ISSN 0012-9593

quatrième série - tome 54 fascicule 1 janvier-février 2021

ANNALES SCIENTIFIQUES de L'ÉCOLE NORMALE SUPÉRIEURE

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

Annales Scientifiques de l'École Normale Supérieure

Publiées avec le concours du Centre National de la Recherche Scientifique

Rédaction / *Editor*

Annales Scientifiques de l'École Normale Supérieure,
45, rue d'Ulm, 75230 Paris Cedex 05, France.
Tél. : (33) 1 44 32 20 88. Fax : (33) 1 44 32 20 80.
annales@ens.fr

Édition et abonnements / *Publication and subscriptions*

Société Mathématique de France
Case 916 - Luminy
13288 Marseille Cedex 09
Tél. : (33) 04 91 26 74 64
Fax : (33) 04 91 41 17 51
email : abonnements@smf.emath.fr

Tarifs

Abonnement électronique : 420 euros.
Abonnement avec supplément papier :
Europe : 551 €. Hors Europe : 620 € ($ 930). Vente au numéro : 77 €.

ISSN 0012-9593 (print) 1873-2151 (electronic)

Directeur de la publication : Fabien Durand
Périodicité : 6 nos / an

Ann. Scient. Éc. Norm. Sup.
4^{e} série, t. 54, 2021, p. 1 à 68

REGULARITY OF ENTROPY, GEODESIC CURRENTS AND ENTROPY AT INFINITY

BY Barbara SCHAPIRA AND Samuel TAPIE

ABSTRACT. – In this work, we introduce a notion of *entropy at infinity* for the geodesic flow of negatively curved manifolds. We introduce the class of noncompact manifolds which admit a critical gap between entropy at infinity and topological entropy. We call them *strongly positively recurrent manifolds* (SPR), and provide many examples. We show that dynamically, they behave as compact manifolds. In particular, they admit a finite measure of maximal entropy.

Using the point of view of currents at infinity, we show that on these SPR manifolds the topological entropy of the geodesic flow varies in a $\mathcal{C}^1$-way along $\mathcal{C}^1$-uniform perturbations of the metric. This result generalizes former work of Katok (1982) and Katok-Knieper-Weiss (1991) in the compact case.

RÉSUMÉ. – Dans ce travail, nous introduisons une notion d'*entropie à l'infini* pour les flots géodésiques des variétés à courbure négative. Nous introduisons la classe des variétés, dites *fortement positivement récurrentes* (SPR), dont l'entropie à l'infini est strictement inférieure à l'entropie topologique. Nous donnons de nombreux exemples de telles variétés. Nous montrons que d'un point de vue dynamique, ces variétés ressemblent à des variétés compactes. En particulier, elles admettent une mesure finie maximisant l'entropie.

À l'aide du point de vue des courants à l'infini, nous montrons que sur ces variétés SPR, l'entropie topologique varie de manière $\mathcal{C}^1$ le long de perturbations $\mathcal{C}^1$-uniformes de la métrique. Ceci généralise des résultats passés de Katok (1982) et Katok-Knieper-Weiss (1991) dans le cas compact.

1. Introduction

1.1. Variation of the topological entropy: An overview

The initial motivation of this work was to answer the following simple question. Consider a hyperbolic surface of finite volume and a smooth compact perturbation of the metric. Does the topological entropy of the geodesic flow vary regularly? More generally, what happens for a smooth perturbation of the metric of a noncompact negatively curved Riemannian manifold?

0012-9593/01/© 2021 Société Mathématique de France. Tous droits réservés doi:10.24033/asens.2455

The answer has been known on compact manifolds since almost thirty years [31, 30, 21], and has been extended to the convex-cocompact case in [52]. A similar argument gives the regularity of the topological entropy for a perturbation of an Anosov flow, cf [31].

Compactness of the underlying space is crucial in the above results, and no result was known until now for manifolds with a non-compact non-wandering set. Even the case of a smooth compact perturbation of the metric of a finite volume hyperbolic surface was not accessible with their arguments. Let us recall the two main steps of their argument to understand why.

The key step is the following inequality, due to Katok in [29] for surfaces, extended in [31] to all dimensions.

THEOREM 1.1 ([29]; [31]). – *Let g_1, g_2 be Riemannian metrics with negative sectional curvature on the same compact manifold M. Then the entropies of their geodesic flows satisfy*

$$h_{\mathrm{top}}(g_1) \leq h_{\mathrm{top}}(g_2) \times \int_{S^{g_1}M} \|v\|^{g_2} \, d\overline{m}^{g_1}_{\mathrm{BM}}(v), \tag{1}$$

where $\|v\|^{g_2} = \sqrt{g_2(v,v)}$ and $\overline{m}^{g_1}_{\mathrm{BM}}$ is the normalized Bowen-Margulis measure on the g_1-unit tangent bundle $S^{g_1}M$ for the g_1-geodesic flow.

Reversing the role of g_1 and g_2 also provides a lower bound for $h_{\mathrm{top}}(g_1)$, and a first order power expansion gives the following smoothness result.

THEOREM 1.2 ([31]). – *Let $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ be a $\mathcal{C}^2$-family of $\mathcal{C}^2$ Riemannian metrics with negative sectional curvature on the same compact manifold M. Then $\varepsilon \mapsto h_{\mathrm{top}}(g_\varepsilon)$ is $\mathcal{C}^1$, and its derivative is given by*

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} h_{\mathrm{top}}(g_\varepsilon) = -h_{\mathrm{top}}(g_0) \times \int_{S^{g_0}M} \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \|v\|^{g_\varepsilon} \, d\overline{m}^{g_0}_{\mathrm{BM}}(v), \tag{2}$$

where $\overline{m}^{g_0}_{\mathrm{BM}}$ is the normalized Bowen-Margulis measure on the g_0-unit tangent bundle $S^{g_0}M$ for the g_0-geodesic flow.

In the previously quoted works, the proofs of (1) strongly use the compactness of the non-wandering set. In the first part of our paper, we use a different approach to generalize it to the non-compact setting. This improves it even in the compact case, providing an explicit transformation rule for the entropies, equality which immediately implies (1), and has other interesting consequences.

The previously known proofs of (2) use the compactness of M for a crucial point: to ensure the finiteness and the continuity of the normalized Bowen-Margulis measures $\overline{m}^{g_\varepsilon}_{\mathrm{BM}}$ in the weak-$*$ topology as ε varies. Neither finiteness of the Bowen-Margulis measure nor its continuity under a variation of the metric can be ensured in general. Maybe the most striking fact of our work is that we introduce a new wide class of manifolds, which we call *SPR manifolds*, SPR meaning *strongly / stably positively recurrent*. The terminology *Stably positively recurrent* has been introduced by Gurevic-Savchenko [26] in the context of countable Markov shifts. Sarig [48] modified it, in the same context, into *strongly positively recurrent*, terminology which has been used later by other authors as Buzzi [8]. See also the very recent work of Velozo [53], who follows also this terminology. Both terminologies are meaningful, and had not yet been considered in our context. It turns out that the same property also appeared recently and independently in the context of geometric group theory in [3] under the name of growth gap.

The class of SPR manifolds that we define here has the remarkable property that the Bowen-Margulis measure is finite, and moreover stays finite and varies continuously along small perturbations. In particular, under $\mathcal{C}^1$-uniform variation of such SPR Riemannian metrics, the topological entropy is $\mathcal{C}^1$ and its derivative is given by (2).

These SPR manifolds include finite volume hyperbolic manifolds, and more generally almost all known examples where the geodesic flow admits a (finite) measure of maximal entropy, as geometrically finite negatively curved manifolds with spectral gap [15], Schottky product examples from [39], and unpublished examples of Ancona [2]. The class of SPR manifolds is much larger than only the above mentioned examples. We postpone the extensive study of SPR manifolds to a later paper [25]. Therefore, the second half of our paper will be devoted to the presentation of a geometrical setting, as large as possible, where this finiteness and continuity of Bowen-Margulis measures can be ensured.

Let us now present our main results with more details.

1.2. Invariant measures and change of Riemannian metrics

Let (M, g_1) be a complete Riemannian manifold, and g_2 be another Riemannian metric on M such that there exists $C > 1$ with $\frac{1}{C} g_1 \leq g_2 \leq C g_1$. We assume moreover that both g_1 and g_2 have pinched negative sectional curvatures with uniformly bounded first derivatives: this implies that g_1-geodesics are g_2-quasi-geodesics and the visual boundary of the universal cover $(\widetilde{M}, g_1)$ is canonically identified with the visual boundary of $(\widetilde{M}, g_2)$; we will denote it by $\partial\widetilde{M}$. We will use extensively this correspondance to compare the dynamics of the geodesic flows on $S^{g_1}M$ and $S^{g_2}M$.

Let $\Gamma = \pi_1(M)$ acting on the universal cover $\widetilde{M}$, let m be a locally finite measure on $S^{g_1}M$, invariant by the geodesic flow $(g_1^t)_{t\in\mathbb{R}}$, and $\widetilde{m}$ its lift to $S^{g_1}\widetilde{M}$. We write $\partial^2\widetilde{M} = (\partial\widetilde{M} \times \partial\widetilde{M})\backslash\mathrm{Diag}$. In g_1-Hopf coordinates (cf Section 2), $S^{g_1}\widetilde{M} \simeq \partial^2\widetilde{M} \times \mathbb{R}$, and $\widetilde{m}$ has a local product structure of the form $d\widetilde{m} = d\mu \times dt$, where μ is a Γ-invariant *geodesic current* on $\partial^2\widetilde{M}$. We write therefore $m = m_\mu^{g_1}$.

We can now define a measure $\widetilde{m}_\mu^{g_2}$ on $S^{g_2}\widetilde{M}$, given in g_2-Hopf parametrization by the same local product formula $\widetilde{m}_\mu^{g_2} = d\mu \times dt$: by Γ-invariance, this induces a locally finite measure $m_\mu^{g_2}$ on $S^{g_2}M$, which is invariant for the geodesic flow $(g_2^t)_{t\in\mathbb{R}}$. The ergodic properties of $(S^{g_1}M, g_1^t, m_\mu^{g_1})$ and $(S^{g_2}M, g_2^t, m_\mu^{g_2})$ are strongly related.

Well known facts imply that if $m_\mu^{g_1}$ and $m_\mu^{g_2}$ are finite then one is ergodic or conservative if and only if the other is. The reader may believe that, since $\frac{1}{C} g_1 \leq g_2 \leq C g_1$, then $m_\mu^{g_1}$ is finite if and only if $m_\mu^{g_2}$ is. We will indeed show that it is the case and relate the masses and entropies of these measures.

In this purpose, let us introduce the *instantaneous geodesic stretch* $\mathcal{E}^{g_1\to g_2} : S^{g_1}\widetilde{M} \to \mathbb{R}$ defined for all $v \in S^{g_1}\widetilde{M}$ by

$$\mathcal{E}^{g_1\to g_2}(v) = \frac{d}{dt}\Big|_{t=0^+} \mathcal{B}^{g_2}_{v_+^{g_1}}(\pi v, \pi g_1^t v) = \frac{d}{dt}\Big|_{t=0^+} \mathcal{B}^{g_2}_{v_+^{g_1}}(o, \pi g_1^t v),$$

where $\pi : S^{g_1}\widetilde{M} \to \widetilde{M}$ is the canonical projection, and $\mathcal{B}^{g_2}_{v_+^{g_1}}(.,.)$ is the Busemann function for g_2 based at the endpoint of the g_1-geodesic generated by v. By Γ-invariance, it induces a map $\mathcal{E}^{g_1\to g_2} : S^{g_1}M \to \mathbb{R}$. We will see in Section 2.3 that this is the derivative along g_1-geodesics of a natural *Morse correspondance* $\Psi^{g_1\to g_2} : S^{g_1}M \to S^{g_2}M$ between the g_1 and g_2 geodesic flows. This Morse correspondence is a global homeomorphism, which sends g_1-geodesics to g_2-geodesics up to a time rescaling. In particular, it induces a homeomorphism from the non-wandering set Ω^{g_1} of the geodesic flow on $S^{g_1}M$ to the non-wandering set Ω^{g_2} of the geodesic flow on $S^{g_2}M$; see next section for precise definitions. This implies the following.

Proposition 1.3. – *For every $m_\mu^{g_2}$-measurable map $G : S^{g_2}M \to \mathbb{R}$, the map $G\circ\Psi^{g_1\to g_2}$ is $m_\mu^{g_1}$-measurable and*

$$\int_{S^{g_2}M} G\, dm_\mu^{g_2} = \int_{S^{g_1}M} G\circ\Psi^{g_1\to g_2}\times\mathcal{E}^{g_1\to g_2}\, dm_\mu^{g_1}.$$

In particular, the masses of $m_\mu^{g_2}$ satisfy

$$\|m_\mu^{g_2}\| = \int_{S^{g_1}M}\mathcal{E}^{g_1\to g_2}\, dm_\mu^{g_1}.$$

Some other versions of the geodesic stretch have already been considered in [20] or [32]; we explain in Section 2.3 the relationship with these references and the interest of our new definition. We then introduce in Section 3 a notion of *local entropy* for invariant measures, which is an analogous in the non-compact setting to Brin-Katok entropy, and which coincides with the classical measure-theoretic entropy for Gibbs measures [(1)]. This also allows us to relate the local entropies of $(S^{g_1}M, g_1^t, m_\mu^{g_1})$ and $(S^{g_2}M, g_2^t, m_\mu^{g_2})$.

Theorem 1.4 (See Theorem 3.11). – *Let (M, g_i), $i = 1,2$ be two equivalent Riemannian metrics on M with pinched negative curvature and uniformly bounded derivatives. Let μ be any geodesic current and $m_\mu^{g_i}$ the associated invariant measure on $S^{g_i}M$ under the geodesic flow (g_i^t). Assume that these measures are finite and ergodic. Then the local entropies of $(g_1^t, m_\mu^{g_1})$ and $(g_2^t, m_\mu^{g_2})$ are related as follows.*

$$h_{\mathrm{loc}}(m_\mu^{g_2}, g_2) = I_\mu(g_2, g_1)\times h_{\mathrm{loc}}(m_\mu^{g_1}, g_1),$$

where

$$I_\mu(g_2,g_1) = \frac{1}{\|m_\mu^{g_2}\|}\int_{S^{g_2}M}\mathcal{E}^{g_2\to g_1}(v)\, dm_\mu^{g_2}(v).$$

The combination of the previous theorem with the variational principle for entropy implies the following result, which is an optimal improvement of (1).

Theorem 1.5 (See Corollary 3.18 and Theorem 3.19). – *Let (M, g_i), $i = 1,2$ be two equivalent Riemannian metrics on M whose curvature is negatively pinched and has uniformly bounded derivatives. Assume that the geodesic flow on $S^{g_2}M$ has a finite measure Bowen-Margulis measure $m_{\mathrm{BM}}^{g_2}$, i.e., a finite measure with maximal entropy. Then*

$$h_{\mathrm{top}}(g_2)\le I_{m_{\mathrm{BM}}^{g_2}}(g_2,g_1)\times h_{\mathrm{top}}(g_1).$$

Moreover, equality holds if and only if the geodesic flow on $S^{g_1}M$ also has a finite Bowen-Margulis measure and there exists a Morse correspondence $F^{g_1\to g_2} : S^{g_1}M\to S^{g_2}M$ which

(1) Riquelme showed recenty [44] that these entropies coincide for all ergodic measures.

conjugates the flows on the non-wandering sets of $S^{g_1}M$ and $S^{g_2}M$ up to a global time scaling by $\frac{h_{\mathrm{top}}(g_1)}{h_{\mathrm{top}}(g_2)}$: for all $v \in \Omega^{g_1}$ and all $t \in \mathbb{R}$,

$$g_2^{h_{\mathrm{top}}(g_2)t} \circ F^{g_1 \to g_2}(v) = F^{g_1 \to g_2} \circ g_1^{h_{\mathrm{top}}(g_1)t}(v).$$

This has been shown by Knieper in [32] for compact manifolds. The relation between the Morse correspondences $F^{g_1 \to g_2}$ and $\Psi^{g_1 \to g_2}$ will be precised in Theorem 3.19. Note that in general, when two negatively curved metrics are equivalent, one may have a finite Bowen-Margulis measure whereas the other may not. The previous theorem has the following striking corollary.

Corollary 1.6. – *Let (M, g_i), $i = 1, 2$ be two equivalent Riemannian metrics on M whose curvature is negatively pinched and has uniformly bounded derivatives. Assume that the geodesic flow on $S^{g_2}M$ has a finite measure Bowen-Margulis measure $m_{\mathrm{BM}}^{g_2}$ and that*

$$h_{\mathrm{top}}(g_2) = I_{m_{\mathrm{BM}}^{g_2}}(g_2, g_1) \times h_{\mathrm{top}}(g_1).$$

Then the marked length spectra of g_1 and g_2 coincide up to a global scaling by $\frac{h_{\mathrm{top}}(g_1)}{h_{\mathrm{top}}(g_2)}$.

Section 4 is devoted to the study of Gibbs measures and their behavior under change of metrics. It happens to be crucial in the proof of the above Theorem 1.5. We show that a (g_1^t)-invariant measure $m_\mu^{g_1}$ is a Gibbs measure for the potential $G : S^{g_1} \to \mathbb{R}$ if and only if the associated (g_2^t)-invariant measure $m_\mu^{g_2}$ is a Gibbs measure for the potential $G \circ \Psi^{g_2 \to g_1} \times \mathcal{E}^{g_2 \to g_1}$. We also give some applications of this last fact to a comparison between the length spectra of (M, g_1) and (M, g_2), see Corollary 4.4.

1.3. Entropy at infinity, SPR manifolds and Bowen-Margulis measures

Let (M, g) be a Riemannian manifold with pinched negative sectional curvatures whose derivatives are uniformly bounded. We introduce a notion of *entropy at infinity* (see Section 7), which measures the highest possible complexity of the (topological) dynamics outside a compact set in the manifold. Note that another definition of entropy at infinity appears in [8, 45, 53], which is somehow the maximal entropy of a sequence of invariant probability measures diverging to infinity. See also [18, 19] for related works in finite volume rank one homogeneous dynamics. It follows from [45] in the geometrically finite case and [53, 25] more generally that this entropy at infinity coincides with our definition.

We call the Riemannian manifold (M, g) *strongly positively recurrent*, shortly SPR, if the entropy at infinity is strictly smaller than the topological entropy of the geodesic flow. This SPR property implies that the geodesic flow admits a finite Bowen-Margulis measure, which is then the unique measure of maximal entropy, according to [36], see Theorem 3.16, and also that this fact remains true under a nice small perturbation of the metric, and that these measures vary continuously in the narrow topology (i.e., in the dual of bounded continuous functions).

Let us comment on this terminology of strong positive recurrence. It comes from the world of symbolic dynamics with the works of Gurevich-Savchenko [26] and Sarig [48]. In Pit-Schapira [41], it is shown that their notion of *recurrence* is equivalent to the conservativity and ergodicity of the Bowen-Margulis measure, whereas *positive recurrence* corresponds to finiteness (and therefore conservativity and ergodicity) of this Bowen-Margulis measure.

Let us summarize the main results that we establish here on the SPR property. A more detailed study of this property and its consequences is the aim of paper [25].

THEOREM 1.7. – *Let (M, g) be a Riemannian manifold with pinched negative curvature.*

1. *The SPR property implies that the geodesic flow admits an invariant probability measure of maximal entropy m^g_{BM}, the so-called* Bowen-Margulis measure. *In the terminology of* [41]*, the SPR property implies that the geodesic flow is positively recurrent.*
2. *Geometrically finite manifolds with critical gap (see* [15]*) have the SPR property,*
3. *Topologically infinite examples of* [2] *presented in Section* 7.3.3 *have the SPR property,*
4. *Schottky product examples of* [39] *have the SPR property.*

Typical examples of manifolds which do not have the SPR property are infinite covers of compact negatively curved manifolds, or geometrically finite manifolds without critical gap (see [15] once again).

As mentioned above, this SPR property is stable in the following sense.

THEOREM 1.8. – *Let (M, g_0) be a SPR manifold with pinched negative curvature and bounded derivatives of the curvature. Let $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ be a $\mathcal{C}^1$-uniform variation of the metric. Then there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, the manifold (M, g_0) is SPR. Moreover, the Bowen-Margulis measures $(m^{g_\varepsilon}_{\mathrm{BM}})$ vary continuously at $\varepsilon = 0$ in the narrow topology.*

Let us recall here that the narrow topology is the dual topology of bounded continuous functions, whereas the vague topology is the dual topology of continuous compactly supported functions. In the above theorem, continuity in the vague topology is not a big problem, whereas noncompactness of the manifold creates huge difficulties to get convergence of the total mass of the measures, and therefore continuity in the narrow topology. It is the key place of the paper where we really absolutely need the SPR property to get convergence of the masses of measures, whereas at several other places the assumption is either not needed, or could be slightly weakened.

This allows us to show the following regularity property for the topological entropy, which answers our initial question. We refer to Section 7 for technical details on the assumptions.

THEOREM 1.9. – *Let (M, g_0) be a SPR manifold with pinched negative curvature and bounded derivatives of the curvature. Let $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ be a $\mathcal{C}^1$-uniform variation of the metric with negative sectional curvatures. Then the map $\varepsilon \mapsto h_{\mathrm{top}}(g_\varepsilon)$ is $\mathcal{C}^1$ near $\varepsilon = 0$, with derivative at* 0

$$\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} h_{\mathrm{top}}(g_\varepsilon) = -h_{\mathrm{top}}(g_0) \times \int_{S^{g_0}M} \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \|v\|^{g_\varepsilon}\, d\overline{m}^{g_0}_{\mathrm{BM}}(v),$$

the normalized Bowen-Margulis measure $\overline{m}^{g_0}_{\mathrm{BM}} = \frac{m^{g_0}_{\mathrm{BM}}}{\|m^{g_0}_{\mathrm{BM}}\|}$ *being the invariant probability measure of maximal entropy for the g_0-geodesic flow.*

Let us emphasize the fact that this theorem is valid in a much greater generality than what we thought initially possible. On the one hand, SPR manifolds are a very general and interesting class of manifolds, much larger than the well known and well studied class of finite volume, or even geometrically finite hyperbolic manifolds, as illustrated by Theorem 1.7. It may be an optimal class to get such result in the sense that we guess that phase transitions for the entropy can happen when the manifold is not SPR, analogous to those obtained by Riquelme-Velozo [45] for the pressure when varying a potential on geometrically finite manifolds.

On the other hand, we allow much more general perturbations than only compact ones since we deal with noncompact $\mathcal{C}^2$-perturbations of our metric, as soon as they are not too wild at infinity.

The paper is organized as follows. In Section 2, we develop the point of view of geodesic currents at infinity, which allows us to associate to an invariant measure $m_\mu^{g_1}$ for the geodesic flow for (M, g_1) an invariant measure $m_\mu^{g_2}$ for the geodesic flow on (M, g_2), and compare their ergodic properties.

In Section 3, we introduce different notions of entropy and develop methods of Section 2 to relate the entropies of $m_\mu^{g_1}$ and $m_\mu^{g_2}$.

In Section 4, we recall general facts about Gibbs measures on non-compact manifolds, we show that $m_\mu^{g_1}$ is a Gibbs measure if and only if $m_\mu^{g_2}$ is and give applications to the length spectrum.

In Section 5 we show some continuity results for geodesics, Busemann functions and non-normalized Bowen-Margulis measures which will be needed in the sequel.

In Section 6, we first show that for a fixed geodesic current μ on $\partial^2\widetilde{M}$, the measure-theoretic entropy $\varepsilon \mapsto h\left(g_\varepsilon^t, m_\mu^{g_\varepsilon}\right)$ is $\mathcal{C}^1$ under a $\mathcal{C}^1$-uniform variation of the Riemannian metrics g_ε. We then show in a very similar proof that, if under a $\mathcal{C}^1$-uniform variation of Riemmanian metrics the normalized Bowen-Margulis measures $\overline{m}_{\mathrm{BM}}^{g_\varepsilon}$ vary continuously in the narrow topology, then the topological entropy is also $\mathcal{C}^1$.

Eventually, in Section 7, we introduce entropy at infinity and SPR manifolds, we show that they have finite Bowen-Margulis measure, and that under a small $\mathcal{C}^1$-uniform variation of Riemannian metrics they remain SPR. On the way, we give some properties of the entropy at infinity of independent interest.

Theorem 1.7 follows from results of Section 7.3, where we provide many examples of SPR manifolds. Theorem 1.8 is a reformulation of the second part of Theorem 7.1. At last, our main variational formula for the topological entropy, Theorem 1.9, follows from Theorems 6.3 and 1.8 (or 7.1).

Acknowledgments

The authors thank warmly S. Gouëzel and J. Buzzi for several discussions during this work, S. Alvarez for his very useful comments on a preliminary version of this text. We also thank warmly the anonymous referee for suggesting to us Theorem 1.5 and Corollary 1.6.

The authors acknowledge the support of the ANR grant ANR JCJC 0108-GEODE (2010-2015), the grant PEPS CNRS Egalité CHATS (Integer Project 2015), and the Centre Henri Lebesgue ANR-11-LABX-0020-01.

2. Hopf parametrization and geodesic currents

2.1. Hopf parametrization and geodesic flow

Let (M, g_0) be a complete manifold with pinched negative sectional curvatures satisfying $-b^2 \leq K_{g_0} \leq -a^2 < 0$, and first derivatives of the curvature bounded. The bounds on the curvature are crucial, particularly the upper bound, at many places of the paper. But the assumption on the derivative will be used (implicitely) only when speaking about the Bowen-Margulis measure and its entropy. Indeed, this allows to get regularity of strong (un)stable foliations, which is used in the paper [36] that we shall use later, see Theorem 3.16.

Let $\widetilde{M}$ be the universal cover of M, equipped with the lifted metric which we will still denote by g_0, and let $\partial_{g_0}\widetilde{M}$ be its visual boundary. Let $\Gamma = \pi_1(M)$ be the fundamental group, acting properly by diffeomorphisms on $\widetilde{M}$. Denote by p_Γ indistinctly the projection $\widetilde{M} \to M$ and its linear tangent map $T\widetilde{M} \to TM$. A metric g on M (or equivalently, a Γ-equivariant metric g on $\widetilde{M}$) will be called *admissible* if it has pinched negative sectional curvature, if the derivatives of the curvature are bounded and if there exists a constant $C_1(g_0, g) > 1$ such that at all $x \in M$,

$$\frac{1}{C_1(g_0,g)}\, g_0 \leq g \leq C_1(g_0,g)\, g_0. \tag{3}$$

This implies that g-geodesics are g_0-quasi-geodesic. By Morse-Klingenberg lemma (see for example [9, Th. 1.7 p. 401] for a proof), they are contained in the $C_2(g_0, g)$-neighborhood of g_0-geodesics, where $C_2(g_0, g)$ only depends on $C_1(g_0, g)$.

In particular the visual boundary $\partial_g\widetilde{M}$ of $(\widetilde{M}, g)$ is canonically identified to the visual boundary of $(\widetilde{M}, g_0)$, and they will therefore both be denoted by $\partial\widetilde{M}$. Moreover, this identification is Hölder continuous w.r.t the visual distances induced by both g_0 and g, so that $\partial\widetilde{M}$ has a natural Hölder structure.

The *limit set* $\Lambda_\Gamma \subset \partial\widetilde{M}$ is the set of accumulation points of any orbit $\Gamma.x$ on the boundary. The *radial limit set* $\Lambda_\Gamma^r \subset \Lambda_\Gamma$ is the set of endpoints of geodesics which, on the quotient manifold M, return infinitely often to some compact set. None of these limit sets depend on the chosen admissible metric.

Let us fix once for all a point $o \in \widetilde{M}$. Let g be any admissible metric on M, and d^g the distance induced by g on M and $\widetilde{M}$. Denote by $S^g M$ (resp. $S^g\widetilde{M}$) the unit tangent bundle of (M, g) (resp. $(\widetilde{M}, g)$), and $\partial^2\widetilde{M} = (\partial\widetilde{M} \times \partial\widetilde{M})\backslash\mathrm{Diag}$. We write $\pi : TM \to M$ and $\pi : T\widetilde{M} \to \widetilde{M}$ the projections from the tangent bundle to its base, and by $(g^t)_{t\in\mathbb{R}}$ the geodesic flow on $S^g M$ or $S^g\widetilde{M}$. For any $v \in S^g\widetilde{M}$, write v_-^g and v_+^g for the negative and positive endpoints in $\partial\widetilde{M}$ of the geodesic $\{\pi g^t v; t \in \mathbb{R}\}$.

REMARK 2.1. – We keep track in our notations of the metric g since we will soon compare these quantities for two different admissible metrics g_1 and g_2.

For all $\xi \in \partial\widetilde{M}$, let $\mathcal{B}_\xi^g$ be the Busemann function at ξ defined, for any $x, y \in \widetilde{M}$, by

$$\mathcal{B}_\xi^g(x, y) = \lim_{z\to\xi} d^g(x, z) - d^g(y, z).$$

The map

$$H^g : v \mapsto \left(v_-^g, v_+^g, \mathcal{B}_{v_+^g}(o, \pi v)\right)$$

is a Hölder homeomorphism from $S^g\widetilde{M}$ to $\partial^2\widetilde{M}\times\mathbb{R}$, called the *Hopf parametrization* of the unit tangent bundle.

The action of Γ by (differentials of) isometries on $S^g\widetilde{M}$ can be written in these coordinates as

$$\gamma.(v_-^g, v_+^g, t) = \left(\gamma.v_-^g, \gamma.v_+^g, t + \mathcal{B}^g_{v_+^g}(\gamma^{-1}.o, o)\right).$$

Let us emphasize the fact that this action of Γ on $\partial^2\widetilde{M}\times\mathbb{R}$, and more specifically on the third factor, depends strongly on the cocycle $\mathcal{B}^g$, and therefore on the metric g.

2.2. Geodesic currents and invariant measures

In the coordinates given by the Hopf parametrization of $S^g\widetilde{M}$, the geodesic flow (g^t) acts by translation on the last factor: for all $v\in S^g\widetilde{M}$, and $s\in\mathbb{R}$,

$$\text{if}\quad H^g(v) = (v_-, v_+, t)\quad\text{then}\quad H^g(g^s v) = (v_-, v_+, t+s).$$

Therefore, any positive Radon measure m on $S^g M$ invariant by the flow lifts to a measure $\widetilde{m}$ on $S^g\widetilde{M}$ of the form $\widetilde{m} = (H^g)^*(\mu\times dt)$, where dt is the Lebesgue measure on $\mathbb{R}$, and μ is a Γ-invariant locally finite positive measure on $\partial^2\widetilde{M}$.

Definition 2.2 (Geodesic current). – *A* Γ-invariant geodesic current*, or simply* geodesic current*, is a* Γ*-invariant positive Radon measure on* $\partial^2\widetilde{M}$.

Given any geodesic current μ and any admissible metric g on M, we will denote by m^g_μ the unique measure on $S^g M$ invariant by the geodesic flow (g^t) whose lift on $S^g\widetilde{M}$ is $\widetilde{m}^g_\mu = (H^g)^*(d\mu\times dt)$. The *non-wandering set* $\Omega^g\subset S^g M$ of the geodesic flow (g^t) is the image on $S^g M$ of the Γ-invariant set $\widetilde{\Omega}^g$ on $S^g\widetilde{M}$ defined by

$$\widetilde{\Omega}^g = (H^g)^{-1}\left((\Lambda_\Gamma\times\Lambda_\Gamma)\backslash\mathrm{Diag}\times\mathbb{R}\right).$$

It was shown by Eberlein [17] that for the geodesic flow of a negatively curved manifold, this definition coincides with the usual definition of the nonwandering set of a flow.

It follows from (3) and [9, Thm 1.7 p. 401] that Ω^g is compact (i.e., (M,g) is *convex-cocompact*) if and only if Ω^{g_0} is. We will mainly be interested in the case where Ω^g is *not compact*.

The measure m^g_μ is locally finite, but may have infinite mass as soon as (M,g) is not convex-cocompact. We will use all over this paper the fact that many properties of the measure m^g_μ only depend on the geodesic current μ and not on the chosen admissible metric g.

Recall first that an invariant measure is *ergodic* if every invariant set either has measure zero or its complementary set has measure zero.

Recall also that a sequence of measures (m_n) converges to m_∞ in the *vague* (respectively *narrow*) *topology* if for every continuous compactly supported (respectively bounded) function f, we have $\int f\,dm_n\to\int f\,dm_\infty$.

An invariant measure is *periodic* if it is (proportional to) the Lebesgue measure on a periodic orbit. The measure m is *conservative* if it satisfies the conclusion of Poincaré recurrence Theorem: for all sets A of positive measure $m(A)>0$, and m-almost all vectors v, the orbit $(g^t v)$ returns infinitely often in A. The measure m has a *product structure* if the associated geodesic current is equivalent to a product of measures on $\partial\widetilde{M}$. The measure m is *strongly mixing* if it is finite and satisfies $m(A \cap g^t B) \to m(A)m(B)$ when $t \to \pm\infty$ for all Borel sets A, B. It is *weakly mixing* if it is finite and $\frac{1}{T}\int_0^T \left|m(A \cap g^t B) - m(A)m(B)\right|$ goes to 0 when $T \to \pm\infty$ for all Borel sets A, B.

First well known properties are given in the following proposition.

PROPOSITION 2.3. – *Let μ be a geodesic current, let g_1 and g_2 be two admissible metrics on M. Then*

1. *the measure $m_\mu^{g_1}$ is supported by a (finite number of) closed geodesic(s) if and only if $m_\mu^{g_2}$ is;*
2. *the measure $m_\mu^{g_1}$ is ergodic for the geodesic flow (g_1^t) if and only if $m_\mu^{g_2}$ is ergodic for the geodesic flow (g_2^t);*
3. *the measure $m_\mu^{g_1}$ is conservative for the geodesic flow (g_1^t) if and only if $m_\mu^{g_2}$ is conservative for the geodesic flow (g_2^t);*
4. *the measure $m_\mu^{g_1}$ has a local product structure iff the measure $m_\mu^{g_2}$ has a local product structure.*

Proof. – The measure $m_\mu^{g_1}$ is supported by a closed geodesic if and only if μ is carried by the Γ-orbit of a couple $(\xi_-, \xi_+) \in \partial^2\widetilde{M}$ where ξ_- and ξ_+ are the fixed points of a hyperbolic element $\gamma \in \Gamma$. Since this property does not depend on g_1, it shows 1.

The measure $m_\mu^{g_1}$ is ergodic for the geodesic flow (g_1^t) if and only if μ is ergodic under the action of Γ on $\partial^2\widetilde{M}$ (cf for instance [46, p. 19]). This property only depends on μ, which shows 2.

The measure $m_\mu^{g_1}$ is conservative for the geodesic flow (g_1^t) if and only if μ gives full measure to $\Lambda_\Gamma^r \times \Lambda_\Gamma^r$ [46, Proof of (b) page 19] where Λ_Γ^r is the radial limit set, which does not depend on the (admissible) metric g_i. This shows 3. □

One should note that in general an invariant measure m_μ^g, even with finite total mass, has no reason to be a probability measure.

We will see further nontrivial relationships between $m_\mu^{g_1}$ and $m_\mu^{g_2}$ later. It would be interesting to know if this kind of result can be extended to (strong) mixing property. All known explicit examples of strongly mixing measures have a local product structure. But there exist mixing measures without such a product structure, for which the above question is relevant.

2.3. Geodesic stretches

Let g_1 and g_2 be two admissible metrics. For all $v \in S^{g_1}M$, define the quantity

$$e^{g_1 \to g_2}(v) = \liminf_{t \to +\infty} \frac{d^{g_2}(\pi\widetilde{v}, \pi g_1^t \widetilde{v})}{t}, \tag{4}$$

where $\widetilde{v}$ is a lift of v to $S^g\widetilde{M}$. This does not depend on the choice of $\widetilde{v}$. Knieper showed in [32] that if m is any invariant measure for (g_1^t), then for m-almost every $v \in S^{g_1}M$,

$$e^{g_1 \to g_2}(v) = \lim_{t \to +\infty} \frac{d^{g_2}(\pi\widetilde{v}, \pi g_1^t \widetilde{v})}{t}. \tag{5}$$

This asymptotic geodesic stretch has been studied by many authors, among which [20], [32], [23]. Sambarino uses a different point of view of reparametrization of the geodesic flow (see for example [47]) which is very close to our point of view below.

Recall that, $\xi \in \partial\widetilde{M}$ being fixed, the Busemann function $\mathcal{B}^g_\xi(.,.)$ is $\mathcal{C}^2$ on $\widetilde{M}^2$ [28, Prop. 3.1]. Therefore, for all $v \in S^{g_1}M$, we can define

$$\mathcal{E}^{g_1\to g_2}(v) = \left.\frac{d}{dt}\right|_{t=0^+} \mathcal{B}^{g_2}_{v^{g_1}_+}(\pi\widetilde{v}, \pi g_1^t\widetilde{v}) = \left.\frac{d}{dt}\right|_{t=0^+} \mathcal{B}^{g_2}_{v^{g_1}_+}(o, \pi g_1^t\widetilde{v}), \tag{6}$$

where $\widetilde{v} \in S^{g_1}\widetilde{M}$ is any lift of v, $v^{g_1}_+$ is[(2)] the positive endpoint in the boundary of the g_1-geodesic associated to $\widetilde{v}$, and $\mathcal{B}^{g_2}_{v^{g_1}_+}(.,.)$ is the Busemann function for g_2 based at the endpoint of the g_1-geodesic generated by v. This definition was inspired by Ledrappier's paper [34]. In his notations, our geodesic stretch satisfies $\mathcal{E}^{g_1\to g_2}(v) = \alpha^{g_2}(v)$, where α^{g_2} is the harmonic 1-form on the g_1-stable foliation associated to the Busemann cocycle of the metric g_2.

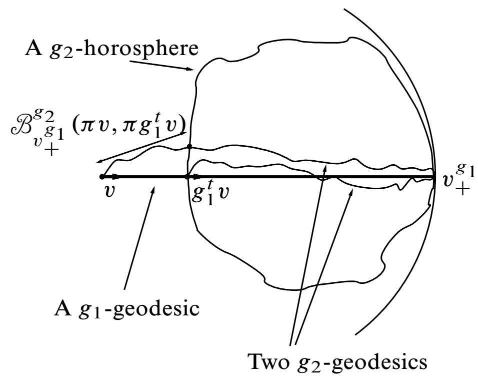


FIGURE 1. Geodesic stretch

DEFINITION 2.4 (Geodesic stretch). – *The maps* $e^{g_1\to g_2} : S^{g_1}M \to \mathbb{R}$ *and* $\mathcal{E}^{g_1\to g_2} : S^{g_1}M \to \mathbb{R}$ *will be called respectively the* asymptotic *and* instantaneous geodesic stretch *of* g_2 *with respect to* g_1.

Anyway, we will most of the time call them both without distinction *geodesic stretch*.

By construction, for all $v \in S^{g_1}M$, $\mathcal{E}^{g_1\to g_1}(v) = e^{g_1\to g_1}(v) = 1$. Observe that there is no obvious relation from the definition between $e^{g_1\to g_2}$ (resp. $\mathcal{E}^{g_1\to g_2}$) and $e^{g_2\to g_1}$ (resp. $\mathcal{E}^{g_2\to g_1}$).

If m is ergodic, then $e^{g_1\to g_2}$ is m-almost everywhere constant. Of course its value strongly depends on the measure m. On the opposite, the map $\mathcal{E}^{g_1\to g_2}$ is defined everywhere and does not depend on the chosen measure. It is in general non-constant, globally Hölder on $S^{g_1}M$ [4, Appendix of Brin], [38, Thm. 7.3], and $\mathcal{C}^1$ along g_1-geodesics (as Busemann functions are $\mathcal{C}^2$, see [28]). We will need the following basic estimate.

LEMMA 2.5. – *Let* g_1 *and* g_2 *be two admissible metrics, and* m *any* g_1*-invariant measure. For* m*-almost all* $v \in S^{g_1}\widetilde{M}$,

$$e^{g_1\to g_2}(v) \le \int_{S^{g_1}M} \|v\|^{g_2}\, dm,$$

(2) We omit the tilde for boundary points to avoid too heavy notations.

whereas for all $v \in S^{g_1}\widetilde{M}$,

$$\mathcal{E}^{g_1 \to g_2}(v) \le \|v\|^{g_2}.$$

Proof. – The first estimate was shown in [32, p. 44]. The second follows from triangular inequality. Indeed, for all $t \ge 0$, $\mathcal{B}^{g_2}_{v_+^{g_1}}(\pi(v), \pi(g_1^t v)) \le d^{g_2}(\pi(v), \pi(g_1^t v))$, and these two quantities vanish at $t = 0$ so that their derivatives at $t = 0$ satisfy the same inequality. Moreover, $d^{g_2}(\pi(v), \pi(g_1^t v))$ is smaller than the g_2-length of the curve $(\pi(g_1^s v))_{0 \le s \le t}$, whose derivative at zero is exactly $\|v\|_{g_2}$. □

Lemma 2.6 and Corollary 2.8 below justify the common name of *geodesic stretch* given to the two maps $e^{g_1 \to g_2}$ and $\mathcal{E}^{g_1 \to g_2}$. Before stating them, recall a well known feature of negative curvature. On a geodesic space X, each triangle (x, y, z) admits an interior triangle (p, q, r) such that $d(r, x) = d(q, x)$, $d(q, z) = d(p, z)$ and $d(p, y) = d(r, y)$. If g is a metric with negative curvature, there exists a universal constant $\Delta(g)$ such that for any geodesic triangle (x, y, z) in $\widetilde{M}$, the associated interior triangle has sides smaller than $\Delta(g)$ (see for example [9, p. 399] for a proof).

Lemma 2.6. – *There exists* $C_3 = C_3(g_1, g_2) > 0$*, depending only on the constant* $C_2(g_1, g_2)$ *(defined just below* (3)*) and the hyperbolicity constant* $\Delta(g_2)$*, such that for all* $\widetilde{v} \in S^{g_1}\widetilde{M}$ *and for all* $T > 0$,

$$\left| d^{g_2}(\pi\widetilde{v}, \pi g_1^T \widetilde{v}) - \mathcal{B}^{g_2}_{v_+^{g_1}}(\pi\widetilde{v}, \pi g_1^T \widetilde{v}) \right| = \left| d^{g_2}(\pi\widetilde{v}, \pi g_1^T \widetilde{v}) - \int_0^T \mathcal{E}^{g_1 \to g_2}(g_1^t \widetilde{v}) dt \right| \le C_3(g_1, g_2).$$

Proof. – Let $\widetilde{v} \in S^{g_1}\widetilde{M}$ and $T > 0$ be fixed. We write $x = \pi\widetilde{v}$, $x_T = \pi g_1^T \widetilde{v}$, and z_T is the intersection between the g_2-geodesic $(x, v_+^{g_1})^{g_2}$ and the g_2-horosphere centered at $v_+^{g_1}$ passing through x_T.

We will need at several occasions the following estimate.

Fact 2.7. – *With the above notations,* $d^{g_2}(x_T, z_T) \le 2C_2(g_1, g_2) + \Delta(g_2)$.

Let us first prove this fact. Consider the g_2-geodesic triangle $x, x_T, v_+^{g_1}$ and its interior triangle, say $p \in (x_T, v_+^{g_1})$, $q \in (x, x_T)$, $r \in (x, v_+^{g_1})$.

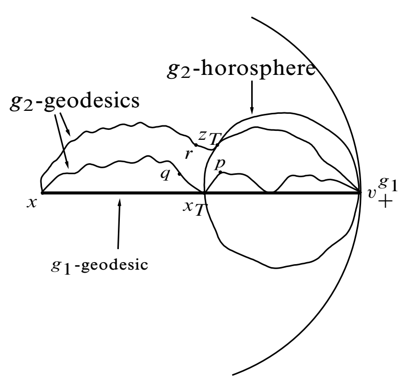


Figure 2. Proof of Lemma 2.6

Then by definition of z_T, $d^{g_2}(z_T,r)=d^{g_2}(x_T,q)$, so that

$$d^{g_2}(x_T,z_T)\le 2d^{g_2}(x_T,q)+d^{g_2}(q,r).$$

Now, the definition of (p,q,r) and Morse-Klingenberg lemma (see Section 2.1) imply $d^{g_2}(x_T,q)\le d^{g_2}(x_T,(x,v_+^{g_1}))\le C_2(g_1,g_2)$. The fact follows.

By definition (6),

$$\int_0^T \mathcal{E}^{g_1\to g_2}(g_1^t\widetilde{v})dt=\mathcal{B}^{g_2}_{v_+^{g_1}}(\pi\widetilde{v},\pi g_1^T\widetilde{v})=d^{g_2}(x,z_T).$$

Thanks to the above fact, we get

$$|d^{g_2}(x,x_T)-d^{g_2}(x,z_T)|\le d^{g_2}(x_T,z_T)\le 2C_2(g_1,g_2)+\Delta(g_2).$$

The result of the lemma follows, with $C_3(g_1,g_2)=2C_2(g_1,g_2)+\Delta(g_2)$. □

Corollary 2.8. – *Let m be a (g_1^t)-invariant probability measure on $S^{g_1}M$. Then*

$$\int_{S^{g_1}M} e^{g_1\to g_2}(v)dm(v)=\int_{S^{g_1}M}\mathcal{E}^{g_1\to g_2}(v)dm(v).$$

Moreover, when m is ergodic, for m-almost every $v\in S^{g_1}M$ and all lifts $\widetilde{v}\in S^{g_1}\widetilde{M}$ of v,

$$\lim_{T\to+\infty}\frac{d^{g_2}(\pi\widetilde{v},\pi g_1^T\widetilde{v})}{T}=\lim_{T\to+\infty}\frac{1}{T}\int_0^T\mathcal{E}^{g_1\to g_2}(g_1^t\widetilde{v})dt=\int_{S^{g_1}M}\mathcal{E}^{g_1\to g_2}(w)dm(w).$$

Proof. – It follows from the previous lemma that for all $\varepsilon>0$, there exists $T_0>0$ such that for all $T\ge T_0$ and all $\widetilde{v}\in S^{g_1}\widetilde{M}$,

$$\frac{1}{T}\left|d^{g_2}(\pi\widetilde{v},\pi g_1^T\widetilde{v})-\int_0^T\mathcal{E}^{g_1\to g_2}(g_1^t\widetilde{v})dt\right|\le\varepsilon.$$

It yields the first equality.

When m is ergodic, for m-almost all vectors $v\in S^{g_1}$,

$$\int_{S^{g_1}M}\mathcal{E}^{g_1\to g_2}(v)dm(v)=\lim_{T\to+\infty}\frac{1}{T}\int_0^T\mathcal{E}^{g_1\to g_2}(g_1^t\widetilde{v})dt$$

and

$$\int_{S^{g_1}M}e^{g_1\to g_2}(v)dm(v)=e^{g_1\to g_2}(v)=\lim_{T\to+\infty}\frac{d^{g_2}(\pi\widetilde{v},\pi g_1^T\widetilde{v})}{T},$$

which concludes the proof of the corollary. □

Let us emphasize the fact that the measures that we will consider will usually have finite mass, but may not be probability measures. We will denote by $\|m\|$ the mass of a finite measure m on TM.

Definition 2.9 (Geodesic stretch with respect to a geodesic current).

Let μ be a geodesic current on $\partial^2\widetilde{M}$ such that m_μ^g is finite. We will call (average) geodesic stretch of g_2 relative to g_1 with respect to μ *the quantity*

$$I_\mu(g_1,g_2)=\frac{1}{\|m_\mu^{g_1}\|}\int_{S^{g_1}M}\mathcal{E}^{g_1\to g_2}(v)dm_\mu^{g_1}(v)=\frac{1}{\|m_\mu^{g_1}\|}\int_{S^{g_1}M}e^{g_1\to g_2}(v)dm_\mu^{g_1}(v).$$

By Corollary 2.8, $I_\mu(g_1, g_2)$ coincides with the definition of the geodesic stretch studied in [32] (note that Knieper only considers invariant *probability* measures).

When (M, g) has finite volume and μ is the Liouville geodesic current of g_1, then

$$I_\mu(g_1, g_2).\mathrm{Vol}(S^{g_1}M) = i(g_1, g_2),$$

where $i(g_1, g_2)$ is the *intersection* between the metrics g_1 and g_2 studied in [20].

It follows from the definition that for all geodesic currents μ such that m^g_μ is finite, $I_\mu(g_1, g_1) = 1$.

Remark 2.10 (Geodesic stretches and Thurston metric). – Given two negatively curved metrics g_1 and g_2 on a compact surface S, the Thurston distance $d_{Th}(g_1, g_2)$ is defined as the supremum over all periodic orbits of the ratios of their lengths:

$$d_{Th}(g_1, g_2) = \sup_\gamma \left(\frac{\ell^{g_2}(\gamma)}{\ell^{g_1}(\gamma)}, \frac{\ell^{g_1}(\gamma)}{\ell^{g_2}(\gamma)} \right).$$

With our notations, this distance could also be defined as the following supremum

$$d_{Th}(g_1, g_2) = \sup_\mu \left(I_\mu(g_1, g_2), I_\mu(g_2, g_1) \right)$$

over all currents μ associated to ergodic measures. Indeed, considering periodic measures immediately shows that Thurston distance is smaller than the above supremum. In the other direction, the density of periodic measures in the set of ergodic measures, see [13], gives the above equality.

2.4. Morse correspondences and geodesic stretches

To compare dynamics of the geodesic flows on $S^{g_1}M$ and $S^{g_2}M$, it is natural to consider their dynamics modulo the Γ-action on $S^{g_1}\widetilde{M}$ and $S^{g_2}\widetilde{M}$. Hopf coordinates are a good motivation to consider the map

$$\widetilde{\Phi}^{g_1\to g_2} := (H^{g_2})^{-1} \circ H^{g_1} : S^{g_1}\widetilde{M} \to S^{g_2}\widetilde{M}.$$

It is a Hölder homeomorphism, but it is unfortunately not Γ-equivariant, as both Γ-actions on each unit tangent bundle $S^{g_i}\widetilde{M}$ are different. In other words, as said earlier, on $\partial^2\widetilde{M} \times \mathbb{R}$, these Γ-actions involve different cocycles on the $\mathbb{R}$ component.

Despite its non-invariance, this map is sometimes useful, because it has the nice property to commute with both geodesic flows. But we need to find another map from $S^{g_1}\widetilde{M}$ to $S^{g_2}\widetilde{M}$ which will be Γ-equivariant. We proceed as follows. For all $v \in S^{g_1}\widetilde{M}$, let $w = \widetilde{\Psi}^{g_1\to g_2}(v)$ be the unique vector in $S^{g_2}\widetilde{M}$ on the g_2-geodesic joining $v^{g_1}_-$ to $v^{g_1}_+$ satisfying $\mathcal{B}^{g_2}_{v^{g_1}_+}(\pi(v), \pi(w)) = 0$.

Lemma 2.11. – *The map $\Psi^{g_1\to g_2}$ is Hölder continuous. Moreover, for all $v \in S^{g_1}\widetilde{M}$, we have*

$$d^{g_2}(\pi v, \pi\Psi^{g_1\to g_2}(v)) \le C_3(g_1, g_2),$$

where $C_3(g_1, g_2)$ is the constant given by Lemma 2.6.

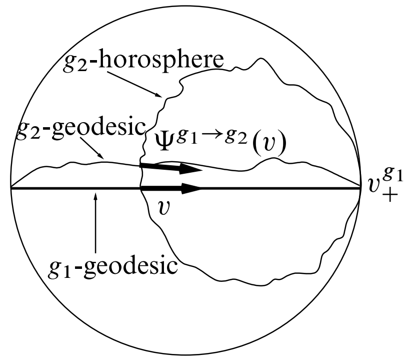


FIGURE 3. Morse correspondence

Proof. – It is Hölder continuous as composition of the maps $\Phi^{g_1\to g_2}$ and some time g_2^τ of the geodesic flow, with $\tau = \tau^{g_1\to g_2}(v)$ depending Hölder-continuously on v (see formula in Lemma 2.12 (4) below).

The bound on $d^{g_2}(\pi v, \pi\Psi^{g_1\to g_2}(v))$ has already been proved in Fact 2.7. □

By construction, the correspondence $\widetilde{\Psi}^{g_1\to g_2}$ is Γ-invariant. We denote by $\Psi^{g_1\to g_2}$ the induced map from $S^{g_1}M$ to $S^{g_2}M$. It is a homeomorphism homotopic to identity sending (g_1^t)-orbits to (g_2^t)-orbits, i.e., a (g_1, g_2)-*Morse correspondence* in the sense of [20].

By definition of both correspondences, the following lemma holds. It says that the geodesic flows (g_1^t) and (g_2^s) on the unit tangent bundles $S^{g_i}\widetilde{M}$ are conjugated by $\Phi^{g_1\to g_2}$, and conjugated up to reparametrization by the Morse correspondance $\Psi^{g_1\to g_2}$.

LEMMA 2.12. – *With the above notations, we have for all* $v \in S^{g_1}\widetilde{M}$

1. $\Phi^{g_1\to g_2} \circ g_1^t(v) = g_2^t \circ \Phi^{g_1\to g_2}(v)$.
2. $\Phi^{g_2\to g_1} = \left(\Phi^{g_1\to g_2}\right)^{-1}$.
3. $\Psi^{g_1\to g_2} \circ g_1^t(v) = g_2^{s^{g_1\to g_2}(t,v)} \circ \Psi^{g_1\to g_2}(v)$, *with* $s^{g_1\to g_2}(t, v) = \mathcal{B}^{g_2}_{v_+^{g_1}}(\pi(v), \pi(g_1^t v))$.
4. $\Psi^{g_1\to g_2}(v) = g_2^{\tau^{g_1\to g_2}(v)} \circ \Phi^{g_1\to g_2}(v)$, *with*
$$\tau^{g_1\to g_2}(v) = \mathcal{B}^{g_2}_{v_+^{g_1}}(o, \pi(v)) - \mathcal{B}^{g_1}_{v_+^{g_1}}(o, \pi(v)).$$
5. $\Psi^{g_2\to g_1} \circ \Psi^{g_1\to g_2}(v) = g_1^{\sigma^{g_1\to g_2}(v)}(v)$, *with* $\sigma^{g_1\to g_2}(v) = \mathcal{B}^{g_1}_{v_+^{g_1}}\left(\pi(v), \pi(\Psi^{g_1\to g_2}v)\right)$.

Let us emphasize that $\Phi^{g_1\to g_2}$ and its inverse are not Γ-invariant, $\Psi^{g_1\to g_2}$ and its inverse are Γ-invariant, the map $\tau^{g_1\to g_2}$ is not Γ-invariant, whereas $\sigma^{g_1\to g_2}$ and the cocycle $s^{g_1\to g_2}(t, v)$ are Γ-invariant.

Proof. – The fact that $\Phi^{g_1\to g_2}$ commutes with the geodesic flows of g_1 and g_2 is immediate by definition of Hopf coordinates. The property about its inverse is also obvious.

By definition of $\Psi^{g_1\to g_2}$, the vectors $\Psi^{g_1\to g_2}(g_1^t v)$, for $t \in \mathbb{R}$, all lie on the g_2-geodesic joining $v_-^{g_1}$ to $v_+^{g_2}$. The only question is to compute
$$s^{g_1\to g_2}(t, v) = \mathcal{B}^{g_2}_{v_+^{g_1}}(\pi(\Psi^{g_1\to g_2}(v)), \pi(\Psi^{g_1\to g_2}(g_1^t v))).$$

By definition of $\Psi^{g_1\to g_2}$,

$$\mathcal{B}^{g_2}_{v_+^{g_1}}(\pi\Psi^{g_1\to g_2}(g_1^t v), \pi(g_1^t v)) = 0 = \mathcal{B}^{g_2}_{v_+^{g_1}}(\pi\Psi^{g_1\to g_2}(v), \pi(v)).$$

Using the cocycle properties of $\mathcal{B}^{g_2}_{v_+^{g_1}}$, we deduce immediately that $s^{g_1\to g_2}(t,v)$ is the algebraic g_2-distance $\mathcal{B}^{g_2}_{v_+^{g_1}}(\pi(v), \pi(g_1^t v))$.

The next affirmation follows from the computation

$$\begin{aligned}\tau^{g_1\to g_2}(v) = \mathcal{B}^{g_2}_{v_+^{g_1}}(\Phi^{g_1\to g_2}(v), \Psi^{g_1\to g_2}(v)) &= \mathcal{B}^{g_2}_{v_+^{g_1}}(o, \pi(v)) - \mathcal{B}^{g_2}_{v_+^{g_1}}(o, \Phi^{g_1\to g_2}(v))\\ &= \mathcal{B}^{g_2}_{v_+^{g_1}}(o, \pi(v)) - \mathcal{B}^{g_1}_{v_+^{g_1}}(o, v).\end{aligned}$$

The last statement follows easily from the previous one. □

2.5. Change of mass

We will need the following variant of Lemma 2.6, which shows once more that $\mathcal{E}^{g_1\to g_2}$ behaves asymptotically as the infinitesimal reparametrization of the flow given by Morse correspondance $\Psi^{g_1\to g_2} : S^{g_1}\widetilde{M} \to S^{g_2}\widetilde{M}$.

Proposition 2.13. – *Let $G : S^{g_2}M \to \mathbb{R}$ be a continuous map and $\widetilde{G} : S^{g_2}\widetilde{M} \to \mathbb{R}$ be its (Γ-invariant) lift to $S^{g_2}\widetilde{M}$. Then for all $v \in S^{g_1}\widetilde{M}$, $T \geq 0$, and $w = \Psi^{g_1\to g_2}(v)$, we have*

$$\int_0^{s^{g_1\to g_2}(T,v)} \widetilde{G}(g_2^s w)\,ds = \int_0^T \widetilde{G}\circ\Psi^{g_1\to g_2}(g_1^t v) \times \mathcal{E}^{g_1\to g_2}(g_1^t v)\,dt,$$

with $s^{g_1\to g_2}(T,v) = \mathcal{B}^{g_2}_{v_+^{g_1}}(\pi(v), \pi(g_1^T v))$ as in Lemma 2.12.

If moreover G is bounded, then there exists $C = C(G, g_1, g_2)$ such that for all $v \in S^{g_1}\widetilde{M}$, $T \geq 0$, and $w = \Psi^{g_1\to g_2}(v)$, we have

$$\left|\int_0^{d^{g_2}(v, g_1^T v)} \widetilde{G}(g_2^s w), ds - \int_0^T \widetilde{G}\circ\Psi^{g_1\to g_2}(g_1^t v) \times \mathcal{E}^{g_1\to g_2}(g_1^t v)\,dt\right| \leq C.$$

If G is not bounded, then for all compact sets $K \subset S^{g_1}M$ there exists another constant $C' = C'(G, K, g_1, g_2)$ such that for all $v \in S^{g_1}\widetilde{M}$ and $T \in \mathbb{R}$ such that both v and $g_1^T v$ belong to $\widetilde{K} = p_\Gamma^{-1}(K) \subset S^{g_1}\widetilde{M}$, we have

$$\left|\int_0^{d^{g_2}(v, g_1^T v)} \widetilde{G}(g_2^s w), ds - \int_0^T \widetilde{G}\circ\Psi^{g_1\to g_2}(g_1^t v) \times \mathcal{E}^{g_1\to g_2}(g_1^t v)\,dt\right| \leq C'.$$

The geodesic stretch $\mathcal{E}^{g_1\to g_2}$ can therefore be understood as the instantaneous reparametrization of the flow (g_1^t) in the correspondance $\Psi^{g_1\to g_2}$.

Proof. – The first equality is a simple change of variable using Lemma 2.12. The second follows using Lemma 2.6 and the fact that G is bounded. Indeed,

$$\left|\int_0^{d^{g_2}(v, g_1^T v)} \widetilde{G}(g_2^s w), ds - \int_0^T \widetilde{G}\circ\Psi^{g_1\to g_2}(g_1^t v) \times \mathcal{E}^{g_1\to g_2}(g_1^t v)\,dt\right|$$

$$
\begin{aligned}
&= \left| \int_{s(T,v)}^{d^{g_2}(\pi(v),\pi(g_1^T v))} \widetilde{G}(g_2^s w)\, ds \right| \\
&\leq \|G\|_\infty \times \left| d^{g_2}(\pi(v), \pi(g_1^T v)) - \mathcal{B}^{g_2}_{v_+^{g_1}}(\pi(v), \pi(g_1^t v)) \right| \\
&= \|G\|_\infty \times \left| d^{g_2}(\pi(v), \pi(g_1^T v)) - \int_0^T \mathcal{E}^{g_1\to g_2}(g_1^t v)\, dt \right| \\
&\leq C_1 \|G\|_\infty .
\end{aligned}
$$

The last assertion is a variation on the second one. If v and $g_1^T v$ are in a compact set K, for any parameter s such that $|s| \leq C_3(g_1, g_2)$, $g_1^{T\pm s} v$ belongs to the $C_3(g_1, g_2)$-neighborhood of K, on which G is bounded. The above computation therefore applies verbatim. □

Remark 2.14. – Proposition 1.3 follows immediately. Given any $m_\mu^{g_2}$-measurable map $G : S^{g_2}M \to \mathbb{R}$, the map $G \circ \Psi^{g_1\to g_2}$ is $m_\mu^{g_1}$-measurable and G on $S^{g_2}M$, and we have

$$
\int_{S^{g_2}M} G\, dm_\mu^{g_2} = \int_{S^{g_1}M} G \circ \Psi^{g_1\to g_2} \times \mathcal{E}^{g_1\to g_2}\, dm_\mu^{g_1}.
$$

The corollary below follows immediately from the above remark. It gives a nice interpretation of the geodesic stretch $I_\mu(g_1, g_2)$.

Corollary 2.15 (Mass transformation law). – *Let μ be a geodesic current such that $m_\mu^{g_1}$ has finite total mass, denoted by $\|m_\mu^{g_1}\|$. Then*

$$
\|m_\mu^{g_2}\| = I_\mu(g_1, g_2) \times \|m_\mu^{g_1}\|.
$$

In particular $m_\mu^{g_1}$ has finite mass if and only if $m_\mu^{g_2}$ has finite mass. Moreover, when it is the case,

$$
I_\mu(g_1, g_2) = \frac{1}{I_\mu(g_2, g_1)} = \frac{\|m_\mu^{g_2}\|}{\|m_\mu^{g_1}\|}.
$$

Remark 2.16. – The previous formula is very natural if $I_\mu(g_2, g_1)$ is interpreted as the average dilation of the reparametrization of the flow via the Morse correspondance $\Psi^{g_1\to g_2}$. Indeed, in the case where (g_1^t) and (g_2^t) are suspension flows over a (fixed) compact basis for distinct ceiling functions, the above formula is well known [1].

2.6. Periodic orbits and geodesic stretch

In this section we relate geodesic stretch and lengths of periodic orbits. The results will not be useful in the sequel of the paper, but are enlightening about the geodesic stretch.

For $i = 1, 2$, for any hyperbolic element $\gamma \in \Gamma$, let γ^{g_i} be the closed g_i-geodesic associated to the conjugacy class of γ. Let $\ell^{g_i}(\gamma)$ be its g_i-length, and $d\ell_\gamma^{g_i}$ be the Lebesgue measure along the geodesic γ^{g_i}. Observe that, up to normalizing constants, the periodic measure $d\ell_\gamma^{g_i}$, $i = 1, 2$, induce the same current at infinity.

Since $m_\mu^{g_1}$ is finite and ergodic, there exists a sequence $(\gamma_k)_{k\in\mathbb{N}}$ of hyperbolic elements such that in the weak topology,

$$
\lim_{k\to\infty} \frac{d\ell_{\gamma_k^{g_1}}}{\ell^{g_1}(\gamma_k)} = \frac{m_\mu^{g_1}}{\|m_\mu^{g_1}\|},
$$

see for instance [13, Lemma 2.2]. This convergence holds a priori in the dual of continuous functions with compact support. But as all measures involved above are probability measures, this convergence also holds in the dual of bounded continuous functions of $S^{g_1}M$.

We can moreover suppose that $\lim_{k\to\infty} \ell^{g_1}(\gamma_k) = +\infty$.

The following proposition shows that the same happens on $S^{g_2}M$, and that the ratio of lengths of periodic orbits in both metrics allows to recover the geodesic stretch.

PROPOSITION 2.17. – *Let (M, g_i), $i = 1, 2$, be two admissible Riemannian structures with pinched negative curvature. Let μ be a geodesic current such that both measures $m^{g_i}_\mu$ are finite. Let (γ_k) be a sequence of hyperbolic elements such that $\frac{d\ell^{g_1}_{\gamma_k}}{\ell^{g_1}(\gamma_k)}$ converges weakly to $\frac{m^{g_1}_\mu}{\|m^{g_1}_\mu\|}$ in the dual of bounded continuous functions. Then $\frac{d\ell^{g_2}_{\gamma_k}}{\ell^{g_2}(\gamma_k)}$ converges weakly to $\frac{m^{g_2}_\mu}{\|m^{g_2}_\mu\|}$ in the dual of bounded continuous functions.*

Moreover, the ratios of lengths satisfy

$$\lim_{k\to+\infty} \frac{\ell^{g_2}(\gamma_k)}{\ell^{g_1}(\gamma_k)} = I_\mu(g_1, g_2).$$

The proof is separated in two lemmas. The first one asserts that viewed on $S^{g_2}M$, the sequence of periodic probability measures associated to (γ_k) also converges to $\frac{m^{g_2}_\mu}{\|m^{g_2}_\mu\|}$ in the dual of bounded continuous functions. The second says that the ratio of lengths $\ell^{g_2}(\gamma_k)/\ell^{g_1}(\gamma_k)$ converges to the average geodesic stretch $I_\mu(g_1, g_2)$.

LEMMA 2.18. – *With the previous notations, for the same sequence (γ_k), in the dual of continuous bounded functions of $S^{g_2}M$,*

$$\lim_{k\to\infty} \frac{d\ell^{g_2}_{\gamma_k}}{\ell^{g_2}(\gamma_k)} = \frac{m^{g_2}_\mu}{\|m^{g_2}_\mu\|}.$$

Proof. – First, as the sequence of probability measures $\frac{d\ell^{g_1}_{\gamma_k}}{\ell^{g_1}(\gamma_k)}$ converges to the probability measure $\frac{m^{g_1}_\mu}{\|m^{g_1}_\mu\|}$, the Γ-invariant lift of $\frac{d\ell^{g_1}_{\gamma_k}}{\ell^{g_1}(\gamma_k)}$ to $S^{g_1}\widetilde{M}$ converges in the dual of continuous functions with compact support towards $\frac{\widetilde{m}^{g_1}_\mu}{\|m^{g_1}_\mu\|}$. Using Hopf coordinates, we deduce that the geodesic current on $\partial^2\widetilde{M}$ associated through H^{g_1} to $\frac{d\ell_{\gamma_k^{g_1}}}{\ell^{g_1}(\gamma_k)}$ converges weakly (in the dual of continuous functions with compact support) to μ. Using the same argument in the other direction, we obtain that the sequence of probability measures $\frac{d\ell^{g_2}_{\gamma_k}}{\ell^{g_2}(\gamma_k)}$ converges weakly (in the dual of continuous functions with compact support) to some multiple of $m^{g_2}_\mu/\|m^{g_2}_\mu\|$.

It is not exactly the desired result. To get the convergence towards the probability measure $m^{g_2}_\mu/\|m^{g_2}_\mu\|$, and in the dual of bounded continuous functions, we need to avoid a possible loss of mass at infinity. To establish this convergence, it is necessary and sufficient to prove that $\frac{d\ell^{g_2}_{\gamma_k}}{\ell^{g_2}(\gamma_k)}$ does not diverge. In other words, we want to check that for all $\varepsilon > 0$, there exists a compact set $\mathscr{K}_\varepsilon \subset S^{g_2}M$, such that for all $k \geq 0$ large enough,

$$\frac{d\ell^{g_2}_{\gamma_k}}{\ell^{g_2}(\gamma_k)}(\mathscr{K}_\varepsilon) \geq 1 - \varepsilon.$$

It follows easily from the fact that there exists a constant $C = C(g_1, g_2)$ such that any g_2-geodesic of $\widetilde{M}$ stays in a $C(g_1, g_2)$-neighborhood of the g_1-geodesic with same endpoints at infinity. Let us write the detail of the argument.

Choose first some $\varepsilon > 0$, and some compact set $K_1 \subset M$ such that $\frac{m_\mu^{g_1}(S^{g_1}K_1)}{\|m_\mu^{g_1}\|} \geq 1-\varepsilon/2$. By convergence of $\frac{d\ell_{\gamma_k}^{g_1}}{\ell^{g_1}(\gamma_k)}$, for all $k \geq k_0$ large enough, we also have $\frac{\ell_{\gamma_k}^{g_1}(S^{g_1}K_1)}{\ell^{g_1}(\gamma_k)} \geq 1-\varepsilon$.

Now, choose a relatively compact preimage $\widetilde{K}_1 \Subset \widetilde{M}$, its g_2-convex closure $\widetilde{K}_2$ and $\widetilde{K}_3 \supset \widetilde{K}_2$ a larger compact convex set of $\widetilde{M}$ containing a $2C(g_1, g_2)$-neighborhood of $\widetilde{K}_2$ for both metrics g_1 and g_2.

Consider a lift $\widetilde{\gamma}_k^{g_1}$ of the g_1-geodesic $\gamma_k^{g_1}$ which intersects $\widetilde{K}_1$, and the associated lift $\widetilde{\gamma}_k^{g_2}$ of the g_2-geodesic $\gamma_k^{g_2}$, at distance at most $C(g_1, g_2)$ from $\widetilde{\gamma}_k^{g_1}$. Let a, b be two points on $\widetilde{\gamma}_k^{g_2}$ such that the length $\ell_{\gamma_k}^{g_2}((a,b)) = \ell^{g_2}(\gamma_k)$. We want to estimate the proportion of g_2-length of $[a,b]$ outside $\Gamma.\mathscr{K} = \Gamma.S^{g_2}K_3$.

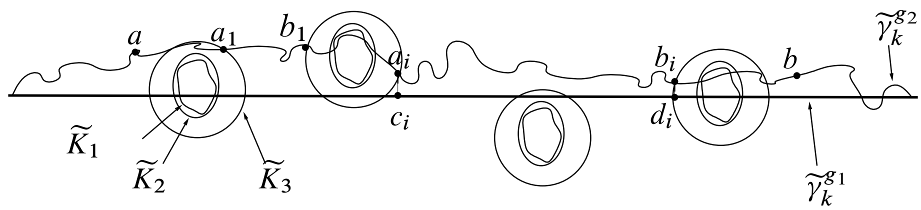


FIGURE 4. Proof of Lemma 2.18

By convexity of $\widetilde{K}_3$, we can write $(a,b) \cap (\Gamma.\widetilde{K}_3)^c$ as the disjoint union $\sqcup(a_i, b_i)$ of finitely many intervals. Thus, we have to show that

$$\frac{\ell_{\gamma_k}^{g_2}(\mathscr{K}^c)}{\ell^{g_2}(\gamma_k)} = \frac{\sum_i \ell_{\gamma_k}^{g_2}(a_i, b_i)}{\ell^{g_2}(\gamma_k)} \leq \varepsilon.$$

Choose two points c_i and d_i on $\widetilde{\gamma}_k^{g_1}$ whose projections (for the metric g_2) on the g_2-geodesic $\widetilde{\gamma}_k^{g_2}$ are exactly a_i and b_i. Such points are not necessarily unique but always exist: take c_i in the intersection of $\widetilde{\gamma}_k^{g_1}$ with the hyperplane orthogonal to $\widetilde{\gamma}_k^{g_2}$ at a_i. Denote by (c_i, d_i) the g_1-geodesic segment on $\widetilde{\gamma}_k^{g_1}$, and let $\lambda > 0$ be such that $\frac{1}{\lambda} g_2 \leq g_1 \leq \lambda g_2$. We have

$$\ell_{\gamma_k}^{g_2}(a_i, b_i) = d^{g_2}(a_i, b_i) \leq d^{g_2}(c_i, d_i) \leq \ell^{g_2}(c_i, d_i) \leq \sqrt{\lambda}\,\ell^{g_1}(c_i, d_i).$$

We deduce that

$$\frac{\ell_{\gamma_k}^{g_2}(\mathscr{K}^c)}{\ell^{g_2}(\gamma_k)} \leq \lambda \sum_i \frac{\ell^{g_1}(c_i, d_i)}{\ell^{g_1}(\gamma_k)} \leq \lambda \frac{\ell^{g_1}((S^{g_1}\Gamma.\widetilde{K}_1)^c) \cap \widetilde{\gamma}_k^{g_1})}{\ell^{g_1}(\gamma_k)},$$

the last inequality coming from the fact that, as a_i and b_i are in the boundary of $\Gamma.\widetilde{K}_3$, and c_i and d_i are at distance at most $C(g_1, g_2)$ resp. from a_i and d_i, they cannot belong to $\Gamma.\widetilde{K}_2$, so that the segment (c_i, d_i) does not intersect $\Gamma.\widetilde{K}_1$. This proves that

$$\frac{\ell_{\gamma_k}^{g_2}(\mathscr{K}^c)}{\ell^{g_2}(\gamma_k)} \leq \lambda\varepsilon,$$

which concludes the proof (up to changing ε in ε/λ). $\square$

Moreover, the lengths $\ell^{g_1}(\gamma_k)$ and $\ell^{g_2}(\gamma_k)$ are related as follows.

LEMMA 2.19. – *With the previous notations, for the same sequence (γ_k),*

$$\lim_{k\to+\infty}\frac{\ell^{g_2}(\gamma_k)}{\ell^{g_1}(\gamma_k)} = I_\mu(g_1,g_2).$$

Proof. – For all $k\in\mathbb{N}$, let $v_k^{g_1}$ (resp. $v_k^{g_2}$) be a tangent vector to $\gamma_k^{g_1}$ (resp. $\gamma_k^{g_2}$) such that $d^{g_2}(\pi v_k^{g_1},\pi v_k^{g_2}) \le C_2(g_1,g_2)$, where $C_2(g_1,g_2) > 0$ is the constant defined just after (3). Let $\widetilde{v}_k^{g_1} \in S^{g_1}\widetilde{M}$ and $\widetilde{v}_k^{g_2} \in S^{g_2}\widetilde{M}$ be lifts of $v_k^{g_1}$ and $v_k^{g_2}$ such that again, $d^{g_2}(\pi\widetilde{v}_k^{g_1},\pi\widetilde{v}_k^{g_2}) \le C_2(g_1,g_2)$. It follows from Proposition 2.13 applied to $F\equiv 1$ that there exists $c_1>0$, only depending on $C_2(g_1,g_2)$ and the bounds on the curvature, such that

$$\left|\ell^{g_2}(\gamma_k) - \int_0^{\ell^{g_1}(\gamma_k)} \mathcal{E}^{g_1\to g_2}(g_1^t\widetilde{v})dt\right| \le c_1.$$

Therefore,

$$\left|\frac{\ell^{g_2}(\gamma_k)}{\ell^{g_1}(\gamma_k)} - \frac{1}{\ell^{g_1}(\gamma_k)}\int_0^{\ell^{g_1}(\gamma_k)} \mathcal{E}^{g_1\to g_2}(g_1^t\widetilde{v})dt\right| \le \frac{c_1}{\ell^{g_1}(\gamma_k)}.$$

By Lemma 2.18, as $\mathcal{E}^{g_1\to g_2}$ is bounded and continuous, we know that

$$\frac{1}{\ell^{g_1}(\gamma_k)}\int_0^{\ell^{g_1}(\gamma_k)} \mathcal{E}^{g_1\to g_2}(g_1^t\widetilde{v})dt \to I_\mu(g_1,g_2),$$

so that the conclusion follows. □

3. Entropy of finite measures

In this section, given two admissible metrics g_1 and g_2 as before, and a geodesic current μ on $\partial^2\widetilde{M}$, we wish to compare the entropies of the measures $m_\mu^{g_1}$ and $m_\mu^{g_2}$. Theorem 3.11 establishes that their ratio is the average geodesic stretch between g_1 and g_2 w.r.t μ, but in the reverse direction compared to the relation between their masses, which leads to Corollary 3.12, which states that the product of the entropy of $m_\mu^{g_i}$ by its mass $\|m_\mu^{g_i}\|$ remains constant under an admissible change of metric.

First, we will recall some definitions and relations between dynamical balls (Subsection 3.1). In Subsection 3.2, we compare two notions of entropy of a measure, the Kolmogorov-Sinai entropy and the local Brin-Katok entropy, recalling well and less known results of Brin-Katok and Riquelme. It allows us to prove Theorem 3.11 and Corollary 3.12 in Subsection 3.3.

3.1. Dynamical balls and shadows

If (φ^t) is a continuous dynamical system on a metric space (X,d), a dynamical ball is a ball for the dynamical distance $d_T(x,y)=\sup_{0\le t\le T} d(\varphi^t x,\varphi^t y)$.

We will restrict ourselves to geodesic flows associated to a Riemannian metric g on S^gM. For such geometric dynamical systems, it is more convenient to work with the Riemannian distance induced by the metric g on M or $\widetilde{M}$ instead of the distance coming from the Sasaki metric on TM or $T\widetilde{M}$. We refer to [4, p. 70] and [38, p. 19–20] for a discussion about the fact that it is the good thing to do in this case.

For all $\varepsilon, T > 0$ and $v \in S^g\widetilde{M}$, we will call *dynamical ball* of center v, diameter ε and length T the set

$$B^g(v,T,\varepsilon) = \{w \in S^g\widetilde{M},\ d^g(\pi(g^t v),\pi(g^t w)) \leq \varepsilon, \text{ for all } 0 \leq t \leq T\}.$$

Note that $B^g(v,0,\varepsilon)$ is the ε-ball with center v for the distance d^g defined above.

REMARK 3.1. – On the quotient, for $v \in S^g M$, one can either consider the quotient dynamical ball $B^g(v,T,\varepsilon) = p_\Gamma(B^g(\widetilde{v},T,\varepsilon))$, $\widetilde{v}$ being any lift of v to $S^g\widetilde{M}$. There is also a more dynamical definition, as

$$B^g_{\mathrm{dyn}}(v,T,\varepsilon) = \{w \in S^g M,\ d^g(\pi(g^t v),\pi(g^t w)) \leq \varepsilon, \text{ for all } 0 \leq t \leq T\}.$$

Of course, if $\widetilde{v} \in S^g\widetilde{M}$ and $v = p_\Gamma(\widetilde{v}) \in S^g M$, one has the obvious inclusion

$$B^g(v,T,\varepsilon)) = p_\Gamma(B^g(\widetilde{v},T,\varepsilon)) \subset B^g_{\mathrm{dyn}}(v,T,\varepsilon). \tag{7}$$

One can easily see that this inclusion is an equality when the injectivity radius of M is uniformly bounded from below, as soon as ε is small enough. However, when the injectivity radius of M is not bounded from below, one can build examples where this inclusion is not an equality [6].

It turns out that in many cases, the most natural dynamical ball to consider is the small ball $p_\Gamma(B^g(\widetilde{v},T,\varepsilon)$. Therefore, we will call it the *small dynamical ball* and denote it by $B^g(v,T,\varepsilon)$.

This problem has not been emphasized in [38], where only these small dynamical balls are considered (see [38, 3.15]). However, in various definitions of local entropies, the large dynamical balls have to be considered.

We will also need the following variant, for $v \in S^g\widetilde{M}$ and $T, T' > 0$:

$$B^g(v;T,T',\varepsilon) = \{w \in S^g\widetilde{M}, d^g(\pi(v),\pi(w)) \leq \varepsilon,\ \text{for all } -T' \leq t \leq T\}.$$

Observe that $B^g(v;T,T',\varepsilon) = g^{T'}(B^g(g^{-T'}v, T+T',\varepsilon))$. As mentioned in the above Remark 3.1, we consider on $S^g M$ the small dynamical balls

$$B^g(v;T,T',\varepsilon) = p_\Gamma(B^g(\widetilde{v};T,T',\varepsilon)).$$

Recall the following well known fact in negative curvature.

LEMMA 3.2. – *Let (M,g) be a manifold with pinched negative curvature. For all $0 < a < b$, there exists a constant $c = c(a,b) > 0$ such that for all vectors $v, w \in S^g\widetilde{M}$, and all $T > 2c$, if $d^g(\pi(g_t v),\pi(g_t w)) \leq b$ for all $0 < t < T$, then $d^g(g_t v, g_t w) \leq a$ for all $c < t < T - c$.*

Proof. – This is an exercise using standard comparison results. Note that the constant $c(a,b)$ also depends on the upper bound of the curvature. □

LEMMA 3.3. – *Let (M,g) be a manifold with pinched negative curvature. For all $0 < \varepsilon_1 < \varepsilon_2$, there exists $C(g,\varepsilon_1,\varepsilon_2) > 0$ such that for all $v \in S^g\widetilde{M}$ and $T, T' > 0$, we have*

$$B^g(v;T + C(g,\varepsilon_1,\varepsilon_2), T' + C(g,\varepsilon_1,\varepsilon_2),\varepsilon_2) \subset B^g(v;T,T',\varepsilon_1) \subset B^g(v;T,T',\varepsilon_2).$$

Proof. – The right inclusion is obvious. The left one comes from Lemma 3.2 above. □

The shadow $\mathcal{O}^g_x(B^g(y,R))$ of the ball $B^g(y,R)$ viewed from x w.r.t. the metric g is the set of positive endpoints in $\partial\widetilde{M}$ of g-geodesic rays starting from x and intersecting $B^g(y,R)$.

Recall Lemma 3.17 from [38].

Lemma 3.4 ([38]). – *For all $r,\alpha>0$ and $T,T'>0$, and $v\in S^g\widetilde{M}$ such that $\mathcal{B}^g_{v^g_+}(\pi(v),o)=0$, if x_t denotes the footpoint of $g_t(v)$, we have*

$$B^g(v;T,T',r)\subset (H^g)^{-1}\left(\mathcal{O}^g_{x_{-T'}}(B^g(x_T,2r)\times\mathcal{O}^g_{x_T}(B^g(x_{-T'},2r))\times\left]-r,r\right[\right),\quad \textit{and}$$

$$(H^g)^{-1}\left(\mathcal{O}^g_{x_{-T'}}(B^g(x_T,r)\times\mathcal{O}^g_{x_T}(B^g(x_{-T'},r))\times\left]-\alpha,\alpha\right[\right)\subset B^g(v;T,T',2r+2\alpha).$$

When g_1 and g_2 are two admissible negatively curved metrics on M, recall that any g_1-geodesic between any two points is at distance at most $C_2(g_1,g_2)$ of the g_2-geodesic joining the same endpoints, and vice versa, for some constant $C_2(g_1,g_2)$ depending only on g_1 and g_2. This leads immediately to the following lemma.

Lemma 3.5. – *Let g_1 and g_2 be two admissible negatively curved metrics on M, and x,y two points on $\widetilde{M}$. Then*

$$\mathcal{O}^{g_1}_x(B^{g_1}(y,R))\subset\mathcal{O}^{g_2}_x(B^{g_2}(y,R+C_2(g_1,g_2)))\subset\mathcal{O}^{g_1}_x(B^{g_1}(y,R+2C_2(g_1,g_2))).$$

These lemmas will have the following very convenient corollary.

Corollary 3.6. – *Let g_1 and g_2 be two admissible negatively curved metrics on M. For all $\varepsilon>0$, there exists $C>0$ and $\varepsilon'=\varepsilon'(\varepsilon)>0$ such that for all $v\in S^{g_1}M$, we have*

$$B^{g_2}(\Psi^{g_1\to g_2}(v),S+C,S'+C,\varepsilon)\subset\Psi^{g_1\to g_2}(B^{g_1}(v,T,T',\varepsilon))\subset B^{g_2}(\Psi^{g_1\to g_2}(v),S,S',\varepsilon'),$$

where

$$\begin{aligned}S&=\mathcal{B}^{g_2}_{v^{g_1}_+}(\pi(v),\pi(g_1^Tv)),\\ S'&=\mathcal{B}^{g_2}_{v^{g_1}_+}(\pi(v),\pi(g_1^{-T'}v))\qquad\textit{and}\\ \varepsilon'&=\varepsilon(5+C_1(g_1,g_2))+C_2(g_1,g_2)+2C_3(g_1,g_2).\end{aligned}$$

Proof. – As the sets considered in the above statement are typically small, we can prove them on $T\widetilde{M}$ instead of TM. Without loss of generality, we can assume that $\pi(v)=o$. Indeed, all lemmas stated above are valid with o an arbitrary point, for example the basepoint of v. In particular, we have $\Psi^{g_1\to g_2}(v)=\Phi^{g_1\to g_2}(v)$.

We start with the right inclusion. Given $u\in B^{g_1}(v,T,T',\varepsilon)$, we want to control the distance $d^{g_2}(g_2^s\Psi^{g_1\to g_2}u,g_2^s\Psi^{g_1\to g_2}v)$. As $u\in B^{g_1}(v,T,T',\varepsilon)$, $\tau^{g_1,g_2}(u)\le\varepsilon(1+C_1(g_1,g_2))$, where τ^{g_1,g_2} was defined in Lemma 2.12, so that

$$d^{g_2}(\Phi^{g_1\to g_2}(u),\Psi^{g_1\to g_2}(u))\le\varepsilon(1+C_1(g_1,g_2)).$$

Therefore, $\Psi^{g_1\to g_2}B^{g_1}(v;T,T',\varepsilon)$ is included in the $\varepsilon(1+C_1(g_1,g_2))$-neighborhood of $\Phi^{g_1\to g_2}(B^{g_1}(v;T,T',\varepsilon)$.

Let $w = \Psi^{g_1\to g_2}(v) = \Phi^{g_1\to g_2}(v)$. Denote by w_s (resp. z_s) the basepoint $\pi(g_2^s w)$, for $s\in\mathbb{R}$, of $g_2^s w$. (resp. of $g_2^s z$). Let $S = \mathcal{B}^{g_2}_{v_+^{g_1}}(o,\pi(g_1^T v)$ and $S' = -\mathcal{B}^{g_2}_{v_+^{g_1}}(o,\pi(g_1^{-T'}v))$. By Lemma 2.6, we know that

$$|S - d^{g_2}(\pi(v),\pi(g_T^1 v))| \le C_3(g_1,g_2) \quad\text{and}\quad |S' - d^{g_2}(\pi(v),\pi(g_{-T'}^1 v))| \le C_3(g_1,g_2).$$

Moreover, the distances $d^{g_2}(w_S, v_T)$ and $d^{g_2}(w_{-S'}, v_{-T'})$ are uniformly bounded. Indeed, by Lemma 2.12, $\Psi^{g_1\to g_2}(g_1^T v) = g_2^S\Psi^{g_1\to g_2}(v)$ so that

$$d^{g_2}(w_S, v_T) = d^{g_2}(\pi(\Psi(g_1^T v), \pi(g_1^T v)) \le C_3(g_1,g_2).$$

Lemma 3.5 and elementary geometric considerations in negative curvature give the inclusion

$$\begin{aligned}\mathcal{O}^{g_1}_{v_{-T'}}(B^{g_1}(v_T,2\varepsilon)) &\times \mathcal{O}^{g_1}_{v_T}(B^{g_1}(x_{-T'},2\varepsilon)) \times]-\varepsilon,\varepsilon[\\ &\subset \mathcal{O}^{g_2}_{w_{-S'}}(B^{g_2}(w_S, 2\varepsilon + C_2(g_1,g_2) + 2C_3(g_1,g_2)))\\ &\quad\times \mathcal{O}^{g_2}_{y_S}(B^{g_2}(y_{-S'}, 2\varepsilon + C_2(g_1,g_2) + 2C_3(g_1,g_2))) \times]-\varepsilon,\varepsilon[.\end{aligned}$$

Lemma 3.4 implies the right inclusion

$$\Phi^{g_1\to g_2}B^{g_1}(v;T,T',\varepsilon) \subset B^{g_2}(w;S,S',4\varepsilon + C_2(g_1,g_2) + 2C_3(g_1,g_2)).$$

The relation between $\Phi^{g_1\to g_2}$ and $\Psi^{g_1\to g_2}$ gives

$$\Psi^{g_1\to g_2}B^{g_1}(v;T,T',\varepsilon) \subset B^{g_2}(\Psi^{g_1\to g_2}(v),S,S',\varepsilon(5 + C_1(g_1,g_2)) + C_2(g_1,g_2) + 2C_3(g_1,g_2)).$$

We proceed in the same way for the left inclusion, but we need in addition the help of Lemma 3.3.

Reasoning similarly as above gives the inclusion

$$B^{g_2}(w;S,S',\varepsilon) \subset \Psi^{g_1\to g_2}\left(B^{g_1}(v;T,T',(4\varepsilon + C_2(g_1,g_2) + 2C_3(g_1,g_2))(1 + C_1(g_1,g_2))\right).$$

As T, T', ε are arbitrary, using Lemma 3.3, we obtain easily the existence of a constant $C>0$ such that

$$B^{g_2}(w;S+C,S'+C,\varepsilon) \subset \Psi^{g_1\to g_2}\left(B^{g_1}(v;T,T',\varepsilon)\right). \qquad \square$$

3.2. Kolmogorov-Sinai, Brin-Katok and topological entropies

The Kolmogorov-Sinai entropy, or measure-theoretical entropy, of a dynamical system T w.r.t an invariant probability measure μ is the supremum over all measurable partitions of the exponential growth rate of the complexity of a partition, when iterated by T, and measured by μ. By Shannon-McMillan-Breiman Theorem, it also equals (the supremum over all partitions of) the exponential decay rate of a typical atom of the iterated partition.

Instead of iterating a measurable partition, when X is a metric space, endowed with the Borel σ-algebra, one can consider exponential decay rate of the measure of typical dynamical balls, which will give us a notion of local entropy, introduced by [10].

When T is a continuous map on a *compact* space X, Brin-Katok [10] showed that for ergodic measures this Kolmogorov-Sinai entropy coincides with the exponential decay of dynamical balls, also called the local entropy. This equality also holds when T is a lipschitz map of a noncompact manifold, as has been verified in [42, Thm. 1.32].

We shall not define the classical Kolmogorov-Sinai entropy, denoted by $h_{KS}(T,m)$, because we do not really use it in this work. But we recall below some definitions of local entropy and the statements of Brin-Katok and Riquelme.

For $(\varphi^t): X \to X$ a dynamical system and m a finite invariant measure, define the lower local entropy

$$\underline{h}_{\mathrm{loc}}(T,m)) = \operatorname*{ess\,inf}_{x\in X} \lim_{\varepsilon\to 0} \liminf_{T\to\infty} -\frac{1}{T}\log m(B_{\mathrm{dyn}}(x,T,\varepsilon)), \tag{8}$$

and the upper local entropy relative to compact sets

$$\overline{h}_{\mathrm{loc}}^{\mathrm{comp}}(T,m)) = \sup_{K} \operatorname*{ess\,sup}_{x\in K} \lim_{\varepsilon\to 0} \limsup_{T\to\infty,\varphi^T x\in K} -\frac{1}{T}\log m(B_{\mathrm{dyn}}(x,T,\varepsilon)). \tag{9}$$

For the geodesic flow in negative curvature, dynamical balls should be defined relatively to a distance on S^gM, but, as mentioned in the above subsection, the "natural" Sasaki distance on S^gM is equivalent to the distance $d(v,w) = \sup_{-1\le t\le 0} d^g(\pi(g_t v), \pi(g_t w))$, so that, when studying asymptotic quantities as entropy, we can use the distance d^g on M instead of the Sasaki distance on S^gM.

The following result is essentially due to Brin-Katok and Riquelme.

Theorem 3.7 ([10],[42],[43]). – *Let (M,g) be a Riemannian manifold with pinched negative curvature. Let m be an invariant ergodic measure under the geodesic flow on S^gM.*

$$h_{KS}(m,g) = \underline{h}_{\mathrm{loc}}(m,g) = \overline{h}_{\mathrm{loc}}^{\mathrm{comp}}(m,g). \tag{10}$$

Proof. – This result is due to Brin-Katok in the compact case. Their proof of the inequality $h_{KS}(m,g) \le \underline{h}_{\mathrm{loc}}(m,g)$ extends verbatim to the noncompact case. In [42, Thm. 1.32], Riquelme proved the equality $h_{KS}(m,g) = \underline{h}_{\mathrm{loc}}(m,g)$ for any Lipschitz dynamical system. In [42, Th. 1.41], he established the inequality $\underline{h}_{\mathrm{loc}}(m,g) \le \overline{h}_{\mathrm{loc}}^{\mathrm{comp}}(m,g)$, and the inequality $\overline{h}_{\mathrm{loc}}^{\mathrm{comp}}(m,g) \le h_{KS}(m,g)$ is established in the proof of [42, Thm. 1.42]. □

As observed in Remark 3.1 there are two notions of dynamical balls and the small ones are more relevant for us. Therefore, we define what we will call the local entropy, denoted by $h^{\Gamma}_{\mathrm{loc}}(m,g)$ in the sequel, as follows.

$$h^{\Gamma}_{\mathrm{loc}}(m,g) = \sup_{K\subset S^gM} \operatorname*{sup\,ess}_{v\in K} \lim_{\varepsilon\to 0} \limsup_{T\to\infty, g^T v\in K} -\frac{1}{T}\log m(B^g(v,T,\varepsilon)). \tag{11}$$

It follows from Theorem 3.7 and inclusion (7) that

$$h_{KS}(m,g) = \overline{h}_{\mathrm{loc}}^{\mathrm{comp}}(m,g) \le h^{\Gamma}_{\mathrm{loc}}(m,g), \tag{12}$$

with equality as soon as M has an injectivity radius bounded from below or m has compact support. We learned recently that Riquelme [44] proved that the last inequality above is also an equality

$$h_{KS}(m,g) = \overline{h}_{\mathrm{loc}}^{\mathrm{comp}}(m,g) = h^{\Gamma}_{\mathrm{loc}}(m,g)$$

in the case of the geodesic flow of a pinched negatively curved manifold.

REMARK 3.8. – Let us emphasize that all definitions of entropies above are sensitive to the scaling of the metric but not sensitive to the scaling of the measure. In particular, if m is finite but not a probability measure, then

$$h^{\Gamma}_{\mathrm{loc}}(m,g) = h^{\Gamma}_{\mathrm{loc}}(\lambda m,g) \quad \text{and} \quad h^{\Gamma}_{\mathrm{loc}}(m,\lambda g) = \frac{h^{\Gamma}_{\mathrm{loc}}(m,g)}{\sqrt{\lambda}}.$$

REMARK 3.9. – Observe that contrarily to Kolmogorov-Sinai entropy, the above definitions of local entropy make perfectly sense for an infinite invariant ergodic and conservative Radon measure. In particular, the Bowen-Margulis measure (see Section 3.4) which, when finite, is the measure of maximal entropy of the geodesic flow, always has a local entropy with respect to small dynamical balls and return times into compact sets which coincides with the topological entropy of the geodesic flow, see Proposition 3.17.

Lemma 3.3 allows us to choose some $\varepsilon > 0$ without need to take the limit when $\varepsilon \to 0$. Moreover, the invariance of the measure allows to consider shifted dynamical balls. It is the result below.

LEMMA 3.10. – *Let (M,g) be a manifold with pinched negative curvature, and μ a geodesic current. Let m^g_μ be the g-invariant measure associated to μ on $S^g M$. One can compute its local entropy as*

$$h^{\Gamma}_{\mathrm{loc}}(m^g_\mu,g) = \sup_K \operatorname*{sup\,ess}_{v\in K} \limsup_{T+T'\to\infty, g^T v\in K, g^{-T'}v\in K} -\frac{1}{T+T'}\log m^g_\mu(B^g(v;T,T',\varepsilon)).$$

Geometers usually are more interested in topological entropy than measure-theoretic entropy. We shall not define topological entropy topologically, but through the variational principle. Denote by $\mathcal{M}^1(g)$ the set of invariant probability measures for the metric g.

The topological entropy of the geodesic flow (g^t), denoted by $h_{\mathrm{top}}(g)$, satisfies

$$h_{\mathrm{top}}(g) = \sup_{m\in\mathcal{M}^1(g)} h_{KS}(m,g). \tag{13}$$

This variational principle is due first to [16, 24, 35] and later Handel-Kitchen [27] on noncompact spaces. It follows from [36] that this supremum is achieved iff the so-called Bowen-Margulis measure is finite (see later Subsection 3.4 for details). In this case, it is the unique measure maximizing entropy.

3.3. Entropy transformation law

Our goal is to prove the following result.

THEOREM 3.11. – *Let (M,g_i), $i=1,2$ be two admissible Riemannian metrics with pinched negative curvature on M. Let μ be a geodesic current and $m^{g_i}_\mu$ the associated invariant measures on $S^{g_i}M$ under the geodesic flow (g^t_i). Assume that these measures are finite and ergodic. Then their local entropies are related as follows:*

$$h^{\Gamma}_{\mathrm{loc}}(m^{g_2}_\mu,g_2) = I_\mu(g_2,g_1)\times h^{\Gamma}_{\mathrm{loc}}(m^{g_1}_\mu,g_1).$$

Thanks to Corollary 2.15, the corollary below follows.

COROLLARY 3.12. – *Under the same assumptions, we have*

$$h^{\Gamma}_{\mathrm{loc}}(m^{g_2}_{\mu}, g_2) \times \|m^{g_2}_{\mu}\| = h^{\Gamma}_{\mathrm{loc}}(m^{g_1}_{\mu}, g_1) \times \|m^{g_1}_{\mu}\|.$$

As mentioned above, in the case of geodesic flows in pinched negative curvature, Riquelme [44] proved that for ergodic probability measures, Kolmogorov Sinai entropy and local entropy coincide. We deduce the following corollary.

COROLLARY 3.13. – *Under the same assumptions, we have*

$$h_{KS}\left(\frac{m^{g_2}_{\mu}}{\|m^{g_2}_{\mu}\|}, g_2\right) \times \|m^{g_2}_{\mu}\| = h_{KS}\left(\frac{m^{g_1}_{\mu}}{\|m^{g_1}_{\mu}\|}, g_1\right) \times \|m^{g_1}_{\mu}\|.$$

Let us prove Theorem 3.11.

Proof. – It follows from Lemma 3.10 that for $i = 1, 2$, the entropy may be computed as

$$h^{\Gamma}_{\mathrm{loc}}(m^{g_i}_{\mu}, g_i) = \sup_K \operatorname*{sup\,ess}_{v \in K} \limsup_{T+T' \to \infty, g_i^T v \in K, g_i^{-T'} v \in K} -\frac{1}{T+T'} \log m^{g_i}_{\mu}(B^{g_i}(v; -T, T', \varepsilon))$$

for some fixed $\varepsilon > 0$, the essential supremum being relative to $m^{g_i}_{\mu}$. The above limsup is constant along (g_i)-orbits, so that by ergodicity, it is $m^{g_i}_{\mu}$-almost surely constant. Observe also that when K grows, the quantity on the right also grows.

Choose some compact set $K \subset TM$ large enough to contain an open subset of $\Omega^{g_i} \cap S^{g_i}M$ for $i = 1, 2$, and to have positive $m^{g_i}_{\mu}$-measure. Choose it large enough so that it allows to estimate entropies $h^{\Gamma}_{\mathrm{loc}}(m^{g_i}_{\mu}, g_i)$, up to some small arbitrary α. In other words,

$$\left| h^{\Gamma}_{\mathrm{loc}}(m^{g_i}_{\mu}, g_i) - \operatorname*{sup\,ess}_{v \in K \cap S^{g_i}M} \limsup_{T+T' \to \infty, g_i^T v \in K, g_i^{-T'} v \in K} -\frac{1}{T+T'} \log m^{g_i}_{\mu}(B^{g_i}(v; -T, T', \varepsilon)) \right| \leq \alpha.$$

Choose a typical $v \in S^{g_1}M \cap K$, which realizes the above essential supremum on K, and the almost sure conclusion of Corollary 2.8 when $T \to \pm\infty$. With the notations of Corollary 3.6, let $w = \Psi^{g_1 \to g_2}(v)$. As observed in the preceding section, we have $m^{g_2}_{\mu} = \mathscr{E}^{g_1 \to g_2} \times \Psi^{g_1 \to g_2}_* m^{g_1}_{\mu}$. But $\mathscr{E}^{g_1 \to g_2}$ is uniformly close to 1 on $B^{g_1}(v, \varepsilon)$.

Thus, up to some constants $e^{\pm c(v,\varepsilon)}$, by Corollary 3.6, we have

$$\begin{aligned} e^{-c(v,\varepsilon)} m^{g_2}_{\mu}(B^{g_2}(w; S+C, S'+C, \varepsilon')) &\leq m^{g_1}_{\mu}\left(B^{g_1}(v; T, T', \varepsilon)\right) \\ &\leq e^{c(v,\varepsilon)} m^{g_2}_{\mu}\left(B^{g_2}(w; S, S', \varepsilon')\right) \end{aligned}$$

with

$$w = \Psi^{g_1 \to g_2}(v), \qquad S = d^{g_2}(\pi(w), \pi(g_1^T w)) \pm C_3(g_1, g_2) \qquad \text{and}$$
$$S' = d^{g_2}(\pi(v), \pi(g_1^{-T'} w)) \pm C_3(g_1, g_2).$$

Observe also that the condition $g_1^T v \in K$ (resp. $g_1^{-T'} v \in K$) implies that $g_2^S w$ (resp. $g_2^{-S'} w$) belongs to the $C_3(g_1, g_2)$-neighborhood of K for any of the two metrics g_1 or g_2. It remains true for $g_2^{S+C} w$ and $g_2^{-S'-C} w$ inside the $C_3(g_1, g_2) + C$-neighborhood of K for the metric g_2. Set $K' = \mathscr{V}_{C_3(g_1,g_2)+C}(K) \supset K$.

By definition of T, T', S, S', we also have $T + T' \to +\infty$ iff $S + S' \to \infty$.

Therefore, taking the limsup of $\frac{1}{S+S'}$ log of the above quantity, we get

$$\limsup_{S+S'\to\infty, g_2^S w\in K', g_2^{-S'} w\in K'} -\frac{1}{S+S'}\log m_\mu^{g_2}(B^{g_2}(w,S,S',\varepsilon'))$$
$$= \limsup_{T+T'\to\infty, g_1^T v\in K, g_1^{-T'} v\in K} \frac{T+T'}{S+S'}\times\frac{-1}{T+T'}\log m_\mu^{g_1}(B^{g_1}(v,T,T',\varepsilon)).$$

By Corollary 2.8 we know that

$$\frac{T+T'}{S+S'} = \frac{T+T'}{\mathcal{B}^{g_2}_{v_+^{g_1}}(\pi(g_1^{-T'}v),\pi(g_1^T v))} = \frac{T+T'}{\int_{-T'}^{T}\mathcal{E}^{g_1\to g_2}(g_1^t v)\,dt}$$

converges when $T+T'\to+\infty$ to

$$\frac{1}{\int_{S^{g_1}M}\mathcal{E}^{g_1\to g_2}\frac{dm_\mu^{g_1}}{\|m_\mu^{g_1}\|}} = \int_{S^{g_2}M}\mathcal{E}^{g_2\to g_1}\frac{dm_\mu^{g_2}}{\|m_\mu^{g_2}\|}.$$

We deduce easily, by taking the supremum in K, that

$$h^\Gamma_{\mathrm{loc}}(m_\mu^{g_2},g_2) = \int_{S^{g_2}M}\mathcal{E}^{g_2\to g_1}\frac{dm_\mu^{g_2}}{\left\|m_\mu^{g_2}\right\|}\,h^\Gamma_{\mathrm{loc}}(m_\mu^{g_1},g_1) = I_\mu(g_2,g_1)\times h^\Gamma_{\mathrm{loc}}(m_\mu^{g_1},g_1). \quad \square$$

3.4. Bowen-Margulis measures and comparison of topological entropies

We define now the so-called *Bowen-Margulis measure*, and use it to deduce from Theorem 3.11 a corollary about the comparison of topological entropies of two metrics g_1 and g_2. The construction below is due to Patterson [37] for compact surfaces, to Sullivan [49, 51] for geometrically finite hyperbolic manifolds, and Yue [56] extended Sullivan's work in variable negative curvature.

Let (M,g) be a negatively curved manifold, with pinched negative curvature. Choose some point $o\in\widetilde{M}$. Consider the Poincaré series

$$P^g_\Gamma(s) = \sum_{\gamma\in\Gamma} e^{-sd^g(o,\gamma o)}.$$

Let $\delta(g)$ be its critical exponent. This exponent is finite, and when Γ is nonelementary, it is positive. The pair (Γ,g) is said to be *divergent* when the above series diverges when $s=\delta(g)$.

The following lemma is immediate from the definition of δ.

LEMMA 3.14. – *Let $(g_\varepsilon)_{-1\le\varepsilon\le1}$ be a family of negatively curved metrics on $M=\widetilde{M}/\Gamma$, such that $e^{-\varepsilon}g_0\le g_\varepsilon\le e^\varepsilon g_0$. Then $e^{-\varepsilon/2}\delta(g_0)\le\delta(g_\varepsilon)\le e^{\varepsilon/2}\delta(g_0)$.*

We need to ensure that the above series diverges at $s=\delta(g)$, which could be false. We will modify $P^g_\Gamma(s)$ into $\widetilde{P}^g_\Gamma(s)$ as follows. The Patterson trick [37] is the following. Define a continuous map $h:(0,+\infty)\to(0,+\infty)$ as the exponential of continuous piecewise affine maps with slope ε_k on the interval I_k, with $\varepsilon_k\to0$ and I_k a sequence of adjacent intervals of increasing length. It is possible to do it in such a way that h is positive, increasing, continuous, with slow growth, and $\frac{h(t_0+t)}{h(t_0)}$ is bounded by $\exp(\varepsilon_k t)$. Moreover, ε_k and I_k can be chosen in order to ensure that

$$\widetilde{P}^g_\Gamma(s) = \sum_{\gamma\in\Gamma} h(d^g(o,\gamma o))e^{-sd^g(o,\gamma o)}$$

has exponent $\delta(g)$ but now diverges at $s = \delta(g)$.

Define for all $x \in \widetilde{M}$ and $s > \delta(g)$ a probability measure

$$\nu_x^s = \frac{1}{\widetilde{P}_\Gamma^g(s)} \sum_{\gamma \in \Gamma} h(d^g(o, \gamma o)) e^{-s d^g(x, \gamma o)} \Delta_{\gamma o}$$

on $\overline{M} = \widetilde{M} \cup \partial \widetilde{M}$, where Δ_x denotes the Dirac mass at the point x. Choose a decreasing sequence $s_k \to \delta(g)$ such that $\nu_o^{s_k}$ converges to a probability measure ν_o^g on $\overline{M}$. Choose for all $x \in \widetilde{M}$ a subsequence s_{k_j} of s_k such that $\nu_x^{s_{k_j}}$ converges to a measure ν_x^g on $\overline{M}$. By construction, as $\widetilde{P}(\delta(g))$ diverges, all these measures are equivalent finite measures supported on $\Lambda_\Gamma \subset \partial \widetilde{M}$, the measure ν_o^g is a probability measure, and this family $(\nu_x^g)_{x \in \widetilde{M}}$ satisfies two crucial properties for all $x, y \in \widetilde{M}$, almost all $\xi \in \partial \widetilde{M}$ and all $\gamma \in \Gamma$:

$$\frac{d\nu_x^g}{d\nu_y^g}(\xi) = \exp\left(-\delta(g) \mathscr{B}_\xi(x, y)\right) \quad \text{and} \quad \gamma_* \nu_x^g = \nu_{\gamma x}^g. \tag{14}$$

A family of measures satisfying (14) is a Γ-*invariant* δ-*conformal density on the boundary*. From these properties follows the Sullivan's Shadow Lemma.

Proposition 3.15 ([49]). – *Let* (ν_x^g) *be a* Γ-*invariant* δ-*conformal density on* Λ_Γ. *Then for all* $R > 0$ *large enough, there exists a constant* $c = c(R) > 0$ *such that*

$$\frac{1}{c} \exp\left(-\delta(g) d^g(o, \gamma o)\right) \le \nu_o^g\left(\mathscr{O}_o(B(\gamma o, R))\right) \le c \exp\left(-\delta(g) d^g(o, \gamma o)\right).$$

A *Bowen-Margulis measure* on $S^g M$ is a measure obtained from such a family (ν_x^g) by the following formula on $S^g \widetilde{M}$, with $v = (H^g)^{-1}(v_-^g, v_+^g, t)$

$$d\widetilde{m}_{\mathrm{BM}}^g(v) = \exp\left(\delta(g) \mathscr{B}_{v_+^g}^g(o, \pi(v)) + \delta(g) \mathscr{B}_{v_-^g}^g(o, \pi(v))\right) d\nu_o^g(v_+^g) d\nu_o^g(v_-^g) dt. \tag{15}$$

This formula being Γ-invariant, it induces on the quotient a Bowen-Margulis measure m_{BM}^g on $S^g M$.

It is well known (see the above references, or Roblin [46] for the most general version) that P_Γ^g diverges at $s = \delta(g)$ iff the Bowen-Margulis measure is ergodic and conservative, and in this case, the family of measures (ν_x^g) is in fact unique. In particular, when this measure m_{BM}^g is finite, it is ergodic and conservative and P_Γ^g diverges at $\delta(g)$.

Otal-Peigné proved the following result, due to Sullivan in the case of geometrically finite hyperbolic manifolds.

Theorem 3.16 ([51], [36]). – *Let* (M, g) *be a manifold with pinched negative curvature and bounded derivatives of the curvature. Then*

$$\delta(g) = h_{\mathrm{top}}(g)$$

is the topological entropy of (g^t). *Moreover, when* m_{BM}^g *is finite and normalized into a probability measure, it is the unique measure maximizing entropy in the sense that* $h_{KS}(\frac{m_{\mathrm{BM}}^g}{\|m_{\mathrm{BM}}^g\|}, g) = h_{\mathrm{top}}(g)$. *When* m_{BM}^g *is infinite, there is no probability measure maximizing entropy.*

It follows from [38, Prop. 3.16] and [36] that, finite or not, the Bowen-Margulis measure satisfies the following equality.

Proposition 3.17. – *Let (M, g) be a negatively curved manifold with pinched negative curvature and bounded derivatives of the curvature. Let m^g_{BM} be a Bowen-Margulis measure. Then*

$$h^{\Gamma}_{\mathrm{loc}}(m^g_{\mathrm{BM}}, g) = \delta(g). \tag{16}$$

Moreover, when m^g_{BM} is finite, it satisfies

$$h^{\Gamma}_{\mathrm{loc}}(m^g_{\mathrm{BM}}, g) = \delta(g) = h_{KS}(\overline{m}^g_{\mathrm{BM}}, g). \tag{17}$$

Proof. – The first equality is a computation done in [38, Prop. 3.16], the second is one of the main results of [36]. □

This equality suggests that we could be able to prove a variational principle for infinite measures, using local entropies instead of Kolmogorov-Sinai entropies. We postpone this study to a further paper.

Corollary 3.18. – *Let (M, g_i), $i = 1, 2$ be two admissible Riemannian metrics on M whose curvature is negatively pinched and has bounded derivatives. Assume that $S^{g_2}M$ has a finite Bowen-Margulis measure $m^{g_2}_{\mathrm{BM}}$ and let $\mu^{g_2}_{\mathrm{BM}}$ be its geodesic current. Then*

$$\begin{aligned} h_{\mathrm{top}}(g_2) = \delta(g_2) = h^{\Gamma}_{\mathrm{loc}}(m^{g_2}_{\mathrm{BM}}, g_2) &= I_{\mu^{g_2}_{\mathrm{BM}}}(g_2, g_1) \times h^{\Gamma}_{\mathrm{loc}}(m^{g_1}_{\mu^{g_2}_{\mathrm{BM}}}, g_1) \\ &\leq I_{\mu^{g_2}_{\mathrm{BM}}}(g_2, g_1) \times h_{\mathrm{top}}(g_1). \end{aligned}$$

Proof. – Let us first note that by Theorem 4.2, the measure $m^{g_1}_{\mu^{g_2}_{\mathrm{BM}}}$ is a Gibbs measure. Moreover, [38, Thm. 1.3] ensures that the Gibbs measure associated to a given potential, when finite, is the unique equilibrium measure of this potential. Therefore $h^{\Gamma}_{\mathrm{loc}}(m^{g_1}_{\mu^{g_2}_{\mathrm{BM}}}) = h_{KS}(m^{g_1}_{\mu^{g_2}_{\mathrm{BM}}})$, and the variational principle ensures that $h_{KS}(m^{g_1}_{\mu^{g_2}_{\mathrm{BM}}}) \leq h_{\mathrm{top}}(g_1)$, which gives the last inequality. □

In the compact case, the inequality $h_{\mathrm{top}}(g_2) \leq I_{\mu^{g_2}_{\mathrm{BM}}}(g_2, g_1) \times h_{\mathrm{top}}(g_1)$ is due to Knieper [32]. Katok had a similar weaker inequality [29], proving that

$$h_{\mathrm{top}}(g_2) \leq \int_{S^{g_2}M} \|v\|^{g_1}\, dm^{g_2}_{\mathrm{BM}} \times h_{\mathrm{top}}(g_1).$$

Our inequality above is valid on any manifold, compact or not, with finite Bowen-Margulis measure. It follows from Lemma 2.5 that it implies Katok's inequality. Let us mention however that it is this weaker version which is really used in the proof of our main theorem of differentiability of entropy.

Let us now study the equality case in Corollary 3.18, as was done in Theorem 1.2 of [32].

Theorem 3.19. – *Let (M, g_i), $i = 1, 2$ be two admissible Riemannian metrics on M whose curvature is negatively pinched and has bounded derivatives. Assume that the geodesic flow on $S^{g_2}M$ has a finite Bowen-Margulis measure $m^{g_2}_{\mathrm{BM}}$ and*

$$h_{\mathrm{top}}(g_2) = I_{\mu^{g_2}_{\mathrm{BM}}}(g_2, g_1) \,.\, h_{\mathrm{top}}(g_1).$$

Then we have the following facts.

1. *The geodesic flow on $S^{g_1}M$ also has a finite Bowen-Margulis measure $m^{g_1}_{\mathrm{BM}}$.*

2. *The geodesic currents* $\mu_{\mathrm{BM}}^{g_1}$ *and* $\mu_{\mathrm{BM}}^{g_2}$ *associated to the Bowen-Margulis measures of* g_1 *and* g_2 *coincide up to normalization. In particular the Patterson-Sullivan densities of* g_1 *and* g_2 *are equivalent.*
3. *For all* $v \in S^{g_1}\widetilde{M}$ *with* $v_+^{g_1} \in \Lambda_\Gamma$, *there exists a unique real number* $\tau(v)$ *such that*
$$\frac{d\nu_{\pi v}^{g_1}}{d\nu^{g_2}_{\pi g_2^{\tau(v)} \circ \Psi^{g_1\to g_2}(v)}}(v_+^{g_1}) = 1.$$
Moreover, the map $F^{g_1\to g_2}$ *defined on* $S^{g_1}\widetilde{M}$ *by*
$$F^{g_1\to g_2}(v) = g_2^{\tau(v)} \circ \Psi^{g_1\to g_2}(v)$$
is Γ*-invariant and induces a Hölder-continuous Morse correspondance between the non-wandering sets* $F^{g_1\to g_2} : \Omega^{g_1} \to \Omega^{g_2}$.
4. *The map* $F^{g_1\to g_2}$ *conjugates the flows on the non-wandering sets of* $S^{g_1}M$ *and* $S^{g_2}M$ *up to a global time scaling by* $\frac{h_{\mathrm{top}}(g_1)}{h_{\mathrm{top}}(g_2)}$*: for all* $v \in \Omega^{g_1}$ *and all* $t \in \mathbb{R}$,
$$g_2^{h_{\mathrm{top}}(g_2)t} \circ F^{g_1\to g_2}(v) = F^{g_1\to g_2} \circ g_1^{h_{\mathrm{top}}(g_1)t}(v).$$

Proof. – Assume that the geodesic flow on $S^{g_2}M$ has finite Bowen-Margulis measure $m_{\mathrm{BM}}^{g_2}$, with geodesic current $\mu_{\mathrm{BM}}^{g_2}$.

Since $m_{\mathrm{BM}}^{g_2}$ is a Gibbs measure with maximal entropy, we have $h_{\mathrm{top}}(g_2) = h_{KS}(m_{\mathrm{BM}}^{g_2}) = h_{\mathrm{loc}}^{\Gamma}(m_{\mathrm{BM}}^{g_2})$. Since by Theorem 4.2 the measure $m^{g_1}_{\mu_{\mathrm{BM}}^{g_2}}$ is also a Gibbs measure, we have
$$h_{KS}\left(m^{g_1}_{\mu_{\mathrm{BM}}^{g_2}}\right) = h_{\mathrm{loc}}^{\Gamma}\left(m^{g_1}_{\mu_{\mathrm{BM}}^{g_2}}\right).$$

Therefore, if $h_{\mathrm{top}}(g_2) = I_{\mu_{\mathrm{BM}}^{g_2}}(g_2, g_1)\,.\,h_{\mathrm{top}}(g_1)$, it follows from Theorem 3.11 that
$$h_{KS}\left(m^{g_1}_{\mu_{\mathrm{BM}}^{g_2}}\right) = h_{\mathrm{top}}(g_1).$$
Therefore, Theorem 3.16 by Otal and Peigné implies that $m^{g_1}_{\mu_{\mathrm{BM}}^{g_2}}$ (which is a finite measure by Corollary 2.15) is, up to normalization, the unique Bowen-Margulis probability measure $m_{\mathrm{BM}}^{g_1}$ of g_1. This shows Item 1.

It implies that there exists $\lambda > 0$ such that
$$\mu_{\mathrm{BM}}^{g_1} = \lambda \mu_{\mathrm{BM}}^{g_2}.$$
By definition of these currents, see (15), it follows immediately that the Patterson-Sullivan measure $\nu_o^{g_1}$ is absolutely continuous with respect to $\nu_o^{g_2}$. This shows Item 2. Moreover, (15) also furnishes an explicit expression of the Radon-Nikodym derivative of $\nu_x^{g_1}$ w.r.t. $\nu_y^{g_2}$, for any two points $x, y \in \widetilde{M}$, which is therefore not only defined almost surely but is a positive Hölder continuous function defined everywhere on Λ_Γ.

The rest of the proof is inspired from [32, Prop. 3.8 p. 52], with the adaptations needed due to the non-compactness of M.

Let $v \in S^{g_1}\widetilde{M}$. Since the Patterson-Sullivan measure for g_2 is $\delta(g_2)$-conformal (see (14)), for all $v \in S^{g_1}M$ and all $t \in \mathbb{R}$, we have

$$\frac{d\nu^{g_1}_{\pi v}}{d\nu^{g_2}_{\pi g_2^t \circ \Psi^{g_1\to g_2}(v)}}(v_+^{g_1}) = e^{-\delta(g_2)t}\frac{d\nu^{g_1}_{\pi v}}{d\nu^{g_2}_{\pi \Psi^{g_1\to g_2}(v)}}(v_+^{g_1}).$$

In particular, for all $v \in S^{g_1}\widetilde{M}$ such that $v_+ = v_+^{g_1} \in \Lambda_\Gamma$ there exists a unique $\tau(v)$ such that

$$\frac{d\nu^{g_1}_{\pi v}}{d\nu^{g_2}_{\pi g_2^{\tau(v)} \circ \Psi^{g_1\to g_2}(v)}}(v_+^{g_1}) = 1.$$

It follows from the Γ-invariance of ν^{g_1}, ν^{g_2} and $\Psi^{g_1\to g_2}$ that the map $v \mapsto \tau(v)$ is also Γ-invariant. Therefore the map

$$F^{g_1\to g_2}(v) = g_2^{\tau(v)} \circ \Psi^{g_1\to g_2}(v)$$

is well defined for all $v \in S^{g_1}\widetilde{M}$ with $v_+^{g_1} \in \Lambda_\Gamma$, is Hölder continuous and is Γ-invariant. Since any non-wandering vector $v \in \Omega^{g_1}$ is the image of a vector $v \in S^{g_1}\widetilde{M}$ with $v_+^{g_1} \in \Lambda_\Gamma$ by the universal covering map, this shows Item 3.

The end of the proof follows [32, p. 53]. For all $t \in \mathbb{R}$ and all $v \in S^{g_1}\widetilde{M}$ with $\xi = v_+^{g_1}$,

$$e^{\delta(g_1)t} = \frac{d\nu^{g_1}_{\pi g_1^t v}}{d\nu^{g_2}_{\pi v}}(\xi) = \frac{d\nu^{g_1}_{\pi F^{g_1\to g_2}(g_1^t v)}}{d\nu^{g_2}_{\pi F^{g_1\to g_2}(v)}}(\xi) = e^{\delta(g_2)s},$$

where $s \in \mathbb{R}$ is such that

$$F^{g_1\to g_2}(g_1^t v) = g_2^s \circ F^{g_1\to g_2}(v).$$

Therefore $s = \frac{\delta(g_1)}{\delta(g_2)}t$ and we get for all $t \in \mathbb{R}$ and all $v \in S^{g_1}\widetilde{M}$ with $v_+^{g_1} \in \Lambda_\Gamma$

$$g_2^{\delta(g_2)t} \circ F^{g_1\to g_2}(v) = F^{g_1\to g_2}(g_1^{\delta(g_1)t}v).$$

This concludes the proof of Theorem 3.19. □

Corollary 1.6 is an immediate consequence of the above few lines.

4. Gibbs measures

This section, particularly Theorem 4.2, is crucial in the proof of Corollary 3.18, and therefore in our approach of Theorem 1.9.

Theorem 4.2 is new on noncompact manifolds, the explicit change of potential being new even on compact manifolds. Corollary 4.4 is new even on compact manifolds.

Gibbs measures are, for a hyperbolic dynamical system, a family of measures with strong stochastic properties, each one associated to a weight, i.e., a Hölder continuous potential, describing somehow that all possible dynamical behaviors typically happen w.r.t one of these measures. For the geodesic flow on the unit tangent bundle of a compact manifold, their geometric construction, adapted from the Patterson-Sullivan construction described in the above section, has been done by Ledrappier in [33]. He proved there, on compact manifolds, that being a Gibbs measure does not depend on the metric. In other words, if g_1 and g_2 are negatively curved metrics on M, an invariant measure $m_\mu^{g_1}$ on $S^{g_1}M$ is a Gibbs measure iff the measure $m_\mu^{g_2}$ on $S^{g_2}M$ is also a Gibbs measure. However, his proof strongly relies on the compactness of M. Our goal in this section is to prove this result differently on noncompact manifolds.

4.1. Definitions

We refer to [38] for details on all notions presented here. Let (M, g) be a negatively curved manifold, with pinched negative curvatures and bounded derivatives of the curvature. Let $F : S^g M \to \mathbb{R}$ be a Hölder continuous map. The *pressure of* F is the quantity

$$P^g(F) = \sup_{m \in \mathscr{M}^1(g)} \left(h_{KS}(m, g) + \int_{S^g M} F \, dm \right), \tag{18}$$

the supremum being considered over all invariant probability measures $m \in \mathscr{M}^1(g)$ such that $\int \max(-F, 0)\, dm < \infty$. An invariant probability measure m is *an equilibrium state for* F if it realizes the above supremum.

Assume that $P^g(F)$ is finite. An invariant measure m under the geodesic flow (g^t) satisfies the *Gibbs property* for the potential F if for all compact sets $K \subset S^g M$ and $\varepsilon > 0$ there exists a constant $C(K, \varepsilon) > 0$ such that for all $v \in K$ and $T > 0$ with $g^T v \in K$, we have

$$\frac{1}{C(K,\varepsilon)} \exp\left(\int_0^T F(g^t v)\, dt - T P^g(F) \right) \le m\left(B^g(v, T, \varepsilon)\right) \tag{19}$$

$$\le C(K,\varepsilon) \exp\left(\int_0^T F(g^t v)\, dt - T P^g(F) \right). \tag{20}$$

A variant of the Patterson-Sullivan construction presented in Subsection 3.4 provides a measure m_F which satisfies (19) see [38, Prop. 3.16]. Moreover, when finite and normalized into a probability measure, it is the unique equilibrium state, i.e., the unique measure realizing the supremum in (18) (see [38, Thm. 6.1]). When this measure m_F is infinite, there is no equilibrium state for F. Let us summarize what is useful for the present work in the following proposition.

