ASTÉRISQUE 287

GEOMETRIC METHODS IN DYNAMICS (II)

VOLUME IN HONOR OF JACOB PALIS

edited by

Welington de Melo
Marcelo Viana
Jean-Christophe Yoccoz

Société Mathématique de France 2003
Publié avec le concours du Centre National de la Recherche Scientifique

W. de Melo

IMPA, Estrada Dona Castorina, 110, Jardim Botânico, Rio de Janeiro 22460-320, Brazil.

E-mail : demelo@impa.br

Url : www.impa.br/~demelo

M. Viana

IMPA, Estrada Dona Castorina, 110, Jardim Botânico, Rio de Janeiro 22460-320, Brazil.

E-mail : viana@impa.br

Url : www.impa.br/~viana

J.-C. Yoccoz

Collège de France, 11, Place Marcelin Berthelot, 75005 Paris, France.

E-mail : jean-c.yoccoz@college-de-france.fr

Url : www.college-de-france.fr/site/equ_dif/p999000715275.htm

2000 *Mathematics Subject Classification.* — 37-XX, 34-XX, 60-XX, 35-XX.

***Key words and phrases.* —** Dynamical systems, ergodic theory, bifurcation theory, differential equations.

GEOMETRIC METHODS IN DYNAMICS (II)

VOLUME IN HONOR OF JACOB PALIS

edited by Welington de Melo, Marcelo Viana, Jean-Christophe Yoccoz

Abstract. **—** This is the second of two volumes collecting original research articles, on several aspects of dynamics, mostly by participants in the International Conference on Dynamical Systems held at IMPA (Rio de Janeiro), in July 2000, to celebrate Jacob Palis' 60th birthday.

Résumé **(Méthodes géométriques en dynamique (II). Volume en l'honneur de Jacob Palis)**

Ceci est le second de deux volumes regroupant des articles originaux de recherche concernant des aspects variés de la théorie des systèmes dynamiques, écrits par certains des participants à la Conférence Internationale sur les Systèmes Dynamiques qui s'est tenue à l'IMPA (Rio de Janeiro), en juillet 2000 pour commémorer le 60^{e} anniversaire de Jacob Palis.

CONTENTS

CONTENTS OF VOLUME I

ABSTRACTS

We consider unitarily invariant probability measures on $\mathbb{GL}_n(\mathbb{C})$ and compare the mean of the logs of the moduli of the eigenvalues of the matrices to the Lyapunov exponents of random matrix products independently drawn with respect to the measure. We prove that the former is always greater or equal to the latter.

For two-dimensional quasi-geostrophic equation in Fourier space, we propose a new type approximation representing itself some quasi-linear equation. Natural finite dimensional approximations of this equation are investigated in the article.

We prove that in the space of all C^r $(r \geqslant 1)$ partially hyperbolic diffeomorphisms, there is a C^1 open and dense set of accessible diffeomorphisms. This settles the C^1 case of a conjecture of Pugh and Shub. The same result holds in the space of volume preserving or symplectic partially hyperbolic diffeomorphisms. Combining this theorem with results in [**Br**], [**Ar**] and [**PugSh3**], we obtain several corollaries. The first states that in the space of volume preserving or symplectic partially hyperbolic diffeomorphisms, topological transitivity holds on an open and dense set. Further, on a symplectic n-manifold $(n \leqslant 4)$ the C^1-closure of the stably transitive symplectomorphisms is precisely the closure of the partially hyperbolic symplectomorphisms. Finally, stable ergodicity is C^1 open and dense among the volume preserving, partially hyperbolic diffeomorphisms satisfying the additional technical hypotheses of [**PugSh3**].

In this paper we embed a high genus surface in $\mathbb{R}^3$ so that its geodesic flow has no conjugate points and is Anosov, despite the fact that its curvature cannot be everywhere negative.

We present a class of random cellular automata with multiple invariant measures which are all non-Gibbsian. The automata have configuration space $\{0,1\}^{\mathbb{Z}^d}$, with $d > 1$, and they are noisy versions of automata with the "eroder property". The noise is totally asymmetric in the sense that it allows random flippings of "0" into "1" but not the converse. We prove that all invariant measures assign to the event "a sphere with a large radius L is filled with ones" a probability μ_L that is too large for the measure to be Gibbsian. For example, for the NEC automaton $(-\ln \mu_L) \asymp L$ while for any Gibbs measure the corresponding value is $\asymp L^2$.

Let $X : \mathbb{R}^2 \smallsetminus \overline{D}_\sigma \to \mathbb{R}^2$ be a C^1 map, where $\sigma > 0$ and $\overline{D}_\sigma = \{p \in \mathbb{R}^2 : ||p|| \leqslant \sigma\}$.
(i) If for some $\varepsilon > 0$ and for all $p \in \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$, no eigenvalue of $DX(p)$ belongs to $(-\varepsilon, \infty)$, there exists $s \geqslant \sigma$, such that $X|_{\mathbb{R}^2 \smallsetminus \overline{D}_s}$ is injective;
(ii) If for some $\varepsilon > 0$ and for all $p \in \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$, no eigenvalue of $DX(p)$ belongs to $(-\varepsilon, 0] \cup \{z \in \mathbb{C} : \Re(z) \geqslant 0\}$, there exists $p_0 \in \mathbb{R}^2$ such that the point ∞, of the Riemann sphere $\mathbb{R}^2 \cup \{\infty\}$, is either an attractor or a repellor of $x' = X(x) + p_0$.

The averaging setup arises in the study of perturbations of parametric families of dynamical systems when parameters start changing slowly in time. Usually, averaging methods are applied to systems of differential equations which combine slow and fast motions. This paper deals with difference equations case which leads to wider class of models and examples. The averaging principle is justified here under a general condition which is verified when unperturbed transformations either preserve smooth measures or they are hyperbolic. The convergence speed in the averaging principle is estimated for some cases, as well.

We consider Axiom A diffeomorphisms g in the isotopy class of a pseudoanosov map f. It is shown that they have a unique "large" basic piece Λ, and necessary and sufficient conditions for g to be semiconjugated to f, that only involve conditions on Λ, are obtained. As a consequence, it is proved that if Λ is exteriorly situated, stable and unstable half-leaves of points of Λ boundedly deviate from geodesics.

We show a positive Livsic type theorem for $\mathcal{C}^2$ Anosov diffeomorphisms f on a compact boundaryless manifold M and Hölder observables A. Given $A : M \to \mathbb{R}$, α-Hölder, we show there exist $V : M \to \mathbb{R}$, β-Hölder, $\beta < \alpha$, and a probability measure μ, f-invariant such that

$$A \leqslant V \circ f - V + \int A \, d\mu.$$

We apply this inequality to prove the existence of an open set $\mathcal{G}_\beta$ of β-Hölder functions, β small, which admit a unique maximizing measure supported on a periodic orbit. Moreover the closure of $\mathcal{G}_\beta$, in the β-Hölder topology, contains all α-Hölder functions, α close to one.

Let $p > 1$ be a prime number, $\mathbb{Q}_p$ the field of p-adic numbers and let $\mathbb{C}_p$ be the smallest complete extension of $\mathbb{Q}_p$ that is algebraically closed. This work is dedicated to the study of the dynamics of rational functions on the projective line $\mathbb{P}(\mathbb{C}_p)$.
To each rational function $R \in \mathbb{C}_p(z)$ we associate its *quasi-periodicity domain*, which is equal to the interior of the set of points in $\mathbb{P}(\mathbb{C}_p)$ that are recurrent by R. We give several caracterizations of the quasi-periodicity domain and we describe its local and global dynamics.
We prove that analytic components of the domain of quasi-periodicity (which are the p-adic analogues of Siegel discs and Herman rings) are open affinoïds (that is, they have simple geometry) and we describe the dynamics on a given component.
Like in the complex case there is a partition of the line $\mathbb{P}(\mathbb{C}_p)$ in the Fatou set and the Julia set. By analogy to the complex case we make the following non-wandering conjecture: every wandering disc is attracted to an attracting cycle. We prove that this holds if and only if every point in the Fatou set is either attracted to an attracting cycle or is mapped to the quasi-periodicity domain under forward iteration.

Let (M, g) be a compact Riemannian manifold without conjugate points. Suppose that the horospheres in $(\widetilde{M}, g)$ depend continuously on their normal directions. Then we show that geodesic rays diverge uniformly in the universal covering $(\widetilde{M}, g)$. We give some applications of this result to the study of the dynamics of the geodesic flow and the global geometry of manifolds without conjugate points.

In this work we study a certain type of discrete groups acting on higher dimensional complex projective spaces. These generalize the classical Schottky groups acting on the Riemann sphere. We study the limit sets of these actions, which turn out to be solenoids. We also look at the compact complex manifols obtained as quotient of the region of discontinuity, divided by the action. We determine their topology and the dimension of the space of their infinitesimal deformations. We show that every such deformation arises from a deformation of the embedding of the group in question into the group of automorphisms of the corresponding complex projective space, which is a reminiscent of the classical Teichmüller theory.

RÉSUMÉS DES ARTICLES

Étant donné une mesure de probabilité sur $\mathbb{GL}_n(\mathbb{C})$ qui est unitairement invariante, nous comparons la moyenne des logarithmes des modules des valeurs propres des matrices aux exposants de Lyapunov des produits de matrices aléatoires indépendantes pour cette mesure. Nous montrons que celui-là est toujours plus grand que celui-ci.

Pour l'équation quasi-géostrophique en deux dimensions dans l'espace de Fourier, nous proposons une nouvelle approximation représentant elle-même une équation quasi-linéaire. On étudie dans cet article des approximations de dimension finie naturelles de cette équation.

Nous montrons que, dans l'espace de tous les difféomorphismes partiellement hyperboliques de classe C^r ($r \geqslant 1$), il existe un ensemble C^1 ouvert et dense de difféomorphismes accessibles. Ceci établit le cas C^1 d'une conjecture de Pugh et Shub. Le même résultat vaut dans l'espace des difféomorphismes partiellement hyperboliques préservant le volume ou symplectiques. En combinant ce théorème avec des résultats de [**Br**], [**Ar**] et [**PugSh3**], nous obtenons plusieurs corollaires. Le premier énonce que, dans l'espace des difféomorphismes partiellement hyperboliques préservant le volume ou symplectiques, la transitivité topologique a lieu sur un ensemble ouvert et dense. Puis, sur une variété symplectique de dimension n ($n \leqslant 4$), l'adhérence C^1 des symplectomorphismes

stablement transitifs est précisément celle des symplectomorphismes partiellement hyperboliques. Enfin, l'ergodicité stable est C^1 ouverte et dense dans l'espace des difféomorphismes partiellement hyperboliques préservant le volume satisfaisant l'hypothèse technique additionnelle de [**PugSh3**].

Dans cet article, nous plongeons une surface de grand genre dans $\mathbb{R}^3$ de telle manière que son flot géodésique n'ait aucun point conjugé et soit Anosov, malgré le fait que la courbure ne puisse être partout négative.

Nous présentons une classe d'automates cellulaires aléatoires avec plusieurs mesures invariantes qui sont toutes non gibbsiennes. Les automates ont $\{0,1\}^{\mathbb{Z}^d}$, avec $d > 1$, comme espace de configuration, et ce sont des versions avec bruit d'automates ayant la "propriété d'érodeur". Le bruit est totalement asymétrique dans le sens qu'il permet des sauts aléatoires de "0"en "1"mais pas le contraire. Nous montrons que toutes les mesures invariantes attachent à l'événement "une sphère de grand rayon L est remplie de 1"une probabilité μ_L qui est trop grande pour qu'une mesure soit gibbsienne. Par exemple, pour l'automate NEC, $(-\ln \mu_L) \asymp L$ alors que pour toute mesure gibbsienne la valeur correspondante est $\asymp L^2$.

Soit $X : \mathbb{R}^2 \setminus \overline{D}_\sigma \to \mathbb{R}^2$ une application C^1, où $\sigma > 0$ et $\overline{D}_\sigma = \{p \in \mathbb{R}^2 : ||p|| \leqslant \sigma\}$.
(i) Si pour un $\varepsilon > 0$ et pour tout $p \in \mathbb{R}^2 \setminus \overline{D}_\sigma$, aucune valeur propre de $DX(p)$ n'appartient à $(-\varepsilon, \infty)$, alors il existe $s \geqslant \sigma$ tel que $X|_{\mathbb{R}^2 \setminus \overline{D}_s}$ est injective.
(ii) Si pour un $\varepsilon > 0$ et pour tout $p \in \mathbb{R}^2 \setminus \overline{D}_\sigma$, aucune valeur propre de $DX(p)$ n'appartient à $(-\varepsilon, 0] \cup \{z \in \mathbb{C} : \Re(z) \geqslant 0\}$, alors il existe $p_0 \in \mathbb{R}^2$ tel que le point ∞ de la sphère de Riemann $\mathbb{R}^2 \cup \{\infty\}$ soit un attracteur ou un repulseur de $x' = X(x) + p_0$.

La moyennisation apparaît dans l'étude des perturbations d'une famille paramétrée de systèmes dynamiques, lorsque les paramètres varient lentement avec le temps. D'habitude, les méthodes de moyennisation sont appliquées aux systèmes d'équations différentielles qui combinent des mouvements lents et rapides. Cet article traite le cas des équations aux différences, qui conduit à un ensemble plus grand de modèles et d'exemples. Le principe de moyennisation est justifié ici sous une condition générale qui est vérifiée lorsque les transformations non perturbées ou bien préservent des mesures lisses ou bien sont hyperboliques. On estime aussi la vitesse de convergence dans le principe de moyennisation.

On considère les difféomorphismes g Axiom A, isotopes à des transformations pseudo-Anosov f. On montre qu'ils ont une unique "grande"partie basique Λ, et on trouve des conditions nécessaires et suffisantes pour que g soit semi-conjugué a f. Ces conditions s'expriment seulement en terme des propriétés de Λ. Comme conséquence on obtient que si Λ est située extérieurement, les demi-feuilles stables et instables des points de Λ sont à distance bornée des géodésiques.

Nous montrons un théorème de type Livsic positif pour les C^2-difféomorphismes Anosov f sur une variété compacte sans bord M et des observables A höldériennes. Étant donnée $A : M \to \mathbb{R}$, α-höldérienne, nous montrons qu'il existe $V : M \to \mathbb{R}$, β-höldérienne, $\beta < \alpha$, et une mesure de probabilité μ, f-invariante, telles que

$$A \leqslant V \circ f - V + \int A \, d\mu.$$

Nous appliquons cette inégalité pour montrer l'existence d'un ouvert $\mathcal{G}_\beta$ de fonctions β-höldériennes, β petit, qui admet une unique mesure maximisante supportée par une orbite périodique. De plus, l'adhérence de $\mathcal{G}_\beta$ dans la topologie β-höldérienne contient toutes les fonctions α-höldériennes, avec α proche de 1.

Soit $p > 1$ un nombre premier, $\mathbb{Q}_p$ le corps des nombres p-adiques et soit $\mathbb{C}_p$ la plus petite extension complète et algébriquement close de $\mathbb{Q}_p$. Ce travail est consacré à l'étude de la dynamique des fonctions rationnelles sur la droite projective $\mathbb{P}(\mathbb{C}_p)$.

À chaque fonction rationnelle $R \in \mathbb{C}_p(z)$ on associe son *domaine de quasi-périodicité*, qui est égal à l'intérieur de l'ensemble des points dans $\mathbb{P}(\mathbb{C}_p)$ qui sont récurrents par R. On donne plusieurs caractérisations du domaine de quasi-périodicité et on décrit sa dynamique locale et globale.

On montre que les composantes du domaine de quasi-périodicité (qui sont les analogues p-adiques des disques des Siegel et des anneaux de Herman) sont des

affinoïdes ouverts (c'est-à-dire que leur géométrie est simple) et on décrit la dynamique sur une composante donnée.
Comme dans le cas complexe on a une partition de la droite $\mathbb{P}(\mathbb{C}_p)$ en l'ensemble de Fatou et l'ensemble de Julia. Par analogie au cas complexe on fait la conjecture de non-errance suivante : tout disque errant est attiré par un cycle attractif. On montre que ceci a lieu si et seulement si tout point dans l'ensemble de Fatou est soit attiré par un cycle attractif, soit rencontre le domaine de quasi-périodicité par itération positive.

Soit (M, g) une variété riemannienne compacte sans points conjugués. Supposons que les horosphères dans $(\widetilde{M}, g)$ dépendent de façon continue de ses vecteurs normaux. Alors, les rayons géodésiques divergent uniformément dans le revêtement universel $(\widetilde{M}, g)$. Nous présentons quelques applications de ce résultat à l'étude de la dynamique du flot géodésique et la géométrie globale des variétés sans points conjugués.

Dans ce travail, nous étudions un certain type de groupes discrets agissant sur les espaces projectifs complexes de dimensions supérieures. Ces actions généralisent les actions classiques de type Schottky sur la sphère de Riemann. Nous étudions les ensembles limites de ces actions, qui se trouvent être des solénoïdes. Nous considérons aussi les variétés complexes compactes obtenues comme quotient de la région de discontinuité par l'action du groupe. Nous déterminons leur topologie et la dimension de l'espace des déformations infinitésimales. Une telle déformation provient d'une déformation du groupe initial dans le groupe des automorphismes projectifs correspondants, ce qui est une réminiscence de la théorie classique de Teichmüller.

PREFACE

These two volumes collect original research articles submitted by participants of the International Conference on Dynamical Systems held at IMPA, Rio de Janeiro, in July 19-28, 2000 to commemorate the 60th birthday of Jacob Palis.

These articles cover a wide range of subjects in Dynamics, reflecting the Conference's broad scope, itself a tribute to the diversity and influence of Jacob's contributions to the mathematical community worldwide, and most notably in Latin America, through his scientific work, his role as an educator of young researchers, his responsibilities in international scientific bodies, and the efforts he has always devoted to fostering the development of Mathematics in all regions of the globe.

His own mathematical work, which extends for more than 80 publications, is described in Sheldon Newhouse's opening article. It is, perhaps, best summarized by the following quotation from Jacob's recent nomination for the French Academy of Sciences: "sa vision, en constante évolution, a considérablement élargi le sujet".

As Jacob does not seem willing to slow down, we should expect much more from him in the years to come...

Rio de Janeiro and Paris,
May 20, 2003
Welington de Melo, Marcelo Viana, Jean-Christophe Yoccoz

Astérisque
287, 2003, p. 1–18

ON RANDOM AND MEAN EXPONENTS FOR UNITARILY INVARIANT PROBABILITY MEASURES ON $\mathbb{GL}_n(\mathbb{C})$

by

Jean-Pierre Dedieu & Mike Shub

Dedicated to Jacob Palis for his sixtieth birthday.

Abstract. — We consider unitarily invariant probability measures on $\mathbb{GL}_n(\mathbb{C})$ and compare the mean of the logs of the moduli of the eigenvalues of the matrices to the Lyapunov exponents of random matrix products independently drawn with respect to the measure. We prove that the former is always greater or equal to the latter.

1. Introduction

Given a probability measure μ on the space of invertible $n \times n$ complex matrices satisfying a mild integrability condition, we have, by Oseledec's Theorem, n random exponents $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_n \geqslant -\infty$ such that for almost every sequence $\ldots g_k \ldots g_1 \in \mathbb{GL}_n(\mathbb{C})$ the limit $\lim \frac{1}{k} \log \|g_k \ldots g_1 v\|$ exists for every $v \in \mathbb{C}^n \smallsetminus \{0\}$ and equals one of the r_i, $i = 1 \ldots n$, see Gol'dsheid and Margulis [**4**] or Ruelle [**8**] or Oseledec [**7**]. The numbers $r_1, \ldots, r_n$ are called Lyapunov exponents. In our context we may call them random Lyapunov exponents or even just random exponents. If the measure is concentrated on a point A, these numbers $\lim \frac{1}{n} \log \|A^n v\|$ are $\log |\lambda_1|, \ldots, \log |\lambda_n|$ where $\lambda_i(A) = \lambda_i$, $i = 1 \ldots n$, are the eigenvalues of A written with multiplicity and $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n|$.

The integrability condition for Oseledec's Theorem is

$$g \in \mathbb{GL}_n(\mathbb{C}) \to \log^+(\|g\|) \text{ is } \mu - \text{integrable}$$

where for a real valued function f, $f^+ = \max[0, f]$. Here we will assume more so that all our integrals are defined and finite, namely:

$$(*) \qquad g \in \mathbb{GL}_n(\mathbb{C}) \to \log^+(\|g\|) \text{ and } \log^+(\|g^{-1}\|) \text{ are } \mu\text{-integrable.}$$

2000 *Mathematics Subject Classification.* — 37Axx, 37Hxx.
Key words and phrases. — Random matrix, Lyapunov exponents, unitarily invariant probability measure, co-area formula.

We will prove:

Theorem 1. — *If μ is a unitarily invariant measure on $\mathbb{GL}_n(\mathbb{C})$ satisfying $(*)$ then, for $k = 1, \ldots, n$,*

$$\int_{A\in\mathbb{GL}_n(\mathbb{C})} \sum_{i=1}^{k} \log|\lambda_i(A)| d\mu(A) \geqslant \sum_{i=1}^{k} r_i.$$

By unitary invariance we mean $\mu(U(X)) = \mu(X)$ for all unitary transformations $U \in \mathbb{U}_n(\mathbb{C})$ and all μ-measurable $X \subseteq \mathbb{GL}_n(\mathbb{C})$.

Corollary 2

$$\int_{A\in\mathbb{GL}_n(\mathbb{C})} \sum_{i=1}^{n} \log^+|\lambda_i(A)| d\mu(A) \geqslant \sum_{i=1}^{n} {r_i}^+.$$

Theorem 1 is not true for general measures on $\mathbb{GL}_n(\mathbb{C})$ or $\mathbb{GL}_n(\mathbb{R})$ even for $n = 2$. Consider

$$A_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

and give probability $1/2$ to each. Then the left hand integral is zero but as is easily seen the right hand sum is positive. So, in this case the inequality goes the other way. We do not know a characterization of measures which make Theorem 1 valid. We would find such a characterization interesting.

The numbers $\sum_{i=1}^{k} r_i$ have a direct geometric interpretation. Let $\mathbb{G}_{n,k}(\mathbb{C})$ denote the Grassmannian manifold of k dimensional vector subspaces in $\mathbb{C}^n$, $A|G_{n,k}$ the restriction of A to the subspace $G_{n,k}$ and ν the natural unitarily invariant probability measure on $\mathbb{G}_{n,k}(\mathbb{C})$.

Theorem 3. — *If μ is a unitarily invariant probability measure on $\mathbb{GL}_n(\mathbb{C})$ satisfying $(*)$ then,*

$$\sum_{i=1}^{k} r_i = \int_{A\in\mathbb{GL}_n(\mathbb{C})} \int_{G_{n,k}\in\mathbb{G}_{n,k}(\mathbb{C})} \log|\operatorname{Det}(A|G_{n,k})| d\nu(G_{n,k}) d\mu(A).$$

We may then restate Theorem 1 in the form we prove it.

Theorem 4. — *If μ is a unitarily invariant probability measure on $\mathbb{GL}_n(\mathbb{C})$ satisfying $(*)$ then, for $k = 1, \ldots, n$*

$$\int_{A\in\mathbb{GL}_n(\mathbb{C})} \sum_{i=1}^{k} \log|\lambda_i(A)| d\mu(A)$$
$$\geqslant \int_{A\in\mathbb{GL}_n(\mathbb{C})} \int_{G_{n,k}\in\mathbb{G}_{n,k}(\mathbb{C})} \log|\operatorname{Det}(A|G_{n,k})| d\nu(G_{n,k}) d\mu(A).$$

There is a considerable literature on random Lyapunov exponents and quite general criteria which guarantee that they are non-zero and even distinct. According to Bougerol and Lacroix in 1985 in [**2**] "The subject matter initiated by Bellman was fully developed by Furstenberg, Guivarc'h, Kesten, Le Page and Raugi." We refer to [**2**] for references prior to 1985 and to three others: Gol'dsheid and Margulis [**4**], Guivarc'h and Raugi [**5**] and Ledrappier [**6**].

Our interest in Theorem 1 and Theorem 4 was motivated by some questions in dynamical systems theory, see Burns, Pugh, Shub and Wilkinson [**3**]. Theorem 1 for $k=1$, the orthogonal group and $\mathbb{GL}_n(\mathbb{R})$ was raised there.

We also get a version of Theorem 4 without the logarithms.

Theorem 5. — *Let μ be a unitarily invariant probability measure on $\mathbb{GL}_n(\mathbb{C})$ satisfying $(*)$ and $1 \leqslant k \leqslant n$. Then*

$$\int_{A\in\mathbb{GL}_n(\mathbb{C})} \prod_{i=1}^{k} |\lambda_i(A)| d\mu(A) \geqslant \int_{A\in\mathbb{GL}_n(\mathbb{C})} \int_{G_{n,k}\in\mathbb{G}_{n,k}(\mathbb{C})} |\operatorname{Det}(A|G_{n,k})| d\nu(G_{n,k}) d\mu(A).$$

There is a special case of Theorems 4 and 5 that is good to keep in mind. Our proof relies it.

Let $A \in \mathbb{GL}_n(\mathbb{C})$ and μ be the Haar measure on $\mathbb{U}_n(\mathbb{C})$ (the unitary subgroup of $\mathbb{GL}_n(\mathbb{C})$) normalized to be a probability measure. In this case Theorem 5 becomes:

Theorem 6. — *Let $A \in \mathbb{GL}_n(\mathbb{C})$. Then, for $1 \leqslant k \leqslant n$,*

$$\int_{U\in\mathbb{U}_n(\mathbb{C})} \sum_{i=1}^{k} \log|\lambda_i(UA)| d\mu(U) \geqslant \int_{G_{n,k}\in\mathbb{G}_{n,k}(\mathbb{C})} \log|\operatorname{Det}(A|G_{n,k})| d\nu(G_{n,k})$$

and

$$\int_{U\in\mathbb{U}_n(\mathbb{C})} \prod_{i=1}^{k} |\lambda_i(UA)| d\mu(U) \geqslant \int_{G_{n,k}\in\mathbb{G}_{n,k}(\mathbb{C})} |\operatorname{Det}(A|G_{n,k})| d\nu(G_{n,k}).$$

When $k=1$, $|\lambda_1(UA)| = \rho(UA)$ is the spectral radius of UA. The Grassmannian manifold is identical to the complex projective space $\mathbb{P}_{n-1}(\mathbb{C})$. Integration on this manifold can be reduced to the unit sphere $\mathbb{S}^{2n-1}$ in $\mathbb{R}^{2n}$ so that

Corollary 7. — *Let $A \in \mathbb{GL}_n(\mathbb{C})$. Then*

$$\int_{U\in\mathbb{U}_n(\mathbb{C})} \log|\rho(UA)| d\mu(U) \geqslant \int_{x\in\mathbb{S}^{2n-1}} \log\|Ax\| d\nu(x)$$

and

$$\int_{U\in\mathbb{U}_n(\mathbb{C})} |\rho(UA)| d\mu(U) \geqslant \int_{x\in\mathbb{S}^{2n-1}} \|Ax\| d\nu(x).$$

We expect a similar result for orthogonally invariant probability measures on $\mathbb{GL}_n(\mathbb{R})$ but we have not proven it. Here we content ourselves with the case $n=2$.

Theorem 8. — *Let μ be a probability measure on $\mathbb{GL}_2(\mathbb{R})$ satisfying*

$$g \in \mathbb{GL}_2(\mathbb{R}) \to \log^+(\|g\|) \text{ and } \log^+(\|g^{-1}\|) \text{ are } \mu\text{-integrable.}$$

(a) *If μ is a $\mathbb{SO}_2(\mathbb{R})$ invariant measure on $\mathbb{GL}_2^+(\mathbb{R})$ then,*

$$\int_{A\in\mathbb{GL}_2^+(\mathbb{R})} \log|\lambda_1(A)|d\mu(A) = \int_{A\in\mathbb{GL}_2^+(\mathbb{R})}\int_{x\in\mathbb{S}^1} \log\|Ax\|d\mathbb{S}^1(x)d\mu(A).$$

(b) *If μ is a $\mathbb{SO}_2(\mathbb{R})$ invariant measure on $\mathbb{GL}_2^-(\mathbb{R})$, whose support is not contained in $\mathbb{RO}_2(\mathbb{R})$ i.e. in the set of scalar multiples of orthogonal matrices, then*

$$\int_{A\in\mathbb{GL}_2^-(\mathbb{R})} \log|\lambda_1(A)|d\mu(A) > \int_{A\in\mathbb{GL}_2^-(\mathbb{R})}\int_{x\in\mathbb{S}^1} \log\|Ax\|d\mathbb{S}^1(x)d\mu(A).$$

Here $\mathbb{GL}_2^+(\mathbb{R})$ (resp. $\mathbb{GL}_2^-(\mathbb{R})$) is the set of invertible matrices with positive (resp. negative) determinant. Theorem 8 is proved in section 5.

2. A More General Theorem

Theorem 4 is actually a special case of the much more general Theorem 11 below. Before we state Theorem 11 we need some preliminaries.

A flag F in $\mathbb{C}^n$ is a sequence of vector subspaces of $\mathbb{C}^n$: $F = (F_1, F_2, \ldots, F_n)$, with $F_i \subset F_{i+1}$ and Dim $F_i = i$. The space of flags is called the flag manifold and we denote it by $\mathbb{F}_n(\mathbb{C})$. Now it is easy to see that $\mathbb{F}_n(\mathbb{C})$ may be represented by $\mathbb{GL}_n(\mathbb{C})/\mathbb{R}_n(\mathbb{C})$ or by $\mathbb{U}_n(\mathbb{C})/\mathbb{T}^n(\mathbb{C})$, where $\mathbb{R}_n(\mathbb{C})$ is the subgroup of $\mathbb{GL}_n(\mathbb{C})$ of upper triangular matrices and $\mathbb{T}^n(\mathbb{C})$ is the subgroup of $\mathbb{GL}_n(\mathbb{C})$ consisting of diagonal matrices with complex numbers of modulus 1, so $\mathbb{T}^n(\mathbb{C}) = \mathbb{U}_n(\mathbb{C}) \cap \mathbb{R}_n(\mathbb{C})$. Regarding $\mathbb{F}_n(\mathbb{C})$ as $\mathbb{U}_n(\mathbb{C})/\mathbb{T}^n(\mathbb{C})$ we see that $\mathbb{F}_n(\mathbb{C})$ has a natural $\mathbb{U}_n(\mathbb{C})$-invariant probability measure.

An invertible linear map $A : \mathbb{C}^n \to \mathbb{C}^n$ naturally induces a map $A_\sharp$ on flags by

$$A_\sharp(F_1, F_2, \ldots, F_n) = (AF_1, AF_2, \ldots, AF_n).$$

The flag manifold and the action of a linear map A on $\mathbb{F}_n(\mathbb{C})$ is closely related to the QR algorithm, see Shub and Vasquez [**9**] for a discussion of this. In particular if F is a fixed flag for A i.e. $A_\sharp F = F$, then A is upper triangular in a basis corresponding to the flag F, with the eigenvalues of A appearing on the diagonal in some order: $\lambda_1(A, F), \ldots, \lambda_n(A, F)$.

Let

$$\mathbb{G} = \{A \in \mathbb{GL}_n(\mathbb{C}) \ : \ |\lambda_1(A)| > |\lambda_2(A)| > \cdots > |\lambda_n(A)|\}.$$

Then, there is a unique flag F such that $A_\sharp(F) = F$ and such that $\lambda_i(A, F) = \lambda_i(A)$ for $i = 1, \ldots, n$. We call this flag the QR flag of A and let $QR : \mathbb{G} \to \mathbb{F}_n(\mathbb{C})$ be the map which associates to $A \in \mathbb{G}$ its QR flag. It follows from Shub-Vasquez [**9**] and the discussion of fixed point manifolds below that QR is a smooth mapping.

Now fix $A \in \mathbb{GL}_n(\mathbb{C})$, define $\mathbb{U}_n(\mathbb{C})A = \{UA \ : \ U \in \mathbb{U}_n(\mathbb{C})\}$ and consider $\mathbb{G}_A = \mathbb{G}\cap(\mathbb{U}_n(\mathbb{C})A)$. Assume that $\mathbb{G}_A \neq \varnothing$. If we restrict QR to $\mathbb{G}_A$ then $QR : \mathbb{G}_A \to \mathbb{F}_n(\mathbb{C})$

is in fact a locally trivial fibration whose fibers are the orbits of a $\mathbb{T}^n(\mathbb{C})$ action we now describe.

Let $D \in \mathbb{T}^n(\mathbb{C})$ and $U \in \mathbb{U}_n(\mathbb{C})$ and let $QR(UA) = U_1\mathbb{R}_n(\mathbb{C})$ where $U_1 \in \mathbb{U}_n(\mathbb{C})$. Let

$$\Phi_A : \mathbb{T}^n(\mathbb{C}) \times \mathbb{G}_A \to \mathbb{G}_A$$

be defined by $\Phi_A(D, UA) = U_1 D U_1^{-1} UA$. In section 4 we establish

Proposition 1

(1) $\Phi_A(D, UA)$ *is well defined.*

(2) $QR(\Phi_A(D, UA)) = QR(UA)$.

(3) $\Phi_A : \mathbb{T}^n(\mathbb{C}) \times \mathbb{G}_A \to \mathbb{G}_A$ *is an action of* $\mathbb{T}^n(\mathbb{C})$ *on* $\mathbb{G}_A$ *whose orbits are the fibers of* $QR : \mathbb{G}_A \to \mathbb{F}_n(\mathbb{C})$.

(4) *If* $D = Diag(d_1, \dots, d_n)$ *then* $\lambda_i(\Phi_A(D, UA)) = d_i \lambda_i(UA)$ *and in particular* $|\lambda_i|$ *is constant on the fibers of* $QR : \mathbb{G}_A \to \mathbb{F}_n(\mathbb{C})$ *for* $i = 1, \dots, n$.

Let

$$\mathbb{V}_A = \{(U, F) \in \mathbb{U}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \ : \ (UA)_\sharp F = F\}.$$

We denote by Π_1 and Π_2 the restrictions to $\mathbb{V}_A$ of the projections $\mathbb{U}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \to \mathbb{U}_n(\mathbb{C})$ and $\mathbb{U}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \to \mathbb{F}_n(\mathbb{C})$. We define an action of $\mathbb{T}^n(\mathbb{C})$ on $\mathbb{V}_A$ denoted $\Psi_A : \mathbb{T}^n(\mathbb{C}) \times \mathbb{V}_A \to \mathbb{V}_A$ by

$$\Psi_A(D)(U, U_1\mathbb{T}^n(\mathbb{C})) = (U_1 D U_1^{-1} U, U_1 \mathbb{T}^n(\mathbb{C})).$$

Proposition 2

(1) Ψ_A *is well defined and smooth.*

(2) *The orbits of* Ψ_A *are the fibers of* $\Pi_2 : \mathbb{V}_A \to \mathbb{F}_n(\mathbb{C})$.

We consider the manifold

$$\mathbb{V} = \{(A, F) \in \mathbb{GL}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \ : \ A_\sharp F = F\}$$

and the restrictions to $\mathbb{V}$ of the two projections $\mathbb{GL}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \to \mathbb{GL}_n(\mathbb{C})$ and $\mathbb{GL}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \to \mathbb{F}_n(\mathbb{C})$ which we again denote by Π_1 and Π_2. By the Jordan Canonical Form Theorem the map Π_1 is surjective. Except on a set of positive codimension, ${\Pi_1}^{-1}(A)$ consists of $n!$ points corresponding to the permutations of the eigenspaces of $A \in \mathbb{GL}_n(\mathbb{C})$. The fibers of the map Π_2 are more complicated.

For $c \in \mathbb{C} \setminus \{0\}$ we write $c\mathbb{U}_n(\mathbb{C})$ to mean $\{cU \ : \ U \in \mathbb{U}_n(\mathbb{C})\}$.

Definition 9. — Let $f : \mathbb{GL}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \to \mathbb{R}$ be continuous.

(1) f is $\mathbb{U}_n(\mathbb{C})$ or unitarily invariant if $f(UA, F) = f(A, F)$ for all $(A, F) \in \mathbb{GL}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C})$ and $U \in \mathbb{U}_n(\mathbb{C})$, and if $f|c\mathbb{U}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C})$ is constant for every $c \in \mathbb{C} \setminus \{0\}$.

(2) For $B \in \mathbb{GL}_n(\mathbb{C})$ let $g(B) = \max_{(B,F)\in\mathbb{V}} f(B, F)$. We say that f is $\mathbb{T}^n(\mathbb{C})$ or torally invariant if $g(\Phi_A(D, B)) = g(B)$ for all $A \in \mathbb{G}$, $B \in \mathbb{G}_A$ and $D \in \mathbb{T}^n(\mathbb{C})$.

Examples of $\mathbb{U}_n(\mathbb{C})$ and $\mathbb{T}^n(\mathbb{C})$ invariant functions are

(1) For $1 \leqslant k \leqslant n$ let $f_k(A, F) = |\operatorname{Det}(A|F_k)|$ where $F = (F_1, F_2, \ldots, F_n) \in \mathbb{F}_n(\mathbb{C})$.

(2) $\log f_k(A, F)$ where $f_k(A, F)$ is as in 1).

Remark 10. — If $A_\sharp F = F$ then $|\operatorname{Det}(A|F_k)| = \prod_{i=1}^k |\lambda_i(A, F)|$.

Given a continuous $f : \mathbb{GL}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \to \mathbb{R}$, let $g : \mathbb{GL}_n(\mathbb{C}) \to \mathbb{R}$ be defined by $g(B) = \max_{(B,F)\in\mathbb{V}} f(B, F)$.

Theorem 11. — *Let $f : \mathbb{GL}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \to \mathbb{R}$ be continuous, unitarily and torally invariant. Let μ be a unitarily invariant probability measure on $\mathbb{GL}_n(\mathbb{C})$ satisfying $(*)$. If f is $\mu \otimes \nu$-integrable then g is μ-integrable and*

$$\int_{A\in\mathbb{GL}_n(\mathbb{C})} g(A)d\mu(A) \geqslant \int_{A\in\mathbb{GL}_n(\mathbb{C})} \int_{F\in\mathbb{F}_n(\mathbb{C})} f(A, F)d\nu(F)d\mu(A).$$

It is now fairly simple to see how Theorem 11 implies Theorem 4. If $f_k(A, F) = \log|\operatorname{Det}(A|F_k)|$ then, by Remark 10, $g(A) = \sum_{i=1}^k \log|\lambda_i(A)|$ where $|\lambda_1(A)| \geqslant |\lambda_2(A)| \geqslant \cdots \geqslant |\lambda_n(A)|$ are the absolute values of the eigenvalues of A. So the left hand integrals in Theorem 4 and 11 are the same. To see that the right hand integrals are the same consider the natural fibration $\Pi_k : \mathbb{F}_n(\mathbb{C}) \to \mathbb{G}_{n,k}(\mathbb{C})$ given by $\Pi_k(F_1, \ldots, F_n) = F_k$. Then $|\operatorname{Det}(A|\Pi_k F)| = |\operatorname{Det}(A|F_k)|$ and it is easy to see that

$$\int_{F\in\mathbb{F}_n(\mathbb{C})} \log|\operatorname{Det}(A|F_k)|d\nu(F) = \int_{G_{n,k}\in\mathbb{G}_{n,k}(\mathbb{C})} \log|\operatorname{Det}(A|G_{n,k})|d\nu(G_{n,k}).$$

We will say more about this in section 4. So we are done.

We now turn to the proof of Theorem 11 which follows from the consideration of a special case.

Let $A \in \mathbb{GL}_n(\mathbb{C})$. We put Haar measure μ on $\mathbb{U}_n(\mathbb{C})$ normalized to be a probability measure. Thus the next proposition is a special case of Theorem 11.

Proposition 3. — *Let $f : \mathbb{GL}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \to \mathbb{R}$ be continuous, unitarily and torally invariant. Let*

$$\mathbb{V}_A = \{(U, F) \in \mathbb{U}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \ : \ (UA)_\sharp F = F\}$$

and $g(B) = \max_{(B,F)\in\mathbb{V}_A} f(B, F)$. If f is $\mu \otimes \nu$-integrable then g is μ-integrable and

$$\int_{U\in\mathbb{U}_n(\mathbb{C})} g(UA)d\mu(U) \geqslant \int_{U\in\mathbb{U}_n(\mathbb{C})} \int_{F\in\mathbb{F}_n(\mathbb{C})} f(UA, F)d\nu(F)d\mu(U).$$

We now see that Proposition 3 implies Theorem 11. Disintegrate the measure μ of Theorem 11 along the orbits of $\mathbb{U}_n(\mathbb{C})$ obtaining $\mathbb{U}_n(\mathbb{C})$ invariant probability measures on each orbit. Identifying an orbit with $\mathbb{U}_n(\mathbb{C})$ we see that these measures are left invariant on $\mathbb{U}_n(\mathbb{C})$ hence they are Haar measures. Now Proposition 3 applies orbit by orbit. Integrating the inequality over the space of orbits proves Theorem 11.

Note that it is sufficient to prove Proposition 3 when A is not a constant times a unitary matrix, for otherwise $g(UA)$ and $f(UA,F)$ are both equal to the constant in the definition of unitary invariance. Thus the integrals are equal since they are equal to this constant. We will assume below that A is not a constant times a unitary matrix i.e. A is not conformal.

Note that in Proposition 3 the right hand integral does not depend on U since f is unitarily invariant. Thus it is not necessary to integrate over $\mathbb{U}_n(\mathbb{C})$, the first integral is constant.

Now we restate Proposition 3 in its simpler form.

Proposition 4. — *Let $f : \mathbb{F}_n(\mathbb{C}) \to \mathbb{R}$ be continuous and torally invariant, suppose A is not unitary or a scalar times a unitary. Let*

$$\mathbb{V}_A = \{(U,F) \in \mathbb{U}_n(\mathbb{C}) \times \mathbb{F}_n(\mathbb{C}) \;:\; (UA)_\sharp F = F\}.$$

Let $g(B) = \max_{(B,F)\in\mathbb{V}_A} f(F)$. Then

$$\int_{U\in\mathbb{U}_n(\mathbb{C})} g(U)d\mu(U) \geqslant \int_{F\in\mathbb{F}_n(\mathbb{C})} f(F)d\nu(F).$$

Now we outline the proof of Proposition 4. We use the diagram

$$\begin{tikzcd} & \mathbb{V}_A \arrow[dl, "\Pi_1"'] \arrow[dr, "\Pi_2"] & \\ \mathbb{U}_n(\mathbb{C}) & & \mathbb{F}_n(\mathbb{C}) \end{tikzcd}$$

to transfer the right hand integral over $\mathbb{F}_n(\mathbb{C})$ to an integral over $\mathbb{U}_n(\mathbb{C})$. First we identify a subset of $\mathbb{U}_n(\mathbb{C})$ over which we will integrate.

Let $\mathbb{G}_1$ be the open subset of $\mathbb{U}_n(\mathbb{C})$ consisting of those U such that the eigenvalues of UA are of distinct modulus. In this case we write them as

$$\lambda_i = \lambda_i(UA), \quad 1 \leqslant i \leqslant n,$$

where $|\lambda_1| > \cdots > |\lambda_n|$.

Proposition 5. — *$\mathbb{G}_1$ is an open set of full measure in $\mathbb{U}_n(\mathbb{C})$, i.e. $\mu(\mathbb{G}_1) = 1$.*

Lemma 1. — *Let $f : \mathbb{F}_n(\mathbb{C}) \to \mathbb{R}$ be continuous and torally invariant. Let $g(B) = \max_{(B,F)\in\mathbb{V}_A} f(F)$. Then*

$$\begin{aligned} \int_{F\in\mathbb{F}_n(\mathbb{C})} f(F)d\nu(F) &= \int_{U\in\mathbb{G}_1} \sum_{(U,F)\in\mathbb{V}_A} f(F) \prod_{j<i} \left|1 - \frac{\lambda_i(UA,F)}{\lambda_j(UA,F)}\right|^{-2} d\mu(U) \\ &\leqslant \int_{U\in\mathbb{G}_1} g(U) \sum_{\sigma\in\Sigma_n} \prod_{j<i} \left|1 - \frac{\lambda_{\sigma(i)}}{\lambda_{\sigma(j)}}\right|^{-2} d\mu(U) \end{aligned}$$

with Σ_n the group of permutations over the set $\{1, 2, \ldots, n\}$.

Proposition 5 and Lemma 1 are proved in section 4. Proposition 3 and 4 follow from Proposition 5, Lemma 1 and from the next proposition.

Proposition 6

$$\int_{U\in\mathbb{G}_1} g(U)\sum_{\sigma\in\Sigma_n}\prod_{j<i}\left|1-\frac{\lambda_{\sigma(i)}}{\lambda_{\sigma(j)}}\right|^{-2} d\mu(U) = \int_{U\in\mathbb{G}_1} g(U)d\mu(U).$$

We will prove Proposition 6 in Section 4 by decomposing the two integrals along the fibers of the QR fibration on which $g(U)$ is constant.

Proposition 7. — *The normal Jacobian of the QR fibration is* $\prod_{j<i}|1-\lambda_i/\lambda_j|^{-2}$ *where* $\lambda_i = \lambda_i(UA)$ *are the eigenvalues of* UA *with* $|\lambda_1| > \cdots > |\lambda_n|$. *Hence*

$$\int_{U\in\mathbb{G}_1} g(U)\sum_{\sigma\in\Sigma_n}\prod_{j<i}\left|1-\frac{\lambda_{\sigma(i)}}{\lambda_{\sigma(j)}}\right|^{-2} d\mu(U)$$
$$= \int_{F\in\mathbb{F}_n(\mathbb{C})} g(U)\int_{U\in QR^{-1}(F)}\sum_{\sigma\in\Sigma_n}\prod_{j<i}\left|\frac{1-\lambda_{\sigma(i)}/\lambda_{\sigma(j)}}{1-\lambda_i/\lambda_j}\right|^{-2} d\mu(QR^{-1}(F))(U)d\nu(F)$$

and

$$\int_{U\in\mathbb{G}_1} g(U)d\mu(U) = \int_{F\in\mathbb{F}_n(\mathbb{C})} g(U)\int_{U\in QR^{-1}(F)}\prod_{j<i}\left|1-\frac{\lambda_i}{\lambda_j}\right|^{2} d\mu(QR^{-1}(F))(U)d\nu(F).$$

Proposition 7 is proved in Section 4. Finally in Section 4 we prove

Proposition 8

$$\int_{U\in QR^{-1}(F)}\prod_{j<i}\left|1-\frac{\lambda_i}{\lambda_j}\right|^{2} d\mu(QR^{-1}(F))(U)$$
$$= \int_{U\in QR^{-1}(F)}\sum_{\sigma\in\Sigma_n}\prod_{j<i}\left|\frac{1-\lambda_{\sigma(i)}/\lambda_{\sigma(j)}}{1-\lambda_i/\lambda_j}\right|^{-2} d\mu(QR^{-1}(F))(U).$$

Now Proposition 7 and Proposition 8 prove Proposition 6 and we are done. To summarize it remains to prove Theorem 3, Proposition 1, Proposition 2, Proposition 5, Lemma 1, Proposition 7 and Proposition 8.

3. Manifolds of fixed points

The manifolds $\mathbb{V}$ and $\mathbb{V}_A$ are manifolds of fixed points. In this section we discuss integration formulas for manifolds of fixed points and prove Lemma 1 and Proposition 7. We begin by recalling the co-area formula.

3.1. The Co-area Formula.— Let $\mathbb{X}$ and $\mathbb{Y}$ be real Riemannian manifolds. We denote by $d\mathbb{X}$ and $d\mathbb{Y}$ the associated volume forms. Suppose $F : \mathbb{X} \to \mathbb{Y}$ is a smooth surjective map and suppose that the derivative $DF(x) : T_x\mathbb{X} \to T_{f(x)}\mathbb{Y}$ is surjective for almost all $x \in \mathbb{X}$. The horizontal space H_x of $T_x\mathbb{X}$ is defined as the orthogonal complement to Ker $DF(x)$. The horizontal derivative of F at x is the restriction of $DF(x)$ to H_x. The normal Jacobian $NJ(F(x))$ is the absolute value of the determinant of the horizontal derivative defined almost everywhere on X:

$$NJ(F(x)) = |\operatorname{Det}\ (DF(x)|_{H_x})\,|.$$

The map F defines a fibration of $\mathbb{X}$ with base $\mathbb{Y}$ and fibers $F^{-1}(y)$, $y \in \mathbb{Y}$. Integration over X with respect to this fibration generalizes Fubini's formula:

Theorem 12 **(Co-area Formula)**. — *Let $F : \mathbb{X} \to \mathbb{Y}$ be a smooth map of real Riemannian manifolds satisfying the preceeding surjectivity conditions. Then, for any integrable $f : \mathbb{X} \to \mathbb{R}$*

$$\int_{x\in\mathbb{X}} f(x)d\mathbb{X}(x) = \int_{y\in\mathbb{Y}}\int_{x\in F^{-1}(y)} \frac{f(x)}{NJ(F(x))} dF^{-1}(y)(x)d\mathbb{Y}(y).$$

Remark 13. — In the co-area formula, $d\mathbb{X}$ and $d\mathbb{Y}$ are the volume forms associated with the Riemannian structures over $\mathbb{X}$ and $\mathbb{Y}$, $dF^{-1}(y)$ is the volume form on $F^{-1}(y)$ equipped with the induced metric.

Remark 14. — The co-area formula also extends to complex Riemannian manifolds. In that case the normal jacobian is equal to

$$NJ(F(x)) = |\operatorname{Det}\ (DF(x)|_{H_x})\,|^2.$$

This follows immediately from the fact that if $A : \mathbb{C}^n \to \mathbb{C}^n$ is a complex linear map and $A_{\mathbb{R}} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ the real map it defines, then

$$|\operatorname{Det}\ A_{\mathbb{R}}| = |\operatorname{Det}\ A|^2.$$

Remark 15. — When $DF(x) : T_x\mathbb{X} \to T_{f(x)}\mathbb{Y}$ is onto, the normal Jacobian is equal to

$$NJ(F(x)) = (\operatorname{Det}\ DF(x)DF(x)^*)^{1/2}$$

so that

$$\int_{x\in\mathbb{X}} f(x)d\mathbb{X}(x) = \int_{y\in\mathbb{Y}}\int_{x\in F^{-1}(y)} \frac{f(x)}{(\operatorname{Det}\ DF(x)DF(x)^*)^{1/2}} dF^{-1}(y)(x)d\mathbb{Y}(y)$$

and in the complex case (see Remark 14)

$$\int_{x\in\mathbb{X}} f(x)d\mathbb{X}(x) = \int_{y\in\mathbb{Y}}\int_{x\in F^{-1}(y)} \frac{f(x)}{\operatorname{Det}\ DF(x)DF(x)^*} dF^{-1}(y)(x)d\mathbb{Y}(y).$$

Remark 16. — The co-area formula also extends to the case of maps $F : \mathbb{X} \to \mathbb{Y}$ between algebraic varieties by considering the restriction of F to the smooth part of X.

3.2. Manifolds of Fixed Points.— Let $\mathcal{F}$ and $\mathcal{M}$ be compact Riemannian manifolds and a smooth map $\Phi : \mathcal{F} \times \mathcal{M} \to \mathcal{M}$ be given. Let

$$\Psi : \mathcal{F} \times \mathcal{M} \to \mathcal{M} \times \mathcal{M}$$

be defined by $\Psi(f, m) = (\Phi(f, m), m)$. Suppose Ψ is transversal to

$$\Delta = \{(m, m) \ : \ m \in \mathcal{M}\} \subset \mathcal{M} \times \mathcal{M}.$$

Then

$$\mathcal{V} = \Psi^{-1}(\Delta) = \{(f, m) \in \mathcal{F} \times \mathcal{M} \ : \ \Phi(f, m) = m\}$$

is a submanifold in $\mathcal{F} \times \mathcal{M}$. We denote by $\Pi_{\mathcal{F}}$ and $\Pi_{\mathcal{M}}$ the restrictions to $\mathcal{V}$ of the projections $\mathcal{F} \times \mathcal{M} \to \mathcal{F}$ and $\mathcal{F} \times \mathcal{M} \to \mathcal{M}$. By Sard's Theorem, almost all $f \in \mathcal{F}$ are regular values of $\Pi_{\mathcal{F}} : \mathcal{V} \to \mathcal{F}$. For these $f \in \mathcal{F}$ the corresponding fixed points $m \in \mathcal{M}$, i.e. $(f, m) \in \mathcal{V}$, are isolated in $\mathcal{M}$. Since $\mathcal{M}$ is compact these fixed points are finite.

Theorem 17. — *Let $\mathcal{F}_{\pitchfork}$ denote the set of $f \in \mathcal{F}$ which are regular values of $\Pi_{\mathcal{F}}$. Let $G : \mathcal{M} \to \mathbb{R}$ be a continuous function. Then*

$$\int_{m \in \mathcal{M}} G(m) d\mathcal{M}(m) = \int_{f \in \mathcal{F}_{\pitchfork}} \sum_{m \in \Pi_{\mathcal{F}}^{-1}(f)} G(m) \frac{NJ(\Pi_{\mathcal{M}}(f, m))}{\operatorname{Vol} \Pi_{\mathcal{M}}^{-1}(m) NJ(\Pi_{\mathcal{F}}(f, m))} d\mathcal{F}(f).$$

Remark 18. — The integral is taken over the set $\mathcal{F}_{\pitchfork}$ of regular values of $\Pi_{\mathcal{F}}$. We note that $f \in \mathcal{F}_{\pitchfork}$ if and only if for all $m \in \mathcal{M}$, such that $(f, m) \in \mathcal{V}$, $id_{T_m\mathcal{M}} - D_{\mathcal{M}}\Phi(f, m)$ is invertible.

Proof. — We apply the co-area formula to the function

$$\frac{G(m) NJ(\Pi_{\mathcal{M}}(f, m))}{\operatorname{Vol} \Pi_{\mathcal{M}}^{-1}(m)}$$

defined over $\mathcal{V}$ with respect to the projection $\Pi_{\mathcal{M}}$. This gives

$$\begin{aligned}
&\int_{(f,m) \in \mathcal{V}} \frac{G(m) NJ(\Pi_{\mathcal{M}}(f, m))}{\operatorname{Vol} \Pi_{\mathcal{M}}^{-1}(m)} d\mathcal{V}(f, m) \\
&\quad = \int_{m \in \mathcal{M}} \int_{(f,m) \in \Pi_{\mathcal{M}}^{-1}(m)} \frac{G(m) NJ(\Pi_{\mathcal{M}}(f, m))}{\operatorname{Vol} \Pi_{\mathcal{M}}^{-1}(m) NJ(\Pi_{\mathcal{M}}(f, m))} d\Pi_{\mathcal{M}}^{-1}(m)(f, m) d\mathcal{M}(m) \\
&\qquad = \int_{m \in \mathcal{M}} G(m) d\mathcal{M}(m).
\end{aligned}$$

We now apply the same formula to the same function with respect to the the projection $\Pi_{\mathcal{F}}$. We notice that the fiber $\Pi_{\mathcal{F}}^{-1}(f)$ consists in a finite number of fixed points so that:

$$\int_{(f,m)\in\mathcal{V}} \frac{G(m)NJ(\Pi_{\mathcal{M}}(f,m))}{\mathrm{Vol}\,\Pi_{\mathcal{M}}^{-1}(m)} d\mathcal{V}(f,m) = \int_{f\in\mathcal{F}_{\pitchfork}} \sum_{m\in\Pi_{\mathcal{F}}^{-1}(f)} G(m)\frac{NJ(\Pi_{\mathcal{M}}(f,m))}{\mathrm{Vol}\,\Pi_{\mathcal{M}}^{-1}(m)NJ(\Pi_{\mathcal{F}}(f,m))} d\mathcal{F}(f)$$

and we are done. □

Now we compute the normal Jacobians in terms of the partial derivatives of $\Phi : \mathcal{F}\times\mathcal{M} \to \mathcal{M}$. The Riemannian structure we put on $\mathcal{V}$ is the restriction of the product structure on $\mathcal{F}\times\mathcal{M}$.

Lemma 2. — *Let $f \in \mathcal{F}_{\pitchfork}$ and $(f,m) \in \mathcal{V}$. Then the tangent space of $\mathcal{V}$ at (f,m) is*

$$T_{(f,m)}\mathcal{V} = \{(\dot f, \dot m) \in T_f\mathcal{F} \times T_m\mathcal{M} \ : \ \dot m = (id_{T_m\mathcal{M}} - D_{\mathcal{M}}\Phi(f,m))^{-1} D_{\mathcal{F}}\Phi(f,m)\dot f\}.$$

Proof. — This is a consequence of Remark 18. □

If we put together Lemma 2, Theorem 17, and Blum-Cucker-Shub-Smale [**1**] Lemma 3, page 242, we have:

Theorem 19. — *Let $G : \mathcal{M} \to \mathbb{R}$ be a continuous function. Then, for real manifolds*

$$\int_{m\in\mathcal{M}} G(m)d\mathcal{M}(m) = \int_{f\in\mathcal{F}_{\pitchfork}} \sum_{m\in\Pi_{\mathcal{F}}^{-1}(f)} G(m)\frac{|\,\mathrm{Det}\,(D_{\mathcal{F}}\Phi(f,m)D_{\mathcal{F}}\Phi(f,m)^*)|^{1/2}}{\mathrm{Vol}\,\Pi_{\mathcal{M}}^{-1}(m)|\,\mathrm{Det}\,(id_{T_m\mathcal{M}} - D_{\mathcal{M}}\Phi(f,m))|} d\mathcal{F}(f).$$

For complex manifolds this formula becomes

$$\int_{m\in\mathcal{M}} G(m)d\mathcal{M}(m) = \int_{f\in\mathcal{F}_{\pitchfork}} \sum_{m\in\Pi_{\mathcal{F}}^{-1}(f)} G(m)\frac{|\,\mathrm{Det}\,(D_{\mathcal{F}}\Phi(f,m)D_{\mathcal{F}}\Phi(f,m)^*)|}{\mathrm{Vol}\,\Pi_{\mathcal{M}}^{-1}(m)|\,\mathrm{Det}\,(id_{T_m\mathcal{M}} - D_{\mathcal{M}}\Phi(f,m))|^2} d\mathcal{F}(f).$$

Similarly we may also evaluate integrals defined on $\mathcal{F}$ using the fibration over $\mathcal{M}$. Suppose that $S : \mathcal{F}_{\pitchfork} \to \mathcal{V}$ is a smooth section of $\mathcal{V}$ defined on $\mathcal{F}_{\pitchfork}$ or on an open set of $\mathcal{F}_{\pitchfork}$ i.e. $\Pi_{\mathcal{F}}S = id_{\mathcal{F}_{\pitchfork}}$.

Theorem 20. — *Let $H : \mathcal{F}_{\pitchfork} \to \mathbb{R}$ be an integrable function defined on $\mathcal{F}_{\pitchfork}$ or on an open set in $\mathcal{F}_{\pitchfork}$. Then, for real manifolds*

$$\int_{f\in\mathcal{F}_{\pitchfork}} H(f)d\mathcal{F}(f) = \int_{m\in\mathcal{M}} \int_{(\Pi_{\mathcal{M}}S)^{-1}(m)} H(f)\frac{|\,\mathrm{Det}\,(id_{T_m\mathcal{M}} - D_{\mathcal{M}}\Phi(f,m))|}{|\,\mathrm{Det}\,(D_{\mathcal{F}}\Phi(f,m)D_{\mathcal{F}}\Phi(f,m)^*)|^{1/2}} d\mathcal{F}(f)$$

and for complex manifolds

$$\int_{f\in\mathcal{F}_{\mathfrak{m}}} H(f)d\mathcal{F}(f)$$
$$= \int_{m\in\mathcal{M}}\int_{(\Pi_{\mathcal{M}}S)^{-1}(m)} H(f)\frac{|\operatorname{Det}(id_{T_m\mathcal{M}} - D_{\mathcal{M}}\Phi(f,m))|^2}{|\operatorname{Det}(D_{\mathcal{F}}\Phi(f,m)D_{\mathcal{F}}\Phi(f,m)^*)|}d\mathcal{F}(f).$$

4. Proofs of Theorem 3, Propositions 1, 2, 5, Lemma 1 and of Propositions 7 and 8

4.1. Proof of Theorem 3.— If not explicitely stated this Theorem is inherent in the works of Furstenberg, Guivarc'h, Raugi, Gol'dsheid, Margulis and possibly other sources. See also Bougerol-Lacroix. We sketch a proof.

We consider two auxilliary spaces and maps:

(1) $\prod_{i=1}^{\infty}\mathbb{GL}_n(\mathbb{C})$ equipped with the product measure $\widehat{\mu}$, and

$$\sigma : \prod_{i=1}^{\infty}\mathbb{GL}_n(\mathbb{C}) \hookleftarrow$$

the one sided shift:

$$\sigma(\ldots g_p\ldots g_1) = (\ldots g_p \ldots g_2).$$

(2) $\prod_{i=1}^{\infty}\mathbb{GL}_n(\mathbb{C})\times\mathbb{G}_{n,k}(\mathbb{C})$ with the measure $\widehat{\mu}\times\nu$ and the map

$$\tau : \prod_{i=1}^{\infty}\mathbb{GL}_n(\mathbb{C})\times\mathbb{G}_{n,k}(\mathbb{C}) \hookleftarrow$$

defined by

$$\tau((\ldots g_p\ldots g_1), G_{n,k}) = (\sigma(\ldots g_p\ldots g_1), g_1(G_{n,k})).$$

$\widehat{\mu}$ is invariant and ergodic for σ and $\widehat{\mu}\times\nu$ is invariant for τ (here we use the unitary invariance of μ). It follows from Birkoff's Ergodic Theorem and the invariance of the measure $\widehat{\mu}\times\nu$ for the map τ that $\lim\frac{1}{p}\log|\operatorname{Det}(g_p\ldots g_1|G_{n,k})|$ exists a.e. in $\prod_{i=1}^{\infty}\mathbb{GL}_n(\mathbb{C})\times\mathbb{G}_{n,k}(\mathbb{C})$, and the integral of $\lim\frac{1}{p}\log|\operatorname{Det}(g_p\ldots g_1|G_{n,k})|$ equals

$$\int_{A\in\mathbb{GL}_n(\mathbb{C})}\int_{G_{n,k}\in\mathbb{G}_{n,k}(\mathbb{C})}\log|\operatorname{Det}(A|G_{n,k})|d\nu(G_{n,k})d\mu(A).$$

Now by Oseledec's theorem for almost all $\widehat{g} = (\ldots g_p\ldots g_1)$ the limit

$$\lim\frac{1}{p}\log|\operatorname{Det}(g_p\ldots g_1|G_{n,k})|$$

exists for almost all $G_{n,k}$ and equals $\sum_{i=1}^{k} r_i$. So

$$\sum_{i=1}^{k} r_i = \int_{A\in\mathbb{GL}_n(\mathbb{C})}\int_{G_{n,k}\in\mathbb{G}_{n,k}(\mathbb{C})}\log|\operatorname{Det}(A|G_{n,k})|d\nu(G_{n,k})d\mu(A).$$

4.2. Proofs of Propositions 1 and 2.— We now turn, in section 3, to the case that $\mathcal{F} = \mathbb{U}_n(\mathbb{C})$, $\mathcal{M} = \mathbb{F}_n(\mathbb{C})$, $\mathcal{V} = \mathbb{V}_A$ and $\Phi(U, F) = (UA)_\sharp(F)$.

Lemma 3. — *Suppose $(UA)_\sharp(U_1\mathbb{T}^n(\mathbb{C})) = U_1\mathbb{T}^n(\mathbb{C})$. Then, for any $V \in \mathbb{U}_n(\mathbb{C})$ one has*

$$(VA)_\sharp(U_1\mathbb{T}^n(\mathbb{C})) = U_1\mathbb{T}^n(\mathbb{C})$$

if and only if there exists $D \in \mathbb{T}^n(\mathbb{C})$ such that $U_1DU_1^{-1}U = V$.

Proof. — If $(VA)_\sharp(U_1\mathbb{T}^n(\mathbb{C})) = U_1\mathbb{T}^n(\mathbb{C})$ then

$$U_1\mathbb{R}_n(\mathbb{C}) = VAU_1\mathbb{R}_n(\mathbb{C}) = VU^{-1}UAU_1\mathbb{R}_n(\mathbb{C}) = VU^{-1}U_1\mathbb{R}_n(\mathbb{C}).$$

So $U_1^{-1}UV^{-1}U_1\mathbb{R}_n(\mathbb{C}) = \mathbb{R}_n(\mathbb{C})$ and $U_1^{-1}UV^{-1}U_1$ is in $\mathbb{R}_n(\mathbb{C}) \cap \mathbb{U}_n(\mathbb{C}) = \mathbb{T}^n(\mathbb{C})$. So there is a $D \in \mathbb{T}^n(\mathbb{C})$ with $UV^{-1} = U_1DU_1^{-1}$ and $V = U_1D^{-1}U_1^{-1}U$.

On the other hand for, $D \in \mathbb{T}^n(\mathbb{C})$,

$$U_1DU_1^{-1}UAU_1\mathbb{R}_n(\mathbb{C}) = U_1DU_1^{-1}U_1\mathbb{R}_n(\mathbb{C}) = U_1D\mathbb{R}_n(\mathbb{C}) = U_1\mathbb{R}_n(\mathbb{C}).$$

So we are done. □

Proof of Proposition 1

(1) If $QR(UA) = U_1\mathbb{R}_n(\mathbb{C}) = U_1'\mathbb{R}_n(\mathbb{C})$ then $U_1' = U_1D'$ for some $D' \in \mathbb{T}^n(\mathbb{C})$. Thus

$$U_1'DU_1'^{-1}UA = U_1D'DD'^{-1}U_1^{-1}UA = U_1DU_1^{-1}UA.$$

From $QR(UA) = U_1\mathbb{R}_n(\mathbb{C})$ we get $(UA)_\sharp U_1\mathbb{R}_n(\mathbb{C}) = U_1\mathbb{R}_n(\mathbb{C})$ so that $UA = U_1RU_1^{-1}$ for some $R \in \mathbb{R}_n(\mathbb{C})$. This gives

$$\Phi_A(D, UA) = U_1DU_1^{-1}UA = U_1DU_1^{-1}U_1RU_1^{-1} = U_1DRU_1^{-1}.$$

Thus the eigenvalues of $\Phi_A(D, UA)$ have distinct modulus and Φ_A is well defined.

(2) Using $UA = U_1RU_1^{-1}$ we get

$$\Phi_A(D, UA)U_1 = U_1DU_1^{-1}UAU_1 = U_1DR$$

so that

$$QR(\Phi_A(D, UA)) = QR(UA) = U_1\mathbb{R}_n(\mathbb{C}).$$

(3) This assertion is exactly Lemma 3.

(4) $\lambda_i(\Phi_A(D, UA)) = d_i\lambda_i(UA)$ is proved in (1). and $|\lambda_i|$ constant on the fibers of QR described in (3) and we are done. □

Proof of Proposition 2. — Similar to the proof of Proposition 1. it also uses Lemma 3. □

4.3. Proof of Lemma 1.— Lemma 3 has an immediate consequence:

Lemma 4

(a) *The volume of the fibers* $\Pi_2^{-1}(F)$, *for* $F \in \mathbb{F}_n(\mathbb{C})$, *with* $\Pi_2 : \mathbb{V}_A \to \mathbb{F}_n(\mathbb{C})$, *is constant and equal to* $\operatorname{Vol} \mathbb{T}^n(\mathbb{C})$.

(b) *The volume of the fibers* $QR^{-1}(F)$, *for* $F \in \mathbb{F}_n(\mathbb{C})$, *is constant and equals* $\operatorname{Vol} \mathbb{T}^n(\mathbb{C})$.

Next we turn our attention to the term $|\operatorname{Det} D_{\mathbb{U}_n(\mathbb{C})}\Phi(U,F) D_{\mathbb{U}_n(\mathbb{C})}\Phi(U,F)^*|$. If we fix a flag F then $D_{\mathbb{U}_n(\mathbb{C})}\Phi(U,F) = D_{\mathbb{U}_n(\mathbb{C})}\Phi_F(U)$ where $\Phi_F(U) = UU_1\mathbb{T}^n(\mathbb{C})$ and U_1 defined by $A_\sharp F = U_1\mathbb{T}^n(\mathbb{C})$. Next we prove that the normal Jacobian of $\Phi_F(U)$ is constant.

Proposition 9. — *Let* $\mathbb{U}_n(\mathbb{C})$ *act on* $\mathbb{U}_n(\mathbb{C})/\mathbb{T}^n(\mathbb{C})$ *by* $\Phi_F(U) = UU_1\mathbb{T}^n(\mathbb{C})$. *Then the normal jacobian of* $\Phi_F(U)$ *is independent of* F, U_1 *and* U *and equals* $\operatorname{Vol} \mathbb{T}^n(\mathbb{C})$.

Proof. — First consider the case $U_1 = I_n$. Then $\Phi_F(U) = U\mathbb{T}^n(\mathbb{C})$ is the projection from $\mathbb{U}_n(\mathbb{C})$ to $\mathbb{U}_n(\mathbb{C})/\mathbb{T}^n(\mathbb{C})$. Before normalizing the Riemannian metric on $\mathbb{U}_n(\mathbb{C})/\mathbb{T}^n(\mathbb{C})$ to make the volume 1, the normal to the fiber is mapped isometrically to the tangent space of $\mathbb{U}_n(\mathbb{C})/\mathbb{T}^n(\mathbb{C})$. Now $R_{U_1} : \mathbb{U}_n(\mathbb{C}) \to \mathbb{U}_n(\mathbb{C})$ defined by $R_{U_1}(U) = UU_1$ is an isometry of $\mathbb{U}_n(\mathbb{C})$ and the fibers of Φ_F are the reciprocal images by R_{U_1} of the fibers of Φ_{I_n}. So the normal jacobians are constant. After normalization, the normal jacobians must equal $\operatorname{Vol} \mathbb{T}^n(\mathbb{C})$ to make $\operatorname{Vol} \mathbb{U}_n(\mathbb{C})$ equal 1. □

Corollary 21. — $|\operatorname{Det} D_{\mathbb{U}_n(\mathbb{C})}\Phi(U,F) D_{\mathbb{U}_n(\mathbb{C})}\Phi(U,F)^*| = \operatorname{Vol} \mathbb{T}^n(\mathbb{C})$ *for any* $F \in \mathbb{F}_n(\mathbb{C})$ *and* $U \in \mathbb{U}_n(\mathbb{C})$.

Proof. — By Remark 15 $|\operatorname{Det} D_{\mathbb{U}_n(\mathbb{C})}\Phi(U,F) D_{\mathbb{U}_n(\mathbb{C})}\Phi(U,F)^*|$ is equal to the normalized Jacobian of $\Phi_F(U)$ and we apply Proposition 9. □

Finally we have from Lemma 4 of Shub-Vasquez [**9**]

Proposition 10. — $|\operatorname{Det}(id - D_{\mathbb{F}_n(\mathbb{C})}\Phi(U,F)| = \prod_{j<i} \left|1 - \dfrac{\lambda_{\sigma(i)}}{\lambda_{\sigma(j)}}\right|$ *where* $\lambda_{\sigma(i)} = \lambda_i(UA, F)$ *and* $|\lambda_1| > \cdots > |\lambda_n|$.

Making the substitutions in Theorem 19 given by Corollary 21 and Proposition 10 we have

Theorem 22. — *Let* $f : \mathbb{F}_n(\mathbb{C}) \to \mathbb{R}$ *be continuous. Then*

$$\int_{F \in \mathbb{F}_n(\mathbb{C})} f(F)\, d\nu(F) = \int_{U \in \mathbb{G}_1} \sum_{(U,F) \in \Pi^{-1}_{\mathbb{U}_n(\mathbb{C})}} f(F) \prod_{j<i} \left|1 - \frac{\lambda_{\sigma(i)}}{\lambda_{\sigma(j)}}\right|^{-2} d\mu(U).$$

This proves Lemma 1.

4.4. Proof of Proposition 7.— Similarly substituting in Theorem 20 gives

Theorem 23. — *Let $g : \mathbb{G}_1 \to \mathbb{R}$ be integrable. Then*

$$\int_{U\in\mathbb{G}_1} g(U)d\mu(U)$$
$$= \int_{F\in\mathbb{F}_n(\mathbb{C})} \int_{(U,F)\in\Pi^{-1}_{\mathbb{F}_n(\mathbb{C})}(F)} g(U) \prod_{j<i} \left|1 - \frac{\lambda_{\sigma(i)}}{\lambda_{\sigma(j)}}\right|^2 d\Pi^{-1}_{\mathbb{F}_n(\mathbb{C})}(F)(U)d\nu(F).$$

This theorem proves Proposition 7.

4.5. Proof of Proposition 8.— Since the fibers $QR^{-1}(F)$ for a given $F \in \mathbb{G}_1$ are isometric to $\mathbb{T}^n(\mathbb{C})$ we have to prove the equality

$$\int_{\mathbb{T}^n(\mathbb{C})} \prod_{j<i} \left|1 - \frac{\lambda_i}{\lambda_j}\right|^2 d\mu(\mathbb{T}^n(\mathbb{C})) = \int_{\mathbb{T}^n(\mathbb{C})} \sum_{\sigma\in\Sigma_n} \prod_{j<i} \left|\frac{1-\lambda_{\sigma(i)}/\lambda_{\sigma(j)}}{1-\lambda_i/\lambda_j}\right|^{-2} d\mu(\mathbb{T}^n(\mathbb{C})).$$

Let us denote the Van der Monde determinant

$$V(\lambda_1,\dots,\lambda_n) = \begin{vmatrix} 1 & \lambda_1 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{n-1} \\ \dots & \dots & \dots & \dots \\ 1 & \lambda_n & \dots & \lambda_n^{n-1} \end{vmatrix} = \prod_{j<i} (\lambda_i - \lambda_j).$$

The first integral is equal to

$$\int_{\mathbb{T}^n(\mathbb{C})} \prod_{j<i} \left|1 - \frac{\lambda_i}{\lambda_j}\right|^2 d\mu(\mathbb{T}^n(\mathbb{C})) = \int_{\mathbb{T}^n(\mathbb{C})} \frac{|V(\lambda_1,\dots,\lambda_n)|^2}{\prod_{j<i}|\lambda_j|^2} d\mu(\mathbb{T}^n(\mathbb{C})).$$

The Van der Monde is equal to

$$V(\lambda_1,\dots,\lambda_n) = \sum_{\sigma\in\Sigma_n} \varepsilon(\sigma)\lambda_1^{\sigma(1)-1}\dots\lambda_n^{\sigma(n)-1}.$$

Here the sum is taken for any permutation σ in the symmetric group and $\varepsilon(\sigma) = \pm 1$ denotes its signature. The square of the absolute value of this Van der Monde is

$$|V(\lambda_1,\dots,\lambda_n)|^2 = \sum_{\sigma,\tau\in\Sigma_n} \varepsilon(\sigma)\varepsilon(\tau)\lambda_1^{\sigma(1)-1}\overline{\lambda}_1^{\tau(1)-1}\dots\lambda_n^{\sigma(n)-1}\overline{\lambda}_n^{\tau(n)-1}.$$

Now we integrate these products over a product of circles:

$$\int_{0<\theta_k<2\pi} \lambda_k^{\sigma(k)-1}\overline{\lambda}_k^{\tau(k)-1}d\theta_k = |\lambda_k|^{\sigma(k)+\tau(k)-2} \int_{0<\theta_k<2\pi} \exp(i\theta_k(\sigma(k)-\tau(k)))d\theta_k.$$

Since $d\theta_k$ is a probability measure, this last integral is equal to 1 when $\sigma(k) = \tau(k)$ and 0 otherwise. For this reason

$$\int_{\mathbb{T}^n(\mathbb{C})} \prod_{j<i} \left|1 - \frac{\lambda_i}{\lambda_j}\right|^2 d\mu(\mathbb{T}^n(\mathbb{C})) = \sum_{\sigma\in\Sigma_n} \frac{|\lambda_1|^{2\sigma(1)-2}\dots|\lambda_n|^{2\sigma(n)-2}}{\prod_{j<i}|\lambda_j|^2}.$$

The second integral is equal to

$$\int_{\mathbb{T}^n(\mathbb{C})} \sum_{\sigma\in\Sigma_n} \prod_{j<i} \left| \frac{1-\lambda_{\sigma(i)}/\lambda_{\sigma(j)}}{1-\lambda_i/\lambda_j} \right|^{-2} d\mu(\mathbb{T}^n(\mathbb{C}))$$
$$= \int_{\mathbb{T}^n(\mathbb{C})} \sum_{\sigma\in\Sigma_n} \frac{|V(\lambda_1,\dots,\lambda_n)|^2}{|V(\lambda_{\sigma(1)},\dots,\lambda_{\sigma(n)})|^2} \prod_{j<i} \left| \frac{\lambda_{\sigma(j)}}{\lambda_j} \right|^2 d\mu(\mathbb{T}^n(\mathbb{C}))$$
$$= \sum_{\sigma\in\Sigma_n} \prod_{j<i} \left| \frac{\lambda_{\sigma(j)}}{\lambda_j} \right|^2 .$$

The first and second integral are equal if and only if

$$\sum_{\sigma\in\Sigma_n} |\lambda_1|^{2\sigma(1)-2} \dots |\lambda_n|^{2\sigma(n)-2} = \sum_{\sigma\in\Sigma_n} \prod_{j<i} |\lambda_{\sigma(j)}|^2$$

or, in other terms, if and only if

$$\sum_{\sigma\in\Sigma_n} |\lambda_1|^{2\sigma(1)-2} \dots |\lambda_n|^{2\sigma(n)-2} = \sum_{\sigma\in\Sigma_n} |\lambda_{\sigma(1)}|^{2(n-1)} |\lambda_{\sigma(2)}|^{2(n-2)} \dots |\lambda_{\sigma(n-1)}|^2.$$

This last equality is obvious.

4.6. Proof of Proposition 5.— $\mathbb{G}_1$ is clearly open and semi-algebraic in $\mathbb{U}_n(\mathbb{C})$. For this reason, “full measure in $\mathbb{U}_n(\mathbb{C})$” is equivalent to “dense in $\mathbb{U}_n(\mathbb{C})$”. We shall prove now this last property.

Consider $\mathbb{V}_{1,A} \subset \mathbb{U}_n(\mathbb{C}) \times \mathbb{U}_n(\mathbb{C})$ defined by $(U_1,U_2) \in \mathbb{V}_{1,A}$ when $(U_2^* U_1 A U_2)_{i,j} = 0$ for $i > j$, that is the flag defined by U_2 is fixed by $(U_1 A)_\#$. $\mathbb{V}_{1,A}$ is a connected smooth real algebraic manifold. It is a locally trivial bundle over $\mathbb{V}_A$ with fiber $\mathbb{T}^n(\mathbb{C})$. Since the map $(U_1,U_2) \to U_2^* U_1 A U_2$ taking $\mathbb{U}_n(\mathbb{C}) \times \mathbb{U}_n(\mathbb{C})$ into $\mathbb{GL}_n(\mathbb{C})$ is transversal to the upper triangular matrices, which can be seen by varying U_1 alone, it follows that $\mathbb{V}_{1,A}$ is also a smooth variety. So a polynomial which vanishes on an open set in $\mathbb{V}_{1,A}$ vanishes identically. It will suffice to prove that the set of $(U_1,U_2) \in \mathbb{V}_{1,A}$ such that $U_1 A$ has distinct eigenvalue modules is dense in $\mathbb{V}_{1,A}$. Now the eigenvalues of $U_1 A$ are the diagonal elements of $U_2^* U_1 A U_2$. The set of $(U_1,U_2) \in \mathbb{V}_{1,A}$ where there are equal modulus eigenvalues on the diagonal is given by the equations

$$(\mathcal{P}_{i,k}) \qquad (U_2^* U_1 A U_2)_{i,i} \overline{(U_2^* U_1 A U_2)_{i,i}} = (U_2^* U_1 A U_2)_{k,k} \overline{(U_2^* U_1 A U_2)_{k,k}}.$$

So, if we show for each (i,k) that there are (U_1,U_2) such that the equality fails, then the variety defined by $\mathcal{P}_{i,k}$ is nowhere dense and the finite union of nowhere dense sets is nowhere dense. Let $A = V_1 D V_2$ be a singular decomposition of A: V_1 and V_2 are in $\mathbb{U}_n(\mathbb{C})$ and $D = \mathrm{Diag}(d_1,\dots,d_n)$ with $0 < d_1 \leqslant \dots \leqslant d_n$. We know by the hypothesis that there are at least two distinct d_i. This gives two unitary matrices U_1 and U_2 such that

$$U_2^* U_1 A U_2 = \mathrm{Diag}(d_1,\dots,d_n)$$

with some pair (d_{i_1}, d_{i_2}) of different moduli. By composing U_2 with a permutation matrix P, $P^*U_2^*U_1AU_2P$ permutes d_{i_1}, d_{i_2} to any two positions we wish, so we are done.

5. Proof of Theorem 8

We may decompose the measure μ along $\mathbb{SO}_2(\mathbb{R})$ orbits. Then we are reduced to comparing the integrals

$$\int_{\mathbb{SO}_2(\mathbb{R})} \log|\lambda_1(R_\theta A)| d\mu(\theta) = \int_{\mathbb{S}^1} \log\|A(\theta)\| d\theta$$

for Det $A > 0$ and

$$\int_{\mathbb{SO}_2(\mathbb{R})} \log|\lambda_1(R_\theta A)| d\mu(\theta) > \int_{\mathbb{S}^1} \log\|A(\theta)\| d\theta$$

for Det $A < 0$ unless A is a constant times a reflection in which case equality holds.

