Products of random matrices: convergence theorems, in *Random Matrices and their Applications* (Brunswick, ME, 1984), Contemp. Math., 50, pp. 31–54. Amer. Math. Soc., Providence, RI, 1986.

[Ko]   KONTSEVICH, M., Lyapunov exponents and Hodge theory, in *The Mathematical Beauty of Physics* (Saclay, 1996), Adv. Ser. Math. Phys., 24, pp. 318–332. World Scientific, River Edge, NJ, 1997.

[KZ]   KONTSEVICH, M. & ZORICH, A., Connected components of the moduli spaces of
          Abelian differentials with prescribed singularities. *Invent. Math.*, 153 (2003), 631–
          678.

[Le]   LEDRAPPIER, F., Positivity of the exponent for stationary sequences of matrices, in
          Lyapunov Exponents (Bremen, 1984), Lecture Notes in Math., 1186, pp. 56–73.
          Springer, Berlin, 1986.

[MMY]   MARMI, S., MOUSSA, P. & YOCCOZ, J.-C., The cohomological equation for Roth-type
          interval exchange maps. *J. Amer. Math. Soc.*, 18 (2005), 823–872.

[Ma]   MASUR, H., Interval exchange transformations and measured foliations. *Ann. of Math.*,
          115 (1982), 169–200.

[Os]   OSELEDETS, V. I., A multiplicative ergodic theorem: Lyapunov characteristic numbers
          for dynamical systems. *Trudy Moskov. Mat. Obshch.*, 19 (1968), 179–210 (Russian);
          English translation in *Trans. Moscow Math. Soc.*, 19 (1968), 197–231.

[Ra]   RAUZY, G., Échanges d'intervalles et transformations induites. *Acta Arith.*, 34 (1979),
          315–328.

[Sc]   SCHWARTZMAN, S., Asymptotic cycles. *Ann. of Math.*, 66 (1957), 270–284.

[Ve1]   VEECH, W. A., Gauss measures for transformations on the space of interval exchange
          maps. *Ann. of Math.*, 115 (1982), 201–242.

[Ve2]   — The Teichmüller geodesic flow. *Ann. of Math.*, 124 (1986), 441–530.

[Vi]   VIANA, M., Almost all cocycles over any hyperbolic system have non-vanishing Lya-
          punov exponents. To appear in *Ann. of Math.*

[Zo1]   ZORICH, A., Asymptotic flag of an orientable measured foliation on a surface, in *Geo-
          metric Study of Foliations* (Tokyo, 1993), pp. 479–498. World Scientific, River Edge,
          NJ, 1994.

[Zo2]   — Finite Gauss measure on the space of interval exchange transformations. Lyapunov
          exponents. *Ann. Inst. Fourier* (Grenoble), 46 (1996), 325–370.

[Zo3]   — Deviation for interval exchange transformations. *Ergodic Theory Dynam. Systems*,
          17 (1997), 1477–1499.

[Zo4]   — How do the leaves of a closed 1-form wind around a surface?, in *Pseudoperiodic
          Topology*, Amer. Math. Soc. Transl., 197, pp. 135–178. Amer. Math. Soc., Provi-
          dence, RI, 1999.

ARTUR AVILA   MARCELO VIANA
CNRS UMR 7599   IMPA – Instituto Nacional
Laboratoire de Probabilités   de Matemática Pura e Aplicada
et Modèles Aléatoires   Estrada Dona Castorina 110
Université Pierre et Marie Curie   22460-320 Rio de Janeiro
Boîte Postale 188   Brazil
FR-75252 Paris Cedex 05   viana@impa.br
France
artur@ccr.jussieu.fr