PROPOSITION 4.1. – *Let (M, g) be a negatively curved manifold with pinched negative curvature and bounded derivatives of the curvature. Let $F : S^g M \to \mathbb{R}$ be a Hölder potential. If the measure m_F is finite and normalized, then*

$$P^g(m_F) = h_{KS}(m_F, g) + \int_{S^g M} F \, dm_F = h^{\Gamma}_{\mathrm{loc}}(m_F, g) + \int_{S^g M} F \, dm_F.$$

4.2. Being a Gibbs measure does not depend on the metric

THEOREM 4.2. – *Let (M, g_i) be two admissible metrics with pinched negative curvature and bounded first derivatives of the curvature on M. Let $F : S^{g_1} M \to \mathbb{R}$ be a Hölder map, and $m_F^{g_1}$ the associated Gibbs measure. We assume that $m_F^{g_1}$ is ergodic and conservative. Let $\mu_F^{g_1}$ be the associated current on $\partial^2 \widetilde{M}$. Let $m^{g_2}_{\mu_F^{g_1}}$ be the g_2-invariant measure associated to the same current.*

Then $m^{g_2}_{\mu_F^{g_1}}$ is also ergodic and conservative, and satisfies the Gibbs property (19) *for the Hölder potential*

$$G = (F - P^{g_1}(F)) \circ \Psi^{g_2 \to g_1} \times \mathcal{E}^{g_2 \to g_1}.$$

Moreover, $P^{g_2}(G) = 0$*. In other words, for all compact subsets* $K \subset S^{g_2}M$ *and* $\varepsilon > 0$ *there exists* $C > 0$ *such that for all* $w \in K$ *and* $S > 0$ *with* $g^S w \in K$*, we have*

$$\frac{1}{C}\, e^{\int_0^S G(g_2^s w)ds} \le m^{g_2}_{\mu_F^{g_1}}(B^{g_2}_\Gamma(w, S, \varepsilon)) \le C e^{\int_0^S G(g_2^s w)ds}.$$

If we assume moreover that the measure $m_F^{g_1}$ *is finite, and is therefore the equilibrium measure associated to* F*, then* $m^{g_2}_{\mu_F^{g_1}}/\|m^{g_2}_{\mu_F^{g_1}}\|$ *is the equilibrium measure associated to* G*.*

REMARK 4.3. – Reversing the role of g_1 and g_2, we observe that the same result holds with the potential $H = \big((F - P^{g_1}(F)) \times (\mathcal{E}^{g_1\to g_2})^{-1}\big) \circ (\Psi^{g_1\to g_2})^{-1}$. Therefore, they must be cohomologous.

Proof. – Conservativity and ergodicity depend only on the current at infinity and not on the (admissible) metric, as said in Proposition 2.3.

Gibbs property for the potential G follows from Corollary 3.6. Let us explain it more in details. We stated Theorem 4.2 in the most natural way, starting from g_1 and going to g_2, but in view of all the statements proved above that we shall use, we will reverse the role of g_1 and g_2, F and G, in the proof below. Assume that $m_G^{g_2}$ is a Gibbs measure w.r.t. the potential G on $S^{g_2}M$, let $\mu = \mu_G^{g_2}$ be its current at infinity, and let us prove that $m_\mu^{g_1}$ is a Gibbs measure w.r.t. the potential $F = (G - P^{g_2}(G)) \circ \Psi^{g_1\to g_2} \times \mathcal{E}^{g_1\to g_2}$.

First choose some compact set $K^{g_1} \subset S^{g_1}M$ and some $\varepsilon > 0$. Let $v \in K^{g_1}$ and $T > 0$ such that $g^T v \in K^{g_1}$. Define a compact set K^{g_2} as the C-neighborhood of $\Psi^{g_1\to g_2}(K^{g_1}) \cup (\Psi^{g_2\to g_1})^{-1}K^{g_1}$, where C is given by Corollary 3.6.

We will use Corollary 3.6 and first part of Proposition 2.13, and the fact that $m_\mu^{g_2} = \Psi_*^{g_1\to g_2}(\mathcal{E}^{g_1\to g_2} \times m_\mu^{g_1})$.

As $\mathcal{E}^{g_1\to g_2}$ is continuous, it is uniformly continuous on K^{g_1} so that for all $v \in K^{g_1}$ and $u \in B^{g_1}(v, \varepsilon)$, $\mathcal{E}^{g_1\to g_2}(u) = e^{\pm c(K^{g_1},\varepsilon)}\mathcal{E}^{g_1\to g_2}(v)$. We deduce that

$$\frac{e^{-c(K^{g_1},\varepsilon)}}{\mathcal{E}^{g_1\to g_2}(v)}\, m_\mu^{g_2}(\Psi^{g_1\to g_2}(B^{g_1}(v, T, \varepsilon))) \le m_\mu^{g_1}(B^{g_1}(v, T, \varepsilon))$$
$$\le \frac{e^{c(K^{g_1},\varepsilon)}}{\mathcal{E}^{g_1\to g_2}(v)}\, m_\mu^{g_2}(\Psi^{g_1\to g_2}(B_\Gamma^{g_1}(v, T, \varepsilon))).$$

Now, using Corollary 3.6, with $w = \Psi^{g_1\to g_2}v$, and $S = \mathcal{B}^{g_2}_{v_+^{g_1}}(\pi(v), \pi(g_1^T v))$, we get

$$\frac{e^{-c(K^{g_1},\varepsilon)}}{\mathcal{E}^{g_1\to g_2}(v)}\, m_\mu^{g_2}(B^{g_2}(w : S + C, C, \varepsilon)) \le m_\mu^{g_1}(B^{g_1}(v, T, \varepsilon)$$
$$\le \frac{e^{c(K^{g_1},\varepsilon)}}{\mathcal{E}^{g_1\to g_2}(v)} m_\mu^{g_2}(B^{g_2}(w, S, \varepsilon')).$$

As $m_\mu^{g_2}$ is a Gibbs measure, and w, $g_2^S w$, but also $g_2^{-C} w$ and $g_2^{S+C} w$ belong to K^{g_2}, there exists a constant $C(G, K^{g_2}, \varepsilon')$ coming from the Gibbs property, such that

$$\frac{e^{-c(K^{g_1},\varepsilon)}}{\mathcal{E}^{g_1\to g_2}(v)}\, \frac{e^{\int_{-C}^{S+C}(G - P^{g_2}(G))(g_2^s w)\,ds}}{C(G, K^{g_2}, \varepsilon')} \le m_\mu^{g_1}(B^{g_1}(v, T, \varepsilon))$$
$$\le \frac{e^{c(K^{g_1},\varepsilon)}}{\mathcal{E}^{g_1\to g_2}(v)} C(G, K^{g_2}, \varepsilon') e^{\int_0^S (G - P^{g_2}(G))(g_2^s w)\,ds}.$$

As G is (Hölder) continuous, it is bounded on K^{g_2}, so that the integral

$$\int_{-C}^{S+C} (G - P^{g_2}(G))(g_2^s w)\, ds$$

is, up to a constant c, uniformly close to $\int_0^S (G - P^{g_2}(G))(g_2^s w)\, ds$. The next ingredient is Proposition 2.13, which gives

$$\begin{aligned}\frac{e^{-c(K^{g_1},\varepsilon)}}{\mathcal{E}^{g_1\to g_2}(v)}\,\frac{e^{-c}}{C(G,K^{g_2},\varepsilon')}e^{\int_0^T F(g_1^t w)\,dt} &\le m_\mu^{g_1}(B^{g_1}(v,T,\varepsilon))\\ &\le \frac{e^{c(K^{g_1},\varepsilon)}}{\mathcal{E}^{g_1\to g_2}(v)}C(G,K^{g_2},\varepsilon')e^{\int_0^T F(g_1^t w)\,dt},\end{aligned}$$

with $F = (G - P^{g_2}(G))\circ\Psi^{g_1\to g_2}\times\mathcal{E}^{g_1\to g_2}$. It is exactly the Gibbs property for $m_\mu^{g_1}$ w.r.t. F.

It remains to show that $P^{g_1}(F) = 0$. To simplify notations, let us assume that $P^{g_2}(G) = 0$. Let ρ be any geodesic current on $\partial^2\widetilde{M}$. By definition,

$$P^{g_1}(F) = \sup_\rho \left(h_{KS}(m_\rho^{g_1}, g_1) + \int_{S^{g_1}M} F dm_\rho^{g_1} \right),$$

the supremum being taken over all currents ρ such that $m_\rho^{g_1}$ is an invariant probability measure. The change of mass and change of entropy (Corollary 2.15 and Theorem 3.11) give

$$P^{g_1}(F) = \sup_\rho I_\rho(g_2,g_1)\left(h(m_\rho^{g_2}/\|m_\rho^{g_2}\|, g_2) + \int G dm_\rho^{g_2}/\|m_\rho^{g_2}\| \right) \le 0.$$

The same computations with $\rho = \mu = \mu_G^{g_2}$ give $P^{g_1}(F) = 0$. □

4.3. Length spectrum and change of metrics

Let g_1 and g_2 be two quasi-isometric negatively curved metrics. There is a particular case where the above results have an easy but striking illustration.

Corollary 4.4. – *Let $(M, g_i)_{i=1,2}$ be two quasi-isometric complete negatively curved metrics on the same connected manifold M. Assume that the Bowen-Margulis measure of g_1 is ergodic and conservative, and let $\mu_{\mathrm{BM}}^{g_1}$ be the associated geodesic current. Then the measure $m^{g_2}_{\mu^{g_1}_{\mathrm{BM}}}$ is also ergodic and conservative. It is a Gibbs measure associated with the potential $G = -h_{\mathrm{top}}(g_1)\,\mathcal{E}^{g_2\to g_1}$.*

Moreover, for all primitive hyperbolic elements $\gamma\in\Gamma$, if w_γ is a periodic vector of $S^{g_2}M$ associated to γ, for all $\varepsilon > 0$ there exists $C\ge 1$ such that for all $T > 0$, we have

$$\frac{1}{C}e^{-h_{\mathrm{top}}(g_1)T\frac{\ell^{g_1}(\gamma)}{\ell^{g_2}(\gamma)}} \le m^{g_2}_{\mu^{g_1}_{\mathrm{BM}}}\left(B^{g_2}(w_\gamma,T,\varepsilon)\right) \le Ce^{-h_{\mathrm{top}}(g_1)T\frac{\ell^{g_1}(\gamma)}{\ell^{g_2}(\gamma)}}.$$

Proof. – It is an immediate application of Theorem 4.2 with $F = 0$. First write T as $T = n\ell^{g_2}(\gamma) + r$, with $0\le r < \ell^{g_2}(\gamma)$. The only thing to notice is that

$$\int_0^{\ell^{g_2}(\gamma)} \mathcal{E}^{g_2\to g_1}(g_2^s w_\gamma)\, ds = \ell^{g_2}(\gamma)\times e^{g_2\to g_1}(\gamma)$$

so that

$$-\int_0^T h_{\mathrm{top}}(g_1)\,\mathcal{E}^{g_2\to g_1}(g_2^s w_\gamma)\, ds = -h_{\mathrm{top}}(g_1)\times T\times\frac{\ell^{g_1}(\gamma)}{\ell^{g_2}(\gamma)} \pm \text{ constant},$$

the error term in the above inequality being smaller than $h_{\mathrm{top}}(g_1)\ell^{g_2}(\gamma)\|\mathcal{E}^{g_2\to g_1}\|_\infty$. $\square$

5. Convergence of geodesics, Busemann functions and invariant measures

In this section, we study the continuity of geodesics, Busemann functions, and Bowen-Margulis measures under a Lipschitz perturbation of the metric with uniform negative curvatures.

Let $(g_\varepsilon)_{-1\le\varepsilon\le 1}$ be a family of metrics on $\widetilde{M}$ with sectional curvatures satisfying $K_{g_\varepsilon}\le -a^2$, such that for all $\varepsilon>0$, at all $x\in\widetilde{M}$, $e^{-\varepsilon}g_0\le g_\varepsilon\le e^{\varepsilon}g_0$.

We first show that the g_ε-geodesic between two points at infinity converge uniformly in the Hausdorff topology of $\widetilde{M}$ to the g_0-geodesic with same extremities, and that the Busemann functions of g_ε converge uniformly on compact sets to the Busemann functions of g_0.

When the variation of metrics is continuous in $\mathcal{C}^1$-topology, this also implies that the Morse-correspondances $\Phi^{g_0\to g_\varepsilon}$ and $\Psi^{g_0\to g_\varepsilon}$ converge to the identity uniformly on compact sets in the $\mathcal{C}^0$-topology of $S^g\widetilde{M}$, and that the geodesic stretch $\mathcal{E}^{g_0\to g_\varepsilon}$ converges to 1.

Eventually, we show that under suitable assumptions, the Bowen-Margulis measures vary continuously in the weak-* topology.

5.1. Convergence of geodesics and Busemann functions

The following lemma is a classical and very useful consequence of the uniform upper bound on the curvature.

Lemma 5.1. – *Let $a>0$ and $(\widetilde{M},g)$ be a complete simply connected manifold with sectional curvatures satisfying $K_g\le -a^2$.*

1. *For all $C>0$, all $\xi\in\partial\widetilde{M}$, $x,y\in\widetilde{M}$ with $d^g(x,y)\le C$, and $t\ge C$, if $x_t=\gamma_{x,\xi}(t)$, we have*
$$\left|\mathcal{B}^g_\xi(x,y)-(d^g(x,x_t)-d^g(y,x_t))\right|\le 2Ce^{-at}.$$
2. *For all $T,K,\alpha>0$, for all $R\ge R_0=T-\frac{1}{a}\ln\frac{\alpha}{4Ke^K}$, if $(\gamma_1(t))_{t\in\mathbb{R}}$ and $(\gamma_2(t))_{t\in\mathbb{R}}$ are g-geodesics with*
$$d^g(\gamma_1(-R),\gamma_2(-R))\le K\quad\textit{and}\quad d^g(\gamma_1(R),\gamma_2(R))\le K,$$
then for all $t\in[-T,T]$,
$$d^g(\gamma_1(t),\gamma_2)\le\alpha.$$

Proof. – We will omit the subscript g in the proof. Let us first prove 1.

Assume $d(x,y)\le C$. We can also assume that $\mathcal{B}_\xi(x,y)\ge 0$. Denote by x' the unique point on $[x,\xi)$ such that $\mathcal{B}_\xi(x',y)=0$. By convexity of the horoball, $d(x,x')\le C$ and $d(x',y)\le C$. Let x_s (resp. y_s) be the points on $[x',\xi)$ (resp. $[y,\xi)$) at distance s of x' (resp. y). It follows from [28] that for all $s\ge C$,
$$d(x_s,y_s)\le d(x',y)e^{-as}\le Ce^{-as}.$$
Observe also that $\left|\mathcal{B}^g_\xi(x,y)-(d^g(x,x_s)-d^g(y,y_s))\right| = |\mathcal{B}_\xi(x_s,y_s)|\le d(x_s,y_s)$, so that $\left|\mathcal{B}^g_\xi(x,y)-(d^g(x,x_s)-d^g(y,x_s))\right|\le 2d(x_s,y_s)\le 2Ce^{-as}$.

To prove 2, denote by x_s the point of $[\gamma_1(-R), \gamma_1(R)]$ at distance s from $\gamma_1(-R)$, y_s the point of $[\gamma_1(-R), \gamma_2(R)]$ at distance s from $\gamma_1(-R)$ and distance say d_s from $\gamma_2(R)$ and z_s the point of $[\gamma_2(-R), \gamma_2(R)]$ at distance d_s from $\gamma_2(R)$. Observe immediately that $|d_s - 2R + s| \leq K$.

By the above, we have $d(x_s, y_s) \leq d(x_{2R}, y_{2R})e^{-as}$. But elementary considerations in the triangle $(x_{2R}, y_{2R}, \gamma_2(R))$ lead to

$$d(x_{2R}, y_{2R}) = d(\gamma_1(R), y_{2R}) \leq d(\gamma_1(R), \gamma_2(R)) + d(\gamma_2(R), y_{2R}) \leq 2K.$$

Thus $d(x_s, y_s) \leq 2Ke^{-as}$.

Similarly we get $d(y_s, z_s) \leq 2Ke^{-ad_s} \leq 2Ke^K e^{-a(2R-s)}$. We deduce that

$$d(x_s, \gamma_2) \leq d(x_s, z_s) \leq 2Ke^K(e^{-as} + e^{-a(2R-s)}).$$

Now, choose $R_0 = T - \frac{1}{a}\ln\frac{\alpha}{4Ke^K}$. For $t \in [-T, T]$, we have $\gamma_1(t) = x_{R+t}$ and $R+t \geq R_0 - T$ and $2R - (R+t) \geq R_0 - T$, so that

$$d(\gamma_1(t), \gamma_2) \leq d(\gamma_1(t), z_{R+t}) \leq 4Ke^K e^{-a(R_0-T)} \leq \alpha. \qquad \square$$

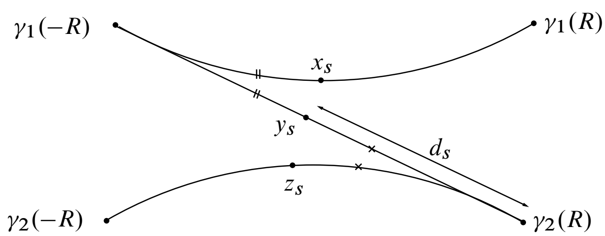


FIGURE 5. Proof of Lemma 5.1

Let us now show that the g_ε-geodesic segments converge to the g_0-geodesic segments in the Hausdorff topology of $\widetilde{M}$.

PROPOSITION 5.2. – *Let g_0 be a complete metric on $\widetilde{M}$ with $K_{g_0} \leq 0$. For all $0 < \varepsilon \leq 1$ small enough let g_ε be a complete metric on $\widetilde{M}$ such that at all $x \in \widetilde{M}$, $e^{-\varepsilon}g_0 \leq g_\varepsilon \leq e^\varepsilon g_0$.*

Then for all $x, y \in \widetilde{M}$, any minimizing g_ε-geodesic γ_ε joining x to y is contained in the D_ε-neighborhood of the g_0-geodesic $[x, y]_0$ from x to y, with $D_\varepsilon \leq \sqrt{\varepsilon}d^{g_0}(x, y)$.

Proof. – Let g_0 and g_ε as above, and $x, y \in \widetilde{M}$. Set $L_0 = d^{g_0}(x, y)$ and $L_\varepsilon = d^{g_\varepsilon}(x, y)$. Let $\gamma_0 : [0, L_0] \to \widetilde{M}$ and $\gamma_\varepsilon : [0, L_\varepsilon] \to \widetilde{M}$ be minimizing geodesics from x to y respectively for g_0 and g_ε, parametrized with unit speed. Note that γ_0 is unique. Let $l \in [0, L_\varepsilon]$ be such that

$$d^{g_0}(\gamma_\varepsilon(l), [x, y]_0) = \max_{t \in [0, L_\varepsilon]} d^{g_0}(\gamma_\varepsilon(t), [x, y]_0) = D_\varepsilon.$$

We call $z = \gamma_\varepsilon(l) \in \widetilde{M}$. Consider the g_0-geodesic triangle with vertices x, y, z. Set $l_1 = d^{g_0}(x, z)$ and $l_2 = d^{g_0}(z, y)$.

We have

$$l_1 \le \int_0^l \|\dot\gamma_\varepsilon(t)\|_{g_0}\,dt \quad \text{and} \quad l_2 \le \int_l^{L_\varepsilon} \|\dot\gamma_\varepsilon(t)\|_{g_0}\,dt. \tag{21}$$

Since $e^{-\varepsilon}g_0 \le g_\varepsilon \le e^{\varepsilon}g_0$, we have $\|\dot\gamma_\varepsilon(t)\|_{g_0} \le e^{\varepsilon/2}$ for all $t \in [0, L_\varepsilon]$ and $\|\dot\gamma_0(t)\|_{g_\varepsilon} \le e^{\varepsilon/2}$ for all $t \in [0, L_0]$. Therefore, by equation (21),

$$L_\varepsilon \le \int_0^L \|\dot\gamma_0(t)\|_{g_\varepsilon}\,dt \le e^{\varepsilon/2}L \quad \text{and} \quad l_1 + l_2 \le \int_0^{L_\varepsilon} \|\gamma_\varepsilon(t)\|_{g_0}\,dt \le e^{\varepsilon}L.$$

Since $K_{g_0} \le 0$, the distance d^{g_0} satisfies $CAT(0)$-triangle comparison property (cf [9] p161): D_ε is less than the height $\bar D$ from $\bar z$ of the comparison triangle $(\bar x, \bar y, \bar z)$ in the Euclidean plane with side lengths $d^{\mathrm{eucl}}(\bar x, \bar y) = L_0$, $d^{\mathrm{eucl}}(\bar x, \bar z) = l_1$ and $d^{\mathrm{eucl}}(\bar y, \bar z) = l_2$. Moreover, for all such Euclidean triangles with $l_1 + l_2 \le e^{\varepsilon}L_0$, the height $\bar D$ is maximal if and only if $l_1 = l_2 = \frac{e^{\varepsilon}L_0}{2}$. Therefore,

$$D_\varepsilon^2 \le \bar D^2 \le \frac{e^{2\varepsilon}L_0^2}{4} - \frac{L_0^2}{4} \le \varepsilon L_0^2$$

as soon as $e^{2\varepsilon} - 1 \le 4\varepsilon$. It proves $D_\varepsilon \le \sqrt{\varepsilon}\, d^{g_0}(x, y)$ and ends the proof of Proposition 5.2. □

Proposition 5.2 together with Morse-Klingenberg Lemma and Lemma 5.1 imply that when the curvatures have a uniform negative upper bound, the complete geodesics on $\widetilde M$ converge uniformly for the g_0-Hausdorff topology under a variation of the metric. Let $a > 0$ be fixed.

PROPOSITION 5.3. – *Let $(g_\varepsilon)_{-1<\varepsilon<1}$ be a family of metrics on $\widetilde M$ with sectional curvatures satisfying $K_{g_\varepsilon} \le -a^2$, such that for all $\varepsilon \in (-1, 1)$, $e^{-\varepsilon}g_0 \le g_\varepsilon \le e^{\varepsilon}g_0$ on each tangent space $T_x\widetilde M$, $x \in \widetilde M$. Then there exists $\alpha : (-1, 1) \to [0, +\infty)$, with $\lim_{\varepsilon\to 0}\alpha(\varepsilon) = 0$, such that for all $\varepsilon \in (-1, 1)$ and all $(\eta, \xi) \in \partial^2\widetilde M$, the g_ε-geodesic with extremities η and ξ is contained in the $\alpha(\varepsilon)$-neighborhood of the g_0-geodesic with extremities η and ξ.*

Proof. – First, recall (see Section 2.1) that the geodesics for g_0 and g_ε are at uniform bounded distance $C_2(g_0, g_\varepsilon) \le C_2(g_0, g_1)$. Let γ_0 be the g_0-geodesic from ξ to η. Choose its origin $\gamma_0(0)$ arbitrarily. For any large $\rho > 0$, we have $d(\gamma_0(\pm\rho), \gamma_\varepsilon) \le C_2(g_0, g_1)$. Consider the g_0-geodesic segment γ_1 joining the nearest point to $\gamma_0(\rho)$ on γ_ε with the nearest point to $\gamma_0(-\rho)$ on γ_ε. This geodesic segment has g_0-length equal to $2R = 2\rho \pm 2C_2(g_0, g_1)$. Choose its origin in such a way that $d^{g_0}(\gamma_0(\pm R), \gamma_1(\pm R)) \le 2C_2(g_0, g_1)$.

For all $\alpha > 0$, Lemma 5.1 applied with $K = 2C_2(g_0, g_1)$, $\alpha/2$ and $T = 1$ gives some $R_0 > 0$ such that when $R \ge R_0$, for all $t \in [-1, 1]$, $d^{g_0}(\gamma_1(t), \gamma_0(t)) \le \alpha/2$.

By Proposition 5.2, $d^{g_0}(\gamma_1(0), \gamma_\varepsilon)) \le 2R\sqrt{\varepsilon}$.

Therefore,

$$d^{g_0}(\gamma_0(0), \gamma_\varepsilon) \le d^{g_0}(\gamma_0(0), \gamma_1(0)) + d^{g_0}(\gamma_1(0), \gamma_\varepsilon)) \le \alpha/2 + 2R\sqrt{\varepsilon}.$$

Choose $R \ge R_0$ and $\varepsilon > 0$ such that $2R_0\sqrt{\varepsilon} \le \alpha/2$ to get $d^{g_0}(\gamma_0(0), \gamma_\varepsilon) \le \alpha$. As the origin on γ_0 is arbitrary, the result follows. □

Observe that, in the above proof, the dependence between α and ε can be made relatively explicit. For $K = C_2(g_0, g_1)$, $T = 1$ and $\alpha/2$ in Lemma 5.1 we get

$$R_0 = 1 + \frac{2C_2(g_0, g_1)}{a} \ln \frac{2C_2(g_0, g_1)}{\alpha} \text{ and } \varepsilon = \frac{\alpha^2}{16R_0^2}.$$

Moreover, our proof only uses $K_{g_0} \leq -a^2 < 0$ and the fact that for all g_ε, the g_ε-geodesic between two points at infinity is unique. The negative lowerbound on the sectional curvatures K_{g_ε} does not need to be uniform.

Proposition 5.4. – *Let $(g_\varepsilon)_{-1\leq\varepsilon\leq 1}$ be a family of complete metrics on $\widetilde{M}$ with $K_{g_\varepsilon} \leq -a^2$, such that for all $\varepsilon > 0$, and all $x \in \widetilde{M}$, $e^{-\varepsilon} g_0 \leq g_\varepsilon \leq e^{\varepsilon} g_0$ on the tangent space $T_x\widetilde{M}$.*

Then the map $\mathcal{B}^{g_\varepsilon} : (x, y, \xi) \mapsto \mathcal{B}_\xi^{g_\varepsilon}(x, y)$ converges to $\mathcal{B}^{g_0}$ as $\varepsilon \to 0$, uniformly on compact sets of $\widetilde{M} \times \widetilde{M} \times \partial\widetilde{M}$.

Proof. – Any compact set $K \subset \widetilde{M} \times \widetilde{M} \times \partial\widetilde{M}$ is contained in some (noncompact) set of the form $H_C = \{(x, y, \xi) \in \widetilde{M} \times \widetilde{M} \times \partial\widetilde{M}; d^{g_0}(x, y) \leq C\}$, for some $C > 0$. It is enough to show that $\mathcal{B}^{g_\varepsilon} \to \mathcal{B}^{g_0}$ as $\varepsilon \to 0$, uniformly on each H_C.

Let $C > 0$ be fixed. For all $\varepsilon \in (-1, 1)$ and all $(x, y, \xi) \in H_C$,

$$d^{g_\varepsilon}(x, y) \leq 2d^{g_0}(x, y) \leq 2C.$$

Let $\eta > 0$ be fixed. Choose x_t at distance t from $x = x_0$ on the g_0-geodesic (x, ξ), and let y_t be the point on the g_0-geodesic (y, ξ) such that $\mathcal{B}_\xi^{g_0}(x, y) = 0$. Let x_t^ε be the projection of x_t on the g_ε-geodesic from x to ξ. Proposition 5.3 ensures that $d^{g_\varepsilon}(x_t, x_t^\varepsilon) \leq \alpha(\varepsilon)$. Let us write

$$\begin{aligned}\left|\mathcal{B}_\xi^{g_\varepsilon}(x, y) - \mathcal{B}_\xi^{g_0}(x, y)\right| \leq & \left|\mathcal{B}_\xi^{g_\varepsilon}(x, y) - d^{g_\varepsilon}(x, x_t^\varepsilon) + d^{g_\varepsilon}(y, x_t^\varepsilon)\right| \\ & + |d^{g_\varepsilon}(x, x_t^\varepsilon) - d^{g_\varepsilon}(y, x_t^\varepsilon) - d^{g_\varepsilon}(x, x_t) + d^{g_\varepsilon}(y, x_t)| \\ & + |d^{g_\varepsilon}(x, x_t) - d^{g_0}(x, x_t) - d^{g_\varepsilon}(y, x_t) + d^{g_0}(y, x_t)| \\ & + \left|\mathcal{B}_\xi^{g_0}(x, y) - d^{g_0}(x, x_t) + d^{g_0}(y, x_t)\right|.\end{aligned}$$

For $t \geq 2C$, by Lemma 5.1, the last term on the right hand side is bounded from the above by $4Ce^{-at}$. For $t \geq 2Ce^{\varepsilon} + \alpha(\varepsilon)$, we also have $d^{g_\varepsilon}(x, x_t^\varepsilon) \geq 2C$ so that again by Lemma 5.1, the first term is bounded from the above by $4Ce^{-ad^{g_\varepsilon}(x,x_t^\varepsilon)} \leq 4Ce^{\alpha(\varepsilon)}e^{-at/2}$. By triangular inequality, the second term is bounded from the above by $2\alpha(\varepsilon)$. The inequality $e^{-\varepsilon}g_0 \leq g_\varepsilon \leq e^{\varepsilon}g_0$ allows to bound the third term by $2(e^{\varepsilon} - 1)(t + C)$.

At last, we get

$$\left|\mathcal{B}_\xi^{g_\varepsilon}(x, y) - \mathcal{B}_\xi^{g_0}(x, y)\right| \leq 4Ce^{\alpha(\varepsilon)}e^{-at/2} + 2\alpha(\varepsilon) + 2(e^{\varepsilon} - 1)(t + C) + 4Ce^{-at}.$$

Let $\eta > 0$ be fixed. Choose first ε_0 so that for $\varepsilon \leq \varepsilon_0$, $\alpha(\varepsilon) \leq 1$. Chose $t \geq 2C$ large enough to guarantee that the first and the last term are each bounded from the above by $\eta/4$. Choose $\varepsilon_1 \leq \varepsilon_0$ small enough to guarantee that for $\varepsilon \leq \varepsilon_1$, $\alpha(\varepsilon) \leq \eta/4$ and $2(e^{\varepsilon} - 1)(t + C) \leq \eta/4$. Thus, $|\mathcal{B}_\xi^{g_\varepsilon}(x, y) - \mathcal{B}_\xi^{g_0}(x, y)| \leq \eta$. This gives the desired result. □

Remark 5.5. – Even though this section is written in a Riemannian setting, all the previous proofs apply verbatim to a family of distances $(d_\varepsilon)_{-1<\varepsilon<1}$ on X such that for all $\varepsilon \in (-1, 1)$, the metric space (X, d_ε) is CAT(−1) and $e^{-\varepsilon}d_0 \leq d_\varepsilon \leq e^{\varepsilon}d_0$.

5.2. Higher regularity, Morse correspondances and geodesic stretch

In this section, we consider metrics $g_\varepsilon \to g_0$ in the $\mathcal{C}^1$-topology. To emphasize the necessity of this assumption, observe that $g_\varepsilon \to g_0$ in the $\mathcal{C}^0$-topology does not imply the convergence of the curvatures nor the convergence of the geodesic flow.

In particular, one can "add mushrooms" on a hyperbolic manifold, and make the mushrooms as small as we want, and build a sequence of manifolds with many points of nonnegative curvature converging to a hyperbolic manifold. The geodesic flow of such g_ε will not converge in general to the geodesic flow of g_0.

In view of its importance in the sequel, recall the convergence that we shall use.

Definition 5.6. – *A family* $(g_\varepsilon)_{-1\le\varepsilon\le 1}$ *of complete Riemannian metrics on* $\widetilde{M}$ *(or* M*) converges in the* $\mathcal{C}^1$*-topology, uniformly on compact sets, to* g_0 *if:*

1. (g_ε) *converges to* g_0 *uniformly on compact sets, i.e., for all compact sets* $K \subset T\widetilde{M}$,
$$\lim_{\varepsilon\to 0}\sup_{v\in K}|g_\varepsilon(v,v) - g_0(v,v)| = 0;$$
2. *the first derivatives of* g_ε *also converge uniformly on compact sets to those of* g_0.

By [22, Thm. 2.79], it implies for all fixed $T > 0$ the uniform convergence on compact sets of the geodesic flows $v \mapsto g_\varepsilon^T v$. As a consequence, we get the following result.

Theorem 5.7. – *Let* $(g_\varepsilon)_{-1<\varepsilon<1}$ *be a family of metrics on* $\widetilde{M}$ *with sectional curvatures satisfying* $K_{g_\varepsilon} \le -a^2$*, such that for all* $\varepsilon \in (-1,1)$*, at all* $x \in \widetilde{M}$*,* $e^{-\varepsilon}g_0 \le g_\varepsilon \le e^{\varepsilon}g_0$*, and* $g_\varepsilon \to g_0$ *in the* $\mathcal{C}^1$ *topology, uniformly on compact sets.*

Let $\widetilde{\Phi}^{g_0\to g_\varepsilon}$ *and* $\widetilde{\Psi}^{g_0\to g_\varepsilon}$ *be the Morse correspondances between* $S^{g_0}\widetilde{M}$ *and* $S^{g_\varepsilon}\widetilde{M}$ *defined in Section 2.4. Then* $\widetilde{\Phi}^{g_0\to g_\varepsilon} \to \mathrm{Id}$ *and* $\widetilde{\Psi}^{g_0\to g_\varepsilon} \to \mathrm{Id}$ *uniformly on all compact sets* $K \subset S^{g_0}\widetilde{M}$ *in the uniform topology of* $\mathcal{C}^0(K, T\widetilde{M})$.

Proof. – Let K be a fixed compact set of $S^{g_0}\widetilde{M}$ and $v \in K$, with $v_\pm^{g_0}$ the endpoints of its g_0-geodesic in $\partial\widetilde{M}$. Denote by $(\gamma_0(t))_{t\in\mathbb{R}}$ the parametrization of this geodesic such that $\gamma_0'(0) = v$. Let γ_ε be the parametrization of the g_ε-geodesic with the same endpoints, with $v_\varepsilon = \gamma_\varepsilon'(0) = \Phi^{g_0\to g_\varepsilon}(v)$.

By Proposition 5.4 and definitions from Section 2.4, uniform convergence of $\widetilde{\Psi}^{g_0\to g_\varepsilon}$ on compact sets will follow from the convergence of $\widetilde{\Phi}^{g_0\to g_\varepsilon}$. So let us prove the latter.

We will use the distance $d(w,w') = \sup_{t\in[0,1]} d^{g_0}(\pi(g_0^t w, g_0^t w')$ on $T\widetilde{M}$ and show that for all $\alpha > 0$, if ε is small enough, for all $v \in K$ and $t \in [0,1]$, $d^{g_0}(\pi(g_0^t v), \pi(g_0^t v_\varepsilon)) \le \alpha$.

Choose some $\alpha > 0$. By Propositions 5.3 and 5.4, for ε small enough, uniformly in $v \in K$, and $t \in [-1,1]$, we know that γ_ε is in the $\alpha/2$-neighborhood of γ_0, and $\gamma_\varepsilon(t)$ is uniformly close to $\gamma_0(t)$. It implies that $v_\varepsilon = \gamma_\varepsilon'(0)$ and $v_0 = \gamma_0'(0)$ are uniformly close. As $g_\varepsilon \to g_0$ in the $\mathcal{C}^1$-topology, uniformly on compact sets, it implies that for ε small enough, for all $t \in [-1,1]$, $\pi(g_\varepsilon^t(v_\varepsilon))$ and $\pi(g_0^t(v_\varepsilon))$ will stay $\alpha/2$-close. In particular, $\pi(g_0^t(v_\varepsilon))$ will stay α-close from $\gamma_0(t)$, for $t \in [-1,1]$. That is the desired convergence. □

Remark 5.8. – Adapting Theorem 5.7 and the definition of the geodesic stretch in the setting of CAT(−1) spaces would require a careful definition of the tangent bundle on such spaces with its topology, which we will not do here.

Let us conclude this section by a key technical ingredient.

THEOREM 5.9. – *Let $(g_\varepsilon)_{-1<\varepsilon<1}$ be a family of metrics on $\widetilde{M}$ with sectional curvatures satisfying $K_{g_\varepsilon} \leq -a^2$, such that for all $\varepsilon \in (-1,1)$, at all $x \in \widetilde{M}$, $e^{-\varepsilon} g_0 \leq g_\varepsilon \leq e^{\varepsilon} g_0$, and $g_\varepsilon \to g_0$ in the $\mathcal{C}^1$-topology, uniformly on compact sets.*

Then uniformly on compact sets of $S^{g_0}\widetilde{M}$, we have

$$\limsup_{\varepsilon\to 0} \mathcal{E}^{g_0\to g_\varepsilon}(v) \leq 1.$$

Moreover, for any geodesic current μ such that $m_\mu^{g_0}$ is finite, we have

$$\mathcal{E}^{g_0\to g_\varepsilon} \to 1 \quad m_\mu^{g_0} - \textit{almost surely}.$$

Proof. – Observe that Lemma 2.5 gives the obvious upper bound $\limsup_{\varepsilon\to 0} \mathcal{E}^{g_0\to g_\varepsilon} \leq 1$, uniformly on $S^{g_0}\widetilde{M}$. For the same reason, $\limsup_{\varepsilon\to 0} \mathcal{E}^{g_\varepsilon\to g_0} \leq 1$, uniformly on $S^{g_\varepsilon}\widetilde{M}$. By Corollary 2.15, one easily deduces that

$$\frac{\|m_\mu^{g_\varepsilon}\|}{\|m_\mu^{g_0}\|} \to 1 \quad \text{when} \quad \varepsilon \to 0. \tag{22}$$

Combined with the fact that $\limsup_{\varepsilon\to 0} \mathcal{E}^{g_0\to g_\varepsilon} \leq 1$, this implies in turn that $\mathcal{E}^{g_0\to g_\varepsilon} \to 1$ $m_\mu^{g_0}$-almost surely. □

5.3. Narrow convergence of measures associated to a fixed geodesic current

Recall that if μ is a Γ-invariant geodesic current and g an admissible metric on M, we denote by m_μ^g the locally finite Radon measure on $S^g M$ whose lift to $S^g\widetilde{M}$ is given by

$$d\widetilde{m}_\mu^g(v) = (H^g)^*(\mu \times dt).$$

The results of the previous paragraph imply the following fact.

PROPOSITION 5.10. – *Let $(g_\varepsilon)_{-1<\varepsilon<1}$ be a family of metrics on $\widetilde{M}$ whose sectional curvatures satisfy $K_{g_\varepsilon} \leq -a^2$, and such that for all $\varepsilon \in (-1,1)$, at all $x \in \widetilde{M}$, $e^{-\varepsilon} g_0 \leq g_\varepsilon \leq e^{\varepsilon} g_0$, and $g_\varepsilon \to g_0$ in the $\mathcal{C}^1$-topology, uniformly on compact sets. Let μ be a Γ-invariant geodesic current such that $\|m_\mu^{g_0}\| < \infty$. Then the measures $m_\mu^{g_\varepsilon}$ converge to $m_\mu^{g_0}$ in the dual of bounded continuous maps on TM (i.e., in the narrow topology).*

Proof. – By definition, for all $\varepsilon \in (-1,1)$ we have $m_\mu^{g_\varepsilon} = (\Psi^{g_0\to g_\varepsilon})_*\left(\mathcal{E}^{g_0\to g_\varepsilon} \times m_\mu^{g_0}\right)$. Therefore the weak-* convergence (in the dual of continuous compactly supported functions) is an immediate consequence of Theorem 5.7 and the dominated convergence theorem.

We also showed that $\|m_\mu^{g_\varepsilon}\| \to \|m_\mu^{g_0}\|$, see equation (22). It is classical that it implies the convergence of the above measures in the dual of bounded continuous functions. The result follows. □

5.4. Weak convergence of Bowen-Margulis measures

We now show that, provided they are unique, the Bowen-Margulis measures are continuous in the weak-* topology under Lipschitz deformations of the metric.

PROPOSITION 5.11. – *Let $(g_\varepsilon)_{-1\le\varepsilon\le 1}$ be a family of metrics on M with sectional curvatures satisfying $K_{g_\varepsilon} \le -a^2$, such that (Γ, g_0) is divergent and for all ε, at all $x \in \widetilde{M}$, $e^{-\varepsilon} g_0 \le g_\varepsilon \le e^{\varepsilon} g_0$.*

Then for all $x \in \widetilde{M}$, any Patterson-Sullivan measure for g_ε normalized at o converges: $\lim_{\varepsilon\to 0} \nu_x^{g_\varepsilon} = \nu_x^{g_0}$ in the weak- topology, uniformly in x on compact sets of $\widetilde{M}$.*

Proof. – For all $\varepsilon \in (-1,1) \setminus \{0\}$, let $\nu_o^{g_\varepsilon}$ be any Patterson-Sullivan measure on Λ_Γ, normalized into a probability measure. (Observe that such a measure is not necessarily unique for $\varepsilon \neq 0$, because only (Γ, g_0) is assumed to be divergent.) Let $\widetilde{\nu}_o = \lim_{\varepsilon_i \to 0} \nu_0^{g_{\varepsilon_i}}$ be any of its weak limits. Define for all $x \in \widetilde{M}$ a measure $\widetilde{\nu}_x$ on Λ_Γ by

$$\frac{d\widetilde{\nu}_x}{d\widetilde{\nu}_o}(\xi) = e^{-\delta(g_0)\mathcal{B}_\xi^{g_0}(o,x)}.$$

It is a Γ-invariant, $\delta(g_0)$-conformal family of measures as defined in (14), normalized at o. By uniqueness of such a family, it coincides with $(\nu_x^{g_0})_{x\in\widetilde{M}}$. □

Recall that μ_{BM}^g denotes the *g-Bowen-Margulis geodesic current* on $\partial^2\widetilde{M}$ given by

$$d\mu_{\mathrm{BM}}^g(\eta,\xi) = d\nu_x^g(\eta)d\nu_x^g(\xi) = e^{-\delta(g)(\mathcal{B}_\xi^g(o,x)+\mathcal{B}_\eta^g(o,x))}d\nu_o^g(\eta)d\nu_o^g(\xi),$$

where x is any point on the g-geodesic with endpoints (η,ξ). We get the immediate corollary of Propositions 5.4 and 5.11.

COROLLARY 5.12. – *Under the same assumptions, in the weak-* topology of $\partial^2\widetilde{M}$, $\lim_{\varepsilon\to 0} \mu_{\mathrm{BM}}^{g_\varepsilon} = \mu_{\mathrm{BM}}^{g_0}$.*

REMARK 5.13. – Once again, Proposition 5.11 and Corollary 5.12 are still valid if we consider a family of Γ-invariant distances $(d_\varepsilon)_{\varepsilon\in(-1,1)}$ on $\widetilde{M}$ such that $(\widetilde{M}, d_\varepsilon)$ is a CAT(-1) and $e^{-\varepsilon}d_0 \le d_\varepsilon \le e^{\varepsilon}d_0$ for all $\varepsilon \in (-1,1)$.

We end this section by the convergence of Bowen-Margulis measures.

THEOREM 5.14 (Convergence of Bowen-Margulis measures). – *Let $(g_\varepsilon)_{-1<\varepsilon<1}$ be a family of metrics on $\widetilde{M}$ with sectional curvatures satisfying $K_{g_\varepsilon} \le -a^2$, such that for all $\varepsilon \in (-1,1)$, at all $x \in \widetilde{M}$, $e^{-\varepsilon}g_0 \le g_\varepsilon \le e^{\varepsilon}g_0$ and $g_\varepsilon \to g_0$ in the $\mathcal{C}^1$-topology, uniformly on compact sets. Assume that Γ is divergent for all metrics g_ε. Then in the weak-* topology of TM,*

$$\lim_{\varepsilon\to 0} m_{\mathrm{BM}}^{g_\varepsilon} = m_{\mathrm{BM}}^{g_0}.$$

Proof. – Let φ be a continuous map with compact support on TM. Write the difference $\int_{TM}\varphi dm^{g_\varepsilon}_{\mu^{g_\varepsilon}_{\mathrm{BM}}} - \int_{TM}\varphi dm^{g_0}_{\mu^{g_0}_{\mathrm{BM}}}$ as

$$\left(\int_{TM}\varphi dm^{g_\varepsilon}_{\mu^{g_\varepsilon}_{\mathrm{BM}}} - \int_{TM}\varphi dm^{g_0}_{\mu^{g_\varepsilon}_{\mathrm{BM}}}\right) + \left(\int_{TM}\varphi dm^{g_0}_{\mu^{g_\varepsilon}_{\mathrm{BM}}} - \int_{TM}\varphi dm^{g_0}_{\mu^{g_0}_{\mathrm{BM}}}\right).$$

By Corollary 5.12, the second difference converges to 0.

Proposition 2.13 allows to rewrite the first difference as

$$\int_{TM} \varphi dm^{g_\varepsilon}_{\mu^{g_\varepsilon}_{\mathrm{BM}}} - \int_{TM} \varphi dm^{g_0}_{\mu^{g_\varepsilon}_{\mathrm{BM}}} = \int_{TM} \left(\varphi \circ \Psi^{g_0\to g_\varepsilon} \times \mathcal{E}^{g_0\to g_\varepsilon} - \varphi\right) dm^{g_0}_{\mu^{g_\varepsilon}_{\mathrm{BM}}}.$$

By Corollary 5.12, $m^{g_0}_{\mu^{g_\varepsilon}_{\mathrm{BM}}}$ converges weakly to $m^{g_0}_{\mu^{g_0}_{\mathrm{BM}}}$ in the dual of continuous functions with compact support.

By Theorem 5.7, as $\limsup_{\varepsilon\to 0} \mathcal{E}^{g_0\to g_\varepsilon} \le 1$, if $\varphi \ge 0$, we have

$$\limsup_{\varepsilon\to 0} \varphi \circ \Psi^{g_0\to g_\varepsilon} \times \mathcal{E}^{g_0\to g_\varepsilon} - \varphi \le 0.$$

As the support of these maps is included in a fixed compact set, we deduce that

$$\limsup_{\varepsilon\to 0} \int_{TM} \left(\varphi \circ \Psi^{g_0\to g_\varepsilon} \times \mathcal{E}^{g_0\to g_\varepsilon} - \varphi\right) dm^{g_0}_{\mu^{g_\varepsilon}_{\mathrm{BM}}} \le 0.$$

Now, rewrite this first difference as

$$-\left(\int_{TM} \varphi dm^{g_0}_{\mu^{g_\varepsilon}_{\mathrm{BM}}} - \int_{TM} \varphi dm^{g_\varepsilon}_{\mu^{g_\varepsilon}_{\mathrm{BM}}}\right) = -\int_{TM} \left(\varphi \circ \Psi^{g_\varepsilon\to g_0} \times \mathcal{E}^{g_\varepsilon\to g_0} - \varphi\right) dm^{g_\varepsilon}_{\mu^{g_\varepsilon}_{\mathrm{BM}}}.$$

Observe first that, by the same arguments used in the proof of equation (22), the ratios of masses $\frac{\|m^{g_0}_{\mu^{g_\varepsilon}_{\mathrm{BM}}}\|}{\|m^{g_\varepsilon}_{\mu^{g_\varepsilon}_{\mathrm{BM}}}\|}$ go to 1 when $\varepsilon \to 0$.

For the same reason as above, $\limsup_{\varepsilon\to 0} \mathcal{E}^{g_\varepsilon\to g_0} \le 1$, so that for $\varphi \ge 0$, by Theorem 5.7, uniformly on TM, the limsup of $\varphi \circ \Psi^{g_\varepsilon\to g_0} \times \mathcal{E}^{g_\varepsilon\to g_0} - \varphi$ is nonpositive. By convergence of the ratio of masses mentioned above, and by convergence of $m^{g_0}_{\mu^{g_\varepsilon}_{\mathrm{BM}}}$ to $m^{g_0}_{\mu^{g_0}_{\mathrm{BM}}}$, its integral also has a nonpositive limsup, and the sign minus in the above expression gives

$$\liminf_{\varepsilon\to 0} \int_{TM} \varphi dm^{g_\varepsilon}_{\mu^{g_\varepsilon}_{\mathrm{BM}}} - \int_{TM} \varphi dm^{g_0}_{\mu^{g_\varepsilon}_{\mathrm{BM}}} \ge 0.$$

The result follows. □

6. Differentiability of the metric and topological entropies

In this section, we show differentiability of topological and measure theoretic entropies at $\varepsilon = 0$, along a variation $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ of metrics of a negatively curved Riemannian manifold $(M = \widetilde{M}/\Gamma, g_0)$. We will focus on two distinct situations.

First, let μ be a Γ-invariant geodesic current on $\partial^2\widetilde{M}$, and for all $\varepsilon \in (-1,1)$, let $m^{g_\varepsilon}_\mu$ be the associated invariant measure for the geodesic flow (g^t_ε) (see Section 2). Assume that the total mass of $m^{g_0}_\mu$ is finite. We will show that the measure theoretic entropy $\varepsilon \mapsto h(m^{g_\varepsilon}_\mu, g_\varepsilon)$ is $\mathcal{C}^1$, with explicit derivatives.

We then focus on the topological entropy. Provided that Bowen-Margulis measures of each geodesic flow (g^t_ε) are finite, and that their masses vary continuously, we show that the topological entropy is also $\mathcal{C}^1$, with a similar formula for its derivative. The proofs are similar in both situations, and inspired from [31] and [52].

Definition 6.1. – *Let M be a (non-compact) manifold. We say that a family of complete Riemannian metrics $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ on M converges to g_0* in the $\mathcal{C}^1$-uniform topology *if:*

1. *$g_\varepsilon \to g_0$ in the $\mathcal{C}^1$ topology, uniformly on compact sets, as in Definition 5.6;*

2. *there exists* $\kappa > 0$ *such that for all* $\varepsilon \in (-1,1)$ *and all* $v \in TM$ *with* $\|v\|_{g_0} \leq 1$,

$$\left| \frac{d}{ds}\bigg|_{s=\varepsilon} g_s(v,v) \right| \leq \kappa.$$

A $\mathcal{C}^2$ variation of metric with compact support, or with non-compact support but uniformly bounded first and second derivatives, is a typical example of such a uniformly $\mathcal{C}^1$ family. If $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ is such a $\mathcal{C}^1$-uniform family of complete metrics on M, one immediately sees that there exists $B = B(C_1, \varepsilon) > 0$ such that at all $x \in M$ and for all $\varepsilon \in (-1,1)$,

$$e^{-B\varepsilon} g_0 \leq g_\varepsilon \leq e^{B\varepsilon} g_0,$$

which allows us to apply the results shown in the previous section.

6.1. Variation of metric entropy

This paragraph is devoted to the proof of the following result, which seems to us new even in the compact case.

Theorem 6.2. – *Let* $b > a > 0$, $\varepsilon > 0$ *and let* $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ *be a family of complete metrics on* $M = \widetilde{M}/\Gamma$ *whose curvatures and first derivatives of curvatures are uniformly bounded, and moreover such that for all* $\varepsilon \in (-1,1)$ *and at all points,* $-b^2 \leq K_{g_\varepsilon} \leq -a^2$. *Assume that* $g_\varepsilon \to g_0$ *in the* $\mathcal{C}^1$*-uniform topology. Let* μ *be a* Γ*-invariant geodesic current on* $\partial^2\widetilde{M}$ *such that* $m_\mu^{g_0}$ *is finite.*

Then the local entropy $\varepsilon \mapsto h^\Gamma_{\mathrm{loc}}(m_\mu^{g_\varepsilon}, g_\varepsilon)$ *of the* (g^t_ε)*-invariant measures* $m_\mu^{g_\varepsilon}$ *is differentiable at* $\varepsilon = 0$ *with derivative given by*

$$\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} h^\Gamma_{\mathrm{loc}}(m_\mu^{g_\varepsilon}, g_\varepsilon) = -h^\Gamma_{\mathrm{loc}}(m_\mu^{g_0}, g_0) \times \int_{S^{g_0}M} \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \|v\|^{g_\varepsilon} \frac{dm_\mu^{g_0}(v)}{\|m_\mu^{g_0}\|}.$$

Proof. – Let μ be a Γ-invariant geodesic current on $\partial^2\widetilde{M}$ such that $m_\mu^{g_0}$ is finite. It follows from Proposition 2.15 that for all $\varepsilon \in (-1,1)$, the measures $m_\mu^{g_\varepsilon}$ are finite. Moreover, by Corollary 5.10, $\lim_{\varepsilon\to 0} m_\mu^{g_\varepsilon} = m_\mu^{g_0}$ and $\lim_{\varepsilon\to 0} \frac{m_\mu^{g_\varepsilon}}{\|m_\mu^{g_\varepsilon}\|} = \frac{m_\mu^{g_0}}{\|m_\mu^{g_0}\|}$ in the narrow topology.

By Theorem 3.7, if g_1 and g_2 are admissible metrics on M, we know that

$$h^\Gamma_{\mathrm{loc}}(m_\mu^{g_2}, g_2) = \int_{S^{g_2}M} \mathcal{E}^{g_1\to g_2}(v)\, d\bar{m}_\mu^{g_2}(v) \times h^\Gamma_{\mathrm{loc}}(m_\mu^{g_1}, g_1). \tag{23}$$

By Theorem 5.9, this implies that the local entropy $h^\Gamma_{\mathrm{loc}}(m_\mu^{g_\varepsilon}, g_\varepsilon)$ converges to $h^\Gamma_{\mathrm{loc}}(m_\mu^{g_0}, g_0)$ when $\varepsilon \to 0$. Moreover, (23) and Lemma 2.5 also imply that

$$h^\Gamma_{\mathrm{loc}}(m_\mu^{g_2}, g_2) \leq \int_{S^{g_2}M} \|v\|_{g_1}\, d\overline{m}_\mu^{g_2}(v) \times h^\Gamma_{\mathrm{loc}}(m_\mu^{g_1}, g_1).$$

Applying it with $g_1 = g_0$ and $g_2 = g_\varepsilon$ first, and second with $g_1 = g_\varepsilon$ and $g_2 = g_0$, we get

$$\begin{aligned} h^\Gamma_{\mathrm{loc}}(m_\mu^{g_0}, g_0)\left(\frac{1}{\int_{S^{g_0}M}\|v\|^{g_\varepsilon}\, d\overline{m}_\mu^{g_0}(v)} - 1\right) &\leq h^\Gamma_{\mathrm{loc}}(m_\mu^{g_\varepsilon}, g_\varepsilon) - h^\Gamma_{\mathrm{loc}}(m_\mu^{g_0}, g_0)) \\ &\leq h^\Gamma_{\mathrm{loc}}(m_\mu^{g_0}, g_0)\left(\int_{S^{g_\varepsilon}M}\|v\|^{g_0}\, d\overline{m}_\mu^{g_\varepsilon}(v) - 1\right), \end{aligned}$$

which yields to

$$h^{\Gamma}_{\mathrm{loc}}(m^{g_0}_{\mu}, g_0)\frac{\int_{S^{g_0}M}\frac{\|v\|^{g_0}-\|v\|^{g_\varepsilon}}{\varepsilon}d\overline{m}^{g_0}_{\mu}(v)}{\int_{S^{g_0}M}\|v\|^{g_\varepsilon}\,d\overline{m}^{g_0}_{\mu}(v)} \leq \frac{h^{\Gamma}_{\mathrm{loc}}(m^{g_\varepsilon}_{\mu}, g_\varepsilon) - h^{\Gamma}_{\mathrm{loc}}(m^{g_0}_{\mu}, g_0)}{\varepsilon}$$

$$\leq h^{\Gamma}_{\mathrm{loc}}(m^{g_0}_{\mu}, g_0)\int_{S^{g_\lambda}M}\frac{\|v\|^{g_0}-\|v\|^{g_\varepsilon}}{\varepsilon}d\overline{m}^{g_\varepsilon}_{\mu}(v).$$

Now, dominated convergence theorem, continuity of $\varepsilon \to h^{\Gamma}_{\mathrm{loc}}(m^{g_\varepsilon}_{\mu}, g_\varepsilon)$ at $\varepsilon = 0$, and narrow convergence of $\frac{m^{g_\varepsilon}_{\mu}}{\|m^{g_\varepsilon}_{\mu}\|}$ towards $\frac{m^{g_0}_{\mu}}{\|m^{g_0}_{\mu}\|}$ give

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} h^{\Gamma}_{\mathrm{loc}}(m^{g_\varepsilon}_{\mu}, g_\varepsilon) = -h^{\Gamma}_{\mathrm{loc}}(m^{g_0}_{\mu}, g_0)\times\int_{S^{g_0}M}\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\|v\|^{g_\varepsilon}\,d\overline{m}^{g_0}_{\mu}(v).$$

This is the desired result. □

6.2. Variation of topological entropy

We now show differentiability of the topological entropy $h_{\mathrm{top}}(g_\varepsilon)$ at $\varepsilon = 0$. It is not a corollary of Theorem 6.2 since we have to consider Bowen-Margulis geodesic currents $\mu^{g_\varepsilon}_{\mathrm{BM}}$ depending on the metric g_ε. However, the strategy of proof is very similar, as by Theorem 5.14, $m^{g_\varepsilon}_{\mathrm{BM}} \to m^{g_0}_{\mathrm{BM}}$ in the weak-* topology. The only missing ingredient is the convergence of Bowen-Margulis measures in the dual of bounded continuous functions. It is therefore required in the assumptions of Theorem 6.3. We refer to Section 7 for the study of the large class of the so-called SPR manifolds, which will satisfy this assumption.

THEOREM 6.3. – *Let $b > a > 0$, and let $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ be a family of complete metrics on M such that*

1. *for all $\varepsilon \in (-1, 1)$ and at all point, $-b^2 \leq K_{g_\varepsilon} \leq -a^2$;*
2. *$g_\varepsilon \to g_0$ uniformly in the $\mathcal{C}^1$ topology as in Definition 6.1;*
3. *for all $\varepsilon \in (-1, 1)$, the Bowen-Margulis measure $m^{g_\varepsilon}_{\mathrm{BM}}$ of the geodesic flow $(g^t_\varepsilon)_{t\in\mathbb{R}}$ on $S^{g_\varepsilon}M$ has finite mass;*
4. *the map $\varepsilon \to \|m^{g_\varepsilon}_{\mathrm{BM}}\|$ is continuous at $\varepsilon = 0$.*

Then the entropy $\varepsilon \mapsto h_{\mathrm{top}}(g_\varepsilon)$ is $\mathcal{C}^1$ at $\varepsilon = 0$ with derivative given by

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} h_{\mathrm{top}}(g_\varepsilon) = -h_{\mathrm{top}}(g_0)\int_{S^{g_0}M}\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0}\|v\|^{g_\varepsilon}\,\frac{dm^{g_0}_{\mathrm{BM}}(v)}{\|m^{g_0}_{\mathrm{BM}}\|}.$$

Proof. – As the preceding one, our strategy of proof is inspired from [31] and [52]. Corollary 3.18 shows that if g_1 and g_2 are admissible metrics M with finite Bowen-Margulis measures, then

$$h_{\mathrm{top}}(g_2) \leq \int_{S^{g_2}M}\mathcal{E}^{g_1\to g_2}(v)\,\frac{dm^{g_2}_{\mathrm{BM}}(v)}{\|m^{g_2}_{\mathrm{BM}}\|}h_{\mathrm{top}}(g_1) \leq \int_{S^{g_2}M}\|v\|_{g_1}\,\frac{dm^{g_2}_{\mathrm{BM}}(v)}{\|m^{g_2}_{\mathrm{BM}}\|}h_{\mathrm{top}}(g_1),$$

where the last inequality follows from Lemma 2.5. Applying it to g_ε and g_0 on both sides, we get for all $\varepsilon \in (-1,1)$,

$$h_{\mathrm{top}}(g_0) \times \frac{\int_{S^{g_0}M} \frac{\|v\|^{g_0}-\|v\|^{g_\varepsilon}}{\varepsilon} \frac{dm_{\mathrm{BM}}^{g_0}(v)}{\|m_{\mathrm{BM}}^{g_0}\|}}{\int_{S^{g_0}M} \|v\|^{g_\varepsilon} \frac{dm_{\mathrm{BM}}^{g_0}(v)}{\|m_{\mathrm{BM}}^{g_0}\|}} \leq \frac{h_{\mathrm{top}}(g_\varepsilon)-h_{\mathrm{top}}(g_0)}{\varepsilon}$$
$$\leq h_{\mathrm{top}}(g_0) \times \int_{S^{g_\varepsilon}M} \frac{\|v\|^{g_0}-\|v\|^{g_\varepsilon}}{\varepsilon} \frac{dm_{\mathrm{BM}}^{g_\varepsilon}(v)}{\|m_{\mathrm{BM}}^{g_\varepsilon}\|}.$$

The assumptions of the theorem are now exactly done to make the above integrals converge. We deduce that topological entropy is differentiable at $\varepsilon = 0$, with

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} h_{\mathrm{top}}(g_\varepsilon) = -h_{\mathrm{top}}(g_0) \times \int_{S^{g_0}M} \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \|v\|^{g_\varepsilon}\, d\bar{m}_{\mathrm{BM}}^{g_0}(v). \qquad \square$$

7. Entropy at infinity and Strongly Positively Recurrent groups

In this section, our goal is to propose a wide class of manifolds and metrics to which Theorem 6.3 will apply. In view of this goal, proving differentiability of entropy, this section is apparently technical. However, the definition of this class of manifolds, and the related concepts studied here, is probably one of the main novelties in our paper. We refer to [3, 14, 55, 25] for further results on these manifolds.

We define the *entropy at infinity* $\delta_\infty(M,g)$ of a negatively curved manifold (M,g) (see Definition 7.12), as the maximal exponential growth of the dynamics away from any given (large) compact set. In particular, it is invariant under any $\mathcal{C}^2$ compact perturbation of a negatively curved metric.

We introduce the class of *strongly positively recurrent* manifolds (M,g), defined as those negatively curved manifolds whose entropy at infinity is strictly smaller than the total topological entropy of the geodesic flow.

As said in the introduction, the notion of *strong positive recurrence* appeared in [48] in the context of symbolic dynamics over an infinite alphabet, and has been used later by some other authors among which [8]. A former terminology due to [26] was *stable positive recurrence*. This terminology could be more adapted to the kind of results that we prove here. In any case, as will be seen below and in [25], the acronym SPR is perfectly adapted to the concept.

The simplest nontrivial examples are geometrically finite hyperbolic manifolds, but this class also includes most known examples of non-compact manifolds with negative curvature whose geodesic flow has a finite Bowen-Margulis measure, and many new ones (see Section 7.3).

Our main result is the following.

THEOREM 7.1. – *Let (M,g_0) be a manifold with pinched negative curvature and bounded derivatives of the curvature.*

If (M,g_0) is a strongly positively recurrent manifold, then the Bowen-Margulis measure of its geodesic flow is finite.

Moreover, if $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ is a uniformly $\mathcal{C}^1$-variation of smooth complete metrics on M with pinched negative curvature and bounded derivatives of metrics, then the following holds.

1. *For $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ small enough, all metrics g_ε are strongly positively recurrent.*
2. *The mass of the associated (finite) Bowen-Margulis $m_{\mathrm{BM}}^{g_\varepsilon}$ varies continuously on $(-\varepsilon_0, \varepsilon_0)$.*

The first part of this theorem (finiteness of Bowen-Margulis) has been proven independently and simultaneously by A. Velozo [53] by a different approach.

As a corollary, all assumptions of Theorem 6.3 hold for such a variation of metrics, so that we get the following result, which answers positively the question at the origin of this work.

Corollary 7.2. – *Let $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ be a uniformly $\mathcal{C}^1$ family of complete metrics on the manifold M with pinched negative curvature and bounded derivatives of the curvature. Assume that (M, g_0) is strongly positively recurrent. Then the entropy $\varepsilon \mapsto h_{\mathrm{top}}(g_\varepsilon)$ is $\mathcal{C}^1$ around $\varepsilon = 0$, and its derivative is given by*

$$\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} h_{\mathrm{top}}(g_\varepsilon) = -h_{\mathrm{top}}(g_0) \times \int_{S^{g_0}M} \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \|v\|^{g_\varepsilon} \, \frac{dm_{\mathrm{BM}}^{g_0}(v)}{\|m_{\mathrm{BM}}^{g_0}\|}.$$

In view of the length of this section, let us present the strategy of the proof.

Heuristically, the SPR assumption allows to neglect the dynamical contribution of the complement of a large compact set to the dynamics. We develop this idea in two introductive parts 7.1 and 7.2, defining the growth of the fundamental group outside a compact set, the entropy at infinity and the class of Strongly Positively Recurrent manifolds.

In Subsection 7.3 we provide an illustration of this concept, by describing different families of examples of SPR manifolds.

A criterion of finiteness of the Bowen-Margulis measure from [41] is used to prove the first part of Theorem 7.1. Subsection 7.4 is devoted to this proof.

All entropies considered here are continuous for a negatively curved perturbation $(g_\varepsilon)_{-1\le\varepsilon\le1}$ satisfying $e^{-\varepsilon}g_0 \le g_\varepsilon \le e^{\varepsilon}g_0$. Thus, the SPR assumption, that is the existence of a critical gap between the entropy at infinity and the topological entropy is stable under such small perturbations. And the existence of a large compact set concentrating the most part of the dynamics allows to prove that its complement is of small Bowen-Margulis measure, uniformly in the perturbation. These ideas are developped in Subsection 7.5, where we prove that for a variation of a SPR metric as above, the mass of the Bowen-Margulis measures varies continuously.

As said in the introduction, these results imply all theorems stated in the introduction. Theorem 1.7 is an immediate consequence of Section 7.3 and the first part of Theorem 7.1. Theorem 1.8 is a reformulation of the second part of Theorem 7.1. At last, our main result, Theorem 1.9, follows from Theorems 6.3 and 1.8 (or 7.1).

7.1. Fundamental group outside a given compact set

Let (M, g) be a complete Riemannian manifold with pinched negative curvature $-b^2 \leq K_g \leq -a^2 < 0$, whose fundamental group $\Gamma = \pi_1(M)$ is non-elementary. Let $p_\Gamma : \widetilde{M} \to M$ be the universal covering map. Let $o \in \widetilde{M}$ be a point, fixed once for all. For any set $W \subset M$, we will write $W^c = M \backslash W$.

DEFINITION 7.3. – *Let $W \subset M$ be a compact pathwise connected set which is the closure of its interior, and whose boundary is piecewise $\mathcal{C}^1$. A* nice preimage of W *is a compact set $\widetilde{W} \subset \widetilde{M}$ such that*

1. *$p_\Gamma(\widetilde{W}) = W$ and the restriction of p_Γ to the interior of $\widetilde{W}$ is injective;*
2. *$\widetilde{W}$ has a piecewise $\mathcal{C}^1$ boundary.*

REMARK 7.4. – We will often refer to and use results of [41]. In this reference, $\mathcal{W}$ is a subset of $S^g M$ and $\widetilde{\mathcal{W}}$ is an open set inside $p_\Gamma^{-1}(\mathcal{W})$ such that $p_\Gamma : \widetilde{\mathcal{W}} \to \mathcal{W}$ is onto. As we deal with several metrics and several unit tangent bundles, it is better here to work with $W \subset M$. The reader can think to $\mathcal{W}$ as $S^g W$. The fact that W is compact here, and $\mathcal{W}$ open in [41] is just a matter of taste in some arguments.

We gather in the following lemma elementary useful facts.

LEMMA 7.5. – *Let W be a compact pathwise connected set with piecewise $\mathcal{C}^1$ boundary, which is the closure of its interior.*

1. *A nice preimage $\widetilde{W}$ of W exists.*
2. *If $W_2 \supset W_1$, then they admit nice preimages $\widetilde{W_2} \supset \widetilde{W_1}$.*
3. *If $\gamma \neq$ id then $\gamma \mathring{\widetilde{W}} \cap \mathring{\widetilde{W}} = \emptyset$.*
4. *The set $\{\gamma \in \Gamma\ ;\ \gamma\widetilde{W} \cap \widetilde{W} \neq \emptyset\}$ is finite. We call such $\gamma\widetilde{W}$ the* adjacent elements *of $\widetilde{W}$.*

Proof. – Choose some $w \in W$, lift it to $\widetilde{w} \in p_\Gamma^{-1}(W)$ and construct the *Dirichlet domain*

$$\widetilde{W} = \{z \in p_\Gamma^{-1}(W),\ \forall \gamma \in \Gamma,\ d^g(z, \widetilde{w}) \leq d^g(z, \gamma\widetilde{w})\}.$$

It is a compact set with $\mathcal{C}^1$-boundary which satisfies the properties stated in the lemma. If $W_1 \subset W_2$, choose some $w \in p_\Gamma^{-1}(W_1) \subset p_\Gamma^{-1}(W_2)$. For $i = 1, 2$ the Dirichlet domains $\widetilde{W_1} \subset \widetilde{W_2} \subset p_\Gamma^{-1}(W_i)$ satisfy Fact 2. □

The following notion was introduced in [41].

DEFINITION 7.6. – *Let $W \subset M$ be a compact set and $\widetilde{W}$ a nice preimage of W. The* fundamental group of M out of $\widetilde{W}$ *is the set $\Gamma^g_{\widetilde{W}}$ of elements $\gamma \in \Gamma$ such that there exists $x, y \in \widetilde{W}$ and a g-geodesic segment c_γ joining x to γy such that for all $h \in \Gamma$,*

$$c_\gamma \cap p_\Gamma^{-1} W = c_\gamma \cap \Gamma\widetilde{W} \subset \widetilde{W} \cup \gamma\widetilde{W}.$$

By compactness of $\widetilde{W}$ we will always assume that $x, y \in \partial\widetilde{W}$.

Heuristically, as explained in [41], $\Gamma^g_{\widetilde{W}}$ represents loops $p_\Gamma([x,\gamma y])$ which go outside W at the beginning, and come back to W only at the end. This heuristics does not work perfectly, depending on the topology of W, for example when it has holes.

The set $\Gamma^g_{\widetilde{W}}$ will help controlling what happens far at infinity. In particular it follows immediately from the definition that it is not sensitive to small compact perturbations of the metric g, as stated in the proposition below.

Proposition 7.7. – *Let (M, g_0) be a complete negatively curved metric and $W \subset M$ be a compact set, with nice preimage $\widetilde{W}$. For any proper compact subset $K \subset \overset{\circ}{W}$ and any metric g such that $g_1 = g_2$ outside K, we have $\Gamma^{g_1}_{\widetilde{W}} = \Gamma^{g_2}_{\widetilde{W}}$.*

By definition, $\mathrm{id}_\Gamma \in \Gamma^g_{\widetilde{W}}$, and $\gamma \in \Gamma^g_{\widetilde{W}}$ iff $\gamma^{-1} \in \Gamma^g_{\widetilde{W}}$. When (M, g) is a geometrically finite manifold, for suitable choice of W, $\Gamma^g_{\widetilde{W}}$ is a union of groups. But in general, $\Gamma^g_{\widetilde{W}}$ is not a group at all, as shown in the following proposition.

Proposition 7.8. – *With the previous notations, let $W \subset M$ be a compact pathwise connected set with piecewise $\mathcal{C}^1$ boundary and $\widetilde{W}$ be a nice preimage of W. If $\gamma \in \Gamma_{\widetilde{W}}$ is a hyperbolic element whose axis A_γ intersects the interior of $\widetilde{W}$, then there exists $N = N(\gamma) > 0$ such that for all $n \geq N$, $\gamma^n \notin \Gamma^g_{\widetilde{W}}$.*

Proof. – Let $\gamma \in \Gamma^g_{\widetilde{W}}$ be such an hyperbolic element. Its axis A_γ intersects $\widetilde{W}$, and therefore also $\gamma\widetilde{W}$ and all iterates $\gamma^n\widetilde{W}$. Choose some $x_0 \in A_\gamma \cap \overset{\circ}{\widetilde{W}}$ and let $d_0 = d^g(x_0, \partial\widetilde{W}) > 0$. Let $x, y \in \widetilde{W}$. By Lemma 5.1 (2), with $K = \mathrm{diam}(\widetilde{W})$, $\alpha = d_0/2$, we know that if $d^g(x, \gamma^n y) = n\ell^g(\gamma) \pm 2\mathrm{diam}\widetilde{W} \geq 2R_0$, all points in the middle interval of length $2T = \ell^g(\gamma)$ of the g-geodesic segment from x to $\gamma^n y$ would be at distance less than $d_0/2$ from A_γ, and therefore some of them would be inside $\gamma^k\widetilde{W}$, for some $1 \leq k \leq n-1$. This proves the proposition. □

The set $\Gamma_{\widetilde{W}}$ depends on W and the choice of its preimage $\widetilde{W}$, but not too strongly as illustrated by the following proposition.

Proposition 7.9. – 1. *Let $W \subset M$ be a compact set (with piecewise $\mathcal{C}^1$ boundary), and $\widetilde{W}$ be a nice preimage. Let $\alpha \in \Gamma$. Then $\Gamma^g_{\alpha\widetilde{W}} = \alpha\Gamma^g_{\widetilde{W}}\alpha^{-1}$.*

2. *If W_1 and W_2 are compact sets of M (with piecewise $\mathcal{C}^1$ boundary) such that $W_1 \subset \overset{\circ}{W_2}$ with respective nice preimages $\widetilde{W}_1 \subset \widetilde{W}_2$, there exists $k \geq 1$ and $\alpha_1, \ldots, \alpha_k \in \Gamma$ such that*

$$\Gamma_{\widetilde{W}_2} \subset \bigcup_{i,j=1}^{k} \alpha_i \Gamma_{\widetilde{W}_1}(\alpha_j)^{-1}.$$

3. *If $\widetilde{W}_1$ and $\widetilde{W}_2$ are nice preimages of W, then there exists a finite set $\{\alpha_1, \ldots, \alpha_k\} \subset \Gamma$ such that*

$$\Gamma_{\widetilde{W}_2} \subset \bigcup_{i,j=1}^{p} \alpha_i \Gamma_{\widetilde{W}_1}(\alpha_j)^{-1}.$$

Proof. – The first item of the proposition is obvious. Let us show 2. Set

$$D = 2\mathrm{diam}(W_2) \quad \text{and} \quad \eta = \inf\{d^g(w, \partial W_2)\ ;\ w \in W_1\} > 0.$$

Let $\gamma \in \Gamma^g_{\widetilde{W}_2}$. There exist $x_2, y_2 \in \partial\widetilde{W}_2$ such that the g-geodesic segment $[x_2, \gamma y_2]$ intersects $\Gamma\widetilde{W}_2$ only in $\widetilde{W}_2 \cup \gamma\widetilde{W}_2$. Now, choose some $x_1, y_1 \in \partial\widetilde{W}_1$.

By Lemma 5.1, there exists $L = L(D,\eta) > 0$ and $R = R(D,\eta) > 2L$ such that for all $x_1, y_1, x_2, y_2 \in \widetilde{M}$ with $d^g(x_1,x_2) \leq D$, $d^g(y_1,y_2) \leq D$ and $d^g(x_2,y_2) \geq R$, the g-geodesic segment (x_1,y_1) is contained in the $\frac{\eta}{2}$-neighborhood of (x_2,y_2) except inside the balls $B^g(x_1,L)$ and $B^g(y_1,L)$.

Let $\alpha_1,\dots,\alpha_k \in \Gamma$ be the (finitely many) elements such that

$$d^g(\widetilde{W}_1, \alpha_i\widetilde{W}_1) = \inf\{d^g(a,b),\ a \in \widetilde{W}_1, b \in \alpha_i\widetilde{W}_1\} \leq L.$$

Let $\widetilde{W}_1 \subset \widetilde{W}_2$ be included nice preimages of W_1 and W_2, and let $\gamma \in \Gamma_{\widetilde{W}_2}$ such that $d^g(o,\gamma o) \geq R + 2D$. Then there exists $x_2, y_2 \in \partial\widetilde{W}_2$ such that $(x_2, \gamma y_2)$ does not intersect $p_\Gamma^{-1}(W_2)$. By construction there exists $x_1, y_1 \in \widetilde{W}_1$ such that $d^g(x_1,x_2) \leq D$ and $d^g(y_1,y_2) \leq D$. The geodesic $(x_1,\gamma y_1)$ is $\frac{\eta}{2}$-close to the geodesic $(x_2,\gamma y_2)$ outside the balls $B^g(x_1,L)$ and $B^g(\gamma y_1, L)$, hence does not intersect $p_\Gamma^{-1}(W_1)$ except maybe in these balls. Thus, there exist α_i, α_j in the above finite set, such that the geodesic segment $(x_1,\gamma y_1)$ does not intersect $\Gamma\widetilde{W}_1$ between $\alpha_i\widetilde{W}_1$ and $\gamma\alpha_j\widetilde{W}_1$. Therefore, $\alpha_i^{-1}\gamma\alpha_j \in \Gamma_{\widetilde{W}_1}$ or in other words,

$$\gamma \in \alpha_i \Gamma_{\widetilde{W}_1} \alpha_j^{-1}.$$

The proof of the last item is similar, and we let it to the reader. □

7.2. Entropy at infinity

Proposition 7.10. – *Let $W \subset M$ be a compact set and $\widetilde{W}$ a nice preimage of W. The critical exponent $\delta_W(g)$ of the Poincaré series $\sum_{\gamma\in\Gamma_{\widetilde{W}}} e^{-sd^g(o,\gamma o)}$ is equal to*

$$\delta_W(g) = \limsup_{R\to\infty} \frac{\log \#\left\{\gamma \in \Gamma_{\widetilde{W}},\ R-1 < d^g(o,\gamma o) \leq R\right\}}{R}$$

and does not depend on the choice of a nice preimage $\widetilde{W} \subset \widetilde{M}$ of W nor $o \in \widetilde{M}$. We call it the entropy out of W *of (M,g).*

Proof. – It follows from the triangular inequality that $\delta_W(g)$ does not depend on the choice of o. Let us show that it does not depend on the choice of preimage. Let $\widetilde{W}_1$ and $\widetilde{W}_2$ be two nice preimages of W. By Proposition 7.9, there exists $k \geq 0$ and $\alpha_1,\dots,\alpha_k \in \Gamma$ such that

$$\Gamma_{\widetilde{W}_2} \subset \bigcup_{i,j=1}^{k} \alpha_i \Gamma_{\widetilde{W}_1} \alpha_j^{-1}. \tag{24}$$

Set

$$D = \max\left\{ d^g(w,o)\ ;\ w \in \widetilde{W}_2 \cup \bigcup_{i,j=1}^{k} \alpha_i\Gamma_{\widetilde{W}_1}(\alpha_j)^{-1}\right\}.$$

Define for $i = 1, 2$ and $R > 0$,

$$\Gamma_{\widetilde{W}_i}(R) = \left\{\gamma \in \Gamma_{\widetilde{W}_i}\ ;\ d^g(o,\gamma o) \leq R\right\}.$$

It follows from (24) and triangular inequality that for all $R > 0$,

$$\Gamma_{\widetilde{W_2}}(R) \subset \bigcup_{i,j=1}^{k} \alpha_i \Gamma_{\widetilde{W_1}}(R+2D)(\alpha_j)^{-1},$$

and therefore $\# \Gamma_{\widetilde{W_2}}(R) \leq k^2 \# \Gamma_{\widetilde{W_1}}(R+2D)$. This gives immediately

$$\limsup_{R\to+\infty} \frac{1}{R} \log \#\Gamma_{\widetilde{W_2}}(R) \leq \limsup_{R\to+\infty} \frac{1}{R} \log \#\Gamma_{\widetilde{W_1}}(R).$$

By symmetry, the reverse inequality also holds, and the result follows. □

Proposition 7.11. – *Let (M, g) be a complete negatively curved metric.*

1. *For any proper compact subset $K \subset \mathring{W}$ and any metric g_2 such that $g_1 = g_2$ outside K, we have $\delta_W(g_1) = \delta_W(g_2)$.*
2. *For all compact sets W_1, W_2 such that $W_1 \subset \mathring{W_2} \subset M$, we have $\delta_{W_1} \geq \delta_{W_2}$.*

Proof. – Item 1 follows from Proposition 7.7. Item 2 can be proven similarly to Proposition 7.10, thanks to Proposition 7.9. □

For a global variation of the metric (i.e., beyond W), even small, the behavior of $\delta_W(g)$ is not clear since the set $\Gamma_{\widetilde{W}}$ depends on the metric.

Definition 7.12. – *The* entropy at infinity *of (M, g) is*

$$\delta_\infty(g) = \inf\{\, \delta_W(g),\ W \subset M \quad \textit{compact set}\, \}.$$

Proposition 7.11 implies the following natural characterization of the entropy at infinity.

Proposition 7.13. – *Let (M, g) be a complete negatively curved manifold and $(W_i)_{i\in\mathbb{N}}$ be an increasing exhaustion of M by compact sets. Then*

$$\delta_\infty(g) = \lim_{i\to\infty} \delta_{W_i}(g).$$

Moreover, it is invariant under any negatively curved perturbation of the metric with compact support.

This entropy at infinity is a dynamical analogous to the bottom of the essential spectrum of the Laplacian in spectral geometry. We will use this fact in some of the examples given in Section 7.3.

Definition 7.14. – *The complete manifold (M, g) is called* strongly/stably positively recurrent (SPR)*, if $\delta_\infty(g) < \delta_\Gamma(g)$. We will also call this property a* critical gap at infinity.

By definition, if (M, g) is strongly positively recurrent, there exists a compact set $W \subset M$ such that $\delta_W < \delta_\Gamma$.

REMARK 7.15. – *The reader may have noticed that the definition of* $\Gamma_{\widetilde{U}}$ *given in* [41, p. 4] *is slightly different from ours, since it is written for an open set* $\widetilde{U}$ *which projects onto* $U = \overset{\circ}{W}$. *Nevertheless, these definitions almost coincide in the following sense. Let* $W \subset M$ *be a compact set with nice preimage* $\widetilde{W} \subset \widetilde{M}$, *let* $\widetilde{U} \subset \widetilde{M}$ *be an open set which projects onto* $U = \overset{\circ}{W}$. *Let* $\Gamma_{\widetilde{U}}$ *be defined as in* [41], *and* $\Gamma_{\widetilde{W}}$ *be defined as above. Then there exists* $\alpha_1, \ldots, \alpha_k \in \Gamma$ *such that*

$$\Gamma_{\widetilde{U}} \subset \bigcup_{i,j=1}^{k} \alpha_i \Gamma_{\widetilde{W}} (\alpha_j)^{-1} \quad \text{and} \quad \Gamma_{\widetilde{W}} \subset \bigcup_{i,j=1}^{k} \alpha_i \Gamma_{\widetilde{U}} (\alpha_j)^{-1}.$$

Therefore $\Gamma_{\widetilde{U}}$ *and* $\Gamma_{\widetilde{W}}$ *have the same critical exponent and all results stated in* [41] *to characterize the finiteness of Gibbs measures in terms of* $\Gamma_{\widetilde{U}}$ *are also valid for our definition of* $\Gamma_{\widetilde{W}}$.

7.3. Examples of SPR manifolds

We present here three classes of SPR manifolds. The first examples are geometrically finite manifolds with critical gap studied in [15]. Schottky products furnish also plenty of examples, generalizing the examples of [39]. At last, we describe examples inspired by Ancona's examples in [2].

These examples are almost the only known examples of non-compact manifolds with finite Bowen-Margulis measure. To our knowledge, the only exception is a construction of Peigné of geometrically finite manifolds with finite Bowen-Margulis measure but without critical gap, see [40, 54].

7.3.1. *Geometrically finite manifolds with critical gap.* – The convex core $CC(M) \subset M$ is the image on M of the convex hull of the limit set Λ_Γ inside $\widetilde{M}$. The nonwandering set $\Omega \subset S^g M$ of the geodesic flow is the set of vectors $v \in S^g M$ such that $v^\pm \in \Lambda_\Gamma$. By definition, $\Omega \subset S^g CC(M)$. A parabolic subgroup $\mathscr{P}$ of Γ is a subgroup which fixes a point at infinity, and therefore stabilizes any horoball $\mathscr{H}$ centered at this point.

A cusp is the image on M of such a horoball.

The manifold M is *geometrically finite* if its convex core can be written as a finite union

$$CC(M) = C_0 \sqcup C_1 \sqcup \cdots \sqcup C_K,$$

where C_0 is a compact set and the C_i are finitely many cusps, images through p_Γ of horoballs $\mathscr{H}_i$ stabilized by parabolic subgroups $\mathscr{P}_i$ of Γ. The complete reference on such manifolds is [7]. Parabolic subgroups have a positive critical exponent. The preimage on $\widetilde{M}$ of a cusp C_i is the orbit of a horoball $\mathscr{H}_i$, and the stabilizer of any horoball $\gamma\mathscr{H}_i$ is conjugated to the stabilizer $\mathscr{P}_i$ of $\mathscr{H}_i$ in Γ.

A *convex-cocompact manifold* is a geometrically finite manifold without cusps; in other words, it is a manifold whose convex core is compact.

PROPOSITION 7.16. – *Let* (M, g) *be a manifold with pinched negative curvature. If* (M, g) *is convex-cocompact, its entropy at infinity is* $-\infty$. *If* (M, g) *is geometrically finite with* k *cusps represented by parabolic subgroups* $\mathscr{P}_1, \ldots, \mathscr{P}_k \subset \Gamma$, *then*

$$\delta_\infty(g) = \max\left\{\delta_{\mathscr{P}_1}(g), \ldots, \delta_{\mathscr{P}_k}(g)\right\}.$$

In particular, a geometrically finite manifold is strongly positively recurrent if and only if

$$\max\left\{\delta_{\mathscr{P}_1}(g),\ldots,\delta_{\mathscr{P}_k}(g)\right\}<\delta_\Gamma.$$

This condition is precisely the *critical gap* criterion introduced by Dalbo, Otal and Peigné in [15]. It is satisfied in particular by locally symmetric geometrically finite manifolds and their small compact $\mathcal{C}^2$ perturbations. The notion of SPR manifold allows to generalize many results of [15] and others on geometrically finite manifolds to all strongly positively recurrent manifolds.

Proposition 7.16 follows immediately from Proposition 7.17 below.

PROPOSITION 7.17. – *Let (M,g) be a manifold with pinched negative curvature.*

If (M,g) is convex-cocompact and W is a compact set such that $CC(M)\subset\overset{\circ}{W}$, then $\Gamma_{\widetilde{W}}$ is finite.

If (M,g) is geometrically finite with k cusps, then there exists a compact set $W\subset M$ with nice preimage $\widetilde{W}$, a finite set $\Gamma^0_{\widetilde{W}}$, finitely many elements $\alpha_1,\ldots,\alpha_N\in\Gamma$, and parabolic subgroups $\mathscr{P}_1,\ldots,\mathscr{P}_k\subset\Gamma$ such that

$$\Gamma^g_{\widetilde{W}}=\Gamma^0_{\widetilde{W}}\cup\bigcup_{i,j}\alpha_i\left(\mathscr{P}_1\cup\cdots\cup\mathscr{P}_k\right)\alpha_j^{-1}.$$

Proof. – Assume first that (M,g) is convex-cocompact and $CC(M)\subset\overset{\circ}{W}$. Let D be the diameter of W and $\eta=\inf\{d^g(w,\partial W)\ ;\ w\in \mathrm{CC}(M)\}>0$. Let $\gamma\in\Gamma^g_{\widetilde{W}}$, $x,y\in\partial\widetilde{W}$ and choose $x_1,y_1\in\widetilde{CC(M)}$ such that $d(x,x_1)\le D$ and $d(y,y_1)\le D$. By Lemma 5.1, there exists some R_0 depending on D,η such that if $\ell^g(\gamma)\ge R_0$, there exists some $z\in(x,\gamma y)$, $z_1\in(x_1,\gamma y_1)$ such that $d^g(z,z_1)\le\eta/2$. But $\widetilde{CC(M)}$ is convex, so that $z_1\in\widetilde{CC(M)}$ and z is at distance $\eta/2$ of $\widetilde{CC(M)}$ and therefore inside $\Gamma\widetilde{W}$. Thus, $\gamma\notin\Gamma_{\widetilde{W}}$. Therefore, all elements of $\Gamma_{\widetilde{W}}$ have bounded length less than R_0, so that $\Gamma_{\widetilde{W}}$ is finite, included in $\{\gamma\in\Gamma,\ell^g(\gamma)\le R_0\}$.

Assume now that M is geometrically finite with cusps, and let $CC(M)=C_0\sqcup(\bigsqcup_{i=1}^k C_i)$ be a decomposition of the convex core into a compact part and finitely many disjoint cusps. Let $W\subset M$ be a compact set such that $\overset{\circ}{W}\supset CC(M)$. Choose some nice preimage $\widetilde{W}$ and disjoint horoballs $\mathscr{H}_i$, $1\le i\le k$ whose boundary intersects $\widetilde{W}$. Let $\mathscr{P}_i$ be the stabilizer of $\mathscr{H}_i$ in Γ.

Let $\gamma\in\Gamma_{\widetilde{W}}$ be such that $\ell^g(\gamma)\ge R_0$ and $x,y\in\partial\widetilde{W}$. As noticed above, by Lemma 5.1, the geodesic segment $(x,\gamma y)$ is (except at the beginning and the end, inside balls $B^g(x,L)$ and $B^g(\gamma y,L)$) in the $\eta/2$ neighborhood of $\widetilde{CC(M)}$. As already said in [41], if $\gamma\in\Gamma_{\widetilde{W}}$, except for a bounded amount of time at the beginning and the end, the geodesic segment $p_\Gamma(x,\gamma y)$ has to leave the compact part C_0 and enter in some cusp C_i. Therefore, there exists a finite set $\{\alpha_1,\ldots\alpha_N\}$ such that for some $1\le i,j\le N$, the geodesic segment $(\alpha_i x,\gamma\alpha_j y)$ stays in some horoball $\mathscr{H}_l$. As in the proof of Proposition 7.9, one deduces that $\Gamma_{\widetilde{W}}\subset\Gamma^0_{\widetilde{W}}\cup\bigcup_{i,j,l}\alpha_l\mathscr{P}_i\alpha_j^{-1}$ with $\Gamma^0_{\widetilde{W}}\subset\{\gamma\in\Gamma,\ell^g(\gamma)\le R_0\}$ as in the convex-cocompact case. □

7.3.2. *Schottky products.* – We present now a family of geometrically infinite examples first studied in [39]. Let G and H be discrete groups of isometries of a complete manifold $(\widetilde{M}, g)$ with pinched negative curvature. They are *in Schottky position* if there exist disjoint compact sets $U_G, U_H \subset \widetilde{M} \cup \partial\widetilde{M}$ such that for all $g \in G\backslash\{\mathrm{id}\}$ and all $h \in H\backslash\{\mathrm{id}\}$, we have

$$g\left((\widetilde{M} \cup \partial\widetilde{M})\backslash U_G\right) \subset U_G \quad \text{and} \quad h\left((\widetilde{M} \cup \partial\widetilde{M})\backslash U_H\right) \subset U_H.$$

In particular, by Klein's ping-pong argument, they generate a free product: $\Gamma = \langle G, H\rangle = G * H$. The entropy at infinity behaves nicely under Schottky products, as shown by the following theorem.

THEOREM 7.18. – *Let G and H be discrete groups of isometries of a complete manifold $(\widetilde{M}, g)$ with pinched negative curvature which are in Schottky position. Let $\Gamma = \langle G, H\rangle = G * H$. Denote respectively by $M_\Gamma = \widetilde{M}/\Gamma$, $M_G = \widetilde{M}/G$ and $M_H = \widetilde{M}/H$ the associated quotient manifolds endowed with the quotient metric induced by g. Then*

$$\delta_\infty(M_\Gamma) = \max\{\delta_\infty(M_G), \delta_\infty(M_H)\}.$$

As an immediate corollary, we get the following result.

COROLLARY 7.19. – *Let G and H be discrete groups of isometries of a complete manifold $(\widetilde{M}, g)$ with pinched negative curvature which are in Schottky position. Let M_G, M_H, and M_{G*H} be the quotient manifolds. Their critical exponents satisfy*

$$\delta_{G*H} \geq \max\{\delta_G, \delta_H\} \geq \max\{\delta_\infty(M_G), \delta_\infty(M_H)\} = \delta_\infty(M_{G*H}). \tag{25}$$

In particular,

1. *if G and H are Strongly Positively Recurrent, then $G * H$ is also;*
2. *if $\delta_{G*H} > \max\{\delta_\infty(M_G), \delta_\infty(M_H)\}$, then $G * H$ is strongly positively recurrent.*

In both cases $(\widetilde{M}/\Gamma, g)$ has a finite Bowen-Margulis measure.

It was originally shown by M. Peigné in [40] that if $\delta_\Gamma > \max\{\delta_G, \delta_H\}$ then $(\widetilde{M}/\Gamma, g)$ has a finite Bowen-Margulis measure. The above corollary with Theorem 7.1 guarantees this finiteness under a weaker condition.

It was shown in [15] that if $G \subset \Gamma$ is a *divergent* subgroup, then $\delta_G < \delta_\Gamma$. We get therefore the following corollary.

COROLLARY 7.20. – *Let G, H be discrete* divergent *groups of isometries of a complete manifold $(\widetilde{M}, g)$ with pinched negative curvature which are in Schottky position. Then $\Gamma = \langle G, H\rangle = G * H$ is strongly positively recurrent.*

This last corollary allows a lot of topologically infinite examples. For instance, if G and H are discrete subgroups of the group of isometries of the hyperbolic space, whose limit sets are not the whole boundary, they can be settled in Schottky position by taking suitable conjugation with hyperbolic elements. If G and H are $\mathbb{Z}$-covers of convex-cocompact groups, they are divergent and their Schottky product gives a SPR manifold, hence with finite Bowen-Margulis measure, whose fundamental group is not even finitely generated.

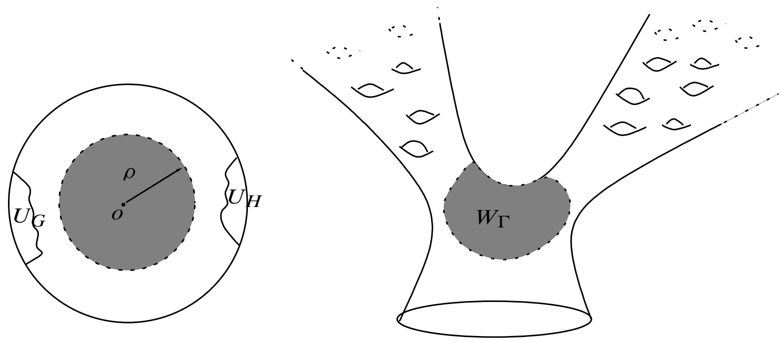


FIGURE 6. Schottky manifold

Proof. – Our proof relies on the ideas of Section 9 of [41]. Let U_G and U_H be the sets ensuring the Schottky position of G and H. Since they are compact in $\widetilde{M} \cup \partial\widetilde{M}$ and since $K_g \leq -a^2 < 0$, a key point is that there exists $\rho > 0$ such that all geodesics from U_G to U_H intersect the ball $B^g(o,\rho)$. Moreover, without loss of generality, we can assume that the point o is neither in U_G nor in U_H.

Let $M_\Gamma = \widetilde{M}/\Gamma$, $M_G = \widetilde{M}/G$ and $M_H = \widetilde{M}/H$. Let $p_\Gamma : \widetilde{M} \to M_\Gamma$, $p_G : \widetilde{M} \to M_G$ and $p_H : \widetilde{M} \to M_H$ be the associated covering maps.

For all $R \geq \rho$, define $W_\Gamma^R = p_\Gamma(B^g(o,R)) \subset M_\Gamma$, $W_G^R = p_G(B^g(o,R)) \subset M_G$ and $W_H^R = p_H(B^g(o,R)) \subset M_H$. Let $\widetilde{W}_\Gamma^R, \widetilde{W}_G^R, \widetilde{W}_H^R \subset \widetilde{M}$ be nice preimages (Dirichlet domains viewed from o) of W_Γ^R, W_G^R, W_H^R, respectively for the actions of Γ, G, H. By definition of W_Γ^R, W_G^R, W_H^R and of a Dirichlet domain, one easily checks that they all lie inside $B^g(o,R)$. Moreover, as p_G and p_H are intermediate covers between $\widetilde{M}$ and M_Γ, we have $o \in \widetilde{W}_\Gamma^R \subset \widetilde{W}_G^R \cap \widetilde{W}_H^R \subset B^g(o,R)$.

Let $\Gamma_{\widetilde{W}_\Gamma^R} \subset \Gamma$, $G_{\widetilde{W}_G^R} \subset G$ and $H_{\widetilde{W}_H^R} \subset H$ be the fundamental groups respectively of Γ, G and H respectively out of $\widetilde{W}_\Gamma^R, \widetilde{W}_G^R, \widetilde{W}_H^R$, according to Definition 7.6.

A key fact is the following.

LEMMA 7.21 ([41]). – *For all $R > 0$, there exists a finite set $S \subset \Gamma$ such that*

$$\Gamma_{\widetilde{W}_G^R} \subset G \cup H \cup S.$$

It implies that $\delta_\infty(M_\Gamma) \leq \delta_{W_\Gamma^R}(\Gamma) \leq \max\{\delta_G, \delta_H\}$, and therefore, if $\delta_\Gamma > \max\{\delta_G, \delta_H\}$ then Γ is strongly positively recurrent.

We precise this inclusion in the following lemma, which implies immediately Theorem 7.18.

LEMMA 7.22. – *With the previous notations, for all $R \geq \rho + 1$, there exists a finite set F such that*

$$\Gamma_{\widetilde{W}_\Gamma^{2R}} \subset S \cup G_{\widetilde{W}_G^{2R}} \cup H_{\widetilde{W}_H^{2R}} \subset S \cup F \cup \bigcup_{\alpha,\beta \in F} \alpha \Gamma_{\widetilde{W}_\Gamma^R} \beta.$$

Proof. – Let us first show the left inclusion. It follows from the previous lemma that $\Gamma_{\widetilde{W}_\Gamma^{2R}} \subset G \cup H \cup S$. Moreover,

$$G\widetilde{W}_G^{2R} = GB^g(o,2R) \subset \Gamma\widetilde{W}_\Gamma^{2R} = \Gamma B^g(o,2R).$$

For each $\gamma \in \Gamma_{\widetilde{W}_\Gamma^{2R}} \cap G$, there exist $x, y \in \widetilde{W}_\Gamma^{2R} \subset \widetilde{W}_G^{2R}$ such that

$$[x,\gamma y] \cap \Gamma.\widetilde{W}_\Gamma^{2R} \subset \widetilde{W}_\Gamma^{2R} \cup \gamma.\widetilde{W}_\Gamma^{2R}, \quad \text{whence} \quad [x,\gamma y] \cap G\widetilde{W}_G^{2R} \subset \widetilde{W}_G^{2R} \cup \gamma\widetilde{W}_G^{2R},$$

so that $\gamma \in G_{\widetilde{W}_G^{2R}}$. It shows that $\Gamma_{\widetilde{W}_\Gamma^{2R}} \cap G \subset G_{\widetilde{W}_G^{2R}}$.

Similarly, $\Gamma_{\widetilde{W}_\Gamma^{2R}} \cap H \subset H_{\widetilde{W}_H^{2R}}$.

Let us now prove the right inclusion. We want to show that there exists a finite set $F \subset \Gamma$ such that $G_{\widetilde{W}_G^{2R}} \subset F \cup \bigcup_{\alpha,\beta\in F} \alpha\Gamma_{\widetilde{W}_\Gamma^R}\beta$, the case of $H_{\widetilde{W}_H^{2R}}$ being similar.

Define $F_{\lambda R}$ as $F_{\lambda R} = \{\gamma \in \Gamma, \gamma B^g(o,\lambda R) \cap B^g(o,\lambda R) \neq \emptyset\}$.

First observe that for $\lambda \geq 2$, we have

$$\widetilde{W}_\Gamma^{2R} \subset \widetilde{W}_G^{2R} \subset \bigcup_{\alpha\in F_{\lambda R}} \alpha.\widetilde{W}_\Gamma^R. \tag{26}$$

Let $g \in G_{\widetilde{W}_G^{2R}}$, $g \notin F$. By definition, there exist $x, y \in \widetilde{W}_G^{2R}$ such that (x, gy) intersects $G.\widetilde{W}_G^{2R}$ only in $\widetilde{W}_G^{2R}$ and $g\widetilde{W}_G^{2R}$. We will show that (x, gy) intersects $\Gamma.\widetilde{W}_\Gamma^R = \Gamma.\widetilde{W}_G^R$ only inside $\widetilde{W}_G^{2R}$ and $g\widetilde{W}_G^{2R}$. By equation (26), as in the proof of Proposition 7.9, it will imply that $g \in \bigcup_{\alpha,\beta\in F_{\lambda R}} \alpha\Gamma_{\widetilde{W}_\Gamma^R}\beta$. In fact, we will show that if (x, gy) intersects some $\gamma.\widetilde{W}_\Gamma^R$, then either γ or $g^{-1}\gamma$ is in the finite set F, so that by the same argument, $g \in \bigcup_{\alpha,\beta\in F_{\lambda R}} \alpha\Gamma_{\widetilde{W}_\Gamma^R}\beta$.

By contradiction, assume that the geodesic segment (x, gy) intersects $\gamma\widetilde{W}_\Gamma^R$, with $\gamma \neq id, g$, and $\gamma, g^{-1}\gamma \notin F_{\lambda R}$. In particular, $d(o,\gamma o) > 2\lambda R$ and $d(go,\gamma o) > 2\lambda R$. As $g \in G_{\widetilde{W}_G^{2R}}$, we know that $\gamma \notin G$. Denote by $z_\gamma \in (x, gy)$ the closest point to γo in $(x,y) \cap \widetilde{W}_\Gamma^R$. By the above, we have $d(x,z_\gamma) \geq d(o,\gamma o) - 3R \geq (2\lambda - 3)R$.

By definition of a Schottky product, as $o \notin U_G \cup U_H$, either $\gamma o \in U_G$ or $\gamma o \in U_H$. Assume first that $\gamma o \in U_H$. Recall that $go \in U_G$. Therefore, the geodesic segment $(\gamma o, go)$ intersects the ball $B(o,\rho)$. As $d(\gamma o, go) \geq 2\lambda R$ and $d(\gamma o, z_\gamma) \leq R$, $d(go, gy) \leq 2R$, the geodesic segment (z_γ, gy) intersects the ball $B(o,\rho + 2R)$. Let w_γ be a point in this intersection. Therefore, we get $d(x,w_\gamma) \leq d(x,o) + d(o,w_\gamma) \leq 4R + \rho \leq 5R$. However, $d(x,w_\gamma) \geq d(x,z_\gamma) > (2\lambda - 3)R$, which leads to a contradiction as soon as $\lambda \geq 4$.

Therefore, the first case holds, $\gamma o \in U_G$, so that γ has a reduced form as $\gamma = g'h'\gamma'$, with $g' \in G \setminus \{id\}$, $h' \in H \setminus \{id\}$, $\gamma' \in \Gamma$. We will distinguish the cases $g' \in F$ and $g' \notin F$.

If $g' \notin F$, consider the segment $[(g')^{-1}o, h'\gamma' o]$. It goes from U_G to U_H so that it intersects the ball $B^g(o,\rho)$. It follows that $[o,\gamma o]$ intersects $g'.B(o,\rho)$ at a point y with $d(o,y) \geq 2\lambda R - \rho$. By Lemma 5.1, for λ large enough, the point y is at distance less than ρ from the geodesic segment $(x,\gamma o)$, and therefore at distance less than $R+1$ from the geodesic segment (x, z_γ). Thus, we deduce that (x, z_γ) intersects the ball $g'B(o,\rho + R + 1)$. As we assumed $R \geq \rho + 1$, this ball is included in $g'B(o,2R) \subset G.\widetilde{W}_G^{2R}$. Moreover, as $\gamma' \notin F$, this intersection $(x,z_\gamma) \cap g'B(o,2R)$ is disjoint from $\widetilde{W}_G^{2R}$, and as $\gamma \notin F$, and the intersection is between x and z_γ, this intersection is also disjoint from $g.\widetilde{W}_G^{2R}$. This is a contradiction with the hypothesis $g \in G_{\widetilde{W}_G^{2R}}$.

It remains the case $g' \in F$, which implies in particular $g' \neq g$. Consider in this case the geodesic segment $[h'\gamma' o, (g')^{-1} go]$. It goes from U_H to U_G, so that it intersects the ball $B^g(o,\rho)$. It follows that $[\gamma o, go]$ intersects $g'B(o,\rho)$. The same arguments on $[z_\gamma, gy]$ instead of $[x, z_\gamma]$ lead once again to a contradiction with the hypothesis $g \in G_{\widetilde{W}_G^{2R}}$.

It concludes the proof, for $F = F_{\lambda R}$, for some $\lambda \geq 4$ determined by the use of Lemma 5.1. $\square$

7.3.3. *Ancona-like examples*. – We present now a family of surfaces inspired by examples of Ancona [2], which is particularly easy to handle using the entropy at infinity introduced before.

THEOREM 7.23. – *Any non-elementary hyperbolic surface $S = \mathbb{H}^2/\Gamma$ with $\delta_\Gamma < 1$ admits a compact perturbation which is strongly positively recurrent: there exists a hyperbolic surface $S' = \mathbb{H}^2/\Gamma'$, homeomorphic to S, which is isometric to S outside a compact set and such that $\delta_\infty(S') < \delta_{\Gamma'}$.*

By Theorem 7.1, all these examples have finite Bowen-Margulis measure.

Note that for *topologically finite surfaces* (i.e., when Γ is finitely generated), this theorem is an immediate consequence of Proposition 7.16 since all topologically finite hyperbolic surfaces are geometrically finite with critical gap (hence SPR), see [15] for a proof. Theorem 7.23 is only interesting for *topologically infinite* hyperbolic surfaces $S = \mathbb{H}^2/\Gamma$ with $\delta_\Gamma < 1$. For instance, by [11], any nonamenable regular cover S of a compact hyperbolic surface S_0 satisfies these hypotheses.

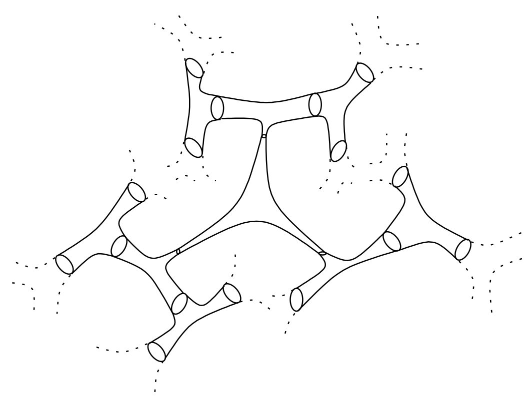


FIGURE 7. SPR surface

Before beginning the proof, recall that on hyperbolic manifolds, the dynamics is strongly related to the spectrum of the Laplacian. In particular, a well-known theorem of Patterson and Sullivan relates the entropy δ_Γ of $M = \mathbb{H}^{n+1}/\Gamma$ with the bottom of the spectrum of the Laplacian $\lambda_0(M)$:

THEOREM 7.24 ([37], [49], [50], [51]). – *Let $M = \mathbb{H}^{n+1}/\Gamma$ be a complete hyperbolic manifold. If $\delta_\Gamma < \frac{n}{2}$, then $\lambda_0(M) = \frac{n^2}{4}$. If $\delta_\Gamma > \frac{n}{2}$, then $\lambda_0(M) = \delta_\Gamma(n - \delta_\Gamma)$.*

Proof of Theorem 7.23. – Let $S = \mathbb{H}^2/\Gamma$ be a complete topologically infinite hyperbolic surface with $\delta_\Gamma < 1$. Denote by g_0 its metric. In any pair of pants decomposition of S, choose finitely many pairs of pants $P_1, \dots, P_K$. Change the metric of S to a metric g_ε, which is equal to g_0 far from the pants P_i, and modified in the neighborhood of the P_i by shrinking the lengths of the boundary geodesics of the pants P_i to a length ε. Let Γ_ε be a discrete group such that the new hyperbolic surface (S, g_ε) is isometric to $\mathbb{H}^2/\Gamma_\varepsilon$.

As the perturbation is compact, for all $\varepsilon > 0$, $\delta_\infty(g_\varepsilon) = \delta_\infty(g_0) < 1$. An elementary computation (see for example [12, Prop. II.2 (ii)]) gives $\lim_{\varepsilon\to 0} \lambda_0(S, g_\varepsilon) = 0$, therefore $\lim_{\varepsilon\to 0} \delta_{\Gamma_\varepsilon} = 1$. This implies that for $\varepsilon > 0$ small enough, (S, g_ε) has a critical gap at infinity : $\delta_{\Gamma_\varepsilon} > \delta_{\Gamma_0} \geq \delta_\infty(g_\varepsilon)$. □

7.4. SPR manifolds have finite Bowen-Margulis measure

This paragraph is devoted to the proof of the first part of Theorem 7.1: if (M, g) is a strongly positively recurrent manifold, then the Bowen-Margulis measure of its geodesic flow has finite mass.

This finiteness result had been shown in [15] on geometrically finite manifolds, under the assumption that $\max\left\{\delta_{\mathscr{P}_1}(g), \dots, \delta_{\mathscr{P}_k}(g)\right\} < \delta_\Gamma$, which is exactly the SPR assumption in the geometrically finite context, although they did not introduce this concept.

As said earlier, this result (finiteness of Bowen-Margulis measure) has been obtained independently, by a different approach, in [53].

Our proof will rely on the following theorem shown in [41].

THEOREM 7.25 ([41]). – *Let (M, g) be a complete manifold with negative curvatures. Then the Bowen-Margulis measure of (M, g) is finite if and only if $\Gamma = \pi_1(M)$ is divergent and there exists a compact set $W \subset M$ with nice preimage $\widetilde{W}$ such that $\Gamma_{\widetilde{W}}$ satisfies*

$$\sum_{\gamma\in\Gamma_{\widetilde{W}}} d(o, \gamma o) e^{-\delta_\Gamma d(o,\gamma o)} < +\infty.$$

Let (M, g) be a complete strongly positively recurrent manifold: there exists a compact set $W \subset M$ such that $\delta_W(g) < \delta_\Gamma(g)$. The second condition $\sum_{\gamma\in\Gamma_{\widetilde{W}}} d(o, \gamma o) e^{-\delta_\Gamma d(o,\gamma o)} < +\infty$ is then automatically satisfied for $\widetilde{W}$ a nice lift of W. Therefore, Theorem 7.1 follows immediately from the following.

THEOREM 7.26. – *Let (M, g) be a strongly positively recurrent manifold. Then its fundamental group Γ is divergent.*

We give first the strategy of the proof. Let (M, g) be a SPR manifold, with $\Gamma = \pi_1(M)$. It follows from Hopf-Tsuji-Sullivan theorem (see [46, p. 18]) that Γ is divergent if and only if any Patterson-Sullivan measure ν_o^g (cf Section 5.4) gives full measure to the radial limit set Λ_Γ^r.

Theorem 7.26 follows from a careful study of Λ_Γ^r. More precisely, if W is a nice set with $\delta_W < \delta_\Gamma$ and nice lift $\widetilde{W}$, we introduce a kind of limit set $\mathscr{L}_{W^c}$ of the subset $\Gamma_{\widetilde{W}}$ of Γ, see Definition 7.27 and Proposition 7.28. We show in Proposition 7.31 that $\nu_o^g(\mathscr{L}_{W^c}) = 0$. By definition, $\partial\widetilde{M} \setminus \mathscr{L}_{W^c}$ consists in asymptotic directions of geodesics returning infinitely often in the compact set W. In particular, it is included in the radial limit set. We deduce that $\nu_o^g(\Lambda_\Gamma^r) = 1$, which implies that Γ is divergent by Hopf-Tsuji-Sullivan Theorem.

Definition 7.27. – *Let $W \subset M$ be a compact subset and $\widetilde{W} \subset \widetilde{M}$ a nice preimage of W. Introduce the set*

$$\Lambda_{\widetilde{W}} = \left\{\xi \in \Lambda_\Gamma \textit{ s.t. } \exists x \in \widetilde{W}, [x,\xi) \cap \Gamma\widetilde{W} \subset \widetilde{W}\right\}.$$

We call the limit set of Γ out of W *the set* $\mathscr{L}_{W^c} = \Gamma\Lambda_{\widetilde{W}}$.

The following proposition shows that all elements of $\Lambda_{\widetilde{W}}$ are limit points of $\Gamma_{\widetilde{W}}o$ in the boundary at infinity, and that the only limit points of $\Gamma_{\widetilde{W}}o$ which are not in $\Lambda_{\widetilde{W}}$ are endpoints of geodesic rays which do not come back inside the interior $\Gamma\ \mathring{\widetilde{W}}$, after leaving $\widetilde{W}$ but touch the boundary $\partial(\Gamma\widetilde{W})$.

Proposition 7.28. – *Let $W \subset M$ be a compact subset and $\widetilde{W} \subset \widetilde{M}$ a nice preimage of W. Then*

$$\Lambda_{\widetilde{W}} \subset \overline{\Gamma_{\widetilde{W}}o}\setminus\Gamma_{\widetilde{W}}o \subset \left\{\xi \in \Lambda_\Gamma \textit{ s.t. } \exists x \in \widetilde{W}, [x,\xi) \cap \Gamma\ \mathring{\widetilde{W}} \subset \widetilde{W}\right\}.$$

Proof. – Without loss of generality, assume that $o \in \widetilde{W}$. We show first the left inclusion. Let $\xi \in \Lambda_{\widetilde{W}} \subset \Lambda_\Gamma$. There exists a sequence (γ_n) of elements of Γ such that $\gamma_n o \to \xi$. Moreover, by definition of $\Lambda_{\widetilde{W}}$, there exists $x \in \widetilde{W}$ such that the geodesic $[x,\xi)$ does not intersect $\Gamma\widetilde{W}$ after leaving $\widetilde{W}$. Thus, for n large enough, the geodesic segment $[x,\gamma_n o]$ also leaves $\widetilde{W}$ before returning to $\gamma_n\widetilde{W}$. Let $\widetilde{\gamma}_n\widetilde{W}$ be the first image of $\widetilde{W}$ crossed by the geodesic segment $[x,\gamma_n o]$ after leaving $\widetilde{W}$. By construction, $\widetilde{\gamma}_n \in \Gamma_{\widetilde{W}}$. Moreover, we have

$$\lim_{n\to+\infty} d^g(x,\widetilde{\gamma}_n o) = +\infty. \tag{27}$$

Indeed, for all $R > 0$, there exists $\eta > 0$ such that inside the (compact) ball $B^g(x,R)$, the distance between $[x,\xi) \cap B^g(x,R)$ and $\Gamma\widetilde{W}\setminus\widetilde{W}$ is at least η. Moreover, it follows from Lemma 5.1 that the sequence of geodesic segments $([x,\gamma_n o])_{n\in\mathbb{N}}$ converges to the half geodesic $[x,\xi]$ uniformly on $B^g(x,R)$. Thus, for all n large enough, $[x,\gamma_n o] \cap B^g(x,R)$ and $[x,\xi] \cap B^g(x,R)$ are $\frac{\eta}{2}$-close, so that $[x,\gamma_n o]$ does not meet $\Gamma\widetilde{W}\setminus\widetilde{W}$ on $B^g(x,R)$, whence $d^g(x,\widetilde{\gamma}_n o) \geq R$.

It follows from the above that the sequence of geodesic segments $([x,\widetilde{\gamma}_n o])_{n\in\mathbb{N}}$ also converges to the half geodesic $[x,\xi]$, so that

$$\xi = \lim_{n\to+\infty} \widetilde{\gamma}_n o \in \overline{\Gamma_{\widetilde{W}}o}\setminus\Gamma_{\widetilde{W}}o.$$

Let us now show that

$$\overline{\Gamma_{\widetilde{W}}o}\setminus\Gamma_{\widetilde{W}}o \subset \left\{\xi \in \Lambda_\Gamma \text{ s.t. } \exists x \in \widetilde{W}, [x,\xi) \cap \Gamma\ \mathring{\widetilde{W}} \subset \widetilde{W}\right\}.$$

Let $\xi \in \overline{\Gamma_{\widetilde{W}}o}\setminus\Gamma_{\widetilde{W}}o$. There exists a sequence $(\gamma_n)_{n\in\mathbb{N}}$ of elements of $\Gamma_{\widetilde{W}}$ such that $\gamma_n o \to \xi$ and $d^g(o,\gamma_n o) \to +\infty$. By definition of $\Gamma_{\widetilde{W}}$, for all $n \geq 0$ there exist $x_n, y_n \in \widetilde{W}$ such that the geodesic segment $[x_n,\gamma_n y_n]$ intersects $\Gamma\widetilde{W}$ only in $\widetilde{W}$ and $\gamma_n\widetilde{W}$. Up to taking a subsequence, we can assume that $x_n \to x_\infty \in \widetilde{W}$ and $y_n \to y_\infty \in \widetilde{W}$ as $n \to +\infty$. Once again, it follows from the compactness of $\widetilde{W}$ and Lemma 5.1 that the sequence of geodesic segments

$([x_\infty, \gamma_n y_\infty])_{n\in\mathbb{N}}$ converges to $[x_\infty, \xi]$ uniformly on compact sets. Therefore $[x_\infty, \xi)$ cannot intersect the interior of $\Gamma\widetilde{W}$. □

We gather in the following proposition elementary properties of the sets $\Lambda_{\widetilde{W}}$ and $\mathcal{L}_{W^c}$.

Proposition 7.29. – *Let (M, g) be a manifold with pinched negative curvature. Let $W \subset M$ be a nice compact set, and $\widetilde{W}$ a nice preimage. With the above notations,*

— *the set $\mathcal{L}_{W^c} = \Gamma\Lambda_{\widetilde{W}}$ is the set of endpoints of geodesics which eventually leave $\Gamma\widetilde{W}$:*

$$\mathcal{L}_{W^c} = \Gamma\Lambda_{\widetilde{W}} = \left\{v_+ \in \Lambda_\Gamma \ ; \ \exists v \in S^g\widetilde{M}, \exists T > 0 \ s.t. \ \forall t \geq T, \pi g^t v \notin \Gamma\widetilde{W}\right\}.$$

— *The* limit set out of W*, $\mathcal{L}_{W^c} = \Gamma\Lambda_{\widetilde{W}}$ does not depend on the choice of a nice preimage $\widetilde{W}$.*

— *If $W_1 \subset W_2$, we have $\mathcal{L}_{W_2^c} \subset \mathcal{L}_{W_1^c}$.*

— *$\Lambda_\Gamma \backslash \left(\mathcal{L}_{W^c}\right) \subset \Lambda_\Gamma^r$, where Λ_Γ^r is the radial limit set.*

Proof. – The first property is left to the reader.

The set $\left\{v_+ \in \Lambda_\Gamma \ ; \ \exists v \in S^g\widetilde{M}, \exists T > 0 \text{ s.t. } \forall t \geq T, \pi g^t v \notin \Gamma\widetilde{W}\right\}$ only depends on $\Gamma\widetilde{W} = p_\Gamma^{-1}(W)$, which is independent of the choice of $\widetilde{W}$.

If $W_1 \subset W_2$, then for all nice preimages $\widetilde{W}_1$ and $\widetilde{W}_2$, we have

$$\Gamma\widetilde{W}_1 = p_\Gamma^{-1}(W_1) \subset p_\Gamma^{-1}(W_2) = \Gamma\widetilde{W}_2,$$

which shows the third point.

The radial limit set is the set of $\xi \in \Lambda_\Gamma$ such that there exists $x \in \widetilde{M}$ and a compact set $K \subset M$ such that the geodesic ray $[x, \xi)$ intersects infinitely often the preimage $p_\Gamma^{-1}(K)$. If $\xi \in \Lambda_\Gamma \backslash \left(\Lambda_{\widetilde{W}}\right)$, by the above proposition, the geodesic ray $[x, \xi]$ intersects infinitely often the set $\Gamma\widetilde{W} = p_\Gamma^{-1}(W)$, which shows the last claim. □

As seen in Section 7.3, basic examples are given by geometrically finite manifolds. The following proposition is an immediate consequence of Propositions 7.17 and 7.28.

Proposition 7.30. – *Let (M, g) be a geometrically finite manifold with pinched negative curvature, with k cusps $C_1, \ldots, C_k$. Let $W = B^g(x, R)$ be a large ball. It admits a nice preimage $\widetilde{W}$ such that*

$$\Lambda_{\widetilde{W}} = \{\xi_1, \ldots, \xi_i\},$$

each point $\xi_1, \ldots, \xi_k \in \Lambda_\Gamma$ being a parabolic point fixed by a parabolic group $\mathcal{P}_i < \Gamma$ representing the cusp C_i.

The following proposition is a detailed version of Theorem 7.26, with additional properties which will be useful in Section 7.5.

If $x \in \widetilde{M}$ and $\xi \in \widetilde{M} \cup \partial\widetilde{M}$, we write $[x, \xi]_T = (\pi g^t v)_{t\in[0,T]}$, where $v \in S_x^g\widetilde{M}$ is the tangent vector at x of the geodesic $[x, \xi]$.

PROPOSITION 7.31. – *Let* $(M = \widetilde{M}/\Gamma, g)$ *be a complete manifold with* $K_g \leq -a^2$*, with* $\Gamma = \pi_1(M)$ *its fundamental group. Assume that* (M, g) *is SPR. Then* Γ *is divergent.*

Moreover, for all compact sets $W \subset M$ *such that* $\delta_W < \delta_\Gamma$*, for all* $\eta \in (0, \delta_\Gamma - \delta_W)$*, there exists* $C = C(g, W, \eta, a) > 0$ *such that for all nice preimages* $\widetilde{W}$ *of* W *and all* $T \geq 4\mathrm{diam}_g(W)$*, if*

$$U_T = U_T(\widetilde{W}, g) = \left\{\xi \in \widetilde{M} \cup \partial\widetilde{M} \ ;\ \exists x \in \widetilde{W} \text{ s.t. } \forall t \in [0, T], [x, \xi]_T \cap \Gamma W \subset \widetilde{W}.\right\},$$

then the unique Patterson-Sullivan density $(\nu_x^g)_{x \in \widetilde{M}}$ *on* Λ_Γ *such that* $\nu_o^g(\Lambda_\Gamma^r) = 1$ *satisfies*

$$\nu_o^g(U_T) \leq C e^{-(\delta_\Gamma - \delta_{W^c} - \eta)T}.$$

In particular,

$$\nu_o^g(\Lambda_{\widetilde{W}}) = \nu_o^g\Big(\bigcap_{T>0} U_T\Big) = 0.$$

Proof. – We start with any Patterson-Sullivan density (ν_x^g) on Λ_Γ obtained as a weak limit of an average as in Section 3.4. We will show that there exists $C > 0$ such that for all $T > 0$ large enough,

$$\nu_o^g(U_T) \leq C e^{-(\delta_\Gamma - \delta_{W^c} - \eta)T}. \tag{28}$$

By definition, U_T is the (open) set of points joined by a geodesic from $\widetilde{W}$ which, after exiting $\widetilde{W}$, does not enter $\Gamma\widetilde{W}$ before time T, so that

$$\Lambda_{\widetilde{W}} = \bigcap_{T>0} U_T.$$

Therefore, (28) implies

$$\nu_o^g\left(\Lambda_{\widetilde{W}}\right) = 0 \quad \text{so that} \quad \nu_o^g(\mathcal{Z}_{W^c}) = 0 \quad \text{and} \quad \nu_o^g(\Lambda_\Gamma^r) = 1.$$

By Hopf-Tsuji-Sullivan Theorem, it will imply that Γ is divergent and the Patterson-Sullivan density is unique.

Recall notations from Section 3.4. We omit the mention of the metric g here. As in [37], choose a positive increasing map $h : \mathbb{R}^+ \to \mathbb{R}^+$ such that for all $\eta > 0$, there exists $C_\eta > 0$ such that

$$\forall r \geq 0, \quad \forall t \geq 0, \quad h(t + r) \leq C_\eta e^{\eta t} h(r), \tag{29}$$

and the series $\widetilde{P_\Gamma}(s) = \sum_{\gamma \in \Gamma} h(d(o, \gamma o) e^{-sd(o,\gamma o)}$ diverges at the critical exponent δ_Γ. Construct a Patterson-Sullivan density (ν_x) s.t. for all $x \in \widetilde{M}$, the measure ν_x is a weak limit as $s \to \delta_\Gamma^+$ of the positive finite measures

$$\nu_x^s = \frac{1}{\widetilde{P_\Gamma}(o, s)} \sum_{\gamma \in \Gamma} h(d(x, \gamma o)) e^{-sd(x,\gamma o)} \mathcal{D}_{\gamma o}. \tag{30}$$

For all $\gamma \in \Gamma_{\widetilde{W}}$, define $\mathcal{O}_{\widetilde{W}}(\gamma\widetilde{W})$ as the set of $y \in \widetilde{M} \cup \partial\widetilde{M}$ such that there exists $v \in S^g\widetilde{W}$ such that the first intersection of the geodesic ray $(\pi g^t v)_{t \geq 0}$ with $\Gamma\widetilde{W}$, after the first exit of $\widetilde{W}$ is in $\gamma\widetilde{W}$, and the point y belongs to $(\pi g^t v)_{t \geq 0}$.