Without loss of generality we may assume that $|\operatorname{Det} A| = 1$ and hence that $\lambda_1(R_\theta A)\lambda_2(R_\theta A) = \pm 1$ for all θ as Det $A = \pm 1$. Now we consider

$$\mathbb{V}_A = \{(R_\theta, x) \in \mathbb{SO}_2(\mathbb{R}) \times \mathbb{S}^1 \ : \ (R_\theta A)x = x\}$$

and the two projections $\Pi_{\mathbb{SO}_2(\mathbb{R})} : \mathbb{V}_A \to \mathbb{SO}_2(\mathbb{R})$ and $\Pi_{\mathbb{S}^1} : \mathbb{V}_A \to \mathbb{S}^1$. Then

$$\begin{aligned}
&\int_{\mathbb{S}^1} \log\|A(\theta)\| d\theta \\
&= \int_{\mathbb{SO}_2(\mathbb{R})} \log|\lambda_1(R_\theta A)| \left|1 - \frac{\lambda_2(R_\theta A)}{\lambda_1(R_\theta A)}\right|^{-1} + \log|\lambda_2(R_\theta A)| \left|1 - \frac{\lambda_1(R_\theta A)}{\lambda_2(R_\theta A)}\right|^{-1} d\mu(\theta) \\
&\qquad = \int_{\mathbb{SO}_2(\mathbb{R})} \log|\lambda_1(R_\theta A)| \left(\left|1 - \frac{\lambda_2(R_\theta A)}{\lambda_1(R_\theta A)}\right|^{-1} - \left|1 - \frac{\lambda_1(R_\theta A)}{\lambda_2(R_\theta A)}\right|^{-1} \right) d\mu(\theta).
\end{aligned}$$

Now for $\lambda_1\lambda_2 = 1$

$$\left|1 - \frac{\lambda_2}{\lambda_1}\right|^{-1} - \left|1 - \frac{\lambda_1}{\lambda_2}\right|^{-1} = \frac{1}{1 - \frac{\lambda_2}{\lambda_1}} - \frac{1}{\frac{\lambda_1}{\lambda_2} - 1} = 1$$

while for $\lambda_1\lambda_2 = -1$

$$\left|1 - \frac{\lambda_2}{\lambda_1}\right|^{-1} - \left|1 - \frac{\lambda_1}{\lambda_2}\right|^{-1} = \frac{1}{1 - \frac{\lambda_2}{\lambda_1}} - \frac{1}{1 - \frac{\lambda_1}{\lambda_2}} = \frac{\lambda_1 + \lambda_2}{\lambda_1 - \lambda_2} < 1.$$

This proves Theorem 8 except for the possibility that Det $A = -1$ and $\log\|A(\theta)\|$ is identically zero, i.e. A is a reflection.

References

[1] Blum L., F. Cucker, M. Shub, S. Smale *Complexity and Real Computation*, Springer, 1998.

[2] Bougerol, P. and J. Lacroix *Products of Random Matrices with Applications to Schrödinger Operators*, Progress in Probability and Statistics, Vol. 8, Birkhauser, 1985.

[3] Burns K., C. Pugh, M. Shub and A. Wilkinson, *Recent Results about Stable Ergodicity*, to appear in: Proceedings on Symposia in Pure Mathematics, the Seattle Conference of Smooth Ergodic Theory, AMS.

[4] Gol'shied I.Ya. and G.A. Margulis, *Lyapunov Indices of a Product of Random Matrices*, Russian Math. Surveys **44:5** (1989), pp. 11-71.

[5] Guivarc'h Y. and A. Raugi, *Propriétés de contraction d'un semi-groupe de matrices inversibles. Coefficients de Liapunoff d'un produit de matrices aléatoires indépendantes*, Israel J. Math. **65** (1989), pp. 165-196.

[6] Ledrappier, F., *Positivity of the Exponent for Stationary Sequences of Matrices*, in: Lyapunov Exponents, Proceedings, Bremen 1984, Arnold L. and V. Wihstutz Eds., Lectures Notes in Math. Vol. 1186.

[7] Oseledec, V.I., *A Multiplicative Ergodic Theorem. Lyapunov Characteristic Numbers for Dynamical Systems*, Trans. Moscow Math. Soc., 19, (1968), pp. 197-231.

[8] Ruelle, D., *Ergodic Theory of Differentiable Dynamical Systems*, Publications Mathématiques de l'IHES, Volume 50, (1979), pp. 27-58.

[9] Shub M. and A. Vasquez, *Some Linearly Induced Morse-Smale Systems, the QR Algorithm and the Toda Lattice*, in: The Legacy of Sonia Kovalevskaya, Linda Keen Ed., Contemporary Mathematics, Vol. 64, AMS, (1987), pp. 181-193.

J.-P. Dedieu, MIP, Département de Mathématique, Université Paul Sabatier, 31062 Toulouse Cedex 04, France • *E-mail :* `dedieu@@mip.ups-tlse.fr`

M. Shub, Department of Mathematical Sciences, IBM T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598, USA • *E-mail :* `mshub@@us.ibm.com`

Astérisque
287, 2003, p. 19–32

ON SOME APPROXIMATIONS OF THE QUASI-GEOSTROPHIC EQUATION

by

Efim I. Dinaburg, Vladimir S. Posvyanskii & Yakov G. Sinai

Abstract. — For two-dimensional quasi-geostrophic equation in Fourier space, we propose a new type approximation representing itself some quasi-linear equation. Natural finite dimensional approximations of this equation are investigated in the article.

1. Introduction

The main difficulty in the proof of existence and uniqueness of solutions of hydrodynamical equations is the lack of understanding of the role played by non-linear, or Eulerian, terms. In Fourier space these terms describe the expansion of initial excitations of Fourier modes but the way how this process goes is in general unclear.

In this paper we propose an approach which leads to some simplifications of the original equations with the belief that the processes of expansion remain the same. Our equations have natural finite-dimensional approximations which are systems of ODE and are easier to tackle.

We restrict ourselves to the two-dimensional quasi-geostrophic equation (QGE) for an unknown function $u(k,t), k = (k_1, k_2) \in R^2$ which in Fourier space has the form (see [**1**], [**2**])

$$\frac{\partial u(k,t)}{\partial t} = \int_{R^2} \frac{((k')^{\perp}, k-k')}{|k-k'|} u(k',t)u(k-k',t)dk' - \nu |k|^{2\alpha} u(k,t) \tag{1}$$

Here $|k| = (k_1^2 + k_2^2)^{1/2}$, $k^{\perp} = (-k_2, k_1)$, $\nu \geqslant 0$ is the viscosity and we are interested in even solutions $u(-k,t) = u(k,t)$. It is well-known that the mathematical difficulties

2000 ***Mathematics Subject Classification.*** — 76D05, 35Q30.
Key words and phrases. — Quasi-geostrophic equation, quasi-linear approximation.

The first and the third authors acknowledge the financial support from RFFI, Grant 99-01-00314. The third author thanks NSF for the support, Grant DMS-0070698.

related to (1) are in many respects similar to the well-known difficulties for the $3D$-Navier-Stokes system.

The main case is $\alpha = 1$. For $0 < \alpha < 1$ we obtain the so-called generalized QGE, which we also consider in this paper.

We are interested in solutions which are smooth in k and decay at infinity rather slowly. Our main assumption says that for such solutions the main contribution to the integral in (1) comes from $|k'| \ll |k|$ or $|k-k'| \ll |k|$. For $|k'| \ll |k|$ we can write

$$\frac{((k')^{\perp}, k-k')}{|k-k'|} = \left((k')^{\perp}, k\right) \left(\frac{1}{|k|} - (\nabla \frac{1}{|k|}, k') \right) + \cdots$$

where dots mean terms of a smaller order of magnitude. Thus for $|k'| \ll |k|$ we keep the term

$$\begin{aligned} &\int_{R^2} ((k')^{\perp}, k) \left(\frac{1}{|k|} - (\nabla \frac{1}{|k|}, k') \right) u(k',t)(u(k,t) - (\nabla u(k,t), k'))dk' \\ &\quad = \int_{R^2} \frac{((k')^{\perp}, k)}{|k|} u(k',t)u(k,t)dk' - u(k,t) \int_{R^2} ((k')^{\perp}, k)(\nabla \frac{1}{|k|}, k')u(k',t)dk' \\ &- \int_{R^2} \frac{((k')^{\perp}, k)}{|k|} (\nabla u(k,t), k')u(k',t)dk' + \int_{R^2} ((k')^{\perp}, k)(\nabla \frac{1}{|k|}, k')(\nabla u(k,t), k'))u(k',t)dk' \end{aligned}$$

The first and the last integrals are zero because the integrands are odd functions of k'. For $|k-k'| \ll k$ put $k^{''} = k - k'$. Then

$$\begin{aligned} &\int_{R^2} \frac{((k-k^{''})^{\perp}, k)}{|k^{''}|} u(k-k^{''},t)u(k^{''},t)dk^{''} \\ &\qquad = \int_{R^2} \frac{((k^{\perp}), k^{''})}{|k^{''}|} (u(k,t) - (\nabla u(k,t), k^{''}))u(k^{''},t)dk^{''} + \cdots \\ &= \int_{R^2} \frac{((k^{\perp}), k^{''})}{|k^{''}|} u(k,t)u(k^{''},t)dk^{''} - \int_{R^2} \frac{((k^{\perp}), k^{''})}{|k^{''}|} (\nabla u(k,t), k^{''})u(k^{''},t)dk^{''} + \cdots \end{aligned}$$

Again dots mean terms of a smaller order of magnitude. The first integral is zero by the same reasons as above, i.e. the parity of the integrand. Thus our approximating equation takes the form

$$\begin{aligned} (2) \quad \frac{\partial u(k,t)}{\partial t} = -u(k,t) &\int_{R^2} ((k')^{\perp}, k)(\nabla \frac{1}{|k|}, k')u(k',t)dk' \\ &- \int_{R^2} \frac{((k')^{\perp}, k)}{|k|} u(k',t) \left(k', \nabla u(k,t)\right) dk' \\ &\quad - \int_{R^2} \frac{((k)^{\perp}, k')}{|k'|} u(k',t) \left(k', \nabla u(k,t)\right) dk' - \nu |k|^{2\alpha} u(k,t) \end{aligned}$$

The equation (2) does not satisfy the energy estimate but apparently remains dissipative because of viscosity. Let us rewrite (2) as follows:

$$(3)\quad \frac{\partial u(k,t)}{\partial t} = -u(k,t)\int_{R^2}((k')^\perp,k)(\nabla\frac{1}{|k|},k')u(k',t)dk' \\ -\int_{R^2}((k')^\perp,k)\left[\frac{1}{|k|}+\frac{1}{|k'|}\right]u(k',t)\,(k',\nabla u(k,t))\,dk' - \nu|k|^{2\alpha}u(k,t)$$

The equation (3) is a first order quasi-linear equation whose coefficients are global functions of u. Take the first term in (3):

$$I_1(t) = \int_{R^2}((k')^\perp,k)\left(\nabla\frac{1}{|k|},k'\right)u(k',t)dk'$$

We have

$$(k')^\perp = (-k_2',k_1');\nabla\frac{1}{|k|} = \left(-\frac{k_1}{(k_1^2+k_2^2)^{3/2}},-\frac{k_2}{(k_1^2+k_2^2)^{3/2}}\right).$$

Therefore

$$\begin{aligned}I_1(t) &= -\int(k_1'k_2-k_2'k_1)\frac{k_1k_1'+k_2k_2'}{|k|^3}\,u(k'\,;\,t)\,dk' \\ &= -\frac{k_2^2}{|k|^3}\int k_1'k_2'u(k',t)dk' + \frac{k_1^2}{|k|^3}\int k_1'k_2'u(k',t)dk' \\ &\qquad \frac{k_1k_2}{|k|^3}\int((k_1')^2-(k_2')^2)u(k',t)dk'.\end{aligned}$$

Denote

$$a_1(t) = \int k_1^2u(k,t)dk;\quad a_2(t) = \int k_2^2u(k,t)dk;\quad a_3(t) = \int k_1k_2u(k,t)dk;$$

Then

$$I_1(t) = \frac{k_1^2-k_2^2}{|k|^3}a_3 - \frac{k_1k_2}{|k|^3}(a_1-a_2)$$

Consider

$$I_2(t) = \int((k')^\perp,k)\frac{1}{|k|}u(k',t)(k',\nabla u(k,t))dk'$$

We have

$$\begin{aligned}I_2(t) &= \int dk'\,\frac{k_2k_1'-k_1k_2{}'}{|k|}u(k',t)(k_1'\frac{\partial u(k,t)}{\partial k_1}+k_2'\frac{\partial u(k,t)}{\partial k_2}) \\ &= \frac{\partial u(k,t)}{\partial k_1}\left[-\frac{k_1}{|k|}\int k_1'k_2'u(k',t)dk' + \frac{k_2}{|k|}\int(k_1')^2u(k',t)dk'\right] \\ &\qquad + \frac{\partial u(k,t)}{\partial k_2}\left[-\frac{k_1}{|k|}\int(k_2')^2u(k',t)dk' + \frac{k_2}{|k|}\int k_1'k_2'u(k',t)dk'\right] \\ &= \frac{\partial u(k,t)}{\partial k_1}\left[-\frac{k_1}{|k|}a_3+\frac{k_2}{|k|}a_1\right] + \frac{\partial u(k,t)}{\partial k_2}\left[-\frac{k_1}{|k|}a_2+\frac{k_2}{|k|}a_3\right]\end{aligned}$$

The last term

$$\begin{aligned} I_3(t) &= \int \frac{((k')^\perp, k)}{|k'|} u(k', t) \left(k'_1 \frac{\partial u(k,t)}{\partial k_1} + k'_2 \frac{\partial u(k,t)}{\partial k_2} \right) dk' \\ &= \frac{\partial u(k,t)}{\partial k_1} \left(-k_1 \int \frac{k'_1 k'_2}{|k'|} u(k',t) dk' + k_2 \int \frac{(k'_1)^2}{|k'|} u(k',t) dk' \right) \\ &\quad + \frac{\partial u(k,t)}{\partial k_2} \left(-k_1 \int \frac{(k'_2)^2}{|k'|} u(k',t) dk' + k_2 \int \frac{k'_1 k'_2}{|k'|} u(k',t) dk' \right) \end{aligned}$$

Denote

$$b_1(t) = \int \frac{(k'_1)^2}{|k'|} u(k',t) dk', \quad b_2(t) = \int \frac{(k'_2)^2}{|k'|} u(k',t) dk', \quad b_3(t) = \int \frac{k'_1 k'_2}{|k'|} u(k',t) dk',$$

$$\begin{aligned} h_0(k,t) &= \left[\frac{k_2^2 - k_1^2}{|k|^3} a_3 + \frac{k_1 k_2}{|k|^3} (a_1 - a_2) - \nu |k|^{2\alpha} \right], \\ h_1(k,t) &= -\left[\frac{k_1}{|k|} a_3 - \frac{k_2}{|k|} a_1 + k_1 b_3 - k_2 b_1 \right], \\ h_2(k,t) &= -\left[\frac{k_1}{|k|} a_2 - \frac{k_2}{|k|} a_3 + k_1 b_2 - k_2 b_3 \right]. \end{aligned}$$

Then the equation (3) takes its final form:

$$\frac{\partial u(k,t)}{\partial t} + \frac{\partial u(k,t)}{\partial k_1} h_1(k,t) + \frac{\partial u(k,t)}{\partial k_2} h_2(k,t) = h_0(k,t) u$$

or

$$\frac{du(k,t)}{dt} = h_0(k,t) u(k,t) \tag{4}$$

where

$$\begin{aligned} \frac{dk_1}{dt} &= h_1(k,t) \\ \frac{dk_2}{dt} &= h_2(k,t). \end{aligned} \tag{5}$$

However, we should not forget that the coefficients a_i, b_i are also functions of unknowns k and u. (5) are the equations for characteristics of our quasi-linear equation. We can think about them as curves along which the non-linearity spreads. Denote by S^{t_1,t_2} the family of shifts along solutions of (5). Then

$$u(k, t_2) = u(k(t_1), t_1) \exp\left(\int_{t_1}^{t_2} h_0(k(\tau), \tau)) d\tau \right), \tag{6}$$

where $k(\tau) = S^{t_1,\tau} k(t_1)$.

Corollary 1. — *The sign of u is preserved along the characteristics of* (5).

Proof. — Follows immediately from (6). □

This property is special for our approximation. Presumably it is not true in a general case.

Corollary 2. — *If* $u(-k,t) = u(k,t)$ *then* $h_1(-k,t) = -h_1(k,t)$, $h_2(-k,t) = -h_2(k,t)$.

This property has an important interpretation. Consider $h_1(k,t), h_2(k,t)$ as the components of our vector field (5). These components are odd functions of k. Therefore the trajectories of the symmetric (with respect to the origin) points are symmetric.

A similar approximation can be constructed for the Navier-Stokes system. It will be discussed in another paper.

2. Finite-dimensional Approximations

Assume that $u(k,0)$ is non-zero only for finitely many k, i.e. $u(k^{(i)},0) = u^{(i)}$ for $i = 1,2,\ldots,I$. Then $u(k,t)$ is non-zero at I points $k^{(i)}(t)$. In this case

$$a_1 = a_1(t) = \sum_{i=1}^{I} (k_1^{(i)})^2 u(k^{(i)},t)$$
$$a_2 = a_2(t) = \sum_{i=1}^{I} (k_2^{(i)})^2 u(k^{(i)},t)$$
$$a_3 = a_3(t) = \sum_{i=1}^{I} k_1^{(i)} k_2^{(i)} u(k^{(i)},t)$$

and

$$b_1(t) = \sum_{i=1}^{I} \frac{(k_1^{(i)})^2}{|k^{(i)}|} u(k^{(i)},t)$$
$$b_2(t) = \sum_{i=1}^{I} \frac{(k_2^{(i)})^2}{|k^{(i)}|} u(k^{(i)},t)$$
$$b_3(t) = \sum_{i=1}^{I} \frac{k_1^{(i)} k_2^{(i)}}{|k^{(i)}|} u(k^{(i)},t)$$

The system of equations of dynamics of the points $k^{(i)}$ takes the form

$$(7) \qquad \begin{aligned} \frac{dk_1^{(i)}}{dt} &= -\left[\frac{k_1^{(i)}}{|k^{(i)}|} a_3 - \frac{k_2^{(i)}}{|k^{(i)}|} a_1 + k_1^{(i)} b_3 - k_2^{(i)} b_1\right] \\ \frac{dk_2^{(i)}}{dt} &= -\left[\frac{k_1^{(i)}}{|k^{(i)}|} a_2 - \frac{k_2^{(i)}}{|k^{(i)}|} a_3 + k_1^{(i)} b_2 - k_2^{(i)} b_3\right] \end{aligned}$$

Let $I = 1$. Then

$$a_1 = (k_1^{(1)})^2 u(k^{(1)}, t), \quad a_2 = (k_2^{(1)})^2 u(k^{(1)}, t), \quad a_3 = k_1^{(1)} k_2^{(1)} u(k^{(1)}, t),$$
$$b_1 = \frac{(k_1^{(1)})^2}{|k^{(1)}|} u(k^{(1)}, t), \quad b_2 = \frac{(k_2^{(1)})^2}{|k^{(1)}|} u(k^{(1)}, t), \quad b_3 = \frac{k_1^{(1)} k_2^{(1)}}{|k^{(1)}|} u(k^{(1)}, t).$$

We immediately see that $h_1 = h_2 = 0$, i.e. the point stays fixed and $u(k^{(i)}, t) \to 0$ as $t \to \infty$.

If $I = 2$ and $k^{(2)} = -k^{(1)}, u(k^{(2)}) = u(k^{(1)})$ then in view of the symmetry (see Corollary 2) both points stay fixed. The first non-trivial case arises for an arbitrary configuration of two points. According to the Corollary 2 it is equivalent to the case of four points consisting of two symmetric pairs. Denote $u_1 = u(k^{(1)}), u_2 = u(k^{(2)})$.

We come to the following remarkable system of ODE:

$$\begin{aligned}
\frac{dk_1^{(1)}}{dt} &= -u_2 k_1^{(2)} \left[\frac{1}{|k^{(1)}|} + \frac{1}{|k^{(2)}|}\right] \left(k_1^{(1)} k_2^{(2)} - k_2^{(1)} k_1^{(2)}\right) \\
\frac{dk_2^{(1)}}{dt} &= -u_2 k_2^{(2)} \left[\frac{1}{|k^{(1)}|} + \frac{1}{|k^{(2)}|}\right] \left(k_1^{(1)} k_2^{(2)} - k_2^{(1)} k_1^{(2)}\right) \\
\frac{dk_1^{(2)}}{dt} &= u_1 k_1^{(1)} \left[\frac{1}{|k^{(1)}|} + \frac{1}{|k^{(2)}|}\right] \left(k_1^{(1)} k_2^{(2)} - k_2^{(1)} k_1^{(2)}\right) \\
\frac{dk_2^{(2)}}{dt} &= u_1 k_2^{(1)} \left[\frac{1}{|k^{(1)}|} + \frac{1}{|k^{(2)}|}\right] \left(k_1^{(1)} k_2^{(2)} - k_2^{(1)} k_1^{(2)}\right) \\
\frac{du_1}{dt} &= \left[-\frac{u_2}{|k^{(1)}|^3} \left(k_1^{(1)} k_2^{(2)} - k_2^{(1)} k_1^{(2)}\right) \left(k_1^{(1)} k_2^{(2)} + k_2^{(1)} k_1^{(2)}\right) - \nu |k^{(1)}|^{2\alpha}\right] u_1 \\
\frac{du_2}{dt} &= \left[\frac{u_1}{|k^{(2)}|^3} \left(k_1^{(1)} k_2^{(2)} - k_2^{(1)} k_1^{(2)}\right) \left(k_1^{(1)} k_2^{(2)} + k_2^{(1)} k_1^{(2)}\right) - \nu |k^{(2)}|^{2\alpha}\right] u_2
\end{aligned} \tag{8}$$

Lemma 1. — *$S = k_1^{(1)} k_2^{(2)} - k_2^{(1)} k_1^{(2)}$ is the first integral of* (8).

Proof. — Direct checking. □

It is not difficult to see that

$$\begin{aligned}
\frac{d|k^{(1)}|}{dt} &= -\frac{u_2 S(k^{(1)}, k^{(2)})}{|k^{(1)}|} \left(\frac{1}{|k^{(1)}|} + \frac{1}{|k^{(2)}|}\right) \\
\frac{d|k^{(2)}|}{dt} &= \frac{u_1 S(k^{(1)}, k^{(2)})}{|k^{(2)}|} \left(\frac{1}{|k^{(1)}|} + \frac{1}{|k^{(2)}|}\right) \\
\frac{d(k^{(1)}, k^{(2)})}{dt} &= S \left(-u_2 |k^{(2)}|^2 + u_1 |k^{(1)}|^2\right) \left(\frac{1}{|k^{(1)}|} + \frac{1}{|k^{(2)}|}\right)
\end{aligned} \tag{9}$$

In view of (9) the last two equations of the system (8) can be rewritten in the form

$$
\begin{aligned}
\frac{d\ln|u_1|}{dt} &= -\frac{d}{dt}\left(\frac{1}{|k^{(1)}|}\right)\left(\frac{1}{|k^{(1)}|}+\frac{1}{|k^{(2)}|}\right)^{-1} - \nu|k^{(1)}|^{2\alpha}\\
\frac{d\ln|u_2|}{dt} &= -\frac{d}{dt}\left(\frac{1}{|k^{(2)}|}\right)\left(\frac{1}{|k^{(1)}|}+\frac{1}{|k^{(2)}|}\right)^{-1} - \nu|k^{(2)}|^{2\alpha}
\end{aligned}
\tag{10}
$$

Lemma 2. — *The function* $F(u,k) = \ln\left(|u_1||u_2|\left(\frac{1}{|k^{(1)}|}+\frac{1}{|k^{(2)}|}\right)\right)$ *is a Lyapunov function of the system* (8).

Proof. — With the help of a simple transformation we obtain from(10) that along any trajectory of the system (8) the derivative of $F(t)$ is less than 0:

$$\frac{d}{dt}F(t) = -\nu\left(|k^{(1)}|^{2\alpha}+|k^{(2)}|^{2\alpha}\right). \qquad \square$$

Corollary. — *For* $S \neq 0$ *and* $t > 0$

$$
\begin{aligned}
|u_1(t)u_2(t)|&\left(\frac{1}{|k^{(1)}(t)|}+\frac{1}{|k^{(2)}(t)|}\right)\\
&\leqslant |u_1(0)u_2(0)|\left(\frac{1}{|k^{(1)}(0)|}+\frac{1}{|k^{(2)}(0)|}\right)\exp(-2\nu|S|^{\alpha}t).
\end{aligned}
$$

Proof. — The statement follows from the inequality:

$$|k^{(1)}(t)|^{2\alpha}+|k^{(2)}(t)|^{2\alpha} \geqslant 2|k^{(1)}(t)|^{\alpha}|k^{(2)}(t)|^{\alpha} > 2|S|^{\alpha}. \qquad \square$$

The system (8) has an invariant four-dimensional manifold

$$\Gamma = \{k^{(1)}, k^{(2)}, u_1, u_2 \mid u_1 = 0, u_2 = 0\}$$

which is locally stable. Our next result shows that this manifold is also globally stable.

Theorem 1. — *For any solution of* (8),

$$\lim_{t\to\infty} u_1(t) = \lim_{t\to\infty} u_2(t) = 0, \quad \lim_{t\to\infty} k_j^{(i)} = \overline{k}_j^{(i)}$$

where $k_j^{(i)}$ *are limiting values of* $\overline{k}_j^{(i)}(t)$ *which certainly depend on the initial conditions.*

A priori there can be solutions which escape to infinity in finite time, i.e. $k_j^{(i)} \to \infty$ as $t \to t_0$, which means some blow up. The theorem shows that this does not happen in our case and each solution approaches some point of Γ.

Proof. — For $S = 0$ the statement of the theorem is obvious. Therefore, we can restrict ourselves to the case $S \neq 0$. It is sufficient to consider positive S, because the case of negative S can be reduced to positive S by changing the order of points. First, we prove the theorem for $\alpha = 1$ and in the Appendix we sketch the main steps for $\alpha < 1$.

Put

$$L(t) = |u_1(t)u_2(t)| \left(\frac{1}{|k^{(1)}(t)|} + \frac{1}{|k^{(2)}(t)|} \right) \quad \text{and} \quad \overline{L}(t) = sqn(u_1(0)u_2(0))L(t).$$

Using these notations we can write $u_1(t)$ in the following way:

$$\begin{aligned} u_1(t) &= \left[-\int_{t_0}^{t} \frac{Su_1(\tau)u_2(\tau)}{|k^{(1)}(\tau)|^3} (k^{(1)}(\tau), k^{(2)}(\tau)) \exp\left(\nu \int_{t_0}^{\tau} |k^{(1)}(s)|^2 ds \right) dt + u_1(t_0) \right] \\ &\qquad\qquad \times \exp\left(-\nu \int_{t_0}^{t} |k^{(1)}(s)|^2 ds \right) \\ &= \Bigg[-\int_{t_0}^{t} \overline{L}(\tau) \frac{S|k_2(\tau)|}{|k^{(1)}(\tau)|^2(|k^{(1)}(\tau)| + |k^{(2)}(\tau)|)} (k^{(1)}(\tau), k^{(2)}(\tau)) \\ &\qquad \exp\left(\nu \int_{t_0}^{\tau} |k^{(1)}(s)|^2 ds \right) d\tau + u_1(t_0) \Bigg] \times \exp\left(-\nu \int_{t_0}^{t} |k^{(1)}(s)|^2 ds \right). \end{aligned}$$

Using the relations

$$\begin{gathered} \overline{L}(\tau) = \overline{L}(\tau_0) \exp\left(-\nu \int_{t_0}^{\tau} (|k^{(1)}(s)|^2 + |k^{(2)}(s)|^2) ds \right), \\ |k^{(1)}(\tau)|(|k^{(1)}(\tau)| + |k^{(2)}(\tau)|) > |k^{(1)}(\tau)||k^{(2)}(\tau)| \geqslant S, \\ |(k^{(1)}(\tau), k^{(2)}(\tau))| \leqslant |k^{(1)}(\tau)||k^{(2)}(\tau)| \end{gathered} \tag{11}$$

we obtain from the previous expressions that

$$\begin{aligned} &\left| u_1(t) - u_1(t_0) \exp\left(-\nu \int_{t_0}^{t} |k^{(1)}(s)|^2 ds \right) \right| \\ &< \left(L(t_0) \int_{t_0}^{t} |k^{(2)}(\tau)|^2 \exp\left(-\nu \int_{t_0}^{\tau} |k^{(2)}(s)|^2 ds \right) d\tau \right) \exp\left(-\nu \int_{t_0}^{t} |k^{(1)}(\tau)|^2 d\tau \right) \\ &\qquad \leqslant \nu^{-1} L(t_0) \left(1 - \exp\left(-\nu \int_{t_0}^{t} |k^{(2)}(\tau)|^2 d\tau \right) \right) \exp\left(-\nu \int_{t_0}^{t} |k^{(1)}(\tau)|^2 d\tau \right). \end{aligned} \tag{12}$$

In the same way $u_2(t)$ satisfies the inequality

$$\begin{aligned} &\left| u_2(t) - u_2(t_0) \exp\left(-\nu \int_{t_0}^{t} |k^{(2)}(s)|^2 ds \right) \right| \\ &\qquad < \nu^{-1} L(t_0) \left(1 - \exp\left(-\nu \int_{t_0}^{t} |k^{(1)}(\tau)|^2 d\tau \right) \right) \exp\left(-\nu \int_{t_0}^{t} |k^{(2)}(\tau)|^2 d\tau \right). \end{aligned}$$

It is not difficult to verify that for the functions $v_1(t) = u_1(t)/|k^{(1)}(t)|$ and $v_2(t) = u_2(t)/|k^{(2)}(t)|$ the following inequalities hold:

(13)
$$\begin{aligned}
&\left|v_1(t) - v_1(t_0)\exp\left(-\nu\int_{t_0}^{t}|k^{(1)}(s)|^2ds\right)\right| \\
&\qquad < (2\nu\sqrt{S})^{-1}L(t_0)\left(1-\exp\left(-\nu\int_{t_0}^{t}|k^{(2)}(\tau)|^2d\tau\right)\right)\exp\left(-\nu\int_{t_0}^{t}|k^{(1)}(\tau)|^2d\tau\right),\\
&\left|v_2(t) - v_2(t_0)\exp\left(-\nu\int_{t_0}^{t}|k^{(2)}(s)|^2ds\right)\right| \\
&\qquad < (2\nu\sqrt{S})^{-1}L(t_0)\left(1-\exp\left(-\nu\int_{t_0}^{t}|k^{(1)}(\tau)|^2d\tau\right)\right),\exp\left(-\nu\int_{t_0}^{t}|k^{(2)}(\tau)|^2d\tau\right)
\end{aligned}$$

The inequalities (13) can be obtained from the system of differential equations for $v_1(t)$ and $v_2(t)$:

$$\begin{aligned}
\frac{dv_1}{dt} &= v_1(t)\left(\frac{v_2(t)S(k^{(1)}(t),k^{(2)}(t))}{|k^{(1)}(t)|^2} - \nu|k^{(1)}(t)|^2\right)\\
\frac{dv_2}{dt} &= v_2(t)\left(\frac{-v_1(t)S(k^{(1)}(t),k^{(2)}(t))}{|k^{(2)}(t)|^2} - \nu|k^{(2)}(t)|^2\right).
\end{aligned}\tag{14}$$

with the help of the above mentioned arguments for $u_1(t)$ and $u_2(t)$. The system (14) follows directly from (8).

At least one of the two integrals $\int_{t_0}^{\infty}|k^{(1)}(\tau)|^2d\tau$ and $\int_{t_0}^{\infty}|k^{(2)}(\tau)|^2d\tau$ diverges because their sum diverges. Assume for example that the first one diverges. Then from (12) it follows that $\lim_{t\to\infty}u_1(t) = 0$ because $\lim_{t_0\to\infty}L(t_0) = 0$. If $\int_{t_0}^{\infty}|k^{(2)}(\tau)|^2d\tau$ also diverges then $\lim_{t\to\infty}u_2(t) = 0$.

In the case of convergence of the last integral $\lim_{t\to\infty}u_2(t)$ exists and is nonzero. Indeed, for any given $\varepsilon > 0$ find such t_0 that $\nu^{-1}L(t_0) < \varepsilon/3$ and then choose $\overline{t}(t_0)$ so that for $t_2 > t_1 > \overline{t}$

$$\left|u_2(t_0)\exp\left(-\nu\int_{t_0}^{t_2}|k^{(2)}(\tau)|^2d\tau\right) - u_2(t_0)\exp\left(-\nu\int_{t_0}^{t_1}|k^{(2)}(\tau)|^2d\tau\right)\right| < \varepsilon/3$$

We have

$$\begin{aligned}
|u_2(t_2)-u_2(t_1)| &\leqslant \left|u_2(t_2) - u_2(t_0)\exp\left(-\nu\int_{t_0}^{t_2}|k^{(2)}(\tau)|^2d\tau\right)\right|\\
&\qquad + \left|u_2(t_1) - u_2(t_0)\exp\left(-\nu\int_{t_0}^{t_1}|k^{(2)}(\tau)|^2d\tau\right)\right|\\
&+\left|u_2(t_0)\exp\left(-\nu\int_{t_0}^{t_2}|k^{(2)}(\tau)|^2d\tau\right) - u_2(t_0)\exp\left(-\nu\int_{t_0}^{t_1}|k^{(2)}(\tau)|^2d\tau\right)\right|\\
&\qquad < +2\nu^{-1}L(t_0) + \varepsilon/3 < \varepsilon.
\end{aligned}$$

This gives the existence of the desired limit.

The same statement is true for $\lim_{t\to\infty} v_2(t)$ (see (13) and (14)). Thus $|k^{(2)}(t)|$ tends to a positive number or to $+\infty$ when $t \to \infty$. This contradicts the convergence of $\int_{t_0}^{\infty} |k^{(2)}(\tau)|^2 d\tau$, i.e. $\lim_{t\to\infty} u_2(t) = 0$.

Now we shall study the behavior of the vectors $k^{(i)}(t)$, $i = 1, 2$, for $t \to \infty$. The first equality in (11) implies that there exist positive constants C_1, C_2 and suffiently large t_0 depending on the inital data for which

(15)
$$C_1 \exp\left(-\nu \int_{t_0}^{t} |k^{(i)}(\tau)|^2 d\tau \right) < |u_i(t)| < C_2 \exp\left(-\nu \int_{t_0}^{t} |k^{(i)}(\tau)|^2 d\tau \right), \quad i = 1, 2.$$

Substituting (15) in (11) we obtain after simple calculations that

$$A_1 < \frac{1}{|k^{(1)}(t)|} + \frac{1}{|k^{(2)}(t)|} < A_2 \tag{16}$$

for some positive A_i depending on the initial data.

From (15) and from the system (8) for components of the vectors $k^{(1)}(t), k^{(2)}(t)$ one can conclude that $k_j^{(i)}(t), i, j = 1, 2$ have finite limits when t tends $\to \infty$. Let us check this for $k_1^{(1)}(t)$:

$$k_1^{(1)}(t) - k_1^{(1)}(t_0) = -S \int_{t_0}^{t} u_2(\tau) k_1^{(2)}(\tau) \left(\frac{1}{|k^{(1)}(\tau)|} + \frac{1}{|k^{(2)}(\tau)|} \right) d\tau \tag{17}$$

The integral in the right-hand side converges because its absolute value is less than the integral

$$\int_{t_0}^{\infty} |u_2(\tau)|\, |k^{(2)}(\tau)| \left(\frac{1}{|k^{(1)}(\tau)|} + \frac{1}{|k^{(2)}(\tau)|} \right) d\tau$$

which is finite.

Theorem 1 is proven for $\alpha = 1$. □

3. Numerical experiments: results and discussion

In the previous section we considered the finite dimensional systems for $I = 1, 2$. For $I > 2$ we do not have rigorous results but did only some numerical experiments to understand the behavior of solutions in some cases. Several results are represented in Figures 1-11.

In Figures 1-8 solutions for $I = 4$ are shown, and in Figures 9-11 for $I = 64$. The initial data for $I = 4$ were taken as follows:

1	0.01	0.0	100.0
2	0.0	0.2	200.0
3	0.02	0.3	−50.0
4	0.04	0.5	−1.0

Here in the table the first column represents the number of a point, the second (third) column represents the first (second) coordinate of the point, the forth is $u_i(0)$.

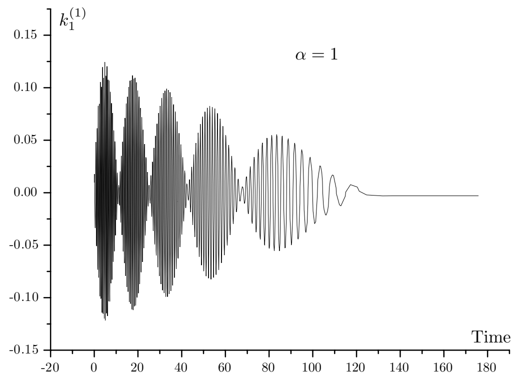

Figure 1

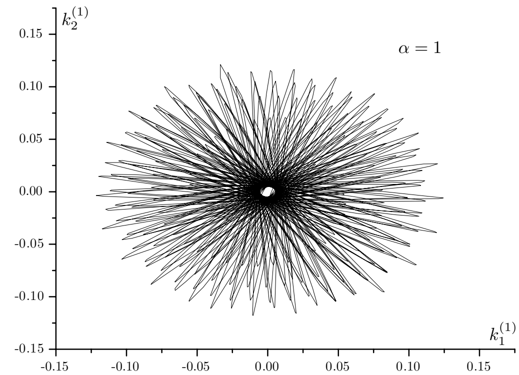

Figure 2

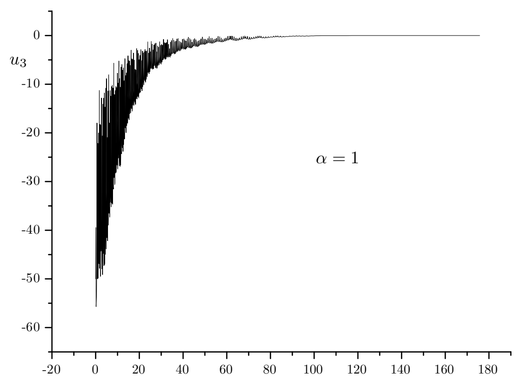

Figure 3

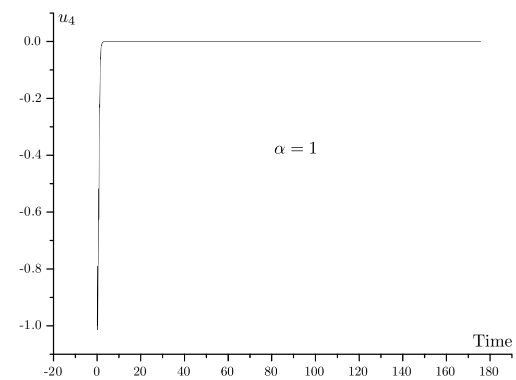

Figure 4

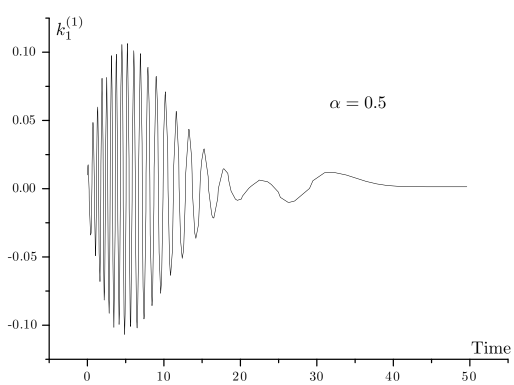

Figure 5

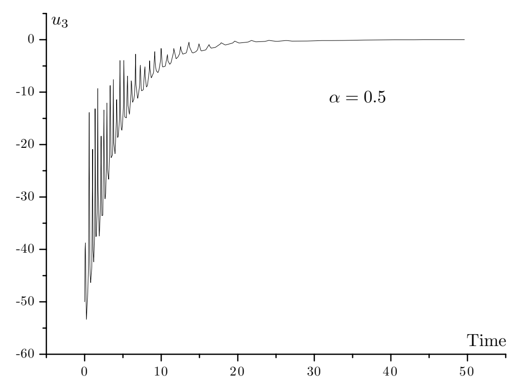

Figure 6

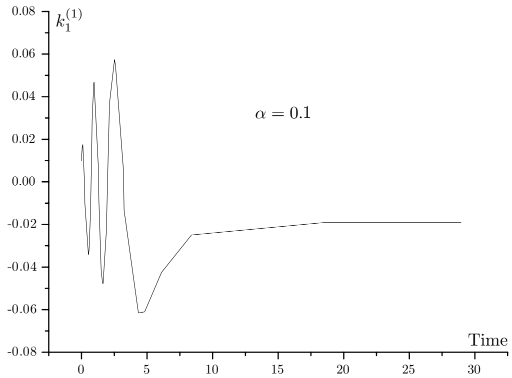

Figure 7

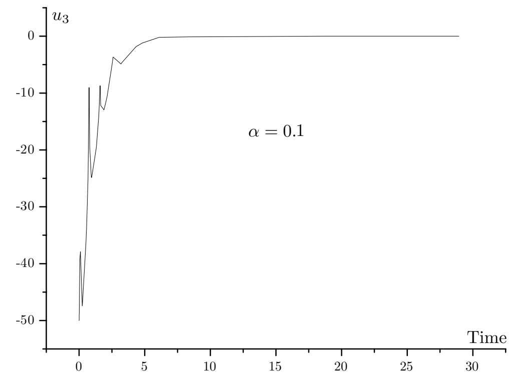

Figure 8

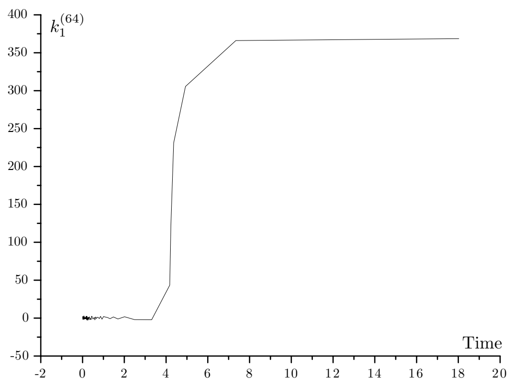

Figure 9

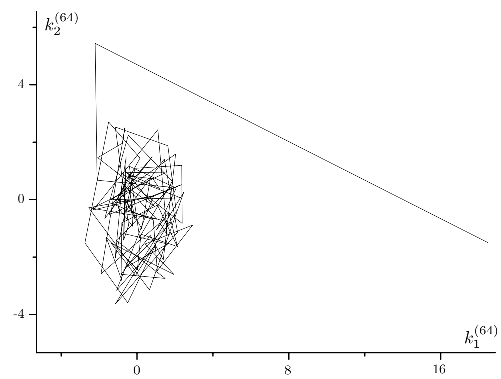

Figure 10

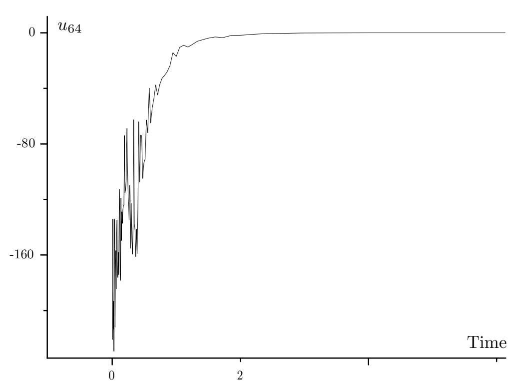

Figure 11

Figures 1, 2, 3, and 4 demonsrate the graphs of the function $k_1^{(1)}(t)$, phase portrait $k_2^{(1)}(k_1^{(1)})$, functions $u_3(t)$ and $u_4(t)$, respectively, for $\alpha = 1$. Figures 5 and 6 show $k_1^{(1)}(t)$ and $u_3(t)$, respectively, for $\alpha = 0.5$. Figures 7 and 8 demonstrate graphs of the same functions for $\alpha = 0.1$. At last on Figures 9, 10 and 11 $k_1^{(64)}(t)$, $k_2^{(64)}(k_1^{(64)})$ and $u_{64}(t)$ respectively are shown. For $I = 64$ initial data were choosen in the following way. Points were distributed uniformely in the square $[0.1, 2] \times [0.1, 2]$ and $u_i(0) = C_1 + C_2 \sin(\beta k_1^{(i)}) + C_3 \sin(\gamma k_2^{(i)})$, where $C_1 = -100$; $C_2 = 50$; $C_3 = 200 : \beta = 5$; $\gamma = 6$. We carried out several hundreds of numerical experiments with different initial data. In all cases the behavior of solutions was similar. Namely, $u_i(t) \to 0$, $k_j^{(i)}(t)$ converge to limiting values as $t \to \infty$.

Appendix. Sketch of the proof of Theorem 1 for $\alpha < 1$

Put

$$w_i(t) = u_i(t) \exp\left(\nu \int_{t_0}^t |k^{(i)}(\tau)|^{2\alpha} d\tau\right) \quad (i = 1, 2)$$

and after simple transformations rewrite the last two equations in (8) in the following form:

$$\begin{aligned}\frac{dw_1}{dt} &= -\frac{\overline{L(t)} S(k^{(1)}, k^{(2)}) |k^{(2)}|}{|k^{(1)}|^2 (|k^{(1)}| + |k^{(2)}|)} \exp\left(\nu \int_{t_0}^t |k^{(1)}(\tau)|^{2\alpha} d\tau\right) \\ \frac{dw_2}{dt} &= \frac{\overline{L(t)} S(k^{(1)}, k^{(2)}) |k^{(1)}|}{|k^{(2)}|^2 (|k^{(1)}| + |k^{(2)}|)} \exp\left(\nu \int_{t_0}^t |k^{(2)}(\tau)|^{2\alpha} d\tau\right)\end{aligned} \tag{18}$$

It is not difficult to see that

$$\begin{aligned}\left|\frac{dw_1}{dt}\right|^\alpha &\leqslant \left(L(t_0) |k^{(2)}|^2 \exp\left(-\nu \int_{t_0}^t |k^{(2)}(\tau)|^{2\alpha} d\tau\right)\right)^\alpha \\ \left|\frac{dw_2}{dt}\right|^\alpha &\leqslant \left(L(t_0) |k^{(1)}|^2 \exp\left(-\nu \int_{t_0}^t |k^{(1)}(\tau)|^{2\alpha} d\tau\right)\right)^\alpha\end{aligned} \tag{19}$$

Integrating both parts of the inequalities (19) from t_0 to t we obtain:

$$\begin{aligned}\int_{t_0}^t \left|\frac{dw_1(\tau)}{d\tau}\right|^\alpha d\tau &\leqslant (L(t_0))^\alpha (\alpha\nu)^{-1} \left(1 - \exp\left(-\alpha\nu \int_{t_0}^t |k^{(2)}(\tau)|^{2\alpha} d\tau\right)\right) \\ \int_{t_0}^t \left|\frac{dw_2(\tau)}{d\tau}\right|^\alpha d\tau &\leqslant (L(t_0))^\alpha (\alpha\nu)^{-1} \left(1 - \exp\left(-\alpha\nu \int_{t_0}^t |k^{(1)}(\tau)|^{2\alpha} d\tau\right)\right)\end{aligned}$$

Since $\lim_{t_0 \to \infty} L(t_0) = 0$ the norms in the $L^\alpha(t_0, \infty)$ of the functions $\dfrac{dw_i}{dt}$ tend to zero in the space $L^\alpha(t_0, \infty)$ when $t_0 \to \infty$.

In the same way, one can show that the norms in the space $L^\alpha(t_0, \infty)$ of the derivatives of functions $w_i(t)/|k^{(i)}(t)|$ also tend to zero when $t_0 \to \infty$.

As in the case $\alpha = 1$, previous arguments give that $\lim_{t \to \infty} u_i(t) = 0, i = 1, 2$. Hence all statments of Theorem 1 follow easily.

References

[1] J. Pedlosky, "Geophysical Fluid Dynamics", Springer, New-York, 1987.

[2] P. Constantin, Wu Jiahong, "Behavior of solutions of 2D quasi-geostrophic equations", *SIAM J. Math. Anal.*, **30**, (1999), no. 5, 937-948.

E.I. Dinaburg, Institute of Earth Physics, RAS, Moscow

V.S. Posvyanskii, N.N. Semenov Institute of Chemical Physics, RAS, Moscow
E-mail : posv@center.chph.ras.ru

Ya.G. Sinai, Mathematics Department, Princeton University
E-mail : sinai@math.princeton.edu

Astérisque
287, 2003, p. 33–60

STABLE ACCESSIBILITY IS C^1 DENSE

by

Dmitry Dolgopyat & Amie Wilkinson

Abstract. — We prove that in the space of all C^r ($r \geqslant 1$) partially hyperbolic diffeomorphisms, there is a C^1 open and dense set of accessible diffeomorphisms. This settles the C^1 case of a conjecture of Pugh and Shub. The same result holds in the space of volume preserving or symplectic partially hyperbolic diffeomorphisms. Combining this theorem with results in [**Br**], [**Ar**] and [**PugSh3**], we obtain several corollaries. The first states that in the space of volume preserving or symplectic partially hyperbolic diffeomorphisms, topological transitivity holds on an open and dense set. Further, on a symplectic n-manifold ($n \leqslant 4$) the C^1-closure of the stably transitive symplectomorphisms is precisely the closure of the partially hyperbolic symplectomorphisms. Finally, stable ergodicity is C^1 open and dense among the volume preserving, partially hyperbolic diffeomorphisms satisfying the additional technical hypotheses of [**PugSh3**].

Introduction

This paper is about the accessibility property of partially hyperbolic diffeomorphisms. We show that accessibility holds for a C^1 open and dense set in the space of all partially hyperbolic diffeomorphisms, thus settling the C^1 version of a conjecture of Pugh and Shub [**PugSh1**]. Partially hyperbolic diffeomorphisms are similar to Anosov diffeomorphisms, in that they possess invariant hyperbolic directions, but unlike Anosov diffeomorphisms, they can also possess invariant directions of non-hyperbolic behavior. Accessibility means that the hyperbolic directions fill up the manifold on a macroscopic scale. Accessibility often provides enough hyperbolicity for a variety of chaotic properties, such as topological transitivity [**Br**] and ergodicity

2000 *Mathematics Subject Classification.* — Primary 37D30, 37C20.
Key words and phrases. — Partial hyperbolicity, accessibility.

This work is partially supported by the NSF.

[**PugSh3**], to hold. As a consequence, we derive several density results about stable ergodicity and stable transitivity among partially hyperbolic diffeomorphisms.

Let M be a smooth compact, connected and boundaryless Riemannian manifold. A diffeomorphism $f : M \to M$ is *partially hyperbolic* if the tangent bundle to M splits as a Tf-invariant sum

$$TM = E^u \oplus E^c \oplus E^s,$$

such that Tf uniformly expands all vectors in E^u and uniformly contracts all vectors in E^s, while vectors in E^c are neither contracted as strongly as any vector in E^s nor expanded as strongly as any vector in E^u. More precisely, for each $p \in M$, there exist $0 < a_p < b_p < 1 < B_p < A_p$ such that:

$$\|T_pf|_{E^s}\| \leqslant a_p < b_p \leqslant m(T_pf|_{E^c}) \leqslant \|T_pf|_{E^c}\| \leqslant B_p < A_p \leqslant m(T_pf|_{E^u}),$$

where $m(T) = \|T^{-1}\|^{-1}$. Throughout this paper we assume that both subbundles E^u and E^s are nontrivial.

A more stringent condition, often called partial hyperbolicity in the literature (cf. [**BrPe**], [**BuPuShWi**]) requires that the constants a_p, b_p, A_p and B_p be chosen independent of p. Since the results in this paper apply to diffeomorphisms satisfying the weaker condition, to avoid excessive terminology, we will use the term partial hyperbolicity in the broader sense.

A partially hyperbolic diffeomorphism f is *accessible* if, for every pair of points $p, q \in M$, there is a C^1 path from p to q whose tangent vector always lies in $E^u \cup E^s$ and vanishes at most finitely many times. We say f is *stably accessible* if every g sufficiently C^1-close to f is accessible. We prove here the following theorem.

Main Theorem. — *For any $r \geqslant 1$, stable accessibility is C^1 dense among the C^r, partially hyperbolic diffeomorphisms of M, volume preserving or not. If M is a symplectic manifold, then stable accessibility is C^1 dense among C^r, symplectic partially hyperbolic diffeomorphisms of M.*

Related to the Main Theorem is the result of Niţică and Török [**NiTö**] that stable accessibility is C^r-dense among partially hyperbolic diffeomorphisms with 1-dimensional, integrable center bundle E^c. Other results about stable accessibility treat more special classes of diffeomorphisms, such as time-one maps of Anosov flows [**BuPuWi**], skew products [**BuWi1**], certain systems where $E^u \oplus E^s$ is integrable [**ShWi**], and systems whose partially hyperbolic splitting is C^1 [**PugSh2**].

The Main Theorem has several corollaries. The first corollary concerns the topological transitivity of partially hyperbolic diffeomorphisms and follows immediately from a theorem of Brin [**Br**]. Denote by $\mathcal{P}H^r(M)$ the set of C^r partially hyperbolic diffeomorphisms of M. If μ and ω are, respectively, Riemannian volume and a symplectic form on M, then set

$$\mathcal{P}H^r_\mu(M) = \{f \in \mathcal{P}H^r(M) \mid f_*(\mu) = \mu\}, \text{ and}$$
$$\mathcal{P}H^r_\omega(M) = \{f \in \mathcal{P}H^r(M) \mid f^*(\omega) = \omega\}.$$

Corollary 0.1. — *For $r \geqslant 1$, there is a C^1-open and dense set of topologically transitive diffeomorphisms in $\mathcal{P}H^r_\mu(M)$. If M has a symplectic form ω, then there is a C^1-open and dense set of transitive diffeomorphisms in $\mathcal{P}H^r_\omega(M)$.*

This corollary is false without the volume preservation assumption. Niţică and Török have shown in [**NiTö**] that there is an open set of accessible non-transitive diffeomorphisms. While it is plausible that for a C^1 open and dense set of diffeomorphisms in the space $\mathcal{P}H^r(M)$, there are only finitely many transitivity components, it is not a direct corollary of the Main Theorem.

M.-C. Arnaud has shown in [**Ar**] that if M is a symplectic 4-manifold, then the stably transitive diffeomorphisms in $\mathrm{Diff}^r_\omega(M)$ are partially hyperbolic. (The same result has been announced by J. Xia in arbitrary dimension). Hence there is a complete picture in dimension 4 of the stably transitive diffeomorphisms, which we summarize in the next corollary.

Corollary 0.2. — *Let M be a symplectic manifold with $\dim(M) \leqslant 4$. The C^1-closure of the stably transitive diffeomorphisms in $\mathrm{Diff}^r_\omega(M)$ coincides with the C^1 closure of the partially hyperbolic ones.*

In other words, invariant tori are essentially the only obstacle for topological transitivity in the symplectic category, at least if $\dim(M) \leqslant 4$. We conjecture that the same is true in the volume preserving case.

Conjecture 0.3. — *In the space of volume preserving diffeomorphisms, the C^1-closure of the stably transitive diffeomorphims coincides with the closure of the diffeomorphisms admitting a dominated splitting.*

For a discussion of the dominated splitting condition and some results related to Conjecture 0.3 see [**Vi**]. Even though the results of this paper could be useful in attacking this conjecture some other ideas (possibly ones from the paper [**BonDi**]) are necessary to solve this problem. Here we note only that in [**BV**] a volume preserving example is presented which is stably transitive yet not partially hyperbolic. A. Tahzhibi has announced a proof that these example are in fact stably ergodic.

Another corollary of the Main Theorem concerns ergodicity of $f \in \mathcal{P}H^r_\mu(M)$. Pugh and Shub proved the following theorem:

***Theorem 0.4* ([PugSh3, Theorem A])**. — *Let $f \in \mathcal{P}H^2_\mu(M)$. If f is center bunched, dynamically coherent, and essentially accessible, then f is ergodic.*

Thus we also have the corollary:

Corollary 0.5. — *Among the center bunched, stably dynamically coherent diffeomorphisms in $\mathcal{P}H^2_\mu(M)$, stable ergodicity is C^1 open and dense.*

Theorem 0.4 refers to partially hyperbolic diffeomorphisms in the stronger sense described earlier, but recently Burns and Wilkinson [**BuWi2**] have shown that these results extend to the larger class of partially hyperbolic diffeomorphisms described in this paper (satisfying additional center bunching conditions). For a description of examples of diffeomorphisms satisfying the conditions "center bunched" and "stably dynamically coherent" see the survey paper [**BuPuShWi**]. In particular, the corollary implies that there is a C^1-open neighborhood $\mathcal{U} \subset \mathcal{P}H^2_\mu(M)$ of f in which stable ergodicity is C^1-open and dense, where f is the time-t map of an Anosov flow, a compact group extension of an Anosov diffeomorphism, an ergodic automorphism of a torus or nilmanifold, or a partially hyperbolic translation on a compact homogeneous space.

This paper arose out of an attempt to prove the following conjecture of Pugh and Shub.

Conjecture 0.6 **([PugSh2**, Conjecture 4] **and [PugSh3**, Conjecture 2]**)**
Stable accessibility is C^r - dense in both $\mathcal{P}H^r(M)$ and $\mathcal{P}H^r_\mu(M)$.

In the spirit of Theorem 0.4, Pugh and Shub also conjectured:

Conjecture 0.7 **([PugSh3**, Conjecture 3]**)**. — *A partially hyperbolic C^2 volume preserving diffeomorphism with the essential accessibility property is ergodic.*

Finally, combining Conjectures 0.6 and 0.7, they conjectured:

Conjecture 0.8 **([PugSh3**, Conjecture 4]**)**. — *Stable ergodicity is C^r-dense in $\mathcal{P}H^r_\mu(M)$.*

As with Theorem 0.4, these conjectures refer to the narrower class of partially hyperbolic diffeomorphisms described above, but in light of the results in [**BuWi2**] and this paper, it seems reasonable to extend them to the class under consideration here.

The question of accessibility is closely related to problems in control theory (see, e.g. [**Lo**]). In fact, analogous density theorems in control theory initially suggested the Conjectures 0.6 and 0.7. The sole reason that the results in control theory cannot be directly transported to this setting is that we do not perturb the bundles E^u and E^s directly, but rather the diffeomorphism f. We'd like to be able to say that a specific perturbation of f has a specific effect on E^u and E^s. What makes this difficult is that $E^s(p)$ and $E^u(p)$ are determined by the entire forward and backward orbit of p, respectively; a perturbation will have various effects along the length of this orbit, some desirable and others not.

The key observation that permits a measure of control is that the effects of the perturbation are greatest along the first few iterates of p. To maximize our control over the bundles E^u and E^s, we isolate regions where we need local accessibility and localize the perturbation to these regions. Choosing the support of the perturbation to be highly non-recurrent then minimizes undesirable "noisy" effects of the perturbations. The trade-off is that the desirable effects of the pertubations are necessarily quite small. Nonetheless, with the right C^1-small perturbation, the desirable effects outweigh the undesirable ones and we obtain accessibility. Similar perturbations are found in [**PP**].

It appears that a localized C^2-small perturbation cannot achieve this, and so the techniques in this paper do not extend to the C^2 setting. New techniques would be required to prove Conjectures 0.6 and 0.7.

Here is how the proof of the Main Theorem goes. Let $f \in \mathcal{P}H^r(M)$ have partially hyperbolic splitting $TM = E^u \oplus E^c \oplus E^s$. By [**BrPe, HiPuSh**] E^u and E^s are tangent to the leaves of continuous foliations which are denoted $\mathcal{W}^u$ and $\mathcal{W}^s$ and are called the unstable and stable foliations respectively. A *us-path for f* is a path $\gamma : [0,1] \to M$ consisting of a finite number of consecutive arcs — called *legs* — each of which is a curve that lies in a single leaf of $\mathcal{W}^u$ or $\mathcal{W}^s$. It is easy to see that f is accessible if and only if for all $p, q \in M$, there is a us-path for f from p to q.

To prove the Main Theorem, we first find a collection of disjoint disks in M. Each disk is approximately tangent to the center direction E^c. We choose this collection large enough so that f is accessible, modulo these disks. More precisely, for every $p, q \in M$, there is a finite sequence of us-paths for f, the first path originating at p and ending in one of the disks, the last path originating in a disk and ending at q. The intermediate paths all begin and end in disks, each path beginning in the disk where the previous path ends. We then perturb f in a small neighborhood of these disks. We can arrange that if this neighborhood and the C^0-size of the pertubation are both sufficiently small, then the perturbed system will still be accessible, modulo the same collection of disks. It is not hard to see that any additional C^1-small perturbation will preserve this property.

Under the C^r-dense assumption that the fixed points of f^k are isolated, for all $k \geqslant 1$, we can choose these disks to be very small and their union highly non-recurrent. This is Lemma 1.2. We show in Lemma 1.1 that it is then possible to perturb f in a neighborhood of these disks by a C^1-small perturbation to obtain a stably accessible g. We prove stable accessibility by showing that any two points in a given disk can be connected by a us-path for g, and for any small perturbation of g. Since any small perturbation of g is already accessible modulo these disks, this gives stable accessibility.

Lemma 1.1 is the only place where it is essential that the perturbation be only C^1-small. When we examine the effect of perturbing f on E^u and E^s, we find that in C^1, the contribution to $E^u(p)$ and $E^s(p)$ of the perturbation *near* p is larger than the *combined* contributions along the rest of the orbit of p; this is not true in C^2. Therefore, a more complicated analysis, taking into account the first several returns, is needed to establish the analogue of our result in the C^2-setting.

1. Proof of the Main Theorem

Proof. — We first prove the Main Theorem in the case where f preserves a smooth volume μ. The proof is easily modified to the non volume preserving case. In the final section we describe how to modify the proof for the symplectic case.

Let $f \in \mathcal{P}H^r_\mu(M)$ and $\delta > 0$ be given. Our goal is to find a stably accessible $g \in \mathrm{Diff}^r_\mu(M)$ with $d_{C^1}(f, g) < \delta$. We give some preliminary definitions and notation.

Let $\mathcal{P}(M)$ be the collection of all subsets of M. We say that f is *accessible on* $X \in P(M)$ if, for every $p, q \in X$, there is a us-path for f from p to q. The diffeomorphism f is *accessible modulo* $\mathcal{X} \subset \mathcal{P}(M)$ if, for every $p, q \in M$, there exist $X_1, \ldots, X_r \in \mathcal{X}$ and us-paths for f: from p to X_1, from q to X_r, and from X_m to X_{m+1}, for $m = 1, \ldots, r-1$. We say that f is *uniformly accessible modulo* $\mathcal{X}$ if f is accessible modulo $\mathcal{X}$ and, further, there is a number N such that the us-paths in the previous definition can be chosen to have length less than N and to have fewer than N legs.

A property of a diffeomorphism f is said to hold *stably* if it also holds for all $g \in \mathrm{Diff}^1(M)$ sufficiently C^1-close to f. Clearly if f is stably accessible modulo $\{X_1, \ldots, X_k\}$ and f is stably accessible on each X_i, then f is stably accessible.

Define a function $R : \mathcal{P}(M) \to \mathbb{N} \cup \{\infty\}$ as follows. For $X \in \mathcal{P}(M)$, let $R(X)$ be the smallest $J \in \mathbb{N} \cup \{\infty\}$ satisfying:

$$f^i(X) \cap X \neq \varnothing, \text{ with } |i| = J + 1. \tag{1}$$

Note that $R(B_\rho(p)) \to \mathrm{per}(p)$, as $\rho \to 0$, where we set $\mathrm{per}(p) = \infty$ if p is not periodic.

We next fix a system of local charts on M. We will on several occasions refer to the orthogonal splitting $\mathbb{R}^n = T_v\mathbb{R}^n = \mathbb{R}^u \oplus \mathbb{R}^c \oplus \mathbb{R}^s$, where $a = \dim(E^a)$.

Let $B^n(v, \rho)$ denote the ball of radius ρ about $v \in \mathbb{R}^n$ with respect to the sup norm on coordinates. More generally, we will use the notation $B^a(v, \rho)$, where $a = u, c, s, c+u, c+s$, or $u+s$, to denote the sup-norm ball of radius ρ about v in the affine space $v + \mathbb{R}^a$.

Applying Moser's theorem on the equivalence of volume forms [**Mo**] we obtain, for any $p \in M$, a C^∞ map

$$\varphi_p : B^n(0,1) \longrightarrow M$$

such that

(1) $\varphi_p(0) = p$,

(2) $T_0\varphi_p$ sends the splitting $T_0\mathbb{R}^n = \mathbb{R}^u \oplus \mathbb{R}^c \oplus \mathbb{R}^s$ to the splitting $T_pM = E^u \oplus E^c \oplus E^s$,

(3) φ_p sends divergence-free vector fields to divergence-free vector fields.

(4) $p \mapsto \varphi_p$ is a uniformly continuous map from M to $C^1(B^n(0,1), M)$. The dependence of φ_p on f is also continuous.

([**Mo**] gives maps satisfying (1), (3) and (4). (2) can be achived by precomposing with a linear map.)

Since we do not assume that E^c is tangent to a foliation, we will work with approximate center manifolds. For $\rho < 1$ and $p \in M$, let

$$V_\rho(p) = \varphi_p(B^c(0,\rho)).$$

We refer to $V_\rho(p)$ as a *c-admissible disk* with *center* p and *radius* ρ and write $r(V_\rho(p)) = \rho$. If D is a c-admissible disk with center p and radius ρ, then for $\beta \in (0,1)$, we denote by βD the c-admissible disk with center p and radius $\beta\rho$.

A *c-admissible family* is a finite collection of pairwise disjoint, c-admissible disks. If $\mathcal{D}$ is a c-admissible family, and $\beta < 1$, then let

$$\beta\mathcal{D} = \{\beta D \mid D \in \mathcal{D}\}, \quad |\mathcal{D}| = \bigcup_{D\in\mathcal{D}} D, \quad r(\mathcal{D}) = \sup_{D\in\mathcal{D}} r(D), \text{ and } \quad R(\mathcal{D}) = R(|\mathcal{D}|).$$

We have the following lemma.

***Lemma 1.1* (Accessibility on central disks)**. — *Let $f \in \mathcal{P}H^r_\mu(M)$ and $\delta > 0$ be given. Then there exists $J > 0$ with the following property.*

If $\mathcal{D}$ is a c-admissible family with $r(\mathcal{D}) < J^{-1}$ and $R(\mathcal{D}) > J$, then for all $\sigma > 0$ and $\beta \in (0,1)$, there exists $g \in \mathrm{Diff}^r_\mu(M)$ such that:

(1) $d_{C^1}(f,g) < \delta$,
(2) $d_{C^0}(f,g) < \sigma$,
(3) *For each $D \in \mathcal{D}$, g is stably accessible on βD.*

We may assume that the fixed points of f^k are isolated, for all $k \geqslant 1$; this property is C^r-dense in $\mathrm{Diff}^r_\mu(M)$. Under this additional assumption we have the following lemma.

***Lemma 1.2* (Accessibility modulo central disks)**. — *Let $f \in \mathcal{P}H^r_\mu(M)$ be given. Assume that the fixed points of f^k are isolated, for all $k \geqslant 1$. Then for every $J > 0$ there exists a c-admissible family $\mathcal{D}$ such that:*

(1) $r(\mathcal{D}) < J^{-1}$,
(2) $R(\mathcal{D}) > J$,
(3) *f is uniformly accessible modulo $\frac{1}{2}\mathcal{D}$.*

***Lemma 1.3* (Persistence of accessibility modulo $\mathcal{D}$)**. — *There exists $\delta_0 > 0$ so that, given $\delta < \delta_0$, a c-admissible family $\mathcal{D}$ with f uniformly accessible modulo $\frac{1}{2}\mathcal{D}$, and $\beta \in (\frac{1}{2}, 1)$, the following holds.*

There exists $\sigma > 0$ such that any g satisfying

(1) $d_{C^1}(f,g) < \delta$,
(2) $d_{C^0}(f,g) < \sigma$,

is accessible modulo $\beta\mathcal{D}$ (and hence g is stably accessible modulo $\beta\mathcal{D}$, since (1) and (2) are open conditions).

The proof of the Main Theorem now follows from Lemmas 1.1, 1.2 and 1.3. Let f and δ be given. After a C^r-small perturbation, we may assume that the fixed points of f^k are isolated, for all k. We may assume that $\delta < \delta_0$, where δ_0 is given by Lemma 1.3.

Choose J according to Lemma 1.1. By Lemma 1.2, there exists a c-admissible family $\mathcal{D}$, with $R(\mathcal{D}) > J$ and $r(\mathcal{D}) < J^{-1}$, such that f is accessible modulo $\frac{1}{2}\mathcal{D}$.

Now fix $\beta \in (\frac{1}{2}, 1)$, and choose σ according to Lemma 1.3. Applying Lemma 1.1 we obtain a diffeomorphism $g \in \mathrm{Diff}^r_\mu(M)$, with $d_{C^1}(f,g) < \delta$ and $d_{C^0}(f,g) < \sigma$, such that g is stably accessible on βD, for each $D \in \mathcal{D}$. By Lemma 1.3, g is also stably accessible modulo $\beta\mathcal{D}$. Thus, g is stably accessible. $\square$

The proofs of Lemmas 1.2 and 1.3 are given in the next section, and the proof of Lemma 1.1 is given in Section 3.

The arguments of Section 2 become simpler if $E^c(f)$ is integrable. In that case, one can work with central disks instead of c-admissible ones. To construct a family of central disks satsifying the conditions of Lemma 1.2 one should take a small $\varepsilon > 0$, choose an ε–net $\{p_j\}$ and let $\mathcal{D}$ be the union of unit central disks centered at p_j. If there are some $i, j < \mathrm{Card}(\mathcal{D})$ and $n < J$ such that $f^n D(p_i) \bigcap D(p_j) \neq \varnothing$ then one can remove this intersection by arbitrary small shift of p_i in an unstable direction (this is possible even if $i = j$ since the periodic leaves of f are isolated by partial hyperbolicity). In case E^c is not integrable the proof becomes significantly more complicated since one has to work with disks which are almost tangent to E^c, but the idea remains the same. As mentioned in the introduction the most difficult part of the proof is Section 3 where the abundance of C^1-perturbations is crucial. Therefore many readers would find it helpful to skip Section 2 during the first reading returning to it after mastering Section 3.

2. Global accessibility

In this section we prove Lemmas 1.2 and 1.3.

Proof of Lemma 1.2. — Let J be given. Let $A = \{p \in M \mid \mathrm{per}(p) \geqslant J+2\}$. Since the fixed points of f^j, $j < J+2$, are isolated, there exist $x_1, \dots, x_m \in M$ such that $A = M \smallsetminus \{x_1, \dots, x_m\}$.

For $\rho > 0$, let $U_\rho(p)$ be the image of the ball $B^n(0,\rho)$ under φ_p. The proof of the following lemma is straightforward.

Lemma 2.1. — *If $r > 0$ is sufficiently small, then every $p \in \bigcup U_r(x_i)$ can be connected to a point in $M \smallsetminus \bigcup U_r(x_i)$ by a us-path with one leg.*

Choose r according to this lemma, and let $A_r = M \smallsetminus \bigcup U_r(x_i)$. Assume r is small enough that A_r is connected. Since A_r is compact and contained in A, there exists $\rho_0 < 1/(2J)$ such that $R(U_{2\rho_0}(q)) > J$, for every $q \in A_r$.

The next lemma follows from the uniform continuity of the invariant splitting for f, and we omit its proof.

Lemma 2.2. — *There exists $K > 1$ so that, for ρ_0 sufficiently small, for all $p \in M$, and for every $q_1, q_2 \in U_{\rho_0/K}(p)$, there is a us-path with $\leqslant 2$ legs from q_1 to some point in $V_{\rho_0}(q_2)$.*

The next lemma is key.

Lemma 2.3. — *Let $K > 1$. If ρ_0 is sufficiently small, then there exist a cover of A_r by finitely many neighborhoods $U_1, \ldots, U_k$ of the form $U_i = U_{\rho_0/K}(q_i)$, and, for $i = 1, \ldots, k$, points $p_i \in U_i$ such that*

$$V_{2\rho_0}(p_i) \cap V_{2\rho_0}(p_j) = \varnothing,$$

for $i \neq j$, and

$$V_{2\rho_0}(p_i) \cap f^m\left(V_{2\rho_0}(p_j)\right) = \varnothing,$$

for all i, j and $0 < |m| \leqslant J$.

Before proving Lemma 2.3, we finish the proof of Lemma 1.2.

Let $U_1, \ldots, U_k$ and $p_1, \ldots, p_k$ be given by Lemma 2.3. For $i = 1, \ldots, c$, let $D_i = V_{2\rho_0}(p_i)$. Note that the collection $\mathcal{D} = \{D_1, \ldots, D_k\}$ satisfies conclusions (1) and (2) of Lemma 1.2, if ρ_0 is sufficiently small.

Lemma 2.2 implies that, for every $p \in B_i$, there is a us-path with $\leqslant 2$ legs from p to some point in $\frac{1}{2}D_i$. It follows that, whenever $B_i \cap B_j \neq \varnothing$, there is a us-path with $\leqslant 4$ legs from some point on $\frac{1}{2}D_i$ to some point in $\frac{1}{2}D_j$. Since A_r is connected and the balls $B_i, \ldots, B_k$ cover A_r, we obtain, for *any* i, j, a sequence of disks $D_{a_0} = D_i, D_{a_1}, \ldots, D_{a_l} = D_j$, such that $\frac{1}{2}D_{a_m}$ is connected to $\frac{1}{2}D_{a_{m+1}}$ by a us-path, for $m = 0, \ldots, l-1$. Then for any $p, q \in A_r$, there are a sequence of disks $D_{b_0}, \ldots, D_{b_s}$ and us-paths: from p to $\frac{1}{2}D_{b_0}$, from q to $\frac{1}{2}D_{b_s}$, and from $\frac{1}{2}D_{b_r}$ to $\frac{1}{2}D_{b_{m+1}}$, for $m = 0, \ldots, s-1$. The length and number of legs of these paths is clearly bounded. Since any point in $M \smallsetminus A_r = \bigcup B_r(x_i)$ can be connected to a point in A_r by a us-path with one leg, it follows that f is uniformly accessible modulo $\{\frac{1}{2}D_1, \ldots, \frac{1}{2}D_k\}$. This proves (3), completing the proof of Lemma 1.2. □

Proof of Lemma 2.3. — We start with a simple Besicovich-type covering lemma.

***Lemma 2.4* (Covering lemma)**. — *For any $C > 0$ there exists an integer $N > 0$ such that, for any compact set $A \subseteq M$, and for $\rho > 0$ sufficiently small, there exist $q_1, \ldots, q_k \in A$ with the following properties, for $i = 1, \ldots, k$:*

– *$A \subseteq B_\rho(q_1) \cup \cdots \cup B_\rho(q_k)$, and*
– *$\#\{j \mid B_{C\rho}(q_i) \cap B_{C\rho}(q_j) \neq \varnothing\} \leqslant N$.*

Proof of Lemma 2.4. — On the manifold M, there exists a constant $K > 0$ such that for every $\rho < 1$ and every $p \in M$, the volume of the ball $B_\rho(p)$ lies between ρ^n/K and $K\rho^n$. Let $N = (4C+2)^n K^2$; this is an upper bound on the number of disjoint balls of radius $\rho/2$ that can fit inside a ball of radius $(2C+1)\rho$.

Let A and ρ be given. Let $S_\rho \subset M$ be a maximal ρ-separated subset of A. Such a set exists by compactness of A. We claim that S_ρ is also ρ-spanning. If not, then there exists $x \in A$ such that $d(x,y) > \rho$, for all $y \in A$. But this contradicts maximality of S_ρ. Hence, if $q_1, \dots, q_k$ are the elements of S_ρ, then

$$A \subseteq B_\rho(q_1) \cup \dots \cup B_\rho(q_k).$$

For $p \in S_\rho$, let $N(p)$ be the set of $q \in S_\rho$ such that

$$B_{C\rho}(p) \cap B_{C\rho}(q) \neq \varnothing.$$

For each $q \in N(p)$, the distance from p to q is less than $2C\rho$, and so the ball $B_{\rho/2}(q)$ is contained in $B_{(2C+1)\rho}(p)$. Since S_ρ is ρ-separated, the balls $B_{\rho/2}(q)$ and $B_{\rho/2}(q')$ are disjoint, for distinct $q, q' \in N(p)$. The cardinality of $N(p)$ is thus bounded by N, which completes the proof. □

The sets $U_\rho(p)$ are uniformly comparable to round balls $B_\rho(p)$, and the maps $\{f^m, |m| \leqslant J\}$ distort distances by a bounded factor. Thus Lemma 2.4 implies the following.

***Corollary 2.5* (Strengthened covering lemma)**. — *Let $C, J > 0$ be given. There exists an integer $N > 0$ such that, for any compact set $A \subseteq M$, and for any $\rho > 0$, there exist $q_1, \dots, q_k \in A$ with the following properties, for $i = 1, \dots, k$:*

- *$A \subseteq U_\rho(q_1) \cup \dots \cup U_\rho(q_k)$, and*
- *$\#\{j \mid U_{C\rho}(q_i) \cap f^m(U_{C\rho}(q_j)) \neq \varnothing,\ \text{for some } |m| < J\} \leqslant N$.*

We now return to the proof of Lemma 2.3. To simplify notation, let $\rho_1 = \rho_0/K$ and let $\rho_2 = 4\rho_0$, where ρ_0 is defined after the statement of Lemma 2.1. Thus $\rho_1 < \rho_0 < \rho_2$ In this notation, we have that $R(U_{\rho_2}(p)) > J$, for all p in A_r.

For $p \in M$ and $\rho > 0$, let $T_\rho(p)$ be the connected component of p in $\varphi_p(B^{u+s}(0,\rho))$. For $d(p,q)$ small enough, the maps $\varphi_q^{-1}\varphi_p$ distort the Euclidean structure by a factor $\leqslant 1.5$. Assume that ρ_2 is small enough that this distortion bound holds whenever $p, q \in U_{\rho_2}(z)$, for any z. From this we obtain that for all $p \in M$ and all $q \in T_{\rho_1}(p)$,

$$V_{2\rho_0}(q) \subset U_{\rho_1 + 3\rho_0}(p) \subset U_{\rho_2}(p). \tag{2}$$

We now apply Corollary 2.5 with $C = 4K$, $A = A_r$, and $\rho = \rho_1$. By Corollary 2.5, there exists $N > 0$ and $q_1, \dots, q_k \in A_r$ such that

- $A_r = U_{\rho_1}(q_1) \cup \dots \cup U_{\rho_1}(q_k)$, and
- $\#\{j \mid U_{\rho_2}(q_i) \cap f^m(U_{\rho_2}(q_j)) \neq \varnothing,\ \text{for some } |m| < J\} \leqslant N$.

For $i = 1, \ldots k$, let $U_i = U_{\rho_1}(q_i)$. The neighborhoods $U_1, \ldots, U_k$ cover A_r.

We choose $p_1, \ldots p_k$ inductively. Let $p_1 = q_1$. Since $V_{2\rho_0}(p_1) \subset U_{\rho_2}(p_1)$, and $R(U_{\rho_2}(p_1)) > J$, we have that

$$V_{2\rho_0}(p_1) \cap f^m(V_{2\rho_0}(p_1)) = \varnothing,$$

for $0 < |m| \leqslant J$.

Fix $i > 1$, and suppose that the points $p_1, \ldots, p_{i-1}$ have already been chosen. We want to choose p_i so that

$$V_{2\rho_0}(p_i) \cap f^m(V_{2\rho_0}(p_i)) = \varnothing,$$

for $0 < |m| \leqslant J$, and

$$V_{2\rho_0}(p_i) \cap f^m(V_{2\rho_0}(p_j)) = \varnothing,$$

for $0 \leqslant |m| \leqslant J$ and $j < i$. The first of these two properties is satisfied if we choose p_i so that $V_{2\rho_0}(p_i) \subseteq U_{\rho_2}(q_i)$. By (2), this is turn will hold if we choose $p_i \in T_{\rho_1}(q_i)$.

Hence, we would like to find $p_i \in T_{\rho_1}(q_i)$ such that

$$V_{2\rho_0}(p_i) \cap f^m(V_{2\rho_0}(p_j)) = \varnothing, \tag{3}$$

for $0 \leqslant |m| \leqslant J$ and $j < i$. The neighborhood $U_{\rho_2}(q_i)$ meets at most N sets of the form $f^m\left(U_{\rho_2}(q_j)\right)$, for m between $-J$ and J. Thus, $\#\mathcal{J}_i \leqslant N$, where $\mathcal{J}_i$ is the collection of all (j, m) with $j < i$, $|m| < J$, and

$$U_{\rho_2}(q_i) \cap f^m(U_{\rho_2}(q_j)) \neq \varnothing.$$

For $q \in M$ and $|m| \leqslant J$, let $V_\rho^m(q)$ be the connected component of $f^m(q)$ in $U_\rho(f^m(q)) \cap f^m(V_1(q))$. There exists $C_0 \geqslant 1$ such that, for all $p, q \in M$,

– $V_{C_0\rho_2}^m(q) \supseteq f^m(V_{2\rho_0}(q))$, and

– if $U_{\rho_2}(p) \cap f^m(U_{\rho_2}(q)) \neq \varnothing$, and ρ_0 is sufficiently small, then $V_{C_0\rho_2}^m(q)$ intersects $T_{C_0\rho_1}(p)$ in exactly one point.

(It is not hard to see that C_0 can be chosen to depend only on J, on the Riemannian structure, and on $\|Tf|_{E^c}\|$.) For $(j, m) \in \mathcal{J}_i$, let $p'_{j,m}$ be the point of intersection of $V_{C_0\rho_2}^m(p_j)$ and $T_{C_0\rho_1}(q_i)$:

$$\{p'_{j,m}\} = V_{C_0\rho_2}^m(p_j) \cap T_{C_0\rho_1}(q_i).$$

Consider the collection of these points

$$\mathcal{P}_i = \{p'_{j,m} \mid (j, m) \in \mathcal{J}_i\} \subset T_{C_0\rho_1}(q_i).$$

The points in $\varphi_{q_i}^{-1}(\mathcal{P}_i)$ lie on $B^{u+s}(0, C_0\rho_1) \subset \mathbb{R}^n$. By elementary Euclidean geometry, there exists $C_1 > 0$ such that, for any $\rho > 0$, and any finite collection of points $\mathcal{Q} \subset B^{u+s}(0, C_0\rho)$, there is a point $v \in B^{u+s}(0, C_0\rho)$ whose distance to the points in $\mathcal{Q}$ is at least $\rho/(C_1\#\mathcal{Q})$.