By definition of U_T and $\Gamma_{\widetilde{W}}$, and triangular inequality, for all $T > 0$ and $\alpha \in \Gamma$, if $\alpha o \in U_T$, there exists $\gamma \in \Gamma_{\widetilde{W}}$ such that $\alpha o \in \mathcal{O}_{\widetilde{W}}(\gamma\widetilde{W})$ and $d(o, \gamma o) \geq T - 2D$, with

$D = \operatorname{diam}(\widetilde{W})$. Indeed, choose γ so that $\gamma\widetilde{W}$ is the first copy of $\widetilde{W}$ intersected by all geodesic segments from $\widetilde{W}$ to αo after exiting $\widetilde{W}$ inside $\Gamma\widetilde{W}$.

In other words, we have

$$\Gamma o \cap U_T \subset \bigcup_{\gamma\in\Gamma_{\widetilde{W}}, d(o,\gamma o)\geq T-2D} \mathcal{O}_{\widetilde{W}}(\gamma\widetilde{W}). \tag{31}$$

Fix $s > \delta_\Gamma$ and recall from (30) that for all $x, y \in \widetilde{M}$ and $\xi \in \Gamma o$,

$$\frac{d\nu_y^s}{d\nu_x^s}(\xi) = e^{-s(d(y,\xi)-d(x,\xi))}\frac{h(d(y,\xi))}{h(d(x,\xi))}.$$

Therefore, for all $\gamma \in \Gamma_{\widetilde{W}}$,

$$\begin{aligned}\nu_o^s\left(\mathcal{O}_{\widetilde{W}}(\gamma\widetilde{W})\right) &= \nu_{\gamma^{-1}o}^s\left(\mathcal{O}_{\gamma^{-1}\widetilde{W}}(\widetilde{W})\right)\\ &= \int_{\mathcal{O}_{\gamma^{-1}\widetilde{W}}(\widetilde{W})} e^{-s(d(\gamma^{-1}o,\xi)-d(o,\xi))}\frac{h(d(\gamma^{-1}o,\xi))}{h(d(o,\xi))}\,d\nu_o^s(\xi).\end{aligned}$$

Moreover, there exists $C > 0$ such that as soon as $d(o,\gamma o) > 2D$, for all $\xi \in \mathcal{O}_{\gamma^{-1}\widetilde{W}}(\widetilde{W})$, $d(\gamma^{-1}o,o) + d(o,\xi) - C \leq d(\gamma^{-1}o,\xi) \leq d(\gamma^{-1}o,o) + d(o,\xi)$.

It implies by (29) that

$$e^{-s(d(\gamma^{-1}o,\xi)-d(o,\xi))}\frac{h(d(\gamma^{-1}o,\xi))}{h(d(o,\xi))} \leq e^{sC}C_\eta e^{-s(d(\gamma^{-1}o,o)}e^{\eta d(\gamma^{-1}o,o)}.$$

Therefore, as $\nu_o^s(\widetilde{M}\cup\partial\widetilde{M}) = 1$, for all $\gamma \in \Gamma_{\widetilde{W}}$ with $d(o,\gamma o) > 2D$ and $2\delta_\Gamma > s > \delta_\Gamma$,

$$\nu_o^s\left(\mathcal{O}_{\widetilde{W}}(\gamma\widetilde{W})\right) \leq C_\eta e^{(-s+\eta)d(o,\gamma o)}.$$

By (31), for all $T > 4D$, we get

$$\nu_o^s(U_T) \leq C_\eta \sum_{\substack{\gamma\in\Gamma_{\widetilde{W}}\\ d(o,\gamma o)\geq T-2D}} e^{(-s+\eta)d(o,\gamma o)}.$$

Taking any weak limit as $s \to \delta_\Gamma^+$, as U_T is an open set, we obtain

$$\nu_o^g(U_T) \leq C_\eta \sum_{\substack{\gamma\in\Gamma_{\widetilde{W}}\\ d(o,\gamma o)\geq T-2D}} e^{(-\delta_\Gamma+\eta)d(o,\gamma o)}.$$

As $\delta_\Gamma - \eta > \delta_{W^c}$, the right hand side decreases exponentially fast as $T \to +\infty$. As mentioned at the beginning of the proof, by Hopf-Tsuji-Sullivan, we deduce that Γ is divergent so that Theorem 7.26 is proven.

Let us prove now the end of the statement of Proposition 7.31. The Patterson-Sullivan ν_o^g is *the* weak limit as $s \to \delta_\Gamma^+$ of $\nu_o^s = \frac{1}{P_\Gamma(o,s)}\sum_{\gamma\in\Gamma} e^{-sd(o,\gamma o)}\mathcal{D}_{\gamma o}$. Repeating exactly the same computations, setting $h \equiv 1$, we get that there exists $C_a > 0$, depending only on the curvature upperbound, such that for all $T \geq 4D$

$$\nu_o^g(U_T) \leq e^{\delta_\Gamma C_a} \sum_{\substack{\gamma\in\Gamma_{\widetilde{W}}\\ d(o,\gamma o)\geq T-2D}} e^{-\delta_\Gamma d(o,\gamma o)}.$$

We get therefore that for all $T \geq 4D$,

$$\nu_o^g(U_T) \leq e^{\delta_\Gamma(C_a+2D)} e^{-(\delta_\Gamma-\delta_W-\eta)T} \sum_{\gamma\in\Gamma_{\widetilde{W}}} e^{-(\delta_W+\eta)d(o,\gamma o)},$$

which is precisely the desired estimate with

$$C(\eta, g, W, a) = e^{\delta_\Gamma(C_a+2D)} \sum_{\gamma\in\Gamma_{\widetilde{W}}} e^{-(\delta_W+\eta)d(o,\gamma o)}. \tag{32}$$ □

Under the above assumption, the Patterson-Sullivan measure ν_o^g gives full mass to the set of endpoints of lifts of geodesics of (M, g) which come back infinitely often in W. This set is in general strictly smaller than the radial limit set. The product structure of the Bowen-Margulis measure (see Section 3.4) implies the following useful fact.

Corollary 7.32. – *Under the same assumptions, let $W \subset M$ be any compact set such that $\delta_W(g) < \delta_\Gamma(g)$. Then the Bowen-Margulis measure of $S^g M$ is finite and gives full mass to the set of bi-infinite geodesics which intersect infinitely often W in the past and in the future.*

7.5. Entropy variation for SPR manifolds

As mentioned earlier, the original motivation of this article was to find reasonable geometric assumptions on non-compact manifolds with negative curvature such that the entropy is regular under a small variation of the metric. In this subsection, our aim is to finish the proof of Theorem 7.1.

Let $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ be a uniformly $\mathcal{C}^1$ family of complete metrics on the manifold M such that for all $\varepsilon \in (-1, 1)$, $-b^2 \leq K_{g_\varepsilon} \leq -a^2$ for some $b > a > 0$, and (M, g_0) is SPR.

Let $W \subset M$ be a compact subset such that $\delta_W(g_0) < \delta_\Gamma(g_0)$, and let $\widetilde{W}$ be a nice preimage of W. For $r > 0$, denote by $W_r = \{x \in M; d^{g_0}(x, W) \leq r\}$ the (g_0, r)-neighborhood of W. Note that $\delta_{W_r}(g_0) \leq \delta_W(g_0) \leq \delta_\infty(g_0) < \delta_\Gamma(g_0)$. Denote by $\widetilde{W}_r$ a nice preimage of W_r such that $\widetilde{W} \subset \widetilde{W}_r$. Observe that $\gamma\widetilde{W}_r$ is the (g_0, r)-neighborhood of $\gamma\widetilde{W}$.

Lemma 7.33. – *For all $r > 0$, there exists a finite set $F \subset \Gamma$ and $\varepsilon_0 > 0$ such that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, we have*

$$\Gamma_{\widetilde{W}_{2r}}(g_0) \subset \bigcup_{\alpha,\beta\in F} \alpha\Gamma_{\widetilde{W}_r}(g_\varepsilon)\beta \quad \textit{and} \quad \Gamma_{\widetilde{W}_r}(g_\varepsilon) \subset \bigcup_{\alpha,\beta\in F} \alpha\Gamma_{\widetilde{W}}(g_0)\beta.$$

Proof. – We prove the right inclusion, the left one is proved similarly.

Let $D = \mathrm{diam}_{g_0}(\widetilde{W})$ and $D' = e^1(D + 1)$, so that for all $\varepsilon \in (-1, 1)$, $\mathrm{diam}_{g_\varepsilon}(\widetilde{W}_r) \leq D'$. It follows from Section 5 that there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (\varepsilon_0, \varepsilon_0)$ $x, y \in \widetilde{W}_r$, and $\gamma \in \Gamma_{\widetilde{W}_r}$, the g_ε-geodesic between x and γy is at distance less than r to the g_0-geodesic between x and y. Reasoning as in Proposition 7.9 leads to the desired result. □

This lemma leads to the following corollary, which implies the first item of Theorem 7.1.

COROLLARY 7.34. – *Let $(g_\varepsilon)_{\varepsilon\in(-1,1)}$ be a uniformly $\mathcal{C}^1$ family of complete metrics on the manifold M such that for all $\varepsilon \in (-1,1)$, $-b^2 \le K_{g_\varepsilon} \le -a^2$ for some $b > a > 0$, and (M, g_0) is SPR. Then for all $\alpha > 0$ and $r > 0$, there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, we have*

$$e^{-\alpha}\delta_{W_{2r}}(g_0) \le \delta_{W_r}(g_\varepsilon) \le e^{\alpha}\delta_W(g_0).$$

In particular, the entropy at infinity $\varepsilon \mapsto \delta_\infty(g_\varepsilon)$ is continuous at $\varepsilon = 0$, and if $\alpha > 0$ is small enough, g_ε is SPR for $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$.

Proof. – Let $r, \alpha > 0$ be fixed, and choose ε_0 as in Lemma 7.33. For all $\gamma \in \Gamma$, we have $d^{g_\varepsilon}(o, \gamma o) \ge e^{-\varepsilon/2} d^{g_0}(o, \gamma o)$. Therefore, for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ we get $\delta_{W_r}(g_\varepsilon) \le e^{\varepsilon/2}\delta_W(g_0) \le e^{\alpha}\delta_W(g_0)$ up to reducing ε_0. The other inequality is proved similarly. □

Let us show now the last item of Theorem 7.1, that is that the mass of the Bowen-Margulis measure of g_ε varies continuously. This will rely on the following estimate, which is a uniform version of Proposition 7.31.

LEMMA 7.35. – *For all $\delta_0 \in (0, \delta_\Gamma(g_0) - \delta_\infty(g_0))$ and $\beta \in (0, \delta_0)$, there exists a compact set $W \subset M$ with nice preimage $\widetilde{W}$, $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, we have $\delta_\Gamma(g_\varepsilon) - \delta_{W^c}(g_\varepsilon) \ge \delta_0$ and*

$$\nu_o^{g_\varepsilon}(U_T(\widetilde{W}, g_\varepsilon)) \le C e^{-\beta T},$$

where $U_T(\widetilde{W}, g_\varepsilon)$ is defined as in Proposition 7.31.

Proof. – Let $\delta_0 \in (0, \delta_\Gamma(g_0) - \delta_\infty(g_0))$ be fixed. By the above corollary, for $|\varepsilon|$ small enough, (M, g_ε) is SPR and has therefore a finite Bowen-Margulis. Choose $\alpha > 0$ small enough and a large enough compact set $W \subset M$ so that $\delta_\infty(g_0) \le \delta_{W^c}(g_0) \le e^{\alpha}\delta_\infty(g_0)$. Let $r > 0$ small enough and $\varepsilon_0 > 0$ given by Corollary 7.34 be such that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$,

$$e^{-\alpha}\delta_{W^c}(g_0) \le \delta_{W_r}(g_\varepsilon) \le e^{\alpha}\delta_{W^c}(g_0).$$

Up to decreasing $\alpha > 0$, we can therefore assume that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$,

$$\delta_\Gamma(g_\varepsilon) - \delta_{W_r}(g_\varepsilon) \ge \delta_0 > 0.$$

Let $\beta \in (0, \delta_0)$ and $\widetilde{W}_r$ nice preimage of W_r be fixed. Define $D > 0$ as $D = \sup_{\varepsilon\in(-\varepsilon_0,\varepsilon_0)} \operatorname{diam}(\widetilde{W}_r)$. For all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, let $U_T^\varepsilon = U_T(\widetilde{W}_r, g_\varepsilon)$ be defined as in Proposition 7.31. By the last estimate in the proof of Proposition 7.31, there exists $C_a > 0$, only depending on the curvature upperbound of the metrics g_ε, such that for all $T > 4D$,

$$\nu_0^{g_\varepsilon}(U_T^\varepsilon) \le e^{\delta_\Gamma(g_\varepsilon)(C_a+2D)} e^{-\beta T} \sum_{\gamma\in\Gamma_{\widetilde{W}_r(g_\varepsilon)}} e^{-(\delta_{W_r}(g_\varepsilon)+\beta) d^{g_\varepsilon}(o,\gamma o)}.$$

Therefore,

$$\nu_0^{g_\varepsilon}(U_T^\varepsilon) \le e^{K} e^{-\beta T} \sum_{\gamma\in\Gamma_{\widetilde{W}(g_0)}} e^{-(e^{-\varepsilon}\delta_W(g_0)+\alpha)e^{-\varepsilon} d^{g_0}(o,\gamma o)},$$

where $K \in \mathbb{R}$ is independent of ε. Up to reducing $\alpha > 0$ and $\varepsilon_0 > 0$, we can suppose that $e^{-\alpha}\delta_W(g_0) + \beta)e^{-\varepsilon} > \delta_W(g_0) + \frac{\beta}{2}$. Therefore, we get that for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$,

$$\nu_0^{g_\varepsilon}(U_T^\varepsilon) \le C e^{-\beta T}, \tag{33}$$

with $C > 0$ being independent of ε. This concludes the proof of Lemma 7.35, the compact set W of the statement being the set W_r of the proof. □

Let us now conclude the proof of Theorem 7.1. Let $W \subset M$, $\widetilde{W} \subset \widetilde{M}$ and $\beta, \varepsilon_0, C > 0$ satisfy the conclusion of Lemma 7.35. For all $R > 0$, set as usual $W_R = \{x \in M; d_{g_0}(x, W) \leq R\}$. We have shown in Theorem 5.14 that, under our current hypotheses, the Bowen-Margulis measure $\varepsilon \mapsto m^{g_\varepsilon}_{\mathrm{BM}}$ varies continuously for the weak-* convergence, i.e., on the dual of compactly supported maps. In particular, for all *fixed compact sets* $K \subset M$ with $m^{g_0}_{\mathrm{BM}}(\partial S^{g_0} K) = 0$, the map $\varepsilon \mapsto m^{g_\varepsilon}_{\mathrm{BM}}(S^{g_\varepsilon} K)$ is continuous at $\varepsilon = 0$. Therefore the following lemma will imply Theorem 7.1.

LEMMA 7.36. – *With the above notations, for all $\alpha > 0$, there exists $R_0 > 0$ such that for all $R \geq R_0$ and all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, we have*

$$m^{g_\varepsilon}_{\mathrm{BM}}(S^{g_\varepsilon}(M \backslash W_R)) \leq \alpha.$$

Proof. – Let $R > 8\,\mathrm{diam}(W)$ be fixed and let $O_R = M \backslash W_R$. By Corollary 7.32, for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, since $\delta_{W^c}(g_\varepsilon) < \delta_\Gamma(g_\varepsilon)$, the Bowen-Margulis measure $m^{g_\varepsilon}_{\mathrm{BM}}$ gives full mass to the set of vectors which hit infinitely often W in the past and in the future. In particular,

$$m^{g_\varepsilon}_{\mathrm{BM}}(S^{g_\varepsilon} O_R) = m^{g_\varepsilon}_{\mathrm{BM}}\left(\coprod_{n \geq R-1} O^\varepsilon_n\right),$$

where O^ε_n is defined for all integers $n \geq R - 1$ by

$$O^\varepsilon_n = \left\{v \in S^{g_\varepsilon} O_R\ ;\ \exists t \in [n, n+1[\text{ s.t. } \forall s \in [0, t), \pi g^{-s} v \notin W \text{ and } \pi g_\varepsilon^{-t} v \in W\right\}.$$

Therefore, since the Bowen-Margulis measure $m^{g_\varepsilon}_{\mathrm{BM}}$ is invariant under the geodesic flow (g^t_ε),

$$m^{g_\varepsilon}_{\mathrm{BM}}(S^{g_\varepsilon} O_R) = \sum_{n \geq R-1} m^{g_\varepsilon}_{\mathrm{BM}}\left(O^\varepsilon_n\right) = \sum_{n \geq R-1} m^{g_\varepsilon}_{\mathrm{BM}}\left(g_\varepsilon^{-n}(O^\varepsilon_n)\right).$$

Now, by definition for all $v \in g_\varepsilon^{-n}(O^\varepsilon_n)$, there exists $t \in [0, 1)$ such that $w = g_\varepsilon^{-n-t} v \in S^{g_\varepsilon} W$ and for all $s \in [0, n]$, we have $\pi g^s w \notin W$.

Let us write

$$\widetilde{A}^\varepsilon_n = \left\{v \in S^{g_\varepsilon}\widetilde{W}; \exists t \in [n, n+1) \text{ s.t. } \forall s \in (0, t), \pi g^s_\varepsilon v \notin \gamma \widetilde{W} \text{ and } \pi g^t v \in \gamma \widetilde{W}\right\}.$$

The reader will easily check that $\bigcup_{s \in [0,1)} g^s_\varepsilon A^\varepsilon_n \subset S^{g_\varepsilon}\widetilde{M}$ projects onto $g^{-n}(O^\varepsilon_n)$. Moreover, as soon as ε_0 is small enough, since $g_\varepsilon \geq e^{-\varepsilon} g_0 \geq \frac{1}{4} g_0$, all vectors $v \in \widetilde{A}^\varepsilon_n$ have a point at infinity v^+ which satisfies $v_+ \in U_{n/2}(\widetilde{W}, g_\varepsilon)$. As the map

$$v \mapsto e^{\delta_\Gamma(g_\varepsilon)(\mathscr{B}_{v_+}(o, \pi v) + \mathscr{B}_{v_-}(o, \pi v)}$$

is uniformly bounded in $v \in W$ and $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, the product structure of the Bowen-Margulis measure (see Section 3.4) implies that

$$m^{g_\varepsilon}_{\mathrm{BM}}(O^\varepsilon_n) = m^{g_\varepsilon}_{\mathrm{BM}}(g_\varepsilon^{-n}(O^\varepsilon_n)) \leq 2K \nu^{g_\varepsilon}_o(\Lambda_\Gamma) \times \nu^{g_\varepsilon}_o(U^\varepsilon_{n/2}),$$

which eventually gives by Lemma 7.35

$$m^{g_\varepsilon}_{\mathrm{BM}}(O^\varepsilon_n) \leq 2KC e^{-\frac{\alpha}{2} n},$$

where C and α do not depend on ε. Therefore, we get

$$m^{g_\varepsilon}_{\mathrm{BM}}(S^{g_\varepsilon} O_R) \leq 2KC \sum_{n \geq R-1} e^{-\frac{\alpha}{2} n} \leq \varepsilon$$

as soon as R is large enough. □

BIBLIOGRAPHY

[1] L. M. Abramov, On the entropy of a flow, *Dokl. Akad. Nauk SSSR* **128** (1959), 873–875.

[2] A. Ancona, Exemples de surfaces hyperboliques de type divergent, de mesure de sullivan associées finies mais non géométriquement finies, preprint.

[3] G. Arzhantseva, C. Cashen, J. Tao, Growth tight actions, *Pacific Journal of Mathematics* **278** (2015), 1–49.

[4] W. Ballmann, *Lectures on spaces of nonpositive curvature*, DMV Seminar **25**, Birkhäuser, 1995.

[5] A. Bellis, On links between horocyclic and geodesic orbits on geometrically infinite surfaces, *J. Éc. polytech. Math.* **5** (2017), 443–454.

[6] A. Bellis, Étude topologique du flot horocyclique: le cas des surfaces géométriquement infinies, Ph.D. Thesis, Université de Rennes, 2018.

[7] B. H. Bowditch, Geometrical finiteness with variable negative curvature, *Duke Math. J.* **77** (1995), 229–274.

[8] M. Boyle, J. Buzzi, R. Gómez, Borel isomorphism of SPR Markov shifts, *Colloq. Math.* **137** (2014), 127–136.

[9] M. R. Bridson, A. Haefliger, *Metric spaces of non-positive curvature*, Grundl. math. Wiss. **319**, Springer, 1999.

[10] M. Brin, A. Katok, On local entropy, in *Geometric dynamics (Rio de Janeiro, 1981)*, Lecture Notes in Math. **1007**, Springer, 1983, 30–38.

[11] R. Brooks, The bottom of the spectrum of a Riemannian covering, *J. reine angew. Math.* **357** (1985), 101–114.

[12] B. Colbois, Y. Colin de Verdière, Sur la multiplicité de la première valeur propre d'une surface de Riemann à courbure constante, *Comment. Math. Helv.* **63** (1988), 194–208.

[13] Y. Coudene, B. Schapira, Generic measures for hyperbolic flows on non-compact spaces, *Israel J. Math.* **179** (2010), 157–172.

[14] R. Coulon, R. Dougall, B. Schapira, S. Tapie, Amenability and covers via twisted patterson-sullivan theory, preprint arXiv:1809.10881.

[15] F. Dal'bo, J.-P. Otal, M. Peigné, Séries de Poincaré des groupes géométriquement finis, *Israel J. Math.* **118** (2000), 109–124.

[16] E. I. Dinaburg, A correlation between topological entropy and metric entropy, *Dokl. Akad. Nauk SSSR* **190** (1970), 19–22.

[17] P. Eberlein, Geodesic flows on negatively curved manifolds. I, *Ann. of Math.* **95** (1972), 492–510.

[18] M. Einsiedler, S. Kadyrov, A. Pohl, Escape of mass and entropy for diagonal flows in real rank one situations, *Israel J. Math.* **210** (2015), 245–295.

[19] M. Einsiedler, E. Lindenstrauss, P. Michel, A. Venkatesh, The distribution of closed geodesics on the modular surface, and Duke's theorem, *Enseign. Math.* **58** (2012), 249–313.

[20] A. Fathi, L. Flaminio, Infinitesimal conjugacies and Weil-Petersson metric, *Ann. Inst. Fourier* **43** (1993), 279–299.

[21] L. Flaminio, Local entropy rigidity for hyperbolic manifolds, *Comm. Anal. Geom.* **3** (1995), 555–596.

[22] S. Gallot, D. Hulin, J. Lafontaine, *Riemannian geometry*, third ed., Universitext, Springer, 2004.

[23] O. Glorieux, Exposant critique des groupes de surface agissant sur $\mathbb{H}^2 \times \mathbb{H}^2$ et $\mathbb{H}^3$, thèse de doctorat, Université Pierre et Marie Curie – Paris VI, 2015, `https://hal.archives-ouvertes.fr/tel-01189128`.

[24] T. N. T. Goodman, Relating topological entropy and measure entropy, *Bull. London Math. Soc.* **3** (1971), 176–180.

[25] S. Gouëzel, C. Noûs, B. Schapira, S. Tapie, F. Riquelme, Pressure at infinity and strong positive recurrence in negative curvature, preprint arXiv:2007.08816.

[26] B. M. Gurevich, S. V. Savchenko, Thermodynamic formalism for symbolic Markov chains with a countable number of states, *Uspekhi Mat. Nauk* **53** (1998), 3–106.

[27] M. Handel, B. Kitchens, Metrics and entropy for non-compact spaces, *Israel J. Math.* **91** (1995), 253–271.

[28] E. Heintze, H.-C. Im Hof, Geometry of horospheres, *J. Differential Geom.* **12** (1977), 481–491.

[29] A. Katok, Entropy and closed geodesics, *Ergodic Theory Dynam. Systems* **2** (1982), 339–365.

[30] A. Katok, G. Knieper, M. Pollicott, H. Weiss, Differentiability and analyticity of topological entropy for Anosov and geodesic flows, *Invent. math.* **98** (1989), 581–597.

[31] A. Katok, G. Knieper, H. Weiss, Formulas for the derivative and critical points of topological entropy for Anosov and geodesic flows, *Comm. Math. Phys.* **138** (1991), 19–31.

[32] G. Knieper, Volume growth, entropy and the geodesic stretch, *Math. Res. Lett.* **2** (1995), 39–58.

[33] F. Ledrappier, Structure au bord des variétés à courbure négative, in *Séminaire de Théorie Spectrale et Géométrie, No. 13, Année 1994–1995*, Sémin. Théor. Spectr. Géom. **13**, Univ. Grenoble I, 1995, 97–122.

[34] F. Ledrappier, Harmonic 1-forms on the stable foliation, *Bol. Soc. Brasil. Mat. (N.S.)* **25** (1994), 121–138.

[35] M. Misiurewicz, A short proof of the variational principle for a $\mathbf{Z}_+^N$ action on a compact space, in *International Conference on Dynamical Systems in Mathematical Physics (Rennes, 1975)*, Astérisque **40**, Soc. Math. France, 1976, 147–157.

[36] J.-P. Otal, M. Peigné, Principe variationnel et groupes kleiniens, *Duke Math. J.* **125** (2004), 15–44.

[37] S. J. Patterson, The limit set of a Fuchsian group, *Acta Math.* **136** (1976), 241–273.

[38] F. Paulin, M. Pollicott, B. Schapira, Equilibrium states in negative curvature, *Astérisque* **373** (2015), 281.

[39] M. Peigné, On the Patterson-Sullivan measure of some discrete group of isometries, *Israel J. Math.* **133** (2003), 77–88.

[40] M. Peigné, On some exotic Schottky groups, *Discrete Contin. Dyn. Syst.* **31** (2011), 559–579.

[41] V. Pit, B. Schapira, Finiteness of Gibbs measures on non-compact manifolds with pinched negative curvature, *Annales Institut Fourier* **68** (2018), 457–510.

[42] F. Riquelme, Autour de l'entropie des difféomorphismes de variétés non compactes, thèse de doctorat, Université Rennes 1, 2016, `https://tel.archives-ouvertes.fr/tel-01386691`.

[43] F. Riquelme, Ruelle's inequality in negative curvature, *Discrete and Continuous Dynamical systems* **38** (2018), 2809–2825.

[44] F. Riquelme, Entropies for geodesic flows, appendix to arXiv:2007.08816, 2019.

[45] F. Riquelme, A. Velozo, Escape of mass and entropy for geodesic flows, *Ergodic Theory Dynam. Systems* **39** (2019), 446–473.

[46] T. Roblin, Ergodicité et équidistribution en courbure négative, *Mém. Soc. Math. Fr. (N.S.)* **95** (2003), 96.

[47] A. Sambarino, The orbital counting problem for hyperconvex representations, *Ann. Inst. Fourier* **65** (2015), 1755–1797.

[48] O. M. Sarig, Phase transitions for countable Markov shifts, *Comm. Math. Phys.* **217** (2001), 555–577.

[49] D. Sullivan, The density at infinity of a discrete group of hyperbolic motions, *Publ. Math. IHÉS* **50** (1979), 171–202.

[50] D. Sullivan, Disjoint spheres, approximation by imaginary quadratic numbers, and the logarithm law for geodesics, *Acta Math.* **149** (1982), 215–237.

[51] D. Sullivan, Entropy, Hausdorff measures old and new, and limit sets of geometrically finite kleinian groups, *Acta Math.* **153** (1984), 259–277.

[52] S. Tapie, A variation formula for the topological entropy of convex-cocompact manifolds, *Ergodic Theory Dynam. Systems* **31** (2011), 1849–1864.

[53] A. Velozo, Entropy theory of geodesic flows, preprint, arXiv:1711.06796.

[54] P. Vidotto, *Ergodic properties of some negatively curved manifolds with infinite measure*, Mém. Soc. Math. Fr. **160**, 2019.

[55] W. Yang, Growth tightness for groups with contracting elements, *Mathematical Proceedings of the Cambridge Philosophical Society* **157** (2014), 297–319.

[56] C. Yue, The ergodic theory of discrete isometry groups on manifolds of variable negative curvature, *Trans. Amer. Math. Soc.* **348** (1996), 4965–5005.

(Manuscrit reçu le 9 mars 2018 ;
accepté le 20 avril 2019.)

Barbara Schapira
Laboratoire Jean Leray
2 rue de la Houssinière - BP92208
44322 Nantes Cedex 3, France

Samuel Tapie
I.R.M.A.R. UMR 7352
UFR de mathématiques
Campus de Beaulieu, bâtiments 22 et 23
263 avenue du Général Leclerc, CS 74205
35042 Rennes Cedex, France

Ann. Scient. Éc. Norm. Sup.
4e série, t. 54, 2021, p. 69 à 131

CONSERVATIVE ANOSOV DIFFEOMORPHISMS OF $\mathbb{T}^2$ WITHOUT AN ABSOLUTELY CONTINUOUS INVARIANT MEASURE

BY ZEMER KOSLOFF

ABSTRACT. – We construct examples of C^1 Anosov diffeomorphisms on $\mathbb{T}^2$ which are of Krieger type III_1 with respect to Lebesgue measure. This shows that the Gurevic Oseledec phenomena that conservative $C^{1+\alpha}$ Anosov diffeomorphisms have a smooth invariant measure does not hold true in the C^1 setting.

RÉSUMÉ. – Sur $\mathbb{T}^2$, on construit des exemples de difféomorphismes C^1 d'Anosov qui sont de type de Krieger III_1 par rapport à la mesure de Lebesgue. Ceci montre que le phénomène de Gurevic Oseledec selon lequel tout difféomorphisme conservatif d'Anosov $C^{1+\alpha}$ a une mesure invariante lisse, n'est pas valable dans le cadre C^1.

1. Introduction

This paper provides the first examples of Anosov diffeomorphisms of $\mathbb{T}^2$ which are conservative and ergodic yet there is no Lebesgue absolutely continuous invariant measure.

Let M be a compact, boundaryless smooth manifold and $f : M \to M$ be a diffeomorphism. A natural question which arises is whether f preserves a measure which is absolutely continuous with respect to the volume measure on M. In order to avoid confusion in what follows, we would like to stress out that in this paper, the term *conservative* means the definition from ergodic theory which is non existence of wandering sets of positive measure. That is f is conservative if and only if for every $W \subset M$ so that $\{f^n W\}_{n\in\mathbb{Z}}$ are disjoint (modulo the volume measure), $\mathrm{vol}(W) = 0$.

It follows from [16] that for a generic C^2 Anosov diffeomorphism there exists no absolutely continuous invariant measure (a.c.i.m.), [6, p. 72, Corollary 4.15.]. Following this result, Sinai asked whether a generic Anosov diffeomorphism will satisfy Poincare recurrence. This question was answered by Gurevic and Oseledec [9] who proved that the set of conservative (Poincare recurrent) C^2 Anosov diffeomorphism is meager in the C^2 topology (restricted to the open set of Anosov diffeomorphisms). Indeed, they have proved that if f is a conservative C^2 Anosov (hyperbolic) diffeomorphism, then f preserves a probability

 doi:10.24033/asens.2456

measure in the measure class of the volume measure which combined with the result of Livsic and Sinai proves the non-genericity result. The proof in [9] uses the absolute continuity of the foliations and existence of SRB measures to show that if the SRB measure for f is not equal to the SRB measure for f^{-1} then there exists a continuous function $g : M \to \mathbb{R}$ and a set $A \subset M$ of positive volume so that,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g\left(f^k(x)\right) \neq \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g\left(f^{-k}(x)\right), \ \forall x \in A.$$

It is then a straightforward argument to construct a set $B \subset A$ of positive volume measure so that for almost every $x \in B$, the set $\left\{k \in \mathbb{N} : \ f^k x \in B\right\}$ is finite, in contradiction with Halmos Recurrence Theorem [1].

This result remains true for $C^{1+\alpha}$, $\alpha > 0$ Anosov diffeomorphisms. However, since there exist C^1 Anosov diffeomorphisms whose stable and unstable foliations are not absolutely continuous [17], this proof can not be generalized for the C^1 setting. This paper is concerned with the question whether every conservative C^1-Anosov diffeomorphism has an absolutely continuous invariant measure.

An easier version of this question was studied before in the context of smooth expanding maps. Every C^2 expanding map of a manifold has an absolutely continuous invariant measure [14]. In contrast to the higher regularity case, Avila and Bochi [4], extending previous results of Campbell and Quas [8], have shown that a generic C^1 expanding map has no a.c.i.m and a generic C^1 expanding map of the circle is not recurrent [8]. It seems natural to argue that these generic statements for expanding maps can be transferred to Anosov diffeormorphisms via the natural extension. However there are several problems with this approach which could be summarized into roughly two parts:

— The natural extension construction is an abstract theorem and in many cases it is not clear if it has an Anosov model. See [22] for constructions of smooth natural extensions.

— In order for the natural extension to be conservative, the expanding map has to be recurrent [19, Th. 4.4] and a generic C^1 expanding map is not recurrent.

Another natural approach in finding C^1 examples with a certain property is to prove that the property is generic in the C^1 topology, see for example [5]. However since by the result of Sinai and Livsic, a generic C^1 Anosov map is dissipative, it is not clear to us how to use this approach to find a conservative example without an a.c.i.m. Nonetheless we prove the following.

Theorem 1. – *There exists a C^1-Anosov diffeomorphism of the two torus $\mathbb{T}^2$ which is ergodic, conservative and there exists no σ-finite invariant measure which is absolutely continuous with respect to the Lebesgue measure on $\mathbb{T}^2$.*

In fact, the ergodic type III transformations (a transformation without an a.c.i.m.) can be further decomposed into the Krieger Araki-Woods classes III_λ, $0 \leq \lambda \leq 1$ [13] , see Section 2, and our examples are of type III_1.

The examples are constructed by modifying a linear Anosov diffeomorphism to obtain a change of coordinates which takes the Lebesgue measure to a measure which is equivalent to a type III Markovian measure (on a Markov partition of the linear diffeomorphism). These examples are greatly inspired by the ideas of Bruin and Hawkins [7] where they modify the map $f(x) = 2x \mod 1$ using the push forward (with respect to the dyadic representation) of a Hamachi product measure on $\{0, 1\}^{\mathbb{N}}$ to the circle. Since by embedding a horseshoe in a linear transformation one loses the explicit formula for the Radon Nykodym derivatives of the modified transformations, we couldn't use measures on a full shift space but rather measures supported on topological Markov shifts. The measures which play the role of the Hamachi measures in our construction are the type III_1 (for the shift) inhomogeneous Markov measures.

This paper is organized as follows. In Section 2 we start by introducing the definitions and background material from nonsingular ergodic theory and smooth dynamics which are used in this paper. We end this section with a discussion on the method of the construction. Section 3 presents the inductive construction of the type III_1 Markov shift examples. In Section 4 we show how to use the one sided Markov measures from the previous section to obtain a modification of the golden mean shift. In Section 5 we show how to embed and modify the one dimensional perturbations of the previous sections to obtain homeomorphisms of the two torus, which when applied as conjugation to a certain total automorphism (the natural extension of the golden mean shift) are examples of type III_1 Anosov diffeomorphisms. Finally in the appendix we prove that these Markovian measures satisfy the aforementioned properties (ergodic, conservative and type III_1).

2. Preliminary definitions and a discussion on the method of construction

2.1. Basics of nonsingular ergodic theory

This subsection is a very short introduction to nonsingular ergodic theory. For more details and explanations please see [1].

Let $(X, \mathcal{B}, \mu)$ be a standard probability space. In what follows equalities (and inclusions) of sets are modulo the measure μ on the space. A measurable map $T : X \to X$ is *nonsingular* if $T_*\mu := \mu \circ T^{-1}$ is equivalent to μ meaning that they have the same collection of negligible sets. If T is invertible one has the Radon Nykodym derivatives

$$(T^n)'(x) := \frac{d\mu \circ T^n}{d\mu}(x) : X \to \mathbb{R}_+.$$

A set $W \subset X$ is *wandering* if $\{T^n W\}_{n\in\mathbb{Z}}$ are pairwise disjoint and as was stated before we say that T is *conservative* if there exists no wandering set of positive measure. By the Halmos' Recurrence Theorem a transformation is conservative if and only if it satisfies Poincare recurrence, that is given a set of positive measure $A \in \mathcal{B}$, almost every $x \in A$ returns to itself infinitely often. A transformation T is *ergodic* if there are no non trivial T invariant sets. That is $T^{-1}A = A$ implies $A \in \{\emptyset, X\}$.

We end this subsection with the definition of the *Krieger ratio set* $R(T)$. We say that $r \geq 0$ is in $R(T)$ if for every $A \in \mathcal{B}$ of positive μ measure and for every $\epsilon > 0$ there exists an $n \in \mathbb{Z}$ such that

$$\mu\left(A \cap T^{-n}A \cap \left\{x \in X : \left|(T^n)'(x) - r\right| < \epsilon\right\}\right) > 0.$$

The ratio set of an ergodic measure preserving transformation is a closed multiplicative subgroup of $[0, \infty)$ and hence it is of the form $\{0\}, \{1\}, \{0, 1\}, \{0\} \cup \{\lambda^n : n \in \mathbb{Z}\}$ for $0 < \lambda < 1$

or $[0,\infty)$. Several ergodic theoretic properties can be seen from the ratio set. One of them is that $0 \in R(T)$ if and only if there exists no σ-finite T-invariant μ-a.c.i.m. Another interesting relation is that $1 \in R(T)$ if and only if T is conservative (Maharams Theorem). If $R(T) = [0,\infty)$ we say that T is of type III_1.

2.2. Anosov diffeomorphisms and topological Markov shifts

Smooth dynamics deals with the case where M is a Riemannian manifold and $f : M \to M$ is a diffeomorphism on M. In this paper we would only talk about the class of Anosov (uniformly hyperbolic) automorphisms. A diffeomorphism f is Anosov if for every $x \in M$ there is a decomposition of the tangent bundle at x, $T_xM = E^s_x \oplus E^u_x$, such that

- The decomposition is D_f-equivariant, here D_f denotes the differential of f. That is $\left(D_f\right)_x\left(E^s_x\right) = E^s_{f(x)}$ and $\left(D_f\right)_x\left(E^u_x\right) = E^u_{f(x)}$.
- There exists $0 < \lambda < 1$ and $C > 0$ so that
$$\left\|\left(D_{f^n}\right)_x v\right\| \le C\lambda^n \|v\|, \text{ for every } v \in E^s_x,\ n \ge 0$$
and
$$\left\|\left(D_{f^{-n}}\right)_x u\right\| \le C\lambda^n \|u\|, \text{ for every } u \in E^u_x,\ n \ge 0.$$

A *topological Markov shift* (TMS) on S is the shift on a shift invariant subset $\Sigma \subset S^{\mathbb{Z}}$ of the form
$$\Sigma_A := \left\{x \in S^{\mathbb{Z}} : A_{x_i,x_{i+1}} = 1\right\},$$
where $A = \{A_{s,t}\}_{s,t\in S}$ is a $\{0,1\}$ valued matrix on S. A TMS is mixing if there exists $n \in \mathbb{N}$ such that $A^n_{s,t} > 0$ for every $s,t \in S$.

Markov partitions of the manifold M as in [3, 20, 6, 2] are an important tool in the study of $C^{1+\alpha}$ Anosov diffeomorphisms. They provide a semiconjugacy between a TMS and the Anosov transformation f. One of the important contributions of this paper is that it uses a connection between inhomogeneous Markov chains supported on a TMS to the Anosov diffeomorphism with the push forward of the Markov measure.

Example 2. – Consider $f : \mathbb{T}^2 \to \mathbb{T}^2$ the toral automorphism defined by
$$f(x,y) = (\{x+y\}, x) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \mod 1,$$
where $\{t\}$ is the fractional part of t. Since $\left|\det\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)\right| = 1$, f preserves the Lebesgue measure on $\mathbb{T}^2$. In addition for every $z \in \mathbb{T}^2$, the tangent space can be decomposed as $\operatorname{span}\{v_s\} \oplus \operatorname{span}\{v_u\}$ where $v_u = (1, 1/\varphi)$ and $v_s = (1, -\varphi)$. Here and throughout the paper φ denotes the golden mean ($\varphi := \frac{1+\sqrt{5}}{2}$).

For every $w \in V_u := \operatorname{span}\{v_u\}$ and $z \in \mathbb{T}^2$
$$D_f(z)w = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} w = \varphi w,$$
For every $u \in V_s := \operatorname{span}\{v_s\}$ and $z \in \mathbb{T}^2$, $D_f(z)u = \left(-\frac{1}{\varphi}\right)u$. These facts can be used (cf. [2, 3]) to construct the Markov partition for f with three elements $\{R_1, R_2, R_3\}$, see Figure 2.1.

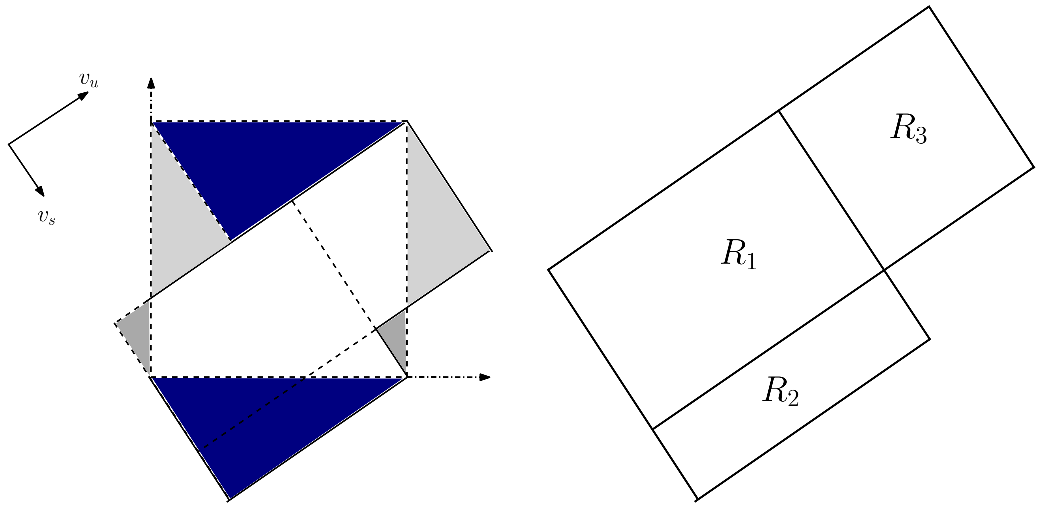


FIGURE 2.1. The construction of the Markov partition

The adjacency matrix of the Markov partition is then defined by $A_{i,j} = 1$ if and only if $\mathrm{int}R_i \cap f^{-1}\left(\mathrm{int}R_j\right) \neq \emptyset$. Here the adjacency matrix is

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Let $\Phi : \Sigma_A \to \mathbb{T}^2$ be the map defined by $\Phi(x) := \bigcap_{n=-\infty}^{\infty} \overline{f^{-n} R_{x_n}}$. Note that by the Baire Category Theorem since $\left\{\bigcap_{n=-N}^{N} \overline{f^{-n} R_{x_n}}\right\}_{N=1}^{\infty}$ is a decreasing sequence of compact sets, $\Phi(x)$ is well defined. The map $\Phi : \Sigma_{\mathbf{A}} \to \mathbb{T}^2$ is continuous, finite to one, and for every $x \in \Sigma_{\mathbf{A}}$,

$$\Phi \circ T(x) = f \circ \Phi(x).$$

In other words, Φ is a semi-conjugacy (topological factor map) between $(\Sigma_{\mathbf{A}}, T)$ to $\left(\mathbb{T}^2, f\right)$. In addition, for every $x \in \mathbb{T}^2 \setminus \bigcup_{n\in\mathbb{Z}} \bigcup_{i=1}^{3} f^{-n}(\partial R_i)$ there exists a unique $w \in \Sigma_{\mathbf{A}}$ so that $\Phi(w) = x$. The Lebesgue measure $m_{\mathbb{T}^2}$ on $\mathbb{T}^2$ is invariant under f. One can check that $m_{\mathbb{T}^2}\left(\bigcup_{i=1}^{3} \partial R_i\right) = 0$ and thus Φ^{-1} defines an isomorphism between $\left(\mathbb{T}^2, m_{\mathbb{T}^2}, f\right)$ and $\left(\Sigma_{\mathbf{A}}, \mu_{\pi_{\mathbf{Q}},\mathbf{Q}}, T\right)$ where $\mu_{\pi_{\mathbf{Q}},\mathbf{Q}}$ is the stationary Markov measure with

$$P_j \equiv \mathbf{Q} := \begin{pmatrix} \frac{\varphi}{1+\varphi} & 0 & \frac{1}{1+\varphi} \\ \frac{\varphi}{1+\varphi} & 0 & \frac{1}{1+\varphi} \\ 0 & 1 & 0 \end{pmatrix} \tag{2.1}$$

and

$$\pi_j = \mathbf{ß_Q} := \left(1/\sqrt{5},\ 1/\varphi\sqrt{5},\ 1/\varphi\sqrt{5}\right) = \left(m_{\mathbb{T}^2}(R_1), m_{\mathbb{T}^2}(R_2), m_{\mathbb{T}^2}(R_3)\right). \tag{2.2}$$

2.2.1. *Nonsingular Markov shifts:–* Let $\{P_n\}_{n=-\infty}^{\infty} \subset M_{S\times S}$ be a sequence of aperiodic and irreducible stochastic matrices on S. In addition let $\{\pi_n\}_{n=-\infty}^{\infty}$ be a sequence of probability distributions on S so that for every $s \in S$ and $n \in \mathbb{Z}$,

$$\sum_{t\in S} \pi_{n-1}(t)\cdot P_n(t,s) = \pi_n(s). \tag{2.3}$$

Then one can define a measure on the collection of cylinder sets,

$$[b]_k^l := \left\{x \in S^{\mathbb{Z}} : \ x_j = b_j \ \forall j \in [k,l]\cap\mathbb{Z}\right\}$$

by

$$\mu\left([b]_k^l\right) := \pi_k(b_k)\prod_{j=k}^{l-1} P_j\left(b_j, b_{j+1}\right).$$

Since the equation (2.3) is satisfied, μ satisfies the consistency condition and therefore by Kolmogorov's extension theorem μ defines a measure on $S^{\mathbb{Z}}$. In this case we say that μ is the Markov measure generated by $\{\pi_n, P_n\}_{n\in\mathbb{Z}}$ and denote $\mu = \mathrm{M}\{\pi_n, P_n : n \in \mathbb{Z}\}$. By $\mathrm{M}\{\pi, P\}$ we mean the measure generated by $P_n \equiv P$ and $\pi_n \equiv \pi$. We say that μ is nonsingular for the shift T on $S^{\mathbb{Z}}$ if $T_*\mu \sim \mu$.

2.3. Local absolute continuity

Let $(X, \mathscr{B})$ be a measure space and $\mathscr{F}_n \subset \mathscr{B}$ be a filtration of X. That is an increasing sequence of σ-algebras such that $\mathscr{F}_n \uparrow \mathscr{B}$. The method in [10, 18] uses ideas from Martingale theory in order to determine whether two Borel probability measures μ, ν are absolutely continuous.

Definition 3. – Given a filtration $\{\mathscr{F}_n\}$, we say that $\nu \ll^{\mathrm{loc}} \mu$ (ν is locally absolutely continuous with respect to μ) if for every $n \in \mathbb{N}$, $\nu_n \ll \mu_n$ where $\nu_n = \nu|_{\mathscr{F}_n}$.

Suppose that $\nu \ll^{\mathrm{loc}} \mu$ w.r.t $\{\mathscr{F}_n\}$, set $z_n := \frac{d\nu_n}{d\mu_n}$. The sequence z_n is a nonnegative martingale with respect to $\mathscr{F}_n$ and thus by the martingale convergence theorem there exists a $[0,\infty]$ valued random variable z_∞ such that $\lim_{n\to\infty} z_n = z_\infty$ a.s. It follows that if $\nu \ll^{\mathrm{loc}} \mu$ then $\nu \ll \mu$ if and only if $z_n \xrightarrow[n\to\infty]{} z_\infty$ in $L^1(\mu)$. The latter holds if and only if the sequence $\{z_n\}_{n=1}^{\infty}$ is uniformly integrable meaning that for all $\epsilon > 0$ there exists $M > 0$ such that for all $n \in \mathbb{N}$, $\int z_n 1_{[z_n>M]} d\mu < \epsilon$.

2.4. Section's overview and explanation of the method of construction

The idea is as follows, let $f(x,y) = (x+y, x)\mathrm{mod}1$, $\{R_1, R_2, R_3\}$ be the corresponding Markov partition for f, Σ_A the resulting topological Markov shift and $\Phi : \Sigma_{\mathbf{A}} \to \mathbb{T}^2$ the topological semiconjugacy with the shift. In addition $\mathbf{Q}$ will always denote the transition matrix corresponding to the Lebesgue measure.

— In Section 3 we present an inductive construction which produces a family of nonatomic inhomogeneous Markov measures which are fully supported on $\Sigma_A \subset \{1,2,3\}^{\mathbb{Z}}$ and are of type III_1.

— Let μ be such a Markov measure generated by $\{\pi_k, P_k : k \in \mathbb{Z}\}$. Since μ is conservative $\Phi_*\mu$ gives zero measure to the images of the boundaries of the rectangles of the Markov partition. The latter property implies that Φ is an isomorphism of $(\mathbb{T}^2, \Phi_*\mu, f)$ and $(\Sigma_{\mathbf{A}}, \mu, Shift)$ and thus $(\mathbb{T}^2, \Phi_*\mu, f)$ is a type III_1 dynamical system.

The type III_1, inhomogeneous Markov measures for the shift on $\Sigma_{\mathbf{A}}$ have the additional property that for every $k \le 0$ the transition matrices of μ at k are the same as the ones arising from the Lebesgue measure ($\forall k \le 0,\ P_k = \mathbf{Q}$). This implies that (after a rotation of the coordinates to the v_u, v_s coordinates) with $\Phi : \Sigma_{\mathbf{A}} \to \mathbb{T}$ being the semiconjugacy map arising from the Markov partition we have

$$d\Phi_*\mu(x, y) = d\nu^+(x)dy.$$

Here ν^+ is the image by the push forward on the stable manifold of the Markov measure on $\{1, 2, 3\}^{\mathbb{N}}$ given by $\{\pi_k, P_k\}_{k=1}^{\infty}$. This property will be used (see Subsection 2.5) to show that there exists an homeomorphism G of $\mathbb{T}^2$ such that $m_{\mathbb{T}^2} \circ G = \Phi_*\mu$ and the transformation $G \circ f \circ G^{-1} : (\mathbb{T}^2, m_{\mathbb{T}^2}) \to (\mathbb{T}^2, m_{\mathbb{T}^2})$ is measure theoretically isomorphic to $(\mathbb{T}^2, \Phi_*\mu, f)$ (1), hence a type III_1 system.

The harder part in the proof of this theorem is to construct a homeomorphism $H : \mathbb{T}^2 \to \mathbb{T}^2$ so that

1. $m \circ H \sim \Phi_*\mu = m \circ G$. Consequently the system $(\mathbb{T}^2, \mathscr{B}_{\mathbb{T}^2}, m_{\mathbb{T}^2}, H \circ f \circ H^{-1})$ is of type III_1 because it is measure theoretically isomorphic to $(\mathbb{T}^2, \mathscr{B}_{\mathbb{T}^2}, m_{\mathbb{T}^2} \circ H, f)$ and the fact that the type III_1 property is invariant upon changing the measure to an equivalent measure.
2. $H \circ f \circ H^{-1}$ is C^1 and Anosov.

In order to obtain this goal and to explain the definition of G it is easier for us to build f as the natural extension of the (non invertible) golden mean shift $Sx = \varphi x \mathrm{mod} 1$.

2.5. The map f as the natural extension of the golden mean shift

The partition $\{J_1 = [0, 1/\varphi^2],\ J_2 = [1/\varphi, 1],\ J_3 = [1/\varphi^2, 1/\varphi]\}$ is a Markov partition for the golden mean shift with $\mathbf{A}$ (the same matrix as the one for f) as its adjacency matrix. See Figure 2.2.

Denote by σ the one sided shift on $\Sigma_{\mathbf{A}}^+$. It can be verified that $(\Sigma_{\mathbf{A}}^+, \nu_{\pi_{\mathbf{Q}}, \mathbf{Q}}, \sigma)$ is isomorphic to $(\mathbb{T}, m_{\mathbb{T}}, S)$ where $m_{\mathbb{T}}$ is the Lebesgue measure on $\mathbb{T}$. The natural extension of $(\Sigma_{\mathbf{A}}^+, \nu_{\pi_{\mathbf{Q}}, \mathbf{Q}}, \sigma)$ is $(\Sigma_{\mathbf{A}}, \mathrm{M}\{\pi_{\mathbf{Q}}, \mathbf{Q}\}, \sigma)$ which is isomorphic to $(\mathbb{T}^2, m_{\mathrm{Leb}}, f)$. This shows that f is indeed the natural extension of the golden mean shift. To see the geometric picture of how S and f are related one can look at the Markov partitions and move to the V_u, V_s coordinates. On those coordinates f acts almost as

$$(u, v) \mapsto (\varphi u \bmod 1, -\varphi^{-1}v) = (Su, -\varphi^{-1}v),$$

(1) The isomorphism $(\mathbb{T}^2, m_{\mathbb{T}^2} \circ G = \Phi_*\mu, f) \xrightarrow{\pi} (\mathbb{T}^2, m_{\mathbb{T}^2}, G \circ f \circ G^{-1})$ is clearly $\pi = G$. Indeed G is a homeomorphism, hence measurable and invertible (and G^{-1} is measurable), $G \circ f = (GfG^{-1}) \circ G$ and $(m_{\mathbb{T}^2} \circ G) \circ G^{-1} = m_{\mathbb{T}^2}$.

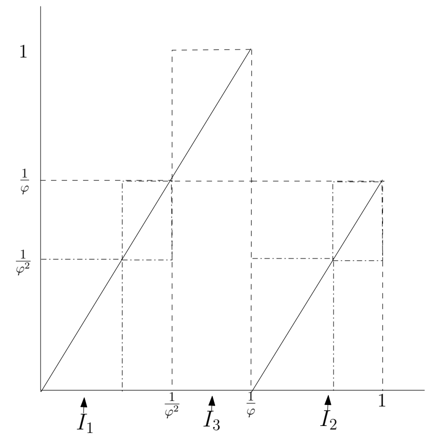


FIGURE 2.2. The Markov partition of $\varphi x \mod 1$

where the mistake is in the second coordinate. To make it precise let

$$\mathbb{M} = [0, 1/\varphi] \times \left[-\varphi/(\varphi+2), \varphi^2/(\varphi+2)\right] \cup [1/\varphi, 1] \times [-\varphi/(\varphi+2), 1/(\varphi+2)].$$

Define $\tilde{f} : \mathbb{M} \to \mathbb{M}$ by

$$\tilde{f}(x, y) = \begin{cases} \left(\varphi x, -\varphi^{-1} y\right), & 0 \le x \le 1/\varphi, \\ \left(\varphi x - 1, -\varphi^{-1}\left(y - \frac{\varphi^2}{\varphi+2}\right)\right), & 1/\varphi \le x \le 1. \end{cases}$$

See Figure 2.3 for the way $\tilde{f}$ maps its 3 rectangles, as can be seen by this picture the action of $\tilde{f}$ is the same as how f acts on its Markov partition.

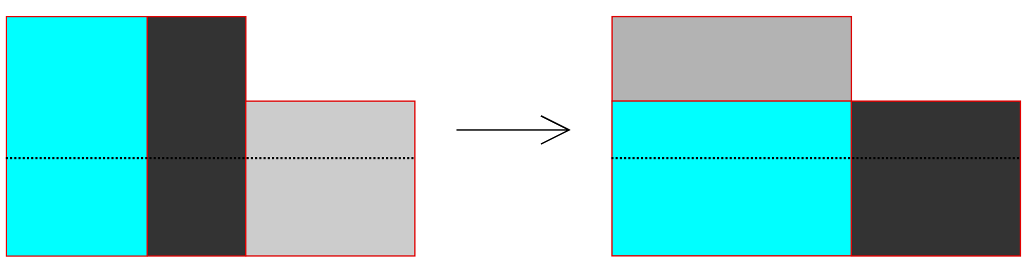


FIGURE 2.3. Action of $\tilde{f}$ on its (soon to be) Markov partition

In order that $\tilde{f}$ will be the same as f, we identify by orientation preserving piecewise translations the following intervals (for a geometric understanding one can see that this identification comes from the way the Markov partition of f tiles the plane):

$$\begin{aligned}
\{0\}\times\left[0,\varphi^2/(\varphi+2)\right] &\simeq \{1\}\times\left[-\varphi/(\varphi+2),1/(\varphi+2)\right]\\
\{0\}\times\left[-\varphi/(\varphi+2),0\right] &\simeq \{1/\varphi\}\times\left[1/(\varphi+2),\varphi^2/(\varphi+2)\right]\\
\left[0,1/\varphi^3\right]\times\left\{\varphi^2/(\varphi+2)\right\} &\simeq \left[1/\varphi^2,1/\varphi\right]\times\left\{-\varphi/(\varphi+2)\right\}\\
\left[1/\varphi^3,1/\varphi\right]\times\left\{\varphi^2/(\varphi+2)\right\} &\simeq \left[1/\varphi,1\right]\times\left\{-\varphi/(\varphi+2)\right\}\\
\left[0,1/\varphi^2\right]\times\left\{-\varphi/(\varphi+2)\right\} &\simeq \left[1/\varphi,1\right]\times\left\{1/(\varphi+2)\right\}.
\end{aligned}$$

The resulting manifold (which is $\mathbb{T}^2$) will be denoted by $\mathbb{M}_\sim$ in order to remind the reader of this change of coordinates and the geometric relation between f and S.

In $\mathbb{M}_\sim$, $d\Phi_*\mu(x,y)=d\nu^+(x)dy$ where ν^+ is a non atomic measure on $\mathbb{T}$. This means that the circle homeomorphism $h(x)=\nu^+[0,x]$ takes the Lebesgue measure on $\mathbb{T}$ to ν^+ and $h(1/\varphi)=\mu(x_1\neq 2)=\frac{1}{\varphi}$. The homeomorphism of $\mathbb{M}_\sim$ defined by $G(x,y)=(h(x),y)$ takes Lebesgue measure of $\mathbb{M}_\sim$ to $\Phi_*\mu$. The perturbed homeomorphism $H:\mathbb{M}_\sim\to\mathbb{M}_\sim$ which will be constructed is of the form $H(x,y)=\left(h_y(x),y\right)$, where for $y\in\left[-\varphi/(\varphi+2),\varphi^2/(\varphi+2)\right]$, $h_y:\mathbb{T}\to\mathbb{T}$ is a circle homeomorphism such that $m_{\mathbb{T}}\circ h_y\sim\nu^+$. This construction is carried out by the following steps:

- the first step is to work on the action of f on the unstable manifold which is the golden mean shift and to construct a circle homeomorphism $\tilde{h}$ such that $\tilde{h}\circ S\circ\tilde{h}^{-1}$ is C^1 expanding and $m_{\mathbb{T}}\circ\tilde{h}\sim\nu^+$. A further important property of the homeomorphisms which we construct is that $\tilde{h}(J_i)=J_i$ for all elements of the Markov partition of S. This will imply for example that $\left(\tilde{h}(x),y\right)$ is an homeomorphism of $\mathbb{M}_\sim$. This step involves adding another parameter for the inductive construction of the measure $\mu=\mathrm{M}\{P_k,\pi_k:\ k\in\mathbb{Z}\}$ and is carried out in Section 4.
- in Section 5 we modify construction of these homeomorphisms $\tilde{h}$ in order to construct the functions h_y in the definition of H. A major challenge in this step is to ensure that $\frac{\partial H\circ\tilde{f}\circ H^{-1}}{\partial y}$ is defined and continuous.

3. The type III_1 Markov shifts supported on Σ_A

Here we present the inductive construction of the inhomogeneous Markov measures.

3.1. Markov Chains

3.1.1. *Basics of Stationary (Homogeneous) Chains.* – Let S be a finite set which we regard as the state space of the chain, $\pi=\{\pi(s)\}_{s\in S}$ a probability vector on S and $\mathrm{P}=(P_{s,t})_{s,t\in S}$ a stochastic matrix. The vector π and P define a Markov chain $\{X_n\}$ on S by

$$\forall n\in\mathbb{Z},\ \ \mathbb{P}_\pi(X_n=t)\,\pi(t)\quad\text{and}\quad \mathbb{P}(X_n=s\mid X_1,\ldots,X_{n-1}):=P_{X_{n-1},s}.$$

P is *irreducible* if for every $s,t\in S$, there exists $n\in\mathbb{N}$ such that $P^n_{s,t}>0$ and P is *aperiodic* if for every $s\in S$, $\gcd\left\{n:\ P^n_{s,s}>0\right\}=1$. Given an irreducible and aperiodic P, there exists a unique stationary probability π_{P} (that is $\pi_P\mathrm{P}=\pi_P$). In addition for every $s,t\in S$,

$P^n_{s,t} \xrightarrow[n\to\infty]{} \pi_P(t)$. Since S is a finite state space, it follows that for any initial distribution π on S,

$$\mathbb{P}_\pi (X_n = t) = \sum_{s\in S} \pi(s) P^n_{s,t} \xrightarrow[n\to\infty]{} \pi_P(t).$$

An important fact which will be used in the sequel is that the stationary distribution is continuous with respect to the stochastic matrix. That is if $\{\mathrm{P_n}\}_{n=1}^\infty$ is a sequence of irreducible and aperiodic stochastic matrices such that

$$\|\mathrm{P}_n - \mathrm{P}\|_\infty := \max_{s,t\in S} \left|(\mathrm{P}_n)_{s,t} - \mathrm{P}_{s,t}\right| \xrightarrow[n\to\infty]{} 0$$

and P is irreducible and aperiodic then $\|\pi_{\mathrm{P}_n} - \pi_{\mathrm{P}}\|_\infty \to 0$.

3.2. Type III_1 Markov Shifts

In this subsection, let $\Omega := \Sigma_{\mathbf{A}}$, $\mathscr{B} := \mathscr{B}_{\Sigma_{\mathbf{A}}}$ and T is the two sided shift on Ω. For two integers $k < l$, write $\mathscr{F}(k,l)$ for the algebra of sets generated by cylinders of the form $[b]_k^l$, $b \in \{1,2,3\}^{l-k}$. That is the smallest σ-algebra which makes the coordinate mappings $\{w_i(x) := x_i : i \in [k,l]\}$ measurable.

3.2.1. *Idea of the construction of the type* III *Markov measure.–* The construction uses the ideas in [11]. For every $j \le 0$

$$P_j \equiv \mathbf{Q} \text{ and } \pi_j \equiv \pi_{\mathbf{Q}},$$

where $\mathbf{Q}$ and $\pi_{\mathbf{Q}}$ are as in (2.1) and (2.2) respectively. On the positive axis one defines on larger and larger chunks the stochastic matrices which depend on a distortion parameter $\lambda_k \ge 1$ where 1 means no distortion. Now a cylinder set $[b]_{-n}^n$ fixes the values of the first n terms in the product form of the Radon Nykodym derivatives. We would like to be able to correct the values in order that we can enforce a given number to be in the ratio set. This corresponds to a lattice condition on λ_k which is less straightforward than the one in [11]. However this is not enough for a Markov measure, since the states are not independent, this forces us to utilize both the convergence to the stationary distribution and the mixing property for stationary chains.

Another difficulty is that the measure of the set $[b]_{-n}^n \cap T^{-N}[b]_{-n}^n \cap \left\{\left(T^N\right)' \approx a\right\}$ could be of very small measure with respect to $\mu\left([b]_{-n}^n\right)$. To remedy this problem, and enable approximation of general sets, we look for many approximately independent such events so that their union covers at least a fixed proportion of $[b]_{-n}^n$.

More specifically the construction goes as follows. We define inductively 5 sequences $\{\lambda_j\}$, $\{m_j\}$, $\{n_j\}$, $\{N_j\}$ and $\{M_j\}$ where

$$\begin{aligned} M_0 &= 1 \\ N_j &:= N_{j-1} + n_j \\ M_j &:= N_j + m_j. \end{aligned}$$

This defines a partition of $\mathbb{N}$ into segments $\left\{\left[M_{j-1}, N_j\right), \left[N_j, M_j\right)\right\}_{j=1}^\infty$. The sequence $\{P_n\}$ equals $\mathbf{Q}$ on the $\left[N_j, M_j\right)$ segments while on the $\left[M_{j-1}, N_j\right)$ segments we have $P_n \equiv \mathbf{Q}_{\lambda_j}$, the λ_j perturbed stochastic matrix. The $\mathbf{Q}$ segments facilitate the form of some of the Radon Nykodym derivatives while the perturbed segments come to ensure that $\mu \perp \mathrm{M}\left\{\pi_{\mathbf{Q}}, \mathbf{Q}\right\}$ and that the ratio set condition is satisfied for cylinder sets.

Notation: By $x = a \pm b$ we mean $a - b \le x \le a + b$.

3.2.2. *The construction.–* For $\lambda \ge 1$ let

$$\mathbf{Q}_{\lambda} := \begin{pmatrix} \frac{\varphi\lambda}{1+\varphi\lambda} & 0 & \frac{1}{1+\varphi\lambda} \\ \frac{\varphi}{1+\varphi} & 0 & \frac{1}{1+\varphi} \\ 0 & 1 & 0 \end{pmatrix}.$$

Choice of the base of induction: Let $M_0 = 1$, $\lambda_1 > 1$, $n_1 = 2$, $N_1 = 3$ and $\mathbf{Q}_1 := \mathbf{Q}_{\lambda_1}$ be the λ_1 perturbed matrix. Set $P_1 = P_2 = \mathbf{Q_1}$ and $\pi_0 = \pi_{\mathbf{Q}}$. The measures π_1, π_2 are then defined by equation (2.3). Let $m_1 = 3$ and thus $M_1 = 6$. Set $P_j = \mathbf{Q}$ for $j \in [N_1, M_1) = [3, 6)$ and π_3, π_4, π_5 be defined by equation (2.3).

Assume that $\{\lambda_j, m_j, n_j, N_j, M_j\}_{j=1}^{l-1}$ have been chosen.

Choice of λ_l. – Notice that the function $f(x) := x\frac{1+\varphi}{1+\varphi x}$ is monotone increasing and continuous in the segment $[1,\infty)$. Therefore we can choose $\lambda_l > 1$ which satisfies the following three conditions:

1. Finite approximation of the Radon-Nykodym derivatives condition:

$$(\lambda_l)^{2m_{l-1}} < e^{\frac{1}{2^l}}. \tag{3.1}$$

   This condition ensures an approximation of the derivatives by a finite product.

2. Lattice condition:

$$\lambda_{l-1} \cdot \frac{1+\varphi}{1+\varphi\lambda_{l-1}} \in \left(\lambda_l \cdot \frac{1+\varphi}{1+\varphi\lambda_l}\right)^{\mathbb{N}}, \tag{3.2}$$

   where $a^{\mathbb{N}} := \{a^n : n \in \mathbb{N}\}$.

3. Let $\mathbf{Q}_l := \mathbf{Q}_{\lambda_l}$ and $\pi_{\mathbf{Q_l}}$ be its unique stationary probability. Notice that when λ_l is close to 1, then $\mathbf{Q_l}$ is close to $\mathbf{Q}$ in the L_∞ sense. Therefore by continuity of the stationary distribution we can demand that

$$\|\pi_{\mathbf{Q}} - \pi_{\mathbf{Q_l}}\|_\infty < \frac{1}{2^l}. \tag{3.3}$$

Choice of n_l. – It follows from the Lattice condition, Equation (3.2), that for each $k \le l-1$,

$$\left(\lambda_k \cdot \frac{1+\varphi}{1+\varphi\lambda_k}\right) \in \left(\lambda_l \cdot \frac{1+\varphi}{1+\varphi\lambda_l}\right)^{\mathbb{N}}.$$

Choose n_l large enough so that for every $k \le l-1$ (notice that the demand on $k = 1$ is enough) there exists $\mathbb{N} \ni p = p(k,l) \le \frac{n_l}{20}$ so that

$$\left(\lambda_l \cdot \frac{1+\varphi}{1+\varphi\lambda_l}\right)^p = \left(\lambda_k \cdot \frac{1+\varphi}{1+\varphi\lambda_k}\right). \tag{3.4}$$

Till now we have defined $\{P_j, \pi_j\}_{j=-\infty}^{M_{l-1}}$. By the mean ergodic theorem for Markov chains [15, Th. 4.16] and (3.3), one can demand by enlarging n_l if necessary that in addition

$$\nu_{\pi_{M_{l-1}},\mathbf{Q_l}}\left(x : \frac{1}{n_l}\sum_{j=1}^{n_l} \mathbf{1}_{[x_j=1]} = \frac{1}{\sqrt{5}} \pm 2^{-l}\right) > 1 - \frac{1}{l}, \tag{3.5}$$

and

$$\nu_{\pi_{M_{l-1}},\mathbf{Q_l}}\left(x:\frac{1}{n_l}\sum_{j=1}^{n_l}\mathbf{1}_{[x_j=2,x_{j+1}=3]}>\frac{1}{15}\right)>1-\frac{1}{l},\tag{3.6}$$

where ν is the Markov measure on $\{1,2,3\}^{\mathbb{N}}$ defined by $\mathbf{Q_l}$ and $\pi_{M_{l-1}}$. The numbers inside the set were chosen so that

$$\left|\pi_{\mathbf{Q_l}}(1)-1/\sqrt{5}\right|<2^{-l},$$

and similarly for l large enough

$$\int\mathbf{1}_{[x_0=2,x_1=3]}(x)d\nu_{\pi_{\mathbf{Q_l}},\mathbf{Q_l}}=\pi_{\mathbf{Q_l}}(2)\,(\mathbf{Q_l})_{2,3}=\left(\frac{1}{\varphi\sqrt{5}}\pm\frac{1}{2^l}\right)\frac{1}{\varphi+1}>\frac{1}{15}.$$

Choice of N_l. – Let $N_l:=M_{l-1}+n_l$. Now set for all $j\in[M_{l-1},N_l)$,

$$P_j=\mathbf{Q_l}$$

and $\{\pi_j\}_{j=M_{l-1}+1}^{N_l}$ be defined by equation (2.3).

Choice of m_l. – Let k_l be the $\left(1\pm\left(\frac{1}{3}\right)^{3N_l}\right)$ mixing time of $\mathbf{Q}$. That is for every $\mathbf{n}>k_l$, $j\in\mathbb{N}$, $A\in\mathscr{F}(0,j)$, $B\in\mathscr{F}(j+\mathbf{n},\infty)$ and initial distribution $\tilde{\pi}$,

$$\nu_{\tilde{\pi},\mathbf{Q}}(A\cap B)=\left(1\pm3^{-3N_l}\right)\nu_{\tilde{\pi},\mathbf{Q}}(A)\,\nu_{\pi_{\mathbf{Q}},\mathbf{Q}}\left(T^{\mathbf{n}+l}B\right).\tag{3.7}$$

Demand in addition that $k_l>N_l$. Let m_l be large enough so that

$$\left(1-9^{-3N_l}\right)^{m_l/4k_l}\le\frac{1}{l},\tag{3.8}$$

and

$$(m_l-N_l)\,\lambda_1^{-2N_l}\ge1.\tag{3.9}$$

To summarize the construction. We have defined inductively sequences $\{n_l\}$, $\{N_l\}$, $\{m_l\}$, $\{M_l\}$ of integers which satisfy

$$M_l<N_{l+1}=M_l+n_l<M_{l+1}=N_{l+1}+m_{l+1}.$$

In addition we have defined a monotone decreasing sequence $\{\lambda_l\}$ which decreases to 1 and using that sequence we defined new stochastic matrices $\{\mathbf{Q_l}\}$. Now we set

$$P_j:=\begin{cases}\mathbf{Q}, & j\le0\\ \mathbf{Q_l}, & M_{l-1}\le j<N_l\\ \mathbf{Q}, & N_l\le j<M_l,\end{cases}\tag{3.10}$$

and $\pi_j=\pi_{\mathbf{P}}$ for $j\le0$. The rest of the π_j's are defined by the consistency condition, equation (2.3). Finally let μ be the Markovian measure on $\{1,2,3\}^{\mathbb{Z}}$ defined by $\{\pi_j,P_j\}_{j=-\infty}^{\infty}$.

Notice that for all $j\in\mathbb{N}$, $\operatorname{supp}P_j\equiv\operatorname{supp}A=\operatorname{supp}\mathbf{Q}$.

THEOREM 4. – *The shift* $\left(\{1,2,3\}^{\mathbb{Z}},\mu,T\right)$ *is nonsingular,conservative, ergodic and of type* III_1.

The proof of Theorem 4 is given in the appendix.

4. Type III perturbation of the golden mean shift arising from Markovian measures

4.1. A perturbation of the golden mean shift

Let $\nu = \mathrm{M}\{\pi_k, \mathrm{P}_k\}_{k=-\infty}^{\infty}$ be the type III_1 (for the shift on $\{1,2,3\}^{\mathbb{Z}}$) Markov measure from Section 3 for the two sided shift. It follows from [19, Thm. 4.4.] that the one sided Markov measure $\nu^+ = \mathrm{M}\{\pi_k, \mathrm{P}_k\}_{k=1}^{\infty}$ on $\{1,2,3\}^{\mathbb{N}}$ is a type III measure for the (one sided) shift.

Let $Sx = \varphi x \mod 1$ and $J_1 := \left(0, 1/\varphi^2\right)$, $J_2 := (1/\varphi, 1)$ and $J_3 := \left(1/\varphi^2, 1/\varphi\right)$ be a Markov partition for S. Denote by

$$\mathrm{Bd}\,(S) := \bigcup_{n=0}^{\infty}\bigcup_{i=1}^{3} \partial\left(S^{-n}J_i\right).$$

The map $\Theta : \Sigma_{\mathbf{A}}^{+} \to [0,1]$, $\Theta(w) = \bigcap_{n=0}^{\infty} \overline{S^{-n}J_{w_n}}$ is a semiconjugacy of $\left(\Sigma_{\mathbf{A}}^{+}, \sigma\right)$ and $(\mathbb{T}, S)$ and for each $x \notin \mathrm{Bd}\,(S)$, $\Theta^{-1}(x)$ consists of one point (point of uniqueness for the Θ representation). Since the support of ν^+ is contained in $\mathrm{G}\,(\sigma) := \Theta^{-1}\left(\mathbb{T}\backslash \mathrm{Bd}\,(S)\right)$, the map Θ is a metric isomorphism between $\left(\Sigma_{\mathbf{A}}^{+}, \nu^+, \sigma\right)$ and $\left(\mathbb{T}, \Theta_*\left(\nu^+\right), S\right)$ and therefore the measure $\mu^+ := \Theta_*\left(\nu^+\right)$ is a type III measure for S. Since μ^+ is a continuous measure, its cumulative distribution function $\mathfrak{h}(x) = \mu^+([0,x])$ is a homeomorphism of $\mathbb{T}$ such that $\mu^+ \circ \mathfrak{h}^{-1}$ is Lebesgue measure on $\mathbb{T}$. It follows that the map $\left(\mathbb{T}, m_{\mathbb{T}}, \mathfrak{h}\circ S\circ \mathfrak{h}^{-1}\right)$ is a type III transformation, where $m_{\mathbb{T}}$ denotes the Lebesgue measure. The problem is that $\mathfrak{h}\circ S\circ\mathfrak{h}^{-1}$ is not necessarily smooth, so we construct $\mathfrak{h}_\epsilon$, as in the idea of the examples of Bruin and Hawkins, close to $\mathfrak{h}$ in the C^0 norm such that

- $\mathfrak{h}_\epsilon \circ S \circ \mathfrak{h}_\epsilon^{-1}$ is C^1 and uniformly expanding.
- $m_{\mathbb{T}} \circ \mathfrak{h}_\epsilon \sim \mu^+$.
- We will have in addition that $\mathfrak{h}(J_i) = J_i$ for every $i \in \{1,2,3\}$, this extra property is crucial for the extension to two dimensions.

Before we go through the construction we would like the reader to recall that the Lebesgue measure on $\mathbb{T}$ is the measure arising from $\mathrm{M}\{\pi, \mathbf{Q}\}$. The main idea is to approximate the change of measure between Lebesgue measure and μ^+ on the semi algebras

$$\mathscr{R}(n) := \left\{ C_{[w]_1^n} := \bigcap_{k=0}^{n-1} S^{-k}J_{w_k} : x\in\Sigma_{\mathbf{A}} \right\} = \left\{C_w :\ w\in\Sigma_{\mathbf{A}}(n)\right\}.$$

The construction goes as follows: We first assume that we are given a type III Markovian measure defined by $\{\lambda_k, M_k, N_k\}_{k=1}^{\infty}$. Then we would like to choose inductively, mostly by continuity arguments a sequence $\underline{\epsilon} = \{\epsilon_k\}$ that will give us the perturbation. However in the end we arrive at a problem, namely that we need that the size of M_k is relatively large with respect to $1/\epsilon_{k-1}$. This problem will be solved by modifying the induction process of Section 3 and adding the choice of the sequence $\underline{\epsilon}$ to the induction. The new induction will be explained in Subsection 4.2.2.

Remark 5. – Before we continue with the construction we would like to remind the reader that at each stage in the inductive construction of the Markovian measure in Section 3 we can take λ_t to be as close to 1 as we like and $n_t,\ M_t/N_t$ to be as large as we want. This is because the conditions on λ_t ((3.1), (3.2) and (3.3)) are that λ_t is small enough whilst the conditions on n_t ((3.4), (3.5) and (3.6)) and $M_t/n_t \sim m_t/n_t$ ((3.8) and (3.9)) are to be large enough.

Special interpolation functions. – Given $\alpha > 0$ we would like to define a Lipschitz function g_α so that $g_\alpha(0) = 0$, $g'_\alpha(0) = 1$, $g_\alpha(1) = \int_0^1 g'_\alpha(x)dx = \alpha$ and $g'_\alpha(1) = \alpha$. We will use the functions $g_\alpha : [0,1] \to [0,\alpha]$ defined by $g_\alpha(0) = 0$ and

$$g'_\alpha(x) = \begin{cases} 1 + 3x \cdot \frac{5\alpha-5}{4}, & 0 \le x \le \frac{1}{3}, \\ \frac{5\alpha-1}{4}, & \frac{1}{3} \le x \le \frac{2}{3}, \\ \frac{5\alpha-1}{4} - (3x-2)\frac{\alpha-1}{4}, & \frac{2}{3} \le x \le 1, \end{cases}$$

which have the additional property that if $\alpha > 1$ then

$$1 = \inf_{x\in[0,1]} g'_\alpha(x) < \sup_{x\in[0,1]} g'_\alpha(x) = \frac{5\alpha-1}{4} < \alpha^2$$

and if $\frac{1}{4} \le \alpha < 1$ then

$$\alpha^2 \le \frac{5\alpha-1}{4} = \inf_{x\in[0,1]} g'_\alpha(x) < \sup_{x\in[0,1]} g'_\alpha(x) = 1.$$

REMARK 6. – For all $\alpha, \epsilon > 0$,

$$\int_0^\epsilon g'_\alpha\left(\frac{x}{\epsilon}\right)dx = \epsilon\int_0^1 g'_\alpha(x)dx = \epsilon\alpha$$

and for all $u > \epsilon$,

$$\int_{u-\epsilon}^u g'_\alpha\left(\frac{u-x}{\epsilon}\right)dx = \epsilon\alpha.$$

4.2. Realization of the homeomorphism of change of measures

For $0 < \epsilon < \frac{1}{\varphi}$ and $\lambda > 1$, let $\psi_{\epsilon,\lambda} : [0,1] \circlearrowleft$ be the function defined by $\psi_{\epsilon,\lambda}(0) = 0$ and

$$\psi'_{\epsilon,\lambda}(x) := \begin{cases} g'_{\left(\frac{\lambda\varphi^2}{1+\lambda\varphi}\right)}\left(\frac{x}{\epsilon}\right), & 0 \le x \le \epsilon, \\ \frac{\lambda\varphi^2}{1+\lambda\varphi}, & \epsilon < x \le \frac{1}{\varphi} - \epsilon, \\ g'_{\left(\frac{\lambda\varphi^2}{1+\lambda\varphi}\right)}\left(\frac{1/\varphi - x}{\epsilon}\right), & \frac{1}{\varphi} - \epsilon < x \le \frac{1}{\varphi}, \\ g'_{\left(\frac{\varphi^2}{1+\lambda\varphi}\right)}\left(\frac{x-1/\varphi}{\epsilon}\right), & \frac{1}{\varphi} \le x \le \frac{1}{\varphi} + \epsilon, \\ \frac{\varphi^2}{1+\lambda\varphi}, & \frac{1}{\varphi} + \epsilon < x \le 1 - \epsilon, \\ g'_{\left(\frac{\varphi^2}{1+\lambda\varphi}\right)}\left(\frac{1-x}{\epsilon}\right), & 1 - \epsilon < x \le 1. \end{cases}$$

If $\epsilon = 0$ then by a rescaling procedure one can use these functions to define the cumulative distribution function of $\Theta_*\left(\nu_{\pi_{\mathbf{Q}_\lambda}, \mathbf{Q}_\lambda}\right)$. The function $\psi_{\epsilon,\lambda}$ is basically an interpolation of a piecewise constant function in order to make it continuous and that the following properties hold:

1. $\psi'_{\epsilon,\lambda}(0) = \psi'_{\epsilon,\lambda}(1) = 1$. This is needed in order to glue $\psi_{\epsilon,\lambda}$ with the identity function and still have a C^1 function.

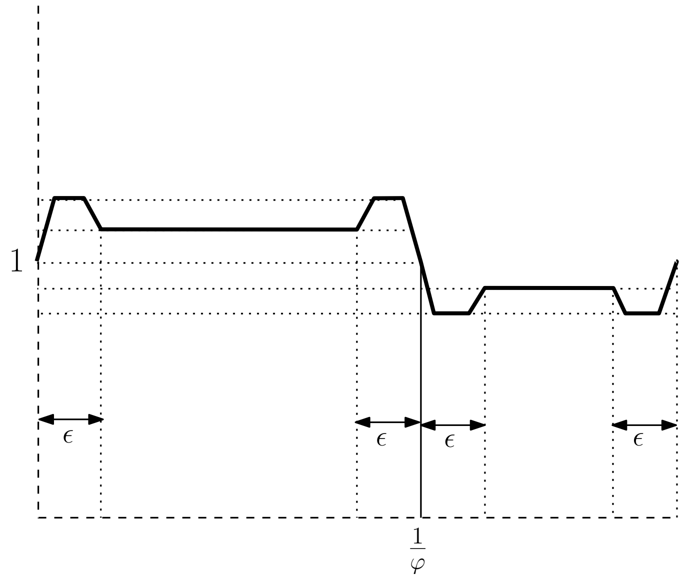


FIGURE 4.1. The graph of $\psi'_{\epsilon,\lambda}$

2. For every ϵ, λ, by Remark 6,

$$\psi_{\epsilon,\lambda}\left(\frac{1}{\varphi}\right) = \left[\int_0^{\epsilon} g'_{\left(\frac{\lambda\varphi^2}{1+\lambda\varphi}\right)}\left(\frac{x}{\epsilon}\right) dx + \frac{\lambda\varphi^2}{1+\lambda\varphi}\left(\frac{1}{\varphi} - 2\epsilon\right)\right.$$
$$\left. + \int_{\frac{1}{\varphi}-\epsilon}^{\frac{1}{\varphi}} g'_{\left(\frac{\lambda\varphi^2}{1+\lambda\varphi}\right)}\left(\frac{1/\varphi - x}{\epsilon}\right) dx\right] = \frac{\lambda\varphi}{1+\lambda\varphi}.$$

Similarly $\psi_{\epsilon,\lambda}(1) = \psi_{\epsilon,\lambda}\left(\frac{1}{\varphi}\right) + \left(\psi_{\epsilon,\lambda}(1) - \psi_{\epsilon,\lambda}\left(\frac{1}{\varphi}\right)\right) = 1.$

3. By Remark 6, $\psi_{\epsilon,\lambda}(\epsilon) = \frac{\lambda\varphi^2}{1+\lambda\varphi} \cdot \epsilon$. Thus for every $\epsilon < x < \frac{1}{\varphi} - \epsilon$,

$$\psi_{\epsilon,\lambda}(x) = \left(\psi_{\epsilon,\lambda}(x) - \psi_{\epsilon,\lambda}(\epsilon)\right) + \psi_{\epsilon,\lambda}(\epsilon) = \frac{\lambda\varphi^2}{1+\lambda\varphi}x,$$

and

$$\frac{\psi_{\epsilon,\lambda}(x)}{\psi_{\epsilon,\lambda}(1/\varphi)} = \varphi x.$$

Similarly, $\psi_{\epsilon,\lambda}\left(\frac{1}{\varphi} + \epsilon\right) - \psi_{\epsilon,\lambda}\left(\frac{1}{\varphi}\right) = \frac{\varphi^2}{1+\lambda\varphi}\epsilon$, thus for every $\frac{1}{\varphi} + \epsilon < x < 1 - \epsilon$,

$$\frac{\psi_{\epsilon,\lambda}(x) - \psi_{\epsilon,\lambda}(1/\varphi)}{\psi_{\epsilon,\lambda}(1) - \psi_{\epsilon,\lambda}(1/\varphi)} = \varphi^2\,(x - 1/\varphi) = \frac{x - 1/\varphi}{1 - 1/\varphi}.$$

4. $\psi'_{\epsilon,\lambda}$ is Lipschitz with Lipschitz constant of the order $1/\epsilon$ when $\epsilon \to 0$ and for every $x \in \mathbb{T}$,

$$\lambda^{-2} \leq \psi'_{\epsilon,\lambda}(x) < \lambda^2. \tag{4.1}$$

Given two sequences $\epsilon_k \geq 0$ and $\lambda_k \geq 1$, let ψ_k denote $\psi_{\lambda_k,\epsilon_k}$.

Define an order on Σ_A^+ in the following way. For $w, z \in \Sigma_A^+$, let

$$j(w,z) := \inf\{n \in \mathbb{N} : \ w_n \neq z_n\}.$$

Then $w \prec z$ if either $w_{j(w,z)} = 1$ or $w_{j(w,z)} = 3$ and $z_{j(w,z)} = 2$ (notice that in the latter case $j(w,z) = 1$). This order has the following property. If $[w]_1^n \neq [y]_1^n$ for some $n \in \mathbb{N}$, then $C_{[w]_1^n}$ is to the left of $C_{[y]_1^n}$ if and only if $w \prec y$.

In addition for $n \in \mathbb{N}$ we write $\bar{x}_n, \underline{x}_n : \Sigma_{\mathbf{A}} \to \mathbb{T}$ to be defined by

$$C_{[w]_1^n} := [\underline{x}_n(w), \bar{x}_n(w)).$$

For $n \in \mathbb{N}$, denote by $\Sigma_{\mathbf{A}}(n)$ the collection of words $w = w_1 w_2 \cdots w_n$ with $[w]_1^n \subset \Sigma_{\mathbf{A}}$.

We will define inductively a sequence $\{h_n\}_{n=1}^{\infty}$ of diffeomorphisms of $\mathbb{T}$. Since $\mathbb{T} = \bigcup_{w \in \Sigma_{\mathbf{A}}(n)} C_{[w]_1^n}$ and each $h_k,\ k < n$ is onto $\mathbb{T}$,

$$\mathbb{T} = \bigcup_{w \in \Sigma_{\mathbf{A}}(n)} H_{n-1}\left(C_{[w]_1^n}\right),$$

where $H_{n-1} := h_{n-1} \circ h_{n-2} \circ \cdots \circ h_1$.

— If $N_t < n < M_t$ for some $t \in \mathbb{N}$, then h_n is the identity.

— If $M_{t-1} < n \leq N_t$ for some $t \in \mathbb{N}$, then h_n is made from $\#\Sigma_{\mathbf{A}}(n)$ scalings of ψ_t or the identity. Let $w(n,1), \ldots, w(n, \#\Sigma_{\mathbf{A}}(n))$ be an enumeration of $\Sigma_{\mathbf{A}}(n)$ with respect to $\prec$. Set $h_n(0) = 0$. Assume we have defined h_n on $\bigcup_{k=1}^{l-1} H_{n-1}\left(C_{w(n,k)}\right)$, we will now define h_n on $H_{n-1}\left(C_{w(n,l)}\right)$.

  — If $w(n,l)_n = 1$, we define for $z \in H_{n-1}\left(C_{w(n,l)}\right)$,

$$h_n(z) := H_{n-1}\left(\underline{x}_n(w)\right) + \mathfrak{l}(n,w)\psi_t\left(\frac{z - H_{n-1}\left(\underline{x}_n(w)\right)}{\mathfrak{l}(n,w)}\right),$$

  where $w = w(n,l)$ and

$$\mathfrak{l}(n,w) := m_{\mathbb{T}}\left(H_{n-1}(w)\right) = H_{n-1}\left(\bar{x}_n(w)\right) - H_{n-1}\left(\underline{x}_n(w)\right).$$

  — If $w(n,l)_n \neq 1$ then for all $z \in H_{n-1}\left(C_{w(n,l)}\right)$,

$$h_n(z) = z.$$

— Note that since we have $\psi_t(1) = 1$ for all $t \in \mathbb{N}$, it follows that $h_n\left(H_{n-1}\left(C_{w(n,l)}\right)\right) = H_{n-1}\left(C_{w(n,l)}\right)$ for all n and l. Consequently, h_n is continuous. The differentiability of h_n at points $\{H_{n-1}\left(\underline{x}_n(w)\right) : w \in \Sigma_{\mathbf{A}}(n)\}$ follows from $\psi_t'(0) = \psi_t'(1) = 1$.

We need to define h_n for all $n \in \{M_t\}_{t=1}^{\infty}$. Here we apply a statistical correction procedure which we will now proceed to describe. In what follows we assume that ϵ_1 is small enough so that

$$m_{\mathbb{T}}\left(\psi_1\left(C_{[w]_1^{N_1}}\right)\right) = m_{\mathbb{T}}\left(\psi_1\left(C_{[w]_1^2}\right)\right) m_{\mathbb{T}}\left(C_{[w]_3^{N_1}} \middle| C_{[w]_1^2}\right).$$

The first equality follows from property 3 of ψ_t provided that ϵ_1 is small enough so that for every $w \in \Sigma_{\mathbf{A}}^+$, the end points of $C_{[w]_1^{N_1}}$ are in $\left[\epsilon_1, \varphi^{-1} - \epsilon_1\right] \cup \left[\varphi^{-1} + \epsilon_1, 1 - \epsilon_1\right] \cup \{0, \varphi^{-1}, 1\}$. The equality then follows from $\psi_1(1/\varphi) = \frac{\lambda_1 \varphi}{1 + \lambda_1 \varphi}$. This relation gives for example that

$$m_{\mathbb{T}}\left(H_{N_1}\left(C_{[w]_1^{N_1}}\right)\right) = \mu^+\left(C_{[w]_1^{N_1}}\right),$$

and we have good knowledge of where the point in $\frac{1}{\varphi}$ proportion in $H_{n-1}\left(C_{[w]_1^n}\right)$ travels. However, since M_t is generally much larger than N_t we loose this control and the useful equality

$$m_{\mathbb{T}}\left(H_{M_t}\left(C_{[w]_1^{N_t+1}}\right)\right) = m_{\mathbb{T}}\left(H_{M_t}\left(C_{[w]_1^{M_t}}\right)\right) m_{\mathbb{T}}\left(C_{[w]_{M_t+1}^{N_t+1}} \middle| C_{[w]_1^{M_t}}\right) \tag{4.2}$$

needs no longer to hold true. The role of h_{M_t} is to take care that equality (4.2) holds true.

The function H'_{N_t} being a product of bounded Lipschitz functions, is a bounded Lipschitz function. Therefore if M_t is large enough with respect to N_t, then (here we use the fact that $h_n = \mathrm{Id}$ for $N_t < n < M_t$) $H'_{N_t} = H'_{M_t-1}$ is almost constant on $H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)$. That means that for every $0 \neq x \in H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)$ in the interior of $H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)$,

$$\left|\frac{m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)}{\int_{C_{[w]_1^{M_t}}} H'_{M_t-1}(s)ds} \cdot H'_{M_t-1}(x) - 1\right| \ll 1.$$

By using a similar idea as in the construction of ψ with the g_α we define h_{M_t} restricted to $H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)$ so that equality (4.2) holds. This is done as follows: for $\alpha_1, \alpha_2 \in \mathbb{R}$, let $G_{\alpha_1,\alpha_2} : [0,1] \to [0,\alpha_2]$ be defined by $G_{\alpha_1,\alpha_2}(0) = 0$ and

$$G'_{\alpha_1,\alpha_2}(x) := \begin{cases} \alpha_1 + \frac{15(\alpha_2-\alpha_1)}{4}x, & 0 \le x \le 1/3, \\ \frac{5\alpha_2-\alpha_1}{4}, & 1/3 \le x \le 2/3, \\ \frac{5\alpha_2-\alpha_1}{4} + \frac{\alpha_2-\alpha_1}{4}(3x-1) & 2/3 \le x \le 1. \end{cases} \tag{4.3}$$

This function is a C^1 function which satisfies $G'_{\alpha_1,\alpha_2}(0) = \alpha_1$ and $G'_{\alpha_1,\alpha_2}(1) = G_{\alpha_1,\alpha_2}(1) = \alpha_2$.

Define $\alpha : \mathbb{N} \times \Sigma_{\mathbf{A}} \to (0,\infty)$ by

$$\alpha(t,w) := \frac{1}{m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)} \int_{C_{[w]_1^{M_t}}} H'_{M_t-1}(s)ds = \frac{m_{\mathbb{T}}\left(H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)\right)}{m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)}.$$

In addition for a finite word $w \in \Sigma_{\mathbf{A}}(M_t)$ we denote by w^- the predecessor of w with respect to $\prec$ restricted on $\Sigma_{\mathbf{A}}(M_t)$. We define $h'_{M_t} \circ H_{M_t-1}(x)$ on $C_{[w]_1^{M_t}}$ to be equal to $\frac{\alpha(t,w)}{H'_{M_t-1}(x)}$ off an $\epsilon_{t+1} m_{\mathrm{Leb}}\left(C_{[w]_1^{M_t}}\right)$ neighborhood of the left end point of the segment $C_{[w]_1^{M_t}}$, $\frac{\alpha(t,w^-)}{H'_{M_{t-1}}(x)}$ on the left endpoint (which is in the boundary of $C_{[w]_1^{M_t}}$) and an interpolation in between by using G_{α_1,α_2} for an appropriately chosen α_1, α_2. Here ϵ_{t+1} has to be small enough so that the end points of $\left\{H_{M_t-1}\left(C_{[w]_1^{N_t+1}}\right) : [w]_1^{N_t+1} \in \Sigma_{\mathbf{A}}(N_{t+1})\right\}$ are not in an ϵ_{t+1} neighborhood of the left end point of $H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)$. Formally $h_{M_t} \circ H_{M_t-1}|_{C_{[w]_1^{M_t}}}$ is defined by

$h_{M_t} \circ H_{M_t-1}\left(\underline{x}_{M_t}(w)\right) = H_{M_t-1}\left(\underline{x}_{M_t}(w)\right)$ and

$$h'_{M_t} \circ H_{M_t-1}(x) = \frac{1}{H'_{M_t-1}(x)} \cdot \begin{cases} G'_{\alpha(t,w^-),\alpha(t,w)}\left(\frac{x-\underline{x}_{M_t}(w)}{\epsilon_{t+1}\, m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)}\right), & \underline{x}_{M_t}(w) \le x < \hat{x}_{M_t}(w), \\ \alpha(t,w), & \hat{x}_{M_t}(w) \le x < \bar{x}_{M_t}(w), \end{cases}$$

where $\hat{x}_{M_t}(w) = \underline{x}_{M_t}(w) + \epsilon_{t+1} m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)$. It follows from the chain rule that for $x \in C_{[w]_1^{M_t}}$,

$$H'_{M_t}(x) = \begin{cases} G'_{\alpha(t,w^-),\alpha(t,w)}\left(\frac{x-\underline{x}_{M_t}(w)}{\epsilon_{t+1}\, m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)}\right), & \underline{x}_{M_t}(w) \le x < \hat{x}_{M_t}(w), \\ \alpha(t,w), & \hat{x}_{M_t}(w) \le x < \bar{x}_{M_t}(w). \end{cases}$$

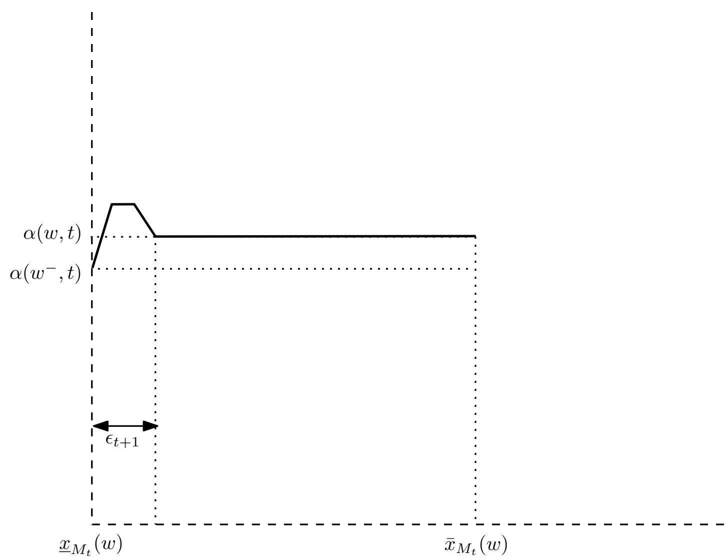


FIGURE 4.2. The graph of H'_{M_t} restricted to $C_{[w]_1^{M_t}}$ when $\alpha(t,w) > \alpha(t,w^-)$

CLAIM 7. – There exists δ_{t+1} such that if $\epsilon_{t+1} < \delta_{t+1}$ then:
(a) for all $w \in \Sigma_{\mathbf{A}}$ and $M_t \le n \le N_{t+1}$

$$H_{M_t}\left(C_{[w]_1^{M_t}}\right) = H_{M_t-1}\left(C_{[w]_1^{M_t}}\right) = H_{N_t}\left(C_{[w]_1^{M_t}}\right).$$

(b) Equation (4.2) holds.

Proof. – Let δ_{t+1} be small enough so that the end points of

$$\left\{H_{M_t-1}\left(C_{[w]_1^{N_{t+1}}}\right) : [w]_1^{N_{t+1}} \in \Sigma_{\mathbf{A}}\left(N_{t+1}\right)\right\}$$

are not in a δ_{t+1} neighborhood of $\left\{H_{M_t-1}\left(\underline{x}_{M_t}(\tilde{w})\right): \tilde{w} \in \Sigma_{\mathbf{A}}(M_t)\right\}$. As a consequence $H'_{M_t}(x) = \alpha(t,w)$ for all $w \in \Sigma_A$ and $x \in \left\{\underline{x}_{N_{t+1}}(w), \bar{x}_{N_{t+1}}(w)\right\}$.

Fix $w \in \Sigma_{\mathbf{A}}$ and $\epsilon_{t+1} \le \delta_{t+1}$, we first prove (a). Write for convenience $\underline{x} = \underline{x}_{M_t}(w)$, $\bar{x} = \bar{x}_{M_t}(w)$ and $\hat{x} = \underline{x} + \epsilon_{t+1}(\bar{x} - \underline{x}) = \underline{x}_{M_t}(w) + \epsilon_{t+1} m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)$. In this notation $\hat{x} - \underline{x} = \epsilon_{t+1} m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)$ and [(2)] we have

$$H_{M_t-1}\left(C_{[w]_1^{M_t}}\right) = \left(H_{M_t-1}(\underline{x}), H_{M_t-1}(\bar{x})\right) = H_{M_t-1}(\underline{x}) + \left(0, H_{M_t-1}(\bar{x}) - H_{M_t-1}(\underline{x})\right).$$

In addition, since $h_{M_t} \circ H_{M_{t-1}}(\underline{x}) = H_{M_{t-1}}(\underline{x})$, then

$$\begin{aligned} H_{M_t}\left(C_{[w]_1^{M_t}}\right) &= H_{M_t}(\underline{x}) + \left(0, H_{M_t}(\bar{x}) - H_{M_t}(\underline{x})\right) \\ &= H_{M_t-1}(\underline{x}) + \left(0, H_{M_t}(\bar{x}) - H_{M_t}(\underline{x})\right). \end{aligned}$$

This shows that (a) is equivalent to showing that

$$H_{M_t}(\bar{x}) - H_{M_t}(\underline{x}) = H_{M_t-1}(\bar{x}) - H_{M_t-1}(\underline{x}).$$

Now

$$\begin{aligned} H_{M_t}(\bar{x}) - H_{M_t}(\underline{x}) &= \int_{\underline{x}}^{\bar{x}} H'_{M_t}(s)ds \\ &= \int_{\underline{x}}^{\hat{x}} G'_{\alpha(t,w^-),\alpha(t,w)}\left(\frac{s-\underline{x}}{\hat{x}-\underline{x}}\right) ds + \alpha(t,w)(\bar{x}-\hat{x}) \\ &= \int_0^{\hat{x}-\underline{x}} G'_{\alpha(t,w^-),\alpha(t,w)}\left(\frac{s}{\hat{x}-\underline{x}}\right) ds + \alpha(t,w)(\bar{x}-\hat{x}). \end{aligned}$$

For all $\alpha_1, \alpha_2, \delta > 0$, $\int_0^\delta G'_{\alpha_1,\alpha_2}\left(\frac{x}{\delta}\right) dx = \delta\alpha_2$. Whence

$$\begin{aligned} H_{M_t}(\bar{x}) - H_{M_t}(\underline{x}) &= \alpha(t,w)(\hat{x}-\underline{x}) + \alpha(t,w)(\bar{x}-\hat{x}) \\ &= \alpha(t,w) m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right) \\ &= H_{M_t-1}(\bar{x}) - H_{M_t-1}(\underline{x}), \end{aligned}$$

we have finished the proof of part (a).

To see part (b) notice that if $\underline{x} \notin C_{[w]_1^{N_{t+1}}}$ then H_{M_t} restricted to $C_{[w]_1^{N_{t+1}}}$ is linear with slope $\alpha(t,w)$. This shows that

$$\begin{aligned} m_{\mathbb{T}}\left(H_{M_t}\left(C_{[w]_1^{M_t}}\right)\right) &= \alpha(t,w) m_{\mathbb{T}}\left(C_{[w]_1^{N_{t+1}}}\right) \\ &= m_{\mathbb{T}}\left(H_{M_t}\left(C_{[w]_1^{M_t}}\right)\right) \frac{m_{\mathbb{T}}\left(C_{[w]_1^{N_{t+1}}}\right)}{m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)} \\ &= m_{\mathbb{T}}\left(H_{M_t}\left(C_{[w]_1^{M_t}}\right)\right) m_{\mathbb{T}}\left(C_{[w]_{M_t+1}^{N_{t+1}}} \,\Big\lfloor\, C_{[w]_1^{M_t}}\right), \end{aligned}$$

(2) For an interval I and a point x, $x + I = \{x + y : y \in I\}$.

as required. If $\underline{x} \in C_{[w]_1^{N_t+1}}$ then $C_{[w]_1^{N_t+1}} = \left[\underline{x}, \overline{x}_{N_{t+1}}(w)\right)$ and thus as in the proof of part (a)

$$\begin{aligned} m_{\mathbb{T}}\left(H_{M_t}\left(C_{[w]_1^{N_t+1}}\right)\right) &= H_{M_t}\left(\bar{x}_{N_{t+1}}(w)\right) - H_{M_t}\left(\underline{x}\right) \\ &= \int_0^{\hat{x}-\underline{x}} G'_{\alpha(t,w^-),\alpha(t,w)}\left(\frac{s}{\hat{x}-\underline{x}}\right) ds + \alpha(t,w)\left(\bar{x}_{N_{t+1}}(w) - \hat{x}\right) \\ &= \alpha(t,w)\left(\bar{x}_{N_{t+1}(w)} - \underline{x}\right) = \alpha(t,w) m_{\mathbb{T}}\left(C_{[w]_1^{N_t+1}}\right). \end{aligned}$$

Continuing as in the case $\underline{x} \notin C_{[w]_1^{N_t+1}}$ one arrives at the conclusion. □

REMARK 8. – An important feature of this construction that will be used in the extension to two dimensions is that for any $1 \leq l \leq \#\Sigma_{\mathbf{A}}(n)$,

$$h_n\left(H_{n-1}\left(C_{w(n,l)}\right)\right) = H_{n-1}\left(C_{w(n,l)}\right). \tag{4.4}$$

This in turn implies that for every $n \in \mathbb{N}$, $\mathscr{H}_n(x,y) := (H_n(x), y)$ is a diffeomorphism of $\mathbb{M}_\sim$ and the Markov partition $\{R_1, R_2, R_3\}$ for $\tilde{f}$ defined by

$$R_i := \begin{cases} J_i \times \left[-\frac{\varphi}{\varphi+2}, \frac{\varphi^2}{\varphi+2}\right], & i \in \{1,3\}, \\ J_2 \times \left[-\frac{\varphi}{\varphi+2}, \frac{1}{\varphi+2}\right], & i = 2. \end{cases}$$

is preserved by $\mathscr{H}_n$.

THEOREM 9. – *There exists a choice of* $\lambda_k \downarrow 1$, $\{n_k, m_k, N_k, M_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ *and* $\underline{\epsilon} = \{\epsilon_k\}_{k\in\mathbb{N}}$ *so that:*

(i) *The Markov measure from the construction of Section 3 is a type* III_1 *measure for the shift on* $\Sigma_{\mathbf{A}}$.

(ii) *The function* $\mathfrak{h}_{\underline{\epsilon}}$ *is a circle homeomorphism and we have* $m_{\mathbb{T}} \circ \mathfrak{h}_{\underline{\epsilon}} \sim \mu^+$, *where* $\mu^+ = \Theta_* M\{P_k, \pi_k\}_{k=1}^{\infty}$.

(iii) *The function* $\mathfrak{g} = \mathfrak{h} \circ S \circ \mathfrak{h}^{-1}$ *is* C^1, *and for every* $x \in \mathbb{T}$,

$$1.6 \leq \mathfrak{g}'(x) \leq 1.7.$$

The proof of this theorem is by showing that we can realize smoothly the inductive construction of Section 3 (with three extra conditions) and include a new sequence $\{\epsilon_k\}$ in it so that the following properties hold:

1. $\mathfrak{h}_\epsilon := \lim_{n\to\infty} H_n$ is a homeomorphism of $\mathbb{T}$.
2. $\mathfrak{g}_n := H_n \circ S \circ H_n^{-1}$ is a convergent sequence in the C^1 topology.
3. The limit function $\mathfrak{g} = \lim_{n\to\infty} \mathfrak{g}_n = \mathfrak{h}_\epsilon \circ S \circ \mathfrak{h}_\epsilon^{-1}$ satisfies $1.6 \leq \mathfrak{g}'(x) \leq 1.7$.
4. $m_{\mathbb{T}} \circ \mathfrak{h}_\epsilon \sim \mu^+$.

4.2.1. *The inductive choice of* $\{\epsilon_l\}_{l=1}^{\infty}$.– Before we continue we would like to set up some notations which will be used.

— Given $\underline{\epsilon} = \{\epsilon_k\}_{k=1}^{t}$ and $n \leq N_t$ we denote by $h_{\underline{\epsilon},n}$ the function in the construction with the sequence $\underline{\epsilon}$ at level n.

— For $j \leq N_t$, $H_{\underline{\epsilon},j} := h_{\underline{\epsilon},j} \circ h_{\underline{\epsilon},j-1} \circ \cdots \circ h_{\underline{\epsilon},1}$. The function $H_{\underline{\epsilon},j}$ only depends on $\{\epsilon_s\}_{s=1}^{t}$ with $j \leq N_t$.

— $H_{0,j}$ will denote the function with $\underline{\epsilon} = \underline{0}$.

LEMMA 10. – *Assume that* $\{\epsilon_s\}_{s=1}^{t}$ *were chosen so that for all* $s < t$ *and* $x \in \mathbb{T}$, $h'_{\underline{\epsilon},M_s}(x) = e^{\pm 2^{-N_s}}$. *If* M_t *is sufficiently large with respect to* N_t *and* ϵ_{t+1} *is small enough then the following two properties hold:*

(i) *For all* $x \in \mathbb{T}$,

$$h'_{\underline{\epsilon},M_t}(x) = e^{\pm 2^{-N_t}}.$$

(ii) *Let* $w \in \Sigma_{\mathbf{A}}$ and $M_t < n \leq N_{t+1}$. Denote by $\xi(n,w) = \underline{x}_n(w) + \frac{1}{\varphi}\left(\overline{x}_n(w) - \underline{x}_n(w)\right)$ the point in $\frac{1}{\varphi}$ proportion in $C_{[w]_1^n}$. Then

$$\frac{H_{\underline{\epsilon},n-1}\left(\xi_n(w)\right) - H_{\underline{\epsilon},n-1}\left(\underline{x}_n(w)\right)}{H_{\underline{\epsilon},n-1}\left(\bar{x}_n(w)\right) - H_{\underline{\epsilon},n-1}\left(\underline{x}_n(w)\right)} = \frac{1}{\varphi}.$$

That is the (reference) point in $\frac{1}{\varphi}$ proportion in $C_{[w]_1^n}$ travels under $H_{\underline{\epsilon},n-1}$ to the reference point in $H_{\underline{\epsilon},n-1}\left(C_{[w]_1^n}\right)$.

Proof. – In the course of the proof we write for $n \leq N_{t+1}$, $h_n = h_{\underline{\epsilon},n}$ and $H_{\underline{\epsilon},n} = H_n$. Let $\delta > 0$. Since h_n is the identity for $N_t < n < M_t$, then $H_{M_t-1} = H_{N_t}$. The function H'_{N_t} is a product of N_t bounded Lipschitz functions and $\inf_{t\in[0,1]} H'_{N_t}(x) > 0$. Therefore there exists $K(t) > 1$, which depends only on $\{\lambda_s, N_s, M_s, \epsilon_s\}_{s=1}^{t-1}$ and $\{N_t, \lambda_t, \epsilon_t\}$, such that for every $x, y \in \mathbb{T}$,

$$\left|H'_{M_t-1}(x) - H'_{M_t-1}(y)\right| = \left|H'_{N_t}(x) - H'_{N_t}(y)\right| \leq K(t)|x-y|$$

and for every $x \in \mathbb{T}$,

$$K(t)^{-1} \leq \left|H'_{M_t-1}(x)\right| < K(t). \tag{4.5}$$

By uniform expansion of S, if M_t is sufficiently large then

$$\sup_{w\in\Sigma_{\mathbf{A}}} m_{\mathbb{T}}\left(H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)\right) \leq \varphi^{-(M_t-1)}K(t) < \frac{\delta}{K(t)^2}.$$

This implies that for every $w \in \Sigma_{\mathbf{A}}$ and $x, y \in H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)$,

$$\left|H'_{M_t-1}(x) - H'_{M_t-1}(y)\right| \leq K(t)|x-y| \leq K(t) m_{\mathbb{T}}\left(H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)\right) < \delta/K(t).$$

Averaging this inequality over all $y \in H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)$, for every $x \in H_{M_t-1}\left(C_{[w]_1^{M_t}}\right)$,

$$\left|H'_{M_t-1}(x) - \alpha(t,w)\right| = \left|H'_{M_t-1}(x) - \frac{1}{m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)} \int_{C_{[w]_1^{M_t}}} H'_{M_t-1}(y)dy\right|$$
$$\leq K(t)^{-1}\delta.$$

If follows from this and the lower bound in (4.5) that for every $x \in C_{[w]_1^{M_t}}$,

$$\left|\frac{\alpha(t,w)}{H'_{M_t-1}(x)} - 1\right| < \delta.$$

A consequence of the latter inequality which is proved by fixing $\mathbf{x}(w) = \underline{x}_{M_t}(w) = \overline{x}_{M_t}(w^-)$ once on w and once on w^-, is that

$$\forall w \in \Sigma_{\mathbf{A}}(M_t), \left|\frac{\alpha(t,w)}{H'_{M_t-1}(\mathbf{x}(w))} - \frac{\alpha(t,w^-)}{H'_{M_t-1}(\mathbf{x}(w))}\right| < 2\delta.$$

Part (i) follows by choosing an appropriate δ and the definition of h_{M_t}.

(ii) By the definition of h_{M_t}, if ϵ_{t+1} is small enough then equation (4.2) holds. Using property 3 of $\psi_{\epsilon_t,\lambda_t}$, a proof by induction shows that for all $M_t < n < J \leq N_{t+1}$,

$$m_{\mathbb{T}}\left(H_n\left(C_{[w]_1^J}\right)\right) = m_{\mathbb{T}}\left(H_n\left(C_{[w]_1^{n+1}}\right)\right) m_{\mathbb{T}}\left(C_{[w]_{n+1}^J} \middle| C_{[w]_1^{n+1}}\right). \tag{4.6}$$

The conclusion follows since if $w_{n+1} \in \{1,2\}$ then $C_{[w1]_1^{n+2}} = \left[\underline{x}_{n+1}(w), \xi_{n+1}(w)\right)$

$$\begin{aligned}\frac{H_n(\xi_{n+1}(w)) - H_n\left(\underline{x}_{n+1}(w)\right)}{H_n(\overline{x}_{n+1}(w)) - H_n\left(\underline{x}_{n+1}(w)\right)} &= \frac{m_{\mathbb{T}}\left(H_n\left(C_{[w1]_1^{n+2}}\right)\right)}{m_{\mathbb{T}}\left(H_n\left(C_{[w]_1^{n+1}}\right)\right)} \\ &= m_{\mathbb{T}}\left(C_{[w1]_{n+2}^{n+2}} \middle| C_{[w]_1^{n+1}}\right) \\ &= m_{\mathbb{T}}(w_2 = 1 | w_1 = 1) = \frac{1}{\varphi}.\end{aligned}$$

If $w_{n+1} = 3$ then $C_{[w21]_1^{n+3}} = \left[\underline{x}_{n+1}(w), \xi_{n+1}(w)\right)$ and then

$$\frac{H_n(\xi_{n+1}(w)) - H_n\left(\underline{x}_{n+1}(w)\right)}{H_n(\overline{x}_{n+1}(w)) - H_n\left(\underline{x}_{n+1}(w)\right)} = m_{\mathbb{T}}(w_2 = 2, w_3 = 1 | w_1 = 3) = \frac{1}{\varphi}. \qquad \square$$

By part (i) of the previous lemma we can choose sequences $\{\lambda_t, n_t, N_t, M_t, \epsilon_t\}_{t\in\mathbb{N}}$ so that $\sup_{x\in\mathbb{T}} h'_{\underline{\epsilon},M_t}(x) \leq e^{2^{-N_t}}$ for all $t \in \mathbb{N}$.

Proposition 11. – *Assume $\varphi/\lambda_1^2 > 1.6$, assume that for all $t \in \mathbb{N}$, $\sup_{x\in\mathbb{T}} h'_{\underline{\epsilon},M_t}(x) \leq e^{2^{-N_t}}$, then*

$$\sup_{n\in\mathbb{N}} \left|h_{\epsilon,n}(x) - x\right| \leq e(1.6)^{-n}$$

and consequently $\lim_{n\to\infty} H_{\underline{\epsilon},n}(x) = \mathfrak{h}_{\underline{\epsilon}}(x)$ *is a homeomorphism of* $\mathbb{T}$.

Proof. – If for some $\tau \leq T+1$, $M_\tau < k < M_{\tau+1}$ then,

$$\sup_{x\in\mathbb{T}} h'_{\underline{\epsilon},k}(x) = \sup_{s\in\mathbb{T}} |\psi_k(s)| \overset{(4.1)}{\leq} \lambda_\tau^2 \leq \lambda_1^2. \tag{4.7}$$

Therefore for every $n \leq M_{T+1}$,

$$\begin{aligned} m_{\mathbb{T}}\left(H_{\epsilon,n-1}\left(C_{[w]_1^n}\right)\right) &\leq \left(\prod_{k=1}^{T} \sup_{x\in\mathbb{T}} \left|h'_{\underline{\epsilon},M_k}(x)\right|\right) \lambda_1^{2n} m_{\mathbb{T}}\left(C_{[x]_1^n}\right) \\ &\leq \exp\left(\sum_{k=1}^{T} 2^{-N_k}\right)\left(\frac{\lambda_1}{\varphi}\right)^n \leq e(1.6)^{-n}. \end{aligned}$$

The invariance of $H_{\underline{\epsilon},n-1}\left(C_{[w]_1^n}\right)$ under $h_{\underline{\epsilon},n}$ implies that

$$\sup_{x\in\mathbb{T}} \left|h_{\underline{\epsilon},n}(x) - x\right| \leq \sup_{w\in\Sigma_{\mathbf{A}}} m_{\mathbb{T}}\left(H_{\underline{\epsilon},n-1}\left(C_{[w]_1^n}\right)\right) \leq e(1.6)^{-n}.$$

Consequently for every $n < m$,

$$\begin{aligned} \left|H_{\underline{\epsilon},m}(z) - H_{\underline{\epsilon},n}(z)\right| &\leq \sum_{k=n}^{m} \left|H_{\underline{\epsilon},k+1}(z) - H_{\underline{\epsilon},k}(z)\right| \\ &= \sum_{k=n}^{m} \left|h_{\underline{\epsilon},k+1}\left(H_{\underline{\epsilon},k}(z)\right) - H_{\underline{\epsilon},k}(z)\right| \\ &\leq e \sum_{k=n}^{m} \sup_{z\in\mathbb{T}} \left|h_{\underline{\epsilon},k+1}(z) - z\right| \leq e\sum_{k=n}^{m}(1.6)^{-k}. \end{aligned}$$

This shows that $\left\{H_{\underline{\epsilon},m}\right\}_{m=1}^{\infty}$ is a Cauchy sequence in $C(\mathbb{T})$. Its limit, being a continuous and strictly increasing function, is a homeomorphism of $\mathbb{T}$. □

LEMMA 12. – *Assume $\{\epsilon_k\}_{k=1}^{t}$ are already chosen so that for all $s < t$ and $x \in \mathbb{T}$, $h'_{\underline{\epsilon},M_s}(x) = e^{\pm 2^{-N_s}}$. If M_t is large enough with respect to N_t then there exists $\tilde{\delta}_{t+1} > 0$ so that for all $\epsilon_{t+1} < \tilde{\delta}_{t+1}$*

$$\mathfrak{g}'_{N_{t+1}}(x) = \lambda_{t+1}^{\pm M_t} e^{\pm 2^{-N_t+2}} \mathfrak{g}'_{N_t}(x). \tag{4.8}$$

Here $\mathfrak{g}_{N_t} = H_{\underline{\epsilon},N_t} \circ S \circ H^{-1}_{\underline{\epsilon},N_t}$.

Proof. – Assume first that $\epsilon_{t+1} = 0$ and since we are not going to vary $\underline{\epsilon}$ we write H_n and h_n to denote $H_{\underline{\epsilon},n}$ and $h_{\underline{\epsilon},n}$. Since $\epsilon_{t+1} = 0$, by Lemma 10 if M_t is large enough then $h'_{M_t}(x) = e^{\pm 2^{-N_t}}$ for all $x \in \mathbb{T}$. We assume that M_t is large enough for this to hold.

Let $z \in \mathbb{T}$, there exists a unique $y = y(z)$ such that $z = H_{N_{t+1}}(y)$. By the chain and differentiation of inverse functions, if $\mathfrak{g}_{N_{t+1}}$ is differentiable at z,

$$\mathfrak{g}'_{N_{t+1}}(z) = \varphi \frac{H'_{N_{t+1}}(Sy)}{H'_{N_{t+1}}(y)}.$$

Therefore since $h_k = \mathrm{id}$ for all $N_t < k < M_t$, $H_{N_t} = H_{M_t-1}$ and

$$\frac{\mathfrak{g}'_{N_{t+1}}(z)}{\mathfrak{g}'_{N_t}(z)} = \frac{H'_{N_{t+1}}(Sy)}{H'_{N_t}(Sy)} \cdot \frac{H'_{N_t}(y)}{H'_{N_{t+1}}(y)}$$

$$= \left(\prod_{k=M_t}^{N_{t+1}} h'_k\left(H_{k-1}(Sy)\right)\right)\left(\prod_{k=M_t}^{N_{t+1}} h'_k\left(H_{k-1}(y)\right)\right)^{-1}.$$

Fix $j \in [M_t, N_{t+1})$. Notice that $H_{j-1}(y) \in H_{j-1}\left(C_{[w]_1^j}\right)$ if and only if

$$H_{j-2}(Sy) \in H_{j-2}\left(C_{[w_2 \cdots w_j]}\right),$$

and that by Lemma 10.(ii), $H_{j-1}(y)$ and $H_{j-2}(Sy)$ are to the right of $\xi\left(C_{[w]_1^j}\right)$ and $\xi\left(C_{[w_2 \cdots w_j]}\right)$ respectively if and only if y is to the right of the reference point in $C_{[w]_1^j}$. Thus under the assumption that $\epsilon_{t+1} = 0$ for all $j \in (M_t + 1, N_{t+1}]$,

$$\frac{h'_{j-1}\left(H_{j-2}(Sy)\right)}{h'_j\left(H_{j-1}(y)\right)} = 1.$$

The last equality together with Lemma 10(i) implies that if M_t is large enough then,

$$\frac{\mathfrak{g}'_{N_{t+1}}(z)}{\mathfrak{g}'_{N_t}(z)} = \frac{h'_{M_t}\left(H_{M_t-1}(Sy)\right)}{h'_{M_t}\left(H_{M_t-1}(y)\right)} \frac{h'_{N_{t+1}}\left(H_{N_{t+1}-1}(Sy)\right)}{h'_{M_t+1}\left(H_{M_t}(y)\right)} \underbrace{\prod_{j=M_t+2}^{N_{t+1}} \frac{h'_{j-1}\left(H_{j-2}(Sy)\right)}{h'_j\left(H_{j-1}(y)\right)}}_{=1}$$

$$= \left(\lambda_{t+1}^8 e^{2^{-N_t+1}}\right)^{\pm 1}$$

The last inequality uses the fact that for $l \in \{M_t + 1, N_{t+1}\}$ and $z \in \mathbb{T}$, $\left|h'_l(z)\right| = \lambda_{t+1}^{\pm 2}$.

In [7] they argue that the estimate on the derivative is continuous (uniformly) with respect to ϵ_{t+1} since $\psi'_{\epsilon,\lambda_{t+1}}$ converges pointwise to $\psi'_{0,\lambda_{t+1}}$ when $\epsilon \to 0$. However this convergence is not uniform (and it can't be as it converges to a step function) and therefore their argument is not sufficient for convergence in the C^1 norm.

We proceed as follows. For $n \in (M_t, N_{t+1}]$ and $w \in \Sigma_{\mathbf{A}}$ with $w_n = 1$ denote by $\mathrm{BS}(n, w)$, the *Bad Set* at stage n for w, to be the following set

$$\left\{y \in C_{[w]_1^n} :\ \forall \delta > 0,\ \exists z \in (y - \delta, y + \delta),\ h'_n \circ H_{n-1}(z) \notin \left\{\frac{\lambda_{t+1}\varphi^2}{1 + \lambda_{t+1}\varphi}, \frac{\varphi^2}{1 + \lambda_{t+1}\varphi}\right\}\right\}.$$

This set, which is a union of four small intervals, is the set of all $y \in C_{[w]_1^n}$ where the derivative of $h'_n \circ H_{n-1}$ is not constant on a neighborhood of y.

First we demand that δ_{t+1} is small enough so that the conclusion of Lemma 10 and equation (4.6) holds for all $\epsilon_{t+1} < \delta_{t+1}$.

Secondly we demand that δ_{t+1} is small enough so that for $M_t < n < m \leq N_{t+1}$, if $\mathrm{BS}(m, w) \cap \mathrm{BS}(n, w) \neq \emptyset$ then one of the end points of $H_{m-1}\left(C_{[w]_1^m}\right)$ is either an end point of $H_{n-1}\left(C_{[w]_1^n}\right)$ or the point in $\frac{1}{\varphi}$ proportion in $H_{n-1}\left(C_{[w]_1^n}\right)$.