Applying this fact to the points in $\varphi_{q_1}^{-1}(\mathcal{P}_i)$, we find that there exists a point $p_i \in T_{\rho_1}(q_i) \subset U_i$, such that, for all $p \in \mathcal{P}_i$,

$$\|\varphi_{q_i}^{-1}(p_i) - \varphi_{q_i}^{-1}(p)\| \geqslant \rho_1/(C_1\#\mathcal{P}_i) \geqslant \rho_1/(C_1N). \tag{4}$$

We claim that if ρ_0 is sufficiently small, then p_i satisfies (3); that is, for all $j < i$ and $|m| \leqslant J$,

$$V_{2\rho_0}(p_i) \cap f^m\left(V_{2\rho_0}(p_j)\right) = \varnothing.$$

Clearly the claim is true for those (m, j) such that $U_{\rho_2}(q_i) \cap f^m\left(U_{\rho_2}(q_j)\right) = \varnothing$, so suppose that $(m, j) \in \mathcal{J}_i$. We show that

$$V_{2\rho_0}(p_i) \cap V^m_{C_0\rho_2}(p_j) = \varnothing,$$

which clearly implies the result.

We shall view everything in $\mathbb{R}^n$. Under the map $\varphi_{p_i}^{-1}$, the sets $V_{2\rho_0}(p_i)$ and $V^m_{C_0\rho_2}(p_j)$ lift, respectively, to $B^c(0, 2\rho_0)$ and a set we'll call W. We show that $B^c(0, 2\rho_0)$ and W are disjoint, for ρ_0 sufficiently small.

The set $V^m_{C_0\rho_2}(p_j)$ is a C^1 disk, tangent at $f^m(p_j)$ to $E^c(f^m(p_j))$. Thus W is a C^1 disk, tangent at a point $w_1 \in W$ to the uniformly continuous distribution $T\varphi_{p_i}^{-1}(E^c)$. Furthermore, the distribution $T\varphi_{p_i}^{-1}(E^c)$ coincides at p_i with $\mathbb{R}^c$.

By the distortion bounds, the diameter of W is on the order of ρ_2:

$$\text{diam}\,(W) \leqslant 3C_0\rho_2. \tag{5}$$

Let $w_2 = \varphi_{p_i}^{-1}(p'_{j,m}) \in W$. Combining the distortion bounds with (4), we obtain that

$$\|w_2\| = \|\varphi_{p_i}^{-1}(p_i) - \varphi_{p_i}^{-1}(p'_{j,m})\| \geqslant 2\rho_1/(3C_1N). \tag{6}$$

All of these statements – about the C^1-smoothness of W, the continuity of the distribution $T\varphi_{p_i}^{-1}E^c$, etc. – hold uniformly over p_i, ρ_0 and $|m| < J$. Thus, to summarize the preceding remarks, we have a constant $C_2 > 0$, and functions $\theta_1, \theta_2 : \mathbb{R}_+ \to \mathbb{R}_+$, all independent of p_i, ρ_0 and m, such that W is contained in the graph of a C^1 function $F : B^c(0, C_2\rho_0) \to \mathbb{R}^{u+s}$, with:

(1) $\|DF(x_1)\| \leqslant \theta_1(\|x_1\|)$, for some $x_1 \in B^c(0, C_2\rho_0)$,
(x_1 corresponds to the point $w_1 \in W$),

(2) $\|F(x_2)\| \geqslant \rho_0/C_2$, for some $x_2 \in B^c(0, C_2\rho_0)$,
(x_2 corresponds to the point w_2),

(3) $\|F(y) - F(x) - DF(x)(y - x)\| \leqslant \theta_2(\|y - x\|)$, for all $x, y \in B^c(0, C_2\rho_0)$,

(4) $\lim_{r\to 0} \theta_1(r) = 0$, and $\lim_{r\to 0} \theta_2(r)/r = 0$.

We claim that if ρ_0 is small enough, then $\|F(x)\| > 0$, for all $x \in B^c(0, C_2\rho_0)$. This implies that W is disjoint from $B^c(0, 2\rho_0)$, which is the desired result. By (3), we have that for all $x \in B^c(0, 2\rho_0)$,

$$\begin{aligned}\|F(x) - F(x_1)\| &\leqslant \|DF(x_1)(x - x_1)\| + \theta_2(\|x - x_1\|)\\ &\leqslant 2C_2\,\rho_0\theta_1(2C_2\rho_0) + \theta_2(2C_2\rho_0).\end{aligned}$$

By the triangle inequality,

$$\begin{aligned}\|F(x)\| &\geqslant \|F(x_2)\| - \|F(x_2) - F(x_1)\| - \|F(x_1) - F(x)\|\\ &\geqslant \rho_0/C_2 - 4C_2\rho_0\,\theta_1(2C_2\rho_0) - 2\theta_2(2C_2\rho_0).\end{aligned}$$

If ρ_0 is sufficiently small, then this quantity is positive. □

Proof of Lemma 1.3. — We want to pass from infinitesimal conditions given in the definition of partial hyperbolicity to local conditions. To this end, let

$$\overline{a}_p(r) = \max_{q\in B_r(p)} \|T_qf|_{E^s(q)}\|,$$

$$\overline{b}_p(r) = \min_{q\in B_r(p)} m(T_qf|_{E^c(q)}),$$

$$\overline{B}_p(r) = \max_{q\in B_r(p)} \|T_qf|_{E^c(q)}\|,$$

$$\overline{A}_p(r) = \min_{q\in B_r(p)} m(T_qf|_{E^u(q)}).$$

By continuity of Tf we can choose $r_0 > 0$ and $\theta < 1$ so that $\overline{a}_p(r_0) < \theta\overline{b}_p(r_0)$ and $\overline{A}_p(r_0) < \theta\overline{B}_p(r_0)$. From now on we will fix this r_0 and write $\overline{a}_p$ etc. instead of $\overline{a}_p(r_0)$ etc.

Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two continuous foliations with C^1 leaves on M. We say that $\mathcal{F}_2$ is ε (C^0–) close to $\mathcal{F}_1$ if given any p, q on the same leaf of $\mathcal{F}_1$ with the leafwise distance at most 1 apart, the $\mathcal{F}_2$ leaf passing through p intersects the ε ball centered at q.

It is clear that there exists $\varepsilon > 0$ such that g is stably accessible modulo $\beta\mathcal{D}$ provided that $\mathcal{W}^u_g$ is ε close to $\mathcal{W}^u_f$ and $\mathcal{W}^s_g$ is ε close to $\mathcal{W}^s_f$. So we need to show that given $\varepsilon > 0$ there is $\sigma > 0$ such that the conditions of Lemma 1.3 with this σ imply ε closeness of dynamical foliations of g to those of f. Namely we prove that for all $p, q \in B^s_{r_0/2}(f,p)$, the intersection $W^s(g,p)\bigcap B_\varepsilon(q)$ is non-empty. (Thus, 1 in the initial definition is replaced by $r_0/2$ but this is sufficient because the unit ball can be covered by a finite number of $r_0/2$-balls.) Let $\alpha_p(n)$ denote

$$\alpha_p(n) = \alpha_p\alpha_{fp}\dots\alpha_{f^{n-1}p}$$

where α is either of $\overline{a}, \overline{A}, \overline{b}, \overline{B}$.

Partial hyperbolicity implies that given $\eta \in (\theta, 1)$ there is a continuous cone family K^{cu} around $E^u \oplus E^c$ such that

(a) $Tf(K^{cu}(p)) \subset K^{cu}(fp)$,
(b) K^{cu} is uniformly transverse to E^s_f, and
(c) for any $v \in K^{cu}(p)$

$$\|Tf(v)\| > \overline{a}_p\eta\|v\|.$$

For δ_0 sufficiently small, K^{cu} will also satisfy (a)–(c) if f is replaced by any g such that $d_{C^1}(f,g) \leqslant \delta_0$. Let $q \in \mathcal{W}^s_f(p)$ and let $d_{\mathcal{W}^s}(p,q) \leqslant r_0/2$. Then $d(f^Np, f^Nq) \leqslant \overline{a}_p(N)$. Let V be a topological disk of dimension $\dim(E^u \oplus E^c)$ passing through q and such that TV belongs to K^{cu} (for example, we could take $V = \varphi_q(B^{u+c}(0,1))$).

Given n we can find σ so small that $d_{C^0}(f,g) < \sigma$ implies that

$$d(g^np, g^nV) < 2\overline{a}_p(n).$$

Since g^nV is uniformly transverse to E^s there exists $C = C(f)$ such that $\mathcal{W}^s_g(g^np)\bigcap g^nV$ contains a point z with $d(g^nq, z) \leqslant C\overline{a}_p(n)$. Hence $g^{-n}z \in \mathcal{W}^s(p)$ and $d(q, g^nz) \leqslant C\eta^n$. Thus, if n is large enough, $\mathcal{W}^s_g$ is $\varepsilon-$close to W^s_f. □

3. Local accessibility

Proof of Lemma 1.1. — Let f and $\delta < \delta_0$ be given. Assume δ_0 is small enough that any g within δ_0 of f in the C^1-metric remains partially hyperbolic, with functions a, b. Since our perturbations are local, it is convenient to adapt the structures we use to a neighborhood of a point p. To each $p \in M$ we shall associate:

(1) a neighborhood $U_p = \varphi_p(B^n(0,1))$,

(2) a C^∞ Riemann structure g_p on U_p with path metric d_p, isometric under φ_p^{-1} to the Euclidean metric on $B^n(0,1)$,

(3) a C^∞ splitting $TU = \widetilde{E}^u \oplus \widetilde{E}^c \oplus \widetilde{E}^s = T\varphi_p(\mathbb{R}^u \oplus \mathbb{R}^c \oplus \mathbb{R}^s)$, that agrees with $E^u \oplus E^c \oplus E^s$ at p,

(4) C^∞ foliations $\widetilde{\mathcal{W}}^u_p, \widetilde{\mathcal{W}}^s_p, \widetilde{\mathcal{W}}^{cu}_p, \widetilde{\mathcal{W}}^{cs}_p$ of U_p, tangent to the corresponding subbundles of the C^∞ splitting in (3),

(5) for $i = 1, \ldots, c$, partial flows $\zeta^i_t : U_p \to U_p$ tangent to the leaves of $\widetilde{\mathcal{W}}^c$,

(6) partial flows $\tau^u_t : U_p \to U_p$ and $\tau^s_t : U_p \to U_p$ tangent to the leaves of $\widetilde{\mathcal{W}}^u$, $\widetilde{\mathcal{W}}^s$, respectively.

We describe the construction of (5) and (6) in more detail. Let $\{e_1, \ldots, e_c\}$ be an orthonormal basis for the $\mathbb{R}^c$ factor in the splitting $\mathbb{R}^n = \mathbb{R}^u \oplus \mathbb{R}^c \oplus \mathbb{R}^s$. For $i = 1, \ldots c$, define the partial flows $\zeta^i_t : B \to B$ by

$$\zeta^i_t(\varphi_p(v)) = \varphi_p(v + te_i).$$

Similarly, fix unit vectors w^u and w^s tangent to the $\mathbb{R}^u$ and $\mathbb{R}^s$ factors in the splitting $\mathbb{R}^n = \mathbb{R}^u \oplus \mathbb{R}^c \oplus \mathbb{R}^s$, and define the partial flows $\tau^u_t, \tau^s_t : B \to B$, by

$$\begin{aligned} \tau^u_t(\varphi_p(v)) &= \varphi_p(v + tw^u), \text{ and} \\ \tau^s_t(\varphi_p(v)) &= \varphi_p(v + tw^s). \end{aligned}$$

Note that τ^u_t (resp. τ^s_t) sends $\widetilde{\mathcal{W}}^{cs}$ leaves (resp. $\widetilde{\mathcal{W}}^{cu}$ leaves) to other $\widetilde{\mathcal{W}}^{cs}$ leaves, and between $\widetilde{\mathcal{W}}^{cs}$ leaves is the exactly the $\widetilde{\mathcal{W}}^u$ (resp. $\widetilde{\mathcal{W}}^s$) holonomy map. Note that, where defined, $\tau^s_{-t}\tau^u_{-t}\tau^s_t\tau^u_t$ is the identity. This expresses the fact that $\widetilde{\mathcal{W}}^u$ and $\widetilde{\mathcal{W}}^s$ are jointly integrable. In Subsection 3.2, we will use the partial flows $\tau^u, \tau^s, \zeta^1_t, \ldots, \zeta^c_t$ to define g.

The next lemma follows directly from the uniform continuity of φ_p.

Lemma 3.1. — *The structures described in (1)-(6) are uniform over $p \in M$ and over g sufficiently C^1-close to f. For all $p \in M$, the structure g_p is uniformly comparable to the original Riemann structure on U_p.*

Since all estimates involving the Riemann structure on M in this paper are local, uniform over $p \in M$, any statement about the Riemann structure becomes valid for g_p by introducing a multiplicative constant. We will therefore be deliberately ambiguous in our notation, using d interchangeably for the Riemannian metric and the local metric d_p. Also, when the point is clear from the context, we will drop the subscript p in describing the various structures.

3.1. A criterion for stable accessibility. — We describe here a condition on the foliations $\mathcal{W}_g^u$, $\mathcal{W}_g^s$ that implies that g is stably accessible on a c-admissible disk for f.

Let D be a c-admissible disk for f centered at $p \in M$, and let $\rho = r(D)$. Let $N_r(D)$ denote the tubular neighborhood of D of radius r. Let $m = m(c, \dim(M))$ be the constant given by Lemma 3.10. Suppose that g is partially hyperbolic. We say that g is *θ-accessible on D* if, for each $i = 1, \ldots, c$, there exists a continuous map

$$H^i : [0,1] \times D \longrightarrow N_{(m-2)\rho}(2D)$$

with $t \mapsto H^i(t,q)$ a 4-legged us-path for g originating at q, and, for some $t_0 \in (0, \rho/2)$, the condition

$$d(H^i(1,q), \zeta^i_{t_0}(q)) < t_0\theta \tag{7}$$

holds for all $q \in D'$. Here ζ^i, d, etc. are the structures described in the previous section, adapted at p.

The next lemma gives a criterion for central accessibility. A basic element of the proof is the "quadrilateral argument" of Brin, in which 4-legged paths are homotoped to the trivial path along 4-legged paths with endpoint in a fixed c-admissible disk. The reader unfamiliar with this argument may consult the survey paper [**BuPuShWi**] for a detailed description; the case $c = 1$ is also proved there. The case $c > 1$ is essentially proved in [**ShWi**], using an index argument.

***Lemma 3.2* (Central accessibility criterion)**. — *Suppose $\beta > 1/2$. For every $\beta' \in (\beta, 1)$, there exist $\theta > 0$, $\delta_1 > 0$ and $\rho_0 > 0$ such that, for every c-admissible disk D of radius $r(D) < \rho_0$, if*

- *$d_{C^1}(f,g) < \delta_1$, and*
- *g is θ-accessible on $\beta' D$,*

then g is stably accessible on βD.

Proof of Lemma 3.2. — Let β, β' be given. Choose $\theta < (\beta' - \beta)/4\beta' c$. By continuity of the bundles $E_g^u(p)$ and $E_g^c(p)$ in (p,g), there exist $\delta_1 > 0$ and $\rho_0 > 0$ such that, if $d_{C^1}(f,g) < \delta_1$, if D is any c-admissible disk of radius $r(D) = \rho \leqslant \rho_0$, and if $s : [0,1] \to N_{(m-2)\rho}(D)$ is any 4-legged us-path for g with $s(0), s(1) \in D$, then

$$d(s(0), s(1)) \leqslant \rho(\beta' - \beta)/4c. \tag{8}$$

Let g, D be any such diffeomorphism and c-admissible disk. Suppose that g is θ-accessible on $\beta' D$. For $i = 1, \ldots c$, we have maps H^i satisfying (7) with $D' = \beta' D$.

We show g is accessible on βD. Since the existence of such H^i is a C^1-open condition, this implies that g is stably accessible on βD.

By varying the lengths of the last 2 legs in the path $t \mapsto H^i(t,q)$, we may arrange that that $H^i(1,q) \in D$, for all $q \in \beta' D$. The reader may verify that it is possible to do so while preserving property (7). (If necessary, the value of θ can be reduced a little).

By a standard argument, the path $t \mapsto H^i(t,q)$, for $q \in \beta' D$, can be homotoped through 4-legged us-paths originating at q to the trivial path so that the endpoints stay in D during the homotopy. The trace of these endpoints along the homotopy gives a curve in D from q to $H^i(1,q)$. More precisely, for $i = 1, \ldots, c$, we obtain

$$\Psi^i : [0,1] \times [0,1] \times \beta' D \longrightarrow N_\theta(D)$$

such that, for all $s \in [0,1]$, $t \mapsto \Psi^i(s,t,q)$ is a us-path for g with $\Psi^i(s,1,q) \in D$, $\Psi^i(0,t,q) = q$ and $\Psi^i(1,t,q) = H^i(t,q)$. Thus, $s \mapsto \Psi^i(s,1,q) =: \Phi^i_s(q)$ is a curve in D, from q to $H^i(1,q)$. Every point on this curve is the endpoint of a us-path originating at q.

By (8), we have, for $q \in \beta' D$:

$$\operatorname{diam}(\Phi^i([0,1] \times \{q\})) < \rho(\beta' - \beta)/4c. \tag{9}$$

For $q \in \beta' D$, we then extend the definition of $\Phi^i_s(q)$ to values of $s > 1$ by the inductive formula

$$\Phi^i_{t+m}(q) = \Phi^i_t(\Phi^i_m(q)),$$

for $t \in (0,1]$ and $m \in \mathbb{N}$. How far $\Phi^i_s(q)$ can be extended in s depends of course on q. Note, however, that (7) gives, for $m \in \mathbb{N}$,

$$\begin{aligned}
d(\Phi^i_m(q), \zeta^i_{mt_0}(q)) &= d(H^i(1, H^i(1, \ldots, H^i(1,q) \cdots)), \zeta^i_{mt_0}(q)) \\
&= d(H^i(1, q_{m-1}), \zeta^i_{t_0}(q'_{m-1})) \\
&\leqslant d(H^i(1, q_{m-1}), \zeta^i_{t_0}(q_{m-1})) + d(\zeta^i_{t_0}(q_{m-1}), \zeta^i_{t_0}(q'_{m-1})) \\
&\leqslant t_0(\beta' - \beta)/4\beta' c + d(q_{m-1}, q'_{m-1}) \\
&\cdots \\
&\leqslant m t_0(\beta' - \beta)/4\beta' c. \qquad (10)
\end{aligned}$$

As before, every point in the image of $\Phi^i_\cdot(q)$ is the endpoint of a us-path originating at q, although this path can have more than 4 legs.

Let $q_0 = \varphi(-\beta'\rho/2(e_1 + \cdots + e_c))$, and define a map $Z : [0, \beta'\rho]^c \to M$ by:

$$Z(a_1, \ldots, a_c) = \zeta^1_{a_1} \cdots \zeta^c_{a_c}(q_0).$$

Note that Z is a homeomorphism onto $\beta' D$.

Next, consider the the map $P : [0, \rho\beta']^c \to D$ defined by:

$$P(a_1, \ldots, a_c) = \Phi^1_{a_1/t_0} \Phi^2_{a_2/t_0} \cdots, \Phi^c_{a_c/r_1 t_0}(q_0).$$

Each point in the image of P is the endpoint of a us-path for g originating at q_0. We claim that D is in the interior of its image. Since Z is a homeomorphism onto $\beta' D$, it suffices to show that $d_{C^0}(P,Z) < d(\partial\beta D, \partial\beta' D) = \rho(\beta'-\beta)/2$.

If $a=(a_1,\dots,a_c)\in[0,\beta'\rho]^c$, with $a_i = t_0(m_i+s_i)$, $m_i\in\mathbb{N}$ and $s_i\in(0,1]$, then, by (9), and (10),

$$\begin{aligned}
d(P(a),Z(a)) &= d(\Phi^1_{a_1/t_0}\Phi^2_{a_2/t_0}\cdots\Phi^c_{a_c/t_0}(q_0), \zeta^1_{a_1}\zeta^2_{a_2}\cdots\zeta^c_{a_c}(q_0))\\
&\leqslant \sum_{i=1}^{c} d(\Phi^i_{a_i/t_0}\Phi^{i+1}_{a_{i+1}/t_0}\cdots\Phi^c_{a_c/t_0}(q_0), \zeta^i_{a_1}\Phi^{i+1}_{a_{i+1}/t_0}\cdots\Phi^c_{a_c/t_0}(q_0))\\
&\leqslant \sum_{i=1}^{c}\operatorname{diam}(\Phi([0,1]\times\{\Phi^{i+1}_{a_{i+1}/t_0}\cdots\Phi^c_{a_c/t_0}(q_0)\})\\
&\qquad+\sum_{i=1}^{c} d(\Phi^i_{m_i}\Phi^{i+1}_{a_{i+1}/t_0}\cdots\Phi^c_{a_c/t_0}(q_0), \zeta^i_{m_i t_0}\Phi^{i+1}_{a_{i+1}/t_0}\cdots\Phi^c_{a_c/t_0}(q_0))\\
&\leqslant c(\beta-\beta')/4c + \sum_{i=1}^{c} m_i t_0(\beta'-\beta)/4\beta' c\\
&< \rho(\beta'-\beta)/2,
\end{aligned}$$

since $t_0 m_i < \rho\beta'$. Then g is accessible on βD. □

3.2. Constructing the perturbation. — Fix $\beta'\in(\beta,1)$. The next lemma completes the proof of Lemma 1.1.

Lemma 3.3. — *For every $\delta,\theta>0$, if $r(\mathcal{D})$ is sufficiently small and $R(\mathcal{D})$ is sufficiently large, then there exists $g\in\operatorname{Diff}^r_\mu(M)$ such that*

(1) $d_{C^1}(f,g)<\delta$,
(2) $d_{C^0}(f,g)<\theta$
(3) *g is θ-accessible on $\beta' D$, for each $D\in\mathcal{D}$.*

The proof of Lemma 1.1 now follows. Choose $\theta<\sigma$ satisfying the hypotheses of Lemma 3.2. Let g be given by Lemma 3.3. Since g is θ-accessible on $\beta' D$, g is stably accessible on D. □

Proof of Lemma 3.3. — Let $\delta,\theta>0$ be given. We will perturb f by composing with a C^∞, volume preserving diffeomorphism $\psi: M\to M$. We first estimate the effect of the composition $\psi\circ f$ on the partially hyperbolic splitting.

Say that $\psi: M\to M$ is *supported on* $X\subset M$ if $\psi = id$ outside of $X\subset M$. The next lemma states that if $R(X)$ is sufficiently large, and $p,q\in X$ are sufficiently close, then for any g, with gf^{-1} supported on X, the subspaces $T\psi^{-1}(E^u_g)(q)$ and $E^s_g(q)$ are very close to $\widetilde{E}^u_p(q)$ and $\widetilde{E}^s_p(q)$, respectively.

***Lemma 3.4* (Bundle Perturbation Lemma)**. — *There exists $\delta_0>0$ such that the following is true. For every $\gamma>0$, there exists $J>0$ such that, if $\psi = g\circ f^{-1}$ is* *supported on a set X with $R(X) > J$, and $d_{C^1}(\psi, id) < \delta_0$, then for any $p, q \in X$ with $d(p,q) < J^{-1}$, we have:*

(1) $\angle_q(E^s_g, \widetilde{E}^s_p) \leqslant \gamma$, *and*

(2) $\angle_q(T\psi^{-1}(E^u_g), \widetilde{E}^u_p) \leqslant \gamma$.

Proof of Lemma 3.4. — Let $\gamma > 0$ be given. Recall that the splittings $TU_p = E^u \oplus E^c \oplus E^s$ and $TU_p = \widetilde{E}^u_p \oplus \widetilde{E}^c_p \oplus \widetilde{E}^s_p$ coincide at p.

By uniform continuity of the splitting $E^u \oplus E^c \oplus E^s$, uniformity of φ_p, and smoothness of ψ, there exists a continuous function $\theta_1 : \mathbb{R}_+ \to \mathbb{R}_+$, with $\theta_1(0) = 0$, such that, for all $p, q \in M$,

$$\angle_q(T\psi(E^u), T\psi(\widetilde{E}^u_p)) \leqslant \theta_1(d(p,q)) \tag{11}$$

$$\angle_q(E^s, \widetilde{E}^s_p) \leqslant \theta_1(d(p,q)), \tag{12}$$

provided $d_{C^1}(\psi, id)$ is small enough.

Let

$$\lambda = \max_p \left(\max\left(\frac{a_p}{b_p}\right), \left(\frac{B_p}{A_p}\right)\right)$$

and note that $\lambda < 1$, because f is partially hyperbolic. There exist $C_1, \theta_0 > 0$ such that, for all all subspaces F^u, F^s with

$$\max\{\angle(F^u, E^u),\ \angle(F^s, E^s)\} \leqslant \theta_0$$

we have:

$$\angle(Tf^{-j}(F^s), Tf^{-j}(E^s)) \leqslant C_1\lambda^j, \text{ and}$$
$$\angle(Tf^{j}(F^u), Tf^{j}(E^u)) \leqslant C_1\lambda^j,$$

for all $j \geqslant 0$. The splitting $E^u_g \oplus E^c_g \oplus E^s_g$ depends continuously on g, and so

$$\max\{\angle(E^u_g, E^u),\ \angle(E^s_g, E^s)\} \leqslant \theta_0,$$

if $d_{C^1}(\psi, id)$ (and so $d_{C^1}(f, g)$) is sufficiently small.

Fix positive $R < R(X)$. If $q \in X$, then $g^i(q) = f^i(q)$, for all i between 0 and R. For these q, we have

$$\begin{aligned}\angle_q(E^s_g, E^s) &= \angle_q(Tg^{-R}E^s_g, Tf^{-R}E^s)\\ &= \angle_q(Tf^{-R}E^s_g, Tf^{-R}E^s)\\ &\leqslant C_1\lambda^R.\end{aligned}$$

Similarly, for $q \in X$, $g^{-i}(q) = f^{-i}\psi^{-1}(q) = f^{-i+1}g^{-1}(q)$, for all i between 1 and $R-1$, and so

$$\begin{aligned}\angle_q(E^u_g, T\psi(E^u)) &= \angle_q(Tg(E^u_g), Tg(E^u))\\ &\leqslant C_2\angle_{g^{-1}(q)}(E^u_g, E^u)\\ &= C_2\angle_{g^{-1}(q)}(Tf^{R-1}E^u_g, Tf^{R-1}E^u)\\ &\leqslant C_1C_2\lambda^{R-1}.\end{aligned}$$

Combining these inequalities with (11), we have shown: there exist λ, $\theta_1 : \mathbb{R}_+ \to \mathbb{R}_+$, and $C > 0$ such that, for any ψ sufficiently close to the identity and supported on X, for all $R < R(X)$, and all $q \in X$, $p \in M$, we have

– $\angle_q(E^s_g, \widetilde{E}^s_p) \leqslant C(\lambda^R + \theta_1(d(p,q)))$, and
– $\angle_q(T\psi^{-1}(E^u_g), \widetilde{E}^u_p) \leqslant 2\angle_q(E^u_g, T\psi(\widetilde{E}^u_p)) \leqslant C(\lambda^R + \theta_1(d(p,q)))$.

Hence if R is sufficiently large and $d(p,q)$ is sufficiently small, these quantities are bounded by γ. □

We will also need the following lemma.

Lemma 3.5. — *There exists $T > 0$ such that, for $\varepsilon > 0$ sufficiently small, for any $p \in M$, and for any c-admissible disk D centered at p, there are C^∞, volume preserving flows $\xi^1_t, \dots, \xi^c_t : U_p \to U_p$ such that, for each i:*

(1) *$\xi^i_t = id$ outside $N_{2\varepsilon}(D)$.*
(2) *For $q \in N_\varepsilon(D)$,*
$$\xi^i_t(q) = \zeta^i_{\varepsilon t}(q).$$
(hence ξ^i_t preserves the leaves of $\widetilde{\mathcal{W}}^c \cap N_\varepsilon(D)$),
(3) $d_{C^1}(id, \xi^i_t) < T|t|$.

Proof of Lemma 3.5. — Let $G = \varphi_p^{-1}(D) = B^c(0,\rho)$, for some $\rho > 0$. Fix i, and let E be the divergence-free vector field on $N_{2\varepsilon}(G) \subset \mathbb{R}^n$ such that, for all v:
$$E(v) = \varepsilon e_i.$$
Let ω be the Euclidean volume form on $\mathbb{R}^n$, and let $\phi_0 = i_E\omega$. Since E is divergence-free, the $(n-1)$-form ϕ_0 is closed: $d\phi_0 = di_E\omega = \operatorname{div}(E)\omega = 0$. Since $N_{2\varepsilon}(G)$ is contractible, there exists an $(n-2)$-form ν on $N_{2\varepsilon}(G)$ such that $d\nu = \phi_0$. We may choose ν so that
$$\|\nu\| \leqslant 2\varepsilon^2, \text{ and } \|\nu\|_{C^1} \leqslant \varepsilon.$$
Let $\sigma : N_{2\varepsilon}(G) \to [0,1]$ be a C^∞ bump function, vanishing on a neighborhood of $\partial N_{2\varepsilon}(G)$ and identically 1 on $N\varepsilon(G)$, such that
$$\|d\sigma\| \leqslant 2/\varepsilon, \text{ and } \|d\sigma\|_{C^1} \leqslant 2/\varepsilon^2.$$
Let $\phi = d(\sigma\nu)$. Then
$$\begin{aligned}\|\phi\|_{C^1} &= \|d\sigma \wedge \nu + \sigma\phi_0\|_{C^1}\\ &\leqslant \|d\sigma\| \cdot \|\nu\|_{C^1} + \|d\sigma\|_{C^1} \cdot \|\nu\| + \|\sigma\|_{C^1} \cdot \|\phi_0\|_{C^1}\\ &\leqslant 8\end{aligned}$$
Hence ϕ has the following properties:

– $\|\phi\|_{C^1} \leqslant T$, where $T = 8$,
– $d\phi = 0$,
– $\phi = \phi_0$ on $N\varepsilon(G)$,
– $\phi = 0$ on $\partial N_{2\varepsilon}(G)$.

Let X be the vector field on $\mathbb{R}^n$ satisfying $i_X\omega = \phi$ and let X_t be the flow generated by X. Let $\xi^i_t = \varphi \circ X_t \circ \varphi^{-1}$. Then ξ^i has the desired properties. $\square$

Returning to the proof of Lemma 3.3, let T be given by Lemma 3.5. Let $\gamma = \theta\delta/100cT$. Choose $J > 0$ according to Lemma 3.4.

Now suppose that $\mathcal{D} = \{D_1, \dots, D_k\}$ is any c-admissible family with $R(\mathcal{D}) > J$ and $r(\mathcal{D}) < J^{-1}$. Choose $\eta < r(\mathcal{D})$ so that the η-neighborhoods of any two c-admissible disks in $\mathcal{D}$ are disjoint.

To prove Lemma 3.3, it suffices to show that for any $D \in \mathcal{D}$, there is a C^∞ volume preserving diffeomorphism ψ, supported on the η-neighborhood $N_\eta(D)$, such that

(1) $d_{C^1}(\psi, id) < \delta$,
(2) $d_{C^0}(\psi, id) < \theta$,
(3) if $\overline{f}$ is any diffeomorphism with $f^{-1}\overline{f}$ supported on $N_\eta(|\mathcal{D}|) \smallsetminus N_\eta(D)$, and $d_{C^1}(\overline{f}, f) < \delta$, then $\psi \circ \overline{f}$ is θ-accessible on $\beta' D$.

To construct the final diffeomorphism g, we proceed disk by disk, constructing for each $D_i \in \mathcal{D}$ a diffeomorphism ψ_i suppported on $N_\eta(D_i)$ so that $\psi_i \circ \psi_{i-1} \cdots \psi_i \circ f$ is θ- accessible on $\beta' D_i$. Then $g = \psi_k \cdots \psi_1 \circ f$ will satisfy the conclusions of Lemma 3.3.

Fix $D \in \mathcal{D}$ centered at p and choose $\varepsilon < \eta/4c$ small enough to satisfy the hypotheses of Lemma 3.5. Let the flows $\xi^1_t, \dots, \xi^c_t$ be given by Lemma 3.5.

For $i = 1, \dots, c$, let $\varepsilon_i = 4i\varepsilon$, let $Z_i = \tau^u_{\varepsilon_i}(D)$, and let

$$N_i = N_{2\varepsilon}(Z_i) = \tau^u_{\varepsilon_i}(N_{2\varepsilon}(D)).$$

The neighborhoods $N_1, \dots, N_c$ are pairwise disjoint. Define $\psi : M \to M$ by

$$\psi(q) = \begin{cases} \tau^s_{-\varepsilon_i}\tau^u_{\varepsilon_i}\tau^s_{\varepsilon_i}\xi^i_{\delta/T}\tau^u_{-\varepsilon_i}(q) & \text{if } q \in N_i, \text{ for some } i, \\ q & \text{otherwise.} \end{cases}$$

Observe that ψ has the following properties:

– ψ preserves μ,
– $\psi = id$ outside $N_1 \cup \cdots \cup N_c \subset N_\eta(D)$,
– ψ preserves the leaves of $\widetilde{\mathcal{W}}^c$ outside of $N_1 \cup \cdots \cup N_c$ and inside of $N_\varepsilon(Z_i)$, for $i = 1, \dots, c$.
– $d_{C^1}(\psi, id) < \delta$.

Let $\overline{f}$ be any diffeomorphism with $d_{C^1}(\overline{f}, f) < \delta$ and $f^{-1}\overline{f}$ supported on $N_\eta(|\mathcal{D}|) \smallsetminus N_\eta(D)$, and let $g = \psi \circ \overline{f}$. It remains to show that g is θ-accessible on $\beta' D$.

We will now examine the behavior of the holonomy maps for g along us-paths whose corners are near the points:

$$\tau^u_{\varepsilon_i}(p), \quad \tau^s_{\varepsilon_i}\tau^u_{\varepsilon_i}(p), \quad \tau^u_{-\varepsilon_i}\tau^s_{\varepsilon_i}\tau^u_{\varepsilon_i}(p).$$

In analogue with τ_t^u and τ_t^s, which translate along leaves of $\widetilde{\mathcal{W}}^u, \widetilde{\mathcal{W}}^s$, respectively, we introduce maps $\pi_t^u, \pi_t^s : N_\eta(\beta' D) \to B$, which translate along $\mathcal{W}_g^u$ and $\mathcal{W}_g^s$ leaves:

$$\{\pi_t^u(q)\} = \mathcal{W}_g^u(q) \cap \widetilde{\mathcal{W}}^{cs}(\tau_t^u(q))$$
$$\{\pi_t^s(q)\} = \mathcal{W}_g^s(q) \cap \widetilde{\mathcal{W}}^{cu}(\tau_t^s(q)).$$

If $d_{C^1}(f,g)$ is sufficiently small, these maps are well-defined for $|t| \leqslant \varepsilon_c$. Notice that if we were to replace $\mathcal{W}_g^u$ and $\mathcal{W}_g^s$ with $\widetilde{\mathcal{W}}^u$ and $\widetilde{\mathcal{W}}^s$, these equations would instead define τ_t^u and τ_t^s, respectively. Between $\widetilde{\mathcal{W}}^{cs}$ leaves, π_t^u is the $\mathcal{W}^u$-holonomy (and similarly for π_t^s). Lemma 3.4 will allow us to predict the behavior of these maps. The upshot is:

On the appropriate domains, $\pi^u \sim \psi\tau^u$ and $\pi^s \sim \tau^s$,

where we will be precise about $\sim$ later.

For $i = 1, \dots c$, let

$$h^i = \pi_{-\varepsilon_i}^s \pi_{-\varepsilon_i}^u \pi_{\varepsilon_i}^s \pi_{\varepsilon_i}^u.$$

Then h^i is a homeomorphism of $N_\varepsilon(\beta' D)$ onto its image. Observe that $h^i(q)$ is the endpoint of a 4-legged us-path for g, originating at q. By construction, these paths depend continuously on q, and so there are continuous maps

$$H^i : [0,1] \times \beta' D \longrightarrow N_\eta(D)$$

with $t \mapsto H^i(t,q)$ a 4-legged us-path for g, such that $H^i(0,q) = q$ and $H^i(1,q) = h^i(q)$. The rest of the argument goes as follows. We will show that $\pi^u \sim \psi\tau^u$ and $\pi^s \sim \tau^s$, which will imply that $h^i \sim \tau_{-\varepsilon_i}^s \psi \tau_{-\varepsilon_i}^u \tau_{\varepsilon_i}^s \psi \tau_{\varepsilon_i}^u$. Since τ^u and τ^s are just translations, $\psi = \zeta_{t_0}^i$ on $\tau_{\varepsilon_i}(D)$, and $\psi = id$ on $\tau_{-\varepsilon_i}^u \tau_{\varepsilon_i}^s \tau_{\varepsilon_i}(D)$, we find that $h^i \sim \zeta_{t_0}^i$, where $t_0 = \varepsilon\delta/T$. The remainder of the argument is devoted to making $\sim$ precise.

***Lemma 3.6* (Holonomy Perturbation Lemma)**. — *For ε and δ sufficiently small,*

$$d_{C^0}(h^i, \zeta_{t_0}^i) \leqslant t_0\theta$$

where $t_0 = \varepsilon\delta/T$, and the C^0-distance is measured on $\beta' D$.

Proof of Holonomy Perturbation Lemma 3.6. — For δ sufficiently small, there exists a neighborhood $Q \subset N_\varepsilon(\beta' D)$ of $\beta' D$ such that, for $i = 1, \dots, c$,

$$\pi_{\varepsilon_i}^u(Q) \subseteq \psi\tau_{\varepsilon_i}^u(N_\varepsilon(\beta' D)),$$
$$\pi_{\varepsilon_i}^s \pi_{\varepsilon_i}^u(Q) \subseteq \tau_{\varepsilon_i}^s \psi \tau_{\varepsilon_i}^u(N_\varepsilon(\beta' D)), \text{ and}$$
$$\pi_{-\varepsilon_i}^u \pi_{\varepsilon_i}^s \pi_{\varepsilon_i}^u(Q) \subseteq \tau_{-\varepsilon_i}^u \tau_{\varepsilon_i}^s \psi \tau_{\varepsilon_i}^u(N_\varepsilon(\beta' D)).$$

We now show that $d(h^i(q), \zeta_{t_0}^i(q)) \leqslant t_0\theta$, for all $q \in Q$ and $i = 1 \dots, c$, which implies the conclusion of the lemma.

From the definition of ψ, we write, for $q \in Q$,

$$\begin{aligned}\zeta^i_{t_0}(q) &= \xi^i_{t_0/\varepsilon}(q)\\ &= \xi^i_{\delta/T}(q)\\ &= \tau^s_{-\varepsilon_i}\tau^u_{-\varepsilon_i}\tau^s_{\varepsilon_i}\psi\tau^u_{\varepsilon_i}(q)\\ &= \tau^s_{-\varepsilon_i}\psi\tau^u_{-\varepsilon_i}\tau^s_{\varepsilon_i}\psi\tau^u_{\varepsilon_i}(q), \qquad (13)\end{aligned}$$

the final equality a consequence of the fact that ψ is supported on $N_1 \cup \cdots \cup N_c$, which is disjoint from $\pi^s_{\varepsilon_i}\pi^u_{\varepsilon_i}(Q)$. On the other hand,

$$h^i(q) = \pi^s_{-\varepsilon_i}\pi^u_{-\varepsilon_i}\pi^s_{\varepsilon_i}\pi^u_{\varepsilon_i}(q). \qquad (14)$$

We show that the corresponding factors in the two compositions (13) and (14) satisfy the desired inequality. More specifically, we show that, restricted to the appropriate domains, the distances $d_{C^0}(\pi^u_{\pm\varepsilon_i}, \psi\tau^u_{\pm\varepsilon_i})$ and $d_{C^0}(\pi^s_{\pm\varepsilon_i}, \tau^s_{\pm\varepsilon_i})$ are bounded by $\theta t_0/4$.

First, consider the maps $\psi \circ \tau^u_{\varepsilon_i}$ and $\pi^u_{\varepsilon_i}$ on the domain Q. Recall that, restricted to $\widetilde{\mathcal{W}}^{cs}$ leaves, $\tau^u_{\varepsilon_i}$ is the $\widetilde{\mathcal{W}}^u$ holonomy map. But $\psi\circ\tau^u_{\varepsilon_i}$ sends $\widetilde{\mathcal{W}}^{cs}$ leaves in $Q \subset N_\varepsilon(\beta' D)$ to $\widetilde{\mathcal{W}}^{cs}$ leaves: $\tau^u_{\varepsilon_i}$ sends $\widetilde{\mathcal{W}}^c$ leaves to $\widetilde{\mathcal{W}}^c$ leaves, and ψ preserves $\widetilde{\mathcal{W}}^{cs}$ leaves in $\tau^u_{\varepsilon_i}(Q) \subset N'_i$. It follows that, restricted to $\widetilde{\mathcal{W}}^{cs}(q) \cap Q$, the map $\psi \circ \tau^u_{\varepsilon_i}$ is the $\psi(\widetilde{\mathcal{W}}^u)$-holonomy map to the transversal $\widetilde{\mathcal{W}}^{cs}(\tau^u_{\varepsilon_i}(q))$, where $\psi(\widetilde{\mathcal{W}}^u)$ is the image of $\widetilde{\mathcal{W}}^u$ under ψ. Recall that $\pi^u_{\varepsilon_i}$ restricts to the $\mathcal{W}^u_g$-holonomy map between $\widetilde{\mathcal{W}}^{cs}(q) \cap N_\varepsilon(\beta' D)$ and $\widetilde{\mathcal{W}}^{cs}(\tau^u_{\varepsilon_i}(q))$.

Thus, between $\widetilde{\mathcal{W}}^{cs}$ leaves, we are comparing the holonomy maps for the foliations $\mathcal{W}^u_g$ and $\psi(\widetilde{\mathcal{W}}^u)$. To compare the holonomies for $\mathcal{W}^u_g$ and $\psi(\widetilde{\mathcal{W}}^u)$, we first apply the smooth change of coordinates $p \mapsto \psi^{-1}(p)$ and compare the holonomies for $\psi^{-1}(\mathcal{W}^u_g)$ and $\widetilde{\mathcal{W}}^u$. Since $d_{C^1}(\psi, id)$ is small, this change of coordinates distorts distances by a factor very close to 1.

The tangent distributions to $\psi^{-1}(\mathcal{W}^u_g)$ and $\widetilde{\mathcal{W}}^u$ are $T\psi^{-1}E^u_g$ and $\widetilde{E}^u$, respectively. According to the Bundle Perturbation Lemma 3.4, the distributions $T\psi^{-1}(E^u_g)$ and $\widetilde{E}^u$ are close; in particular,

$$\angle_q(T\psi^{-1}(E^u_g), \widetilde{E}^u) \leqslant \gamma = \theta\delta/100cT, \qquad (15)$$

for all $q \in N_\eta(D)$. We now apply the next elementary lemma.

Lemma 3.7. — *Let $\mathcal{F}$ be a continuous foliation of $B \subset U_p$ with C^1, u-dimensional leaves, transverse to $\widetilde{E}^s \oplus \widetilde{E}^c$. Let T_1 and T_2 be smooth disks tangent to $\widetilde{E}^s \oplus \widetilde{E}^c$. Assume that both the $\mathcal{F}$- and $\widetilde{\mathcal{W}}^u$-holonomy maps between T_1 and T_2 are defined, and denote them by $h^{\mathcal{F}}$ and $h^{\widetilde{\mathcal{W}}^u}$, respectively. Then, for all $q \in T_1$,*

$$d(h^{\mathcal{F}}(q), h^{\widetilde{\mathcal{W}}^u}(q)) \leqslant dist(T_1, T_2) \cdot \sup_{q \in B} \angle_q(T\mathcal{F}, \widetilde{E}^u).$$

The analogous statement holds for s-dimensional foliations transverse to $\widetilde{E}^s \oplus \widetilde{E}^c$.

Applying Lemma 3.7 to the foliation $\psi^{-1}\mathcal{W}_g^u$, and using inequality (15), we obtain that, for any two transversals $T_1 \subset Q$ and $T_2 = \tau_{\varepsilon_i}^u(T_1)$, and $q \in T_1$,

$$\begin{aligned} d(h^{\psi^{-1}(\mathcal{W}_g^u)}(q), h^{\widetilde{\mathcal{W}}^u}(q)) &\leqslant \operatorname{dist}(T_1, T_2) \cdot \sup_{q \in N_\eta(D)} \angle_q(T\psi^{-1}(E_g^u), \widetilde{E}^u) \\ &\leqslant \varepsilon_i \gamma \\ &\leqslant (4c\varepsilon)\gamma \\ &\leqslant (4c\varepsilon)(\theta\delta/100cT) \\ &< \theta(\varepsilon\delta/8T) \\ &= \theta t_0/8. \end{aligned}$$

But then, for all $q \in Q$,

$$\begin{aligned} d(\pi_{\varepsilon_i}^u(q), \psi\tau_{\varepsilon_i}^u(q)) &= d(\psi h^{\psi^{-1}(\mathcal{W}_g^u)}(q), \psi h^{\widetilde{\mathcal{W}}^u}(q)) \\ &\leqslant \operatorname{Lip}(\psi)\theta t_0/8 \\ &\leqslant \theta t_0/4. \end{aligned}$$

Similarly, for $q \in \psi\tau_{\varepsilon_i}^u(Q)$,

$$\begin{aligned} d(\pi_{\varepsilon_i}^s(q), \tau_{\varepsilon_i}^s(q)) &\leqslant 2\varepsilon_i \angle(E_g^s, \widetilde{E}^s) \\ &< \theta t_0/4. \end{aligned}$$

Combining these inequalities and using the fact that $\tau_{\varepsilon_i}^s$ is an isometry, we have, for all $q \in Q$,

$$\begin{aligned} d(\pi_{\varepsilon_i}^s\pi_{\varepsilon_i}^u(q), \tau_{\varepsilon_i}^s\psi\tau_{\varepsilon_i}^u(q)) &\leqslant d(\pi_{\varepsilon_i}^s\pi_{\varepsilon_i}^u(q), \tau_{\varepsilon_i}^s\pi_{\varepsilon_i}^u(q)) + d(\tau_{\varepsilon_i}^s\pi_{\varepsilon_i}^u(q), \tau_{\varepsilon_i}^s\psi\tau_{\varepsilon_i}^u(q)) \\ &= d(\pi_{\varepsilon_i}^s\pi_{\varepsilon_i}^u(q), \tau_{\varepsilon_i}^s\pi_{\varepsilon_i}^u(q)) + d(\pi_{\varepsilon_i}^u(q), \psi\tau_{\varepsilon_i}^u(q)) \\ &< \theta t_0/4 + \theta t_0/4 \\ &= \theta t_0/2. \end{aligned}$$

Proceeding in this fashion, we obtain that for $q \in Q$,

$$d(\pi_{-\varepsilon_i}^s\pi_{-\varepsilon_i}^u\pi_{\varepsilon_i}^s\pi_{\varepsilon_i}^u(q), \tau_{-\varepsilon_i}^s\psi\tau_{-\varepsilon_i}^u\tau_{\varepsilon_i}^s\psi\tau_{\varepsilon_i}^u(q)) < \theta t_0,$$

which completes the proof. □

This completes the proof of Lemma 3.3. We have now shown that for each i, there exists

$$H^i : [0,1] \times \beta' D \longrightarrow N_\eta(D)$$

with $t \mapsto H^i(t,q)$ a 4-legged us-path for g, such that $H^i(0,q) = q$ and

$$d(H^i(1,q), \zeta_{t_0}^i(q)) = d(h^i(q), \zeta_{t_0}^i(q)) < \theta t_0,$$

where $t_0 = \varepsilon\delta/T$. Hence, g is θ-accessible. □

3.3. The symplectic case. — If f preserves a symplectic form ω, then the perturbation g can also be made symplectic.

As in the proof for the volume preserving case, we begin with a local system of C^∞ charts $\varphi_p : B^{2m}(0,1) \to M$, defined for each $p \in M$, where $2m = n$. Similar to the volume preserving case, these charts can be chosen to have the following properties:

(1) $\varphi_p(0) = p$,

(2) $T_0\varphi_p$ sends the splitting $T_0\mathbb{R}^n = \mathbb{R}^u \oplus \mathbb{R}^c \oplus \mathbb{R}^s$ to the splitting $T_pM = E^u \oplus E^c \oplus E^s$,

(3) the symplectic form $\varphi_p^*\omega$ is a linear pullback of the standard symplectic form on $\mathbb{R}^{2m}$:

$$\varphi_p^*\omega = A_p^*(\sum dp_i \wedge dq_i),$$

for some linear map A_p on $\mathbb{R}^{2m}$,

(4) $p \mapsto \varphi_p$ is a uniformly continuous map from M to $C^1(B^n(0,1), M)$. The dependence of φ_p, A_p on f is also continuous.

By Darboux's theorem, for each $p \in M$, there exists a neighborhood U_p of p and coordinates $\kappa_p : U_p \to \mathbb{R}^{2m}$ such that, in these coordinates, ω takes the standard form $\sum dp_i \wedge dq_i$. For each p, $T_p\kappa_p$ sends the splitting $T_pM = E^u(p) \oplus E^c(p) \oplus E^s(p)$ to a splitting $\mathbb{R}^{2m} = \mathbb{R}^u_p \oplus \mathbb{R}^c_p \oplus \mathbb{R}^s_p$. Let $A_p : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ be a linear map that sends $B^{2m}(0,1)$ into $\kappa_p(U_p)$ and sends the trivial splitting $\mathbb{R}^{2m} = \mathbb{R}^u \oplus \mathbb{R}^c \oplus \mathbb{R}^s$ to $\mathbb{R}^{2m} = \mathbb{R}^u_p \oplus \mathbb{R}^c_p \oplus \mathbb{R}^s_p$, chosen to depend continuously on p, f. Then $\varphi_p = \kappa_p^{-1} \circ A_p$ satisfies properties (1)-(4).

With this modification, the proof of the Main Theorem in the symplectic case proceeds exactly as in the volume preserving one, replacing "μ" by "ω", until the proof on Lemma 3.3. Since we will modify slightly the statement and proof of this lemma, we restate it here in the symplectic case.

Lemma 3.8. — *For every $\delta, \theta > 0$, if $r(\mathcal{D})$ is sufficiently small and $R(\mathcal{D})$ is sufficiently large, then there exists $g \in \mathrm{Diff}^r_\omega(M)$ such that*

(1) *$d_{C^1}(f,g) < \delta$,*

(2) *$d_{C^0}(f,g) < \theta$, and*

(3) *each $D \in \mathcal{D}$ is covered by c-admissible disks $\beta V_1, \ldots, \beta V_k$ such that g is θ-accessible on V_i, for each i.*

Remark. — If θ and δ are sufficiently small, then any $g \in \mathrm{Diff}^r_\omega(M)$ satisfying conditions (1)-(3) in Lemma 3.8 is stably accessible on D; for then Lemma 3.2 implies that g is stably accessible on each βV_i, which implies stabe accessibility on their union, which contains D.

Proof of Lemma 3.8. — Let $\mathcal{D}$ be a c-admissible family. Using Lemma 3.10 below, we will cover each $D \in \mathcal{D}$ with c-disks $\beta V_1, \ldots, \beta V_k$. The lemma associates to each i

an open set (a union of balls) $N(D,i) \subset N_\eta(D)$; for different i, these sets are disjoint. We will then perturb inside of $N(D,i)$ to obtain θ-accessibility on V_i.

Similar to the volume preserving case, we will need to show that if $r(\mathcal{D})$ is sufficiently small and $R(\mathcal{D})$ is sufficiently large, then for each $D \in \mathcal{D}$ and each c-disk V_i in the cover of D, there is a symplectic C^∞ diffeomorphism ψ, supported on $N(D,i)$, with

(a) $d_{C^1}(\psi, id) < \delta$,
(b) $d_{C^0}(\psi, id) < \theta$,
(c) if $\overline{f}$ is any diffeomorphism with $f^{-1}\overline{f}$ supported on $N_\eta(|\mathcal{D}|) \setminus N(D,i)$ and $d_{C^1}(\overline{f}, f) < \delta$, then $\psi \circ \overline{f}$ is θ-accessible on V_i.

Each perturbation ψ is supported on a union of balls (as opposed to a tubular neighborhood); this allows for symplectic perturbations. The next lemma replaces Lemma 3.5 for the symplectic case.

Lemma 3.9. — *There exists $T > 0$ such that, for $\varepsilon > 0$ sufficiently small, for each $p \in M$ and $q \in B_{1/2}(p)$, there are C^∞, symplectic flows $\xi_t^i = \xi_t^{i,q} : M \to M$, $i = 1, \dots, c$, such that, for each i:*

(1) $\xi_t^i = id$ *outside* $B_{2\varepsilon}(q)$
(2) *For* $x \in B_\varepsilon(q)$,
$$\xi_t^i(x) = \zeta_{\varepsilon t}^i(x).$$
(hence ξ_t^i preserves the leaves of $\widetilde{\mathcal{W}}^c \cap B_\varepsilon(q)$),
(3) $d_{C^1}(id, \xi_t^i) < T|t|$.

Here, all balls $B_\rho(q)$ are measured in the d_p-metric, and all other invariant structures $\widetilde{\mathcal{W}}^c, \zeta^i$, etc. are adapted at p.

Proof of Lemma 3.9. — Let p, q, i be given and let $v = \varphi_p^{-1}(q) \in B^n(0, 1/2)$. We will explain how T is chosen later. Since constant vector fields are locally Hamiltonian with respect to $\varphi_p^*\omega = A_p^*(\sum dp_i \wedge dq_i)$, there exists a Hamiltonian vector field X^i, supported on $B^n(v, 1/2)$, such that $X^i = e_i$ on $B^n(v, 1/4)$. Since the C^1-size of φ_p is uniformly controlled, there exists a $T_0 > 0$, independent of p, q, i, such that $\|X^i\|_{C^1} < T_0$. Given $\varepsilon < 1/4$, let
$$Y^i(x) = \varepsilon X^i((x - v)/4\varepsilon).$$
Then Y^i is Hamiltonian (if X^i has Hamiltonian H, then Y^i has Hamiltonian $4\varepsilon^2 H((x-v)/4\varepsilon)$), is supported on $B^n(v, 2\varepsilon)$, and satisfies $\|Y^i\|_{C^1} \leqslant 4T_0$. Furthermore, $Y^i = \varepsilon e_i$ on $B^n(v, \varepsilon)$. The vector field $(\varphi_p)_* Y^i$ generates the desired flow ξ_i. Clearly T can be chosen to depend only on T_0 and other uniform data. $\square$

Next, we choose the balls. The proof of the next lemma is an elementary exercise in Euclidean geometry.

Lemma 3.10. — *There exists $m > 2$, depending only on c and $\dim(M)$, such that, for $\varepsilon > 0$ sufficiently small and all $p \in M$, there exist $k > 0$ and points*

$$\{q_{i,j} \mid i = 1, \ldots, c,\ j = 1, \ldots, k\} \subset N_{(m-2)\varepsilon}(D)$$

with the following properties:

(1) *there exist $p_1, \ldots, p_k \in D$ and $\varepsilon_{i,j} > 0$ such that $\tau^u_{-\varepsilon_{i,j}}(q_{i,j}) = p_j$,*
(2) *the balls in the collection*

$$\{B_{2\varepsilon}(q_{i,j}), B_{2\varepsilon}(\tau^s_{\varepsilon_{i,j}}(q_{i,j})), B_{2\varepsilon}(\tau^u_{\varepsilon_{i,j}}\tau^s_{\varepsilon_{i,j}}(q_{i,j})) \mid i = 1, \ldots, c,\ j = 1, \ldots, k\}$$

are pairwise disjoint,
(3) *the balls*

$$B_{\beta\varepsilon}(p_1), \ldots, B_{\beta\varepsilon}(p_k)$$

cover D.

Given δ and θ, let T be given by Lemma 3.9, let $\gamma = \theta\delta/100cT$, and choose J according to Lemma 3.4. Let $\mathcal{D}$ be any c-admissible family with $R(\mathcal{D}) > J$ and $r(\mathcal{D}) < J^{-1}$. Let $D \in \mathcal{D}$ with center p. Proceeding as in the proof of Lemma 3.3, choose $\varepsilon < \theta/4mc$ satisfying the hypotheses of Lemma 3.9, where m is given by Lemma 3.10. Fix $1 \leqslant i \leqslant c$, and let the points $\{q_{i,j}\}, p_i$ be given by Lemma 3.10. Let $V_i = V_\varepsilon(p_i)$; by Lemma 3.10 the disks $\beta V_1, \ldots, \beta V_k$ cover D. Let

$$N(D,i) = \bigcup\{B_{2\varepsilon}(q_{i,j}), B_{2\varepsilon}(\tau^s_{\varepsilon_{i,j}}(q_{i,j})), B_{2\varepsilon}(\tau^u_{\varepsilon_{i,j}}\tau^s_{\varepsilon_{i,j}}(q_{i,j})) \mid j = 1, \ldots c\}.$$

We show that properties (a)-(c) above are satisfied for this i.

In each ball $B_{2\varepsilon}(q_{i,j})$, let $\xi^{i,j}_t = \xi^{i,q_j}_t$ be the flow given by Lemma 3.9, with $q = q_j$.

Define $\psi : M \to M$ by

$$\psi(q) = \begin{cases} \xi^{i,j}_{\delta/T} & \text{if } q \in B_{2\varepsilon}(q_{i,j}),\ \text{for some } j, \\ q & \text{otherwise.} \end{cases}$$

Then ψ has the following properties:

- $\psi^*\omega = \omega$,
- $\psi = id$ outside $N(D,i)$,
- ψ preserves the leaves of $\widetilde{\mathcal{W}}^c$ outside of $\bigcup_j B_{2\varepsilon}(q_{i,j})$ and inside of $\bigcup_j B_\varepsilon(q_{i,j})$,
- $d_{C^1}(\psi, id) < \delta$.

Let $\overline{f}$ be any diffeomorphism with $f^{-1}\overline{f}$ supported on $N_\eta(|\mathcal{D}|) \smallsetminus N(D,i)$, and let $g = \psi \circ \overline{f}$. Let It remains to show that g is θ-accessible on V_i.

By the same argument as in the proof of the Main Theorem, we obtain that for each $x \in V_i$, there is a 4-legged us-path for g from x to a point $y \in N_{(m-2)\varepsilon}(D)$ such that

$$d(y, \zeta^i_{t_0}(x)) < \theta t_0,$$

In other words, g is θ-accessible on V_i. □

Acknowledgement. — We thank Keith Burns, François Ledrappier, Charles Pugh and Michael Shub for many useful conversations. Some of the ideas in this paper grew out of an attempt of the second author with Burns, Pugh and Shub to prove a C^2 version of the results here.

References

[BV] Bonatti, C. and M. Viana, *SRB measures for partially hyperbolic systems whose central direction is mostly contracting*, Israel J. Math. **115** (2000), 157–194.

[Ar] Arnaud, M.-C., *The generic sympletic C^1 diffeomorphisms of 4-dimensional sympletic manifolds are hyperbolic, partially hyperbolic or have a completely elliptic periodic point*, Prépublications d'Orsay 2000-17.

[BonDi] Bonatti, C. and L.J. Diaz *Persistent nonhyperbolic transitive diffeomorphisms*, Ann. of Math. **143** (1996), 357–396.

[Br] Brin, M., *Topological transitivity of one class of dynamical systems and flows of frames on manifolds of negative curvature*, Func. Anal. Appl. **9** (1975), 9–19.

[BrPe] Brin, M. and Ya. Pesin, *Partially hyperbolic dynamical systems*, Math. USSR Izvestija **8** (1974), 177–218.

[BuPuWi] Burns, K., C. Pugh and A. Wilkinson, *Stable ergodicity and Anosov flows*, Topology **39** (2000), 149–159.

[BuPuShWi] Burns, K., C. Pugh, M. Shub and A. Wilkinson, *Recent results about stable ergodicity*, Proc. Symposia AMS **69** (2001), 327–366.

[BuWi1] Burns, K. and A. Wilkinson, *Stable ergodicity of skew products*, Ann. Sci. École Norm. Sup. **32** (1999), 859–889.

[BuWi2] Burns, K. and A. Wilkinson, *Better center bunching*. in preparation.

[HiPuSh] Hirsch, M., C. Pugh and M. Shub, *Invariant manifolds*, Lecture Notes in Mathematics, **583**, Springer-Verlag, 1977.

[Lo] Lobry, C., *Controllability of nonlinear systems on compact manifolds,* SIAM J. Control **12** (1974) 1–4.

[Mo] Moser, J., *On volume elements on manifolds,* Trans. AMS. **120** (1965) 285–294.

[NiTö] Niţică, V. and A. Török, *An open dense set of stably ergodic diffeomorphisms in a neighborhood of a non-ergodic one*, Topology **40** (2001) 259–278.

[PP] Parry, W. and M. Pollicott, *Stability of mixing for toral extensions of hyperbolic systems.* Tr. Mat. Inst. Steklova **216** (1997), Din. Sist. i Smezhnye Vopr., 354–363.

[PugSh1] Pugh, C. and M. Shub, *Stable ergodicity and partial hyperbolicity*, in International Conference on Dynamical Systems: Montevideo 1995, a Tribute to Ricardo Mañé Pitman Res. Notes in Math. **362** (F. Ledrappier et al, eds.) 182-187.

[PugSh2] Pugh, C. and M. Shub, *Stably ergodic dynamical systems and partial hyperbolicity*, J. of Complexity **13** (1997), 125–179.

[PugSh3] Pugh, C. and M. Shub, *Stable ergodicity and julienne quasiconformality*, J. Eur. Math. Soc. **2** (2000) 1–52.

[ShWi] Shub, M. and A. Wilkinson, *Stably ergodic approximation: two examples*, Ergod. Th. and Dyam. Syst. **20** (2000) 875–893.

[Vi] Viana, M., *Dynamics: a probabilistic and geometric perspective* Proc. ICM–98, Documenta Math. Extra Vol. **I** (1998) 557–578.

D. Dolgopyat, Deparment of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park MD 20742 USA

A. Wilkinson, Deparment of Mathematics, Northwestern University, 2033 Sheridan Road, Evanston IL 60208 USA

Astérisque
287, 2003, p. 61–69

ANOSOV GEODESIC FLOWS FOR EMBEDDED SURFACES

by

Victor J. Donnay & Charles C. Pugh

Abstract. — In this paper we embed a high genus surface in $\mathbb{R}^3$ so that its geodesic flow has no conjugate points and is Anosov, despite the fact that its curvature cannot be everywhere negative.

1. Introduction

At the International Conference on Dynamical Systems held in Rio de Janeiro in July, 2000, Michael Herman asked whether the geodesic flow for an embedded surface in $\mathbb{R}^3$ can be uniformly hyperbolic, i.e., Anosov. Using techniques from our paper [**5**] and a suggestion of John Franks and Clark Robinson, we answer Herman's question affirmatively. The embedded surface looks like a spherical shell with many holes drilled through it. See Figures 1 and 2.

The Lobachevsky-Hadamard Theorem states that if a Riemann manifold has negative sectional curvature then its geodesic flow is Anosov. The celebrated thesis of Anosov [**1**] shows that this implies ergodicity, in fact the Bernoulli property, a stronger form of ergodicity.

In [**2**], Burns and Donnay showed that every surface M embeds in $\mathbb{R}^3$ so that its geodesic flow is Bernoulli; however, this cannot be a consequence of M having negative curvature. For a compact surface $M \subset \mathbb{R}^3$ necessarily has regions of positive curvature, the standard explanation being that there is a smallest sphere S which contains M, and there are points at which S is tangent to M. At these points the

2000 *Mathematics Subject Classification.* — 37D40.
Key words and phrases. — Anosov geodesic flow, no conjugate points, isometrically embedded surface, high genus.

The authors thank the Instituto de Matemática Pura e Aplicada for hospitality and support during part of the time this paper was written.

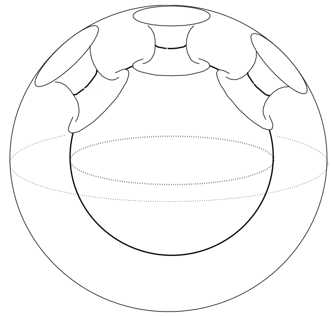

FIGURE 1. An embedded surface formed by connecting two concentric spheres with many tubes.

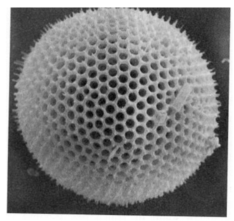

FIGURE 2. The radiolarian *Aulonia hexagona*, a marine micro-organism, as it appears through an electron microscope, by S.A. Kling.

curvature of M is positive. By continuity, the curvature of M is positive at nearby points too. The Bernoulli geodesic flows constructed by Burns and Donnay employ "focusing caps" to control the positive curvature. However, the caps are bounded by closed geodesics on which the curvature is zero, preventing uniform hyperbolicity. If the caps are perturbed to destroy these parabolic orbits the system can become non-ergodic [**3, 4**].

Instead of using caps, we use tubes of negative curvature together with the notion of a finite horizon geometry, which we introduced in [**5**], and are thereby able to show

Theorem A. — *There exist embedded surfaces in $\mathbb{R}^3$ for which the geodesic flows are Anosov.*

As an extension of Theorem A we discuss the immersed case, which has interest when the surface is not orientable.

Theorem B. — *There exist immersed non-orientable surfaces in $\mathbb{R}^3$ for which the geodesic flows are Anosov.*

The basic ingredient in our construction is illustrated in Figure 3; connect two flat tori (they are not embedded in $\mathbb{R}^3$) via a tube of negative curvature. The geodesic flow for this genus two surface is Bernoulli but not uniformly hyperbolic - since there

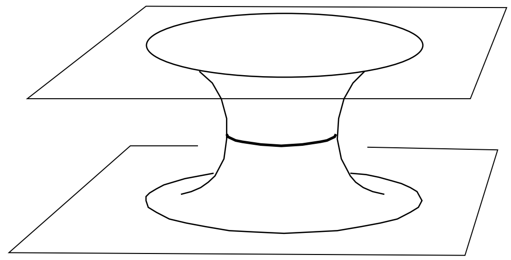

FIGURE 3. Two flat tori joined by a negatively curved tube.

are periodic geodesics lying completely in a flat region. If we now connect the two tori by enough tubes to produce a finite horizon pattern (see Section 2), i.e. every geodesic enters a tube in a bounded time, then the geodesic flow for this high genus surface is Anosov. To make an embedded Anosov example, we follow the suggestion of Franks and Robinson: reproduce the construction using very large and nearly flat concentric spheres instead of tori, again in a finite horizon pattern of tubes.

Remark. — Theorems A and B give the existence of high genus surfaces in $\mathbb{R}^3$ with Anosov geodesic flows, but we do not know a good lower bound on the genus. In [**6**], Wilhelm Klingenberg shows that no surface whose Riemann structure has conjugate points, which are produced by a surfeit of positive curvature, can have an Anosov geodesic flow. Hence our construction also provides examples of embedded surfaces without conjugate points. By Klingenberg's result, the sphere and torus never have Riemann structures whose geodesic flows are Anosov. So in particular, these surfaces cannot embed in $\mathbb{R}^3$ in such a way that their geodesic flows are Anosov. But what about the bitorus? Can it embed in $\mathbb{R}^3$ so that its geodesic flow is Anosov? Is it at least possible to embed the bitorus so that its metric has no conjugate points?

2. Finite Horizon

Let M be a surface equipped with a Riemann structure. A family $\mathcal{C}$ of curves $C_1, \dots, C_k$ in M gives M *ϕ-finite horizon* if every unit length geodesic crosses at least one curve in $\mathcal{C}$ at an angle $\geqslant \phi$. In [**5**] we show in detail how to choose $\mathcal{C}$ that gives M finite horizon, when M is a surface embedded in $\mathbb{R}^3$ and its Riemann structure is the one it inherits from the embedding. Here is an outline of the construction.

We first construct a fine, smooth triangulation of M whose triangles have uniformly bounded eccentricity and nearly geodesic edges. (The eccentricity of a triangle is the reciprocal of its smallest vertex angle.) We then draw small geodesic discs at the vertices of the triangulation, and a string of N "pearl discs" along each edge of the triangulation outside the vertex discs. Finally, we draw $2N + 2$ "wing discs" parallel to the string of pearl discs. Altogether this gives $9(N + 1)$ discs per triangle. The pearl and wing discs have radius r, which is much less than the radius R of the vertex discs, and this makes the pearl and wing discs along one edge of a triangle disjoint from those along a different edge.

Technically, once we have a bound on the eccentricity of the triangles that appear in our triangulations, we choose R and N. We then keep R and N fixed, while we dilate the surface by a factor of 2^n, $n \to \infty$, making ever finer triangulations of the dilated surface that have nearly linear triangles of roughly unit size. The radii r of the pearl and wing disks vary depending on the length of the edge of the triangle but lie in a compact interval.

With respect to the flat Riemann structure, the disc pattern for a triangle is shown in Figure 4. Every unit segment starting inside the flat triangle must cross the boundary circles of these discs at some positive angle. By compactness, they cross at some uniformly positive angle ϕ, a fact that remains true under small perturbations. For example, if we shrink all the discs by a factor $\mu < 1$, where $1-\mu$ is small, they still give the finite horizon property for unit segments. Similarly, the finite horizon property still holds if the flat metric is replaced by a nearly flat metric.

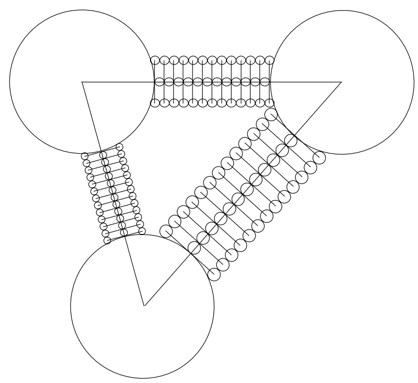

FIGURE 4. The pattern of discs for a linear triangle that gives the finite horizon property.

Denote by $2^n M$ the surface gotten by dilating M by a factor 2^n. The Riemann structure of $2^n M$ restricted to a nearly linear triangle T of roughly unit size is nearly flat. Thus, the geodesic discs of radius μr and μR laid down in the pattern of Figure 4 are disjoint and give the finite horizon property for unit geodesics on $2^n M$ when n is large.

We then flatten these disjoint geodesic discs by pushing each into the tangent plane at its center. Slightly smaller round discs lie in the flattened geodesic discs, and they still give the finite horizon property. The net effect is that the given surface M is replaced by a new one, $2^n M$, with diameter roughly 2^n, and having a great number of disjoint, flat *plateau discs* such that any unit geodesic crosses the boundary of at least one plateau disc at an angle $\geqslant \phi > 0$ for n large.

3. Dispersing Tubes

In [**5**] we glue "focusing caps" in place of the plateau discs above to make the geodesic flow non-uniformly hyperbolic. Here we glue tubes between pairs of plateau discs to make it Anosov.

Definition. — A *dispersing tube* T is a surface of revolution

$$T = \{(r, \theta, z) : r = h(z)\}$$

such that $h : [-1, 1] \to (0, 1]$ satisfies

(a) $h(z) = h(-z)$.

(b) $h(\pm 1) = 1$.

(c) If $|z| < 1$ then h is smooth and $h''(z) > 0$.

(d) The graph of h is infinitely tangent to the lines $z = \pm 1$. In particular $\lim_{z \to \pm 1} h'(z) = \pm\infty$.

Thus, T is a catenoid-like surface with its ends made planar. It has negative curvature. See Figure 3. The geodesics on a dispersing tube are simple to describe. There is the closed geodesic Γ around the "waist" of the tube, and there are geodesics asymptotic to it. Every other geodesic either enters and exits T without meeting Γ, or it crosses Γ once on its way from one end of T to the other. The entry and exit angles are equal because the tube is symmetric.

Note that independent linear scalings of z and r preserve the properties of T: it has negative curvature, it is infinitely tangent to the planes containing its boundary circles, and it contains a unique, closed waist geodesic Γ. Thus, we can make T long and thin, or short and broad. To avoid bending T, which may introduce positive curvature, we must be sure to keep its boundary circles in parallel planes.

4. The Perforated Sphere

Here is the proof of Theorem A. Take a sphere in $\mathbb{R}^3$ and make the finite horizon construction described in Section 2. (Any surface could be used instead of the sphere.) This gives a sequence of spheroids S_n of radius 2^n that contain many disjoint plateau discs of roughly unit radius. Each plateau disc lies in the plane normal to the radius vector from the origin. Then take a concentric copy of S_n, say S_n', which is S_n shrunk by the factor $1 - 1/2^n$. The spheroids have radii that differ by 1. As $n \to \infty$, this makes the plateau discs nearly equal in radius and parallel in pairs. Replace each pair of parallel plateau discs by a dispersing tube. The boundary circles of the dispersing tubes have radius equal to the plateau disc in S_n', which is slightly less than the corresponding radius in S_n. But as $n \to \infty$, the difference tends to 0, and consequently the slightly smaller discs on the outer spheroid continue to give the finite horizon property there. The dispersing tubes are roughly of unit size, so they all have roughly the same effect on geodesics passing through them.

The spheroids S_n, S'_n, with plateau disc pairs replaced by dispersing tubes is the perforated sphere $M = M_n$. There are three types of geodesics on M. The closed geodesic Γ around the waist of each tube, the geodesics that are asymptotic to these closed geodesics, and the geodesics that regularly enter and exit dispersing tubes at an angle $\geqslant \phi$.

Let φ be the geodesic flow for M. Its phase space is the unit tangent bundle SM. To show that φ is Anosov we consider the normal bundle N to the flow direction X. Then $T(SM) = N \oplus X$ is a $T\varphi$-invariant splitting. The normal bundle is given by $N = H \oplus V$, where H is the horizontal subspace and V is the vertical subspace. For $x \in SM$, let $P(x)$ be the standard, closed positive cone that consists of lines through the origin of $N(x)$ lying in the first and third quadrants with respect to $N = H \oplus V$. We claim that the positive cone field P is contracted uniformly into itself by the time one map $T\varphi_1$.

For $\xi \in N_x$, $\xi \neq 0$, let $u(t)$ denote the slope of the vector $T\varphi_t(\xi)$ with respect to the splitting $N = H \oplus V$ at $\varphi_t(x)$. Then u solves the Riccati equation

$$u' = -K(t) - u^2.$$

The vertical edge of the cone field is easily seen to be mapped inside the positive cone by a uniform amount under the time one map. Thus we need only examine the horizontal edge of the cone which corresponds to solutions of the Riccati equation with initial condition $u(0) = 0$. Henceforth, we restrict our attention to Riccati solutions with this initial condition.

Every unit geodesic has the following life. It experiences strictly negative curvature $K \leqslant \nu_0 < 0$ for at least a fixed time $t_0 > 0$ because it enters at least one tube at an angle $\geqslant \phi$, and it experiences curvature $K \leqslant \kappa_0$ for the rest of the time, where κ_0 is the maximum of the curvature on the surface. The positive curvature bound κ_0 becomes uniformly small when we take n large enough, while the negative bound ν_0 stays fixed, and the time bound t_0 stays fixed.

First, let us assume that the curvature on the surface is non-positive, so that $\kappa_0 = 0$. Then we can make the estimate that $u' = -K - u^2 \geqslant 0 - u^2$ which implies that $u(t = 1) > u_0 > 0$ for $u(t)$ any solution of the the Riccati solution along a geodesic on the surface M. The bound u_0 equals the value of the solution at $t = 1$ of the piecewise constant Riccati equation with $K(t) = \nu_0$ for $t \in [0, t_0]$ and $K(t) = 0$, for $t \in (t_0, 1]$.

By continuous dependence on parameters, if $\kappa_0 > 0$ is sufficiently small (i.e., if n is large), then $u(t = 1) > u_0/2 > 0$. Hence, the bottom edge of the cone is mapped into the cone by a uniform amount. We conclude that P is uniformly contracted into itself by the time one map $T\varphi_1$. This implies that

$$x \mapsto E^u(x) = \bigcap_{n=1}^{\infty} T\varphi_n(P(\varphi_{-n}x))$$

is a line field, and the restriction of $T\varphi_1$ to E^u is a uniform expansion. Symmetry implies that $T\varphi_{-1}$ contracts the negative cone field, and that it contains a line field E^s which is contracted by $T\varphi_1$. Thus, $T(SM) = E^u \oplus X \oplus E^s$ is an Anosov splitting for $T\varphi$, which completes the proof of Theorem A.

As a consequence of Theorem A we get the following stability result.

Corollary. — *There is a high genus surface M such that the set $\mathcal{E}$ of embeddings $M \to \mathbb{R}^3$ for which the geodesic flow is Anosov is non-empty and open. In particular there exist such embeddings of M that are analytic.*

Remark. — The earlier examples of ergodic geodesic flows for embedded surfaces [**2**], [**5**] are not stably ergodic as one can perturb the focusing cap to produce a "light-bulb" shaped cap which traps a positive measure set of trajectories and hence prevents ergodicity [**3**]. (Stable ergodicity means that the system and all small pertrurbations of it are ergodic.) Thus, our Anosov example above is the first geodesic flow for an embedded surface known to be stably ergodic.

Proof. — Theorem A asserts that $\mathcal{E} \neq \varnothing$. Uniform hyperbolicity is an open condition, so $\mathcal{E}$ is open with respect to the C^3 topology. The proof is completed by recalling that embeddings are open in the space of mappings, and analytic mappings are dense in the C^∞ topology. □

5. Non-orientable Surfaces

Here we show out how to construct a non-orientable immersed surface in $\mathbb{R}^3$ whose geodesic flow is Anosov, thereby proving Theorem B.

The simplest idea is to attach a Klein bottle or a Klein handle to the surface M constructed in Section 4. Doing so produces a certain amount of positive curvature, and it becomes unclear whether negative curvature continues to dominate.

A second idea is this. Take the previous pair of spheroids with the tubes joining them and select a pair of points p, q such that p is on the outer spheroid, q is on the inner spheroid, the points p, q are not near any of the tubes (they lie in the unused, middle portions of the triangles), and the segment $[p, q]$ passes through the origin. Then make plateaus at p and q and draw a long thin tube T of negative curvature from one plateau to the other, as shown in Figure 5. This causes the new surface to be non-orientable. (With just one perforation and the long tube, the surface is ambiently diffeomorphic to the standard immersion of the Klein bottle.) Every geodesic has the same type of behavior as before, except now it may spend a fair portion of its time in the flank of T where the curvature is barely negative. As $n \to \infty$, the negativity

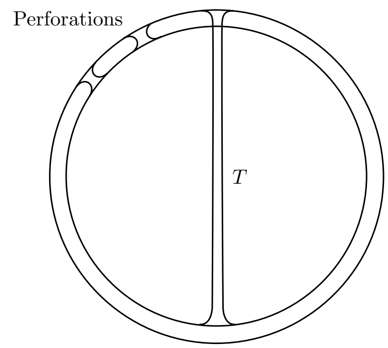

FIGURE 5. A long thin dispersing tube joining plateaus at p and q, shown in cross-section.

in T is on the order of $1/2^n$, which is the same as the worst positivity on the surface. Thus, it is not clear that negativity outweighs positivity enough to make the geodesic flow Anosov. More care may in fact validate this construction.

Here is what we do instead.

Choose a pair of adjacent triangles on the spheroid S_n constructed in Section 4. Form plateaus at their two centers based on a common plane, rather than different tangent planes at each center. Arbitrarily choose one of the two tangent planes to use. Do the same on the parallel spheroid S'_n. Hence all four plateaus are based on parallel planes. Make the tube connections between S_n and S'_n at all but these four new plateaus, and call the resulting surface M_1. Take a copy of M_1, say M_2, and rotate it to line up the four new unconnected plateaus from each surface, in parallel. Then draw tubes from one plateau disc to the other as shown in Figure 6.

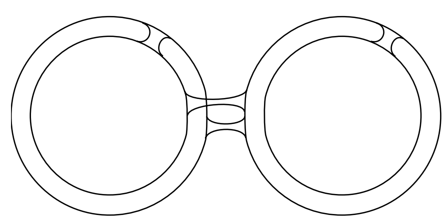

FIGURE 6. Cross-sectional view of connecting plateaus with tubes to make an immersed non-orientable surface whose geodesic flow is Anosov.

The resulting surface is non-orientable and the tubes all have roughly the same size. By the same sort of estimates as in the orientable case, negativity continues to dominate positivity and the geodesic flow remains Anosov.

Acknowledgement. — Figure 2 is reprinted from Haq et al, Editors, Introduction to marine micropaleontology, 1998, pp 203-244, S.A. King: ”Radiolaria”, with permission from Elsevier Science.

References

[1] D.V. Anosov (1967) “Geodesic Flows on Closed Riemann Manifolds with Negative Curvature”, Proceedings of the Steklov Institute of Mathematics, Number 90.

[2] K. Burns and V.J. Donnay (1997) “Embedded surfaces with ergodic geodesic flow”, International Journal of Bifurcation and Chaos, Volume 7, Number 7, pages 1509-1527.

[3] V.J. Donnay (1988) “Geodesic flow on the two-sphere, part I: positive measure entropy”, Ergodic Theory and Dynamical Systems 8, pages 531-553.

[4] V.J. Donnay (1996) “Elliptic Islands in Generalized Sinai Billiards”, Ergodic Theory and Dynamical Systems 16, pages 975-1010.

[5] V.J. Donnay and C.C. Pugh (2003) “Finite Horizon Riemann Structures and Ergodicity”, to appear in Ergodic Theory and Dynamic Systems.

[6] W. Klingenberg (1974) “Riemannian Manifolds With Geodesic Flow of Anosov Type”, The Annals of Mathematics, Second Series, Volume 99, Number 1, pages 1-13.

V.J. Donnay, Mathematics Department, Bryn Mawr College, Bryn Mawr, Pennsylvania, 19010
E-mail : `vdonnay@brynmawr.edu`

C.C. Pugh, Mathematics Department, University of California, Berkeley California, 94720
E-mail : `pugh@math.berkeley.edu` • *Url :* `http://people.brynmawr.edu/vdonnay/`

Astérisque
287, 2003, p. 71–87

NON-GIBBSIANNESS OF THE INVARIANT MEASURES OF NON-REVERSIBLE CELLULAR AUTOMATA WITH TOTALLY ASYMMETRIC NOISE

by

Roberto Fernández & André Toom

Abstract. — We present a class of random cellular automata with multiple invariant measures which are all non-Gibbsian. The automata have configuration space $\{0,1\}^{\mathbb{Z}^d}$, with $d > 1$, and they are noisy versions of automata with the "eroder property". The noise is totally asymmetric in the sense that it allows random flippings of "0" into "1" but not the converse. We prove that all invariant measures assign to the event "a sphere with a large radius L is filled with ones" a probability μ_L that is too large for the measure to be Gibbsian. For example, for the NEC automaton $(-\ln \mu_L) \asymp L$ while for any Gibbs measure the corresponding value is $\asymp L^2$.

1. Introduction

Studies of cellular automata and of their continuous-time counterpart, the spin-flip dynamics, have been successful in determining how many invariant measures the automaton or dynamics have. Much less is known about properties of these measures. A natural question is whether they are Gibbsian, that is whether they could correspond to measures describing the equilibrium state of some statistical mechanical system. There are two categories of evolutions —both with local and strictly positive updating rates— for which the answer is known to be positive: (1) If the updating prescription has a high level of stochasticity —*high noise regime*—, in which case Gibbsianness comes together with uniqueness of the invariant measure [**15, 19, 18**]; and (2) if the updating satisfies a detailed balance condition for some Boltzmann-Gibbs weights [**20**]. Known cases of non-Gibbsianness, on the other hand, refer to automata where the updating rates are either non-strictly positive [**16**], [**30**, Chapter 7] or non-local [**23**].

2000 ***Mathematics Subject Classification.*** — 37B15, 37C40, 60A10.
Key words and phrases. — Gibbs vs. non-Gibbs measures, cellular automata, invariant measures, non-ergodicity, eroders, convex sets.

R.F.: Travel support by an agreement USP-Cofecub (projet UC/68/99, *Comportement à temps intermédiaire des systèmes aléatoires*).
A.T.: Supported by FAPESP, grant # 98/15994-0 and CNPq, grant # 300991/1998-3.

In this paper we present some examples of stochastic *non-reversible automata* — that is, automata not satisfying any form of detailed balance—, with multiple invariant measures, all of them non-Gibbsian. Our class of automata can be seen as a generalization of the North-East-Center (NEC) majority model introduced in [**24**] and discussed in many papers. Its non-ergodicity was first proved in [**28**] (see also the discussion in [**15**]) and later by another method in [**2**]. Also it was simulated more than once [**1, 21, 22**]. Models of this sort are obtained by superimposing stochastic errors (noise) to deterministic automata having the so-called *eroder property*: finite islands of aligned spins, within a sea of spins aligned in the opposite direction, disappear in a finite time.

We allow only *one-sided* noise or stochastic error —a "0" can stochastically be turned into a "1", but not the reverse. Thus some of our transition rates are zeros and therefore the "dichotomy" result of [**20**, Corollary 1] is not applicable. Our work does not settle the long-standing issue of the Gibbsianness of the invariant measures of NEC models with non totally asymmetric noise. There are conflicting arguments and evidences for the model with symmetric noise: An interesting heuristic argument has been put forward [**30**, Chapter 5] pointing in the direction of Gibbsianness, and a couple of pioneer numerical studies yielded findings respectively consistent with Gibbsianness [**21**] and non-Gibbsianness [**22**]. However, we hope that the simple non-Gibbsianness mechanism clearly illustrated by our examples could be a useful guide and reference for the study of the more involved two-way-noise situation.

In our examples, non-Gibbsianness shows up in the same way as in the basic voter model [**16**]: Large droplets of aligned ("unanimous") spins have too large probability for the invariant measures to be Gibbsian. More precisely, we show that once a suitable "spider" of "1" appears, the dynamics causes the alignment of the spins in a neighboring sphere. This sort of damage-spreading property (or error-correcting deficiency) implies that the presence of a sphere of "1" is penalized by the invariant measures only as a sub-volume exponential. This contradicts well known Gibbsian properties. In fact, we can be more precise. Gibbsian measures are characterized by two properties [**13**]: uniform non-nullness and quasilocality. As we comment in Section 3, the large probability of aligned droplets means that the invariant measures can not be uniformly non-null. More generally, such invariant measures can not be the result of block renormalizations of non-null, in particular Gibbsian, measures. Furthermore, known arguments [**7**] (briefly reviewed in Section 3 below), imply that if one of these measures is not a product measure, then its non-Gibbsianness is preserved by further single-site renormalization transformations.

2. Simple examples

Before plunging into the technical and notational details needed to describe our results in full generality, we would like to present some simple examples that contain the essential ideas. The examples are defined on the configuration space $\{0,1\}^{\mathbb{Z}^2}$.

Example 1: The NEC model. — Its deterministic version is defined by a translation-invariant parallel updating defined by the rule

$$x_{\det}^{t+1}(0,0) = major\Big\{x^t(0,1)\,,\, x^t(1,0)\,,\, x^t(0,0)\Big\}, \tag{1}$$

where $x^t(i,j)$ denotes the configuration at site $(i,j) \in \mathbb{Z}^2$ immediately after the t-th iteration of the transformation and $major : \{0,1\}^{2k+1} \to \{0,1\}$ is the majority function, i.e. the Boolean function of any odd number of arguments, which equals "1" if and only if most of its arguments equal "1". This prescription yields an evolution, which is symmetric with respect to the flip $0 \leftrightarrow 1$ [a function with this property is called a *self-spin-flip function* in Section 4 below]. We consider a noisy version, where in addition spins "0" flip into "1" independently with a certain probability ε, while spins "1" remain unaltered. This corresponds to stochastic updating

$$\mathrm{Prob}\Big(x^{t+1}(i,j) = 0 \;\Big|\; x^t\Big) = (1-\varepsilon)\,\Big[1 - x_{\det}^{t+1}(i,j)\Big]. \tag{2}$$

The "all-ones" delta-measure δ_1 is invariant for this automaton. For small ε there is at least another invariant measure, as a consequence of Theorem 4.2 below.

Let us start with the following simple observations which are immediate consequences of the NEC rule (1) and the one-sidedness of the noise:

(i) Horizontal lines (parallel to axis i) filled with spins "1" remain invariant under the evolution.

(ii) The same invariance holds for vertical lines (parallel to axis j) filled with spins "1".

(iii) After one evolution-step (that is, after one parallel updating of all the spins), a line of slope -1 filled with spins "1" moves into the parallel line immediately to the south-west.

(iv) If the (infinite) "spider" formed by the i-axis, the j-axis and the line $i+j=0$ is filled with "1", then after t steps the evolution causes the whole triangle $\{(i,j) : i,j \leqslant 0,\ i+j \geqslant -t\}$ to be filled with "1".

The last observation can be visualized as a displacement, at speed 1, of the "front" formed by the line $i+j=0$, with a simultaneous displacement (here a trivial one), at speed 0, of the "fronts" formed by the i- and j-axis. This combined displacement produces a growing triangle full of "1".

The same observations hold if full lines are replaced by finite segments, except that, depending on the values of neighboring spins, in each iteration each segment can lose one or both of the "1" at its endpoints. We conclude that if at some time the spider

$$\begin{aligned}\mathrm{SP}_{(0,0),L} = \Big\{(i,0) \in \mathbb{Z}^2 : -8L \leqslant i \leqslant 4L\Big\} \bigcup \Big\{(0,j) \in \mathbb{Z}^2 : -8L \leqslant j \leqslant 4L\Big\}\\ \bigcup \Big\{(i,j) \in \mathbb{Z}^2 : i+j = 0\,,\, -6L \leqslant i \leqslant 6L\Big\}\end{aligned} \tag{3}$$

is filled with "1", then after $4L$ iterations the "1" fill a triangular region that contains the sphere of radius L centered at $(-L, L)$, to be denoted $S_{(-L,-L),L}$. Therefore, if μ

is a invariant measure,

$$\mu(1_{S_{(-L,-L),L}}) \geqslant \mu(1_{\mathrm{SP}_{(0,0),L}}) \geqslant \varepsilon^{3(12L+1)}. \tag{4}$$

We have denoted 1_Λ, for $\Lambda \subset \mathbb{Z}^2$, the event $\{x : x(i,j) = 1, (i,j) \in \Lambda\}$. The last inequality in (4) follows from the fact that a "1" has a probability at least ε to appear at a given site because of the noise. As commented in Section 3, such a probability is too large for the invariant measure to be Gibbsian, or block-transformed Gibbsian.

Example 2: North-South maximum of minima (NSMM). — The initial deterministic prescription is defined by

$$x_{\mathrm{det}}^{t+1}(0,0) = \max\Big\{\min\Big(x^t(0,0)\,,\, x^t(1,0)\Big)\,,\, \min\Big(x^t(0,1)\,,\, x^t(1,1)\Big)\Big\} \tag{5}$$

plus translation-invariance. The corresponding evolution is not symmetric under flipping, unlike the previous example. The stochastic version is obtained by adding one-sided noise as in (2). For small ε this automaton has more than one invariant measure (see comment after Theorem 4.2). One of them is, of course, the "all-ones" delta-measure δ_1.

The mechanism for non-Gibbsianness for this model is even simpler to describe than for the NEC model. Indeed, it suffices to observe that whenever a horizontal line is filled with "1", then in the next iteration these "1" survive and in addition the parallel line immediately to the South becomes also filled with "1". The same phenomenon happens for finite horizontal segments, except that each creation of a new segment filled with "1" can be accompanied by shrinkages of up to two sites (the spins at the endpoints) of all the previously created segments. We conclude that if the "spider" (which looks more like a snake in this case)

$$\widetilde{\mathrm{SP}}_{(0,0),L} = \Big\{(i,0) \in \mathbb{Z}^2 : -3L \leqslant i \leqslant 3L\Big\} \tag{6}$$

is filled with "1" at some instant, then $2L$ instants later the "1" will cover at least a square region that includes the sphere $S_{(0,-L),L}$. Arguing as for (4), we obtain for all invariant measures μ the bound

$$\mu(1_{S_{(0,-L),L}}) \geqslant \mu(1_{\widetilde{\mathrm{SP}}_{(0,0),L}}) \geqslant \varepsilon^{6L+1}, \tag{7}$$

which implies that μ is neither Gibbsian nor block-transformed Gibbsian.

A comment by A. van Enter (private communication) gives a colorful description of the mechanism acting in both preceding examples: "the spider fills his stomach faster ($\asymp L$ sites at a time) than his legs shrink ($\asymp 1$ sites at a time)".

Example 3: A non-example. — The automata defined by the deterministic prescription

$$\begin{aligned} x_{\mathrm{det}}^{t+1}(0,0) = \mathit{major}\Big\{&\min\Big(x^t(0,2),\, x^t(-1,2)\Big),\, \min\Big(x^t(2,0),\, x^t(2,-1)\Big),\\ &\min\Big(x^t(0,-1),\, x^t(-1,0)\Big)\Big\} \end{aligned} \tag{8}$$

followed by one-sided noise (2), also has multiple invariant measures; this follows from Theorem 4.2 (see the comment following this theorem). Nevertheless, neither the mechanism of Example 1 (travelling fronts), nor that of Example 2 (growing strips) are present, so the theory of the present paper does not apply.

3. Non-nullness and the probability of aligned spheres

We present in this section the key property used in our paper to detect non-Gibbsianness. To state it in its natural generality we introduce some definitions.

We consider a general space of the form $\Omega = \mathcal{A}^{\mathbb{Z}^d}$ where $\mathcal{A}$ is some finite set, equipped with the usual product σ-algebra. For $\Lambda \subset \mathbb{Z}^d$ and $z \in \Omega$ we denote z_Λ the cylinder

$$z_\Lambda = \{x \in \Omega : x_i = z_i, i \in \Lambda\}. \tag{9}$$

Definition 3.1. — A measure μ in Ω is said to have the *alignment-suppression property (ASP)* if there is a positive number C such that the inequality

$$-\ln \mu(z_\Lambda) \geqslant C \cdot |\Lambda| \tag{10}$$

holds for every configuration $z \in \Omega$ and for every finite set $\Lambda \subset \mathbb{Z}^d$.

All Gibbs measures have the ASP property, but many non-Gibbsian measures too. We construct now a general class of measures with this property by considering renormalized measures having suitable non-nullness features. For this we consider an auxiliary configuration space $\Omega_0 = S^{\mathbb{Z}^d}$. The single-site space S can be very general, not necessarily finite or even compact. We assume that there is a σ-algebra on S and consider the usual product Borel σ-algebra on Ω_0. A *renormalization transformation* from Ω_0 to Ω is a probability kernel $T(\,\cdot\,|\,\cdot\,)$ from Ω_0 to Ω. In words, $T(A\,|\,\omega)$ is the probability that, given a configuration $\omega \in \Omega_0$, the "renormalized" configuration is in A. This represents a general stochastic transformation while deterministic transformations are the special cases obtained via delta-like prescriptions $T(\,\cdot\,|\,\omega)$. A *block-renormalization transformation* is a transformation, for which probabilities factorize in the following sense: to every $i \in \mathbb{Z}^d$ there corresponds a finite set $B(i) \subset \mathbb{Z}^d$, called *block*, with the following properties:

(i) If two points are far enough from each other, the corresponding blocks are disjoint. That is, there is a positive d_0 such that if the distance between $k, \ell \in \mathbb{Z}^d$ is greater than d_0, then $B(k) \cap B(\ell) = \varnothing$ ($d_0 = 1$ for the renormalization transformations used in statistical mechanics, while $d_0 > 1$ for common cellular-automata transformations).

(ii) If $i_1, \dots, i_k$ are sites in $\mathbb{Z}^d$, and $a_1, \dots, a_k$ are values in $\mathcal{A}$, then

$$T\Big(\{x_{i_1} = a_1, \dots, x_{i_k} = a_k\} \;\Big|\; \omega\Big) = \prod_{j=1}^{k} \widehat{T}_{i_j}\Big(\{x_{i_j} = a_j\} \;\Big|\; \omega_{B(i_j)}\Big). \tag{11}$$

Our notation indicates that the functions $\widehat{T}_{i_j}(\{x_{i_j} = a_j\}|\cdot)$ depend only on the values of ω_ℓ for $\ell \in B(i_j)$ (i.e., they are measurable with respect to the σ-algebra generated by the cylinders with base in $B(i_j)$). Examples of such transformations include decimation (deterministic), Kadanoff transformations (stochastic), majority rule, sign fields and transitions of cellular automata (the last three can be deterministic or stochastic, depending on the setting). These transformations are well known in physics, their precise definitions can be found, for instance, in [**6**, Section 3.1.2].

The kernel T naturally induces a transformation at the level of measures: each probability measure ρ on Ω_0 is mapped into a probability measure ρT on Ω —the *renormalized measure*— defined by

$$\int_\Omega f(x)\,(\rho T)(dx) = \int_{\Omega_0} \Big[\int_\Omega f(x)\,T(dx|\omega)\Big]\,\rho(d\omega) \tag{12}$$

for all suitable f (e.g. continuous or non-negative measurable). For each measure ρ on Ω_0 and each block $B(i)$ let us consider the conditional probabilities $\rho(d\omega_{B(i)}\,|\,\omega_{\mathbb{Z}^d\smallsetminus B(i)})$. For a given transformation T we single out the set $\mathcal{P}_T$ of measures on Ω_0 that admit conditional probabilities such that

$$\min_{a\in\mathcal{A}}\ \inf_{i\in\mathbb{Z}^d}\ \inf_{\omega_{\mathbb{Z}^d\smallsetminus B(i)}} \int \widehat{T}\Big(\{x_i = a\}\ \Big|\ \omega_{B(i)}\Big)\cdot\rho\Big(d\omega_{B(i)}\ \Big|\ \omega_{\mathbb{Z}^d\smallsetminus B(i)}\Big) \geqslant \delta, \tag{13}$$

for some $\delta > 0$. We denote $\mathcal{P}$ the union of these families $\mathcal{P}_T$ over all block-renormalization transformations T. Here is our key characterization.

Theorem 3.1. — *Every measure in $\mathcal{P}$ has the alignment-suppression property.*

Proof. — Let T, ρ be such that $\mu = \rho T$. By property (ii) above, there exists a constant $\gamma > 0$ (proportional to d_0) such that for any $\Lambda \subset \mathbb{Z}^d$ there is a family of sites $i_1,\ldots,i_k \in \Lambda$ with $k \geqslant \gamma|\Lambda|$, all of which are far enough from each other and therefore the blocks $B(i_1),\ldots,B(i_k)$ are disjoint. We therefore have that for every $z \in \Omega$

$$\mu(z_\Lambda) = \int \rho\Big(\widehat{T}\big(\{x_{i_1} = z_{i_1}\}\ \big|\ \cdot\,\big)\ \Big|\ \omega_{\mathbb{Z}^d\smallsetminus B(i_1)}\Big) \prod_{j=2}^{k} \widehat{T}_{i_j}\big(\{x_{i_j} = z_{i_j}\}\ \big|\ \omega_{B(i_j)}\big)\,\rho(d\omega)$$

$$\leqslant (1-\delta) \int \prod_{j=2}^{k} \widehat{T}_{i_j}\big(\{x_{i_j} = z_{i_j}\}\ \big|\ \omega_{B(i_j)}\big)\,\rho(d\omega). \tag{14}$$

This inequality is an immediate consequence of condition (13). After k iterations of this procedure we obtain

$$\mu(z_\Lambda) \leqslant (1-\delta)^k \leqslant (1-\delta)^{\gamma|\Lambda|}. \tag{15}$$

□

The class $\mathcal{P}$ of measures is a very large class. It contains practically all block transformations of Gibbs measures with finite alphabet obtained via standard statistical mechanics prescriptions (decimation, Kadanoff, majority rule, etc), plus the measures generated by finite-time evolutions of usual cellular automata prescriptions. There is by now a vast literature about such measures —see, for instance, [**6, 18, 3**]; for recent reviews with many references see [**4, 10, 11, 8**]— showing that many of them are non-Gibbsian. In fact, the family $\mathcal{P}_I$, where I is the identity, includes all *uniformly non-null measures*. These are measures μ that have, for each finite region $\Lambda \subset \mathbb{Z}^d$, uniformly bounded conditional probabilities $\mu(d\omega_\Lambda \,|\, \omega_{\mathbb{Z}^d \smallsetminus \Lambda})$, that is, such that there exist $\delta_\Lambda > 0$ with

$$\min_{a_\Lambda \in \mathcal{A}^\Lambda} \inf_{\omega_{\mathbb{Z}^d \smallsetminus \Lambda}} \mu\Big(\{x_\Lambda = a_\Lambda\} \;\Big|\; \omega_{\mathbb{Z}^d \smallsetminus \Lambda}\Big) \geqslant \delta_\Lambda. \tag{16}$$

Here we have denoted $a_\Lambda = (a_i)_{\,i\in\Lambda}$. Gibbs measures are uniformly non-null — and in addition quasilocal (the finite-volume conditional probabilities are continuous functions of the external conditions $\omega_{\mathbb{Z}^d \smallsetminus \Lambda}$)— hence they also belong to $\mathcal{P}_I$. Property (13) seems to be more general than usual non-nullness, in particular it does not depend on the existence of a whole system of conditional probabilities.

The invariant measures of the automata of the present paper, on the other hand, do not have the alingment-suppression property, hence they do not belong to the class $\mathcal{P}$. They therefore can be neither Gibbsian nor uniformly non-null nor block-transformed Gibbsian. As further examples of measures without the ASP we mention the invariant measures of the basic voter model [**16**], the invariant measure of some non-local dynamics [**23**], and the sign-fields of massless Gaussians [**14, 5**], anharmonic crystals [**6**, Section 4.4] and solid-on-solid (SOS) models [**9, 17**].

For measures μ having a well defined relative entropy density $s(\,\cdot\,|\,\mu)$, the alignment-supression property (10) implies that $s(\delta_z\,|\,\mu) > 0$ for every periodic configuration $z \in \Omega$. The relative entropy density is known to exist for translation-invariant Gibbs measures [**12**, Chapter 15]. Recent work in [**25**] shows that it is also well defined for most translation-invariant measures obtained through block transformations of Gibbs measures. Because of this, the non-Gibbsianness resulting from the lack of ASP has often been interpreted as "too large probabilities of large deviations". The non-Gibbsianness (non-nullness) criterion obtained by falsifying Theorem 3.1, however, is a more general argument that needs neither translation invariance of μ nor the existence of the entropy density.

For completeness, we mention a further result obtained in [**7**].

Theorem 3.2. — *Suppose μ is a measure in Ω such that (i) it violates the ASP property for some periodic configuration $z \in \Omega$, and (ii) it is not a product measure. Then, for every single-site block-renormalization transformation T (i.e. a transformation defined by blocks $B(i)$ formed by only one site), the measure μT is not Gibbsian.*

This result follows from the fact that such a violation implies that $s(\delta_z|\mu) = 0$, which in its turn implies that $s(\delta_z T\,|\,\mu T) = 0$. If μT were Gibbs, then by a well known result [**12**, Theorem 15.37] the measure $\delta_z T$ would be Gibbs for an equivalent interaction. But this is impossible because the latter is a product measure and the former is not. Note that if T corresponds to a not-totally asymmetric noise, the measure μT is uniformly non-null. Hence its non-Gibbsianness would correspond to lack of quasilocality.

For the automata of this paper, we suspect that many of their invariant measures are non-product.

4. General Results

We now describe a large family of automata exhibiting a general version of the non-Gibbsianness mechanism of the first two examples in Section 2. Throughout the article we consider the d-dimensional integer space $\mathbb{Z}^d$ with $d > 1$ embedded into the d-dimensional real space $\mathbb{R}^d$ with the same axes and Euclidean norm $\|\cdot\|$. The configuration space is $\Omega = \{0,1\}^{\mathbb{Z}^d}$. We first need some definitions.

For any $i \in \mathbb{Z}^d$ we denote $\tau_i : \Omega \to \Omega$ the translation of Ω defined by $(\tau_i\, x)_j = x_{j-i}$. Any function $f : \Omega \to \{0,1\}$ will be called a transition function. Given any transition function f, we define the corresponding operator $D_f : \Omega \to \Omega$ by the rule

$$\forall\, i \in \mathbb{Z}^d \;:\; (D_f\, x)_i = f(\tau_i\, x). \tag{17}$$

We call $f : \Omega \to \{0,1\}$ *standard* if it has the following three properties:

1) f is *local*, i.e. there is a finite set $\Delta \subset \mathbb{Z}^d$ —the *support* of f— such that $f(x) \equiv f(x_\Delta)$. Given Δ, we denote $\|\Delta\|$ the maximum of $\|i\|$ for $i \in \Delta$.

2) f is *monotonic*, that is $(\forall\, i : x_i \leqslant y_i) \implies f(x) \leqslant f(y)$.

3) f is not a constant. (Otherwise our theorem is either trivially true if $f \equiv 1$ or trivially false if $f \equiv 0$.)

Since f is monotonic and non-constant,

$$f(\text{“all zeros”}) = 0 \quad\text{and}\quad f(\text{“all ones”}) = 1. \tag{18}$$

For any $x \in \Omega$ we denote its indicator $\mathrm{Ind}(x) = \{i \in \mathbb{Z}^d \mid x_i = 1\}$. Conversely, for any $S \subset \mathbb{Z}^d$ we denote $\mathrm{Conf}(S)$ that configuration, whose indicator is S.

Let us call an element of $\mathbb{R}^d$ a *direction* if its norm equals 1. For any direction p we call a *front* with this direction any configuration whose indicator has the form

$$\{i \in \mathbb{Z}^d \;\big|\; \langle i, p\rangle \leqslant C\}, \tag{19}$$

where C is a real number and $\langle\cdot,\cdot\rangle$ denotes the scalar product in $\mathbb{R}^d$. It is evident that for any standard f the operator D_f transforms any front (19) into a front with the same direction, C being substituted by $C + V_p$, where V_p does not depend on C. We call V_p the *velocity* of D_f in the direction p.

Let us call a configuration $x \in \Omega$ invariant for D_f if $D_f\, x = x$. Given $x,\ y \in \Omega$, we call y a finite deviation of x if the set of those $i \in \mathbb{Z}^d$, for which $y_i \neq x_i$, is finite. We say that an invariant configuration x attracts D_f if for any its finite deviation y there is a time t such that $D_f^{\,t}\, y = x$.

Let $\mathcal{M}$ denote the set of probability measures on Ω (on the σ-algebra generated by cylinder sets). For any $\varepsilon \in [0,1]$ we define one-sided noise $N_\varepsilon : \mathcal{M} \to \mathcal{M}$ as follows: when applied to a measure δ_x concentrated in a configuration $x = (x_i)$, it produces a product measure $N_\varepsilon\, \delta_x$, in which the i-th component equals 1 with a probability 1 if $x_i = 1$ and with a probability ε if $x_i = 0$.

Theorem 4.1. — *Take any standard f, such that "all ones" attracts D_f, and make any one of the following two assumptions:*

a) $V_p + V_{-p} \geqslant 0$ *for all directions* p.

b) *There is a direction p such that* $V_p + V_{-p} > 0$.

Then for any $\varepsilon > 0$ all the invariant measures of $N_\varepsilon\, D_f$ satisfy

$$-\ln \mu(\mathbf{1}_{S_{0,L}}) \prec L^{d-1}. \tag{20}$$

Here and in the sequel $f \prec g$ or $g \succ f$, for f and g positive functions means that there exists a constant $C > 0$ such that $g \geqslant C \cdot f$. If $f \prec g$ and $f \succ g$, we write $f \asymp g$.

If $\varepsilon = 0$, our theorem may be false, for example if D is the identity. Notice also that in the case b) our assumption that "all ones" attracts D_f is redundant because it follows from b).

Let us present some further considerations that clarify the statement of the theorem. Given any non-constant affine function $\phi : \mathbb{R}^d \to \mathbb{R}$ and two numbers $C_1 \leqslant C_2$, we call a *layer* any configuration $\mathrm{Conf}\{i \in \mathbb{Z}^d \,\big|\, C_1 \leqslant \phi(i) \leqslant C_2\}$. We call the thickness of this layer the distance between the hyperplanes $\phi = C_1$ and $\phi = C_2$, that is $(C_2 - C_1) / \|\phi\|$, where $\|\cdot\|$ is the norm. We call a layer thick-enough if its thickness is not less than $2\,\|\Delta\|$. We call the two normal unit vectors to hyperplanes $\phi = \mathrm{const}$ the directions of this layer. If f is standard, D_f transforms any thick-enough layer into a layer with the same directions, the thickness of the layer changing by $V_p + V_{-p}$. The condition a) of our theorem means that thickness of any thick-enough layer does not decrease and the condition b) means that thickness of some layer increases under the action of D_f.

Of the examples of Section 2, the NEC automaton satisfies condition a), while the NSMM automaton satisfies condition b) for $p = (0,1)$. For the non-example, however, $V_p + V_{-p} < 0$ for all directions p. In all the three cases f [given, respectively, by (1), (5) and (8)] is standard, and both "all zeros" and "all ones" attract D_f.

The NEC example is representative of a class of models with a further duality property. For any $x_i \in \{0,1\}$ we denote $\neg x_i = 1 - x_i$. Accordingly, if x is a configuration, $\neg x$ is another configuration such that $(\neg x)_i \equiv \neg(x_i)$. Any transition function f has an associated spin-flip [(1)] function denoted $\neg f$ and defined by the identity $\neg f(x) \equiv f(\neg x)$. Let us call f self-spin-flip if it coincides with its spin-flip. If f is

(1) In the theory of Boolean functions $\neg f$ is called dual, but we have to use another term because in the theory of random processes the word "duality" is used for another purpose.

standard and self-spin-flip, then $V_p + V_{-p} \equiv 0$, so the thickness of all layers does not change under the action of D_f, which provides many examples where out results can be applied. For example, the function *major*$(\cdot)$, described above, is self-spin-flip.

It is evident that under the hypothesis of Theorem 4.1, the measure δ_1 is invariant for any superposition $N_\varepsilon\, D_f$. Hence, the theorem is not trivial only if the automata have more than one invariant measure. This is ensured by the following theorem. Given f, let us call a set $S \subset \mathbb{Z}^d$ a one-set if $f(\mathrm{Conf}(S)) = 1$. Since one-sets belong to $\mathbb{Z}^d$, they belong to $\mathbb{R}^d$, where we can consider their convex hulls, the intersection of which is denoted σ_1. In the analogous way we call a set $S \subset \mathbb{Z}^d$ a zero-set if $f(\mathrm{Conf}(\mathbb{Z}^d - S)) = 0$ and denote σ_0 the intersection of their convex hulls.

Theorem 4.2. — *For any operator D_f defined by (17), where f is standard, the following four statements are equivalent:*

1) *$N_\varepsilon\, D_f$ has more than one invariant measure for some positive ε.*
2) *The configuration "all zeros" attracts D_f.*
3) *σ_0 is empty.*
4) *There are a natural number $m \leqslant d+1$ and m affine functions $\phi_1, \dots, \phi_m \;:\; \mathbb{R}^d \to \mathbb{R}$ such that:*

$$\begin{cases} \text{i) } \textit{for every } j \in [1,m] \textit{ the set } \{p \in \mathbb{Z}^d \;:\; \phi_j(p) \leqslant 0\} \textit{ is a zero-set.} \\ \text{ii) } \phi_1 + \cdots + \phi_m \equiv \mathrm{const} > 0. \\ \text{iii) } \textit{There is a rational point } p \in \mathbb{R}^d \textit{ such that } \phi_j(p) > 0 \textit{ for all } j \in [1,m]. \end{cases}$$

This theorem proves, in particular, that the three examples of Section 2 exhibit multiple invariant measures for ε small. Indeed, in the three cases the configuration "all zeros" attracts D_f.

5. Proof of Theorem 4.2

If we omit the condition iii) in 4), our Theorem 4.2 almost follows from theorems 5 and 6 and lemma 12 of [**29**]. However, there is some difference, so for the reader's convenience we completely deduce 4) from 3).

Suppose that σ_0 is empty. Every zero-set can be represented as an intersection of several zero-half-spaces, i.e. half-spaces, which are zero-sets, where a half-space is a subset of $\mathbb{R}^d$, where some non-constant affine function does not exceed zero. Thus there are several zero-half-spaces, whose intersection is empty. Everyone of them can be represented as $\{p \in \mathbb{R}^d \,\big|\, f_i(p) \leqslant 0\}$, where f_i are affine functions on $\mathbb{R}^d$. We can choose these functions so that they have no common direction of recession (that is, no direction p such that $f_i(p) \leqslant f_i(0)$ for all i), which allows us to apply to them Theorem 21.3 on page 189 of [**26**]. Since the intersection of our zero-half-spaces is empty, the case (a) of this theorem is excluded in the present situation, whence the case (b) takes place, which amounts to our conditions i) and ii) in 4), the products

$\lambda_i f_i$ mentioned in the case (b) serving as our ϕ_i. We may assume that our m is the minimal for which there are functions satisfying i) and ii). Based on this, let us prove statement iii) using the following lemma, which is a direct consequence of Theorem 21.1 on page 186 of [**26**]:

Lemma 5.1. — *Let $\phi_1, \ldots, \phi_m$ be affine functions on R^d. Then one and only one of the following alternatives holds:*

(a) *There exists some $x \in \mathbb{R}^d$ such that $\phi_1(x) > 0, \ldots, \phi_m(x) > 0$;*

(b) *There exist non-negative real numbers $\lambda_1, \ldots, \lambda_m$, not all zero, such that the sum $\lambda_1\phi_1(x) + \cdots + \lambda_m\phi_m(x)$ is a non-positive constant.*

Let us assume that the case (b) takes place in our situation. We may assume that λ_m is the greatest of $\lambda_1, \ldots, \lambda_m$, and therefore positive. From the statement ii) of 4), not all λ_i are equal to λ_m. Let us divide all terms by λ_m:

$$\frac{\lambda_1}{\lambda_m}\phi_1 + \cdots + \frac{\lambda_{m-1}}{\lambda_m}\phi_1 + \phi_m = \text{const} \leqslant 0$$

and subtract this from the statement ii) of 4):

$$\left(1 - \frac{\lambda_1}{\lambda_m}\right)\phi_1 + \cdots + \left(1 - \frac{\lambda_{m-1}}{\lambda_m}\right)\phi_1 = \text{const} \geqslant 0.$$

Here all coefficients are non-negative and not all are zero. Therefore the functions $(1 - \lambda_i/\lambda_m)\phi_i$ for $i = 1, \ldots, m-1$ also satisfy the conditions i) and ii) of 4) with a smaller value of m, which contradicts our assumption. Thus case (b) is excluded, so case (a) takes place, whence there is a point $p \in \mathbb{R}^d$ where all $\phi_j(p) > 0$. Since all ϕ_j are continuous, there is a rational point with this property also, whence condition iii) of 4) follows, $\square$

6. Proof of Theorem 4.1

6.1. Proof of (20) in case a) of the theorem. — Rewording Theorem 4.2 for the case when 0 and 1 are permuted, we see that whenever f is standard and "all ones" attracts D_f, there exist a natural number $m \leqslant d+1$ and m affine functions $\phi_1, \ldots, \phi_m \ : \ \mathbb{R}^d \to R$ such that:

$$(21) \quad \begin{cases} \text{i) for every } j \in [1, m] \text{ the set } \{i \in \mathbb{Z}^d \ : \ \phi_j(i) \leqslant 0\} \text{ is a one-set.} \\ \text{ii) } \phi_1 + \cdots + \phi_m \equiv \text{const} > 0. \\ \text{iii) There is a rational point } p \in \mathbb{R}^d \text{ such that } \phi_j(p) > 0 \text{ for all } j \in [1, m]. \end{cases}$$

For instance, for the NEC example there are $m = 3$ such affine functions, whose level lines are horizontal, vertical and lines of slope -1 respectively.

For every j let us denote $\overline{\phi}_j = \phi_j - \phi_j(0)$, whence $\phi_j = \overline{\phi}_j + \phi_j(0)$, where $\overline{\phi}_j$ is the linear part. Notice that $|\phi_j(0)| \leqslant \|\phi_j\| \cdot \|\Delta\|$ and that $\phi_1(0) + \cdots + \phi_m(0) > 0$. Notice also that if f is standard, "all ones" attracts D_f and $V_p + V_{-p} \geqslant 0$ for all directions p, then for any $j \in [1, m]$ and any thick-enough layer

$$y = \mathrm{Conf}\{i \in \mathbb{Z}^d \,\big|\, C_1 \leqslant \phi_j(i) \leqslant C_2\},$$

$$\mathrm{Ind}(D_f\, y) \supseteq \Big\{i \in \mathbb{Z}^d \,\Big|\, C_1 + \phi_j(0) \leqslant \phi_j(i) \leqslant C_2 + \phi_j(0)\Big\}. \tag{22}$$

Lemma 6.1. — *Take any standard f and assume that "all ones" attracts D_f and that $V_p + V_{-p} \geqslant 0$ for all directions p. Take x^* defined by*

$$\mathrm{Ind}(x^*) = \bigcup_{1 \leqslant j \leqslant m} \Big\{i \in \mathbb{Z}^d \,\Big|\, |\phi_j(i)| \leqslant 2\|\Delta\| \cdot \|\phi_j\|\Big\}. \tag{23}$$

Then for $t = 0, 1, 2, 3, \ldots$ the indicator of $D^t\, x^*$ includes the union $A_t \cup B_t$, where

$$A_t = \bigcup_{1 \leqslant j \leqslant m} \Big\{i \in \mathbb{Z}^d \,\Big|\, |\phi_j(i) - t \cdot \phi_j(0)| \leqslant 2\|\Delta\| \cdot \|\phi_j\|\Big\} \tag{24}$$

and

$$B_t = \bigcap_{1 \leqslant j \leqslant m} \Big\{i \in \mathbb{Z}^d \,\Big|\, \overline{\phi}_j(i) - t \cdot \phi_j(0) \leqslant 0\Big\}. \tag{25}$$

[For the NEC example of Section 2, this lemma corresponds to observation (iv).]

Let us prove this lemma by induction. Base of induction: Since A_0 coincides with $\mathrm{Ind}(x^*)$ and $B_0 \subset A_0$, our statement is true for $t = 0$.

Induction step. — Let us suppose that $\mathrm{Ind}(D^t\, x^*) \supseteq A_t \cup B_t$, take any $i \in A_{t+1} \cup B_{t+1}$ and prove that $i \in \mathrm{Ind}(D^{t+1}\, x^*)$. Let us consider two cases.

Case 1. — Let i belong to A_{t+1}. Then our statement follows from (22).

Case 2. — Let i belong to B_{t+1}, but not to A_{t+1}. Then

$$\overline{\phi}_j(i) - (t+1) \cdot \phi_j(0) \leqslant -2\|\Delta\| \cdot \|\phi_j\|$$

for all $j \in [1, m]$. Notice that

$$\overline{\phi}_j(i + v_k) \leqslant \overline{\phi}_j(i) + \|\phi_j\| \cdot |v_k| \leqslant \overline{\phi}_j(i) + \|\phi_j\| \cdot \|\Delta\|.$$

Therefore

$$\begin{aligned} \overline{\phi}_j(i + v_k) - t \cdot \phi_j(0) &\leqslant \overline{\phi}_j(i) + \|\phi_j\| \cdot \|\Delta\| - (t+1)\phi_j(0) + \phi_j(0) \\ &\leqslant -2\|\Delta\| \cdot \|\phi_j\| + \|\phi_j\| \cdot \|\Delta\| + \phi_j(0) \leqslant 0. \end{aligned}$$

Thus

$$i + \Delta \;\subset\; B_t \;\subset\; \mathrm{Ind}(D^t\, x^*).$$

Hence from (18) $i \in \mathrm{Ind}(D^{t+1}\, x^*)$. Lemma 6.1 is proved.

Lemma 6.2. — *Under the conditions of Lemma 6.1, there is a positive constant $\alpha > 0$ such that for all $t = 0, 1, 2, \ldots$ the set B_t defined by (25) contains a sphere in $\mathbb{Z}^d$ with the radius $\alpha \cdot t$.*

Proof. — In fact we shall prove that

$$\forall\, i \in \mathbb{Z}^d,\ t = 0, 1, 2, \ldots\ :\ |i + t \cdot p| \leqslant \alpha \cdot t \implies i \in B_t,$$

where p is that rational point where all $\phi_j(p) > 0$, whose existence is provided by iii). Let us denote $\kappa_j = \phi_j(p) > 0$ and $\alpha = \min_j(\kappa_j \,/\, \|\phi_j\|)$, that is the minimal distance from p to the hyperplanes $\phi_j = 0$. Let us consider three cases.

Case 1. — Let $p = 0$. Then $\phi_j(0) = \kappa_j > 0$ for all j. Now let us take any point i in the sphere with the radius $\alpha \cdot t$ and center 0. This means that

$$|i| \leqslant \alpha \cdot t = \min_j(\phi_j(0) \,/\, \|\phi_j\|) \cdot t.$$

Then

$$\overline{\phi}_j(i) \leqslant |i| \cdot \|\phi_j\| \leqslant \phi_j(0) \,/\, \|\phi_j\| \cdot t \cdot \|\phi_j\| = t \cdot \phi_j(0)$$

for all j, whence $i \in B_t$.

Case 2. — Let $p \in \mathbb{Z}^d$. Then along with our operator D_f we consider another operator D_g, where $g(x) \equiv f(\tau_p\, x)$. The function g is also standard, D_g is also attracted by "all ones" and the affine functions provided for D_g by iii) of (21) can be obtained from those for D_f by the same translation, so their values at 0 are $\kappa_1, \ldots, \kappa_m > 0$, whence D_g fits our case 1. So the set B_t for D_g contains a sphere with the center 0 and radius $\alpha \cdot t$. Since D_f commutates with all translations, the set B^t for D_g results from the set B_t for D_f by a translation at $t \cdot p$. Thus the set B_t for D_f results from B_t for D_g by the opposite translation, whence it contains a sphere with the center $-t \cdot p$ and the same radius.

Case 3. — Let p be any rational point. Let us denote q the least common denominator of all the coordinates of p and immerse our $\mathbb{Z}^d$ into the set $\mathbb{Z}_q^d$, where $\mathbb{Z}_q = \{n/q \,\big|\, n \in \mathbb{Z}\}$. Let us denote $\Omega_q = \{0,1\}^{\mathbb{Z}_q^d}$. Now f can be considered as a function g from Ω_q to $\{0,1\}$. Now let us "stretch" $\mathbb{Z}_q^d$ to turn it into $\mathbb{Z}^d$. Under this transformation the function g remains standard and "all ones" still attracts D_g. In addition to that, the affine functions for D_g with the properties (21) now can be obtained from those for D_f by a homothety with coefficient q. Therefore their values at the integer point $q \cdot p$ are $\kappa_1, \ldots, \kappa_m > 0$. So D_g fits our case 2, whence the set B_t for D_g contains a sphere with the center $-t \cdot q \cdot p$ and radius $\alpha \cdot q \cdot t$, whence the set B_t for D_f contains a sphere with the center $-t \cdot p$ and radius $\alpha \cdot t$. Lemma 6.2 is proved. □

Now let us prove (20). From monotonicity it is sufficient to prove this inequality for $\mu = (N_\varepsilon\, D_f)^t\ \delta_0$ for some t. Let us choose t_1 such that $\alpha \cdot t_1 \geqslant R + d$. Then, taking x^* defined by (23) as the initial configuration, after t_1 time-steps we obtain a configuration, whose indicator contains a sphere with the radius $R + d$ and therefore contains a sphere with the radius R and center at some integer point p. However, what we actually need is a finite deviation from "all zeros", which coincides with x^* only within a sphere with the radius $R + t_1 \cdot \|\Delta\|$ and has zeros outside it. The cardinality of its indicator does not exceed $C(R^{d-1}+1)$ with an appropriate C. Translating this configuration at the vector $-p$, we obtain another configuration, which fills with ones a sphere with radius R and center at the origin after t_1 time-steps. The probability that the actual configuration's indicator contains this configuration is not less that $\varepsilon^{C(R^{d-1}+1)}$, whence (20) follows. □

6.2. Proof of (20) in case b) of the theorem. — This time we define x^* as follows:

$$\mathrm{Ind}(x^*) = \Big\{ i \in \mathbb{Z}^d \;\Big|\; |\langle i, p\rangle| \leqslant \|\Delta\| \Big\}.$$

Then for all $t = 0, 1, 2, \dots$

$$\mathrm{Ind}(D_f^t\, x^*) \;\supseteq\; \Big\{ i \in \mathbb{Z}^d \;\Big|\; -\|\Delta\| + t \cdot V_{-p} \leqslant |\langle i, p\rangle| \leqslant \|\Delta\| + t \cdot V_p \Big\}.$$

Here the right side is a layer with the thickness $2\|\Delta\| + t(V_p + V_{-p})$. Given any $R \geqslant 0$, let us choose the minimal integer t_1 for which $2\|\Delta\| + t_1(V_p + V_{-p}) \geqslant R + d$. Then indicator of $D_f^{t_1} x^*$ contains a sphere with an integer center and radius R. If we take an initial condition which coincides with x^* within a sphere with the center at the origin and radius $R + d + t_1 \cdot \|\Delta\|$, we shall obtain the same result. This configuration has $C(R^{d-1}+1)$ components that equal 1, where C is an appropriate constant. Further we argue like in case a). □

7. Final notes

Note 1. — Using minoration arguments, is is easy to expand our theorem to some random cellular automata, which cannot be represenred as $N_\varepsilon\, D_f$. Using the same Δ as before and choosing transition probabilities $\theta(x \,|\, y_\Delta)$ for all $x \in \{0,1\}$ and $y \in \{0,1\}^\Delta$, we can define a random cellular automaton as an operator $P \,:\, \mathcal{M} \to \mathcal{M}$ which transforms any δ_y, where $y \in \Omega$, into a product-measure in which the probability that the i-th component equals x is $\theta(x \,|\, y_{i+\Delta})$. This operator majorates $N_\varepsilon\, D_f$ if

$$\theta(x \,|\, y_\Delta) \begin{cases} = 1 & \text{if } f(y_\Delta) = 1, \\ \geqslant \varepsilon & \text{if } f(y_\Delta) = 0. \end{cases}$$

As soon as this condition holds and D_f satisfies conditions of our theorem, all invariant measures of P also satisfy (20) and therefore are non-Gibbs.

Note 2. — In some cases it is possible to obtain a stronger estimation than (20). Let $d > a > 0$ and $f(x)$ equal

$$\min_{i_1,\dots,i_a \in \{0,1\}} \quad \max_{i_{a+1},\dots,i_d \in \{0,1\}} \quad x(i_1, \dots, i_d)$$

where $i_1, \dots, i_d$ are the coordinates of $\mathbb{Z}^d$. In this case

$$-\ln \mu(\mathbf{1}(S_{0,L})) \prec L^a,$$

where μ is any invariant measure of $N_\varepsilon D_f$. If $a < d-1$, this estimation is stronger than (20). This estimation can be proved in the same manner as in the case b), only x^* now is defined by the condition:

$$x_i^* = 1 \text{ if } \quad \max(|i_{a+1}|, \ldots, |i_d|) \leqslant \text{const.}$$

Note 3. — Given a standard f, let us assume that "all zeros" attracts D_f. Then we hope to estimate $-\ln \mu(\mathbf{1}(S_{0,L}))$ from below as follows:

$$-\ln \mu(\mathbf{1}(S_{0,L})) \succ L.$$

If we succeed, this will settle the question of asymptotics of $-\ln \mu(\mathbf{1}(S_{0,L}))$ in some cases, e.g. in our examples 1 and 2.

Note 4. — Those conditions under which our theorem holds and is non-trivial can be satisfied only for $d > 1$. However, a statement similar to our theorem for the one-dimensional case was proved in [**27**]. Namely, it was proved that all non-trivial invariant measures of a class of one-dimensional random cellular automata did not belong to a class, which included all Markov measures.

Acknowledgments. — We thank Aernout van Enter for very useful comments and criticism. We also thank the organizers and funding agencies of the IV Brazilian School of Probability where the final discussions for this paper took place.

References

[1] C. Bennett and G. Grinstein. Role of Irreversibility in Stabilizing Complex and Nonergodic Behavior in Locally Interacting Discrete Systems. *Phys. Rev. Letters*, v. 55 (1985), n. 7, pp. 657-660.

[2] M. Bramson and L. Gray. A useful renormalization argument. In *Random walks, Brownian motion, and interacting particle systems*, 113–152, Birkhäuser Boston, Boston, MA, 1991.

[3] J. Bricmont, A. Kupiainen, and R. Lefevere. Renormalization group pathologies and the definition of Gibbs states. *Comm. Math. Phys.*, 194(2):359–388, 1998.

[4] A.C.D. van Enter. The renormalization-group peculiarities of Griffiths and Pearce: What have we learned? In *Mathematical Results in Statistical Mechanics*, S. Miracle-Sole, J. Ruiz and V. Zagrebnov eds., World Scientific, Singapore, 1999, pp. 509–26.

[5] T.C. Dorlas and A.C.D. van Enter. Non-Gibbsian limit for large-block majority-spin transformations. *J. Stat. Phys.*, 55:171–181, 1989.

[6] A.C.D. van Enter, R. Fernández, and A.D. Sokal. Regularity properties and pathologies of position-space renormalization-group transformations: Scope and limitations of Gibbsian theory. *J. Stat. Phys.*, 72:879–1167, 1993.

[7] A.C.D. van Enter and J. Lörinczi. Robustness of the non-Gibbsian property: some examples. *J. Phys. A*, 29:2465–73, 1996.

[8] A.C.D. van Enter, C. Maes, and S. Shlosman. Dobrushin's program on Gibbsianity restoration: weakly Gibbs and almost Gibbs random fields. In *On Dobrushin's way. From probability theory to statistical physics*, Amer. Math. Soc., Providence, RI, 2000, pp. 59–70.

[9] A.C.D. van Enter and S.B. Shlosman. (Almost) Gibbsian description of the sign fields of SOS fields. *J. Statist. Phys.*, 92:353–368, 1998.

[10] R. Fernández. Random fields in lattices. The Gibbsianness issue. *Resenhas*, 3(4):391–421, 1998.

[11] R. Fernández. Measures for lattice systems. *Phys. A*, 263(1-4):117–130, 1999. STATPHYS 20 (Paris, 1998).

[12] H.-O. Georgii. *Gibbs Measures and Phase Transitions*. Walter de Gruyter (de Gruyter Studies in Mathematics, Vol. 9), Berlin–New York, 1988.

[13] O.K. Kozlov. Gibbs description of a system of random variables. *Probl. Inform. Transmission*, 10:258–65, 1974.

[14] J. L. Lebowitz and C. Maes, The effect of an external field on an interface, entropic repulsion, *J. Stat. Phys.*, v. 46, pp. 39–49, 1987.

[15] J.L. Lebowitz, C. Maes, and E.R. Speer. Statistical mechanics of probabilistic cellular automata. *J. Stat. Phys.*, 59:117–70, 1990.

[16] J.L. Lebowitz and R.H. Schonmann. Pseudo-free energies and large deviations for non-Gibbsian FKG measures. *Prob. Th. Rel. Fields*, 77:49–64, 1988.

[17] J. Lörinczi, Non-Gibbsianness of the reduced SOS-measure, *Stoch. Proc. Appl.*, v. 74, pp. 83–88, 1998.

[18] J.Lőrinczi and C. Maes. Weakly Gibbsian measures for lattice spin systems. *J. Statist. Phys.*, 89(3-4):561–579, 1997.

[19] C. Maes and K. Vande Velde. The interaction potential of a stationary measure of a high-noise spinflip process. *J. Math. Phys.*, 34:3030–1, 1993.

[20] C. Maes and K. Vande Velde. The (non-)Gibbsian nature of states invariant under stochastic transformations. *Physica A*, 206:587–603, 1994.

[21] D. Makowiec. Gibbsian versus non-Gibbsian nature of stationary states for Toom probabilistic cellular automata via simulations. *Phys. Rev. E*, 55:6582–8, 1997.

[22] D. Makowiec. Stationary states of Toom cellular automata in simulations. *Phys. Rev. E*, 60:3787–95, 1999.

[23] F. Martinelli and E. Scoppola. A simple stochastic cluster dynamics: rigorous results. *J. Phys. A*, 24:3135–57, 1991.

[24] M. B. Petrovskaya, I. I. Piatetski-Shapiro and N. B. Vasilyev. Modelling of voting with random errors. *Automatics and telemechanics*, v.10, pp. 103-107, 1969 (in Russian).

[25] C.-E. Pfister. Thermodynamical aspects of classical lattice systems. Preprint. Notes for a minicourse at the IV-th Brazilian School of Probability, Mambucaba, Brazil, August 2000.

[26] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970.

[27] A.L. Toom. On invariant measures in non-ergodic random media. In *Probabilistic Methods of Investigation*, A. Kolmogorov, editor, Moscow University Press, Moscow, 1972, pp. 43–51. (In Russian).

[28] A.L. Toom. Nonergodic multidimensional systems of automata. *Probl. Inform. Transmission*, 10(3):70–79, 1974.

[29] A.L. Toom. Stable and attractive trajectories in multicomponent systems. In *Multicomponent Random Systems. Advances in Probability and Related Topics*, R. Dobrushin and Ya. Sinai eds., Dekker, New York, 1980, v. 6, pp. 549-576.

[30] K. Vande Velde. On the question of quasilocality in large systems of locally interacting components. K. U. Leuven thesis, 1995.

R. Fernández, Mathématiques, site Colbert, Faculté des Sciences Université de Rouen F-76821 Mont Saint-Aignan, France • *E-mail :* `Roberto.Fernandez@univ-rouen.fr`

A. Toom, Departamento de Estatistica, Universidade Federal de Pernambuco, Recife/PE 50740-540, Brazil • *E-mail :* `toom@de.ufpe.br` or `toom@member.amd.org`

Astérisque
287, 2003, p. 89–102

INJECTIVITY OF C^1 MAPS $\mathbb{R}^2 \to \mathbb{R}^2$ AT INFINITY AND PLANAR VECTOR FIELDS

by

Carlos Gutierrez & Alberto Sarmiento

Abstract. — Let $X : \mathbb{R}^2 \smallsetminus \overline{D}_\sigma \to \mathbb{R}^2$ be a C^1 map, where $\sigma > 0$ and $\overline{D}_\sigma = \{p \in \mathbb{R}^2 : ||p|| \leqslant \sigma\}$.
(i) If for some $\varepsilon > 0$ and for all $p \in \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$, no eigenvalue of $DX(p)$ belongs to $(-\varepsilon, \infty)$, there exists $s \geqslant \sigma$, such that $X|_{\mathbb{R}^2 \smallsetminus \overline{D}_s}$ is injective;
(ii) If for some $\varepsilon > 0$ and for all $p \in \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$, no eigenvalue of $DX(p)$ belongs to $(-\varepsilon, 0] \cup \{z \in \mathbb{C} : \Re(z) \geqslant 0\}$, there exists $p_0 \in \mathbb{R}^2$ such that the point ∞, of the Riemann sphere $\mathbb{R}^2 \cup \{\infty\}$, is either an attractor or a repellor of $x' = X(x) + p_0$.

1. Introduction

The study of planar vector fields around singularities has somehow motivated the present work. A sample of this study is the work done by C. Chicone, F. Dumortier, J. Sotomayor, R. Roussarie, F. Takens. See for instance [**Chi, DRS, Rou, Tak**]. Here we study the behavior of a vector field $X : \mathbb{R}^2 \to \mathbb{R}^2$ around infinity. While a C^1 vector field around a singularity is quite regular, we work under conditions that do not imply, a priori, any regularity of the vector field around infinity. Given an open subset U of $\mathbb{R}^2$ and a C^1 map $Y : U \to \mathbb{R}^2$, we shall denote by $\mathrm{Spec}(Y) = \{\text{eigenvalues of } DY(p) : p \in U\}$. Our main result is the following

Theorem 1. — *Let $X = (f, g) : \mathbb{R}^2 \smallsetminus \overline{D}_\sigma \to \mathbb{R}^2$ be a C^1 map, where $\sigma > 0$ and $\overline{D}_\sigma = \{p \in \mathbb{R}^2 : ||p|| \leqslant \sigma\}$. The following is satisfied:*

(i) *if for some $\varepsilon > 0$, $\mathrm{Spec}(X)$ is disjoint of $(-\varepsilon, \infty)$, then there exists $s \geqslant \sigma$, such that $X|_{\mathbb{R}^2 \smallsetminus \overline{D}_s}$ is injective;*

2000 *Mathematics Subject Classification.* — 37E35, 37C10.
Key words and phrases. — Injectivity, Reeb component, vector fields.

C.G.: Supported in part by PRONEX/FINEP/MCT - grant number 76.97.1080.00.

(ii) *if for some* $\varepsilon > 0$, $\operatorname{Spec}(X)$ *is disjoint of* $(-\varepsilon, 0] \cup \{z \in \mathbb{C} : \Re(z) \geqslant 0\}$*, then, there exists* $p_0 \in \mathbb{R}^2$ *such that the point* ∞*, of the Riemann sphere* $\mathbb{R}^2 \cup \{\infty\}$*,is either an attractor or a repellor of* $x' = X(x) + p_0$*.*

To give an idea of the proof of this result, let us introduce the following definition.

Let $X = (f, g) : \mathbb{R}^2 \smallsetminus \overline{D}_\sigma \to \mathbb{R}^2$ be a C^1 map as in Theorem 1. Since $f : \mathbb{R}^2 \smallsetminus \overline{D}_\sigma \to \mathbb{R}$ is a C^1 submersion, $q \in \mathbb{R}^2 \to \nabla f^{\#}(q) = (-f_y(q), f_x(q))$, the Hamiltonian vector field of f, has no singularities. Let $g_0(x, y) = xy$ and consider the set

$$B = \{(x, y) \in [0, 2] \times [0, 2] : 0 < x + y \leqslant 2\}.$$

We will say that $\mathcal{A} \subset \mathbb{R}^2$ is a *HRC* (*Half–Reeb Component*) of $\nabla f^{\#}$ (see figure 1) if there is a homeomorphism $h : B \to \mathcal{A}$ which is a topological equivalence between $\nabla f^{\#}|_{\mathcal{A}}$ and $\nabla {g_0}^{\#}|_B$, and such that

(1) $h(\{(x, y) \in B : x + y = 2\})$ (called *the compact edge of* $\mathcal{A}$) is a smooth segment transversal to $\nabla f^{\#}$ in the complement of $h(1, 1)$, and

(2) both $h(\{(x, y) \in B : x = 0\})$ and $h(\{(x, y) \in B : y = 0\})$ are full half–trajectories of $\nabla f^{\#}$.

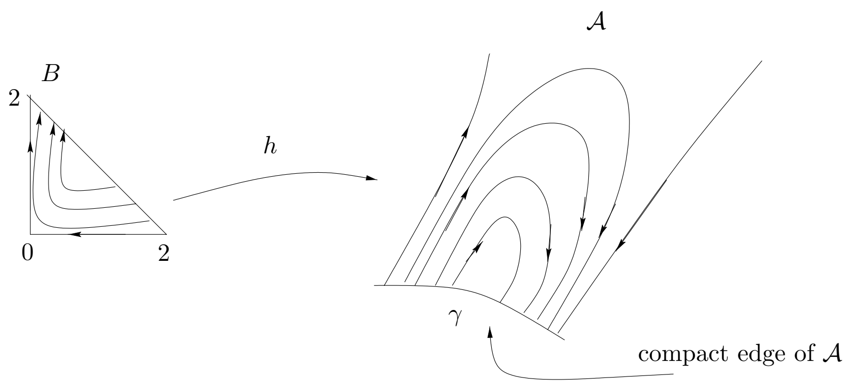

FIGURE 1. A half-Reeb component.

Observe that $\mathcal{A}$ may not be a closed subset of $\mathbb{R}^2$.

Proceed to give an idea of the proof of Theorem 1. First, we shall prove that:

Proposition 1. — *if* $X = (f, g) : \mathbb{R}^2 \smallsetminus \overline{D}_\sigma \to \mathbb{R}^2$ *is a* C^1 *map as in Theorem 1, then any HRC of* $\nabla^{\#} f$ *is a bounded subset of* $\mathbb{R}^2$*.*

This is used to prove

Theorem 2. — *if* $Y = (\widetilde{f}, \widetilde{g}) : \mathbb{R}^2 \to \mathbb{R}^2$ *is a* C^1 *map such that, for some* $\varepsilon > 0$, $\operatorname{Spec}(Y) \cap (-\varepsilon, \varepsilon) = \varnothing$*, then* Y *is injective.*

Roughly speaking about Theorem 2, if the foliation induced by $\nabla \widetilde{f}^{\#}$ has no half-Reeb components then, $\nabla \widetilde{f}^{\#}$ is topologically equivalent to the foliation, on the (x, y)-plane, induced by the form dx (the foliation is made up by all the vertical straight lines). The injectivity of X will follow from the fact that $\nabla \widetilde{f}^{\#}$ and $\nabla \widetilde{g}^{\#}$ are linearly independent everywhere.

Sections 3 and 4 are devoted to prove

Corollary 2. — *if $X = (f, g) : \mathbb{R}^2 \smallsetminus \overline{D}_\sigma \to \mathbb{R}^2$ is a C^1 map as in Theorem 1, then there exists a smooth compact disc E such that $\nabla f^{\#}$, restricted to $\mathbb{R}^2 \smallsetminus E$, is topologically equivalent to the foliation, on $\mathbb{R}^2 \smallsetminus \overline{D}_1$, induced by dx.*

Observe that the foliation, on $\mathbb{R}^2 \smallsetminus \overline{D}_1$, induced by dx has exactly two tangencies with $\partial \overline{D}_1$ (at $(1, 0)$ and $(0, 1)$) which are "quadratic" and "external". Let us say a little more about what is proved in Section 3 and 4: We show, in Section 3, that given any generic smooth compact disc $F \supset \overline{D}_\sigma$ the number of "external" tangencies of ∇f with ∂F is equal to 2 plus the number of "internal" tangencies of ∇f with ∂F. We show, in Section 4, that the disc F can be deformed to a smooth compact disc E so that the referred "external" and "internal" tangencies cancel in pairs yielding exactly 2 tangencies which are "external".

Using Theorem 2 we obtain

Proposition 2. — *Let X be as in Corollary 2. If X takes ∂E diffeomorphically to a circle then $X|_{\mathbb{R}^2 \smallsetminus E}$ may be extended to a map which satisfies conditions of Theorem 2 and so it is injective.*

The proof of item (ii) of Theorem 1 is finished in Sections 5 and 6 by showing that, under conditions of Corollary 2, the disc E can be deformed so that, for the resulting new disc, still denoted by E, $\nabla f^{\#}|_{\mathbb{R}^2 \smallsetminus E}$, is topologically equivalent to the foliation, on $\mathbb{R}^2 \smallsetminus \overline{D}_1$, induced by dx and moreover X takes ∂E diffeomorphically to a circle. Then the result follows from Proposition 2.

The item (ii) of Theorem 1 follows from the corresponding item (i) and some previous Gutierrez and Teixeira work [**G-T**].

Throughout this article, given an embedded circle $C \subset \mathbb{R}^2$, the compact disc (resp. open disc) bounded by C will be denoted by $\overline{D}(C)$ (resp. $D(C)$). Also, we will freely use the fact that the assumptions of the theorem are open in the Whitney C^1–topology. In this way, when possible and necessary, we will assume that X is smooth and that it satisfies some generic property which will be made precise at the proper place.

Acknowledgements. — We wish to thank referee's comments which have been appreciated and incorporated into this work.

2. A global injectivity result

We shall need the following lemma which is contained in the proof of [**Gut**, Lemma 2.5]. For $\theta \in \mathbb{R}$: let R_θ denote the linear rotation

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Lemma 1. — *Let $X = (f,g) : \mathbb{R}^2 \setminus \overline{D}_\sigma \to \mathbb{R}^2$ be a C^1 map as in Theorem 1. Suppose that $\nabla^\# f$ has an HRC which is unbounded (as a subset of $\mathbb{R}^2$) but whose projection on the x-axis is a compact interval. Then, there exists $\varepsilon > 0$ such that, for all $\theta \in (-\varepsilon, 0) \cup (0, \varepsilon)$ $\nabla^\# f_\theta$ has a HRC whose projection on the x-axis is an interval of infinite length; here $(f_\theta, g_\theta) = R_\theta \circ X \circ R_{-\theta}$.*

The proof of Proposition 1 and Theorem 2 can be found in [**CGL**] but, as we have already said and for sake of completeness, they are included here.

Proposition 1. — *Let $X = (f,g) : \mathbb{R}^2 \setminus \overline{D}_\sigma \to \mathbb{R}^2$ be a C^1 map as in Theorem 1. Then any HRC of $\nabla^\# f$ is a bounded subset of $\mathbb{R}^2$.*

Proof. — Let $\mathcal{A}$ be a half Reeb component for f. Let $\Pi : \mathbb{R}^2 \to \mathbb{R}$ be the projection on the first coordinate. By composing with a rotation if necessary, in the way that is stated in Lemma 1, we may suppose that $\Pi(\mathcal{A})$ is an interval of infinite length, say $[b, \infty)$. We may also assume that X is smooth and —by Thom's Transversality Theorem for jets [**G-G**]— that

(a1) the set

$$T = \{(x,y) \in \mathbb{R}^2 : f_y(x,y) = 0\}$$

is made up of regular curves;

(a2) There is a discrete subset Δ of T such that if $p \in T \setminus \Delta$ (resp. $p \in \Delta$), $\nabla^\# f$ has quadratic contact (resp. cubic contact) with the vertical foliation of $\mathbb{R}^2$.

Then, if $a > b$ is large enough,

(b) for any $x \geqslant a$, the vertical line $\Pi^{-1}(x)$ intersects exactly one trajectory $\alpha_x \subset \mathcal{A}$ of $\nabla f^\#|_{\mathcal{A}}$ such that $\Pi(\alpha_x) \cap (x, \infty) = \varnothing$; in other words, x is the maximum for the restriction $\Pi|_{\alpha_x}$.

It follows that

(c) if $x \geqslant a$ and $p \in \alpha_x \cap \Pi^{-1}(x)$ then $p \in T \cap \mathcal{A} \setminus \Delta$.

Let T_m be the set of $p \in \mathcal{A}$ such that, for some $x \geqslant a$, $p \in \alpha_x \cap \Pi^{-1}(x)$. Notice that, for every $x \geqslant a$, $\alpha_x \cap \Pi^{-1}(x)$ is a finite set; nevertheless, by (b), (c) and by using Thom's Transversality Theorem for jets, we may get the following stronger statement:

(d) There is a sequence $F = \{a_1, a_2, \ldots, a_i, \cdots\}$ in $[a, \infty)$, which may be either empty or finite or else countable, such that if $x \in F$ (resp. $x \in [a,\infty) \setminus F$), then $\Pi^{-1}(x) \cap T_m$ is a two-point-set (resp. a one-point-set).

If $x \in [a,\infty) \smallsetminus F$, define $\eta(x) = (x,\eta_2(x)) = \Pi^{-1}(x) \cap T_m$. Observe that $\eta : [a,\infty) \smallsetminus F \to T_m$ is a smooth embedding. As $f|_{\mathcal{A}}$ is bounded,

(e) $F \circ \eta$ extends continuously to a strictly increasing bounded map defined in $[a,\infty)$ such that, for all $x \in [a,\infty) \smallsetminus F$, $f_x(\eta(x))$ has constant sign.

Therefore, there exists a real constant K such that

$$\begin{aligned} K = \int_{a_1}^{\infty} \frac{d}{dx} f(\eta(x))\, dx &= \sum_{i=1}^{\infty} \int_{a_i}^{a_{i+1}} \frac{d}{dx} f(\eta(x))\, dx \\ &= \sum_{i=1}^{\infty} \int_{a_i}^{a_{i+1}} f_x(\eta(x)) \end{aligned}$$

This and (e) imply that, for some sequence $x_n \to \infty$, $\lim_{n\to\infty} f_x(\eta(x_n)) = 0$. This is the required contradiction. □

Theorem 2. — *Let $X : \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 map. Suppose that, for some $\varepsilon > 0$, $\operatorname{Spec}(X) \cap (-\varepsilon, \varepsilon) = \varnothing$. Then X is injective.*

Proof. — By Proposition 1, the Hamiltonian vector fields induced by the coordinate functions of $X = (f,g)$ have no Reeb component. Therefore X is injective. □

3. Index of a vector field along a circle

We shall say that a collar neighborhood U of an embedded circle $C \subset R^2 \smallsetminus \overline{D}_\sigma$ is *interior* (resp. *exterior*) if U is contained in $\overline{D}(C)$ (resp. $\mathbb{R}^2 \smallsetminus D(C)$).

Proposition 2. — *Let $X = (f,g) : \mathbb{R}^2 \smallsetminus \overline{D}_\sigma \to \mathbb{R}^2$ be a C^1 map as in Theorem 1. Let $C \subset R^2 \smallsetminus \overline{D}_\sigma$ be a smooth circle surrounding the origin. Suppose that $X(C)$ is an embedded circle and that there exists an exterior collar neighborhood $U \subset \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$ of C such that $X(U)$ is also an exterior collar neighborhood of $X(C)$. Then X is an embedding.*

Proof. — By the assumptions, X can be extended to a C^1 map $Y : R^2 \to R^2$ which takes $\overline{D}(C)$ diffeomorphically onto the $\overline{D}(X(C))$. See [**Hir**]. Under these conditions we may apply either Theorem 2 or Gutierrez and Fessler Injectivity Theorem [**Gut, Fes**] to conclude that Y is an embedding and, a fortiori, that X is an embedding too. □

The theorem below on indexes of singularities of vector fields will be used to prove theorem 1. The proof can be found in [**Har**, Theorem 9.2]

Let C be a simple closed curve of R^2. A C^1 vector field $Y : R^2 \to R^2$ is said to be *internally (or externally) tangent to* C at $x_0 \in C$, if there exists an $\varepsilon > 0$ such that the solution arc $\phi(t)$ of the equation

$$x' = Y(x), \qquad x(0) = x_0,$$

is interior (or exterior) to C for $0 < |t| \leqslant \varepsilon$. We shall denote by $j_Y(C)$ the index of Y along C.

Theorem 3. — *Let Y be a C^1 vector field on a connected open set $E \subset R^2$. Let C be a positive oriented Jordan curve of class C^1 in E with the property that $Y(x) \neq 0$ on C and that Y is tangent to C at only a finite number of points $x_1, \dots, x_n$ of C. Let n^i (resp. n^e be the number of these points x_j where the solution arc $\phi(t)$ of $x' = Y(x)$, $\phi(0) = x_j$ for small $|t|$ is internally (resp. externally) tangent to C at x_j (so that $n^i + n^e \leqslant n$). Then $2j_Y(C) = 2 + n^i - n^e$.*

Corollary 1. — *Let assume the notation and conditions of Theorem 1. In particular let $X = (f, g) : \mathbb{R}^2 \smallsetminus \overline{D}_\sigma \to \mathbb{R}^2$ be a C^1 map. If $C \subset \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$ is a smooth circle surrounding the origin, then $j_{\nabla f^\#}(C) = j_{\nabla f}(C) = 0$.*

Proof. — If $j_{\nabla f}(C) \neq 0$ there would exist a point $p \in C$ and a real $a > 0$ such that $\nabla f(p) = (a, 0)$. In particular a will be an eigenvalue of $DX(p)$. This contradiction with the assumptions of Theorem 1 proves the corollary. □

4. Avoiding internal tangencies

We say that $\nabla f^\#$ is *in general position with* an embedded circle $C \subset \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$ if there exists a subset F of C, at most finite such that (i) $\nabla f^\#$ is transversal to $C \smallsetminus F$, (ii) $\nabla f^\#$ has a quadratic tangency with C at each point of F, and (iii) a trajectory of $\nabla f^\#$ can meet tangentially C at most at one point.

Lemma 2. — *Suppose that $\nabla f^\#$ is in general position with a smooth circle $C \subset \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$ which surrounds the origin. Suppose also that a trajectory γ of $\nabla f^\#$ meets C transversally somewhere and with an external tangency at a point p. Then the trajectory γ contains a closed subinterval $[p, r]_f$ which meets C exactly at $\{p, r\}$ (doing it transversally at r) and the following is satisfied:*

(i) *If $[p, r]$ denotes the closed subinterval of C such that $\Gamma = [p, r] \cup [p, r]_f$ bounds a compact disc $\overline{D}(\Gamma)$ contained in $\mathbb{R}^2 \smallsetminus D(C)$, then points of $\gamma \smallsetminus [p, r]_f$ nearby p do not belong to $\overline{D}(\Gamma)$;*

(ii) *Let $(\widetilde{p}, \widetilde{r})$ and $[\widetilde{p}, \widetilde{r}]$ be subintervals of C satisfying $[p, r] \subset (\widetilde{p}, \widetilde{r}) \subset [\widetilde{p}, \widetilde{r}]$. If $\widetilde{p}$ and $\widetilde{r}$ are close enough to p and r, respectively, then we may deform C into a smooth circle C_1 in such a way that the deformation fixes $C \smallsetminus (\widetilde{p}, \widetilde{r})$ and takes $[\widetilde{p}, \widetilde{r}] \subset C$ to a closed interval $[\widetilde{p}, \widetilde{r}]_1 \subset C_1$ which is close to $[p, r]_f$. Furthermore, $\nabla f^\#$ is in general position with C_1 and the number of tangencies of $\nabla f^\#$ with C_1 is smaller than that of $\nabla f^\#$ with C.*

Proof. — Certainly, γ contains a closed subinterval $[p, r]_f \subset \gamma$ which meets C exactly at $\{p, r\}$. As $\nabla f^\#$ is in general position with C, γ meet C transversally at r. Let

$[p, r]$ be the closed subinterval of C such that $\Gamma = [p, r] \cup [p, r]_f$ bounds a compact disc $\overline{D}(\Gamma)$ contained in $\mathbb{R}^2 \setminus D(C)$.

If $[p, r]_f$ does not satisfy (i) then the points of $\gamma \setminus [p, r]_f$ nearby p belong to $\overline{D}(\Gamma)$. Hence, as $\nabla f^{\#}$ has no singularities in $\mathbb{R}^2 \setminus D(C)$, points of $\gamma \setminus [p, r]_f$ nearby p belong to $\overline{D}(\Gamma)$. Therefore, there must be a closed subinterval $[q, p]_f \subset \gamma \cap \overline{D}(\Gamma)$ (see fig. 2.a) such that:

(a1) the union Γ_1 of the closed interval $[p, q] \subset [p, r]$ and $[p, q]_f$ bounds a compact disc $\overline{D}(\Gamma_1)$ contained in $(\mathbb{R}^2 \setminus D(C)) \cap \overline{D}(\Gamma)$;

(a2) $[q, p]_f$ meets C exactly at $\{p, q\}$, doing it transversally at q (with an external tangency at p); also, points of $\gamma \setminus [q, p]_f$ nearby p do not belong to $\overline{D}(\Gamma_1)$.

Summarizing either $[p, r]_f$ or $[q, p]_f$ satisfies (i). Therefore we may denote by $[p, r]_f$ the arc which satisfies (i).

We claim that (i) implies (ii). In fact, let us choose a small flow box B of $\nabla^{\#} f$ whose interior contains $[p, r]_f$. By the assumptions, we may suppose that $\widetilde{p}, \widetilde{r} \in B \cap C$. We may see that there exists a closed interval $[\widetilde{p}, \widetilde{r}]_T \subset B$ transversal to $\nabla f^{\#}$ (drawn as a dotted line in fig. 2.b) and such that $C_1 = (C \setminus [\widetilde{p}, \widetilde{r}]) \cup [\widetilde{p}, \widetilde{r}]_T$ is a smooth circle, surrounding the origin, contained in $\mathbb{R}^2 \setminus D(C)$. Moreover, $\nabla f^{\#}$ meets C_1 with a smaller number of tangencies than it does it with C. The remaining conclusions of (ii) can easily be checked. Therefore (i) implies (ii). $\square$

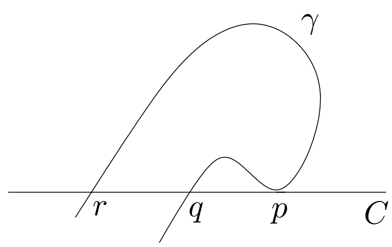

FIGURE 2.a

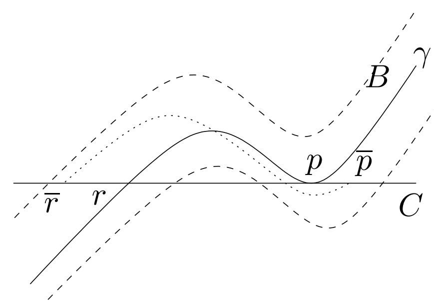

FIGURE 2.b

Remark 1. — Let suppose that $\nabla f^{\#}$ is in general position with a smooth circle $C \subset \mathbb{R}^2 \setminus \overline{D}_\sigma$ which surrounds the origin. Suppose also that $\nabla f^{\#}$ has an internal tangency with C at the point q; then, by observing the trajectories of $\nabla f^{\#}$ around q, we may see that there exist closed subintervals $[p, q]$ $[q, r]$ of C, with $[p, q] \cap [q, r] = \{q\}$, and a homeomorphism $T : [p, q] \to [q, r]$ such that,

(a1) $Tp = r, Tq = q$ and, for every $x \in (p, q]$, there is an arc of trajectory $[x, Tx]_f \subset \mathbb{R}^2 \setminus D(C)$ of $\nabla f^{\#}$, starting at x, ending at Tx and meeting C exactly and transversally at $\{x, Tx\}$, (a2) the family $[x, Tx]_f : x \in (p, q]$ depends continuously on x and tends to $\{q\}$ as $x \to q$.

Lemma 3. — *Let suppose that $\nabla f^{\#}$ is in general position with a smooth circle $C \subset \mathbb{R}^2 \setminus \overline{D}_\sigma$ which surrounds the origin. Suppose also that $\nabla f^{\#}$ has an internal tangency with C at the point q. Given any pair of subintervals $[p, q]$, $[q, r]$ of C (generated by q) as in Remark 1, the family $\{[x, Tx]_f : x \in (p, q]\}$ tends continuously to the compact arc of trajectory $[p, Tp]_f \subset \mathbb{R}^2 \setminus D(C)$ which either meets C exactly and transversally at $\{p, Tp\}$ or meets C with a quadratic external tangency; this second alternative happens if, and only if, $(p, q]$ is the maximal interval with properties (a1) and (a2) of Remark 1.*

Proof. — If $(p, q]$ is not the maximal interval with properties (a1) and (a2) of Remark 1, then $[p, Tp]_f \subset \mathbb{R}^2 \setminus D(C)$ meets C exactly and transversally at $\{p, Tp\}$.

Otherwise, as $\mathbb{R}^2 \setminus \overline{D}(C)$ is not bounded, the closure of $(p, q] \cup [p, r)$ cannot be the whole circle. Therefore, there are two possibilities. The first one is that the positive (resp. negative) half-trajectory γ_p^+ (resp. γ_r^-) of $\nabla f^{\#}$ starting at p (resp. at r) does not meet C and so it must accumulate at the point ∞ of the Riemann sphere $\mathbb{R}^2 \cup \infty$. Under these circumstances, the subinterval $[p, q] \cup [q, r]$ is the compact edge of non-bounded *HRC* of $\nabla f^{\#}$ made up of $\gamma_p^+ \cup \gamma_r^-$ together with the union of the arcs $[x, Tx]_f$, with $x \in (p, q]$. This contradiction with Proposition 1 shows that the second possibility must happen: this lemma is true. □

Lemma 4. — *There exists a smooth circle $C \subset \mathbb{R}^2 \setminus \overline{D}_\sigma$, surrounding the origin, in general position with $\nabla f^{\#}$, and such that*

(i) *if a trajectory γ of $\nabla f^{\#}$ meets C, with an external tangency, say p, then $\gamma \cap C = \{p\}$;*

(ii) *As a consequence of (i), every tangency of the Hamiltonian vector field $\nabla f^{\#}$ with C is quadratic and external. In particular, there exists a correspondence between tangencies and HRCs (which —by Proposition 1— are contained in the disc of $\mathbb{R}^2$ bounded by C).*

Proof. — By a small C^2 perturbation of f, we may assume that $\nabla f^{\#}$ is in general position with C; in particular, every tangency of $\nabla f^{\#}$ with C is quadratic. If (i) of this lemma is not satisfied, we may use Lemma 2 to obtain a new circle C_1 such that $\nabla f^{\#}$ is in general position with C_1 and the number of tangencies of $\nabla f^{\#}$ with C_1 is smaller than that of $\nabla f^{\#}$ with C. Using this procedure, as many times as necessary, we will be able to obtain a circle as required to prove (i)

As (i) is true, (ii) follows from Lemma 3. □

Corollary 2. — *There exists a smooth circle $C \subset \mathbb{R}^2 \setminus \overline{D}_\sigma$, surrounding the origin and there are two points $a, b \in C$, with $f(a) < f(b)$, such that $\nabla f^{\#}$ is tangent to C exactly at a and b; moreover, these tangencies are quadratic and external.*

Proof. — In fact, by Lemma 4, we may take a smooth circle $C \subset \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$, surrounding the origin, such that every tangency of the Hamiltonian vector field $\nabla f^\#$ with C is quadratic and external. Therefore, by the index formula of Theorem 3 and by Corollary 1, there are two points $a, b \in C$, with $f(a) < f(b)$, satisfying this corollary. □

5. Main Proposition

This section is devoted to the proof of the following

Proposition 3. — *There exists a smooth circle $C \subset \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$, surrounding the origin, such that $X(C)$ is also an embedded circle and, for some exterior collar neighborhood $U \subset \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$ of C, $X(U)$ is also an exterior collar neighborhood of $X(C)$.*

The proof of this proposition will be completed at the end of this section after some preparatory lemmas.

We say that a smooth circle $C \subset \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$ is *of ETT (i.e. external tangency type) for* $\nabla f^\#$ if the following is satisfied: C surrounds the origin, there are two points $a, b \in C$, with $f(a) < f(b)$, and there are points $a_1, a_2, \dots, a_n \in C_-$ and $b_1, b_2, \dots, b_n \in C_+$, where C_- and C_+ are the connected components of $C \smallsetminus \{a, b\}$, such that:

(a1) $\nabla f^\#$ is tangent to C exactly at a and b; also, these tangencies are quadratic and external;

(a2) $f(a) = \inf\{f(x) : x \in C\} < \sup\{f(x) : x \in C\} = f(b)$;

(a3) f takes diffeomorphically each C_i, with $i \in \{-, +\}$, onto the open interval $(f(a), f(b))$ (i.e., $X(C_i)$ is the graph of a map $(f(a), f(b)) \to \mathbb{R}$);

(a4) X restricted to $C \smallsetminus \{a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n\}$ is an embedding, and also, $X(C_-)$ and $X(C_+)$ meet transversally to each other

(a5) $(X(a_1), X(a_2), \dots, X(a_n)) = (X(b_1), X(b_2), \dots, X(b_n))$ and
$f(a) < f(a_1) = f(b_1) < f(a_2) = f(b_2) < \cdots < f(a_n) = f(b_n) < f(b)$.

(a6) there are sequences $x_n \to a$ and $y_n \to b$ of points x_n and y_n in $\mathbb{R}^2 \smallsetminus \overline{D}(C)$ such that, for all n, $f(x_n) < f(a) < f(b) < f(y_n)$. This means that the local exterior of C around a (resp. around b) is taken to the unbounded connected component of $\mathbb{R}^2 \smallsetminus X(C)$. In particular, $n \geqslant 0$ is an even number.

(a7) If $x \in \mathbb{R}^2 \smallsetminus D(C)$ is close enough to $y \in C_+$ (resp. $y \in C_-$) and $f(x) = f(y)$, then $g(y) < g(x)$ (resp. $g(y) > g(x)$).

(a8) If $\overline{a}_1, \overline{a}_n \in C_-$ and $\overline{b}_1, \overline{b}_n \in C_+$ are close enough to a_1, a_n and b_1, b_n, respectively, and $[a_1, a_n] \subset (\overline{a}_1, \overline{a}_n)$, $[b_1, b_n] \subset (\overline{b}_1, \overline{b}_n)$ then, $X([\overline{a}_1, a_1) \cup (a_n, \overline{a}_n])$ is below $X([\overline{b}_1, b_1) \cup (b_n, \overline{b}_n])$ (i.e. if $a' \in [\overline{a}_1, a_1) \cup (a_n, \overline{a}_n]$ and $b' \in [\overline{b}_1, b_1) \cup (b_n, \overline{b}_n]$ are such that $f(a') = f(b')$ then $g(a') < g(b')$).