To understand why we choose these points, notice that in those marked endpoints

$$h'_n(x) = 1.$$

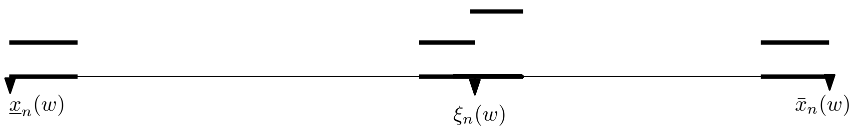


FIGURE 4.3. The small intervals demonstrate the possibilities of locations of $BS(m, w)$.

This can be done if for example [3]

$$\frac{m_{\mathbb{T}}\left(H_{\epsilon,m-1}\left(C_{[w]_1^m}\right)\right)}{m_{\mathbb{T}}\left(H_{\underline{\epsilon},n-1}\left(C_{[w]_1^n}\right)\right)} \geq \left(\frac{1}{\varphi^2\lambda_{t+1}^2}\right)^{n-m} \geq \frac{1}{5^{-n_{t+1}}} \gg \delta_{t+1}.$$

Indeed, if $\epsilon_{t+1} \leq \delta_{t+1}$, then $BS(n, w)$ is the union of four subintervals of $H_{\underline{\epsilon},n-1}\left(C_{[w]_1^n}\right)$ of considerably smaller length than $H_{\underline{\epsilon},m-1}\left(C_{[w]_1^m}\right)$ and thus their bad sets can only intersect in a unique interval if either $\bar{x}_m(w) \in \{\bar{x}_n(w), \xi_n(w)\}$ or $\underline{x}_m(w) \in \{\underline{x}_m(w), \xi_n(w)\}$.

In fact with such a choice of δ_{t+1} one has that for all $w \in \Sigma_{\mathbf{A}}$ and $M_t < n < m \leq N_{t+1}$, $\mathrm{BS}(n, w) \cap \mathrm{BS}(m, w)$ is always one interval for which one of its end points satisfies

$$h_n'(H_{n-1}(x)) = 1. \tag{4.9}$$

In addition,

$$\begin{aligned}\frac{m_{\mathbb{T}}\left(\mathrm{BS}(n, w) \cap \mathrm{BS}(m, w)\right)}{m_{\mathbb{T}}\left(\mathrm{BS}(n, w)\right)} &\leq \sup_{x\in\mathbb{T}}\left(h_n^{-1} \circ \cdots \circ h_m^{-1}\right)'(x)\frac{m_{\mathbb{T}}\left(C_{[w]_1^m}\right)}{m_{\mathbb{T}}\left(C_{[w]_1^n}\right)} \\ &\leq \left(\frac{\lambda_1^2}{\varphi}\right)^{n-m} \leq (1.6)^{n-m}.\end{aligned} \tag{4.10}$$

By the definition of $h_{\epsilon,n}$, $h_{\underline{\epsilon},n}' \circ H_{n-1}$ is a Lipschitz function with a Lipschitz constant of order $\mathrm{Const.}/m_{\mathbb{T}}\left(\mathrm{BS}(n, w)\right)$.

It follows from (4.9) and (4.10) that there exists a constant $B > 0$ such that for all $y \in \mathrm{BS}(n, w) \cap \mathrm{BS}(m, w)$,

$$h_n'(H_n(y)) = e^{\pm B(1.6)^{n-m}}.$$

The final argument is as follows: given $x \in \mathbb{T}$ there is a unique $y \in \mathbb{T}$ such that $x = H_{N_{t+1}}(y)$. Let w be such that $y \in C_{[w]_1^{N_{t+1}}}$. If $y \notin \bigcup_{n=M_t+1}^{N_{t+1}} \mathrm{BS}(n, w)$ then a similar analysis as in the case $\epsilon_{t+1} = 0$ yields the conclusion. Otherwise there exists a maximal $M_t < \mathbf{J} = J(y) \leq N_{t+1}$ such that $y \in \mathrm{BS}(\mathbf{J}, w)$. A similar argument as in the case $\epsilon_{t+1} = 0$ yields

$$g_{N_{t+1}}'(z) = \prod_{k=\mathbf{J}}^{N_{t+1}} \frac{h_k\left(H_{k-1}\left(Sy\right)\right)}{h_k\left(H_{k-1}(y)\right)} = \lambda_{t+1}^{\pm 4} g_{\mathbf{J}}' \circ H_{\mathbf{J}-1}(y). \tag{4.11}$$

[3] Here notice that $\sup_{n\leq k\leq N_{t+1}, x\in\mathbb{T}} \frac{1}{h_{\underline{\epsilon},k}'(x)} \leq \lambda_{t+1}^{-2}$ irrespectible of the choice of $\underline{\epsilon}$.

For $M_t + 2 \leq n < \mathbf{J} - M_t/4$, either $y \notin \mathrm{BS}(k, w)$ for all $k \leq n$ and then we proceed as in the case $\epsilon_{t+1} = 0$ or $y \in \mathrm{BS}(n, w) \cap \mathrm{BS}(\mathbf{J}, w)$ and then,

$$h'_n \circ H_{n-1}(y) = e^{\pm B(1.6)^{n-\mathbf{J}}}.$$

In addition, $S(\mathrm{BS}(n, w) \cap \mathrm{BS}(\mathbf{J}, w))$ is an interval of size $\varphi m_{\mathbb{T}}(\mathrm{BS}(n, w) \cap \mathrm{BS}(\mathbf{J}, w))$ with one point x for which [(4)] $h'_n \circ H_{n-1}(x) = 1$. Therefore as before,

$$h'_n \circ H_{n-1}(Sy) = e^{\pm \varphi B(1.6)^{m-\mathbf{J}}}.$$

Thus, using that for all $M_t \leq k \leq N_{t+1}$, $\frac{h'_k(H_{k-1}(Sy))}{h'_j(H_{k-1}(y))} \leq \lambda_{t+1}^4$

$$\begin{aligned} &g'_{\mathbf{J}} \circ H_{J-1}(y) \\ &\quad \leq \mathfrak{g}'_{N_t}(H_{M_t-1}(y)) \frac{h'_{M_t}(H_{M_t-1}(Sy))}{h'_{M_t}(H_{M_t-1}(y))} \prod_{n=M_t+1}^{\mathbf{J}-M_t/4} e^{3B(1.6)^{n-\mathbf{J}}} \prod_{k=J-M_t/4}^{J} \frac{h'_k(H_{k-1}(Sy))}{h'_j(H_{k-1}(y))} \\ &\quad \leq \left(\mathfrak{g}'_{N_t}(z)\, e^{2^{-N_t+2}}\right) e^{C(1.6)^{-M_t/4}} \lambda_{t+1}^{M_t}. \end{aligned}$$

The upper bound follows from the last equation together with (4.11) since $M_t/4 \gg N_t$. The lower bound is similar. □

A consequence of Lemma 12 is that we can choose $\underline{\epsilon} = \{\epsilon_k\}_{k=1}^{\infty}$ so that $\mathfrak{g}_{N_t}$ and $D_{\mathfrak{g}_{N_t}}$ converge uniformly to a map $\mathfrak{g}$ with

$$D_{\mathfrak{g}}(x) = \varphi \cdot \left(\prod_{t \in \mathbb{N}} \lambda_t^{\pm M_{t-1}}\right) \cdot e^{\sum_{t=1}^{\infty} 2^{-N_t+4}}. \tag{4.12}$$

By taking care that for each $t \in \mathbb{N}$, $\lambda_t^{M_{t-1}}$ is small enough and the N_t are large enough,

$$1.6 \leq \varphi \cdot \left(\prod_{t \in \mathbb{N}} \lambda_t^{\pm M_{t-1}}\right) \cdot \exp\left(\sum_{t=1}^{\infty} 2^{-N_t+4}\right) \leq 1.7,$$

thus the limiting transformation $\mathfrak{g}$ is uniformly expanding. What remains to be shown before we can explain the modified inductive construction of $\{\lambda_k, M_k, N_k, \epsilon_k\}_{k=1}^{\infty}$ is that we can choose $\underline{\epsilon}$ so that $m_{\mathbb{T}} \circ \mathfrak{h}_{\underline{\epsilon}} \sim \mu^+$.

LEMMA 13. – *Assume that μ^+ is a push forward via Θ of the Markovian type* III_1 *measure for the shift defined by $\{\lambda_k, m_k, n_k, M_k, N_k\}_{k=1}^{\infty}$. Then there exists a sequence $\underline{\varepsilon} = \{\varepsilon_k\}_{k=1}^{\infty}$ such that for every $\underline{\epsilon} = \{\epsilon_k\}_{k=1}^{\infty}$ which satisfies $\forall k \in \mathbb{N}$, $\epsilon_k \leq \varepsilon_k$, the function $\mathfrak{h}_{\underline{\epsilon}}$ defined previously satisfies*

$$m_{\mathbb{T}} \circ \mathfrak{h}_{\underline{\epsilon}} \sim \mu^+.$$

Proof. – The proof of the lemma will be done by applying the theory of local absolute continuity of Shiryaev with $\mathscr{F}_t$ the sigma algebra generated by $\left\{C_{[w]_1^{N_t}} : w \in \Sigma_{\mathbf{A}}\right\}$. For $\underline{\epsilon} = \{\epsilon_k\}_{k=1}^{\infty}$, we will use the notation $\varrho_{\underline{\epsilon},n}(x) := h'_{\underline{\epsilon},n}\left(H_{\underline{\epsilon},n-1}(x)\right)$.

Given $\underline{\epsilon} = \{\epsilon_k\}_{k=1}^{t}$,

$$\left(m_{\mathbb{T}} \circ \mathfrak{h}_{\underline{\epsilon}}\right)_t := m_{\mathbb{T}} \circ \mathfrak{h}_{\underline{\epsilon}}\big|_{\mathscr{F}_t} = m_{\mathbb{T}} \circ H_{\underline{\epsilon},N_t},$$

[(4)] x is either an end point or the point in $\frac{1}{\varphi}$ proportion in $C_{[w_2,\ldots,w_{n+1}]}$.

and

$$\left(\mu^{+}\right)_{t} := \mu|_{\mathscr{F}_t} = m_{\mathbb{T}} \circ H_{0,N_t}.$$

A calculation shows that

$$z_t(x) := \frac{d\left(m_{\mathbb{T}} \circ \mathfrak{h}_{\underline{\epsilon}}\right)_t}{d\left(\mu^{+}\right)_t}(x) = \frac{H'_{\underline{\epsilon},N_t}(x)}{H'_{0,N_t}(x)}.$$

Writing $\tilde{H}_{\underline{\epsilon},k,t}$ for the function $H_{\delta(\underline{\epsilon},k),N_t}$ with

$$\delta\left(\underline{\epsilon},k\right)_j := \begin{cases} \epsilon_j, & 1 \leq j \leq k \\ 0, & j > k, \end{cases}$$

and noticing that $\tilde{H}_{\delta(\underline{\epsilon},0),N_t} = H_{0,N_t}$ we get

$$z_t(x) = \prod_{k=1}^{t} \frac{\tilde{H}'_{\underline{\epsilon},k,t}(x)}{\tilde{H}'_{\underline{\epsilon},k-1,t}(x)}. \tag{4.13}$$

By [18, p. 527 Remark 2] it remains to show that we can choose $\underline{\varepsilon}$ such that if for all $k \in \mathbb{N}$, $\epsilon_k < \varepsilon_k$, then $\{z_t\}_{t=1}^{\infty}$ is uniformly integrable with respect to μ. We proceed to show how to choose $\underline{\varepsilon}$. Let $x \in \mathbb{T}\backslash \mathrm{Bd}\,(S)$.

Fix $k \in \mathbb{N}$. By the chain rule and the fact that $H_{\delta(\underline{\epsilon},k),M_{k-1}} = H_{\delta(\underline{\epsilon},k-1),M_{k-1}}$ one sees that

$$\frac{\tilde{H}'_{\underline{\epsilon},k,t}(x)}{\tilde{H}'_{\underline{\epsilon},k-1,t}(x)} = \left(\prod_{l=M_{k-1}+1}^{N_k} \frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)}\right) \cdot \left(\prod_{l=N_k}^{N_t} \frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)}\right).$$

First we will want to prove that if ϵ_k is small enough, then

$$\prod_{l=N_k}^{N_t} \frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)} \leq e^{3(1.6)^{-N_k}}. \tag{4.14}$$

To see (4.14), first notice that since for every $s \geq k$ and $N_s < n < M_s$,

$$h_{\delta(\underline{\epsilon},k),n} = h_{\delta(\underline{\epsilon},k-1),n} = \mathrm{id},$$

then for all $k \leq s \leq t-1$,

$$\prod_{l=N_s}^{M_s-1} \frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)} = 1.$$

Secondly, for $s \geq k$ and $M_s < n \leq N_{s+1}$ there exists $w \in \Sigma_{\mathbf{A}}$ such that $x \in C_{[w]_1^n}$. If $w_n \neq 1$ then $\rho_{\delta(\underline{\epsilon},k),l}(x) = \rho_{\delta(\underline{\epsilon},k),l}(x) = 1$. Otherwise notice that for $\eta \in \{\delta\left(\underline{\epsilon},k\right), \delta\left(\underline{\epsilon},k-1\right)\}$, $\eta_{s+1} = 0$ and $H_{\eta,n-1}(x)$ is to the right of the point in $\frac{1}{\varphi}$ proportion in $H_{\eta,n-1}\left(C_{[w]_1^n}\right)$ if and only if x is to the right of the point in $\frac{1}{\varphi}$ in $C_{[w]_1^n}$. Therefore for all $s \geq k$ and $M_s < l \leq N_{s+1}$, $\rho_{\delta(\underline{\epsilon},k),l}(x) = \rho_{\delta(\underline{\epsilon},k-1),l}(x)$ and by Lemma 10.(i),

$$\begin{aligned}\prod_{s=k}^{t-1} \prod_{l=M_s}^{N_{s+1}} \frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)} &= \prod_{s=k}^{t-1} \frac{\rho_{\delta(\underline{\epsilon},k),M_s}(x)}{\rho_{\delta(\underline{\epsilon},k-1),M_s}(x)} \\ &\leq \prod_{s=k}^{t-1} \frac{e^{(1.6)^{-N_t}}}{e^{-(1.6)^{-N_t}}} \leq e^{3(1.6)^{-N_k}}.\end{aligned}$$

We remark here that similarly one can get that

$$\prod_{s=k}^{t-1}\prod_{l=M_s}^{N_{s+1}}\frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)}\geq e^{-3(1.6)^{-N_k}},$$

which in turn shows that there exists $c>1$ such that

$$z_t(x)=c^{\pm 1}\prod_{k=1}^{t}\prod_{l=M_{k-1}}^{N_k}\frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)}. \tag{4.15}$$

If for every $t<k$, we have chosen M_t to be large enough so that Lemma 10.(i) holds then there exists $c>0$ such that

$$z_t(x)=c^{\pm 1}\prod_{k=1}^{t}\prod_{l=M_{k-1}+1}^{N_k}\frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)}.$$

As $\epsilon_k\to 0$,

$$\psi_k'(x):=\psi_{\epsilon_k,\lambda_k}'(x)\to\psi_{0,\lambda_k}'(x)\ \mu\text{ a.e. } x.$$

It follows that

$$\prod_{l=M_{k-1}+1}^{N_k}\frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)}\xrightarrow[\epsilon_k\to 0]{}1\ \mu\ a.e.\ x.$$

By Egorov's Theorem there exists $A_k\in\mathscr{B}_{\mathbb{T}}$, with $\mu(A_k)>1-\frac{1}{2^k\prod_{r=1}^{k}(\lambda_r)^{4nr}}$ such that

$$\prod_{l=M_{k-1}+1}^{N_k}\frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)}\xrightarrow[\epsilon_k\to 0]{}1,\ \text{uniformly in } x\in A_k.$$

The lower bound on the measure of A_k is chosen because for every $\epsilon_k>0$

$$\max_{x,y\in\mathbb{T}}\prod_{l=M_{k-1}+1}^{N_k}\frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(y)}\lessapprox\left(\max_{x,y\in\mathbb{T}}\frac{\psi_{\epsilon_k,\lambda_k}'(x)}{\psi_{0,\lambda_k}'(y)}\right)^{n_k}=(\lambda_k)^{4n_k}. \tag{4.16}$$

Now we are finally in a position to define the sequence $\underline{\varepsilon}$. Let ε_k be small enough so that for every $\underline{\epsilon}$ with $\epsilon_k<\varepsilon_k$ and $x\in A_k$,

$$1-\frac{1}{k^2}\leq\prod_{l=M_{k-1}+1}^{N_k}\frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)}\leq 1+\frac{1}{k^2}.$$

Let $\underline{\epsilon}$ which satisfies for every $k\in\mathbb{N}$, $\epsilon_k<\varepsilon_k$. For large M, if for some $n\in\mathbb{N}$ and $x\in\mathbb{T}$, $z_n(x)>M$, then there exists $q=q(M)\leq n$ such that $x\in\bigcup_{r=q}^{n}A_r^c$. Therefore by (4.15) and decomposing the set $[z_n>M]$ by the last $r\leq n$ for which $x\in A_r^c$,

$$\begin{aligned}\int_{[z_n>M]}z_n(x)d\mu(x)&\leq c\int_{[z_n>M]}\left(\prod_{k=1}^{n}\prod_{l=M_{k-1}+1}^{N_k}\frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)}\right)d\mu(x)\\&\leq c\sum_{r=q(M)}^{n}\int_{[A_r^c\cap(\bigcap_{j=r+1}^{n}A_j)]}\left(\prod_{k=1}^{n}\prod_{l=M_{k-1}+1}^{N_k}\frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)}\right)d\mu(x).\end{aligned}$$

Since for $x \in A_j$, $\prod_{l=M_{j-1}+1}^{N_j} \frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)} \leq 1 + \frac{1}{j^2}$,

$$\begin{aligned}
&\sum_{r=q(M)}^{n} \int_{[A_r^c \cap (\bigcap_{j=r+1}^{n} A_j)]} \left(\prod_{k=1}^{n} \prod_{l=M_{k-1}+1}^{N_k} \frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)} \right) d\mu(x) \\
&\quad \leq \sum_{r=q(M)}^{n} \prod_{j=r+1}^{n} \left(1 + \frac{1}{j^2}\right) \int_{A_r^c} \left(\prod_{k=1}^{r} \prod_{l=M_{k-1}+1}^{N_k} \frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)} \right) d\mu\,(x) \\
&\quad \leq \left[\prod_{j=1}^{\infty} \left(1 + \frac{1}{j^2}\right) \right] \sum_{r=q(M)}^{n} \mu\left(A_r^c\right) \max_{x \in \mathbb{T}} \left(\prod_{k=1}^{r} \prod_{l=M_{k-1}+1}^{N_k} \frac{\rho_{\delta(\underline{\epsilon},k),l}(x)}{\rho_{\delta(\underline{\epsilon},k-1),l}(x)} \right) \\
&\quad \overset{(4.16)}{\lessapprox} \left[\prod_{j=1}^{\infty} \left(1 + \frac{1}{j^2}\right) \right] \sum_{r=q(M)}^{n} \frac{1}{2^r} \\
&\quad \lessapprox 2^{-q(M)} \prod_{j=1}^{\infty} \left(1 + \frac{1}{j^2}\right).
\end{aligned}$$

When $M \to \infty$ then $q(M) \to \infty$ and therefore

$$\sup_{n \in \mathbb{N}} \int_{[z_n > M]} z_n(x) d\mu(x) \lessapprox 2^{-q(M)} \prod_{j=1}^{\infty} \left(1 + \frac{1}{j^2}\right) \to 0 \text{ as } M \to \infty.$$

This shows that $\{z_n\}$ is uniformly integrable and hence $m_{\mathbb{T}} \circ \mathfrak{h}_{\underline{\epsilon}} \sim \mu^+$. □

4.2.2. *The modified induction process for choosing $\{\lambda_k, N_k, M_k, n_k, m_k, \epsilon_k\}$ and the proof of Theorem 9.–* In the course of the construction here we arrived at two conditions on $\{\epsilon_k\}$ and two extra conditions on $\{\lambda_k, M_k\}$. In order to show the existence of these sequences one has to modify the induction process of Section 3 as follows and insert the choice of $\{\epsilon_t\}$ in the induction.

In the proof of the previous lemmas we have an extra condition on the size of M_t (or $m_t = M_t - N_t$) which is determined by $\{N_s, \lambda_s, M_{s-1}, \epsilon_s\}_{s=1}^{t}$.

The choice of ε_{t+1} in Lemma 13, $\tilde{\delta}_{t+1}$ in Lemma 12 and ϵ_{t+1} in Proposition 10 is determined by $\{N_s, \lambda_s, M_{s-1}, \epsilon_s\}_{s=1}^{t}$ and $\{N_{t+1}, M_t\}$. We also need to take care that

$$1.6 \leq \varphi \cdot \left(\prod_{t \in \mathbb{N}} \lambda_t^{\pm 2M_{t-1}} \right) \cdot \exp\left(\pm \sum_{t=1}^{\infty} 2^{-N_t+4} \right) \leq 1.7.$$

This shows now that the order of choice in the induction is as follows

$$\{\lambda_s, n_s, N_s, m_s, M_s, \epsilon_s\}_{s=1}^{t} \Rightarrow \lambda_{t+1} \Rightarrow \{n_{t+1}, N_{t+1}\} \Rightarrow \epsilon_{t+1} \Rightarrow \{m_{t+1}, M_{t+1}\}.$$

The modifications needed to be done in the inductive construction are: first change the condition (3.1) on λ_{t+1} with the condition

$$\lambda_{t+1}^{2M_t} \leq \exp\left(2^{-N_t}\right),$$

as this involves making λ_{t+1} smaller this choice is valid. This gives that

$$\varphi \cdot \left(\prod_{t \in \mathbb{N}} \lambda_t^{\pm 2M_{t-1}} \right) \cdot \exp\left(\pm \sum_{t=1}^{\infty} 2^{-N_t+4} \right) = \varphi \cdot \exp\left(\pm \sum_{t=1}^{\infty} 2^{-N_t+5} \right).$$

By demanding now that $N_1 > 20$, we get

$$\varphi \cdot \left(\prod_{t\in\mathbb{N}} \lambda_t^{\pm 2M_{t-1}}\right) \cdot \exp\left(\pm \sum_{t=1}^{\infty} 2^{-N_t+4}\right) = \varphi \cdot \exp\left(\pm \sum_{t=1}^{\infty} 2^{-N_t+5}\right) \in (1.6, 1.7)$$

as we required. There is no further change in the inductive choice of λ_t, n_t, N_t as they will not depend on $\underline{\epsilon}$.

Given $\{\lambda_s, n_s, N_s, m_s, M_s, \epsilon_s\}_{s=1}^{t}$ and N_{t+1} we choose ϵ_{t+1} to be small enough so that the conclusions of Lemma 10.(ii), Lemma 12 and Lemma 13 hold true.

Then we choose m_{t+1} based on the original constraints from Section 3 together with the restriction that $M_{t+1} = m_{t+1} + N_{t+1}$ is large enough so that the conclusion of Lemma 10.(i) is true. Since this involves perhaps enlarging m_{t+1} it is consistent with the other constraints of the induction.

Proof of Theorem 9. – Choose $\{\lambda_k, N_k, M_k, n_k, m_k, \epsilon_k\}_{k=1}^{\infty}$ as in the inductive construction. Build the Markovian measure $\eta = \mathrm{M}\{\mathrm{P}_k, \pi_k : k \in \mathbb{Z}\}$ determined by $\{\lambda_k, N_k, M_k, n_k, m_k\}_{k=1}^{\infty}$, $\mu := \Phi_*(\eta)$ and $\mu^+ = \Theta_*(\mathrm{M}\{\mathrm{P}_k, \pi_k : k \in \mathbb{N}\})$.

Part (i) follows from Theorem 4 since $\{\lambda_k, N_k, M_k, n_k, m_k\}_{k=1}^{\infty}$ satisfy the constraints of the inductive construction in Section 3 hence it is a type III_1 measure for the shift.

(ii) and (iii): Since we chose $\underline{\epsilon} = \{\epsilon_k\}$ so that the conclusion of Lemma 13 holds, it follows that $m_{\mathbb{T}} \circ \mathfrak{h}_{\underline{\epsilon}} \sim \mu^+$. As we chose the sequences so that the conditions of Lemma 12 hold, for all $t \in \mathbb{N}$ and $x \in \mathbb{T}$,

$$\mathfrak{g}'_{N_{t+1}}(x) = \exp\left(\pm 2^{-N_t+4}\right) \lambda_{t+1}^{\pm 2M_t} \mathfrak{g}'_{N_t}(x).$$

Therefore $\left\{\mathfrak{g}'_{N_t}\right\}_{t=1}^{\infty}$ is a Cauchy sequence in $C(\mathbb{T})$, its limit function satisfies

$$1.6 \le \mathfrak{g}'(x) = \varphi \cdot \left(\prod_{t\in\mathbb{N}} \lambda_t^{\pm 2M_{t-1}}\right) \cdot \exp\left(\pm \sum_{t=1}^{\infty} 2^{-N_t+4}\right) \le 1.7. \qquad \square$$

5. Type III_1 Anosov Diffeomorphisms

Let $\{\lambda_k, m_k, n_k, M_k, N_k\}_{k=1}^{\infty}$ and $\epsilon = \{\epsilon_k\}_{k=1}^{\infty}$ as in Theorem 9 and let $\mathfrak{h}_\epsilon$ be the resulting function. Set $\mathfrak{H}_\epsilon(x, y) := (\mathfrak{h}_\epsilon(x), y)$ and

$$\mathfrak{G}(x, y) := \mathfrak{H}_\epsilon \circ \tilde{f} \circ \mathfrak{H}_\epsilon^{-1}(x, y) = \begin{cases} \left(\mathfrak{g}(x), -\varphi^{-1} y\right), & 0 \le x \le 1/\varphi, \\ \left(\mathfrak{g}(x), -\varphi^{-1}\left(y - \frac{\varphi^2}{\varphi+2}\right)\right), & 1/\varphi \le x \le 1. \end{cases}$$

In the construction of Section 3, $P_k = \mathbf{Q}$ for all $k < 0$. Writing $\mathbf{m}_{\mathbb{M}}$ for the Lebesgue measure on $\mathbb{M}_\sim$ one then has

$$d\mathbf{m}_{\mathbb{M}} \circ \mathfrak{H}_{\underline{0}}(x, y) = d\mu^+(x)dy = d\mu(x, y),$$

or in other words $\mathbf{m}_{\mathbb{M}} \circ \mathfrak{H}_{\underline{0}} = \Phi_* \mathrm{M}\{\mathrm{P}_k, \pi_k : k \in \mathbb{Z}\}$. Therefore since $m_{\mathbb{T}} \circ \mathfrak{h}_{\underline{\epsilon}} \sim \mu^+$,

$$\mu := \Phi_* \mathrm{M}\{\mathrm{P}_k, \pi_k : k \in \mathbb{Z}\} \sim \mathbf{m}_{\mathbb{M}} \circ \mathfrak{H}_{\underline{\epsilon}}.$$

Consequently $\left(\mathbb{M}_\sim, \mathscr{B}_{\mathbb{M}_\sim}, \mathbf{m}_{\mathbb{M}}, \mathfrak{G}\right)$ is a type III_1 transformation. This is because $(\mathbb{M}_\sim, \mathscr{B}_\sim, \mathbf{m}_{\mathbb{M}}, \mathfrak{G})$ is measure theoretically equivalent to $(\mathbb{M}_\sim, \mathscr{B}_\sim, \mathbf{m}_{\mathbb{M}} \circ \mathfrak{H}_{\underline{\epsilon}}, \tilde{f})$ which is orbit equivalent to $(\mathbb{M}_\sim, \mathscr{B}_\sim, \mu, \tilde{f})$.

By Remark 8, $\mathfrak{G}_\epsilon$ is one to one and onto. In addition, for every $(x,y) \notin \partial\mathbb{M}$, $\mathfrak{G}$ is differentiable in a neighborhood of (x,y) as all the partial derivatives are continuous in $\mathbb{M}\backslash\partial\mathbb{M}$, and

$$D_{\mathfrak{G}}(x,y) = \begin{pmatrix} \mathfrak{g}'(x) & 0 \\ 0 & -\varphi^{-1} \end{pmatrix}.$$

The problem is that $\mathfrak{G}$ when viewed as a transformation of $\mathbb{M}_\sim$ is not even continuous on the horizontal lines of $\partial\mathbb{M}$.

We define a sequence of functions $\mathfrak{r}_n(x,y) : \mathbb{T} \times \left[-\varphi/(\varphi+2), \varphi^2/(\varphi+2)\right] \to \mathbb{T}, n \in \mathbb{N}$ using the construction of the previous section. This defines a sequence $h_{n,y}(\cdot) := \mathfrak{r}_n(\cdot, y) : (\mathbb{T} \text{ or } [0,1/\varphi]) \to \mathbb{T}$ and

$$\mathfrak{h}_y(x) := \lim_{n\to\infty} h_{n,y} \circ h_{n-1,y} \circ \cdots \circ h_{1,y}(x),$$

where we will take care that the limit exists. The new examples will then be of the form

$$\mathfrak{Z}(x,y) := \begin{cases} \left(\mathfrak{h}_{-y/\varphi} \circ S \circ \mathfrak{h}_y^{-1}(x), -y/\varphi\right), & x \le 1/\varphi \\ \left(\mathfrak{h}_{-y/\varphi + \frac{\varphi}{\varphi+2}} \circ S \circ \mathfrak{h}_y^{-1}(x), -y/\varphi + \frac{\varphi}{\varphi+2}\right), & 1/\varphi \le x \le 1 \end{cases} : \mathbb{M}_\sim \to \mathbb{M}_\sim.$$

Particular care in the definition of $\mathfrak{h}_y$ is taken in order to ensure that if $(x,y) \sim (\hat{x},\hat{y})$ then $\mathfrak{h}_y(x) = \mathfrak{h}_{\hat{y}}(\hat{x})$ as this is needed for the continuity of $\mathfrak{Z}$ on $\partial\mathbb{M}$.

5.1. Definition of the coupling time on the horizontal boundary of $\mathbb{M}$

Denote by

$$U_1 := \left([0,1/\varphi] \times \left(\frac{\varphi^2}{\varphi+2} - \frac{1}{\varphi^{10}}, \frac{\varphi^2}{\varphi+2}\right]\right) \cup \left([1/\varphi,1] \times \left[-\frac{\varphi}{\varphi+2}, -\frac{\varphi}{\varphi+2} + \frac{1}{\varphi^{10}}\right)\right),$$

$$U_2 := \left((1/\varphi,1) \times \left(\frac{1}{\varphi+2} - \frac{1}{\varphi^{10}}, \frac{1}{\varphi+2}\right]\right) \cup \left([1/\varphi,1] \times \left[-\frac{\varphi}{\varphi+2}, -\frac{\varphi}{\varphi+2} + \frac{1}{\varphi^{10}}\right)\right)$$

and $\mathbb{M}\backslash U := U_1 \cup U_2$. Then $\mathbb{M}\backslash U$ is a neighborhood of the horizontal lines of $\partial\mathbb{M}$.

In our construction for any $(x,y) \in U$,

$$\mathfrak{r}_n(x,y) = h_n(x),$$

with h_n the functions in the one dimensional example in Section 4. This means that for any $(x,y) \in U$,

$$\mathfrak{h}_y(x) = \mathfrak{h}_{\underline{\epsilon}}(x).$$

We now will proceed to specify the construction of $\mathfrak{h}_y(x)$ for $(x,y) \in \mathbb{M}\backslash U$.

On the horizontal lines there is a problem that there are points $(x,y), (\hat{x},\hat{y}) \in \partial\mathbb{M}$ that are equivalent in $\mathbb{M}_\sim$ and

$$(h_{\underline{\epsilon}}(x), y) \neq (h_{\underline{\epsilon}}(\hat{x}), \hat{y}).$$

For example, consider the case $x = 1/\varphi^3$, $y = \frac{\varphi^2}{\varphi+2}$ and $\hat{x} = 1/\varphi$, $\hat{y} = -\frac{\varphi}{\varphi+2}$. The point $\frac{1}{\varphi}$ is a fixed point for $\mathfrak{h}_{\underline{\epsilon}}$ meaning $\mathfrak{h}_{\underline{\epsilon}}(1/\varphi) = 1/\varphi$. Since $\mathfrak{h}_{\underline{\epsilon}}\left(1/\varphi^3\right) = \frac{\lambda_1}{\varphi(1+\lambda_1\varphi)} \neq \frac{1}{\varphi^3}$ we get

$$\left(\mathfrak{h}_{\underline{\epsilon}}(x), y\right) \neq \left(\mathfrak{h}_{\underline{\epsilon}}(\hat{x}), \hat{y}\right) = (\hat{x}, \hat{y}).$$

However if we took care that $\frac{1}{\varphi^3}$ is a fixed point for $\mathfrak{h}_y$ then we will have the desired equality. It turns out that the correct way to do this will be by setting $h_{1,y}|_{[0,1/\varphi^3)} = h_{2,y}|_{[0,1/\varphi^3)} =$

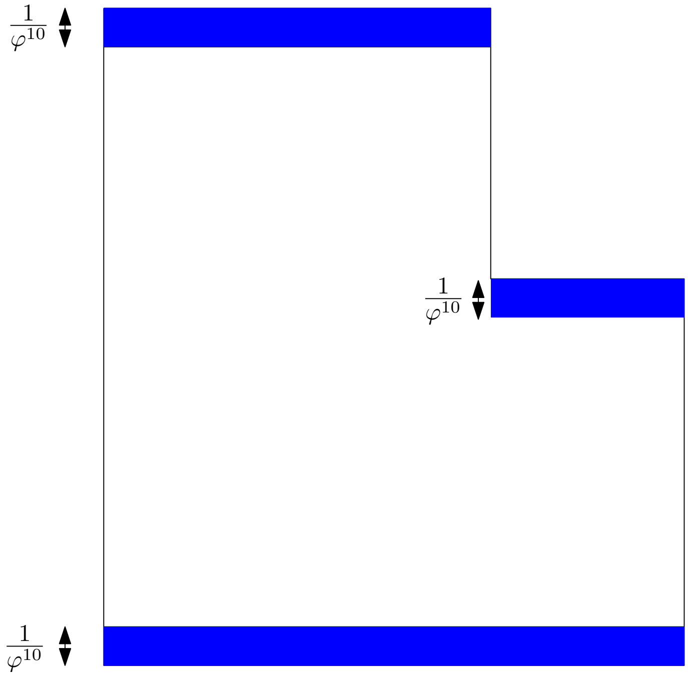


FIGURE 5.1. The bands are $\mathbb{M}\backslash U$

$\mathrm{Id}|_{[0,1/\varphi^3)}$ and to start perturbing (similarly as in the definition of h_n from the previous section) from $n \geq 3$. In general we will have a decomposition of the horizontal lines of $\partial\mathbb{M}$ to $\{V_i\}_{i=1}^{\infty}$ and we will start perturbing at V_i from $n \geq i+1$.

To be more precise the horizontal boundary consists of the lines $[0,1/\varphi)\times\left\{\varphi^2/(\varphi+2)\right\}$, $[1/\varphi,1)\times\{1/(\varphi+2)\}$ and $\mathbb{T}\times\{-\varphi/(\varphi+2)\}$. We look at a countable partition of the horizontal lines $\partial\mathbb{M}$ which are identified by $\sim$ and couple them in a time $T\in\mathbb{N}$ such that in the symbolic space on $\mathbb{T}$, the move $w(T)_T\to 1$ is possible for both pieces identified.

5.1.1. *The partition of horizontal subsegments of* $\partial\mathbb{M}$

1. $V_1 := \left[0,1/\varphi^2\right)\times\left\{-\frac{\varphi}{\varphi+2}\right\}\sim[1/\varphi,1)\times\left\{\frac{1}{\varphi+2}\right\}$. In this case $[0,1/\varphi^2)=C_{[1]_1^1}$ and $[1/\varphi,1)=C_{[2]_1^1}$ and $T(V_1)=2$.

2. $V_2 := \left[0,1/\varphi^3\right)\times\left\{\frac{\varphi^2}{\varphi+2}\right\}\sim\left[1/\varphi^2,1/\varphi\right)\times\left\{-\frac{\varphi}{\varphi+2}\right\}$. Here $[0,1/\varphi^3)=C_{[11]_1^2}$ and $\left[1/\varphi^2,1/\varphi\right)=C_{[32]_1^2}$ and $T(V_2)=3$.

3. $V_3 := \left[1/\varphi^3,1/\varphi^2\right)\times\left\{\frac{\varphi^2}{\varphi+2}\right\}\sim\left[1/\varphi,1/\varphi+1/\varphi^4\right)\times\left\{-\frac{\varphi}{\varphi+2}\right\}$. Here $[1/\varphi,1/\varphi^2)=C_{[132]_1^3}$ and $\left[1/\varphi,1/\varphi+1/\varphi^4\right)=C_{[211]_1^3}$ and $T(V_3)=4$.

4. $V_4 := \left[1/\varphi^2,1/\varphi^2+1/\varphi^5\right)\times\left\{\frac{\varphi^2}{\varphi+2}\right\}\sim\left[1/\varphi+1/\varphi^4,1/\varphi+1/\varphi^3\right)\times\left\{-\frac{\varphi}{\varphi+2}\right\}$. Here $C_{[3211]}=\left[1/\varphi^2,1/\varphi^2+1/\varphi^5\right)$ and $C_{[2132]}=\left[1/\varphi+1/\varphi^4,1/\varphi+1/\varphi^3\right)$.

5. For general $j > 4$, $V_j := C_{[\mathbf{w}(j)]_1^j} \times \left\{\frac{\varphi^2}{\varphi+2}\right\} \sim C_{[\tilde{\mathbf{w}}(j)]_1^j} \times \left\{-\frac{\varphi}{\varphi+2}\right\}$ where $\mathbf{w}(j), \tilde{\mathbf{w}}(j)$ are the following words of length j,

$$\mathbf{w}(j) = \begin{cases} 32\cdots 32132 & j \text{ odd} \\ 32\cdots 3211, & j \text{ even} \end{cases}$$

$$\tilde{\mathbf{w}}(j) = \begin{cases} 23\cdots 23211, & j \text{ odd} \\ 23\cdots 232132 & j \text{ even.} \end{cases}$$

As is expected for all $j \geq 2$, $T\left(V_j\right) = j + 1$. The following is immediate from the definition.

Claim 14. – For any $j \geq 2$,

$$\tilde{f}\left(V_j\right) = V_{j-1},$$

and

$$\tilde{f}\left(V_1\right) = [0, 1/\varphi) \times \{\frac{1}{\varphi+2}\} \subset U.$$

5.1.2. *Definition of the perturbation maps $h_{n,y}$.* – For $w \in \Sigma_{\mathbf{A}}$ and $n \in \mathbb{N}$, we write again $C_{[w]_1^n} := [\underline{x}_n(w), \bar{x}_n(w))$. Let

$$\mathbf{u}(x,y) := \begin{cases} \min\left\{\frac{\varphi^2}{\varphi+2} - y, y + \frac{\varphi}{\varphi+2}\right\}, & 0 \leq x \leq \frac{1}{\varphi} \\ \min\left\{\frac{1}{\varphi+2} - y, y + \frac{\varphi}{\varphi+2}\right\} & \frac{1}{\varphi} \leq x \leq 1 \end{cases}$$

be the minimal distance of (x, y) to the horizontal lines of $\partial\mathbb{M}$. In addition we will write $\mathbf{y}(x,y) : \mathbb{M} \to \left\{-\frac{\varphi}{\varphi+2}, \frac{1}{\varphi+2}, \frac{\varphi^2}{\varphi+2}\right\}$ to be the value so that

$$\mathbf{u}(x,y) = |\mathbf{y}(x,y) - y|.$$

Under that notation $(x, \mathbf{y}(x,y))$ is the closest point to (x, y) in the horizontal boundary. Let $(x,y) \in \mathbb{M}_\sim$.

Case 1: $(x,y) \in U$. – We do the regular construction as in Section 4. That is for any $N_t < n < M_t$, h_n is the identity. For any $M_t < n \leq N_t$, if $w_n = 1$ then $h_n|_{H_{n-1}\left(C_{[w]_1^n}\right)}$ is a rescaling of ψ_t to the interval $H_{n-1}\left(C_{[w]_1^n}\right)$ and if $w_n \neq 1$ then $h_n|_{H_{n-1}\left(C_{[w]_1^n}\right)}$ is the identity. If for some t, $n = M_t$ then $h_{M_t} \circ H_{M_t-1}|_{C_{[w]_1^{M_t}}}$ is the distribution correction function in the construction. Finally we set

$$\mathtt{r}_n(x,y) = h_{n,y}(x) := h_n(x)$$

and

$$K_{n,y}(x) := h_{n,y} \circ h_{n-1,y} \circ \cdots \circ h_{1,y}(x) = H_n(x).$$

Case 2: $(x,y)\notin U$. – In this case $\mathbf{u}(x,y)<\frac{1}{\varphi^{10}}$. Let $(x,\mathbf{y}(x,y))\in\partial\mathbb{M}$ be the closest point on the horizontal lines of $\partial\mathbb{M}$ to (x,y). Let $\mathbf{j}(x,y)\in\mathbb{N}$ be the integer so that

$$(x,\mathbf{y}(x,y))\in V_{\mathbf{j}(x,y)}.$$

This means that either $x\in C_{[\mathbf{w}(j)]_1^j}$ (if $x\le 1/\varphi$) or $x\in C_{[\tilde{\mathbf{w}}(j)]_1^j}$ $(1/\varphi\le x<1)$. We will define the construction for $x\in C_{[\mathbf{w}(j)]_1^j}$, the other case being similar. First we define for any $x\in C_{[\mathbf{w}(j)]_1^j}$,

$$K_{\mathbf{j}(x,y),\mathbf{y}(x,y)}(x)=x.$$

Then for any $(x,y)\in\mathbb{M}\backslash U$ such that $x\in C_{[\mathbf{w(j)}]_1^{\mathbf{j}}}$ we set

$$K_{\mathbf{j},y}(x):=(H_{\mathbf{J}}(x)-x)\left[3\varphi^{20}\left(\mathbf{u}(x,y)\right)^2-2\varphi^{30}\mathbf{u}(x,y)^3\right]+x.$$

For $n>\mathbf{j}(x,y)$, assume that we have defined for all $\mathbf{j}<k<n$, $h_{k,y}:=\mathtt{r}_k(\cdot,y)$ and $x\in C_{[w]_1^n}\subset C_{[\mathbf{w(j)}]_1^{\mathbf{j}}}$. We set

$$K_{n-1,y}|_{C_{[w]_1^n}}=h_{n-1,y}\circ\cdots\circ h_{\mathbf{j}+1,y}\circ K_{\mathbf{j},y}(x),$$

and

$$\mathfrak{l}_n(y,w):=m_{\mathbb{T}}\left(K_{n-1,y}\left(C_{[w]_1^n}\right)\right)=K_{n-1,y}\left(\bar{x}_n\left(w\right)\right)-K_{n-1,y}\left(\underline{x}_n\left(w\right)\right).$$

If $w_n\ne 1$ or $N_t<n<M_t$ for some $t\in\mathbb{N}$, then for all $x\in K_{n-1,y}\left(C_{[w]_1^n}\right)$, $\mathtt{r}_n(x,y):=x$. If $w_n=1$ and $M_t<n\le N_t$ and $\mathbf{j}<n$ then

$$\mathtt{r}_n(x,y):=K_{n-1,y}\left(\underline{x}_n(w)\right)+\mathfrak{l}_n(y,w)\psi_t\left(\frac{x-K_{n-1,y}\left(\underline{x}_n(w)\right)}{\mathfrak{l}_n(y,w)}\right).$$

Finally if $j\le M_t=n$ then $r_{M_t,y}$ is the distribution correction function with H_{M_t-1} replaced by $K_{M_t-1,y}$.

REMARK 15. – The 2 variable function

$$\mathfrak{q}_{\mathbf{J}}(x,\mathbf{u}):=(H_{\mathbf{J}}(x)-x)\left[3\varphi^{20}\mathbf{u}^2-2\varphi^{30}\mathbf{u}^3\right]+x$$

was chosen because of its following properties:

1. $\mathfrak{q}_{\mathbf{J}}\left(x,\varphi^{-10}\right)=H_{\mathbf{J}}(x)$ and $\mathfrak{q}_{\mathbf{J}}(x,0)=x$. This means that for $y\in\partial\mathbb{M}$, $\mathtt{r}_n(x,y)=x$ and therefore $K_{\mathbf{j},y}$ interpolates between the identity map and $H_{\mathbf{J}}|_{C_{[\mathbf{w(j)}]_1^{\mathbf{j}}}}$.
2. A consequence of the previous property is that $\frac{\partial\mathfrak{q}_{\mathbf{j}}}{\partial x}(x,0)=1$ and $\frac{\partial\mathfrak{q}_{\mathbf{J}}}{\partial x}\left(x,\varphi^{-10}\right)=\frac{\partial H_{\mathbf{J}}}{\partial x}(x)$. This is needed in order that $\frac{\partial K_{\mathbf{J},y}}{\partial x}$ will be continuous in y.
3. $\frac{\partial\mathfrak{q}_{\mathbf{J}}}{\partial u}(x,0)=\frac{\partial\mathfrak{q}_{\mathbf{J}}}{\partial u}\left(x,\varphi^{-10}\right)=0$ which is necessary for continuity of $\frac{\partial K_{\mathbf{j},y}}{\partial y}$.
4. $\sup_{0\le z\le\varphi^{-10}}\left|\frac{\partial\mathfrak{q}_{\mathbf{j}}}{\partial u}(x,z)\right|=\frac{3}{2}\varphi^{10}\left|H_{\mathbf{j}}(x)-x\right|$. We will show that the right hand side is uniformly exponentially small when $\mathbf{j}\to\infty$. The control of the derivatives in the y direction is to our opinion the hardest part in this section.

The idea behind this construction can be summarized as follows: for a fixed (x, y) which is close enough to the horizontal segment on the boundary we first look at the coupling time of the interval which contains the point closest to (x, y) on the boundary. On the boundary we start to apply the rescaling after the coupling time to ensure that the resulting map will be a map of $\mathbb{M}_\sim$ (respects the equivalence relation). Inside U we just start perturbing from the start and in what remains we do an interpolation using $q_{\mathbf{J}}$, of the map on the boundary and the map on U.

5.1.3. *Definition of $\mathfrak{Z}_N$ and the new examples of Anosov diffeomorphisms*

Define $\mathfrak{Z}_N : \mathbb{M}_\sim \to \mathbb{M}_\sim$,

$$\mathfrak{Z}_N(x,y) = \begin{cases} \left(K_{N,-y/\varphi} \circ S \circ K^{-1}_{N,y}(x), -y/\varphi\right), & (x,y) \in R_1 \cup R_3, \\ \left(K_{N,-y/\varphi+\frac{\varphi}{\varphi+2}} \circ S \circ K^{-1}_{N,y}(x), -y/\varphi + \frac{\varphi}{\varphi+2}\right), & (x,y) \in R_2. \end{cases}$$

REMARK 16. – In the construction of the previous subsection for every $n \in \mathbb{N}$, $x \in \{0, 1/\varphi^2, 1/\varphi\}$ are fixed points for $h_{\epsilon,n}$ (Remark 8). This remains true for $\mathfrak{r}_n$ in the sense that for all y and $x \in \{0, 1/\varphi^2, 1/\varphi\}$, $\mathfrak{r}_n(x,y) = \hat{r}_n(x,y) = x$. This shows that $\mathfrak{Z}_N$ is continuous. In addition if x is an endpoint of the segment $K_{n,y}\left(C_{[w]_1^n}\right)$ for some $w \in \Sigma_{\mathbf{A}}$ and y, then $\frac{\partial \mathfrak{r}_n}{\partial x}(x,y) = 1$. This gives that $\mathfrak{Z}_N$ is C^1. The invariance of the Markov partition $\{J_1, J_2, J_3\}$ of S under $K_{y,n}$ gives that $\mathfrak{Z}_N$ is one to one and onto and

$$\mathfrak{Z}_n^{-1}(x,y) = \begin{cases} \left(K_{n,-\varphi y} \circ \left(S|_{J_1 \cup J_3}\right)^{-1} \circ K^{-1}_{n,y}(x), -\varphi y\right), & -\frac{\varphi}{\varphi+2} \le y \le \frac{1}{\varphi+2} \\ \left(K_{n,-\varphi y+\varphi^2/(\varphi+2)} \circ \left(S|_{J_2}\right)^{-1} \circ K^{-1}_{n,y}(x), -\varphi y + \frac{\varphi^2}{\varphi+2}\right), & \frac{1}{\varphi+2} < y \le \frac{\varphi^2}{\varphi+2}. \end{cases}$$

Here $\left(S|_{J_1 \cup J_3}\right)^{-1}(x) := \frac{x}{\varphi} : [0,1] \to [1, 1/\varphi]$ is the inverse branch of S to the segment $[0, 1/\varphi]$ and $\left(S|_{J_2}\right)^{-1}(x) = \frac{x+1}{\varphi} : [0, 1/\varphi] \to [1/\varphi, 1]$ is the inverse branch of S to the segment $[1/\varphi, 1]$. Since $\mathfrak{Z}_n^{-1}$ is C^1, $\mathfrak{Z}_n$ is a diffeomorphism.

THEOREM 17. – *The sequence $\mathfrak{Z}_{N_t}$ converges in the C^1 topology to a type* III_1 *Anosov diffeomorphism.*

The proof of this theorem consists in a series of lemmas. The first step is to show that $K_{N_t,y}(x)$ converges uniformly in $\mathbb{M}$ as $t \to \infty$.

LEMMA 18. – *If $x \in C_{[\mathbf{w}(J)]_1^J} \cup C_{[\tilde{\mathbf{w}}(J)]_1^J}$ and $J \in [M_{t-1}, N_t)$,*

$$|H_J(x) - x| \le m_{\mathbb{T}}\left(C_{[w(x)]_1^{J-3}}\right) \le \varphi^{-(J-3)},$$

where $w(x) \in \{\mathbf{w}(J), \tilde{\mathbf{w}}(J)\}$ is such that $x \in C_{[w(x)]_1^J}$.

Proof. – By the form of $\mathbf{w}(J)$ and $\tilde{\mathbf{w}}(J)$ one has that for all $l \le J-3$, $w(x)_l \in \{2, 3\}$, hence

$$H_{J-3}(x)|_{C_{[w(x)]_1^{J-3}}} = x.$$

Since $\underline{x}_{J-3}(w(x))$, $\bar{x}_{J-3}(w(x))$ are fixed points of h_{j-2}, h_{j-1} and h_j, $H_J\left(C_{[w(x)]_1^{J-3}}\right) = C_{[w(x)]_1^{J-3}}$, the lemma follows. □

COROLLARY 19. – *The limit*

$$\lim_{n\to\infty} K_{n,y}(x) =: \mathfrak{h}_y(x)$$

exists uniformly in $\mathbb{M}$ *and is a continuous function and the function* $\mathscr{H}(x,y) = \left(\mathfrak{h}_y(x), y\right)$ *is a homeomorphism of* $\mathbb{M}_\sim$.

Proof. – The proof is similar to the proof of Lemma 11. First we claim that for every $n \in \mathbb{N}$,

$$\sup_{(x,y)\in\mathbb{M}_\sim} |\mathtt{r}_n(x,y) - x| \leq (1.5)^{-n}. \tag{5.1}$$

This is true since for every $w \in \Sigma_{\mathbf{A}}$,

$$\mathtt{r}_n\left(K_{n-1,y}\left(C_{[w]_1^n}\right), y\right) = K_{n-1,y}\left(C_{[w]_1^n}\right),$$

and consequently

$$|\mathtt{r}_n(x,y) - x| \leq m_{\mathbb{T}}\left(K_{n-1,y}\left(C_{[w]_1^n}\right)\right) \leq \left(\frac{\lambda_1^2}{\varphi}\right)^{-n}.$$

The last inequality follows since $\left|\frac{\partial \psi_l}{\partial x}(x)\right| \leq \lambda_1^2$ for every $l \in \mathbb{N}$ and thus $\left|\frac{\partial \mathtt{r}_k}{\partial x}\right|_\infty \leq \lambda_1^2$. Proceeding as in Lemma 11, it follows that for every y, $\left\{K_{n,y}(x)\right\}_{n=1}^\infty$ is a Cauchy sequence in the uniform topology. Thus, $\mathfrak{h}_y(x)$ is a continuous function in $\mathbb{M}$ as it is a uniform limit of continuous functions. Notice that $\mathfrak{h}_y$ is a homeomorphism of the circle for $y \in [-\varphi/(\varphi+2), 1/(\varphi+2)]$ or of $[0, 1/\varphi]$ if $y \in \left[1/(\varphi+2), \varphi^2/(\varphi+2)\right]$.

It remains to show that if $(x,y) \sim (\hat{x},\hat{y})$ (for points on $\partial\mathbb{M}$) then $\left(\mathfrak{h}_y(x), y\right) \sim \left(\mathfrak{h}_{\hat{y}}(\hat{x}), \hat{y}\right)$. Let $(x,y), (\hat{x},\hat{y}) \in \partial\mathbb{M}$ with $(x,y) \sim (\hat{x},\hat{y})$. There exists $\mathbf{j}(x,y) \in \mathbb{N}$ such that $(x,y), (\hat{x},\hat{y}) \in V_{\mathbf{j}}$. Since $(x,y) \sim (\hat{x},\hat{y})$, it follows that for every $n > \mathbf{j}$ and a word $w \in \Sigma_{\mathbf{A}}(n)$,

$$x \in C_{[\mathbf{w}(\mathbf{j})w]_1^{n+\mathbf{j}}} \Leftrightarrow \hat{x} \in C_{[\tilde{\mathbf{w}}(\mathbf{j})w]_1^{n+\mathbf{j}}}.$$

Since for all $n \leq \mathbf{j}$, $\mathtt{r}_n|_{V_{\mathbf{J}}} = Id_{\mathbb{T}}$, this property and the definition of $r_n(\cdot,\cdot)$ yield that for all $n \in \mathbb{N}$,

$$\left(K_{n,y}(x), y\right) \sim \left(K_{n,\hat{y}}(\hat{x}), \hat{y}\right).$$

The lemma follows by taking $n \to \infty$. □

Denote the function of the first coordinate by $\mathfrak{z}_n(x,y)$. Our goal is to prove that the limit

$$\mathfrak{z}(x,y) := \lim_{t\to\infty} \mathfrak{z}_{N_t}(x,y)$$

exists for all (x,y) and $\mathfrak{z}$ is a $C^1(\mathbb{M}_\sim)$ function with

$$1.6 \leq \frac{\partial \mathfrak{z}}{\partial x}(x,y) \leq 1.7.$$

The conclusion of hyperbolicity of $\mathfrak{Z}$ will follow from a standard lemma in the theory of Lyapunov exponents.

LEMMA 20. – *If in addition*

$$1.6 \le \varphi \cdot \lambda_1^6 \left(\prod_{t \in \mathbb{N}} \lambda_t^{\pm 2M_{t-1}} \right) \cdot \exp\left(\pm \sum_{t=1}^{\infty} 2^{-N_t+4} \right) \le 1.7,$$

then $\frac{\partial \mathfrak{z}}{\partial x}$ *is a continuous function in* $\mathbb{M}_{\sim}$ *and*

$$1.6 \le \frac{\partial \mathfrak{z}}{\partial x}(x, y) \le 1.7.$$

REMARK. – The extra condition in this lemma can easily be inserted into the inductive construction of the sequence $\{\lambda_k, M_k, N_k, m_k, n_k, \epsilon_k\}_{k=1}^{\infty}$.

Proof. – Let $(x, y) \in \mathbb{M}_{\sim}$. For the convenience of the reader, we will first show that $\frac{\partial \mathfrak{z}_{N_t}}{\partial x}(x, y)$ converges pointwise and $1.6 \le \frac{\partial \mathfrak{z}}{\partial x}(x, y) \le 1.7$ and then argue that the convergence is in fact uniform.

Let $t \in \mathbb{N}$ and $(x, y) \in \mathbb{M}_{\sim}$ be fixed. There exists a $w = w(x) \in \Sigma_{\mathbf{A}}$ such that for all $t \in \mathbb{N}$, $x \in K_{N_t,y}\left(C_{[w]_1^{N_t}}\right)$. As in the proof of Lemma 12, we write $z_y(x)$ to be the unique point in $C_{[w]_1^{N_t}}$ such that $x = K_{N_t,y}\left(z_y(x)\right)$. Recall that

$$\begin{aligned} \mathfrak{z}_{N_t}(x, y) &= \begin{cases} K_{N_t,-y/\varphi}\left(\varphi K_{N_t,y}^{-1}(x)\right), & 0 \le x \le 1/\varphi \\ K_{N_t,-y/\varphi+\varphi/\varphi+2}\left(\varphi K_{N_t,y}^{-1}(x) - 1\right), & 1/\varphi \le x \le 1 \end{cases} \\ &= \begin{cases} K_{N_t,-y/\varphi}\left(S z_y(x)\right), & 0 \le x \le 1/\varphi \\ K_{N_t,-y/\varphi+\varphi/\varphi+2}\left(S z_y(x)\right), & 1/\varphi \le x \le 1. \end{cases} \end{aligned}$$

By the chain rule, the lemma will follow once we show that uniformly in $(x, y) \in \mathbb{M}_{\sim}$ with $0 \le x \le 1/\varphi$,

$$\lim_{t \to \infty} \left(\frac{\partial K_{N_t,-y/\varphi}}{\partial x}\left(S z_y(x)\right) \right) \cdot \left(\frac{\partial K_{N_t,y}}{\partial x}\left(z_y(x)\right) \right)^{-1} \in \left(\frac{1.6}{\varphi}, \frac{1.7}{\varphi} \right), \tag{5.2}$$

and for every $(x, y) \in \mathbb{M}_{\sim}$ with $1/\varphi \le x \le 1$,

$$\lim_{t \to \infty} \left(\frac{\partial K_{N_t,-y/\varphi+\varphi/\varphi+2}}{\partial x}\left(S z_y(x)\right) \right) \cdot \left(\frac{\partial K_{N_t,y}}{\partial x}\left(z_y(x)\right) \right) \in \left(\frac{1.6}{\varphi}, \frac{1.7}{\varphi} \right).$$

We will separate the proof for three cases: we assume that $(x, y) \in R_1 \cup R_3$, equivalently $0 \le x \le 1/\varphi$, the proof when $(x, y) \in R_2$ is similar and just involves changing the appearance of $-y/\varphi$ by $-y/\varphi + \varphi/(\varphi + 2)$.

Case 1: $(x, y) \in U \cap \mathfrak{Z}^{-1}U$. – In this case

$$\mathfrak{z}_{N_t}(x, y) = H_{N_t} \circ S \circ H_{N_t}^{-1}(x),$$

and the conclusion is true by Lemma 12.

Case 2: $(x,y)\in U\cap \mathfrak{Z}^{-1}U^c$. – Firstly since $(x,y)\in U$ then $K_{n,y}(x)=H_n(x)$. In addition, because $\mathfrak{Z}(x,y)\notin U$ there exists $\mathbf{J}\in\mathbb{N}$ such that

$$[w_2,...,w_{\mathbf{J}+1}]_1^{\mathbf{J}}=[\mathbf{w}(\mathbf{J})]_1^{\mathbf{J}}\ \text{ or }\ [\tilde{\mathbf{w}}(\mathbf{J})]_1^{\mathbf{J}},$$

and consequently $w_n\neq 1$ for all $2\le l<\mathbf{J}-2$. This shows that $K_{n,y}\big|_{C_{[w]_1^n}}=h_n\circ\cdots\circ h_{\mathbf{J}-2}\circ h_1$ and writing $z_y(x)\in\mathbb{T}$ for the point such that $K_{N_t,y}\left(z_y(x)\right)=x$,

$$\frac{\partial K_{N_t,y}}{\partial x}\left(z_y(x)\right)=\frac{\partial h_1}{\partial x}\left(z_y(x)\right)\cdot\prod_{k=\mathbf{J}-2}^{N_t}\frac{\partial h_l}{\partial x}\left(H_{l-1}\left(z_y(x)\right)\right).$$

By the definition of the construction

$$\frac{\partial K_{N_t,-y/\varphi}}{\partial x}\left(Sz_y(x)\right)=\prod_{l=\mathbf{J}}^{N_t}\frac{\partial h_{l,-y/\varphi}}{\partial x}\left(K_{l-1,-y/\varphi}(Sz_y(x))\right),$$

and, here $\lambda_{t(\mathbf{J})}=\lambda_k$ if $N_{k-1}<\mathbf{J}\le N_k$ or 1 otherwise,

$$\left(\frac{\partial K_{N_t,-y/\varphi}}{\partial x}\left(Sz_y(x)\right)\right)\cdot\left(\frac{\partial K_{N_t,y}}{\partial x}\left(z_y(x)\right)\right)^{-1}=\lambda_{t(\mathbf{J})}^{\pm 6}\left(\frac{\partial h_1}{\partial x}\left(z_y(x)\right)\right)^{-1}\cdot\mathbf{I},$$

where

$$\mathbf{I}:=\left(\prod_{l=\mathbf{J}}^{N_t}\frac{\partial h_{l,-y/\varphi}}{\partial x}\left(K_{l-1,-y/\varphi}\left(Sz_y(x)\right)\right)\right)\cdot\left(\prod_{l=\mathbf{J}}^{N_t}\frac{\partial h_{l,y}}{\partial x}\left(H_l\left(z_y(x)\right)\right)\right)^{-1}.$$

As in the proof of Lemma 12, assuming that $\underline{\epsilon}\equiv 0$, one has that for $l\ge\mathbf{J}+1$, $K_{l-1,-y/\varphi}\left(Sz_y(x)\right)$ is to the right of the point in $1/\varphi$ proportion in $K_{l-1,-y/\varphi}\left(C_{[w_2,\ldots,w_{l+1}]}\right)$ if and only if $K_{l-1,y}\left(z_y(x)\right)$ is to the right of the point in $1/\varphi$ proportion in $K_{l,-y/\varphi}\left(C_{[w]_1^{l+1}}\right)$. This means that in the case $\underline{\epsilon}=0$,

$$\left(\frac{\partial h_{l-1,-y/\varphi}}{\partial x}\left(K_{l-2,-y/\varphi}(Sz_y(x))\right)\right)\cdot\left(\frac{\partial h_{l,y}}{\partial x}\left(K_{l-1,y}(z_y(x))\right)\right)^{-1}=1.$$

By proceeding with the analysis of the bad sets as in Lemma 12 one proves that

$$\mathbf{I}=\left(\prod_{k=t(\mathbf{J})}^{t}\lambda_k^{\pm 2M_{k-1}}\right)\exp\left(\pm\sum_{k=t(\mathbf{J})}^{t}2^{-N_k+4}\right),$$

and thus

$$\begin{aligned}&\left(\frac{\partial K_{N_t,-y/\varphi}}{\partial x}\left(Sz_y(x)\right)\right)\cdot\left(\frac{\partial K_{N_t,y}}{\partial x}\left(z_y(x)\right)\right)^{-1}\\&\quad=\left[\left(\frac{\partial h_1}{\partial x}\left(z_y(x)\right)\right)^{-1}\left(\prod_{k=t(\mathbf{J})}^{t}\lambda_k^{\pm 2M_{k-1}}\right)\cdot\exp\left(\pm\sum_{k=t(\mathbf{J})}^{t}2^{-N_k+4}\right)\right].\end{aligned}$$

This shows (5.2). In fact, because

$$\lim_{s\to\infty}\left(\prod_{k=s}^{\infty}\lambda_k^{\pm 2M_{k-1}}\right)\exp\left(\pm\sum_{k=s}^{\infty}2^{-N_k+4}\right)=1,$$

the convergence is uniform as $t\to\infty$.

Case 3: $(x, y) \in U^c$. – In this case let $\tilde{\mathbf{J}} \in \mathbb{N}$ be such that the closest point to (x, y) on the horizontal segments of $\partial\mathbb{M}$ is in $V_{\tilde{\mathbf{J}}}$. If $\tilde{\mathbf{J}} = 1$ then $\left(Sz_y(x), -y/\varphi\right) \in U$. Otherwise $\left(Sz_y(x), -y/\varphi\right)$ is in U^c and the closest point to it on the horizontal segments of $\partial\mathbb{M}$ is in $V_{\tilde{\mathbf{J}}-1}$. Consequently

$$\left(\frac{\partial K_{N_t,-y/\varphi}}{\partial x}\left(Sz_y(x)\right)\right)\cdot\left(\frac{\partial K_{N_t,y}}{\partial x}\left(z_y(x)\right)\right)^{-1}$$
$$= \left(\prod_{l=\tilde{\mathbf{J}}-1}^{N_t}\frac{\partial h_{l,-y/\varphi}}{\partial x}\left(K_{l-1,-y/\varphi}\left(Sz_y(x)\right)\right)\right)\cdot\left(\prod_{l=\tilde{\mathbf{J}}}^{N_t}\frac{\partial h_{l,y}}{\partial x}\left(K_{l-1,y}\left(z_y(x)\right)\right)\right)^{-1}.$$

Similarly as in case 2, one has

$$\left(\frac{\partial K_{N_t,-y/\varphi}}{\partial x}\left(Sz_y(x)\right)\right)\cdot\left(\frac{\partial K_{N_t,y}}{\partial x}\left(z_y(x)\right)\right)^{-1}$$
$$= \lambda_{t(\tilde{\mathbf{J}})}^{\pm 2}\left(\prod_{k=t(\tilde{\mathbf{J}})}^{t}\lambda_k^{\pm 2M_{k-1}}\right)\exp\left(\pm\sum_{k=t(\tilde{\mathbf{J}})}^{t}2^{-N_k+4}\right)$$

and the convergence is uniform. □

5.1.4. *Proving differentiability in the y-direction.*– Again we will prove differentiability in the y direction for $(x, y) \in R_1 \cup R_3$. The idea of the proof here is as follows. If $(x, y) \in U$ then $K_{n,\tilde{y}}(x) = H_{\underline{\epsilon},n}(x)$ for all $\tilde{y}$ in a neighborhood of (x, y), hence $\frac{\partial K_{n,y}}{\partial y}(\cdot) \equiv 0$. Otherwise, for $(x, y) \in \mathbb{M}\backslash U$, $K_{\mathbf{j}(x,y)-2,\tilde{y}}(x) = x$ and the first (major) change between $K_{n,y}(x)$ and $K_{n,\tilde{y}}(x)$ appears at time $n = \mathbf{j}(x, y)$. We will show that for our construction the y derivative of $K_{n,y}(x)$ can be bounded above by a (bounded) constant times $\frac{\partial K_{\mathbf{j}(x,y),y}(x)}{\partial y}$, the uniform convergence of $\partial\mathfrak{z}_n/\partial y$ will follow from the chain rule and simple arithmetic.

The following notation will be used in this subsection. Usually we will consider $x \in [0, 1/\varphi]$ and work constantly with a fixed $w \in \Sigma_{\mathbf{A}}$ such that $x \in C_{[w]_1^n}$ for all $n \in \mathbb{N}$. If that is the case we will write $[\underline{x}_n, \bar{x}_n)$ to denote $C_{[w]_1^n}$.

For $-\frac{\varphi}{\varphi+2} \leq y \leq \frac{\varphi^2}{\varphi+2}$ and $n \geq \mathbf{N}(y)$, let $\mathrm{BS}(n, w, y) \subset C_{[w]_1^n}$ to be the bad set as in the proof of Lemma 12 with $h'_n \circ H_n$ replaced by $\left(\frac{\partial r_{n,y}}{\partial x}\right) \circ K_{n-1,y}$.

For an $w \in \Sigma_{\mathbf{A}}$ we denote by $w_1^n = w_1 w_2 \cdots w_n$ the finite word derived by w up to time n. Given a finite word w_1^n, $[w_1^n]$ denotes the n-periodic word defined by w_1^n. Finally given two words w and $\tilde{w}$ (in which case w is a finite word), the word $w\tilde{w}$ denotes the concatenation of w and $\tilde{w}$.

Recall the definition of $K_{\mathbf{j}(x,y),y}(x) = \mathfrak{r}_{\mathbf{j}(x,y)}(x, y)$ which is defined by

$$\mathfrak{r}_{\mathbf{j}(x,y)}(x, y) := \left(H_{\mathbf{j}(x,y)}(x) - x\right)\mathscr{P}(x, y) + x,$$

where

$$\mathscr{P}(x, y) = \left[3\varphi^{20}\left(\mathbf{u}(x, y)\right)^2 - 2\varphi^{30}\mathbf{u}(x, y)^3\right].$$

In the following proof if $\mathbf{j}(x,y)=\mathbf{J}$ we will need a different definition of the bad set for $n=\mathbf{J}$. Let

$$\mathbb{BS}(\mathbf{J}):=\begin{cases}\mathrm{BS}(\mathbf{J}-2,\mathbf{w}(\mathbf{J})), & \mathbf{J}\text{ odd}\\ \mathrm{BS}(\mathbf{J}-1,\mathbf{w}(\mathbf{J}))\cup\mathrm{BS}(\mathbf{J},\mathbf{w}(\mathbf{J})) & \mathbf{J}\text{ even}\end{cases}$$

if $(x,y)\in R_1\cup R_3$ (For $(x,y)\in R_2$ change the odd to even and even to odd).

Lemma 21. – *Assume that $M_t\le\mathbf{j}\le N_{t+1}$, and $x\in C_{[w]_1^{N_{t+1}}}$. If $x\notin\mathbb{BS}(\mathbf{J})$, for every $\mathbf{j}<n\le N_{t+1}$, there exists $0\le\beta_n(x)\le\varphi$ such that for every (x,y) with $\mathbf{j}(x,y)=\mathbf{j}$,*

$$K_{n,y}(x)=K_{n,y}(\underline{x}_n)+\beta_n(x)l_n(y,w),$$

in addition $\beta_{N_{t+1}}(x)$ is continuous in x.

Proof. – Let $(x,y)\in(\mathbb{M}\backslash U)\cup(R_1\cup R_3)$ so that $\mathbf{j}(x,y)=\mathbf{j}$ (the case $(x,y)\in R_2$ is similar). The proof is by induction on n. Since $x\notin\mathbb{BS}(\mathbf{J})$, $\underline{x}_{\mathbf{J}+1}\in\left\{\underline{x}_{\mathbf{J}},\underline{x}_{\mathbf{J}}+\varphi^{-1}\left(\bar{x}_{\mathbf{J}}-\underline{x}_{\mathbf{J}}\right)\right\}$ and

$$\mathbf{w}(\mathbf{J})=\begin{cases}32\cdots32132, & \mathbf{J}\text{ odd,}\\ 32\cdots3211, & \mathbf{J}\text{ even,}\end{cases}$$

it follows that if $\mathbf{J}$ is even then by property (3) of ψ_t,

$$\begin{aligned}H_{\mathbf{J}}(x)-H_{\mathbf{J}}\left(\underline{x}_{\mathbf{J}+1}(w)\right)&=\begin{cases}(\varphi\psi_{t+1}(1/\varphi))^2\left(x-\underline{x}_{\mathbf{J}+1}\right), & w_{\mathbf{J}+1}=1\\ (\varphi\psi_{t+1}(1/\varphi))\left(\varphi^2\left(1-\psi_{t+1}(1/\varphi)\right)\right)\left(x-\underline{x}_{\mathbf{J}+1}\right), & w_{\mathbf{N}+1}=3\end{cases}\\ &=\frac{P_{\mathbf{J}-1}(w_{\mathbf{J}-1},w_{\mathbf{J}})}{\mathrm{Q}(w_{\mathbf{J}},w_{\mathbf{J}+1})}\frac{P_{\mathbf{J}}(w_{\mathbf{J}},w_{\mathbf{J}+1})}{\mathrm{Q}(w_{\mathbf{J}},w_{\mathbf{J}+1})}\left(x-\underline{x}_{\mathbf{J}+1}\right)\\ &:=b\left(x-\underline{x}_{\mathbf{J}+1}\right),\end{aligned}$$

and if $\mathbf{J}$ is odd then

$$\begin{aligned}H_{\mathbf{J}}(x)-H_{\mathbf{J}}\left(\underline{x}_{\mathbf{J}+1}\right)&=\varphi^2\left(1-\psi_{t+1}(1/\varphi)\right)\left(x-\underline{x}_{\mathbf{J}+1}\right)\\ &=\frac{P_{\mathbf{J}-2}(w_{\mathbf{J}-2},w_{\mathbf{J}-1})}{\mathrm{Q}(w_{\mathbf{J}-2},w_{\mathbf{J}-1})}\left(x-\underline{x}_{\mathbf{J}+1}\right)\\ &:=b\left(x-\underline{x}_{\mathbf{J}+1}\right).\end{aligned}$$

It then follows that

$$\begin{aligned}K_{\mathbf{J},y}(x)-K_{\mathbf{J},y}\left(\underline{x}_{\mathbf{J}+1}\right):&=\mathtt{r}_{\mathbf{J}}(x,y)-\mathtt{r}_{\mathbf{J}}\left(\underline{x}_{\mathbf{J}+1},y\right)\\ &=\left(x-\underline{x}_{\mathbf{J}+1}\right)\left[(b-1)\mathscr{P}(x,y)+1\right],\end{aligned}$$

and

$$\begin{aligned}\mathfrak{l}_{\mathbf{j}+1}(y,w):&=K_{\mathbf{j},y}(\bar{x}_{\mathbf{J}+1})-K_{\mathbf{j},y}\left(\underline{x}_{\mathbf{j}+1}\right)\\ &=\left(\bar{x}_{\mathbf{j}+1}-\underline{x}_{\mathbf{J}+1}\right)\left[(b-1)\mathscr{P}(x,y)+1\right].\end{aligned}$$

This implies that

$$\frac{K_{\mathbf{J},y}(x)-K_{\mathbf{J},y}\left(\underline{x}_{\mathbf{J}+1}\right)}{\mathfrak{l}_{\mathbf{J}+1}(y,w)}=\frac{x-\underline{x}_{\mathbf{J}+1}}{\bar{x}_{\mathbf{J}+1}-\underline{x}_{\mathbf{J}+1}}.\tag{5.3}$$

Therefore

$$K_{\mathbf{J}+1,y}(x) = K_{\mathbf{J+1},y}\left(\underline{x}_{\mathbf{J}+1}\right) + \underbrace{\psi_{t+1}\left(\frac{x-\underline{x}_{\mathbf{j+1}}}{\bar{x}_{\mathbf{J}+1}-\underline{x}_{\mathbf{j+1}}}\right)}_{:=\beta_{\mathbf{j}}(x)} \mathfrak{l}_{\mathbf{J}+1}(y,w)$$

and the base of induction is proved.

For the inductive step notice that if the conclusion of the lemma is true for $n \in \mathbb{N}$, then

$$\frac{K_{n,y}(x) - K_{n,y}\left(\underline{x}_{n+1}\right)}{\mathfrak{l}_{n+1}(y,w)} = \frac{\beta_n(x) - \beta_n\left(\underline{x}_{n+1}\right)}{\beta_n\left(\bar{x}_{n+1}\right) - \beta_n\left(\underline{x}_{n+1}\right)}$$

does not depend on y. The conclusion then follows for $n+1$ with

$$\beta_{n+1}(x) := \psi_{t+1}\left(\frac{\beta_n(x) - \beta_n\left(\underline{x}_{n+1}\right)}{\beta_n\left(\bar{x}_{n+1}\right) - \beta_n\left(\underline{x}_{n+1}\right)}\right)$$

and the continuity of β_{n+1} follows from the continuity of β_n and ψ_{t+1}. $\square$

The last lemma shows the importance of knowing how $\frac{\partial \mathfrak{l}_n}{\partial y}$ decays when $\mathbf{N}(y) < n \leq N_{t+1}$. We will now show that it is exponential in n.

LEMMA 22. – *Let $M_t \leq \mathbf{j} < n \leq N_{t+1}$, a $w \in \Sigma_{\mathbf{A}}$ with $w_1^{\mathbf{j}} = \mathbf{w}(\mathbf{j})$ and $-\frac{\varphi}{\varphi+2} \leq y \leq \frac{\varphi^2}{\varphi+2}$, then*

$$\left|\frac{\partial \mathfrak{l}_n}{\partial y}(y,w)\right| \leq (1.6)^{\mathbf{j}-n}\left|\frac{\partial l_{\mathbf{j}}}{\partial y}(y,w)\right|$$

and

$$\left|\frac{\partial \mathfrak{l}_{\mathbf{j}+1}}{\partial y}(y,w)\right| \lesssim (1.6)^{-\mathbf{j}}.$$

Proof. – We assume $\left(\underline{x}_{N_{t+1}}, y\right) \notin \partial U$, the proof for the case $\left(\underline{x}_{N_{t+1}}, y\right) \in \partial U$ is similar. In this case for small $|h|$, $\left(\underline{x}_{N_{t+1}}, y+h\right) \in U^c$.

Since $\bar{x}_{N_{t+1}}(w)$ is not in the bad set $\mathbb{BS}\left(\mathbf{j}(x,y)\right)$, it follows from (5.3) that for small $|h|$,

$$\begin{aligned}
\frac{m_{\mathbb{T}}\left(K_{\mathbf{j},y+h}\left(C_{[w]_1^{N_{t+1}}}\right)\right)}{m_{\mathbb{T}}\left(K_{\mathbf{j},y+h}\left(C_{[w]_1^{\mathbf{j}+1}}\right)\right)} &:= \frac{K_{\mathbf{j},y+h}\left(\bar{x}_{N_{t+1}}\right) - K_{\mathbf{j},y+h}\left(\underline{x}_{N_{t+1}}\right)}{K_{\mathbf{j},y+h}\left(\bar{x}_{\mathbf{j}+1}\right) - K_{\mathbf{j},y+h}\left(\underline{x}_{\mathbf{j}+1}\right)} \\
&= \frac{\bar{x}_{N_{t+1}} - \underline{x}_{N_{t+1}}}{\bar{x}_{\mathbf{j}+1} - \underline{x}_{\mathbf{j}+1}} \\
&= \frac{m_{\mathbb{T}}\left(C_{[w]_1^{N_{t+1}}}\right)}{m_{\mathbb{T}}\left(C_{[w]_1^{\mathbf{j}+1}}\right)}.
\end{aligned}$$

It then follows by definition of $h_{n,y}$ for $n > \mathbf{j}$ that for $|h|$ small,

$$\mathfrak{l}_{N_{t+1}}(y+h,w) = l_{\mathbf{j}+1}(y+h,w)\prod_{k=\mathbf{j}+1}^{N_{t+1}} P_k\left(w_k, w_{k+1}\right),$$

hence

$$\frac{\mathfrak{l}_{N_{t+1}}(y+h,w)}{\mathfrak{l}_{N_{t+1}}(y,w)}=\frac{\mathfrak{l}_{\mathbf{j}+1}(y+h,w)}{\mathfrak{l}_{\mathbf{j}+1}(y,w)}.$$

This yields that

$$\mathfrak{l}_{N_{t+1}}(y+h,w)-\mathfrak{l}_{N_{t+1}}(y,w)=\frac{\mathfrak{l}_{N_{t+1}}(y,w)}{\mathfrak{l}_{\mathbf{j}(x,y)+1}(y,w)}\left[\mathfrak{l}_{\mathbf{j}(x,y)+1}(y+h,w)-\mathfrak{l}_{\mathbf{j}(x,y)+1}(y,w)\right],$$

dividing by h and taking limit $h\to 0$ we get

$$\left|\frac{\partial \mathfrak{l}_{N_{t+1}}(y,w)}{\partial y}\right|=\frac{\mathfrak{l}_{N_{t+1}}(y,w)}{\mathfrak{l}_{\mathbf{j}+1}(y,w)}\left|\frac{\partial \mathfrak{l}_{\mathbf{j}+1}(y,w)}{\partial y}\right|\leq (1.6)^{\mathbf{j}-N_{t+1}}\left|\frac{\partial \mathfrak{l}_{\mathbf{j}+1}(y,w)}{\partial y}\right|.$$

The last inequality follows from, for all $N_t\leq \mathbf{j}<n\leq N_{t+1}$,

$$\prod_{k=\mathbf{j}}^{n-1}P_k\left(w_k,w_{k+1}\right)\leq\left(\frac{\lambda_{t+1}\varphi}{1+\lambda_{t+1}\varphi}\right)^{\mathbf{j}-n}\leq\left(\frac{\lambda_{t+1}}{\varphi}\right)^{\mathbf{j}-n}<(1.6)^{\mathbf{j}-n}.$$

For the proof of the second part notice that for $x\in\left\{\underline{x}_{\mathbf{j}},\bar{x}_{\mathbf{j}}\right\}$,

$$\begin{aligned}\left|\frac{\partial K_{\mathbf{j},y}}{\partial y}(x)\right|&=\left|\frac{\partial \mathfrak{r}_{\mathbf{j}}}{\partial y}(x,y)\right|\\&\leq\left|\frac{\partial \mathscr{P}}{\partial y}(x,y)\right|\left|H_{\epsilon,\mathbf{j}}(x)-x\right|\\&\leq\frac{3}{2}\varphi^{10}\left|m\left(C_{[w]_1^{\mathbf{j}-2}}\right)\right|\leq\frac{3}{2}\varphi^{12}\varphi^{-\mathbf{j}}\leq\frac{1}{2}(1.62)^{-\mathbf{j}},\end{aligned}\tag{5.4}$$

for all large $\mathbf{j}$. Thus (recall $\mathfrak{l}_{\mathbf{j}+1}(y,w)=\mathfrak{r}_{\mathbf{j}(x,y)}\left(\bar{x}_{\mathbf{j}+1}\right)-\mathfrak{r}_{\mathbf{j}}\left(\underline{x}_{\mathbf{j}+1}\right)$)

$$\left|\frac{\partial \mathfrak{l}_{\mathbf{j}+1}}{\partial y}(y,w)\right|\leq(1.6)^{-\mathbf{j}},\quad\text{for all large } \mathbf{j}.\qquad\square$$

Lemma 21 shows that if $x\in C_{[\mathbf{w}(\mathbf{j})]_1^n}$ is not in $x\in\mathbb{BS}(\mathbf{j})$ then the y-derivative of $K_{N_{t+1},y}(x)$ (here t is the number such that $N_t<\mathbf{j}<N_{t+1}$) is controlled by the derivative on a finite collection of points plus the evolution of the lengths of the intervals. We would like to point out that there is actually no bad set if $N_t<\mathbf{j}\leq M_t$ because then $K_{\mathbf{j},y}|_{C_{[\mathbf{w}(\mathbf{j})]_1^{\mathbf{j}}}}=id_{\mathbb{T}}$. This idea will be reiterated with a slight modification for the derivatives $\partial K_{n,y}/\partial y$ for $M_s<n<N_s$ where $M_s>\mathbf{j}(x,y)$.

For points in the bad set we will apply a correction point procedure which we call the x-delta method. Assume that $x\in\mathbb{BS}(\mathbf{j})$. For Δ-small (so that $(x,y\pm\Delta)\in U^c$)) there exists a unique $\mathbf{x}(\Delta)$ such that

$$\frac{K_{\mathbf{j}(x,y),y+\Delta}(\mathbf{x}(\Delta))-K_{\mathbf{j}(x,y),y+\Delta}\left(\underline{x}_{\mathbf{j}(x,y)+1}\right)}{\mathfrak{l}_{\mathbf{j}(x,y)+1}(y+\Delta,w)}=\frac{K_{\mathbf{j}(x,y),y}(x)-K_{\mathbf{j}(x,y),y}\left(\underline{x}_{\mathbf{j}(x,y)+1}\right)}{\mathfrak{l}_{\mathbf{j}(x,y)+1}(y,w)}.\tag{5.5}$$

We will use Lemma 22 to obtain a first order approximation for $\mathbf{x}(\Delta)$ when Δ is small.

In the next lemma, if $M_t<\mathbf{j}+1<N_{t+1}$ $\beta_{\mathbf{j}+1}(x):=\psi_{t+1}\left(\frac{K_{\mathbf{j},y}(x)-K_{\mathbf{j},y}\left(\underline{x}_{\mathbf{j}+1}\right)}{\mathfrak{l}_{\mathbf{j}+1}(y,w)}\right)$ and for $M_t<\mathbf{j}+1<n\leq N_{t+1}$

$$\beta_{n+1}(x):=\psi_{t+1}\left(\frac{\beta_n(x)-\beta_n\left(\underline{x}_{n+1}\right)}{\beta_n\left(\bar{x}_{n+1}\right)-\beta_n\left(\underline{x}_{n+1}\right)}\right).$$

LEMMA 23. – *Assume that* $M_t \leq \mathbf{j} \leq N_{t+1}$, *and* $x \in C_{[w]_1^{N_{t+1}}}$ *with* $w_1^{\mathbf{j}} = \mathbf{w}(\mathbf{j})$. *The following holds:*

(i) *For every* y *so that* $\mathbf{j}(x,y) = \mathbf{j}$ *and* Δ *so that* $(x, y \pm \Delta) \in \mathbb{M}\backslash U$,

$$K_{N_{t+1},y+\Delta}(\mathbf{x}(\Delta)) = K_{N_{t+1},y+\Delta}\left(\underline{x}_{N_{t+1}}\right) + \beta_{N_{t+1}}(x) l_{N_{t+1}}(y+\Delta, w),$$

(ii) $|\mathbf{x}(\Delta) - x| \leq 4(1.6)^{-N_{t+1}}\Delta + o(\Delta)$ *as* $\Delta \to 0$.

Proof. – (i) This is the same as the proof of Lemma 21 by using (5.5) as the starting point.

(ii) If $x \notin \mathbb{BS}(\mathbf{j}) \cap C_{[w]_1^{N_t}}$ then by Lemma 21, $\mathbf{x}(\Delta) = x$. Since $\underline{x}_{N_{t+1}} \notin \mathbb{BS}(\mathbf{j})$,

$$\begin{aligned}\frac{K_{\mathbf{j},y+\Delta}\left(\underline{x}_{N_{t+1}}\right) - K_{\mathbf{j},y+\Delta}\left(\underline{x}_{\mathbf{j}(x,y)+1}\right)}{\mathfrak{l}_{\mathbf{j}+1}(y+\Delta, w)} &= \frac{\underline{x}_{N_{t+1}} - \underline{x}_{\mathbf{j}(x,y)+1}}{\overline{x}_{\mathbf{j}+1} - \underline{x}_{\mathbf{j}+1}} \\ &= \frac{K_{\mathbf{j},y}\left(\underline{x}_{N_{t+1}}\right) - K_{\mathbf{j},y}\left(\underline{x}_{\mathbf{j}+1}\right)}{\mathfrak{l}_{\mathbf{j}+1}(y, w)}.\end{aligned}$$

Therefore by adding and subtracting $K_{\mathbf{j},y+\Delta}\left(\underline{x}_{N_{t+1}}\right)/\mathfrak{l}_{\mathbf{j}+1}(y+\Delta, w)$ on the right hand side and $K_{\mathbf{j},y}\left(\underline{x}_{N_{t+1}}\right)/\mathfrak{l}_{\mathbf{j}+1}(y, w)$ on the left hand side of Equation 5.5, if follows that Equation (5.5) is equivalent to

$$K_{\mathbf{j},y+\Delta}(\mathbf{x}(\Delta)) - K_{\mathbf{j},y+\Delta}\left(\underline{x}_{N_{t+1}}\right) = \frac{\mathfrak{l}_{\mathbf{j}+1}(y+\Delta, w)}{\mathfrak{l}_{\mathbf{j}+1}(y, w)}\left(K_{\mathbf{j},y}(x) - K_{\mathbf{j},y}\left(\underline{x}_{N_{t+1}}\right)\right). \tag{5.6}$$

For the ease of notation we will write $\underline{X} := \underline{x}_{N_{t+1}}$ and $H_{\mathbf{j}}(z) := H_{\underline{\epsilon},\mathbf{j}}(z)$. Since by Lemma 22,

$$\mathfrak{l}_{\mathbf{j}+1}(y+\Delta, w) = \mathfrak{l}_{\mathbf{j}+1}(y, w) \pm (1.6)^{-\mathbf{J}}\Delta + o(\Delta)$$

we have

$$\frac{\mathfrak{l}_{\mathbf{j}+1}(y+\Delta, w)}{\mathfrak{l}_{\mathbf{j}+1}(y, w)}\left(K_{\mathbf{j},y}(x) - K_{\mathbf{j},y}(\underline{x})\right) = \left[1 \pm \frac{(1.6)^{-\mathbf{j}}\Delta}{\mathfrak{l}_{\mathbf{j}+1}(y, w)}\right]\left(K_{\mathbf{j},y}(x) - K_{\mathbf{j},y}(\underline{X})\right) + o(\Delta).$$

In addition for all $|\Delta|$ small,

$$K_{\mathbf{j},y+\Delta}(\mathbf{x}(\Delta)) - K_{\mathbf{j},y+\Delta}(\underline{X}) = \left(H_{\mathbf{j}}(\mathbf{x}(\Delta)) - H_{\mathbf{j}}(\underline{X}) - (\mathbf{x}(\Delta) - \underline{X})\right)\mathscr{P}(x, y+\Delta) + (\mathbf{x}(\Delta) - \underline{X}).$$

By Taylor expansion

$$\mathscr{P}(x, y+\Delta) = \mathscr{P}(x, y) + \Delta\frac{\partial\mathscr{P}}{\partial y}(x, y) + o(\Delta).$$

Using this one can show that (5.6) yields,

$$(\mathbf{x}(\Delta) - x) + \left(H_{\mathbf{j}}(\mathbf{x}(\Delta)) - H_{\mathbf{j}}(x) - (\mathbf{x}(\Delta) - x)\right)\mathscr{P}(x, y) = \Delta(\mathbf{I} + \mathbf{II}) + o(\Delta),$$

where

$$|\mathbf{I}| := \left|\frac{(1.6)^{-\mathbf{j}}}{\mathfrak{l}_{\mathbf{j}+1}(y, w)}\left(K_{\mathbf{j},y}(x) - K_{\mathbf{j},y}(\underline{X})\right)\right| \leq (1.6)^{-\mathbf{j}}\varphi^{-N_{t+1}+\mathbf{j}}$$

and

$$
\begin{aligned}
|\mathbf{II}| :&= \left|\left(H_{\mathbf{j}}(\mathbf{x}(\Delta)) - H_{\mathbf{j}}\left(\underline{X}\right) - (\mathbf{x}(\Delta) - \underline{X})\right)\frac{\partial \mathscr{P}}{\partial y}(x,y)\right| \\
&\leq \left|\lambda_{t+1}^4 - 1\right| \max_{(x,y)\in U^c}\left|\frac{\partial \mathscr{P}}{\partial y}(x,y)\right| \left|\mathbf{x}(\Delta) - \underline{X}\right| \\
&\leq \left|\lambda_{t+1}^4 - 1\right| \left(\frac{3}{2}\varphi^{10}\right)\varphi^{-N_{t+1}} \\
&\leq \varphi^{-N_{t+1}}.
\end{aligned}
$$

For both inequalities we used the fact that

$$\max\left\{\left|\mathbf{x}(\Delta) - \underline{x}_{N_{t+1}}\right|, \left|x - \underline{x}_{N_{t+1}}\right|\right\} \leq \varphi^{-N_{t+1}}.$$

Since

$$\left|\left(H_{\mathbf{j}}(\mathbf{x}(\Delta)) - H_{\mathbf{j}}(x) - (\mathbf{x}(\Delta) - x)\right)\mathscr{P}(x,y)\right| \leq \left|\lambda_{t+1}^4 - 1\right| \left|\mathbf{x}(\Delta) - x\right|$$

we get by the triangle inequality that

$$
\begin{aligned}
|\mathbf{x}(\Delta) - x|\left(1 - \left|\lambda_{t+1}^4 - 1\right|\right) &\leq \left|\mathbf{x}(\Delta) - x + \left(H_{\mathbf{j}}(\mathbf{x}(\Delta)) - H_{\mathbf{j}}(x) - (\mathbf{x}(\Delta) - x)\right)\mathscr{P}(x,y)\right| \\
&\leq \Delta\left(|\mathbf{I}| + |\mathbf{II}|\right) + o(\Delta) \\
&\leq 2\Delta(1.6)^{-N_{t+1}} + o(\Delta).
\end{aligned}
$$

As $1 - |\lambda_{t+1} - 1| \geq \frac{1}{2}$ the conclusion of part (ii) follows. □

From now on, we work under the assumption that $(1.62)/\lambda_1^2 \geq 1.6$. As λ_1 can be made arbitrarily small this is compatible with the inductive procedure.

Corollary 24. – (i) *For every $(x,y) \in R_1 \cup R_3$, if $N_t < \mathbf{j}(x,y) \leq N_{t+1}$ then*

$$\left|\frac{\partial K_{N_{t+1},y}(x)}{\partial y}\right| \leq 6(1.6)^{-\mathbf{j}(x,y)}.$$

In addition if $(x,y) \in \partial U^c \cup \partial \mathbb{M}$ then

$$\frac{\partial K_{N_{t+1},y}(x)}{\partial y} = 0.$$

(ii) *Assume that $\{N_k, M_{k-1}, \epsilon_k, \lambda_k\}_{k=1}^s$ are chosen, there exists a choice of $M_s, \lambda_{s+1}, N_{s+1}$ and ϵ_{s+1} (compatible with the inductive procedure) such that*

$$\left|\frac{\partial \mathfrak{l}_{M_{t+1}}}{\partial y}(y,w)\right| \leq 3(1.6)^{-N_{t+1}}.$$

Proof. – (i) First we claim that for all $w \in \Sigma_{\mathbf{A}}$ such that $[w]_1^{\mathbf{j}(x,y)} = \mathbf{w}(\mathbf{j}(x,y))$,

$$\frac{\partial K_{N_{t+1},y}\left(\underline{x}_{N_{t+1}}\right)}{\partial y} = \pm 4(1.6)^{-\mathbf{j}(x,y)}. \tag{5.7}$$

This is true because of the following argument. For each $\mathbf{j}(x,y) < n \leq N_{t+1}$, either $\underline{x}_{n-1} = \underline{x}_n$ and then

$$K_{n,y}\left(\underline{x}_n\right) = K_{n-1,y}\left(\underline{x}_{n-1}\right)$$

or $\underline{x}_n = \underline{x}_{n-1} + \varphi^{-1}(\bar{x}_{n-1} - \underline{x}_{n-1})$ and then

$$K_{n,y}(\underline{x}_n) = K_{n-1,y}(\underline{x}_n) = K_{n-1,y}(\underline{x}_{n-1}) + \frac{\lambda_{t+1}\varphi}{1+\lambda_{t+1}\varphi}\mathfrak{l}_n(y,w).$$

This equality remains true in a neighborhood of y. Therefore for all $\mathbf{j}(x,y) < n \leq N_{t+1}$,

$$\begin{aligned}\left|\frac{\partial K_{n,y}(\underline{x}_n(w))}{\partial y}\right| &\leq \left|\frac{\partial K_{n-1,y}(\underline{x}_{n-1}(w))}{\partial y}\right| + \frac{2}{3}\max_{[w]_1^{N_{t+1}} \subset [\mathbf{w}(\mathbf{j}(x,y))]}\left|\frac{\partial \mathfrak{l}_n}{\partial y}(y,w)\right| \\ &\overset{\text{Lem. 22}}{\leq} \left|\frac{\partial K_{n-1,y}(\underline{x}_{n-1}(w))}{\partial y}\right| + \frac{2}{3}(1.6)^{-n}\end{aligned}$$

and so

$$\begin{aligned}\left|\frac{\partial K_{N_{t+1},y}\left(\underline{x}_{N_{t+1}}(w)\right)}{\partial y}\right| &\leq \left|\frac{\partial K_{\mathbf{j}(x,y),y}\left(\underline{x}_{\mathbf{j}(x,y)}\right)}{\partial y}\right| + \frac{2}{3}\sum_{n=\mathbf{j}(x,y)+1}^{N_{t+1}}(1.6)^{-n} \\ &\overset{(5.4)}{\leq} 4(1.6)^{-\mathbf{j}(x,y)}.\end{aligned}$$

Now for a general $0 \leq x \leq 1/\varphi$,

$$\begin{aligned}\frac{\partial K_{N_{t+1},y}(x)}{\partial y} &= \lim_{\Delta\to 0}\frac{K_{N_{t+1},y+\Delta}(x) - K_{N_{t+1},y}(x)}{\Delta} \\ &= \lim_{\Delta\to 0}\frac{K_{N_{t+1},y+\Delta}(\mathbf{x}(\Delta)) - K_{N_{t+1},y}(x)}{\Delta} \\ &\quad + \lim_{\Delta\to 0}\frac{K_{N_{t+1},y+\Delta}(\mathbf{x}(\Delta)) - K_{N_{t+1},y+\Delta}(x)}{\Delta}.\end{aligned}$$

As

$$\left|\frac{\partial K_{N_{t+1},y+\Delta}(x)}{\partial x}\right| \leq \lambda_{t+1}^{2(N_{t+1}-\mathbf{j}(x,y))},$$

it follows that

$$\begin{aligned}\lim_{\Delta\to 0}\left|\frac{K_{N_{t+1},y+\Delta}(\mathbf{x}(\Delta)) - K_{N_{t+1},y+\Delta}(x)}{\Delta}\right| &\leq \lambda_{t+1}^{2N_{t+1}}\lim_{\Delta\to 0}\left|\frac{\mathbf{x}(\Delta)-x}{\Delta}\right| \\ &\overset{\text{Lem 23}}{\leq} \left(\frac{1.62}{\lambda_1^2}\right)^{-N_{t+1}}.\end{aligned} \tag{5.8}$$

By Lemma 23.(i) if $x \in C_{[w]_1^{N_{t+1}}}$ then,

$$\begin{aligned}\lim_{\Delta\to 0}\left|\frac{K_{N_{t+1},y+\Delta}(\mathbf{x}(\Delta)) - K_{N_{t+1},y}(x)}{\Delta}\right| &\leq \lim_{\Delta\to 0}\left|\frac{K_{N_{t+1},y+\Delta}\left(\underline{x}_{N_{t+1}}(w)\right) - K_{N_{t+1},y}\left(\underline{x}_{N_{t+1}}(w)\right)}{\Delta}\right| \\ &\quad + \beta_n(x)\lim_{\Delta\to 0}\left|\frac{\mathfrak{l}_{N_{t+1}}(y+\Delta,w) - \mathfrak{l}_{N_{t+1}}(y,w)}{\Delta}\right| \\ &\leq \left|\frac{\partial K_{N_{t+1},y}\left(\underline{x}_{N_{t+1}}\right)}{\partial y}\right| + \left|\frac{\partial \mathfrak{l}_{N_{t+1}}}{\partial y}(y,w)\right| \\ &\leq 4(1.6)^{-\mathbf{j}(x,y)} + (1.6)^{-N_{t+1}}\end{aligned}$$

and the conclusion follows.

The second part of (i) in the corollary is true since if $(x,y)\in\partial U^c\cup\partial\mathbb{M}$, then $\frac{\partial\mathscr{P}}{\partial y}(x,y)=0$. Therefore $\frac{\partial \mathfrak{l}_{\mathbf{j}(x,y)+1}}{\partial y}(y,w)=0$ and $\mathbf{x}(\Delta)=x+o(\Delta)$.

(ii) Let $w\in\Sigma_{\mathbf{A}}(M_{t+1})$. As for all y, $\left\{K_{N_{t+1},y}\left(\underline{x}_{M_{t+1}}\right), K_{N_{t+1},y}\left(\bar{x}_{M_{t+1}}\right):\ w\in\Sigma_{\mathbf{A}}(M_{t+1})\right\}$ are fixed points for $h_{M_{t+1},y}$ it follows that for all Δ,

$$\begin{aligned}\mathfrak{l}_{M_{t+1}}(y+\Delta,w)&=K_{M_{t+1}-1,y+\Delta}\left(\bar{x}_{M_{t+1}}\right)-K_{M_{t+1}-1,y+\Delta}\left(\underline{x}_{M_{t+1}}\right)\\&=K_{N_{t+1},y+\Delta}\left(\bar{x}_{M_{t+1}}\right)-K_{N_{t+1},y+\Delta}\left(\underline{x}_{M_{t+1}}\right).\end{aligned}$$

The last line follows from $h_{n,y+\Delta}=id_{\mathbb{T}}$ for $N_{t+1}<n<M_{t+1}$. Writing $\underline{\mathbf{x}}(\Delta)$ (respectively $\bar{\mathbf{x}}(\Delta)$) for the x-delta point of $\underline{x}_{M_{t+1}}(w)$ (respectively $\bar{x}_{M_{t+1}}(w)$), by Lemma 23, for $|\Delta|$ small,

$$K_{N_{t+1},y+\Delta}\left(\underline{\mathbf{x}}(\Delta)\right)=K_{N_{t+1},y+\Delta}\left(\underline{x}_{N_{t+1}}\right)+\beta_{N_{t+1}}\left(\underline{x}_{M_{t+1}}\right)\mathfrak{l}_{N_{t+1}}(y+\Delta,w),$$

and

$$K_{N_{t+1},y+\Delta}\left(\bar{\mathbf{x}}(\Delta)\right)=K_{N_{t+1},y+\Delta}\left(\underline{x}_{N_{t+1}}\right)+\beta_{N_{t+1}}\left(\bar{x}_{M_{t+1}}\right)\mathfrak{l}_{N_{t+1}}(y+\Delta,w).$$

It then follows that for $|\Delta|$ small,

$$\mathfrak{l}_{M_{t+1}}(y+\Delta,w)=\left\{\beta_{N_{t+1}}\left(\bar{x}_{M_{t+1}}\right)-\beta_{N_{t+1}}\left(\underline{x}_{M_{t+1}}\right)\right\}\mathfrak{l}_{N_{t+1}}(y+\Delta,w)+\overline{I}_\Delta+\underline{I}_\Delta,$$

where by (5.8),

$$\left|\overline{I}_\Delta\right|:=\left|K_{N_{t+1},y+\Delta}\left(\bar{\mathbf{x}}(\Delta)\right)-K_{N_{t+1},y+\Delta}\left(\bar{x}_{M_{t+1}}\right)\right|\le(1.1)\,\Delta(1.59)^{-N_{t+1}}$$

and

$$\left|\underline{I}_\Delta\right|:=\left|K_{N_{t+1},y+\Delta}\left(\underline{\mathbf{x}}(\Delta)\right)-K_{N_{t+1},y+\Delta}\left(\underline{x}_{M_{t+1}}\right)\right|\le(1.1)\,\Delta(1.59)^{-N_{t+1}}.$$

It then follows that

$$\begin{aligned}\left|\frac{\partial\mathfrak{l}_{M_{t+1}}}{\partial y}(y,w)\right|&\le\underbrace{\left\{\beta_{N_{t+1}}\left(\bar{x}_{M_{t+1}}\right)-\beta_{N_{t+1}}\left(\underline{x}_{M_{t+1}}\right)\right\}}_{\ll 1}\left|\frac{\partial\mathfrak{l}_{N_{t+1}}}{\partial y}(y,w)\right|+\lim_{\Delta\to0}\frac{\left|\overline{I}_\Delta\right|+\left|\underline{I}_\Delta\right|}{\Delta}\\&\lesssim 3(1.6)^{-N_{t+1}}\ \text{as } M_{t+1}\to\infty.\end{aligned}$$ □

So far we have managed to show how to control $\frac{\partial K_{N_{t+1},y}(x)}{\partial y}$ by a constant times the derivative at level $\mathbf{j}(x,y)$ where $t=t(y)=\min\{t\in\mathbb{N}:N_{t+1}\ge\mathbf{j}(x,y)\}$. The next step is for $s>t(y)$, to obtain a relation between $\frac{\partial K_{s+1,y}(x)}{\partial y}$ and $\frac{\partial K_{N_{s,y}}}{\partial y}$.

Definition 25. – For $M_s<n\le N_{s+1}$, $x\in C_{[w]_1^{N_{s+1}}}$ and $|\Delta|$ small we define $\mathbf{x}_s(\Delta)$ to be the unique point such that

$$\frac{K_{M_s,y+\Delta}\left(\mathbf{x}_s(\Delta)\right)-K_{M_s,y+\Delta}\left(\underline{x}_{M_s+1}\right)}{\mathfrak{l}_{M_s+1}(y+\Delta,w)}=\frac{K_{M_s,y}(x)-K_{M_s,y}\left(\underline{x}_{M_s+1}\right)}{\mathfrak{l}_{M_s+1}(y,w)}.$$

Setting similarly to before for $0\le x\le1/\varphi$ and y such that $\mathbf{j}(x,y)<M_s$,

$$\beta_{M_s+1}(x):=\psi_{s+1}\left(\frac{K_{M_s,y}(x)-K_{M_s,y}\left(\underline{x}_{M_s+1}\right)}{\mathfrak{l}_{M_s+1}(y,w)}\right)$$

and for $M_s + 1 < n \leq N_{s+1}$,

$$\beta_n(x) := \psi_{s+1}\left(\frac{\beta_{n-1}(x) - \beta_{n-1}(\underline{x}_n)}{\beta_{n-1}(\bar{x}_n) - \beta_{n-1}(\underline{x}_n)}\right).$$

LEMMA 26. – *For all $(x, y) \in R_1 \cup R_3$ with $\mathbf{j}(x, y) < N_s$ the following holds:*

1. *for every $|\Delta|$ small,*

$$K_{N_{s+1},y+\Delta}(\mathbf{x}_s(\Delta)) = K_{N_{s+1},y+\Delta}\left(\underline{x}_{N_{s+1}}\right) + \beta_{N_{s+1}}(x) \mathfrak{l}_{N_{s+1}}(y + \Delta, w).$$

2. (i) *For every $M_s < n \leq N_{s+1}$,*

$$\left|\frac{\partial \mathfrak{l}_n}{\partial y}(y, w)\right| \leq (1.6)^{M_s - n}\left|\frac{\partial \mathfrak{l}_{M_s}}{\partial y}(y, w)\right|.$$

(ii) $\left|\frac{\partial \mathfrak{l}_{M_s}}{\partial y}(y, w)\right| \leq (1.6)^{-n_s}$.

(iii) *Assume that $\{N_k, M_{k-1}, \epsilon_k, \lambda_k\}_{k=1}^s$ are chosen, there exists a choice of $M_s, \lambda_{s+1}, N_{s+1}$ and ϵ_{s+1} (compatible with the inductive procedure) such that*

$$|\mathbf{x}_s(\Delta) - x| \leq \varphi^{-N_{s+1}}\Delta + o(\Delta)$$

as $\Delta \to 0$.

Proof. – This is done by induction on s. The base of induction is the first $s \in \mathbb{N}$ such that $N_s > \mathbf{j}(x, y)$.

1. This is similar to the proof of Lemma 21 and Lemma 23.(i).

2. (i) Let $w \in \Sigma_{\mathbf{A}}$. The starting point is that by the definition of $h_{M_t,y}$ (as a distribution correcting function), Equation 4.2 holds for $K_{M_t,y}$. Therefore for all $w \in \Sigma_{\mathbf{A}}$ and $h > 0$ small,

$$m_{\mathbb{T}}\left(K_{M_s,y+h}\left(C_{[w]_1^{N_{s+1}}}\right)\right) = \mathfrak{l}_{M_s}(y + h, w)\frac{m_{\mathbb{T}}\left(C_{[w]_1^{N_{s+1}}}\right)}{m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)}.$$

The rest is similar to the proof of the first part of Lemma 22 with $\mathbf{j}(x, y)$ replaced by M_s.

2. (ii) Since for all y, $\left\{K_{N_s,y}\left(\underline{x}_{M_s}(w)\right), K_{N_s,y}(\bar{x}_{M_s}(w)) : w \in \Sigma_{\mathbf{A}}(M_s)\right\}$ are fixed points for $h_{M_s,y}$ it follows that (here $\underline{x}_{M_s} = \underline{x}_{M_s}(w)$)

$$\mathfrak{l}_{M_s}(y + \Delta, w) = K_{N_s,y+\Delta}(\bar{x}_{M_s}) - K_{N_s,y+\Delta}\left(\underline{x}_{M_s}\right).$$

The base of induction is Corollary 24.(ii).

The proof of the inductive step is the same as the proof of the base of induction where we use the induction hypothesis that

$$\left|\frac{\partial \mathfrak{l}_{N_s}}{\partial y}(y, w)\right| \leq (1.6)^{M_{s-1} - N_s}\left|\frac{\partial \mathfrak{l}_{M_{s-1}}}{\partial y}(y, w)\right| \leq (1.6)^{-N_{s-1} - n_s},$$

and

$$|\mathbf{x}_s(\Delta) - x| \leq \varphi^{-N_{s+1}}\Delta + o(\Delta).$$

Therefore,

$$\left|\overline{I}_\Delta(s)\right| := \left|K_{N_{t+1},y+\Delta}(\bar{\mathbf{x}}_s(\Delta)) - K_{N_{t+1},y+\Delta}(\bar{x}_{M_s})\right| \leq \Delta(1.6)^{-N_{s+1}}$$

and

$$|\underline{I}_\Delta| := \left|K_{N_{t+1},y+\Delta}(\underline{\mathbf{x}}_s(\Delta)) - K_{N_{t+1},y+\Delta}\left(\underline{x}_{M_s}\right)\right| \leq \Delta(1.6)^{-N_{s+1}}.$$

It then follows that

$$\begin{aligned}\left|\frac{\partial \mathfrak{l}_{M_s}}{\partial y}(y,w)\right| &\le \underbrace{\left\{\beta_{N_s}\left(\bar{x}_{M_s}\right)-\beta_{N_s}\left(\underline{x}_{M_s}\right)\right\}}_{\ll 1}\left|\frac{\partial \mathfrak{l}_{N_s}}{\partial y}(y,w)\right| + \lim_{\Delta\to 0}\frac{\left|\overline{I}_{\Delta}(s)\right|+\left|\underline{I}_{\Delta}(s)\right|}{\Delta} \\ &\lesssim (1.6)^{-N_{s-1}}(1.6)^{-n_s} + 2(1.6)^{-N_{s+1}} \\ &\le (1.6)^{-n_s}.\end{aligned}$$

(iii) We first recall the definition of $K_{M_s,y}$. Define $\alpha : \mathbb{N}\times\Sigma_{\mathbf{A}}\times\left[-\varphi/(\varphi+2),\varphi^2/(\varphi+2)\right]$ by

$$\alpha(s,w,y) := \frac{\int_{C_{[w]_1^{M_s}}}\frac{\partial K_{M_s-1,y}}{\partial x}(x)dx}{m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)} = \frac{\mathfrak{l}_{M_s}(y,w)}{m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)}.$$

It follows that for $s\in\mathbb{N}$ such that $N_s > \mathbf{j}(x,y)$, the definition of $K_{M_s,y}$ restricted to $C_{[w]_1^{M_s}}$ is the function defined by $K_{M_s,y}\left(\underline{x}_{M_s}\right) = K_{N_s,y}\left(\underline{x}_{M_s}\right)$ and x-derivative

$$\frac{\partial K_{M_s,y}}{\partial x}(x) := \begin{cases}\frac{\partial G_{\alpha(s,w^-,y),\alpha(s,w,y)}}{\partial x}\left(\frac{x-\underline{x}_{M_s}}{\epsilon_{t+1}m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)}\right) & 0\le x-\underline{x}_{M_s}\le \epsilon_{s+1}m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)\\ \alpha(s,w,y) & x-\underline{x}_{M_s}\ge \epsilon_{s+1}m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right),\end{cases}$$

where w^- is the predecessor of w in $\Sigma_{\mathbf{A}}(M_t)$ and $G_{\alpha_1,\alpha_2} : [0,1]\to[0,1]$ is the function defined by (4.3).

Therefore the function

$$\mathscr{K}_{t,y}(x) := \frac{K_{M_s,y}(x)-K_{M_s,y}\left(\underline{x}_{M_s}\right)}{\mathfrak{l}_{M_s}(y,w)} = \left(m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)\right)^{-1}\left[\frac{K_{M_s,y}(x)-K_{M_s,y}\left(\underline{x}_{M_s}\right)}{(s,y,w)}\right]$$

satisfies $\mathscr{K}_{t,y}\left(\underline{x}_{M_t}(w)\right)=0$ and

$$\begin{aligned}&\frac{\partial \mathscr{K}_{t,y}}{\partial x}(x)\\ &:= m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)^{-1}\cdot\begin{cases}\frac{\partial G_{\Upsilon(s,w,y),1}}{\partial x}\left(\frac{x-\underline{x}_{M_s}}{\epsilon_{s+1}m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)}\right), & 0\le x-\underline{x}_{M_s}\le \epsilon_{s+1}m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)\\ 1 & x-\underline{x}_{M_s}\ge \epsilon_{s+1}m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right),\end{cases}\end{aligned}$$

where $\Upsilon(s,w,y) = \frac{\alpha(s,w^-,y)}{\alpha(s,w,y)}$. The function $\mathscr{K}_{s,y}$ is important since the definition of $\mathbf{x}_s(\Delta)$ is as the unique point so that

$$\mathscr{K}_{s,y+\Delta}\left(\mathbf{x}_t(\Delta)\right) = \mathscr{K}_{t,y}(x).$$

Since for all $0 < a$, $\int_0^1 G'_{a,1}(x)dx = 1$, it follows that if $x - \underline{x}_{M_s} \geq \epsilon_{s+1} m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)$ then writing $\hat{x} = \underline{x}_{M_s} + \epsilon m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)$

$$\begin{aligned}\mathscr{K}_{s,y+\Delta}(x) &= \int_{\underline{x}_{M_s}}^{x} \frac{\partial \mathscr{K}_{s,y+\Delta}}{\partial x}(x)\\ &= \left(m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)\right)^{-1}\left(\int_{\underline{x}_{M_s}}^{\hat{x}} \frac{\partial G_{\Upsilon(s,w,y+\Delta),1}}{\partial x}\left(\frac{x-\underline{x}_{M_s}}{\epsilon_{s+1} m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)}\right)dx + (x-\hat{x})\right)\\ &= \frac{x-\underline{x}_{M_s}}{m_{\mathbb{T}}\left(C_{[w]_1^{M_s}}\right)} = \mathscr{K}_{s,y}(x) \quad \text{by a similar reasoning.}\end{aligned}$$

If $x - \underline{x}_{M_s} \leq \epsilon_{s+1} m\left(C_{[w]_1^{M_s}}\right)$ then using the fact that for all $0 \leq x \leq 1$, $\left|G_{a,1}(x) - G_{b,1}(x)\right| \leq x\,|b-a|$,

$$\begin{aligned}\mathscr{K}_{s,y+\Delta}(x) &= \mathscr{K}_{s,y}(x) \pm |\Upsilon(s,w,y+\Delta) - \Upsilon(s,w,y)| \cdot \left(x - \underline{x}_{M_s}(w)\right)\\ &= \mathscr{K}_{s,y}(x) \pm \epsilon_{s+1}\,|\Upsilon(s,w,y+\Delta) - \Upsilon(s,w,y)|\,.\end{aligned}$$

$$\begin{aligned}|\Upsilon(s,w,y+\Delta) - \Upsilon(s,w,y)| &= \frac{m\left(C_{[w]_1^{M_s}}\right)}{m\left(C_{[w^-]_1^{M_s}}\right)}\left|\frac{\mathfrak{l}_{M_s}(y+\Delta,w^-)}{\mathfrak{l}_{M_s}(y+\Delta,w)} - \frac{\mathfrak{l}_{M_s}(y,w^-)}{\mathfrak{l}_{M_s}(y,w)}\right|\\ &\leq \Delta\frac{\varphi\left(\left|\frac{\partial \mathfrak{l}_{M_s}}{\partial y}(y,w)\right| + \left|\frac{\partial \mathfrak{l}_{M_s}}{\partial y}(y,w^-)\right|\right)}{\min\left(\mathfrak{l}_{M_s}(y,w), \mathfrak{l}_{M_s}(y,w^-)\right)} + o(\Delta).\end{aligned}$$

Thus if ϵ_{s+1} is sufficiently small (this choice depends on $\{N_t, M_t, \lambda_t : t \leq t\}$ and N_{s+1}) then

$$\mathscr{K}_{s,y+\Delta}(x) = \mathscr{K}_{s,y}(x) \pm \frac{1}{2}\Delta\varphi^{-N_{s+1}-M_s} + o(\Delta),$$

and for all Δ sufficiently small

$$\begin{aligned}\mathscr{K}_{s,y}(x) &= \mathscr{K}_{s,y+\Delta}(\mathbf{x}_s(\Delta))\\ &= \mathscr{K}_{s,y}(\mathbf{x}_s(\Delta)) \pm \frac{1}{2}\Delta\varphi^{-N_{s+1}} + o(\Delta).\end{aligned}$$

Since

$$\frac{\partial \mathscr{K}_{s,y}}{\partial x}(x) \geq \left(2m_{\mathbb{T}}\left(C_{[w]_1^{M_t}}\right)\right)^{-1} \geq \frac{\varphi^{-M_s}}{2},$$

it follows that

$$\begin{aligned}|x - \mathbf{x}_s(\Delta)| &\leq 2\varphi^{M_s}\left|\mathscr{K}_{s,y}(x) - \mathscr{K}_{s,y}(\mathbf{x}_t(\Delta))\right|\\ &\leq \Delta\varphi^{-N_{s+1}} + o(\Delta)\end{aligned}$$

as required. □

The next corollary is the final ingredient for the proof of Theorem 17.

Corollary 27. – *There exists a choice of* $\{\lambda_t, n_t, N_t, m_t, M_t, \epsilon_t\}_{t\in\mathbb{N}}$ *such that:*

(i) *For all* $(x,y)\in\mathbb{M}_\sim$ *and for all* $t\in\mathbb{N}$ *such that* $\mathbf{j}(x,y)<N_t$,

$$\left|\frac{\partial K_{N_{t+1},y}(x)}{\partial y}\right|\leq\left|\frac{\partial K_{N_t,y}\left(\underline{x}_{M_t}(w)\right)}{\partial y}\right|+2(1.6)^{-n_t}.$$

(ii) $\frac{\partial K_{N_{t+1},y}(x)}{\partial y}$ converges uniformly in $R_1\cup R_3$ as $t\to\infty$. For every $(x,y)\in\partial(R_1\cup R_3)$ or $(x,y)\in U$,

$$\lim_{t\to\infty}\left|\frac{\partial K_{N_{t+1},y}(x)}{\partial y}\right|=0.$$

Proof. – We assume that $\{\lambda_t,n_t,N_t,m_t,M_t,\epsilon_t\}$ are chosen so that Lemma 26 holds.

(i) Similarly as in the proof of Corollary 24.(i) one can use the fact that

$$\left|\frac{\partial K_{N_{t+1},y}\left(\underline{x}_{N_{t+1}}(w)\right)}{\partial y}\right|\leq\left|\frac{\partial K_{M_t,y}\left(\underline{x}_{M_t}(w)\right)}{\partial y}\right|+\sum_{n=M_t+1}^{N_{t+1}}\left|\frac{\partial \mathfrak{l}_n}{\partial y}(y,w)\right|,$$

and

$$K_{M_t,y}\left(\underline{x}_{M_t}(w)\right)=K_{N_t,y}\left(\underline{x}_{M_t}(w)\right)$$

to show that

$$\begin{aligned}\left|\frac{\partial K_{N_{t+1},y}\left(\underline{x}_{N_{t+1}}\right)}{\partial y}\right|&\overset{\text{Lem.26.2.i}}{\leq}\left|\frac{\partial K_{N_t,y}\left(\underline{x}_{M_t}\right)}{\partial y}\right|+\sum_{n=M_t+1}^{N_{t+1}}(1.6)^{M_t-n}\left|\frac{\partial \mathfrak{l}_{M_t}}{\partial y}(y,w)\right|\\&\overset{\text{Lem.26.2.ii}}{\leq}\left|\frac{\partial K_{N_t,y}\left(\underline{x}_{M_t}\right)}{\partial y}\right|+(1.6)^{-n_t}\sum_{n=M_t+1}^{N_t}(1.6)^{M_t-n}\\&\leq\left|\frac{\partial K_{N_t,y}\left(\underline{x}_{M_t}\right)}{\partial y}\right|+\frac{5}{3}(1.6)^{-n_t}.\end{aligned}$$

Therefore by Lemma 26,

$$\begin{aligned}&\lim_{\Delta\to0}\left|\frac{K_{N_{t+1},y+\Delta}\left(\mathbf{x}_t(\Delta)\right)-K_{N_{t+1},y}(x)}{\Delta}\right|\\&\quad\leq\left(\left|\frac{\partial K_{N_t,y}\left(\underline{x}_{M_t}\right)}{\partial y}\right|+\frac{5}{3}(1.6)^{-n_t}\right)+\beta_n(x)\frac{\partial l_{N_{t+1}}}{\partial y}(y,w)\\&\quad\leq\left(\left|\frac{\partial K_{N_t,y}\left(\underline{x}_{M_t}\right)}{\partial y}\right|+\frac{5}{3}(1.6)^{-n_t}\right)+\beta_n(x)\,(1.6)^{-n_{t+1}}\end{aligned}$$

and by Lemma 26.2.(iii),

$$\begin{aligned}\left|\frac{K_{N_{t+1},y+\Delta}\left(\mathbf{x}_t(\Delta)\right)-K_{N_{t+1},y+\Delta}(x)}{\Delta}\right|&\leq\left(\sup_{(x,y)\in\mathbb{M}_\sim}\left|\frac{\partial K_{N_{t+1},y+\Delta}(x)}{\partial x}\right|\right)\left(\frac{\left|\mathbf{x}_t(\Delta)-x\right|}{\Delta}\right)\\&\leq\lambda_1^{2N_{t+1}}\varphi^{-N_{t+1}}+o(\Delta)\leq(1.6)^{-N_t}+o(\Delta).\end{aligned}$$

A combination of the previous two inequalities and $n_t=o(N_t)$, $N_t=o\left(n_{t+1}\right)$ shows that

$$\left|\frac{\partial K_{N_{t+1},y}(x))}{\partial y}\right|\leq\left|\frac{\partial K_{N_t,y}\left(\underline{x}_{M_t}(w)\right)}{\partial y}\right|+2(1.6)^{-n_t}.$$

(ii) Let $t \in \mathbb{N}$, and $(x, y) \in R_1 \cup R_3$. By applying part (i) of this corollary repeatedly one has that with $t(x, y) = \min\{t : N_t > \mathbf{j}(x, y)\}$. For $t > t(x, y) + 1$ and w such that $x \in C_{[w]_1^{N_t+1}}$, by the first part of the corollary

$$\left|\frac{\partial K_{N_t,y}\left(\underline{x}_{M_t}\right)}{\partial y}\right| \leq \left|\frac{\partial K_{N_{t-1},y}\left(\underline{x}_{M_{t-1}}\right)}{\partial y}\right| + 2\,(1.6)^{-n_t}.$$

For $t = t(x, y)$ by Corollary 24

$$\left|\frac{\partial K_{N_{t(x,y)},y}\left(\underline{x}_{M_{t(x,y)}}\right)}{\partial y}\right| \leq (1.6)^{-\mathbf{j}(x,y)}.$$

A combination of these two observations shows that for $t > t(x, y)$

$$\begin{aligned}\left|\frac{\partial K_{N_t,y}(x)}{\partial y}\right| &\leq \left|\frac{\partial K_{N_{t(x,y)},y}\left(\underline{x}_{M_{t(x,y)}}\right)}{\partial y}\right| + 2\sum_{k=t(x,y)}^{t}(1.6)^{-n_k}.\\ &\leq (1.6)^{-\mathbf{j}(x,y)} + 2(1.6)^{-n_{t(x,y)}}.\end{aligned}$$

This is enough to show that $\left\{\frac{\partial K_{N_{s+1},y}(x)}{\partial y}\right\}_{s=1}^{\infty}$ is a Cauchy sequence in the uniform topology. Indeed if $s, t > t(x, y)$, then

$$\left|\frac{\partial K_{N_t,y}(x)}{\partial y} - \frac{\partial K_{N_s,y}(x)}{\partial y}\right| \leq \sum_{k=t}^{s}(1.6)^{-n_k} \leq 2(1.6)^{-n_t}.$$

If $N_t < \mathbf{j}(x, y) - 3 \leq N_{s+1}$ then $K_{N_{t+1},y}(x) = x$ in a neighborhood of y and hence

$$\begin{aligned}\left|\frac{\partial K_{N_{t+1},y}(x)}{\partial y} - \frac{\partial K_{N_{s+1},y}(x)}{\partial y}\right| &= \left|\frac{\partial K_{N_{s+1},y}(x)}{\partial y}\right|\\ &\leq (1.6)^{-\mathbf{j}(x,y)} + 2(1.6)^{-n_t}\\ &\leq 3(1.6)^{-n_t}.\end{aligned}$$

We leave the bound on the easier cases $t = t(x, y)-1 < s, s, t < t(x, y)-1$ to the reader. □

REMARK 28. – The latter corollary shows that $\lim_{t\to\infty} \frac{\partial K_{N_t,y}}{\partial y}$ is uniformly continuous in x as a uniform limit of continuous functions. As a consequence, since for all $x \in [0, 1/\varphi]$, the sequence $K^{-1}_{N_t,y}(x)$ converges uniformly to $\mathfrak{h}_y^{-1}(x)$ and $\mathfrak{h}_y^{-1}$ is a homeomorphism of $[0, 1/\varphi]$ then for all $x \in [0, 1/\varphi]$,

$$\frac{\partial K_{N_t,y}\left(K^{-1}_{N_t,y}(x)\right)}{\partial y} \xrightarrow[t\to\infty]{} \lim_{t\to\infty} \frac{\partial K_{N_t,y}\left(\mathfrak{h}_y^{-1}(x)\right)}{\partial y}$$

and the convergence is uniform in $t \in \mathbb{N}$.