Lemma 5. — *There is a smooth circle $C \subset \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$ of ETT for $\nabla f^\#$.*

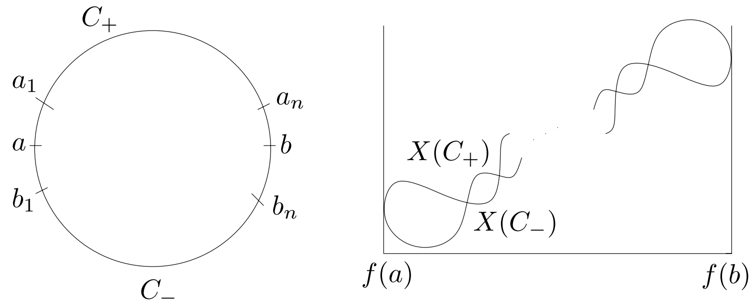

FIGURE 3

Proof. — In fact, by Corollary 2, we may take a smooth circle $C \subset \mathbb{R}^2 \setminus \overline{D}_\sigma$, surrounding the origin, such that there are two points $a, b \in C$, with $f(a) < f(b)$, and so that (a1) above is satisfied. This implies that (a2) and (a3) of definition above are also satisfied. Furthermore, by a small perturbation of X, we may assume that (a4) and (a5) of definition above are satisfied too. Item (a6) follows directly from the preceding properties. As $X(C)$ is tangent to the vertical foliation at the points $X(a)$ and $X(b)$, and by using (a6), The connected components C_- and C_+ can be named to satisfy (a7). Item (a7) implies (a8). $\square$

In the following of this section, C will be a smooth circle of ETT for $\nabla f^\#$ and we shall use all corresponding introduced notation.

Given $\alpha, \beta \in C_-$ (resp. $\alpha, \beta \in C_+$), $[\alpha, \beta], (\alpha, \beta), [\alpha, \beta)$ will denote subintervals of C_- (resp. of C_+) with endpoints α, β. Let L denote the straight line which passes through the points $X(a_1)$ e $X(a_n)$. Let $\mathcal{L}$ be the foliation of $\mathbb{R}^2$ made up by all the straight lines parallel to the line L. By a small perturbation of $X(C)$ with support in $X([a_1, a_n] \cup [b_1, b_n])$, we may assume that

(b) every point of tangency of $X([a_1, a_n])$ with $\mathcal{L}$ is quadratic, $X([a_1, a_n])$ and $X([b_1, b_n])$ are transversal to L.

Also, by taking $\overline{a}_1, \overline{a}_n \in C_-$ and $\overline{b}_1, \overline{b}_n \in C_+$ close to a_1, a_n and b_1, b_n, respectively, and $[a_1, a_n] \subset (\overline{a}_1, \overline{a}_n)$, $[b_1, b_n] \subset (\overline{b}_1, \overline{b}_n)$, we may suppose as well that

(c) $X([\overline{a}_1, a_1) \cup (a_n, \overline{a}_n])$ and $X([\overline{b}_1, b_1) \cup (b_n, \overline{b}_n])$ are disjoint of L.

Let $\theta \in \mathbb{R}$ be such that $R_\theta(\mathcal{L})$ is made up of vertical lines, where R_θ is the rigid rotation of angle θ. Recall that $(f_\theta, g_\theta) = X_\theta = R_\theta \circ X \circ R_\theta^{-1}$. By means of a small C^2−perturbation of f, we may assume that

(d) $\nabla f_\theta{}^\#$ is in general position with $R_\theta(C)$.

Then we have that

Lemma 6. — *Remark 1 and Lemma 3 are also valid when referred to the vector field*

$$Y = (R_{-\theta})_* \nabla f_\theta{}^{\#} = R_{-\theta} \circ \nabla f_\theta{}^{\#} \circ R_\theta.$$

Also X takes any trajectory of Y into a subinterval of a leaf of $\mathcal{L}$.

Proof. — If γ is a trajectory of Y then $R_\theta(\gamma)$ is a trajectory of $\nabla f_\theta{}^{\#}$. Therefore, $X_\theta \circ R_\theta(\gamma)$ is a subinterval of a vertical line and so $R_{-\theta} \circ X_\theta \circ R_\theta(\gamma)$ is a subinterval of a leaf of $\mathcal{L}$. However,

$$R_{-\theta} \circ X_\theta \circ R_\theta = X.$$

This implies that X takes any trajectory of Y into a subinterval of a leaf of $\mathcal{L}$. On the other hand, as $(f_\theta, g_\theta) = X_\theta = R_\theta \circ X \circ R_\theta^{-1}$ satisfies the assumptions of Theorem 1, Remark 1 and Lemma 3 are valid for the vector field $\nabla f_\theta{}^{\#}$. By definition of Y, we obtain the remaining conclusion of this lemma. $\square$

We claim that

Lemma 7. — *If $X([\overline{a}_1, a_1) \cup (a_n, \overline{a}_n])$ is below L and $X([\overline{b}_1, b_1) \cup (b_n, \overline{b}_n])$ is above L, then there is a smooth circle $C_1 \subset \mathbb{R}^2 \smallsetminus D(C)$, surrounding the origin, obtained from C by a deformation which fixes $C \smallsetminus ((\overline{a}_1, \overline{a}_n) \cup (\overline{b}_1, \overline{b}_n))$ and takes $[\overline{a}_1, \overline{a}_n] \subset C$ and $[\overline{b}_1, \overline{b}_n] \subset C)$ to the closed sub-intervals $[\overline{a}_1, \overline{a}_n]_{C_1} \subset C_1$ and $[\overline{b}_1, \overline{b}_n]_{C_1} \subset C_1$, respectively, which satisfy $X([\overline{a}_1, \overline{a}_n]_{C_1})$ is below L and $X([\overline{b}_1, \overline{b}_n]_{C_1})$ is above L. In particular, C_1 is as requested to prove Proposition 3.*

Proof. — Suppose that Y has an internal tangency with C at the point $q \in (a_1, a_n)$. By (d) and Lemma 6 we may proceed as in Remark 1 (applied to Y and considering the notation introduced there) to obtain sub-intervals $[p, q]$, $[q, r]$ of C (generated by q), determined by the condition that $(p, q]$ is the maximal subinterval of $[a_1, a_n]$ satisfying properties (a1) and (a2) of Remark 1. By Lemma 3, every element of the family $\{[x, Tx]_Y : x \in (p, q]\}$ is an arc of trajectory of Y. Notice that the maximality criterium for $(p.q]$ right above is different from that of Lemma 3.

To perform a sequence of adequate deformations, we meet three possible cases:

The first one is that $\{p, r\} \cap \{a_1, a_n\} \neq \varnothing$. Consider only the case in which $p = a_1$ and $r \neq a_n$. We may deform C into a new circle C_1 in such a way that: the deformation fixes $C \smallsetminus (\overline{a}_1, \overline{a}_n)$ and takes $[\overline{a}_1, \overline{a}_n] \subset C$ to a closed sub-interval $[\overline{a}_1, \overline{a}_n]_{C_1} \subset C_1$ such that

(e) the cardinality of $L \cap [\overline{a}_1, \overline{a}_n]_{C_1}$ is less than that of (the finite set) $L \cap [\overline{a}_1, \overline{a}_n]$; and, concerning tangencies with $\mathcal{L}$, that are above L, $[\overline{a}_1, \overline{a}_n]_{C_1}$ has less ones than $[\overline{a}_1, \overline{a}_n]$.

In this deformation the arc $[p, Tp] \subset C$ has been taken to an interval whose image by X is below L. This deformation takes place inside a small neighborhood of $\{[x, Tx]_Y : x \in [p, q]\}$) and so C_1 surrounds the origin. Also as both $[\overline{a}_1, \overline{a}_n] \subset C$ and $\mathcal{L}$ are transversal to the vertical foliation, $[\overline{a}_1, \overline{a}_n]_{C_1} \subset C_1$ can be obtained to be transversal to the vertical foliation. We do not care if C_1 has more self-intersections than C.

The second case happens when $\{p,r\} \subset (a_1,a_n)$ and the arc of trajectory $[p,r]_Y$ of Y meets C according to the conditions (i) of Lemma 3. The arguments of such lemma imply that we may deform of C into a new circle C_1 according to the following conditions. The deformation fixes $C \setminus (\overline{a}_1,\overline{a}_n)$ and takes $[\overline{a}_1,\overline{a}_n] \subset C$ to a closed sub-interval $[\overline{a}_1,\overline{a}_n]_{C_1} \subset C_1$ such that

(f) $L \cap [\overline{a}_1,\overline{a}_n]_{C_1}$ has the same number of elements than $L \cap [\overline{a}_1,\overline{a}_n]$; and, concerning tangencies with $\mathcal{L}$, that are above L, $X([\overline{a}_1,\overline{a}_n]_{C_1}])$ has one less than $X([\overline{a}_1,\overline{a}_n])$.

As above, this deformation takes place inside a small neighborhood of $\cup\{[x,Tx]_Y : x \in [p,q]\}$) and so C_1 surrounds the origin. Also as $X([\overline{a}_1,\overline{a}_n])$ and $\mathcal{L}$ are transversal to the vertical foliation, $[\overline{a}_1,\overline{a}_n]_{C_1} \subset C_1$ can be obtained so that $X([\overline{a}_1,\overline{a}_n]_{C_1})$ is transversal to the vertical foliation. Again, as in case above, we do not care if C_1 has more self- intersections than C.

The third case occurs when $\{p,r\} \subset (a_1,a_n)$ and the arc of trajectory $[p,r]_Y$ of Y does not meet C according to the condition (i) of Lemma 3. We shall show now that this case is not possible. In fact, otherwise, this supposition and (d) imply that the open subinterval of trajectory (p,r) meets tangentially C exactly once, say at s.

Let $[p,s]$ and $[s,r]$ be the subintervals of C such that $[p,s]_Y \cup [p,s]$ and $[s,r]_Y \cup [s,r]$ bound discs $\overline{D}([p.s]_Y \cup [p,s])$ and $\overline{D}([s,r]_Y \cup [s,r])$ contained in $\mathbb{R}^2 \setminus D(C)$. Then, either $[s,r]_Y$ is contained in $\overline{D}([p.s]_Y \cup [p,s])$ or $[p,s]_Y$ is contained in $\overline{D}([s,r]_Y \cup [s,r])$. If $[s,r]_Y$ is contained in $\overline{D}([p.s]_Y \cup [p,s])$, then the circle $(C \setminus [p,s]) \cup [p,s]_Y$ can be approximated by a circle C_1 such that $X(C_1)$ has exactly two tangencies with the vertical foliation: $\{X(a), X(s)\}$. It follows from (a7) that X meets C_1 with an internal tangency at s. As X meets C_1 with an external tangency at a, we conclude, by Theorem 3 that $j_X(C_1) = 1$. This contradiction with Corollary 1 shows that $[s,r]_Y$ is not contained in $\overline{D}([p.s]_Y \cup [p,s])$. Similarly, $[p,s]_Y$ is not contained in $\overline{D}([s,r]_Y \cup [s,r])$. This contradiction implies that the third case is not possible.

As cases 1 and 2 above are the only possible ones, and thanks to (e)-(f), we only need to perform finitely many times the process (just described above) of obtaining new circles, of ETT for $\nabla f^{\#}$, in order to finally obtain a smooth circle, say C_2, such that $X([\overline{a}_1,\overline{a}_n]_{C_2})$ is below L. Similarly, by a deformation that fixes $C_2 \setminus [\overline{b}_1,\overline{b}_n]$ we shall finally obtain one circle as requested in this lemma. □

Proof of Proposition 3. — By (a8), $X([\overline{a}_1,a_1) \cup (a_n,\overline{a}_n])$ is below $X([\overline{b}_1,b_1) \cup (b_n,\overline{b}_n])$. It is easy to see that we may deform C, locally around $\{a_1,b_1,a_n,b_n\}$ so that the new circle of ETT for $\nabla f^{\#}$ satisfies the conditions of Lemma 7 and so it can be deformed into one as requested to prove this proposition. □

As a direct consequence of this proposition and Proposition 2 we obtain:

Corollary 3. — *Under the assumptions of Theorem 1, there exists an embedded circle $C \subset \mathbb{R}^2 \setminus \overline{D}_\sigma$ such that X restricted to $\mathbb{R}^2 \setminus \overline{D}(C)$ can be extended to an orientation preserving embedding from $\mathbb{R}^2$ into $\mathbb{R}^2$.*

6. Proof of Theorem 1

Item (i) of Theorem 1 follows directly from Corollary 3.

We shall need the following result of Gutierrez and Teixeira [**G-T**]

Theorem 4. — *Let $\sigma > 0$ and $Y : \mathbb{R}^2 \to \mathbb{R}^2$ be a C^1 vector field satisfying the following conditions:*

(i) *Y has a singularity, say S;*
(ii) *for all $p \in \mathbb{R}^2 \smallsetminus \overline{D}_\sigma$, no eigenvalue of $DY(p)$ belongs to $\{z \in \mathbb{C} : \Re(z) \geqslant 0\}$;*
(iii) *for all $p \in \mathbb{R}^2$, $\mathrm{Det}(DY(p)) > 0$.*

If $\mathcal{I}(Y) = \displaystyle\int_{\mathbb{R}^2} \mathrm{Trace}(DY)$ is less than 0 (resp. greater of equal than 0), then the point ∞ of the Riemann sphere $\mathbb{R}^2 \cup \{\infty\}$ is a repellor (resp. an attractor) of Y.

Proceed to prove item (ii) of Theorem 1. Under the terms of Corollary 3, X can be extended to an orientation preserving embedding $\widetilde{X} : \mathbb{R}^2 \to \mathbb{R}^2$. Choose $p_0 \in \mathbb{R}^2$ such that $\widetilde{X} + p_0$ has a singularity. By applying Theorem 4, we shall obtain that the point ∞ of the Riemann sphere $\mathbb{R}^2 \cup \{\infty\}$ is either a repellor or an attractor of $\widetilde{X} + p_0$. This proves item (ii) of Theorem 1 because, around infinity, $\widetilde{X} + p_0$ and $X + p_0$ coincide.

We should comment that there are vector fields of $\mathbb{R}^2$, as in Theorem 4, having either attracting or repelling behavior at ∞ [**G-T**]. For sake of completeness we present in next section the example of Gutierrez and Teixeira of a vector field, as in Theorem 4, having attracting behavior at ∞.

7. An example

The purpose of this section is to exhibit a vector field X satisfying the conditions of Theorem 4 and such that the unstable manifold $W^u(0)$, of 0, is $\mathbb{R}^2$. In particular "∞" is an attractor of X. The required vector field is given by:

$$X(x, y) = g(r)(e^{-r}x - y, x + e^{-r}y)$$

where $r = \sqrt{x^2 + y^2}$ and $g(r) = \dfrac{1 - e^{-r}}{r\sqrt{1 + e^{-2r}}}$.

The following expressions can be obtained by a symbolic computer system like Mathematica:

(a) $\mathrm{Det}(DX) = \dfrac{e^r - 1}{re^{2r}}$

(b) $\mathrm{Trace}(DX) = \dfrac{e^r + (r-1)(1 + 2e^{2r} - e^{3r})}{re^{4r}(1 + e^{-2r})^{3/2}}$

(c) $\mathcal{I}(X) = \displaystyle\int_{\mathbb{R}^2} \mathrm{Trace}(DX) = 0.$

It is clear that $\mathrm{Det}(DX) > 0$ everywhere and that $\mathrm{Trace}(DX) < 0$ in the region $\{r > K\}$ for some large K.

It follows, from (a)–(c) and Theorem 4, that "∞" is an attractor of X. To obtain a stronger conclusion, we may observe that the inner product

$$\langle (x,y), X(x,y)\rangle = g(r)r^2e^{-r}$$

is greater than 0, for all $\sigma > 0$; therefore, the vector field X points outside all discs whose boundary has the form $\{r = \text{constant}\}$. This implies that $W^u(0) = \mathbb{R}^2$.

References

[CGL] M. Cobo, C. Gutierrez & J. Llibre. *On injectivity of C^1–maps of the real plane.* Preprint.

[Chi] C. Chicone. *Ordinary Differential Equations with Applications.* Springer Graduate Texts in Applied Mathematica, 34 (1999).

[DRS] F. Dumortier, R. Roussarie & J. Sotomayor. *Bifurcations of Planar Vector Fields: Nilpotent Singularities and Abelian Integrals.* Springer Lecture Notes in Mathematics, V. 1480 (1991).

[Fes] R. Fessler. *A proof of the two dimensional Markus–Yamabe Stability Conjecture and a generalization.* Annales Polinici Mathematici LXII.1, pp. 45–74 (1995).

[G-G] G. Golubitsky and V. Guillemin. *Stable Mappings and Their Singularities.* Springer-Verlag New York Inc. (1980).

[Gut] C. Gutierrez. *A solution to the bidimensional Global Asymptotic Stability Conjecture.* Ann. Inst. Henri Poincaré. V. 12 No 6, 1995, p 627-671.

[G-T] C. Gutierrez and M. Teixeira. *Asymptotic stability at infinity of planar vector fields.* Bol. Soc. Bras. Mat. Vol. 26, No. 1, pp 57-66 (1995).

[Har] P. Hartman. *Ordinary Differential Equations*, John Wiley and Sons, Inc. (1964).

[Hir] M.W. Hirsch. *Differential Topology* Graduate texts in mathematics, vol. 33; Springer-Verlag New York Inc. (1976)

[Rou] R. Roussarie. *Bifurcation of Planar Vector Fields and Hilbert's* 16^{th} *Problem.* Birkhauser Verlag, Basel (1998).

[Tak] F. Takens. *Singularities of Vector Fields.* Publ. Math. IHES, 43 (1974).

C. Gutierrez, ICMC-USP, São Carlos & IMPA, Rio de Janeiro, ICMC-USP: Av. Dr. Carlos Botelho, 1465; Caixa Postal 668, CEP- 13560-970, Sao Carlos-SP, Brazil
E-mail : gutp@icmc.sc.usp.br

A. Sarmiento, UFMG-ICEx-Dpto. de Matemática, Av. Antônio Carlos 6627, Pampulha, CEP 30161-970 - Belo Horizonte-MG-Brazil • *E-mail :* sarmient@mat.ufmg.br
Url : www.mat.ufmg.br/~sarmiento/

Astérisque
287, 2003, p. 103–123

AVERAGING IN DIFFERENCE EQUATIONS DRIVEN BY DYNAMICAL SYSTEMS

by

Yuri Kifer

Dedicated to Jacob Palis for his sixtieth birthday

Abstract. — The averaging setup arises in the study of perturbations of parametric families of dynamical systems when parameters start changing slowly in time. Usually, averaging methods are applied to systems of differential equations which combine slow and fast motions. This paper deals with difference equations case which leads to wider class of models and examples. The averaging principle is justified here under a general condition which is verified when unperturbed transformations either preserve smooth measures or they are hyperbolic. The convergence speed in the averaging principle is estimated for some cases, as well.

1. Introduction

In the study of evolution of many real systems we can usually observe only few parameters while other less significant ones are regarded as constant in time. A more precise investigation may reveal that these parameters change, as well, but much slower than the others. These leads to complicated double scale equations describing slow and fast motions which are difficult to solve directly. Such problems were encountered with already long ago in celestial mechanics in the study of perturbations of planetary motion. People noticed that good approximations of the slow motion on long time intervals can be obtained by averaging coefficients of its equation in fast variables. This averaging principle was applied in celestial mechanics long before it was rigourously justified in some cases in the middle of the 20th century (see [**18**] and historical remarks there).

Traditionally, averaging methods were employed in the study of two scale ordinary differential equations describing a continuous time motion. On the other hand, it is well known that the study of discrete time dynamical systems, i.e. of iterates of

2000 ***Mathematics Subject Classification.*** — Primary: 34C05; Secondary: 39A11, 60J05, 37D20.
Key words and phrases. — Averaging, difference equations, dynamical systems.

transformations (not necessarily invertible), enables us to deal with a wider class of models and examples and to reveal new effects. Suppose that an idealized physical system can be described by a transformation F_0 of a $(d+m)$-dimensional space and there exist functions $x_1, \dots, x_d$ which do not change along orbits of F_0 (integrals of motion). Then, generically, F_0 can be written as a transformation of a locally trivial fiber bundle $\mathcal{M} = \{(x,y) : x \in \mathbb{R}^d, y \in M_x\}$ with base $\mathbb{R}^d$ and fibers M_x being m-dimensional manifolds acting by the formula $F_0(x,y) = (x, f_x y)$ where $f_x = f(x,\cdot) : M_x \to M_x$ is a transformation of M_x. It is natural to view a real physical system as a perturbation of the above idealized one, and so it should be described by a transformation

$$F_\varepsilon(x,y) = (x + \varepsilon\Phi(x,y,\varepsilon), f(x,y,\varepsilon)) \tag{1.1}$$

where $\Phi(\cdot,\cdot,\varepsilon) : \mathcal{M} \to \mathbb{R}^d$ and $f(x,\cdot,\varepsilon) : M_x \to M_x$. Since locally $\mathcal{M}$ has a product structure $U \times M$, where U is an open subset of $\mathbb{R}^d$ and M is an m-dimensional manifold, and iterates $F_\varepsilon^n(x,y)$ of any point (x,y) in $U \times M$ stay there for all $n \leqslant \delta/\varepsilon$ with small but fixed $\delta = \delta(x) > 0$ we conclude that it suffices to study the evolution on time intervals of order $1/\varepsilon$ only on product spaces and then glue pieces of orbits together.

In this paper we consider difference equations of the form

$$\begin{aligned} X^\varepsilon(n+1) - X^\varepsilon(n) &= \varepsilon\Phi(X^\varepsilon(n), Y^\varepsilon(n), \varepsilon), \quad X^\varepsilon(0) = x, \\ Y^\varepsilon(n+1) &= f(X^\varepsilon(n), Y^\varepsilon(n), \varepsilon), \qquad Y^\varepsilon(0) = y \end{aligned} \tag{1.2}$$

where $X^\varepsilon(n) = X^\varepsilon_{x,y}(n) \in \mathbb{R}^d$, $Y^\varepsilon(n) = Y^\varepsilon_{x,y}(n)$ runs on a compact m-dimensional Riemannian manifold M, $\Phi = \Phi(x,y,\varepsilon)$ is a Lipschitz in x, y, ε vector function, $f_x(\cdot,\varepsilon) = f(x,\cdot,\varepsilon)$ is a family of smooth maps (usually, endomorphisms or diffeomorphisms) of M close to f_x. Thus $(X^\varepsilon_{x,y}(n), Y^\varepsilon_{x,y}(n)) = F^n_\varepsilon(x,y)$. The equations (1.2) usually cannot be solved explicitly and it is desirable to approximate its solutions for small ε. Returning back to the unperturbed $\varepsilon = 0$ case eliminates the slow motion X^ε completely and gives a rather pure approximation valid only for bounded time intervals. The averaging principle is supposed to give a prescription how to approximate the slow motion X^ε on time intervals of order $1/\varepsilon$. Recurrent relations (1.2) can be regarded as a more general than usual setup for perturbations of dynamical systems where not only the transformation itself is perturbed but also we begin to take into account evolution of some parameters whose change was disregarded before.

We note that the standard continuous time averaging setup (see [**13**]) can be always reduced by discretizing time to a model described by difference equations of type (1.2). On the other hand, an attempt to go the other way around faces substantial difficulties since the standard suspension construction should be implemented now for different transformations f_x and it is not clear how to glue everything together in an appropriate way. Observe, that (1.2) can be generalized adding some randomness in the right hand sides there so that $f_x(\cdot,\varepsilon)$ become random endomorphisms, but we will not discuss this setup here.

Assume, first, that the fast motion $Y^\varepsilon(n)$ is independent of the slow variables, i.e. $f(x,y,\varepsilon)=fy$, and so $Y^\varepsilon_{x,y}(n)=f^ny$. For an ergodic f-invariant probability measure μ the limit

$$\overline{\Phi}_\mu(x)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\Phi(x,f^ny)=\int\Phi(x,y)d\mu(y) \tag{1.3}$$

exists for μ-almost all y. For such y's uniformly in n the solution $X^\varepsilon_{x,y}$ of (1.2) is close on any time interval of order $1/\varepsilon$ to the solution $\overline{X}^\varepsilon=\overline{X}^\varepsilon_\mu=\overline{X}^\varepsilon_{x,\mu}$, taken at integer times, of the differential equation

$$\frac{d\overline{X}^\varepsilon(t)}{dt}=\varepsilon\overline{\Phi}(\overline{X}^\varepsilon(t)),\quad \overline{X}^\varepsilon(0)=x \tag{1.4}$$

where $\overline{\Phi}=\overline{\Phi}_\mu$ (see similar continuous time results in [**18**]). Already in this case the averaging principle works only for μ-almost all initial points y and for different y's averaged solutions may be different. In the particular case when f is uniquely ergodic the convergence in (1.3) is uniform in y and for all y, whence the averaged equation (1.4) and its solution are unique and the latter approximates $X^\varepsilon(n)$, $n\in[0,N/\varepsilon]$ uniformly.

The general case (1.2) when the fast and the slow motions are fully coupled is much more complicated. The averaging principle suggests here to approximate X^ε_x by $\overline{X}^\varepsilon_x$ satisfying (1.4) but with $\overline{\Phi}$ given by

$$\overline{\Phi}(x)=\overline{\Phi}_y(x)=\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\Phi(x,f^k_xy) \tag{1.5}$$

provided the last limit exists for "most" x and y. If μ_x is an ergodic invariant measure of f_x then the limit (1.5) exists for μ_x-almost all y's and

$$\overline{\Phi}(x)=\overline{\Phi}_{\mu_x}(x)=\int\Phi(x,y)d\mu_x(y). \tag{1.6}$$

Observe that Lipschitz continuity of $\overline{\Phi}$ cannot be guaranteed now without further assumptions even for smooth Φ, and so we do not have automatically existence and, especially, uniqueness of solutions in (1.4) in these general circumstances. On the other hand, consider the recurrent relation for $\overline{\overline{X}}^\varepsilon(n)=\overline{\overline{X}}^\varepsilon_x(n)$,

$$\overline{\overline{X}}^\varepsilon(n+1)=\overline{\overline{X}}^\varepsilon(n)+\varepsilon\overline{\Phi}(\overline{\overline{X}}^\varepsilon(n)),\quad \overline{\overline{X}}^\varepsilon(0)=x \tag{1.7}$$

which determines $\overline{\overline{X}}^\varepsilon(n)$ without any conditions on $\overline{\Phi}$ and it is easy to see that if $\overline{\Phi}$ is Lipschitz continuous and bounded then

$$\max_{0\leqslant n\leqslant T/\varepsilon}|\overline{X}^\varepsilon_x(n)-\overline{\overline{X}}^\varepsilon_x(n)|\leqslant C_T\varepsilon \tag{1.8}$$

for some $C_T > 0$ independent of ε. Thus we may discuss the approximation of $X^\varepsilon(n)$ by $\overline{\overline{X}}^\varepsilon(n)$ under more general conditions when we even do not have uniquely defined solutions of (1.4).

In general, there exists no natural family of invariant measures μ_x, $x \in \mathbb{R}^d$, since the transformations f_x may have rather different properties for different x's and the averaging principle can be justified here only under substantial restrictions. First, the averaging prescription relies here on existence of a family of probability measures μ_x such that the limit (1.5) exists μ_x-almost everywhere (a.e.) and it is given by (1.6) (at least, Lebesgue a.e. in x). Of course, in addition, we need sufficiently good dependence of Φ and f in (1.2) on ε but still, this does not seem to be enough, in general. The problem here is that the average in (1.5) is taken along orbits of the unperturbed fast motion but in the perturbed evolution (1.2) we cannot disregard now changes in the slow variable parameter of the fast motion, and so we have to study the interplay between unperturbed and perturbed dynamics. Namely, the method of this paper relies on measure estimates of sets of pairs (x, y) which arrive under the action of F_ε^k to sets of points with a specified behavior of averages for the unperturbed evolution. Then we will show that the slow motion is close to the averaged one in certain L^1-sense. Required estimates can be done assuming, for instance, that each f_x is a smooth endomorphism or a diffeomorphism of M preserving a smooth measure μ_x on M which is ergodic for Lebesgue almost all (a.a.) x. This result is a discrete time version of Anosov's theorem [**1**] which is one of few general results about fully coupled averaging. Actually, we prove our result under a general condition which is satisfied in essentially all known cases where the averaging principle holds true and it does not rely on existence of smooth invariant measures as in Anosov's approach.

Recently, quite a few papers dealt with a class of diffeomorphisms called stably ergodic (see, for instance, [**5**]) which are volume preserving ergodic diffeomorphisms having a C^2-neighborhood of volume preserving ergodic diffeomorphisms. If each f_x from our parametric family belongs to such a neighborhood then our results yield an L^1-convergence in the averaging principle. Moreover, we need ergodicity only for almost all x's which suggests to study parametric families of volume preserving diffeomorphisms which are ergodic for almost all parameter values. When convergence in the averaging principle in a fully coupled setup (1.2) holds true for any reasonable Φ we can naturally regard this as a manifestation of compatibility of f_x's or their stability within our parametric family.

Observe that our result works in the case when all f_x's are C^2 expanding transformations of M which always possess fast mixing smooth invariant measures μ_x. On the other hand, close relatives of expanding transformations Anosov and Axiom A diffeomorphism do not possess, generically, smooth invariant measures. Still, relying on specific properties of Axiom A system in a neighborhood of an attractor we will be able to carry out necessary estimates for μ_x being either Lebesgue or corresponding

Sinai-Ruelle-Bowen (SRB) measures, and so the averaging principle will be justified in this case, as well. Moreover, using moderate deviations estimates from [**12**] for this case we will give an estimate of deviation of the slow motion from the averaged one. More delicate limit theorems (large deviations, central limit theorem etc.) for these deviations will be studied in another paper. Some relevant results in this direction were obtained recently in [**3**].

Our conditions need ergodicity of measures μ_x only for a.a. and not all x's which is important in the presence of resonances. For instance, let f_x be a parametric family of toral translations. All of them preserve the Lebesgue measure but only translations with rationally independent mod 1 frequencies are ergodic. Assuming that these frequencies depend only on the slow variable x we see that, generically, they will be rationally independent mod 1 for Lebesgue almost all and not all x's. For such translations and also for some skew translations of the torus (which are both uniquely ergodic) we will be able to estimate the speed of convergence in the averaging principle deriving a discrete time version of Neistadt's theorem (see a comprehensive exposition of Anosov's and Neistadt's theorems in [**13**]).

The author is grateful to the anonimous referee for several useful suggestions improving the exposition.

2. Preliminaries and main results

Assume that the right hand sides in (1.2) satisfy

$$|\Phi(x,y,\varepsilon)-\Phi(z,v)|+d_M(f(x,y,\varepsilon),f(z,v))\leqslant L(\varepsilon+|x-z|+d_M(y,v))\\ \text{and } |\Phi(x,y,\varepsilon)|\leqslant L \tag{2.1}$$

for some $L>0$ independent of $\varepsilon>0$, $x,z\in\mathbb{R}^d$ and $y,v\in M$, where $\Phi(z,v)=\Phi(z,v,0)$, $f(z,v)=f_zv=f(z,v,0)$, $f_x=f(x,\cdot):M\to M$ for each $x\in\mathbb{R}^d$ is a Lipschitz map and d_M is the Riemannian metric on M. In Corollaries 2.2 and 2.3 below we will assume also that

$$\|\Phi(\cdot,\cdot,\varepsilon)\|_{C^1}+\|f(\cdot,\cdot,\varepsilon)\|_{C^1}\leqslant L \quad \text{and} \quad \|D_{x,y}f(\cdot,\cdot,\varepsilon)-D_{x,y}f(\cdot,\cdot)\|\leqslant L\varepsilon \tag{2.2}$$

where $\|\cdot\|_{C^k}$ is the C^k norm of the corresponding map and $D_{x,y}f$ is the differential at (x,y) of the map $f:\mathbb{R}^d\times M\to M$. Our setup includes also a family of probability measure μ_x, $x\in\mathbb{R}^d$ on M depending measurably on x. For each $n\in\mathbb{N}$ and $\delta>0$ set

$$E(n,\delta)=\Big\{(x,y):\Big|\frac{1}{n}\sum_{k=0}^{n-1}\Phi(x,f_x^ky)-\overline{\Phi}(x)\Big|>\delta\Big\}$$

where $\overline{\Phi}(x)=\int_M\Phi(x,y)d\mu_x(y)$. Assume that for all $x,z\in\mathbb{R}^d$,

$$\Big|\overline{\Phi}(x)-\int_M\Phi(x,y)d\mu_z(y)\Big|\leqslant L^2|x-z| \tag{2.3}$$

and there exist $\alpha, \varepsilon_0 > 0$ such that for any $T, \delta > 0, k \in \mathbb{N}$, and a compact $K \subset \mathbb{R}^d$ we can find $d_{T,K}(k,\delta) \to 0$ as $k \to \infty$ and $\eta(\varepsilon) \to \infty$ as $\varepsilon \to 0$ such that for any $\varepsilon \leqslant \varepsilon_0$ and $k \leqslant \eta(\varepsilon)$,

$$\mu\big((K \times M) \cap F_\varepsilon^{-n} E(k,\delta)\big) \leqslant d_{T,K}(k,\delta) \quad \text{if } n \leqslant T/\varepsilon - k, \tag{2.4}$$

where $d\mu(x,y) = d\mu_x(y) d\ell(x)$ and ℓ is the Lebesgue measure on $\mathbb{R}^d$.

Theorem 2.1. — *Suppose that (2.1), (2.3) and (2.4) hold true. Then for any $T > 0$ and a compact set $K \subset \mathbb{R}^d$,*

$$\lim_{\varepsilon \to 0} \int_K \int_M \sup_{0 \leqslant n \leqslant T/\varepsilon} |X_{x,y}^\varepsilon(n) - \overline{X}_x^\varepsilon(n)| d\mu_x(y) d\ell(x) = 0 \tag{2.5}$$

where $\overline{X}_x^\varepsilon(t)$ is the solution of (1.4).

The conditions (2.1) and (2.2) are clear and rather standard, the condition (2.3) is less straightforward, in general, while the assumption (2.4) is far from being transparent. We will provide in Corollaries 2.2 and 2.3 two important classes of transformations f_x such that (2.3) and (2.4) hold true for any perturbation satisfying (2.1) and (2.2). It is instructive to verify these conditions in the simpler well known setup when the fast motion does not depend on ε and on the slow one, i.e. when $F_\varepsilon(x,y) = (x + \varepsilon\Phi(x,y,\varepsilon), fy)$ where f is a map of M. Suppose that all measures μ_x coincide with the same ergodic f-invariant probability measure μ_0 on M so that $\mu = \ell \times \mu_0$. Then (2.3) follows automatically and

$$\mu\big((K \times M) \cap F_\varepsilon^{-n} E(k,\delta)\big) = \int_K \mu_0(E_x) d\ell(x)$$

where

$$E_x = E_x(n,k,\delta) = \{y \in M : (X_{x,y}^\varepsilon(n), f^n y) \in E(k,\delta)\}.$$

Set $E_z(k,\delta) = \{y : (z,y) \in E(k,\delta)\}$. By (2.1) we have that $E_u(k,\delta) \subset E_z(k,\delta/2)$ provided $|u - z| \leqslant \delta/4L$. Hence,

$$\begin{aligned} E_{x,z} = E_{x,z}(n,k,\delta) = \{y \in M : |X_{x,y}^\varepsilon(n) - z| \leqslant \delta/4L \text{ and } f^n y \in E_{X_{x,y}^\varepsilon(n)}(k,\delta)\} \\ \subset f^{-n} E_z(k,\delta/2). \end{aligned}$$

Let K_r denotes the closed r-neighborhood of K. By (2.1), $X_{x,y}^\varepsilon(n) \in K_{LT}$ if $x \in K$ and $n \leqslant T/\varepsilon$. Thus, if $z_1, \ldots, z_l$ is a minimal $\delta/4L$-net in K_{LT} then $X_{x,y}^\varepsilon(n)$ belongs to a $\delta/4L$-ball around some z_i provided $n \leqslant T/\varepsilon$. Then, for any $x \in K$,

$$E_x(n,k,\delta) \subset \cup_{i=1}^l E_{x,z_i}(n,k,\delta) \subset \cup_{i=1}^l f^{-n} E_{z_i}(k,\delta/2).$$

Since μ_0 is f-invariant and ergodic we obtain from here that

$$\mu_0(E_x(n,k,\delta)) \leqslant \sum_{i=1}^l \mu_0(E_{z_i}(k,\delta/2)) \to 0 \quad \text{as } k \to \infty$$

and (2.4) follows.

We note that though the assumption (2.4) does not seem to be weakest possible it is rather clear that without some compatibility between measures μ_x, $x \in \mathbb{R}^d$ the averaging principle is not going to work, in general. Consider, for instance, the following simplest example where $d = 1$, M is a circle $\mathbb{T}^1$ of length 1, all f_x coincide with the identity transformation of $\mathbb{T}^1$, $\Phi(x, y, \varepsilon) = \Phi(y)$ is a C^1 function depending only on y. We define μ_x to be the unit mass at x (mod 1) regarded as a point of $\mathbb{T}^1$ which is identified with the unit interval whose end points are glued together. Of course, each μ_x is an ergodic invariant measure of the identity transformation and (2.3) holds true, as well. Extending Φ as a 1-periodic function to the whole $\mathbb{R}^1$ we can write $\overline{\Phi}(x) = \int \Phi(y) d\mu_x(y) = \Phi(x)$ for any $x \in \mathbb{R}^1$. Then $d\overline{X}^\varepsilon(t)/dt = \varepsilon\Phi(\overline{X}^\varepsilon(t))$ and $\overline{Z}(t) = \overline{X}^\varepsilon(t/\varepsilon)$ satisfies $d\overline{Z}(t)/dt = \Phi(\overline{Z}(t))$. Clearly, $X^\varepsilon_{x,y}(n) = x + \varepsilon n\Phi(y)$, and so

$$\int_0^1 \int_{\mathbb{T}^1} \sup_{0 \leqslant n \leqslant T/\varepsilon} |X^\varepsilon_{x,y}(n) - \overline{X}^\varepsilon_x(n)| d\mu_x(y) d\ell(x) = \int_0^1 \sup_{0 \leqslant t \leqslant T} |t\Phi(x) + x - \overline{Z}_x(t)| d\ell(x) + O(\varepsilon).$$

The last integral is positive, in general, (take, for instance, $\Phi(x) = \cos^2 2\pi x$ obtaining $\overline{Z}_x(t) = (2\pi)^{-1} \arctan(2\pi t + \tan 2\pi x)$) and it does not depend on ε, so we do not have (2.5) in this case. More substantial examples of nonconvergence in (2.5) can be constructed, as well, but the whole question is not yet completely understood.

Next, we will provide more specific conditions which ensure that (2.4) is satisfied. We assume now that (2.1) and (2.2) hold true and that for each $x \in \mathbb{R}^d$ we are given an f_x-invariant probability measure μ_x where f_x is supposed to be now a C^2-endomorphism of M, i.e. its differential $D_y f_x$ is nondegenerate at any point $y \in M$.

Corollary 2.2. — *Suppose that each measure μ_x has a Radon-Nikodim derivative $q(x, y) = q_x(y) = d\mu_x(y)/d\rho(y)$ with respect to the normalized Riemannian volume ρ on M such that*

$$\|q\|_{C^1} + \|1/q\| \leqslant L \tag{2.6}$$

where $\|\cdot\|_{C^1}$ and $\|\cdot\|$ are corresponding C^1 and supremum norms. Then there exists $C = C_{T,K}$ such that for all $n \leqslant T/\varepsilon$ and $k \in \mathbb{N}$,

$$\mu((K \times M) \cap F_\varepsilon^{-n} E(k, \delta)) \leqslant C\mu((K_{LT} \times M) \cap E(k, \delta)) \tag{2.7}$$

were, again, $d\mu(x, y) = d\mu_x(y) d\ell$. Assume, in addition, that for ℓ-a.a. x the limit (1.5) exists μ_x-a.e. and it is given by (1.6). Then $\mu((K_{LT} \times M) \cap E(k, \delta)) \to 0$ as $k \to \infty$ and (2.4) follows. Since (2.1) and (2.6) imply (2.3) then (2.5) holds true, as well. Clearly, the measures μ_x can be replaced there by the Riemannian volume ρ.

We claim that (2.4) is also satisfied in the setup of hyperbolic diffeomorphisms. Namely, we assume now that f_x, $x \in \mathbb{R}^d$ are diffeomorphisms and for each x there exists a compact f_x-invariant set $\Lambda_x \subset M$ which is a basic hyperbolic attractor for f_x (see [**11**]). Moreover, we assume that there exists an open set $W \subset M$ such that for all $x \in \mathbb{R}^d$,

$$\Lambda_x \subset W, \quad f_x\overline{W} \subset W, \quad \text{and} \quad \cap_{n>0} f_x^n W = \Lambda_x. \tag{2.8}$$

Denote by μ_x the Sinai-Ruelle-Bowen (SRB) invariant measure of f_x on Λ_x. Recall, that μ_x can be obtained as a weak limit of $f_x^n \rho_W$ as $n \to \infty$ where ρ_W is the normalized restriction of the Riemannian volume ρ on M to W (see [**11**]). There are several other important characterizations of the SRB measure μ_x, in particular, it is the unique equilibrium state of f_x for the function

$$\varphi_x(y) = -\log J_x^u(y) \tag{2.9}$$

where $J_x^u(y)$ is the absolute value of the Jacobian with respect to the Riemannian inner products of the linear map $D_y f_x : \Gamma_{x,y}^u \to \Gamma_{x,f_x y}^u$ where $T_{\Lambda_x} M = \Gamma_x^s \oplus \Gamma_x^u$ is the hyperbolic splitting. The measure μ_x sits on Λ_x and, in general, even when $\Lambda_x = M$ (Anosov diffeomorphism case) is singular with respect to the Riemannian volume ρ so that Corollary 2.2 is not applicable here.

Corollary 2.3. — *Suppose that (2.1) and (2.2) hold true and for each $x \in \mathbb{R}^d$ a C^2 diffeomorphism f_x of M is given which C^2 depends on x and possesses a basic hyperbolic attractor Λ_x satisfying (2.8). Then (2.3)–(2.5) hold true if each μ_x is taken to be the corresponding SRB measure. This remains true if instead of SRB measures we take in Theorem 2.1 μ_x coinciding for each x with the Riemannian volume ρ_W restricted to the set W satisfying (2.8). Moreover, for each $\gamma > 0$ and a compact set K there exists $C_{K,\gamma} > 0$ such that*

$$\int_K \int_M \sup_{0 \leqslant n \leqslant T/\varepsilon} |X_{x,y}^\varepsilon(n) - \overline{X}_x^\varepsilon(n)| d\rho_W(y) d\ell(x) \leqslant C_{K,\gamma} (\log 1/\varepsilon)^{-\frac{1}{2}+\gamma}. \tag{2.10}$$

Note that the convergence (2.5) can be derived from the results announced in [**3**] but we consider it useful to have an independent direct proof based on Theorem 2.1. The bound (2.10) will be derived from estimates of the next section and the moderate deviations asymptotics obtained in [**12**]. The estimate (2.10) holds true also when f_x, $x \in \mathbb{R}^d$ are C^2 expanding endomorphisms of M but (2.5) follows for them already from Corollary 2.2 since they preserve smooth ergodic invariant measures (see, for instance, [**14**]). Moreover, it is possible to extend Corollary 2.3 to the continuous time case of flows with hyperbolic attractors. If f_x do not depend on x then methods from [**12**] yield easily much better estimate of order $\sqrt{\varepsilon}$ for the left hand side of (2.10) (cf. [**8**]). In the general case $Y_{x,y}^\varepsilon(n)$ and $f_x^n y$ diverge exponentially fast and the arguments of the next section yield only a logarithmic estimate of speed of convergence in (2.10). Still, a more precise study of normalized deviations $\varepsilon^{-1/2}(X_{x,y}^\varepsilon([t/\varepsilon]) - \overline{X}_x^\varepsilon([t/\varepsilon]))$ which should lead also to the central limit theorem here is likely to provide the order $\sqrt{\varepsilon}$ estimate for the left hand side of (2.10) in the general case, as well.

Next, we consider two types of specific uniquely ergodic diffeomorphisms f_x of an m-dimensional torus $\mathbb{T}^m$ falling into the framework of Corollary 2.2 for which we will be able to obtain good estimates of speed of convergence in (2.5). First type of these diffeomorphisms consists of translations of $\mathbb{T}^m$ defined by

$$f_x(y_1,\ldots,y_m) = \big(\{y_1+\omega_1(x)\},\ldots,\{y_m+\omega_m(x)\}\big) \tag{2.11}$$

where $\omega(x) = (\omega_1(x),\ldots,\omega_m(x))$ is a vector function of frequencies and $\{a\}$ denotes the fractional part of a. All these f_x preserve the Lebesgue measure ρ on $\mathbb{T}^m$ and it is well known (see, for instance, [**7**]) that f_x is ergodic (and even uniquely ergodic) if and only if the vector $\omega(x)$ has rationally independent mod 1 components. The second class of diffeomorphisms consists of skew translations of $\mathbb{T}^m$ having the form

$$\begin{aligned} &f_x(y_1,\ldots,y_m)\\ &\quad = \big(\{y_1+\alpha(x)\},\{y_2+p_{21}y_1\},\ldots,\{y_m+p_{m1}y_1+\cdots+p_{m,m-1}y_{m-1}\}\big)\end{aligned} \tag{2.12}$$

where p_{ij} are positive integers and α is a function on $\mathbb{R}^d$. Again, each f_x preserves the Lebesgue measure ρ and it is ergodic (and uniquely ergodic) if and only if $\alpha(x)$ is irrational.

Since M is now the torus $\mathbb{T}^m$ we can regard $\Phi(x,y)$, $y=(y_1,\ldots,y_m)$ as a vector function on $\mathbb{R}^d\times\mathbb{R}^m$ 1-periodic in each y_j, $j=1,\ldots,m$. Furthermore, we assume that $\Phi(x,y)$ can be extended as an analytic function $\Phi(x,y+iz)$ to a strip

$$\{y+iz:\, y\in\mathbb{R}^m,\, |z_i|<\kappa,\, i=1,2,\ldots,m\}\subset\mathbb{C}^m,\quad \kappa>0$$

with $|\Phi|\leqslant L$. The latter condition can be relaxed to finite differentiability similar to [**13**] and it will be used only to get appropriate estimates on remainders of Fourier series. Following, [**13**] we say that a map $\xi:\mathbb{R}^d\to\mathbb{R}^l$ satisfies Kolmogorov's nondegeneracy condition if its Jacobi matrix $(\partial\xi_j/\partial x_k)$ has rank l at any point x which means that $d\geqslant l$ and the maximal absolute value $\lambda_\xi(x)$ of determinants of $l\times l$ submatrices of $(\partial\xi_j/\partial x_k)$ is positive.

Theorem 2.4. — *Suppose that Φ and f satisfy (2.1) and (2.2) and, in addition, Φ satisfies the above analyticity condition and f_x, $x\in\mathbb{R}^d$ are either all translations of $\mathbb{T}^m$ defined by (2.11) or all skew translations defined by (2.12) with $\omega(x):\mathbb{R}^d\to\mathbb{R}^m$ and $\alpha(x):\mathbb{R}^d\to\mathbb{R}$ satisfying Kolmogorov's nondegeneracy condition. Then there exists $c_0>0$ such that for every $c<c_0$, each compact set $K\subset\mathbb{R}^d$ and any $T>0$ there exists $C_{T,K,c}>0$ such that*

$$\int_K\int_M \sup_{0\leqslant n\leqslant T/\varepsilon}|X^\varepsilon_{x,y}(n)-\overline{X}^\varepsilon_x(n)|d\rho(y)d\ell(x)\leqslant C_{T,K,c}\varepsilon^c. \tag{2.13}$$

Moreover, if f_x, $x\in\mathbb{R}^d$ are given by (2.11) then we can take $c_0=1/5$ and if they are given by (2.12) then $c_0=1/(3m+7)$ will do.

Actually, employing the approach from [**15**] it is possible to prove (2.13) for f_x given by (2.11) with $c=1/2$.

3. General estimates and convergence

We begin with a general basic estimate which will be used in the proof of both Theorems 2.1 and 2.4. Set $R_k(x,y) = d_M(Y^\varepsilon_{x,y}(k), f^k_x y)$.

Proposition 3.1. — *Suppose that (2.1) and (2.3) hold true. There exists $C > 0$ such that if $1 \leqslant n(\varepsilon) \leqslant T/\varepsilon$ and $N(\varepsilon)$ is the integral part of $T(\varepsilon n(\varepsilon))^{-1}$ then for all $x \in \mathbb{R}^d$, $y \in M$, $\varepsilon \in (0,1)$ and $\delta > 0$,*

$$\begin{aligned}(3.1)\qquad \sup_{0\leqslant n\leqslant T/\varepsilon} |X^\varepsilon_{x,y}(n) - \overline{X}^\varepsilon_x(n)| \leqslant Ce^{CT} \times \Big(&T\delta + \varepsilon n(\varepsilon)(T+1)\\ &+ \varepsilon T \textstyle\sum_{j=0}^{N(\varepsilon)-1} \big(\sum_{k=0}^{n(\varepsilon)-1} R_k(X^\varepsilon_{x,y}(jn(\varepsilon)), Y^\varepsilon_{x,y}(jn(\varepsilon))\big)\\ &+ n(\varepsilon)\mathbb{I}_{E(n(\varepsilon),\delta)}(X^\varepsilon_{x,y}(jn(\varepsilon)), Y^\varepsilon_{x,y}(jn(\varepsilon)))\Big)\end{aligned}$$

where $\mathbb{I}_\Gamma(v) = 1$ if $v \in \Gamma$ and $= 0$, otherwise.

Proof. — By (1.4)–(1.6) and (2.1),

$$\sup_{k\leqslant s\leqslant k+1} |\overline{X}^\varepsilon_x(s) - \overline{X}^\varepsilon_x(k)| \leqslant L\varepsilon,$$

and so by (2.3),

$$\Big|\int_k^{k+1} \overline{\Phi}(\overline{X}^\varepsilon(s))ds - \overline{\Phi}(\overline{X}^\varepsilon(k))\Big| \leqslant \varepsilon L^2(L+1).$$

Hence, by (1.2) and (1.4) for $X^\varepsilon(n) = X^\varepsilon_{x,y}(n)$ and $\overline{X}^\varepsilon(n) = \overline{X}^\varepsilon_x(n)$ we have

$$\begin{aligned}|X^\varepsilon(n) - \overline{X}^\varepsilon(n)| &\leqslant \varepsilon^2 nL^2(L+1) + \varepsilon\Big|\textstyle\sum_{k=0}^{n-1}\big(\Phi(X^\varepsilon(k), Y^\varepsilon(k), \varepsilon) - \overline{\Phi}(\overline{X}^\varepsilon(k))\big)\Big|\\ &\leqslant \varepsilon^2 nL^2(L+1) + \varepsilon\textstyle\sum_{k=0}^{n-1}\big|\big(\Phi(X^\varepsilon(k), Y^\varepsilon(k), \varepsilon) - \Phi(X^\varepsilon(k), Y^\varepsilon(k))\big|\\ (3.2)\qquad &+ \varepsilon\Big|\textstyle\sum_{k=0}^{n-1}\big(\Phi(X^\varepsilon(k), Y^\varepsilon(k)) - \overline{\Phi}(X^\varepsilon(k))\big)\Big| + \varepsilon\textstyle\sum_{k=0}^{n-1}\big|\overline{\Phi}(X^\varepsilon(k)) - \overline{\Phi}(\overline{X}^\varepsilon(k))\big|\\ &\leqslant \varepsilon^2 nL(L^2+L+1) + \varepsilon\Big|\textstyle\sum_{k=0}^{n-1}\big(\Phi(X^\varepsilon(k), Y^\varepsilon(k)) - \overline{\Phi}(X^\varepsilon(k))\big)\Big|\\ &\qquad + \varepsilon L^2(L+1)\textstyle\sum_{k=0}^{n-1}|X^\varepsilon(k) - \overline{X}^\varepsilon(k)|.\end{aligned}$$

By a version of the discrete Gronwall inequality (see, for instance, Lemma 4.20 in [**9**]) we derive from (3.2) that

$$\begin{aligned}(3.3)\quad |X^\varepsilon(n) - \overline{X}^\varepsilon(n)| \leqslant (1+\varepsilon L(L+1))^{n-1}\Big(&\varepsilon^2 nL(L^2+L+1)\\ &+ \varepsilon\big|\textstyle\sum_{k=0}^{n-1}\big(\Phi(X^\varepsilon(k), Y^\varepsilon(k)) - \overline{\Phi}(X^\varepsilon(k))\big)\big|\Big).\end{aligned}$$

Next, setting $x^\varepsilon_j = X^\varepsilon_{x,y}(jn(\varepsilon))$ and $y^\varepsilon_j = Y^\varepsilon_{x,y}(jn(\varepsilon))$ we obtain by (2.1) that

$$\begin{aligned}(3.4)\quad &\sup\nolimits_{0\leqslant n\leqslant T/\varepsilon}\Big|\textstyle\sum_{k=0}^{n-1}\big(\Phi(X^\varepsilon_{x,y}(k), Y^\varepsilon_{x,y}(k)) - \overline{\Phi}(X^\varepsilon_{x,y}(k))\big)\Big|\\ &\leqslant 2Ln(\varepsilon) + \textstyle\sum_{j=0}^{N(\varepsilon)-1}\Big|\sum_{k=0}^{n(\varepsilon)-1}\big(\Phi(X^\varepsilon_{x^\varepsilon_j,y^\varepsilon_j}(k), Y^\varepsilon_{x^\varepsilon_j,y^\varepsilon_j}(k)) - \overline{\Phi}(X^\varepsilon_{x^\varepsilon_j,y^\varepsilon_j}(k))\big)\Big|,\end{aligned}$$

$$\Big|\sum_{k=0}^{n(\varepsilon)-1}\big(\Phi(X^\varepsilon_{x^\varepsilon_j,y^\varepsilon_j}(k),Y^\varepsilon_{x^\varepsilon_j,y^\varepsilon_j}(k))-\Phi(X^\varepsilon_{x^\varepsilon_j,y^\varepsilon_j}(k),f^k_{x^\varepsilon_j}y^\varepsilon_j)\big)\Big| \leqslant L\sum_{k=0}^{n(\varepsilon)-1}R_k(x^\varepsilon_j,y^\varepsilon_j), \tag{3.5}$$

$$\Big|\sum_{k=0}^{n(\varepsilon)-1}\big(\Phi(X^\varepsilon_{x^\varepsilon_j,y^\varepsilon_j}(k),f^k_{x^\varepsilon_j}y^\varepsilon_j)-\Phi(x^\varepsilon_j,f^k_{x^\varepsilon_j}y^\varepsilon_j)\big)\Big|\leqslant L\sum_{k=0}^{n(\varepsilon)-1}|X^\varepsilon_{x^\varepsilon_j,y^\varepsilon_j}(k)-x^\varepsilon_j|, \tag{3.6}$$

and

$$|X^\varepsilon_{z,v}(k)-z|\leqslant\varepsilon Lk \tag{3.7}$$

for any $z\in\mathbb{R}^d$ and $v\in M$. In addition, by (2.1) and (2.3),

$$\Big|\sum_{k=0}^{n(\varepsilon)-1}\overline{\Phi}(X^\varepsilon_{x^\varepsilon_j,y^\varepsilon_j}(k))-\overline{\Phi}(x^\varepsilon_j)n(\varepsilon)\Big|\leqslant L(L+1)\sum_{k=0}^{n(\varepsilon)-1}|X^\varepsilon_{x^\varepsilon_j,y^\varepsilon_j}(k)-x^\varepsilon_j|. \tag{3.8}$$

Observe also that

$$\Big|\sum_{k=0}^{n(\varepsilon)-1}\Phi(z,f^k_zv)-\overline{\Phi}(z)n(\varepsilon)\Big|\leqslant n(\varepsilon)(\delta+2L\mathbb{I}_{E(n(\varepsilon),\delta)}(z,v)). \tag{3.9}$$

Finally, Proposition 3.1 follows from (3.3)–(3.9). □

Now we can complete the proof of Theorem 2.1. Observe that by (2.1) and (3.7),

$$\begin{aligned}R_k(x,y)&=d_M(f(X^\varepsilon_{x,y}(k-1),Y^\varepsilon_{x,y}(k-1),\varepsilon),f^k_xy)\\&\leqslant d_M(f(X^\varepsilon_{x,y}(k-1),Y^\varepsilon_{x,y}(k-1),\varepsilon),f(X^\varepsilon_{x,y}(k-1),Y^\varepsilon_{x,y}(k-1))\\&\qquad+d_M(f(X^\varepsilon_{x,y}(k-1),Y^\varepsilon_{x,y}(k-1),\varepsilon),f(x,Y^\varepsilon_{x,y}(k-1))\\&\qquad+d_M(f_x(Y^\varepsilon_{x,y}(k-1)),f_x(f^{k-1}_xy))\\&\leqslant L(\varepsilon+L\varepsilon k+R_{k-1}(x,y))\\&=\varepsilon L\sum_{l=0}^{k-1}L^l(1+L(k-l))\\&\leqslant\varepsilon L(1+Lk)(L^k-1)(L-1)^{-1}.\end{aligned} \tag{3.10}$$

Integrating (3.1) against μ over $K\times M$ and taking $n(\varepsilon)=\min([\eta(\varepsilon)],[(log\frac{1}{\varepsilon})^{1-\alpha}])$ for some $\alpha\in(0,1)$ we derive from (2.4), (3.1), and (3.10) that for any $T,\gamma>0$ there exists $C(T,\gamma)>0$ such that

$$\int_K\int_M\sup_{0\leqslant n\leqslant T/\varepsilon}|X^\varepsilon_{x,y}(n)-\overline{X}^\varepsilon_x(n)|d\mu(x,y)\leqslant C(T,\gamma)(\delta+\varepsilon^{1-\gamma}+d_{T,K}(n(\varepsilon),\delta)). \tag{3.11}$$

Letting, first, $\varepsilon\to0$ and then $\delta\to0$ we obtain (2.5).

4. Proof of Corollaries

4.1. We deal first with Corollary 2.2. Denote by Jac_yf_x the Jacobian of the linear map $D_yf_x:T_yM\to T_{f_xy}M$ with respect to the Riemannian norms. Since f_x is an endomorphism Jac_yf_x is bounded away from zero uniformly in $y\in M$ and in x belonging to a compact set. The density q_x of the f_x-invariant measure μ_x satisfies

$$q_x(y)=\sum_{v\in f_x^{-1}y}\frac{q_x(v)}{|\mathrm{Jac}_vf_x|}. \tag{4.1}$$

By perturbation arguments, for any compact set $V \subset \mathbb{R}^d$ there exists $\varepsilon(V) > 0$ such that if $\varepsilon \leqslant \varepsilon(V)$ then the differential $D_{z,v}F_\varepsilon : \mathbb{R}^d \times M \to \mathbb{R}^d \times M$ is nondegenerate on $V \times M$ and, moreover, its Jacobian $\mathrm{Jac}_{z,v}F_\varepsilon$ is uniformly bounded away from zero there. Then for any bounded Borel function g on $\mathbb{R}^d \times M$,

$$\begin{aligned}\int_{V\times M} g\circ F_\varepsilon(x,y)d\mu_x(y)d\ell(x) &= \int_{V\times M} g\circ F_\varepsilon(x,y)q_x(y)\times d\rho(y)d\ell(x)\\ &= \int_{F_\varepsilon(V\times M)} g(z,v)\sum_{(x,y)\in F_\varepsilon^{-1}(z,v)} \frac{q_x(y)}{|\mathrm{Jac}_{x,y}F_\varepsilon|}d\rho(v)d\ell(z).\end{aligned} \tag{4.2}$$

It follows from (2.1), (2.2), and from the implicit function theorem arguments that there exists $C_1 > 0$ depending only on V such that if $x \in V$, $F_\varepsilon(x,y) = (z,v)$, and $f_z^{-1}v = (w_1,\dots,w_k)$ then $F_\varepsilon^{-1}(z,v) = ((x_1,y_1),\dots,(x_k,y_k))$ with $|x_i - z| \leqslant C_1\varepsilon$ and $|w_i - y_i| \leqslant C_1\varepsilon$, $i = 1,\dots,k$. Thus by (2.1), (2.2), (2.6), and (4.1),

$$\sum_{(x,y)\in F_\varepsilon^{-1}(z,v)} \frac{q_x(y)}{|\mathrm{Jac}_{x,y}F_\varepsilon|} \leqslant \sum_{y\in f_z^{-1}v} \frac{q_z(y)}{|\mathrm{Jac}_y f_z|} + C_2\varepsilon = q_z(v) + C_2\varepsilon \tag{4.3}$$

for some $C_2 > 0$ depending only on V. Suppose that $g \geqslant 0$ then substituting (4.3) into (4.2) and taking into account (2.6) we obtain

$$\begin{aligned}\int_{V\times M} g\circ F_\varepsilon(x,y)d\mu_x(y)d\ell(x) &\leqslant \int_{F_\varepsilon(V\times M)} g(z,v)(q_z(v)+C_2\varepsilon)d\rho(v)d\ell(z)\\ &\leqslant (1+C_2L\varepsilon)\int_{F_\varepsilon(V\times M)} g(z,v)d\mu_x(v)d\ell(z).\end{aligned} \tag{4.4}$$

If $(x,y) \in K \times M$ then by (2.1) we see that $F_\varepsilon^n(x,y) \in K_{LT} \times M$ for all $n \in [0, T/\varepsilon]$ where, recall, K_r is the closed r-neighborhood of K. It follows that there exists $C_{T,K} > 0$ such that

$$\int_{K\times M} g\circ F_\varepsilon^n(x,y)d\mu_x(y)d\ell(x) \leqslant C_{T,K}\int_{K_{LT}\times M} g(z,v)d\mu_x(v)d\ell(z) \tag{4.5}$$

for all $n \in [0, T/\varepsilon]$ and taking $g = \mathbb{I}_{E(k,\delta)}$ we derive (2.7).

4.2. Next, we consider the setup of Corollary 2.3. First, observe that (2.3) follows from § 14 in [**2**] (see also [**17**]). If x belongs to a compact set K then by (2.1), $X_{x,y}^\varepsilon(n) \in \mathcal{X} = K_{LT}$ for all $n \leqslant T/\varepsilon$ so we will have to consider x-coordinates in $\mathcal{X}$ only. Any vector $\xi \in T(\mathbb{R}^d \times M) = \mathbb{R}^d \oplus TM$ can be uniquely written as $\xi = \xi^{\mathcal{X}} + \xi^W$ where $\xi^{\mathcal{X}} \in T\mathbb{R}^d$ and $\xi^W \in TM$ and it has the Riemannian norm $|||\xi||| = |\xi^{\mathcal{X}}| + \|\xi^W\|$ where $|\cdot|$ is the usual Euclidean norm on $\mathbb{R}^d$ and $\|\cdot\|$ is the Riemannian norm on M. The corresponding metrics on M and on $\mathbb{R}^d \times M$ will be denoted by d_M and d, respectively, so that if $z_1 = (x_1,w_1)$, $z_2 = (x_2,w_2) \in \mathbb{R}^d \times M$ then $d(z_1,z_2) = |x_1 - x_2| + d_M(w_1,w_2)$. It is known (see [**10**] and [**16**]) that the hyperbolic splitting $T_{\Lambda_x}M = \Gamma_x^s \oplus \Gamma_x^u$ over Λ_x can be continuously extended to the splitting $T_W M = \Gamma^s_x \oplus \Gamma^u_x$ over W which is forward invariant with respect to Df_x and uniformly in $x \in \mathcal{X}$ satisfies exponential estimates

$$\|Df^n_x \xi\| \leqslant e^{-\kappa n}\|\xi\| \quad \text{and} \quad \|Df^{-n}_x \eta\| \leqslant e^{-\kappa n}\|\eta\| \tag{4.6}$$

for all $\xi \in \Gamma^s_x$, $\eta \in \Gamma^u_x$, and $n \geqslant n_0$. Moreover, by [**6**] (see also [**17**]) we can choose these extensions so that Γ^s_x and Γ^u_x will be C^1 in x in the corresponding Grassmann manifold. Each vector $\xi \in T_{x,w}(\mathcal{X} \times W) = T_x\mathcal{X} \oplus T_w W$ can be represented uniquely in the form $\xi = \xi^{\mathcal{X}} + \xi^u + \xi^s$ with $\xi^{\mathcal{X}} \in T_x\mathcal{X}$, $\xi^u \in \Gamma^u_{x,w}$ and $\xi^s \in \Gamma^s_{x,w}$. For any small $\varepsilon, \beta > 0$ set $\mathcal{C}^u(\varepsilon,\beta) = \{\xi \in T(\mathcal{X}\times W) : \|\xi^s\| \leqslant \varepsilon\beta^{-2}\|\xi^u\| \text{ and } |\xi^{\mathcal{X}}| \leqslant \varepsilon\beta^{-1}\|\xi^u\|\}$ and $\mathcal{C}^u_{x,w}(\varepsilon,\beta) = \mathcal{C}^u(\varepsilon,\beta) \cap T_{x,w}(\mathcal{X} \times W)$ which are cones around Γ^u and $\Gamma^u_{x,w}$, respectively. Similarly, we define $\mathcal{C}^s(\varepsilon,\beta) = \{\xi \in T(\mathcal{X} \times W) : \|\xi^u\| \leqslant \varepsilon\beta^{-2}\|\xi^s\| \text{ and } |\xi^{\mathcal{X}}| \leqslant \varepsilon\beta^{-1}\|\xi^s\|\}$ and $\mathcal{C}^s_{x,w}(\varepsilon,\beta) = \mathcal{C}^s(\varepsilon,\beta) \cap T_{x,w}(\mathcal{X} \times W)$. We claim that there exist $n_1, \beta_0, \varepsilon(\beta) > 0$ such that if $F^k_\varepsilon z \in \mathcal{X} \times W\, \forall\, k = 0, 1, \dots, n$, $n \geqslant n_1$, $\beta \leqslant \beta_0$, $\varepsilon \leqslant \varepsilon(\beta)$ then

$$D_z F^n_\varepsilon \mathcal{C}^u_z(\varepsilon,\beta) \subset \mathcal{C}^u_{F^n_\varepsilon z}(\varepsilon,\beta), \quad \mathcal{C}^s_z(\varepsilon,\beta) \supset D_z F^{-n}_\varepsilon \mathcal{C}^s_{F^n_\varepsilon z}(\varepsilon,\beta) \tag{4.7}$$

and for any $\xi \in \mathcal{C}^u_z(\varepsilon,\beta)$, $\eta \in \mathcal{C}^s_{F^n_\varepsilon z}(\varepsilon,\beta)$,

$$|||D_z F^n_\varepsilon \xi||| \geqslant e^{\frac{1}{2}\kappa n}|||\xi|||, \quad |||D_z F^{-n}_\varepsilon \eta||| \geqslant e^{\frac{1}{2}\kappa n}|||\eta|||. \tag{4.8}$$

Before deriving (4.7) and (4.8) we explain how to use them in order to obtain (2.4). For any linear subspace Ξ of $T_z(\mathcal{X} \times W)$ denote by $J^\Xi_\varepsilon(z)$ absolute value of the Jacobian of the linear map $D_z F_\varepsilon : \Xi \to D_z F_\varepsilon \Xi$ with respect to inner products induced by the Riemannian metric and set $J^\Xi_\varepsilon(n,z) = \prod_{k=0}^{n-1} J_\varepsilon^{D_z F^k_\varepsilon \Xi}(F^k_\varepsilon z)$. Denote also by $J^u_x(y)$ absolute value of the Jacobian of the linear map $D_y f_x : \Gamma^u_{x,y} \to \Gamma^u_{x,f_x y}$ and set $J^u_x(\varepsilon,n,y) = \prod_{k=0}^{n-1} J^u_{X^\varepsilon_{x,y}(k)}(Y^\varepsilon_{x,y}(k))$. Let n^u and n^s be the dimensions of $\Gamma^u_{x,y}$ and $\Gamma^s_{x,y}$, respectively, which do not depend on x, y by continuity considerations. If Ξ is an n^u-dimensional subspace of $T_z(\mathcal{X} \times W)$, $z = (x,y)$, and $\Xi \subset \mathcal{C}^u_{x,y}(\varepsilon,\beta)$ then it follows easily from (2.1), (2.2), and (4.7) that there exists a constant $C_1 > 0$ independent of $x \in \mathcal{X}$ and $y \in M$, such that for any small $\varepsilon > 0$ and $n \in \mathbb{N}$,

$$(1 - C_1\varepsilon)^n \leqslant J^\Xi_\varepsilon(n,z)(J^u_x(\varepsilon,n,y))^{-1} \leqslant (1 + C_1\varepsilon)^n. \tag{4.9}$$

For each $y \in \Lambda_x$ and $\gamma > 0$ small enough set

$$W^s_x(y,\gamma) = \{v \in W : d_M(f^k_x y, f^k_x v) \leqslant \gamma\ \forall\, k \geqslant 0\}$$

and

$$W^u_x(y,\gamma) = \{v \in W : d_M(f^k_x y, f^k_x v) \leqslant \gamma\ \forall\, k \leqslant 0\}$$

which are local stable and unstable manifolds for f_x at y. According to [**10**] these families can be included into continuous families of n^s and n^u-dimensional stable and unstable discs $W^s_x(y,\gamma)$ and $W^u_x(y,\gamma)$, respectively, defined for all $y \in W$ and such that $W^s_x(y,\gamma)$ is tangent to Γ^s_x, $W^u_x(y,\gamma)$ is tangent to Γ^u_x, $f_x W^s_x(y,\gamma) \subset W^s_x(f_x y,\gamma)$ and $W^u_x(y,\gamma) \supset f^{-1}_x W^u_x(f_x y,\gamma)$. For any $z, \widetilde{z} \in \mathbb{R}^d \times M$ set

$$d^\varepsilon_n(z,\widetilde{z}) = \max\{d(F^k_\varepsilon z, F^k_\varepsilon \widetilde{z}) : 0 \leqslant k < n\}$$

and

$$B_x^\varepsilon(y,\gamma,n)=\{v\in W:\, d_n^\varepsilon((x,y),(x,v))\leqslant\gamma\}.$$

Let $v\in W_x^u(y,\gamma)$ and assume that $B_x^\varepsilon(v,\gamma,n)$ does not intersect the boundary $\partial W_x^u(y,\gamma)$ of $W_x^u(y,\gamma)$. Set

$$V_x^u(v)=V_x^u(v,\gamma)=W_x^u(y,\gamma)\cap B_x^\varepsilon(v,\gamma,n)$$

and

$$V_x^u(v,k)=V_x^u(v,\gamma,k)=F_\varepsilon^k(\{x\}\times V_x^u(v)),\quad k=1,2,\dots,n.$$

By (4.7), $TV_x^u(v,k)\subset\mathcal{C}^u(\varepsilon,\beta)$, $k=1,2,\dots,n$ (where TV is the tangent bundle of V) and we conclude by (4.8), (4.9) and the volume lemma type arguments (see Appendix in [**4**]) that

$$C^{-1}\leqslant\rho^u(V_x^u(v))J_x^u(\varepsilon,n,v)\leqslant C \tag{4.10}$$

where ρ^u is the induced Riemannian volume on $W_x^u(y,\gamma)$ and we denote by C here and below different positive constants depending on γ but not on $x\in\mathcal{X}$, y,ε, and $n\leqslant T/\varepsilon$. By the volume lemma arguments we derive also that

$$\rho^u(V_x^u(v)\cap F_\varepsilon^{-n}E(k,\delta))\leqslant C\big(J_x^u(\varepsilon,n,v)\big)^{-1}\rho_V\big(V_x^u(v,n)\cap E(k,\delta)\big) \tag{4.11}$$

where ρ_V is the induced n^u-dimensional Riemannian volume on $V_x^u(v,n)$. By (4.7) and (4.8) it is easy to see that if $w\in V_x^u(v)$ then

$$d(F_\varepsilon^k(x,v),F_\varepsilon^k(x,w))\leqslant C\exp(-C^{-1}(n-k)) \tag{4.12}$$

for all $k=0,1,\dots,n$ which together with (1.2) yield that

$$|X_{x,v}^\varepsilon(n)-X_{x,w}^\varepsilon(n)|\leqslant C\varepsilon, \tag{4.13}$$

i.e. the x-coordinate of points in $V_x^u(v,n)$ may differ at most by $2C\varepsilon$. It follows by (2.1) that if $(z,u)\in V_x^u(v,n)\cap E(k,\delta)$ then $(X_{x,v}^\varepsilon(n),u)\in E(k,\delta/2)$ provided $k\leqslant(\log 1/\varepsilon)^{1-\alpha}$, $\alpha>0$ and ε is small enough. Then, we will have also that for such k,

$$\{X_{x,v}^\varepsilon(n)\}\times W_{X_{x,v}^\varepsilon(n)}^s(u,\gamma)\subset E(k,\delta/3)$$

provided γ is small enough. This together with (4.7) yield that

$$\begin{aligned}\rho_V\big(V_x^u(v,n)\cap E(k,\delta)\big)&\leqslant C\rho^u\big(W_{X_{x,v}^\varepsilon(n)}^u(Y_{x,v}^\varepsilon(n),\gamma))\cap E_{X_{x,v}^\varepsilon(n)}(k,\delta/2)\big)\\&\leqslant C^2\rho(E_{X_{x,v}^\varepsilon(n)}(k,\delta/3))\end{aligned} \tag{4.14}$$

where $E_z(k,r)=\{v:\,(z,v)\in E(k,r)\}$ and $k\leqslant(\log 1/\varepsilon)^{1-\alpha}$. In the right hand side of (4.14) we can write also $\mu_{X_{x,v}^\varepsilon(n)}$ in place of ρ. It follows from the upper moderate deviations bound in [**12**] that

$$r_k=\sup_{z\in\mathcal{X}}\rho(E_z(k,k^{a-1}))\leqslant\exp(-c_ak^{2a-1}) \tag{4.15}$$

for each $a\in(\frac12,1)$ and some $c_a>0$. Observe that (4.15) remains true with μ_z in place of ρ.

Now, choose a maximal set of points $v_i \in W^u_x(y,\gamma)$ such that $V^u_x(v_i,\gamma)$ does not intersect the boundary of $W^u_x(y,2\gamma)$ and $d^\varepsilon_n(v_i,v_j) \geqslant \gamma$ if $i \neq j$. Then $\cup_i V^u_x(v_i,\gamma) \supset W^u_x(y,\gamma)$ and $V^u_x(v_i,\gamma/2)$ are disjoint for different i's. Applying the volume lemma style arguments as above to $V^u_x(v_i,\gamma/2)$ we conclude from (4.10) that

$$\sum_i \left(J^u_x(\varepsilon,n,v_i)\right)^{-1} \leqslant C\rho^u(W^u_x(y,\gamma)). \tag{4.16}$$

It follows from (4.11) and (4.14)–(4.16) that

$$\rho^u(W^u_x(y,\gamma)\cap F^{-n}_\varepsilon E(k,\delta)) \leqslant \sum_i \rho^u(V^u_x(v_i,\gamma)\cap F^{-n}_\varepsilon E(k,\delta)) \leqslant Cr_k \tag{4.17}$$

provided $\delta \geqslant k^{a-1}$ and $k \leqslant (\log\frac{1}{\varepsilon})^{1-\alpha}$. Let $B((x,y),c\gamma)$ be a ball in $\mathbb{R}^d \times M$ of radius $c\gamma$ centered at (x,y) for some small constant c. The family $W^u_z(v,\gamma)\cap B((x,y),c\gamma)$, $(z,v)\in B((x,y),c\gamma)$ forms a measurable partition of $B((x,y),c\gamma)$ (even a foliation) and conditional measures of $\ell\times\rho$ relative to this partition are equivalent to the corresponding measures ρ^u. Hence, (4.17) implies

$$\ell\times\rho(B((x,y),c\gamma)\cap F^{-n}_\varepsilon E(k,\delta)) \leqslant Cr_k \tag{4.18}$$

provided γ is small enough. Choose a minimal finite cover of $\mathcal{X}\times M$ by balls of radius $c\gamma$ centered at some points (x_j,y_j), making the above construction for each point (x_j,y_j) and applying (4.18) we arrive at (2.4) and (2.10) follows from (3.1), (3.10), (4.15) and (4.18). Since essential estimates concern only volumes on unstable and close to unstable manifolds where the SRB measures are equivalent to the volume there we see that (4.19) remains true if we replace $\ell\times\rho$ by μ such that $d\mu(x,y) = d\mu_x(y)d\ell(x)$, and so (2.4) and (2.10) remain true with such μ, as well.

Finally, we prove (4.7) and (4.8). Let $\xi = \xi^{\mathcal{X}} + \xi^u + \xi^s \in T_z(\mathcal{X}\times M)$, $D_zF^n\xi^{\mathcal{X}} = \zeta = \zeta^{\mathcal{X}} + \zeta^u + \zeta^s \in T_{F^n z}(\mathcal{X}\times M)$, $z=(x,w)$, $D_w f^n_x\xi^u = \eta^u$, and $D_w f^n_x \xi^s = \eta^s$. Then $D_z f^n_x\xi = \zeta^{\mathcal{X}} + (\zeta^u+\eta^u) + (\zeta^s+\eta^s)$ and $|\xi^{\mathcal{X}}| = |\zeta^{\mathcal{X}}|$, $\|\zeta^u\| \leqslant Ce^{Cn}|\xi^{\mathcal{X}}|$, $\|\zeta^s\| \leqslant Ce^{Cn}|\xi^{\mathcal{X}}|$, $\|\eta^s\| \leqslant C\|\xi^s\|$ for some $C>0$ independent of ξ and $\|\eta^u\| \geqslant e^{\kappa n}\|\xi^u\|$ if $n \geqslant n_0$. Hence, for $n\geqslant n_0$,

$$\|\xi^u+\eta^u\| \geqslant \|\eta^u\| - \|\zeta^u\| \geqslant e^{\kappa n}\|\xi^u\| - Ce^{Cn}|\xi^{\mathcal{X}}| \tag{4.19}$$

and

$$\|\zeta^s+\eta^s\| \leqslant \|\zeta^s\| + \|\eta^s\| \leqslant Ce^{Cn}|\xi^{\mathcal{X}}| + C\|\xi^s\|. \tag{4.20}$$

Suppose that $\|\xi^u\| \geqslant \beta\varepsilon^{-1}|\xi^{\mathcal{X}}|$ and $\|\xi^u\| \geqslant \beta^2\varepsilon^{-1}\|\xi^s\|$. Then by (4.19),

$$\|\zeta^u+\eta^u\| \geqslant (\tfrac12 e^{\kappa n}\beta\varepsilon^{-1} - Ce^{Cn})|\xi^{\mathcal{X}}| + \tfrac12 e^{\kappa n}\beta^2\varepsilon^{-1}\|\xi^s\|. \tag{4.21}$$

Put $n_1 = [\kappa^{-1}\ln(8C+5)]+1$, choose $\beta_0>0$ so that $e^{\kappa n} \geqslant 4\beta Ce^{Cn}$ for any $\beta\leqslant\beta_0$ and all $n\in[n_1,2n_1]$, and set $\varepsilon(\beta) = \frac{\beta}{4}\min(\beta, C^{-1}e^{-2Cn_1})$. Then by (4.21) we see that $D_zF^n_0\xi \in \mathcal{C}^u_{F^n_0 z}(\varepsilon,2\beta)$ and since we have to check (4.7) only for $n\in[n_1,2n_1]$ we obtain easily from (2.1) and (2.2) that the perturbation F_ε of F_0 satisfies the first part of (4.7) provided β_0 and $\varepsilon(\beta)$ are chosen sufficiently small. The second part of (4.7) follows in the same way.

Next, for $n \geqslant n_0$,

$$\begin{aligned}|||D_zF_0^n\xi||| &\geqslant \|\eta^u\| - |\zeta^{\mathcal{X}}| - \|\zeta^u\| - \|\zeta^s\| - \|\eta^s\| \\ &\geqslant e^{\kappa n}\|\xi^u\| - (1+2Ce^{Cn})|\xi^{\mathcal{X}}| - C\|\xi^s\| \\ &\geqslant (e^{\kappa n} - \beta^{-1}\varepsilon(1+2Ce^{Cn}) - \beta^{-2}\varepsilon C)\|\xi^u\| \\ &\geqslant (e^{\kappa n} - \beta^{-1}\varepsilon(1+2Ce^{Cn}) - \beta^{-2}\varepsilon C)(1+\varepsilon\beta^{-1}+\varepsilon\beta^{-2})^{-1}|||\xi|||.\end{aligned}\tag{4.22}$$

Choose $\varepsilon(\beta)$ so small (for instance, $\varepsilon(\beta) = \beta^3$) that for all $\varepsilon \leqslant \varepsilon(\beta)$ and $\beta \leqslant \beta_0$,

$$e^{\kappa n} - \varepsilon\beta^{-1}(1+2Ce^{Cn}) - \varepsilon\beta^{-2}C \geqslant (1+\varepsilon\beta^{-1}+\varepsilon\beta^{-2})e^{\frac{2}{3}\kappa n}$$

for any $n \in [n_1, 2n_1]$. Then, $|||D_zF_0^n\xi||| \geqslant e^{\frac{2}{3}\kappa n}|||\xi|||$ for all such n, and so if ε small enough we have also $|||D_zF_\varepsilon^n\xi||| \geqslant e^{\frac{1}{2}\kappa n}|||\xi|||$. Using (4.7) and repeating this argument for $D_zF_\varepsilon^{in_1}\xi$, $i = 1, 2, \ldots.$ in place of ξ we derive the first assertion in (4.8) for all $n \geqslant n_1$ and the second one follows in the same way.

5. Toral translations and skew translations

First, we write the Fourier series for Φ,

$$\Phi(x,y) = \sum_{k\in\mathbb{Z}^m} \Phi_k(x)\exp(2\pi i(k,y))\tag{5.1}$$

where $(k,y) = \sum_{j=1}^m k_jy_j$, the vector coefficients Φ_k are given by

$$\Phi_k(x) = \int \Phi(x,y)\exp(-2\pi i(k,y))d\rho(y)\tag{5.2}$$

where ρ is the Lebesgue measure on $\mathbb{T}^m$. Set

$$r_N(x,y) = \sum_{k:\,|k|>N} \Phi_k(x)\exp(2\pi i(k,y))$$

then by the well known estimates of tails of Fourier expansions of analytic functions

$$r_N(x,y) \leqslant c_\Phi^{-1}e^{-c_\Phi N}\tag{5.3}$$

for some $c_\Phi > 0$ which can be written explicitly via the supremum norm of the analytic extension of Φ (see Appendix 1 in [**13**]). By (5.2), $\overline{\Phi}(x) = \Phi_0(x)$, and so by (2.1) and (5.1)–(5.3),

$$\Big|\frac{1}{n}\sum_{l=0}^{n-1}\Phi(x, f_x^ky) - \overline{\Phi}(x)\Big| \leqslant c_\Phi^{-1}e^{-c_\Phi N} + \frac{L}{n}\sum_{k:\,|k|\leqslant N}\Big|\sum_{l=0}^{n-1}\exp(2\pi i(k, f_x^ly))\Big|.\tag{5.4}$$

Next, we will deal separately with translations and skew translations.

5.1. Consider, first, toral translations f_x defined by (2.11). Then we have to estimate

$$\Big|\sum_{l=0}^{n-1}\exp(2\pi i(k,y+l\omega(x)))\Big| \leqslant 2|\exp(2\pi i(k,\omega(x)))-1|^{-1}. \tag{5.5}$$

Let $V\subset\mathbb{R}^d$ be a compact set and

$U_V(\eta,N)=\{x\in V:\,|\exp(2\pi i(k,\omega(x)))-1|\leqslant\eta|k|^{-m}\text{for some } k \text{ with } 0<|k|<N\}.$

It follows by Diophantine approximations type arguments (cf. Appendix 4 in [**13**]) that for some constant $C>0$,

$$\ell(U_V(\eta,N))\leqslant C\eta(\inf_{x\in V}\lambda_\omega(x))^{-1}\ell(V) \tag{5.6}$$

with λ_ω defined before the statement of Theorem 2.4. Set $N=N(\varepsilon)=[-c_1\log\varepsilon]$, $n=n(\varepsilon)=\varepsilon^{-c_2}$, $\delta=\delta(\varepsilon)=\varepsilon^{c_3}$, and $\eta=\eta(\varepsilon)=\varepsilon^{c_4}$ where $c_1,c_2,c_3,c_4>0$ will be picked up later. Then by (5.4) and (5.5) for any $x\in V\smallsetminus U_V(\eta,N)$,

$$\Big|\frac{1}{n}\sum_{l=0}^{n-1}\Phi(x,f_x^k y)-\overline{\Phi}(x)\Big|\leqslant C(\varepsilon^{c_\Phi c_1}+(\log\varepsilon)^{2m}\varepsilon^{c_2-c_4}) \tag{5.7}$$

for some $C>0$ independent of ε. Hence, if

$$\min(c_\Phi c_1,c_2-c_4)>c_3 \tag{5.8}$$

then

$$V\cap E(\delta(\varepsilon),n(\varepsilon))\subset U_V(\eta(\varepsilon),N(\varepsilon)) \tag{5.9}$$

provided ε is small enough.

Next, we improve the estimate (3.10) in our particular case. Since f_x is an isometry then $d_M(f_x(Y^\varepsilon_{x,y}(l-1)),f_x(f_x^{l-1}y))=R_{l-1}(x,y)$ and (3.10) becomes

$$R_l(x,y)\leqslant\varepsilon L(1+Ll)+R_{l-1}(x,y)\leqslant\varepsilon Ll(1+Ll). \tag{5.10}$$

It follows from (2.7), (3.1), (5.6), (5.7), and (5.10) that (2.13) holds true with $c=\min(c_3,1-2c_2,c_4)$ provided (5.8) is satisfied. Since c_1 can be chosen arbitrarily large we have to choose only $c_2,c_3,c_4>0$ so that $\min(c_3,1-2c_2,c_4)$ is maximal possible assuming that $c_2-c_4>c_3$. Set $c_5=\min(c_2-c_4,1-2c_2,c_4)$ then $c_5\geqslant c$. Since we can increase one term in the last minimum only by decreasing another term there, it is clear that c_5 will be maximized when all three terms are equal. Solving emerging then two equations we obtain $c_5=c_4=1/5$ and $c_2=2/5$. Taking c_3 arbitrary close but less than $1/5$ we conclude that (2.13) holds true with any $c<c_0=1/5$.

5.2. Next, we consider skew translations f_x defined by (2.12). Identifying $y\in\mathbb{T}^m$ with a vector of $\mathbb{R}^m$ having coordinates $y_i\in[0,1)$, $i=1,\dots,m$ we define recursively vectors $\widetilde{f}_x^k y$, $\widetilde{Y}^\varepsilon_{x,y}(k)\in\mathbb{R}^m$ by $\widetilde{f}_x^0 y=y$, $\widetilde{Y}^\varepsilon_{x,y}(0)=y$ and $\widetilde{f}_x^n y=(I+P)\widetilde{f}_x^{n-1}y+a(x)$, $\widetilde{Y}^\varepsilon_{x,y}(n)=(I+P)\widetilde{Y}^\varepsilon_{x,y}(n-1)+a(X^\varepsilon_{x,y}(n-1))$ where I is the $m\times m$ identity matrix, $P=(p_{jl})$ is the matrix whose elements for $l<j$ appear in the definition (2.12) and

for $l > j$ are equal to zero, and $a(x)$ is the m-vector whose first coordinate is $\alpha(x)$ and all other coordinates are zero. Then

$$\widetilde{f}_x^n y = (I+P)^n y + \sum_{l=1}^{n} (I+P)^{n-l} a(x) \tag{5.11}$$

and

$$\widetilde{Y}_{x,y}^{\varepsilon}(n) = (I+P)^n y + \sum_{l=1}^{n} (I+P)^{n-l} a(X_{x,y}^{\varepsilon}(l-1)). \tag{5.12}$$

The crucial fact in using the formulas (5.11) and (5.12) is that P is nilpotent, and so

$$P^j = 0 \quad \text{for all } j \geqslant m. \tag{5.13}$$

For any vector $v = (v_1, \ldots, v_m) \in \mathbb{R}^m$ denote by $\{v\}$ the vector in $\mathbb{T}^m$ whose coordinates are fractional parts $\{v_1\}, \ldots, \{v_m\}$ of $v_1, \ldots, v_m$, i.e. $\{\cdot\} : \mathbb{R}^m \to \mathbb{T}^m$ is the natural projection. It is clear that

$$f_x^n y = \{\widetilde{f}_x^n y\} \quad \text{and} \quad Y_{x,y}^{\varepsilon}(n) = \{\widetilde{Y}_{x,y}^{\varepsilon}(n)\}. \tag{5.14}$$

Expanding binomials in (5.11) the coordinates of $f_x^n y$ can be written more explicitly. Namely, (see [**7**], § 2 in Ch. 7) $f_x^n(y_1, \ldots, y_m) = (y_1^{(n)}(\alpha(x)), \ldots, y_m^{(n)}(\alpha(x)))$ where

$$\begin{aligned} & y_1^{(n)}(\beta) = y_1 + n\beta, \\ & y_l^{(n)} = y_l + \sum_{j=1}^{l-1} y_j \sum_{q=1}^{l-j} \binom{n}{q} p_{lj}^{(q)} + \beta \sum_{q=1}^{l-1} \binom{n}{q+1} p_{l1}^{(q)}, \quad 1 \leqslant l \leqslant m, \end{aligned} \tag{5.15}$$

where $p_{lj}^{(q)}$ are elements of the qth power of the matrix P.

Observe, that the projection $\{\cdot\}$ does not increase distances, and so by (3.7) and (5.11)–(5.14),

$$\begin{aligned} R_n(x,y) = d_M(Y_{x,y}^{\varepsilon}(n), f_x^n y) & \leqslant |\widetilde{Y}_{x,y}^{\varepsilon} - \widetilde{f}_x^n y| \\ & \leqslant CLn^m \sum_{l=1}^{n} |X_{x,y}^{\varepsilon}(l-1) - x| \leqslant \varepsilon CL^2 n^{m+2} \end{aligned} \tag{5.16}$$

for some $C > 0$ depending only on P and m.

Let $k = (k_1, \ldots, k_m) \in \mathbb{Z}^m$, $|k| \neq 0$ and set $l_0 = l_0(k) = \max\{l : k_l \neq 0\}$. Define the map $\Psi_k : \mathbb{R}^{l_0} \to \mathbb{R}^{l_0}$ acting by $\Psi_k(y_1, \ldots, y_{l_0-1}, \beta) = (\gamma_1, \ldots, \gamma_{l_0})$ so that

$$\sum_{l=1}^{l_0} k_l y_l^{(n)}(\beta) = G_\gamma(n) = (k, y) + \gamma_1 n + \gamma_2 n^2 + \cdots + \gamma_{l_0} n^{l_0}$$

for any $n \in \mathbb{N}$ where $y_l^{(n)}(\beta)$, $l = 1, \ldots, l_0$ are given by (5.15). Observe, that Ψ_k is one-to-one and the coordinates of $\Psi_k^{-1}(\gamma_1, \ldots, \gamma_{l_0})$ can be obtained recursively from the formulas

$$
\begin{aligned}
\beta k_{l_0} p^{(l_0-1)}_{l_0 1} &= l_0! \gamma_{l_0}, \quad y_1 k_{l_0} p^{(l_0-1)}_{l_0 1} = (l_0-1)!(\gamma_{l_0-1} + c_{11}\beta), \\
y_2 k_{l_0} p^{(l_0-2)}_{l_0 2} &= (l_0-2)!(\gamma_{l_0-2} + c_{21}\beta + c_{22}y_1), \dots, \\
y_{l_0-1} k_{l_0} p^{(1)}_{l_0, l_0-1} &= (\gamma_1 + c_{l_0-1,1}\beta + \textstyle\sum_{j=1}^{l_0-2} c_{l_0-1,j+1} y_j)
\end{aligned}
\tag{5.17}
$$

where the coefficients c_{jl}, $j \geqslant l$ satisfy

$$
|c_{jl}| \leqslant c(l_0)|k| \tag{5.18}
$$

for some $c(m) > 0$ depending only on m. Next, we will employ an elementary estimate of Weyl's sums which can be found on p.p. 215–216 in [**19**]. Namely, let

$$
W_\gamma(n) = \sum_{l=1}^{n} \exp(2\pi i G_\gamma(l))
$$

then it follows easily that for any compact $V \subset \mathbb{R}^{l_0}$,

$$
\int_V |W_\gamma(n)|^2 d\ell(\gamma) \leqslant n \ell(V_{\sqrt{l_0}})
$$

where, recall, V_r is an r-neighborhood of V. By Chebyshev's inequality we conclude that for any $n \in \mathbb{N}$ and $\widetilde{c} \in (0, 1/2)$ there exists a Borel set $U = U_{V,n,\widetilde{c}} \subset V$ such that

$$
\ell(U) \leqslant \ell(V_{\sqrt{l_0}}) n^{2\widetilde{c}-1} \tag{5.19}
$$

and

$$
|W_\gamma(n)| \leqslant n^{1-\widetilde{c}} \quad \text{provided } \gamma \notin U. \tag{5.20}
$$

It follows from (5.17)–(5.20) and the nondegeneracy condition on $\alpha(x)$ that for any $T > 0$ and a compact $K \subset \mathbb{R}^d$ there exists a constant $C(K, T) > 0$ such that

$$
\Big| \sum_{l=0}^{n-1} \exp(2\pi i (k, f_x^l y)) \Big| \leqslant n^{1-\widetilde{c}} \tag{5.21}
$$

provided $x \in K_{LT}$, $y \in \mathbb{T}^m$ and $(x, y) \notin U_k \subset \mathbb{R}^d \times \mathbb{T}^m$ with

$$
\ell \times \rho(U_k) \leqslant C(K, T)|k|^{l_0} n^{2\widetilde{c}-1} \tag{5.22}
$$

where U_k depends on n, k and $\widetilde{c}$. Set now $N = N(\varepsilon) = [-c_1 \log \varepsilon]$, $n = n(\varepsilon) = \varepsilon^{-c_2}$, and $\delta = \delta(\varepsilon) = \varepsilon_3^c$. Then by (5.4) for any $(x, y) \notin U = \cup_{k:|k| \leqslant N} U_k$,

$$
\Big| \frac{1}{n} \sum_{l=0}^{n-1} \Phi(x, f_x^k y) - \overline{\Phi}(x) \Big| \leqslant C(\varepsilon^{c_\Phi c_1} + (\log \varepsilon)^{3m} \varepsilon^{\widetilde{c} c_2}) \tag{5.23}
$$

for some $C > 0$ independent of ε. Hence, if

$$
\min(c_\Phi c_1, \widetilde{c} c_2) > c_3 \tag{5.24}
$$

then

$$
K_{LT} \cap E(\delta(\varepsilon), n(\varepsilon)) \subset U. \tag{5.25}
$$

It follows from (2.7), (3.1), (5.16), (5.22), and (5.25) that (2.13) holds true with

$$c = \min(c_3, 1-(m+2)c_2, c_2(1-2\widetilde{c})) \tag{5.26}$$

provided $\widetilde{c} < 1/2$ and (5.24) is satisfied. Since c_1 can be chosen arbitrarily large we have to choose only $\widetilde{c}$, c_2, $c_3 > 0$ so that the right hand side of (5.26) is maximal possible assuming that $\widetilde{c}c_2 > c_3$. Set $c_4 = \min(\widetilde{c}c_2, 1-(m+2)c_2, c_2(1-2\widetilde{c}))$, then $c_4 \geqslant c$. Again, c_4 will be maximized when all three term in the minimum there will be equal which gives $\widetilde{c} = 1/3$, $c_2 = 3/(3m+7)$, and $c_4 = 1/(3m+7)$. It follows that (2.13) holds true with any $c < c_0 = 1/(3m+7)$.

References

[1] D.B. Anosov, *Averaging in systems of ordinary differential equations with fast oscillating solutions,* Izv. Acad. Nauk SSSR Ser. Mat., **24** (1960), 731–742 (in Russian).

[2] V.I. Bakhtin, *A direct method of constructing an invariant measure on a hyperbolic attractor,* Russian Acad. Sci. Izv. Math., **41** (1993), 207–227.

[3] V.I. Bakhtin, *On the averaging method in a system with fast hyperbolic motions,* Proc. Math. Inst. of Belarus Acad. Sci., **6** (2000), 23–26 (in Russian).

[4] R. Bowen and D. Ruelle, *The ergodic theory of Axiom A flows,* Invent. Math., **29** (1975), 181–202.

[5] K. Burns, C. Pugh, and A. Wilkinson, *Stable ergodicity and Anosov flows,* Topology, **39** (2000), 149–159.

[6] G. Contreras, *Regularity of topological and metric entropy of hyperbolic flows,* Math. Z., **210** (1992), 97–111.

[7] I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai, *Ergodic Theory,* (1982), Springer-Verlag, Berlin.

[8] H.S. Dumas and F. Golse, *The averaging method for perturbations of mixing flows,* Ergod. Th. & Dynam. Sys. **17** (1997) 1339–1358.

[9] S.N. Elaydi, *An Introduction to Difference Equations,* (1996), Springer–Verlag, New York.

[10] M. Hirsch, J. Palis, G. Pugh and M. Shub, *Neighborhoods of hyperbolic sets,* Invent. Math., **9** (1970), 121–134.

[11] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems,* (1995), Cambridge Univ. Press, Cambridge.

[12] Yu. Kifer, *Limit theorems in averaging for dynamical systems,* Ergod. Th.& Dynam. Sys., **15** (1995), 1143–1172.

[13] P. Lochak and C. Meunier, *Multiple Averaging for Classical Systems,* (1988), Springer-Verlag, New York.

[14] R. Mañé, *Ergodic Theory and Differentiable Dynamics,* (1987), Springer-Verlag, Berlin.

[15] A.I. Neishtadt, *Averaging in multifrequency systems. II,* Sov. Phys. Dokl., **21**, No. 2 (1976), 80–82.

[16] C. Robinson, *Structural stability of C^1 diffeomorphisms,* J. Diff. Equat., **22** (1976), 28–73.

[17] D. Ruelle, *Differentiation of SRB states,* Comm. Math. Phys., **187** (1997), 227–241.

[18] J.A. Sanders and F.Verhurst, *Averaging Methods in Nonlinear Dynamical Systems,* (1985), Springer-Verlag, Berlin.

[19] I.M. Vinogradov, *Selected Works,* (1985) Springer–Verlag, Berlin.

Y. Kifer, Institute of Mathematics, Hebrew University of Jerusalem, Jerusalem 91904, Israel
E-mail : kifer@math.huji.ac.il

Astérisque
287, 2003, p. 125–134

ON BASIC PIECES OF AXIOM A DIFFEOMORPHISMS ISOTOPIC TO PSEUDOANOSOV MAPS

by

Jorge Lewowicz & Raúl Ures

Abstract. — We consider Axiom A diffeomorphisms g in the isotopy class of a pseudoanosov map f. It is shown that they have a unique "large" basic piece Λ, and necessary and sufficient conditions for g to be semiconjugated to f, that only involve conditions on Λ, are obtained. As a consequence, it is proved that if Λ is exteriorly situated, stable and unstable half-leaves of points of Λ boundedly deviate from geodesics.

1. Introduction

In this paper we consider Axiom A diffeomorphisms in the isotopy class of pseudoanosov maps. It is known (see [**H1, L1**]) that any pseudoanosov map f is persistent in its isotopy class i.e. for any homeomorphism g isotopic to f there exists a closed invariant set J_g such that g restricted to J_g is semiconjugated to f. When g verifies Axiom A, the first author (in [**L2**]) introduced the definition of "small" and "large" basic pieces of g. Roughly speaking, a basic piece B is small if an adequate lift to the universal cover of a B-stable (unstable) manifold is bounded. Otherwise, B is large. In [**L2**] it is proved that it is necessary and sufficient for a large basic piece Λ to be contained in J_g that the above mentioned lift lies at a bounded distance of an f-stable set of a point. In this work we prove that an Axiom A diffeomorphism g has a unique large basic piece Λ and, using that if g is not semiconjugated to f the Nielsen classes of g^n grow exponentially faster than the Nielsen classes of f^n (see [**H2**]), we prove that it is necessary and sufficient for g being semiconjugated to f that $\Lambda \subset J_g$. These results are proved in sections 3 and 4.

2000 *Mathematics Subject Classification.* — 37E30, 37C15, 37C50, 37C75, 37D20.
Key words and phrases. — Basic pieces, Axiom A, Pseudoanosov, Nielsen class, Isotopy class.

The second author was supported in part by a CONICYT-Clemente Estable grant.

In section 5 we lead with the case when the large basic set Λ is exteriorly situated i.e. there are no nul-homotopic loops that consist in the union of a stable and an unstable arc of a point of Λ. We show that, in this case, g is semiconjugated to f. As a consequence we obtain that half-leaves of stable and unstable manifolds of Λ boundedly deviates from geodesics that have the same asymptotic direction. Let us say that V. Grines (see [**G1**] and the survey [**G2**]) obtained this bounded deviation property for unstable (stable) manifolds of exteriorly situated nontrivial attractors (resp. repellers) of any diffeomorphism of a closed surface of genus larger than one and also for both, unstable and stable manifolds, provided the attractor (repeller) is a basic piece of a diffeomorphism that verifies Axiom A and that satisfies the strong transversality condition. In [**RW**] R. C. Robinson and R. Williams constructed an example of an Axiom A diffeomorphism on a surface of genus > 1 such that its nonwandering set consists exactly of an exteriorly situated attractor and an exteriorly situated repeller and that does not verify the strong transversality condition. V. Grines ([**G3**]) also showed that the stable manifold of the attractor (unstable for the repeller) of this example does not have the bounded deviation from geodesics property. Our results show that this kind of behaviour is impossible in the isotopy class of pseudoanosov maps (the fact that the Robinson-Williams example is not isotopic to a pseudoanosov map can be checked directly)

Finally, let us say that to understand the proofs, some familiarity with the theory of pseudoanosov maps is needed. We do not include this background material that the reader may find in [**FLP, CB, M, HT, T**].

2. Preliminaries

Let f be a pseudoanosov map of a compact connected oriented boundaryless surface M, let F be a lift of f to the universal cover $\mathcal{M}$ of M and $\pi : \mathcal{M} \to M$ the covering projection. Then, there exist (see [**T, FLP**]) equivariant pseudometrics D_S, D_U, and $\lambda > 1$ such that, for $\xi, \eta \in \mathcal{M}$,

$$D_S(F^{-1}(\xi), F^{-1}(\eta)) = \lambda D_S(\xi, \eta)$$
$$D_U(F(\xi), F(\eta)) = \lambda D_U(\xi, \eta)$$

and $D = D_S + D_U$ is an equivariant metric on $\mathcal{M}$.

For the remainder of this paper g will be a homeomorphism isotopic to f.

Definition 2.1. — *We say that $x, y \in M$ are equivalent iff for some (and then for any) lift G of g there exist $\xi \in \pi^{-1}(x)$, $\eta \in \pi^{-1}(y)$ and $K > 0$ such that $D(G^n(\xi), G^n(\eta)) \leqslant K$ for all $n \in \mathbb{Z}$.*

Obviously, this is an equivalence relation. In the following proposition we state some properties of this relation. For completeness we include the proofs that are essentially contained in [**H1, L1, CS**].

Proposition 2.2. — *The following statements hold:*

(1) *The constant K of the definition above only depends on g i.e. $\exists K_g$ such that if x is equivalent to y then $D(G^n(\xi), G^n(\eta)) < K_g$ for all $n \in \mathbb{Z}$. Moreover, K_g tends to 0 as g tends to f (see* [**L1, H1**]*).*

(2) *Let $[x]_g$ be the equivalence class of $x \in M$. Then, $[x]_g$ is a compact set and $g([x]_g) = [g(x)]_g$.*

(3) *The quotient space under the equivalence relation, $\overline{M}_g$, is a compact metrizable space and, by part 2, g induces a homeomorphism $\overline{g} : \overline{M}_g \to \overline{M}_g$. When $g = f$, $\overline{M}_f = M$ and $\overline{f} = f$, due to the infinite expansivity of any lift of f.*

Proof. — In order to prove the first part of the proposition take F, G to be adequate lifts of f, g and let $R > 0$ be such that $D(F, G) < R$ and $D(F^{-1}, G^{-1}) < R$.

Suppose that $D(\xi, \eta) > 2K$, then $D_U(\xi, \eta) > K$ or $D_S(\xi, \eta) > K$. With no loose of generality we can assume that the first inequality holds. If $K > 2R(\lambda - 1)^{-1}$ there exists $1 < \alpha < \lambda$ such that $K > 2R(\lambda - \alpha)^{-1} > 2R(\lambda - 1)^{-1}$ which implies $D_U(G(\xi), G(\eta)) > \lambda K - 2R > \alpha K$.

Thus, $D_U(G^n(\xi), G^n(\eta)) > \alpha^n K$. Since $\alpha^n K$ tends to infinite with n, this implies that if x and y are equivalent there exist ξ, η so that

$$D(G^n(\xi), G^n(\eta)) \leqslant 4R(\lambda - 1)^{-1} \quad \forall\, n \in \mathbb{Z}.$$

Choose $K_g = 4R(\lambda - 1)^{-1}$.

To prove the second part, take x_k equivalent to x for all $k \in \mathbb{N}$ and $x_k \to_{k\to+\infty} x^*$. Then, given $\xi \in \pi^{-1}(x)$, there are $\xi_k \in \pi^{-1}(x_k)$ such that $D(G^n(\xi_k), G^n(\xi)) < K_g$ $\forall\, n \in \mathbb{Z}$.

By taking, if necessary, a convergent subsequence, we may assume that $\xi_k \to_{k\to+\infty} \xi^* \in \pi^{-1}(x^*)$. Then, $D(G^n(\xi^*), G^n(\xi)) < K_g$ $\forall\, n \in \mathbb{Z}$.

Now we prove the third part of the proposition. The proof is similar to the one included in [**CS**] where the same result for g C^0-close enough to an expansive homeomorphisms is shown.

It is not difficult to see that if $\{x_n\} \subset M$ is a sequence such that $x_n \to x$, $\limsup [x_n]_g \subset [x]_g$. Then, given an open set $U \subset M$, the set $\{y \in M; [y]_g \subset U\}$ is open. This easily implies that $\overline{M}_g$ is Haussdorf and metrizable. □

We want to study the connection between the dynamics of $\overline{g}$ and f; to this end we find conditions on g to be semiconjugated to f.

Define

Definition 2.3. — *The g-orbit of x is shadowed by the f-orbit of y iff there exist $\xi \in \pi^{-1}(x), \eta \in \pi^{-1}(y)$ such that $\{D(G^n(\xi), F^n(\eta)) : n \in \mathbb{Z}\}$ is bounded, for G, F equivariantly homotopic lifts of g, f.*

Observe that if the g-orbit of x is shadowed by the f-orbit of y, then y is unique and every g-orbit of a point of $[x]_g$ is shadowed by the f-orbit of y. Moreover, there is a uniform bound (independent of x and y) for $\{D(G^n(\xi), F^n(\eta)) : n \in \mathbb{Z}\}$ that tends to 0 as g approaches f (see [**H1, L1**]).

It is known that for any $y \in M$ there exists $x \in M$ such that the g-orbit of x is shadowed by the f-orbit of y(see [**H1, L1**]).

OK? ***Definition 2.4***. — *Call J_g the set of $x \in M$ such that the g-orbit of x is shadowed by the f-orbit of some y.*

We remark that J_g consists of *all* points that are f-shadowed.

It is not difficult to see that $J_g = J_{g^n}$, if we consider f^n instead of f. J_g is a compact g-invariant set that not necessarily equals M. Moreover, there exists a continuous surjection homotopic to the inclusion, $h : J_g \to M$, such that $f \circ h = h \circ g|_{J_g}$ (see [**L1, H1**]). This implies that, as the equivalence classes on J_g coincide with $h^{-1}(y)$, $y \in M$, the quotient of J_g under the equivalence relation we are interested in, is homeomorphic to M and $\overline{g}$ restricted to this set is conjugated to f.

Assume now that g is Axiom A; we look for conditions in order to have semiconjugacy to f ($J_g = M$).

Given $\xi \in \mathcal{M}$, let

$$W_S^F(\xi) = \{\eta \in \mathcal{M}; D(F^n(\xi), F^n(\eta)) \to 0 \textit{ as } n \to +\infty\}$$
$$W_U^F(\xi) = \{\eta \in \mathcal{M}; D(F^n(\xi), F^n(\eta)) \to 0 \textit{ as } n \to -\infty\}$$

denote the F-stable and unstable sets of ξ.

Analogously, denote by $W_S^G(\xi)$ and $W_U^G(\xi)$ the G-stable and unstable manifolds of ξ.

When g is an Axiom A diffeomorphism, for $\xi \in \mathcal{M}$, B a basic piece of g and $\pi(\xi) \in B$, we shall denote

$${}^B W_S^G(\xi) = \{\eta \in W_S^G(\xi); \pi(\eta) \in B\}$$
$${}^B W_U^G(\xi) = \{\eta \in W_U^G(\xi); \pi(\eta) \in B\}.$$

Definition 2.5. — *We shall say that B is "small" iff ${}^B W_S^G(\xi)$ is bounded for some (and then for every) $\xi \in \pi^{-1}(B)$. In case that ${}^B W_S^G(\xi)$ is unbounded we say that B is "large".*

Observe that B being small implies that the quotient of B under the equivalence relation is a periodic orbit for $\overline{g}$ (all the points in a topologically mixing component of B are equivalent).

In [**L2**] was proved that large basic pieces for g always exist in the case that g is C^0-near enough to f. The same proof works for g isotopic to f because it is sufficient to take $x \in J_g$ such that its g-orbit is shadowed by an f-orbit dense in the future. Then, the basic piece containing the ω-limit set of x ($\omega_g(x)$) is large.

We will say that a closed curve is essential if it is not nul-homotopic.

Lemma 2.6. — *Let Λ be a large basic piece of g. Then, there exists an essential closed curve consisting of an arc of $W^s(x)$ and an arc of $W^u(x)$, where $x \in \Lambda$.*

Proof. — Suppose that $\gamma = \gamma_s \cup \gamma_u$ is a nul-homotopic closed curve such that $\gamma_s \subset W^s(x)$, $\gamma_u \subset W^u(x)$, $x \in \Lambda$ (we may suppose that $x \in \gamma_s \cap \gamma_u$).

Then, there exists a closed curve $\widetilde{\gamma} \subset \mathcal{M}$ such that $\pi(\widetilde{\gamma}) = \gamma$. On the other hand, $\widetilde{\gamma} = \widetilde{\gamma}_s \cup \widetilde{\gamma}_u$ where $\widetilde{\gamma}_s$ and $\widetilde{\gamma}_u$ are arcs of $W^G_S(\xi)$ and $W^G_U(\xi)$ with $\xi \in \pi^{-1}(x)$. This implies that any point in $\gamma_s \cap \gamma_u$ is equivalent to x.

Thus, if there are no essential curves as in the statement, all points in the closure of $W^s(x)$ are equivalent, and since Λ is a finite union of iterates of this closure it would be small. $\square$

In the next two sections we will use some basic notions about Nielsen classes of fixed points that we define below.

Definition 2.7. — *Let h be an homeomorphism of M and $p, q \in Fix(h)$; p is Nielsen equivalent to q iff for some (and then for any) lift H of h to $\mathcal{M}$ there exist a covering transformation τ and P, Q lifts of p, q such that $H(P) = \tau P$ and $H(Q) = \tau Q$. A Nielsen class of h is essential iff the Lefschetz index of the class is different from 0.*

The Nielsen class of a point $p \in Fix(h)$ is *not removable* if for any homeomorphism k isotopic to h there exist $p' \in Fix(k)$, equivariantly homotopic lifts H, K of h, k, P, P' lifts of p, p' and a covering transformation τ such that $H(P) = \tau P$ and $K(P') = \tau P'$. Essential Nielsen classes are not removable (see [**B**] chapter IV for more details).

Lemma 2.8. — *Two fixed points of g^n are equivalent iff they are Nielsen equivalent (see* [**H1**]*)*

Proof. — The if part follows easily from the definition of Nielsen equivalence.

Now, if p and q are equivalent we have that $G^n(P) = \sigma P$ and $G^n(Q) = \tau Q$ for σ, τ covering transformations and

$$D(P, (\sigma^{-1}G^n)^k Q) = D((\sigma^{-1}G^n)^k P, (\sigma^{-1}G^n)^k Q) < K_g$$

Then, there exists k_0 such that $(\sigma^{-1}G^n)^{k_0} Q = Q$ and, on the other hand, $\sigma^{-1}\tau Q = \sigma^{-1}G^n Q$.

This implies that $(\sigma^{-1}G^n)^{k_0}$ commutes with $\sigma^{-1}\tau$ and, as g is in the isotopy class of a pseudoanosov map, we obtain that $\sigma^{-1}\tau$ is the identity. $\square$

We obtain the following corollary.

Corollary 2.9. — *Let g satisfy Axiom A then there exists $T > 0$ such that for all $n \in \mathbb{Z}$ the number of Nielsen classes of fixed points of g^n included in small basic pieces is less than T.*

Proof. — Since the set of equivalence classes of points in small basic pieces is finite, the result follows from Lemma 2.8. $\square$

3. Uniqueness of large basic pieces

In this section we will use some results of the type of those obtained by Nielsen. The proofs may be found in [**CB, HT, M**].

In what follows we identify $\mathcal{M}$ with the Poincaré disk $\mathbb{D}$. We call S_∞ the boundary of $\mathbb{D}$ as an euclidean subset of $\mathbb{R}^2$ and $\overline{\mathbb{D}} = \mathbb{D} \cup S_\infty$.

***Proposition 3.1* ([CB, HT, M]).** — *Let g be a homeomorphism isotopic to a pseudoanosov map and G any lift of g.*

i) *G has an even number of fixed points in S_∞, alternatively attracting and repelling in $\overline{\mathbb{D}}$.*

ii) *Each of these fixed points has the property that given any simple closed geodesic τ in M it has a nested base of neighbourhoods bounded by lifts of τ.*

This follows from the fact that, associated to the isotopy class of a pseudoanosov map f, there exist two transversal minimal geodesic laminations $\mathcal{F}^s$ and $\mathcal{F}^u$ such that both of them transversely intersects infinitely many times any simple closed geodesic. As a consequence, it is not difficult to prove that any end-point of a lift to $\mathbb{D}$ of a geodesic of $\mathcal{F}^s$ or $\mathcal{F}^u$, has a nested sequence of neighbourhoods bounded by lifts of τ converging to this end-point. Finally let us say that the attracting (repeller) fixed points mentioned above are end-points of geodesics of $\mathcal{F}^u$ (resp. $\mathcal{F}^s$). See [**CB**] Chapter 5 and [**HT**] Proposition 4.3.

***Theorem 3.2*.** — *Let g be an Axiom A diffeomorphism in the isotopy class of a pseudoanosov map then, g has a unique large basic piece.*

Proof. — Let $NE_n(g)$ be the number of essential Nielsen classes of g^n. By [**H1**] all the periodic points in essential classes of g^n are in J_g for all n and for each essential class of g^n, there exists a unique periodic point of f that shadows all the points of this Nielsen class.

As

$$\lim_{n\to\infty} \frac{\log NE_n(g)}{n} = \log \lambda$$

(the topological entropy of f, see [**FLP**]), by using Corollary 2.9 we obtain that there are periodic points of essential Nielsen classes in some large basic pieces.

In order to prove the theorem, we shall show that if Λ is a large basic piece and p is a periodic point of an essential Nielsen class of g^n contained in a large basic piece then $p \in \Lambda$. Obviously, this implies the theorem.

Let Λ be a large basic piece, then by Lemma 2.6 there exists an essential simple closed curve γ consisting of an arc of $W^s(q)$ and an arc of $W^u(q)$ where q is a periodic point of Λ.

Let G_n be a lift of g^{2n} such that $G_n(P) = P$ ($\pi(P) = p$) and fixes (setwise) the separatices of $W^U_G(P)$, where G is a lift of g. Now, if p is an essential fixed point of g^n

contained in a large basic set, by 3.1, $W^U_{G_n}(P) = W^U_G(P)$ accumulates in an attracting fixed point (sink) of the extension of G_n to S_∞ (we also call G_n its extension to $\overline{\mathbb{D}}$).

Since γ is freely homotopic to a simple closed geodesic, any lift of it has the same end-points at S_∞ of the corresponding lift of the geodesic. Then, again by 3.1, $W^U_G(P)$ cuts a lift of γ and this implies that $W^u(p)$ cuts $W^s(q)$. Analogously, $W^s(p)$ cuts $W^u(q)$. Thus $p \in \Lambda$. □

Corollary 3.3. — *$\Omega(\overline{g})$ consists of the quotient of Λ and a finite set of periodic orbits.*

Proof. — First of all observe that, by the density of periodic points, the quotient of $\Omega(g)$ is contained in $\Omega(\overline{g})$.

Now, if $[x]_g \subset \Omega(g)^c$, on account of the fact that each y in $[x]_g$ has a neighbourhood W such that $g^n(W)$ lies very close to the nonwandering set of g for sufficiently large n, it is easy to show that there exists an open set U, $[x]_g \subset U$ such that $g^n(U) \cap U = \varnothing$, $\forall\, n \in \mathbb{N}$. As in the proof of the third part of Proposition 2.2, by using that $V = \{y \in M; [y]_g \subset U\}$ is an open set, it is easy to see that $[x]_g$ is a wandering point of $\overline{g}$. □

4. Conditions for semiconjugacy

From now on let Λ be the large basic set of g given by Theorem 3.2.

Theorem 4.1. — *$\Lambda \subset J_g$ if and only if g is semiconjugated to f.*

Proof. — The "if" part is obvious because, since J_g consists of all points f-shadowed, $J_g = M$.

Let $N_n(h)$ be the number of distinct Nielsen classes of h^n, $h = f, g$. Then, if $\Lambda \subset J_g$, the number of Nielsen classes of g^n intersecting Λ is less or equal to $N_n(f)$ (see [**H1**]). Since, by Lemma 2.9, $N_n(g) \leqslant N_n(f) + T$, this implies that

$$\limsup \frac{\log N_n(g)}{n} \leqslant \log \lambda$$

(in fact the equality holds because f minimizes this number).

Then, by Corollary 3.6 of [**H2**], g is semiconjugated to f. □

Corollary 4.2. — *g is semiconjugated to f iff there exist $\xi, \eta \in \mathcal{M}$, $\pi(\xi) \in \Lambda$, and $V > 0$ such that for every $\zeta \in {}^{\Lambda}W^G_S(\xi)$, $D(\zeta, W^F_S(\eta)) < V$.*

Proof. — See Theorem 4.1 of [**L2**] and observe that the proof works for g in the isotopy class of f. □

5. Exteriorly situated basic pieces

Lemma 5.1. — *Let $p \in Fix(g^n)$ and G_n be a lift of g^n to $\mathbb{D}$ such that $G_n(P) = P$, $\pi(P) = p$ and $\overline{W_U^{G_n}(P)}$ contains two different fixed points of $G_n|_{S_\infty}$. Then, $p \in J_g$.*

Proof. — By [**H1**] it is enough to prove that the Nielsen class of p is essential. In order to prove this, first observe that, since the fixed points of G_n at S_∞ are sinks or sources, the fixed points accumulated by $W_U^{G_n}(P)$ are sinks. Then, $G_n|_{S_\infty}$ has at least four fixed points and, since the Lefschetz index of any of them is $1/2$ (when the fixed point is in S_∞ we have to take into account half of the index of the homeomorphisms on the double of $\overline{\mathbb{D}}$, which agrees with G_n on each copy of $\overline{\mathbb{D}}$) and the Euler's characteristic of $\mathbb{D}$ is 1, we obtain that the index of the Nielsen class of p is strictly negative. ☐

To state Theorem 5.3 we need another definition.

Definition 5.2. — *A nontrivial basic set (different from a periodic point) is called exteriorly situated if there is no nul-homotopic closed curve consisting of an arc of the stable and an arc of the unstable manifold of some point of it (see, for instance,* [**G1**]*).*

Observe that it is not difficult to prove that exteriorly situated basic pieces are large, since otherwise the whole stable (unstable) manifold of a periodic point would be included in a disk centered at the periodic point.

Theorem 5.3. — *If Λ is exteriorly situated then g is semiconjugated to f.*

Proof. — By Theorem 3.2 it is enough to show that $\Lambda \subset J_g$.

Take a periodic $p \in \Lambda$ such that it is not a boundary point of Λ. Now, take a simple closed curve $\gamma = \gamma_U \cup \gamma_S$ formed by oriented arcs of stable and unstable manifolds of Λ (not containing boundary points) intersecting at two points of Λ with the same oriented intersection number. The lift of γ is formed by a disjoint union of curves with different end-points at S_∞ which corresponds to the end-points of the connected lifts of the simple closed geodesic homotopic to γ. Let $P \in \pi^{-1}(p)$. Then, we claim that $W_U^G(P)$ cuts a connected lift of γ, say $\widetilde{\gamma}$, in at most one point. To see this we first observe that, as Λ is exteriorly situated, $W_U^G(P)$ cannot intersect twice the same lift of γ_S. Thus, if $W_U^G(P)$ cuts twice $\widetilde{\gamma}$, an arc of it and an arc of $\widetilde{\gamma}$ bounds a disk D. The boundary of D contains at least one lift $\widetilde{\gamma}_U$ of γ_U and l arcs contained in lifts of γ_S. By the form we choose γ we can continue $\widetilde{\gamma}_U$ inside D along the lift of the unstable manifold containing γ_U. As this curve is unbounded, it should cut the boundary of D at a lift of γ_S. In this way we obtain a new disk bounded by an arc of a lift of a unstable manifold of Λ, arcs of lifts of γ_U and, at most, $l-1$ arcs of lifts of γ_S. Repeating this argument, we will obtain a disk bounded by an arc of unstable manifold of Λ and one arc in a lift of γ_S which contradicts the fact of Λ being exteriorly situated (see [**G1**] for a similar argument). As $W^u(p)$ intersects infinitely many times γ, it is not difficult to see that $W_U^G(P)$ has exactly two end-points at S_∞.

Now, if G_n is a lift of g^n that fixes P and the separatrices of $W_U^G(P)$ we obtain that the end-points of $W_U^{G_n}(P) = W_U^G(P)$ are fixed and, by Lemma 5.1, $p \in J_g$. The density of non-boundary periodic points in Λ implies the statement. $\square$

Definition 5.4. — *Suppose that a curve $\gamma : [0, +\infty) \to M$ has a lift $\widetilde{\gamma}$ with one end-point $\lim_{t \to +\infty} \widetilde{\gamma}(t) = \widetilde{\gamma}^+ \in S_\infty$. Let τ be a geodesic that also has $\widetilde{\gamma}^+$ as end-point. From $\widetilde{\gamma}(t)$ take the perpendicular τ_t^{per} to τ and call $\tau_t = \tau \cap \tau_t^{per}$. We say that γ has the property of bounded deviation from geodesic if $\{d(\widetilde{\gamma}(t), \tau_t); t \in [0, +\infty)\}$ is bounded. Here d is the metric of constant negative curvature (see* [**G1**]*) Observe that the property is independent of the geodesic with end-point $\widetilde{\gamma}^+$.*

Corollary 5.5. — *If x beelongs to the exteriorly situated basic piece Λ, the connected components of $W_g^s(x) \smallsetminus \{x\}$ and $W_g^u(x) \smallsetminus \{x\}$ not containing boundary points verifies the property of bounded deviation from geodesics.*

Proof. — Call an f-stable (unstable) half-leaf of x an arcwise connected component of $W_f^s(x) \smallsetminus \{x\}$ ($W_f^u(x) \smallsetminus \{x\}$) when it does not contain singularities. In case that $W_f^s(x)$ ($W_f^u(x)$) contains a singularity all the half-leaves of x are defined, in a natural way, by choosing one at each time the half-leaves of the singularity.

Half-leaves of f have the property of bounded deviation from geodesics (see, for instance, the construction of pseudoanosov maps of [**M**] or [**HT**] or the results about foliations of [**A**])

Then, as images of the g-stable and unstable manifolds by the semiconjugation are contained, respectively, in the stable and unstable sets of f and the semiconjugation is homotopic to the identity, the corollary is proved. $\square$

References

[A] S. Aranson, *Qualitative properties of foliations of closed surfaces*, J. Dynam. & Control Syst. **6** (2000) 127–157.

[B] R. Brown, *The Lefschetz Fixed Point Theorem*, Scott, Foresman and Company (1971).

[CB] A. Casson, S. Bleiler, *Automorphisms of surfaces after Nielsen and Thurston*, London Math. Soc. Student Text **9** (1988).

[CS] M. Cerminara, M. Sambarino, *Stable and unstable sets for C^0 perturbations of expansive homeomorphisms of surfaces* , Nonlinearity **12** (1999) 321–332.

[FLP] A. Fathi, F. Laudenbach, V. Poenaru, *Travaux de Thurston sur les surfaces* , (Séminaire Orsay) Astérisque **66-67** (1979).

[G1] V. Grines, *Structural stability and asymptotic behaviour of invariant manifolds of A-diffeomorphisms of surfaces*, J. Dynam. & Control Syst. **3** (1997) 91–110.

[G3] V. Grines, *Flows on two dimensional manifolds* , In:Ch.4, Encyclopedia of Math.Sci., **1**, Springer-Verlag, Berlin-New York,(1988) 151–242.

[G2] V. Grines, *On topological classification of A-diffeomorphisms of surfaces* , J. Dynam. & Control Syst. **6** (2000) 97–126.

[H1] M. Handel, *Global shadowing of Pseudo Anosov homeomorphisms*, Erg. Th. & Dyn. Sys. **5** (1985) 373–377.

[H2] M. Handel, *Entropy and semiconjugacy in dimension two*, Erg. Th. & Dyn. Sys. **8** (1988) 585–596.

[HT] M. Handel, W. Thurston, *New proofs of some results of Nielsen* , Adv. in Math **56** (1985) 173–191.

[L1] J. Lewowicz, *Persistence in Expansive Systems*, Erg. Th. & Dyn. Sys. **3** (1983) 567–578.

[L2] J. Lewowicz, *On Axiom A diffeomorphisms C^0-close to Pseudo Anosov maps*, to appear in Proceedings Int. Conf. on Topology and its Appl., Top. & its Appl. (special issue).

[M] R. Miller, *Geodesic laminations from Nielsen's viewpoint* , Adv. in Math **45** (1982) 189–212.

[RW] R. Robinson, R. Williams, *Finite Stability is not generic* , Proc. Symp. Univ. of Brazil (1971), Academic Press, New York, London (1973) 451–462.

[T] W. Thurston *On the geometry and dynamics of diffeomorphisms of surfaces*, preprint and Bull. Am. Math. Soc. **19** (1988), 417–431.

J. Lewowicz, CC 30, IMERL - Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay • *E-mail :* lew@fing.edu.uy

R. Ures, CC 30, IMERL - Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay • *E-mail :* ures@fing.edu.uy

Astérisque
287, 2003, p. 135–146

SUB-ACTIONS FOR ANOSOV DIFFEOMORPHISMS

by

Artur O. Lopes & Philippe Thieullen

Dedicated to Jacob Palis

Abstract. — We show a positive Livsic type theorem for $\mathcal{C}^2$ Anosov diffeomorphisms f on a compact boundaryless manifold M and Hölder observables A. Given $A : M \to \mathbb{R}$, α-Hölder, we show there exist $V : M \to \mathbb{R}$, β-Hölder, $\beta < \alpha$, and a probability measure μ, f-invariant such that

$$A \leqslant V \circ f - V + \int A \, d\mu.$$

We apply this inequality to prove the existence of an open set $\mathcal{G}_\beta$ of β-Hölder functions, β small, which admit a unique maximizing measure supported on a periodic orbit. Moreover the closure of $\mathcal{G}_\beta$, in the β-Hölder topology, contains all α-Hölder functions, α close to one.

1. Introduction

We consider a compact riemannian manifold M of dimension $d \geqslant 2$ without boundary and a $\mathcal{C}^2$ transitive Anosov diffeomorphism $f : M \to M$. The tangent bundle TM admits a continuous Tf-invariant splitting $TM = E^u \bigoplus E^s$ of expanding and contracting tangent vectors. We assume M is equiped with a riemannian metric and there exists a constant $C(M)$, depending only on M and the metric and constants depending on f

$$0 < \Lambda_s < \lambda_s < 1 < \lambda_u < \Lambda_u$$

such that for all $n \in \mathbb{Z}$

$$\begin{cases} C(M)^{-1}\lambda_u^n \leqslant \|T_x f^n \cdot v\| \leqslant C(M)\Lambda_u^n & \text{for all } v \text{ in } E_x^u, \\ C(M)^{-1}\Lambda_s^n \leqslant \|T_x f^n \cdot v\| \leqslant C(M)\lambda_s^n & \text{for all } v \text{ in } E_x^s. \end{cases}$$

2000 *Mathematics Subject Classification.* — 37D20.
Key words and phrases. — Anosov diffeomorphisms, minimizing measures.

A.L.: Partially supported by PRONEX-CNPq - Sistemas Dinâmicas and Institute of Millenium - IMPA.
Ph.T.: Partially supported by CNRS URA 1169.

Livsic theorem [**5**] asserts that, if $A : M \to M$ is a given Hölder function and satisfies $\int A\, d\mu = 0$ for all f-invariant probability measure μ, then A is equal to a coboundary V (which is Hölder too), that is:

$$A = V \circ f - V.$$

What hapens if we only assume $\int A\, d\mu \geqslant 0$ for all f-invariant probability measure μ? We denote by $\mathcal{M}(f)$, the set of f-invariant probability measures and $m(A, f) = \sup\big\{ \int A\, d\mu \mid \mu \in \mathcal{M}(f)\big\}$.

For a β-Holder function V

$$\text{Höld}_\beta(V) = \sup_{0<d(x,y)} \Big\{ \frac{|V(x) - V(y)|}{d(x,y)^\beta} \Big\}.$$

We prove the following:

Theorem 1. — *Let $f : M \to M$ be a $\mathcal{C}^2$ transitive Anosov diffeomorphism on a compact manifold M without boundary. For any given α-Hölder function $A : M \to \mathbb{R}$, there exists a β-Hölder function $V : M \to \mathbb{R}$, that we call sub-action, such that:*

$$A \leqslant V \circ f - V + m(A, f),$$

and

$$\beta = \alpha \frac{\ln(1/\lambda_s)}{\ln(\Lambda_u/\lambda_s)}, \quad \text{Höld}_\beta(V) \leqslant \frac{C(M)}{\min(1 - \lambda_u^{-\alpha}, 1 - \lambda_s^\alpha)^2} \,\text{Höld}_\alpha(A)$$

where $C(M)$ is some constant depending only on M and the metric.

By analogy with Hamiltonian mechanics and the way we define V from A, we may interpret A as a lagrangian and V as a sub-action. This result extends a similar one we obtained in [**4**] for expanding maps of the circle (see [**2**] [**6**] for related results). The same techniques of [**4**] also apply for the one-directional shift as it is mentioned in [**4**].

The proof we give here is for bijective smooth systems, and we obtain V continuous in all M. Our result can not be derived (via Markov partition) directly from an analogous result for the bi-directional shift.

Corollary 2. — *The hypothesis are the same as in theorem 1. The following statements are equivalent:*

(i) *$A \geqslant V \circ f - V$ for some bounded measurable function V,*
(ii) *$\int A\, d\mu \geqslant 0$ for all f-invariant probability measure μ,*
(iii) *$\sum_{k=0}^{p-1} A \circ f^k(x) \geqslant 0$ for all $p \geqslant 1$ and point x periodic of period p,*
(iv) *$A \geqslant V \circ f - V$ for some Hölder function V.*

The proof of that corollary is straightforward and uses (for (iii) $\Rightarrow$ (ii)) the fact that the convex hull of periodic measures is dense in the set of all f-invariant probability measures for topological dynamical systems satisfying the shadowing lemma (see Lemma 5). F. Labourie suggested to us the following corollary:

Corollary 3. — *The hypothesis are the same as in theorem 1. If A satisfies $\int A\,d\mu \geqslant 0$ for all $\mu \in \mathcal{M}(f)$ and $\sum_{k=0}^{p-1} A \circ f^k(x) > 0$ for at least one periodic orbit x of period p then $\int A\,d\lambda > 0$ for all probability measure λ giving positive mass to any open set.*

Again the proof is straitforward: $R = A - V \circ f + V \geqslant 0$ for some continuous V and $\int R\,d\lambda = 0$ for such a measure λ implies $R = 0$ everywhere and in particular $\sum_{k=0}^{p-1} A \circ f(x) = 0$ for all periodic orbit x.

Any measure μ satisfying $\int A\,d\mu = m(A,f)$ is called a maximizing measure and since A is continuous, such a measure always exists. It is then natural to ask the following two questions: For which A, the set of maximizing measures is reduced to a single measure ? In the case there exists a unique maximizing measure, to what kind of compact set, the support of this measure looks like ?

The following theorem gives a partial answer for "generic" functions A.

Theorem 4. — *Let $f : M \to M$ be a $\mathcal{C}^2$ transitive Anosov diffeomorphism and $\beta < \ln(1/\lambda_s)/\ln(\Lambda_u/\lambda_s)$. Then there exists an open set $\mathcal{G}_\beta$ of β-Hölder functions (open in the $\mathcal{C}^\beta$-topology) such that:*

(i) *any A in $\mathcal{G}_\beta$ admits a unique maximizing measure μ_A;*

(ii) *the support of μ_A is equal to a periodic orbit and is locally constant with respect to $A \in \mathcal{G}_\beta$;*

(iii) *any α-Hölder function with $\alpha > \beta\ln(\Lambda_u/\lambda_s)/\ln(1/\lambda_s)$ is contained in the closure of $\mathcal{G}_\beta$ (the closure is taken with respect to the $\mathcal{C}^\beta$-topology).*

The proof of Theorem 4 is a simplification of what we gave in [**4**] in the one-dimensional setting. The existence of sub-actions is in both cases the main ingredient of the proof.

Now we will concentrate in one of our main results, namely, Theorem 1; the basic idea is the following: given a finite covering of M by open sets $\{U_1,\dots,U_l\}$ with sufficiently small diameter, we construct a Markov covering (and not a Markov partition) $\{R_1,\dots,R_l\}$ of rectangles: each R_i contains U_i and satisfies

$$x \in U_i \cap f^{-1}(U_j) \Longrightarrow f(W^s(x,R_i)) \subset W^s(f(x),R_j),$$

where $W^s(x,R_i)$ denotes the local stable leaf through x restricted to R_i. We then associate to each R_i a local sub-action V_i, defined on R_i by:

$$V_i(x) = \sup\big\{S_n(A-m)\circ f^{-n}(y) + \Delta^s(y,x) \mid n \geqslant 0,\ y \in W^s(x,R_i)\big\}$$

where $\Delta^s(y,x)$ is a kind of cocycle along the stable leaf $W^s(x)$:

$$\Delta^s(y,x) = \sum_{n\geqslant 0}(A\circ f^n(y) - A\circ f^n(x)),$$

and where $S_n(A-m) = \sum_{k=0}^{n-1}(A-m)\circ f^k$.

This family $\{V_1,\dots,V_l\}$ of local sub-actions satisfies the inequality:

$$x \in U_i \cap f^{-1}(U_j) \Longrightarrow V_i(x) + A(x) - m \leqslant V_j \circ f(x)$$

and enable us to construct a global sub-action V:

$$V(x)=\sum_{i=1}^{l}\theta_i(x)V_i(x)$$

where $\{\theta_1,\dots,\theta_l\}$ is a smooth partition of unity associated to the covering $\{U_1,\dots,U_l\}$. The main difficulty is to prove that each V_i is Hölder on R_i.

2. Existence of sub-actions

We continue our description of the dynamics of transitive Anosov diffeomorphisms (for details information, see Bowen's monography [**3**]). All the results we are going to use depend on a small constant of expansiveness $\varepsilon^*>0$ (by definition this constant says that any pseudo-orbit can be followed by true orbits) depending on f and M in the following way:

$$\varepsilon^* = C(M)^{-1}\min\Big(\frac{\lambda_u-1}{\|D^2f\|_\infty},\frac{1-\lambda_s}{\|D^2f\|_\infty}\Big)$$

where $C(M)\geqslant 1$ is a constant depending only on M and the riemannian metric. At each point x, one can define its local stable manifold $W^s_\varepsilon(x)$ for every $\varepsilon<\varepsilon^*$:

$$W^s_\varepsilon(x)=\big\{y\in M \mid d(f^n(x),f^n(y))\leqslant\varepsilon\ \forall\, n\geqslant 0\big\}$$

which are $\mathcal{C}^2$ embeded closed disks of dimension $d^s=\dim E^s_x$ and tangent to E^s_x. In the same manner, $W^u_\varepsilon(x)$ is defined replacing f by f^{-1}. If two points x,y are close enough, $d(x,y)<\delta$, then $W^s_\varepsilon(x)$ and $W^u_\varepsilon(y)$ have a unique point in common, called $[x,y]$:

$$[x,y]=W^s_\varepsilon(x)\cap W^u_\varepsilon(y)=W^s_{\varepsilon^*}(x)\cap W^u_{\varepsilon^*}(y),$$

where $\varepsilon=K^*\delta$ and K^* is again a large constant depending on M and f:

$$K^*=\frac{C(M)}{\min\left(1-\lambda_u^{-1},1-\lambda_s\right)}.$$

This estimate is in fact a particular case of Bowen's shadowing lemma:

***Lemma 5* (Bowen)**. — *If δ is small enough, $\delta<\varepsilon^*/K^*$, if $(x_n)_{n\in\mathbb{Z}}$ is a bi-infinite δ-pseudo-orbit, that is, $d(f(x_n),x_{n+1})<\delta$ for all $n\in\mathbb{Z}$, then there exists a unique true orbit $\{f^n(x)\}_{n\in\mathbb{Z}}$ which ε-shadow $(x_n)_{n\in\mathbb{Z}}$, that is $d(f^n(x),x_n)<\varepsilon$ for all $n\in\mathbb{Z}$ with $\varepsilon=K^*\delta$.*

This lemma (see [**3**]) for proof) is the main ingredient for constructing (dynamical) rectangles. A rectangle R is a closed set of diameter less than ε^*/K^* satisfying:

$$x,y\in R\Longrightarrow [x,y]\in R.$$

We will not use the notion of proper rectangles but will use instead the notion of Markov covering.

Definition 6. — Let $\mathcal{U} = \{U_1, \dots, U_l\}$ be a covering of M by open sets of diameter less than $\varepsilon^*/(K^*)^2$. We call a Markov covering associated to $\mathcal{U}$, a finite set $\mathcal{R} = \{R_1, \dots, R_l\}$ of rectangles of diameter less than ε^*/K^* satisfying:

$$\begin{aligned}
&U_i \subset R_i \\
&x \in U_i \cap f^{-1}(U_j) \Longrightarrow f(W^s(x, R_i)) \subset W^s(f(x), R_j) \\
&y \in f(U_i) \cap U_j \Longrightarrow f^{-1}(W^u(y, R_j)) \subset W^u(f^{-1}(y), R_i) \\
&\forall\, j,\ \exists\, i,\ f(U_i) \cap U_j \neq \varnothing
\end{aligned}$$

where $W^s(x, R_i) = W^s_{\varepsilon^*}(x) \cap R_i$ and $W^u(y, R_j) = W^u_{\varepsilon^*}(y) \cap R_j$.

An easy consequence of the shadowing lemma shows there always exist such Markov coverings:

Proposition 7. — *For every covering $\mathcal{U}$ of M by open sets such that the diameter of each U_i is less than $\varepsilon^*/(K^*)^2$, there exists a Markov covering $\mathcal{R}$ by rectangles of diameter less than ε^*/K^*.*

Proof. — Given $\mathcal{U} = \{U_1, \dots, U_l\}$ such a covering, we define the following compact space of $\varepsilon^*/(K^*)^2$ pseudo-orbits:

$$\Sigma = \{\omega = (\dots, \omega_{-2}, \omega_{-1} \mid \omega_0, \omega_1, \dots) \quad \text{s.t.} \quad U_{\omega_n} \cap f^{-1}(U_{\omega_{n+1}}) \neq \varnothing\}.$$

Here ω is a sequence of indices in $\{1, \dots, l\}$ and Σ is a subshift of finite type where $i \to j$ is a possible transition iff $U_i \cap f^{-1}(U_j)$ is not empty. Given such $\omega \in \Sigma$, we choose for all $n \in \mathbb{Z}$, $x_n \in U_{\omega_n}$ so that $f(x_n) \in U_{\omega_{n+1}}$. Then $(x_n)_{n\in\mathbb{Z}}$ is a $\varepsilon^*/(K^*)^2$ pseudo-orbit which corresponds to a unique true orbit $(f^n(x))_{n\in\mathbb{Z}}$ satisfying:

$$d(f^n(x), U_{\omega_n}) < \varepsilon^*/K^* \quad \forall\, n \in \mathbb{Z}.$$

Since ε^* is a constant of expansiveness, there can exists at most one point x satisfying the previous inequality for all n. We call that point $\pi(\omega)$ and notice that the map

$$\pi : \Sigma \to M$$

is surjective (for $\mathcal{U}$ is a covering), commutes with the left shift σ, $f \circ \pi = \pi \circ \sigma$, is continuous by expansiveness (in fact Hölder if Σ is equiped with the standard metric). Also notice that π may not be finite-to-one. We first construct a Markov cover on Σ as usual by the braket

$$[\omega, \omega'] = (\dots, \omega'_{-2}, \omega'_{-1} \mid \omega_0, \omega_1, \dots)$$

where $\omega = (\omega_n)_{n\in\mathbb{Z}}$, $\omega' = (\omega'_n)_{n\in\mathbb{Z}}$ and $\omega'_0 = \omega_0$. By uniqueness in the construction of $\pi(\omega)$, we get

$$\begin{aligned}
&\pi([\omega, \omega']) = [\pi(\omega), \pi(\omega')] \\
&\pi([i]) = R_i \text{ is a rectangle of } M \text{ containing } U_i \\
&\pi(W^s(\omega, [i])) = W^s(\pi(\omega), R_i) \quad \text{whenever } \omega_0 = i
\end{aligned}$$

where $[i]$, $i = 1, \dots, l$, is the cylinder $\{\omega \in \Sigma \mid \omega_0 = i\}$ and $W^s(\omega, [i])$ is the symbolic stable set $\{\omega' \in \Sigma \mid \omega'_n = \omega_n \ \forall n \geqslant 0\}$. (For the proof of the last equality, we just notice: if $x = \pi(\omega)$, $y \in W^s(x, R_i)$ and $y = \pi(\omega')$ then $\pi([\omega, \omega']) = y$ and $[\omega, \omega'] \in W^s(\omega, [i])$.) To finish the proof we only show

$$x \in U_i \cap f^{-1}(U_j) \Longrightarrow f(W^s(x, R_i)) \subset W^s(f(x), R_j).$$

Indeed, $x = \pi(\omega)$ for some $\omega = (\cdots, \omega_{-1} \mid i, j, \omega_2, \cdots)$ and

$$\sigma(W^s(\omega, [i]) \subset W^s(\sigma(\omega), [j]).$$

To conclude, we apply π on both sides. □

Definition 8. — Let $\mathcal{R} = \{R_1, \dots, R_l\}$ be a Markov covering of M associated to some open covering $\mathcal{U} = \{U_1, \dots, U_l\}$. We define a local sub-action by

$$V_i(x) = \sup\{\, S_n(A - m) \circ f^{-n}(y) + \Delta^s(y, x) \mid n \geqslant 0,\ y \in W^s(x, R_i),\, \}$$

for $x \in U_i$, and where $S_n B = \sum_{k=0}^{n-1} B \circ f^k$, $\Delta^s(y, x) = \sum_{k\geqslant 0}(A \circ f^k(y) - A \circ f^k(x))$ and the supremum is taken over all $n \geqslant 0$ and points $y \in W^s(x, R_i)$.

Before showing V_i is a (finite!) Hölder function on each R_i, let's conclude the proof of Theorem 1:

Proof of Theorem 1. — Let $\mathcal{U} = \{U_1, \dots, U_l\}$ be an open covering of M, $\{R_1, \dots, R_l\}$ a Markov covering associated to $\mathcal{U}$ and $\{\theta_1, \dots, \theta_l\}$ a partition of unity adapted to $\mathcal{U}$. Let $\{V_1, \dots, V_l\}$ constructed as above and

$$V = \sum_i \theta_i V_i.$$

Suppose we have proved that $x \in U_i \cap f^{-1}(U_j)$ implies

$$V_i(x) + (A - m)(x) \leqslant V_j \circ f(x).$$

Multiplying this inequality by $\theta_i(x)\theta_j \circ f(x)$ and summing over i and j (whether or not $i \to j$ is a possible transition), we get

$$V(x) + (A - m)(x) \leqslant V \circ f(x) \quad (\forall\, x \in M).$$

We now prove the local sub-cohomological equation: if $x \in U_i \cap f^{-1}(U_j)$ and $y \in W^s(x, R_i)$, then $f(y) \in W^s(f(x), R_j)$ and

$$\begin{aligned} S_n(A - m) \circ f^{-n}(y) &+ \Delta^s(y, x) + (A - m)(x) \\ &= S_{n+1}(A - m) \circ f^{-(n+1)} \circ f(y) + \Delta^s(f(y), f(x)) \leqslant V_j \circ f(x). \end{aligned}$$

Taking the supremum over all $n \geqslant 0$ and all $y \in W^s(x, R_i)$, we get indeed

$$V_i(x) + (A - m)(x) \leqslant V_j \circ f(x).$$

That finishes the proof of theorem 1. □

We now come to our main technical lemma. We notice that, even in the case where A is Lipschitz, we only obtain a Hölder sub-action.

Lemma 9. — *If A is α-Hölder on M, R is a rectangle and V is defined as in Definition 8, then V is β-Hölder on R with exponent*

$$\beta = \alpha \frac{|\ln \lambda_s|}{\ln \Lambda_u + |\ln \lambda_s|} < \alpha.$$

Proof. — We divide the proof into four steps:

Step one. — If $d(x, x') < \varepsilon^*$ and x, x' are on the same stable leaf, then

$$\Delta^s(x, x') \leqslant \sum_{n \geqslant 0} |A \circ f^n(x) - A \circ f^n(x')| \leqslant C(M) \frac{\text{Höld}_\alpha(A)}{1 - \lambda_s^\alpha} d(x, x')^\alpha,$$

for some constant $C(M)$ depending only on M and the metric.

Indeed, it follows from the contraction $d(f^k(x), f^k(x')) \leqslant C(M)\lambda_s^k d(x, x')$ for $k \geqslant 0$ and the fact that A is α-Hölder.

Step two. — For every $n \geqslant 1$, $x, x' \in M$ such that $d(f^k(x), f^k(x')) < \varepsilon^*/K^*$ for all $0 \leqslant k \leqslant n$, then

$$\sum_{k=0}^{n-1} |A \circ f^k(x) - A \circ f^k(x')| \leqslant K(M, f) \max(d(x, x')^\alpha, d(f^n(x), f^n(x'))^\alpha),$$

$$\text{where} \quad K(M, f) = C(M) \frac{\text{Höld}_\alpha(A)}{\min(1 - \lambda_u^{-\alpha}, 1 - \lambda_s^\alpha)^2}.$$

Indeed, one can build $w = [x, x']$; then on the one hand, $d(x, w) \leqslant \varepsilon^*$ and x, w are on the same stable leaf; on the other hand, $d(f^n(w), f^n(x')) \leqslant \varepsilon^*$ and $f^n(w)$ and $f^n(x')$ are on the same unstable leaf. We conclude by applying step one and the estimates:

$$d(x, w) \leqslant K^* d(x, x'), \quad d(f^n(w), f^n(x')) \leqslant K^* d(f^n(x), f^n(x')).$$

Step three. — We show that $V(x)$ is finite for every $x \in R$. It is precisely here that the choice of the normalizing constant $m(A, f)$ is important.

Indeed, since a transitive Anosov diffeomorphism is mixing (see [**3**]), there exists an integer $\tau^* \geqslant 1$ such that, for every finite orbit $\{f^{-n}(y), \dots, f^{-1}(y), y\}$, n arbitrary, $f^{\tau^*}(B(y, \varepsilon^*/K^*))$ contains $f^{-n}(y)$. Thanks to the shadowing lemma, there exists a periodic orbit z, of period $n + \tau^*$, satisfying

$$d(f^{-k}(z), f^{-k}(y)) \leqslant \varepsilon^* \qquad (\forall\, k = 0, 1, \dots, n).$$

Using step two, $\sum_{k=1}^{n} (A \circ f^{-k}(y) - A \circ f^{-k}(z))$ is uniformly bounded in n by some constant $C(M, f)$. As any periodic orbit is associated to an invariant probability, then, $\sum_{k=1}^{n+\tau^*} (A \circ f^{-k}(z) - m(A, f)) \leqslant 0$.

Without lost of generality we can assume $m(A,f)=0$. Therefore, we get

$$\begin{aligned}\sum_{k=1}^{n} A\circ f^{-k}(y) &\leqslant C(M,f)+\sum_{k=1}^{n+\tau^*} A\circ f^{-k}(z)+\tau^*\|A\|_\infty\\ &\leqslant C(M,f)+\tau^*\|A\|_\infty.\end{aligned}$$

Step four. — We finally prove that V is Hölder on R. Let $n\geqslant 0$, $x,x'\in R$, $y\in W^s(x,R)$ and define $y'=[x',y]$ belonging to R since R is a rectangle and to the same local unstable manifold as y. Then for some N we are going to choose soon: let $B=A-m(A,f)$,

$$\begin{aligned}S_nB\circ f^{-n}(y)+\Delta^s(y,x)\leqslant S_nB\circ f^{-n}(y')+\Delta^s(y',x')&\\ +\sum_{k=-n}^{N-1}|A\circ f^k(y)-A\circ f^k(y')| &\qquad (=\Sigma_1)\\ +\sum_{k=0}^{N-1}|A\circ f^k(x)-A\circ f^k(x')| &\qquad (=\Sigma_2)\\ +|\Delta^s(f^N(y),f^N(x))| &\qquad (=\Sigma_3)\\ +|\Delta^s(f^N(y'),f^N(x'))| &\qquad (=\Sigma_4)\end{aligned}$$

We now bound from above each Σ_i with respect to $d(x,x')$:

$$\begin{aligned}\Sigma_1 &\leqslant C(M)\frac{\text{Höld}_\alpha(A)}{1-\lambda_u^{-\alpha}}d(f^N(y),f^N(y'))^\alpha,\\ \Sigma_2 &\leqslant C(M)\frac{\text{Höld}_\alpha(A)}{\min(1-\lambda_u^{-\alpha},1-\lambda_s^\alpha)^2}\max(d(x,x')^\alpha,d(f^N(x),f^N(x'))^\alpha),\\ \Sigma_3 &\leqslant C(M)\frac{\text{Höld}_\alpha(A)}{1-\lambda_s^\alpha}d(f^N(y),f^N(x)),\\ \Sigma_4 &\leqslant C(M)\frac{\text{Höld}_\alpha(A)}{1-\lambda_s^\alpha}d(f^N(y'),f^N(x'))^\alpha.\end{aligned}$$

We now choose $N=N(x,x')$ by $\lambda_s^t\varepsilon^*=\Lambda_u^t d(x,x')$, $N=[t]+1$ and then choose $\widetilde{\varepsilon}\geqslant\varepsilon^*$ so that $\lambda_s^N\widetilde{\varepsilon}=\Lambda_u^N d(x,x')$. Then

$$\begin{gathered}d(f^N(x),f^N(x'))\leqslant C(M)\Lambda_u^N d(x,x')\leqslant C(M)\lambda_s^N\widetilde{\varepsilon},\\ d(f^N(y),f^N(x)) \text{ or } (f^N(y'),f^N(x'))\leqslant C(M)\lambda_s^N\varepsilon^*\leqslant C(M)\lambda_s^N\widetilde{\varepsilon}.\end{gathered}$$

In particular, we get first $d(f^N(y),f^N(y'))\leqslant 3C(M)\lambda_s^N\widetilde{\varepsilon}$ and next:

$$\begin{gathered}\Sigma_1+\cdots+\Sigma_4\leqslant 6C(M)\frac{\text{Höld}_\alpha(A)}{\min(1-\lambda_u^{-\alpha},1-\lambda_s^\alpha)^2}(\lambda_s^N\widetilde{\varepsilon})^\alpha=K(M,f)(\lambda_s^N\widetilde{\varepsilon})^\alpha,\\ S_nB\circ f^{-n}(y)+\Delta^s(y,x)\leqslant S_nB\circ f^{-n}(y')+\Delta^s(y',x')+K(M,f)(\lambda_s^N\widetilde{\varepsilon})^\alpha,\\ V(x)\leqslant V(x')+K(M,f)(\lambda_s^N\widetilde{\varepsilon})^\alpha.\end{gathered}$$

But

$$\lambda_s^N \widetilde{\varepsilon} = d(x,x')^{\ln(1/\lambda_s)/\ln(\Lambda_u/\lambda_s)}. \qquad \square$$

Remark 10. — We have not used explicitly the fact that the stable foliation W^s is Hölder but our proof (step four) is close to showing W^s is Hölder of exponent $\gamma = \ln(\lambda_u/\lambda_s)/\ln(\Lambda_u/\lambda_s)$.

Proof. — We show that if $\varepsilon < \varepsilon^*/K^*$, $d(x,x') \leqslant \varepsilon$, $y \in W^s_\varepsilon(x)$, $y' \in W^s_\varepsilon(x')$ and $y \in W^u_{\varepsilon^*}(y')$ then

$$d(y,y') \leqslant 3C(M)^2 d(x,x')^\gamma$$

where $\gamma = \ln(\lambda_u/\lambda_s)/\ln(\Lambda_u/\lambda_s)$.

Indeed we choose $t > 0$ real such that $\lambda_s^t \varepsilon = \Lambda_u^t d(x,x')$, $N = [t]+1$, and $\widetilde{\varepsilon}$ close to ε so that $\lambda_s^N \widetilde{\varepsilon} = \Lambda_u^N d(x,x')$ where $\widetilde{\varepsilon}/\varepsilon$ varies between 1 and Λ_u/λ_s. Then

$$d(f^N(x), f^N(y)) \text{ or } d(f^N(x'), f^N(y')) \text{ or } d(f^N(x), f^N(x')) \leqslant C(M)\lambda_s^N \widetilde{\varepsilon},$$
$$d(f^N(y), f^N(y')) \leqslant 3C(M)\lambda_s^N \widetilde{\varepsilon},$$
$$d(y,y') \leqslant 3C(M)^2(\lambda_s/\lambda_u)^N \widetilde{\varepsilon} = 3C(M)^2 d(x,x')^\gamma. \qquad \square$$

3. Maximizing periodic measures

The proof of Theorem 4 requires two ingredients: the first one is the notion of sub-actions we have already studied, the second is the notion of strongly non-wandering points we are going to explain.

Definition 11. — Given $A \in \mathcal{C}^\beta(M)$ and $m = m(A,f)$, a point $x \in M$ is said to be strongly non-wandering with respect to A, if for any $\varepsilon > 0$, there exist $n \geqslant 1$ and $y \in M$ such that

$$y \in B(x,\varepsilon), \quad f^n(y) \in B(x,\varepsilon) \quad \text{and} \quad \Big|\sum_{k=0}^{n-1}(A-m)\circ f^k(y)\Big| < \varepsilon$$

where $B(x,\varepsilon)$ denotes the ball centered at x and radius ε. We call $\Omega(A,f)$ the set of strongly non-wandering points.

The first non-trivial but easy observation is that $\Omega(A,f)$ is non-empty; more precisely:

Lemma 12. — *The set $\Omega(A,f)$ is compact forward and backward f-invariant and contains the support of any maximizing measure.*

Proof. — If μ is maximizing, by Atkinson's theorem [**1**], for almost μ-point x, the Birkhoff's sums $\sum_{k=0}^{n-1}(A-m)\circ f^k$ are recurrent (in the sense of random walk theory) to $\int (A-m)\,d\mu = 0$: that is, for any Borel set B of positive μ-measure and for any $\varepsilon > 0$, the set

$$\Big\{x \in B \mid \exists\, n \geqslant 1 \ f^n(x) \in B \text{ and } \big|\textstyle\sum_{k=0}^{n-1}(A-m)\circ f^k(x)\big| < \varepsilon\Big\}$$

has positive μ-measure. Since by definition of the support of a measure, any ball $B(x,\varepsilon)$ has positive μ-measure, we have proved that $\mathrm{supp}(\mu)$ is included in $\Omega(A,f)$. □

The second observation is that any Hölder function A is cohomologuous to $m(A,f)$ on $\Omega(A,f)$, more precisely:

Lemma 13. — *Let A be a $\mathcal{C}^0$-function and assume A admits a $\mathcal{C}^0$ sub-action V, then*

$$\Omega(A,f) \subseteq \Sigma_V(A,f) = \big\{x \in M \mid A - m = V\circ f - V\big\}$$

and any f-invariant measure μ whose support in contained in $\Omega(A,f)$ is maximizing.

The set $\Sigma_V(A,f)$ will play an important role later and it is convenient to to give it a name:

Definition 14. — Let A be a $\mathcal{C}^0$-function and V be a sub-action of A.

(i) We call the set $\Sigma_V(A,f) = \big\{x \in M \mid A - m = V\circ f - V\big\}$, the V-action-set of A.

(ii) Two points x, y of the V-action-set are said to be V-connected and we shall write $x \overset{V}{\to} y$, if for every $\varepsilon > 0$, there exist $n \geqslant 1$ and $z \in M$ (not necessarily in $\Sigma_V(A,f)$) such that

$$x \in B(z,\varepsilon), \quad y \in B(f^n(z),\varepsilon), \quad |S_N(A-m)(z) - (V(y) - V(x))| < \varepsilon.$$

Notice that, if V is β-Hölder for some $\beta > 0$, using the shadowing lemma, one can prove that $x \overset{V}{\to} y$ and $y \overset{V}{\to} u$ imply $x \overset{V}{\to} u$. This is so, because if z_x and n_x are the ones for $x \overset{V}{\to} y$ in (ii) above, and if z_y and n_y are the ones for $y \overset{V}{\to} u$ in (ii) above, then considering the pseudo-orbit $z_x, \dots, f^{n_x}(z_x), z_y, \dots, f^{n_y}(z_y)$, we can find by shadowing the z for $x \overset{V}{\to} u$ in (ii) above.

Proof of Lemma 13. — Define $R = V\circ f - V - A + m$ and choose $x \in \Omega(A,f)$. Then $\sum_{k=0}^{n_i-1}(A-m)\circ f^k(y_i)$ converges to 0 for a sequence of points y_i and a sequence of integers n_i such that y_i converges to x, n_i converges to $+\infty$ and $f^{n_i}(y_i)$ converges to x. Since R is non-negative,

$$0 \leqslant R(y_i) \leqslant \sum_{k=0}^{n_i-1} R\circ f^k(y_i) = V\circ f^{n_i}(y_i) - V(y_i) - \sum_{k=0}^{n_i-1}(A-m)\circ f^k(y_i)$$

converges to 0 and by continuity of R: $R(x) = 0$. □

Definition 15. — For any $\beta > 0$, define

$$\mathcal{G}_\beta = \{A \in \mathcal{C}^\beta(M) \mid \Omega(A,f) \text{ is a periodic orbit}\}.$$

Our next goal is to show that $\mathcal{G}_\beta$ is open in the $\mathcal{C}^\beta$ topology. We could have choosen a bigger set: the set of A in $\mathcal{C}^\beta(M)$ such that $\Omega(A,f)$ is minimal and is dynamically isolated (i.e. there exists U, open, containing $\Omega(A,f)$ as the only f-invariant compact set inside U) and the proof below would again be the same.

Lemma 16. — *For any $\beta > 0$, $\mathcal{G}_\beta$ is open in the $\mathcal{C}^\beta$ topology and $\Omega(A,f)$ is locally constant as a function of A in $\mathcal{G}_\beta$.*

Proof. — Let $A \in \mathcal{G}_\beta$. We want to show that $\Omega(A,f) = \Omega(B,f)$ whenever B is sufficiently close to A in the $\mathcal{C}^\beta$ topology. By contradiction: let U be an isolating open set of the periodic orbit $\Omega(A,f) = \mathrm{orb}(p)$ and $\{A_n\}$ be a sequence of β-Hölder observables converging to A in the $\mathcal{C}^\beta$ topology such that $\Omega(A,f)$ is not included in U for each n.

Each A_n admits (Theorem 1) a γ-Hölder subaction V_n with γ-Hölder norm uniformly bounded and $\gamma = \beta \ln(1/\lambda_s)/\ln(\Lambda_u/\lambda_s)$. By Ascoli, $\{V_n\}$ admits a subsequence converging in the $\mathcal{C}^0$ topology to some γ-Hölder function V. Since the set of non-empty compact sets is compact with respect to the Hausdorff topology, we may assume that $\{\Omega(A_n,f)\}$ has a sub-sequence converging to some compact invariant set K. Each A_n satisfies:

$$A_n - m(A_n,f) \leqslant V_n \circ f - V_n \qquad (\forall\, x \in M),$$
$$A_n - m(A_n,f) = V_n \circ f - V_n \qquad (\forall\ x \in \Omega(A_n,f)).$$

By continuity of $m(A,f)$ with respect to A (for the $\mathcal{C}^0$ topology),

$$A - m(A,f) \leqslant V \circ f - V \qquad (\forall\, x \in M)$$
$$A - m(A,f) = V \circ f - V \qquad (\forall\, x \in K).$$

We have assumed that each $\Omega(A_n,f) \smallsetminus U$ is not empty, then $K \smallsetminus U$ is not empty too. Let $x_0 \in K \smallsetminus U$, the ω-limit set $\omega(x_0)$ and the α-limit set $\alpha(x_0)$ of x_0 are compact invariant sets included in $\Omega(A,f)$, necessarily:

$$\omega(x_0) = \alpha(x_0) = \mathrm{orb}(p) \subset \overline{\mathrm{orb}(x_0)} \subset \Sigma_V(A,f).$$

Since p is V-connected to x_0 and x_0 is V-connected to p, x_0 is V-connected to itself which is equivalent to $x_0 \in \Omega(A,f)$. We just have obtained a contradiction. □

Proof of Theorem 4. — Let β given and A, α-Hölder with:

$$\beta < \widetilde{\beta} = \alpha \frac{\ln(1/\lambda_s)}{\ln(\Lambda_u/\lambda_s)}.$$

According to Theorem 1, there exists V, $\widetilde{\beta}$-Hölder, satisfying:

$$A - m \leqslant V \circ f - V \qquad (\forall\, x \in M).$$

Define $R = V \circ f - V - A + m$, $\phi_n = \min(R, 1/n)$ and $B_n = A + \phi_n$. Then ϕ_n is $\widetilde{\beta}$-Hölder with $\text{Höld}_{\widetilde{\beta}}(\phi_n) \leqslant \text{Höld}_{\widetilde{\beta}}(R)$ and

$$A - m \leqslant B_n - m \leqslant V \circ f - V \qquad (\forall\, x \in M)$$
$$B_n - m = V \circ f - V \qquad (\forall\, x \in \{R < 1/n\}).$$

In particular $m(B_n, f) = m(A, f)$ and the V-action set of B_n contains a neighborhood $\{R < 1/n\}$ of $\Omega(A, f)$. Using the shadowing lemma, we construct a periodic orbit $\text{orb}(p)$ inside $\{R < 1/n\}$ and we just have proved a perturbation B_n of A satisfies

$$\text{orb}(p) \cup \Omega(A, f) \subset \Omega(B_n, f).$$

Let ψ_n be any $\widetilde{\beta}$-Hölder function with small $\widetilde{\beta}$-Hölder norm satisfying:

$$\psi_n(x) = 0 \qquad (\forall\, x \in \text{orb}(p))$$
$$\psi_n(x) > 0 \qquad (\forall\, x \in M \smallsetminus \text{orb}(p)).$$

Then $A_n = B_n - \psi_n = A + \phi_n - \psi_n$ is such that the minimizing measure has support on $\text{orb}(p)$, and A_n has small $\mathcal{C}^0$ norm and (possibly large) uniform $\widetilde{\beta}$-Hölder norm. Therefore (A_n) converges to A in the $\mathcal{C}^\beta$-topology and each A_n has a unique maximizing measure which is supported on a periodic orbit. □

References

[1] G. Atkinson, Recurrence of cocycles and random walks, *J. London Math. Soc.*, **(2)**, **13** (1976) 486–488.

[2] Th. Bousch. Le Poisson n'a pas d'arêtes, *Ann. Inst. Henri Poincaré*, **36** (2000), 489-508.

[3] R. Bowen. Equilibrium states and the ergodic theory of Anosov diffeomorphisms, *Lecture Notes in Mathematics*, vol. **470**. Springer Verlag, Berlin, Heidelberg, New York, (1975).