Proof of Theorem 17. – Lemma 20 shows that $\frac{\partial \mathfrak{z}}{\partial x} = \lim_{t\to\infty} \frac{\partial \mathfrak{z}_{N_t}}{\partial x}(x, y)$ exists and is a continuous function of $\mathbb{M}_\sim$. It remains to show that $\frac{\partial \mathfrak{z}}{\partial y} = \lim_{t\to\infty} \frac{\partial \mathfrak{z}_{N_t}}{\partial y}(x, y)$ exists and is a continuous function of $\mathbb{M}_\sim$. To this end, write

$$B_N(x, y) = \begin{pmatrix} \frac{\partial K_{N,y}(x)}{\partial x} & \frac{\partial K_{N,y}(x)}{\partial y} \\ 0 & 1 \end{pmatrix}$$

for the differential of the map $(x, y) \mapsto \big(K_{N,y}(x), y\big)$. By the chain rule

$$D\mathfrak{z}_{N_t}(x, y) = B_{N_t}\left(SK_{N_t,y}^{-1}(x), -y/\varphi)\right)\begin{pmatrix}\varphi & 0\\ 0 & -1/\varphi\end{pmatrix} B_{N_t}^{-1}\left(K_{N_t,y}^{-1}(x), y\right).$$

This yields that

$$\begin{aligned}\frac{\partial\mathfrak{z}_{N_t}}{\partial y}(x, y) &= \varphi\frac{\partial K_{N_t,-y/\varphi}(SK_{N_t,y}^{-1}(x))}{\partial x}\\ &\quad\cdot\left(-\left(\frac{\partial K_{N_t,-y/\varphi}\left(K_{N_t,y}^{-1}(x)\right)}{\partial x}\right)^{-1}\left(\frac{\partial K_{N_t,y}\left(K_{N_t,y}^{-1}(x)\right)}{\partial y}\right)\right)\\ &\quad-\frac{1}{\varphi}\frac{\partial K_{N_t,-y/\varphi}\left(SK_{N_t,y}^{-1}(x)\right)}{\partial y}\\ &= -\frac{\partial\mathfrak{z}_{N_t}}{\partial x}(x, y)\cdot\frac{\partial K_{N_t,y}\left(K_{N_t,y}^{-1}(x)\right)}{\partial y}-\frac{1}{\varphi}\frac{\partial K_{N_t,-y/\varphi}\left(SK_{N_t,y}^{-1}(x)\right)}{\partial y}.\end{aligned}$$

Since all the terms on the right hand side converge uniformly as $t \to \infty$, the theorem is proved. ☐

5.2. Proof of the Anosov property for $\mathfrak{z}$.

So far we have shown that $\mathfrak{z}_{N_t}$ converges uniformly to $\mathfrak{z}$ and we have estimated the derivatives. We are going to use the following well-known lemma, the proof of which can be found in [21]. A function $\mathscr{A} : \mathbb{M}_\sim \times \mathbb{Z} \to SL(2, \mathbb{R})$ is *linear cocycle* over a homeomorphism $\mathfrak{f} : \mathbb{M}_\sim \to M_\sim$ if for any $m.n \in \mathbb{Z}$ and $x \in \mathbb{T}^2$,

$$\mathscr{A}_{m+n}(x) = \mathscr{A}_m \circ \mathfrak{f}^n(x)\,\mathscr{A}_n(x).$$

We say that the cocycle is hyperbolic if there are $\sigma > 1$ and $C > 0$ so that for every $x \in \mathbb{M}_\sim$ there exist transverse lines E_x^s and E_x^u in $\mathbb{R}^2$ such that

1. $\mathscr{A}(x)E_x^s = E_{\mathfrak{f}(x)}^s$ and $\mathscr{A}(x)E_x^u = E_{\mathfrak{f}(x)}^u$.
2. $|\mathscr{A}_n(x)v^s| \le C\sigma^n|v^s|$ and $|\mathscr{A}_{-n}(x)v^u| \le C\sigma^n|v^u|$ for every $v^s \in E_x^s$, $v^u \in E_x^u$ and $n \ge 1$.

PROPOSITION ([21, Prop. 2.1]). – *Let $A : \mathbb{M}_\sim \times \mathbb{Z} \to SL(2, \mathbb{R})$ be a linear cocycle over a homeomorphism $\mathfrak{f} : M_\sim \to M_\sim$. If there exists $v \in \mathbb{R}^2$, constants $c > 0$ and $\sigma > 1$ such that $|A_n(x)v| \ge c\sigma^n$ then $\mathscr{A}$ is hyperbolic. The transverse lines $E_x^s, E_{x,}^u$ in $\mathbb{R}^2$ satisfy that for any $\sigma_0 < \sigma$, there exists $C > 0$ so that for any $v^s \in E_x^s$, $v^u \in E_x^u$ and $n \ge 1$,*

$$|\mathscr{A}_n(x)v^s| \le C\sigma_0^n|v^s| \text{ and } |\mathscr{A}_{-n}(x)v^u| \le C\sigma_0^n|v^u|.$$

Proof that $\mathfrak{z}$ is Anosov. – Define $\mathscr{A} : \mathbb{M}_\sim \times \mathbb{Z} \to SL\,(2, \mathbb{R})$ by

$$\mathscr{A}(x) = \frac{1}{\det\left(D\mathfrak{z}(x, y)\right)}D\mathfrak{z}(x, y).$$

Since $D\mathfrak{z}$ is of the form

$$\begin{pmatrix}\frac{\partial\mathfrak{z}}{\partial x}(x, y) & *\\ 0 & -1/\varphi\end{pmatrix}$$

and

$$1.6 \leq \frac{\partial \mathfrak{z}}{\partial y}(x, y) \leq 1.7$$

one has that for all $(x, y) \in M_\sim$,

$$\frac{1.6}{\varphi} \leq |\det(D\mathfrak{z}(x, y))| \leq \frac{1.7}{\varphi}$$

and

$$\left| D\mathfrak{z}^n(x, y) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right| = \left| \begin{pmatrix} \prod_{k=0}^{n-1} \frac{\partial \mathfrak{z}}{\partial x} \circ \mathfrak{z}^k(x, y) \\ 0 \end{pmatrix} \right| \geq (1.6)^n \left| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right|.$$

It then follows that with $v = (1, 0)^{tr}$

$$|\mathscr{A}_n(x, y)v| = (\det(D\mathfrak{z}^n(x, y)))^{-1} |D\mathfrak{z}^n(x, y)v| \geq \left(\frac{\varphi \cdot 1.6}{1.7}\right)^n |v| \geq (1.52)^n |v|$$

there exist transverse lines E_x^s and E_x^u in $\mathbb{R}^2 \simeq T_x M_\sim$ and $C > 0$ so that for any $v^u \in E_x^u$,

$$|\mathscr{A}_n(x)v^u| \geq (1.5)^n |v^u|.$$

It follows that

$$\begin{aligned} |D\mathfrak{z}^n(x, y)v^u| &= |\mathscr{A}_n(x)v^u| \prod_{k=0}^{n-1} \left|\det\left(D\mathfrak{z}\left(\mathfrak{z}^k(x, y)\right)\right)\right| \\ &\geq C(1.5)^n \left(\frac{1.6}{\varphi}\right)^n |v^u| \\ &\geq C(1.48)^n |v^u|. \end{aligned}$$

Similarly one has for every $v^s \in E_x^s$,

$$\begin{aligned} |D\mathfrak{z}^n(x, y)v^s| &\leq C(1.5)^{-n} \left(\frac{1.7}{\varphi}\right)^n |v^s| \\ &\leq C\,(0.7)^n \end{aligned}$$

and so $\mathfrak{z} : \mathbb{M}_\sim \to \mathbb{M}_\sim$ is Anosov. □

5.3. Proof of the type III_1 property for $\mathfrak{z}$.

For $-\frac{\varphi}{\varphi+2} < y < \frac{1}{\varphi+2}$ let,

$$\mathfrak{K}_y(x) := \lim_{n\to\infty} K_{n,y}(x) : \mathbb{T} \to \mathbb{T}$$

and for $\frac{1}{\varphi+2} \leq y \leq \frac{\varphi^2}{\varphi+2}$,

$$\mathfrak{K}_y(x) := \lim_{n\to\infty} K_{n,y}(x) : [0, 1/\varphi] \to [0, 1/\varphi].$$

In both cases it is an orientation preserving homeomorphism.

We will show that the measures $m_{\mathrm{Leb}(\mathbb{T})} \circ \mathfrak{K}_y$ and $m_{\mathrm{Leb}([0,1/'])} \circ \mathfrak{K}_y$ are equivalent measures to μ^+, the measure on $\mathbb{T}$ arising from $\{\lambda_k, m_k.M_k, n_k.N_k\}$ in the previous section.

In addition, the Radon Nykodym derivative

$$\frac{d\eta_y}{d\mu^+}(x) : \mathbb{M} \to [0, \infty)$$

defined by

$$\frac{d\eta_y}{d\mu^+}(x) := \frac{dm_{\mathbb{T}} \circ \mathfrak{K}_y}{d\mu^+}(x),$$

is a $(\mathbb{M}_\sim, \mathscr{B}(\mathbb{M}_\sim), \mu)$ measurable function. This means that the measure η on $\mathbb{M}_\sim$ defined by

$$\begin{aligned}\int_{M_\sim} u(x,y)d\eta &= \int_{\mathbb{M}_\sim} u(x,y)d\eta_y(x)dy \\ &= \int_{\mathbb{M}_\sim} u(x,y)\frac{d\eta_y}{d\mu^+}(x)d\mu^+(x)dy\end{aligned}$$

is equivalent to $\mu = \mathbf{m}_{\mathbb{M}} \circ \mathfrak{H}_{\underline{0}} = \mu^+ \otimes dy$.

Since $\left(\mathbb{M}_\sim, \mathscr{B}_{\mathbb{M}}, \mu, \tilde{f}\right)$ is a type III_1 transformation and $\mu \sim \eta$, $\left(\mathbb{M}, \mathscr{B}_{\mathbb{M}}, \eta, \tilde{f}\right)$ is a type III_1 transformation. Thus $(\mathbb{M}, \mathscr{B}_{\mathbb{M}}, \mathrm{m}_{\mathbb{M}}, \mathfrak{Z})$ is a type III_1 transformation since $\pi(x,y) := (\mathfrak{K}_y(x), y) : (\mathbb{M}, \mathscr{B}_{\mathbb{M}}, \mathrm{m}_{\mathbb{M}}, \mathfrak{Z}) \to \left(\mathbb{M}, \mathscr{B}_{\mathbb{M}}, \eta, \tilde{f}\right)$ is an isomorphism. Therefore what is left to prove is the following.

LEMMA 29. – (i) *For all* $-\frac{\varphi}{\varphi+2} < y < \frac{1}{\varphi+2}$, $\left(\mathfrak{K}_y^{-1}\right)_* m_{\mathbb{T}}$ *is an equivalent measure to* μ^+ *(the measure on* $\mathbb{T}$ *arising from* $\{\lambda_k, m_k.M_k, n_k.N_k, \epsilon_k\}$ *in the previous section).*

(ii) *For all* $\frac{1}{\varphi+2} < y < \frac{\varphi^2}{\varphi+2}$, $\left(\mathfrak{K}_y\right)_* m_{\mathbb{T}}|_{[0,1/\varphi]}$ *is an equivalent measure to* $\mu^+|_{[0,1/\varphi]}$.

(iii) *The Radon Nykodym derivatives* $\frac{d\eta_y}{d\mu^+}(x) : \mathbb{M} \to [0,\infty)$ *are measurable in* $(\mathbb{M}_\sim, \mathscr{B}_{\mathbb{M}_\sim}, \mu)$.

Proof. – Fix $-\frac{\varphi}{\varphi+2} < y < \frac{1}{\varphi+2}$. The proof is the same as in Lemma 13 by using the theory of local absolute continuity of Shiryaev with $\mathscr{F}_t := \left\{C_{[w]_1^{N_t}} : w \in \Sigma_{\mathbf{A}}\right\}$. By the construction

$$\left(m_{\mathbb{T}} \circ \mathfrak{K}_y\right)_t := m_{\mathbb{T}} \circ \mathfrak{K}_y\big|_{\mathscr{F}_t} = m_{\mathbb{T}} \circ K_{N_t,y},$$

and

$$\left(\mu^+\right) = m_{\mathbb{T}} \circ H_{\underline{0},N_t}.$$

Therefore,

$$Z_{t,y}(x) := \frac{d\left(m_{\mathbb{T}} \circ \mathfrak{K}_y\right)_t}{d\left(\mu^+\right)_t}(x) = \frac{\left(\frac{\partial K_{N_t,y}}{\partial x}\right)}{\left(\frac{\partial H_{\underline{0},N_t}}{\partial x}\right)}(x).$$

The rest of the proof that $Z_{t,y}(x)$ is uniformly integrable and hence converges a.s. as $t \to \infty$ is the same as in Lemma 13. This proves (i) and (ii).

To see (iii), notice that the function

$$(x,y) \mapsto \frac{d\eta_y}{d\mu^+}(x) = \lim_{t\to\infty} Z_{t,y}(x)$$

is almost surely a limit of continuous functions, hence measurable. □

Appendix

Proof of Theorem 4

Assume that $\{\lambda_k, n_k, N_k, m_k, M_k\}_{k\geq 1}$ and M_0 are chosen via the inductive construction, $\{\pi_k, P_k\}_{k\in\mathbb{Z}}$ are defined by (3.10) and $\mu = \mathrm{M}\{\pi_k, P_k : k \in \mathbb{Z}\}$ .Again T denotes the shift on $\Sigma_{\mathbf{A}}$. The proof of non-singularity, the K-property of the shift with respect to μ and that

$$T'(x) = \frac{d\mu\circ T}{d\mu}(x) = \prod_{k=1}^{\infty} \frac{P_{k-1}(x_k, x_{k+1})}{P_k(x_k, x_{k+1})}$$

appears in [12, Thm. 6]. In order to show the other properties of the Markov Shift, we will need a more concrete expression of the Radon Nykodym derivatives. The measure μ, or more concretely its transition matrices, differs from the stationary $\{\pi_{\mathbf{Q}}, \mathbf{Q}\}$ measure only when one moves inside state 1 in the segments $[M_j, N_{j+1})$. Denote by

$$L_j(x) := \#\left\{k \in [M_{j-1}, N_j) : x_k = 1\right\}$$

and

$$V_j(x) = \#\left\{k \in [M_{j-1}, N_j) : x_k = x_{k+1} = 1\right\}.$$

LEMMA 30. – *For every $\epsilon > 0$, there exists $t_0 \in \mathbb{N}$ s.t for every $t > t_0$, $N_t \leq n < m_t$ and $x \in \{1,2,3\}^{\mathbb{Z}}$,*

$$(T^n)'(x) = (1\pm\epsilon)\prod_{k=1}^{t}\left[\left(\frac{1+\varphi}{1+\varphi\lambda_k}\right)^{L_k\circ T^n(x)-L_k(x)}\cdot\lambda_k^{V_k\circ T^n(x)-V_k(x)}\right].$$

Proof. – Let $\epsilon > 0$, $t\in\mathbb{N}$ and $N_t \leq n < m_t$. Canceling out all the $k's$ such that $P_{k-n} = P_k$ one can see that

$$(T^n)'(x) = I_t \cdot \tilde{I}_t$$

where (notice in the definition of I_t that $n > N_t$)

$$I_t = \prod_{u=1}^{t}\left[\left(\prod_{k=M_{u-1}}^{N_u}\frac{P_{k-n}(x_k,x_{k+1})}{P_k(x_k,x_{k+1})}\right)\cdot\left(\prod_{k=M_{u-1}+n}^{N_u+n-1}\frac{P_{k-n}(x_k,x_{k+1})}{P_k(x_k,x_{k+1})}\right)\right]$$

and (here notice that $M_t > N_t + m_t > N_t + n$)

$$\tilde{I}_t = \prod_{u=t+1}^{\infty}\left[\left(\prod_{k=M_{u-1}}^{M_{u-1}+n-1}\frac{P_{k-n}(x_k,x_{k+1})}{P_k(x_k,x_{k+1})}\right)\cdot\left(\prod_{k=N_u}^{N_u+n-1}\frac{P_{k-n}(x_k,x_{k+1})}{P_k(x_k,x_{k+1})}\right)\right].$$

We will analyze the two terms separately. Since for every $M_{u-1} \leq k < M_{u-1} + n$, $P_k = \mathbf{Q_u}$ and $P_{k-n} = \mathbf{Q}$,

$$\frac{P_{k-n}(x_k,x_{k+1})}{P_k(x_k,x_{k+1})} \leq \frac{\mathbf{Q}_{1,3}}{(\mathbf{Q_u})_{1,3}} = \frac{1+\varphi\lambda_u}{1+\varphi} \leq \lambda_u.$$

Similarly for $N_u \leq k < N_u + n$, $P^{(k)} = \mathbf{Q}$ and $P_{k-n} = \mathbf{Q_u}$. Therefore

$$\frac{P_{k-n}(x_k,x_{k+1})}{P_k(x_k,x_{k+1})} \leq \frac{(\mathbf{Q_u})_{1,1}}{\mathbf{Q}_{1,1}} \leq \lambda_u,$$

and

$$\lambda_u^{-2n} \le \left(\prod_{k=M_{u-1}}^{M_{u-1}+n-1} \frac{P_{k-n}(x_k, x_{k+1})}{P_k(x_k, x_{k+1})}\right) \cdot \left(\prod_{k=N_u}^{N_u+n-1} \frac{P_{k-n}(x_k, x_{k+1})}{P_k(x_k, x_{k+1})}\right) \le \lambda_u^{2n},$$

here the lower bound is achieved by a similar analysis. This gives

$$\begin{aligned}\tilde{I}_t &= \prod_{u=t+1}^{\infty} \left[\lambda_u^{\pm 2n}\right] = \prod_{u=t+1}^{\infty} \left[\lambda_u^{\pm 2m_{u-1}}\right] \quad (\text{since } \forall u > t, n < m_t < m_u) \\ &\overset{(3.1)}{=} e^{\pm \sum_{n=t+1}^{\infty} \frac{1}{2^n}} \xrightarrow[t\to\infty]{} 1.\end{aligned}$$

Consequently there exists $t_0 \in \mathbb{N}$ so that for all $x \in \Sigma_{\mathbf{A}}$, $t > t_0$ and $N_t \le n \le m_t$,

$$(T^n)'(x) = (1 \pm \epsilon) I_t.$$

By noticing that for $k \in \bigcup_{j=1}^{t} \left([M_{j-1}, N_j) \cup [M_{j-1}+n, N_j+n)\right)$,

$$P_{k-n}(x_k, x_{k+1}) \neq P_k(x_k, x_{k+1})$$

if and only if $x_k = 1$, one can check that

$$I_t = \prod_{k=1}^{t} \left[\left(\frac{1+\varphi}{1+\varphi\lambda_k}\right)^{L_k \circ T^n(x) - L_k(x)} \lambda_k^{V_k \circ T^n(x) - V_k(x)}\right]. \qquad \square$$

COROLLARY 31. – *The shift* $(\{1,2,3\}^{\mathbb{Z}}, \mu, T)$ *is conservative and ergodic.*

Proof. – Since the shift is a K-automorphism it is enough to prove conservativity.

For every $j \in \mathbb{N}$, $0 \le L_k(x), D_k(x) \le n_k$. Whence

$$\begin{aligned}\left(\frac{1+\varphi}{1+\varphi\lambda_k}\right)^{L_k \circ T^n(x) - L_k(x)} \lambda_k^{V_k \circ T^n(x) - V_k(x)} &\ge \left(\frac{1+\varphi}{1+\varphi\lambda_k}\right)^{L_k \circ T^n(x)} \lambda_k^{-V_k(x)} \\ &\ge \lambda_k^{-2n_k} \ge \lambda_1^{-2n_k},\end{aligned}$$

and for every $t \in \mathbb{N}$,

$$\prod_{k=1}^{t} \left[\left(\frac{1+\varphi}{1+\varphi\lambda_k}\right)^{L_k \circ T^n(x) - L_k(x)} \lambda_k^{V_k \circ T^n(x) - V_k(x)}\right] \ge \lambda_1^{-2\sum_{k=1}^{t} n_k} \ge \lambda_1^{-2N_t}.$$

By Lemma 30 there exists $t_0 \in \mathbb{N}$ such that for all $t \ge t_0$, $N_t \le n \le m_t$ and $x \in \Sigma_{\mathbf{A}}$, $(T^n)'(x) \ge \frac{\lambda_1^{-2N_t}}{2}$. Therefore for all $x \in \Sigma_{\mathbf{A}}$,

$$\sum_{n=1}^{\infty} (T^n)'(x) \ge \sum_{t=1}^{\infty} \sum_{n=N_t}^{m_t} (T^n)'(x) \ge \sum_{t=t_0}^{\infty} \frac{1}{2}(m_t - N_t)\lambda_1^{-2N_t} \overset{(3.9)}{=} \infty.$$

By Hopfs criteria the shift is conservative. $\square$

A.0.1. *Proof of the type* III$_1$ *property*. – In order to prove that the ratio set is $[0,\infty)$ we are going to use the following principle: since $R(T)$ is a multiplicative subset it is enough to show that there exists $y_n \in R(T)\setminus\{1\}$ with $y_n \to 1$ as $n \to \infty$.

THEOREM 32. – *Let μ be the Markov measure constructed in Subsection* 3.2.2. *For every $n \in \mathbb{N}$, $\lambda_n \cdot \frac{1+\varphi}{1+\varphi\lambda_n} \in R(T)$ and therefore the shift is type* III$_1$.

Fix $n \in \mathbb{N}$. The first stage in proving that $\lambda_n \cdot \frac{1+\varphi}{1+\varphi\lambda_n} \in R(T)$ is to show that the ratio set condition is satisfied for all cylinders with a positive proportion of the measure of the cylinder set. Then for a general $A \in \mathscr{B}_+$, we use the density of cylinder sets in $\mathscr{B}$.

Given $t \in \mathbb{N}$, denote by $\mathcal{C}(t)$ the collection of all $[c]_0^{N_t}$ cylinder sets such that

$$L_t(c) = \sum_{k=M_{t-1}}^{N_t-1} \mathbf{1}_{[c_k=1]} \in \left(\frac{n_t}{4}, \frac{n_t}{2}\right) \text{ and } \sum_{k=M_{t-1}}^{N_t-1} \mathbf{1}_{[c_k=2, c_{k+1}=3]} \geq \frac{n_t}{15}. \tag{A.1}$$

Since

$$\mu\left([c]_{M_{t-1}}^{N_t}\right) = \nu_{\pi_{M_{t-1}}, \mathbf{Q}_t}\left([c]_0^{n_t}\right),$$

it follows from (3.5) and (3.6) that for all t large enough,

$$\mu\left(\bigcup_{C\in\mathcal{C}(t)} C\right) \geq 1 - \frac{1}{2t}.$$

In order to shorten the notation, given $M, j \in \mathbb{N}$, $B \in \mathscr{B}$ and $\epsilon > 0$, let

$$\mathfrak{RSC}(M, B, j, \epsilon) := B \cap T^{-M} B \cap \left[\left(T^M\right)' = \lambda_j \cdot \frac{1+\varphi}{1+\varphi\lambda_j} \cdot (1 \pm \epsilon)\right],$$

and for $M \in \mathbb{N}$,

$$\Sigma_{\mathbf{A}}(M) := \{1,2,3\}^M \cap \Sigma_{\mathbf{A}}.$$

LEMMA 33. – *For every $[b]_{-n}^n$ cylinder set, $\epsilon > 0$ and $j \in \mathbb{N}$, there exists a $t_0 \in \mathbb{N}$ so that for all $t > t_0$ the following holds:*

For every $C = [c]_0^{N_t-1} \in \mathcal{C}(t)$ there exists $d = d(b, C) \in \Sigma_{\mathbf{A}}(N_t + n)$ such that for every $\mathbb{N} \ni l \leq m_t/k_t$,

$$C \cap [d]_{lk_t-n}^{lk_t+N_t-1} \subset T^{-lk_t}[b]_{-n}^n \cap \left[\left(T^{lk_t}\right)' = \lambda_j \cdot \frac{1+\varphi}{1+\varphi\lambda_j} \cdot (1 \pm \epsilon)\right]. \tag{A.2}$$

Recall that $k_t > N_t$ is defined as a $\left(1 \pm 3^{-3N_t}\right)$ mixing time for $\mathbf{Q}$.

Proof. – Let $[b]_{-n}^n, \epsilon > 0$ and $j \in \mathbb{N}$ be given. By Lemma 30 there exists τ such that for every $t \geq \tau$ and $1 \leq l \leq m_t/k_t$ (here $lk_t \in [N_t, m_t)$),

$$\left(T^{lk_t}\right)'(x) = (1 \pm \epsilon) \prod_{k=1}^t \left[\left(\frac{1+\varphi}{1+\varphi\lambda_k}\right)^{L_k \circ T^{lk_t}(x) - L_k(x)} \cdot \lambda_k^{V_k \circ T^{lk_t}(x) - V_k(x)}\right].$$

Choose t_0 to be any integer which satisfies $t_0 > \max(\tau, j)$ and $M_{t_0} > n$.

Let $t > t_0$ and choose a cylinder set $[c]_0^{N_t} \in \mathcal{C}(t)$ which intersects $[b]_{-n}^n$. That is $c_i = b_i$ for $i \in [0,n]$. We need now to choose $d \in \Sigma_{\mathbf{A}}(N_t + n)$ which satisfies (A.2). Notice that for $x \in [d]_{lk_t-n}^{lk_t+N_t} \cap [c]_0^{N_t}$,

$$\prod_{k=1}^{t}\left[\left(\frac{1+\varphi}{1+\varphi\lambda_k}\right)^{L_k\circ T^{lk_t}(x)-L_k(x)}\cdot\lambda_k^{V_k\circ T^{lk_t}(x)-V_k(x)}\right] = \prod_{k=1}^{t}\left[\left(\frac{1+\varphi}{1+\varphi\lambda_k}\right)^{L_k(d)-L_k(c)}\cdot\lambda_k^{V_k(d)-V_k(c)}\right],$$

in this representation we look at $[d]_{-n}^{N_t}$. For all $k \in [0, M_{t-1}]$, let

$$d_k = c_k$$

and for all $k \in [-n, 0)$,

$$d_k = b_k.$$

Notice that this means that for $k \in [-n, n]$, $d_k = b_k$ and thus

$$[d]_{lk_t-n}^{lk_t+N_t} \subset T^{-lk_t}[b]_{-n}^n.$$

Let $p(j,t) \leq \frac{n_t}{20}$ be the integer (condition (3.4)) such that

$$\left(\lambda_t\cdot\frac{1+\varphi}{1+\varphi\lambda_t}\right)^{p(j,t)} = \lambda_j\cdot\frac{1+\varphi}{1+\varphi\lambda_j}.$$

Set $d_k = 1$ for all $k \in [M_{t-1}, M_{t-1} + V_t(c) + p(j,t)]$ and then continue repeatedly with the sequence "321," $L_t(c) - V_t(c)$ times. Since c satisfies (A.1), this construction is well defined (e.g., we have not reached yet $k = N_t - 1$). Continue with sequences of 32 till $k = N_t - 1$.

Thus we have defined d in such a way that

$$L_t(d) - L_t(c) = p(j,t)$$

and

$$V_t(d) - V_t(c) = p(j,t).$$

In addition for all $0 \leq k < t$, $L_k(d) = L_k(c)$ and $V_k(c) = V_k(d)$. Thus for all $x \in [d]_{lk_t-n}^{lk_t+N_t} \cap [c]_0^{N_t}$,

$$\begin{aligned}\left(T^{lk_t}\right)'(x) &= (1\pm\epsilon)\prod_{k=1}^{t}\left[\left(\frac{1+\varphi}{1+\varphi\lambda_k}\right)^{L_k(d)-L_k(c)}\cdot\lambda_k^{V_k(d)-V_k(c)}\right]\\ &= (1\pm\epsilon)\left(\lambda_t\cdot\frac{1+\varphi}{1+\varphi\lambda_t}\right)^{p(j,t)}\\ &= (1\pm\epsilon)\left(\lambda_j\frac{1+\varphi}{1+\varphi\lambda_j}\right).\end{aligned}$$

This proves the lemma. □

In the course of the proof one sees that the event

$$\left([b]_{-n}^n \cap [c]_0^{N_t}\right) \cap \left(T^{-lk_t}[b]_{-n}^n \cap \left\{\left(T^{lk_t}\right)' = \lambda_j\cdot\frac{1+\varphi}{1+\varphi\lambda_j}\cdot(1\pm\epsilon)\right\}\right)^c$$

is $\mathscr{F}\left(lk_t - n, lk_t + N_t\right)$ measurable and does not depend on $\mathscr{F}\left(lk_t + N_t, lk_t + 2N_t\right)$.

REMARK 34. – Given $[c]_0^{N_t} \in \mathcal{C}_t$ we have defined $d = d(c) \in \Sigma_{\mathbf{A}}(N_t + n)$. The definition of d is not necessarily one to one. This is because if $[\tilde{c}]_0^{M_t-1} = [c]_0^{M_t-1}$, $V_t(c) = V_t(\tilde{c})$ and $L_t(c) = L_t(\tilde{c})$ then $d(c) = d(\tilde{c})$. In order to make it one to one we will use

$$[d(c), c]_{lk_t-n}^{lk_t+2N_t}$$

instead of $[d(c)]_{lk_t}^{lk_t+N_t}$ where by $[a,b]_l^{l+\mathrm{length}(a)+\mathrm{length}(b)}$ we mean the concatenation of a and b. This can be thought of as putting a marker on $d(c)$. In order that the concatenation will be in Σ_A we need that

$$\mathbf{Q}\left(d(c)_{N_t-1}, c_0\right) > 0.$$

This can be done by possibly changing the last two coordinates of $d(c)$. This will change the value of $\left(T^{lk_t}\right)'$ by at most a factor of $\lambda_t^{\pm 4}$, which is close enough to one. We will denote by $\mathbf{d}(c) := (\tilde{d}(c), c)$. We still have

$$[\mathbf{d}(c)]_{lk_t-n}^{lk_t+2N_t} \subset T^{-lk_t}[b]_{-n}^n \cap \left[\left(T^{lk_t}\right)' = \lambda_j \cdot \frac{1+\varphi}{1+\varphi\lambda_j} \cdot (1 \pm \epsilon)\right],$$

but now the map $c \mapsto \mathbf{d}(c)$ is one to one.

In the proof of the next lemma we will make use of the fact that for every cylinder set $\left([a]_m^l\right)^c$ is $\mathscr{F}(m,l)$ measurable.

LEMMA 35. – *For every $[b]_{-n}^n$ cylinder set, $\epsilon > 0$ and $j \in \mathbb{N}$ there exists $t_0 \in \mathbb{N}$ such that for all $t > t_0$,*

$$\mu\left(\bigcup_{l=1}^{m_t/4k_t} \mathfrak{RSC}\left(4lk_t, [b]_{-n}^n, j, \epsilon\right)\right) \geq 0.8\mu\left([b]_{-n}^n\right).$$

Proof. – Let $[b]_{-n}^n$ be a cylinder set and t_0 be as in Lemma 33. For all $t \geq t_0$, $[c]_0^{N_t} \in \mathcal{C}(t)$ which intersects $[b]_{-n}^n$ and $1 \leq l \leq m_t/4k_t$,

$$\left([c]_0^{N_t} \cap [b]_{-n}^n\right) \cap \left(\mathfrak{RSC}\left(4lk_t, [b]_{-n}^n, j, \epsilon\right)\right)^c \subset [c]_0^{N_T} \cap [b]_{-n}^n \cap \left([\mathbf{d}(c)]_{4lk_t-n}^{4lk_t+N_t}\right)^c.$$

As $\mathbf{Q}_{1,3} = \min\left\{\mathbf{Q}_{i,j} : 1 \leq i, j \leq 3\right\}$,

$$\nu_{\pi_{\mathbf{Q}},\mathbf{Q}}\left([d(c)]_{4lk_t-n}^{4lk_t+N_t}\right) \geq \pi_{\mathbf{Q}}(3)\,\mathbf{Q}_{1,3}^{2N_t+n-1} \gtrsim \frac{1}{3^{3N_t}}.$$

Therefore, one has by repeated applications of (3.7) (mixing time condition),

$$\begin{aligned}
&\mu\left(\left([b]_{-n}^{n}\cap[c]_{0}^{N_t}\right)\cap\left(\bigcup_{l=1}^{m_t/4k_t}\mathfrak{RSC}\left(lk_t,[b]_{-n}^{n},j,\epsilon\right)\right)^{c}\right)\\
&\quad\leq\mu\left(\left([b]_{-n}^{n}\cap[c]_{0}^{N_t}\right)\cap\left\{\bigcap_{l=1}^{m_t/4k_t}\left([d_c]_{4lk_t-n}^{4lk_t+N_t}\right)^{c}\right\}\right)\\
&\quad\leq\mu\left(\left([b]_{-n}^{n}\cap[c]_{0}^{N_t}\right)\right)\prod_{1=1}^{m_t/4k_t}\left[\left(1+3^{-3N_t}\right)\left(1-\nu_{\pi_{\mathbf{Q}},\mathbf{Q}}\left([d(c)]_{4lk_t}^{4lk_t+N_t}\right)\right)\right]\\
&\quad\leq\mu\left(\left([b]_{-n}^{n}\cap[c]_{0}^{N_t}\right)\right)\left[\left(1+3^{-3N_t}\right)\left(1-3^{-3N_t}\right)\right]^{m_t/4k_t}\\
&\quad\overset{(3.8)}{\leq}\frac{1}{t}\mu\left(\left([b]_{-n}^{n}\cap[c]_{0}^{N_t}\right)\right).
\end{aligned}$$

Notice that in the application of the mixing time condition we used that

$$(4\,(l+1)\,k_t-n)-(4lk_t+N_t)>(4l+3)k_t-(4l+1)\,k_t=2k_t.$$

If t is large enough then

$$\mu\left(\Sigma_{\mathbf{A}}\setminus\bigcup_{C\in\mathcal{C}(t)}C\right)<0.1\mu\left([b]_{-n}^{n}\right),$$

and for all $[c]_0^{N_t}=C\in\mathcal{C}(t)$,

$$\begin{aligned}
\mu\left([b]_{-n}^{n}\cap[c]_{0}^{N_t}\cap\left(\bigcup_{l=1}^{m_t/4k_t}\mathfrak{RSC}\left(lk_t,[b]_{-n}^{n},j,\epsilon\right)\right)\right)&>\left(1-\frac{1}{t}\right)\mu\left([b]_{-n}^{n}\cap[c]_{0}^{N_t}\right)\\
&\geq 0.9\mu\left([b]_{-n}^{n}\cap[c]_{0}^{N_t}\right).
\end{aligned}$$

The lemma follows from

$$\begin{aligned}
&\mu\left(\bigcup_{l=1}^{m_t/4k_t}\mathfrak{RSC}\left(lk_t,[b]_{-n}^{n},j,\epsilon\right)\right)\\
&\quad\geq\mu\left(\biguplus_{[c]_0^{N_t}\in\mathcal{C}(t)}[b]_{-n}^{n}\cap[c]_{0}^{N_t}\cap\bigcup_{l=1}^{m_t/4k_t}\mathfrak{RSC}\left(lk_t,[b]_{-n}^{n},j,\epsilon\right)\right)\\
&\quad\geq 0.9\sum_{[c]_0^{NT}\in\mathcal{C}(t)}\mu\left([b]_{-n}^{n}\cap[c]_{0}^{N_t}\right)\\
&\quad\geq 0.8\mu\left([b]_{-n}^{n}\right).
\end{aligned}$$

□

Proof of Theorem 32. – This is a standard approximation technique. Let $j\in\mathbb{N}$, $A\in\mathcal{B}$, $\mu(A)>0$ and $\epsilon>0$. Since the ratio set condition on the derivative is monotone with respect to ϵ and

$$1<\lambda_j\cdot\frac{1+\varphi}{1+\varphi\lambda_j}<2,$$

we can assume that

$$1 \le \lambda_j \cdot \frac{1+\varphi}{1+\varphi\lambda_j}(1 \pm \epsilon) \le 2. \tag{A.3}$$

Since $\mathcal{F}(-n,n) \uparrow \mathcal{B}$ as $n \to \infty$, there exists a cylinder set $\mathfrak{b} = [b]_{-n}^{n}$ such that

$$\mu\left(A \cap \mathfrak{b}\right) > 0.99\mu\left(\mathfrak{b}\right).$$

By Lemma 35 there exists $t \in \mathbb{N}$ for which

$$\mu\left(\mathfrak{b} \cap \left\{\bigcup_{l=1}^{m_t/4k_t} T^{-4lk_t}\mathfrak{b} \cap \left[\left(T^{4lk_t}\right)' = \lambda_j \cdot \frac{1+\varphi}{1+\varphi\lambda_j}\cdot(1\pm\epsilon)\right]\right\}\right) > 0.8\mu(\mathfrak{b}).$$

Denote by

$$B = \mathfrak{b} \cap \left\{\bigcup_{l=1}^{m_t/4k_t} T^{-4lk_t}\mathfrak{b} \cap \left[\left(T^{4lk_t}\right)' = \lambda_j \cdot \frac{1+\varphi}{1+\varphi\lambda_j}\cdot(1\pm\epsilon)\right]\right\}.$$

We can assume that for $x \in B$, there exists $C(x) = [c]_0^{N_t} \in \mathcal{C}_t$ so that $x \in C(x)$. Then by the proof of Lemma 33 there exists $d(C(x)) \in \Sigma_{\mathbf{A}}(2N_t + n)$ such that if $x \in [d(C(x))]_{4lk_t-n}^{4lk_t+2N_t}$, then

$$\left(T^{4lk_t}\right)'(x) = \lambda_j \cdot \frac{1+\varphi}{1+\varphi\lambda_j}\cdot(1\pm\epsilon) \text{ and } x \in T^{-4lk_t}\mathfrak{b}. \tag{A.4}$$

Define $\phi : B \to \mathbb{N}$

$$\phi(x) := \inf\left\{l \le m_t/4k_t : [x]_{4lk_t-n}^{4lk_t+2N_t} = [\mathbf{d}(C(x))]_{4lk_t-n}^{4lk_t+2N_t}\right\}$$

and $S = T^{\phi} : B \to S(B) \subset \mathfrak{b}$. We claim that S is one to one. Indeed, since the map $[c]_0^{N_t} \mapsto \mathbf{d}(c)$ is one to one, for every $x, y \in B$ such that $C(x) \neq C(y)$,

$$[Sy]_{-n}^{2N_t} = [\mathbf{d}(C(y))]_{-n}^{2N_t} \neq [\mathbf{d}(C(x))]_{-n}^{2N_t} = [Sx]_{-n}^{2N_t},$$

consequently $Sx \neq Sy$. In addition, by the definition of ϕ, if $x \neq y$ and $C(x) = C(y)$ then $Sx \neq Sy$.

It follows from (A.4) and (A.3), that for all $x \in B$,

$$S'(x) := \frac{d\mu \circ S}{d\mu}(x) = \lambda_j \cdot \frac{1+\varphi}{1+\varphi\lambda_j}\cdot(1\pm\epsilon) \in [1,2].$$

Therefore $\frac{d\mu\circ S^{-1}}{d\mu}(y) \ge \frac{1}{2}$ for all $y \in S(B)$. A calculation shows that

$$\begin{aligned}\mu(S(B)\cap A) &> \mu(S(B)) - \mu(\mathfrak{b}\backslash A)\\ &> \mu(B) - \mu(\mathfrak{b}\backslash A)\\ &= 0.79\mu\left(\mathfrak{b}\right),\end{aligned}$$

and

$$\mu\left(S^{-1}\left(S(B)\cap A\right)\right) > \frac{\mu\left(S(B)\cap A\right)}{2} > 0.39\mu(\mathfrak{b}).$$

So

$$\sum_{l=1}^{m_t/4lk_t} \mu\left(A \cap \left\{T^{-4lk_t} A \cap \left[\left(T^{4lk_t}\right)' = \lambda_j \cdot \frac{1+\varphi}{1+\varphi\lambda_j} \cdot (1 \pm \epsilon)\right]\right\} \cap [\phi = 4lk_t]\right)$$

$$\geq \mu\left((B \cap A) \cap S^{-1}\left(S(B) \cap A\right)\right) \quad \{\text{Notice that } B,\ S(B) \subset \mathfrak{b}\}$$

$$\geq \mu\left(B \cap A\right) - \mu\left(\mathfrak{b} \backslash S^{-1}\left(S(B) \cap A\right)\right)$$

$$\geq 0.18\mu\left(\mathfrak{b}\right),$$

and thus there exists $l \in \mathbb{N}$ such that

$$\mu\left(A \cap T^{-4lk_t} A \cap \left[\left(T^{4lk_t}\right)' = \lambda_j \cdot \frac{1+\varphi}{1+\varphi\lambda_j} \cdot (1 \pm \epsilon)\right]\right) > 0.$$

This proves the theorem. □

Acknowledgement

Section 3 is part of the author's PhD thesis in Tel Aviv University done under the supervision of Jon Aaronson. I would like to thank many people among them Jon Aaronson, Omri Sarig and Ian Melbourne for many helpful discussions and their continuous support. This research was supported in part by the European Advanced Grant StochExtHomog (ERC AdG 320977).

BIBLIOGRAPHY

[1] J. Aaronson, *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs **50**, Amer. Math. Soc., 1997.

[2] R. L. Adler, Symbolic dynamics and Markov partitions, *Bull. Amer. Math. Soc. (N.S.)* **35** (1998), 1–56.

[3] R. L. Adler, B. Weiss, *Similarity of automorphisms of the torus*, Memoirs of the American Mathematical Society **98**, Amer. Math. Soc., 1970.

[4] A. Avila, J. Bochi, A generic C^1 map has no absolutely continuous invariant probability measure, *Nonlinearity* **19** (2006), 2717–2725.

[5] C. Bonatti, S. Crovisier, A. Wilkinson, The C^1 generic diffeomorphism has trivial centralizer, *Publ. Math. Inst. Hautes Études Sci.* **109** (2009), 185–244.

[6] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, revised ed., Lecture Notes in Math. **470**, Springer, 2008.

[7] H. Bruin, J. Hawkins, Examples of expanding C^1 maps having no σ-finite invariant measure equivalent to Lebesgue, *Israel J. Math.* **108** (1998), 83–107.

[8] J. T. Campbell, A. N. Quas, A generic C^1 expanding map has a singular S-R-B measure, *Comm. Math. Phys.* **221** (2001), 335–349.

[9] B. M. Gurevič, V. I. Oseledec, Gibbs distributions, and the dissipativity of C-diffeomorphisms, *Dokl. Akad. Nauk SSSR* **209** (1973), 1021–1023.

[10] J. M. Kabanov, R. v. Lipcer, A. N. Širjaev, On the question of the absolute continuity and singularity of probability measures, *Mat. Sb. (N.S.)* **104(146)** (1977), 227–247, 335.

[11] Z. Kosloff, On a type III_1 Bernoulli shift, *Ergodic Theory Dynam. Systems* **31** (2011), 1727–1743.

[12] Z. Kosloff, On manifolds admitting stable type III_1 Anosov diffeomorphisms, *J. Mod. Dyn.* **13** (2018), 251–270.

[13] W. Krieger, On non-singular transformations of a measure space. I, II, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **11** (1969), 11.

[14] K. Krzyżewski, On expanding mappings, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **19** (1971), 23–24.

[15] D. A. Levin, Y. Peres, E. L. Wilmer, *Markov chains and mixing times*, Amer. Math. Soc., 2009.

[16] A. N. Livšic, J. G. Sinaĭ, Invariant measures that are compatible with smoothness for transitive C-systems, *Dokl. Akad. Nauk SSSR* **207** (1972), 1039–1041.

[17] C. Robinson, L. S. Young, Nonabsolutely continuous foliations for an Anosov diffeomorphism, *Invent. math.* **61** (1980), 159–176.

[18] A. N. Shiryayev, *Probability*, Graduate Texts in Math. **95**, Springer, 1984.

[19] C. E. Silva, P. Thieullen, A skew product entropy for nonsingular transformations, *J. London Math. Soc.* **52** (1995), 497–516.

[20] J. G. Sinaĭ, Markov partitions and y-diffeomorphisms, *Funkcional. Anal. i Priložen* **2** (1968), 64–89.

[21] M. Viana, *Lectures on Lyapunov exponents*, Cambridge Studies in Advanced Math. **145**, Cambridge Univ. Press, 2014.

[22] M. Viana, (Dis)continuity of Lyapunov exponents, *Ergodic Theory Dynam. Systems* **40** (2020), 577–611.

(Manuscrit reçu le 19 août 2017 ;
accepté le 11 juin 2019.)

Zemer Kosloff
Mathematics institute
University of Warwick
Gibbet Hill Road, Coventry, CV47AL, UK
Einstein institute of Mathematics
Hebrew University of Jerusalem
Kiryat Safra campus, Jerusalem
E-mail: `zemer.kosloff@mail.huji.ac.il`

Ann. Scient. Éc. Norm. Sup.
4ᵉ série, t. 54, 2021, p. 133 à 233

ANALYTIC NORMAL FORMS AND INVERSE PROBLEMS FOR UNFOLDINGS OF 2-DIMENSIONAL SADDLE-NODES WITH ANALYTIC CENTER MANIFOLD

BY C. ROUSSEAU AND L. TEYSSIER

ABSTRACT. – We give normal forms for generic k-dimensional parametric families $(Z_\varepsilon)_\varepsilon$ of germs of holomorphic vector fields near $0 \in \mathbb{C}^2$ unfolding a saddle-node singularity Z_0, under the condition that there exists a family of invariant analytic curves unfolding the weak separatrix of Z_0. These normal forms provide a moduli space for these parametric families. In our former 2008 paper, a modulus of a family was given as the unfolding of the Martinet-Ramis modulus, but the realization part was missing. We solve the realization problem in that partial case and show the equivalence between the two presentations of the moduli space. Finally, we completely characterize the families which have a modulus depending analytically on the parameter. We provide an application of the result in the field of non-linear, parameterized differential Galois theory.

RÉSUMÉ. – Nous donnons des formes normales pour les familles génériques $(Z_\varepsilon)_\varepsilon$ à k paramètres de germes de champs de vecteurs holomorphes au voisinage de $0 \in \mathbb{C}^2$, et déployant une singularité Z_0 de type col-nœud, sous la condition qu'il existe une famille de courbes analytiques invariantes déployant la séparatrice faible de Z_0. Ces formes normales donnent un espace de modules pour ces familles génériques. Dans notre article de 2008, nous avions donné un module de classification pour ces familles génériques, lequel consistait en un déploiement du module de Martinet-Ramis, mais la partie réalisation était manquante. Dans cet article, nous donnons la réalisation dans ce cas spécial, et nous montrons l'équivalence entre les deux présentations de l'espace des modules. Finalement, nous caractérisons complètement les familles dont le module dépend analytiquement des paramètres. Nous donnons une application du résultat en théorie de Galois paramétrique non linéaire.

The first author was supported by NSERC in Canada. The second author was supported by the French National Research Agency grant ANR-13-JS01-0002-01. He would also like to express his gratitude for the welcome and support he received during his stay at the CNRS unit UMI-3457 located in the CRM / Université de Montréal.

 doi:10.24033/asens.2457

1. Introduction

Heuristically, moduli spaces of holomorphic dynamical systems not only encode but also describe qualitatively the dynamics itself, and to some extent allow a better understanding of remarkable dynamical phenomena. This paper is part of a large program aimed at studying the conjugacy classes of dynamical systems in the neighborhood of stationary points (up to local changes of analytic coordinates). Stationary points and their invariant manifolds organize the global dynamics while degenerate stationary points organize the bifurcation diagrams in families of dynamical systems. Stationary points of discrete dynamical systems correspond to fixed-points of the iterated map(s), while for continuous dynamical systems they correspond to singularities in the underlying differential equation(s).

A natural tool for studying conjugacy classes is the use of normal forms. For hyperbolic stationary points (generic situation), the system is locally conjugate to its linear part so that the quotient space of (local) hyperbolic systems is given by the space of linear dynamical systems. However, for most non-hyperbolic stationary points the normalizing change of coordinates (sending *formally* the system to a normal form) is given by a divergent power series. Divergence is very instructive: it tells us that the dynamics of the original system and that of the normal form are qualitatively different. In that respect, a subclass of singularities that has been thoroughly studied in the beginning of the 80's is that of 1-resonant singularities: these include parabolic fixed-points of germs of 1-dimensional diffeomorphisms, resonant-saddle singularities and saddle-node singularities of 2-dimensional vector fields, as well as non-resonant irregular singular points of linear differential systems. These various resonant dynamical systems share a lot of common properties, among which is the finite-determinacy of their formal normal forms (e.g., polynomial expressions in the case of vector fields). Another property they share is that they can be understood as the *coalescence* of special "geometric objects," either of stationary points or of a singular point with a limit cycle in the case of the Hopf bifurcation at a weak focus.

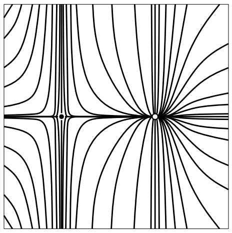
1.1. Scope of the paper

The present work is the follow-up of [41] in which we described a set of functional moduli for unfoldings of codimension k saddle-node vector fields $Z = (Z_\varepsilon)_\varepsilon$ depending on a finite-dimensional parameter $\varepsilon \in (\mathbb{C}^k, 0)$. Here we focus mainly on the inverse problem and on the question of finding (almost unique) normal forms, as we explain below.

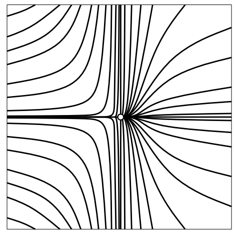
The most basic example of such an unfolding is given by the codimension 1 unfolding (expressed in the canonical basis of $\mathbb{C}^2$)

$$Z_\varepsilon(x, y) := \begin{bmatrix} x^2 + \varepsilon \\ y \end{bmatrix}, \quad \varepsilon \in \mathbb{C}. \tag{1.1}$$

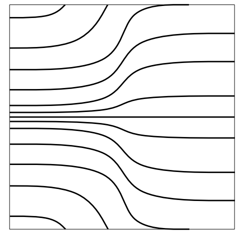
Real slices of the phase-portraits are shown in Figure 1.1. The merging (bifurcation) occurs at $\varepsilon = 0$: for $\varepsilon \neq 0$ the system has two stationary points located at $(\pm\sqrt{-\varepsilon}, 0)$ which collide as ε reaches 0.



(A)
$\varepsilon <$
0



(B)
$\varepsilon =$
0



(C)
$\varepsilon >$
0

FIGURE 1.1. Typical members of the simplest saddle-node bifurcation.

1.2. Modulus of classification

Each merging stationary point organizes the dynamics in its own neighborhood in a rigid way. The local models of these rigid dynamics seldom agree on overlapping areas and in general cannot be glued together. If this incompatibility persists as the confluence happens, then we have divergence of the normalizing series at the limit. In the case of 1- or 2-dimensional resonant systems the normalizing series is k-summable. The divergence is then quantified by the *Stokes phenomenon*: there exists a formal normalizing transformation, and a covering of a punctured neighborhood of the singularity by $2k$ sectors over which there exist unique sectorial normalizing transformations that are Gevrey-asymptotic to the formal normalization. Comparing the normalizing transformations on intersections of consecutive sectors provides a modulus of analytic classification. This modulus takes the form of Stokes matrices for irregular singularities of linear differential systems and functional moduli for singularities of nonlinear dynamical systems (see for instance [20]).

The classification of resonant systems may seem rather mysterious. But if we remember that we are studying the merging of "simple" singularities, then it becomes natural to unfold the situation and study the "multiple" singularity as a limiting case. Indeed, analyzing unfoldings sheds a new light on the "complicated" dynamics of the limiting systems. The idea was suggested by several mathematicians, including V. Arnold, A. Bolibruch and J. Martinet [30]. It was put in practice for unfoldings of saddle-node singularities by A. Glutsyuk [15] on regions in parameter space over which the confluent singularities are all hyperbolic. The system can be linearized in the neighborhood of each singularity, and the mismatch in the normalizing changes of coordinates tends to the components of the saddle-node's Martinet-Ramis modulus [31] when the singularities merge. But the tools were still missing for a full classification of unfoldings of multiple singularities, in particular on a full neighborhood in parameter space of the bifurcation value.

The thesis of P. Lavaurs [24] on parabolic points of diffeomorphisms opened the way for such classifications, for he studied the complementary regions in parameter space. The first classification of generic unfoldings of codimension 1 fixed-point of diffeomorphisms regarded the parabolic point [29], and then the resonant-saddle and saddle-node singularities of differential equations [37, 38]. The first classification of generic unfoldings of codimension k saddle-nodes was done by the authors [41] using the visionary ideas of

A. Douady, J.F. Estrada and P. Sentenac [11, 2] that R. Oudkerk had used on some regions in parameter space in his thesis [34]. Then followed classifications of generic unfoldings of codimension k parabolic points [40] and of non-resonant irregular singular points of Poincaré rank k differential systems [18].

In the spirit of this general context we obtained in [41] a (family of) functional data

$$(\mathfrak{m}_\varepsilon)_\varepsilon = \left(f_\varepsilon^{\mathrm{s}}, \psi_\varepsilon^{\mathrm{s}}, \psi_\varepsilon^{\mathrm{n}}\right)_\varepsilon .$$

For $\varepsilon = 0$ this data coincides with the saddle-node's modulus [31, 49, 44]. Although the original work of J. Martinet and J.-P. Ramis already covered parametric cases, it was then assumed that the (formal) type of the singularity remained constant. On the contrary we were interested in bifurcations, which are deformations where the additional parameters change the type (or number) of singularities. Our main contribution was to reconcile Glutsyuk's and Lavaurs's viewpoint and devise a uniform framework valid for a complete neighborhood of the bifurcation value of the parameter. That being said, the very nature of our geometric construction prevented the modulus to be continuous on the whole parameter space. This space needs to be split into a finite number of *cells* whose closures cover a neighborhood of the bifurcation value, on which the modulus is analytic on ε with continuous extension to the closure.

1.3. The inverse (or realization) problem

At the time of the first works on the question, identifying the moduli space was still out of reach. Performing this identification is called the inverse problem. It was first solved for codimension 1 parabolic fixed-points and resonant-saddle singularities [9, 39], as well as for the irregular singularities of linear differential systems with Poincaré rank 1 [23]. For codimension k the realization problem was first solved for unfoldings of non-resonant irregular singular points of Poincaré rank k [19]. But the realization question is still open for unfoldings of codimension k parabolic points.

Let us formulate the inverse problem in the case at hands.

INVERSE PROBLEM. – *Among all elements of the vector space $\mathcal{M}$ to which $\mathfrak{m} = (\mathfrak{m}_\varepsilon)_\varepsilon$ belongs, to identify those coming as moduli of a saddle-node bifurcation.*

The present paper answers completely this challenge in the case of bifurcations with a persistent analytic center manifold. The common feature to that case and the one studied in [19] is that solving the inverse problem ultimately provides unique normal forms (privileged representative in each analytic class).

Having persistent analytic center manifold can be read in the modulus as the condition $\psi^{\mathrm{n}} = \mathrm{Id}$. Although any element of the specialization of $\mathcal{M}$ at $\varepsilon = 0$ can be realized as the modulus of a saddle-node vector field [31, 44], this property does not hold anymore for bifurcations: the typical element of $\mathcal{M} \cap \{\psi^{\mathrm{n}} = \mathrm{Id}\}$ can never be realized as a modulus of saddle-node bifurcation. Let us explain how this is so. It is rather easy to get convinced that there is no obstruction to realize any given deformation $(\mathfrak{m}_\varepsilon)_{\varepsilon \in \mathrm{cl}(\mathcal{E})}$ of a saddle-node's modulus $\mathfrak{m}_0$ over any given cell $\mathcal{E}$ in parameter space. By this we mean that for each fixed $\varepsilon \in \mathrm{cl}(\mathcal{E})$ it is possible to find a holomorphic vector field Z_ε on a neighborhood $\mathcal{U}$ of $(0, 0)$

such that comparisons between its sectorial normalizing maps coincide with $\mathfrak{m}_\varepsilon$. Furthermore the dependence $\varepsilon \mapsto Z_\varepsilon$ has the expected regularity on the cell's closure, and the neighborhood $\mathcal{U}$ is independent on ε. The sole obstacle lies therefore in gluing these cellular realizations together over cellular intersections in order to obtain a genuine analytic parametric family Z whose modulus agrees with $\mathfrak{m}$. Favorable situations can be characterized by a strong criterion imposed on $\mathfrak{m}$, called *compatibility condition*. A necessary and sufficient condition is that two realizations over different cells in parameter space be conjugate over the intersection of the two cells, thus allowing correction to a uniform family. One difficulty is to express this condition on the abstractly encoded dynamics $\mathfrak{m}$ (that is, before performing the cellular realization). The compatibility condition takes the simple form that the abstract holonomy pseudogroups generated by $\mathfrak{m}$ be conjugate, a condition which can easily be expressed in terms of the modulus. The general case of a bifurcation without analytic center manifold remains open, and we hope to address it in the near future.

1.4. Summary of the paper's content

Here we review the content of the present work. For precise statements of our main results, as for more detailed proof techniques, we refer to Section 2. Recall that one can associate two dynamical data to a vector field $X = A\frac{\partial}{\partial x} + B\frac{\partial}{\partial y}$:

— the trajectories of X parametrized by the complex time in the associated flow-system

$$\begin{cases} \dot{x} &= A\left(x, y\right), \\ \dot{y} &= B\left(x, y\right); \end{cases}$$

— the underlying foliation $\mathcal{F}_X$ whose leaves coincide with orbits of X, obtained by forgetting about a particular parametrization of the trajectories. The foliation really is attached to the underlying non-autonomous differential equation

$$A\left(x, y\right) y' = B\left(x, y\right)$$

rather than to the vector field itself.

The action of (analytic or formal) changes of variables Ψ on vector fields X by conjugacy is obtained as the pullback

$$\Psi^* X := \mathrm{D}\Psi^{-1}\left(X \circ \Psi\right).$$

The vector fields X and $\Psi^* X$ are then (analytically or formally) *conjugate*. When two foliations $\mathcal{F}_X$ and $\mathcal{F}_{\widetilde{X}}$ are conjugate (when X is conjugate to a scaling of $\widetilde{X}$ by a non-vanishing function) it is common to say that X and $\widetilde{X}$ are *orbitally equivalent*. While for unfoldings we also allow parameter changes, we restrict our study to parameter / coordinates changes of the form

$$\Psi : \left(\varepsilon, x, y\right) \longmapsto \left(\phi\left(\varepsilon\right), \Psi_\varepsilon\left(x, y\right)\right).$$

In this paper we focus on families $Z = \left(Z_\varepsilon\right)_{\varepsilon \in (\mathbb{C}^k, 0)}$ unfolding a codimension k saddle-node singularity for $\varepsilon = 0$ and the study of their conjugacy class (*resp.* orbital equivalence class) under local analytic changes of variables and parameter (*resp.* and scaling by non-vanishing functions). Such families can always be brought by a formal change of variables and parameter into the formal normal form [(1)]

$$u_\varepsilon(x)\left(P_\varepsilon(x)\frac{\partial}{\partial x}+y\left(1+\mu_\varepsilon x^k\right)\frac{\partial}{\partial y}\right),$$

where

$$\begin{aligned}P_\varepsilon(x)&=x^{k+1}+\varepsilon_{k-1}x^{k-1}+\cdots+\varepsilon_1 x+\varepsilon_0,\quad k\in\mathbb{N}\\ u_\varepsilon(x)&=u_{0,\varepsilon}+u_{1,\varepsilon}x+\cdots+u_{k,\varepsilon}x^k,\quad u_{0,\varepsilon}\neq 0\end{aligned}$$

and $\varepsilon\mapsto(\mu_\varepsilon,u_{0,\varepsilon},\ldots,u_{k,\varepsilon})$ is holomorphic near 0. A proof of this widely accepted result seems to be missing in the literature, hence we provide one.

The first step in our previous work [41] consisted in preparing the unfolding $(Z_\varepsilon)_\varepsilon$ by bringing it in a form where the polynomial P_ε determines the $\frac{\partial}{\partial x}$-component. Formal and analytic equivalences between such forms must consequently preserve the coefficients of P_ε, which then become privileged *canonical parameters*. This process eliminates the difficulty of dealing with changes of parameters and allows to work for fixed values of ε. Then we established a complete classification. The modulus was composed of two parts: the *formal part* given by the formal normal form above, and the *analytic part* given by an unfolding of the saddle-node's functional modulus. The formal / analytic part of the modulus itself consists in the Martinet-Ramis orbital part (characterizing the vector field up to orbital equivalence) and an additional part classifying the time. For example μ_ε is the formal orbital class while u_ε is the formal temporal class.

We completely solve the realization problem for orbital equivalence (i.e., for foliations) when each Z_ε admits a single analytic invariant manifold passing through every singularity. But we do more: we provide almost unique "normal forms" (the only degree of freedom being linear transformations in y), which are polynomial in x when $\mu_0\notin\mathbb{R}_{\leq 0}$. In that generic situation, an unfolding is orbitally equivalent to an unfolding over $\mathbb{P}_1(\mathbb{C})\times(\mathbb{C},0)$ of the form

$$P_\varepsilon(x)\frac{\partial}{\partial x}+y\left(1+\mu_\varepsilon x^k+\sum_{j=1}^{k}x^j R_{j,\varepsilon}(y)\right)\frac{\partial}{\partial y},$$

where the R_j are analytic in both the geometric variable y and the parameter ε. In this generic case the construction is a direct generalization of that of F. Loray's [26, Theorems 2 and 4] for $\varepsilon=0$ and $k=1$, and only involves tools borrowed from complex geometry. In the non-generic case (when $\mu_0\leq 0$) we also provide almost unique "normal forms," which are in some sense global in x: in this case the foliation is defined on a fiber bundle of negative degree $-\tau(k+1)<\mu_0$ for some positive τ over $\mathbb{P}_1(\mathbb{C})$ and is induced by vector fields of the form

$$X_\varepsilon(x,y):=P_\varepsilon(x)\frac{\partial}{\partial x}+y\left(1+\mu_\varepsilon x^k+\sum_{j=1}^{k}x^j R_{j,\varepsilon}\left(P_\varepsilon^\tau(x)y\right)\right)\frac{\partial}{\partial y}. \tag{1.2}$$

(1) As is customary we write vector fields in the form of derivations, by identifying the canonical basis of $\mathbb{C}^2$ with $\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y}\right)$.

This result offers a new presentation of the moduli space which has the advantage over that of [41] to be made up of functions analytic in the parameter (it does not require the splitting of the parameter space into cells).

As far as normal forms are concerned, we provide some also for the family of vector fields. This requires normalizing the "temporal part". The method used is an unfolding of the construction of R. Schäfke and L. Teyssier [43] performed for $\varepsilon = 0$. As a by-product we provide an explicit section of the cokernel of the derivation X_ε (i.e., a linear complement of the image of X_ε acting as a Lie derivative on the space of analytic germs).

An important observation is that the normalization we just described does not involve classification moduli in any way (nor does it rely on the analytical classification for that matter), at least in the generic case $\mu_0 \notin \mathbb{R}_{\leq 0}$. Therefore it does not answer the inverse (or realization) problem which is posed in terms of classification moduli. This leads us to discuss the compatibility condition.

As we mentioned earlier we can realize any unfolding $\mathfrak{m} = (\mathfrak{m}_\varepsilon)_{\varepsilon \in \mathrm{cl}(\mathcal{E})}$ of a saddle-node's modulus $\mathfrak{m}_0$ over a cell $\mathcal{E}$ in parameter space, but we have such control of the construction that we can guarantee this realization is an unfolding in normal form (1.2), save for the fact that the functions $\varepsilon \mapsto R_{j,\varepsilon}$ are merely analytic on $\mathcal{E}$ with continuous extension to the closure. It is possible to express the holonomy group of X_ε with respect to the analytic center manifold (the geometrical dynamics) as a representation of an abstract group of words formed with elements of the modulus $\mathfrak{m}$ (acting in orbits space). The compatibility condition simply states that the holonomy pseudogroups over the intersection of two neighboring cells are conjugate by a tangent-to-identity mapping. If the condition is satisfied then two cellular realizations are conjugate for values of the parameter in the cells' intersection. Usually when such a situation occurs, we need to apply a conjugacy to the vector fields so that they match in the new coordinates. Here no need for it. Indeed, since the realizations over the different cells are in normal form, they necessarily are conjugate by a linear map. The additional hypothesis in the compatibility condition that the conjugating map is tangent-to-identity allows to conclude that the cellular unfoldings actually agree and therefore define a genuine unfolding analytic in $\varepsilon \in \left(\mathbb{C}^k, 0\right)$.

Our analysis presents in an effective way the relationship between Rousseau-Teyssier classification moduli and the coefficients of the normal forms, so that numerical, and in some cases symbolic, computations can be performed. Also, we have refined our understanding of the modulus compared to the presentation in [41]. The number of cells is now the optimal number $C_k = \frac{1}{k+1}\binom{2k}{k}$ (the k-th Catalan's number) given by the Douady-Estreda-Sentenac classification [11, 10]. Moreover we have reduced the degrees of freedom: instead of having the modulus given up to conjugacy by linear functions depending both on ε and the cell, now the modulus is given up to conjugacy by linear functions depending only on ε in an analytic way. This new equivalence relation in the presentation of the modulus was essential in getting the realizations over the different cells to match when the compatibility condition is satisfied.

Last but not least we were able to completely characterize the moduli that depend analytically on the parameter. These only occur when $k = 1$ and their normal forms are given by particular Bernoulli unfoldings (Definition 2.11)

$$P_\varepsilon(x)\frac{\partial}{\partial x} + y\left(1 + \mu_\varepsilon x^k + x r_\varepsilon(x)\left(P_\varepsilon(x)^\tau y\right)^d\right)\frac{\partial}{\partial y} \tag{1.3}$$

with $d \in \mathbb{N}$ and $d\mu \in \mathbb{Z}$ (in particular μ must be a rational constant, which is seldom the case). This proves that the compatibility condition is not trivially satisfied by every element of $\mathcal{M} \cap \{\psi^{\mathrm{n}} = \mathrm{Id}\}$. On the contrary, the typical situation is that of moduli which are analytic and bounded only on single cells. This reminds us of the setting of Borel-summable divergent power series, in particular in the case $k = 1$ where the cells are actual sectors and it can be proved that the moduli are sectorial sums of $\frac{1}{2}$-summable power series (as in [9]). When $k > 1$ the lack of a theory of summation in more than one variable prevents us from reaching similar conclusions, although the moduli are natural candidates for such sums and a general summation theory should probably contain the case we studied here. We reserve such considerations for future works, perhaps using the theory of polynomial summability recently introduced by J. Mozo and R. Schäfke [33, 5].

1.5. Applications

Our main results can be used to solve problems outside the scope of finding normal forms or addressing the *local* inverse problem. Let us mention two applications, the second of which we develop in Section 2.3.

The first (and most straightforward) one concerns the global inverse problem, also known as non-linear Riemann-Hilbert problem, posed by Y. Ilyashenko and S. Yakovenko in [20, Chapter IV]. Being given a (germ of a) complex surface $\mathcal{M}$ seen as the total space of a fiber bundle over a divisor $\mathbb{P}_1(\mathbb{C}) \subset \mathcal{M}$, the problem is to characterize the holonomy representations of complex foliations on $\mathcal{M}$ tangent to (and regular outside) the divisor and transverse to the fibers, except over $k+2$ singularities (which are all assumed non-degenerate) where the fibers are invariant by the foliation. Using a sibling of Loray's technique, they solve it for fiber bundles of degree 0 and -1, although they only provide details for the former case. Our results open the way to generalizations in several directions:

- allowing saddle-node(s) with central manifold along the divisor and adding to the holonomy representation the components of the modulus of the saddle-nodes, similarly to the generalized linear Riemann-Hilbert problem when irregular singularities are allowed;
- allowing foliations depending analytically on the parameter;
- considering realizations on fiber bundles of negative degree: we obtain here realizations on bundles with degree given by an arbitrary non-positive multiple of $k+1$ (see Conjecture 8.5 for a brief discussion of possible improvement to any non-positive degree);
- allowing resonant nodes: in our paper all nodes were linearizable because their Camacho-Sad index was greater than 1. But nodes with smaller Camacho-Sad index pose no additional problem.

We propose to address this matter in the near future.

The other application regards differential Galois theory: heuristically, classification invariants carry Galoisian information (pertaining to the integrability in Liouvillian closed-form). For instance, in the limiting case of a saddle-node singularity it is well-known that Martinet-Ramis moduli play the same role for non-linear equations as Stokes matrices do for linear systems near an irregular singularity. A Galoisian formulation of this fact in terms of Malgrange groupoid [28, 27] can be found in the work of G. Casale [6]. When the differential equation depends on a parameter ε, the recent thesis of D. Davy describes a form of "semi-continuity" for specializations of its parametrized Malgrange groupoid $\mathfrak{M}$. Davy proves that the size of the groupoid $\mathfrak{M}_\varepsilon$ is constant if ε is generic, more precisely if the parameter does not belong to a (maybe empty) countable union Ω of hypersurfaces, while for $\varepsilon \in \Omega$ the groupoid $\mathfrak{M}_\varepsilon$ can only get smaller. The present study illustrates and refines this phenomenon.

Consider the extreme case $P_\varepsilon(x)\frac{\partial}{\partial x} + y\left(1 + \mu_\varepsilon x^k + \varepsilon R_\varepsilon(x,y)\right)\frac{\partial}{\partial y}$ for R arbitrary: the vector field X_0 is surely "not less integrable" (it is the formal normal form) than for $\varepsilon \neq 0$. This is actually the only possible kind of degeneracy near the saddle-node bifurcation, for we will establish that $\Omega \cap \left(\mathbb{C}^k, 0\right)$ is either empty or a germ of an analytic variety. We obtain the latter property by unfolding a result by M. Berthier and F. Touzet [1], characterizing vector fields admitting a local non-trivial Liouvillian first integral near an elementary singularity. We deduce that normal forms of integrable unfoldings are necessarily a Bernoulli unfolding (1.3). Both proofs are very different in nature, and we obtain a particularly short one by framing the problem for normal forms, revealing the usefulness of their simple expression and of the explicit section of their cokernel.

2. Statement of the main results

In all that follows ε is the parameter, belonging to some $\left(\mathbb{C}^k, 0\right)$ for $k \in \mathbb{N}$, and we study (holomorphic germs of) a parametric family of (germs at $0 \in \mathbb{C}^2$ of) vector fields $Z = (Z_\varepsilon)_{\varepsilon \in (\mathbb{C}^k, 0)}$ for which a saddle-node bifurcation occurs at $\varepsilon = 0$. That is to say, when $\varepsilon = 0$ the vector field Z_0 is of *saddle-node* type near the origin of $\mathbb{C}^2$:

— 0 is an isolated singularity of Z_0,

— the differential at 0 of the vector field has exactly one non-zero eigenvalue (the singularity is elementary degenerate).

The family $Z = (Z_\varepsilon)_\varepsilon$ is called a holomorphic germ of an *unfolding* of Z_0. We study in details only "generic" unfoldings, those which possess the "right number" of parameters to encode the bifurcation structure. Roughly speaking we require that for an open and dense set of parameters the vector field Z_ε have $k + 1$ distinct non-degenerate singular points. The latter merge into a saddle-node singularity of multiplicity $k + 1$ (codimension k) as $\varepsilon \to 0$. Let us make these statements precise.

Definition 2.1. – An unfolding Z of a codimension $k \in \mathbb{N}$ saddle-node Z_0 is *generic* if there exists a biholomorphic change of coordinates and parameter such that, in the new coordinates (x, y) and new parameter ε, the singular points of each Z_ε are given by $P_\varepsilon(x) = y = 0$, where

$$P_\varepsilon(x) := x^{k+1} + \varepsilon_{k-1}x^{k-1} + \cdots + \varepsilon_1 x + \varepsilon_0.$$

REMARK 2.2. – Generic families are essentially *universal*. In particular, the bifurcation diagram of singular points is the elementary catastrophe of codimension k (in the complex domain).

The analytic unstable manifold of Z_0, tangent at 0 to the eigenspace associated to the non-zero eigenvalue of its differential, is called the *strong separatrix*. The other eigenspace corresponds to a "formal separatrix" $\{y = \widehat{s}_0(x)\}$ called the *weak separatrix* (generically divergent [36], always summable in the sense of Borel [17]). We say that a saddle-node is convergent or divergent according to the nature of its weak separatrix.

DEFINITION 2.3. – We say that the generic unfolding Z is *purely convergent* when there exists a holomorphic function

$$\begin{aligned} s : \left(\mathbb{C}^k, 0\right) \times (\mathbb{C}, 0) &\longrightarrow \mathbb{C} \\ (\varepsilon, x) &\longmapsto s_\varepsilon(x) \end{aligned}$$

such that:

— each graph $\mathcal{S}_\varepsilon$ of s_ε is tangent to Z_ε and contains $\operatorname{Sing}(Z_\varepsilon)$ (the singular set of Z_ε, consisting in all zeros of Z_ε),

— $\mathcal{S}_0$ is the weak separatrix of Z_0 (in particular the latter is convergent).

We call $\mathrm{Convergent}_k$ the set of all such unfoldings.

REMARK 2.4. – 1. By applying beforehand the change of variables

$$(\varepsilon, x, y) \mapsto (\varepsilon, x, y + s_\varepsilon(x))$$

to the unfolding we can always assume that $\{y = 0\}$ is invariant by Z_ε for all $\varepsilon \in \left(\mathbb{C}^k, 0\right)$.

2. There exist unfoldings Z of a convergent saddle-node Z_0 such that, for all ε close enough to 0, no analytic invariant curve $\mathcal{S}_\varepsilon$ exist. We use the term "purely convergent" to insist that in the present case *every* vector field Z_ε for $\varepsilon \in \left(\mathbb{C}^k, 0\right)$ must admit an analytic invariant curve.

2.1. Normalization of purely convergent unfoldings

For $\mathbf{z}$ a finite-dimensional complex multivariable we write $\mathbb{C}\{\mathbf{z}\}$ the algebra of convergent power series in $\mathbf{z}$, naturally identified with the space of germs of a holomorphic function at $\mathbf{0}$. We extend this notation in the obvious manner so that $\mathbb{C}\{\varepsilon, x\}$ is the space of convergent power series in the $k + 1$ complex variables $\varepsilon_0, \ldots, \varepsilon_{k-1}$ and x.

2.1.1. *Formal classification.* – We first give an unfolded version of the well-known Bruno-Dulac-Poincaré normal forms [3, 13, 12]. Here we do not assume that Z is purely convergent.

Formal Normalization Theorem. – *Let $k \in \mathbb{N}$ be given. For $\varepsilon = (\varepsilon_0, \ldots, \varepsilon_{k-1}) \in \mathbb{C}^k$ define the polynomial*

$$P_\varepsilon(x) := x^{k+1} + \sum_{j=0}^{k-1} \varepsilon_j x^j.$$

Take a generic unfolding Z of a saddle-node of codimension k. There exists $(\mu, u) \in \mathbb{C}\{\varepsilon\} \times \mathbb{C}\{\varepsilon, x\}$ with $(\varepsilon, x) \mapsto u_\varepsilon(x)$ polynomial in x of degree at most k and satisfying $u_0(0) \neq 0$, such that Z is formally conjugate to the formal normal form

$$\widehat{Z} := u\,\widehat{X}, \tag{2.1}$$

where

$$\widehat{X}_\varepsilon(x, y) := P_\varepsilon(x)\frac{\partial}{\partial x} + y\left(1 + \mu_\varepsilon x^k\right)\frac{\partial}{\partial y} \tag{2.2}$$

defines the formal orbital normal form. *Notice that these vector fields are polynomial in (x, y) and holomorphic in $\varepsilon \in \left(\mathbb{C}^k, 0\right)$.*

In general the parameter of the normal form $\widehat{Z}$ differs from the original parameter of Z. However the formal change of parameter $\varepsilon \mapsto \phi(\varepsilon)$ happens to be actually analytic (as proved in [37, Theorem 3.5] and recalled in Theorem 4.1). Moreover such normal forms are essentially unique, in the sense that among all formal conjugacies only some linear changes of variables and parameter preserve the whole family. For example, transforming x into αx for some nonzero $\alpha \in \mathbb{C}$ in $P_\varepsilon \frac{\partial}{\partial x}$ yields the vector field

$$\frac{1}{\alpha} P_\varepsilon(\alpha x)\frac{\partial}{\partial x} = \alpha^k P_{\widetilde{\varepsilon}}(x)\frac{\partial}{\partial x}$$

where $\widetilde{\varepsilon} := \left(\varepsilon_j \alpha^{1-j}\right)_{j<k}$. Therefore by taking $\alpha^k = 1$ the linear change $(\widetilde{\varepsilon}, x) \mapsto (\varepsilon, \alpha x)$ transforms $\left(\widehat{X}_\varepsilon\right)_\varepsilon$ into $\left(\widehat{X}_{\widetilde{\varepsilon}}\right)_{\widetilde{\varepsilon}}$. It turns out this is the only degree of freedom for formal changes of parameters (see Section 4), which makes the parameter of the normal form special.

Definition 2.5. – The parameter of the normal form $\widehat{Z}$ (*modulo* the action of $\mathbb{Z}/k\mathbb{Z}$ on (ε, x)) is called the *canonical parameter* of the original unfolding Z. In all the following a representative ε of the canonical parameter is always implicitly fixed and forbidden to change.

As a consequence, two formal normal forms with formal invariants (μ, u) and $(\widetilde{\mu}, \widetilde{u})$ as above are (for fixed canonical parameter ε):

1. orbitally formally equivalent if, and only if, they have the same *formal orbital invariant* $\mu = \widetilde{\mu}$;
2. formally conjugate if, and only if, they have the same *formal invariant* $(\mu, u) = (\widetilde{\mu}, \widetilde{u})$.

2.1.2. *Analytical normalization*

DEFINITION 2.6. – For $k \in \mathbb{N}$ a positive integer let us introduce the functional space in the complex multivariable $(\varepsilon, x, v) \in \mathbb{C}^{k+2}$:

$$\mathrm{Section}_k\{v\} := \left\{ f_\varepsilon(x,v) = v\sum_{j=1}^{k} f_{\varepsilon,j}(v)\,x^j \;:\; f_{\varepsilon,j}(v) \in \mathbb{C}\{\varepsilon, v\} \right\}.$$

We let v figure explicitly in the notation $\mathrm{Section}_k\{v\}$ since this variable (and this variable only) will be subject to further specification.

NORMALIZATION THEOREM. – *For a given $k \in \mathbb{N}$ we fix a formal orbital invariant $\mu \in \mathbb{C}\{\varepsilon\}$ and choose $\tau \in \mathbb{Z}_{\geq 0}$ such that $\mu_0 + (k+1)\tau \notin \mathbb{R}_{\leq 0}$. For every $Z \in$ $\mathrm{Convergent}_k$ with formal invariant (μ, u), there exist $Q, R \in$ $\mathrm{Section}_k\{P^\tau y\}$ such that Z is analytically conjugate to*

$$\mathcal{Z} := \frac{u}{1+uQ}\mathcal{X} \tag{2.3}$$

where

$$\mathcal{X} := \widehat{\mathrm{X}} + Ry\frac{\partial}{\partial y}. \tag{2.4}$$

REMARK 2.7. – In case $\tau = 0$ (which can be enforced whenever the generic condition $\mu_0 \notin \mathbb{R}_{\leq 0}$ holds) normal forms induce foliations with holomorphic extension to $\mathbb{P}_1(\mathbb{C}) \times (\mathbb{C}, 0)$. This is no longer true if $\tau > 0$ and if R is not polynomial in the y-variable.

Specializing the theorem to $\varepsilon = 0$, we recover the earlier results [43, 26]. Let us briefly present the unfolded geometric construction of F. Loray (performed at an orbital level in [26] when $k = 1$) to get the gist of the argument. We define a holomorphic family of abstract foliated complex surfaces $(\mathcal{M}, \mathcal{F}) = (\mathcal{M}_\varepsilon, \mathcal{F}_\varepsilon)_{\varepsilon \in (\mathbb{C}^k, 0)}$ given by two charts. The first one is a domain $\mathcal{U}^0 := \left\{0 \leq |x| < \rho^0\right\} \times (\mathbb{C}, 0)$ together with some arbitrary convergent unfolding Z, provided the following non-restrictive properties (see [41]) are fulfilled for all $\varepsilon \in \left(\mathbb{C}^k, 0\right)$:

- Z_ε is holomorphic on the domain and has at most $k+1$ singular points in $\mathcal{U}^0$ (counted with multiplicity in case of saddle-nodes) each one located within $\mathcal{U}^0 \cap \{0 \leq |x| < {}^1/_{\rho^\infty}\}$ for some $\rho^\infty > 1/\rho_0$,
- Z_ε is transverse to the lines $\{x = c\}$ whenever $P_\varepsilon(c) \neq 0$,
- Z_ε leaves $\{y = 0\}$ invariant.

The other chart is a domain $\mathcal{U}^\infty := \{{}^1/_{\rho^\infty} < |x| \leq \infty\} \times (\mathbb{C}, 0)$ equipped with a foliation $\mathcal{F}_\varepsilon^\infty$

- having a single, reduced singularity at $(\infty, 0)$,
- otherwise transverse to the lines $\{x = \mathrm{cst}\}$,
- leaving $\{y = 0\}$ invariant.

Biholomorphic fibered transition maps fixing $\{y=0\}$ exist on the annulus $\mathcal{U}^0 \cap \mathcal{U}^\infty$ precisely when Z_ε and $\mathcal{F}_\varepsilon^\infty$ have (up to local conjugacy) mutually inverse holonomy maps above, say, the invariant circle $\frac{\rho_0\rho_\infty+1}{2\rho_\infty}\mathbb{S}^1\times\{0\}$. The resulting complex surface $\mathcal{M}_\varepsilon$ is naturally a holomorphic fibration by disks over the divisor $\mathcal{L} \simeq \mathbb{P}_1(\mathbb{C})$. In other words $\mathcal{M}_\varepsilon$ is a germ of a Hirzebruch surface, classified at an analytic level [22, 48, 14] by the self-intersection $-\widehat{\tau} \in \mathbb{Z}_{\leq 0}$ of $\mathcal{L}$ in $\mathcal{M}_\varepsilon$. From the compactness of $\mathcal{L}$ stems the polynomial-in-x nature of the foliation $\mathcal{F}_\varepsilon$. Other considerations then allow to recognize that $\mathcal{F}$ is (globally conjugate to a family of foliations) in normal form (2.3).

Let us explain where $\mathcal{F}_\varepsilon^\infty$ comes from, and at the same time how the Hirzebruch class $\widehat{\tau} = (k+1)\,\tau$ is involved. When the construction of $(\mathcal{M}, \mathcal{F})$ is possible, the global holomorphic foliation $\mathcal{F}_\varepsilon$ leaves the compact divisor $\mathcal{L}$ invariant and Camacho-Sad index formula [4] applies. The sum of indices of Z_ε at its $k+1$ singularities, with respect to $\mathcal{L}$, is μ_ε so $\mathcal{F}_\varepsilon^\infty$ must have index $-(\mu_\varepsilon + \widehat{\tau})$. By assumption the singularity at $(\infty, 0)$ can therefore never be a (saddle-)node. Invoking the realization result of [43, Section 4.4] (more precisely in the chart near $(\infty, 0)$) it is always possible to find a foliation $\mathcal{F}_\varepsilon^\infty$ with the desired properties. On the contrary when $\mu_\varepsilon + \widehat{\tau} \leq 0$ then no such $\mathcal{F}_\varepsilon^\infty$ may exist at all except in very special cases (detailed in [26, Theorem 2]) since, for instance, the holonomy along $\mathcal{L}$ of a node is always linearizable while the weak holonomy of Z_ε has no reason to be linearizable. We discuss this problem in more details while dealing with the non-linear Riemann-Hilbert problem below.

Therefore one can always take $\tau := 0$ except when $\mu_0 \leq 0$, which accounts for the "twist" $P_\varepsilon(x)^\tau\, y \sim_{x\to\infty} x^{\widehat{\tau}} y$ in normal forms (2.3).

2.1.3. *Normal forms uniqueness*. – To fully describe the quotient space (moduli space) of $\mathrm{Convergent}_k$ by analytical conjugacy / orbital equivalence, the Normalization Theorem must be complemented with a description of equivalence classes within the family of normal forms (2.3), leading us to discuss its uniqueness clause.

Definition 2.8. – 1. For $Z \in \mathrm{Convergent}_k$ we denote

$$\mathfrak{n}(Z) := (\mu, u, R, Q) \tag{2.5}$$

$$\mathfrak{o}(Z) := (\mu, R) \tag{2.6}$$

respectively the *normal invariant* of Z and its *normal orbital invariant*, where the functional tuples on the right-hand side are given by the Normalization Theorem.

2. For $c \in \mathbb{C}\{\varepsilon\}^\times$ and $f \in \mathbb{C}\{\varepsilon, x, y\}$ define

$$c^* f := (\varepsilon, x, y) \longmapsto f_\varepsilon(x, c_\varepsilon y)\,.$$

We extend component-wise this action of $\mathbb{C}\{\varepsilon\}^\times$ to tuples of functions such as $\mathfrak{n}$ and $\mathfrak{o}$ above.

Uniqueness Theorem. – 1. *Two normal forms* (2.3) *associated to the same fixed* τ *and moduli* (2.5) $\mathfrak{n}$ *and* $\widetilde{\mathfrak{n}}$ *are analytically conjugate (by a change of coordinates fixing the parameter) if, and only if, there exists* $c \in \mathbb{C}\{\varepsilon\}^\times$ *such that* $c^*\mathfrak{n} = \widetilde{\mathfrak{n}}$. *For any conjugacy* $\Psi : (\varepsilon, x, y) \mapsto (\varepsilon, \Psi_\varepsilon(x, y))$ *there exists a unique* $t \in \mathbb{C}\{\varepsilon\}$ *such that*

$$\Psi = c^* \Phi_{\mathcal{Z}}^t,$$

where $\Phi^t_{\mathcal{Z}}$ is the local flow of $\mathcal{Z}$ at time $t \in \mathbb{C}$. Moreover it is fibered in the x-variable if, and only if, $t = 0$. In that case Ψ is linear:

$$\Psi = c^*\mathrm{Id} : (\varepsilon, x, y) \longmapsto (\varepsilon, x, c_\varepsilon y) .$$

2. *Let $\mathfrak{o}$ and $\widetilde{\mathfrak{o}}$ be the corresponding orbital invariants. The normal forms are analytically orbitally equivalent (by a change of coordinates fixing the parameter) if, and only if, there exists $c \in \mathbb{C}\{\varepsilon\}^\times$ such that $c^*\mathfrak{o} = \widetilde{\mathfrak{o}}$. For any orbital equivalence Ψ there exists a unique $F \in \mathbb{C}\{\varepsilon, x, y\}$ such that*

$$\Psi = c^*\Phi^F_{\mathcal{Z}}.$$

Moreover Ψ is fibered in the x-variable if, and only if, $F = 0$. In that case Ψ is linear.

REMARK 2.9. – In particular normal forms (2.3) are unique when only tangent-to-identity in the y-variable, fibered in the x-variable conjugacies are allowed.

Again the proof is largely based on the strategy of F. Loray introduced in [26], although the actual implementation in the parametric case calls for subtle adaptations. The idea is to extend any local and fibered conjugacy between normal forms to a global conjugacy on a "big" neighborhood of $\mathcal{L}$, from which it easily follows that only linear maps can do that.

2.2. Inverse problem

For given $k \in \mathbb{N}$ we can split the parameter space $\left(\mathbb{C}^k, 0\right)$ into $C_k = \frac{1}{k+1}\binom{2k}{k}$ open cells $\mathcal{E}_\ell$ such that

$$\bigcup_\ell \mathcal{E}_\ell = \left(\mathbb{C}^k, 0\right) \setminus \Delta_k,$$

where Δ_k is the set of parameters ε for which P_ε has at least a multiple root (Δ_k is the discriminant curve). We recall that we can associate [41] an orbital modulus to a purely convergent unfolding Z

$$\mathfrak{m}(Z) := (\mathfrak{m}_\ell)_{1 \le \ell \le C_k}$$
$$\mathfrak{m}_\ell := \left(\phi^{j,\mathrm{s}}_\ell\right)_{j \in \mathbb{Z}/k\mathbb{Z}}$$

where for each $j \in \mathbb{Z}/k\mathbb{Z}$ and each ℓ the map

$$(\varepsilon, h) \in \mathcal{E}_\ell \times (\mathbb{C}, 0) \longmapsto \phi^{j,\mathrm{s}}_{\ell,\varepsilon}(h)$$

is holomorphic, vanishes along $\{h = 0\}$ and admits a continuous extension to $\mathrm{cl}(\mathcal{E}_\ell) \times (\mathbb{C}, 0)$.

REMARK 2.10. – 1. The upper index "s" is purely notational and refers to the fact that the function $\phi^{j,\mathrm{s}}_{\ell,\varepsilon}$ comes from the j-th "s"addle intersection, where the dynamics behaves very much like a saddle point.

2. The diffeomorphisms $\psi^{j,\mathrm{s}}_{\ell,\varepsilon}$, which unfold the components $\psi^{j,\mathrm{s}}_0$ of the (classical) Martinet-Ramis modulus, are given by $\psi^{j,\mathrm{s}}_{\ell,\varepsilon}(h) = h \exp\left(\frac{2\mathrm{i}\pi\mu_\varepsilon}{k} + \phi^{j,\mathrm{s}}_{\ell,\varepsilon}(h)\right)$.

Let us write $\mathscr{H}_\ell\{h\}$ the vector space of all such functions, so that

$$\mathfrak{m}(Z) \in \prod_\ell \mathscr{H}_\ell\{h\}^k .$$

The data $\mathfrak{m}(Z)$ is a complete orbital invariant for the local analytic classification of purely convergent unfoldings.

2.2.1. *Orbital realization.* – The definition of the *compatibility condition* involves notions going beyond the scope of the present summarized statements. We refer to Section 7.3 for a precise definition. Instead let us use the following terminology.

Definition. – We say that $(\mu, \mathfrak{m}) \in \mathbb{C}\{\varepsilon\} \times \prod_\ell \mathscr{H}_\ell\{h\}^k$ is *realizable* if there exists a generic convergent unfolding Z with formal orbital class μ and orbital modulus $\mathfrak{m} = \mathfrak{m}(Z)$.

For the sake of completeness, let us state the following fundamental result even though all material was not properly introduced.

Orbital Realization Theorem. – *Let $\mu \in \mathbb{C}\{\varepsilon\}$ be given. A functional data $\mathfrak{m} \in \prod_\ell \mathscr{H}_\ell\{h\}^k$ yields a realizable $(\mu, \mathfrak{m})$ if, and only if, $(\mu, \mathfrak{m})$ satisfies the compatibility condition (presented in Definition 7.16).*

Although it is not directly used in the present paper, considerations akin to those from [43] show that the map sending a normal form to its orbital modulus

$$\mathfrak{o} = (\mu, R) \longmapsto (\mu, \mathfrak{m})$$

is upper-triangular, in the sense that the n-th-jet of $\mathfrak{m}_\ell$ with respect to h is completely determined by μ and the n-th-jet of R with respect to y. In that sense passing from modulus to normal form is a (non-effective) computable process. In the case $k = 1$ we show how to compute the diagonal entries. More details on this topic can be found in Section 10.1.

2.2.2. *Moduli which are analytic with respect to the parameter.* – Our final main result proves that the compatibility condition defines a proper subset of the vector space $\mathbb{C}\{\varepsilon\} \times \prod_\ell \mathscr{H}_\ell\{h\}^k$. Let us start with a definition.

Definition 2.11. – 1. A *Bernoulli vector field* of index $d \in \mathbb{Z}_{\geq -1}$ takes the form

$$X(x, y) = a(x)\frac{\partial}{\partial x} + \left(b(x) + c(x)y^d\right)y\frac{\partial}{\partial y}$$

for holomorphic $a, b, c \in \mathbb{C}\{x\}$.

2. A *Bernoulli unfolding* is a saddle-node unfolding $(X_\varepsilon)_\varepsilon$ with members of the special form

$$X_\varepsilon = \widehat{X}_\varepsilon + g_\varepsilon(x)y^{d+1}\frac{\partial}{\partial y}$$

for some analytic germ $(\varepsilon, x) \mapsto g_\varepsilon(x) \in \mathbb{C}\{\varepsilon, x\}$. In particular each X_ε is a Bernoulli vector field.

Remark 2.12. – Such a vector field is named that way because the underlying non-autonomous differential equation is Bernoulli:

$$a(x)y' = b(x)y + c(x)y^{d+1}.$$

Parametrically Analytic Orbital Moduli Theorem. – *Let there be given* $\mu \in \mathbb{C}\{\varepsilon\}$ *and* $\mathfrak{m} = (\mathfrak{m}_\ell)_\ell \in \prod_\ell \mathscr{H}_\ell \{h\}^k$. *Assume* $\mathfrak{m}$ *is holomorphic, in the sense that* $\mathfrak{m}_\ell = M|_{\mathcal{E}_\ell \times (\mathbb{C},0)}$ *for some* $M \in h\mathbb{C}\{\varepsilon, h\}^k$. *The following conditions are equivalent:*

1. $(\mu, \mathfrak{m})$ *satisfies the compatibility condition,*
2. *either* $\mathfrak{m} = 0$, *or* $k = 1$ *and there exists* $d \in \mathbb{N}$, $\alpha \in \mathbb{C}\{\varepsilon\} \setminus \{0\}$ *such that*
   - $d\mu \in \mathbb{Z}$ *(in particular* μ *is a rational constant),*
   - $M(h) = -\frac{1}{d}\log\left(1 - \alpha h^d\right)$.

If one of the conditions is satisfied and $\mathfrak{m} \neq 0$, *then* $(\mathfrak{m}_\ell)_\ell$ *can be realized by a Bernoulli unfolding:*

$$\mathscr{X}_\varepsilon = \widehat{\mathrm{X}}_\varepsilon + r_\varepsilon x P_\varepsilon(x)^{\tau d}\, y^{d+1}\frac{\partial}{\partial y}$$

for some $r \in \mathbb{C}\{\varepsilon\} \setminus \{0\}$.

Remark 2.13. – By letting $\mathbb{C}\{\varepsilon\}^\times$ act linearly through $(\varepsilon, x, y) \mapsto (\varepsilon, x, c_\varepsilon y)$ we may normalize further r to some ε^κ for $\kappa \in \mathbb{Z}_{\geq 0}$. See also Section 10.1. Observe that r_ε in the normal form does not depend on x because $k = 1$.

2.3. Application: non-linear differential Galois theory

M. Berthier and F. Touzet [1] proved that the Martinet-Ramis modulus of a convergent saddle-node vector field admitting non-trivial Liouvillian first integrals [35] must be a ramified homography $h \mapsto \alpha h\left(1 + \beta h^d\right)^{-1/d}$, from which they deduce that the vector field is conjugate to a Bernoulli vector field. It is indeed straightforward to compute the modulus of a Bernoulli vector field (by solving explicitly the underlying differential equation) and observe that it is a ramified homography, and that all such ramified homographies are reached this way. (We refer to Lemma 9.3 for the computation.)

Roughly speaking Liouvillian integrability corresponds to differential equations admitting "closed-form" solutions obtained by iteratively taking quadrature (or exponential thereof) of elements of (algebraic extensions of) the base-field (here, meromorphic functions on a polydisk containing $P_\varepsilon^{-1}(0) \cap \{y = 0\}$).

Integrability Theorem. – *Let* $(Z_\varepsilon)_\varepsilon$ *be a generic, purely convergent saddle-node unfolding and denote by* $\mathscr{L}$ *the germ of set consisting in those* $\varepsilon \in \left(\mathbb{C}^k, 0\right)$ *for which* Z_ε *admits a Liouvillian first integral. The following statements are equivalent.*

1. *The locus of integrability* $\mathscr{L}$ *is full:* $\mathscr{L} = \left(\mathbb{C}^k, 0\right)$.
2. *Its (analytic) Zariski closure is full:* $\overline{\mathscr{L}}^{\mathtt{Zar}} = \left(\mathbb{C}^k, 0\right)$.
3. *Its orbital normal form* $\mathscr{X}$ *is a Bernoulli unfolding.*

Remark 2.14. – 1. The case $\overline{\mathscr{L}}^{\mathtt{Zar}} \neq \left(\mathbb{C}^k, 0\right)$ corresponds to $\mathscr{L}$ being a germ at 0 of a proper analytic subvariety. Then $\mathscr{L}$ is the locus of parameters for which the normal form is Bernoulli. For instance in case of the normal form given by $R_\varepsilon(t) := t^d + L(\varepsilon)\, t^{d+1}$ we have $\mathscr{L} = L^{-1}(0)$, as we discuss after the proof of the theorem.

2. In the case $k = 1$ the second condition is equivalent to any of the following three conditions: the germ $\mathcal{L}$ accumulates on 0, $\mathcal{L}$ is infinite, $\mathcal{L} \neq \{0\}$.

Proof. – The property of having a non-trivial Liouvillian first integral is both orbital and invariant by change of analytic coordinates, so we do not lose generality by taking $Z = \mathcal{X}$ in normal form (2.4). Integrability is equivalent to the existence of a Godbillon-Vey sequence [16] of length at most 2, that is to the existence of two non-zero meromorphic 1-forms ω and η for which

$$\begin{aligned} \mathrm{d}\eta &= 0 \\ \mathrm{d}\omega &= \delta\omega \wedge \eta, \quad \delta \in \{0,1\} \\ \omega\left(\mathcal{X}\right) &= 0. \end{aligned}$$

(The multivalued map $H := \int \exp\left(\delta \int \eta\right)\omega$ is indeed a Liouvillian first integral of $\mathcal{X}$, obtained by quadrature of closed 1-forms.) This in turn is (almost) equivalent to solving for meromorphic, transverse $Y \neq 0$ in the Lie-bracket equation

$$\left[\mathcal{X}, Y\right] = \delta Y, \quad \delta \in \{0,1\}, \tag{2.7}$$

since the dual basis (η, ω) of $\left(\mathcal{X}, Y\right)$ is a Godbillon-Vey sequence and vice versa. There is a subtlety here, because $\mathcal{X}$ may fail to meet this condition while there could exist an integrating factor V for which $V\mathcal{X}$ does. We deliberately ignore this eventuality, because the case $V \neq 1$ can be deduced from the particular case $V = 1$ by a direct (albeit cumbersome) adaptation. For the same reason we only deal with the case $\mu_0 \notin \mathbb{R}_{\leq 0}$.

The implication (1) $\Rightarrow$ (2) is trivial. The implication (3) $\Rightarrow$ (1) comes from Remark 9.4, where a Liouvillian first-integral is computed explicitly for all $\varepsilon \in \left(\mathbb{C}^k, 0\right)$. Alternately one can compute the Godbillon-Vey sequence as the dual basis of

$$(V\mathcal{X}, Y) := \left(\frac{1-\mu x^k}{d}\mathcal{X}, \exp\left(d\int_0^x \frac{\mu^2 z^{2k}}{1-\mu z^k} \times \frac{\mathrm{d}z}{P(z)}\right) y^{d+1}\frac{\partial}{\partial y}\right).$$

Observe that $V\mathcal{X}$ is in normal form and its temporal modulus is trivial.

Let us prove (2) $\Rightarrow$ (3). The strategy is the following: we first show that the vector field is Bernoulli for each $\varepsilon \in \mathcal{L}$, then we invoke the analyticity of the normal form and the fact that $\overline{\mathcal{L}}^{\mathtt{Zar}}$ is full to cover a whole neighborhood of 0 in parameter space. Hence, let us fix $\varepsilon \in \mathcal{L}$ and drop the index ε altogether. According to the above discussion one can find $\delta \in \{0,1\}$ and a vector field

$$Y = A(x,y)\frac{\partial}{\partial y} + B(x,y)\,\mathcal{X}$$

solving (2.7) for two functions $A \neq 0$ and B meromorphic on a polydisk containing $P^{-1}(0) \cap \{y = 0\}$. From (2.7) we deduce the relations

$$\begin{cases} \mathcal{X}\cdot B &= \delta B, \\ \mathcal{X}\cdot A &= \left(\delta + 1 + \mu x^k + \frac{\partial_y R}{\partial y}\right) A. \end{cases} \tag{2.8}$$

The second equation tells us that $\{A = 0\} \cup \{A = \infty\}$ is a union of separatrices of $\mathcal{X}$, therefore of the form

$$A(x, y) = y^{d+1} U(x, y) \prod_{P(z)=0} (x - z)^{\ell(z)},$$

for some choice of $d, \ell(z) \in \mathbb{Z}$ and for some holomorphic and never vanishing function U. Let us prove that $R = r(x) y^d$, from which follows either $d \in \mathbb{N}$ or $R = 0$.

The last equation of (2.8) becomes

$$\mathcal{X} \cdot \log U = \delta - d\left(1 + \mu x^k\right) - \sum_{P(z)=0} \ell(z) \frac{P(x)}{x - z} + \left(y \frac{\partial R}{\partial y} - dR\right),$$

because $\mathcal{X} \cdot \log y = 1 + \mu x^k + R$. Evaluating this identity at any one of the $k + 1$ points $(x, y) = (x, 0)$ such that $P(x) = 0$ yields $0 = \delta - d\left(1 + \mu x^k\right) - \ell(x) P'(x)$, since on the one hand R and $y \frac{\partial R}{\partial y}$ vanish when $y = 0$ while on the other hand $\ell(z) \frac{P(x)}{x-z}$ evaluates to 0 if $z \neq x$ and to $\ell(x) P'(x)$ otherwise. As a consequence we have equality of the polynomials $\sum_{P(z)=0} \ell(z) \frac{P(x)}{x-z} = \delta - d\left(1 + \mu x^k\right)$ of degree k. Therefore

$$\mathcal{X} \cdot \log U = y \frac{\partial R}{\partial y} - dR.$$

In the course of Section 6 we show that $\operatorname{im}\left(\mathcal{X}\cdot\right) \cap \mathrm{Section}_k\{y\} = \{0\}$ (see Remark 6.3). Hence, the fact that $y \frac{\partial R}{\partial y} - dR \in xy\mathbb{C}[x]_{<k}\{y\} = \mathrm{Section}_k\{y\}$ lies in the image of $\mathcal{X}\cdot$ can only mean $y \frac{\partial R}{\partial y} - dR = 0$. From this we deduce at once that

$$R(x, y) = x r(x) y^d, \quad r \in \mathbb{C}[x]_{<k}.$$

The condition that, for a specific ε, the vector field $\mathcal{X}_\varepsilon$ be Bernoulli corresponds to the vanishing of all coefficients in R_ε of y^n but for $n = d$. Since $(\varepsilon, y) \mapsto R_\varepsilon(y)$ is analytic with respect to ε and $\overline{\mathcal{L}}^{\mathrm{Zar}} = \left(\mathbb{C}^k, 0\right)$, if d is independent on ε then the identity principle implies that $R_\varepsilon(x, y) = x r_\varepsilon(x) y^d$ for all $(\varepsilon, x, y) \in \left(\mathbb{C}^{k+2}, 0\right)$. The fact that d is indeed independent on ε stems from Baire's category theorem. □

REMARK 2.15. – The proof relies in an essential way on the analyticity of the orbital normal form $\mathcal{X}$ with respect to ε near 0. Compare with the method of proof of [1]: for $\varepsilon = 0$ the argument is based on the fact that the existence of a Godbillon-Vey sequence forces the Martinet-Ramis modulus to be a ramified homography. This argument works as well for $\varepsilon \neq 0$, but we could not have argued on from there since the modulus is in general not analytic at $\varepsilon = 0$: although being a ramified homography is an analytic condition, an accumulation of zeros of this relation as $\varepsilon \to 0$ could arise without holding for all ε (for $k = 1$, say). This situation cannot occur, and our shorter argument does not involve the unfolded modulus of classification.

The Galoisian characterization of the existence of Godbillon-Vey sequences of length at most 2 is performed in [6], and for fixed ε its length equals the (transverse) rank $\mathrm{rk}(\mathfrak{M}_\varepsilon)$ of the Galois-Malgrange groupoid $\mathfrak{M}_\varepsilon$. This rank takes values in $\{0, 1, 2, \infty\}$, integrability corresponding to finite values. For the normal forms (2.4) with $R_\varepsilon(x,t) = \sum_{n>0} r_{\varepsilon,n}(x)\, t^n$, we have

$$\operatorname{rk}(\mathfrak{M}_\varepsilon) = \begin{cases} 1 + \#\{n : r_{\varepsilon,n} \neq 0\} & \text{if it is } \leq 2, \\ \infty & \text{otherwise.} \end{cases}$$

Therefore $\varepsilon \mapsto \operatorname{rk}(\mathfrak{M}_\varepsilon)$ is lower semi-continuous: accidental values of the rank can only correspond to more integrable systems.

Example. – Taking into account Remark 2.13, in the case $k = 1$ and $R \neq 0$ the vector field $\mathscr{X}_0$ is "more integrable" (transverse rank 1) than the generic $\mathscr{X}_\varepsilon$ (transverse rank 2) if and only if the exponent κ is positive.

This is a special instance of a general result on parametrized Galois-Malgrange groupoids obtained recently by G. Casale and D. Davy [7]. They show that for rather general deformations of foliations $(\mathscr{F}_\varepsilon)_\varepsilon$, the rank $\operatorname{rk}(\mathfrak{M}_\varepsilon)$ of the specialization $\mathfrak{M}_\varepsilon$ of the Galois-Malgrange groupoid of the family is lower semi-continuous in ε. Moreover, the locus of discontinuity is contained in a countable union of proper analytic subvarieties. We showed that in the case of purely convergent saddle-node bifurcations, the locus of discontinuity is at most a proper analytic subvariety.

2.4. Structure of the paper

- We begin with fixing notations and providing precise definitions in Section 3. Readers familiar with complex foliations may skip this section.
- The Formal Normalization Theorem is proved in Section 4.

We first present the generic case (for which one can take $\tau = 0$), since it is easier to highlight the ideas than in the case $\tau > 0$.

- The orbital part of the Normalization and Uniqueness Theorems are established in Section 5 when $\tau = 0$.
- The temporal part of the Normalization and Uniqueness Theorems are established in Section 6 when $\tau = 0$.
- In Section 7 one finds the definition of compatibility condition, and the proof of the Orbital Realization Theorem in the generic case $\tau = 0$.
- In Section 8 we prove the Orbital Realization Theorem in the case $\tau > 0$. This provides a posteriori a proof of the orbital part of the Normalization and Uniqueness Theorems when $\tau > 0$.
- In Section 9 we discuss the Bernoulli unfoldings and prove the Parametrically Analytic Orbital Moduli Theorem.
- Finally, in Section 10, we conclude with a few words on computations.

3. Preliminaries

3.1. Notations

3.1.1. *General notations*

- We let the set $\mathbb{N} := \{1, 2, \ldots\}$ stand for all positive integers, whereas the set of non-negative integers will be written $\mathbb{Z}_{\geq 0} = \{0, 1, \ldots\}$.
- For $n \in \mathbb{N}$ we let $(\mathbb{C}^n, 0)$ stand for any small enough domain in $\mathbb{C}^n$ containing 0.
- The domain $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$ is the standard open unit disk.
- The unit circle of $\mathbb{R}^2 \simeq \mathbb{C}$ is denoted by $\mathbb{S}^1 := \{z \in \mathbb{C} : |z| = 1\}$.
- The closure of a subset A of a topological space is written $\mathrm{cl}(A)$.
- $k \in \mathbb{N}$ is fixed, $\varepsilon = (\varepsilon_0, \ldots, \varepsilon_{k-1}) \in \left(\mathbb{C}^k, 0\right)$ is the parameter and

$$P_\varepsilon(x) = x^{k+1} + \sum_{j=0}^{k-1} \varepsilon_j x^j. \tag{3.1}$$

- The parameter space $\left(\mathbb{C}^k, 0\right)$ is covered by the closure of $C_k = \frac{1}{k+1}\binom{2k}{k}$ open and contractible cells $\mathcal{E}_\ell$.
- The period operator $\mathfrak{T} = \left(\mathfrak{T}^j\right)_{j \in \mathbb{Z}/k\mathbb{Z}}$ is built near Definition 6.10.
- The very nature of constructions involves using more sub- and super-scripts than one is generally comfortable with. To alleviate this downside we stick to a single convention: *subscripts are* always *parameter-related*, while superscripts are in general related to the geometric variables (x, y) or to indices in power series expansions. Example: we write $V_{\ell,\varepsilon}^{j,\mathrm{s}}$ for the "s"addle part of the j-th sector in x-variable, relatively to the parameter ε being taken in the ℓ-th parametric cell. In the course of the text we try to drop indices whenever possible.
- *The dependency on the parameter ε is implicit in most instances*. For example, $\mu \in \mathbb{C}\{\varepsilon\}$ stands for the formal orbital modulus while μ_ε stands for the value of μ at the particular value of the parameter ε. Yet in many places where ε is fixed we do use μ instead of μ_ε in order to help reducing the notational footprint. This also applies for other parametric objects.

3.1.2. *Functional spaces*. – In the following $\mathcal{R}$ is a commutative ring with a multiplicative action by complex numbers.

- $\mathcal{R}^\times$ is the multiplicative group of its invertible elements.
- $\mathcal{R}[\mathbf{z}]$ is the commutative ring of polynomials in the complex finite-dimensional (multi)variable $\mathbf{z} = (z_1, \ldots, z_n)$ with coefficients in $\mathcal{R}$.
- After choosing a binary relation $\prec$ among $\{=, <, \leq, \ldots\}$ we let $\mathcal{R}[\mathbf{z}]_{\prec d}$ be the subset of $\mathcal{R}[\mathbf{z}]$ consisting of polynomials P such that $\deg P \prec d$.
- The projective limit $\mathcal{R}[[\mathbf{z}]] := \lim_{d\to\infty} \mathcal{R}[\mathbf{z}]_{\leq d}$ is the ring of formal power series in $\mathbf{z}$ with coefficients in $\mathcal{R}$.
- $\mathbb{C}\{\mathbf{z}\}$ is the algebra of convergent formal power series in the complex multivariable $\mathbf{z} \in \mathbb{C}^n$, naturally identified to the set of germs of a holomorphic function near $0 \in \mathbb{C}^n$.

REMARK 3.1. – We will mostly use the spaces:

— $\mathbb{C}[[\varepsilon]]$, $\mathbb{C}[[\varepsilon, x]]$ and $\mathbb{C}[[\varepsilon, x, y]]$,
— $\mathbb{C}\{\varepsilon\}$, $\mathbb{C}\{\varepsilon\}^{\times}$, $\mathbb{C}\{\varepsilon, x\}$, and $\mathbb{C}\{\varepsilon, x, y\}$
— $\mathbb{C}\{\varepsilon\}[x]$, $\mathbb{C}\{\varepsilon\}[x]_{\leq k}^{\times}$ and

$$\mathrm{Section}_k\{v\} := xv\mathbb{C}\{\varepsilon, v\}[x]_{<k}\,.$$

Let $\mathscr{D} \subset \mathbb{C}^n$ be a domain containing 0 equipped with the affine coordinates $\mathbf{z} = (z_1, \ldots, z_n)$.

— $\mathrm{Holo}(\mathscr{D})$ is the algebra of complex-valued functions holomorphic on $\mathscr{D}$.
— $\mathrm{Holo}_{\mathrm{c}}(\mathscr{D})$ is the Banach subalgebra of $\mathrm{Holo}(\mathscr{D})$ of all holomorphic functions $f : \mathscr{D} \to \mathbb{C}$, with bounded continuous extension to $\mathrm{cl}(\mathscr{D})$, equipped with the norm

$$\|f\|_{\mathscr{D}} := \sup_{\mathbf{z} \in \mathscr{D}} |f(\mathbf{z})|\,.$$

— $\mathrm{Holo}_{\mathrm{c}}(\mathscr{D})'$ is the Banach space of all holomorphic functions $f : \mathscr{D} \to \mathbb{C}$ vanishing on $\{z_n = 0\}$ with the norm

$$\|f\|'_{\mathscr{D}} := \sup_{\mathbf{z} \in \mathscr{D}} \left|\frac{f(\mathbf{z})}{z_n}\right|.$$

Notice that $\|f\|'_{\mathscr{D}} \leq \left\|\frac{\partial f}{\partial z_n}\right\|_{\mathscr{D}}$ whenever $\frac{\partial f}{\partial z_n} \in \mathrm{Holo}_{\mathrm{c}}(\mathscr{D})$ and $\mathscr{D}$ is convex in the variable z_n.

— We let

$$\mathscr{H}_\ell\{\mathbf{z}\} := \bigcup_{\mathscr{D} = (\mathbb{C}^n, 0)} \mathrm{Holo}_{\mathrm{c}}(\mathscr{E}_\ell \times \mathscr{D})',$$

where $\mathscr{E}_\ell$ is a parametric cell.

3.1.3. *Vector fields and Lie derivative.* – Let $Z = \sum_{j=1}^n A_j \frac{\partial}{\partial z_j}$ be a germ of a holomorphic vector field at the origin of $\mathbb{C}^n$ (or formal vector field at this point).

— If f is a formal power series or a holomorphic function in $\mathbf{z} = (z_1, \ldots, z_n) \in (\mathbb{C}^n, 0)$, we denote by $Z \cdot f$ the directional Lie derivative of f along Z

$$Z \cdot f := \sum_{j=1}^n A_j \frac{\partial f}{\partial z_j} = \mathrm{D}f(Z)\,.$$

The operator is extended component-wise on vectors of power series or functions.

— We define recursively for $n \in \mathbb{Z}_{\geq 0}$ the n-th iterate of the Lie derivative, the operator written $Z^{\cdot n}$, by

$$Z^{\cdot 0} := \mathrm{Id}$$
$$Z^{\cdot n+1} := Z \cdot (Z^{\cdot n})\,.$$

— The flow of Z at time t starting from $\mathbf{z}$ is the formal n-tuple of power series $\Phi_Z^t(\mathbf{z})$ solving the flow-system

$$\frac{\partial \Phi_Z^t(\mathbf{z})}{\partial t} = Z \circ \Phi_Z^t(\mathbf{z}),$$

which is a convergent power series in $(t, \mathbf{z})$ if, and only if, Z is holomorphic. At some point we invoke the classical formal identity of Lie

$$f \circ \Phi_Z^t = \sum_{n \geq 0} \frac{t^n}{n!} Z \cdot^n f. \tag{3.2}$$

— Two vector fields Z and $\widetilde{Z}$ are formally / locally conjugate when there exists a n-tuple of formal / convergent power series Ψ with invertible derivative at 0 such that

$$\widetilde{Z} \cdot \Psi = Z \circ \Psi.$$

In that case we write $\widetilde{Z} = \Psi^* Z$.

— Two vector fields Z and $\widetilde{Z}$ are formally / locally orbitally equivalent when there exists a formal power series / holomorphic function U with $U(0) \neq 0$ such that UZ and $\widetilde{Z}$ are conjugate (in the same convergence class).

3.2. Conjugacy and orbital equivalence

Definition 3.2. – Two unfoldings $Z = (Z_\varepsilon)_\varepsilon$ and $\widetilde{Z} = \left(\widetilde{Z}_{\widetilde{\varepsilon}}\right)_{\widetilde{\varepsilon}}$ are locally *conjugate* (resp. *orbitally equivalent*) if there exists a holomorphic mapping

$$\Psi : (\varepsilon, x, y) \longmapsto (\phi(\varepsilon), \Psi_\varepsilon(x, y))$$

such that:

1. $\varepsilon \in \left(\mathbb{C}^k, 0\right) \mapsto \widetilde{\varepsilon} = \phi(\varepsilon)$ has invertible derivative at 0,
2. for each $\varepsilon \in \left(\mathbb{C}^k, 0\right)$ the component Ψ_ε is a local conjugacy (resp. orbital equivalence) between Z_ε and $\widetilde{Z}_{\phi(\varepsilon)}$.

If the above conditions are fulfilled we write

$$\Psi^* Z = \widetilde{Z}.$$

We extend in the obvious way the definition for formal conjugacy / orbital equivalence.

Remark 3.3. – The very first step of any construction performed here consists in recalling the preparation of the generic unfolding Z (Theorem 4.1). For unfoldings in prepared form (4.1) the parameter ε becomes a formal invariant. Hence we only use conjugacies fixing ε, that is $\Psi : (\varepsilon, x, y) \mapsto (\varepsilon, \Psi_\varepsilon(x, y))$. In that setting one can always deduce Ψ knowing Ψ_ε, therefore when we use the notation Ψ we generally refer to the map $(\varepsilon, x, y) \mapsto \Psi_\varepsilon(x, y)$, except when the context is ambiguous.

Definition 3.4. – Consider a formal transform $\Psi : (\varepsilon, x, y) \mapsto (\varepsilon, \Psi_\varepsilon(x, y))$. We say that Ψ is *fibered* when $\Psi_\varepsilon(x, y) = (x, \psi_\varepsilon(x, y))$.

Definition 3.5. – Ψ is a *symmetry* (resp. *orbital symmetry*) of Z when Ψ is a self-conjugacy (resp. orbital self-equivalence) of Z.

REMARK 3.6. – Hence, to determine the orbital symmetries of Z it suffices to determine the changes Ψ such that $\Psi^* Z = UZ$ for some U with $U_0(0,0) \neq 0$.

4. Formal normalization

The formal normalization is based on three ingredients, each one corresponding to a step of the construction:

- a preparation *à la* Dulac of unfoldings: for $\varepsilon = 0$ one recovers the Dulac prepared form [12, 13];
- the existence of a formal "family of weak separatrices" which we can straighten to $\{y = 0\}$;
- a variation on Lie's identity (3.2) already used in [41, 44] to perform the analytic classification of saddle-nodes vector fields and their unfoldings. The formula reduces the problem of finding changes of variables to solving an uncoupled system of cohomological equations.

4.1. Preparation

Take $\theta \in \mathbb{Z}/k\mathbb{Z}$ and set $\alpha := \exp{}^{2i\pi\theta}/k$. For $\varepsilon := (\varepsilon_j)_{j<k} \in (\mathbb{C}^k, 0)$ we define

$$\theta^* \varepsilon := \left(\varepsilon_j \alpha^{j-1}\right)_{j<k}.$$

THEOREM 4.1 ([41, Proposition 3.1 and Theorem 3.5]). – *Any generic unfolding is analytically conjugate to an unfolding of the form*

$$Z = UX \tag{4.1}$$

$$X = \widehat{\mathrm{X}} + A\frac{\partial}{\partial y} \tag{4.2}$$

$$A(x,y) = P(x)\,a(x) + yR(x,y)$$

where $\widehat{\mathrm{X}}$ *and* P *are defined in* (2.2) *and* (3.1)*, while* $a \in \mathbb{C}\{\varepsilon, x\}$*,* $R \in y\mathbb{C}\{\varepsilon, x, y\}$ *and* $U \in \mathbb{C}\{\varepsilon, x, y\}$ *with* $U_0(0,0) \neq 0$*. In the particular case of an analytic weak separatrix one can take* $a := 0$.

Besides if two such prepared forms $(Z_\varepsilon)_\varepsilon$ *and* $\left(\widetilde{Z}_{\widetilde{\varepsilon}}\right)_{\widetilde{\varepsilon}}$ *are formally orbitally equivalent then there exists* $\theta \in \mathbb{Z}/k\mathbb{Z}$ *such that* $\widetilde{\varepsilon} = \theta^*\varepsilon$*: the parameter is unique* modulo *this action and is called* canonical.

REMARK 4.2. – Although the original result is stated in [41] at an analytic level, the proof that ε becomes an invariant *modulo* the action of $\mathbb{Z}/k\mathbb{Z}$ stems from a formal computation and is therefore valid for formal orbital equivalence too. The idea of the proof is that the parameter completely determines the data of local eigenratios and vice versa, which are well-known orbital invariants.

From now on we only deal with unfoldings in prepared form (4.1) and only consider transforms fixing the canonical parameter ε.

4.2. Straightening weak separatrices

PROPOSITION 4.3 ([21, Proposition 2]). – *For any unfolding X in prepared form* (4.2) *there exists a formal power series*

$$\widehat{s} \in \mathbb{C}\,[[\varepsilon, x]]\, P$$

solving the parametric family of differential equations

$$P_\varepsilon(x)\,\frac{\mathrm{d}\widehat{s}_\varepsilon}{\mathrm{d}x}(x) = \widehat{s}_\varepsilon\left(1 + \mu_\varepsilon x^k\right) + A_\varepsilon\left(x, \widehat{s}_\varepsilon(x)\right). \tag{4.3}$$

Performing the transform $(\varepsilon, x, y) \mapsto (\varepsilon, x, y + \widehat{s}_\varepsilon(x))$ sends Z to a prepared unfolding

$$Z = \widehat{U}\left(\widehat{X} + \widehat{A}y\frac{\partial}{\partial y}\right) \tag{4.4}$$

for some formal *power series $\widehat{U}$ and $\widehat{A}$ in $\mathbb{C}\,[[\varepsilon, x, y]]$ with*

$$\widehat{U}(\varepsilon, x, 0) =: \widehat{U}_\varepsilon(x, 0) = u_\varepsilon(x) + \mathrm{O}\left(P_\varepsilon(x)\right)$$
$$\widehat{A}(\varepsilon, x, 0) =: \widehat{A}_\varepsilon(x, 0) = \mathrm{O}\left(P_\varepsilon(x)\right),$$

with $u \in \mathbb{C}\,[[\varepsilon]]\,[x]$ a polynomial (in x) of degree at most k such that $u_0(0) \neq 0$. In the particular case when Z is purely convergent the latter power series (and the coefficients of u) are convergent.

The proof of [21] is done for $k = 1$ but the general case is similar. It is based first on the following classical lemma, the proof of which is included for the sake of completeness.

LEMMA 4.4. – *Let $\widetilde{g} \in \mathbb{C}\,[[x, y]]$ and $\overline{g} \in x^p\mathbb{C}\,[[x]]$ be given with $p \in \mathbb{Z}_{\geq 0}$, such that either $p > 0$* or *$\widetilde{g}(x, y) = \mathrm{O}(x)$. Let $h \in \mathbb{C}\,[[x]]$ be such that $h(0) \neq 0$ and define $g(x, y) := \overline{g}(x) + y^2\widetilde{g}(x, y)$.*

The differential equation

$$x^{k+1} f'(x) + h(x)\, f(x) + g\left(x, f(x)\right) = 0 \tag{4.5}$$

has a unique formal solution f, which moreover belongs to $x^p\mathbb{C}\,[[x]]$.

REMARK 4.5. – Note that Equation (4.5) is nothing else than the differential equation determining the center manifold of the saddle-node vector field

$$x^{k+1}\frac{\partial}{\partial x} - \left(yh(x) + g(x, y)\right)\frac{\partial}{\partial y}$$

when g and h are holomorphic germs.

Proof. – Letting $C := h(0) \neq 0$ and $g(x, 0) = \overline{g}(x) =: \sum_{m=p}^{\infty} b^m x^m$; then substituting $f(x) =: \sum_{m=p}^{\infty} a^m x^m$ into (4.5) and grouping terms of same degree $m \geq p$, we get

$$Ca^m + b^m + F^m(a^p, \ldots, a^{m-1}) = 0$$

for some polynomial F^m depending on the m-jet of g and h. Hence, we can solve uniquely for each a^m. □

We then derive Proposition 4.3 from the following technical lemma which we will also use later on.