[4] G. Contreras, A.O. Lopes, Ph. Thieullen. Lyapunov minimizing measures for expanding maps of the circle, *Ergod. Th. & and Dynam. Sys*, **21**, (2001), 1379–1409.

[5] A. Livsic. Some homology properties of Y-systems. *Mathematical Notes of the USSR Academy of Sciences*, **10** (1971), 758–763.

[6] S.V. Savchenko. Cohomological inequalities for finite topological Markov chains. *Funct. Anal. and its Appl.*, **33(3)**, 236–238 (1999)

A.O. Lopes, Instituto de Matemática, UFRGS, Porto Alegre 91501-970, Brasil
E-mail : alopes@mat.ufrgs.br

Ph. Thieullen, Département de Mathématiques, Université Paris-Sud, 91405 Orsay cedex, France
E-mail : Philippe.Thieullen@math.u-psud.fr

Astérisque
287, 2003, p. 147–230

DYNAMIQUE DES FONCTIONS RATIONNELLES SUR DES CORPS LOCAUX

par

Juan Rivera-Letelier

Em homenagem a Jacob, na acasião de seu sexagésimo aniversário.

Résumé. — Soit $p > 1$ un nombre premier, $\mathbb{Q}_p$ le corps des nombres p-adiques et soit $\mathbb{C}_p$ la plus petite extension complète et algébriquement close de $\mathbb{Q}_p$. Ce travail est consacré à l'étude de la dynamique des fonctions rationnelles sur la droite projective $\mathbb{P}(\mathbb{C}_p)$.
À chaque fonction rationnelle $R \in \mathbb{C}_p(z)$ on associe son *domaine de quasi-périodicité*, qui est égal à l'intérieur de l'ensemble des points dans $\mathbb{P}(\mathbb{C}_p)$ qui sont récurrents par R. On donne plusieurs caractérisations du domaine de quasi-périodicité et on décrit sa dynamique locale et globale.
On montre que les composantes du domaine de quasi-périodicité (qui sont les analogues p-adiques des disques des Siegel et des anneaux de Herman) sont des affinoïdes ouverts (c'est-à-dire que leur géométrie est simple) et on décrit la dynamique sur une composante donnée.
Comme dans le cas complexe on a une partition de la droite $\mathbb{P}(\mathbb{C}_p)$ en l'ensemble de Fatou et l'ensemble de Julia. Par analogie au cas complexe on fait la conjecture de non-errance suivante : tout disque errant est attiré par un cycle attractif. On montre que ceci a lieu si et seulement si tout point dans l'ensemble de Fatou est soit attiré par un cycle attractif, soit rencontre le domaine de quasi-périodicité par itération positive.

Introduction

Ce travail est essentiellement ma thèse de doctorat, laquelle a été soutenue à Orsay pendant l'année 2000.

Soit p un nombre premier et $\mathbb{Q}_p$ le corps des nombres p-adiques. Ce travail est consacré à l'étude de la dynamique des fonctions rationnelles à coefficients dans des extensions de $\mathbb{Q}_p$.

***Classification mathématique par sujets* (2000).** — 32H50, 37F10, 14G20, 39B12.
Mots clefs. — Systèmes dynamiques, théorie itérative, corps locaux, quasi-périodique.

Je remercie la Fundación Andes, Beca Presidente de la Republica de Chile et la Fundación Rivera Letelier pour son soutien financier.

Il y a des questions en théorie des nombres qui sont reliées aux systèmes dynamiques et les gens ont étudié, plus ou moins implicitement, des systèmes dynamiques depuis longtemps. Par exemple, la définition de la fraction continue d'un nombre réel est reliée à l'itération de l'application de Gauss $x \to 1/x - [1/x]$. Notamment on a la relation suivante entre la théorie des courbes elliptiques et l'itération des fonctions rationnelles. On considère une courbe elliptique

$$E \,:\, y^2 = x^3 + ax + b,$$

où a et b appartiennent à un corps de nombres K (extension finie des rationnels). Pour toute extension L de K l'ensemble $E(L)$ de points dans E à coordonnées dans L a une structure de groupe abélien, avec $\mathcal{O} = (\infty, \infty)$ comme élément neutre. Cette structure est définie par,

$$(x_0, y_0) + (x_1, y_1) = (\lambda^2 - x_0 - x_1, \lambda(\lambda^2 - x_0 - x_1) + \nu)$$

où $\lambda = (y_1 - y_0)/(x_1 - x_0)$ si $x_1 \neq x_0$ et $\lambda = (3x_0^2 + a)/2y_0$ si $x_1 = x_0$ et $\nu = y_0 - \lambda x_0 = y_1 - \lambda x_1$. Géométriquement la somme de trois points est égal à $\mathcal{O}$ si et seulement si les trois points sont colinéaires. On dit qu'un point de E est de *torsion* s'il existe $n \geqslant 1$ tel que

$$[n]P = \underbrace{P + \cdots + P}_{n} = \mathcal{O}.$$

Ceci a lieu si et seulement si P est pré-périodique pour l'application de doublement $P \to [2]P$, c'est-à-dire qu'il existe $k \geqslant 0$ et $l \geqslant 1$ tels que $[2^k]P = [2^{k+l}]P$. La première coordonnée de l'application de doublement est donnée par

$$x([2](x_0, y_0)) = \frac{x_0^4 - 2ax_0^2 - 8bx_0 + a^2}{4x_0^3 + 4ax_0 + 4b},$$

qui ne dépend que de x_0. On voit alors qu'un point (x, y) de E est un point de torsion si et seulement si x est pré-périodique pour cette fonction rationnelle de degré 4. Ainsi l'étude des points de torsion K-rationnels d'une courbe elliptique est étroitement reliée à l'étude des points pré-périodiques sur K d'une fonction rationnelle.

Il y a beaucoup de résultats dans ce dernier contexte ; on renvoie le lecteur aux références de [**MS2**] et [**Be**]. Northcott a montré une propriété fondamentale :

***Théorème* (Northcott [No])**. — *Soit K un corps de nombres. Alors toute fonction rationnelle à coefficients dans K, de degré au moins deux, a un nombre fini de points pré-périodiques sur la droite projective $\mathbb{P}(K)$.*

Ce théorème vaut en dimension quelconque ; voir aussi Lewis [**Le**]. On donne une nouvelle démonstration de ce théorème (en dimension 1) dans un cadre légèrement plus général (Section 4.5).

Morton et Silverman ont conjecturé qu'il existe une borne pour le nombre de points pré-périodiques qui ne dépend que du degré de $[K : \mathbb{Q}]$ de K et du degré de la fonction rationnelle ; voir [**MS1**].

Principe de Hasse. — Dans ce genre de problème il y a un principe appelé *principe de Hasse*. Hasse a montré le théorème suivant ; voir e.g. [**Ca**].

***Théorème* (Hasse).** — *La forme quadratique,*

$$a_1 x_1^2 + a_2 x_2^2 + \cdots + a_n x_n^2, \quad \textit{où } a_1, \ldots, a_n \in \mathbb{Q},$$

a un zéro non-trivial sur $\mathbb{Q}^n$ si et seulement s'il en a sur $\mathbb{Q}_{\mathcal{P}}^n$ pour toute valeur absolue $\mathcal{P}$, où $\mathbb{Q}_{\mathcal{P}}$ dénote la complétion de $\mathbb{Q}$ par $\mathcal{P}$.

Rappelons que une valeur absolue $\mathcal{P}$ sur un corps K est une fonction $\mathcal{P} : K \to \mathbb{R}$ telle que $\mathcal{P}(x) \geqslant 0$, $\mathcal{P}(x) = 0$ si et seulement si $x = 0$, $\mathcal{P}(xy) = \mathcal{P}(x)\mathcal{P}(y)$ et $\mathcal{P}(x+y) \leqslant \mathcal{P}(x) + \mathcal{P}(y)$, pour tout $x, y \in K$.

L'importance de ce théorème est que souvent il est plus facile de déterminer si une forme quadratique a des zéros non-triviaux sur $\mathbb{Q}_{\mathcal{P}}^n$ que sur $\mathbb{Q}^n$, car $\mathbb{Q}_{\mathcal{P}}$ est complet.

Le principe de Hasse consiste à : d'abord, traiter le problème sur les différentes complétions du corps de base et après, essayer de recoller l'information. Mais on n'a pas toujours une équivalence comme dans le théorème de Hasse. Par exemple l'équation $(x^2 - 2)(x^2 - 17)(x^2 - 34) = 0$ a une solution sur toute complétion de $\mathbb{Q}$, mais n'a pas des solutions sur $\mathbb{Q}$. Il en est de même pour la courbe elliptique $2y^2 = x^4 - 17$; voir [**Ca**].

Dynamique des fonctions rationnelles. — Si l'on veut appliquer le principe de Hasse pour étudier la dynamique d'une fonction rationnelle à coefficients dans un corps de nombres, on doit étudier d'abord la dynamique des fonctions rationnelles sur les différentes complétions du corps de base.

Il y a deux types de valeurs absolues : les *non-archimédiennes* qui satisfont l'inégalité triangulaire forte : $|x+y| \leqslant \max\{|x|, |y|\}$, et les *archimédiennes* qui ne la satisfont pas. D'après un théorème de Orstrowski la complétion d'un corps de nombres par rapport à une valeur absolue archimédienne est isomorphe à $\mathbb{C}$ ou à $\mathbb{R}$; voir par exemple [**Ca**].

Parfois il est plus facile d'étudier la dynamique d'une fonction rationnelle sur $\mathbb{C}$ que sur $\mathbb{R}$ car on profite du fait que $\mathbb{C}$ est algébriquement clos. L'étude de la dynamique d'une fonction rationnelle complexe est classique et elle a été initiée par Fatou et Julia dans les années 1910. Depuis ces premiers travaux, et notamment dans les vingt dernières années, la dynamique des fonctions rationnelles complexes a connu un grand développement, même si des questions fondamentales comme la densité de l'hyperbolicité sont encore ouvertes.

L'étude de la dynamique des fonctions rationnelles sur des corps non-archimédiens est plus récente ; voir [**MS1**], [**Hs**] et [**Be**]. Ce travail est consacré à l'étude de la dynamique d'une fonction rationnelle dont les coefficients appartiennent à $\mathbb{C}_p$, qui est la plus petite extension complète et algébriquement close de $\mathbb{Q}_p$.

Généralités sur le corps $\mathbb{C}_p$. — La *norme p-adique* sur $\mathbb{Q}$, notée $|\cdot|_p$, est définie par $|p^\nu r/s|_p = p^{-\nu}$, pour $\nu \in \mathbb{Z}$ et pour tous les entiers r et s qui ne sont pas divisibles par p. On note $\mathbb{Q}_p$ la complétion de $\mathbb{Q}$ par la norme p-adique et on l'appelle le *corps des nombres p-adiques.* On peut étendre $|\cdot|_p$ à la clôture algébrique $\overline{\mathbb{Q}}$ de $\mathbb{Q}$. La complétion de $\overline{\mathbb{Q}}$ est le corps $\mathbb{C}_p$, qui est aussi algébriquement clos.

Comme p est fixe, on note $|\cdot|_p$ simplement par $|\cdot|$. Une *boule fermée* (resp. *ouverte*) de $\mathbb{C}_p$ est un ensemble de la forme $B_r(a)^+ = \{z \in \mathbb{C}_p \mid |z-a| \leqslant r\}$ (resp. $B_r(a) = \{z \in \mathbb{C}_p \mid |z-a| < r\}$) où $r \in |\mathbb{C}_p| = \{|z| \mid z \in \mathbb{C}_p - \{0\}\}$. Notons que tout élément $w \in B_r^+(a)$ (resp. $w \in B_r(a)$) est un centre ; $B_r^+(w) = B_r^+(a)$ (resp. $B_r(w) = B_r(a)$) et si deux boules fermées (resp. ouvertes) s'intersectent, alors l'une est contenue dans l'autre. Dans la droite projective on appelle boule fermée (resp. ouverte) le complémentaire d'une boule ouverte (resp. fermée) de $\mathbb{C}_p$. La notion de boule de $\mathbb{P}(\mathbb{C}_p)$ est alors invariante par automorphismes projectifs.

Une propriété des fonctions rationnelles à coefficients dans $\mathbb{C}_p$ qui simplifiera la géométrie est que les fonctions rationnelles préservent les affinoïdes. Un *affinoïde fermé* (resp. *ouvert*) *connexe* est une intersection finie non-vide de boules fermées (resp. ouvertes) de $\mathbb{P}(\mathbb{C}_p)$ et un *affinoïde fermé* (resp. *ouvert*) est une union finie d'affinoïdes fermés (resp. ouvertes) connexes. Alors l'image et la préimage d'un affinoïde fermé (resp. ouvert) par une fonction rationnelle est aussi un affinoïde fermé (resp. ouvert) (Proposition 2.6). Une union d'affinoïdes connexes fermés qui contient un point donné est appelé un *espace analytique connexe.* Alors la classe des espaces analytiques, qui sont les unions finies de espaces analytiques connexes, est aussi invariante par les fonctions rationnelles.

On prendra cette notion pour définir une notion de composante connexe : la *composante analytique* d'une partie ouverte X de $\mathbb{P}(\mathbb{C}_p)$ qui contient $z_0 \in X$ est l'union de tous les espaces analytiques connexes qui contient z_0 et qui sont contenus dans X ; voir [**Be**]. Notons qu'une composante analytique est un espace analytique connexe non-vide.

Les *fonctions holomorphes* définies sur les affinoïdes fermés connexes sont les limites uniformes de fonctions rationnelles sans pôles sur l'affinoïde. Une telle fonction est constante ou a un nombre fini de zéros. Alors une fonction définie sur un espace analytique est *holomorphe* si sa restriction à tout affinoïde fermé connexe est holomorphe ; voir [**FvP**] et Sections 1.2 et 1.3.3.

Points périodiques. — Rappelons que si z_0 est un point périodique de R de période minimale k alors $\{z_0, R(z_0), \ldots, R^{k-1}(z_0)\}$ est le *cycle* de z_0 ; $\lambda = (R^k)'(z_0)$ ne dépend que du cycle et est appelé le *multiplicateur* du cycle. On considère la classification suivante des points périodiques, en analogie avec le cas complexe.

• Si $|\lambda| < 1$ on dit que z_0 est *attractif.* Si de plus $\lambda = 0$, alors on dit que z_0 est *super-attractif.*

• Si $|\lambda| = 1$ on dit que z_0 est *indifférent*. Si λ est une racine de l'unité on dit que z_0 est *parabolique* ou *indifférent rationnel*. Sinon on dit que z_0 est *indifférent irrationnel*.

• Si $|\lambda| > 1$ on dit que z_0 est *répulsif*.

Un théorème de Fatou dit qu'une fonction rationnelle à coefficients complexes, de degré au moins deux, est telle que tout cycle attractif ou parabolique attire au moins un point critique ; voir [**Fa**]. Comme une fonction rationnelle R a au plus $2\deg(R)-2$ points critiques on a la propriété suivante.

***Théorème* (Fatou [Fa])**. — *Soit $R \in \mathbb{C}(z)$ une fonction rationnelle de degré au moins deux. Alors le nombre de cycles attractifs ou paraboliques de R est au plus* $2\deg(R)-2$.

Cette propriété est loin d'être vraie dans le cas p-adique. Par exemple tous les points périodiques du polynôme $z^p \in \mathbb{C}_p[z]$ sont attractifs. Mais on montre que *si une fonction rationnelle à coefficients dans $\mathbb{C}_p$ et degré $d > 1$, a plus de $3d-3$ cycles attractifs, alors elle en a une infinité* (Corollaire 4.9). On conjecture que pour une telle fonction rationnelle, soit tout cycle attractif attire un point critique (et alors il y a au plus $2d-2$ cycles attractifs), soit il y a une infinité de cycles attractifs, voir Section 4.1.1.

D'autre part, notons que la notion de cycle super-attractif et cycle parabolique ne dépend pas d'une valuation. On obtient alors le théorème suivant par le principe de Lefschetz.

Théorème 1. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré $d \geqslant 2$. Alors le nombre de cycles super-attractifs et paraboliques est au plus $2d-2$.*

Dynamique au voisinage des points fixes (Sections 3.1 et 3.3). — Comme dans le cas complexe on peut décrire, dans la plupart des cas, la dynamique au voisinage d'un point fixe à partir de son multiplicateur. Une des différences avec le cas complexe est que, comme $\{|\lambda| = 1\} \subset \mathbb{C}_p$ est un ensemble ouvert, la condition d'avoir un point fixe indifférent est ouverte. Par conséquent il n'y a pas d'analogie non-archimédienne aux très riches phénomènes liés à la bifurcation d'un point fixe indifférent, comme par exemple l'implosion parabolique ; voir [**La**].

En général la dynamique locale des points fixes est assez simple et en particulier la dynamique au voisinage d'un point fixe qui n'est ni parabolique ni super-attractif est localement linéarisable. Il y a aussi des formes locales canoniques pour les points fixes paraboliques et super-attractifs. De plus, deux germes sont conjugués si et seulement s'ils le sont formellement ; voir [**HY**], [**Lu1**], [**TVW**], [**AV**] et Sections 3.1 et 3.3.1. Les mêmes assertions dans le cas complexe ne sont pas toujours vraies et sont reliées à des problèmes de petits diviseurs assez subtils ; voir [**Y**], [**Ec1**] et [**Vo**].

Ensembles de Fatou et de Julia (Section 4.3). — On peut définir les ensembles de Fatou et de Julia comme dans le cas complexe, mais on remplace la normalité du complexe par l'*uniformité lipschitzienne*.

L'*ensemble de Fatou* d'une fonction rationnelle $R \in \mathbb{C}_p(z)$, noté $F(R)$, est l'ensemble de tous les points $z_0 \in \mathbb{P}(\mathbb{C}_p)$ tels qu'il existe un voisinage U de z_0 tel que la famille $\{R^n|_U\}_{n\geqslant 0}$ est uniformément lipschitzienne ; voir [**Hs**].

On a les propriétés usuelles : $F(R)$ est ouvert, $R^{-1}(F(R)) = F(R)$ et $F(R^n) = F(R)$, pour $n \geqslant 1$. Mais l'*ensemble de Julia* $J(R) = \mathbb{P}(\mathbb{C}_p) - F(R)$ n'est pas nécessairement compact (rappelons que $\mathbb{C}_p$ n'est pas localement compact). De plus l'ensemble de Julia peut être vide (par exemple il n'est pas difficile de voir que $J(z^p) = \varnothing$). On a aussi propriété suivante.

***Théorème* (Benedetto [Be])**. — *L'ensemble de Fatou d'une fonction rationnelle est non-vide.*

Dans le cas complexe on a au contraire : l'ensemble de Julia est non-vide et peut être égal à $\mathbb{P}(\mathbb{C})$. Comme conséquence de ce théorème (et d'un théorème de Hsia), on obtient que l'ensemble de Julia est toujours d'intérieur vide.

Cas complexe. — Dans le cas complexe l'étude de la dynamique de l'ensemble de Fatou, est faite en étudiant les composantes connexes, car il n'est pas difficile de voir que la fonction rationnelle envoie une composante connexe sur une composante connexe. Fatou et Julia ont classifié les composantes connexes en trois types (voir par exemple [**CG**]) :

- *Composantes attractives.* Ce sont les composantes connexes qui sont attirées par un cycle, qui peut être attractif ou parabolique.
- *Préimages des composantes de quasi-périodicité.* C'est-à-dire qu'il existe $k \geqslant 0$ et une suite $n_j \to \infty$ telle que R^{n_j+k} converge uniformément vers R^k sur les parties compactes de la composante en question. Dans ce cas, un itéré de la composante est périodique pour R et peut être, soit un disque appelé *disque de Siegel*, soit un anneau appelé *anneau de Herman*.
- *Composantes errantes.* Ce sont les composantes qui ne sont pas pré-périodiques pour la fonction rationnelle.

En 1985 Sullivan a montré que toutes les composantes connexes de l'ensemble de Fatou sont pré-périodiques et par conséquent, il ne peut y avoir que des composantes attractives ou des préimages des composantes de quasi-périodicité ; voir [**Sul**].

Sur la structure de l'ensemble de Fatou (Section 4.4). — Dans le cas p-adique à chaque fonction rationnelle $R \in \mathbb{C}_p(z)$ on associe son *domaine de quasi-périodicité*, que l'on note $\mathcal{E}(R)$, qui est égal à l'intérieur de l'ensemble de points de $\mathbb{P}(\mathbb{C}_p)$ qui sont récurrents par R. On montre le théorème suivant.

Théorème de Classification. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Alors l'ensemble de Fatou de R se décompose dans les ensembles disjoints suivants.*

(1) *Bassins d'attraction. (L'ensemble de points de $\mathbb{P}(\mathbb{C}_p)$ qui sont attirés par un cycle attractif.)*

(2) $\mathcal{E}'(R) = \cup_{n\geqslant 0} R^{-n}(\mathcal{E}(R))$.
(3) *L'union des disques errants qui ne sont pas attirés par un cycle attractif.*

On dit qu'un disque $D \subset \mathbb{P}(\mathbb{C}_p)$ est *errant* si pour tous $k > l \geqslant 0$ on a $R^k(D) \cap R^l(D) = \varnothing$. Par analogie au cas complexe on fait la conjecture suivante.

Conjecture de Non-Errance. — *Tout disque errant est attiré par un cycle attractif.*

D'après le Théorème de Classification la Conjecture de Non-Errance est équivalente à la conjecture suivante.

Structure Conjecturale de l'Ensemble de Fatou. — *Tout point de l'ensemble de Fatou appartient à $\mathcal{E}'(R)$ ou est attiré par un cycle attractif.*

Benedetto a étudié les composantes analytiques de l'ensemble de Fatou et il a fait une conjecture de non-errance dans ce contexte ; voir [**Be**]. Cependant il arrive que l'ensemble de Fatou soit égal à $\mathbb{P}(\mathbb{C}_p)$, même s'il y a plusieurs comportements attractifs et quasi-périodiques ; voir exemples 6.6 et 6.3.

Dynamique attractive (Section 4.1). — L'étude des bassins d'attraction des cycles attractifs peut être faite comme dans le cas complexe et ne présente pas de surprises. Une fonction rationnelle envoie une composante analytique d'un bassin d'attraction sur une composante analytique. De plus un *bassin d'attraction immédiat* (c'est-à-dire une composante analytique du bassin qui contient un point du cycle) est soit un disque rationnel ouvert ; soit un ensemble qui ressemble au complémentaire d'un ensemble de Cantor dans $\mathbb{P}(\mathbb{C}_p)$ (Théorème 2).

Dynamique quasi-périodique (Section 4.2). — Par analogie au cas complexe on définit le *domaine de quasi-périodicité* d'une fonction rationnelle $R \in \mathbb{C}_p(z)$ comme

$$\mathcal{E}(R) = \{z_0 \in U \mid \text{il existe } n_j \to \infty \text{ tel que } R^{n_j} \to \mathrm{id}, \text{ sur un voisinage de } z_0\}.$$

En particulier $\mathcal{E}(R)$ est ouvert et tout point dans $\mathcal{E}(R)$ est récurrent par R. On montre que $\mathcal{E}(R)$ est égal à l'intérieur de l'ensemble de points récurrents par R (Corollaire 4.27). Une définition équivalente (Proposition 3.14) est :

$$\mathcal{E}(R) = \{z_0 \in X \mid \text{il existe } k = k(z_0) \geqslant 1 \text{ tel que } \left\{\frac{R^{nk} - \mathrm{id}}{nk}\right\}_{n\geqslant 0}$$
$$\text{est uniformément convergent sur un voisinage de } z_0, \text{ quand } |n|_p \to 0\}.$$

Il n'est pas difficile de voir que R est injective sur $\mathcal{E}(R)$, $R(\mathcal{E}(R)) = \mathcal{E}(R)$ et $\mathcal{E}(R^n) = \mathcal{E}(R)$, pour $n \geqslant 1$ (Proposition 3.9). En général $\mathcal{E}(R)$ n'est pas fermé ; voir exemple dans [**R-L**].

On peut décrire la dynamique locale sur le domaine de quasi-périodicité comme suit : soit $z_0 \in \mathcal{E}(R)$ n'est pas périodique et alors il existe un entier $k \geqslant 1$ tel que R^k est localement conjugué à la translation $z \mapsto z + k$ (Proposition 3.16) ; soit $z_0 \in \mathcal{E}(R)$ est périodique et par conséquent R est localement linéarisable si z_0 n'est pas parabolique ;

sinon R est localement de la forme $\lambda z(1 + z^n + az^{2n})$, où λ est une racine de l'unité, $n \geqslant 1$ et $a \in \mathbb{C}_p$; (Section 3.3.1).

Logarithme itératif (Section 3.2). — Pour montrer ces propriétés on considère la fonction R_* définie au voisinage d'un point $z_0 \in \mathcal{E}(R)$ par,

$$R_* = \lim_{|n|_p \to 0} \frac{R^{nk} - \mathrm{id}}{nk}, \quad \text{où } k = k(z_0),$$

qui est appelé le *logarithme itératif* de R. Cette fonction a aussi été considérée par Lubin dans [**Lu1**] au voisinage des points fixes. Par définition R_* est localement holomorphe et on montre que $R_*(z_0) = 0$ si et seulement si $z_0 \in \mathcal{E}(R)$ est périodique par R ; voir aussi [**Lu1**].

Ceci nous donne deux corollaires intéressants. Le premier est que les points périodiques indifférents sont isolés, ce qui n'était pas tellement clair ; voir [**Hs**] et [**Be**]. L'autre corollaire est une nouvelle démonstration du théorème de Northcott unidimensionnel (Section 4). On peut aussi utiliser le logarithme itératif pour expliciter les conjugaisons locales décrites dans la Section 3.3.1.

Composantes du domaine de quasi-périodicité (Section 5). — On montre qu'une fonction rationnelle permute les composantes analytiques du domaine de quasi-périodicité et on montre la propriété fondamentale suivante.

Théorème 3. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins deux et C une composante analytique de $\mathcal{E}(R)$. Alors C est un affinoïde ouvert, c'est-à-dire :*

$$C = \mathbb{P}(\mathbb{C}_p) - B_0 \cup \cdots \cup B_n,$$

où $n \geqslant 0$ et $B_0, \ldots, B_n$ sont des boules fermées.

Par analogie avec le cas complexe on dit que C est un *disque de Siegel* si $n = 0$ et on dit que C est un *n-anneau de Herman* si $n > 0$ (on suppose que les boules $B_0, \ldots, B_n$ sont disjointes).

On montre que pour tout affinoïde ouvert il existe une fonction rationnelle telle que chaque composante connexe de l'affinoïde est une composante analytique de son domaine de quasi-périodicité.

De plus on montre que pour tout $n \geqslant 1$ il existe une fonction rationnelle de degré 2 ayant un n-anneau de Herman. En revanche une fonction rationnelle complexe de degré deux ne peut pas avoir un anneau de Herman ; voir [**Sh**].

Itérés fractionnaires (Section 4.2). — Étant donnée une composante analytique C de $\mathcal{E}(R)$, et quitte à remplacer R par un itéré, on a une action $(w, z) \mapsto R^{\circ w}(z)$ de $\mathbb{Z}_p = \{w \in \mathbb{Q}_p \mid |z|_p \leqslant 1\}$ sur C telle que $R^{\circ 1} = R$ et telle que pour chaque $w \in \mathbb{Z}_p$ l'application $R^{\circ w} : C \to C$ est un automorphisme holomorphe de C (Proposition 5.6).

En effet pour chaque affinoïde fermé contenu dans C on a que pour chaque $w \in \mathbb{C}_p$ de norme assez petite la série

$$R^{\circ w} = T_0 + wT_1 + \frac{w(w-1)}{2}T_2 + \cdots ,$$

converge uniformément sur l'affinoïde et on a $R^{\circ n} = R^n$, pour chaque entier $n \geqslant 1$, et on a

$$R^{\circ(w_1+w_2)} = R^{\circ w_1} \circ R^{\circ w_2},$$

quand $R^{\circ(w_1+w_2)}$, $R^{\circ w_1}$ et $R^{\circ w_2}$ sont définis. Par conséquent $R_* = \lim_{w\to 0}(R^{\circ w} - \mathrm{id})/w$ et alors on peut considérer le logarithme itératif R_* comme un champ de vecteurs tel que le temps w du flot engendré par R_* est $R^{\circ w}$.

Composantes et points périodiques (Section 5.2). — On obtient que chaque composante analytique du domaine de quasi-périodicité contient une infinité de points périodiques, mais chaque affinoïde fermé contenu dans la composante rencontre au plus un nombre fini d'eux. De plus, quitte à prendre un itéré, les périodes de ces points périodiques sont de la forme p^n, pour $n \geqslant 0$.

On montre ensuite qu'à chaque bout d'une composante analytique du domaine de quasi-périodicité sont attachés une infinité de disques de Siegel (Section 5.3). Par conséquent si $R \in \mathbb{C}_p(z)$ est une fonction rationnelle de degré au moins deux, alors les affirmations suivantes sont équivalentes.

– *R a un point périodique indifférent.*
– *R a une infinité de points périodiques indifférents.*
– *$\mathcal{E}(R) \neq \varnothing$.*
– *Il y a une infinité de disques de Siegel.*

Un exemple. — Un exemple intéressant est la fonction rationnelle

$$R(z) = \lambda z^2 \frac{1-bz}{z-a},$$

où $|\lambda| = 1$ et $|a| = |b| < 1$ (exemple 5.15). La couronne

$$C = \{z \in \mathbb{C}_p \mid |a| < |z| < |b|^{-1}\}$$

est une composante analytique de $\mathcal{E}(R)$ et si $|\lambda - 1| < 1$ alors tous les points périodiques de R sur C sont de période de la forme p^n. De plus pour $p > 2$ on montre (à l'aide d'un théorème de Keating) que si R n'a pas de cycles paraboliques, alors pour tout $n \geqslant 0$, la fonction rationnelle R a exactement 2 cycles de période primitive p^n dans C.

Sur la rectification des applications à allure polynomiale. — On finit avec une remarque. Dans le complexe la théorie des applications quasi-conformes est un outil très utile. Notamment, la preuve du théorème de non-errance de Sullivan dépend de façon essentielle de cette théorie ; voir [**Sul**]. Une autre application est le *théorème de rectification* de Douady et Hubbard qui dit que toute application à allure polynomiale est quasi-conformément conjuguée à un polynôme ; voir [**DH**].

Il semble difficile qu'on puisse avoir un analogue du théorème de rectification de Douady et Hubbard. Par exemple soit $i \in \mathbb{C}_5$ la racine de $z^2 = -1$ telle que $|i-2| < 1$, et considérons le polynôme

$$P(z) = -z + (5+i)z^2 - (i+3)z^3 \in \mathbb{C}_5[z].$$

Alors $(P, \{|z| < 5\})$ est une application à allure polynomiale de degré 2 qui a 0 et 1 comme points fixes paraboliques, car $P'(0) = -1$ et $P'(1) = -i$. Mais par le Théorème 1, un polynôme de degré 2 a au plus un cycle parabolique.

Organisation et remarques. — Ce travail est divisé en 6 sections. Souvent la démonstration la plus longue d'une (sous)section est à la fin de la section. Maintenant on décrit brièvement le contenu de chaque section.

La *Section* 1 contient les préliminaires. On considère d'abord les objets géométriques avec leurs propriétés basiques (*Section* 1.2). A part ceux mentionnés dans cette introduction on définit les *bouts*, les *systèmes projectifs* et l'arbre associé à un espace analytique connexe. La *Section* 1.3 contient des outils d'analyse.

Dans la *Section* 2 on considère l'action d'une fonction rationnelle sur les différents objets géométriques. On définit aussi la notion de *composante d'injectivité* et on montre que une composante d'injectivité est un affinoïde ouvert (Proposition 2.9). Ceci est le premier pas dans la démonstration du Théorème 3.

La *Section* 3 est consacrée à la dynamique locale. Cette section ne dépend pas de la Section 2. Dans les *Sections* 3.1 et 3.3 on décrit la dynamique au voisinage des points fixes et dans la *Section* 3.2 on introduit le *domaine de quasi-périodicité* et on décrit ses propriétés locales.

La *Section* 4 est consacrée à la dynamique des fonctions rationnelles. On commence par le Théorème 1 sur la borne du nombre de cycles super-attractifs et paraboliques. Dans la *Sections* 4.1 on étudie les bassins d'attraction et dans la *Section* 4.2 on considère des propriétés globales du domaine de quasi-périodicité. Dans la *Section* 4.3 on considère les ensembles de Fatou et de Julia avec leurs propriétés basiques. Dans la Section 4.4 on montre le Théorème de Classification et on énonce les conjectures décrites dans cette introduction. Dans la *Section* 4.5 on considère la dynamique de fonctions rationnelles dites *simples*, qui ont été introduites par Morton et Silverman dans [**MS2**].

Dans la *Section* 5 on étudie les composantes analytiques du domaine de quasi-périodicité. On montre le Théorème 3 sur la géométrie des composantes analytiques et dans la *Section* 5.1 on considère la dynamique sur une composante analytique donnée. Dans la *Section* 5.2 on étudie les points périodiques dans une composante donnée et dans la *Section* 5.3 on étudie la dynamique aux bouts des composantes analytiques.

Dans la *Section* 6 on considère des exemples.

Remerciements. — Je remercie J.-C. Yoccoz pour les nombreuses discussions et pour l'éclaircissement de plusieurs idées, ainsi que pour la patience avec laquelle il a lu les différentes versions préliminaires de ce travail et pour les nombreuses corrections et améliorations qu'il a apportées. Je remercie aussi S. Crovisier, S. Senti et I. de la Cruz qui ont fait des nombreux corrections d'orthographe pour ce travail.

1. Préliminaires

1.1. Généralités. — On note $\mathbb{Z}$ l'anneau des nombres entiers, $\mathbb{Q}$ le corps des nombres rationnels, $\mathbb{R}$ le corps des nombres réels, $\mathbb{C}$ le corps des nombres complexes et, si $q \geqslant 1$ est une puissance d'un nombre premier, on note $\mathbb{F}_q$ l'unique corps avec q éléments. On appelle *corps de nombres* une extension finie de $\mathbb{Q}$. Étant donné un corps K on note $\overline{K}$ la clôture algébrique de K.

Un *corps valué* $(K, |\cdot|)$ est un corps K muni d'une fonction $|\cdot| : K \to \mathbb{R}$ qui vérifie les propriétés suivantes.

(i) $|x| \geqslant 0$, pour tout $x \in K$.
(ii) $|x| = 0$ si et seulement si $x = 0$.
(iii) $|xy| = |x| \cdot |y|$, pour tous $x, y \in K$.
(iv) $|x + y| \leqslant |x| + |y|$, pour tous $x, y \in K$.

On dit que $(K, |\cdot|)$ est *ultramétrique*, ou *non-archimédien*, si de plus on a

(iv)' $|x + y| \leqslant \max\{|x|, |y|\}$, pour tous $x, y \in K$.

Si K est complet pour la distance induite par la valeur absolue $|\cdot|$, on dit que $(K, |\cdot|)$ est *complet.* Le groupe $|K| = \{|x| \mid x \in K - \{0\}\}$ s'appelle le *groupe de valuation* de K. Si K est algébriquement clos, alors $|K|$ est dense dans $\mathbb{R}$.

Lorsque K est non-archimédien, l'ensemble $\mathcal{O}_K = \{x \in K \mid |x| \leqslant 1\}$ est un anneau appelé *l'anneau de valuation* de K et l'idéal $\mathcal{P}_K = \{x \in K \mid |x| < 1\}$ est appelé *l'idéal de valuation* de K ; c'est un idéal maximal de $\mathcal{O}_K$. Alors $\widetilde{K} = \mathcal{O}_K/\mathcal{P}_K$ est un corps, qui est appelé le *corps résiduel* de K ; pour $x \in \mathcal{O}_K$, on note $\widetilde{x} \in \widetilde{K}$ l'image de x dans $\widetilde{K}$.

Étant donné un nombre premier $p > 1$ on considère la norme *p-adique* $|\cdot|_p$ sur $\mathbb{Q}$ définie par $|p^\nu r/s|_p = p^{-\nu}$, pour tout $\nu \in \mathbb{Z}$ et tous les entiers r et s qui ne sont pas divisibles par p. On note $\mathbb{Q}_p$ la complétion de $\mathbb{Q}$ par $|\cdot|_p$, qu'on appelle le *corps de nombres p-adiques.* Alors $(\overline{\mathbb{Q}}_p, |\cdot|_p)$ n'est pas complet, on note $\mathbb{C}_p$ sa complétion, qui est algébriquement close. On note $\mathbb{D}_p = \mathcal{P}_{\mathbb{C}_p} = \{z \in \mathbb{C}_p \mid |z| < 1\}$.

Notons que $\mathbb{F}_p$ est le corps résiduel de $\mathbb{Q}_p$ et que $\overline{\mathbb{F}}_p$ est celui de $\overline{\mathbb{Q}}_p$ et $\mathbb{C}_p$. De plus $\log_p |\mathbb{C}_p| = \mathbb{Q}$.

1.1.1. La droite projective et la métrique chordale. — Étant donné un corps K on considère la *droite projective* de K notée $\mathbb{P}(K)$, qui est l'ensemble des droites dans $K^2 = K \times K$ passant par $(0, 0)$. On notera $[x, y] \in \mathbb{P}(K)$, pour $(x, y) \in K^2 - \{0, 0\}$, le point correspondant à la droite $\{(\lambda x, \lambda y) \mid \lambda \in K\}$. Le groupe d'automorphismes de $\mathbb{P}(K)$ est isomorphe à $GL(2,K)/K^*$, où $K^* = K - \{0\}$. En effet, à $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2,K)/K^*$ on associe l'automorphisme

$$\varphi([x,y]) = [ax+by, cx+dy].$$

On identifie $\mathbb{P}(K) - [1,0]$ avec K par l'application $[\lambda,1] \mapsto \lambda$ et on note ∞ le point $[1,0]$. Alors on a une identification naturelle $\mathbb{P}(K) = K \cup \{\infty\}$. On étend l'application $x \in \mathcal{O}_K \mapsto \widetilde{x} \in \widetilde{K}$ à $\mathbb{P}(K)$, par $\widetilde{x} = \infty \in \mathbb{P}(\widetilde{K})$ pour tout $x \in \mathbb{P}(K) - \mathcal{O}_K$.

Une *coordonnée*, ou plus précisément, une *coordonnée globale* w de $\mathbb{P}(K)$ est $w \in GL(2,K)/K^*$ et l'identification de $\mathbb{P}(K) - \{w^{-1}([1,0])\}$ à K donnée par $[x,y] \mapsto x_1/y_1$ où $[x_1,y_1] = w([x,y])$. On dit qu'une coordonnée w est *affine* si w est représentée par une matrice de la forme $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right) \in GL(2,K)$.

Si K est un corps valué ultramétrique, alors on considère, comme dans [**Ru1**], la *distance chordale* d sur $\mathbb{P}(K)$ donnée par

$$d([x_0,y_0],[x_1,y_1]) = \frac{|x_0y_1 - x_1y_0|}{\max\{|x_0|,|y_0|\} \cdot \max\{|x_1|,|y_1|\}},$$

ou en coordonnées

$$d(z_0,z_1) = \frac{|z_0 - z_1|}{\max\{|z_0|,1\} \cdot \max\{|z_1|,1\}},$$

où $|\cdot|$ désigne la norme ultramétrique sur K correspondante ; voir aussi [**Ru2**] et [**MS2**]. Notons que la distance chordale coïncide avec la distance induite par $|\cdot|$ sur $\{|z| \leqslant 1\}$. De plus il est facile de voir que la distance chordale est invariante par les automorphismes représentés par les matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2,\mathcal{O}_K)$ telles que $|ad - bc| = 1$.

On peut considérer la distance chordale comme l'analogue de la distance sur $\mathbb{P}(\mathbb{C}) \equiv \mathbb{C} \cup \{\infty\}$ que l'on obtient quand on identifie $\mathbb{C} \cup \{\infty\}$ à $\{x \in \mathbb{R}^3 \mid |x| = 1\}$ par la projection stéréographique et en considérant la distance sur $\{x \in \mathbb{R}^3 \mid |x| = 1\}$ induite par celle de $\mathbb{R}^3$.

1.1.2. Corps ultramétriques localement compacts. — On considère un corps valué $(K, |\cdot|)$ ultramétrique et complet.

Proposition. — *Le corps K est localement compact si et seulement si le groupe des valuations est discret et le corps résiduel est fini.*

Dans ce cas il existe $\pi \in \mathcal{P}_K$ tel que pour tout système de représentants A de $\widetilde{K}$ et tout $x \in K - \{0\}$, il existe $N \in \mathbb{Z}$ et $\{a_n\}_{n \geqslant N} \subset A$ tels que,

$$x = \sum_{n=N}^{\infty} a_n \pi^n \text{ avec } a_N \neq 0.$$

On appelle π une *uniformisante* de $(K, |\cdot|)$.

Par exemple $\mathbb{Q}_p$ est localement compact et on peut prendre $\pi = p$ comme uniformisante et $A = \{0, 1, \ldots, p-1\}$ comme système des représentants de $\widetilde{\mathbb{Q}}_p = \mathbb{F}_p$.

Il est facile de voir que toute extension finie de $\mathbb{Q}_p$ est complète et localement compacte. D'autre part $\mathbb{C}_p$ n'est pas localement compact, car son corps résiduel est infini (et son groupe des valuations n'est pas discret).

1.2. Objets géométriques. — Dans cette section on introduit différents objets géométriques et on montre les propriétés qui nous seront utiles dans la suite.

1.2.1. Boules, affinoïdes et espaces analytiques. — Étant donné $x \in \mathbb{C}_p$ et $r > 0$ on pose

$$B_r(x) = \{z \in \mathbb{C}_p \mid |z - x| < r\} \quad \text{et} \quad B_r^+(x) = \{z \in \mathbb{C}_p \mid |z - x| \leqslant r\}.$$

Si $r \in |\mathbb{C}_p|$ on appelle $B_r^+(x)$ (resp. $B_r(x)$) *boule fermée* (resp. *ouverte*) de $\mathbb{C}_p$. Si $r \notin |\mathbb{C}_p|$ alors $B_r(x) = B_r^+(x)$ et on l'appelle *boule irrationnelle* de $\mathbb{C}_p$.

Une boule fermée (resp. ouverte, irrationnelle) de $\mathbb{P}(\mathbb{C}_p)$ est soit une boule fermée (resp. ouverte, irrationnelle) de $\mathbb{C}_p$, soit le complémentaire dans $\mathbb{P}(\mathbb{C}_p)$ d'une boule ouverte (resp. fermée, irrationnelle) de $\mathbb{C}_p$. La classe des boules fermées (resp. ouvertes, irrationnelles) de $\mathbb{C}_p$ (resp. $\mathbb{P}(\mathbb{C}_p)$) est invariante par changement de coordonnée affine (resp. projectif).

Si B_0, B_1 sont des boules fermées, ouvertes ou irrationnelles de $\mathbb{C}_p$, alors $B_0 \cap B_1 = \varnothing$, $B_0 \subset B_1$ ou $B_1 \subset B_0$. Une boule fermée (ouverte, irrationnelle) est topologiquement ouverte et fermée, donc la notion de boule ouverte et fermée n'est pas topologique. On appelle une boule fermée (resp. ouverte) simplement *boule* (resp. *disque*).

Un *affinoïde fermé* (resp. *ouvert*) *connexe* est une intersection finie non-vide de boules fermées (resp. ouvertes). Un *affinoïde fermé* (resp. *ouvert*) est une union finie d'affinoïdes connexes fermés (resp. ouverts). Alors $\mathbb{P}(\mathbb{C}_p)$ est un affinoïde fermé (resp. ouvert) connexe, car c'est l'intersection vide de boules fermées (resp. ouvertes).

Une intersection ou une union finie d'affinoïdes fermés (resp. ouverts) est un affinoïde fermé (resp. ouvert). Une intersection finie non-vide d'affinoïdes fermés (resp. ouverts) connexes est un affinoïde fermé (resp. ouvert) connexe. L'union de deux affinoïdes fermés (resp. ouverts) connexe dont l'intersection est non-vide est un affinoïde fermé (resp. ouvert) connexe. Le complémentaire d'un affinoïde fermé (resp. ouvert) est un affinoïde ouvert (resp. fermé).

Proposition 1.1. — *Soit $X = X_1 \cup X_2 \cup \cdots \cup X_n$ où les X_i sont des affinoïdes connexes, et soit $x \in X$. L'union des affinoïdes connexes contenant x et contenus dans X est un affinoïde connexe qui contient les X_i qu'il rencontre. On l'appelle la* composante connexe *de x dans X. X est l'union disjointe des ses composantes connexes.*

Démonstration. — Soit $J \subset \{1, \ldots, n\}$ l'ensemble des j tels qu'il existe un affinoïde connexe contenu dans X, contenant x et rencontrant X_j. Posons $Y = \cup_J X_j$. Soit Z un affinoïde connexe contenant x et contenu dans X ; si Z rencontre un X_k, on a $k \in J$; on conclut que $Z \subset Y$.

Inversement, soit $j \in J$; notons Y_j un affinoïde connexe contenu dans X contenant x, et rencontrant X_j, et posons $Z_j = Y_j \cup X_j$: c'est à nouveau un affinoïde connexe contenant x et contenu dans X, et c'est aussi le cas pour $Z := \cup_J Z_j$. On a donc $Y = \cup_J X_j \subset \cup_J Z_j \subset Z$ mais aussi $Z \subset Y$ d'après ci-dessus, donc Z est le plus grand affinoïde connexe contenant x et contenu dans X. □

Un *espace analytique connexe* est une union croissante d'affinoïdes connexes. Les unions (finies ou infinies) d'affinoïdes connexes contenant un même point sont des espaces analytiques connexes. Un *espace analytique* est une union finie d'espaces analytiques connexes. Les affinoïdes fermés et ouverts sont des espaces analytiques ; c'est aussi le cas du complémentaire d'un ensemble compact dans un espace analytique. Les espaces analytiques sont les domaines des fonctions holomorphes ; voir Section 1.3 et [**FvP**].

Suivant Benedetto on appelle *composante analytique* d'une partie ouverte X de $\mathbb{P}(\mathbb{C}_p)$ qui contient $x \in X$, l'union de tous les espaces analytiques connexes contenus dans X et contenant x ; voir [**Be**]. Une composante analytique est alors un espace analytique connexe non-vide.

Une *couronne* $C \subset \mathbb{P}(\mathbb{C}_p)$ est un ensemble qui après changement de coordonnée approprié, est de la forme

$$C(I) = \{z \in \mathbb{C}_p \mid |z| \in I\},$$

où $I \subset (0, +\infty)$ est un intervalle non-vide. Si l'intervalle I est ouvert, alors on dit que C est une *couronne ouverte* ou simplement *couronne* quand il est clair qu'il s'agit d'une couronne ouverte.

De façon équivalente une couronne ouverte est un ensemble de la forme $\mathbb{P}(\mathbb{C}_p) - B_0 \sqcup B_1$, où B_0 et B_1 sont des ensembles disjoints qui sont des points, des boules fermées ou des boules irrationnelles, qu'on appelle *composantes* de $\mathbb{P}(\mathbb{C}_p) - C$. On appelle

$$\mathrm{mod}(C) = \mathrm{mod}(C(I)) = \log_p |I| \in (0, \infty]$$

le *module* de C. Les couronnes sont des espaces analytiques connexes.

1.2.2. Bouts et systèmes projectifs

Définition 1.2. — Soit B une boule fermée de $\mathbb{P}(\mathbb{C}_p)$. Une *chaîne évanescente* associée à B est une suite décroissante $\{A_i\}_{i \geqslant 1}$ de couronnes ouvertes de la forme $\mathbb{P}(\mathbb{C}_p) - B \sqcup B_i$, vérifiant $\cap A_i = \varnothing$.

On considère la relation d'équivalence $\sim$ entre chaînes évanescentes définie par : $\{A_i\}_{i \geqslant 1} \sim \{\widetilde{A}_i\}_{i \geqslant 1}$ si et seulement si, pour tout $n \geqslant 1$ il existe m tel que $A_m \subset \widetilde{A}_n$ et $\widetilde{A}_m \subset A_n$. Notons que deux chaînes évanescentes sont équivalentes si et seulement s'ils ont la même boule associée.

Un *bout* est une classe d'équivalence de chaînes évanescentes par la relation d'équivalence $\sim$. Notons que l'on a des bijections naturelles entre les bouts, les boules fermées et les boules ouvertes. Étant donné un bout $\mathcal{P}$, on note $B_{\mathcal{P}}$ (resp. $D_{\mathcal{P}}$) la boule (resp. le disque) associée à $\mathcal{P}$.

On dit qu'un bout $\mathcal{P}$ est *inclus* dans un ensemble X et on note $\mathcal{P} \prec X$, si pour toute chaîne évanescente $\{A_i\}_{i\geqslant 1}$ qui représente $\mathcal{P}$ il existe $n \geqslant 1$ tel que $A_n \subset X$. Un *bout d'un espace analytique* X est un bout $\mathcal{P} \prec X$ tel que $\mathcal{P} \not\prec Y$ pout tout affinoïde fermé $Y \subset X$. L'ensemble $\mathcal{T}$ des bouts d'un affinoïde ouvert connexe X est fini, les boules $B_{\mathcal{P}}$, pour $\mathcal{P} \in \mathcal{T}$, sont disjointes et on a $X = \mathbb{P}(\mathbb{C}_p) - \cup_{\mathcal{T}} B_{\mathcal{P}}$.

Étant donné un bout $\mathcal{P}$ considérons une coordonnée w telle que $D_{\mathcal{P}} = \{|w| < 1\}$. Pour $\xi \in \mathbb{P}(\widetilde{\mathbb{C}}_p)$ soit $\mathcal{P}(\xi)$ le bout tel que $D_{\mathcal{P}(\xi)} = \{w \mid \widetilde{w} = \xi\}$. On appelle $\mathcal{S} = \{\mathcal{P}(\xi)\}_{\mathbb{P}(\widetilde{\mathbb{C}}_p)}$ le *système projectif de bouts* associé à $\mathcal{P}$ ou simplement un *système projectif.* On dit que la coordonnée w est *compatible* avec $\mathcal{P}$ ou $\mathcal{S}$. Chaque boule ou disque est associé alors à un unique système projectif. Notons que le paramétrage $\mathcal{P}(\xi)$ dépend aussi de la coordonnée w. A chaque coordonnée w correspond un unique *système projectif canonique* : celui associé à $\{|w| < 1\}$.

Étant donné un espace analytique connexe X on dit que $\mathcal{S}$ est *inclus* dans X, et on note $\mathcal{S} \prec X$, si X n'est pas contenu dans un disque associé à $\mathcal{S}$. De façon équivalente $\mathcal{S} \prec X$ si et seulement s'il existe $\mathcal{P}_0, \mathcal{P}_1 \in \mathcal{S}$ distincts, tels que $\mathcal{P}_0, \mathcal{P}_1 \prec X$. Dans ce cas l'ensemble des bouts $\mathcal{P} \in \mathcal{S}$ tels que $D_{\mathcal{P}} \not\subset X$ est fini.

L'ensemble de tous les systèmes projectifs de $\mathbb{P}(\mathbb{C}_p)$ est appelé *l'espace hyperbolique p-adique*, que l'on note $\mathbb{H}_p$; voir remarque ci-dessous. Étant donnés $\mathcal{S}_0$ et $\mathcal{S}_1 \in \mathbb{H}_p$ distincts il existe un unique $\mathcal{P}_i \in \mathcal{S}_i$ tel que $\mathcal{S}_{1-i} \prec D_{\mathcal{P}_i}$, pour $i = 0, 1$; on pose $C(\mathcal{S}_0, \mathcal{S}_1) = D_{\mathcal{P}_0} \cap D_{\mathcal{P}_1}$. Si $z_0, z_1 \in \mathbb{P}(\mathbb{C}_p)$ sont distincts on définit les couronnes $C(z_0, z_1) = \mathbb{P}(\mathbb{C}_p) - \{z_0, z_1\}$ et $C(S_0, z_1) = D_{\mathcal{P}} - \{z_1\}$, où $\mathcal{P} \in \mathcal{S}_0$ est le bout tel que $z_1 \in D_{\mathcal{P}}$. Pour $\mathcal{S}_0, \mathcal{S}_1 \in \mathbb{H}_p \cup \mathbb{P}(\mathbb{C}_p)$ on pose,

$$(\mathcal{S}_0, \mathcal{S}_1) = \{\mathcal{S} \in \mathbb{H}_p - \{\mathcal{S}_0, \mathcal{S}_1\} \mid C(\mathcal{S}_0, \mathcal{S}) \cap C(\mathcal{S}, \mathcal{S}_1) = \varnothing\},$$

$[\mathcal{S}_0, \mathcal{S}_1) = \{\mathcal{S}_0\} \cup (\mathcal{S}_0, \mathcal{S}_1)$ et $[\mathcal{S}_0, \mathcal{S}_1] = [\mathcal{S}_0, \mathcal{S}_1) \cup \{\mathcal{S}_1\}$. Chaque triplet $\mathcal{S}_0$, $\mathcal{S}_1$, $\mathcal{S}_2 \in \mathbb{H}_p \cup \mathbb{P}(\mathbb{C}_p)$ détermine $\mathcal{S} \in \mathbb{H}_p \cup \mathbb{P}(\mathbb{C}_p)$ tel que $[\mathcal{S}_i, \mathcal{S}_j] \cap [\mathcal{S}_i, \mathcal{S}_k] = [\mathcal{S}_i, \mathcal{S}]$, où $\{i, j, k\} = \{1, 2, 3\}$. Si $\mathcal{S}_0$, $\mathcal{S}_1$ et $\mathcal{S}_2$ sont distincts, alors $\mathcal{S} \in \mathbb{H}_p$.

On dit qu'un système projectif $\mathcal{S}$ est *entre* $\mathcal{S}_0$ et $\mathcal{S}_1$ si $\mathcal{S} \in (\mathcal{S}_0, \mathcal{S}_1)$. De plus on dit que qu'un ensemble $\mathcal{A} \subset \mathbb{H}_p \cup \mathbb{P}(\mathbb{C}_p)$ est *convexe* si pour tous $\mathcal{S}_0, \mathcal{S}_1 \in \mathcal{A}$ distincts on a $(\mathcal{S}_0, \mathcal{S}_1) \subset \mathcal{A}$. Pour $\mathcal{T} \subset \mathbb{H}_p \cup \mathbb{P}(\mathbb{C}_p)$ l'ensemble $[\mathcal{T}] = \cup_{\mathcal{T}} [\mathcal{S}_0, \mathcal{S}_1]$ est la *clôture convexe* de $\mathcal{T}$. On pose $(\mathcal{T}) = \cup_{\mathcal{T}} (\mathcal{S}_0, \mathcal{S}_1)$ qui est aussi un ensemble convexe.

Pour $\mathcal{S}_0$, $\mathcal{S}_1 \in \mathbb{H}_p \cup \mathbb{P}(\mathbb{C}_p)$ différents, la formule

$$d(\mathcal{S}_0, \mathcal{S}_1) = \mathrm{mod}(C(\mathcal{S}_0, \mathcal{S}_1)) \in \log_p |\mathbb{C}_p| = \mathbb{Q},$$

définit une distance sur $\mathbb{H}_p$. Alors $(\mathcal{S}_0, \mathcal{S}_1) \subset \mathbb{H}_p$ est isométrique à l'intervalle $(0, d(\mathcal{S}_0, \mathcal{S}_1)) \cap \mathbb{Q} \subset \mathbb{R}$ si $\mathcal{S}_0$ ou $\mathcal{S}_1$ n'appartient pas à $\mathbb{P}(\mathbb{C}_p)$ et est isométrique à $\mathbb{Q} \subset \mathbb{R}$ sinon.

Remarque. — La complétion $\overline{\mathbb{H}}_p$ de $(\mathbb{H}_p, d)$ est un espace métrique géodésique au sens que pour tous $\mathcal{S}_0$, $\mathcal{S}_1 \in \overline{\mathbb{H}}_p$ il existe un sous-ensemble $\overline{[\mathcal{S}_0, \mathcal{S}_1]}$ de $\overline{\mathbb{H}}_p$ isométrique à l'intervalle $[0, d(\mathcal{S}_0, \mathcal{S}_1)] \subset \mathbb{R}$, ayant $\mathcal{S}_0$ et $\mathcal{S}_1$ comme extrémités ; voir [**GH**] p. 16. Si $\mathcal{S}_0$, $\mathcal{S}_1 \in \mathbb{H}_p$ alors $\overline{[\mathcal{S}_0, \mathcal{S}_1]}$ est la complétion de $[\mathcal{S}_0, \mathcal{S}_1] \subset \mathbb{H}_p$.

Le segment géodésique $\overline{[\mathcal{S}_0, \mathcal{S}_1]}$ est uniquement déterminé par $\mathcal{S}_0$ et $\mathcal{S}_1$. On a aussi que $\overline{\mathbb{H}}_p$ est un arbre réel au sens que si deux segments géodésiques ont exactement une extrémité en commun, leur réunion est un segment géodésique ; voir [**GH**] p. 31. Par conséquent $\overline{\mathbb{H}}_p$ est 0-hyperbolique au sens de Gromov. Le bord à l'infini $\partial_\infty \overline{\mathbb{H}}_p$ (défini comme dans [**GH**]) est égal au bord à l'infini de $\mathbb{H}_p$ et est canoniquement identifié à $\mathbb{P}(\mathbb{C}_p)$.

La distance sur $\partial_\infty \mathbb{H}_p = \mathbb{P}(\mathbb{C}_p)$ relative à $\mathcal{S} \in \mathbb{H}_p$, induite par la distance d est définie par

$$d_{\mathcal{S}}(z_0, z_1) = p^{-d(\mathcal{S}, \mathcal{S}')}, \quad \text{pour } z_0, z_1 \in \mathbb{P}(\mathbb{C}_p) \text{ avec } z_0 \neq z_1,$$

où $\mathcal{S}' \in \mathbb{H}_p$ est déterminé par $[\mathcal{S}, \mathcal{S}'] = [\mathcal{S}, z_0) \cap [\mathcal{S}, z_1)$; voir [**GH**]. Il est facile de voir que la distance $d_{\mathcal{S}}$ coïncide avec la distance chordale, dans une coordonnée compatible avec $\mathcal{S}$.

1.2.3. Arbres affines. — Fixons un espace analytique connexe $X \subset \mathbb{P}(\mathbb{C}_p)$. On va associer à X un arbre $\mathcal{A}_X$ et une partition canonique de X. Cet arbre est une adaptation de l'arbre défini par Motzkin en [**Mo**] pour les quasiconnexes ; voir aussi [**GvP**].

Étant donné un système projectif $\mathcal{S} \prec X$ on pose

$$n_{\mathcal{S}} = \#\{\mathcal{P} \in \mathcal{S} \mid D_{\mathcal{P}} \cap X = \varnothing\} \text{ et}$$
$$m_{\mathcal{S}} = \#\{\mathcal{P} \in \mathcal{S} \mid D_{\mathcal{P}} \cap X \neq \varnothing \text{ et } D_{\mathcal{P}} \not\subset X\}$$

qui sont finis. Alors $\mathcal{S}$ est un *point* (resp. *sommet*) de $\mathcal{A}_X$ si et seulement si $m_{\mathcal{S}} \geqslant 2$ ou $n_{\mathcal{S}} > 0$ (resp. $m_{\mathcal{S}} \geqslant 3$ ou $n_{\mathcal{S}} > 0$). Dans ce cas on associe à $\mathcal{S}$ l'affinoïde fermé connexe,

$$X_{\mathcal{S}} = \mathbb{P}(\mathbb{C}_p) - \cup_{\mathcal{P} \in \mathcal{S},\, D_{\mathcal{P}} \not\subset X} D_{\mathcal{P}} \subset X.$$

La *projection* $\pi_X : X \to \mathcal{A}_X$ est définie par $\pi_X^{-1}(\mathcal{S}) = X_{\mathcal{S}}$.

Les arêtes de $\mathcal{A}_X \subset \mathbb{H}_p$ sont les sous-ensembles convexes de $\mathcal{A}_X$ qui ne contiennent pas de sommet, maximaux pour ces propriétés. Si I est une arête de $\mathcal{A}_X$ alors,

$$X_I = \cup_{\mathcal{S} \in I} X_{\mathcal{S}} \subset X,$$

est une couronne ouverte. Les deux composantes de $\mathbb{P}(\mathbb{C}_p) - X_I$ rencontrent $\mathbb{P}(\mathbb{C}_p) - X$ et la couronne X_I est maximale pour ces propriétés. Un sommet $\mathcal{S}$ de $\mathcal{A}_X$ est une *extrémité* de I si l'une des composantes de $\mathbb{P}(\mathbb{C}_p) - X_I$ est une boule fermée ayant $\mathcal{S}$ comme système projectif associé.

On a la partition canonique,

$$X = \left(\sqcup_{\mathcal{S} \text{ sommet de } \mathcal{A}_X} X_{\mathcal{S}}\right) \sqcup \left(\sqcup_{I \text{ arête de } \mathcal{A}_X} X_I\right).$$

Si $\mathcal{S}$ est un sommet de $\mathcal{A}_X$, alors tout bout $\mathcal{P} \in \mathcal{S}$ tel que $D_{\mathcal{P}} \cap X \neq \varnothing$ et $D_{\mathcal{P}} \not\subset X$ est un bout d'une couronne associée à une arête de $\mathcal{A}_X$ dont $\mathcal{S}$ est une extrémité (il y a donc $m_{\mathcal{S}}$ arêtes dont $\mathcal{S}$ est extrémité). Il peut y avoir des sommets $\mathcal{S}$ de $\mathcal{A}_X$ qui ne sont pas des points de ramification, dans ce cas $n_{\mathcal{S}} > 0$.

Si $\mathcal{B} \subset \mathcal{A}_X$ est convexe, alors $Y = \pi_X^{-1}(\mathcal{B})$ est un espace analytique connexe et on a $\mathcal{A}_Y = \mathcal{B}$ et $Y_{\mathcal{S}} = X_{\mathcal{S}}$ pour tout point $\mathcal{S}$ de $\mathcal{A}_Y$.

La complétion $\overline{\mathcal{A}}_X$ de $(\mathcal{A}_X, d)$ est un arbre ; on note $\mathcal{A}_X^{\mathbb{R}}$ le plus petit connexe dans $\overline{\mathcal{A}}_X$ contenant $\mathcal{A}_X$ qui est aussi égal à la clôture convexe de $\mathcal{A}_X$ dans $\overline{\mathcal{A}}_X \subset \overline{\mathbb{H}}_p$.

Exemple 1.3

(i) $\mathcal{A}_X = \varnothing$ si et seulement si X est une boule ouverte ou une boule irrationnelle ou le complémentaire d'un point ou $\mathbb{P}(\mathbb{C}_p)$.

(ii) X est une couronne ouverte si et seulement si $\mathcal{A}_X$ n'a pas de sommets et a une seule arête ; $\mathcal{A}_X$ (resp. $\mathcal{A}_X^{\mathbb{R}}$) est alors isométrique a un segment de $\mathbb{Q}$ (resp. $\mathbb{R}$) de longueur $\mathrm{mod}(X)$.

(iii) $\mathcal{A}_X$ est réduit à un sommet $\mathcal{S}$ si et seulement si $X = X_{\mathcal{S}} = \mathbb{P}(\mathbb{C}_p) - \cup_{\mathcal{T}} D_{\mathcal{P}}$ où $\mathcal{T}$ est un partie finie non-vide de $\mathcal{S}$ (on a $\#\mathcal{T} = n_{\mathcal{S}}$). Par exemple $n_{\mathcal{S}} = 1$ si et seulement si X est une boule fermée et $n_{\mathcal{S}} = 2$ si et seulement si X est une sphère ; c'est-à-dire $X = \{z \mid |z| = 1\}$ pour un certain choix de coordonnée.

(iv) Pour tout ensemble fini $\mathcal{T} \subset \mathbb{P}(\mathbb{C}_p)$ on a $\mathcal{A}_{\mathbb{P}(\mathbb{C}_p)-\mathcal{T}} = (\mathcal{T}) \subset \mathbb{H}_p$; voir Section 1.2.2.

(v) $\mathcal{A}_X$ est un *arbre fini*, c'est-à-dire $\mathcal{A}_X$ a un nombre fini de sommets (et donc un nombre fini d'arêtes) si et seulement si $\mathbb{P}(\mathbb{C}_p) - X = \cup_{i=0}^n B_i$, où les B_i sont disjoints et chaque B_i est soit un point, soit une boule, ouverte, fermée ou irrationnelle.

Dans ce cas il existe une immersion très élégante de $\mathcal{A}_X$ dans $\mathbb{Q}^n \subset \mathbb{R}^n$ due à Yoccoz : on suppose que $\infty \in B_0$ et pour chaque B_i on choisit $z_i \in B_i$. Alors $i : \mathcal{A}_X \to \mathbb{Q}^n$ est déterminée par,

$$i \circ \pi_X : X \longrightarrow (\log_p |z - z_1|, \dots, \log_p |z - z_n|) \in \mathbb{Q}^n \subset \mathbb{R}^n.$$

La restriction de i à chaque arête est une isométrie sur son image, où l'on considère $\mathbb{R}^n$ avec la norme du maximum.

(vi) $X \neq \mathbb{P}(\mathbb{C}_p)$ est un affinoïde fermé si et seulement si $\mathcal{A}_X$ est fini et $\mathcal{A}_X^{\mathbb{R}}$ est compact ; ceci est équivalente à que $\mathcal{A}_X$ soit la clôture convexe dans $\mathbb{H}_p$ d'un ensemble fini. X est un affinoïde ouvert si et seulement s'il existe un ensemble fini $\mathcal{T} \subset \mathbb{H}_p$ tel que $\mathcal{A}_X = (\mathcal{T}) \subset \mathbb{H}_p$; voir Section 1.2.2.

(vii) L'arbre de $X_p = \mathbb{C}_p - \mathcal{O}_{\mathbb{Q}_p}$. Le système projectif canonique $\mathcal{S}$ est un sommet de $\mathcal{A}_{X_p}$, sauf si $p = 2$. A cause de cela on considère d'abord le cas $p > 2$; voir figure 1.

Les sommets de $\mathcal{A}_{X_p}$, différents de $\mathcal{S}$, sont les systèmes projectifs $\mathcal{S}_{a_0 \dots a_n}$, où $a_i \in \{0, 1, \dots, p-1\}$, associés aux boules

$$B_{a_0 \dots a_n} = \{|z - (a_0 + \cdots + a_n p^n)| \leqslant p^{-(n+1)}\}$$

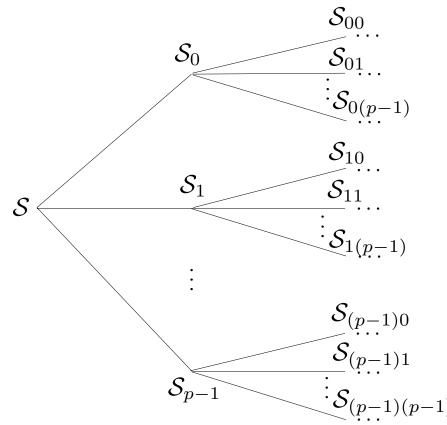

FIGURE 1. Arbre de $\mathbb{C}_p - \mathcal{O}_{\mathbb{Q}_p}$, pour $p > 2$.

de telle sorte que $\mathcal{S}_{a_0\ldots a_{n-1}}$ est lié à $\mathcal{S}_{a_0\ldots a_{n-1}a_n}$ par l'arête associée à

$$\{p^{-(n+1)} < |p - (a_0 + \cdots + a_n p^n)| < p^{-n}\}.$$

De plus $\mathcal{S}$ est lié à $\mathcal{S}_{a_0}$, pour $a_0 \in \{0, \ldots, p-1\}$. Notons que toutes les arêtes de $\mathcal{A}_{X_p}$ sont de longueur 1.

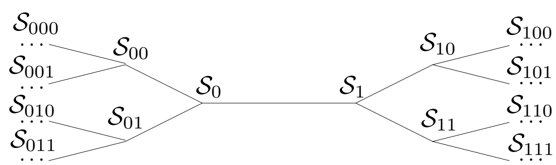

FIGURE 2. Arbre de $\mathbb{C}_2 - \mathcal{O}_{\mathbb{Q}_2}$.

Supposons $p = 2$. Considérons le système projectif $\mathcal{S}_i$ associé à la boule $B_i = \{|z - i| < \frac{1}{2}\}$, pour $i = 0, 1$. Alors $\mathcal{S}_0$ et $\mathcal{S}_1$ sont des sommets de A_{X_2} et la couronne $\mathbb{P}(\mathbb{C}_p) - B_0 \sqcup B_1$ est associée à une arête de longueur 2 qui joint $\mathcal{S}_0$ à $\mathcal{S}_1$. Les autres arêtes sont de longueur 1. En général les sommets de $\mathcal{A}_{X_2}$ sont les systèmes projectifs $\mathcal{S}_{a_0a_1\ldots a_n}$, où $a_i \in \{0, 1\}$, associés aux boules

$$B_{a_0a_1\ldots a_n} = \{|z - (a_0 + a_1 2^1 + \cdots + a_n 2^n)| \leqslant 2^{-(n+1)}\},$$

de telle façon que $\mathcal{S}_{a_0\ldots a_{n-1}}$ est lié à $\mathcal{S}_{a_0\ldots a_{n-1}a_n}$ par l'arête associée à

$$\{2^{-(n+1)} < |z - (a_0 + \cdots + a_n 2^n)| < 2^{-n}\}.$$

1.3. Analyse ultramétrique. — Dans cette section on considère les outils d'analyse qu'on va utiliser dans ce travail. Pour des références on renvoie le lecteur à [**Ca**], [**Ro**] et [**FvP**].

1.3.1. Séries convergentes. — Une série

$$f(z) = a_0 + a_1 z + \cdots \in \mathbb{C}_p((z)),$$

est convergente sur $\{|z| < r\}$ si et seulement si $\limsup_{k\to\infty} |a_k|^{1/k} r \leqslant 1$. Si $r \in |\mathbb{C}_p|$ alors f est convergente sur $\{|z| \leqslant r\}$ si et seulement si $\lim_{k\to\infty} |a_k| r^k = 0$.

Une fonction rationnelle $R \in \mathbb{C}_p(z)$ admet un développement en série en chaque point z_0, qui n'est pas un pôle. Le rayon de convergence est égal à la plus petite distance entre z_0 et un pôle.

On note $\mathcal{H}(B)$ l'anneau des séries convergentes sur $B = \{|z| \leqslant r$ ou $\{|z| < r\}$. On munit $\mathcal{H}(B)$ de la norme uniforme $\|\cdot\|_B$. Si $B = \{|z| \leqslant r\}$ alors

$$\|f\|_B = \sup_{i \geqslant 0} |a_i| r^i,$$

et on a $\|fg\|_B = \|f\|_B \|g\|_B$ et $\|f+g\|_B \leqslant \max\{\|f\|_B, \|g\|_B\}$, pour tous $f, g \in \mathcal{H}(B)$. Donc $(\mathcal{H}(B), \|\cdot\|_B)$ est un anneau valué, ultramétrique et complet.

Principe du Maximum. — *Soit $B = \{|z - a| \leqslant r\}$ avec $r \in |\mathbb{C}_p|$. Alors pour tout $f \in \mathcal{H}(B)$ on a*

$$\sup_B |f(z)| = \|f\|_B = \sup_{|z-a|=r} |f(z)|.$$

Lemme de Hensel. — *Soit $f \in \mathcal{O}_{\mathbb{C}_p}[[z]]$ une série convergente sur $\{|z| < 1\}$ à coefficients entiers. S'il existe z_0 tel que $|f(z_0)| < |f'(z_0)|^2$ alors il existe un unique zéro w de f tel que $|w - z_0| \leqslant |f(z_0)|/|f'(z_0)|$.*

Pour $w \in \mathbb{C}_p$ considère la translation $T_w(z) = z - w$. Notons que T_w préserve la distance sur $\mathbb{C}_p$ induite par la norme $|\cdot|$ et $T_w' \equiv 1$. De plus $T_w(B_r(0)) = B_r(0)$ pour chaque $r > |w|$.

Lemme de Schwarz. — *Considérons une série $f \in \mathcal{H}(B_r(0))$ de la forme $f(z) = z(a_0 + a_1 z + \cdots)$. Alors les propriétés suivantes sont équivalentes.*

(i) $|f(z)| \leqslant |z|$ *pour tout* $z \in B_r(0)$. *En particulier* $|f'(0)| \leqslant 1$.
(ii) $|a_i| \leqslant r^{-i}$, *pour* $i \geqslant 0$.
(iii) $f(B_r(0)) \subset B_r(0)$.

Considérons maintenant une série $f \in \mathcal{H}(B_r(0))$ telle que $f(B_r(0)) \subset B_r(0)$. Alors $|f(z_0) - f(z_1)| \leqslant |z_0 - z_1|$ et $|f'(z)| \leqslant 1$ pour tous z_0, z_1 et $z \in B_r(0)$. De plus les propriétés suivantes sont équivalentes.

(iv) $f : B_r(0) \to B_r(0)$ *est un automorphisme.*
(v) *Il existe* $z \in B_r(0)$ *tel que* $|f'(z)| = 1$.
(v)' $|f'(z)| = 1$ *pout tout* $z \in B_r(0)$.
(vi) *Il existe* $z_0, z_1 \in B_r(0)$ *différents tels que* $|f(z_0) - f(z_1)| = |z_0 - z_1|$.

(vi)' *Pour tous* $z_0, z_1 \in B_r(0)$ *on a* $|f(z_0) - f(z_1)| = |z_0 - z_1|$.

Le corollaire suivant est immédiat.

Corollaire 1.4. — *Si* $f \in \mathcal{H}(B_r(0))$ *alors pour tous* x, $y \in B_r(0)$,

$$|f(x) - f(y)| \leqslant \frac{|x-y|}{r} \|f\|_{B_r(0)}.$$

En particulier notons que si $\|f\|_{B_r(0)} < r$, alors $B_r(0)$ contient un unique point fixe de f.

Démonstration du Lemme de Schwarz. — Clairement (ii) $\Rightarrow$ (i) $\Rightarrow$ (iii). Supposons (iii). Par le Principe du Maximum pour tout $r_0 \in |\mathbb{C}_p| \cap (0, r)$ on a

$$|a_i| r^i \leqslant \sup_{|z|=r_0} \left| \frac{f(z)}{z} \right| \leqslant \frac{r}{r_0}, \text{ pour } i \geqslant 1.$$

Donc $|a_i| \leqslant r^{-i}$, pour $i \geqslant 1$, et (iii) $\Rightarrow$ (ii).

Soit $f \in \mathcal{H}(B_r(0))$ une série telle que $f(B_r(0)) \subset B_r(0)$. Notons que pour chaque $w \in B_r(0)$ on a $T_{f(w)} \circ f \circ T_{-w}(0) = 0$. Donc par (i) on a $|f(z_0) - f(z_1)| \leqslant |z_0 - z_1|$ et $|f'(z)| \leqslant 1$ pour tous z_0, z_1 et $z \in B_r(0)$.

Évidemment (vi)$'$ $\Rightarrow$ (vi)$'$ et (vi)$'$ $\Rightarrow$ (v)$'$ $\Rightarrow$ (v). Supposons $f(0) = 0$. Alors par (ii) on a $|f(z) - f'(0)z| < |z|$ et $|f'(z) - f'(0)| < 1$ pour tout $z \in B_r(0) - \{0\}$. Donc (vi) $\Rightarrow$ (v) $\Rightarrow$ (v)$'$.

Supposons (iv). Fixons $w \in B_r(0)$ et soit $g = T_{f(w)} \circ f \circ T_{-w}$, de telle façon que $g(0) = 0$. Donc par (i) appliqué à g et à g^{-1} on a $|g'(0)| = 1$ et $|g(z)| = |z|$ pour tout $z \in B_r(0)$. Donc (iv) $\Rightarrow$ (v)$'$ et (iv) $\Rightarrow$ (vi)$'$.

Il suffit de montrer (v) $\Rightarrow$ (iv). Supposons alors (v) et on peut supposer de plus $f(0) = 0$. Alors $f(z) = z(a_0 + a_1 z + \cdots)$ où $|a_0| = 1$ et on a $|a_i| \leqslant r^{-i}$, pour $i \geqslant 0$, par (ii).

On définit par induction $b_i \in \mathbb{C}_p$, pour $i \geqslant 0$, tel que $|b_i| \leqslant r^{-i}$ et tel que

$$\sum_{0 \leqslant j \leqslant i} b_j (f(z))^j = z(1 + \cdots + c_{i+1} z^{i+1} + \cdots).$$

On pose $b_1 = a_1^{-1}$ et supposons b_i déjà définie. Alors notons que c_{i+1} comme plus haut satisfait $|c_{i+1}| \leqslant r^{-(i+1)}$. Donc $b_{i+1} = c_{i+1} a_1^{-(i+1)}$ satisfait les hypothèses d'induction. Par conséquent la série $z(b_0 + b_1 z + \cdots) \in \mathcal{H}(B_r(0))$ est l'inverse de f et donc f est un automorphisme de $B_r(0)$. □

1.3.2. Degré de Weierstrass, polygone de Newton et fonction de valuation. — On considère $\mathcal{O}_{\mathbb{C}_p}[[z]]$, l'anneau des séries à coefficients dans $\mathcal{O}_{\mathbb{C}_p}$. Pour une telle série $f(z) = a_0 + a_1 z + \cdots \in \mathcal{O}_{\mathbb{C}_p}[[z]]$, le plus petit entier $d \geqslant 0$ tel que $|a_d| = 1$ est appelé le *degré de Weierstrass* de f que l'on note wideg(f). Si wideg(f) est fini alors il est égal au nombre de zéros de f dans $\{|z| < 1\}$, comptés avec multiplicité.

Plus généralement considérons

$$f(z) = a_0 + a_1 z + \cdots \in \mathbb{C}_p((z)), \text{ avec } a_0 \neq 0.$$

Le *polygone de Newton* de f, que l'on note $P(f)$, est l'enveloppe convexe dans $\mathbb{R}^2$ des points $(j, \log_p |a_j|)$, pour $a_j \neq 0$, et de $\{(0,t) \mid t \leqslant \log_p |a_0|\}$. Soient $(i_j, \log_p |a_{i_j}|)$, pour $j \geqslant 0$, les sommets de $P(f)$ (on peut avoir un nombre fini ou infini des sommets). Comme $a_0 \neq 0$ on a $i_0 = 0$.

Proposition. — *Soit $m \geqslant 1$ tel que $i_0 < i_1 < \cdots < i_m$ sont définis. Pour $1 \leqslant k \leqslant m$ on pose*

$$r_k = \left(\frac{|a_{i_{k-1}}|}{|a_{i_k}|} \right)^{1/(i_k - i_{k-1})} \in |\mathbb{C}_p|.$$

Alors $r_1 < \cdots < r_m$, f est convergente sur $\{|z| \leqslant r_m\}$ et f a $i_k - i_{k-1}$ zéros de norme r_k, comptés avec multiplicité, et ces zéros sont tous les zéros de f dans $\{r_{k-1} < |z| \leqslant r_k\}$.

En particulier notons que si f est une série convergente dans $B_r(0)$, alors $f(B_r(0))$ est de la forme $B_s(f(0))$ et il existe $d \in \{1, 2, \ldots\} \cup \{\infty\}$ tel que pour tout $w \in f(B_r(0))$, la série $f - w$ a d zéros de f dans $B_r(0)$, comptés avec multiplicité. On dit alors que $f : B_r(0) \to f(B_r(0))$ *est de degré d.*

Étant donné un polynôme $P \in \mathbb{C}_p[z]$, $y \in \mathbb{C}_p$ et $r \in |\mathbb{C}_p|$ on définit

$$|P|_y(r) = \sup_{|x-y|=r} |P(x)|.$$

Soit $R = P/Q \in \mathbb{C}_p(z)$, avec P, $Q \in \mathbb{C}_p[z]$. On définit,

$$|R|_y = \frac{|P|_y}{|Q|_y},$$

qui s'étend en une fonction continue définie sur $[0, \infty)$, monomiale par morceaux. On appelle $|R|_y$ la *fonction de valuation* de R en y. Notons que si y, $z \in \mathbb{C}_p$ et $r \geqslant |y - z|$ alors $|R|_y(r) = |R|_z(r)$. Si R n'a pas ni zéros ni pôles dans $\{x \mid |x - y| = r\}$ alors $|R|_y(r) = |R(x)|$ pour tout x tel que $|x - y| = r$.

Le graphe de l'application $\log_p r \mapsto \log_p(|R|_y(r))$ est appelé le *polygone de valuation* de R en y. La pente du polygone à gauche (resp. à droite) du point d'abscisse $\log_p r$ représente la différence entre le nombre de zéros et de pôles de R dans $B_r(y)$ (resp. $B_r^+(y)$). Donc le changement de pente au point d'abscisse $\log_p r$ représente la différence entre le nombre de zéros et les pôles situées sur $B_r^+(y) - B_r(y)$. En particulier le polygone de valuation d'un polynôme est convexe.

1.3.3. Fonctions holomorphes. — Comme dans le cas complexe on dit qu'une fonction est *analytique* si elle a un développement en série en chaque point. Cette notion est trop général : la fonction qui vaut 1 sur $\{|z| \leqslant 1\}$ et 0 dans le complémentaire est analytique dans ce sens. C'est pour cette raison qui on considère la notion plus restreinte de *fonction holomorphe* définie sur un espace analytique.

Une fonction *holomorphe* définie sur un affinoïde fermé X est la limite uniforme de fonctions rationnelles sans pôles dans X. Les fonctions holomorphes définies sur un espace analytique $X \subset \mathbb{P}(\mathbb{C}_p)$ sont les fonctions dont la restriction à tout affinoïde fermé contenu dans X est holomorphe.

On note $\mathcal{H}(X)$ l'anneau des fonctions holomorphes définies sur X. Si X est une boule fermée alors $\mathcal{H}(X)$ coïncide avec la définition de la Section 1.3.1. On munit $\mathcal{H}(X)$ de la norme ultramétrique

$$\|f\|_X = \sup_{x\in X} |f(x)|.$$

Si X est un espace analytique connexe différent de $\mathbb{P}(\mathbb{C}_p)$, alors on dit qu'une fonction définie sur X est *méromorphe* si elle est de la forme $[f,g]$ où $f,g \in \mathcal{H}(X)$ n'ont pas de zéros en commun. Un affinoïde fermé $Y \subset X$ rencontre au plus un nombre fini de zéros et de pôles d'une fonction méromorphe non identiquement nulle.