LEMMA 4.6. – (See [21] for the case $k = 1$.) *Let* $\widetilde{g}_\varepsilon \in \mathbb{C}[[\varepsilon, x, y]]$ *and* $\overline{g}_\varepsilon \in \mathbb{C}[[\varepsilon, x]]\, P$ *be given, i.e.,* $\overline{g}_\varepsilon(x) = P_\varepsilon(x) \sum_{|\mathbf{m}|\geq 0} \overline{g}^{\mathbf{m}}(x)\varepsilon^{\mathbf{m}}$ *where each* $\overline{g}^{\mathbf{m}}(x)$ *is itself a formal power series in* x, *and let* $h_\varepsilon \in \mathbb{C}[[\varepsilon, x]]$ *be such that* $h_0(0) \neq 0$. *Define* $g_\varepsilon(x, y) := \overline{g}_\varepsilon(x) + y^2\widetilde{g}_\varepsilon(x, y)$.

The family of differential equations

$$P_\varepsilon(x) f_\varepsilon'(x) + h_\varepsilon(x) f_\varepsilon(x) + g_\varepsilon(x, f_\varepsilon(x)) = 0 \tag{4.6}$$

has a unique formal solution f, *which moreover belongs to* $\mathbb{C}[[\varepsilon, x]]\, P$.

Proof. – Let $g_\varepsilon(x, y) =: P_\varepsilon(x) \sum_{|\mathbf{m}|\geq 0} b^{\mathbf{m},0}(x)\varepsilon^{\mathbf{m}} + \sum_{|\mathbf{m}|\geq 0} \left(\sum_{n\geq 2} b^{\mathbf{m},n}(x) y^n\right)\varepsilon^{\mathbf{m}}$. Substituting $f_\varepsilon(x) = P_\varepsilon(x) \sum_{|\mathbf{m}|\geq 0} a^{\mathbf{m}}(x)\varepsilon^{\mathbf{m}}$ into (4.6) and setting $\varepsilon := 0$ we first get

$$x^{k+1}\frac{\mathrm{d}a^{\mathbf{0}}}{\mathrm{d}x}(x) + \left(h^{\mathbf{0}}(x) + kx^k\right) a^{\mathbf{0}}(x) + x^{-(k+1)} g^{\mathbf{0}}\left(x, x^{k+1}a^{\mathbf{0}}(x)\right) = 0,$$

which admits a formal solution in $x^{k+1}\mathbb{C}[[x]]$ by direct application of Lemma 4.4 in the case $p = 0$. Likewise, by grouping terms with same $\varepsilon^{\mathbf{m}}$ for $|\mathbf{m}| \geq 1$ we obtain

$$x^{k+1}\frac{\mathrm{d}a^{\boldsymbol{m}}}{\mathrm{d}x}(x) + \left(h^{\mathbf{0}}(x) + kx^k + \ell^{\mathbf{m}}(x)\right) a^{\mathbf{m}}(x) + \left(b^{\mathbf{m},0}(x) + F^{\mathbf{m}}(x)\right) = 0, \tag{4.7}$$

where

$$\ell^{\mathbf{m}}(x) = \sum_{n\geq 2} n x^{(n-1)(k+1)} a^{\mathbf{0}}(x)^{n-1} b^{\mathbf{0},n}(x) = \mathrm{O}\left(x^{k+1}\right),$$

and $F^{\mathbf{m}} \in \mathbb{C}[[x]]$ is some formal power series depending polynomially on $(a^{\mathbf{n}}(x))_{|\mathbf{n}|<|\mathbf{m}|}$ and on the $|\mathbf{m}|$-jet of g and h. By induction on $|\mathbf{m}|$, we recursively find formal solutions $a^{\mathbf{m}} \in \mathbb{C}[[x]]$, for Equation (4.7) has the same type as (4.5) with $\widetilde{g} := 0$, and hence, has a formal solution given by Lemma 4.4. Uniqueness is straightforward. □

4.3. Normalization and cohomological equations

The tool for proving the Formal Normalization Theorem is the following.

PROPOSITION 4.7 ([46, 45]). – *Let* W *and* Y *be commuting, formal (resp. holomorphic) planar vector fields. Let* $F \in \mathbb{C}[[x, y]]$ *(resp. a germ of a holomorphic function) be given with* $F(0, 0) = 0$. *Then* $\Psi := \Phi_Y^F$ *is a formal (resp. analytic) change of variables near* $(0, 0)$ *and*

$$\Psi^* W = W - \frac{W \cdot F}{1 + Y \cdot F} Y.$$

This tool is used in the following manner.

— First if we could find a formal solution T of the (parametric families of) cohomological equations

$$X \cdot T = \frac{1}{U} - \frac{1}{u} \tag{4.8}$$

for a convenient choice of $u \in \mathbb{C}\{\varepsilon, x\}^\times$, then uX would be formally conjugate to Z by the tangential change of variables $\mathcal{T}$ given by

$$\mathcal{T} := \Phi_{uX}^T. \tag{4.9}$$

This is the content of the proposition for $Y := W := uX$ and $F := T$.

— From Proposition 4.3 we built the formal, fibered transform $\mathcal{S}$ given by

$$\mathcal{S}\,:\,(x,y)\longmapsto(x,y-\widehat{s}(x))$$

such that $\mathcal{S}^*\left(\widehat{\mathrm{X}}+\widehat{A}y\frac{\partial}{\partial y}\right)=X$.

— Finally, since $y\frac{\partial}{\partial y}$ commutes with the normal form $\widehat{\mathrm{X}}$, if we could solve formally the cohomological equation

$$\left(\widehat{\mathrm{X}}+\widehat{A}y\frac{\partial}{\partial y}\right)\cdot O=-\widehat{A} \tag{4.10}$$

then $\widehat{\mathrm{X}}$ would be formally conjugate to $\widehat{\mathrm{X}}+\widehat{A}y\frac{\partial}{\partial y}$ by the fibered, transverse change of variables $\mathcal{O}$ given by

$$\mathcal{O}:=\Phi^{O}_{y\frac{\partial}{\partial y}}\,:\,(x,y)\longmapsto(x,y\exp O(x,y))\,. \tag{4.11}$$

We explain below how those formal power series are built and to which extent they are unique. We consequently obtain a formal conjugacy $\mathcal{O}\circ\mathcal{S}\circ\mathcal{T}$ between $\widehat{\mathrm{Z}}$ and Z (notice that u is left invariant by the fibered $\mathcal{O}\circ\mathcal{S}$, so that it also conjugates $\widehat{\mathrm{Z}}$ to uX).

Lemma 4.8. – *Let X be in the form* (4.2) *for $A\in\mathbb{C}[[\varepsilon,x,y]]$, and take $G\in\mathbb{C}[[\varepsilon,x,y]]$. There exists a formal solution $F\in\mathbb{C}[[\varepsilon,x,y]]$ of the cohomological equation*

$$X\cdot F=G \tag{4.12}$$

if, and only if, $G(x,0)$ belongs to the ideal generated by P. In that case F is unique up to the free choice of $F(0,0)\in\mathbb{C}[[\varepsilon]]$.

Proof. – Let

$$F(x,y)=:\sum_{n\geq0}F^n(x)\,y^n\quad\text{and}\quad G(x,y)=:\sum_{n\geq0}G^n(x)\,y^n.$$

We proceed by induction on $n\geq0$ by identifying coefficients of powers of y in (4.12). For each $n\in\mathbb{Z}_{\geq0}$ we must therefore solve

$$P\frac{\partial F^n}{\partial x}+n\left(1+\mu x^k\right)F^n=G^n+\mathrm{o}(n)\,, \tag{4.13}$$

where $\mathrm{o}(n)$ stands for terms containing F^m for $m<n$ only, and are thus already known.

— The case $n=0$ outlines the formal obstruction (notice that the choice of $F^0(0)$ is free).

— For $n>0$ no additional obstruction appears and F^n is uniquely determined. Then Lemma 4.6 provides the unique formal solution of the family of differential Equations (4.13). □

We finally derive the Formal Normalization Theorem by writing

$$U(x,0)=u(x)+\mathrm{O}(P(x))\,,\quad u\in\mathbb{C}\{\varepsilon\}[x]^{\times}_{\leq k}\,,$$

and finding a (unique with $T(0,0)=0$) formal solution T of (4.8) by Lemma 4.8. As for the power series O, a (unique with $O(0,0)=0$) formal solution of (4.10) exists by Proposition 4.3, and Lemma 4.8, for X given in (4.2).

DEFINITION 4.9. – Let Z be an unfolding in prepared form (4.1). We write $\mathcal{N} := \mathcal{O} \circ \mathcal{S} \circ \mathcal{T}$ the canonical *formal normalization* of Z satisfying $\mathcal{N}^* \widehat{\mathrm{Z}} = Z$ where $\mathcal{O}$, $\mathcal{S}$ and $\mathcal{T}$ are built above.

4.4. Uniqueness

Addressing the uniqueness clause in the Formal Normalization Theorem boils down to studying the case of the normal forms, because of the canonical choice of normalization maps $\mathcal{N}$ done in Definition 4.9.

LEMMA 4.10. – *Let Ψ be a formal orbital symmetry of the formal normal form $\widehat{\mathrm{Z}}$ (fixing the canonical parameter).*

1. *There exist unique $F \in \mathbb{C}[[\varepsilon, x, y]]$ and $c \in \mathbb{C}[[\varepsilon]]^\times$ such that*
$$\Psi = \left(c^* \mathrm{Id}\right) \circ \Phi^F_{\widehat{\mathrm{Z}}}$$
where $c^*\mathrm{Id}$ is the linear mapping $(x, y) \mapsto (x, cy)$. (The converse statement clearly holds.)
2. *Ψ is a symmetry of $\widehat{\mathrm{Z}}$ if, and only if, $F \in \mathbb{C}[[\varepsilon]]$.*
3. *Ψ is fibered if, and only if, $F = 0$.*

Proof. – 1. By Remark 3.6 we want to determine $V \in \mathbb{C}[[\varepsilon, x, y]]^\times$ such that $\Psi^* \widehat{\mathrm{Z}} = V \widehat{\mathrm{Z}}$. Because ε is a formal invariant governing the eigenvalues of (the differential of) the vector fields at the singularities, Ψ cannot change the eigenvalues, so that $V(x, y) = 1 + \mathrm{O}(P(x)) + \mathrm{O}(y)$. According to Lemma 4.8 there exists a (unique) formal solution F with $F(0, 0) = 0$ to the cohomological equation
$$\widehat{\mathrm{X}} \cdot F = \frac{1}{uV} - \frac{1}{u}.$$
Therefore $\widehat{\Psi} := \Psi \circ \left(\Phi^F_{\widehat{Z}}\right)^{\circ -1}$ induces a symmetry of $\widehat{\mathrm{Z}}$.

Write $\widehat{\Psi} : (\varepsilon, x, y) \mapsto (\varepsilon, \phi_\varepsilon(x, y), \psi_\varepsilon(x, y))$. By considering the $\frac{\partial}{\partial x}$-component of $\widehat{\mathrm{Z}}$ one obtains the relation
$$(uP) \circ \phi = \widehat{\mathrm{Z}} \cdot \phi.$$
Setting $y := 0$ yields
$$(u_\varepsilon P_\varepsilon) \circ \phi_\varepsilon(x, 0) = u_\varepsilon(x) P_\varepsilon(x) \frac{\partial \phi_\varepsilon}{\partial x}(x, 0)$$
so that
$$\phi_\varepsilon(x, 0) = \Phi^{t_\varepsilon}_{u_\varepsilon P_\varepsilon \frac{\partial}{\partial x}}(x) = \Phi^{t_\varepsilon}_{\widehat{\mathrm{Z}}}(x, 0)$$
for some $t \in \mathbb{C}[[\varepsilon]]$. Hence we may assume without loss of generality that $F_\varepsilon(0, 0) = t_\varepsilon$ and $\phi_\varepsilon(x, 0) = x$. Writing $\phi_\varepsilon(x, y) = x + \sum_{n \geq \nu} \phi^n_\varepsilon(x) y^n$ with $\nu > 0$ we obtain for the term of y-degree ν
$$P' \phi^\nu = P \frac{\partial \phi^\nu}{\partial x} + \nu \left(1 + \mu x^k\right) \phi^\nu,$$

whose unique formal solution is $\phi^{\nu} = 0$, since it is the equation of the weak separatrix of $P\frac{\partial}{\partial x}+y\left(\nu(1+\mu x^{k})+P'\right)\frac{\partial}{\partial y}$. As a matter of consequence $\phi_{\varepsilon}(x,y)=x$ and $\widehat{\Psi}$ is fibered. Lastly, by considering the $\frac{\partial}{\partial y}$-component of $\widehat{Z}$ one obtains the relation

$$\left(1+\mu x^{k}\right)\psi = \widehat{X}\cdot\psi.$$

Setting $y := 0$ yields

$$\psi_{\varepsilon}(x,0)=0$$

so that $\psi_{\varepsilon}(x,y) = y\exp N_{\varepsilon}(x,y)$ for some $N \in \mathbb{C}[[\varepsilon,x,y]]$. The corresponding cohomological equation reads

$$0=\widehat{X}\cdot N$$

and only admits $N \in \mathbb{C}[[\varepsilon]]$ as formal solution (uniqueness clause of Lemma 4.8). We then set $c := \exp N$. (2) and (3) are clear from the previous arguments. □

We derive the following precise statement. Item (2) plays an essential role in proving the (analytic) Uniqueness Theorem.

Corollary 4.11. – *Consider two unfoldings Z and $\widetilde{Z}$ in prepared form* (4.1).

1. *Let Ψ be a formal conjugacy between Z and $\widetilde{Z}$ (fixing the canonical parameter), namely $\Psi^{*}Z=\widetilde{Z}$. Let $\mathscr{N}=\mathscr{O}\circ\mathscr{T}\circ\mathscr{S}$ and $\widetilde{\mathscr{N}}=\widetilde{\mathscr{O}}\circ\widetilde{\mathscr{T}}\circ\widetilde{\mathscr{S}}$ be the respective canonical tangent-to-identity formal normalizations as in Definition 4.9.*
   (a) *There exists unique $c\in\mathbb{C}[[\varepsilon]]^{\times}$ and $t\in\mathbb{C}[[\varepsilon]]$ such that*
   $$\Psi=\mathscr{N}^{\circ-1}\circ\left(c^{*}\mathrm{Id}\right)\circ\widetilde{\mathscr{N}}\circ\Phi^{t}_{\widetilde{Z}}.$$
   (The converse statement clearly holds.)
   (b) *If Ψ is analytic then so are t and c. (The converse statement does not generally hold.)*
2. *If Z and $\widetilde{Z}$ are analytically orbitally equivalent (by an orbital equivalence fixing the canonical parameter) then there exists a fibered analytic orbital equivalence between them*
   $$\mathscr{S}^{\circ-1}\circ\mathscr{O}^{\circ-1}\circ\left(c^{*}\mathrm{Id}\right)\circ\widetilde{\mathscr{O}}\circ\widetilde{\mathscr{S}}$$
   for some $c\in\mathbb{C}\{\varepsilon\}$.

Remark 4.12. – The partial conclusion "there exists a fibered orbital equivalence" in Claim (2) was proved in [41, Lemma 3.4] by unfolding the homotopy technique of [31, Lemma 2.2.2]. We give here an alternate proof. In the other part of the conclusion, pay attention that $\mathscr{O}\circ\mathscr{S}$ and $\widetilde{\mathscr{O}}\circ\widetilde{\mathscr{S}}$ are only formal power series, but the composition is a convergent power series.

Proof. – 1. (a) follows from Lemma 4.10: the formal map $\mathscr{N}\circ\Psi\circ\widetilde{\mathscr{N}}^{\circ-1}$ is a symmetry of the normal form $\widehat{Z}$, and $\widetilde{\mathscr{N}}$ is a formal conjugacy between $\widetilde{Z}$ and $\widehat{Z}$, hence conjugating their flow (as formal power series):

$$\Phi^{t}_{\widehat{Z}}\circ\widetilde{\mathscr{N}}=\widetilde{\mathscr{N}}\circ\Phi^{t}_{\widetilde{Z}}.$$

1. (b) Here we assume that Ψ is analytic. Following (a) we have

$$\Psi = \mathscr{T}^{\circ -1} \circ \left(\mathscr{S}^{\circ -1} \circ \mathscr{O}^{\circ -1} \circ \left(c^* \mathrm{Id} \right) \circ \widetilde{\mathscr{O}} \circ \widetilde{\mathscr{S}} \circ \Phi^t_{u\widetilde{X}} \right) \circ \widetilde{\mathscr{T}}.$$

Using both facts that $\widehat{\Psi} := \mathscr{S}^{\circ -1} \circ \mathscr{O}^{\circ -1} \circ (c^* \mathrm{Id}) \circ \widetilde{\mathscr{O}} \circ \widetilde{\mathscr{S}}$ is fibered, and that

$$\Phi^t_{u\widetilde{X}} (x, y) = \left(\Phi^t_{uP\frac{\partial}{\partial x}} (x), y \left(\exp t + y\phi (x, y, t) \right) \right),$$

we first derive

$$\Psi = \mathscr{T}^{\circ -1} \circ \left(\Phi^t_{uP\frac{\partial}{\partial x}}, \psi \right) \circ \widetilde{\mathscr{T}}.$$

Because $T(0,0) = \widetilde{T}(0,0) = 0$, we have

$$\Psi_\varepsilon (0,0) = (t_\varepsilon \varepsilon_0, \ldots)$$

from which we deduce the convergence of t. We also have the identity

$$\frac{\partial \psi}{\partial y} (0,0) = c \exp t,$$

from which the convergence of c follows also.

2. It is sufficient to assume that $Z := X$ is analytically conjugate by some Ψ (fixing the canonical parameter) to $\widetilde{Z} := \widetilde{U}\widetilde{X}$ for some $\widetilde{U} \in \mathbb{C}\{\varepsilon, x, y\}^\times$. In that setting we have $u = 1$ and $\mathscr{T} = \mathrm{Id}$, so that according to (1)

$$\Psi = \mathscr{S}^{\circ -1} \circ \mathscr{O}^{\circ -1} \circ \left(c^* \mathrm{Id} \right) \circ \widetilde{\mathscr{O}} \circ \widetilde{\mathscr{S}} \circ \widetilde{\mathscr{T}} \circ \Phi^t_{\widetilde{Z}},$$

where $t \in \mathbb{C}\{\varepsilon\}$. As a matter of consequence the mapping $\Phi^t_{\widetilde{Z}}$ is analytic, and so is $\widehat{\Psi} := \Psi \circ \left(\Phi^t_{\widetilde{X}} \right)^{\circ -1}$. Because $\widehat{\Psi} \circ \widetilde{\mathscr{T}}^{\circ -1}$ is fibered, the x-component of $\widehat{\Psi}$ (which is analytic) is equal to the x-component of $\widetilde{\mathscr{T}}$. The former is of the form $(x, y) \mapsto A\left(x, y, \widetilde{T}(x, y)\right)$ for some holomorphic function $A \in \mathbb{C}\{\varepsilon, x, y, t\}$ with $\frac{\partial A}{\partial t} \neq 0$, and where $\widetilde{T}$ is the solution of (4.8) for $U := \widetilde{U}$. Thus $\widetilde{T}$ is a convergent power series, and so is $\widehat{\Psi} \circ \widetilde{\mathscr{T}}^{\circ -1}$. □

5. Geometric orbital normalization

Here we prove the orbital part of the Normalization and Uniqueness Theorems for $\tau = 0$. Sections 5.2–5.5 are devoted to the construction of the normal form conjugacy, while its uniqueness is thoroughly studied in Section 5.6. Before going into the details we start with a brief description of the general strategy. Let us call $\mathbb{D}$ the unit disk.

For fixed

$$0 < \frac{1}{\rho^\infty} < \rho^0$$

we introduce two analytic charts:

— the original coordinates

$$(x, y) \in \mathscr{U}^0 := \rho^0 \mathbb{D} \times (\mathbb{C}, 0),$$

— the coordinates at infinity

$$(u,v)\in\mathscr{U}^{\infty}:=\rho^{\infty}\mathbb{D}\times(\mathbb{C},0)$$

with (involutive) standard transition map on $\mathbb{C}^{\times}\times(\mathbb{C},0)$

$$(u,v)\longmapsto\left(\frac{1}{u},v\right)=(x,y)\,.$$

For convenience we write O^0 and O^∞ the respective expressions of a holomorphic object O in the charts $\mathscr{U}^0$ and $\mathscr{U}^\infty$ respectively.

Start from an arbitrary $X^0\in\text{Convergent}_k$ in prepared form (4.2), with $A\in y^2\mathbb{C}\{\varepsilon,x,y\}$ holomorphic and bounded on $\mathscr{U}^0$, and such that $\mu_0\notin\mathbb{R}_{\leq 0}$. It is always possible to make this assumption thanks to Theorem 4.1, since $a=0$ in that case. Notice in particular that A is bounded since we can always take a smaller ρ^0 and decrease similarly the size of the neighborhood of $\{y=0\}$: hence $\mathscr{U}^0$ can be taken inside a compact set on which A is defined.

In the following we assume that ε is so small that the $k+1$ singularities of X^0_ε lie in $\{0\leq|x|<{}^1\!/\!_{\rho^\infty}\}\times\{0\}$. The following steps constitute what we refer to as the *unfolded Loray construction*.

Gluing. – We find a vector field family X^∞ on $\mathscr{U}^\infty$ whose holonomy over $\rho^\infty\mathbb{S}^1\times\{0\}$ is the inverse of $\mathfrak{h}_\varepsilon$, the corresponding "weak" holonomy of X^0 (Section 5.2). Therefore foliations induced by each vector field can be glued one to the other over the annulus $\mathscr{U}^0\cap\mathscr{U}^\infty$ by an identification of the form

$$(u,v)=\left(\frac{1}{x},y\exp\phi\left(x,y\right)\right)$$

(Section 5.3). This operation results in a family of foliated abstract complex surfaces $(\mathscr{M},\mathscr{F})$.

Normalizing. – We construct a fibered biholomorphic equivalence between $\mathscr{M}$ and a standard neighborhood of $\{y=0\}\simeq\mathbb{P}_1(\mathbb{C})$, that is a complex surface with charts $\mathscr{U}^0$, $\mathscr{U}^\infty$ and transition map *exactly* $(u,v)=\left(\frac{1}{x},y\right)$ (Section 5.4). Because $\mathbb{P}_1(\mathbb{C})\times\{0\}$ is compact the expression of the new X^0 is polynomial in x with controlled degree, thus in orbital normal form (2.4) as expected by the Normalization Theorem (Section 5.5).

Uniqueness. – From the special form of the normalized vector field, it can be seen that the closure of the saturation of any small neighborhood of $(0,0)$ contains a whole $\mathbb{P}_1(\mathbb{C})\times r\mathbb{D}$. Therefore any local conjugacy between normal forms (which we choose fibered thanks to Corollary 4.11 (2)) can be analytically continued by a construction *à la* Mattei-Moussu on $\mathbb{P}_1(\mathbb{C})\times r\mathbb{D}$. But this manifold has very few fibered automorphisms, allowing to conclude (Section 5.6).

In the unfolded Loray construction, only what happens in the first chart (x,y) is of a different nature than when $\varepsilon=0$. As seen from the other chart (u,v), the only important ingredient for the construction is the "weak" holonomy $\mathfrak{h}_\varepsilon$ of the unfolding (see Section 5.1 below). Hence the original arguments do not need to be unfolded near $(\infty,0)$, although we must take care that everything remains holomorphic in the parameter. The first two steps of the unfolded Loray construction require external results that need to be parametrically controlled:

1. the realization of the weak holonomy $\mathfrak{h}$ by a foliation near $(\infty, 0)$, obtained by the construction of [43];
2. the normalization of the transition map between the charts (x, y) and (u, v) on the annulus

$$\mathscr{A} := \{1/\rho^0 < |u| < \rho^\infty\} \times (\mathbb{C}, 0)\,,$$

as done in [42].

Both proofs are similar in spirit and only rely on complex (holomorphic) analysis and (what amounts to) a fixed-point method. Parametric holomorphy follows from the explicit integral formulas. Because normalizing transition maps is relatively easy, we prove a parametric version of Savelev's Theorem in Section 5.4. It contains the main steps and ideas upon which are based the respective proofs of the Normalization Theorem for vector fields (Section 6) and of the Realization Theorem (Section 7). The latter is nothing but an unfolded version of the main result of [43], retrospectively making the present article more self-contained.

5.1. Weak holonomy

We name

$$\Pi : \ (x, y) \longmapsto x$$

the projection on the invariant line $\{y = 0\}$ and let

$$\Sigma \subset \Pi^{-1}(x_*)$$

be a germ of a transverse disk endowed with the coordinate $y \in (\mathbb{C}, 0)$. Starting from $y \in \Sigma$ there exists a unique path

$$\begin{aligned} \gamma_y : \ [0, 1] &\longrightarrow \mathscr{U}^0 \\ \gamma_y(0) &= (x_*, y) \end{aligned}$$

tangent to X_ε^0 such that

$$\gamma := \Pi \circ \gamma_y \ = \ s \longmapsto x_* \exp 2\mathrm{i}\pi s.$$

We define $\mathfrak{h}_\varepsilon(y)$ as the y-component of the final value $\gamma_y(1)$. The *weak holonomy mapping* $\mathfrak{h}_\varepsilon$ as described is a germ of a biholomorphism near the fixed-point 0 whose linear part is governed by the formal orbital invariant μ in the following way:

$$\mathfrak{h}_\varepsilon(y) = y \exp 2\mathrm{i}\pi\mu_\varepsilon + \mathrm{o}(y) \in \mathrm{Diff}(\Sigma, 0)\,.$$

The analytic dependence of trajectories of X_ε^0 on the parameter ε ensures that $\mathfrak{h} \in \mathbb{C}\{\varepsilon, y\}$.

5.2. Parametric holonomy realization at $(\infty, 0)$

THEOREM 5.1 ([43, Main theorem and Section 4.4]). – *Let $(\Delta_\eta)_{\eta\in(\mathbb{C}^n,0)}$ be an analytic family of elements of* $\mathrm{Diff}(\mathbb{C},0)$*, that is $(\eta, v) \mapsto \Delta_\eta(v) \in \mathbb{C}\{\eta, v\}$ and $\Delta_0'(0) \neq 0$. Let $\lambda \in \mathbb{C}\{\eta\}$ be given such that $\Delta_\eta'(0) = \exp(-2i\pi\lambda_\eta)$ and $\lambda_0 \notin \mathbb{R}_{\leq 0}$. There exists an analytic family of vector fields $(X_\eta^\infty)_{\eta\in(\mathbb{C}^n,0)}$ of the form*

$$X_\eta^\infty(u,v) = -u\frac{\partial}{\partial u} + v\left(\lambda_\eta + u\left(1 + f_\eta(v)\right) + g_\eta(v)\right)\frac{\partial}{\partial v}, \quad f, g \in v\mathbb{C}\{\eta, v\}, \tag{5.1}$$

holomorphic on the domain $\mathcal{U}^\infty$ and satisfying for all $\eta \in (\mathbb{C}^n, 0)$:

1. $(0,0)$ *is the only singularity of X_η^∞ in $\mathcal{U}^\infty$,*
2. *the holonomy of X_η^∞ above the circle $w_*\mathbb{S}^1 \times \{0\}$, computed on a germ of transverse disk $\{u = u_*\}$ with respect to the projection $(u,v) \mapsto u$, is exactly Δ_η.*

REMARK 5.2. – The result of [43] asserts the existence of a vector field of the form (5.1) with $f := 0$ whose holonomy on Σ is conjugate to Δ by some analytic family Ψ. The conjugacy $(u,v) \mapsto (u, \Psi(v))$ transforms the vector field into the form (5.1) for different f, g but its holonomy is exactly Δ on Σ.

In the generic case $\lambda_0 \notin \mathbb{R}$ the theorem is (almost) trivial. All holonomy maps

$$\Delta_\eta : v \longmapsto v\exp\left(-2i\pi\lambda_\eta + \delta_\eta(v)\right), \quad \delta_\eta(0) = 0,$$

are hyperbolic and locally analytically linearizable for that matter (Koenig's theorem), the unique tangent-to-identity linearization being given by $\Psi_\eta : v \mapsto v\exp\psi_\eta(v)$, where

$$\psi_\eta := \sum_{n=0}^{\infty} \delta_\eta \circ \Delta_\eta^{\circ n}.$$

Local uniform convergence ensures that ψ is analytic in both t and η. The fibered mapping $(u,v) \mapsto (u, \Psi_\eta(v))$ transforms the linear vector field $-u\frac{\partial}{\partial u} + \lambda v\frac{\partial}{\partial v}$ into a vector field X_η^∞ fulfilling the conclusions (1)-(2) of the theorem (but not of the form (5.1)). However if $\lambda_0 \in \mathbb{R}$ this construction fails: the linearization domain may shrink to a point (if Δ_0 is not analytically linearizable). The form (5.1) has the advantage of being valid for all cases, analytically in the parameter. Notice that the presence of the term $-uv\frac{\partial}{\partial v}$ in (5.1) discards any linear realization even when $\lambda_0 \notin \mathbb{R}$.

We define $\eta := \varepsilon \in (\mathbb{C}^k, 0)$,

$$\lambda_\varepsilon := \mu_\varepsilon \notin \mathbb{R}_{\leq 0},$$
$$\Delta_\varepsilon : v \longmapsto \mathfrak{h}_\varepsilon^{\circ -1}(v),$$

and apply Theorem 5.1, to obtain an analytic family X^∞ in the chart (u,v). In order to stitch the induced foliation with that of X_ε^0 we need to prepare it by changing slightly the coordinates on $\mathcal{U}^{\infty}$. Let $\widetilde{X}_{\varepsilon}^{0}$ be the vector field corresponding to X_{ε}^{0} in the coordinates $(u,v)=\left(\frac{1}{x},y\right)$, that is

$$\begin{aligned}\widetilde{X}_{\varepsilon}^{0}(u,v) &= -u^{2}P_{\varepsilon}\left(\frac{1}{u}\right)\frac{\partial}{\partial u}+v\left(\mu_{\varepsilon}u^{-k}+1+\mathrm{O}(v)\right)\frac{\partial}{\partial v}\\ &= uP_{\varepsilon}\left(\frac{1}{u}\right)\times\left(-u\frac{\partial}{\partial u}+v\left(\lambda_{\varepsilon}+h_{\varepsilon}(u)+\mathrm{O}(v)\right)\frac{\partial}{\partial v}\right),\end{aligned}$$

where

$$h_{\varepsilon}:\ u\longmapsto\frac{u^{k}+\mu_{\varepsilon}}{u^{k+1}P_{\varepsilon}\left(\frac{1}{u}\right)}-\mu_{\varepsilon}\quad\in\mathrm{Holo}\left(\left(\mathbb{C}^{k},0\right)\times\rho^{\infty}\mathbb{D}\right) \tag{5.2}$$

vanishes at 0. Notice indeed that the polynomial $u^{k+1}P_{\varepsilon}\left(\frac{1}{u}\right)\in\mathbb{C}[u]_{\leq k+1}$ has its roots outside the closed disk $\mathrm{cl}(\rho^{\infty}\mathbb{D})$, whereas it takes the value 1 at 0. Remark also that the quantity $uP_{\varepsilon}\left(\frac{1}{u}\right)$ needs to be factored out in order to recognize a vector field alike to X_{ε}^{∞} near $(\infty,0)$. This function is non-vanishing on the annulus $\mathcal{A}$. Let $\widetilde{X}_{\varepsilon}^{\infty}$ be the vector field corresponding to X_{ε}^{∞} through the inverse transform

$$(u,v)\longmapsto\left(u,v\exp\int_{u_{*}}^{u}\left(h_{\varepsilon}(z)-z\right)\frac{\mathrm{d}z}{z}\right). \tag{5.3}$$

By construction we have

$$\widetilde{X}_{\varepsilon}^{\infty}(u,v)=-u\frac{\partial}{\partial u}+v\left(\lambda_{\varepsilon}+h_{\varepsilon}(u)+\mathrm{O}(v)\right)\frac{\partial}{\partial v},$$

which glues with $\widetilde{X}_{\varepsilon}^{0}$ through $(u,v)=\left(\frac{1}{x},y\right)$ as presented in the next paragraph.

5.3. Gluing

Both transformed vector fields $\widetilde{X}^{0}$ and $\widetilde{X}^{\infty}$ built in the previous section have same holonomy Δ_{ε} on Σ. We glue the (foliations defined by the) vector fields $\widetilde{X}_{\varepsilon}^{0}$ and $\widetilde{X}_{\varepsilon}^{\infty}$ over the fibered annulus $\mathcal{A}$ through a fibered map Φ_{ε} fixing Σ and (classically) obtained by foliated path-lifting, as we explain now. For $(u,v)\in\mathcal{A}$ we join u_{*} to u in $\mathcal{A}\cap\{v=0\}$ by some path γ and define

$$\Phi_{\varepsilon}(u,v):=\left(u,\mathfrak{h}_{\varepsilon,\gamma}^{\infty}\circ\left(\mathfrak{h}_{\varepsilon,\gamma}^{0}\right)^{\circ-1}(v)\right),$$

where $\mathfrak{h}_{\varepsilon,\gamma}^{0}$ (resp. $\mathfrak{h}_{\varepsilon,\gamma}^{\infty}$) is the holonomy map obtained by lifting the path γ through Π in the foliation induced by $\widetilde{X}_{\varepsilon}^{0}$ (resp. $\widetilde{X}_{\varepsilon}^{\infty}$). The map Φ_{ε} is well-defined because when γ is a loop both mappings $\mathfrak{h}_{\varepsilon,\gamma}^{\infty}$ and $\mathfrak{h}_{\varepsilon,\gamma}^{0}$ coincide with the same corresponding iterate of Δ_{ε}. Clearly Φ_{ε} depends analytically on $\varepsilon\in\left(\mathbb{C}^{k},0\right)$ and is a germ of a fibered biholomorphism near $\mathcal{A}\cap\{v=0\}$ satisfying

$$\begin{aligned}\Phi^{*}\widetilde{X}^{0} &= uP\left(\frac{1}{u}\right)\widetilde{X}^{\infty},\\ \Phi(u,v) &= \left(u,v\exp\phi(u,v)\right),\\ \phi(u,0) &= \phi(u_{*},v)=0.\end{aligned}$$

5.4. Normalizing

So far the construction yields an analytic family of complex foliated surfaces, written $(\mathcal{M}, \mathcal{F})$, defined by the charts $\left(\mathcal{U}^0, \mathcal{F}^0\right)$ and $\left(\mathcal{U}^\infty, \mathcal{F}^\infty\right)$ with transition map

$$(u, v) \longmapsto \left(\frac{1}{u}, v \exp \phi\,(u, v)\right) = (x, y)\,. \tag{5.4}$$

REMARK 5.3. – The foliation $\mathcal{F}_\varepsilon$ is transverse to the fibers of the global fibration by disks $\Pi : \mathcal{M}_\varepsilon \to \mathcal{L}$ given in the first chart by $(x, y) \mapsto x$, except along the $k+2$ invariant disks $\{P_\varepsilon (x) = 0\}$ and $\{x = \infty\}$.

Each manifold $\mathcal{M}_\varepsilon$ is a neighborhood of the invariant divisor $\mathcal{L} \simeq \mathbb{P}_1 (\mathbb{C})$, corresponding to $\{y = 0\}$ and $\{v = 0\}$ in the respective chart, while the natural embedding $\mathbb{P}_1 (\mathbb{C}) \hookrightarrow \mathcal{M}$ has self-intersection 0 according to Camacho-Sad index formula [4] (the singularities near $(0, 0)$ contribute for a sum of Camacho-Sad indices equal to μ_ε while the singularity at $(\infty, 0)$ does for $-\lambda_\varepsilon = -\mu_\varepsilon$).

DEFINITION 5.4. – For $r > 0$ we define the *standard neighborhood of radius r of the Riemann sphere*

$$\mathrm{Sphere}\,(r) := \mathbb{P}_1 (\mathbb{C}) \times r\mathbb{D},$$

the complex surface equipped with the (global) affine coordinates

$$(u, v) \in \mathbb{C} \times r\mathbb{D}$$

and transition map on $\mathbb{C}^\times \times r\mathbb{D}$ given by $(u, v) = \left(\frac{1}{x}, y\right)$, i.e., by (5.4) with $\phi := 0$. The other chart of $\mathrm{Sphere}\,(r)$ is the domain $(x, y) \in \mathbb{C} \times r\mathbb{D}$. When speaking of a *standard neighborhood of the sphere* we actually refer to $\mathrm{Sphere}\,(r)$ for some $r > 0$ small enough.

THEOREM 5.5. – *Let $\mathcal{M}$ be an analytic family of complex surfaces with transition maps* (5.4)*. There exists a standard neighborhood $\mathcal{V} = \mathrm{Sphere}\,(r)$ of $\mathcal{L}$, for some $r > 0$, and an analytic family of fibered holomorphic injective mappings*

$$\Psi : \mathcal{V} \longrightarrow \mathcal{M}$$

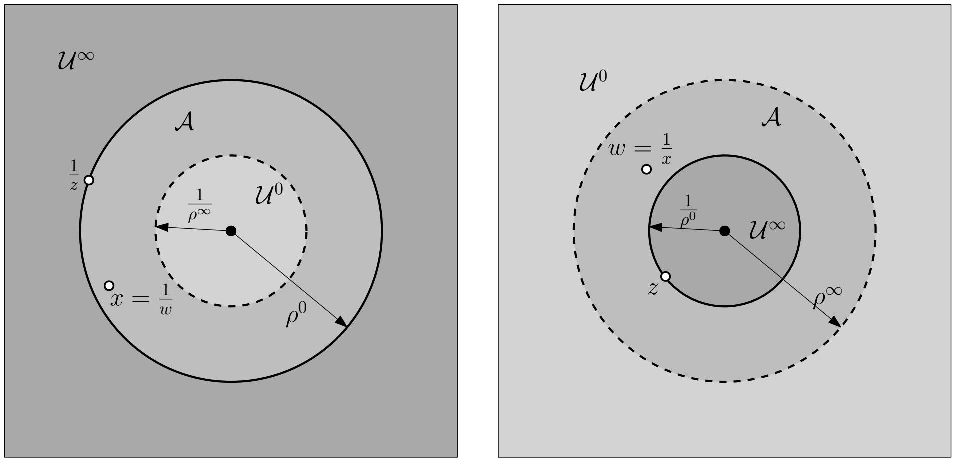
agreeing with the identity on $\mathcal{L}$.

The rest of the subsection is devoted to the proof of this theorem. We refer to Section (3.1.2) for the definitions of the functional spaces in use. We are looking for Ψ, or rather its expression in the charts $\mathcal{U}^0$ and $\mathcal{U}^\infty$, in the form

$$\begin{aligned}\Psi^0 (x, y) &= \left(x, y \exp \psi^0 (x, y)\right)\\ \Psi^\infty (u, v) &= (u, v \exp \psi^\infty (u, v))\,.\end{aligned}$$

The normalization equation becomes a non-linear additive Cousin problem on $\mathcal{A}$:

$$\psi^0 \left(\frac{1}{u}, v\right) - \psi^\infty (u, v) = \phi \circ \Psi^\infty (u, v)\,.$$



(A) For F^0 in (x, y) coordinates.

(B) For F^∞ in (u, v) coordinates.

FIGURE 5.1. Integration contours in the respective charts.

Starting from $\psi_0^0 := 0$ and $\psi_0^\infty := 0$ we build iteratively two bounded sequences of holomorphic functions

$$\psi_n^\sharp \in \mathrm{Holo_c}\left(\left(\mathbb{C}^k, 0\right) \times \rho^\sharp\mathbb{D} \times r\mathbb{D}\right), \quad \sharp \in \{0, \infty\}$$

solution of the linearized additive Cousin problem (or discrete cohomological equation)

$$\psi_{n+1}^0\left(\frac{1}{u}, v\right) - \psi_{n+1}^\infty(u, v) = \phi\left(u, v \exp \psi_n^\infty(u, v)\right). \tag{5.5}$$

The Cousin problem has explicit solutions given by a Cauchy-Heine transform. From these solutions we obtain a priori bounds on the norm of $\psi_n^\sharp$, allowing to fix the radius $r > 0$ beforehand. We let

$$\mathcal{U}_r^0 := \left\{(x, y) : |x| < \rho^0, |y| < r\right\},$$
$$\mathcal{U}_r^\infty := \{(u, v) : |u| < \rho^\infty, |v| < r\},$$

be an atlas for Sphere (r). We postpone the proof of the next main lemma to the end of the section.

LEMMA 5.6. – *Assume that* $\phi \in \mathrm{Holo_c}\left(\mathscr{A}_\eta\right)'$ *for some domain* $\mathscr{A}_\eta := \left\{\frac{1}{\rho^0} < |u| < \rho^\infty\right\} \times \eta\mathbb{D}$. *Let* $\psi \in \mathrm{Holo_c}\left(\mathcal{U}_r^\infty\right)$ *be such that the image of* $\mathscr{A}_r$ *by* $(u, v) \mapsto (u, v \exp \psi(u, v))$ *is included in* $\mathscr{A}_\eta$. *Define*

$$\begin{cases} F^\infty(u, v) := \frac{1}{2i\pi} \oint_{\rho^\infty \mathbb{S}^1} \phi(z, v \exp \psi(z, v)) \frac{\mathrm{d}z}{z-u}, \\ F^0(x, y) := \frac{x}{2i\pi} \oint_{\frac{1}{\rho^0}\mathbb{S}^1} \phi(z, y \exp \psi(z, y)) \frac{\mathrm{d}z}{xz-1}. \end{cases} \tag{5.6}$$

Then the following properties hold.

1. $F^0 \in \mathrm{Holo_c}\left(\mathscr{U}_r^0\right)$ *and* $F^\infty \in \mathrm{Holo_c}\left(\mathscr{U}_r^\infty\right)$. *Moreover for* $\sharp \in \{0, \infty\}$

$$\left\|F^\sharp\right\|_{\mathscr{U}_r^\sharp} \leq rK \left\|\phi\right\|'_{\mathscr{A}_\eta} \exp \left\|\psi\right\|_{\mathscr{U}_r^\infty}, \tag{5.7}$$

where

$$K := \left(1 + \frac{2\rho^0}{\rho^\infty\rho^0 - 1}\right).$$

2. *For all* $(u, v) \in \mathscr{A}_r$ *we have*

$$F^0\left(\frac{1}{u}, v\right) - F^\infty(u, v) = \phi\left(u, v \exp \psi(u, v)\right). \tag{5.8}$$

3. *These are the only holomorphic solutions of the previous equation which are bounded, up to the addition of a function* $v \mapsto f^\infty(v)$ *with* $f^\infty \in \mathrm{Holo_c}(r\mathbb{D})$.

REMARK 5.7. – The integral Formula (5.6) shows right away that $F^\sharp$ depends holomorphically on any extraneous parameter on which ϕ were to depend holomorphically.

It is straightforward to check that fixing some

$$0 < r \leq \eta \exp\left(-\eta K \left\|\phi\right\|'_{(\mathbb{C}^k,0)\times\eta\mathbb{D}}\right)$$

inductively produces well-defined sequences $\left(\psi_n^\sharp\right)_{n\in\mathbb{N}}$ of $\mathrm{Holo_c}\left(\mathscr{U}_r^\sharp\right)$, for we have the implication

$$\begin{aligned}&\left\|\psi_{\varepsilon,n}^\infty\right\|_{\mathscr{U}_r^\infty} < \eta K \left\|\phi_\varepsilon\right\|'_{\eta\mathbb{D}}\\ &\qquad \Longrightarrow \left|v \exp \psi_{\varepsilon,n}^\infty(u, v)\right| < r \left|\exp \psi_{\varepsilon,n}^\infty(u, v)\right| < r \exp\left(\eta K \left\|\phi_\varepsilon\right\|'_{\eta\mathbb{D}}\right) \leq \eta\end{aligned}$$

for all $(u, v) \in \mathscr{A}_r$. Using (5.7) with $\psi := \psi_{\varepsilon,n}^\infty$ finally yields

$$\left\|\psi_{\varepsilon,n+1}^\infty\right\|_{\mathscr{U}_r^\infty} < \eta K \left\|\phi_\varepsilon\right\|'_{\eta\mathbb{D}}.$$

We establish now that both sequences converge in $\mathrm{Holo_c}\left(\mathscr{U}_r^\sharp\right)$. The hypothesis $\phi(u, 0) = 0$ guarantees that $\psi_{n+1}^\sharp(u, v) = \psi_n^\sharp(u, v) + \mathrm{O}\left(v^{n+1}\right)$, hence the bounded sequence $\left(\psi_n^\sharp\right)_{n\in\mathbb{N}}$ converges for the projective topology on $\mathbb{C}[[\varepsilon, u]][[v]]$ (for the Krull distance actually). Therefore the sequences converge towards holomorphic and bounded functions

$$\psi^\sharp := \lim_{n\to\infty} \psi_n^\sharp \in \mathrm{Holo_c}\left(\left(\mathbb{C}^k, 0\right) \times \rho^\sharp\mathbb{D} \times r\mathbb{D}\right)$$

according to the next lemma.

LEMMA 5.8. – [43, Lemma 2.10] *Let* $\mathscr{D}$ *be a domain in* $\mathbb{C}^m$ *and consider a bounded sequence* $\left(f_p\right)_{p\in\mathbb{N}}$ *of* $\mathrm{Holo_c}(\mathscr{D})$ *satisfying the additional property that there exists some point* $\mathbf{z}_0 \in \mathscr{D}$ *such that the corresponding sequence of Taylor series* $\left(T_p\right)_{p\in\mathbb{N}}$ *at* $\mathbf{z}_0$ *is convergent in* $\mathbb{C}[[\mathbf{z} - \mathbf{z}_0]]$ *(for the projective topology). Then* $\left(f_p\right)_p$ *converges uniformly on compact sets of* $\mathscr{D}$ *towards some* $f_\infty \in \mathrm{Holo_c}(\mathscr{D})$.

REMARK 5.9. – The limiting functions $\psi^{\sharp}$ are *not* obtained through the use of a fixed-point theorem, although they are a fixed-point of (5.5). The method used here, based on Lemma 5.8, does not use the fact that $\mathrm{Holo}_{\mathrm{c}}\left(\mathscr{D}\right)$ is a Banach space, only that it is a Montel space (any bounded subset is sequentially compact). Also the estimate (5.7) obtained in Lemma 5.6 (1) is easier to derive than a sharper estimate aimed at establishing that $\psi_n^{\sharp} \mapsto \psi_{n+1}^{\sharp}$ is a contraction mapping.

5.4.1. *Proof of Lemma 5.6 (2)*. – This is nothing but Cauchy formula.

5.4.2. *Proof of Lemma 5.6 (1)*. – Clearly the function $F^{\sharp}$ is holomorphic on the domain $\mathscr{U}_r^{\sharp}$. Notice also that modifying slightly the integration path does not change the value of the function, so that $F^{\sharp}$ is bounded on $\mathscr{U}_r^{\sharp}$. Let us evaluate its norm.

Set $\Psi : (u,v) \mapsto (u, v\exp\psi(u,v))$ and define $\rho > 0$ by $2\rho := \rho^{\infty} + \frac{1}{\rho^0}$. We prove the estimate on $\|F^{\infty}\|_{\mathscr{U}_r^{\infty}}$ and $\left\|F^0\right\|_{\mathscr{U}_r^0}$ in two steps: first we bound $|F^{\infty}(u,v)|$ when $|u| \leq \rho$ (resp. $\left\|F^0\right\|_{\mathscr{U}_r^0}$ when $|x| \leq \frac{1}{\rho}$), then when $\rho \leq |u| < \rho^{\infty}$ (resp. $\frac{1}{\rho} \leq |x| < \rho^0$).

— For $|u| \leq \rho$ and $|v| < r$ one has

$$|F^{\infty}(u,v)| \leq \|\phi \circ \Psi\|_{\mathscr{A}_r} \times \frac{1}{2\pi}\oint_{\rho^{\infty}\mathbb{S}^1}\left|\frac{\mathrm{d}z}{z-u}\right|.$$

On the one hand

$$\frac{1}{2\pi}\oint_{\rho^{\infty}\mathbb{S}^1}\left|\frac{\mathrm{d}z}{z-u}\right| \leq \frac{1}{\rho^{\infty}-\rho} = \frac{2\rho^0}{\rho^{\infty}\rho^0-1} < K,$$

while on the other hand, for all $(u,v) \in \mathscr{A}_r$,

$$|\phi(u, v\exp\psi(u,v))| \leq |v|\,\|\phi\|'_{\mathscr{A}_\eta}\exp\|\psi\|_{\mathscr{U}_r^{\infty}}.$$

Taking both bounds together completes the first step of the proof.

— This gives a corresponding bound for F^0 when $|x| \leq \frac{1}{\rho}$ since

$$\frac{|x|}{2\pi}\oint_{\frac{1}{\rho^0}\mathbb{S}^1}\left|\frac{\mathrm{d}z}{xz-1}\right| \leq \frac{1}{\rho-\frac{1}{\rho^0}} = \frac{2\rho^0}{\rho^{\infty}\rho^0-1}.$$

Taking (5.8) into account, one therefore deduces for $\frac{1}{\rho^0} < |u| \leq \rho$ the estimate

$$\begin{aligned}|F^{\infty}(u,v)| &\leq |\phi(u, v\exp\psi(u,v))| + \left|F^0\left(\frac{1}{u}, v\right)\right| \\ &\leq |v|\,\|\phi\|'_{\mathscr{A}_\eta}\exp\|\psi\|_{\mathscr{U}_r^{\infty}}\left(1+\frac{2\rho^0}{\rho^{\infty}\rho^0-1}\right)\end{aligned}$$

as expected.

— The bound for F^0 when $\frac{1}{\rho} \leq x < \rho^0$ is obtained similarly.

5.4.3. *Proof of Lemma 5.6 (3).* – Assume that $\left(\widetilde{F}^0, \widetilde{F}^\infty\right)$ is another pair of solution. Then for all $\left(\frac{1}{x}, y\right) = (u, v) \in \mathscr{A}_r$ we have

$$f^0(x, y) := F^0(x, y) - \widetilde{F}^0(x, y) = F^\infty(u, v) - \widetilde{F}^\infty(u, v) =: f^\infty(u, v),$$

defining a bounded and holomorphic function f on Sphere(r). The next lemma ends the proof.

LEMMA 5.10. – *If $f \in \mathrm{Holo_c}(\mathrm{Sphere}(r))$ then $\frac{\partial f^\infty}{\partial u} = 0$. In other words there is a natural isometry of Banach spaces*

$$\mathrm{Holo_c}(\mathrm{Sphere}(r)) \simeq \mathrm{Holo_c}(r\mathbb{D}).$$

Proof. – In the chart $\mathscr{U}_r^\infty$ expand f^∞ into a power series $f^\infty(u, v) = \sum_{n \geq 0} f_n(u) v^n$ convergent on $\mathbb{C} \times r\mathbb{D}$. By assumption f is bounded so that from Cauchy's estimate we get

$$|f_n(u)| \leq \|f\|_{\mathrm{Sphere}(r)} r^{-n}$$

for all $u \in \mathbb{C}$. Liouville Theorem tells us that each f_n is constant. □

5.5. Normal form recognition (proof of orbital Normalization Theorem)

The aim of this subsection is to shortly prove that the vector field $\Psi^* X_\varepsilon^0$ resulting from Theorem 5.5 is in normal form (2.4). Because Savelev's normalizing fibered mapping Ψ agrees with the identity on $\mathscr{L}$, each $\mathscr{F}_\varepsilon$ is induced in the chart $\mathscr{U}_r^0$ by a holomorphic vector field of the form

$$\mathscr{X}_\varepsilon^0(x, y) := \Psi^* X_\varepsilon^0 = P_\varepsilon(x) \frac{\partial}{\partial x} + y\left(1 + \mu_\varepsilon x^k + A_\varepsilon(x, y)\right) \frac{\partial}{\partial y},$$

where $A \in \mathrm{Holo}\left((\mathbb{C}^k, 0) \times \rho^0\mathbb{D} \times r\mathbb{D}\right)$ and $A(x, 0) = 0$.

We must prove the following result.

LEMMA 5.11. – *There exists a sequence of polynomials $a_n \in \mathbb{C}\{\varepsilon\}[x]_{\leq k}$ such that*

$$A(x, y) = \sum_{n=1}^{\infty} a_n(x) y^n$$

on $\mathscr{U}_r^0$.

Proof. – The expansion for A is valid for $(x, y) \in \mathscr{U}_r^0$ and a_n holomorphic in x. In the other chart $(u, v) = \left(\frac{1}{x}, y\right)$ the vector field $\mathscr{X}_\varepsilon^0$ is orbitally equivalent (conjugate after division by $uP_\varepsilon\left(\frac{1}{u}\right)$) to

$$\mathscr{X}_\varepsilon^\infty(u, v) := -u\frac{\partial}{\partial u} + v\left(\lambda_\varepsilon + h_\varepsilon(u) + \frac{1}{u^{k+1}P_\varepsilon\left(\frac{1}{u}\right)} u^k A_\varepsilon\left(\frac{1}{u}, v\right)\right)\frac{\partial}{\partial v}$$

where h is given by (5.2). This particular vector field must coincide with the holomorphic vector field defining $\mathscr{F}_\varepsilon$ in the chart $\mathscr{U}_r^\infty$ after application of (5.3), because every transform used from the start is fibered so that the factor $uP_\varepsilon\left(\frac{1}{u}\right)$ over $\mathscr{A}_r$ remains the same and no other function can be factored out. Therefore $u^k A_\varepsilon\left(\frac{1}{u}, v\right)$ is holomorphic near $(0, 0)$, and the conclusion follows. □

5.6. Proof of orbital Uniqueness Theorem (2)

Assume that there exists an orbital equivalence between two normal forms $\mathscr{X}$ and $\widetilde{\mathscr{X}}$. Those vector fields are in prepared form (4.2) thus they satisfy the hypothesis of the results presented in Section 4, and in particular there exists a fibered analytical conjugacy Ψ near $(0,0)$ between $\mathscr{X}$ and $\widetilde{\mathscr{X}}$, according to Corollary 4.11 (2).

Let $(\mathrm{Sphere}\,(r)\,,\mathscr{F})$ and $\left(\mathrm{Sphere}\,(r)\,,\widetilde{\mathscr{F}}\right)$ be the families of foliated standard neighborhoods of the sphere induced respectively by $\mathscr{X}$ and $\widetilde{\mathscr{X}}$. The fibered mappings Ψ are holomorphic and injective on a domain $\mathscr{D} \subset \mathscr{U}^0 \subset \mathrm{Sphere}\,(r)$ containing $(0,0)$. By a foliated path-lifting technique (as before) Ψ can be analytically continued on the domain

$$\mathscr{U}_\varepsilon := \mathrm{Sat}_{\mathscr{F}_\varepsilon}\left(\mathscr{D}\right) \subset \mathrm{Sphere}\,(r)\,.$$

Using the special form of the normal form $\mathscr{X}_\varepsilon$ we derive the following lemma in Section 5.6.2.

Lemma 5.12. – *There exists* $r' > 0$ *such that* $\mathrm{Sphere}\,(r') \setminus \{x = \infty\} \subset \mathscr{U}_\varepsilon$ *for all* $\varepsilon \in \left(\mathbb{C}^k, 0\right)$.

This lemma implies that Ψ_ε extends to a fibered, injective and holomorphic mapping $\mathrm{Sphere}\,(r') \setminus \{x = \infty\} \to \mathrm{Sphere}\,(r)$. The fact that Ψ_ε extends analytically to $\{x = \infty\}$ uses a variation on the Mattei-Moussu construction. The proof is standard, but we include it for the sake of completeness.

Lemma 5.13 ([32, Theorem 2]). – *We consider two germs of a holomorphic vector field* X *and* $\widetilde{X}$*, both with a singularity at the origin of same eigenratio* $\lambda \notin \mathbb{R}_{\geq 0}$ *and in the form*

$$x\frac{\partial}{\partial x} + \lambda y(1 + \mathrm{O}\,(x))\frac{\partial}{\partial y}. \tag{5.9}$$

Fix a germ of a transverse disk $\Sigma := \{x = x_*, y \in (\mathbb{C}, 0)\}$*, for* x_* *small enough, and assume that there exists an injective and holomorphic mapping* $\psi : \Sigma \to \{x = x_*\}$ *conjugating the respective holonomies induced by* X *and* $\widetilde{X}$*, computed through the fibration* $(x, y) \mapsto x$ *by turning around* $\{x = 0\}$*. Then there exists a holomorphic and injective, fibered mapping* Ψ *conjugating* X *and* $\widetilde{X}$ *on a connected neighborhood of* $(0,0)$ *containing* Σ*. We can even require that* Ψ *coincides with* ψ *on* Σ.

Proof. – Assume first that $\lambda < 0$. We can consider that the holonomies Γ and $\widetilde{\Gamma}$ are defined on $\Sigma := \{x = x_*\}\times r'\mathbb{D}$ and set $\Psi(x_*, y) := (x_*, \psi\,(y))$ on Σ. We then extend Ψ over the circle $\{|x| = |x_*|\}$ as a map of the form $\Psi(x, y) = (x, \psi(x, y))$, with $\psi(x_*, y) = \psi(y)$: the extension is done by the path-lifting technique detailed in Section 5.3. Ψ is of course well-defined because ψ conjugates the holonomies. To extend Ψ to $\rho\mathbb{D}\times r'\mathbb{D}$, we use the path-lifting along rays $\{\arg|x| = \mathrm{cst}\}$. Starting at (x_0, y) we lift the ray through x_0 up to $|x| = \rho$ in the leaf of X. We apply Ψ to the resulting point and then lift the ray back in the leaf of $\widetilde{X}$. The corresponding point is called $\Psi(x_0, y)$. We must show that

$$\{x_0\} \times C_1 r'\mathbb{D} \subset \Psi(\{x_0\} \times r'\mathbb{D}) \subset \{x_0\} \times C_2 r'\mathbb{D}$$

for some positive constants C_1, C_2 independent of x_0. For this purpose we can suppose that the $\mathrm{O}\,(x)$ part in (5.9) is bounded by $1/2$ (this is the case if ρ is sufficiently small). Then

$$|\lambda|\,|y|\left(1 - \frac{1}{2}\,|x_0|\exp t\right) < \frac{\mathrm{d}\,|y|}{\mathrm{d}t} < |\lambda|\,|y|\left(1 + \frac{1}{2}\,|x_0|\exp t\right),$$

yielding by Gronwall inequality that

$$|y(0)|\exp\left(\lambda t-\int_0^t\frac{1}{2}x_0\exp t\ \mathrm{d}t\right)\leq|y(t)|\leq|y(0)|\exp\left(\lambda t+\int_0^t\frac{1}{2}|x_0|\exp t\ \mathrm{d}t\right).$$

The conclusion follows since $\exp\left(\int_0^t\frac{1}{2}|x_0|\exp t\ \mathrm{d}t\right)\in\left]\exp\frac{-|x_0|}{2},1\right[$ is bounded and bounded away from 0 for $t<0$.

The previous argument remains valid when λ is not real. It suffices to replace $|\lambda|$ by $|\Re(\lambda)|$. □

REMARK 5.14. – The proof clearly shows that Ψ depends analytically on ε were X and $\widetilde{X}$ to depend analytically on ε.

The following lemma proved in Section 5.6.1 allows to complete the proof of the Uniqueness Theorem (2) by observing that injective holomorphic mappings on some standard neighborhood of the sphere are of a rather special kind.

LEMMA 5.15. – *Take some analytic family of maps* $\Psi:$ Sphere$(r')\to$ Sphere(r) *satisfying the following properties:*

- Ψ *is fibered,*
- Ψ_ε *is injective and holomorphic on* Sphere(r') *for every* $\varepsilon\in(\mathbb{C}^k,0)$.

Then

$$\Psi_\varepsilon^0(x,y)=\left(x,y\sum_{n=0}^{\infty}\psi_n y^n\right),\tag{5.10}$$

where, for all $n\in\mathbb{Z}_{\geq 0}$,

$$\psi_n\in\mathbb{C}\{\varepsilon\}$$

with a common radius of convergence, and ψ_0 *does not vanish for* $\varepsilon=0$. *Conversely, any convergent power series* Ψ *as above defines an analytic family satisfying the above properties for some* $r'>0$ *small enough.*

As a matter of consequence for every $\varepsilon\in(\mathbb{C}^k,0)$ and for any $(x,y)\in\mathscr{U}_r^0$

$$\Psi_\varepsilon(x,y)=(x,y\psi_\varepsilon(y)),\quad \psi_\varepsilon(0)\neq 0.$$

To preserve globally orbital normal forms (2.4) is so demanding that ψ_ε ends up being constant. Indeed, from

$$\Psi_\varepsilon^*\mathscr{X}_\varepsilon(x,y)=\widehat{\mathrm{X}}(x,y)+y\frac{A_\varepsilon(x,y)}{y\psi_\varepsilon'(y)+\psi_\varepsilon(y)}\frac{\partial}{\partial y}=\widetilde{\mathscr{X}}_\varepsilon(x,y),$$

where

$$A_\varepsilon(x,y):=x\psi_\varepsilon(y)R_\varepsilon(x,y\psi_\varepsilon(y))-y\psi_\varepsilon'(y)\left(1+\mu x^k\right),$$

we deduce by setting $x:=0$ that

$$0=A_\varepsilon(0,y)=-y\psi_\varepsilon'(y)$$

so that ψ_ε is constant, for otherwise $\widetilde{\mathcal{X}}_\varepsilon$ would not be in normal form. In each case we obtain finally

$$\psi_\varepsilon(v) = c_\varepsilon \in \mathbb{C}^\times$$

as expected. The remaining claim is a straightforward consequence of the study performed in Section 4.

5.6.1. *Proof of Lemma 5.15.* – The expansion (5.10) is valid on $\mathcal{U}^0_{r'}$ provided ψ_n depend holomorphically on x. Let us show that ψ_n is constant. Applying the transition mapping $(x, y) \mapsto \left(\frac{1}{x}, y\right)$ we obtain the expression of Ψ in the other chart:

$$\Psi^\infty(u, v) = \left(u, v \sum_{n=0}^{\infty} \psi_n\left(\frac{1}{u}\right) v^n\right),$$

holomorphic in $(u, v) \in \mathcal{U}^\infty_{r'}$. This implies in particular that each function $u \mapsto \psi_n\left(\frac{1}{u}\right)$ must be holomorphic at 0; the conclusion follows. The converse statement is straightforward.

5.6.2. *Proof of Lemma 5.12.* – We can find $\rho, r' > 0$ such that $\mathrm{cl}(\rho\mathbb{D} \times r'\mathbb{D}) \subset \mathcal{D}$, where $\mathcal{D}$ is the domain of Ψ. We show that, for some convenient choice of $r'' \leq r'$ every point $(x_*, y_*) \in \{|y| < r''\}$ can be linked to a point of $\rho\mathbb{D} \times r'\mathbb{D}$ by a path contained in a leaf of $\mathcal{F}^0_\varepsilon$. Only the case $|x_*| > \rho$ is not trivial. Since the singularity at $(\infty, 0)$ is neither a node nor a saddle-node, every small germ of a disk $\{u = u_*\}$ sufficiently close to $\{u = 0\}$, which is transverse to the separatrix $\mathcal{L}$, saturates a full pointed neighborhood $(\mathbb{C}, 0)^2 \setminus \{u = 0\} \subset \mathcal{U}^\infty_r$ under $\mathcal{F}^\infty_\varepsilon$. Therefore there exists $r''' > 0$ such that $\{0 < |u| \leq |u_*|, |v| < r'''\} \subset \mathcal{U}_\varepsilon$. Because $\mathcal{L}$ is invariant by $\mathcal{F}_\varepsilon$ and $\mathcal{L} \setminus (\{|x| < \rho\} \cup \{|u| < |u_*|\})$ is compact we may reduce r''' to some r'' in such a way that $\rho\mathbb{S}^1 \times r''\mathbb{D} \subset \mathcal{U}_\varepsilon$ (flow-box argument), which settles the proof.

6. Temporal normal forms

This section is devoted to proving the temporal part of the Normalization Theorem and of the Uniqueness Theorem in the case $\tau = 0$ (which particularly implies $\mu_0 \notin \mathbb{R}_{\leq 0}$). Recall how in Section 4 we obtained formal normal forms. The time-component U of any unfolding in orbital normal form (2.4)

$$Z = U\,\mathcal{X}$$

can be written as

$$\frac{1}{U} = C + I,$$

where

$$I \in \mathrm{im}\left(\mathcal{X}\cdot\right)$$

$$C \in \mathrm{coker}\left(\mathcal{X}\cdot\right), \quad C(0, 0) = \frac{1}{U(0, 0)} \neq 0,$$

for a given (arbitrary for now) choice of $\mathrm{coker}\left(\mathcal{X}\cdot\right)$, an algebraic supplementary in $\mathbb{C}[[\varepsilon, x, y]]$ to the image $\mathrm{im}\left(\mathcal{X}\cdot\right)$ of the (formal) Lie derivative $\mathcal{X}\cdot : \mathbb{C}[[\varepsilon, x, y]] \to \mathbb{C}[[\varepsilon, x, y]]$. According to the discussion following Proposition 4.7, Z is (formally) conjugate to $\frac{1}{C}\,\mathcal{X}$.

We have shown in Lemma 4.8 that

$$\mathbb{C}\left[\left[\varepsilon, x, y\right]\right] = \operatorname{im}\left(\mathcal{X}\cdot\right) \oplus \mathbb{C}\left[\left[\varepsilon\right]\right]\left[x\right]_{<k},$$

or more precisely that the following sequence of $\mathbb{C}\left[\left[\varepsilon\right]\right]$-linear operators is exact:

$$0 \longrightarrow \mathbb{C}\left[\left[\varepsilon\right]\right] \longrightarrow \mathbb{C}\left[\left[\varepsilon, x, y\right]\right] \xrightarrow{\mathcal{X}\cdot} \mathbb{C}\left[\left[\varepsilon, x, y\right]\right] \xrightarrow{\widehat{\mathfrak{T}}} \mathbb{C}\left[\left[\varepsilon\right]\right]\left[x\right]_{\leq k} \longrightarrow 0, \tag{6.1}$$

where $\widehat{\mathfrak{T}}$ maps G to the remainder of the Euclidean division of its partial function $x \mapsto G(x,0)$ by P. As a consequence we may take

$$\operatorname{coker}\left(\mathcal{X}\cdot\right) := \mathbb{C}\left[\left[\varepsilon\right]\right]\left[x\right]_{<k},$$

so that Z is formally conjugate to $\frac{1}{\widehat{\mathfrak{T}}\left(\frac{1}{U}\right)}\mathcal{X}$.

Remark 6.1. – The additional fact that

$$\widehat{\mathfrak{T}}\left(\frac{1}{U}\right) = \frac{1}{\widehat{\mathfrak{T}}\left(U\right)} + \mathrm{O}\left(P\right)$$

finally implies that Z is formally conjugate to $u\mathcal{X}$ where $u := \widehat{\mathfrak{T}}\left(U\right)$, as in the Formal Normalization Theorem. This is because one can write (for $u_0\left(0\right) \neq 0$)

$$\begin{aligned}\frac{1}{U\left(x,y\right)} &= \frac{1}{u\left(x\right) + \mathrm{O}\left(P\left(x\right)\right) + \mathrm{O}\left(y\right)} = \frac{1}{u\left(x\right)} \times \frac{1}{1 + \mathrm{O}\left(P\left(x\right)\right) + \mathrm{O}\left(y\right)} \\ &= \frac{1}{u\left(x\right)} + \mathrm{O}\left(P\left(x\right)\right) + \mathrm{O}\left(y\right).\end{aligned}$$

The previous argument still works for convergent power series, by replacing $\mathbb{C}\left[\left[\varepsilon, x, y\right]\right]$ with $\mathbb{C}\left\{\varepsilon, x, y\right\}$: if we provide an explicit cokernel in $\mathbb{C}\left\{\varepsilon, x, y\right\}$ of $\mathcal{X}\cdot|_{\mathbb{C}\left\{\varepsilon,x,y\right\}}$ then we can describe an explicit family of temporal normal forms.

Theorem 6.2. – *Assume $\tau = 0$ (which particularly implies $\mu_0 \notin \mathbb{R}_{\leq 0}$). Let an orbital normal form $\mathcal{X}$ be given. It acts by directional derivative on the linear space $\mathbb{C}\left\{\varepsilon, x, y\right\}$ in such a way that*

$$\mathbb{C}\left\{\varepsilon, x, y\right\} = \operatorname{im}\left(\mathcal{X}\cdot\right) \oplus \mathbb{C}\left\{\varepsilon\right\}\left[x\right]_{\leq k} \oplus \operatorname{Section}_k\left\{y\right\}.$$

(We refer to Section 3.1.2 *for the definition of the functional spaces.)*

Remark 6.3. – The construction of the cokernel of $\mathcal{X}\cdot$ is eventually performed for ε fixed. Therefore the theorem can also be specialized in the following way: for every $\varepsilon \in \left(\mathbb{C}^k, 0\right)$ such that $\mu_\varepsilon \notin \mathbb{R}_{\leq 0}$ and every disk $D \supset P_\varepsilon^{-1}\left(0\right)$ not containing any root of $1 + \mu_\varepsilon x^k$, we have the $\mathbb{C}$-linear decomposition

$$\operatorname{Holo_c}\left(D\right)\left\{y\right\} = \operatorname{im}\left(\mathcal{X}_\varepsilon\cdot\right) \oplus \mathbb{C}\left[x\right]_{\leq k} \oplus xy\mathbb{C}\left[x\right]_{<k}\left\{y\right\}.$$

If $\mu_\varepsilon \in \mathbb{R}_{\leq 0}$ a section of the cokernel is given by $xP_\varepsilon^\tau y\mathbb{C}\left[x\right]_{<k}\left\{P_\varepsilon{}^\tau y\right\}$.

The aim of this section is to prove this theorem but, before doing so, let us explain how it helps completing the proofs of the Normalization and Uniqueness Theorems. Every function $U \in \mathbb{C}\{\varepsilon, x, y\}^{\times}$ can be written uniquely as

$$U = \frac{u}{1 + uG}$$

where $\widehat{\mathfrak{T}}(G) = 0$, by simply taking $u := \widehat{\mathfrak{T}}(U)$ as in Remark 6.1. Then Theorem (6.2) allows decomposing G uniquely as

$$G = Q + I$$

with $Q \in \mathrm{Section}_k\{y\}$ and $I \in \mathcal{X} \cdot \mathbb{C}\{\varepsilon, x, y\}$, so that Z is analytically conjugate to some $\frac{u}{1+uQ}\mathcal{X}$, unique up to the action of linear transforms $(x, y) \mapsto (x, cy)$ as expected (as follows from Uniqueness Theorem (2) which has been proved in the previous section). This yields Uniqueness Theorem (1).

6.1. Reduction of the proof

We must study the obstructions to solve analytically cohomological equations of the form

$$\mathcal{X} \cdot F = G, \quad G \in \mathbb{C}\{\varepsilon, x, y\} \cap \ker \widehat{\mathfrak{T}}.$$

First observe that this equation, restricted to the invariant line $\{y = 0\}$, is always satisfied by a holomorphic function $f : x \mapsto F(x, 0)$ solving

$$f'(x) = \frac{G(x,0)}{P(x)} \in \mathbb{C}\{\varepsilon, x\}.$$

By subtracting f from F and $x \mapsto G(x, 0)$ from G, we may always assume without loss of generality that

$$G(x,0) = F(x,0) = 0,$$

i.e., $G \in \mathbb{C}\{\varepsilon, x, y\}'$ as defined in Section 3.1.2.

Let

$$\Delta_k := \left\{\varepsilon \in \left(\mathbb{C}^k, 0\right) : \#P_\varepsilon^{-1}(0) \leq k\right\}$$

be a germ at 0 of the discriminant hypersurface of P_ε, so that each open set $\left(\mathbb{C}^k, 0\right) \backslash \Delta_k$ consists in generic values of the parameter for which P_ε has only simple roots. Proving Theorem 6.2 will require to work in the functional spaces

$$\mathcal{H}_\ell\{\mathbf{z}\} := \bigcup_{\mathcal{D}=(\mathbb{C}^n,0)} \mathrm{Holo_c}\left(\mathcal{E}_\ell \times \mathcal{D}\right)', \quad \mathbf{z} := (z_1, \ldots, z_n)$$

for some decomposition $(\mathcal{E}_\ell)_\ell$ of $\left(\mathbb{C}^k, 0\right) \backslash \Delta_k$ into finitely many (germs of) open cells as explained in Section 6.3. (We recall that the definition of the space $\mathrm{Holo_c}\left(\mathcal{D}\right)'$ is given in Section 3.1.2.) We choose these spaces because of the next property.

Lemma 6.4. – *We have*

$$\mathbb{C}\{\varepsilon, \mathbf{z}\}' = \bigcap_\ell \mathcal{H}_\ell\{\mathbf{z}\}.$$

(By the intersection on the right hand side we of course mean the functions that have an extension on the unions of the different domains.)

Proof. – We certainly have

$$\mathbb{C}\{\varepsilon,\mathbf{z}\}' \subset \bigcap_{\ell} \mathcal{H}_\ell\{\mathbf{z}\}\,.$$

Conversely if $f \in \bigcap_\ell \mathcal{H}_\ell\{\mathbf{z}\}$ then f defines a bounded, holomorphic function on $\left((\mathbb{C}^k,0)\setminus\Delta_k\right)\times(\mathbb{C}^n,0)$, which extends holomorphically to $\left(\mathbb{C}^{k+n},0\right)$ according to Riemann's Theorem on removable singularities. □

Working over a fixed cell germ $\mathcal{E}_\ell$ is easy as compared to working on a full neighborhood of the parameter space $\left(\mathbb{C}^k,0\right)$.

Proposition 6.5 ([41]). – *Let $\mathcal{E}_\ell$ be a parameter cell. There exists $\mathfrak{T}_\ell$, called the* period operator *over $\mathcal{E}_\ell$, such that the sequence of* $\mathrm{Holo_c}(\mathcal{E}_\ell)$*-linear operators is exact:*

$$0 \longrightarrow \mathrm{Holo_c}(\mathcal{E}_\ell) \longrightarrow \mathcal{H}_\ell\{x,y\} \xrightarrow{\mathcal{X}\cdot} \mathcal{H}_\ell\{x,y\} \xrightarrow{\mathfrak{T}_\ell} \prod_{\mathbb{Z}/k\mathbb{Z}} \mathcal{H}_\ell\{h\} \tag{6.2}$$

where h is a one-dimensional variable (meant to take the values of a first integral).

The surjectivity of the period operator $\mathfrak{T}_\ell$ has not been established in the cited reference, but it would have followed from an immediate adaptation of the argument of [44, Lemma 3.4]. Here, though, we prove a stronger result by producing an explicit section to the period operator (Proposition 6.6 to come). The construction of the period operator over $\mathcal{E}_\ell$ is explained in Section 6.2 below. It involves cutting up $\left(\mathbb{C}^2,0\right)\setminus\left(P_\varepsilon^{-1}(0)\times\{0\}\right)$ into k open (bounded) spiraling sectors and building sectorial solutions of the cohomological equation. The period operator measures how much solutions on neighboring sectors disagree on intersections. Contrary to what would have make things easier

$$\mathfrak{T}_\ell\left(\mathbb{C}\{\varepsilon,x,y\}'\right) \neq \bigcap_p \prod_{\mathbb{Z}/k\mathbb{Z}} \mathcal{H}_p\{h\} = \prod_{\mathbb{Z}/k\mathbb{Z}} \mathbb{C}\{\varepsilon,h\}'\,,$$

so that $\mathfrak{T}_\ell$ is neither onto nor into the natural candidate $\prod_{\mathbb{Z}/k\mathbb{Z}}\mathbb{C}\{\varepsilon,h\}'$. This situation differs drastically from the case $\varepsilon=0$, and can be explained. It turns out that the variable h in the j-th factor of $\prod_{j\in\mathbb{Z}/k\mathbb{Z}}\mathcal{H}_p\{h\}$ stands for values of the canonical first integral of $\mathcal{X}$ on the j-th sector (see the discussion preceding Definition 6.10). Different sectorial decompositions for fixed ε, corresponding to different cells $\mathcal{E}_\ell$ containing ε, lead to incommensurable sectorial dynamics: there is no correspondence between h-variables coming from different overlapping cells (see also Section 9). Therefore we need to relocate the obstructions in geometrical space (x,y), by introducing a well-chosen section $\mathfrak{S}_\ell$ of $\mathfrak{T}_\ell$.

Proposition 6.6. – *Let $\mathcal{E}_\ell$ be a parameter cell and assume $\tau=0$ (which particularly implies $\mu_0\notin\mathbb{R}_{\leq 0}$). There exists a linear isomorphism*

$$\mathfrak{S}_\ell : \prod_{\mathbb{Z}/k\mathbb{Z}} \mathcal{H}_\ell\{h\} \longrightarrow x\,\mathcal{H}_\ell\{y\}[x]_{<k}$$

such that $\mathfrak{T}_\ell\circ\mathfrak{S}_\ell=\mathrm{Id}$. This particularly means that we recover a cellular cokernel of $\mathcal{X}\cdot$ as follows:

$$\mathcal{H}_\ell\{x,y\} = \left(\mathcal{X}\cdot\mathcal{H}_\ell\{x,y\}\right)\,\oplus\,x\,\mathcal{H}_\ell\{y\}[x]_{<k}\,.$$

This proposition is showed later in Section 6.4 using a refinement of the Cauchy-Heine transform, this time on unbounded sectors in the x-variable. Theorem 6.2 is proved once we establish the next gluing property, as done in Section 6.5.

PROPOSITION 6.7. – *For every parameter cells $\mathcal{E}_\ell$ and $\mathcal{E}_{\widetilde{\ell}}$ with non-empty intersection we have*

$$\mathfrak{S}_\ell \circ \mathfrak{T}_\ell = \mathfrak{S}_{\widetilde{\ell}} \circ \mathfrak{T}_{\widetilde{\ell}}$$

on $\mathcal{H}_\ell \{x, y\} \cap \mathcal{H}_{\widetilde{\ell}} \{x, y\}$.

From Lemma 6.4 we deduce the identity

$$\mathrm{Section}_k \{y\} = \bigcap_\ell x \mathcal{H}_\ell \{y\} [x]_{<k} ,$$

hence the proposition actually provides us with a well-defined, surjective operator

$$\begin{aligned} \mathfrak{K} :\ \mathbb{C}\{\varepsilon, x, y\}' &\longrightarrow \mathrm{Section}_k \{y\} \\ G &\longmapsto \mathfrak{S}_\ell \left(\mathfrak{T}_\ell (G)\right), \end{aligned} \tag{6.3}$$

whose kernel coincides with $\mathcal{X} \cdot \mathbb{C}\{\varepsilon, x, y\}'$, i.e., the sequence of $\mathbb{C}\{\varepsilon\}$-linear operators

$$0 \longrightarrow \mathbb{C}\{\varepsilon, x, y\}' \xrightarrow{\mathcal{X}\cdot} \mathbb{C}\{\varepsilon, x, y\}' \xrightarrow{\mathfrak{K}} \mathrm{Section}_k \{y\} \longrightarrow 0 \tag{6.4}$$

is exact, as required to establish Theorem 6.2.

6.2. Cohomological equation and period operator

THEOREM 6.8 ([41]). – *For every $\rho > 0$ there exists:*

— *a covering of $\left(\mathbb{C}^k, 0\right) \setminus \Delta_k$ by finitely many open, contractible cells $(\mathcal{E}_\ell)_\ell$,*

— *for every $\varepsilon \in \mathcal{E}_\ell$, a covering of*

$$V_\varepsilon := \rho\mathbb{D} \backslash P_\varepsilon^{-1} (0)$$

into k open, contractible squid sectors

$$V_{\ell,\varepsilon}^j ,\ j \in \mathbb{Z}/k\mathbb{Z},$$

for which the following properties are satisfied. Recall that the closure of a subset A of a topological space is written $\mathrm{cl}\,(A)$.

1. *Each map $\varepsilon \mapsto \mathrm{cl}\left(V_{\ell,\varepsilon}^j\right)$ is continuous for the Hausdorff distance on compact sets and*

$$\lim_{\substack{\varepsilon \to 0 \\ \mathcal{E}_\ell}} \mathrm{cl}\left(V_{\ell,\varepsilon}^j\right) = \mathrm{cl}\left(V_0^j\right)$$

coincides with (the closure of) a usual sector of the limiting saddle-node, namely

$$V_0^j := \left\{ x :\ 0 < |x| < \rho, \arg x \in \left]-\frac{3\pi}{2k} + \eta + j\frac{2\pi}{k},\quad \frac{3\pi}{2k} - \eta + j\frac{2\pi}{k}\right[\right\}$$

for some $\eta \in \left]0, \frac{\pi}{2k}\right[$.

2. *We let*

$$V_{\ell}^{j}:=\bigcup_{\varepsilon\in\mathcal{E}_{\ell}}\{\varepsilon\}\times V_{\ell,\varepsilon}^{j}.$$

For every $G\in\mathrm{Holo_c}\left(\mathcal{E}_{\ell}\times\rho\mathbb{D}\times(\mathbb{C},0)\right)'$ *there exists a unique family* $\left(F_{\ell}^{j}\right)_{j\in\mathbb{Z}/k\mathbb{Z}}$ *such that* F_{ℓ}^{j} *is the unique solution of*

$$\mathcal{X}\cdot F=G$$

in the space $\mathrm{Holo_c}\left(V_{\ell}^{j}\times(\mathbb{C},0)\right)'$. *Moreover*

$$\lim_{\substack{\varepsilon\to 0\\ \mathcal{E}_{\ell}}}F_{\ell,\varepsilon}^{j}=F_{0}^{j}$$

uniformly on compact sets of $V_{0}^{j}\times(\mathbb{C},0)$, *where* F_{0}^{j} *is the canonical sectorial solution of the limiting cohomological equation* [45].

3. *There exists a solution* $F\in\mathrm{Holo_c}\left(\mathcal{E}_{\ell}\times\rho\mathbb{D}\times(\mathbb{C},0)\right)$ *of* $\mathcal{X}\cdot F=G$ *if, and only if, for every* $\varepsilon\in\mathcal{E}_{\ell}$ *and* $j\in\mathbb{Z}/k\mathbb{Z}$

$$F_{\ell,\varepsilon}^{j+1}=F_{\ell,\varepsilon}^{j}$$

on corresponding pairwise intersections of sectors $V_{\ell,\varepsilon}^{j}\times(\mathbb{C},0)$.

We provide details regarding how squid sectors and parameter cells are obtained in Section 6.3 below. The way sectorial solutions $\left(F_{\ell}^{j}\right)_{j\in\mathbb{Z}/k\mathbb{Z}}$ are built is explained in [41, Section 7]. The third property encodes all we need to know in order to characterize algebraically the obstructions to solve analytically cohomological equations. It is, as usual, eventually a consequence of Riemann's Theorem on removable singularities.

REMARK 6.9. – 1. A usual saddle-node sector is divided by rays separated by an angle slightly larger than $\frac{\pi}{k}$: allowing an extra $\frac{\pi}{2k}$ on each side yields sectors of opening between $\frac{\pi}{k}$ and $\frac{2\pi}{k}$. However we are in the particular case of a saddle-node with analytic center manifold, meaning that we need twice less sectors to describe the singularity structure. Hence the angle between the dividing rays can be taken as big as $\frac{2\pi}{k}$: allowing an extra $\frac{\pi}{2k}$ on each side yields an opening between $\frac{2\pi}{k}$ and $\frac{3\pi}{k}$.

2. A corollary to this theorem is the fact that any generic convergent unfolding is conjugate to its formal normal form over every region $V_{\ell}^{j}\times(\mathbb{C},0)$. In particular each $\mathcal{X}$ is conjugate over $V_{\ell,\varepsilon}^{j}\times(\mathbb{C},0)$ to $\widehat{\mathrm{X}}$ by a fibered mapping

$$(x,y)\longmapsto\left(x,y\exp N_{\ell,\varepsilon}^{j}(x,y)\right)$$

built upon a sectorial solution of

$$\mathcal{X}_{\varepsilon}\cdot N_{\ell,\varepsilon}^{j}=-R_{\varepsilon}$$

as in Proposition 4.7.

3. A really important property of the construction: it is performed [41, Section 7] for each *fixed* $\varepsilon \in \mathcal{E}_\ell$, the holomorphic / continuous dependence on ε of resulting objects being a by-product. This greatly simplifies understanding what happens on overlapping cells. This is also the reason why we omit to include the subscripts ℓ and ε in the sequel, whenever doing so does not introduce ambiguity.

The period operator $\mathfrak{T}_\ell$ is obtained as follows. Fix $\varepsilon \in \mathcal{E}_\ell$ and $\rho > 0$ as in the previous theorem. Starting from any $G \in \mathrm{Holo}\,(\rho\mathbb{D} \times (\mathbb{C},0))'$ we can find a unique collection $\left(F^j\right)_{j\in\mathbb{Z}/k\mathbb{Z}} \in \prod_j \mathrm{Holo}\left(V^j \times (\mathbb{C},0)\right)'$ of bounded functions solving the equation $\mathcal{X} \cdot F = G$ over a squid sector. On each intersection we have $\mathcal{X} \cdot F^{j+1} = G = \mathcal{X} \cdot F^j$ so that $F^{j+1} - F^j$ is a first integral of $\mathcal{X}$. Therefore it factors as

$$F^{j+1} - F^j = T^j \circ H^j, \quad T^j \in \mathbb{C}\{h\}' \tag{6.5}$$

where $H^j = H^j_{\ell,\varepsilon}$ is the *canonical sectorial first integral* with connected fibers

$$H^j := \widehat{H}^j \exp N^j, \tag{6.6}$$

obtained from that of the formal normal form

$$\widehat{H}^j\,(x,y) := y \exp \int^x -\frac{1+\mu z^k}{P\,(z)}\mathrm{d}z \tag{6.7}$$

by composition with the sectorial normalization (Remark 6.9). We can fix once and for all a determination of each first integral $\widehat{H}^j = \widehat{H}^j_\ell$ on V^j_ℓ in such a way that

$$\widehat{H}^{j+1} = \widehat{H}^j \exp{2\mathrm{i}\pi\mu/k} \tag{6.8}$$

in $V^{j,\mathrm{s}}$. The linear factor appearing on the right-hand side is here to accommodate the multivaluedness of $\exp \int^x -\frac{1+\mu z^k}{P(z)}\mathrm{d}z = x^{-\mu} \times \mathrm{holo}\,(x)$ near ∞, so that $\widehat{H}^{j+k} = \widehat{H}^j$.

Definition 6.10. – Consider a parameter cell $\mathcal{E}_\ell$ and $\rho > 0$ as in Theorem 6.8. For $G \in \mathrm{Holo_c}\,(\mathcal{E}_\ell \times \rho\mathbb{D} \times (\mathbb{C},0))$ define the *period* of G with respect to $\mathcal{X}$ as the k-tuple

$$\mathfrak{T}_\ell\,(G) := \frac{1}{2\mathrm{i}\pi}\left(T^j\right)_{j\in\mathbb{Z}/k\mathbb{Z}} \in \prod_{\mathbb{Z}/k\mathbb{Z}} \mathcal{H}_\ell\,\{h\}$$

where $T^j_\varepsilon := T^j$ is build as above in (6.5) for $G := G_\varepsilon$ and $\varepsilon \in \mathcal{E}_\ell$. We define $\mathfrak{T}^j_\ell\,(G) := \frac{1}{2\mathrm{i}\pi}T^j$ to be the j-th component of $\mathfrak{T}_\ell\,(G)$.

Remark 6.11. – Following up on Remark 6.9 (1), it seems that the period of R must play a special role regarding classification, since it measures the discrepancy between sectorial orbital conjugacies to the formal normal form $\widehat{\mathrm{X}}$. It is actually the case that the unfolded Martinet-Ramis modulus is linked to this period through the relationship

$$\psi^{j,\mathrm{s}}_\ell\,(h) = h \exp\left(\frac{2\mathrm{i}\pi\mu}{k} + \phi^{j,\mathrm{s}}_\ell\,(h)\right) = h \exp\left(\frac{2\mathrm{i}\pi\mu}{k} - \mathfrak{T}^j_\ell\,(R)\,(h)\right).$$

A similar formula holds for the temporal modulus, namely $f^{j,\mathrm{s}}_\ell = \mathfrak{T}^j_\ell\left(\frac{1}{U} - 1\right)$. We refer to [41] for a more detailed discussion regarding these integral representations of the modulus of classification.

We sum up the relevant results needed in the sequel as a corollary to Theorem 6.8.

COROLLARY 6.12. – *Pick* $\varepsilon \in \left(\mathbb{C}^k, 0\right) \setminus \Delta_k$ *and* $\rho > 0$ *such that* $P_\varepsilon^{-1}(0) \subset \rho\mathbb{D}$*, as well as some holomorphic function* $G \in \mathrm{Holo}\left(\rho\mathbb{D} \times (\mathbb{C}, 0)\right)'$*. The following assertions are equivalent.*

1. *There exists* $F \in \mathrm{Holo}\left(\rho\mathbb{D} \times (\mathbb{C}, 0)\right)'$ *such that* $\mathcal{X}_\varepsilon \cdot F = G$.
2. *There exists* ℓ *with* $\varepsilon \in \mathcal{E}_\ell$ *such that*
$$\mathfrak{T}_\ell\left(G\right)_\varepsilon = 0.$$
3. *For all* ℓ *with* $\varepsilon \in \mathcal{E}_\ell$ *we have*
$$\mathfrak{T}_\ell\left(G\right)_\varepsilon = 0.$$

If moreover $G \in \mathcal{H}_\ell\{x, y\}$ *then*

$$\lim_{\substack{\varepsilon \to 0 \\ \mathcal{E}_\ell}} \mathfrak{T}_\ell\left(G\right)_\varepsilon = \mathfrak{T}\left(G_0\right)$$

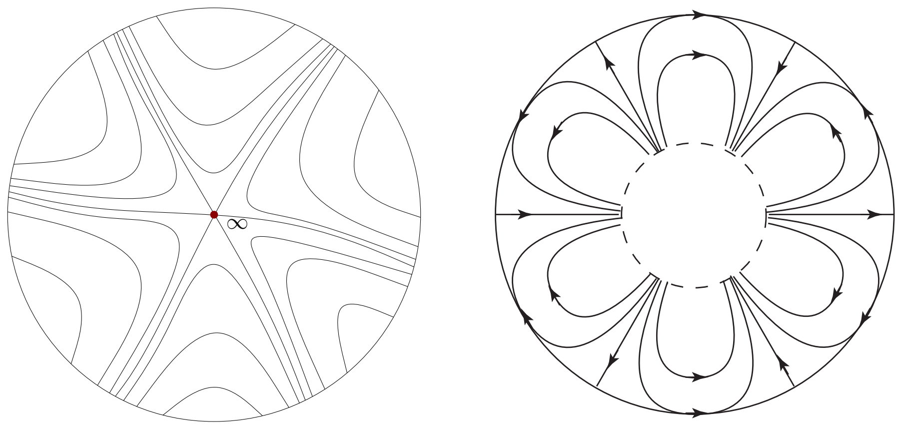
uniformly on $(\mathbb{C}, 0)$*, where* $\mathfrak{T} : \mathbb{C}\{x, y\}' \to \prod_{\mathbb{Z}/k\mathbb{Z}} \mathbb{C}\{h\}'$ *is the period operator of the limiting saddle-node* [45].

Proof. – For fixed ε and ℓ Theorem 6.8 asserts the equivalence between existence of an analytic solution F of the cohomological equation $X_\varepsilon \cdot F = G$ and vanishing of the period $\mathfrak{T}_\ell\left(G\right)_\varepsilon$. But the analyticity of F has nothing to do with the way the underlying squid sectors are cut, therefore $\mathfrak{T}_{\widetilde{\ell}}(G)_\varepsilon = 0$ as soon as $\varepsilon \in \mathcal{E}_{\widetilde{\ell}}$. □

6.3. Description of (unbounded) squid sectors and parameter cells

To characterize the dynamics, describe the modulus of analytic classification and more generally build the period operator, we need to work over k open *squid sectors* in x-space covering either $\rho\mathbb{D} \setminus P_\varepsilon^{-1}(0)$ (bounded case) or $\mathbb{C} \setminus P_\varepsilon^{-1}(0)$ (unbounded case). Since $\{y = 0\}$ is an analytic center manifold, each sector in this paper is the union of two consecutive sectors described originally in [41]. The cited reference also guarantees that it is sufficient to limit ourselves to the complement of the discriminant hypersurface $\Delta_k \ni 0$ in parameter space. Although we only reach parameters for which all roots of P_ε are simple, the construction passes without difficulty to the limit $\varepsilon \to \Delta_k$. For $\varepsilon \notin \Delta_k$ the squid sectors are attached to two or three roots. When $\varepsilon \to 0$ they converge to the sectors used in the description of the Martinet-Ramis modulus for convergent saddle-nodes.

The singular points depend analytically on $\varepsilon \in \left(\mathbb{C}^k, 0\right) \setminus \Delta_k$. To obtain a family of squid sectors suiting our needs, we must ensure that the sectors vary continuously as ε does. This is however not achievable on a full pointed neighborhood of Δ_k in parameter space, for reasons we are about to explain (we particularly refer to Remark 6.16). Even so, we manage to deal with all values of ε by covering the space $\left(\mathbb{C}^k, 0\right)$ with the closure of finitely many contractible domains $\left(\mathcal{E}_\ell\right)_\ell$ in ε-space, which we call *cells*, on which admissible families of squid sectors exist.



(A) Neighborhood of ∞ for $k = 3$

(B) The neighborhood of $\rho\mathbb{D}$

FIGURE 6.1. The separatrices of the pole at ∞ and the petals along the boundary of the disk $\rho\mathbb{D}$.

6.3.1. *The dynamics of* $\dot{x} = P_\varepsilon(x)$. – Let us recall the main features of the vector field $P_\varepsilon \frac{\partial}{\partial x}$. When P_ε has distinct roots x_ε, each singular point x_ε has an associated nonzero eigenvalue $\lambda_\varepsilon = P'_\varepsilon(x_\varepsilon)$.

- The point x_ε is a *radial node* if $\lambda_\varepsilon \in \mathbb{R}$. It is attracting (resp. repelling) if $\lambda_\varepsilon < 0$ (resp. $\lambda_\varepsilon > 0$).
- The point x_ε is a *center* if $\lambda_\varepsilon \in i\mathbb{R}$.
- The point x_ε is a *focus* if $\lambda_\varepsilon \notin \mathbb{R} \cup i\mathbb{R}$. It is attracting (resp. repelling) if $\Re(\lambda_\varepsilon) < 0$ (resp. $\Re(\lambda_\varepsilon) > 0$).

The point $x = \infty$ serves as an organizing center; indeed, the vector field $P_\varepsilon \frac{\partial}{\partial x}$ has a pole of order $k-1$ with $2k$ separatrices at $x = \infty$, alternately attracting and repelling (see Figure 6.1), thus limiting $2k$ saddle sectors at ∞. The system is structurally stable in the neighborhood of ∞ for ε small. These saddle sectors give a phase portrait resembling $2k$ petals along the boundary of any (sufficiently large) disk centered at the origin. The relationship between the magnitude of the parameter and the size of the disk will be detailed in Section 6.3.4.

The dynamics is completely determined by the separatrices of ∞. Because all roots of P_ε are simple, only two types of behavior occur.

- For generic values of ε, following the separatrices from ∞ (either in backward or forward direction) one lands at repelling ($t \to -\infty$) or attracting ($t \to \infty$) singular points x_ε of focus or radial node type. In that case, each singular point is attached to at least one separatrix and the system is structurally stable among polynomial systems of degree $k + 1$. See Figure 6.2 for a phase portrait with generic ε.
- The sets of generic ε are separated by bifurcation hypersurfaces of (real) codimension 1. For these non-generic values of ε a homoclinic connection occurs between an attracting separatrix and a repelling separatrix of infinity: there is then a real integral curve

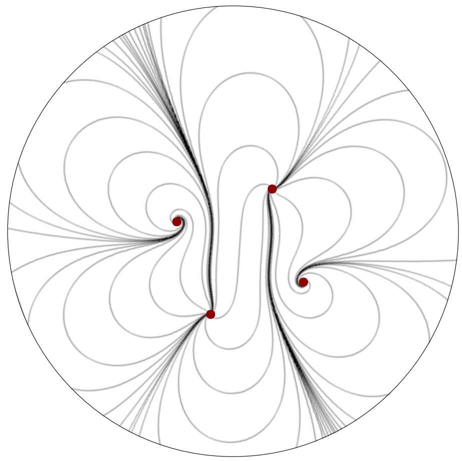


FIGURE 6.2. An example of a structurally stable real foliation induced by a complex polynomial vector field of degree 4 for ε in some K_ℓ.

flowing out from infinity in the x-plane and flowing back to infinity in finite time. For these bifurcation sets, the singular points $P_\varepsilon^{-1}(0)$ can be split into two (nonempty) subsets I_1 and I_2 satisfying

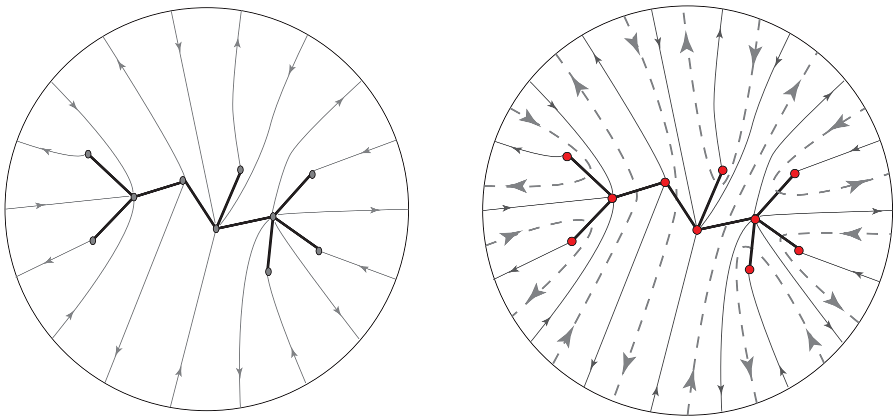
$$\sum_{x\in I_m} \frac{1}{P_\varepsilon'(x)} \in \mathrm{i}\mathbb{R}, \qquad m = 1, 2. \tag{6.9}$$

This can be seen by integrating the 1-form $\mathrm{d}t = \frac{\mathrm{d}x}{P_\varepsilon(x)}$ along a homoclinic orbit, and evaluating residues. When I_m is a singleton, the corresponding singular point is a center.

The union of the $2k$ separatrices of ∞ is called the *separating graph* in [11] (see Figure 6.3(A)). It splits $\mathbb{C}$ into k simply connected regions. In each of these regions we can draw a curve γ_j connecting the interior of a saddle sector at ∞ to the interior of another saddle sector (see Figure 6.3(B)). There are exactly $C_k = \frac{1}{k+1}\binom{2k}{k}$ ways of pairing two by two the saddle sectors of ∞ by non-intersecting curves, thus providing a topological invariant for the vector field (we also refer to [10]).

6.3.2. *Rough description of the cells $\mathcal{E}_\ell$ in parameter space.* – The non-generic values of ε form a set of (real) codimension 1 which partitions a convenient neighborhood of 0 in parameter space (to be described slightly later) into C_k open regions K_ℓ, corresponding to structurally stable vector fields with the same topological invariant. In each region K_ℓ, the topology of the phase portrait is completely determined by the topological way of attaching the $2k$ separatrices to the $k+1$ singular points. If x_ε is a root of P_ε (depending continuously on ε) then $\Re\left(P_\varepsilon'(x_\varepsilon)\right)$ has a constant sign for all $\varepsilon \in K_\ell$. Each cell $\mathcal{E}_\ell$ in parameter space will be a small enlargement of K_ℓ, so that the cells cover the complement of Δ_k.

A. Douady, J.F. Estrada and P. Sentenac have also provided a very clever parametrization of the domains K_ℓ, thus showing that they are simply connected.



(A) The separating graph (B) The curves γ_i

FIGURE 6.3. On the left, the separating graph formed by the separatrices landing at the singular points and flow lines (in bold) between the singular points. On the right, the curves γ_i (in dotted lines) used to calculate the τ_i.

THEOREM 6.13 ([11]). – *Let K_ℓ be a maximal domain corresponding to structurally stable vector fields. Then, there exists a biholomorphism $\Phi_\ell : K_\ell \to \mathbb{H}^k$, where $\mathbb{H}$ is the upper half-plane. In particular, K_ℓ is contractible. The set $\Phi_\ell^{-1}\left((\mathrm{i}\mathbb{R}_{\geq 0})^k\right)$, which we call the* spine *of K_ℓ, corresponds to polynomial vector fields with real eigenvalues at each singular point.*

The map Φ_ℓ is defined as follows: let $\left(\gamma_\varepsilon^j\right)_{j\in\{1,\dots,k\}}$ be k disjoint loops attached to ∞ and pairing the saddle sectors of ∞, without intersecting the separating graph. Then $\Phi_\ell(\varepsilon) = (\tau_\varepsilon^1, \dots, \tau_\varepsilon^k)$, where

$$\tau_\varepsilon^j := \int_{\gamma_\varepsilon^j} \mathrm{d}t = \int_{\gamma_\varepsilon^j} \frac{\mathrm{d}x}{P_\varepsilon(x)},$$

the orientation of γ_ε^j being chosen so that $\Im\left(\tau_\varepsilon^j\right) > 0$.

Since $\tau_\varepsilon^j = 2\mathrm{i}\pi \sum_{x\in I} \frac{1}{P_\varepsilon'(x)}$, where I is the set of singular points in a domain bounded by γ_ε^j, the sum τ_ε^j admits an analytic continuation outside K_ℓ. In particular, when ε is a boundary point of K_ℓ for which there is a homoclinic loop through ∞, some of the τ_ε^j become real.

The cells have a very useful conic structure, induced by a multiplicative action of $\mathbb{R}_{>0} \ni \lambda$ through linear rescaling

$$(\varepsilon_{k-1}, \dots, \varepsilon_0, x, t) \longmapsto \left(\lambda^{-(k-2)}\varepsilon_{k-1}, \dots, \varepsilon_1, \lambda\varepsilon_0, \lambda x, \lambda^{-k} t\right), \tag{6.10}$$

as indeed the differential equation $\dot{x} = P_\varepsilon(x)$ is invariant under this action. The cones we use are of the form

$$\left\{\left(\lambda^2\varepsilon_{k-1}, \dots, \lambda^k\varepsilon_1, \lambda^{k+1}\varepsilon_0\right) : \lambda \in]0,1[\,, \varepsilon \in K\right\},$$

where K is a relative domain within a sphere-like real hypersurface. This compact hypersurface takes the form $\{\|\varepsilon\| = \mathrm{cst}\}$ with

$$\|\varepsilon\| := \max\left(|\varepsilon_{k-1}|^{\frac{1}{2}}, \ldots, |\varepsilon_1|^{\frac{1}{k}}, |\varepsilon_0|^{\frac{1}{k+1}}\right). \tag{6.11}$$

REMARK 6.14. – The expression 6.11 does not define a norm because the homogeneity axiom is not satisfied. However, if we take into account that the ε_j are the symmetric functions in the roots $(x_0, \ldots, x_k) \in \mathbb{C}^{k+1}$ of P_ε, it lifts to a norm on $\mathbb{C}^{k+1}$. Thus $\|\varepsilon\|$ measures the magnitude of the parameter ε and the $\|\bullet\|$-balls form a fundamental basis of neighborhood of 0. In the following we consider only these parametric neighborhoods.

The regions K_ℓ of structural stability defined above are cones of this form, and so will be their enlargements to cells $\mathcal{E}_\ell$ covering the complement of Δ_k. Also, when considering limits for $\varepsilon \to 0$ it will be natural to consider limits for $\lambda \to 0$ along orbits of the $\mathbb{R}_{>0}$-action

$$\left\{\left(\lambda^2\varepsilon_{k-1}, \ldots, \lambda^k\varepsilon_1, \lambda^{k+1}\varepsilon_0\right) : \lambda \in]0,1[\right\}. \tag{6.12}$$

6.3.3. *Saddle- and node-like singular points, admissible angles*. – We want to stress that a singular point x_ε of $\dot{x} = P_\varepsilon(x)$ with non-real eigenvalue $\lambda = a + \mathrm{i}b$ can be both attracting and repelling depending on how we approach it along logarithmic spirals. Making sense of this statement entails complexifying the time. Let us explain how.

— Consider the linear equation $\dot{x} = \lambda x$. Its solutions are $x(t) = x_0 \exp(\lambda t)$. Now, let us allow complex values of t along some slanted real line $t = (c + \mathrm{i}d)T = T\exp(\mathrm{i}\theta)$ in $\mathbb{C}$-space for some fixed $c + \mathrm{i}d \in \mathbb{S}^1$, with $c > 0$ (corresponding to $\theta \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$) and $T \in \mathbb{R}$. Then

$$x(t(T)) = x_0 \exp\left(((ac - bd) + \mathrm{i}(ad + bc))\, T\right),$$

and $\lim_{T\to+\infty} x(t(T)) = 0$ (*resp.* $\lim_{T\to-\infty} x(t(T)) = 0$) when $ac - bd < 0$ (*resp.* $ac - bd > 0$).

— Since $b \neq 0$, it is always possible to find $c_1 > 0, d_1$ (*resp.* $c_2 > 0, d_2$) such that $ac_1 - bd_1 > 0$ (*resp.* $ac_2 - bd_2 < 0$).

— Note that approaching the singular point along a line $t = (c + id)T$ in t-space is the same as approaching it along a real trajectory of the rotated equation $\frac{\mathrm{d}x}{\mathrm{d}T} = \lambda \exp(\mathrm{i}\theta) \times x$. Such a trajectory is a logarithmic spiral.

— All these properties hold for the original system too, since the vector field $P_\varepsilon \frac{\partial}{\partial x}$ is analytically linearizable near the singular point (Poincaré's theorem).

Locally around each root x_ε the squid sectors will coincide with domains bounded by asymptotic logarithmic spirals, given by trajectories of rotated vector fields $\exp(\mathrm{i}\theta_\varepsilon) P_\varepsilon \frac{\partial}{\partial x}$. The angular function $(\varepsilon, x) \in \mathcal{E}_\ell \times \mathbb{C} \mapsto \theta_\varepsilon(x) \in \left]-\frac{\pi}{4}, \frac{\pi}{4}\right[$ will be piecewise constant and zero outside a neighborhood of ∂K_ℓ, and for x far from the singular points.

DEFINITION 6.15. – Let $\mathcal{E}$ be a domain in the complement of Δ_k.

1. An *admissible angle* on $\mathcal{E}$ is a piecewise constant function $\theta:\ \mathcal{E}\times\mathbb{C}\to\left]-\frac{\pi}{4},\frac{\pi}{4}\right[$ such that for any analytic family of roots $(x_\varepsilon)_{\varepsilon\in\mathcal{E}}$ of $(P_\varepsilon)_{\varepsilon\in\mathcal{E}}$, the function $\varepsilon\in\mathcal{E}\mapsto\Re\left(P'_\varepsilon(x_\varepsilon)\exp\left(\mathrm{i}\theta_\varepsilon(x_\varepsilon)\right)\right)$ keeps a constant sign. In the following we use the notation

$$\vartheta:=\exp\left(\mathrm{i}\theta\right). \tag{6.13}$$

2. We say that an analytic family $(x_\varepsilon)_{\varepsilon\in\mathcal{E}}$ of singular points of $\dot{x}=P_\varepsilon(x)$ is of *node type* on $\mathcal{E}$ if there exists an admissible angle such that

$$\Re\left(P'_\varepsilon(x_\varepsilon)\,\vartheta_\varepsilon(x_\varepsilon)\right)>0\qquad(\forall\varepsilon\in\mathcal{E})$$

and of *saddle type* on $\mathcal{E}$ if

$$\Re\left(P'_\varepsilon(x_\varepsilon)\,\vartheta_\varepsilon(x_\varepsilon)\right)<0\qquad(\forall\varepsilon\in\mathcal{E}).$$

We use the notation $\left(x^{\mathrm{n}}_\varepsilon\right)_\varepsilon$ (resp. $\left(x^{\mathrm{s}}_\varepsilon\right)_\varepsilon$) for a family of roots of node (resp. saddle) type on the domain $\mathcal{E}$.

REMARK 6.16. – 1. The cells $\mathcal{E}_\ell$ in parameter space will be small contractible enlargements of the cones K_ℓ, on which there exist admissible angles. Additional constraints will be demanded to these angular functions in order to guarantee that the cells and sectors meet all technical requirements.

2. The choice of $\frac{\pi}{4}$ for an upper bound in the size of an admissible angle θ is arbitrary as any bound $\alpha\in\left]0,\frac{\pi}{2}\right[$ would do. However the larger α, the smaller the bound ρ on $\|\varepsilon\|$. Indeed we approach each singular point along a trajectory of some vector field $\vartheta_\varepsilon(x)\,P_\varepsilon(x)$. When θ is large and the singular points are not far enough from $r\mathbb{S}^1$, the trajectory follows wide spirals and may escape $r\mathbb{D}$ before landing at the singular point. An "absolute" (i.e., independent of the bound α) necessary condition for the existence of an admissible angle such that $(x_\varepsilon)_\varepsilon$ has node- (*resp.* saddle-) type on a neighborhood of K_ℓ is that $P'_\varepsilon(x_\varepsilon)\notin\mathbb{R}_{<0}$ (*resp.* $P'_\varepsilon(x_\varepsilon)\notin\mathbb{R}_{>0}$) for $\varepsilon\in K_\ell$. Therefore no admissible angle exists on a full pointed neighborhood of Δ_k.

3. We can illustrate on the formal normal form why admissible angles are of paramount importance. In the flow system of $\vartheta\widehat{\mathrm{X}}$ for real time

$$\begin{cases}\dot{x} &= \vartheta_\varepsilon(x)\,P_\varepsilon(x)\\ \dot{y} &= \vartheta_\varepsilon(x)\,y\left(1+\mu_\varepsilon x^k\right)\end{cases}$$

the variation of the modulus $\phi:=|y|^2=y\overline{y}$ of a solution follows the law

$$\dot{\phi}=2\phi\Re\left(\vartheta_\varepsilon(x)\left(1+\mu_\varepsilon x^k\right)\right).$$

Close enough to the singularity $(x_\varepsilon,0)$ all non-zero solutions therefore accumulate backwards exponentially fast on $(x_\varepsilon,0)$ if $x_\varepsilon=x^{\mathrm{n}}_\varepsilon$ is of node type or, on the contrary, diverge forwards exponentially fast for a saddle type root $x^{\mathrm{s}}_\varepsilon$. This behavior mimics that of a node / saddle planar foliation near a point with real residue $\vartheta_\varepsilon(x_\varepsilon)\,P'_\varepsilon(x_\varepsilon)$. This dynamical dichotomy is the cornerstone of the construction of the period operator (the modulus of classification) in [41].

6.3.4. *Size of sectors and of the parameter*. – The diameter ρ of the bounded part of the sectors is such that $|1+\mu x^k| > \frac{1}{2}$ when $|x| < \rho$. Note that the roots of P_ε all lie within $\sqrt{k}\,\|\varepsilon\|\operatorname{cl}(\mathbb{D})$. Indeed it suffices to show that if $|x| > \sqrt{k}\,\|\varepsilon\|$, then $P_\varepsilon(x) \neq 0$. On the one hand $\left|x^{k+1}\right| > k^{\frac{k+1}{2}}\,\|\varepsilon\|^{k+1}$. On the other hand

$$\left|\sum_{j=0}^{k-1}\varepsilon_j x^j\right| \leq \|\varepsilon\|^{k+1}\sum_{j=0}^{k-1}k^{\frac{j}{2}} \leq \|\varepsilon\|^{k+1}\,k^{\frac{k+1}{2}}.$$

In fact outside the disk $\sqrt{k}\,\|\varepsilon\|\operatorname{cl}(\mathbb{D})$ the trajectories of $P_\varepsilon\frac{\partial}{\partial x}$ are petals as depicted in Figure 6.1 (B). Set

$$\rho_\varepsilon := 2\sqrt{k}\,\|\varepsilon\|\,. \tag{6.14}$$

Then we choose ε sufficiently small so that $\rho_\varepsilon < \frac{\rho}{2}$. Later in Lemma 6.23 we will further reduce ρ and ε so that

$$\left|\mu x^k\right| + 2\rho\left|P''(x)\right| \leq \frac{3}{4} \tag{6.15}$$

for $|x| < \rho$.

6.3.5. *The ideal construction of sectors*. – Let us now choose a cone K_ℓ and describe the corresponding open squid sectors $\left(V^j_{\ell,\varepsilon}\right)_{j\in\mathbb{Z}/k\mathbb{Z}}$ covering $\rho\mathbb{D}\setminus P_\varepsilon^{-1}(0)$. On a "large" neighborhood of the spine of K_ℓ (to be made precise below), i.e., not too close to the boundary of K_ℓ, they are limited by real trajectories of $P_\varepsilon\frac{\partial}{\partial x}$ chosen as follows (see also Figure 6.4).

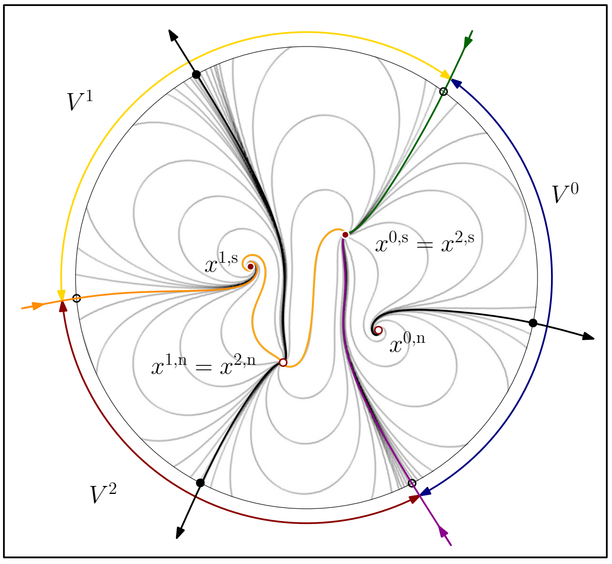


FIGURE 6.4. Curves involved in the ideal decomposition. Stable separatrices at ∞ in black, unstable ones in green, orange and purple.

1. The unstable separatrices of $P_\varepsilon \frac{\partial}{\partial x}$ through ∞ split $\rho\mathbb{S}^1$ into k arcs. We enlarge slightly these arcs to an open covering of the circle. Each arc is one piece of the boundary of a sector V_ε^j.
2. Two other pieces of the boundary of V_ε^j are given by the forward trajectories of $P_\varepsilon \frac{\partial}{\partial x}$ through the endpoints of the arc, which land in singular points $x^{j-1,\mathrm{s}}$ and $x^{j,\mathrm{s}}$ (not necessarily distinct) such that $\Re\left(P_\varepsilon'\left(x^{j,\mathrm{s}}\right)\right) < 0$ (i.e., the roots are of saddle type). These trajectories spiral as soon as $\Im\left(P_\varepsilon'\left(x^{j,\mathrm{s}}\right)\right) \neq 0$ (which is the generic situation).
3. Suppose $x^{j,\mathrm{s}} \neq x^{j-1,\mathrm{s}}$. For a given boundary arc of $\rho\mathbb{S}^1$ there exists one stable separatrix through ∞ which cuts it at one point and lands at root $x^{j,\mathrm{n}}$ of node type. This singular point belongs to the boundary of V_ε^j. The last two pieces of the boundary are two complete trajectories of $P_\varepsilon \frac{\partial}{\partial x}$, one joining $x^{j,\mathrm{n}}$ to $x^{j-1,\mathrm{s}}$ and the other joining $x^{j,\mathrm{n}}$ to $x^{j,\mathrm{s}}$. These trajectories are chosen in such a way that $\left(V_\varepsilon^j\right)_{j\in\mathbb{Z}/k\mathbb{Z}}$ cover $\rho\mathbb{D}\setminus P_\varepsilon^{-1}(0)$.
4. When $x^{j,\mathrm{s}} = x^{j-1,\mathrm{s}}$, we introduce two trajectories between $x^{j,\mathrm{s}}$ and $x^{j,\mathrm{n}}$, thus introducing a self-intersection of V_ε^j. This is motivated by the need of dealing with ramified functions near $x^{j,\mathrm{n}}$. See Figure 6.13 (A).

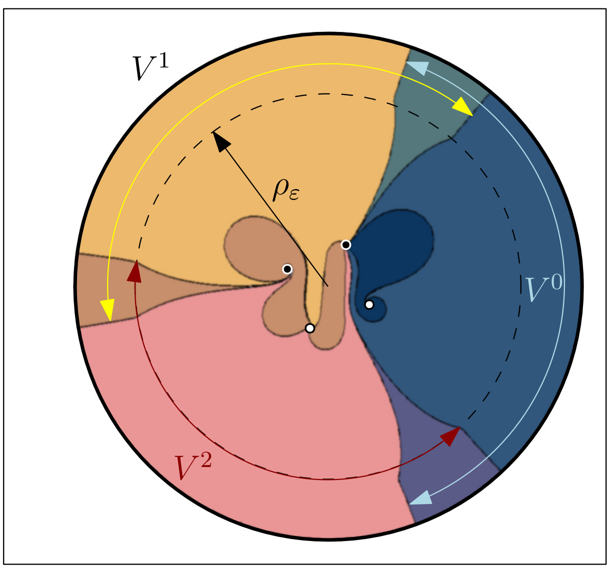


FIGURE 6.5. Decomposition into bounded, overlapping squid sectors induced by the flow depicted in Figures 6.2 and 6.4.

6.3.6. *The problem with the ideal construction of sectors.* – Of course the ideal construction will not always work. It can fail for the following reasons. For a set $I \subset P_\varepsilon^{-1}(0)$ and $\varepsilon \notin \Delta_k$ define

$$\nu_\varepsilon(I) := \sum_{x\in I} \frac{1}{P_\varepsilon'(x)}. \tag{6.16}$$

— The first one is when ε is not generic: the separatrices may form a homoclinic loop preventing them to land at singular points. A homoclinic loop γ through ∞ partitions the set of singular points $P_\varepsilon^{-1}(0)$ into I and I' such that

$$\Re\left(\nu_\varepsilon(I)\right) = \Re\left(\nu_\varepsilon\left(I'\right)\right) = 0. \tag{6.17}$$

— When ε is close to a hypersurface corresponding to a homoclinic loop, it can also occur that the trajectories through the endpoints of the arc first exit the disk $\rho\mathbb{D}$ before landing at a singular point.

— When ε crosses a hypersurface corresponding to a homoclinic loop, then $\Re\left(P_\varepsilon'(x_\varepsilon)\right)$ can change sign, thus preventing the above construction to be continuous in $\varepsilon \in \mathcal{E}_\ell$.

— As ε approaches 0 (or, more generally, Δ_k) we would like the sectors to converge to usual sectors associated to saddle-node singularities.

6.3.7. *The remedy in the construction of sectors.* – The remedy to all these problems is the same. We want to keep the above picture all over the cell $\mathcal{E}_\ell$ and we want the cells to cover the complement of Δ_k. The boundary of K_ℓ is composed of real hypersurfaces corresponding to homoclinic loop bifurcations. On each such hypersurface we have (6.17) for some I, while on K_ℓ the real part of the corresponding $\nu_\varepsilon(I)$ has a fixed sign and so does $\Im\left(\tau_j\right)$. But we have seen in Section 6.3.3 that this is not an obstruction for having the points remaining of node- or saddle-type: we just need to be sufficiently careful on how we approach them, by adjusting the spiraling of the sectors. In practice, this boils down to replacing the piece of a trajectory of $P_\varepsilon \frac{\partial}{\partial x}$ inside the disk $\rho_\varepsilon\mathbb{D}$ by the piece of a trajectory of $\exp(\mathrm{i}\theta) \times P_\varepsilon \frac{\partial}{\partial x}$ for some admissible angle θ as in Definition 6.15 (with some additional specifications).

PROPOSITION 6.17. – *Being given $\delta \in \left]0, \frac{\pi}{4}\right[$ and $\rho > 0$, there exists $\eta > 0$ such that the following properties hold.*

1. *Let $\mathcal{E}_\ell$ be the open set in $\{\|\varepsilon\| < \eta\} \setminus \Delta_k$ defined by the next conditions:*
   — *for each homoclinic-loop bifurcation hypersurface on the boundary of K_ℓ, separating the singular points in two nonempty groups $I \cup I'$ as in* (6.17)*, we have*
   $$\begin{cases} \arg \nu_\varepsilon(I) \in \left]-\frac{\pi}{2} - \delta, \frac{\pi}{2} + \delta\right[& \text{if } \Re\left(\nu_\varepsilon(I)\right) > 0 \text{ on } K_\ell, \\ \arg \nu_\varepsilon(I) \in \left]\frac{\pi}{2} - \delta, \frac{3\pi}{2} + \delta\right[& \text{if } \Re\left(\nu_\varepsilon(I)\right) < 0 \text{ on } K_\ell, \end{cases}$$
   — *for the τ_ε^j defined in Theorem 6.13 we have*
   $$\arg \tau_\varepsilon^j \in \left]-\delta, \pi + \delta\right[\quad \text{for all } j \in \{1, \ldots, k\}.$$
   Then $\mathcal{E}_\ell$ is a conic contractible neighborhood of K_ℓ and $\bigcup_\ell \mathcal{E}_\ell = \{\|\varepsilon\| < \eta\} \setminus \Delta_k$.
2. *There exists an admissible angle θ (corresponding to a direction $\vartheta = \exp(\mathrm{i}\theta)$) on $\mathcal{E}_\ell$ such that for each homoclinic-loop bifurcation hypersurface on the boundary of K_ℓ, separating the singular points in two nonempty groups $I \cup I'$, we have*

$$\begin{cases} \arg \sum_{x \in I} \frac{1}{\vartheta_\varepsilon(x) P_\varepsilon'(x)} \in \left]-\frac{\pi}{2} + \delta, \frac{\pi}{2} - \delta\right[& \text{if } \Re\left(\nu_\varepsilon(I)\right) > 0 \text{ on } K_\ell, \\ \arg \sum_{x \in I} \frac{1}{\vartheta_\varepsilon(x) P_\varepsilon'(x)} \in \left]\frac{\pi}{2} + \delta, \frac{3\pi}{2} - \delta\right[& \text{if } \Re\left(\nu_\varepsilon(I)\right) < 0 \text{ on } K_\ell. \end{cases} \tag{6.18}$$

3. *Any trajectory of $\vartheta_\varepsilon P_\varepsilon \frac{\partial}{\partial x}$, starting from a point of $\rho\mathbb{S}^1$ and entering the disk, does not exit the disk $\rho\mathbb{D}$ before landing at a singular point.*

Proof. – 1. This is clear.

2. We build the angle θ (piece-wise constant in x) in such a way that $\vartheta_\varepsilon = 1$ when ε is on the spine of K_ℓ. When we approach a component of ∂K_ℓ corresponding to a homoclinic loop separating the roots of P_ε as $I \cup I'$, we can rotate the vector field by an angle $|\theta| \leq 2\delta < \frac{\pi}{4}$ so that $\arg \sum_{x\in I} \frac{1}{\vartheta_\varepsilon(x) P'_\varepsilon(x)}$ belongs to the given interval.

3. A more precise quantitative description of the sectors is needed to show that the magnitude of ρ (in x-space) together with the choice of δ give constraints on the size η of the $\|\bullet\|$-ball in ε-space, and that taking $|\theta|$ large enough is sufficient to secure the conclusion. All this is done in the time coordinate $t = \int \frac{\mathrm{d}x}{P_\varepsilon(x)}$. We come back to this below in Section 6.3.9. $\square$

Definition 6.18. – 1. The contractile, conic domain $\mathscr{E}_\ell$ given by the previous proposition is called a *cell* in parameter space.

2. The k domains in x-space built like ideal sectors but bounded by trajectories of $\vartheta_\varepsilon P_\varepsilon \frac{\partial}{\partial x}$ instead of $P_\varepsilon \frac{\partial}{\partial x}$ are called *squid sectors*.

6.3.8. *Pairing sectors*

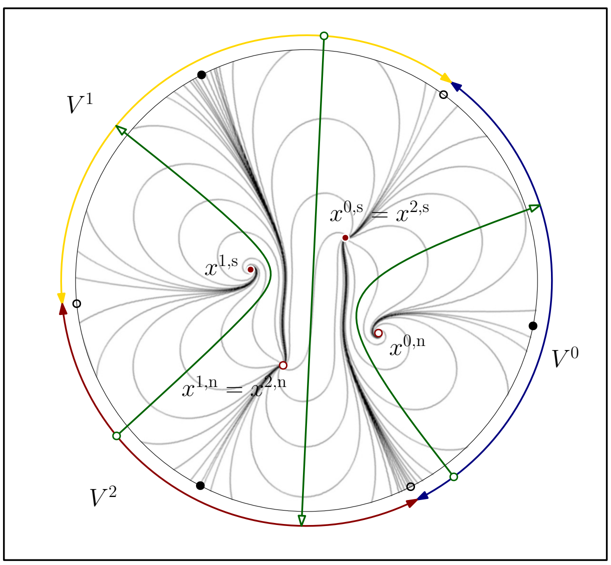


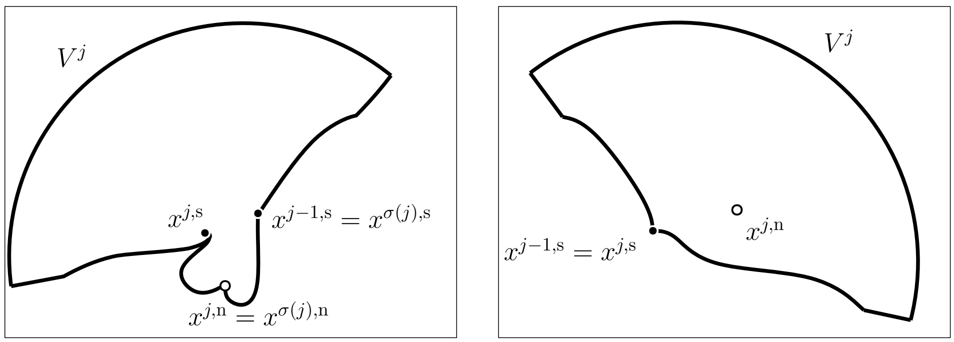
Figure 6.6. Construction of the non-crossing permutation σ; here $\sigma = \left(\begin{smallmatrix} 0 & 1 & 2 \\ 0 & 2 & 1 \end{smallmatrix}\right)$.

Definition 6.19. – Recall that a *non-crossing permutation* $\sigma \in \mathfrak{S}_k$ is a permutation such that if $p_0, \ldots, p_{k-1}$ are circularly ordered points on a circle, there exist pairwise non-intersecting curves within the inscribed disk joining p_j and $p_{\sigma(j)}$ for all j.

1. There exists a (non-crossing) permutation $\sigma = \sigma_\varepsilon$ on $\{0, \ldots, k-1\}$ yielding a pairing of the sector V_ε^j with $V_\varepsilon^{\sigma(j)}$ (see Figure 6.6) in the following way. If the sector V_ε^j shares its vertices $x^{j-1,\mathrm{s}}$ and $x^{j,\mathrm{n}}$ with a distinct sector $V_\varepsilon^{j'}$, then we define $\sigma(j) := j'$. Otherwise we let $\sigma(j) = j$.

2. The squid sector V_ε^j is *introvert* if $\sigma(j) = j$, and *extrovert* otherwise (see Figure 6.7).

The permutation σ is a complete topological invariant [11, 2] for structurally stable vector field $P\frac{\partial}{\partial x}$ (i.e., for generic ε) and any non-crossing permutation can be realized in this way. In particular $\varepsilon \mapsto \sigma_\varepsilon$ is constant on the conic domains K_ℓ.



(A) Extrovert squid sector: $\sigma(j) \neq j$

(B) Introvert squid sector

FIGURE 6.7. The two kinds of bounded squid sector for $k > 1$.

6.3.9. *Practical description and quantitative estimates.* – Here we end the proof of Proposition 6.17. As discussed earlier, finding an admissible angular function is equivalent to finding suitable piecewise affine real curves in the complex time coordinates. Studying $\dot{x} = P_\varepsilon(x)$ for complex values of the time t is the natural point of view taken by [11, 2]. In that setting we could view the whole x-line as a single complex trajectory of the flow of $P_\varepsilon\frac{\partial}{\partial x}$. Although one might consequently try to parametrize points in the x-variable by values of the time $t(x) \in \mathbb{C}$ this is too simplistic: the time function is multivalued at ∞. Nonetheless, the idea is very powerful and fruitful if we limit ourselves to simply connected domains in time space. Let us define the time function by

$$t(x) := \int_\infty^x \frac{\mathrm{d}z}{P_\varepsilon(z)}. \tag{6.19}$$

When ε is generic we obtain

$$t(x) = \sum_{x_\varepsilon \in P_\varepsilon^{-1}(0)} \frac{1}{P_\varepsilon'(x_\varepsilon)} \log(x - x_\varepsilon).$$

We are interested both in the map t and in its inverse $t^{\circ -1}$. As such the map (6.19) is not well defined, since it depends on the homotopy type of the path from ∞ to x in the integral. Hence its natural domain $\mathcal{S}$ is the Riemann surface given by the universal cover of $\mathbb{CP}^1$ punctured at the roots of P_ε. Since the integral starts at $x = \infty$ then $t(\infty) = 0$. We have to remember that ∞ is a pole of order $k-1$ of the vector field: there the time function has locally the form $t = -\frac{1}{kx^k}$ (and hence ∞ can be reached in finite time.) At all points of $\mathcal{S}$ different from ∞ the map t is locally biholomorphic, giving a structure of Riemann surface to the image $t(\mathcal{S} \setminus \{\infty\})$. Also, if we turn around ∞ once in x-space, then t will make k

turns around 0. Hence, the natural domain of $t^{\circ -1}$ is a ramified Riemann surface over the t-space with branch points of degree k at each of the images of ∞ by t over $\mathcal{S}$. We can consider it as a k-sheeted Riemann surface, and we change sheet when we turn around a branch point. There is a unique branch point when $\varepsilon = 0$. However, when $\varepsilon \neq 0$, there are *periods*, which correspond to the different times to go from ∞ to ∞ along paths circling some singular points. The distance between two images of ∞ (two branch points) on one sheet is a period of a loop around singular points. These periods are all greater than some $C\,\|\varepsilon\|^{-k}$ (see Lemma 6.20 below).

The image of the complement in $\mathbb{C}$ of $\rho\mathbb{D}$ under the map t is therefore, for small ε, a union of holes (topological disks) of approximate radius $\frac{1}{k\rho^k}$ in the k-sheeted Riemann surface (Figure 6.8) over $\mathbb{C}_t$, with one central hole around 0. The ramifications (branch points) occur at the images of ∞. Each hole contains an image of ∞ by t. A *half-sheet* around the central hole in t-space, i.e., a sector of opening π centered at the center of the hole and bounded by two horizontal half-lines, corresponds to an approximate angle of $\frac{\pi}{k}$ on $\partial\rho\mathbb{D}$ (or to a saddle sector of ∞). Hence one τ_j of Theorem 6.13 is associated to each half-sheet, thus pairing the half-sheets two by two. Since τ_j is a period in t-space, it is a distance between centers of holes and, on each half-sheet, the next hole is obtained by translating the current hole by τ_j.

LEMMA 6.20. – *There exists $C > 0$, depending only on k, such that*

$$|\tau_j| > C\,\|\varepsilon\|^{-k}\,.$$

Proof. – It suffices to show that there exists $C > 0$ such that $|\tau_j| > C$ when $\|\varepsilon\| = 1$, and then to use the rescaling (6.10). This is done as follows. Changing the time $t \mapsto t' := \mathrm{e}^{-\mathrm{i}\arg\tau_j}\,t$, then $\tau_j' = \mathrm{e}^{-\mathrm{i}\arg\tau_j}\,\tau_j$ is the time along a homoclinic loop between two separatrices of ∞ for the transformed vector field. In Section 6.3.4, all roots have been shown to belong to $\sqrt{k}\,\mathrm{cl}\,(\mathbb{D})$. The time τ_j' is then larger than twice the minimum time to go from ∞ to $\{|x| = 2k\}$, and this minimum is positive on the compact set $\|\varepsilon\| = 1$. □

Let us first describe what happens on the spine of the cell. There, holes are aligned vertically (the τ_j are pure imaginary) and each sector (which is an ideal sector) corresponds to a horizontal strip as in Figure 6.8. If we want to cover $\rho\mathbb{D}\setminus\{P_\varepsilon^{-1}(0)\}$ then we should cover a little more than a full turn around one hole. The width of the strip should be a little over $\frac{\tau_j}{2}$ on the top side and over $\frac{\tau_{j+1}}{2}$ on the bottom side. When moving to t-space the singular points have been sent to ∞, to the left (*resp.* right) for the singular points of node (*resp.* saddle) type. In such a picture we see the connected parts of the intersections of two consecutive sectors that go to the boundary.

The internal intersection parts (that we later call gate parts) can only be seen by using the periodicity of t. There are similar half strips on the $\sigma(j)$-th sheet, with a hole at a distance τ_j and on the $\sigma(j+1)$-th sheet, with a hole at a distance τ_{j+1}. Their translations by the corresponding period τ_j and $-\tau_{j+1}$ brings them on the j-th sheet where they intersect the initial strip (Figure 6.8).

If we now move away from the spine of K_ℓ, then two things happen.

— On the one hand, the τ_j bend. When they approach the real line (horizontal direction), then it is no more possible to pass a horizontal strip because the holes block the way: the remedy is to slant the strip so that it avoids the hole altogether.

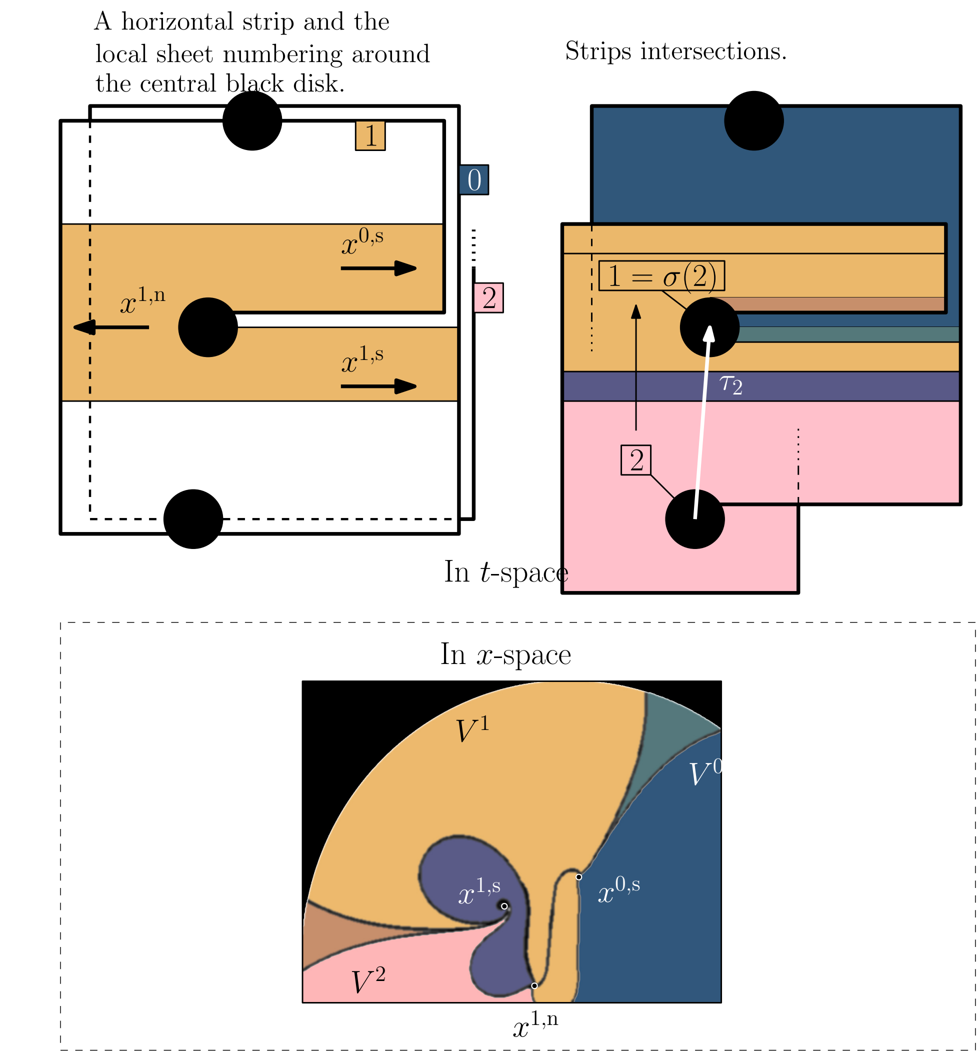


FIGURE 6.8. The images of these horizontal strips in t-space are sectors V_ε^j in x-space.

— On the other hand, in the t-space, each singular point x_ε turns, since it is located at infinity in the direction of $-\frac{1}{P_\varepsilon'(x_\varepsilon)}$. An infinite half-strip in the direction $\vartheta = \exp(\mathrm{i}\theta)$ can only be sent to a sector with vertex at x_ε if

$$\Re\left(-\frac{\vartheta}{P_\varepsilon'(x_\varepsilon)}\right) > 0 \tag{6.20}$$

(corresponding to the scalar product of $-\frac{1}{P_\varepsilon'(x_\varepsilon)}$ and ϑ being positive). This forces giving an angle to the strip in the infinite end of the half-strip approaching a singular point.

The choice of δ in (6.18) guarantees that τ_j cannot turn of an angle larger than $\frac{\pi}{2} + \delta$. The size of the holes is of the order of $\frac{1}{k\rho^k}$, which is very small compared to the τ_j and $\frac{1}{|P'_\varepsilon(x_\varepsilon)|}$ if ε is sufficiently small.

Now the strip has three infinite ends, a wide one on the left side attached to a point of node type $x_\varepsilon^{j,\mathrm{n}}$, and two thinner ones attached to $x_\varepsilon^{j-1,\mathrm{s}}$ and $x_\varepsilon^{j,\mathrm{s}}$. The slope ϑ_ε for each infinite end should be chosen so that (6.20) be satisfied for the singular point corresponding to that end of the strip.

This is how it is done. In the ideal situation the curves γ_j, used to pair the saddle sectors (permutation σ) and to define the τ_j, split the disk into $k+1$ regions, each containing a singular point. When we are no more in the ideal situation, then several of the curves γ_j have disappeared, corresponding to the fact that some strips are either too thin to pass a trajectory or have disappeared. Then there remains only a few γ_j dividing the disk in $m < k+1$ regions. Each of these regions contains some singular points. In a given region, we have two possibilities:

1. either there are several singular points: then they have kept their saddle or node type and are linked by trajectories that form a tree;
2. or there is a unique singular point, which is a center or a very widely spiraling focus.

For each γ_j that has disappeared because $\Im\left(\tau_j\right)$ is too small, we bend the strip between the holes while keeping its width a little more that $\frac{\tau_j}{2}$ (resp. $\frac{\tau_{j+1}}{2}$) (see Figure 6.9). This process restores that part of the strip and forces the bent separatrices to stay inside the disk.

Just before the disappearance of γ_j, each separatrix was attached to a singular point. If the singular point is close to a center as in (2) above, then the bent separatrix will spiral to the singular point: we may add a little more bending so that it does not escape the disk before doing so. In (1) the bent separatrix has no choice but to cross one of the trajectories of the tree between two singular points, one of which is the singularity to which it was attached before. When it does so, we turn to follow a parallel trajectory going to the singular point then bringing back the strip to the horizontal direction. We make the same thing for the three infinite ends of each strip. When doing so, we pay attention to take the same slope at all infinite ends attached to a given singular point.

Remark 6.21. – When $\varepsilon \to 0$ along a curve (6.12) then $P'_\varepsilon(x_\varepsilon) \to 0$ and the half-strips are replaced by half-planes. More generally when ε tends to a point of Δ_k, some half-strips are replaced by half-planes.

6.3.10. *Large (unbounded) squid sectors.* – When $\mu_0 \notin \mathbb{R}_{\leq 0}$, we will also need a covering of the whole of $\mathbb{C}$ by k sectors. For that purpose, we append to the sectors V_ε^j an infinite part obtained in the following way: if x_1 and x_2 are the endpoints of the boundary arc of V_ε^j along $\rho\mathbb{D}$, then we follow geometric spirals $x_m \exp\left((1+\mathrm{i}\nu)\mathbb{R}_{\geq 0}\right)$ for $m \in \{1,2\}$ and some ν such that

$$\Re\left(\mu_0\right) > \nu\Im\left(\mu_0\right).$$

If we come back to the representation of the sector in t-space, this amounts to appending some spiraling sector inside the holes (a neighborhood of ∞ in x-space is covered by a sector of opening $2k\pi$ in t-space).

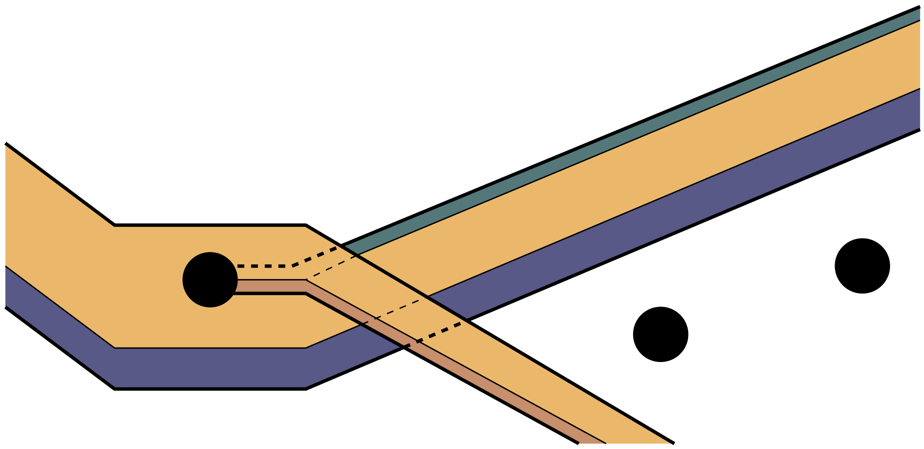


FIGURE 6.9. A slanted strip in t-space whose image is a sector V_ε^j in x-space.

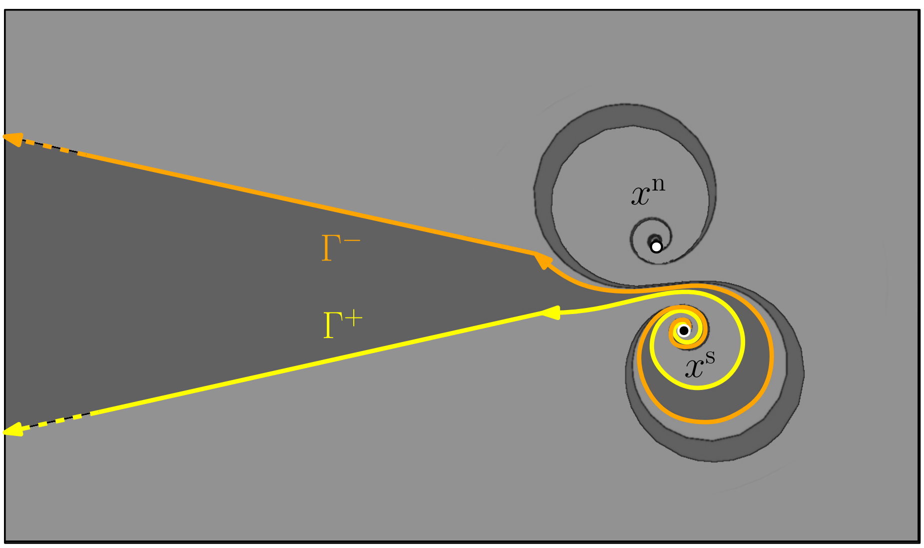


FIGURE 6.10. Unbounded squid sector for $k = 1$ and $\Re(\mu) > 0$. When $\Re(\mu) \leq 0$ and $\mu \notin \mathbb{R}$ the shaded area bends to form a geometric spiral near ∞. See also Figure 6.14 for the case $k > 1$.

We still denote by V_ε^j the resulting unbounded sectors, since the context will never be ambiguous.

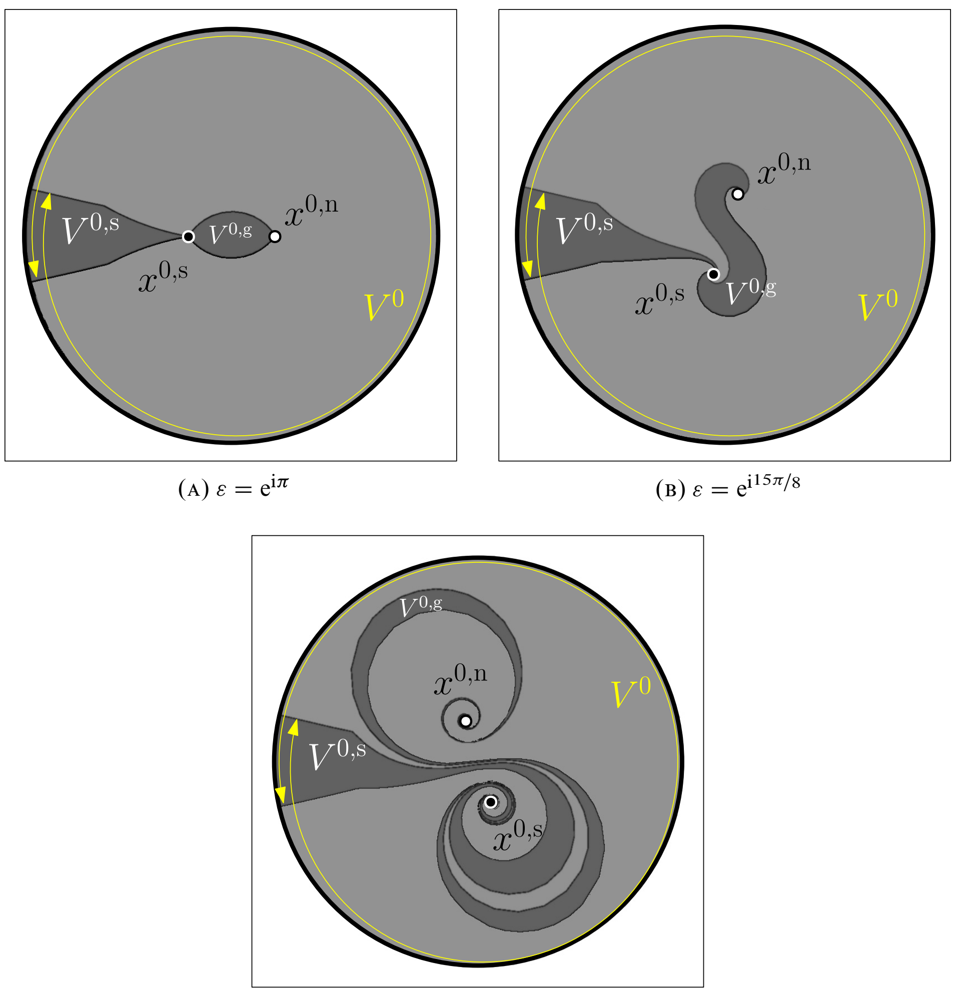
6.3.11. *Intersections of squid sectors*

DEFINITION 6.22. – We let $\Gamma^{j,+}$ (resp. $\Gamma^{j-1,-}$) be the part of the boundary of the unbounded sector V_ε^j joining $x^{j,\mathrm{s}}$ (resp. $x^{j-1,\mathrm{s}}$) to ∞ with this orientation. The intersection of two squid sectors V_ε^j and $V_\varepsilon^{j'}$ is made of up to three parts in general, and up to four parts when $k = 2$ (see Figure 6.13).

— If $j' = j + 1$ (resp. $j' = j - 1$) a (connected) *saddle part* $V_\varepsilon^{j,\mathrm{s}}$ (resp. $V_\varepsilon^{j-1,\mathrm{s}}$), bounded by the two curves $\Gamma^{j\pm}$ (resp. $\Gamma^{j-1,\pm}$) to the common point $x^{j,\mathrm{s}}$ (resp. $x^{j-1,\mathrm{s}}$) of saddle type. When $k = 1$, the saddle-part corresponds to a self-intersection.

— If $j' = \sigma(j)$ a *gate part* $V_\varepsilon^{j,\mathrm{g}}$ included in $\rho_\varepsilon\mathbb{D}$ and adherent to the two singular points $x^{j,\mathrm{s}}$ and $x^{j,\mathrm{n}}$. When $j = \sigma(j)$, the gate part of an introvert sector corresponds to a self-intersection.

— If $j' = \sigma(j)$ and $j = \sigma(j')$ for $j \neq j'$, a second gate part $V_\varepsilon^{j',\mathrm{g}}$ adherent to the singular points $x^{j-1,\mathrm{s}} = x^{j',\mathrm{s}}$ and $x^{j,\mathrm{n}}$ (Figure 6.13 (B)).



(A) $\varepsilon = \mathrm{e}^{\mathrm{i}\pi}$ (B) $\varepsilon = \mathrm{e}^{\mathrm{i}15\pi/8}$

(C) $\varepsilon = \mathrm{e}^{\mathrm{i}17\pi/8}$

FIGURE 6.11. Squid sectors for different values of ε when $k = 1$.

6.3.12. *Non-equivalent decompositions.* – For the same value of the parameter ε in the intersection of two cells (or a cell's self-intersection), the disk $\rho\mathbb{D}$ is split in non-equivalent ways into bounded squid sectors (see Figures 6.12 and 6.13). By "non-equivalent" we mean that at least one boundary of a squid sector is attached to another root of P_ε when passing from one cell to the other.

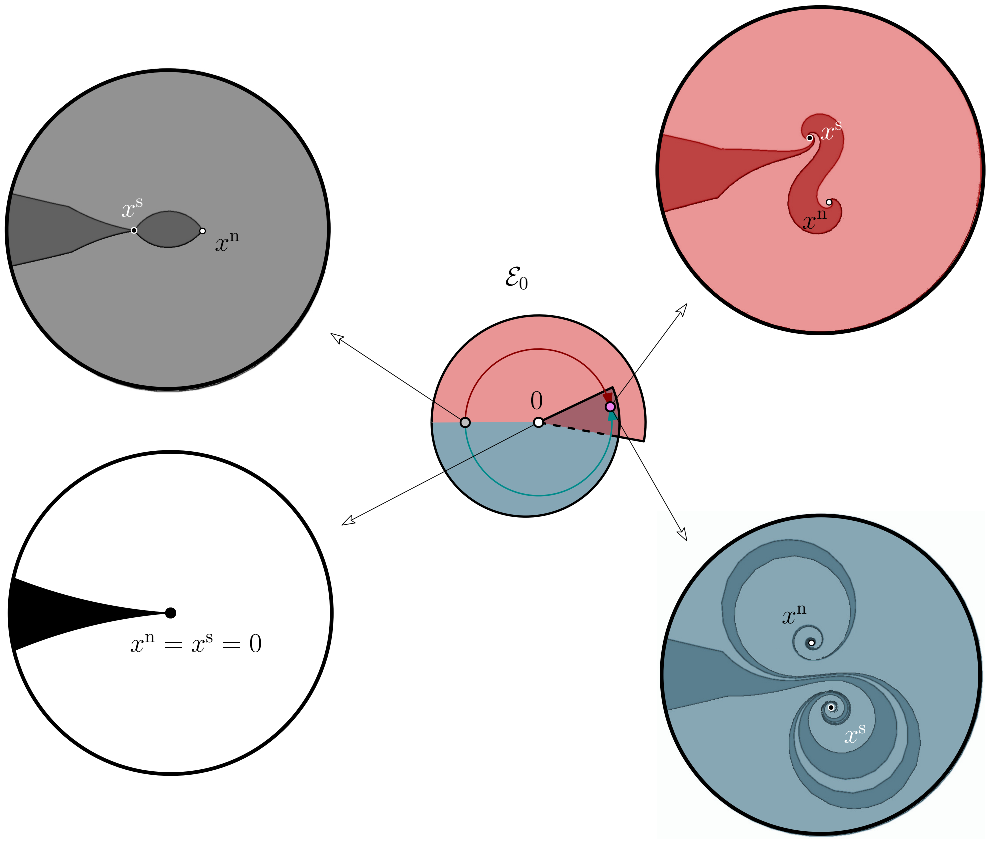
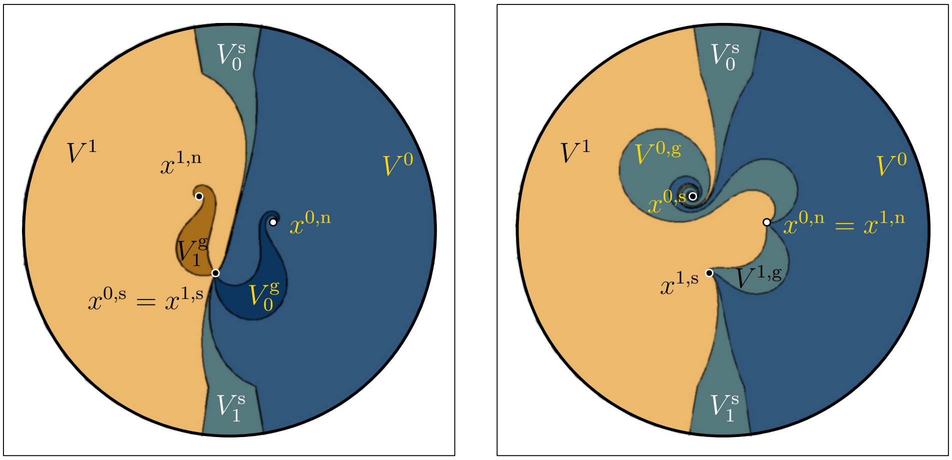


FIGURE 6.12. The single (self-overlapping) cell $\mathcal{E}$ with diverse configurations when $k = 1$. Non-equivalent decompositions are shown on the right. In each picture the location of the node-like singularity x^{n} is given by the analytic continuation of the principal determination of $\sqrt{-\varepsilon}$.

6.3.13. *Some useful estimates.* – We shape the squid sectors in this way because in doing so we gain control on the convergence and on the magnitude of integrals involved in the Cauchy-Heine transform appearing in the next section, in the wake of Remark 6.16. In the following lemma we use the boundary $\Gamma^{j,\pm}$ of saddle-parts of unbounded sectors as depicted in Figure 6.10.

LEMMA 6.23. – *Assume $\tau = 0$ (which particularly implies $\mu_0 \notin \mathbb{R}_{\leq 0}$). One can take ρ and $\mathcal{E}_\ell$ sufficiently small so that the following properties hold.*



(A) With two introvert squid sectors: $\sigma = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$

(B) With two extrovert squid sectors: $\sigma = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

FIGURE 6.13. Non-equivalent decompositions for same ε when $k = 2$.

1. *For all $r > 0$ the model first integral* (6.7) *is bounded on $V_\varepsilon^{j,\mathrm{s}} \times r\mathbb{D}$, more precisely there exists $C > 0$ such that*
$$(\forall \varepsilon \in \mathcal{E}_\ell) \qquad \sup_{V_\varepsilon^{j,\mathrm{s}} \times r\mathbb{D}} \left| \widehat{H}^j \right| \leq rC.$$

2. *Also $\widehat{H}^j$ is $\frac{\mathrm{d}z}{z-x}$-absolutely integrable over any component $\Gamma = \Gamma^{j,\pm}$ of the boundary of saddle part intersections (given the outgoing orientation): for all $x \in V_\varepsilon^j \backslash \Gamma$ and $y \in \mathbb{C}$ we have*
$$\int_\Gamma \widehat{H}^j(z, y) \frac{\mathrm{d}z}{z - x} =: y I^j(x) \in \mathbb{C}.$$

3. *There exists a constant $C > 0$ such that for all $\varepsilon \in \mathcal{E}_\ell$ and all $x \in V_\varepsilon^j \backslash \Gamma$*
$$\left| I^j(x) \right| \leq \frac{C}{|z_* - x_*|},$$
where $z_* = \Gamma \cap \rho\mathbb{S}^1$ and x_* is likewise the intersection of $\rho\mathbb{S}^1$ and the curve passing through x built in the same way as Γ.

Proof. – Because $\widehat{H}^j$ is linear in y we may only consider the case $y := 1$. Let
$$\widehat{h} : x \mapsto \widehat{H}^j(x, 1)$$
be the corresponding partial function. The proof is done in two steps, corresponding to the two different components "inner" (inside $\rho\mathbb{D}$) and "outer" ($|x| \geq \rho$). We parametrize Γ by a piecewise analytic curve $z : \mathbb{R} \to \mathbb{C}$ detailed below, such that (with the obvious abuse of

notations)

$$\begin{cases} z(-\infty) & = \infty \\ z(0) & = z_{\pm}^{j} \in \rho\mathbb{S}^{1} \\ z(\infty) & = x^{j,\mathrm{s}} \end{cases}.$$

In what follows, $C > 0$ indicates a real constant (independent on ε) whose value varies according to the place where it appears.

1. We invoke again the variational argument presented in Remark (6.16). Over $]0,\infty[$ we follow the flow of $\vartheta P\frac{\partial}{\partial x}$ and we can indeed estimate the modulus
$$\phi(t) := \left|\widehat{h}(z(t))\right|,$$
as $\widehat{h}$ is solution of
$$\frac{\mathrm{d}\widehat{h}}{\widehat{h}} = -\left(1+\mu z^{k}\right)\frac{\mathrm{d}z}{P_{\varepsilon}},$$
so that
$$\frac{\dot{\phi}}{\phi}(t) = -\Re\left(\vartheta\left(1+\mu z^{k}\right)\right).$$
Since $\frac{1}{2|\mu_0|} > \rho^{k}$, and taking the hypothesis $|\arg\vartheta| < \frac{\pi}{4}$ into account we obtain
$$\frac{\dot{\phi}}{\phi} \leq -C < 0$$
and
$$\left|\widehat{h}(z(t))\right| \leq \left|\widehat{h}(z(t_{\varepsilon}))\right|. \tag{6.21}$$
Over $]-\infty,0[$ we follow the flow of
$$\dot{z} = -(1+\mathrm{i}\nu)z,$$
above which the modulus of $\widehat{h}$ is governed by
$$\begin{aligned}\frac{\dot{\phi}}{\phi} &= \Re\left(\frac{\left(1+\mu z^{k}\right)(1+\mathrm{i}\nu)z}{P_{\varepsilon}(z)}\right)\\ &= \Re\left(\frac{\left(1+\mu z^{k}\right)(1+\mathrm{i}\nu)}{z^{k}}\times\frac{z^{k+1}}{P_{\varepsilon}(z)}\right)\\ &\geq C\Re(\mu+\mathrm{i}\mu\nu) =: \alpha > 0\end{aligned}$$
for $|z|$ sufficiently large since ν is chosen in such a way that $\Re(\mu+\mathrm{i}\nu\mu) > 0$ for all $\varepsilon \in \mathcal{E}_{\ell}$. To conclude the proof we only have to remark that $\sup_{|z|=\rho}\left|\widehat{h}(z)\right| \geq \left|\widehat{h}(z(0))\right|$ converges uniformly towards $\sup_{|z|=\rho}\left|\widehat{H}_0^{j}(z,1)\right| < \infty$ as $\varepsilon \to 0$.
2. and (3) We use the following trick. We work with the integral
$$J^{j}(x) = \int_{-\infty}^{\infty}\widehat{h}(z(t))\frac{z'(t)dt}{z(t)-x(t)},$$

where $t \mapsto x(t)$ is defined similarly as $t \mapsto z(t)$ except for the fact that it passes through x, uniquely defining $x_* = x(0)$. To conclude we will need to bound away from 0 (uniformly in ε) the quantity $\left|\frac{z(t)-x}{z(t)-x(t)}\right|$. But this is clear from the pictures because if $\operatorname{dist}(x, \Gamma)$ is realized for $z = z(t)$ then $x \simeq x(t)$.

Now, to study $J^j(x)$ we repeat the above argument but with the function

$$\phi_\sharp(t) := \left|\frac{\widehat{h}(z(t))\, A_\sharp(z(t))}{z(t) - x(t)}\right|, \quad \sharp \in \{0, \infty\},$$

where $A_0(z) := \vartheta P_\varepsilon(z)$ and $A_\infty(z) := -(1 + i\nu)z$. The variations of $\phi_\sharp$ are governed by

$$\frac{\dot{\phi}_\sharp}{\phi_\sharp}(t) = \Re\left(-\frac{1 + \mu z^k}{P_\varepsilon(z(t))} A_\sharp(z(t)) + A'_\sharp(z(t)) - \frac{A_\sharp\big(z(t) - A_\sharp(x(t))\big)}{z(t) - x(t)}\right)$$

for t in the corresponding interval so that $\dot{z} = A_\sharp(z)$ and $\dot{x} = A_\sharp(x)$. In the case $\sharp = \infty$, the sum of the last two terms vanishes and then

$$\frac{\dot{\phi}_\infty}{\phi_\infty} \geq C > 0$$

for large z (hence t close to $-\infty$) from the choice of ν. Let us now deal with the case $\sharp = 0$. We have chosen $\rho > \rho_\varepsilon$ so that

$$\sup_{|z|<\rho}\left(\left|\mu z^k\right| + 2\rho\left|P''(z)\right|\right) \leq \frac{3}{4}.$$

Because for all $x, z \in \rho\mathbb{D}$

$$\left|P(x) - P(z) - (x - z)P'(z)\right| \leq |x - z|^2 \sup_{\rho\mathbb{D}}\left|P''\right|$$

we obtain

$$\frac{\dot{\phi}_0}{\phi_0} \leq -C < 0$$

and

$$|\phi_0(t)| \leq |\phi_0(0)| \exp(-Ct)$$

for $t \geq 0$.

Therefore the integral

$$\int_{z(0)}^{z(t)} \widehat{h}(z)\frac{\mathrm{d}z}{z - x} = \vartheta \int_0^t \widehat{h}(z(t))\frac{P_\varepsilon(z(t))}{z(t) - x}\mathrm{d}t$$

is absolutely convergent as $t \to \infty$ and

$$\left|\int_{z(0)}^{z(\infty)} \widehat{h}(z)\frac{\mathrm{d}z}{z - x}\right| \leq C\,|\phi_0(0)|.$$

But $C\,|\phi_0(0)| \leq \frac{C}{|x(0) - z(0)|}$ as expected. □

6.4. Cellular section of the period: proof of Proposition 6.6

The cellular section $\mathfrak{S}_\ell$ of the period operator is obtained from a variation on the method introduced in Section 5.4 to normalize the glued abstract manifold by solving a linear Cousin problem. It is an unfolding of the technique used in [43] for $\varepsilon = 0$. The initial data is a k-tuple

$$T = \left(T^j\right)_j \in \prod_{\mathbb{Z}/k\mathbb{Z}} \mathscr{H}_\ell \{h\}$$

and we seek $Q \in x\,\mathscr{H}_\ell \{y\} [x]_{<k}$, that is

$$Q(x, y) = x \sum_{n>0} Q_n(x)\, y^n$$

for some polynomial $Q_n \in \mathrm{Holo_c}(\mathcal{E}_\ell)[x]_{<k}$ in x of degree less than k, such that

$$\mathfrak{T}_\ell(Q) = T.$$

We then define the section as

$$\mathfrak{S}_\ell(T) := Q.$$

The construction goes along the following steps. They are performed for fixed ε in a fixed $\mathcal{E}_\ell$, with explicit control on the parametric regularity. Hence we omit mentioning explicitly the dependence on ε and ℓ. For $r > 0$ define

$$\mathcal{V}_r^j := \left\{(x, y) \in V^j \times \mathbb{C} :\ |y| < r\right\}.$$

We define in a similar fashion the fibered intersections $\mathcal{V}_r^{j,\sharp}$ for $\sharp \in \{\mathrm{s}, \mathrm{g}\}$.

— Build sectorial, bounded functions F^j on $\mathcal{V}_r^j$ such that

$$F^{j+1} - F^j = 2\mathrm{i}\pi T^j \circ H^j \tag{6.22}$$

on $\mathcal{V}_r^{j,\mathrm{s}}$, where H^j is the j-th canonical sectorial first integral of $\mathscr{X}$, as in (6.6). This is done again by a Cauchy-Heine transform (Section 6.4.1).

— Because of the functional Equation (6.22) the identity $\mathscr{X} \cdot F^{j+1} = \mathscr{X} \cdot F^j$ holds and allows to patch together a holomorphic function $Q := \mathscr{X} \cdot F^j$ on a whole $\mathbb{C} \times (\mathbb{C}, 0)$ which, by construction, satisfies

$$\begin{aligned} \mathfrak{T}^j(Q) \circ H^j &= F^{j+1} - F^j \\ &= T^j \circ H^j \end{aligned}$$

(Section 6.4.2).

— Growth control near $x = \infty$ and a final normalization allows concluding that $Q \in x\mathbb{C}\{y\}[x]_{<k}$ (Section 6.4.3).

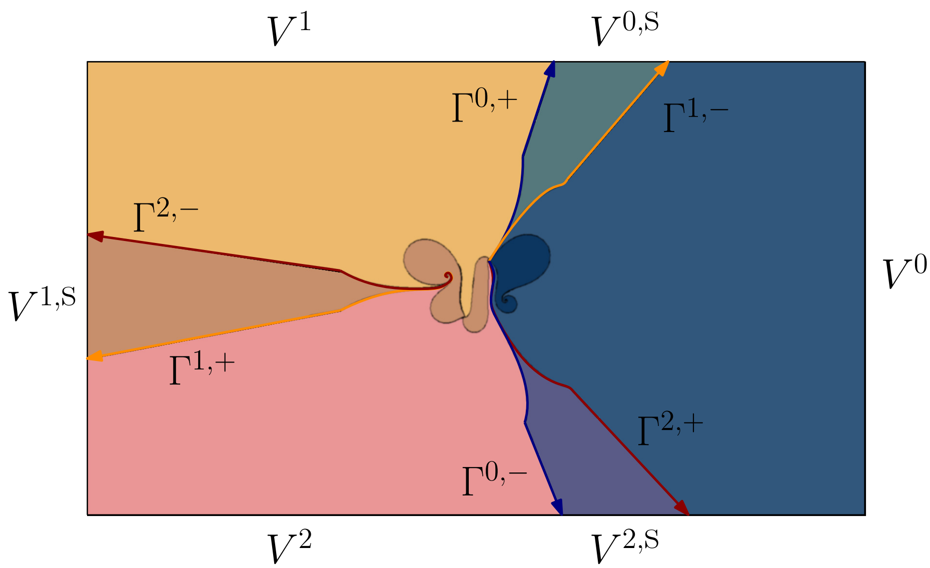


FIGURE 6.14. Unbounded squid sectors and paths of integration ($k > 1$).

6.4.1. *Cauchy-Heine transform*

DEFINITION 6.24. – In the following we fix a collection $N = \left(N^j\right)_j \in \prod_{j\in\mathbb{Z}/k\mathbb{Z}} \mathrm{Holo_c}\left(\mathcal{V}_r^j\right)'$, which is a k-tuple of functions with an expansion

$$N^j(x, y) = \sum_{n>0} N^{j,n}(x)\, y^n$$

uniformly absolutely convergent on every $\mathcal{V}_{r'}^j$ for all $0 \leq r' < r$, whose norm is given by

$$\|N\| := \max_j \sup_{\mathcal{V}_r^j} \left|N^j\right|.$$

1. We define the j-th *sectorial first integral associated to* N as the holomorphic function
$$\begin{aligned} H_N^j :\ \mathcal{V}_r^j &\longrightarrow \mathbb{C} \\ (x, y) &\longmapsto \widehat{H}^j(x, y) \exp N^j(x, y), \end{aligned}$$
where $\widehat{H}^j$ is the sectorial canonical model first integral (6.7) continued over unbounded squid sectors.
2. For a given $\eta > 0$ we say that N is η-*adapted* if $H_N^j\left(\mathcal{V}_r^{j,s}\right) \subset \eta\mathbb{D}$.

Of course we prove in due time (Corollary 7.7) that $N := N_\varepsilon$, defined as the collection of sectorial solutions of the normalizing equation $\mathcal{X}_\varepsilon \cdot N_\varepsilon^j = -R_\varepsilon$, satisfies the hypothesis of the definition and that $\sup\left|H_N^j\left(\mathcal{V}_r^{j,s}\right)\right| \to 0$ as $r \to 0$ (uniformly in $\varepsilon \in \mathcal{E}_\ell$), mainly because it is already the case for the model first integral (Lemma 6.23 (1)). Therefore, for given $\eta > 0$, it will always be possible to find r (independently on ε) such that N is η-adapted, allowing us to use the next result, genuinely the key point in building the cellular section of the period.

Proposition 6.25. – *Assume $\tau = 0$ (which particularly implies $\mu_0 \notin \mathbb{R}_{\leq 0}$). Let $\mathcal{E}_\ell$ be a fixed cell as in Section 6.3. For every $T \in \prod_{\mathbb{Z}/k\mathbb{Z}} \mathrm{Holo_c}\,(\eta\mathbb{D})'$ holomorphic on a disk of radius $\eta > 0$, for every η-adapted collection N, the k-tuple of functions*

$$\mathfrak{F} = \mathfrak{F}(T, N) := \left(F^j\right)_j \in \prod_{j \in \mathbb{Z}/k\mathbb{Z}} \mathrm{Holo_c}\left(\mathcal{V}_r^j\right)'$$

defined by

$$F^j(x,y) := \sum_{p \neq j+1} \int_{\Gamma^{p,-}} \frac{T^{p-1}\left(H_N^{p-1}(z,y)\right)}{z-x}\mathrm{d}z + \int_{\Gamma^{j,+}} \frac{T^j\left(H_N^j(z,y)\right)}{z-x}\mathrm{d}z \tag{6.23}$$

fulfills the next conclusions. The paths of integration $\Gamma^{j,\pm}$ bound the unbounded squid sectors in the following way: the boundary of the saddle part $V^{j,\mathrm{s}}$ of (unbounded) squid sectors is $\Gamma^{j,+} \cup \Gamma^{j+1,-}$, as in Figures 6.10 and 6.14, and we set

$$\left\|T'\right\| := \max_j \sup_{\eta\mathbb{D}} \left|\frac{\mathrm{d}T^j}{\mathrm{d}h}\right|.$$

1. *For every $(x,y) \in \mathcal{V}_r^{j,\mathrm{s}}$*

$$F^{j+1}(x,y) - F^j(x,y) = 2\mathrm{i}\pi T^j\left(H_N^j(x,y)\right) \tag{6.24}$$

   while for every $(x,y) \in \mathcal{V}_r^{\sigma(j),\mathrm{g}}$

$$F^j(x,y) = F^{\sigma(j)}(x,y).$$

   (When $k = 1$ we refer to (3) of the following remark for a fuller explanation.)

2. $F^j \in \mathrm{Holo_c}\left(\mathcal{V}^j\right)'$.
3. *There exists $K > 0$ independent on T, N, r and ε such that the following estimates hold.*

   (a) $$\|\mathfrak{F}\| \leq rK\left\|T'\right\| \exp\|N\|.$$

   (b) $$\left\|y\frac{\partial\mathfrak{F}}{\partial y}\right\| \leq rK\left\|T'\right\| \left\|1 + y\frac{\partial N}{\partial y}\right\| \exp\|N\|.$$

   (c) $$\left\|x\frac{\partial\mathfrak{F}}{\partial x}\right\| \leq rK\left\|T'\right\| \left\|1 + x\frac{\partial N}{\partial x}\right\| \exp\|N\|.$$

Remark 6.26. – 1. The absolute convergence of the integrals involved in (6.23) is established in the course of the proof, mainly thanks to the estimates given by Lemma 6.23. Notice also that for fixed ε and y the mapping $x \mapsto F^j(x,y)$ is holomorphic on V^j since the squid sector does not contain any of the curves $\Gamma^{p,-}$ except for $p = j+1$.

2. The integral expression (6.23) and Item (3) above clearly show that $\mathfrak{F}$, as a function of $\varepsilon \in \mathcal{E}_\ell$, has the same regularity as T.

3. In the case $k = 1$ the expression (6.23) yields $F(x,y) = \int_{\Gamma^+}(\cdots)\,\mathrm{d}z$, which can be analytically continued in the x-variable on the self-overlapping squid sector (Figure 6.10). As x reaches Γ^- "from below" the analytic continuation coincides with $\int_{\Gamma^-}(\cdots)\,\mathrm{d}z$, because the difference of determination is given by

$$D(x,y) := \int_{\Gamma^+ - \Gamma^-} \frac{T(H_N(z,y))}{z-x}\mathrm{d}z,$$

and Cauchy's formula asserts that $D(x,y) = 0$ whenever x is outside the saddle-part V^{s} enclosed by $\Gamma^{+} \cup \Gamma^{-}$. On the contrary if $x \in V^{\mathrm{s}}$ then $D(x,y) = 2\mathrm{i}\pi T(H_N(x,y))$, which is the way to understand (6.24).

Proof. – This proposition follows the general lines of [43, Theorem 2.5] for $\varepsilon = 0$. A simpler instance of the strategy can be found in Lemma 5.6. Except when necessary we drop every sub- and super-scripts.

1. This is nothing but Cauchy residue formula. We indeed compute (omitting to include the integrand for the sake of readability)

$$\begin{aligned} F^{j+1}(x,y) - F^{j}(x,y) &= \int_{\Gamma^{j+1,+}} - \int_{\Gamma^{j,+}} + \sum_{p \neq j+2} \int_{\Gamma^{p,-}} - \sum_{p \neq j+1} \int_{\Gamma^{p,-}} \\ &= \left(\int_{\Gamma^{j+1,-}} - \int_{\Gamma^{j,+}} \right) - \left(\int_{\Gamma^{j+2,-}} - \int_{\Gamma^{j+1,+}} \right). \end{aligned}$$

The candidate singularity in the common integrand $\frac{T^p(H_N^p(z,y))}{z-x}$ in $\int_{\Gamma^{p+1,-}} - \int_{\Gamma^{p,+}}$ is $z = x$. This happens only when $x \in V^{p,\mathrm{s}}$. By hypothesis $x \in V^{j,\mathrm{s}}$ hence (6.24) holds.

Actually one needs to use a growing family of compact loops within $V^{j,\mathrm{s}}$ converging toward $\partial V^{j,\mathrm{s}}$, then to apply Cauchy formula to each one of them and take the limit. The only possible choice for the connected component of $\mathbb{C} \backslash \left(\Gamma^{j+1,-} \cup \Gamma^{j,+} \right)$ for which this construction works is $V^{j,\mathrm{s}}$, since in that sector we can establish tame estimates for the growth of the integrand (see (3) below), and we can also establish untamed estimates outside a neighborhood of $\mathrm{cl}\left(V^{j,\mathrm{s}}\right)$.

2. Taking for granted that the integrand defining $F(x,y)$ for $(x,y) \in \mathrm{cl}\left(\mathcal{V}_r\right)$ is bounded from above by a real-analytic, integrable function on ∂V^{s}, the analyticity of F on $\mathcal{V}_r$ is clear from the Definition (6.23). Integration paths used to evaluate F can be slightly deformed outwards without changing the value of the integral, which shows that F can be analytically continued to any point (x,y) with $x \in \partial V \backslash P_{\varepsilon}^{-1}(0)$ and $|y| \leq r$. Concluding that F extends as a continuous function on $\mathrm{cl}\left(\mathcal{V}_r\right) \backslash P_{\varepsilon}^{-1}(0)$ is again a consequence of (6.23) for y is an extraneous parameter. Dominated convergence of $F(x,y)$, continuity on $\mathrm{cl}(\mathcal{V}_r) \cap P_{\varepsilon}^{-1}(0)$ and boundedness of F are established in (3).

3. We begin with proving (a). Since, for $p \in \mathbb{Z}/k\mathbb{Z}$,

$$\left| T^p(h) \right| \leq |h| \left\| T' \right\|,$$

we deduce

$$\left| \frac{T^p(H(z,y))}{z-x} \right| \leq \frac{\left| \widehat{H}(z,y) \right|}{|z-x|} \left\| T' \right\| \exp \|N\|.$$

We then invoke the estimates derived for the model family in Lemma 6.23, showing dominated convergence for $F(x,y)$. In order to bound F it is sufficient to consider only the problem of bounding F near a single $\Gamma := \Gamma^{j,+}$. A uniform bound K for the rightmost sum of integrals simply requires bounding uniformly $\frac{1}{|z_* - x_*|}$ where $z_*, x_* \in \rho\mathbb{S}^1$. Of course no uniform bound in x exists when x tends to Γ (i.e., x_* tends to z_*). To remedy this problem we bisect V^{s} with a curve $\widehat{\Gamma}$ parallel to Γ and passing through the middle of the arc $\rho\mathbb{S}^1 \cap V^{\mathrm{s}}$. When x is taken in the component of $V^{j} \backslash \widehat{\Gamma}$ not accumulating on Γ the value of $\frac{1}{|z_* - x_*|}$ is uniformly bounded. When x is taken in the other part we use the functional relation (6.24):

in that configuration x is understood as an element of V^{j+1} far from $\Gamma^{j+1,-}$ and we are back to the situation we just solved.

A little bit more detailed analysis allows proving that $x \mapsto F(x,y)$ is Cauchy [(2)] near $x^{j,s}$, so that F extends continuously to $\left\{x^{j,s}\right\} \times r\mathbb{D}$. Items (b) and (c) are obtained much in the same way, the details are straightforward adaptations of (a). $\square$

6.4.2. *Holomorphy of Q_ε.* – Now all functions $\mathcal{X}_\varepsilon \cdot \mathfrak{F}^j$ patch on intersecting squid sectors to define

$$Q \in \operatorname{Holo}\left(\left(\mathbb{C} \backslash P_\varepsilon^{-1}(0)\right) \times r\mathbb{D}\right).$$

If we show that Q is bounded near each disk $\left\{x^{j,s}\right\} \times r\mathbb{D}$ then Riemann's theorem on removable singularities guarantees the holomorphic extension of Q to $\mathbb{C} \times \mathbb{D}$. But

$$|Q(x,y)| \leq |P_\varepsilon(x)| \left\| \frac{\partial F}{\partial x} \right\| + \left(1 + |\mu|\,|x|^k + |R(x,y)|\right) \left\| y \frac{\partial F}{\partial y} \right\| \tag{6.25}$$

so that taking Proposition 6.25 (3) into account brings the conclusion.

6.4.3. *Growth control of Q_ε near $x = \infty$.* – In Section 7.2 we prove that the k-tuple of sectorial solutions N of the cohomological equation of normalization $\mathcal{X} \cdot N^j = -R$ satisfies the conditions $\left\| x \frac{\partial N}{\partial x} \right\| \leq \frac{1}{3}$ and $\left\| y \frac{\partial N}{\partial y} \right\| \leq \frac{1}{3}$ if r is chosen small enough (Corollary 7.7).

LEMMA 6.27. – *For every fixed $y \in r\mathbb{D}$ the entire function $x \mapsto Q(x,y)$ is actually a polynomial of degree at most k, and*

$$Q(x,y) = \sum_{n>0} q_n(x)\, y^n, \quad q_n \in \mathbb{C}[x]_{\leq k} \tag{6.26}$$

on $\mathbb{C} \times r\mathbb{D}$.

Proof. – Since $x \mapsto R(x,y)$ is a polynomial of degree at most k, there exists a constant $C > 0$ such that $1 + |\mu|\,|x|^k + |R(x,y)| \leq C\,|x|^k$ for every $|x| \geq \rho$. The bound (6.25) on $x \mapsto Q(x,y)$ also holds near ∞ so that

$$|Q(x,y)| \leq \left| \frac{P_\varepsilon(x)}{x} \right| \left\| x \frac{\partial F^j}{\partial x} \right\| + C\,|x|^k \left\| y \frac{\partial F}{\partial y} \right\|.$$

From Prop. 6.25(3)(b,c) and the control on $\left\| x \frac{\partial N^j}{\partial x} \right\|, \left\| y \frac{\partial N^j}{\partial y} \right\|$ we infer $\left\| x \frac{\partial F^j}{\partial x} \right\|, \left\| y \frac{\partial F^j}{\partial y} \right\| < +\infty$, from which we deduce $\frac{P_\varepsilon(x)}{x} \left\| x \frac{\partial F^j}{\partial x} \right\| = \mathrm{O}\left(x^k\right)$ and finally $Q(x,y) = \mathrm{O}\left(x^k\right)$ as well. $\square$

To complete the proof of Proposition 6.6 we need to modify Q so that $Q(0,y) = 0$. In order not to change the period of Q we can only subtract from Q a function of the form $\mathcal{X}_\varepsilon \cdot F$ with F holomorphic. This is done by setting

$$F(y) := \int_0^y \frac{Q(0,v)}{v} \mathrm{d}v,$$

so that $Q - \mathcal{X}_\varepsilon \cdot F$ vanishes on $\{x = 0\}$ while still admitting an expansion of the form (6.26).

(2) A function f from a metric space E to another one F is Cauchy at a if for all $\varepsilon > 0$ there exists $\delta > 0$ such that $x, y \in \mathrm{B}(a,\delta)$ implies $d(f(x), f(y)) < \varepsilon$.

6.5. Stitching cellular sections together: proof of Proposition 6.7

Fix $(\mathbb{C}^k,0)\backslash\Delta_k$ and $\rho>0$ not larger than what is allowed in Lemma 6.23, and take $G\in\mathrm{Holo_c}(\rho\mathbb{D}\times(\mathbb{C},0))'$. We prove now that for any fixed $\varepsilon\in\mathcal{E}_\ell$, at most one $Q\in x\mathbb{C}[x]_{<k}\{y\}'$ exists such that $G-Q\in\mathrm{im}(\mathcal{X}_\varepsilon\cdot)$, that is $\mathfrak{T}(G)=\mathfrak{T}(Q)$. This amounts to showing that $\mathrm{im}(\mathcal{X}_\varepsilon\cdot)\cap x\mathbb{C}[x]_{<k}\{y\}'=\{0\}$ for all fixed $\varepsilon\in(\mathbb{C}^k,0)\backslash\Delta_k$.

Let $G\in\mathrm{im}(\mathcal{X}_\varepsilon\cdot)\cap x\mathbb{C}[x]_{<k}\{y\}'$ and write

$$G(x,y)=\mathcal{X}_\varepsilon\cdot\sum_{n\geq d}F_n(x)\,y^n=\sum_{n\geq d}G_n(x)\,y^n\in\mathrm{Holo}(r\mathbb{D}\times(\mathbb{C},0)),\quad d\in\mathbb{N};$$

we claim that $G_d=0$, which is sufficient to establish the result. It turns out that for its part of least degree in y the cohomological equation only depends on its formal normal form:

$$\widehat{\mathrm{X}}_\varepsilon\cdot\left(y^dF_d(x)\right)=y^dG_d(x).$$

Such a relation holds if and only if the period of y^dG_d along the formal normal form vanishes: $\widehat{\mathfrak{T}}\left(y^dG_d\right)=0$. Therefore we need to prove that

$$\begin{aligned}\widehat{\mathfrak{T}}:\ x\mathbb{C}[x]_{<k}\,y^d&\longrightarrow\mathbb{C}^kh^d\\ y^dG_d&\longmapsto\widehat{\mathfrak{T}}\left(y^dG_d\right)\end{aligned}$$

is injective if ε is small enough. As recalled in Corollary 6.12 we know that for every $a\in\mathbb{N}$

$$\lim_{\substack{\varepsilon\longrightarrow0\\ \mathcal{E}_\ell}}\widehat{\mathfrak{T}}\left(x^ay^b\right)=\widehat{\mathfrak{T}}_0\left(x^ay^b\right),$$

where $\widehat{\mathfrak{T}}_0$ is the period of the model saddle-node $\widehat{\mathrm{X}}_0$. The auxiliary result [46, Proposition 2] states precisely that $\widehat{\mathfrak{T}}_0$ is invertible, and therefore so is $\widehat{\mathfrak{T}}$ for small ε as expected.

7. Orbital Realization Theorem

In this section we address the inverse problem for the classification of unfoldings performed in [41], in the special case of convergent unfoldings of formal invariant μ with

$$\mu_0\notin\mathbb{R}_{\leq0}$$

and $\tau=0$. The residual cases $\mu_0\leq0$ or $\tau>0$ are dealt with in Section 8. Also notice that we only carry this study for the orbital part, the case of the temporal realization is explained in [47] when $k=1$. Generalizing this approach for $k>1$ by using the tools introduced in Section 6 should not be difficult.

We summarize in Section 7.1 how the invariants of classification are built. They unfold Martinet-Ramis's invariants [31] for the limiting saddle-node, obtained as transition maps between sectorial spaces of leaves. Yet the construction can only be carried out analytically on a given parametric cell $\mathcal{E}_\ell$, yielding a cellular invariant $\mathfrak{m}_\ell\in\prod_{\mathbb{Z}/k\mathbb{Z}}\mathcal{H}_\ell\{h\}$ (see Section 6.1 for the definition of the functional spaces $\mathcal{H}_\ell$ and Section 7.1 for the definition of m_ℓ). The orbital modulus $\mathfrak{m}(X)$ of an unfolding X consists in the whole collection $(\mathfrak{m}_\ell)_\ell$.

Definition 7.1. – We say that $(\mu,\mathfrak{m})\in\mathbb{C}\{\varepsilon\}\times\prod_\ell\mathcal{H}_\ell\{h\}^k$ is *realizable* if there exists a generic convergent unfolding X with formal orbital class μ and orbital modulus $\mathfrak{m}=\mathfrak{m}(X)$.

In Section 7.2 we prove the next result.

THEOREM 7.2. – *Assume* $\tau = 0$ *(which particularly implies* $\mu_0 \notin \mathbb{R}_{\leq 0}$*). Fix a germ at* $0 \in \left(\mathbb{C}^{k+1}, 0\right)$ *of a cell* $\mathcal{E}_\ell$*. Given* $\mathfrak{m}_\ell \in \prod_{\mathbb{Z}/k\mathbb{Z}} \mathcal{H}_\ell(h)$ *and* μ *with* $\mu_0 \notin \mathbb{R}_{\leq 0}$*, there exists a unique* $R_\ell \in x\,\mathcal{H}_\ell\{y\}[x]_{<k}$ *such that*

$$\mathcal{X}_{\ell,\varepsilon} := \widehat{\mathrm{X}} + yR_{\ell,\varepsilon}\frac{\partial}{\partial y}$$

has $\mathfrak{m}_\ell$ *for transition maps in sectorial space of leaves (i.e., for modulus).*

The fact that this "analytical synthesis" gives unique forms of the same kind as those given by Loray's "geometric" construction bolsters the naturalness of the normal forms presented here. Indeed the next corollary provides an indirect solution of the inverse problem.

COROLLARY 7.3. – *A couple* $(\mu, \mathfrak{m})$ *with* $\mu_0 \notin \mathbb{R}_{\leq 0}$ *is realizable if and only if* $R_{\ell,\varepsilon} = R_{\widetilde{\ell},\varepsilon}$ *for all* $\varepsilon \in \mathcal{E}_\ell \cap \mathcal{E}_{\widetilde{\ell}}$ *and all* $\left(\ell, \widetilde{\ell}\right)$*.*

Proof. – The equality $R_\ell = R_{\widetilde{\ell}}$ on $\mathcal{E}_\ell \cap \mathcal{E}_{\widetilde{\ell}}$ defines a bounded, holomorphic function R in the parameter $\varepsilon \in \left(\mathbb{C}^{k+1}, 0\right) \backslash \Delta_k$, which extends holomorphically to a whole neighborhood $\left(\mathbb{C}^{k+1}, 0\right)$ by Riemann's theorem on removable singularities. The corresponding unfolding $\mathcal{X}$ has modulus $\mathfrak{m}\left(\mathcal{X}\right) = (\mathfrak{m}_\ell)_\ell$ by construction.

Conversely, the Normalization Theorem tells us that we can as well assume that the vector field is in normal form $\mathcal{X}$ (2.4), without changing the orbital modulus $\mathfrak{m} = \mathfrak{m}\left(\mathcal{X}\right)$. Moreover, the normalization can be performed by tangent-to-identity mappings in the y-variable. According to Theorem 7.2, R_ℓ is uniquely determined by the component $\mathfrak{m}_\ell$ of $\mathfrak{m}$, hence $R = R_\ell$ on $\mathcal{E}_\ell$. □

Somehow this characterization is not satisfactory since it involves the auxiliary unfolding $\mathcal{X}_\ell$. In Section 7.3 we present an intrinsic characterization of realizable $(\mu, \mathfrak{m})$ as a *compatibility condition* imposed on the different dynamics induced by each pair $(\mu, \mathfrak{m}_\ell)$ on the sectorial space of leaves (Definition 7.16). Roughly speaking the condition requires that the abstract holonomy groups be conjugate over cells overlaps. In case of an actual unfolding X (i.e., realizable $(\mu, \mathfrak{m})$) these groups represent in the space of leaves the actual weak holonomy group induced by X in (x, y)-space.

7.1. Classification moduli

Starting from a generic convergent unfolding X of codimension k in prepared form (4.2) with given orbital formal invariant μ (with no restriction on μ_0), we can build the following k-tuple of periods (Definition 6.10) on a germ of a cellular decomposition $(\mathcal{E}_\ell)_{1\leq\ell\leq C_k}$, called the *orbital modulus* of X:

$$\begin{aligned} \mathfrak{m}(X) &:= \left(\mathfrak{m}_\ell(X)\right)_{1\leq\ell\leq C_k}, \\ \mathfrak{m}_\ell(X) &:= \left(\phi_\ell^{j,\mathrm{s}}\right)_{j\in\mathbb{Z}/k\mathbb{Z}}, \\ \phi_\ell^{j,\mathrm{s}} &:= 2\mathrm{i}\pi\mathfrak{T}_\ell^j(-R) \in \mathcal{H}_\ell\{h\}. \end{aligned} \tag{7.1}$$

We state the main result of [41] in the specific context of convergent unfoldings.

DEFINITION 7.4. – 1. Fix a germ of a cell $\mathcal{E}_\ell$. For $c \in \mathbb{C}\{\varepsilon\}^\times$, $\theta \in \mathbb{Z}/k\mathbb{Z}$ and $f = (f^j)_{j \in \mathbb{Z}/k\mathbb{Z}} \in \mathcal{H}_\ell\{h\}^k$ define

$$(c,\theta)^* f : (\varepsilon, h) \longmapsto \left(f_\varepsilon^{j+\theta}(c_\varepsilon h)\right)$$

and extend component-wise this action to tuples.

2. We say that two collections $\mathfrak{m}, \widetilde{\mathfrak{m}} \in \prod_\ell \mathcal{H}_\ell\{h\}^k$ are *equivalent* if there exists $c \in \mathbb{C}\{\varepsilon\}^\times$ and $\theta \in \mathbb{Z}/k\mathbb{Z}$ such that

$$(c,\theta)^* \mathfrak{m} = \widetilde{\mathfrak{m}}. \tag{7.2}$$

REMARK 7.5. – The presentation of Definition 7.4 is equivalent to that of [31] for $\varepsilon = 0$. The transition functions there are simply given by $\psi^{j,s}(h) = h \exp\left(\frac{2i\pi\mu}{k} + \phi^{j,s}\right)$. This fact will be explained in more details in Section 7.3.

THEOREM 7.6 ([41]). – *Two generic, prepared convergent unfoldings X and $\widetilde{X}$, in the same formal orbital class μ with respective orbital moduli $\mathfrak{m}(X)$ and $\mathfrak{m}(\widetilde{X})$, are equivalent by some local analytic diffeomorphism if and only if their respective orbital moduli $\mathfrak{m}(X)$ and $\mathfrak{m}(\widetilde{X})$ are equivalent. Moreover X is locally equivalent to its formal normal form $\widehat{\mathrm{X}}$ if and only if $\mathfrak{m}(X) = 0$.*

The pair (c,θ) involved in the equivalence between moduli has a geometrical interpretation. First set $\lambda := \exp 2i\pi\theta/k$ and apply the diagonal mapping

$$(\varepsilon_0, \ldots, \varepsilon_{k-1}, x) \longmapsto \left(\varepsilon_0\lambda^{-1}, \ldots, \varepsilon_j\lambda^{j-1}, \ldots, \varepsilon_{k-1}\lambda^{k-2}, x\lambda\right)$$

to X so that the moduli of the new unfolding, still written X, differs from the original by a shift in the indices j of offset θ, as explained in Section 4.1. According to Corollary 4.11 we may as well restrict our study now to fibered conjugacies Ψ between X and $\widetilde{X}$ fixing $\{y = 0\}$. Under these assumptions we have

$$\Psi : (\varepsilon, x, y) \mapsto (\varepsilon, x, y\,(c + \mathrm{o}(1))).$$

This very fact explains why c is independent on the cell $\mathcal{E}_\ell$ in the equivalence relation (7.2).

7.2. Parametric normalization: proof of Theorem 7.2

In this section we solve the inverse problem on a given parametric cell $\mathcal{E}_\ell$ when μ_0 is not in $\mathbb{R}_{\leq 0}$. Given any collection

$$\mathfrak{m}_\ell := \left(\phi^{j,s}\right)_j \in \prod_{\mathbb{Z}/k\mathbb{Z}} \mathcal{H}_\ell\{h\}$$

we can fix $\eta > 0$ such that every $\phi^{j,s}$ belongs to $\mathrm{Holo_c}(\mathcal{E}_\ell \times \eta\mathbb{D})'$. The strategy is to synthesize a k-tuple of sectorial functions $\left(H^j\right)_j$ whose transition maps over saddle parts are determined by $\mathfrak{m}_\ell$ as in (7.3) below, then to recognize that they actually are sectorial first-integrals of a holomorphic vector field X_ε in normal form.

We repeat the recipe of Theorem 5.5 in order to solve the nonlinear equation

$$H^{j+1} = H^j \exp\left(2i\pi\mu/k + \phi^{j,s} \circ H^j\right), \tag{7.3}$$

by successively solving the linear Cousin problem of Proposition 6.25 in the way we explain now. For $\varepsilon := 0$ this is precisely the technique of [43].

We want to find a solution of $N = \frac{1}{2\pi i}\mathfrak{F}(\mathfrak{m}_\ell, N)$ with $N \in \mathrm{Holo_c}\left(V_\ell^j \times r\mathbb{D}\right)'$, where $\mathfrak{F}$ is given in (6.23), and we build one through an iterative process. We start from

$$N_0 := (0)_j$$

and build

$$N_{n+1} := \frac{1}{2\mathrm{i}\pi}\mathfrak{F}(\mathfrak{m}_\ell, N_n)$$

given by Proposition 6.25. The fact that each sequence $\left(N_n^j\right)_n$ converges uniformly to some $N^j \in \mathrm{Holo_c}\left(V_\ell^j \times r\mathbb{D}\right)'$ for some $r > 0$ follows in every other respect the argument presented in the proof of Theorem 5.5, thus we shall not repeat it here.

So far we have built a k-tuple of bounded, holomorphic functions $N = \left(N^j\right)_j$ satisfying the next properties.

Corollary 7.7. – *Assume $\tau = 0$ (which particularly implies $\mu_0 \notin \mathbb{R}_{\leq 0}$). Let*

$$H^j := \widehat{H}^j \exp N^j$$

be the canonical first-integral associated with N^j.

1. $\left(H^j\right)_j$ *is a solution of* (7.3).
2. *Up to decrease $r > 0$ we can assume that:*
   (a) *N is η-adapted (as in Definition 6.24), more precisely:*
   $$\left\|H^j\right\| \leq rC$$
   for some constant $C > 0$,
   (b) $\left|x\frac{\partial N^j}{\partial x}\right| \leq \frac{1}{3}$ *and* $\left|y\frac{\partial N^j}{\partial y}\right| \leq \frac{1}{3}$ *on $V^j \times r\mathbb{D}$.*

Proof. – 1. Because $H^j = \widehat{H}^j \exp N^j$ and $\widehat{H}^{j+1} = \widehat{H}^j \exp^{2\mathrm{i}\pi\mu/k}$ (see (6.8)) we have

$$\frac{H^{j+1}}{H^j} = \exp\left(2\mathrm{i}\pi\mu/k + N^{j+1} - N^j\right).$$

Because $\left(N^j\right)_j$ is obtained as the fixed-point of the Cauchy-Heine operator

$$\left(N^j\right)_j \longmapsto \frac{1}{2\mathrm{i}\pi}\mathfrak{F}\left(\left(\phi^{j,\mathrm{s}}\right)_j, \left(N^j\right)_j\right),$$

according to Proposition 6.25 (1) the identity $N^{j+1} - N^j = \phi^{j,\mathrm{s}} \circ H^j$ holds, which validates the claim.

2. We have:

(a) This is clear thanks to Proposition 6.25.

(b) Up to decrease slightly η we can assume that the derivative of each component of $\mathfrak{m}_\ell$ is bounded on $\eta\mathbb{D}$. From the construction of N^j and Proposition 6.25 (3) we have

$$\frac{\left\|y\frac{\partial N_{n+1}}{\partial y}\right\|}{1 + \left\|y\frac{\partial N_n}{\partial y}\right\|} \leq \frac{rK}{2\mathrm{i}\pi}\left\|\mathfrak{m}'\right\| \exp\left\|N_n\right\| \leq \frac{1}{4},$$

if r is taken small enough. The conclusion follows by taking the limit $n \to \infty$. The argument for $x\frac{\partial N}{\partial x}$ is identical. ☐

Now define

$$X^j := \widehat{\mathrm{X}} + yR^j \frac{\partial}{\partial y}$$

with

$$R^j := -\frac{P\frac{\partial N}{\partial x} + y\left(1+\mu x^k\right)\frac{\partial N}{\partial y}}{1 + y\frac{\partial N^j}{\partial y}}. \tag{7.4}$$

LEMMA 7.8. – *We have*

1. $X^j \cdot N^j = -R^j$ *or, equivalently,* $X^j \cdot H^j = 0$.
2. $R^{j+1} = R^j$ *on* $\mathcal{V}^{j,\mathrm{s}}$.

Proof. – This is formally the same proof as for $\varepsilon = 0$: we refer to [43] for details.

1. It follows from elementary calculations.
2. It is equivalent to showing $X^j \cdot H^{j+1} = 0$. But this condition is met because of (1) and the fact that H^{j+1} is a function of H^j, as per (7.3). □

The lemma indicates that all pieces of $\left(R^j\right)_j$ glue together into a holomorphic function R. From (7.4) and the estimates on the derivatives of N^j obtained in Corollary 7.7 we conclude that R is bounded near the roots of P_ε (hence Riemann's theorem on removable singularities applies). The argument of Section 6.4.3 can now be invoked identically with $Q := R$ to obtain

$$R(x,y) = \sum_{n>0} r_n(x)\, y^n$$

for some polynomials r_n in x of degree at most k. We can simplify R further by applying to $\widehat{\mathrm{X}} + Ry\frac{\partial}{\partial y}$ the change of coordinates

$$(x,y) \longmapsto (x, y\exp N(y)), \quad N \in y\mathbb{C}\{y\},$$

where

$$N' = -\frac{R(0,y)}{y(1+R(0,y))}.$$

The new vector field $\widehat{\mathrm{X}} + \widetilde{R}y\frac{\partial}{\partial y}$ satisfies $\widetilde{R} \in x\mathcal{H}_\ell\{y\}[x]_{<k}$, as sought.

REMARK 7.9. – Notice that Lemma 7.8 asserts $(x,y) \mapsto \left(x, N_\ell^j(x,y)\right)$ is a fibered normalization of $\mathcal{X}$ over squid sectors.

7.3. Compatibility condition

Here we impose no restriction on μ_0.

7.3.1. *Node-leaf coordinates.* – To each squid sector V_ℓ^j we attach a unique natural coordinate h which parametrizes the space of leaves Ω_ℓ^j over that sector: this coordinate corresponds to values taken by the canonical first-integral H_ℓ^j (with connected fibers) as defined in Corollary 7.7. Moreover,

$$H_\ell^j\left(V_\ell^j \times (\mathbb{C},0)\right) = \mathbb{C}.$$

This comes from the fact that the sector's shape adheres to the point in a node-like configuration, forcing the model first integral $\widehat{H}_\ell^j$ to be surjective: a complete proof of the above statement can be found in [41]. This space of leaves is customarily compactified as the Riemann sphere Ω_ℓ^j by adding the point ∞ corresponding to the "vertical separatrices" $\left\{x = x^{j,\mathrm{n}}\right\}$ of the node-type singularity.

Because we deal with convergent unfoldings, this coordinate is completely determined by the space of leaves of the singular point $x^{j,\mathrm{n}}$ of node type attached to V^j, with two distinguished leaves corresponding to 0 (along $\{y = 0\}$) and ∞ (along $\left\{x = x^{j,\mathrm{n}}\right\}$). In particular, it remains the same when we change the point(s) of saddle type $x^{j,\mathrm{s}}$ and $x^{\sigma(j),\mathrm{s}}$ attached to a sector V^j but leave the point of node type $x^{j,\mathrm{n}}$ unchanged, while passing from one cell to another.

Let us prove briefly the result on which the compatibility condition is built. We recall that ρ_ε is the radius of a disk containing all roots of P_ε, as defined by (6.14).

Lemma 7.10. – *For every $x_* \in V_\ell^j \backslash \rho_\varepsilon \mathbb{D}$ the partial mapping*

$$h_\ell^j : y \longmapsto H_\ell^j(x_*, y)$$

is a local diffeomorphism near 0 whose multiplier at 0 does not depend on ℓ. In particular for any $\widetilde{\ell}$ such that $\mathcal{E}_\ell \cap \mathcal{E}_{\widetilde{\ell}} \neq \emptyset$, the diffeomorphism

$$\delta := h_{\widetilde{\ell}}^j \circ \left(h_\ell^j\right)^{\circ -1}$$

is tangent-to-identity. Moreover there exists $\eta_1, \eta_2, r > 0$ such that for all $\varepsilon \in \mathrm{cl}\left(\mathcal{E}_\ell \cap \mathcal{E}_{\widetilde{\ell}}\right)$

$$\eta_1 \mathbb{D} \subset \delta_\varepsilon(r\mathbb{D}) \subset \eta_2 \mathbb{D}$$

and δ_ε is injective on $r\mathbb{D}$.

In the sequel we write this map $\delta_{\widetilde{\ell} \leftarrow \ell}$.

Proof. – According to Corollary 7.7 we have

$$H_\ell^j(x_*, y) = y \widehat{H}_\ell^j(x_*, 1) + \mathrm{o}(y).$$

Since x_* lies outside the disk containing the roots of P the value of $\widehat{H}_\ell^j(x_*, 1)$, as fixed by the determination chosen in (6.7), does not depend on ℓ (but it does on j). The existence of $\eta_1, \eta_2, r > 0$ satisfying the expected properties is a consequence of [41, Corollary 8.8] and Lemma 6.23 (1). □

Definition 7.11. – For a choice of $x_*^j \in V^j \backslash \rho_\varepsilon \mathbb{D}$ we call h_ℓ^j the *node-leaf coordinate* of the unfolding X_ℓ above x_* in the sector V_ℓ^j and relative to the cell $\mathcal{E}_\ell$.

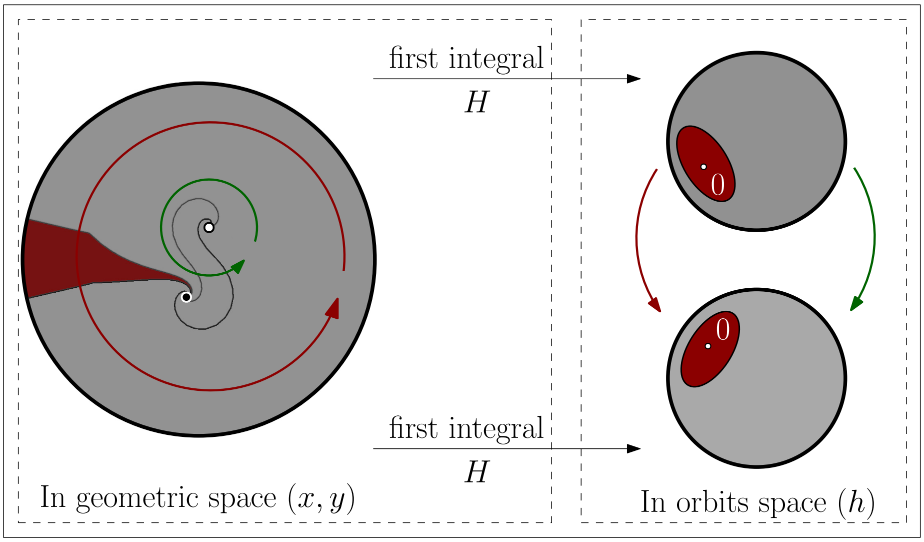


FIGURE 7.1. Passing from geometric to orbits space *via* the sectorial first integral H^j. Colored arrows show how the change of determination in H^j takes place as a mapping between two sectorial spaces of leaves: the necklace dynamics.

7.3.2. *Necklace dynamics.* – Here we work in a fixed germ of a cell $\mathcal{E}_\ell$ for fixed $\varepsilon \in \mathcal{E}_\ell$; we drop the ℓ and ε indices whenever not confusing. According to the constructions performed in [41], and hinted at by Theorem 7.2, the orbital modulus $(\mu, \mathfrak{m}(X))$ of a convergent unfolding encodes the way the different node-leaf coordinates glue above the intersection of squid sectors:

$$\begin{cases} H^{j+1} = H^j \exp\left(2\mathrm{i}\pi\mu/k + \phi^{j,\mathrm{s}} \circ H^j\right) & \text{above } V^{j,\mathrm{s}}, \\ H^{\sigma(j)} = L_{\nu^j} \circ H^j & \text{above } V^{j,\mathrm{g}}, \end{cases}$$

where

$$L_c : h \longmapsto ch, \quad c \neq 0,$$

and $\nu^j = \nu^j_\ell \in \mathbb{C}^\times$ relates to the dynamical invariants μ and the residues $\left(\frac{1}{P'_\varepsilon(x^m)}\right)_m$ at the roots $(x^m)_{0 \leq m \leq k}$: indeed, the ramification at the linear level of the first integral at a singular point, given by $\exp\left(-2\mathrm{i}\pi\frac{1+\mu_\varepsilon(x^m)^k}{P'_\varepsilon(x^m)}\right)$, is equal to the product of all ramifications when crossing sectors while turning around the point, i.e., to the product of one factor $\exp 2\mathrm{i}\pi\mu/k$ for each crossed sector $V^{j,\mathrm{s}}$ and one factor ν_j for each crossed sector $V^{j,\mathrm{g}}$. It is therefore rather natural to consider the germs of diffeomorphisms in node-leaf coordinate

$$\begin{aligned} \psi^{j,\mathrm{s}}_\ell &: h \longmapsto h \exp\left(2\mathrm{i}\pi\mu/k + \phi^{j,\mathrm{s}}_\ell(h)\right), \\ \psi^{j,\mathrm{g}}_\ell &: h \longmapsto \nu^j_\ell h, \end{aligned} \tag{7.5}$$

where $(\mathfrak{m}_\ell)_\ell = \mathfrak{m}(X)$ and $\mathfrak{m}_\ell = \left(\phi^{j,\mathrm{s}}_\ell\right)_j$. Obviously one can do the same construction starting from any tuple $\mathfrak{m} \in \prod_\ell \mathscr{H}_\ell\{h\}^k$.

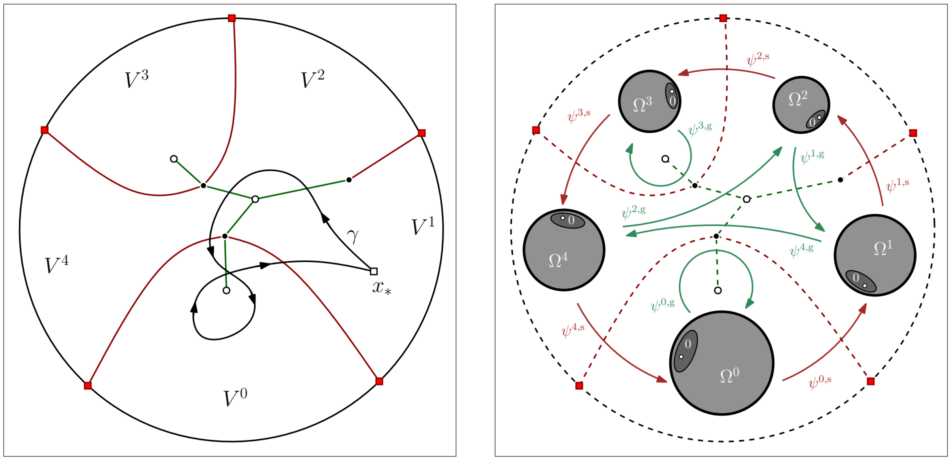


FIGURE 7.2. Schematics of the necklace dynamics and of the corresponding sectorial decomposition for $k=5$ and $\sigma=\left(\begin{smallmatrix}0&1&2&3&4\\0&4&1&3&2\end{smallmatrix}\right)$. The loop $\gamma\in\pi_1\left(\rho\mathbb{D}\backslash P_\varepsilon^{-1}(0),x_*\right)$ corresponds to the word $\mathfrak{w}(\gamma)=\mathrm{s}_0^+\mathrm{g}_0^+\mathrm{g}_0^+\mathrm{s}_4^+\mathrm{g}_2^-\mathrm{g}_1^-$ in necklace dynamics.

REMARK 7.12. – For some value of the parameter ε in a given cell $\mathcal{E}_\ell$, the saddle mappings ψ^{s} are entirely determined by μ and $\mathfrak{m}$, while the gate mappings ψ^{g} are entirely determined by μ.

The dynamics induced by these germs is of interest to us only if it encodes the underlying dynamics of the unfolding (weak holonomy group). A necessary condition is that the latter group does not depend on ℓ, i.e., on the peculiar way of slicing the space into sectors which is imposed by our construction. Therefore we only want to consider the "abstract" holonomy representation of $\pi_1\left(\rho\mathbb{D}\backslash P_\varepsilon^{-1}(0),x_*\right)$ in the space of leaves. Let us describe this representation (see Figure 7.2 for an example).

DEFINITION 7.13. – We fix a base-sector V^{j_*} and a base-point $x_*\in V^{j_*}\backslash\rho_\varepsilon\mathbb{D}$, as well as some $\mathfrak{m}_\ell=\left(\phi_\ell^{j,\mathrm{s}}\right)_j\in\mathscr{H}_\ell\left\{h\right\}^k$.

1. To any loop $\gamma\in\pi_1\left(\rho\mathbb{D}\backslash P_\varepsilon^{-1}(0),x_*\right)$ we associate the multiplicative word $\mathfrak{w}_\ell(\gamma)$ in the $4k$ letters $\left\{\mathrm{s}_j^\pm,\mathrm{g}_j^\pm\,:\,j\in\mathbb{Z}/k\mathbb{Z}\right\}$ obtained by keeping track of bounded squid sectors boundaries crossed successively when traveling along γ. The superscript $+$ (resp. $-$) is given to s_j according to whether one crosses the saddle boundary from V^j to V^{j+1} (*resp.* from V^{j+1} to V^j), "in the same direction" as $\psi^{j,\mathrm{s}}$ (*resp.* $\left(\psi^{j,\mathrm{s}}\right)^{\circ-1}$) applies. For g_j we take the same convention for gate transitions $\psi^{j,\mathrm{g}}$ and postulate the algebraic relations $\left(\mathrm{s}_j^\pm\right)^{-1}=\mathrm{s}_j^\mp$, $\left(\mathrm{g}_j^\pm\right)^{-1}=\mathrm{g}_j^\mp$.
2. To any word $\mathfrak{w}=\prod_n\omega_{j_n}^\pm$ we associate the germ

$$\psi_\ell\left[\mathfrak{w}\right]:\ h\longmapsto\bigcirc_n\left(\psi_\ell^{j_n,\omega}\right)^{\circ\pm1}.$$

For instance

$$\psi\left[\mathrm{s}_0^+\mathrm{g}_0^+\mathrm{g}_0^+\mathrm{s}_4^+\mathrm{g}_2^-\mathrm{g}_1^-\right] = \psi^{0,\mathrm{s}} \circ \left(\psi^{0,\mathrm{g}}\right)^{\circ 2} \circ \psi^{4,\mathrm{s}} \circ \left(\psi^{2,\mathrm{g}}\right)^{\circ -1} \circ \left(\psi^{1,\mathrm{g}}\right)^{\circ -1}.$$

3. We write

$$\mathcal{W}_\ell := \mathfrak{w}_\ell\left(\pi_1\left(\rho\mathbb{D}\backslash P_\varepsilon^{-1}(0), x_*\right)\right)$$

the image group of *admissible words*, that is all the words corresponding to all the encodings (1) of a loop with given base-point x_* in a disk of given radius ρ punctured with the roots of P_ε.

4. Let $\mathfrak{m} = (\mathfrak{m}_\ell)_\ell \in \prod_\ell \mathcal{H}_\ell\{h\}$. The collection of image groups $\mathcal{G}(\mathfrak{m}) = \left(\mathcal{G}_\ell\right)_\ell$ of germs of a biholomorphism fixing 0 given by

$$\mathcal{G}_\ell := \psi_\ell\left[\mathcal{W}_\ell\right],$$

is called the *necklace dynamics* associated to $(\mu, \mathfrak{m})$ based at the sector V^{j_*}.

Remark 7.14. – 1. To keep notations light we write $\psi_\ell[\gamma]$ instead of $\psi_\ell[\mathfrak{w}_\ell(\gamma)]$ for $\gamma \in \pi_1\left(\rho\mathbb{D}\backslash P^{-1}(0), x_*\right)$. The context will never be ambiguous.

2. Obviously the morphisms $\mathfrak{w}_\ell$ and $\mathfrak{w}_{\widetilde{\ell}}$ are distinct. The change of cell in $\mathcal{E}_\ell \cap \mathcal{E}_{\widetilde{\ell}}$ can be translated algebraically as a group isomorphism $\mathcal{W}_\ell \longrightarrow \mathcal{W}_{\widetilde{\ell}}$. For instance when $k = 1$ the isomorphism acts on generators as

$$\begin{cases} \mathrm{g}^+ \longmapsto \mathrm{g}^-\mathrm{s}^+ \\ \mathrm{s}^+\mathrm{g}^- \longmapsto \mathrm{g}^+ \end{cases}$$

with notations of Figure 7.3.

Remark 7.15. – 1. The groups $\mathcal{W}_\ell$ and $\mathcal{G}_\ell$ do not depend on the particular choice of the base-point $x_* \in V^{j_*}$, but do depend on the base-sector V^{j_*}.

2. Changing the base-sector from V^{j_*} to another sector V^j induces an inner conjugacy between respective necklace dynamics.

7.3.3. *Compatibility condition*

Definition 7.16. – Let $\mathfrak{m} \in \prod_\ell \mathcal{H}_\ell\{h\}^k$ and $\mu \in \mathbb{C}\{\varepsilon\}$. We say that $(\mu, \mathfrak{m})$ satisfies the *compatibility condition* if the different necklace dynamics (i.e., abstract holonomy pseudogroups) combined to form $\mathcal{G}(\mathfrak{m})$ are conjugate, in the sense that there exists $x_* \in \rho\mathbb{D}\backslash\rho_\varepsilon\mathbb{D}$ in a fixed base sector V^{j_*} such that for every $\ell, \widetilde{\ell}$ and any connected component C of $\mathcal{E}_\ell \cap \mathcal{E}_{\widetilde{\ell}} \neq \emptyset$ there exists a (perhaps small) subdomain $\Lambda \subset C$ such that for all $\varepsilon \in \Lambda$ there exists $\delta_{\widetilde{\ell}\leftarrow\ell,\varepsilon} \in \mathrm{Diff}(\mathbb{C}, 0)$ satisfying:

- $\delta'_{\widetilde{\ell}\leftarrow\ell,\varepsilon}(0) = 1$,
- for all $\gamma \in \pi_1\left(\rho\mathbb{D}\backslash P^{-1}(0), x_*\right)$,

$$\delta^*_{\widetilde{\ell}\leftarrow\ell,\varepsilon}\psi_{\ell,\varepsilon}[\gamma] = \psi_{\widetilde{\ell},\varepsilon}[\gamma], \tag{7.6}$$

where $\delta^*\psi = \delta^{-1} \circ \psi \circ \delta$ is the usual conjugacy for diffeomorphisms.

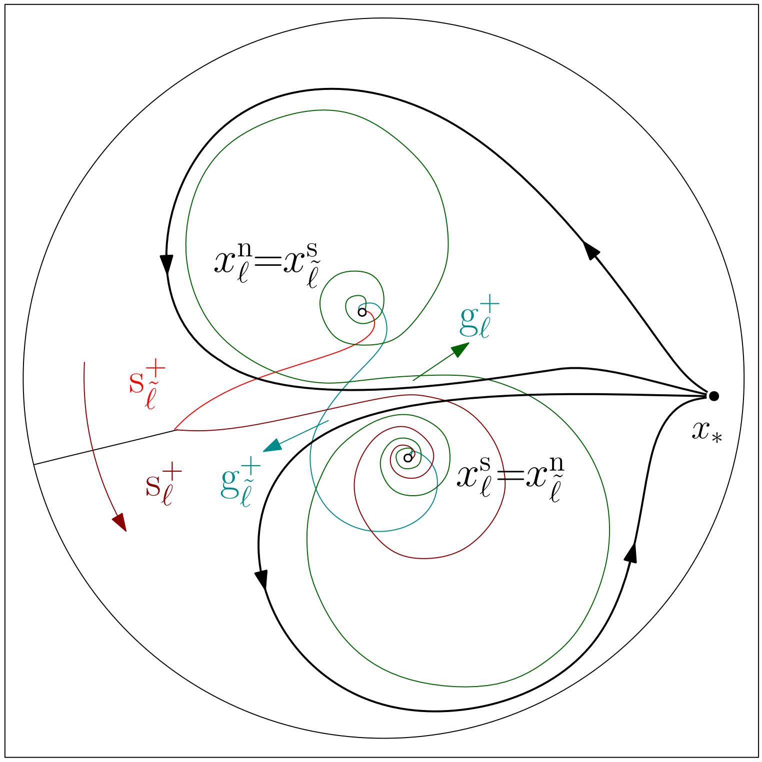


FIGURE 7.3. The generators of the two holonomies on the self intersection of the unique cell $\mathscr{E}$.

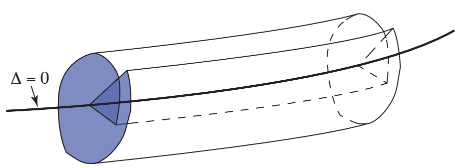


FIGURE 7.4. A cell $\mathscr{E}_\ell$ having self-intersection around a regular part of Δ_k.

REMARK 7.17. – Notice that the compatibility condition also applies when $\widetilde{\ell} = \ell$, i.e., $\mathscr{E}_\ell$ is a self-intersecting cell with self-intersection $\mathscr{E}^{\cap}_{\ell}$ around a regular part of Δ_k as in Figure 7.4, with the obvious adaptations. To avoid confusion we denote by $\overline{\varepsilon}$ and $\widetilde{\varepsilon}$ the "distinct points" corresponding to the same parameter $\varepsilon \in \mathscr{E}^{\cap}_{\ell}$ seen from two different overlapping parts of the cell. More generally we decorate objects with corresponding signs, like $\overline{\psi}$ or $\widetilde{\psi}$ in order to really stand for $\psi_{\ell,\overline{\varepsilon}}$ and $\psi_{\ell,\widetilde{\varepsilon}}$ respectively.

LEMMA 7.18. – *If* $(\mu, \mathfrak{m})$ *is realizable then the compatibility condition holds.*

Proof. – Fix some point $x_* \in V^0 \backslash \rho_\varepsilon \mathbb{D}$ and take $\delta_{\widetilde{\ell}\leftarrow\ell} := h^0_{\widetilde{\ell}} \circ \left(h^0_\ell\right)^{\circ -1}$ on $\mathscr{E}_\ell \cap \mathscr{E}_{\widetilde{\ell}}$ as in Lemma 7.10. $\square$

REMARK 7.19. – 1. Although we do not impose that the mappings $\delta_{\widetilde{\ell}\leftarrow\ell}$ exist on the connected component C of $\mathcal{E}_\ell \cap \mathcal{E}_{\widetilde{\ell}}$, nor depend analytically on $\varepsilon \in \Lambda \subset C$, it will be true retrospectively and the dynamical conjugacies $\delta_{\widetilde{\ell}\leftarrow\ell}$ are always of the form described in Lemma 7.10. In particular the collection $\left(\delta_{\widetilde{\ell}\leftarrow\ell}\right)_{\ell,\widetilde{\ell}}$ is a cocycle:

$$\delta_{\ell_2\leftarrow\ell_1} \circ \delta_{\ell_1\leftarrow\ell_0} = \delta_{\ell_2\leftarrow\ell_0}$$

whenever all three mappings are simultaneously defined.

2. The compatibility condition could be weakened further. The existence of $\delta_{\widetilde{\ell}\leftarrow\ell}$ as above is only needed for ε belonging to a set Λ of full analytic Zariski closure, i.e., such that if a holomorphic function f on C satisfies $f|_\Lambda = 0$ then $f = 0$. The cornerstone of the proof of the Realization Theorem consists indeed in applying Corollary 7.3: it suffices to check whether the identity $R_\ell - R_{\widetilde{\ell}} = 0$ holds on every connected component C of $\mathcal{E}_\ell \cap \mathcal{E}_{\widetilde{\ell}}$.

7.4. Normal forms stitching: proof of Orbital Realization Theorem when $\mu_0 \notin \mathbb{R}_{\leq 0}$ and $\tau = 0$.

Thanks to Lemma 7.18, only the converse direction of the Realization Theorem still requires a full proof at this stage. Assume then that the compatibility condition holds. Let us fix a base point x_* in a base sector V^{j_*} and pick $\varepsilon \in \Lambda \subset C \subset \mathcal{E}_\ell \cap \mathcal{E}_{\widetilde{\ell}}$ as in Definition 7.16. Recalling Lemma 7.10, the tangent-to-identity mapping

$$\Psi : (x_*, y) \longmapsto \left(x_*, \left(h^0_{\widetilde{\ell}}\right)^{\circ -1} \circ \delta_{\widetilde{\ell}\leftarrow\ell} \circ h^0_\ell\right) \tag{7.7}$$

conjugates the weak holonomy pseudogroups given by the representation

$$\mathfrak{h}_\ell : \pi_1\left(\rho\mathbb{D}\backslash P^{-1}(0), x_*\right) \longrightarrow \mathrm{Diff}\left(\{x = x_*\}, 0\right).$$

Let us formulate a direct consequence of the main results of [11] (see [2]) in a manner adapted to our setting.

LEMMA 7.20. – *The map $\varepsilon \in \mathcal{E}_\ell \mapsto \left(\nu^j_\varepsilon\right)_{j\in\mathbb{Z}/k\mathbb{Z}}$ is holomorphic and locally injective. In particular there exists a subdomain $\Lambda' \subset \Lambda$ such that for all $\varepsilon \in \Lambda'$, every singular point of X_ℓ and $X_{\widetilde{\ell}}$ is hyperbolic.*

Using an extension of the Mattei-Moussu construction for hyperbolic singularities (see below) we can analytically continue Ψ (defined in (7.7)) on a whole neighborhood of $\{y = 0\}$ as a fibered equivalence between $\mathcal{X}_\ell$ and $\mathcal{X}_{\widetilde{\ell}}$. The argument developed in Section 5.6 (to prove uniqueness of the normal form) is performed for fixed ε, therefore there exists

$$c \in \mathbb{C}^\times$$

such that

$$R_{\ell,\varepsilon}(x, cy) = R_{\widetilde{\ell},\varepsilon}(x, y).$$

But the conjugacy Ψ is tangent to the identity in the y-variable thus $c = 1$. Therefore $R_{\ell,\varepsilon} = R_{\widetilde{\ell},\varepsilon}$ on Λ, thus on C by analytic continuation. Since this argument can be carried out for any connected component C of any cellular intersection, Corollary 7.3 yields the conclusion.

REMARK 7.21. – In fact Ψ itself must be the identity, therefore

$$\delta_{\widetilde{\ell}\leftarrow\ell} = h^0_{\widetilde{\ell}} \circ \left(h^0_\ell\right)^{\circ-1}$$

as in Lemma 7.10.

There only remains a single gap in the above argument, namely that of extending Ψ near each hyperbolic singularity. Let $\mathcal{F}_\ell$ be the foliation induced by $\mathcal{X}_\ell$ and take a germ $\Sigma \subset \{x = x_*\}$ of a transverse disk at $(x_*, 0)$ in such a way that Ψ is holomorphic and injective on Σ. The union of the saturation $\mathrm{Sat}_{\mathcal{F}_\ell}(\Sigma)$ and the vertical separatrices $P^{-1}(0)$ is a full neighborhood of $\{y = 0\}$ since no singular point of $\mathcal{F}_\ell$ is a node. Therefore Ψ can be extended as a fibered, injective mapping by the usual path-lifting technique except along the separatrices $P^{-1}(0)$. Up to divide $\mathcal{X}_\ell$ and $\mathcal{X}_{\widetilde{\ell}}$ by a local holomorphic unit near each singularity, we can assume that the hypotheses of Lemma 5.13 are met. This completes the proof of the Realization Theorem when $\mu_0 \notin \mathbb{R}_{\leq 0}$.

8. General case $\tau > 0$

In this section we fix $\tau \in \mathbb{N}$ such that

$$\mu_0 + \tau(k+1) \notin \mathbb{R}_{\leq 0}.$$

8.1. End of proof of (orbital) Normalization, Uniqueness and Realization Theorems

We explain now how to reduce the case $\tau > 0$ to the case $\tau = 0$ already dealt with. We exploit the observation that formally $\mathrm{Section}_k\{P^\tau y\}$ is the pullback of $\mathrm{Section}_k\{y\}$ by the mapping

$$T : (\varepsilon, x, y) \longmapsto \left(\varepsilon, x, P^\tau_\varepsilon(x)\, y\right). \tag{8.1}$$

Albeit not invertible along the lines $\{P_\varepsilon(x) = 0\}$ (its image is not a neighborhood of $\{y = 0\}$), the mapping T transforms the model unfolding

$$\widehat{\mathrm{X}}(x, y) = P_\varepsilon(x)\frac{\partial}{\partial x} + y\left(1 + \mu_\varepsilon x^k\right)\frac{\partial}{\partial y} \tag{8.2}$$

into

$$\widehat{\mathrm{Y}} := T^*\widehat{\mathrm{X}} = P_\varepsilon\frac{\partial}{\partial x} + (1 + \tau P'_\varepsilon + \mu_\varepsilon x^k)y\frac{\partial}{\partial y}.$$

Observe that

$$\tau P'(x) + \mu x^k \sim_\infty (\tau(k+1) + \mu)\, x^k,$$

so that involving P^τ in this way shifts the formal invariant by $\tau(k+1)$. Apart from the fact that $\widehat{Y}$ is not in prepared form (4.2), all the theory developed before for the Realization Theorem applies in this case too. Let us be more specific. The key property we used intensively was to be able to perform most arguments for fixed ε. This was proved sufficient because automorphisms of prepared forms fixing the x-variable must also fix the canonical parameter ε.

LEMMA 8.1. – 1. *The group of (fibered) symmetries*

$$(\varepsilon, x, y) \mapsto (\eta(\varepsilon), X(\varepsilon, x), Y(\varepsilon, x, y))$$

of (the unfolding of) vector fields defined by (8.2)*, is isomorphic to* $\mathbb{Z}/k\mathbb{Z} \times \mathbb{C}^\times$ *through the linear representation*

$$\zeta_0 : \mathbb{Z}/k\mathbb{Z} \times \mathbb{C}^\times \longrightarrow \mathrm{GL}_{k+2}(\mathbb{C}) \tag{8.3}$$

$$(\theta, c) \longmapsto \Big((\varepsilon_0, \dots, \varepsilon_{k-1}, x, y) \mapsto \Big(\alpha\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{k-1}\alpha^{-(k-2)}, \alpha x, c y\Big)\Big), \tag{8.4}$$

where $\alpha = \exp{}^{2i\pi\theta}/k$.

2. *This statement continues to hold in the more general case of an unfolding*

$$P_\varepsilon(x)\frac{\partial}{\partial x} + (1 + Q_\varepsilon(x))y\frac{\partial}{\partial y}, \tag{8.5}$$

where $Q_\varepsilon \in \mathbb{C}[x]_{\leq k}$ *is a polynomial in* x *of degree at most* k *and* $Q_\varepsilon(0) = 0$*, save for the fact that the representation* $\zeta_\tau : \mathbb{Z}/k\mathbb{Z} \times \mathbb{C}^\times \to \mathrm{Diff}\left(\mathbb{C}^{k+2}, 0\right)$ *has no reason to be linear.*

3. *In particular, any symmetry tangent to the identity is the identity.*

Proof. – (1) is shown in [41]. For (2), there exists a diffeomorphism Ψ of the form $(\varepsilon, x, y) \mapsto (\eta, X, Y)$ transforming a general formal normal form (8.5) to the standard formal normal form (8.2). Then any symmetry of a general formal normal form is given by $\Psi^{-1} \circ \zeta_0(\theta, c) \circ \Psi$ for some $(\theta, c) \in \mathbb{Z}/k\mathbb{Z} \times \mathbb{C}^\times$. (3) follows. □

REMARK 8.2. – 1. In view of Lemma 8.1, we could have replaced (8.2) by some other (8.5) in all our constructions regarding realization. In such a form, the parameters are again canonical, as long as we consider changes of coordinates tangent to the identity.

2. The structure of sectors, and also the decomposition in cells $\mathcal{E}$, are determined from P_ε alone in (8.2): only the size of the neighborhoods of the origin in x-space and in parameter space might need to be slightly adjusted when passing from the coordinates (x, y) to the coordinates $(x, P^\tau(x)\, y)$. Hence, instead of considering (8.2), we could have taken a normal form (8.5) with the same sectors V_ℓ^j and same cells $\mathcal{E}_\ell$.

The rest of our argument relies on the next transport result.

LEMMA 8.3. – 1. $(\mu, \mathfrak{m})$ *satisfies the compatibility condition if and only if* $(\mu + \tau(k+1), \mathfrak{m})$ *does.*

2. *Take* $\mathcal{X}$ *in orbital normal form* (2.4) *with* $\tau := 0$*. Consider the corresponding unfolding*

$$\mathcal{Y} := T^*\mathcal{X} = P_\varepsilon\frac{\partial}{\partial x} + y\left(1 + \tau P'_\varepsilon + \mu x^k + R(x, P^\tau y)\right)\frac{\partial}{\partial y},$$

for T *as in* (8.1)*. Then* $\mathcal{X}$ *and* $\mathcal{Y}$ *have same orbital invariant* $\mathfrak{m}(\mathcal{X}) = \mathfrak{m}(\mathcal{Y})$.

We postpone the proof till Section 8.1.4. In the meantime we finish establishing the main theorems.

8.1.1. *End of proof of Orbital Realization Theorem.* – Let $(\mu, \mathfrak{m})$ satisfy the compatibility condition and let us prove it is realizable as the orbital modulus of some convergent unfolding. Normalization and Realization theorems so far hold when $\tau = 0$ (in particular $\mu_0 \notin \mathbb{R}_{\leq 0}$): in that case $\mathfrak{m}$ is the modulus of an unfolding in normal form

$$P_\varepsilon(x)\frac{\partial}{\partial x} + y\left(1 + \mu(\varepsilon)x^k + y\sum_{j=1}^{k} x^j R_j(y)\right)\frac{\partial}{\partial y}. \tag{8.6}$$

To consider the case $\tau > 0$, we need to use the following remark: the whole proof for $\tau = 0$ would have worked *verbatim* with the formal part and parameters given in some alternate form (8.5). This would have produced a realization of the form

$$P_\varepsilon(x)\frac{\partial}{\partial x} + y\left(1 + Q_\varepsilon(x) + y\sum_{i=j}^{k} x^j R_j(y)\right)\frac{\partial}{\partial y}, \tag{8.7}$$

with new canonical parameters. Let τ be a positive integer such that $\mu_0 + \tau(k+1) > 0$ and consider the new formal normal form

$$\widehat{Y}(x, y) = P_\varepsilon(x)\frac{\partial}{\partial x} + (1 + \tau P_\varepsilon'(x) + \mu(\varepsilon)x^k)y\frac{\partial}{\partial y}$$

corresponding to $Q_\varepsilon := \tau P_\varepsilon' + \mu(\varepsilon)x^k$ in (8.5), with formal invariant

$$\widehat{\mu} := \mu + \tau(k+1).$$

But according to Lemma 8.3:

1. $(\widehat{\mu}, \mathfrak{m})$ is compatible,
2. it is realized in the form (8.7),
3. the change $(x, y) \mapsto \left(x, P_\varepsilon^{-\tau}(x)y\right)$ transforms (8.7) back into an unfolding
$$P_\varepsilon(x)\frac{\partial}{\partial x} + y\left(1 + \mu(\varepsilon)x^k + \sum_{j=1}^{k} x^j R_j(P_\varepsilon^\tau(x)y)\right)\frac{\partial}{\partial y},$$
4. the latter unfolding is holomorphic on a whole neighborhood of $\left(\mathbb{C}^{k+2}, 0\right)$, and is therefore a realization of $(\mu, \mathfrak{m})$.

8.1.2. *End of proof of Normalization Theorem.* – The proof we just finished shows that any realizable $(\mu, \mathfrak{m})$ can be realized in normal form.

8.1.3. *End of proof of Uniqueness Theorem.* – Each vector field $\mathscr{X}_\varepsilon$ of the unfolding in normal form (2.4) is holomorphic on a domain

$$D(r) := \bigcup_{\varepsilon \in (\mathbb{C}^k, 0)} \{(\varepsilon, x, y) : |x| < \rho, |P_\varepsilon^\tau(x)\, y| < r\}.$$

Let E be a neighborhood of 0 in $\mathbb{C}^{k+2}$ and $\Psi : E \to \left(\mathbb{C}^{k+2}, 0\right)$ be a local conjugacy between normal forms $\mathscr{X}$ and $\widetilde{\mathscr{X}}$, which can be assumed fibered thanks to Corollary 4.11 (2). We can use the Uniqueness Theorem in the coordinates $(x, P^\tau(x)\, y)$ (given by the Uniqueness Theorem for $\mu_0 \notin \mathbb{R}_{\leq 0}$, already proved) at the cost of showing that $T^*\Psi = T \circ \Psi \circ T^{\circ -1}$ is holomorphic and injective on some small neighborhood of $(0, 0)$ uniformly in ε. This is not trivial since the image of $E \cap \{\varepsilon = \mathrm{cst}\}$ by T can never be such a uniform neighborhood of $(0,0)$ if E is bounded in the y-variable. But $T\left(D\left(r\right) \cap \{\varepsilon = \mathrm{cst}\}\right)$ is, so we wish to extend Ψ to some $D\left(r'\right) \subset D\left(r\right)$. The usual path-lifting technique in the foliation $\mathcal{F}_\varepsilon$ induced by $\mathcal{X}_\varepsilon$ allows to extend Ψ_ε on

$$\mathcal{U}_\varepsilon := \{\varepsilon\} \times \mathrm{Sat}_{\mathcal{F}_\varepsilon}\left(E\right) \subset D\left(r\right).$$

Using the special form of the normal form $\mathcal{X}_\varepsilon$ we conclude the proof of the Uniqueness Theorem.

Lemma 8.4. – *Assume that $\rho > 0$ is small enough so that $\left|\mu_\varepsilon x^k + \tau P'_\varepsilon\left(x\right)\right| < \frac{1}{4}$ for all $x \in \rho\mathbb{D}$ and all $\|\varepsilon\|$ small enough. There exists $r \geq r' > 0$ such that for $\mathcal{U}_\varepsilon = \mathcal{U}_\varepsilon\left(r\right)$ defined as above one has $D\left(r'\right) \subset \bigcup_{\varepsilon \in (\mathbb{C}^k,0)} \mathcal{U}_\varepsilon \subset D\left(r\right)$.*

Proof. – For a solution of the flow system

$$\begin{cases} \dot{x} &= -P_\varepsilon\left(x\right) \\ \dot{y} &= -y\left(1 + \mu_\varepsilon x^k + R_\varepsilon\left(x,y\right)\right) \end{cases}$$

with $t \in \mathbb{R}$ and initial value (x_*, y_*), the modulus of $\phi\left(t\right) := \left|P_\varepsilon^\tau\left(x\left(t\right)\right) y\left(t\right)\right|$ satisfies

$$\dot{\phi} = -\phi \Re\left(1 + \mu_\varepsilon x^k + R_\varepsilon + \tau P'_\varepsilon\right).$$

Since $R_\varepsilon\left(x,0\right) = 0$ we can choose r so small that $\left|\mu_\varepsilon x^k + R_\varepsilon + \tau P'_\varepsilon\left(x\right)\right| < \frac{1}{2}$ for all $(\varepsilon, x, y) \in D\left(r\right)$, and $\dot{\phi} < -\phi/2$. Hence starting at (x_*, y_*) with $\left|P_\varepsilon^\tau\left(x_*\right) y_*\right| < r$ and $|x_*| < \rho$, the trajectory for positive t never escapes $D\left(r\right)$. But $t \mapsto \left|y\left(t\right)\right|$ is also exponentially decreasing, therefore we eventually reach a point within E.

Again, this is the ideal situation, because it may happen that $x\left(t\right)$ exits $\mathrm{cl}\left(\rho\mathbb{D}\right)$. If $\left|x\left(t_0\right)\right| = \rho$ then we modify the trajectory x by solving $\dot{x} = \pm i P_\varepsilon\left(x\right)$ from t_0 on, the sign being chosen so that $\pm i P_\varepsilon\left(x\left(t_0\right)\right)$ points inside $\rho\mathbb{S}^1$, until we reach a point $x\left(t_1\right)$ through which the solution of $\dot{x} = -P_\varepsilon\left(x\right)$ stays in $\rho\mathbb{D}$ in positive time (i.e., accumulate on an attractive singularity). While for $t \in [t_0, t_1]$ we cannot control the sign of $\dot{\phi} = \pm\phi \Im\left(\mu_\varepsilon x^k + R_\varepsilon + \tau P'_\varepsilon\right)$, resulting in a probable increase in ϕ, the total amount by which $\frac{\phi}{r}$ increases is bounded uniformly in (x_*, y_*) and ε. Therefore there exists a radius $r \geq r' > 0$ for which, if $(x_*, y_*) \in D\left(r'\right)$, the modified trajectory $t \geq 0 \mapsto \left(x\left(t\right), y\left(t\right)\right)$ does not escape from $D\left(r\right)$ and thus eventually enters E. □

8.1.4. *Proof of Lemma 8.3.* – First, as noted in Remark 8.2, we can choose the same sectors in x and same cells in the parameter ε, possibly after adjusting their diameter. Also, we have chosen to take the linear parts of the $\psi_\ell^{j,s}$ of the form $\exp 2i\pi\mu/k$. This choice is arbitrary. What is needed is that the product of these linear parts be equal to $\exp 2i\pi\mu$. Because $(k+1)\tau \in \mathbb{Z}$, so that $\exp 2i\pi\mu = \exp\left(2\pi i\left(\mu + (k+1)\tau\right)\right)$, we are perfectly entitled to take the same linear parts for $\mathfrak{m}\left(\mathcal{X}\right)$ and $\mathfrak{m}\left(\mathcal{Y}\right)$.

The Camacho-Sad index $\widetilde{\lambda}^j$ (resp. λ^j) of the singular point $(z, 0) \in P^{-1}\left(0\right) \times \{0\}$ in $\mathcal{Y}_\varepsilon$ (resp. $\mathcal{X}_\varepsilon$), relatively to the invariant line $\{y = 0\}$, is given by

$$\widetilde{\lambda}^j = \frac{P'_\varepsilon(z)}{1 + \tau P'_\varepsilon(z) + \mu_\varepsilon z^k}, \qquad \lambda_j = \frac{P'_\varepsilon(z)}{1 + \mu_\varepsilon z^k}.$$

Hence, $\frac{1}{\widetilde{\lambda}^j} = \frac{1}{\lambda^j} + \tau$, yielding $\exp {}^{2i\pi/\widetilde{\lambda}^j} = \exp {}^{2i\pi/\lambda^j}$. This means that the gate transition maps are the same for both dynamical necklaces induced by $(\mu, \mathfrak{m})$ and by $(\widehat{\mu}, \mathfrak{m})$. Thus, the holonomies involved in the compatibility condition are the same provided (2) holds. In particular, this means that $(\widehat{\mu}, \mathfrak{m})$ satisfies the compatibility condition, proving (1).

Show now that $\mathfrak{m}$ is the analytic part of the modulus of $\mathcal{Y}$. It suffices to consider a fixed $\varepsilon \in \mathcal{E}_\ell$ and a corresponding saddle part $V^{j,\mathrm{s}}$. Recall how a normalizing map between $\mathcal{Y}$ and its formal model, as in Remark 6.9, defines the canonical sectorial first integral

$$\widetilde{H}(x,y) = y\widetilde{E}(x) \exp \widetilde{N}^j(x,y),$$

where $\widetilde{E}(x) = \prod_{j=0}^k (x - x^j)^{-1/\widetilde{\lambda}_j}$ is the multiplier in the model first integral of $\mathcal{Y}_\varepsilon$. Let $\psi^{j,\mathrm{s}} : h \mapsto h \exp\left({}^{2i\pi\mu}/k + \mathfrak{m}(h)\right)$ be the Martinet-Ramis invariant as in Section 7.1, that is

$$\widetilde{H}^{j+1} = \widetilde{\psi}^{j,\mathrm{s}} \circ \widetilde{H}^j.$$

Let us now move to $\mathcal{X}$. It is clear that a normalizing map over $\mathcal{V}^j$ transforming $\mathcal{X}_\varepsilon$ into its normal form is given by

$$(x,y) \longmapsto \left(x, y \exp N^j(x,y)\right)$$
$$N_j(x,y) = \widetilde{N}_j(x, P_\varepsilon^\tau(x)y).$$

Moreover, the domain of this map is of the form $V^j \times \{|P_\varepsilon^\tau(x)y| < r\}$. Since

$$\prod_{j=0}^k (x - x^j)^{-\frac{1}{\lambda_j}} = \widetilde{E}(x) P^\tau(x)$$

the canonical first integral of $\mathcal{X}$ has the form

$$H^j(x,y) = E(x) y \exp N_j(x,y) = \widetilde{E}(x)\, (P^\tau(x) y) \exp \widetilde{N}^j(x, P^\tau(x)\, y).$$

It follows at once that

$$H^{j+1} = \widetilde{\psi}^{j,\mathrm{s}} \circ H^j,$$

yielding the conclusion $\psi^{j,\mathrm{s}} = \widetilde{\psi}^{j,\mathrm{s}}$ as expected.

8.2. Section of the period operator: end of proof of the Normalization Theorem

Let $\mathcal{X}$ be a generic unfolding in orbital normal form (2.4), understood as a derivation. Theorem 6.8 holds regardless of the value of μ_0 or τ. The study performed in Section 6 to establish Theorem 6.2 can be repeated here but for the fact that the canonical section of the period operator needs to be adapted. The mapping defined in (6.3) becomes

$$\mathfrak{K} : \mathbb{C}\{\varepsilon, x, y\}' \longrightarrow \mathrm{Section}_k \{P^\tau y\}$$
$$G \longmapsto \mathfrak{S}_\ell\left(\mathfrak{T}_\ell(G)\right)$$

whose kernel coincides with $\mathcal{X} \cdot \mathbb{C}\{\varepsilon, x, y\}'$, i.e., the sequence of $\mathbb{C}\{\varepsilon\}$-linear operators

$$0 \longrightarrow \mathbb{C}\{\varepsilon, x, y\}' \xrightarrow{\mathcal{X}\cdot} \mathbb{C}\{\varepsilon, x, y\}' \xrightarrow{\mathfrak{K}} \mathrm{Section}_k \{P^\tau y\} \longrightarrow 0$$

is exact. Up to this modification the temporal part of Realization Theorem is established.

The most obvious reason why one must adapt the target space of the section operator is computational. Proposition 10.5 below recalls the formula for the period of the formal model $\widehat{\mathrm{X}}$ for $k = 1$. For $xy^m \in \mathrm{Section}_k\{y\}$, $m \in \mathbb{N}$, it may happen that $\widehat{\mathfrak{T}}(xy^m)$ vanishes, exactly when $m\mu \in \mathbb{Z}_{\leq 0}$. This situation cannot happen if $\mu_0 \notin \mathbb{R}_{\leq 0}$, of course. Pre-composing xy^m by $P^\tau(x)\,y$ yields

$$\widehat{\mathfrak{T}}\left(xP^{m\tau}(x)\,y^m\right) = \widehat{\mathfrak{T}}\left(x^{m\tau(k+1)+1}y^m\right) + \mathrm{O}(\varepsilon),$$

and by hypothesis $m(\mu_0 + (k+1)\tau) \notin \mathbb{Z}_{\leq 0}$. As already noticed, the presence of P^τ acts as a shift by $(k+1)\tau$ on powers of x. Here it guarantees that $\mathfrak{S}_\ell$ remains invertible. Notice that the map $\mathfrak{S}_\ell$ needs to undertake the same modification as in (8.1); compare (6.23). We will not go into further details.

8.3. Alternate normal forms

The normal forms we propose in the Normalization Theorem are not strictly speaking a generalization of [26, 43], which is what we expected to accomplish in the first place and which we propose as a conjecture.

Conjecture 8.5. – *Fix $k \in \mathbb{N}$, a germ of holomorphic function $\mu \in \mathbb{C}\{\varepsilon\}$, and $\widehat{\tau} \in \mathbb{Z}_{\geq 0}$ such that $\mu_0 + \widehat{\tau} \notin \mathbb{R}_{\leq 0}$. Any generic convergent unfolding of a germ of saddle-node holomorphic vector field with the formal invariant μ is orbitally conjugate to an unfolding of the form*

$$\widehat{\mathrm{X}} + y\widehat{R}\frac{\partial}{\partial y}, \quad \widehat{R} \in x\mathbb{C}[x]_{<k}\left\{x^{\widehat{\tau}}y\right\}.$$

Such a form is unique up to conjugacy by linear maps $(\varepsilon, x, y) \mapsto (\varepsilon, x, c_\varepsilon y)$, $c \in \mathbb{C}\{\varepsilon\}^\times$.

(A similar conjecture can be stated for the temporal part.) This conjecture is very likely to be true as we almost managed to ascertain both the geometric normalization and the cellular realization in that form. In both questions we encountered difficulties of a technical nature, which can surely be overcome by bringing in tedious estimates.

9. Bernoulli unfoldings

The primary aim of this section is to establish that the compatibility condition is not trivially satisfied by proving the Parametrically Analytic Orbital Moduli Theorem. The most difficult direction is (1) ⇒ (2). The whole proof is geared toward using rigidity results of Abelian finitely generated pseudogroups $G < \mathrm{Diff}(\mathbb{C}, 0)$. Let us briefly explain how Abelian pseudogroups come into consideration here. Elements $\psi_\ell[\gamma]$ and $\psi_{\widetilde{\ell}}[\gamma]$ in overlapping cellular necklace dynamics are conjugate by the transition mapping $\delta_{\widetilde{\ell}\leftarrow\ell}$ coming from the compatibility condition. The parametric holomorphy of $\mathfrak{m}$ forces the equality $\psi_\ell[\Gamma] = \psi_{\widetilde{\ell}}[\Gamma]$ for well-chosen loops Γ, from which stems the commutativity relation

$$\psi_\ell[\Gamma] \circ \delta_{\widetilde{\ell}\leftarrow\ell} = \delta_{\widetilde{\ell}\leftarrow\ell} \circ \psi_\ell[\Gamma].$$

Such pseudogroups are completely understood and form now a classical topic of complex dynamical systems, we refer for instance to [8, 25]. "Bernoulli diffeomorphisms" (defined below) play a central role in this theory as archetypal examples of solvable and Abelian pseudogroups.

9.1. Bernoulli diffeomorphism

DEFINITION 9.1. – We say that $\psi \in \mathrm{Diff}(\mathbb{C},0)$ is a *Bernoulli diffeomorphism of index* $d \in \mathbb{N}$ if there exist $\alpha, \beta \in \mathbb{C}$ with $\alpha \neq 0$ such that

$$\psi(h) = \frac{\alpha h}{\left(1+\beta h^d\right)^{1/d}} =: \mathrm{Ber}\left(d, \begin{matrix}\alpha \\ \beta\end{matrix}\right)(h).$$

We define $\mathrm{Ber}(d)$ the set of all such algebraic functions, regardless of the special values of α and β (these are in particular germs of analytic diffeomorphisms at the origin). Of course when $d \neq \widetilde{d}$ the intersection $\mathrm{Ber}(d) \cap \mathrm{Ber}\left(\widetilde{d}\right)$ coincides with the group $\mathrm{GL}_1(\mathbb{C})$.

Let us quickly state without proof the next basic property.

LEMMA 9.2. – *The set* $\mathrm{Ber}(d)$ *is a group equipped with a semi-direct law. More precisely*

$$\mathrm{Ber}\left(d, \begin{matrix}\alpha \\ \beta\end{matrix}\right) \circ \mathrm{Ber}\left(d, \begin{matrix}\widetilde{\alpha} \\ \widetilde{\beta}\end{matrix}\right) = \mathrm{Ber}\left(d, \begin{matrix}\alpha\widetilde{\alpha} \\ \beta\widetilde{\alpha}^d + \widetilde{\beta}\end{matrix}\right).$$

The definition of Bernoulli diffeomorphisms is motivated by the following computation.

LEMMA 9.3. – *The necklace dynamics of an unfolding of Bernoulli vector field* $\mathscr{X} = \widehat{\mathrm{X}} + y^{d+1} r(x) \frac{\partial}{\partial y}$ *consists in Bernoulli diffeomorphisms of index* d. *Moreover*

$$\mathfrak{m}(\mathscr{X}) = -\frac{1}{d}\log\left(1 + 2\mathrm{i}\pi d \widehat{\mathfrak{T}}\left(y^d r\right)\right).$$

Proof. – As in [46, Section 3.3] one tries and finds an expression for the sectorial first integrals H^j in the form

$$H^j(x,y) = \frac{\widehat{H}^j(x,y)}{\left(1 - d f^j(x) y^d\right)^{1/d}}.$$

Because

$$\begin{aligned}\mathscr{X} \cdot H^j &= \frac{\widehat{H}^j}{\left(1-df^j(x)y^d\right)^{1/d+1}}\left(\left(1-df^j(x)y^d\right)y^d r(x) + \mathscr{X}\cdot\left(f^j(x)y^d\right)\right)\\ &= \frac{\widehat{H}^j}{\left(1-df^j(x)y^d\right)^{1/d+1}}\left(y^d r(x) + \widehat{X}\cdot\left(f^j(x)y^d\right)\right),\end{aligned}$$

then H^j is a first integral for $\mathscr{X}$ if and only if

$$\widehat{\mathrm{X}} \cdot \left(y^d f^j(x)\right) = -y^d r(x). \tag{9.1}$$

This equation admits a formal solution (Lemma 4.8) because $\widehat{\mathrm{X}}$ is linear in the y-variable, and the $f^j(x)y^d$ are the sectorial solutions of this equation (Theorem 6.8). In fact $(x,y) \mapsto \left(x, \frac{y}{\left(1-dy^d f^j(x)\right)^{1/d}}\right)$ is the canonical sectorial normalization of $\mathscr{X}$.

First notice that by definition of the period operator for the formal model (Definition 6.10) we have for all $(x,y) \in V^{j,\mathrm{s}} \times \mathbb{C}$:

$$y^d f^{j+1}(x) - y^d f^j(x) = -\widehat{\mathfrak{T}}^j\left(y^d r\right)\left(\widehat{H}^j(x,y)\right).$$

From the special form of H^j we deduce

$$\begin{aligned}\frac{H^{j+1}}{H^j\exp 2\mathrm{i}\pi\mu/k} &= \left(1-dy^d\frac{f^{j+1}-f^j}{1-df^jy^d}\right)^{-1/d}\\ &= \exp\left(-\frac{1}{d}\log\left(1+d2\mathrm{i}\pi\frac{\widehat{\mathfrak{T}}^j\left(y^dr\right)\left(\widehat{H}^j\right)}{\left(\widehat{H}^j\right)^d}\times\left(H^j\right)^d\right)\right).\end{aligned}$$

Because $\widehat{H}$ is linear in the y-variable we know that $\widehat{\mathfrak{T}}\left(y^dr\right)(h) = \alpha h^d$ for some complex coefficients $\alpha = \left(\alpha_\varepsilon^j\right)_{j\in\mathbb{Z}/k\mathbb{Z}}$. Hence $\frac{\widehat{\mathfrak{T}}(y^dr)(\widehat{H})}{\widehat{H}^d}\times H^d = \widehat{\mathfrak{T}}\left(y^dr\right)(H)$. The rest follows from (7.3). $\square$

Remark 9.4. – In the course of the proof we establish in particular that Bernoulli unfoldings admit families of Liouvillian first-integrals of the form

$$H(x,y) = \frac{\widehat{H}(x,y)}{\left(1-dy^df(x)\right)^{1/d}}$$

for the Liouvillian solution f of (9.1) obtained by variation of the constant

$$f(x) := E(x)^{-1}\int^x E(z)\,r(z)\,\frac{\mathrm{d}z}{P(z)}$$

where E is solution of

$$P(x)\,E'(x) = dE(x)\left(1+\mu x^k\right).$$

9.2. Holomorphic modulus: proof of the Parametrically Analytic Orbital Moduli Theorem

The direction (2)⇒(1) is a consequence of Lemma 9.3 above and of Proposition 10.5 below stating that the model period operator $\widehat{\mathfrak{T}}\left(y^dr\right)$ is analytic in the parameter when $k=1$ and $d\mu\in\mathbb{Z}$.

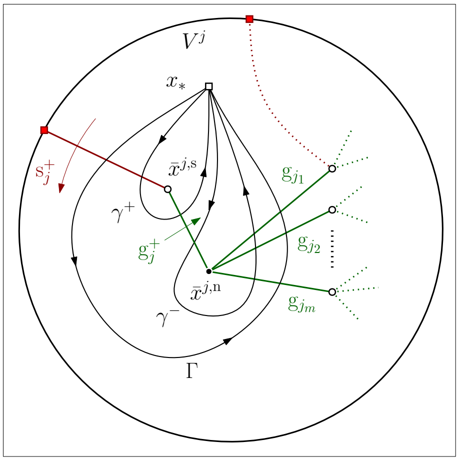
Conversely let us suppose that $(\mu,\mathfrak{m})$ is realizable and that $\mathfrak{m}_\ell = \phi|_{\mathcal{E}_\ell\times(\mathbb{C},0)}$ for some holomorphic k-tuple

$$\phi = \left(\phi^j\right)_j \in h\mathbb{C}\{\varepsilon,h\}^k.$$

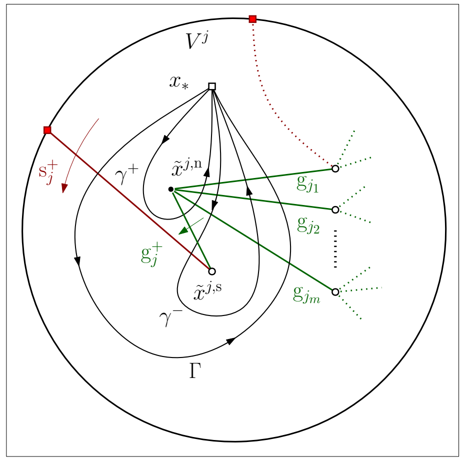
If $\phi=0$ then $\mathfrak{m}=\mathfrak{m}\left(\widehat{\mathrm{X}}\right)$ (Theorem 7.6), so we can as well assume that $\phi\neq 0$. We first establish that $k=1$ by contraposition, and then present the case $k=1$. That case can be found originally in [47, Proposition 6] for $\mu=0$. We generalize here the result to arbitrary μ.

Recall that for $c\in\mathbb{C}^\times$ we write

$$L_c : h\longmapsto ch.$$



(A) In the parameter $\overline{\varepsilon}$



(B) In the parameter $\widetilde{\varepsilon}$

FIGURE 9.1. The construction involved in Lemma 9.5.

9.2.1. *Reduction to the case $k = 1$.* – Assume then that $k > 1$ and prove $\phi = \left(\phi^j\right)_j = 0$. For each $j \in \mathbb{Z}/k\mathbb{Z}$ there exists a cell $\mathcal{E}_\ell$ for which $x^{j,\mathrm{s}}$ is attached to only one saddle sector $V^{j,\mathrm{s}}$. Let $x^{j,\mathrm{n}}$ be the node point attached to $x^{j,\mathrm{s}}$ in the boundary of $V^{j,\mathrm{s}}$. The cell $\mathcal{E}_\ell$ self-intersects around a regular part of Δ_k in such a way that the nature of the points $x^{j,\mathrm{s}}$ and $x^{j,\mathrm{n}}$ is exchanged when seen from one part or the other of the intersection. With the conventions discussed in Remark 7.17, by this we mean

$$\begin{cases} \overline{x}^{j,\mathrm{s}} = \widetilde{x}^{j,\mathrm{n}} \\ \overline{x}^{j,\mathrm{n}} = \widetilde{x}^{j,\mathrm{s}}. \end{cases}$$

We refer to Figures 7.4 and 9.1.

Fix a base-point and base-sector $x_* \in V^j \backslash \rho_\varepsilon \mathbb{D}$ and take γ^-, γ^+ two loops based at x_* of index 1 around respectively $x^{\mathrm{n}}_{\overline{\varepsilon}}$ and $x^{\mathrm{s}}_{\overline{\varepsilon}}$, and index 0 with respect to the other roots of P as in Figure 9.1. Let $\Gamma := \gamma^+\gamma^-$ be a loop encircling only $\{x^{\mathrm{n}}, x^{\mathrm{s}}\}$. The compatibility condition ensures the existence of a tangent-to-identity map

$$\delta := \delta_{\ell \leftarrow \ell}$$

which conjugates the respective necklace dynamics based at x_*. In particular

$$\begin{aligned} \delta^* \overline{\psi}\left[\gamma^\pm\right] &= \widetilde{\psi}\left[\gamma^\pm\right], \\ \delta^* \overline{\psi}\left[\Gamma\right] &= \widetilde{\psi}\left[\Gamma\right]. \end{aligned} \tag{9.2}$$

LEMMA 9.5. – *We follow the notations of Figure 9.1. Let $k > 1$, and let $m \geq 1$ be the number of singular points different from $x^{j,\mathrm{s}}$ and $x^{j,\mathrm{n}}$. Each passage of a gate by Γ in the figure yields a linear gate map $L_{\overline{\nu}^{j_p}}$* (resp. $L_{\widetilde{\nu}^{j_p}}$) *for some $\overline{\nu}^{j_p} \in \mathbb{C}^\times$ (resp. $\widetilde{\nu}^{j_p} \in \mathbb{C}^\times$) and $1 \leq p \leq m$. We also set $\psi^{j,\mathrm{g}} =: L_{\nu^j}$.*

1. *The equality $\overline{\psi}\left[\Gamma\right] = \widetilde{\psi}\left[\Gamma\right]$ holds, defining a germ $\Delta \in \mathrm{Diff}\left(\mathbb{C}, 0\right)$.*

2. $\Delta'(0)\exp\dfrac{-2i\pi\mu}{k} = \prod_{p=1}^{m}\overline{\nu}^{j_p} = \prod_{p=1}^{m}\widetilde{\nu}^{j_p}$.

3. $\delta\circ\Delta = \Delta\circ\delta$.

Proof. – Observe that

$$\widetilde{\psi}\left[\gamma^{+}\right] = L_{\widetilde{c}/\widetilde{\nu}^{j}}\,,\quad \widetilde{c} := \prod_{p=1}^{m}\widetilde{\nu}^{j_p}. \tag{9.3}$$

The linear part is invariant by conjugacy so that $\widetilde{c} = \overline{\nu}^{j}\widetilde{\nu}^{j}\exp{}^{2i\pi\mu/k}$. Similarly, considering γ^{-} yields $\overline{c} = \overline{\nu}^{j}\widetilde{\nu}^{j}\exp{}^{2i\pi\mu/k}$. Hence

$$\overline{c} = \prod_{p=1}^{m}\overline{\nu}^{j_p} = \overline{\nu}^{j}\widetilde{\nu}^{j}\exp{}^{2i\pi\mu/k} = \widetilde{c}.$$

Since $\widetilde{\psi}\left[\Gamma\right] = L_{\widetilde{c}}\circ\psi^{j,\mathrm{s}}$ and $\overline{\psi}\left[\Gamma\right] = L_{\overline{c}}\circ\psi^{j,\mathrm{s}}$ the result follows. □

Recall that the map

$$\varepsilon\in\mathcal{E}_{\ell}\longmapsto\left(\nu_{\varepsilon}^{j}\right)_{j\in\mathbb{Z}/k\mathbb{Z}}$$

is locally injective (Lemma 7.20). In particular $\Delta'(0)$ is not constant and therefore must take non-rational values on a small subdomain $\Lambda\subset\mathcal{E}_{\ell}^{\cap}$. It follows that for $\varepsilon\in\Lambda$ the Abelian group $\langle\delta,\Delta\rangle < \mathrm{Diff}(\mathbb{C},0)$ is non-resonant and therefore formally linearizable [25]. Hence $\delta = \mathrm{Id}$.

LEMMA 9.6. – *If* $\delta = \mathrm{Id}$ *then* $\phi^{j} = 0$.

Proof. – According to (9.2), δ conjugates $\psi_{\overline{\varepsilon}}\left[\gamma^{+}\right] = L_{\overline{\nu}^{j}}\circ\psi^{j,\mathrm{s}}$ to $\psi_{\widetilde{\varepsilon}}\left[\gamma^{+}\right]$, but the latter is linear thanks to (9.3), therefore $\psi^{j,\mathrm{s}}$ itself is linear. It can only mean that $\phi^{j,\mathrm{s}} = 0 = \phi^{j}$ using (7.5), the equality holding on the whole cell $\mathcal{E}_{\ell}$ by analytic continuation. □

Since j is arbitrary we just established

$$(k>1)\Longrightarrow(\phi=0)\,.$$

9.2.2. *The case* $k=1$*: end of the proof of the Parametrically Analytic Orbital Moduli Theorem*

Since $k=1$ we drop the index $j=0$. We work in the self-intersection $\mathcal{E}^{\cap}$ of the single parametric cell, and use the notations and constructions involved just above. In particular Figure 9.1 remains the same except for the fact that there are no gate passages $j_1,\ldots,j_m$ on the right-hand side of the pictures.

Recall that we consider a system with $\mathfrak{m}\neq 0$. Lemma 9.6 forbids $\delta = \mathrm{Id}$, thus $\Delta = \psi^{\mathrm{s}}$ is non-linear (Δ was introduced in Lemma 9.5). Then $\langle\delta,\Delta\rangle$ is an Abelian group. Consequently there exists [31] a formal tangent-to-identity change $\widehat{\varphi}$ in the variable h, unique $d\in\mathbb{N}$, $\lambda\in\mathbb{C}$ and $t\in\mathbb{C}\backslash\{0\}$ such that, writing $\widehat{f} := \widehat{\varphi}^{*}f$ for all $f\in\mathrm{Diff}(\mathbb{C},0)$,

$$\widehat{\delta} = \Phi^{1}_{Z(d,\lambda)}$$
$$\widehat{\Delta} = \alpha\Phi^{t}_{Z(d,\lambda)},\quad \alpha\in\mathbb{C}^{\times}$$
$$Z(d,\lambda) = \frac{h^{d+1}}{1+\lambda h^{d}}\frac{\partial}{\partial h}.$$

Commutativity forces the relation

$$\alpha^d = 1.$$

Since $\alpha = \exp 2i\pi\mu$ this gives $d\mu \in \mathbb{Z}$ as expected. Observe that for all $s \in \mathbb{C}$

$$\Phi^s_{Z(d,0)} \in \mathrm{Ber}\,(d)\,,$$

therefore we aim at showing $\lambda = 0$. This is ultimately done by applying the next lemma.

Lemma 9.7 ([8, Assertions 1.1 to 1.4]). – *In the following ξ is a formal diffeomorphism in the variable h at* 0.

1. *Let $Z, \widetilde{Z}$ be formal vector fields in the variable h at* 0 *belonging to $h^{d+1}\mathbb{C}\,[[h]]^\times \frac{\partial}{\partial h}$. If $\xi^*\Phi^1_Z = \Phi^1_{\widetilde{Z}}$ then $\xi^* Z = \widetilde{Z}$ (the converse is trivial).*
2. *Assume that $\xi^* Z\,(d,\lambda) = aZ\,(d,\lambda)$ with $a \neq 1$. Then $\lambda = 0$ and $\xi \in \mathrm{Ber}\,(d)$ (in particular ξ is analytic).*

Let us show now that $\lambda = 0$ and $\widehat{\varphi} \in \mathrm{Ber}\,(d)$ itself, forcing $\Delta = \psi^{\mathrm{s}} \in \mathrm{Ber}\,(d)$ by application of Lemma 9.2. The key is to exploit the fact (9.2), which can be rewritten as:

$$\delta^* \overline{\psi}\left[\mathrm{g}^+\mathrm{s}^+\right] = \widetilde{\psi}\,[\mathrm{g}^-] = L_{1/\widetilde{\nu}}. \tag{9.4}$$

Indeed, referring to Definition 7.13 for the definition of the letters $\mathrm{g}^\pm, \mathrm{s}^\pm$ and their image by $\psi\,[\bullet]$, and looking at Figure 9.1, we compare the holonomy maps around the upper singular point. On the left, the singular point is of saddle type and the holonomy map is the composition of $\overline{\psi}\left[s^+\right]$ (crossing the saddle sector in the direction of the arrow) with $\overline{\psi}\left[\mathrm{g}^+\right]$ (crossing the gate sector in the direction of the arrow). On the right, the same singular point is of node type. Turning around, it comes to crossing the gate sector in the inverse direction of the arrow. Hence its holonomy map is $\widetilde{\psi}\,[\mathrm{g}^-]$. The last equality in (9.4) follows from the fact that $\widetilde{\psi}\left[\mathrm{g}^+\right] = L_{\widetilde{\nu}}$. Note that (9.4) means that δ linearizes $\overline{\psi}\left[\mathrm{g}^+\mathrm{s}^+\right]$.

Of course the multipliers at the fixed point in (9.4) must be the same. On the left, this multiplier is simply that of $\overline{\psi}\left[\mathrm{g}^+\mathrm{s}^+\right]$, since conjugacy by δ preserves the multiplier. On the one hand the multiplier at the fixed point of $\overline{\psi}\left[\mathrm{s}^+\right]$ is $\exp 2i\pi\mu$, according to (7.5) for $k := 1$, as indeed $\overline{\psi}\left[\mathrm{g}^+\right] = L_{\overline{\nu}}$. On the other hand $\widetilde{\psi}\,[\mathrm{g}^-] = L_{1/\widetilde{\nu}}$ so that

$$\widetilde{\nu}\,\overline{\nu}\exp 2i\pi\mu = 1.$$

We also have $\overline{\psi}\left[\mathrm{s}^+\right] = \Delta$, since it is the holonomy obtained by turning counterclockwise around the two singular points. Hence, replacing in (9.4) yields $L_{\overline{\nu}} \circ \Delta \circ \delta = \delta \circ L_{1/\widetilde{\nu}}$. Composing both sides on the left with $L_{\widetilde{\nu}}$ and taking $\widehat{\varphi}^*$ on both sides yields

$$\widehat{L}^*_{1/\widetilde{\nu}}\widehat{\delta} = \widehat{L}_{\widetilde{\nu}} \circ \widehat{\delta} \circ \widehat{L}_{1/\widetilde{\nu}} = \widehat{L}_{\overline{\nu}\widetilde{\nu}} \circ \widehat{\Delta} \circ \widehat{\delta}.$$

For the sake of simplicity we only deal with the case $\mu \in \mathbb{Z}$, the general case can be adapted by taking into account that $\widehat{L}^{\circ d}_{\overline{\nu}\widetilde{\nu}} = \mathrm{Id}$. Under the current hypothesis $\widehat{L}_{\overline{\nu}\widetilde{\nu}} = \mathrm{Id}$, so that $\widehat{L}_{1/\widetilde{\nu}}$ is a formal conjugacy between $\widehat{\delta} = \Phi^1_{Z(d,\lambda)}$ and $\widehat{\Delta} \circ \widehat{\delta} = \Phi^{1+t}_{Z(d,\lambda)} = \Phi^1_{(1+t)Z(d,\lambda)}$ for some $t = t_\varepsilon \in \mathbb{C}^\times$. According to Lemma 9.7 with $\xi := \widehat{L}_{1/\widetilde{\nu}}$ and $a := 1 + t \neq 1$, we have $\lambda = 0$ and $\widehat{L}_{1/\widetilde{\nu}} \in \mathrm{Ber}\,(d)$.

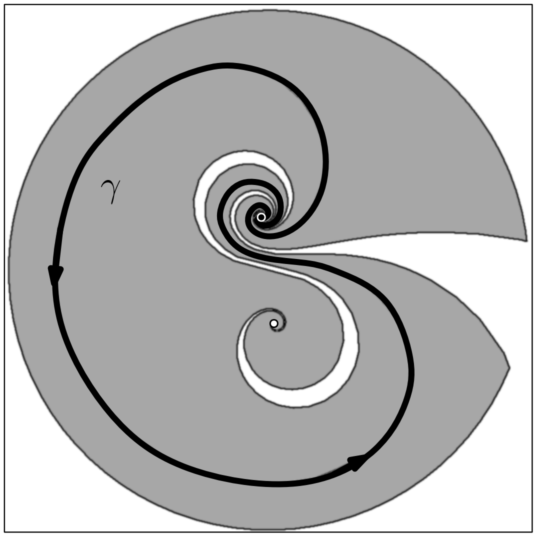


FIGURE 10.1. The (asymptotic) path of integration used to compute the period, which is a cycle when $k = 1$.

So far $\widehat{\varphi}$ is a formal linearization of $\widehat{L}_{1/\widetilde{\nu}}$ which is tangent-to-identity. For values ε of the parameter corresponding to $\widetilde{\nu} \notin \mathbb{R}$ (say $\Im(\widetilde{\nu}) > 0$) the fix-point 0 of $\widehat{L}_{1/\widetilde{\nu}}$ is hyperbolic: the map $\widehat{\varphi}$ is locally holomorphic at 0, unique and therefore given by

$$\widehat{\varphi} := \lim_{n\to\infty} L_{-n/\widetilde{\nu}} \circ \widehat{L}_{1/\widetilde{\nu}}^{\circ n}$$

uniformly on a neighborhood of 0. Lemma 9.2 implies that for every $n \in \mathbb{N}$ we have

$$L_{-n/\widetilde{\nu}} \circ \widehat{L}_{1/\widetilde{\nu}}^{\circ n} \in \mathrm{Ber}(d),$$

therefore $\widehat{\varphi} \in \mathrm{Ber}(d)$ as requested, since the group $\mathrm{Ber}(d)$ is closed for the topology of local uniform convergence. This completes the proof of the Parametrically Analytic Orbital Moduli Theorem.

10. A few words about computations

All the discussion regarding the actual (symbolic or numeric) computations of normal forms and moduli of saddle-nodes, as presented in [43, Section 4] for saddle-nodes, can be repeated *verbatim* in the case of convergent unfoldings: we will not reproduce it here. We nonetheless present in Section 10.1 a consequence of one particular result, thus unfolding the main result of [46], which leads us to try and compute the period associated to the formal orbital normal form $\widehat{X}$ in Section 10.2.

10.1. Computation of the dominant term of the orbital invariant

The next lemma holds for a fixed value of $\varepsilon \in \mathcal{E}_\ell$.

Lemma 10.1 (See [43, Proposition 4.1]). – *Let $r_n \in x\mathbb{C}[x]_{<k}$ be the coefficients of*

$$R(x,y) = \sum_{n>0} r_n(x)(P^\tau(x)y)^n \tag{10.1}$$

in the normal form $\mathcal{X}$. Let $c_\ell^{j,p}(n,m) \in \mathbb{C}$ be the coefficients of the period

$$\mathfrak{T}_\ell^j(x^n y^m)(h) = \sum_{p>0} c_\ell^{j,p}(n,m)h^p$$

relative to $\mathcal{X}$. Then we have the following properties.

Triangularity: $c_\ell^{j,p}(n,m) = 0$, *if $p < m$ and*

$$c_\ell^{j,m}(m,m)h^m = 2\mathrm{i}\pi\widehat{\mathfrak{T}}_\ell^j(x^n y^m)(h)$$

is independent of R.

Algebraicity: *For $p > m$, the coefficient $c_\ell^{j,p}(n,m)$ depends polynomially on the $k(p-m)$ variables given by the coefficients of $r_1, \ldots, r_{p-m}$ and vanishes when $R = 0$.*

Proof. – It is exactly the proof done in [43, Proposition 4.1] since exchanging x^{k+1} for $P_\varepsilon(x)$ does not modify anything in the actual computation. We give some brief elements of the proof.

Let us drop all indexes and let $x \mapsto y(x,h)$ be the sectorial solution of the differential equation induced by the vector field $\mathcal{X}$ with initial value $H(x_*, y(x_*,h)) = h$ (here x_* is fixed once and for all in V^s). Computing $\mathfrak{T}(x^n y^m)(h)$ requires to compute the integral $\int_\gamma x^n y(x,h)^m \frac{\mathrm{d}x}{P(x)}$ for an asymptotic path $\gamma \subset \mathbb{C}\times\{0\}$ joining the two nodes in the closure of the union of consecutive squid sectors (see Figure 10.1). This integral is absolutely convergent because $m > 0$ and γ spirals in the right manner (see also Lemma 6.23). Since

$$H(x,y) = \widehat{H}(x,y)\exp N(x,y),$$

with $\widehat{H}$ linear in the y-variable, we necessary have

$$y(x,h) = \widehat{y}(x,h) + h\mathrm{O}(h) \tag{10.2}$$

where $\widehat{y}(x,h) = \frac{h}{\widehat{H}(x,1)}$ is the solution corresponding to the formal model $\widehat{\mathrm{X}}$. This gives the triangularity. The algebraicity property stems from the fact that the computation can be performed formally in the y-variable. The sought property is true for the expansion (10.2) (by studying the inverse of the normalizing mapping) because it is true for solutions of cohomological equations $\mathcal{X} \cdot N = -R$. □

We extract from this statement useful consequences.

Proposition 10.2. – 1. *The quantity*

$$\begin{aligned}\inf\{n : r_n \neq 0\} &= \inf\left\{n : (\exists j)\ \phi_n^j \neq 0\right\}\\ &=: d \in \overline{\mathbb{N}}\end{aligned}$$

does not depend on the cell.

2. *The valuation d is infinite if and only if the unfolding is analytically conjugate to its formal normal form.*

3. *If $d < \infty$ the dominant term of the invariant is given by the period of the formal model*

$$2\mathrm{i}\pi\widehat{\mathfrak{T}}\left(r_d y^d\right) = \left(h \mapsto \phi_d^j h^d\right)_{j\in\mathbb{Z}/k\mathbb{Z}}.$$

Remark 10.3. – The value of d does not depend on the cell but may differ from the value obtained on the boundary Δ_k. Yet the analytic continuation principle ensures that

$$\min_{\varepsilon\in\Delta_k} \inf\{n : r_{n,\varepsilon} \neq 0\} \geq d$$

because R is analytic.

From this proposition we deduce a final normalization ensuring uniqueness.

Corollary 10.4. – *Assume the generic convergent unfolding X is not analytically conjugate to its formal normal form $\widehat{\mathrm{X}}$ defined in* (2.2)*. There exists a unique $(\kappa, j, d) \in \mathbb{Z}_{\geq 0}\times\{1, 2, \ldots, k\}\times\mathbb{N}$ such that X is conjugate to the normal form $\mathscr{X} = \widehat{\mathrm{X}} + Ry\frac{\partial}{\partial y}$ as in* (10.1) *where:*

$$r_{\varepsilon,d}(x) = \varepsilon^\kappa x^j + \mathrm{o}\left(x^j\right)$$
$$r_{\varepsilon,n} = 0 \quad \textit{if } n < d.$$

Notice that in the case $\kappa > 0$ this normal form may fail to deliver meaningful information at the limit $\varepsilon \to 0$. Take the extreme case $R_\varepsilon(x, y) = \varepsilon^\kappa x^j y^d$ with $\kappa > 0$: for every $\varepsilon \neq 0$ the vector field $\mathscr{X}$ is not equivalent to the model $\widehat{\mathrm{X}}_\varepsilon$ but $\mathscr{X}_0$ is.

10.2. Formula for the period of formal models

Unfortunately only the case $k = 1$ seems tractable enough to obtain closed-form expressions involving the Gamma function. For the case $k = 2$ one could derive a closed-form formula additionally using generalized hypergeometric functions, which is already stretching a bit far what a "closed-form" is. There is no evidence that similar calculations can be performed for $k > 2$.

Proposition 10.5 ([47, Proposition 8]). – *Here $k = 1$. Let us introduce the double covering $\varepsilon = -s^2$ in the parameter space. Then for $m \in \mathbb{N}$ and $n \in \mathbb{Z}_{\geq 0}$:*

$$\widehat{\mathfrak{T}}_s\left(x^n y^m\right)(h) = h^m \times \frac{(-m)^{n+m\mu}}{\Gamma(n+m\mu)} \times t_{s,n,m} \times T_{s,m}$$
$$t_{s,n,m} := \frac{1}{2^n}\sum_{p+q=n}\binom{n}{p}\prod_{j=0}^{p-1}\left(1 - s\left(\mu + \frac{2j}{m}\right)\right)\prod_{j=0}^{q-1}\left(1 + s\left(\mu + \frac{2j}{m}\right)\right)$$
$$T_{s,m} := \frac{\left(-\frac{2s}{m}\right)^{m\mu}}{1+s\mu} \times \frac{\Gamma\left(-\frac{m}{2s} + \frac{m\mu}{2}\right)}{\Gamma\left(-\frac{m}{2s} - \frac{m\mu}{2}\right)}.$$

This period is holomorphic and bounded in the parameter s on the sector

$$S := \left\{0 < |s| < \frac{1}{2|\mu_0|}, \frac{\pi}{4} < \arg s < \frac{7\pi}{4}\right\}$$

and extends continuously at 0 *by*

$$\widehat{\mathfrak{T}}_0 (x^n y^m)(h) = h^m \times \frac{(-m)^{n+m\mu_0}}{\Gamma (n + m\mu_0)}.$$

For given s *small enough, the period is zero if and only if* $n + m\mu_\varepsilon \in \mathbb{Z}_{\leq 0}$. *The period is an even function of* s *(i.e., holomorphic in the parameter* ε*) if and only if* $m\mu \in \mathbb{Z}$. *In that case* μ *is a rational constant.*

The result is shown by using the Pochhammer contour integral formula for the Beta function. Indeed an affine change of coordinates sends $(x-s)^\alpha (x+s)^\beta$ to a multiple of $(1-z)^\alpha z^\beta$. The final expression comes from diverse classical properties of the Gamma function. The eventual lack of evenness of the period comes from the term $T_{s,m}$. If $m\mu$ is not an integer then $T_{s,m}$ is multivalued and has an accumulation of zeros and poles as $s \to 0$ outside the sector S. Only the coincidence of these two infinite sets when $m\mu \in \mathbb{Z}$ allows the period to be holomorphic through lucky root / pole cancelations.

Since $T_{s,m}$ is independent on n, any nonzero period $\widehat{\mathfrak{T}} (y^m g)$ of a germ $g \in \mathbb{C}\{\varepsilon, x\}$ is holomorphic in ε if and only if $m\mu \in \mathbb{Z}$. From Lemma 9.3, Theorem 6.2 and the Parametrically Analytic Orbital Moduli Theorem we can generalize this observation.

Corollary 10.6. – *Let* $G \in \mathbb{C}\{\varepsilon, x, y\}$ *with* $G(\bullet, 0) = \mathrm{O}(P)$. *Let us assume that the period* $\widehat{\mathfrak{T}}(G)$ *is nonzero. Then,* $\widehat{\mathfrak{T}}(G)$ *is holomorphic in the parameter if and only if all three conditions hold:*

— $k = 1$ *,*

— *there exists* $d \in \mathbb{N}$ *such that* $d\mu \in \mathbb{Z}$*,*

— *there exist two germs* $F \in \mathbb{C}\{\varepsilon, x, y\}$ *and* $Q \in \mathrm{Section}_1 \left\{P^{d\tau} y^d\right\} \setminus \{0\}$ *such that*

$$G = Q + \widehat{\mathrm{X}} \cdot F.$$

The fact that the period is never a holomorphic function of the parameter if $k > 1$ is probably a sign that a “simple” formula for $\widehat{\mathfrak{T}} (x^n y^m)$ does not exist.

BIBLIOGRAPHY

[1] M. Berthier, F. Touzet, Sur l’intégration des équations différentielles holomorphes réduites en dimension deux, *Bol. Soc. Brasil. Mat. (N.S.)* **30** (1999), 247–286.

[2] B. Branner, K. Dias, Classification of complex polynomial vector fields in one complex variable, *J. Difference Equ. Appl.* **16** (2010), 463–517.

[3] A. D. Bruno, *Local methods in nonlinear differential equations*, Springer Series in Soviet Mathematics, Springer, 1989.

[4] C. Camacho, P. Sad, Invariant varieties through singularities of holomorphic vector fields, *Ann. of Math.* **115** (1982), 579–595.

[5] M. Canalis-Durand, J. Mozo-Fernández, R. Schäfke, Monomial summability and doubly singular differential equations, *J. Differential Equations* **233** (2007), 485–511.

[6] G. Casale, Feuilletages singuliers de codimension un, groupoïde de Galois et intégrales premières, *Ann. Inst. Fourier* **56** (2006), 735–779.

[7] G. Casale, D. Davy, Specialization of the Galois groupoid, preprint, 2017.

[8] D. Cerveau, R. Moussu, Groupes d'automorphismes de $(\mathbf{C}, 0)$ et équations différentielles $y dy + \cdots = 0$, *Bull. Soc. Math. France* **116** (1988), 459–488.

[9] C. Christopher, C. Rousseau, The moduli space of germs of generic families and analytic diffeomorphisms unfolding a parabolic fixed point, *C. R. Math. Acad. Sci. Paris* **345** (2007), 695–698.

[10] K. Dias, Enumerating combinatorial classes of the complex polynomial vector fields in $\mathbb{C}$, *Ergodic Theory Dynam. Systems* **33** (2013), 416–440.

[11] A. Douady, J. Estrada, P. Sentenac, Champs de vecteurs polynomiaux sur $\mathbb{C}$, preprint, 2005.

[12] H. Dulac, Recherches sur les points singuliers des équations différentielles, *Journal de l'École polytechnique* **2** (1904), 1–125.

[13] H. Dulac, Sur les points singuliers d'une équation différentielle, *Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys.* **1** (1909), 329–379.

[14] W. Fischer, H. Grauert, Lokal-triviale Familien kompakter komplexer Mannigfaltigkeiten, *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **1965** (1965), 89–94.

[15] A. A. Glutsyuk, Confluence of singular points and the nonlinear Stokes phenomenon, *Tr. Mosk. Mat. Obs.* **62** (2001), 54–104.

[16] C. Godbillon, J. Vey, Un invariant des feuilletages de codimension 1, *C. R. Acad. Sci. Paris Sér. A-B* **273** (1971), A92–A95.

[17] M. Hukuhara, T. Kimura, T. Matuda, *Equations différentielles ordinaires du premier ordre dans le champ complexe*, Publications of the Mathematical Society of Japan, 7. The Mathematical Society of Japan, Tokyo, 1961.

[18] J. Hurtubise, C. Lambert, C. Rousseau, Complete system of analytic invariants for unfolded differential linear systems with an irregular singularity of Poincaré rank k, *Mosc. Math. J.* **14** (2014), 309–338, 427.

[19] J. Hurtubise, C. Rousseau, Moduli space for generic unfolded differential linear systems, *Advances in Mathematics* **307** (2017), 1268–1323.

[20] Y. Ilyashenko, S. Yakovenko, *Lectures on analytic differential equations*, Graduate Studies in Math. **86**, Amer. Math. Soc., 2008.

[21] M. Klimeš, Confluence of Singularities of Nonlinear Differential Equations via Borel–Laplace Transformations, *J. Dyn. Control Syst.* **22** (2016), 285–324.

[22] K. Kodaira, On stability of compact submanifolds of complex manifolds, *Amer. J. Math.* **85** (1963), 79–94.

[23] C. Lambert, C. Rousseau, Moduli space of unfolded differential linear systems with an irregular singularity of Poincaré rank 1, *Mosc. Math. J.* **13** (2013), 529–550, 553–554.

[24] P. Lavaurs, Systèmes dynamiques holomorphes : explosion de points périodiques paraboliques, Ph.D. Thesis, Université Paris 11, 1989.

[25] F. Loray, Dynamique des groupes d'automorphismes de $\mathbf{C}, 0$, *Bol. Soc. Mat. Mexicana* **5** (1999), 1–23.

[26] F. Loray, Versal deformation of the analytic saddle-node, in *Analyse complexe, systèmes dynamiques, sommabilité des séries divergentes et théories galoisiennes. II*, Astérisque **297**, 2004, 167–187.

[27] B. Malgrange, Le groupoïde de Galois d'un feuilletage, in *Essays on geometry and related topics, Vol. 1, 2*, Monogr. Enseign. Math. **38**, 2001, 465–501.

[28] B. Malgrange, On nonlinear differential Galois theory, *Chinese Ann. Math. Ser. B* **23** (2002), 219–226, Dedicated to the memory of Jacques-Louis Lions.

[29] P. Mardešić, R. Roussarie, C. Rousseau, Modulus of analytic classification for unfoldings of generic parabolic diffeomorphisms, *Mosc. Math. J.* **4** (2004), 455–502, 535.

[30] J. Martinet, Remarques sur la bifurcation nœud-col dans le domaine complexe, *Astérisque* **150-151** (1987), 131–149, 186, Singularités d'équations différentielles (Dijon, 1985).

[31] J. Martinet, J.-P. Ramis, Problèmes de modules pour des équations différentielles non linéaires du premier ordre, *Inst. Hautes Études Sci. Publ. Math.* **55** (1982), 63–164.

[32] J.-F. Mattei, R. Moussu, Holonomie et intégrales premières, *Ann. Sci. École Norm. Sup.* **13** (1980), 469–523.

[33] J. Mozo Fernández, R. Schäfke, Asymptotic expansions and summability with respect to an analytic germ, *Publ. Mat.* **1** (2019), 3–79.

[34] R. Oudkerk, The parabolic implosion: Lavaurs maps and strong convergence for rational maps, in *Value distribution theory and complex dynamics (Hong Kong, 2000)*, Contemp. Math. **303**, Amer. Math. Soc., 2002, 79–105.

[35] M. van der Put, M. F. Singer, *Galois theory of linear differential equations*, Grundl. math. Wiss. **328**, Springer, 2003.

[36] G. Remoundos, Contribution à la théorie des singularités des équations différentielles du premier ordre, *Bull. Soc. Math. France* **36** (1908), 185–194.

[37] C. Rousseau, Modulus of orbital analytic classification for a family unfolding a saddle-node, *Mosc. Math. J.* **5** (2005), 245–268.

[38] C. Rousseau, Normal forms for germs of analytic families of planar vector fields unfolding a generic saddle-node or resonant saddle, in *Nonlinear dynamics and evolution equations*, Fields Inst. Commun. **48**, Amer. Math. Soc., 2006, 227–245.

[39] C. Rousseau, The moduli space of germs of generic families of analytic diffeomorphisms unfolding of a codimension one resonant diffeomorphism or resonant saddle, *J. Differential Equations* **248** (2010), 1794–1825.

[40] C. Rousseau, Analytic moduli for unfoldings of germs of generic analytic diffeomorphisms with a codimension k parabolic point, *Ergodic Theory Dynam. Systems* **35** (2015), 274–292.

[41] C. Rousseau, L. Teyssier, Analytical moduli for unfoldings of saddle-node vector fields, *Mosc. Math. J.* **8** (2008), 547–614, 616.

[42] V. I. Savel′ev, Zero-type imbedding of a sphere into complex surfaces, *Moscow Univ. Math. Bull* **37** (1982), 34–39.

[43] R. Schäfke, L. Teyssier, Analytic normal forms for convergent saddle-node vector fields, *Ann. Inst. Fourier* **65** (2015), 933–974.

[44] L. Teyssier, Analytical classification of singular saddle-node vector fields, *J. Dynam. Control Systems* **10** (2004), 577–605.

[45] L. Teyssier, Équation homologique et cycles asymptotiques d'une singularité nœud-col, *Bull. Sci. Math.* **128** (2004), 167–187.

[46] L. Teyssier, Examples of non-conjugated holomorphic vector fields and foliations, *J. Differential Equations* **205** (2004), 390–407.

[47] L. Teyssier, Coalescing complex planar stationary points, in *Interdisciplinary Mathematical Research and Applications: In Honor of Professor Christiane Rousseau* (T. Bourama, ed.), Proceedings in Mathematics & Statistics **157**, Springer, 2016, 497–593.

[48] T. Ueda, On the neighborhood of a compact complex curve with topologically trivial normal bundle, *J. Math. Kyoto Univ.* **22** (1982/83), 583–607.

[49] S. M. Voronin, Y. I. Meshcheryakova, Analytic classification of germs of holomorphic vector fields with a degenerate elementary singular point, *Vestnik Chelyab. Univ. Ser. 3 Mat. Mekh. Inform.* **3(9)** (2003), 16–41.

(Manuscrit reçu le 9 juillet 2016 ;
accepté le 14 septembre 2019.)

C. Rousseau
Département de Mathématiques et de Statistique
Université de Montréal
Canada
E-mail: christiane.rousseau@umontreal.ca

L. Teyssier
Laboratoire IRMA
CNRS & Université de Strasbourg
France
E-mail: teyssier@math.unistra.fr

Ann. Scient. Éc. Norm. Sup.
4^e série, t. 54, 2021, p. 235 à 257

A SHARP FREIMAN TYPE ESTIMATE FOR SEMISUMS IN TWO AND THREE DIMENSIONAL EUCLIDEAN SPACES

BY Alessio FIGALLI AND David JERISON

ABSTRACT. – Freiman's theorem is a classical result in additive combinatorics concerning the approximate structure of sets of integers that contain a high proportion of their internal sums. As a consequence, one can deduce an estimate for sets of real numbers: "If $A \subset \mathbb{R}$ and $\left|\frac{1}{2}(A+A)\right| - |A| \ll |A|$, then A is close to its convex hull." In this paper we prove a sharp form of the analogous result in dimensions 2 and 3.

RÉSUMÉ. – Le théorème de Freiman est un résultat classique de la combinatoire additive concernant la structure approximative des ensembles d'entiers qui contiennent une forte proportion de leurs sommes internes. En conséquence, on déduit l'estimée suivante : "Si $A \subset \mathbb{R}$ et $\left|\frac{1}{2}(A+A)\right| - |A| \ll |A|$, alors A est proche de son enveloppe convexe." Dans cet article, nous prouvons une forme optimale du résultat correspondant en dimensions 2 et 3.

1. Introduction

Given a set $A \subset \mathbb{R}^n$, define the semisum by

$$\tfrac{1}{2}(A+A) := \left\{ \tfrac{x+y}{2} \,:\, x \in A,\ y \in A \right\}.$$

Evidently, $\frac{1}{2}(A+A) \supset A$, and for convex sets K, $\frac{1}{2}(K+K) = K$. Also, $\left|\frac{1}{2}(A+A)\right| = |A| > 0$ implies that A is equal to its convex hull $\mathrm{co}(A)$ minus a set of measure zero (see [3, Théorème 6]).

The stability of this statement is a natural question that has already been extensively investigated in the one dimensional case. Indeed, by approximating sets in $\mathbb{R}$ with finite unions of intervals, one can translate the problem to $\mathbb{Z}$ and in the discrete setting the question becomes a well studied problem in additive combinatorics. More precisely, set

$$\delta(A) := \left|\tfrac{1}{2}(A+A)\right| - |A|,$$

doi:10.24033/asens.2458

where $|\cdot|$ denotes the outer Lebesgue measure. The following theorem can be obtained as a corollary of a result of G. Freiman [9] about the structure of additive subsets of $\mathbb{Z}$ (see [5] for more details, and also [11] and the references therein for more recent developments on this one dimensional problem):

THEOREM 1.1. – *Let $A \subset \mathbb{R}$ be a measurable set of positive Lebesgue measure, and assume that $\delta(A) < |A|/2$. Then*

$$|\mathrm{co}(A) \setminus A| \leq 2\,\delta(A).$$

Note that the assumption $\delta(A) < |A|/2$ is necessary, as can be seen by considering the set $A = [0,1] \cup [R, R+1]$ with $R \gg 1$.

In [5, Theorem 1.2] we extended Theorem 1.1 to every dimension, but with a dimensional dependence in the exponent (see also [6] for a stability result when one considers the semisum of two different sets). Our result was as follows.

THEOREM 1.2. – *Let $n \geq 2$. There exist computable dimensional constants $\delta_n, C_n > 0$ such that if $A \subset \mathbb{R}^n$ is a measurable set of positive Lebesgue measure with $\delta(A) \leq \delta_n |A|$, then*

$$\frac{|\mathrm{co}(A) \setminus A|}{|A|} \leq C_n \left(\frac{\delta(A)}{|A|}\right)^{\alpha_n}, \qquad \text{where } \alpha_n := \frac{1}{8 \cdot 16^{n-2} n!(n-1)!}.$$

Note that the dimensional smallness assumption on $\delta(A)$ is necessary. Indeed, consider $t = 1/2$ and the set

$$A := B_1(0) \cup \{Re_1\}, \qquad R \gg 1.$$

Then $|\mathrm{co}(A) \setminus A| \approx R$ is arbitrarily large, while $\delta(A) = \left|B_{1/2}\left(\frac{R}{2}e_1\right)\right| = 2^{-n}|A|$, hence $\delta_n \leq 2^{-n}$.

The proof in [5] is based on induction on dimension and Fubini-type arguments, and it leads to a bad estimate for the exponent α_n. In fact, we believe that $\alpha_n = 1$, which we formulate more precisely in the following conjecture.

CONJECTURE 1.3. – *Suppose that A is a measurable subset of $\mathbb{R}^n$, of positive Lebesgue measure. There exist computable constants C_n and $d_n > 0$, depending only on n, such that the following holds: if $\delta(A) \leq d_n |A|$, then*

$$|\mathrm{co}(A) \setminus A| \leq C_n\,\delta(A).$$

In this paper we introduce a completely new strategy that allows us to prove this *sharp* stability estimate in dimensions 2 and 3.

THEOREM 1.4. – *Conjecture 1.3 is valid for $n \leq 3$.*

The exponent $\alpha_n = 1$ may look surprising at first sight, as most sharp stability results for minimizers of geometric inequalities in dimension $n \geq 2$ hold with the exponent $1/2$. In particular, the best possible stability exponent for the Brunn-Minkowski inequality on convex sets is $1/2$, see [8, 7]. In contrast, our stability inequality with exponent 1 is affine invariant and additive under partitions of the set by convex tilings, and these properties are crucial to the proof. Even though we have stopped at $n = 3$, the proof is by induction on n and is organized with the hope that parts of it will ultimately apply to the case of general n. There is at least one other stability inequality in which the exponent 1 is optimal in all dimensions, namely the one proved in [4]. (Observe that the exponent 1 becomes natural when looking at critical points instead of minimizers, see for instance [2, Theorem 1.2], but this is a consequence of the different definition of the "deficit" δ.)

Acknowledgments. – AF was partially supported by NSF Grants DMS-1262411 and DMS-1361122, and by ERC Grant "Regularity and Stability in Partial Differential Equations (RSPDE)". DJ was partially supported by NSF grant DMS 1500771 and a 2018 Simons Fellowship. The authors thank Emanuel Milman for pointing out to them reference [3]. The authors are very grateful to an anonymous referee, who found a major error in the first version of this paper. In that version, we asserted that the conjecture was true in all dimensions, based on the false claim that inequality (2.4) was valid with the exponent 1 rather than n. In this revision we had to make substantial changes that allow us to prove the conjecture in dimensions 2 and 3. In Remark 2.5 we briefly comment on the case $n \geq 4$.

2. Proof of Theorem 1.4

As the reader will see, many of the arguments for the proof of Theorem 1.4 are valid in any dimension. For this reason we shall work with a generic n for most of the proof, and we shall use some geometric considerations specific to $n = 2$ and $n = 3$ only towards the end.

Basic considerations

Since Theorem 1.4 is known for $n = 1$ (see Theorem 1.1), we can assume that $n \geq 2$ and, by induction on dimension, we can also assume that Theorem 1.4 holds in dimension $n - 1$.

Denote the convex hull of A by $K := \mathrm{co}(A)$. Since the theorem is affine invariant, after dilation we can assume, with no loss of generality, that $|A| = 1$. Assuming that $\delta(A) \ll 1$, it follows by [1] and/or [5, Theorem 1.2] that [(1)]

$$\mu := |K \setminus A| \ll 1. \tag{2.1}$$

In particular, $1 \leq |K| \leq 2$. Therefore, using the lemma of F. John [10], up to an affine transformation with Jacobian bounded from above and below by a dimensional constant, we can assume that K satisfies

$$B_{1/\sqrt{n}} \subset K \subset B_{\sqrt{n}} \tag{2.2}$$

for balls of radius $1/\sqrt{n}$ and $\sqrt{n}$ centered at the origin.

By approximation, [(2)] we can assume the set A is compact and that ∂K consists of finitely many polygonal faces. In particular, $\frac{1}{2}(A + A)$ is compact, hence measurable. Furthermore,

(1) Although this estimate can be deduced as a consequence of [1], that result does not provide computable constants, as the proof is based on a contradiction argument relying on compactness.

(2) One way to define a suitable approximation is to consider a sequence of finite sets $V_k \subset V_{k+1} \subset A$ such that the polyhedra $P_k = \mathrm{co}(V_k)$ satisfy $|P_k| \to |\mathrm{co}(A)|$ as $k \to \infty$, and a sequence of compact subsets $A'_k \subset A$ such that $|A'_k| \to |A|$ as $k \to \infty$. Then let $A_k := V_k \cup [A'_k \cap (1 - 1/k)P_k]$. Since $|A_k| \to |A|$, it suffices to prove the estimate of Theorem 1.4 for A_k and then let $k \to \infty$.

since all vertices of the faces are extreme points, they belong to A. Finally, we may divide each face into simplices without adding any vertices, so that ∂K can be seen as a finite union of simplices, all of whose vertices belong to A.

Reduction to a set A that contains $(1 - C\mu^{1/n})K$

We get started with the proof by showing that points of K that are sufficiently far from the boundary of K are in $\frac{1}{2}(A + A)$. Indeed, since $\|f * g\|_{L^\infty} \leq \|f\|_{L^\infty}\|g\|_{L^1}$ for any pair of functions f and g,

$$\begin{aligned} |\chi_{K/2} * \chi_{K/2}(x) - \chi_{A/2} * \chi_{A/2}(x)| &\leq |\chi_{K/2} * (\chi_{K/2} - \chi_{A/2})|(x) \\ &\quad + |\chi_{A/2} * (\chi_{K/2} - \chi_{A/2})|(x) \\ &\leq 2\|\chi_{(K\setminus A)/2}\|_{L^1} = 2^{1-n}|K \setminus A| \\ &\leq |K \setminus A| = \mu \qquad \forall\, x \in \mathbb{R}^n. \end{aligned} \tag{2.3}$$

Because K satisfies (2.2), there is a dimensional constant $\hat{c} > 0$ such that

$$\chi_{K/2} * \chi_{K/2}(x) \geq \hat{c}\,\mathrm{dist}(x, \partial K)^n \qquad \forall\, x \in K, \tag{2.4}$$

therefore

$$\{x \in K \,:\, \hat{c}\,\mathrm{dist}(x, \partial K)^n > \mu\} \subset \{\chi_{K/2} * \chi_{K/2} > \mu\}. \tag{2.5}$$

Since

$$\begin{aligned} 0 < \chi_{A/2} * \chi_{A/2}(x) &= \int_{\mathbb{R}^n} \chi_{A/2}(y)\chi_{A/2}(x - y)\, dy \\ &\Rightarrow \quad \exists\, y \in A \text{ s.t. } y \in A/2,\ x - y \in A/2 \\ &\Rightarrow \quad x \in \tfrac{1}{2}(A + A), \end{aligned} \tag{2.6}$$

it follows from (2.3), (2.5), and (2.6), that

$$(1 - \hat{C}\mu^{1/n})K \subset \{x \in K \,:\, \hat{c}\,\mathrm{dist}(x, \partial K)^n > \mu\} \subset \{\chi_{K/2} * \chi_{K/2} > \mu\} \subset \tfrac{1}{2}(A + A)$$

for some dimensional constant $\hat{C}$. Consequently, by the definition of $\delta(A)$,

$$\big|[(1 - \hat{C}\mu^{1/n})K] \setminus A\big| \leq \delta(A). \tag{2.7}$$

Denote

$$\rho := 2\hat{C}\mu^{1/n}, \quad A' := [(1 - \rho)K] \cup A.$$

Then, since $A \subset K$ and

$$\max\big\{\tfrac{1}{2}(1 - \rho) + \tfrac{1}{2}, 1 - \rho\big\} = 1 - \rho/2,$$

we have

$$\begin{aligned} \tfrac{1}{2}(A' + A') &= \big[\tfrac{1}{2}(A + A)\big] \cup \big[\tfrac{1}{2}\big((1 - \rho)K + A\big)\big] \cup (1 - \rho)K \\ &\subset \big[\tfrac{1}{2}(A + A)\big] \cup \big[\tfrac{1}{2}\big((1 - \rho)K + K\big)\big] \cup (1 - \rho)K \\ &= \big[\tfrac{1}{2}(A + A)\big] \cup (1 - \rho/2)K. \end{aligned}$$

Therefore, since $\rho/2 = \hat{C}\mu^{1/n}$, thanks to (2.7) we get

$$\delta(A') \leq \delta(A) + \big|[(1 - \rho/2)K] \setminus A\big| \leq 2\delta(A).$$

Also, again by (2.7),

$$|K \setminus A| \leq |K \setminus A'| + \delta(A).$$

Since $\mathrm{co}(A') = K$, if we prove the theorem with A' in place of A, then the result for A will follow immediately. Thus, after replacing A with A', we can assume that

$$A \supset (1-\rho)K \qquad \text{with} \quad \rho := 2\hat{C}\mu^{1/n}. \tag{2.8}$$

Recall that, by choosing d_n small enough, we can ensure that μ (and hence ρ) is arbitrarily small.

Splitting A into "simpler" sets

Denote by $\{\Sigma_i\}_{i=1}^M$ the simplices whose union is ∂K, let K_i be the convex hull of Σ_i with the origin, and define

$$A_i := A \cap K_i.$$

Note that (2.8) implies that

$$(1-\rho)K_i \subset A_i, \qquad \rho = 2\hat{C}\mu^{1/n} \ll 1. \tag{2.9}$$

Also,

$$\sum_i |K_i \setminus A_i| = |K \setminus A|. \tag{2.10}$$

Moreover, since the sets $\{K_i\}_{i=1}^M$ are convex and disjoint, the sets $\left\{\frac{1}{2}(A_i + A_i)\right\}_{i=1}^M$ are also disjoint, therefore

$$\sum_i \left|\tfrac{1}{2}(A_i + A_i)\right| = \left|\bigcup_i \tfrac{1}{2}(A_i + A_i)\right| \leq \left|\tfrac{1}{2}(A + A)\right|.$$

Since $\sum_i |A_i| = |A|$, this proves that

$$\sum_i \delta(A_i) \leq \delta(A). \tag{2.11}$$

Main lemma and conclusion

Our main lemma is the following.

Lemma 2.1. – *Let A_i, K_i, and ρ be as above. Then, for $n \leq 3$, there exist dimensional constants $\bar{C}_n \geq 1$ and $\rho_n > 0$ such that*

$$|K_i \setminus A_i| \leq \bar{C}_n \delta(A_i). \tag{2.12}$$

provided $\rho \leq \rho_n$.

Assuming Lemma 2.1 has been proved, Theorem 1.4 follows immediately. Indeed, choosing d_n sufficiently small, it follows by [5, Theorem 1.2] and the definitions of ρ and μ (see (2.8) and (2.1)) that $\rho \leq \rho_n$ provided $\delta(A) \leq d_n$. Then, adding the inequalities (2.12), (2.10), and (2.11), we find

$$|K \setminus A| = \sum_i |K_i \setminus A_i| \leq \bar{C}_n \sum_i \delta(A_i) \leq \bar{C}_n\, \delta(A),$$

as desired. Thus, we are left with proving Lemma 2.1.

Proof of Lemma 2.1

We begin by writing the lemma in a different normalized form. Fix an index i. Since inequality (2.12) is invariant under affine transformations, we may take Σ_i to be an equilateral simplex of $(n-1)$-Hausdorff measure 1, centered on the x_n-axis and contained in the hyperplane $\{x_n = 0\}$. Moreover, we may move the vertex of K_i from the origin to the point $\big(0, \dots, 0, \frac{1}{2\rho}\big)$, so that (2.8) implies that $K_i \cap \big\{x_n \geq \frac{1}{2}\big\} \subset A_i$. It suffices to prove (2.12) in this normalized situation.

To simplify the notation further, we remove the subscript i, renaming Σ_i, K_i, A_i, with the letters Σ, K, A, respectively. With these changes, we can rewrite Lemma 2.1 as follows. (Note that, in this new normalization, $|K|$ is comparable to $1/\rho$ and (2.2) is not satisfied anymore.) Here and in the sequel, $\mathscr{H}^s$ denotes the s-dimensional Hausdorff measure.

LEMMA 2.2. – *Let Σ be an equilateral $(n-1)$-simplex centered on the x_n-axis satisfying*

$$\mathscr{H}^{n-1}(\Sigma) = 1, \qquad \Sigma \subset \{x_n = 0\}.$$

Let K be the n-simplex with one vertex at $\big(0, \dots, 0, \frac{1}{2\rho}\big)$ and base Σ. Suppose that A is a compact set satisfying

$$K \cap \big\{x_n \geq \tfrac{1}{2}\big\} \subset A \subset K,$$

and that all of the vertices of Σ belong to A. Then, for $n \leq 3$, there exist dimensional constants $\bar{C}_n \geq 1$ and $\rho_n > 0$ such that

$$|K \setminus A| \leq \bar{C}_n\, \delta(A)$$

provided $\rho \leq \rho_n$.

Proof of Lemma 2.2

The rough idea of the proof is to start with the set

$$K \cap \{1 \leq x_n \leq 2\} \subset A$$

and use the fact that the vertices of Σ belong to A in order to apply the semisum operation repeatedly to generate more points of A up to errors estimated by $\delta(A)$. As we shall see, a more refined argument involving several steps will be needed. The first five steps, proving (2.15), are valid in all dimensions, but the sixth step is restricted to dimensions 2 and 3.

Step 1: Setting up an iteration. – Let $\epsilon > 0$ be a small dimensional constant to be fixed later, set $\gamma := \frac{1}{2} + \epsilon$, and define

$$K_j := K \cap \{\gamma^j \leq x_n \leq 2\gamma^j\} \qquad \forall\, j \geq 0. \tag{2.13}$$

The natural idea would be to consider consecutive layers $2^{-j} \leq x_n \leq 2^{-j+1}$, but we need to introduce the ratio $\gamma > 1/2$ to slow down the rate of decrease of x_n for reasons that we will explain after concluding Step 1. Note that, with this definition, consecutive sets K_j are not disjoint, but rather overlap in a fraction of order ϵ of their total volume.

Let the vertices of Σ be denoted by $\{\hat{x}_k\}_{k=1}^n$. We define the following sets iteratively:

$$E_0 := K_0, \qquad E_{j+1} := K_{j+1} \cap \left(\bigcup_{k=1}^{n} \tfrac{1}{2}(\hat{x}_k + E_j) \cup E_j \cup (1-\epsilon)K \right). \tag{2.14}$$

Here $(1-\epsilon)K$ denotes the dilation of K with respect to the origin, namely the n-simplex with one vertex at $\left(0, \dots, 0, \frac{1-\epsilon}{2\rho}\right)$ and base $(1-\epsilon)\Sigma$. We note that $E_j = K_j$ when $n = 2$, while the shape of E_j is much more involved for $n \geq 3$ (see Figure 1).

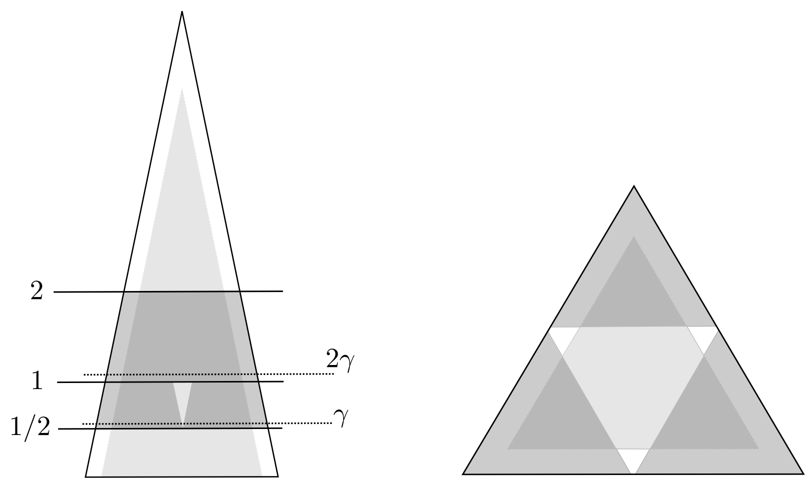


FIGURE 1. On the left, we consider $n = 2$. The overlapping shaded regions are $(1-\epsilon)K$, K_0, $\frac{1}{2}(\hat{x}_1 + K_0)$, and $\frac{1}{2}(\hat{x}_2 + K_0)$. Note that their union covers $K_1 = K \cap \{\gamma \leq x_2 \leq 2\gamma\}$; thus $E_1 = K_1$. On the right, we consider the horizontal slice $K \cap \{x_3 = \gamma\}$ for $n = 3$. The overlapping shaded regions are the intersection of this slice with $(1-\epsilon)K$, $\frac{1}{2}(\hat{x}_1 + K_0)$, $\frac{1}{2}(\hat{x}_2 + K_0)$, and $\frac{1}{2}(\hat{x}_3 + K_0)$. Note that this set does not cover $K \cap \{x_3 = \gamma\}$, hence $E_1 \subsetneq K_1$. For $n \geq 3$ the sets E_j have a fractal structure, described in detail in the proof of Lemma 2.4 for $n = 3$.

Set $E := \bigcup_{j \geq 0} E_j$. We claim that there exists a dimensional constant $C_0 \geq 1$ such that

$$|E \setminus A| \leq C_0\, \delta(A). \tag{2.15}$$

The proof of this claim will be carried out in Steps 2–5 below.

Before proceeding with the second step, we explain some geometric features of the core of the proof in Step 4. If most of K_j belongs to A, then the fact that $\hat{x}_k$ belongs to A implies that most of $\frac{1}{2}(\hat{x}_k + K_j)$, $k = 1, \dots, n$, belongs to A. This relatively easy step is carried out in 4(e) below. On the other hand, these n regions do not cover $K_{j+1} \setminus K_j$. As the picture on the left in Figure 1 shows, even in dimension $n = 2$, the two regions miss a narrow inverted trapezoid in the next layer, K_{j+1}. When $n = 3$, the right-hand picture in Figure 1 of $K \cap \{x_n = \gamma\}$ (the lowest horizontal slice of K_1) shows that the 3 regions cover the three equilateral triangles at the corners of the large triangle. What is missing is a hexagon inside, with very short sides coinciding with the sides of the large triangle. For higher slices of K_1, above the level $x_3 = \gamma$, the missing portion is an even larger hexagon, just as the missing portion for $n = 2$ gets larger as x_n increases.

To fill more of K_{j+1} with elements of A, we consider horizontal slices that are at a lower level than K_j by a factor $2^{-m} \approx \epsilon^2$. In Steps 4(a) and 4(b), under the assumptions that (at the appropriate inductive stage) the volume fraction of A is suitably large (see (2.22)) and the volume fraction of $[(A+A)/2]\setminus A$ is suitably small (see (2.33)), we show that there is a "large" horizontal slice $A_t = A \cap \{x_n = t\}$ at a height $t \approx \epsilon^2\gamma^j$. In Steps 4(c) and 4(d), we then use the semisum between K_j and the slice A_t to show that most of $(K_{j+1}\setminus K_j)\cap(1-\epsilon)K$ belongs to A. This is why we are able to include $K_j \cap (1-\epsilon)K$ in the definition of E_j, which is crucial to conclude the proof in dimension 3. The fact that t can be chosen small enough relative to γ^j is what makes it possible to obtain the inequality $\gamma^{j+1} > (\gamma^j + t)/2$ used in Steps 4(c) and 4(d) (see (2.39)). Such an inequality is essential in our proof, and this is what requires us to slow the rate of descent towards the base Σ from 2^{-j} to γ^j for some $\gamma = \frac{1}{2} + \epsilon > \frac{1}{2}$.

As a consequence of $\gamma > 1/2$, there is an overlap between K_j and K_{j+1}. This overlapping gives rise to an extra term σ_j (see (2.17)) in the bound (2.19), which will then be controlled in the second part of Step 5.

Step 2: Setting the notation. – Define the numbers

$$\nu_j := |E_j \setminus A|, \qquad \delta_j := \Big|\Big(\Big[\tfrac{1}{2}(A+A)\Big]\setminus A\Big)\cap K_j\Big|, \tag{2.16}$$

and

$$\sigma_j := \big|\big[K_j \cap K_{j+1} \cap (1-\epsilon)K\big]\setminus A\big|. \tag{2.17}$$

Note that

$$|K_j| \geq |E_j| \geq |\{\gamma^j \leq x_n \leq 2\gamma^j\} \cap (1-\epsilon)K| = (1-\epsilon)^{n-1}|K_j|. \tag{2.18}$$

We claim that there exist dimensional constants $M, N \geq 1$, with N integer, such that

$$\nu_{j+1} \leq \frac{8}{9}\nu_j + \sigma_j + M\sum_{i=0}^{N}\delta_{j+i} \qquad \forall\, j \geq 0. \tag{2.19}$$

The proof of (2.19) will be split over Step 3 and Steps 4(a)-4(e) below.

Step 3: The case $\nu_j \approx |E_j|$. – Consider first the case

$$\nu_j \geq \frac{2}{3}|E_j|. \tag{2.20}$$

Note that, for $\rho \ll 1$, the sets K_j are almost vertical cylinders of height γ^j, and more precisely (recalling that $\mathscr{H}^{n-1}(\Sigma) = \mathscr{H}^{n-1}(K \cap \{x_n = 0\}) = 1$)

$$\gamma^j \geq |K_j| \geq (1 - C\rho)\gamma^j, \tag{2.21}$$

where $C > 0$ is a dimensional constant. This implies that $|K_{j+1}| = \big(1 + O(\rho)\big)\gamma|K_j|$, so it follows by (2.18) that

$$\nu_j \geq \frac{2}{3}|E_j| \geq \frac{2}{3}(1-\epsilon)^{n-1}|K_j| = \frac{2(1-\epsilon)^{n-1}}{3\gamma\big(1+O(\rho)\big)}|K_{j+1}| \geq \frac{2(1-\epsilon)^{n-1}}{3\gamma\big(1+O(\rho)\big)}\nu_{j+1},$$

which proves (2.19) because $\frac{3\gamma}{2(1-\epsilon)^{n-1}}\big(1+O(\rho)\big) \leq \frac{8}{9}$ provided ϵ and ρ are sufficiently small (recall that $\gamma = \frac{1}{2}+\epsilon$).

Step 4: The case ν_j not too large. – We now consider the case

$$\nu_j \leq \frac{2}{3}|E_j|. \tag{2.22}$$

Step 4(a): Finding some nontrivial fraction of A near the vertices. – Using (2.18), it follows that

$$|A \cap K_j| \geq |A \cap E_j| = |E_j| - \nu_j \geq \frac{1}{3}|E_j| \geq \frac{1}{3}(1-\epsilon)^{n-1}|K_j| \geq \frac{1}{4}|K_j|. \tag{2.23}$$

Now, for any $k = 1, \dots, n$, consider the sets (see Figure 2)

$$A^k_{j,\ell} := (1-2^{-\ell})\hat{x}_k + 2^{-\ell}(A \cap K_j) \qquad \forall\, \ell \geq 0, \tag{2.24}$$

and note that, because of (2.23),

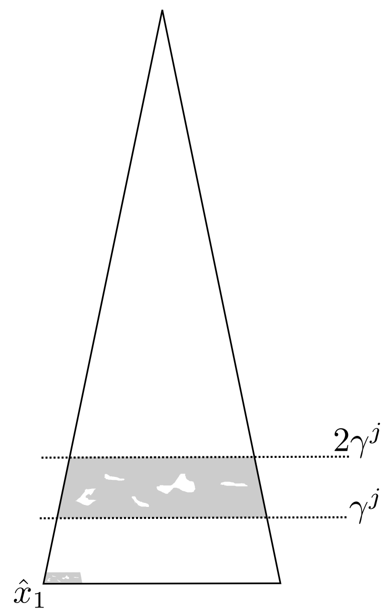


FIGURE 2. The larger grey area represents $A \cap K_j$. Small in the corner is the set $A^1_{j,\ell}$ for some $\ell \gg 1$.

$$|A^k_{j,\ell}| = 2^{-n\ell}|A \cap K_j| \geq 2^{-n\ell-2}|K_j|. \tag{2.25}$$

Our goal is to show that, provided the numbers δ_{j+i} are small enough for sufficiently many indices i, then $A^k_{j,\ell} \cap A$ has almost the same measure as $A^k_{j,\ell}$. To prove this, for convenience we define the auxiliary numbers

$$\delta^k_{j,\ell} := \Big|\Big(\big[\tfrac{1}{2}(A+A)\big] \setminus A\Big) \cap A^k_{j,\ell}\Big| \qquad \forall\, \ell \geq 0.$$

Also, we iteratively define

$$B^k_{j,0} := A \cap K_j, \qquad B^k_{j,\ell+1} := \tfrac{1}{2}\big(\hat{x}_k + (A \cap B^k_{j,\ell})\big).$$

Since $\hat{x}_k \in A$, one can easily see by induction on ℓ that the following inclusion holds:

$$\tfrac{1}{2}\big(\hat{x}_k + (A \cap B^k_{j,\ell})\big) \subset \big[\tfrac{1}{2}(A+A)\big] \cap A^k_{j,\ell+1} \qquad \forall\, \ell \geq 0.$$

Therefore

$$|A \cap B^k_{j,\ell+1}| \geq \Big|\tfrac{1}{2}\big(\hat{x}_k + A \cap B^k_{j,\ell}\big)\Big| - \delta^k_{j,\ell+1} = \frac{1}{2^n}|A \cap B^k_{j,\ell}| - \delta^k_{j,\ell+1} \qquad \forall\, \ell \geq 0.$$

Since $|A \cap B^k_{j,0}| = |A \cap K_j| = 2^{n\ell}|A^k_{j,\ell}|$ and $A \cap B^k_{j,\ell} \subset A^k_{j,\ell}$, we deduce that

$$|A \cap A^k_{j,\ell}| \geq |A \cap B^k_{j,\ell}| \geq 2^{-n\ell}|A \cap K_j| - \sum_{r=1}^{\ell} \delta^k_{j,r} = |A^k_{j,\ell}| - \sum_{r=1}^{\ell} \delta^k_{j,r} \qquad \forall\, \ell \geq 1. \tag{2.26}$$

We now start to fix some parameters. Choose an integer m such that

$$\epsilon^2 \leq 2^{-m} \leq 2\epsilon^2, \tag{2.27}$$

and then choose N large enough so that

$$2^{-(m+1)} \leq \gamma^N \leq 2^{-m}. \tag{2.28}$$

With these definitions, it follows that $\cup_{r=1}^m A^k_{j,r} \subset \cup_{i=0}^N K_{j+i}$. Therefore, since the sets $\{A^k_{j,r}\}_{1\leq r\leq m}$ are disjoint, it follows that

$$\sum_{r=1}^{m} \delta^k_{j,r} \leq \sum_{i=1}^{N} \delta_{j+i}.$$

Hence, by (2.26) applied with $\ell = m$, we get

$$|A \cap A^k_{j,m}| \geq |A^k_{j,m}| - \sum_{i=1}^{N} \delta_{j+i}. \tag{2.29}$$

We are now ready to prove (2.19). Consider first the case in which

$$\sum_{i=1}^{N} \delta_{j+i} \geq \epsilon |A^k_{j,m}|.$$

Then, since $\nu_{j+1} \leq |K_{j+1}| \leq |K_j|$ for ρ small enough (see (2.21)), recalling (2.25) and that $\gamma^{-N} \leq 2^{-m}$ (see (2.28)), we deduce that

$$\sum_{i=1}^{N} \delta_{j+i} \geq \frac{\epsilon}{4}\gamma^{nN}|K_j| \geq \frac{\epsilon}{4}\gamma^{nN}\nu_{j+1},$$

so (2.19) follows immediately with $M = 4\gamma^{-nN}\epsilon^{-1}$.

Next, we must consider the case in which $\sum_{i=1}^N \delta_{j+i} \leq \epsilon |A^k_{j,m}|$. In that case, (2.29) gives

$$|A \cap A^k_{j,m}| \geq (1-\epsilon)|A^k_{j,m}| \qquad \forall\, k = 1, \ldots, n. \tag{2.30}$$

In other words, we proved that A covers almost all the sets $\{A^k_{j,m}\}_{k=1}^n$, which are small rescaled copies of $A \cap K_j$ that live in a ϵ^2 neighborhood of the n vertices $\hat{x}_k$ (recall (2.27)).

Note that whereas the sets $A^k_{j,m}$ for different k are translates of each other, the sets $A \cap A^k_{j,m}$ are not. To enforce this additional property, we first translate them to the same point, intersect them, and then move them back. More precisely, recalling (2.24), we set

$$\hat{A}_{j,m} := \bigcap_{k=1}^{n} \big((A \cap A^k_{j,m}) - (1-2^{-m})\hat{x}_k\big), \qquad \hat{A}^k_{j,m} := \hat{A}_{j,m} + (1-2^{-m})\hat{x}_k.$$

Now, thanks to (2.30),

$$\hat{A}^k_{j,m} \subset A \cap A^k_{j,m}, \qquad |\hat{A}^k_{j,m}| \geq (1-n\epsilon)|A^k_{j,m}| \qquad \forall\, k = 1, \dots, n, \tag{2.31}$$

and $\hat{A}^k_{j,m}$ and $\hat{A}^{k'}_{j,m}$ are the same set for any $k, k' \in \{1, \dots, n\}$, up to a translation orthogonal to the x_n axis. Also, it follows by (2.31), (2.25), and (2.21), that

$$\begin{aligned} &|\hat{A}^k_{j,m}| \leq |A^k_{j,m}| = 2^{-nm}|A \cap K_j| \leq 2^{-nm}|K_j| \leq 2^{-nm}\gamma^j, \\ &|\hat{A}^k_{j,m}| \geq (1-n\epsilon)|A^k_{j,m}| \geq (1-n\epsilon)2^{-nm-2}|K_j| \geq 2^{-nm-3}\gamma^j, \end{aligned} \tag{2.32}$$

provided ϵ and ρ are sufficiently small.

Step 4(b): Finding an almost full slice in A at $\{x_n = t\}$ for some $t \approx 2^{-m}\gamma^j$ using Fubini and induction.– We look at the slab

$$S_{j,m} := K \cap \{2^{-m}\gamma^j \leq x_n \leq 2^{-m+1}\gamma^j\},$$

and define $\delta_{j,m} := \big|\big(\big[\tfrac{1}{2}(A+A)\big] \setminus A\big) \cap S_{j,m}\big|$. Note that $A^k_{j,m} \subset S_{j,m}$ for any $k = 1, \dots, n$.

Recall that d_{n-1} is the dimensional constant corresponding to Theorem 1.4 in dimension $n-1$. It will suffice to prove the existence of a suitable slice inside $S_{j,m}$ assuming

$$\delta_{j,m} \leq \epsilon^{2n+8} d_{n-1} |\hat{A}^k_{j,m}| \tag{2.33}$$

(note that $|\hat{A}^k_{j,m}|$ is independent of k). Indeed, since $S_{j,k} \subset K_{j+N-1} \cup K_{j+N}$ it holds

$$\delta_{j,m} \leq \delta_{j+N-1} + \delta_{j+N}. \tag{2.34}$$

Hence, if (2.33) fails then (recall (2.32) and (2.28))

$$\delta_{j+N-1} + \delta_{j+N} \geq \epsilon^{2n+8} d_{n-1}|\hat{A}^k_{j,m}| \geq \epsilon^{2n+8} d_{n-1}(1-n\epsilon)2^{-nm-2}|K_j| \geq \frac{\epsilon^{2n+8} d_{n-1}\gamma^{nN}}{8}\nu_{j+1},$$

which proves (2.19) with $M = 8\gamma^{-nN}\epsilon^{-2n-8}d_{n-1}^{-1}$.

Now we can proceed under the additional assumption (2.33). Define

$$A_t := A \cap \{x_n = t\} \supset \big(\cup_{k=1}^n (\hat{A}^k_{j,m})\big) \cap \{x_n = t\} =: \hat{A}_t,$$

and consider $\delta(A_t) = \mathscr{H}^{n-1}\big(\tfrac{1}{2}(A_t + A_t) \setminus A_t\big)$. Since $\hat{A}^k_{j,m} \subset A$, it follows by (2.33) and (2.32) that

$$\int_{2^{-m}\gamma^j}^{2^{-m+1}\gamma^j} \delta(A_t)\, dt \leq \delta_{j,m} \leq \epsilon^{2n+8} d_{n-1}|\hat{A}^k_{j,m}| \leq \epsilon^{2n+8} d_{n-1} 2^{-nm}\gamma^j. \tag{2.35}$$

Also, recalling (2.32), it follows that

$$\frac{1}{2^{-m}\gamma^j}\int_{2^{-m}\gamma^j}^{2^{-m+1}\gamma^j} \mathscr{H}^{n-1}(\hat{A}_t)\, dt \geq \frac{1}{2^{-m}\gamma^j}\sum_{k=1}^{n} |\hat{A}^k_{j,m}| \geq n2^{-(n+1)m-3}.$$

Hence, since $\mathcal{H}^{n-1}(\hat{A}_t) \leq \mathcal{H}^{n-1}(A_t) \leq 1$, there exists a set $J \subset [2^{-m}\gamma^j, 2^{-m+1}\gamma^j]$ such that [(3)]

$$\mathcal{H}^1(J) \geq n2^{-(n+2)m-4}\gamma^j, \quad \text{with} \quad \mathcal{H}^{n-1}(\hat{A}_t) \geq n2^{-(n+1)m-4} \quad \forall\, t \in J.$$

Combining this estimate with (2.35), we deduce that

$$\begin{aligned}\frac{1}{\mathcal{H}^1(J)}\int_J \delta(A_t) &\leq \frac{\delta_{j,m}}{n2^{-(n+2)m-4}\gamma^j} \leq \frac{\epsilon^{2n+8}d_{n-1}2^{-nm}\gamma^j}{n2^{-(n+2)m-4}\gamma^j}\\ &\leq \frac{\epsilon^{2n+8}d_{n-1}2^{(n+3)m+8}}{n^2}\,\mathcal{H}^{n-1}(\hat{A}_t)\\ &\leq \frac{\epsilon^{2n+8}d_{n-1}2^{(n+3)m+8}}{n^2}\,\mathcal{H}^{n-1}(A_t) \qquad \forall\, t \in J.\end{aligned}$$

Recalling (2.27), this proves that

$$\frac{1}{\mathcal{H}^1(J)}\int_J \delta(A_t) \leq \frac{2^{n+6}}{n\epsilon^{n+2}\gamma^j}\delta_{j,m} \leq \frac{2^{n+11}}{n^2}\epsilon^2 d_{n-1}\,\mathcal{H}^{n-1}(\hat{A}_t) \qquad \forall\, t \in J.$$

In particular, choosing ϵ sufficiently small, by the Mean Value Theorem we can find $t \in [2^{-m}\gamma^j, 2^{1-m}\gamma^j]$ such that

$$\delta(A_t) \leq \frac{2^{n+6}}{n\epsilon^{n+2}\gamma^j}\delta_{j,m} \leq \epsilon^{3/2}d_{n-1}\,\mathcal{H}^{n-1}(\hat{A}_t), \qquad \mathcal{H}^{n-1}(\hat{A}_t) > 0.$$

Hence, since $\mathcal{H}^{n-1}(\hat{A}_t) \leq \mathcal{H}^{n-1}(A_t)$, we can apply Theorem 1.4 to A_t and we deduce that

$$\mathcal{H}^{n-1}(\mathrm{co}(A_t) \setminus A_t) \leq C_{n-1}\delta(A_t) \leq C_{n-1}\frac{2^{n+6}}{n\epsilon^{n+2}\gamma^j}\delta_{j,m} \leq C_{n-1}\epsilon^{3/2}d_{n-1}. \tag{2.36}$$

Also, because $\mathcal{H}^{n-1}(\hat{A}_t) > 0$, it follows that $\mathrm{co}(A_t)$ contains at least one point in $\hat{A}^k_{j,m} \cap \{x_n = t\}$ for any $k = 1, \ldots, n$. Recalling that $\hat{A}^k_{j,m} \subset (1-2^{-m})\hat{x}_k + 2^{-m}K_j$ and that $2^{-m} \leq 2\epsilon^2$ (see (2.27)), it follows that $\mathrm{co}(A_t)$ contains n points $\{\hat{x}^k_t\}_{k=1}^n$ such that $|\hat{x}^k_t - \hat{x}_k| \leq C\epsilon^2$, thus

$$\mathrm{co}(A_t) \supset \big((1-\epsilon)K\big) \cap \{x_n = t\}. \tag{2.37}$$

In the next steps we use the slice A_t and a semisum to control a large fraction of ν_{j+1}. Because the argument in dimension $n = 2$ is much easier than in higher dimensions, for convenience of the reader we first treat this case.

(3) This estimate follows from a general simple fact: if $f : I \subset \mathbb{R} \to [0,1]$ satisfies $\frac{1}{\mathcal{H}^1(I)}\int_I f(t)\,dt \geq \eta > 0$, then there exists $J \subset I$ such that

$$\mathcal{H}^1(J) \geq \frac{\eta}{2}\mathcal{H}^1(I) \qquad \text{and} \qquad f(t) \geq \frac{\eta}{2} \quad \forall\, t \in J.$$

Indeed, if this was false, we would have that $f \leq \eta/2$ on a set $I' \subset I$ of measure larger than $(1-\eta/2)\,\mathcal{H}^1(I)$, therefore (recall that $0 \leq f \leq 1$)

$$\begin{aligned}\int_I f(t)\,dt \leq \int_{I'} f(t)\,dt + \int_{I\setminus I'} f(t)\,dt &\leq \mathcal{H}^1(I')\frac{\eta}{2} + \mathcal{H}^1(I \setminus I')\\ &\leq \Big(1-\frac{\eta}{2}\Big)\frac{\eta}{2}\mathcal{H}^1(I) + \frac{\eta}{2}\mathcal{H}^1(I) < \eta\,\mathcal{H}^1(I),\end{aligned}$$

a contradiction.

Step 4(c): Use the slice from Step 4(b) and a semisum to control a large fraction of ν_{j+1}: the case $n = 2$. – Thanks to (2.36) and (2.37), there exists a point $z = (z_1, t) \in A \subset \mathbb{R}^2$ with $|z_1| \le C\epsilon^{3/4}$ and $t \in [2^{-m}\gamma^j, 2^{-m+1}\gamma^j]$. In particular, recalling that $t \approx \epsilon^2\gamma^j$ and that $\gamma = \frac{1}{2} + \epsilon$, we have, for ϵ sufficiently small,

$$\big(\tfrac{1}{4}K\big) \cap (K_{j+1} \setminus K_j) \subset \big(\tfrac{1}{4}K\big) \cap \left\{\tfrac{\gamma^j+t}{2} \le x_2 \le \tfrac{2\gamma^j+t}{2}\right\} \subset \tfrac{1}{2}(z + K_j) \subset K_{j+1} \cup K_{j+2}$$

where $\frac{1}{4}K$ denotes the dilation of K by a factor $\frac{1}{4}$ with respect to the origin. Finally, since $K_j = E_j$ for $n = 2$, the definition of ν_j and δ_j (see (2.16)) yields

$$\begin{aligned}\big|\big[\big(\tfrac{1}{4}K\big) \cap (K_{j+1} \setminus K_j)\big] \setminus A\big| &\le \big|\tfrac{1}{2}\big(z + (K_j \setminus A)\big)\big| + \delta_{j+1} + \delta_{j+2}\\ &\le \frac{1}{4}\nu_j + \delta_{j+1} + \delta_{j+2}.\end{aligned} \tag{2.38}$$

Step 4(d): Use the slice from Step 4(b) and a semisum to control a large fraction of ν_{j+1}: the case $n \ge 3$. – Given $s \ge 0$, define $K_{s,\epsilon} := \big((1-\epsilon)K\big) \cap \{x_n = s\}$ and $A_{s,\epsilon} := A \cap K_{s,\varepsilon}$. Then

$$\tfrac{1}{2}\big(A_{s,\epsilon} + A_{t,\epsilon}\big) \setminus A_{\frac{s+t}{2},\epsilon} \subset \big[\tfrac{1}{2}(A + A) \setminus A\big] \cap \big((1-\epsilon)K\big) \cap \big\{x_n = \tfrac{s+t}{2}\big\}.$$

Using the inclusion above for $s \in [\gamma^j, 2\gamma^j]$, and noticing that for ϵ small

$$K_{j+1} \setminus K_j \subset K \cap \left\{\tfrac{\gamma^j+t}{2} \le x_n \le \tfrac{2\gamma^j+t}{2}\right\} \subset K_{j+1} \cup K_{j+2}, \tag{2.39}$$

we get

$$\begin{aligned}\big|\big[(1-\epsilon)K \cap (K_{j+1} \setminus K_j)\big] \setminus A\big| &\le \left|(1-\epsilon)K \cap \left\{\tfrac{\gamma^j+t}{2} \le x_n \le \tfrac{2\gamma^j+t}{2}\right\} \setminus A\right|\\
&\le \left|(1-\epsilon)K \cap \left\{\tfrac{\gamma^j+t}{2} \le x_n \le \tfrac{2\gamma^j+t}{2}\right\} \setminus \tfrac{1}{2}(A + A)\right|\\
&\quad + \left|\big(\big[\tfrac{1}{2}(A + A)\big] \setminus A\big) \cap \left\{\tfrac{\gamma^j+t}{2} \le x_n \le \tfrac{2\gamma^j+t}{2}\right\}\right|\\
&\le \int_{\frac{\gamma^j+t}{2}}^{\frac{2\gamma^j+t}{2}} \mathscr{H}^{n-1}\big(K_{\tau,\epsilon} \setminus \tfrac{1}{2}(A + A)\big)\, d\tau + \delta_{j+1} + \delta_{j+2}\\
&= \frac{1}{2}\int_{\gamma^j}^{2\gamma^j} \mathscr{H}^{n-1}\big(K_{\frac{s+t}{2},\epsilon} \setminus \tfrac{1}{2}(A + A)\big)\, ds + \delta_{j+1} + \delta_{j+2}\\
&\le \frac{1}{2}\int_{\gamma^j}^{2\gamma^j} \mathscr{H}^{n-1}\big(K_{\frac{s+t}{2},\epsilon} \setminus \tfrac{1}{2}(A_{s,\epsilon} + A_{t,\epsilon})\big)\, ds + \delta_{j+1} + \delta_{j+2}.\end{aligned}$$

To estimate the last integral, we define the "vertical semisum with slope ρ" of two sets F_s and F_t contained respectively in two levels $\{x_n = s\}$ and $\{x_n = t\}$ with $0 \le s, t \le 1$ by

$$\tfrac{1}{2}(F_s +_{v,\rho} F_t) := \big\{\tfrac{1}{2}(z + w, s + t) : (z, s) \in F_s,\ (w, t) \in F_t,\ (1 - 2\rho s)w = (1 - 2\rho t)z\big\}.$$

Note that if $\rho = 0$ this is just a trivial one-dimensional semisum in the vertical variable (since in that case $z = w$), namely

$$\tfrac{1}{2}(F_s +_{v,0} F_t) := \big\{\tfrac{1}{2}(z, s + t) : (z, s) \in F_s,\ (z, t) \in F_t\big\},$$

and it is clear that

$$\mathscr{H}^{n-1}\big(\tfrac{1}{2}(F_s +_{v,0} F_t)\big) \le \min\big\{\mathscr{H}^{n-1}(F_s), \mathscr{H}^{n-1}(F_t)\big\} \le \frac{1}{2}\Big(\mathscr{H}^{n-1}(F_s) + \mathscr{H}^{n-1}(F_t)\Big). \tag{2.40}$$

In our case, since K is not quite a vertical cylinder but instead has a small angle 2ρ, we are asking that the points (z,s) and (w,t) be collinear with the vertex $\big(0,\frac{1}{2\rho}\big)$ of K, and the analogue of (2.40) for sets $F_s, F_t \subset K$ becomes

$$\mathscr{H}^{n-1}\big(\tfrac{1}{2}(F_s +_{v,\rho} F_t)\big) \le \frac{1+O(\rho)}{2}\Big(\mathscr{H}^{n-1}(F_s)+\mathscr{H}^{n-1}(F_t)\Big). \tag{2.41}$$

Hence, since $K_{\frac{s+t}{2},\epsilon} = \frac{1}{2}(K_{s,\epsilon} +_{v,\rho} K_{t,\epsilon})$, one can easily check that

$$\mathscr{H}^{n-1}\big(K_{\frac{s+t}{2},\epsilon} \setminus \tfrac{1}{2}(A_{s,\epsilon} +_{v,\rho} A_{t,\epsilon})\big) \le \frac{1+O(\rho)}{2}\Big(\mathscr{H}^{n-1}(K_{s,\epsilon}\setminus A_{s,\epsilon})+\mathscr{H}^{n-1}(K_{t,\epsilon}\setminus A_{t,\epsilon})\Big).$$

Also, we observe that

$$\tfrac{1}{2}(A_{s,\epsilon}+A_{t,\epsilon}) \supset \tfrac{1}{2}(A_{s,\epsilon} +_{v,\rho} A_{t,\epsilon}).$$

Combining together all these bounds, and recalling (2.36), (2.37), and (2.34), we get

$$\begin{aligned}
\big|\big[(1-\epsilon)K\cap(K_{j+1}\setminus K_j)\big]\setminus A\big| &\le \frac{1+O(\rho)}{2}\int_{\gamma^j}^{2\gamma^j}\mathscr{H}^{n-1}\big(K_{s,\epsilon}\setminus A_{s,\epsilon}\big)\,ds\\
&\quad+\frac{1+O(\rho)}{2}\int_{\gamma^j}^{2\gamma^j}\mathscr{H}^{n-1}\big(K_{t,\epsilon}\setminus A_{t,\epsilon}\big)\,ds+\delta_{j+1}+\delta_{j+2}\\
&\le \frac{1+O(\rho)}{2}\big|\big((1-\epsilon)K\big)\cap K_j\big]\setminus A\big|\\
&\quad+\frac{1+O(\rho)}{2}\int_{\gamma^j}^{2\gamma^j}C_{n-1}\frac{2^{n+6}}{n\epsilon^{n+2}\gamma^j}\delta_{j,m}\,ds+\delta_{j+1}+\delta_{j+2}\\
&\le \frac{1+O(\rho)}{2}\big|\big((1-\epsilon)K\big)\cap K_j\big]\setminus A\big|\\
&\quad+C_{n-1}\frac{2^{n+6}}{n\epsilon^{n+2}}(\delta_{j+N-1}+\delta_{j+N})+\delta_{j+1}+\delta_{j+2}.
\end{aligned}$$

Recalling the definitions of E_j, ν_j, and σ_j (see (2.14), (2.16), and (2.17)), this proves that

$$\big|\big[(1-\epsilon)K\cap K_{j+1}\big]\setminus A\big| \le \frac{1+O(\rho)}{2}\nu_j+\sigma_j+C_{n-1}\frac{2^{n+6}}{n\epsilon^{n+2}}(\delta_{j+N-1}+\delta_{j+N})+\delta_{j+1}+\delta_{j+2}. \tag{2.42}$$

Step 4(e): Use a semisum to control the remaining fraction of ν_{j+1}. – Since $\hat{x}_k \in A$, we see that

$$\Big(\bigcup_{k=1}^{n}\tfrac{1}{2}(\hat{x}_k+E_j)\Big)\setminus\tfrac{1}{2}(A+A)\subset\Big(\bigcup_{k=1}^{n}\tfrac{1}{2}(\hat{x}_k+(E_j\setminus A))\Big).$$

Therefore, since $\cup_{k=1}^{n}\frac{1}{2}(\hat{x}_k+E_j)\subset K_{j+1}$, recalling the definition of ν_j and δ_j (see (2.16)) we get

$$\left|\Big(\bigcup_{k=1}^{n}\tfrac{1}{2}(\hat{x}_k+E_j)\Big)\setminus A\right| \le \frac{n}{2^n}|E_j\setminus A|+\delta_{j+1}=\frac{n}{2^n}\nu_j+\delta_{j+1}. \tag{2.43}$$

Note that for $n=2$ we have

$$E_j=K_j \qquad\text{and}\qquad \Big(\bigcup_{k=1}^{2}\tfrac{1}{2}(\hat{x}_k+K_j)\Big)\cup\big[\big(\tfrac{1}{4}K\big)\cap K_{j+1}\big]\supset K_{j+1},$$

while $n2^{-n} \leq 3/8$ for $n \geq 3$. Hence, combining (2.43) with (2.38) and (2.42), for any $n \geq 2$ we obtain

$$\nu_{j+1} \leq \left(\frac{1+O(\rho)}{2} + \frac{3}{8}\right)\nu_j + \sigma_j + M\sum_{i=0}^{N} \delta_{j+i}$$

for some dimensional constant M, concluding the proof of (2.19).

Step 5: Proof of (2.15). – Since $\nu_0 = 0$ (because $K_0 \subset A$ by assumption), by summing (2.19) with respect to j we obtain

$$\sum_{j\geq 0} \nu_j \leq \frac{8}{9}\Big(\sum_{j\geq 0} \nu_j\Big) + \sum_{j\geq 0} \sigma_j + M\sum_{j\geq 1}\sum_{i=0}^{N} \delta_{j+i}.$$

Moreover, the last term can be bounded by

$$MN\sum_{j\geq 0} \delta_j = MN\sum_{j\geq 0} \delta_{2j} + MN\sum_{j\geq 0} \delta_{2j+1}.$$

Noticing that the sets $\{K_{2j}\}_{j\geq 0}$ and the sets $\{K_{2j+1}\}_{j\geq 0}$ are disjoint, it follows that

$$\sum_{j\geq 0} \delta_{2j} \leq \delta(A), \qquad \sum_{j\geq 0} \delta_{2j+1} \leq \delta(A).$$

Hence, combining these estimates together, we proved that

$$\frac{1}{9}\Big(\sum_{j\geq 0} \nu_j\Big) \leq \sum_{j\geq 0} \sigma_j + 2MN\,\delta(A).$$

Since $\sum_{j\geq 0} \nu_j \geq |E \setminus A|$, we get

$$\frac{1}{9}|E \setminus A| \leq \sum_{j\geq 0} \sigma_j + 2MN\,\delta(A).$$

Note that this would prove (2.15) if we did not have the additional term $\sum_{j\geq 0} \sigma_j$. The idea to get rid of this additional term is the following: since the volume of $K_j \cap K_{j+1}$ is only a fraction ϵ of the volume of K_j and K_{j+1}, if A were uniformly distributed inside the sets K_j, then we would have

$$\sigma_j \leq C\epsilon(\nu_j + \nu_{j+1}),$$

from which we would conclude easily. Although A need not be uniformly distributed, we can prove analogous inequalities starting our iteration at many levels, and then add them up so that the average overlap of A with $K_j \cap K_{j+1}$ is sufficiently uniform.

Thus, to handle the terms σ_j, we take $\tau \in [\gamma, 1]$ and define the sets

$$K_j^\tau := K \cap \{\tau\gamma^j \leq x_n \leq 2\tau\gamma^j\},$$

$$E_0^\tau := K_0^\tau,$$

$$E_{j+1}^\tau := \Big(\bigcup_{k=1}^{n} \tfrac{1}{2}(\hat{x}_k + E_j^\tau)\Big) \cup \big((1-\epsilon)K \cap \{-2\tau\gamma^j \leq x_n \leq -\tau\gamma^j\}\big),$$

$$E^\tau := \cup_{j\geq 0} E_j^\tau,$$

and the numbers

$$\nu_j^\tau := |E_j^\tau \setminus A|,$$

$$\delta_j^\tau := \Big|\Big(\big[\tfrac{1}{2}(A+A)\big] \setminus A\Big) \cap K_j^\tau\Big|,$$
$$\sigma_j^\tau := \big|\big[(1-\epsilon)K \cap K_j^\tau \cap K_{j+1}^\tau\big] \setminus A\big|.$$

Now, if we repeat the very same proof as above with these new sets, we obtain

$$\frac{1}{9}|E^\tau \setminus A| \le \sum_{j\ge 0} \sigma_j^\tau + 2MN\,\delta(A)$$

(note that we still have $K_0^\tau \subset A$, therefore $\nu_j^\tau = 0$). Noticing that $E = E^1 \subset E^\tau$ for all $\tau \in (\gamma, 1)$ (in other words, the sets E^τ are monotonically decreasing in τ), this proves that

$$\frac{1}{9}|E \setminus A| \le \sum_{j\ge 0} \sigma_j^\tau + 2MN\,\delta(A). \tag{2.44}$$

We now observe that, since $\gamma = \frac{1}{2} + \epsilon$,

$$K_j^\tau \cap K_{j+1}^\tau = K \cap \{\tau\gamma^{j+1} \le x_n \le 2\tau\gamma^j\} = K \cap \{\tau\gamma^j \le x_n \le (1+2\epsilon)\tau\gamma^j\},$$

hence the sets

$$\Big\{K_j^{\tau_m} \cap K_{j+1}^{\tau_m} \,:\, j \ge 0,\ \tau_m = 1 - 2m\epsilon,\ m = 0, \dots, \lfloor \tfrac{\gamma}{4\epsilon} \rfloor\Big\}$$

are disjoint. This implies that

$$\sum_{m=0}^{\lfloor \frac{\gamma}{4\epsilon} \rfloor} \sum_{j\ge 0} \sigma_j^{\tau_m} \le |E \setminus A|,$$

that combined with (2.44) gives

$$\lfloor \tfrac{\gamma}{4\epsilon} \rfloor\, |E \setminus A| \le 9 \sum_{m=0}^{\lfloor \frac{\gamma}{4\epsilon} \rfloor} \sum_{j\ge 0} \sigma_j^{\tau_m} + 18\cdot MN \lfloor \tfrac{\gamma}{4\epsilon} \rfloor\, \delta(A) \le 9|E \setminus A| + 18\cdot MN \lfloor \tfrac{\gamma}{4\epsilon} \rfloor\, \delta(A).$$

Choosing ϵ sufficiently small that $\lfloor \frac{\gamma}{4\epsilon} \rfloor \ge 10$ proves (2.15).

Step 6: Getting control of A on all of K. – Note that (2.15) provides control on the measure of A inside E. In particular, since $E_j = K_j$ when $n = 2$, this already proves Lemma 2.1 (and therefore Theorem 1.4) in the case $n = 2$. Thus for the remainder of the proof we may assume $n = 3$. In this case, the goal is to enlarge the set E on which we control the measure of A to all of K.

For $0 \le t < 1/2\rho$, set

$$\Sigma(t) = K \cap \{x_3 = t\}.$$

By hypothesis, $\Sigma(t) \cap A = \Sigma(t)$ for $t \ge 1/2$. Our approach to estimating $\Sigma(t) \setminus A$ for $0 \le t < 1/2$ will be to intersect $\Sigma(t) \setminus E$ with segments parallel to sides of the triangle $\Sigma(t)$ near the boundary, and show that these missing parts are sufficiently small and atomized that we can apply the following one-dimensional lemma.

LEMMA 2.3. – *Let $J \subset \mathbb{R}$ be an interval. Suppose that $A \subset J$ and $E \subset J$, and*

$$\chi_{E/2} * \chi_{E/2}(x) \ge \frac{1}{10}\operatorname{dist}(x, \partial J) \quad \textit{for all } x \in J. \tag{2.45}$$

Then

$$|J \setminus A| \le |\tfrac{1}{2}(A+A) \setminus A| + 20|E \setminus A|.$$

Proof. – The proof of (2.3) applies with K replaced by E and shows that

$$|\chi_{E/2} * \chi_{E/2}(x) - \chi_{A/2} * \chi_{A/2}(x)| \le |E \setminus A|.$$

Therefore, if $x \in J$ and $\operatorname{dist}(x, \partial J) > 10\,|E \setminus A|$, we can use (2.45) to obtain

$$\chi_{A/2} * \chi_{A/2}(x) \ge \chi_{E/2} * \chi_{E/2}(x) - |E \setminus A| > |E \setminus A| - |E \setminus A| = 0,$$

thus $x \in \frac{1}{2}(A+A)$ (see (2.6)). Since

$$|\{x \in J \ : \ \operatorname{dist}(x, \partial J) \le 10\,|E \setminus A|\}| \le 20\,|E \setminus A|,$$

it follows that $|J \setminus \frac{1}{2}(A+A)| \le 20\,|E \setminus A|$, and consequently

$$|J \setminus A| \le |\tfrac{1}{2}(A+A) \setminus A| + |J \setminus \tfrac{1}{2}(A+A)| \le |\tfrac{1}{2}(A+A) \setminus A| + 20\,|E \setminus A|,$$

as desired. □

To describe the complement of E in $\Sigma(t)$, we introduce several more notations. Recall that the vertices of $\Sigma = \Sigma(0)$ are $\hat{x}_i$, $i = 1, 2, 3$, so that the vertices of $\Sigma(t)$ are given by $\hat{x}_i(t) = (1 - 2\rho t)\hat{x}_i + (0, 0, \frac{1}{2\rho})$. Denote the sides of $\Sigma(t)$ by $\Sigma_i(t)$, with the convention that the endpoints of $\Sigma_1(t)$ are $\hat{x}_2(t)$ and $\hat{x}_3(t)$, and likewise for permutations of the indices. Since Σ has sidelength

$$s_0 := 2 \cdot 3^{-1/4},$$

the length of the sides of $\Sigma_i(t)$ is given by

$$s(t) := \mathscr{H}^1(\Sigma_i(t)), \qquad s(t) = (1 - 2\rho t)s_0.$$

Let $t \in [0, 1/2]$, and let $m \ge 1$ be such that $2^{-m} \le t < 2^{-m+1}$. We will define, iteratively, the set of open subintervals $I_{j,k}(t)$ of $\Sigma_1(t)$, with $j = 1, \dots m$ and $k = 1, \dots, 2^{j-1}$, whose union is the complement of E in $\Sigma_1(t)$. To begin, set

$$I_{1,1}(t) := \Sigma_1(t) \setminus \left(\tfrac{1}{2}(\hat{x}_2 + \Sigma_1(2t)) \cup \tfrac{1}{2}(\hat{x}_3 + \Sigma_1(2t))\right).$$

Then $I_{1,1}(t)$ is the open interval centered at the midpoint of $\Sigma_1(t)$ of length $s(t) - s(2t) = 2\rho t s_0$. The set $\Sigma_1(t) \setminus I_{1,1}(t)$ consists of two closed segments. Define $I_{2,1}(t)$ and $I_{2,2}(t)$ to be the open intervals with the same length as $I_{1,1}(t)$ centered at the midpoints of these two closed intervals. Continue iteratively, given $2^\ell - 1$ open subintervals of $\Sigma_1(t)$

$$I_{j,k}(t), \qquad j = 1, \dots, \ell, \quad k = 1, \dots, 2^{j-1},$$

of equal length $2\rho t s_0$ and equal spacing. The intervals $\{I_{\ell+1,k}(t)\}_{1 \le k \le 2^\ell}$ are of length $2\rho t s_0$ and centered at the midpoints of the closed intervals complementary to the intervals we have already defined.

Set

$$V_\ell := \{i 2^{-\ell}\hat{x}_2 + (2^\ell - i - 1)2^{-\ell}\hat{x}_3 : i = 0, \dots, 2^\ell - 1\}.$$

Then, by construction,

$$E \cap \Sigma_1(t) = \bigcup_{v \in V_m} (v + 2^{-m}\Sigma_1(2^m t)) = \Sigma_1(t) \setminus \bigcup_{j=1}^{m} \bigcup_{k=1}^{2^{j-1}} I_{j,k}(t). \tag{2.46}$$

There is, of course, a similar description of $E \cap \Sigma_2(t)$ and $E \cap \Sigma_3(t)$.

To describe the rest of $E \cap \Sigma(t)$, we introduce more notations. For any 1-dimensional segment I in $\mathbb{R}^3$, given $h > 0$ and $\alpha \geq 1$, define a "flared neighborhood" of I by

$$\mathcal{F}_{h,\alpha}(I) := \{x \in \mathbb{R}^3 \,:\, \mathrm{dist}(x, I^*) \leq h, \quad \mathrm{dist}(x, I) \leq \alpha\, \mathrm{dist}(x, I^*)\}$$

where I^* denotes the line containing I. Note that $\mathcal{F}_{h,\alpha}(I)$ is symmetric with respect to I^*, and consists of the union of two trapezoids with I as shorter base.

For $2^{-m} \leq t < 2^{-m+1}$, set

$$\mathcal{S}_1^{\alpha}(t) := \left(\bigcup_{j=1}^{m} \bigcup_{k=1}^{2^{j-1}} \mathcal{F}_{s_0 2^{-j+1}\epsilon,\alpha}(I_{j,k}(t)) \right) \cap \Sigma(t),$$

and define $\mathcal{S}_i^{\alpha}(t)$ as the image of $\mathcal{S}_1^{\alpha}(t)$ under any rigid motion of $\mathbb{R}^3$ that maps $\Sigma_1(t)$ to $\Sigma_i(t)$, see Figures 3 and 4.

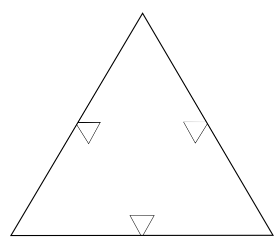


FIGURE 3. $\mathcal{S}_1^{\alpha}(t) \cup \mathcal{S}_2^{\alpha}(t) \cup \mathcal{S}_3^{\alpha}(t) \subset \Sigma(t)$ for $\alpha = \pi/3$ and $t \in [1/2, 1)$.

With these notations we can now estimate the complement of E.

LEMMA 2.4. – *For ρ and ϵ sufficiently small and for all t, $0 < t < 1/2$,*

$$\Sigma(t) \setminus E \subset \bigcup_{i=1}^{3} \mathcal{S}_i^{2}(t).$$

Before proving this lemma we will use it to finish the proof of Lemma 2.1 and hence Theorem 1.4.

Note that $\mathcal{S}_1^2(t) \cap \Sigma(t)$ is a union of "upward" trapezoids whose shorter bases are the $2^m - 1$ intervals $I_{j,k}(t)$ of length $2\rho t s_0$ ($2^{-m} \leq t < 2^{-m+1}$, $1 \leq j \leq m$, $1 \leq k \leq 2^{j-1}$). The complements in $\Sigma_1(t)$ of these bases are 2^m intervals of equal length $\ell(t)$ given by

$$\ell(t) := \mathcal{H}^1(2^{-m}\Sigma_1(2^m t)) = 2^{-m}(1 - 2\rho 2^m t)s_0 > (1 - 4\rho)2^{-m}s_0 .$$

For $i = 1, 2, 3$, let $T_i(t)$ be the isosceles triangle in $\Sigma(t)$ with base $\Sigma_i(t)$ whose equal sides are of slope 4ϵ relative to the base (and hence of height less than $2\epsilon s_0$). We claim that

$$\mathcal{S}_i^2(t) \cap \Sigma(t) \subset T_i(t), \quad i = 1, 2, 3,$$

see Figure 4.

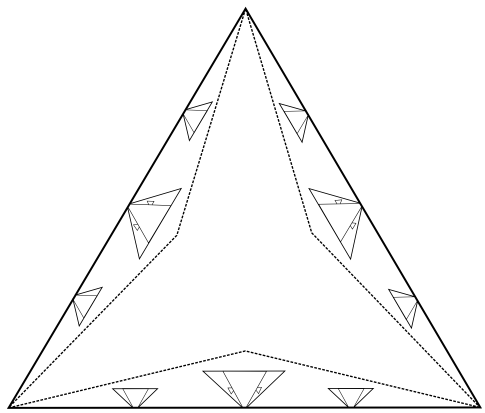


FIGURE 4. Some illustrative part of the fractal set appearing in the proof of Lemma 2.4. Note that the bases of the trapezoids have all the same length, given by $2\rho t s_0$. By widening the trapezoids from $\alpha = 2/\sqrt{3}$ to $\alpha = 2$, we ensure that even when later on in the iteration we may add some additional trapezoids to the lateral sides of a previous one, these will still be included in the wider trapezoid. The dotted lines represent the triangles $T_i(t)$, $i = 1, 2, 3$.

To see this, suppose, without loss of generality, that $i = 1$ and call the direction of $\Sigma_1(t)$ horizontal. The left side of the smallest isosceles triangle with base $\Sigma_1(t)$ that encloses $\mathcal{S}_1^2(t) \cap \Sigma(t)$ starts at the left endpoint of $\Sigma_1(t)$ and passes through the upper left corner of the short trapezoid in $\mathcal{S}_1^2(t) \cap \Sigma(t)$ nearest that corner. That trapezoid has height $h_m = 2^{-m+1}\epsilon s_0$, and horizontal distance from the endpoint of $\Sigma_1(t)$ given by $\ell(t) - \sqrt{3}h_m$. Thus the slope is

$$\frac{h_m}{\ell(t) - h_m} \le \frac{2^{-m+1}\epsilon s_0}{(1-4\rho)2^{-m}s_0 - \sqrt{3}\,2^{-m+1}\epsilon s_0} = \frac{2\epsilon}{1 - 4\rho - 2\sqrt{3}\,\epsilon} \le 4\epsilon,$$

for ρ and ϵ less than $1/100$.

Next, for $0 \le h \le \epsilon s_0$, we consider segments parallel to the side $\Sigma_1(t)$, excluding very short segments at the ends corresponding to the thin triangles $T_2(t)$ and $T_3(t)$, and then remove, in addition, $\mathcal{S}_1^2(t)$:

$$J_1^h(t) := \{x \in \Sigma(t) : \operatorname{dist}(x, \Sigma_1(t)) = h\} \setminus (T_2(t) \cup T_3(t)); \quad E_1^h(t) = J_1^h(t) \setminus \mathcal{S}_1^2(t).$$

(See Figure 5.) We define $E_i^h(t) \subset J_i^h(t)$ analogously for $i = 2, 3$. Lemma 2.4 implies that $E_i^h(t) \subset E$, and hence

$$\int_0^1 \int_0^{\epsilon s_0} \mathcal{H}^1(E_i^h(t) \setminus A)\, dh\, dt \le |E \setminus A|, \quad i = 1, 2, 3. \tag{2.47}$$

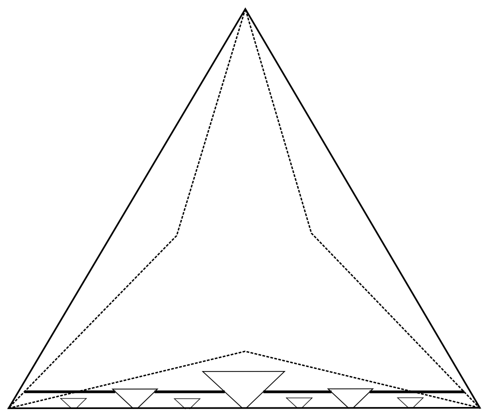


FIGURE 5. The bold line represents the set $E_1^h(t)$. This is obtained by taking an horizontal segment at height h connecting the triangles $T_2(t)$ and $T_3(t)$, and removing the part covered by $\mathcal{S}_1^2(t)$

To confirm that the one-dimensional Lemma 2.3 applies to $E_1^h(t)$ as a subset of the interval $J_1^h(t)$, observe that the set $J_1^h(t) \cap \mathcal{S}_1^2(t)$ that we excluded to form $E_1^h(t)$ consists of equally spaced intervals of equal length

$$\mathscr{H}^1(I_{j,k}) + 2\sqrt{3}\,h = 2\rho t s_0 + 2\sqrt{3}\,h, \quad s_0 2^{-j+1}\epsilon \geq h, \quad 1 \leq k \leq 2^{j-1}.$$

The value of j ranges from 1 to j^* with the maximum value determined by the constraints $j^* \leq m$ and $2^{j^*} \leq 2s_0\epsilon/h$. The total number of intervals is

$$1 + 2 + \cdots + 2^{j^*-1} \;=\; 2^{j^*} - 1 \;<\; 2^{j^*} \;\leq\; \min\left(2^m, \frac{2s_0\epsilon}{h}\right).$$

Since $2^{m-1}t \leq 1$ (by the definition of m), the total length of these complementary intervals is less than

$$(2\rho t s_0 + 2\sqrt{3}\,h)\min\left(2^m, \frac{2s_0\epsilon}{h}\right) \leq 2^{m+1}\rho t s_0 + 4\sqrt{3}\epsilon s_0 \leq 10(\rho+\epsilon)s_0.$$

Note that $J_1^h(t)$ has length $(1 - O(\epsilon+\rho))s_0$, and that $J_1^h(t) \cap \mathcal{S}_1^2(t)$ is at most an $O(\rho+\epsilon)$ fraction of $J_1^h(t)$. It follows that, for all $x \in J_1^h(t)$,

$$\chi_{E_1^h(t)/2} * \chi_{E_1^h(t)/2}(x) \geq (1 - O(\epsilon+\rho))\operatorname{dist}(x, \partial J_1^h(t)),$$

in which we abuse notation by identifying $J_1^h(t)$ with its isometric image in a real line and likewise the subset $E_1^h(t)$. (Note that although the $2^m - 2$ internal intervals of $E_1^h(t)$ have equal length, the two on the ends are slightly longer. This only improves the convolution inequality at the very ends. We excluded the triangles $T_2(t)$ and $T_3(t)$ from $J_1^h(t)$ in order to arrange this favorable situation at the ends: we do not want the interval on which we apply Lemma 2.3 to intersect $\mathcal{S}_2^2(t)$ and $\mathcal{S}_3^2(t)$.)

Having confirmed the hypothesis of Lemma 2.3, and likewise for the analogous sets $E_i^h(t) \subset J_i^h(t)$, we apply the lemma to conclude that

$$\mathscr{H}^1(J_i^h(t) \setminus A) \leq \mathscr{H}^1(J_i^h(t) \cap \tfrac{1}{2}(A+A) \setminus A) + 20\,\mathscr{H}^1(E_i^h(t) \cap A), \quad i = 1, 2\ 3.$$

Since

$$K(t) \subset E(t) \cup \left(\bigcup_{i=1}^{3} \bigcup_{h=0}^{\epsilon s_0} J_i^h(t) \right),$$

these three inequalities, along with (2.47) and Fubini's theorem, imply that

$$\begin{aligned} |K \setminus A| &\leq |E \setminus A| + \sum_{i=1}^{3} \int_0^1 \int_0^{\epsilon s_0} \mathscr{H}^1(J_i^h(t) \setminus A)\, dh\, dt \\ &\leq |E \setminus A| + 3|\tfrac{1}{2}(A+A) \setminus A| + 60|E \setminus A| \leq 64 C_0\, \delta(A) \end{aligned}$$

with the dimensional constant $C_0 \geq 1$ of (2.15). This ends the proof of Lemma 2.1 and Theorem 1.4, except for the proof of Lemma 2.4 that we now provide.

Proof of Lemma 2.4. – The complementary set $\Sigma(t) \setminus E$ is a fractal built iteratively out of (occasionally truncated) trapezoids arising as the complements of sets of scaled equilateral triangles. Figure 3 shows the fractal in its simplest, starting layer $1/2 \leq t < 1$. We will organize the description of a superset of the fractal. Figure 4 shows the widened trapezoids of the superset that we will use to enclose successive generations of smaller and smaller trapezoids in the fractal. Within $T_1(t)$, the triangle with base $\Sigma_1(t)$ defined above (see Figure 4), we will refer to the "first generation" of the complementary set as the set involving semisums with the endpoints $\hat{x}_2$ and $\hat{x}_3$ and trapezoids that touch $\Sigma_1(t)$ only. This first generation is a subset of $\mathscr{S}_1^{\alpha}(t)$ with $\alpha = \alpha_0 = 2/\sqrt{3}$, corresponding to the angle $\pi/3$. The second generation of points in $T_1(t) \setminus E$ arise from first generation points in $T_2(2t)$ and $T_3(2t)$. Consider, for example, the semisum of $\hat{x}_3$ with points of the first generation in $T_3(2t)$. For any $\alpha < 2$,

$$\mathscr{S}_3^{\alpha}(2t) \cap \Sigma(2t) \subset \mathscr{S}_3^{2}(2t) \cap \Sigma(2t) \subset T_3(2t).$$

Therefore,

$$\tfrac{1}{2}\left(\hat{x}_3 + \Sigma(2t) \cap \mathscr{S}_3^{\alpha}(2t)\right) \setminus (1-\epsilon)K \subset \tfrac{1}{2}\left(\hat{x}_3 + T_3(2t)\right) \setminus (1-\epsilon)K$$

is contained in a triangle of base size $O(\epsilon)$ and height $O(\epsilon^2)$. More precisely, the base is a non-parallel side of the trapezoid $\Sigma(t) \cap \mathscr{T}_{s_0\epsilon,\alpha_0}(I_{1,1}(t))$, and the other vertex is on the line parallel to $I_{1,1}(t)$ at distance $s_0\epsilon$. Note the very important shrinkage that comes from subtracting $(1-\epsilon)K$. The set we are translating is contained in a triangle of size $O(1)$ by $O(\epsilon)$ but the part of the translation that is outside of $(1-\epsilon)K$ has diameter $O(\epsilon)$ and width $O(\epsilon^2)$. The second generation exceptional set is covered by opening the neighborhood of $I_{1,1}(t)$ by changing the flare parameter from α_0 to $\alpha_1 = \alpha_0 + 10\epsilon$. The same widening eventually occurs, appropriately scaled, at all of the intervals $I_{j,k}(t)$ at least for sufficiently small t, but no other additions occur if we only use one step with a convex combination involving a vertex and an opposite side. In all, at the second generation, in which at most one such step is used, the exceptional set is contained in the set

$$\bigcup_{i=1}^{3} \mathscr{S}_i^{\alpha_1}(t), \quad \alpha_1 = \alpha_0 + 10\epsilon\,.$$

Repeating this argument, we find that the exceptional set generated using at most k steps involving a vertex and an opposite side is contained in

$$\bigcup_{i=1}^{3} \mathcal{S}_i^{\alpha_k}(t), \quad \alpha_k = \alpha_0 + (10\epsilon) + (10\epsilon)^2 + \cdots + (10\epsilon)^k .$$

Evidently, for sufficiently small ϵ, $\alpha_k < 2$ for all k. This covers the entire complement of E in $\Sigma(t)$ and concludes the proof of Lemma 2.4. □

REMARK 2.5. – *In closing, we note that in our inductive argument for $n = 3$, we proved that the complement of E contains only relatively short one-dimensional segments at all appropriate scales near the boundary of K. When $n = 4$ the set E has nearly full $\mathcal{H}^4$ measure on many suitably scaled subsets, but its complement has too many segments of large diameter near ∂K. Therefore, further arguments are required to enlarge E enough to finish the case $n = 4$ and higher.*

BIBLIOGRAPHY

[1] M. Christ, Near equality in the Brunn-Minkowski inequality, preprint arXiv:1207.5062.

[2] G. Ciraolo, A. Figalli, F. Maggi, M. Novaga, Rigidity and sharp stability estimates for hypersurfaces with constant and almost-constant nonlocal mean curvature, *J. reine angew. Math.* **741** (2018), 275–294.

[3] S. Dubuc, Critères de convexité et inégalités intégrales, *Ann. Inst. Fourier* **27** (1977), 135–165.

[4] D. Ellis, E. Friedgut, G. Kindler, A. Yehudayoff, Geometric stability via information theory, *Discrete Anal.* (2016), Paper No. 10, 29.

[5] A. Figalli, D. Jerison, Quantitative stability for sumsets in $\mathbb{R}^n$, *J. Eur. Math. Soc. (JEMS)* **17** (2015), 1079–1106.

[6] A. Figalli, D. Jerison, Quantitative stability for the Brunn-Minkowski inequality, *Adv. Math.* **314** (2017), 1–47.

[7] A. Figalli, F. Maggi, A. Pratelli, A refined Brunn-Minkowski inequality for convex sets, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26** (2009), 2511–2519.

[8] A. Figalli, F. Maggi, A. Pratelli, A mass transportation approach to quantitative isoperimetric inequalities, *Invent. math.* **182** (2010), 167–211.

[9] G. A. Freĭman, The addition of finite sets. I, *Izv. Vysš. Učebn. Zaved. Matematika* **1959** (1959), 202–213.

[10] F. John, Extremum problems with inequalities as subsidiary conditions, in *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948*, Interscience Publishers, N. Y., 1948, 187–204.

[11] A. DE ROTON, Small sumsets in $\mathbb{R}$: full continuous $3k-4$ theorem, critical sets, *J. Éc. polytech. Math.* **5** (2018), 177–196.

(Manuscrit reçu le 4 mai 2017 ;
accepté le 17 juillet 2019.)

Alessio FIGALLI
ETH Zürich
Department of Mathematics
Rämistrasse 101, 8092 Zürich
Switzerland
E-mail: alessio.figalli@math.ethz.ch

David JERISON
Department of Mathematics
Massachusetts Institute of Technology
77 Massachusetts Ave
Cambridge, MA 02139-4307, USA
E-mail: jerison@math.mit.edu

INSTRUCTIONS AUX AUTEURS

Les Annales scientifiques de l'École normale supérieure publient des articles de mathématique en français, anglais, allemand ou italien. Les manuscrits soumis pour publication doivent être accompagnés d'un résumé en français et en anglais de 150 mots maximum. Les titres en français doivent être traduits en anglais, les titres en anglais, allemand ou italien doivent l'être en français.

Les articles acceptés doivent être composés en TEX (de préférence LATEX). Pour faciliter le travail ultérieur d'édition, on recommande d'éviter d'introduire des macros nouvelles.

Les articles seront soumis sous forme de fichier PDF sur la plateforme `https://annales-ens.centre-mersenne.org/index.php/ASENS/index`

INSTRUCTIONS TO AUTHORS

The Annales scientifiques de l'École normale supérieure publish mathematical papers in French, English, German or Italian. Manuscripts submitted for publication should be sent together with an abstract in French and in English, of maximum length 150 words. Titles in French must be supplied with an English translation. Titles in English, German or Italian should have a translation in French.

Accepted papers must be prepared using the TEX typesetting system (preferably in LATEX). To facilitate the eventual editorial work, we recommend to avoid introducing any new macros.

The PDF files of the papers can be uploaded on the platform `https://annales-ens.centre-mersenne.org/index.php/ASENS/index`

Directeur de la publication : Fabien Durand. Responsable de la rédaction : Yves de Cornulier.

SOMMAIRE / CONTENTS

TOME 54 FASCICULE 1 2021

ARTICLES À PARAÎTRE / FORTHCOMING PAPERS

R. J. Conlon & F. Rochon. – New examples of complete Calabi-Yau metrics on $\mathbb{C}^n$ for $n \geq 3$. *Nouveaux exemples de métriques de Calabi-Yau complètes sur $\mathbb{C}^n$ pour $n \geq 3$.*

E. Gwynne & J. Miller. – Convergence of the self-avoiding walk on random quadrangulations to $\mathrm{SLE}_{8/3}$ on $\sqrt{8/3}$-Liouville quantum gravity. *Convergence des marches auto-évitantes sur des quadrangulations aléatoires vers le processus $SLE_{8/3}$ sur une gravité quantique de Liouville de paramètre $\sqrt{8/3}$.*

S. Popa. – On the vanishing cohomology problem for cocycle actions of groups on II_1 factors. *Sur le problème d'annulation en cohomologie pour les actions à cocycle de groupes sur les facteurs II_1.*

R. Boutonnet, A. Ioana & J. Peterson. – Properly proximal groups and their von Neumann algebras. *Groupes proprement proximaux et leurs algèbres de von Neumann.*

T. Marquis. – Cyclically reduced elements in Coxeter groups. *Éléments cycliquement réduits dans les groupes de Coxeter.*

A. M. Licata & H. Queffelec. – Braid groups of type ADE, Garside monoids, and the categorified root lattice. *Groupes de tresses en type ADE, monoïdes de Garside et le réseau des racines catégorifié.*

Société Mathématique de France
SMF