Soit $X \subset \mathbb{P}(\mathbb{C}_p)$ un espace analytique connexe et $z_0 \in X$. Étant donnée $f \in \mathcal{M}(X)$ considérons une coordonnée telle que $z_0 = 0$. Alors f est de la forme

$$f(z) = a_d z^d + a_{d+1} z^{d+1} + \cdots, \text{ avec } d \in \mathbb{Z} \text{ et } a_d \neq 0.$$

L'entier d ne dépend pas du choix de la coordonnée. On l'appelle l'*ordre* de f en z_0 et on note $\mathrm{ord}_f(z_0)$.

2. Propriétés des fonctions rationnelles

Dans cette section on étudie l'action des fonctions rationnelles sur des objets géométriques. On fixe pour toute cette section une fonction rationnelle $R \in \mathbb{C}_p(z)$ non-constante.

On commence par les propriétés élémentaires des fonctions rationnelles. Étant donné un point $w \in \mathbb{P}(\mathbb{C}_p)$ le *degré* local de R en w, que l'on note $\deg_R(w)$, est définit comme suit. On considère des coordonnées telles que $w = 0$ et $R(0) = 0$. Alors R est localement de la forme $a_d z^d + \cdots$ où $d \geqslant 1$ et $a_d \neq 0$; on défini $\deg_R(w) = d$. Il n'est pas difficile de voir que $\deg_R(w)$ ne dépend pas du choix des coordonnées.

Pour tout $w \in \mathbb{P}(\mathbb{C}_p)$ on a

$$\sum_{R(z)=w} \deg_R(z) = \deg(R)$$

et pour tout $Q \in \mathbb{C}_p(z)$ on a $\deg_{Q\circ R}(w) = \deg_Q(R(w)) \cdot \deg_R(w)$.

Les points *critiques* de $R \in \mathbb{C}_p(z)$ sont les points $w \in \mathbb{P}(\mathbb{C}_p)$ tels que $\deg_R(w) > 1$. Alors la *multiplicité* de w comme point critique de R est $\deg_R(w) - 1$. Une fonction rationnelle R a $2\deg(R) - 2$ points critiques comptés avec multiplicité. C'est-à-dire

$$\sum_{w\in\mathbb{P}(\mathbb{C}_p)} (\deg_R(w) - 1) = 2\deg(R) - 2.$$

Étant donnés X, $Y \subset \mathbb{P}(\mathbb{C}_p)$ tels que $R(X) \subset Y$, on dit que $R : X \to Y$ est *de degré* $d \geqslant 1$ si tout point dans Y a exactement d préimages dans X, comptées avec multiplicité ; de façon équivalente, pour tout $y \in Y$ on a

$$\sum_{x \in X,\, R(x)=y} \deg_R(y) = d.$$

2.1. Fonctions rationnelles et bouts. — Dans cette section on décrit l'action d'une fonction rationnelle sur les bouts.

Lemme 2.1. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Alors on a les propriétés suivantes.*

(i) *Si $\{A_i\}_{i\geqslant 1}$ est une chaîne évanescente alors $\{R(A_i)\}_{i\geqslant n}$ est une chaîne évanescente, pour n assez grand.*

(ii) *Si $\{A_i\}_{i\geqslant 1}$ et $\{\widetilde{A}_i\}_{i\geqslant 1}$ sont deux chaînes évanescentes équivalentes alors $\{R(A_i)\}_{i\geqslant n}$ et $\{R(\widetilde{A}_i)\}_{i\geqslant n}$ sont équivalentes, pour n assez grand. On définit ainsi l'image $R(\mathcal{P})$ d'un bout $\mathcal{P}$ par R.*

(iii) *Soit $\mathcal{P}$ un bout. Alors il existe $d \geqslant 1$ tel que pour tout représentant $\{A_i\}_{i\geqslant 1}$ de $\mathcal{P}$ et pour tout i assez grand $R : A_i \to R(A_i)$ est de degré d. On note $d = \deg_R(\mathcal{P})$ et on l'appelle* degré local de R en $\mathcal{P}$.

(iv) *Soit $\mathcal{P}$ un bout. Alors il existe $N \geqslant 0$ tel que chaque point $y \in \mathbb{P}(\mathbb{C}_p)$ à N préimages de R dans $D_{\mathcal{P}}$ si $y \notin D_{R(\mathcal{P})}$, et $N + \deg_R(\mathcal{P})$ si $y \in D_{R(\mathcal{P})}$.*

La démonstration de ce lemme est à la fin de cette section. Notons que pour $R, Q \in \mathbb{C}_p(z)$ on a $\deg_{R\circ Q}(\mathcal{P}) = \deg_R(Q(\mathcal{P})) \cdot \deg_Q(\mathcal{P})$.

Corollaire 2.2. — *Soit $\mathcal{P}$ un bout et $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Alors $R(D_{\mathcal{P}}) = \mathbb{P}(\mathbb{C}_p)$ ou $R : D_{\mathcal{P}} \to D_{R(\mathcal{P})}$ est de degré* $\deg_R(\mathcal{P})$.

Démonstration. — Soit $N \geqslant 0$ comme en (iv) du lemme. Si $\deg_R(\mathcal{P}) + N < \deg(R)$ alors $R(D_{\mathcal{P}}) = \mathbb{P}(\mathbb{C}_p)$. Mais si $\deg_R(\mathcal{P}) + N = \deg(R)$ alors $R : D_{\mathcal{P}} \to D_{R(\mathcal{P})}$ est de degré $\deg_R(\mathcal{P})$. □

Lemme 2.3. — *Considérons des fonctions rationnelles R, $Q \in \mathbb{C}_p(z)$ et un bout $\mathcal{P}$ tel que la boule fermée $B_{R(\mathcal{P})}$ soit de la forme $\{|z - z_0| \leqslant r\}$. Soit $\{A_n\}_{n\geqslant 1}$ une chaîne évanescente définissant $\mathcal{P}$. Supposons qu'il existe n_0 tel que $|Q(z) - R(z)| \leqslant r$ pour $z \in A_{n_0}$. Alors $Q(\mathcal{P}) = R(\mathcal{P})$ et $\deg_Q(\mathcal{P}) = \deg_R(\mathcal{P})$.*

Démonstration. — Par un changement de coordonnée projectif au départ et affine à l'arrivée, on se ramène à $B_{\mathcal{P}} = B_{R(\mathcal{P})} = \{|z| < 1\}$ et $r = 1$. Soit $d = \deg(\mathcal{P})$. Par le lemme précédent on peut supposer que

$$R : A_{n_0} = \{1 < |z| < r_0\} \longrightarrow \{1 < |z| < r_0^d\}$$

est de degré d. Alors pour tout $z_1 \in A_{n_0}$ on a $|R(z_1)| = |z_1|^d$ et l'image du disque $D_{z_1} = \{|z - z_1| < |z_1|\}$ par R est le disque $\{|z - R(z_1)| < |R(z_1)|\}$. Donc on a

$Q(D_{z_1}) = R(D_{z_1})$; on conclut que $Q : A_n \to R(A_n)$ est de degré d pour n assez grand. Par conséquent $Q(\mathcal{P}) = R(\mathcal{P})$ et $\deg_Q(\mathcal{P}) = d = \deg_R(\mathcal{P})$. □

Démonstration du Lemme 2.1

(i) Il est facile de voir que cette propriété est vraie pour les automorphismes de $\mathbb{P}(\mathbb{C}_p)$, donc on peut faire des changements de coordonnée. Après changement de coordonnée on peut supposer que le disque associé à $\{A_i\}_{i\geqslant 1}$ est $\{|z| < 1\}$. Alors pour i assez grand $A_i = \{r_i < |z| < 1\}$, où $r_i \to 1$ quand $i \to \infty$.

Étant donné $P(z) = a_0 + a_1 z + \cdots + a_k z^k \in \mathbb{C}_p[z]$, soit $n(P)$ le plus petit entier qui maximise $|a_n|$ et on pose $T_P(z) = P(z) - a_n z^n$. Après changement de coordonnée à l'arrivée on suppose que $R = P/Q$ avec $P(z) = z^{n(P)} + T_P(z) \in \mathbb{C}_p[z]$ et $Q(z) = z^{n(Q)} + T_Q(z) \in \mathbb{C}_p[z]$. Quitte à changer R par $R-1$ on peut supposer que $n(P) \neq n(Q)$ et quitte à changer R par $1/R$ on peut supposer que $n(P) > n(Q)$.

Alors pour $w \in B_1(0)$ le polynôme

$$P_w(z) = P(z) - wQ(z) = c_0 + c_1 z + \cdots + c_n z^n,$$

est à coefficients entiers et $|c_{n(P)}| = 1$.

On choisit $r \in (0,1)$ tel que si $r < |w| < 1$ alors $|c_{n(Q)}| = |w|$, $|c_i| \leqslant |w|$ pour $n(Q) < i < n(P)$ et $|c_j| r_w^j < |c_{n(Q)}| r_w^{n(Q)}$ pour $0 \leqslant j < n(Q)$, où $r_w = |w|^{1/(n(P)-n(Q))}$. Par conséquent $(n(Q), \log_p |w|)$ et $(n(P), 0)$ sont des sommets consécutifs du polygone de Newton de P_w. Par la proposition de la Section 1.3.2 P_w à $n(P) - n(Q)$ zéros dans $\{|z| = r_w\}$, comptés avec multiplicité. C'est-à-dire que R a $n(P) - n(Q)$ préimages de w dans $\{|z| = r_w\}$, comptées avec multiplicité. Donc,

$$R : \{r^{1/(n(P)-n(Q))} < |z| < 1\} \longrightarrow \{r < |z| < 1\},$$

est de degré $d = n(P) - n(Q)$.

(ii) Ceci résulte immédiatement de la définition de chaîne évanescente.

(iii) Ceci a été montré dans (i).

(iv) Posons $d = \deg_R(\mathcal{P})$. Après changements de coordonnée on peut supposer que $D_{\mathcal{P}} = D_{R(\mathcal{P})} = \{|z| < 1\}$. Par (i) il existe $r_0 \in (0,1)$ tel que $R : \{r_0 < |z| < 1\} \to \{r_0^d < |z| < 1\}$ est de degré $\deg_R(\mathcal{P})$. En particulier $|R(z)| = |z|^d$ pour $r_0 < |z| < 1$. Soit N le nombre de pôles de R dans $\{|z| < 1\}$. Alors R a $N + d$ zéros dans $\{|z| < 1\}$; voir polygone de valuation dans la Section 1.3.2.

Considérons $y \in \{|z| < 1\}$. Alors $|R(z) - y| = |z|^d$ pour $\max\{r_0, |y|\} < |z| < 1$. Donc $R - y$ a $N + d$ zéros dans $\{|z| < 1\}$.

Considérons maintenant $y \notin \{|z| > 1\}$. Alors $|yR(z)/(R(z) - y)| = |R(z)|$ pour $r_0 < |z| \leqslant 1$. Comme la fonction rationnelle $yR/(R - y)$ a les mêmes zéros que R sur $\{|z| < 1\}$, on conclut que $yR/(R - y)$ a N pôles dans $\{|z| < 1\}$. □

2.1.1. Théorème des résidus pour les fonctions rationnelles. — Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Étant donné un bout $\mathcal{P}$ soit $ord_R(\mathcal{P})$ défini comme suit.

- $\mathrm{ord}_R(\mathcal{P}) = \deg_R(\mathcal{P})$ si $D_{R(\mathcal{P})}$ est de la forme $\{|z| < r\}$.

- $\mathrm{ord}_R(\mathcal{P}) = -\deg_R(\mathcal{P})$ si $D_{R(\mathcal{P})}$ est de la forme $\{|z| > r\} \cup \{\infty\}$.
- $\mathrm{ord}_R(\mathcal{P}) = 0$ si $D_{\mathcal{P}} \subset \mathbb{C}_p - \{0\}$ ou $B_{\mathcal{P}} \subset \mathbb{C}_p - \{0\}$.

Par (iv) du Lemme 2.1 on a que $\mathrm{ord}_R(\mathcal{P})$ est égal à la différence entre les zéros et les pôles de R sur $D_{\mathcal{P}}$. On considère une version du Théorème des Résidus pour les fonctions rationnelles ; voir [**FvP**].

Théorème des Résidus. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Soit X un affinoïde ouvert connexe avec bouts $\mathcal{P}_1, \ldots, \mathcal{P}_k$. Alors*

$$\sum_X \mathrm{ord}_R(z) = \sum_{1 \leqslant i \leqslant k} \mathrm{ord}_R(\mathcal{P}_i).$$

Démonstration. — Soient Z et P le nombre de zéros et de pôles de R sur X respectivement (comptés avec multiplicité). Alors $\sum_X \mathrm{ord}_R(z) = Z - P$.

Soient Z_i et P_i les nombres de zéros et de pôles de R sur $B_{\mathcal{P}_i} = \mathbb{P}(\mathbb{C}_p) - D_{\mathcal{P}_i}$, respectivement. Par la remarque plus haut on a $\mathrm{ord}_R(\mathcal{P}_i) = P_i - Z_i$.

Notons que $\deg(R) = Z + Z_1 + \cdots + Z_k = P + P_1 + \cdots + P_k$. Par conséquent on a

$$0 = Z + Z_1 + \cdots + Z_k - (P + P_1 + \cdots + P_k) = \sum_X \mathrm{ord}_R(z) - \sum_{1 \leqslant i \leqslant k} \mathrm{ord}_R(\mathcal{P}_i). \quad \square$$

2.2. Fonctions rationnelles et systèmes projectifs

Proposition 2.4. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle et $\mathcal{S}_0$ un système projectif. Alos on a les propriétés suivantes.*

(i) *Il existe un système projectif $\mathcal{S}_1$ tel que si $\mathcal{P} \in \mathcal{S}_0$ alors $R(\mathcal{P}) \in \mathcal{S}_1$.*

(ii) *Considérons des paramétrages $\mathcal{S}_i = \{\mathcal{P}_i(\xi)\}_{\xi \in \mathbb{P}(\widetilde{\mathbb{C}}_p)}$, pour $i = 0, 1$. Alors il existe une fonction rationnelle $\widetilde{R} \in \widetilde{\mathbb{C}}_p(z)$ telle que pour tout $\xi \in \mathbb{P}(\widetilde{\mathbb{C}}_p)$ on a*

$$R(\mathcal{P}_0(\xi)) = \mathcal{P}_1(\widetilde{R}(\xi)) \ \textit{ et } \ \deg_R(\mathcal{P}(\xi)) = \deg_{\widetilde{R}}(\xi).$$

Donc pour tout $\mathcal{P}_1 \in \mathcal{S}_1$ on a

$$\sum_{\mathcal{P}_0 \in \mathcal{S}_0, R(\mathcal{P}_0) = \mathcal{P}_1} \deg_R(\mathcal{P}_0) = \deg(\widetilde{R}).$$

(iii) *Il existe un sous-ensemble fini $\mathcal{T} \subset \mathcal{S}_0$ tel que $R(D_{\mathcal{P}}) = \mathbb{P}(\mathbb{C}_p)$ pour tout $\mathcal{P} \in \mathcal{T}$ et tel que $R : D_{\mathcal{P}} \to D_{R(\mathcal{P})}$ est de degré $\deg_R(\mathcal{P})$ pour tout $\mathcal{P} \in \mathcal{S}_0 - \mathcal{T}$; dans ce dernier cas $R(D_{\mathcal{P}}) = D_{R(\mathcal{P})}$.*

La démonstration de cette proposition est à la fin de cette section.

On note $\deg_R(\mathcal{S}_0)$ le degré de $\widetilde{R}$, qui ne dépend pas du choix des coordonnées. Par (iii) du Lemme 2.1 et par (iii) de la Proposition 2.4, l'application $\deg_R : \mathbb{H}_p \to \{1, 2, \ldots\}$ est semi-continue supérieurement et en particulier la condition $\deg_R(\mathcal{S}) = 1$ est ouverte ; donc si $\deg_R(\mathcal{S}) = 1$ alors $R : \mathbb{H}_p \to \mathbb{H}_p$ est une isométrie au voisinage de $\mathcal{S}$.

On pose $R^{-1}(0) \cup R^{-1}(\infty) = \{a_0, \ldots, a_n\}$ et après changement de coordonnée on suppose que $a_0 = \infty$. On considère l'application,

$$I : X = \mathbb{C}_p - \{a_1, \ldots, a_n\} \longrightarrow (\log_p |z - a_1|, \ldots, \log_p |z - a_n|) \in \mathbb{Q}^n \subset \mathbb{R}^n.$$

Alors $\log_p |R| : X \to \mathbb{Q} \subset \mathbb{R}$ se factorise comme $\log_p |R| = L \circ I$ où $L : \mathbb{Q}^n \to \mathbb{Q}$ est un fonctionnelle linéaire à coefficients entiers. De plus I se factorise comme $I = i \circ \pi_X$ où $i : \mathcal{A}_X \to \mathbb{Q}^n$ est une immersion de telle façon que $X_{\mathcal{S}} = \{z \mid I(z) = i(\mathcal{S})\}$ pour chaque point $\mathcal{S} \in \mathcal{A}_X$; voir Section 1.2.3. La restriction de i à chaque arête est une isométrie avec son image et i s'étend à une immersion propre $i^{\mathbb{R}} : \mathcal{A}_X^{\mathbb{R}} \to \mathbb{R}^n$. L'application $L \circ i^{\mathbb{R}} : \mathcal{A}_X^{\mathbb{R}} \to \mathbb{R}$ est aussi propre.

Lemme 2.5. — *Soit $\mathcal{S}^0$ (resp. $\mathcal{P}^0$) un système projectif (resp. bout). L'ensemble $R^{-1}(\mathcal{S}^0)$ (resp. $R^{-1}(\mathcal{P}^0)$) des systémes projectifs $\mathcal{S}$ (resp. bouts $\mathcal{P}$) tels que $R(\mathcal{S}) = \mathcal{S}^0$ (resp. $R(\mathcal{P}) = \mathcal{P}^0$) est fini et on a*

$$\sum_{R(\mathcal{S})=\mathcal{S}^0} \deg_R(\mathcal{S}) \Big(\textit{resp.} \sum_{R(\mathcal{P})=\mathcal{P}^0} \deg_R(\mathcal{P}) \Big) = \deg(R).$$

Démonstration. — Après changement de coordonée à l'arrivée on peut supposer que $\mathcal{P}^0$ est le bout associé à $\{|z| < 1\}$. Soient $\{a_0, a_1, \ldots, a_n\}$, X, I, $i, \ldots$ comme plus haut.

Considérons une suite $\{w_n\}_{n \geqslant 0} \subset B_1(0)$ telle que $|w_n| \to 1$ et soit z_n tel que $R(z_n) = w_n$. Comme $L \circ i^{\mathbb{R}}$ est propre, quitte à prendre une sous-suite on peut supposer que $\pi_X(z_n) \in \mathcal{A}_X$ converge vers $\mathcal{S} \in \mathcal{A}_X^{\mathbb{R}}$.

On montre d'abord que $\mathcal{S} \in \mathcal{A}_X$. Si $\mathcal{S}$ n'est pas un sommet de $\mathcal{A}_X$, alors $\mathcal{S}$ est contenu dans une arête $I^{\mathbb{R}}$ de $\mathcal{A}_X^{\mathbb{R}}$. Comme $i^{\mathbb{R}}(I^{\mathbb{R}})$ est un segment dans $\mathbb{R}^n$ avec extrémités dans $\mathbb{Q}^n$ et $L \circ i^{\mathbb{R}}(\mathcal{S}) = 1$, on a $\mathcal{S} \in \mathcal{A}_X$.

Quitte à prendre une sous-suite, on peut supposer que $\pi_X(z_n) \in (\pi_X(z_0), \mathcal{S})$, pour $n > 0$. Par conséquent le bout $\mathcal{P}$ ayant $\{C(\pi_X(z_n), \mathcal{S})\}_{n \geqslant 0}$ comme représentant est tel que $R(\mathcal{P}) = \mathcal{P}^0$. De plus, pour tout n assez grand $\deg_R(\mathcal{P})$ est égal au nombre de préimages de w_n dans $C(\pi_X(z_0), \mathcal{S})$, comptées avec multiplicité.

Si l'on considère tous les bouts $\mathcal{P}_1, \ldots, \mathcal{P}_n$ obtenus de cette manière on a pour n assez grand,

$$\sum \deg_R(\mathcal{P}_i) = \#\text{ préimages de } w_n = \deg(R).$$

En particulier $R^{-1}(\mathcal{P}) = \{\mathcal{P}_1, \ldots, \mathcal{P}_m\}$. Alors les affirmations sur les systèmes projectifs suivent de la Proposition 2.4. $\square$

Démonstration de la Proposition 2.4

(i) et (ii). Il est facile de vérifier ces propriétés pour les automorphismes de $\mathbb{P}(\mathbb{C}_p)$. Donc on peut faire des changement de coordonnée.

Etant donné un polynôme $P \in \mathbb{C}_p[z]$ et $n \geqslant 0$ on note $a_n(P)$ le coefficient de z^n. De plus on note $n(P) \geqslant 0$ (resp. $m(P) \geqslant 0$) le plus petit (resp. le plus grand) entier qui maximise $|a_i(P)|$, pour $0 \leqslant i \leqslant \deg(P)$ (voir démonstration de (i) du Lemme 2.1).

Après changement des coordonnées on suppose que $\mathcal{S}_0$ est le système projectif canonique. Considérons le paramétrage $\{\mathcal{P}(\xi)\}_{\mathbb{P}(\widetilde{\mathbb{C}}_p)}$ de $\mathcal{S}_0$.

On pose $R = P/Q$ où $P, Q \in \mathbb{C}_p[z]$. Comme dans dans la démonstration de (i) du Lemme 2.1 on peut supposer $m(P) > m(Q)$ et de plus $|a_{m(P)}(P)| = |a_{m(Q)}(Q)| = 1$. Alors $R(\mathcal{P}(\infty)) = \mathcal{P}(\infty) = \mathcal{P}(\widetilde{R}(\infty))$ et $\deg_R(\mathcal{P}(\infty)) = m(P) - m(Q) = \deg_{\widetilde{R}}(\infty)$, où $\widetilde{R}$ est la réduction de R.

On montrera $\mathcal{S}_1 = \mathcal{S}_0$. Soit $\xi \in \widetilde{\mathbb{C}}_p$ et soit $\zeta \in \{|z| \leqslant 1\}$ tel que $\widetilde{\zeta} = \xi$. Posons $T_\zeta(z) = z - \zeta$ et notons que $T_\zeta(\mathcal{P}(\xi)) = \mathcal{P}(0)$. De plus $m(P \circ T_\zeta) = m(P)$ et $m(Q \circ T_\zeta) = m(Q)$.

Si $n(P \circ T_\zeta) > n(Q \circ T_\zeta)$ alors

$$R(\mathcal{P}(\zeta)) = R \circ T_\zeta(\mathcal{P}(0)) = \mathcal{P}(0) = \mathcal{P}(\widetilde{R}(\zeta))$$

et $\deg_R(\mathcal{P}(\xi)) = n(P \circ T_\zeta) - n(Q \circ T_\zeta) = \deg_{\widetilde{R}}(\xi)$. Donc (i) et (ii) sont vérifiés dans ce cas. Le cas $n(P \circ T_\zeta) < n(Q \circ T_\zeta)$ est similaire.

Supposons $n = n(P \circ T_\zeta) = n(Q \circ T_\zeta)$ et posons $b = a_n(P \circ T_\zeta)$ et $c = a_n(Q \circ T_\zeta)$. Alors

$$\widetilde{R}(\xi) = \frac{\widetilde{b}}{\widetilde{c}}, \quad n\big(P \circ T_\zeta - \frac{b}{c} Q \circ T_\zeta\big) > n(Q \circ T_\zeta) \quad \text{et} \quad \big|a_n\big(P \circ T_\zeta - \frac{b}{c} Q \circ T_\zeta\big)\big| = 1$$

car $m(Q \circ T_\zeta) > m(Q \circ T_\zeta)$ et $|a_m(Q \circ T_\zeta)| = |a_m(P \circ T_\zeta)| = 1$. Par conséquent

$$\Big(R \circ T_\zeta - \frac{b}{c}\Big)(\mathcal{P}(0)) = \mathcal{P}(0) \text{ et}$$

$$deg_{R \circ T_\zeta - b/c}(\mathcal{P}(0)) = n\Big(P \circ T_\zeta - \frac{b}{c} Q \circ T_\zeta\Big) - n(Q \circ T_\zeta) = \deg_{\widetilde{R}}(\xi).$$

Donc $R(\mathcal{P}(\xi)) = \mathcal{P}(\widetilde{b}/\widetilde{c}) = \mathcal{P}(\widetilde{R}(\xi))$.

(iii) Soit $\mathcal{P} \in \mathcal{S}_0$. Par le Corollaire 2.2 on a $R(D_\mathcal{P}) = \mathbb{P}(\mathbb{C}_p)$ ou $R : D_\mathcal{P} \to D_{R(\mathcal{P})}$ est de degré $\deg_R(\mathcal{P})$. Comme les ensembles $D_\mathcal{P}$, pour $\mathcal{P} \in \mathcal{S}_0$, sont disjoints deux à deux l'ensemble $\mathcal{T} \subset \mathcal{S}_0$ des bouts tels que $R(D_\mathcal{P}) = \mathbb{P}(\mathbb{C}_p)$ est fini. $\square$

2.3. Fonctions rationnelles et espaces analytiques. — La propriété suivante des fonctions rationnelles simplifiera l'étude de la dynamique ; voir aussi [**Be**].

Proposition 2.6. — *Soit $\mathcal{C}$ la classe des affinoïdes fermés connexes, des affinoïdes ouverts connexes ou des espaces analytiques connexes. Si $R \in \mathbb{C}_p(z)$ est une fonction rationnelle de degré au moins deux, alors $R(X) \in \mathcal{C}$ pour tout $X \in \mathcal{C}$ et il existe $X_1, \ldots, X_n \in \mathcal{C}$ disjoints deux à deux tels que*

$$R^{-1}(X) = X_1 \sqcup \cdots \sqcup X_n.$$

De plus $R : X_i \to X$ est de degré d_i, où les d_i sont entiers positifs et vérifient $d_1 + \cdots + d_n = \deg(R)$.

En particulier pour $r \in |\mathbb{C}_p|$ l'ensemble $\{z \in \mathbb{C}_p \mid |R(z)| \leqslant r\}$ (resp. $\{z \in \mathbb{C}_p \mid |R(z)| < r\}$) est un affinoïde fermé (resp. ouvert).

La démonstration de cette proposition dépend du lemme suivant.

Lemme 2.7. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Considérons r_0, $r_1 \in |\mathbb{C}_p|$ avec $r_1 \geqslant r_0$, tels que $|R|_0 \equiv 1$ sur $[r_0, r_1]$ et $|R|_0 \equiv a_i t^{d_i}$ sur $(r_0 - \varepsilon, r_0]$ et $[r_1, r_1 + \varepsilon)$ pour $i = 0, 1$ respectivement, avec d_0 et $d_1 \neq 0$. Alors*

$$R(\{r_0 \leqslant |z| \leqslant r_1\}) = \mathbb{P}(\mathbb{C}_p),\ \{|z| = 1\},\ \{|z| \leqslant 1\} \text{ ou } \{|z| \geqslant 1\} \cup \{\infty\}.$$

Démonstration. — On suppose d'abord que $r_0 = r_1 = r$. Après changement de coordonnée on suppose $r = 1$. Soit $\mathcal{S} = \{\mathcal{P}(\xi)\}_{\xi \in \widetilde{\mathbb{C}}_p}$ le système projectif canonique. Par hypothèse $R(\mathcal{P}(0))$, $R(\mathcal{P}(\infty)) \in \{\mathcal{P}(0), \mathcal{P}(\infty)\}$. Donc pour tout $\xi \in \widetilde{\mathbb{C}}_p - \{0\}$ il existe $\zeta \in \widetilde{\mathbb{C}}_p - \{0\}$ tel que $R(\mathcal{P}(\zeta)) = \mathcal{P}(\xi)$. Alors le lemme suit du Corollaire 2.2 dans ce cas.

Supposons maintenant $r_1 > r_0$. On pose $R = P/Q$ avec $P(z) = a_0 + \cdots + a_k z^k \in \mathbb{C}_p[z]$ et $Q(z) = b_0 + \cdots + b_l z^l \in \mathbb{C}_p[z]$.

Comme $d_0 \neq 0$ (resp. $d_1 \neq 0$) P ou Q a un zéro dans $\{|z| = r_0\}$ (resp. $\{|z| = r_1\}$) ; voir polygône de valuation dans la Section 1.3.2.

Par la proposition dans la Section 1.3.2, si P a un zéro dans $\{|z| = r_0\}$ (resp. $\{|z| = r_1\}$) alors il existe $i_0 < i_1$ (resp. $i_m < i_{m+1}$) tels que

$$(i_0, \log_p |a_{i_0}|) \text{ et } (i_1, \log_p |a_{i_1}|)$$

$$(\text{resp. } (i_m, \log_p |a_{i_m}|) \text{ et } (i_{m+1}, \log_p |a_{i_{m+1}}|))$$

sont des sommets consecutifs du polygône de Newton $P(P)$ de P tels que $r_0^{i_0}|a_{i_0}| = r_0^{i_1}|a_{i_1}|$ (resp. $r_1^{i_m}|a_{i_m}| = r_1^{i_{m+1}}|a_{i_{m+1}}|$).

Comme $|R|_0 \equiv 1$ sur $[r_0, r_1]$ le point $p_1 = (j_1, \log_p |b_{j_1}|)$ (resp. $p_m = (j_m, \log_p |b_{j_m}|)$) est aussi un sommet de $P(Q)$ et on définit $j_0 = j_1 = i_1$ (resp. $j_{m+1} = j_m = i_m$).

Si P n'a pas de zéros dans $\{|z| = r_0\}$ (resp. $\{|z| = r_1\}$) alors Q en a un et on définit $j_0 < j_1$ (resp. $j_m < j_{m+1}$) de façon analogue et on pose $i_0 = i_1 = j_1$ et $i_{m+1} = i_m = j_m$.

Donc $i_1 = j_1$, $i_m = j_m$, $i_0 \neq j_0$ et $i_{m+1} \neq j_{m+1}$. Soit p_0 (resp. p_{m+1}) égal à $(i_0, \log_p |a_{i_0}|)$ (resp. $(i_{m+1}, \log_p |a_{i_{m+1}}|)$ si $i_0 < j_0$ (resp. $i_m > j_{m+1}$) ou $(j_0, \log_p |b_{j_0}|)$ (resp. $(j_{m+1}, \log_p |b_{j_{m+1}}|)$) sinon.

Alors $P(P) \cup P(Q)$ est au dessous des droites qui passent par p_0 et p_1 et par p_m et p_{m+1}. On choisit l'indice $m \geqslant 0$ de telle façon que $p_n = (i_n, \log_p |a_{i_n}|)$, pour $1 \leqslant n \leqslant m$, soient tous les sommets de $P(P)$ avec $i_0 < i_1 < \cdots < i_m$. Comme $|R|_0 \equiv 1$ sur $[r_0, r_1]$ les points $p_0, \ldots, p_m$ sont tous les sommets de $P(Q)$ avec abscisse entre $j_1 = i_1$ et $j_m = i_m$.

Considérons $w \in \mathbb{C}_p$ et soit $P_w = P - wQ$. Si $|w| = 1$ alors p_0 et p_{m+1} sont des sommets de $P(P_w)$ et $P(P_w)$ est au dessous des droites qui pasent par p_0 et p_1 et par p_m et p_{m+1}. Par la proposition de la Section 1.3.2 P_w a un zéro dans $\{r_0 \leqslant |z| \leqslant r_1\}$. Par conséquent $\{|z| = 1\} \subset R(\{r_0 \leqslant |z| \leqslant r_1\})$.

Supposons $|w| < 1$, le cas $w \in \{|z| > 1\} \cup \{\infty\}$ étant analogue. Les points $p_1, \ldots, p_m$ sont tous les sommets de $P(P_w)$ avec abscisse en $[i_0, i_m]$. De plus $P(P_w)$ est au dessous des droites qui passent par p_0 et p_1 et par p_m et p_{m+1}. Par la proposition de la Section 1.3.2 le polynôme P_w a un zéro dans $\{r_0 \leqslant |z| \leqslant r_1\}$ si et seulement si $i_0 < i_{m+1}$. Donc,

$$\text{soit } \{|z| < 1\} \subset R(\{s_0 \leqslant |z| \leqslant s_1\}, \text{ soit } \{|z| < 1\} \cap R(\{s_0 \leqslant |z| \leqslant s_1\}) = \varnothing. \qquad \square$$

Corollaire 2.8. — *Soient r_0, $r_1 \in |\mathbb{C}_p|$ avec $r_0 < r_1$ (resp. $r_0 \leqslant r_1$) et on pose $C = \{r_0 < |z| < r_1\}$ (resp. $\{r_0 \leqslant |z| \leqslant r_1\}$). Alors $R(C)$ est égal à $\mathbb{P}(\mathbb{C}_p)$, une boule ouverte (resp. fermée) ou après changement de coordonnée $\{s_0 < |z| < s_1\}$ (resp. $\{s_0 \leqslant |z| \leqslant s_1\}$).*

Démonstration. — Après changement de coordonnée on peut supposer que R est tel que $|R|_0(t) = a_0 t^{d_0}$ sur $[r_0, r_0 + \varepsilon)$ (resp. sur $(r_0 - \varepsilon, r_0]$) et tel que $|R|_0(t) = a_1 t^{d_1}$ sur $(r_1 - \varepsilon, r_1]$ (resp. sur $[r_1, r_1 + \varepsilon)$, pour $i = 0, 1$ respectivement, avec d_0 et $d_1 \neq 0$.

Par le lemme précédent

$$\{z \in \mathbb{C}_p \mid |z| \in |R|_0((r_0, r_1))\} \subset R(C)$$
$$(\text{resp. } \{z \in \mathbb{C}_p \mid |z| \in |R|_0([r_0, r_1])\} \subset R(C)).$$

De plus si R a un zéro et un pôle dans C alors $R(C) = \mathbb{P}(\mathbb{C}_p)$. Si R n'a pas de pôles, alors $|R|_0$ est convexe sur (r_0, r_1) et

$$R(C) \subset \{|z| < \max\{|R|_0(r_0), |R|_0(r_1)\}\}$$
$$(\text{resp. } \subset \{|z| \leqslant \max\{|R|_0(r_0), |R|_0(r_1)\}\})$$

avec égalité si R a un zéro dans C. Si de plus R n'a pas de zéros dans C

$$R(C) = \{\min\{|R|_0(r_0), |R|_0(r_1)\} < |z| < \max\{|R|_0(r0), |R|_0(r_1)\}\}$$
$$(\text{resp. } \{\min\{|R|_0(s_0), |R|_0(s_1)\} \leqslant |z| \leqslant \max\{|R|_0(s_0), |R|_0(s_1)\}\}).$$

Le cas où R n'a pas des zéros dans C est similaire. $\square$

Démonstration de la Proposition 2.6. — On va montrer d'abord que les classes en considération sont invariantes. Soit X un affinoïde fermé connexe. Si $X = \mathbb{P}(\mathbb{C}_p)$ on a $R(X) = \mathbb{P}(\mathbb{C}_p)$, donc on suppose $X \neq \mathbb{P}(\mathbb{C}_p)$. Par conséquent $\mathcal{A}_X \neq \varnothing$ et X a une décomposition canonique,

$$X = \left(\sqcup_{\mathcal{S} \text{ sommet de } \mathcal{A}_X} X_{\mathcal{S}}\right) \sqcup \left(\sqcup_{I \text{ arête de } \mathcal{A}_X} X_I\right),$$

donnée dans la Section 1.2.3.

Etant donnée une arête I de $\mathcal{A}_X$ avec extrémités $\mathcal{S}_0$ et $\mathcal{S}_1$, considérons l'affinoïde fermé connexe $\widehat{X}_I = X_{\mathcal{S}_0} \cup X_I \cup X_{\mathcal{S}_1}$. Notons que si I_0 et I_1 ont une extrémité commune alors $X_{I_0} \cap X_{I_1} \neq \varnothing$.

Par la Proposition 2.4 on peut montrer (comme dans le corollaire du Lemme 2.7) que si I est une arête de $\mathcal{A}_X$ avec extrémités $\mathcal{S}_0$ et $\mathcal{S}_1$ alors $R(\widehat{X}_I)$ est un affinoïde fermé connexe. Donc $R(X) = \cup \widehat{X}_I$ est un affinoïde fermé connexe.

Ceci montre aussi que si X est un espace analytique connexe, alors $R(X)$ est aussi un espace analytique connexe. Pour le cas d'affinoïdes ouverts connexes, pour chaque sommet $\mathcal{S}$ de $\mathcal{A}_X$ on considère l'affinoïde ouvert connexe $\widehat{X}_{\mathcal{S}} = X_{\mathcal{S}} \cup (\cup X_I)$, où l'union porte sur toutes les arêtes I de $\mathcal{A}_X$ ayant $\mathcal{S}$ comme extrémité. Alors on procède comme dans le cas des affinoïdes fermés connexes.

(1) Considérons un affinoïde fermé connexe, le cas des affinoïdes ouverts connexes étant similaire. Soit P l'ensemble fini des bouts $\mathcal{P}$ tels que $R(\mathcal{P})$ est un bout de $\mathbb{P}(\mathbb{C}_p) - X$ et pour $y \in R^{-1}(X)$ soit Y_y l'affinoïde fermé connexe

$$Y_y = \cap_{\mathcal{P} \in P,\, y \in B_{\mathcal{P}}} B_{\mathcal{P}}.$$

Par (iv) du Lemme 2.1 pour chaque $\mathcal{P} \in P$ il existe $d_{\mathcal{P}} \geqslant 0$ tel que tout $x \in B_{R(\mathcal{P})}$ a $d_{\mathcal{P}}$ préimages dans $D_{\mathcal{P}}$.

Donc pour tout $y \in R^{-1}(X)$ un point $x \in X = \cap_{\mathcal{P} \in P} B_{R(\mathcal{P})}$ a

$$d_y = \deg(R) - \sum_{\mathcal{P} \in P,\, y \in B_{\mathcal{P}}} d_{\mathcal{P}}$$

préimages par R dans Y_y et par conséquent $R : Y_y \to X$ est de degré d_y.

Soit $x \in X$ et soient $y, z \in R^{-1}(x)$ tels que $Y_y \cap Y_z \neq \varnothing$. Alors $Y_y = Y_z$, car sinon il existe un bout $\mathcal{P}$ de $\mathbb{P}(\mathbb{C}_p) - Y_y$ tel que $\mathcal{P} \prec Y_z$ ou vice-versa. Mais dans ce cas $R(\mathcal{P}) \prec X$, ce qui n'est pas possible car par définition $R(\mathcal{P})$ est un bout de $\mathbb{P}(\mathbb{C}_p) - X$.

Par conséquent si l'on considère la décomposition $Y_1 \sqcup \cdots \sqcup Y_n$ de $Y = \cup_{y \in R^{-1}(x)} Y_y$ en affinoïdes fermés connexes disjoints deux à deux (donnée par la Proposition 1.1) il existe des entiers $d_i \geqslant 1$, tels que $R : Y_i \to X$ est de degré d_i, pour $1 \leqslant i \leqslant n$. Donc $d_1 + \cdots + d_n = \deg(R)$ car $R^{-1}(x) \subset Y = Y_1 \sqcup \cdots \sqcup Y_n$, et par conséquent $R^{-1}(X) = Y_1 \sqcup \cdots \sqcup Y_n$.

(2) Soit $X = \cup_{i \geqslant 1} X_i \in \mathcal{C}$ un espace analytique connexe où $\{X_i\}_{i \geqslant 1}$ est une suite croissante d'affinoïdes fermés connexes. On a montré que $R^{-1}(X_i)$ est un affinoïde fermé. Soit $n_i \geqslant 1$ le nombre de composantes connexes de X_i. Alors notons que $n_{i+1} \leqslant n_i$, pour $i \geqslant 1$. En effet si

$$R^{-1}(X_i) = Y_1 \sqcup \cdots \sqcup Y_{n_i} \text{ et } R^{-1}(X_{i+1}) = Z_1 \sqcup \cdots \sqcup Z_{n_{i+1}},$$

sont les décompositions en composantes connexes alors

$$Y_1 \sqcup \cdots \sqcup Y_{n_i} = R^{-1}(X_i) \subset R^{-1}(X_{i+1}) = Z_1 \sqcup \cdots \sqcup Z_{n_{i+1}},$$

et $Z_j \cap R^{-1}(X) \neq \varnothing$ pour tout $1 \leqslant j \leqslant n_{i+1}$. Donc $n_{i+1} \leqslant n_i$. Par conséquent pour i assez grand $n_i = n$ ne dépend pas de i. Alors on peut supposer que $Y_k^i \subset Y_k^{i+1}$, pour $1 \leqslant k \leqslant n$. De plus il existe $d_1, \ldots, d_n \geqslant 1$ tels que $d_1 + \cdots + d_n = \deg(R)$ et $R : Y_k^i \to X_i$ est de degré d_k, pour $1 \leqslant k \leqslant n$.

Donc $Y_k = \cup_{i \gg 1} Y_k^i$ est un espace analytique connexe, $R : Y_k \to Y$ est de degré d_k et $R^{-1}(Y) = Y_1 \sqcup \cdots \sqcup Y_n$. $\square$

2.4. Composante d'injectivité. — Le but de cette section est de montrer le résultat suivant, qui est le premier pas dans la démonstration du Théorème 3 ; voir Section 5.

Proposition 2.9. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle et $x \in \mathbb{P}(\mathbb{C}_p)$ non-critique. Alors il existe un affinoïde ouvert $I(x)$ qui contient x, tel que R est injective sur $I(x)$ et tel que, pour tout espace analytique connexe X qui contient x sur lequel R est injective, on a $X \subset I(x)$.*

Définition 2.10. — Soient R et x comme dans la proposition. Alors on appelle $I(x)$ la *composante d'injectivité* de x pour R.

La démonstration de la Proposition 2.9 dépend de quelques lemmes.

Lemme 2.11. — *Supposons que la fontion rationnelle $R \in \mathbb{C}_p(z)$ est injective sur l'espace analytique connexe X. Alors pour tout système projectif $\mathcal{S} \prec X$ on a $\deg_R(\mathcal{S}) = 1$.*

Démonstration. — Si $\deg_R(S) > 1$ alors il y a deux cas.

Soit $\deg_R(\mathcal{P}) > 1$ pour tout $\mathcal{P} \in \mathcal{S}$: alors, si $\mathcal{P} \in \mathcal{S}$ est tel que $D_{\mathcal{P}} \subset X_{\mathcal{S}}$, R n'est pas injective sur $D_{\mathcal{P}}$.

Soit on peut trouver $\mathcal{P}_0$, $\mathcal{P}_1 \in \mathcal{S}$ distincts tels que $D_{\mathcal{P}_0}$, $D_{\mathcal{P}_1} \subset X_{\mathcal{S}}$ et $R(\mathcal{P}_0) = R(\mathcal{P}_1)$; on peut alors choisir $x_0 \in D_{\mathcal{P}_0}$ et $x_1 \in D_{\mathcal{P}_1}$ tels que $R(x_0) = R(x_1)$. □

Lemme 2.12. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Pour $r_0, r_1 \in \mathbb{R}$ soit $C(r_0, r_1) = \{r_0 < |z| < r_1\}$ et pour $r \in |\mathbb{C}_p|$ soit $\mathcal{S}_r$ le système projectif associé à $\{|z| < r\}$.*

(i) *Pour chaque $r \in \mathbb{R} - \mathbb{Q}$ il existe $d \geqslant 1$ et r_0, $r_1 > 0$ avec $r \in (r_0, r_1)$, tels que $C = R(C(r_0, r_1))$ est une couronne et $R : C(r_0, r_1) \to C$ est de degré d. On a alors $\mathrm{mod}(C) = d \cdot \mathrm{mod}(C(r_0, r_1))$.*

(ii) *Supposons que $\deg_R(\mathcal{S}_r) = d \geqslant 1$ pour tout $r \in |\mathbb{C}_p|$ dans l'intervalle (r_0, r_1). Alors il existe une coordonnée à l'arrivée tel que $|R|_0(t) = t^d$ sur $[r_0, r_1]$. Si R est injective sur $C(r_0, r_1)$ alors $|R(z)| = |z|$ pour $r_0 < |z| < r_1$ et en particulier $R(C(r_0, r_1)) = C(r_0, r_1)$.*

Corollaire 2.13. — *Si C est une couronne ouverte et $R \in \mathbb{C}_p(z)$ est injective sur C, alors $R(C)$ est aussi une couronne, $\mathrm{mod}(R(C)) = \mathrm{mod}(C)$ et R induit une isométrie entre $\mathcal{A}_C$ et $\mathcal{A}_{R(C)}$.*

Démonstration. — On suppose $C = C(r_0, r_1)$. Par le Lemme 2.11 on a $\deg_R(\mathcal{S}_r) = 1$ pour tout $r \in (r_0, r_1)$ et par (ii) du lemme on peut supposer que $|R(z)| = |z|$ pour $r_0 < |z| < r_1$. Donc $R(C) = C$ et $R(\mathcal{S}_r) = \mathcal{S}_r$ pour tout $r \in (r_0, r_1)$. □

Démonstration du Lemme 2.12

(i) Etant donné $P(z) = a_0 + \cdots + a_n z^n \in \mathbb{C}_p[z]$ soit $n(P)$ le plus petit entier qui maximise $r^{1-i}|a_i|$. Alors on procède comme dans la démonstration de (i) du Lemme 2.1.

(ii) Par (iii) du Lemme 2.1 et par (i), pour tout $r \in (r_0, r_1)$ il existe un voisinage $U_r = (s_r, t_r)$ où R est telle que

$$d(R(\mathcal{S}_u), R(\mathcal{S}_v)) = d \cdot d(\mathcal{S}_u, \mathcal{S}_v) = d|u - v|,$$

pour tout $u, v \in U_r$. On peut aussi choisir U_{r_0} (resp. U_{r_1}) de la forme $(r_0, r_0 + \varepsilon)$ (resp. $(r_1 - \varepsilon, r_1)$). Donc on peut trouver un nombre fini de U_r qui recouvrent (s, t).

Soient $u < v < w$ tels que $u \in U_{s_0}$, $v \in U_{s_0} \cap U_{s_1}$ et $w \in U_{s_1}$. Soit $\mathcal{S}$ le système projectif tel que

$$[R(\mathcal{S}_u), R(\mathcal{S}_v)] \cap [R(\mathcal{S}_v), R(\mathcal{S}_w)] = [R(\mathcal{S}_v), \mathcal{S}]\,;$$

voir Section 1.2.2. Comme $v \in U_{s_0} \cap U_{s_1}$ on a $\mathcal{S} = R(\mathcal{S}_v)$, c'est-à-dire $R(\mathcal{S}_v) \in (R(\mathcal{S}_u), R(\mathcal{S}_w))$. Donc pour tous $u < v$ dans (r_0, r_1) on a $d(R(\mathcal{S}_u), R(\mathcal{S}_v)) = d \cdot |u - v|$. Alors après changement de coordonnée à l'arrivée on a $|R|_0(r) = r^d$ sur $[r_0, r_1]$.

Si R est injective sur $C(r_0, r_1)$ soir $r \in (r_0, r_1)$ et $\mathcal{P} \in \mathcal{S}_r$ tel que $D_{\mathcal{P}} \subset \{|z| = r\}$. Par (iv) du Lemme 2.1 on a $R(D_{\mathcal{P}}) = D_{R(\mathcal{P})}$. Comme R fixe les bouts associés à $\{|z| < r\}$ et $\{|z| \leqslant r\}$ on a $R(\{|z| = r\}) = \{|z| = r\}$. □

Lemme 2.14. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle et soit X un espace analytique connexe tel que $\mathcal{A}_X \neq \varnothing$. Alors R est injective sur X si et seulement si R est injective sur chaque $X_{\mathcal{S}}$, pour $\mathcal{S} \in \mathcal{A}_X$.*

Démonstration. — Supposons que R est injective sur chaque $X_{\mathcal{S}}$, pour $\mathcal{S} \in \mathcal{A}_X$. Si X est une couronne ouverte alors, par (ii) du Lemme 2.12, on peut supposer que $X = \{s < |z| < t\}$ et $|R|_0 \equiv \mathrm{id}$ sur $[s, t]$. Par hypothèse R est injective sur chaque $\{|z| = r\}$, pour $r \in (s, t)$. Donc $R(\{|z| = r\}) = \{|z| = r\}$ et par conséquent R est injective sur X.

Dans le cas général soient $\mathcal{S}^0$, $\mathcal{S}^1 \in \mathcal{A}_X$. Alors il suffit de montrer que $R(X_{\mathcal{S}^0}) \cap R(X_{\mathcal{S}^1}) = \varnothing$. Quitte à remplacer X par $\cup_{\mathcal{S} \in [\mathcal{S}^0, \mathcal{S}^1]} X_{\mathcal{S}}$ on peut supposer que $\mathcal{S}^0$ et $\mathcal{S}^1$ sont des sommets de $\mathcal{A}_X$. Soient $\mathcal{S}_0 = \mathcal{S}^0, \mathcal{S}_1, \ldots, \mathcal{S}_k = \mathcal{S}^1$ tous les sommets de $\mathcal{A}_X$ contenus dans $[\mathcal{S}^0, \mathcal{S}^1]$.

On va trouver inductivement des coordonnées w_i à l'arrivée, telles que R est l'identité sur $[\mathcal{S}_0, \mathcal{S}_i]$, pour $0 \leqslant i \leqslant k$. On peut choisir w_0 sans problème. Supposons que w_i est déjà définie. Comme on a vérifié le lemme pour les couronnes on a $R(C(\mathcal{S}_i, \mathcal{S}_{i+1})) = C(R(\mathcal{S}_i), R(\mathcal{S}_{i+1}))$. Comme $\deg_R(\mathcal{S}_i) = 1$ on a $C(R(\mathcal{S}_{i-1}), R(\mathcal{S}_i)) \cap C(R(\mathcal{S}_i), R(\mathcal{S}_{i+1})) = \varnothing$, donc on peut trouver une coordonnée w_{i+1} qui coïncides avec w_i sur $(\mathcal{S}_0, \mathcal{S}_i)$ et telle que $(R(\mathcal{S}_i), R(\mathcal{S}_{i+1}))$ soit égal à $(\mathcal{S}_i, \mathcal{S}_{i+1})$ dans cette coordonnée.

On fixe la coordonnée w_k à l'arrivée. Alors $R(X_{\mathcal{S}_0})$ et $R(X_{\mathcal{S}_k})$ sont contenus dans de composantes différentes de $\mathbb{P}(\mathbb{C}_p)-C(\mathcal{S}_0,\mathcal{S}_k)$. Par conséquent $R(X_{\mathcal{S}^0})\cap R(X_{\mathcal{S}^1})=\varnothing$. □

Démonstration de la Proposition 2.9. — Fixons une coordonnée telle que $R(x)=0$ et soit $X=\mathbb{P}(\mathbb{C}_p)-R^{-1}(0)\cup R^{-1}(\infty)$. Comme x n'est pas un point critique R est injective sur un disque D_x contenant x et tel que $D_x\cap(R^{-1}(0)\cup R^{-1}(\infty))=\{x\}$. Alors le système projectif $\mathcal{S}_x$ associé à D_x est un point de $\mathcal{A}_X$ et $\deg_R(\mathcal{S})=1$ pour tout $\mathcal{S}\in(x,\mathcal{S}_x]\subset\mathcal{A}_X$.

Soit $\mathcal{B}\subset\mathcal{A}_X$ le plus grand convexe qui contient $\mathcal{S}_x$ et tel que $\deg_R(\mathcal{S})=1$ pour tout $\mathcal{S}\in\mathcal{B}$. Comme la condition $\deg_R(\mathcal{S})=1$ est ouverte $\mathcal{B}$ est ouverte. Par (i) du Lemme 2.12 on peut écrir $\pi_X^{-1}(\mathcal{B})=I(x)\cap X$, où $I(x)$ est un affinoïde ouvert connexe.

Si $\mathcal{S}$ est un point de $\mathcal{A}_X$ et $\mathcal{P}\in\mathcal{S}$ est tel que $D_{\mathcal{P}}\subset X_{\mathcal{S}}\subset X$ alors $R(D_{\mathcal{P}})\subset\mathbb{C}_p-\{0\}$. Par (iv) du Lemme 2.1 $R:D_{\mathcal{P}}\to D_{R(\mathcal{P})}$ est de degré $\deg_R(\mathcal{P})$. Comme $\deg_R(\mathcal{S})=1$, R est injective sur X_S. Par le Lemme 2.14 R est injective sur $\pi_X^{-1}(\mathcal{B})=I(x)\cap X$ et par conséquent R est injective sur $I(x)$.

Si Y est un espace analytique connexe alors $\pi_X(Y\cap X)$ est convexe. Donc si Y contient x et n'est pas contenu dans $I(x)$, alors $\pi_X(Y\cap X)\not\subset\mathcal{B}$. Par maximalité de $\mathcal{B}$ il existe $\mathcal{S}\in\pi_X(Y\cap X)$ tel que $\deg_R(\mathcal{S})>1$. Comme $x\in Y$ et Y rencontre $X_{\mathcal{S}}$ on a $\mathcal{S}\prec Y$. Par le Lemme 2.11 R n'est pas injective sur Y. □

2.5. Invariance des arbres. — Le but de cette section est de montrer la proposition suivante qu'est un cas particulier d'un théorème de Motzkin ; voir [**Mo**].

Proposition 2.15. — *Soit X un espace analytique connexe et $R\in\mathbb{C}_p(z)$ une fonction rationnelle injective sur X.*

(i) *Si $\mathcal{S}\prec X$ est un système projectif, alors $R(\mathcal{S})\prec Y=R(X)$.*

(ii) *Soit $\mathcal{P}\prec X$ un bout. Alors $R:D_{\mathcal{P}}\cap X\to D_{R(\mathcal{P})}\cap Y$ est de degré 1. En particulier, si $D_{\mathcal{P}}\subset X$ alors $R(D_{\mathcal{P}})=D_{R(\mathcal{P})}$.*

(iii) *Le système projectif $\mathcal{S}\prec X$ est un point (resp. sommet) de $\mathcal{A}_X$ si et seulement si $R(\mathcal{S})$ est un point (resp. sommet) de $\mathcal{A}_Y$.*

(iv) *R induit une isométrie de $\mathcal{A}_X$ sur $\mathcal{A}_Y$.*

Corollaire 2.16. — *Soit X un affinoïde ouvert connexe et $R\in\mathbb{C}_p(z)$ une fonction rationnelle telle que $R:X\to X$ est de degré 1. Alors il existe n tel que R^n induit l'identité sur $\mathcal{A}_X$.*

Démonstration. — Soit n tel que les bouts de X soient fixes par R. Si $\mathcal{A}_X=\varnothing$ alors il n'y à rien a montrer ; donc on suppose que $\mathcal{A}_X\neq\varnothing$ et par conséquent X a au moins deux bouts. Donc pour tout point $\mathcal{S}$ de $\mathcal{A}_X$ il existe des systèmes projectifs $\mathcal{S}_0$ et $\mathcal{S}_1$ associés à des bouts de X tels que $\mathcal{S}\prec(\mathcal{S}_0,\mathcal{S}_1)$. Notons que $R^n(\mathcal{S}_i)=\mathcal{S}_i$, pour $i=0,1$. Comme R induit une isométrie sur $\mathcal{A}_X$ on a

$$d(\mathcal{S},\mathcal{S}_i)=d(R^n(\mathcal{S}),R^n(\mathcal{S}_i))=d(R^n(\mathcal{S}),R(\mathcal{S}_i)),$$

pour $i = 0, 1$. Donc $R(\mathcal{S}) = \mathcal{S}$. $\square$

Démonstration de la Propostion 2.15

(i) D'après (iii) de la Proposition 2.4, sauf pour un nombre fini de bouts $\mathcal{P} \in \mathcal{S}$ on a $R(D_{\mathcal{P}}) = D_{R(\mathcal{P})}$. Donc $\mathcal{S} \prec X$ impliques $R(\mathcal{S}) \prec Y$.

(ii) Comme R est injective sur X il existe une composante d'injectivité I de R qui contient X. Rappelons que I est un affinoïde ouvert connexe. Soit Z la composante connexe de $R^{-1}(D_{R(\mathcal{P})})$ ayant $\mathcal{P}$ comme bout. Si $\mathcal{Q} \neq \mathcal{P}$ est un bout de Z alors $\mathcal{Q} \not\prec I$ et comme I est un affinoïde ouvert connexe, $B_{\mathcal{Q}} \cap I = \varnothing$. Par conséquent $D_{\mathcal{P}} \cap X \subset D_{\mathcal{P}} \cap I \subset Z$ et $R(D_{\mathcal{P}} \cap X) \subset R(Z) = D_{R(\mathcal{P})}$. Si $X \subset D_{\mathcal{P}}$ le résultat est clair, donc on suppose que $X \not\subset D_{\mathcal{P}}$. Par conséquent, si $\mathcal{S}$ est le système projectif associé à $\mathcal{P}$, alors $\mathcal{S} \prec X$ et par le Lemme 2.11 on a $\deg_R(\mathcal{S}) = 1$. Donc,

$$R(X - D_{\mathcal{P}}) \subset \cup_{\mathcal{Q} \in \mathcal{S} - \{\mathcal{P}\}} R(X \cap D_{\mathcal{Q}}) \subset \cup_{\mathcal{Q} \in \mathcal{S} - \{\mathcal{P}\}} D_{R(\mathcal{Q})} = \mathbb{P}(\mathbb{C}_p) - D_{R(\mathcal{P})},$$

et par conséquent $R : X \cap D_{\mathcal{P}} \to Y \cap D_{R(\mathcal{P})}$ est de degré 1.

(iii) D'après (i) $\mathcal{S} \prec X$ implique $R(\mathcal{S}) \prec Y$. Dans ce cas on a d'après (ii)

$$\#\{\mathcal{P} \in \mathcal{S} | D_{\mathcal{P}} \cap X = \varnothing\} = \#\{Q \in R(\mathcal{S}) | D_{\mathcal{Q}} \cap Y = \varnothing\} \text{ et}$$

$$\#\{\mathcal{P} \in \mathcal{S} | D_{\mathcal{P}} \cap X \neq \varnothing, D_{\mathcal{P}} \not\subset X\} = \#\{\mathcal{Q} \in R(\mathcal{S}) | D_{\mathcal{Q}} \neq \varnothing, D_{\mathcal{Q}} \not\subset X\}.$$

Donc $\mathcal{S}$ est un point (resp. sommet) de $\mathcal{A}_X$ si et seulement si $R(\mathcal{S})$ est un point (resp. sommet) de $\mathcal{A}_Y$. Par (ii) pour tout bout $\mathcal{P} \prec X$ on a $D_{\mathcal{P}} \subset X$ si et seulement si $D_{R(\mathcal{P})} \subset Y$. Comme $\deg_R(\mathcal{S}) = 1$ on a $R(X_{\mathcal{S}}) = Y_{R(\mathcal{S})}$.

(iv) Soient $\mathcal{S}^0$, $\mathcal{S}^1 \in \mathcal{A}_X$ et soient $\mathcal{S}_0, \ldots, \mathcal{S}_k$ les sommets de $\mathcal{A}_X$ dans $(\mathcal{S}^0, \mathcal{S}^1)$ tels que les $(\mathcal{S}_i, \mathcal{S}_{i+1})$ sont disjoints, pour $0 \leqslant i \leqslant k$, où $\mathcal{S}_0 = \mathcal{S}^0$ et $\mathcal{S}_{k+1} = \mathcal{S}^1$. On va montrer par induction dans i que $d(R(\mathcal{S}_0), R(\mathcal{S}_i)) = d(\mathcal{S}_0, \mathcal{S}_i)$. Par (iii) et par le Lemme 2.11, R induit une isométrie entre $I_i = (\mathcal{S}_i, \mathcal{S}_{i+1})$ et $R(I_i) = (R(\mathcal{S}_i), R(\mathcal{S}_{i+1}))$.

Comme R est injective sur X, R est injective sur $\mathcal{A}_X$ et par conséquent les $R(I_i)$ sont disjoints. Comme $d(R(\mathcal{S}_i), R(\mathcal{S}_{i+1})) = d(\mathcal{S}_i, \mathcal{S}_{i+1})$ on a $d(R(\mathcal{S}_0), R(\mathcal{S}_k)) = d(\mathcal{S}_0, \mathcal{S}_k)$. $\square$

3. Dynamique locale

Soit U un ouvert de $\mathbb{C}_p$ et $f : U \to \mathbb{C}_p$ une application. Un point $z_0 \in U$ est *périodique* s'il existe $k \geqslant 1$ tel que $z_0, \ldots, f^{k-1}(z_0) \in U$ et $f^k(z_0) = z_0$. On appelle k *période* de x ; si k est le plus petit entier avec cette propriété, alors k est appelé la *période primitive* de z_0 et on dit que $\{z_0, f(z_0), \ldots, f^{k-1}(z_0)\}$ est le *cycle* de z_0. Quand f est analytique au voisinage du cycle de z_0, la dérivée $(f^k)'(z_0) = (f^k)'(f^j(z_0))$ est appelé le *multiplicateur* du cycle correspondant.

Le multiplicateur est invariant par conjugaison analytique et sa nature donne une certaine information sur la dynamique de f au voisinage du cycle. On considère la classification suivante des points périodiques, en analogie avec le cas complexe.

Définition 3.1. — Soit z_0 un point périodique et λ son multiplicateur.

- Si $|\lambda| < 1$ on dit que z_0 est *attractif*. Si de plus $\lambda = 0$, alors on dit que z_0 est *super-attractif*.
- Si $|\lambda| = 1$ on dit que z_0 est *indifférent*. Si λ est une racine de l'unité on dit que z_0 est *parabolique* ou *indifférent rationnel*. Sinon on dit que z_0 est *indifférent irrationel*.
- Si $|\lambda| > 1$ on dit que z_0 est *répulsif*.

Dans les Sections 3.1 et 3.3 on considère la dynamique locale des points périodiques, selon la classification précédente. Dans la Section 3.2 on considère la dynamique 'quasi-périodique'.

3.1. Points périodiques attractifs. — Dans cette section on considère la dynamique au voisinage des points périodiques attractifs. On peut étendre la plupart des propriétés aux points périodiques répulsifs en considérant que, si z_0 est un point périodique répulsif pour f, alors z_0 est un point périodique attractif pour f^{-1}.

Si $z_0 \in U$ un point périodique attractif de f. Alors on dit que

$$\mathcal{W}^s_f(z_0) = \{z \in U \mid d(f^n(z), f^n(z_0)) \to 0 \text{ quand } n \to \infty\}$$

est le *bassin d'attraction* de z_0 et que $\cup_{n\geqslant 1}\mathcal{W}^s_f(f^n(z_0))$ est le *bassin d'attraction du cycle* de z_0. Notons que $\mathcal{W}^s_f(z_0)$ est ouvert et invariant par f^k, où k est la période de z_0.

On considère l'ensemble $A(\mathbb{C}_p)$ des séries à coefficients sur $\mathbb{C}_p$ convergentes au voisinage de 0 qui sont de la forme

$$f(z) = \lambda z + a_1 z^2 + a_2 z^3 + \cdots,$$

avec $|\lambda| < 1$. Pour une telle série, différente de λz, on définit

$$r(f) = \left(\sup_{k\geqslant 1} |a_k|^{1/k}\right)^{-1}.$$

Il n'est pas difficile de voir que f est convergente sur $B_{r(f)}(0)$ et que $r(f) < \infty$.

Proposition 3.2. — *Soit $f \in A(\mathbb{C}_p)$ une série différente de λz. Alors on a les propriétés suivantes.*

(i) *Pour $|z| < r(f)$ on a $|f(z)| \leqslant |z| \max\{|\lambda|, |z|/r(f)\}$. Donc $B_{r(f)}(0) \subset \mathcal{W}^s_f(0)$.*

(ii) *$r(f) = \sup\{r > 0 \mid f$ est convergente sur $B_r(0)$ et $f(B_r(0)) \subset B_r(0)\}$.*

Démonstration

(i) On pose $f(z) = \lambda z + a_1 z^2 + \cdots$. Pour $|z| < r(f)$ on a

$$|f(z)| = |z| \cdot |\lambda + a_1 z + a_2 z^2 + \cdots| \leqslant \max\{|\lambda|, |z|/r(f)\},$$

car $|a_k|^k r(f) \leqslant 1$.

(ii) Si f est convergente sur $B_r(0)$ avec $r > r(f)$, alors

$$\limsup_{k\geqslant 1} |a_k|^{1/k} \leqslant r^{-1} < r(f)^{-1} = \sup_{k\geqslant 1} |a_k|^{1/k},$$

donc il existe k tel que $r(f) = |a_k|^{1/k}$. Si $s \in |\mathbb{C}_p| \cap (r(f), r)$ on a

$$\|f\|_{B_s^+(0)} = s \cdot \max\left\{|\lambda|, |a_l|(r(f))^l \left(\frac{s}{r(f)}\right)^l\right\} \geqslant |s| \left(\frac{s}{r(f)}\right)^k > s. \qquad \square$$

Notons que par (ii) on a $f(B_{r(f)}(0)) \subset B_{r(f)}(0)$ et $r(f^n) \geqslant r(f)$ pour tout entier $n \geqslant 1$.

Proposition 3.3. — *Considérons une série $f \in A(\mathbb{C}_p)$.*

(i) *Si $\lambda = f'(0) \neq 0$ alors $\{\lambda^{-n} f^n\}_{n \geqslant 1}$ converge uniformément vers une fonction $\varphi \in \mathcal{H}(B_{r(f)}(0))$ sur $B_r^+(0)$, pour tout $r < r(f)$. Donc $\varphi \circ f = \lambda \varphi$. De plus, si $f(B_{r(f)}(0)) = B_{r(f)}(0)$, alors $\varphi : B_{r(f)}(0) \to \mathbb{C}_p$ est surjective.*

(ii) *Si $f'(0) = 0$ alors il existe une série φ convergente au voisinage de 0, telle que $\varphi \circ f = \varphi^d$, où $d = \deg_f(0)$.*

La démonstration de cette proposition est à la fin de cette section. Le résultat sur la linéarisation est un cas particulier d'un théorème en [**HY**] et est aussi montré dans [**Lu1**]. En considérant (i) de la proposition précédente on fait la définition suivante.

Définition 3.4. — Soit $f \in A(\mathbb{C}_p)$. Si $f(B_{r(f)}(0)) = B_{r(f)}(0)$ alors on dit que $B_{r(f)}(0)$ est le *bassin d'attraction immédiat* de 0 pour f.

En général f peut être injective sur $B_{r(f)}(0)$, par exemple $f(z) = pz/(1-z)$. D'autre part on a les propriétés suivantes.

Proposition 3.5. — *Soit $f \in A(\mathbb{C}_p)$ une série différente de λz.*

(1) *Si $r(f) \notin |\mathbb{C}_p|$ alors $r(f)$ est le rayon de convergence de f et $f : B_{r(f)}(0) \to B_{r(f)}(0)$ est de degré infini.*

(2) *Si $r(f)$ est strictement inférieur au rayon de convergence de f, on a $r(f) \in |\mathbb{C}_p|$ et $B_{r(f)}(0)$ est le bassin d'attraction immédiat de 0.*

(3) *Si $r(f) = 1$ et $B_1(0)$ est le bassin d'attraction immédiat de 0, alors $f : B_1(0) \to B_1(0)$ est de degré d, où $d = \mathrm{wideg}(f) \in \{2, 3, \ldots\} \cup \{\infty\}$. Si de plus le rayon de convergence de f est plus grand que 1, on a $d < \infty$.*

Démonstration

(1) Si $r(f) \notin |\mathbb{C}_p|$ alors il existe $k_i \to \infty$ tel que $|a_{k_i}|^{1/k_i} \to r(f)$, par conséquent $r(f)$ est le rayon de convergence de f et $f : B_{r(f)}(0) \to B_{r(f)}(0)$ est de degré infini.

(2) Si le rayon de convergence est strictement plus grand que $r(f)$ alors il existe k tel que $r(f) = |a_k|^{1/k} \in |\mathbb{C}_p|$. Donc

$$\|f\|_{B_{r(f)}(0)} = r(f) \max\{|\lambda|, |a_k|(r(f))^k\} = r(f) \text{ et } f(B_{r(f)}(0)) = B_{r(f)}(0).$$

(3) Dans ce cas pour tout $z_0 \in B_1(0)$ on a $d = \mathrm{wideg}(f - z_0) = \mathrm{wideg}(f)$, donc $f : B_1(0) \to B_1(0)$ est de degré d. Si le rayon de convergence est plus grand que 1 on a $|a_k| < 1$ pour tout k grand, donc $d = \mathrm{wideg}(f) < \infty$. $\square$

Si $B_{r(f)}(0)$ est le bassin d'attraction immédiat de f on a par la proposition précédente $d > 1$. En général d peut être infini, mais si par exemple f est à coefficients sur un corps à valuation discrète, alors $d < \infty$.

Exemple 3.6. — On considère la fonction rationnelle $f \in \mathbb{C}_p(z)$ définie par $f(z) = \lambda z^2/(z+1)$, avec $|\lambda| \in (0,1)$, qui est conjugué au polynôme $\lambda^{-1}(z^2+z)$. On a $r(f) = 1$, mais $f(B_1(0)) \neq B_1(0)$. De plus il n'est pas difficile de voir que pour tout disque $D \subset \mathbb{P}(\mathbb{C}_p)$ qui contient 0 on a $f(D) \neq D$.

Démonstration de la Proposition 3.3

(i) Soit $r > 0$ tel que $|f(z)| = |\lambda z|$ pour $|z| < r$; on pose $f(z) = \lambda z(1 + g(z))$ avec $g(0) = 0$. Soit $C > 0$ tel que $|g(z)| < C|z|$, pour $|z| < r$. Donc $\|g \circ f^n\|_{B_r^+(0)} < C|\lambda|^n r$ et

$$\varphi_n(z) = \lambda^{-n} f^n(z) = z(1 + g(z))(1 + g \circ f(z)) \cdots (1 + g \circ f^{n-1}(z)),$$

converge uniformément sur $B_r^+(0)$ vers une fonction $\varphi \in \mathcal{H}(B_r^+(0))$. Comme pour tout $n \geqslant 0$ on a $\lambda \varphi_{n+1} = \varphi_n \circ f$, on conclut que $\varphi \circ f = \lambda \varphi$.

Montrons que $\varphi \in \mathcal{H}(B_{r(f)}(0))$. En effet soit $s < r(f)$. Il existe $n \geqslant 1$ tel que $f^n(B_s^+(0)) \subset B_r(0)$, d'après (i) de la Proposition 3.2. La fonction $\lambda^n \varphi = \varphi \circ f^n$ appartient donc à $\mathcal{H}(B_s^+(0))$ et coïncide avec φ dans $\mathcal{H}(B_r^+(0))$. On obtient ainsi que $\varphi \in \mathcal{H}(B_{r(f)}(0))$.

L'image de φ contient une boule $B_t(0)$. L'équation fonctionnelle montre alors, si $f(B_{r(f)}(0)) = B_{r(f)}(0)$, que cette image contient la boule $B_{|\lambda|^{-n}t}(0)$ pour tout $n \geqslant 0$.

(ii) Après conjugaison par λz on peut supposer que $f(z) = z^d + \cdots$, où $d = \deg_f(0) > 1$. On pose $f(z) = z^d(1 + g(z))$ avec $g(0) = 0$. On considère $r > 0$ et $C > 0$ tels que $\|g\|_{B_t^+(0)} < Ct < 1$, pour $t \in (0, r)$. En particulier on a $|f(z)| = |z|^d$ pour $0 < |z| < r$. Notons que si $t < p^{-1/(p-1)}$ alors $|\ln(1+z)| = |z|$. On suppose que $Cd^{n+1}r^{d^n} < p^{-1/(p-1)}$ pour tout $n \geqslant 1$. Donc si $t \in (0, r)$ et $n \geqslant 0$ on a

$$\left\| \frac{1}{d^{n+1}} \ln(1 + g \circ f^n) \right\|_{B_t^+(0)} \leqslant d^{n+1} |g \circ f^n|_0(t) \leqslant d^{n+1} C t^{d^n} < C d^{n+1} r^{d^n} < p^{-1/(p-1)},$$

et $\lim_{n\to\infty} \|\frac{1}{d^{n+1}} \ln(1 + g \circ f^n)\|_{B_t^+(0)} = 0$ pour $t \in (0, r)$ fixé. Par conséquent

$$h = \sum_{n \geqslant 0} \frac{1}{d^{n+1}} \ln(1 + g \circ f^n) \in \mathcal{H}(B_r(0)),$$

satisfait $d \cdot h = \ln(1+g) + h \circ f$. Comme $\|h\|_{B_r(0)} < p^{-1/(p-1)}$ on a $\varphi(z) = z \exp(h(z)) \in \mathcal{H}(B_r(0))$ et

$$\varphi^d(z) = (z \exp(h_{n+1}(z)))^d = f(z) \exp(h_n \circ f(z)) = \varphi \circ f(z). \qquad \square$$

3.2. Dynamique quasi-périodique et logarithme itératif. — Dans cette section on considère les aspects locaux de la dynamique quasi-périodique d'un système dynamique donné.

A chaque système dynamique (f, U) on associe son *domaine de quasi-périodicité*, que l'on note $\mathcal{E}(f)$. Dans cette section on s'intéresse aux aspects locaux de la dynamique de f sur $\mathcal{E}(f)$. Par conséquent on n'utilise que des propriétés élémentaires d'analyse et plus concrètement les outils de la Section 1.3.1. Des aspects globaux de la dynamique sur le domaine de quasi-périodicité sont étudiés dans les Sections 4.2 et 5.

Un des modèles de la dynamique quasi-périodique est la translation $z \mapsto z+1$. Supposons donc que f est localement conjugué à $z \mapsto z+1$; c'est-à-dire qu'il existe une fonction analytique h au voisinage d'un point z_0 telle que $h'(z_0) \neq 0$ et $h \circ f(z) = h(z) + 1$. Alors notons que

$$\lim_{|n|_p \to 0} \frac{f^n(z) - z}{n} = \lim_{|n|_p \to 0} \frac{h^{-1}(h(z)+n) - h^{-1}(h(z))}{n} = \frac{1}{h'(z)}.$$

Ceci explique la définition suivante.

Définition 3.7. — On appelle

$$\mathcal{E}(f) = \Big\{ z_0 \in U \mid \text{ il existe } k = k(z_0) \geqslant 1, \text{ tel que } \Big\{ \frac{f^{nk} - \mathrm{id}}{nk} \Big\}_{n \geqslant 0}$$
$$\text{est uniformément convergente sur un voisinage de } z_0, \text{ quand } |n|_p \to 0 \Big\}$$

le *domaine de quasi-périodicité* de f. De plus la fonction $f_* : \mathcal{E}(f) \to \mathbb{C}_p$, définie par

$$f_*(z_0) = \lim_{|n|_p \to 0} \frac{f^{nk}(z_0) - z_0}{nk},$$

pour $z_0 \in \mathcal{E}(f)$, est appelé le *logarithme itératif* de f.

Ainsi définie f_* est une fonction analytique ; voir aussi Corollaire 4.17. Notons que la définition de f_* ressemble à la définition du nombre de rotation d'un homéomorphisme du cercle. Dans ce dernier cas il faut considérer le relèvement sur $\mathbb{R}$ et $n \to \infty$ au lieu de $|n|_p \to 0$. Donc on peut dire que pour $z_0 \in \mathcal{E}(f)$, $f_*(z_0)$ est le 'nombre de rotation infinitésimal de f en z_0'.

Remarque 3.8. — La notion de logarithme itératif formel est connue dans la littérature sous divers autres noms ; voir les références en [**Ec1**] et aussi la très complète référence de [**Ku**] pour ce sujet et autres. Ici on considère l'approche d'Ecalle ; voir [**Ec1**]. Il a considéré le logarithme itératif pour les séries formelles sur $\mathbb{C}$ tangentes à l'identité et il a étudié des problèmes de convergence. Lubin a considéré le logarithme itératif défini sur les anneaux à valuation discrète au voisinage des points fixes paraboliques et il a montré 1 et 2 de la Proposition 3.16 (ci-dessous), dans ce contexte par une méthode différente ; voir [**Lu1**].

Proposition 3.9. — *On a les propriétés suivantes du domaine de quasi-périodicité.*

(i) $\mathcal{E}(f)$ *est ouvert.*

(ii) f *est injective sur* $\mathcal{E}(f)$.

(iii) $\mathcal{E}(f)$ *est invariant par* f.

(iv) $\mathcal{E}(f^n) = \mathcal{E}(f)$ *et* $(f^n)_* = nf_*$, *pour tout* $n \geqslant 1$.

(v) *Si* $g = h \circ f \circ h^{-1}$ *alors* $\mathcal{E}(g) = h(\mathcal{E}(f))$ *et* $g_* \circ h = h' f_*$. *En particulier pour tout entier* $m \geqslant 1$ *on a* $f_* \circ f^m \equiv f_*(f^m)'$.

Démonstration

(i) Suit directement de la définition.

(ii) Soient z_0 et $z_1 \in \mathcal{E}(f)$ différents. Alors il existe $r > 0$ et $k \geqslant 1$ tels que la suite $\{\frac{f^{nk}-\mathrm{id}}{nk}\}_{n\geqslant 0}$ est uniformément convergente sur $B_r(z_0)$ et $B_r(z_1)$. De plus on peut supposer que $B_r(z_0) \cap B_r(z_1) = \varnothing$. Donc pour n de norme p-adique assez petit on a $f^{nk}(z_0) \in B_r(z_0)$ et $f^{nk}(z_1) \in B_r(z_1)$. Par conséquent $f^{nk}(z_0) \neq f^{nk}(z_1)$ et $f(z_0) \neq f(z_1)$.

(iii) Il suffit de montrer que, si $\{\frac{f^{nk}-\mathrm{id}}{nk}\}_{n>0}$ est uniformément convergente sur un voisinage de 0, alors $\{\frac{f^{kn}-\mathrm{id}}{nk} \circ f\}_{n>0}$ aussi. Notons que pour $n \geqslant 0$ on a

$$\frac{f^{kn} \circ f - f}{kn} = \frac{f \circ f^{kn} - f}{kn} = \frac{f(\mathrm{id} + (f^{kn} - \mathrm{id})) - f}{kn}.$$

Donc $\{\frac{f^{kn}-\mathrm{id}}{kn} \circ f\}_{n\geqslant 0}$ converge uniformément vers $f' \circ f_*$ sur un voisinage de z_0.

(iv) Immédiat.

(v) On observe simplement que

$$g_* = \lim_{|n|_p \to 0} \frac{g^{kn} \circ h - h}{kn} = \lim_{|n|_p \to 0} \frac{h \circ f^{kn} - h}{f^{kn} - \mathrm{id}} \frac{f^{kn} - \mathrm{id}}{kn} = h' f_*. \qquad \square$$

Exemple 3.10. — Soit $f(z) = \dfrac{z}{1 - az}$. Alors

$$\frac{f^n(z) - z}{n} = \frac{az^2}{1 - naz}.$$

Donc $f_*(z) = az^2$ et $\mathcal{E}(f) = \mathbb{P}(\mathbb{C}_p)$.

Soit $f(z) = \lambda z$ avec $|\lambda| = 1$; on a

$$\frac{f^n(z) - n}{n} = \frac{\lambda^n - 1}{n} z.$$

Donc $f_* \equiv (\ln \lambda) z$ et $\mathcal{E}(f) = \mathbb{P}(\mathbb{C}_p)$. Il n'est pas difficile de voir que, si R est une fonction rationnelle de degré un telle que $\mathcal{E}(R) \neq \varnothing$, alors R est conjuguée à une de ces fonctions.

3.2.1. Automorphismes d'un disque et logarithme itératif. — Pour $w \in \mathbb{C}_p$ et $k \geqslant 0$ soit

$$\binom{w}{k} = \frac{w(w-1)\cdots(w-(k-1))}{1 \cdots k}$$

et considérons la fonction continue $\rho : (0, \infty) \to (0, \infty)$ définie par $\rho(s) = sp^{1/(p-1)}$, pour $s \geqslant 1$, $\rho(p^{-n}) = p^{1/p^n(p-1)}$, pour $n \geqslant 0$ et de la forme $a_n s^{b_n}$ si $p^{-(n+1)} \leqslant s \leqslant p^{-n}$. Notons que $\rho(s) \to 1$ quand $s \to 0$.

Lemme 3.11. — *Soit $R > 0$ et $f \in \mathcal{H}(B_R(0))$ un automorphisme de $B_R(0)$ tel que $\| f - \mathrm{id}\,\|_{B_R(0)} < \gamma R$, où $\gamma \in (0,1)$. Alors on a les propriétés suivantes.*

(i) $\mathcal{E}(f) = B_R(0)$.

(ii) *Posons* $T_0 = \mathrm{id}$ *et pour* $n \geqslant 0$ *posons* $T_{n+1} = T_n \circ f - T_n$ *; donc* $T_n = \sum_{i=0}^{n}(-1)^i\binom{n}{i} f^{n-i}$. *Pour* $w \in \mathbb{C}_p$ *tel que* $\rho(|w|) < \gamma^{-1}$ *on a* $\|T_n\|_{B_R(0)} \leqslant R\gamma^n$. *De plus la série*

$$f^{\circ w} = \sum_{i=0}^{m} \binom{w}{i} T_i,$$

converge uniformément dans $B_R(0)$ vers un automorphisme de $B_R(0)$. De plus si $\rho(|w_i|) < \gamma^{-1}$, pour $i = 1, 2$, alors $f^{\circ w_1} \circ f^{\circ w_2} = f^{\circ(w_1+w_2)}$.

(iii) *Pour tout $w \in \mathbb{C}_p$ avec $\rho(|w|) < \gamma^{-1}$ on a*

$$\left\| f_* - \frac{f^{\circ w} - \mathrm{id}}{w} \right\|_{B_R(0)} \leqslant CR|w| \quad et \quad \left\| f_* - \sum_{1 \leqslant i \leqslant k} (-1)^{i-1} \frac{T_i}{i} \right\|_{B_R(0)} \leqslant k\gamma^k R,$$

où la constante C ne dépend que de $\gamma\rho(|w|)$. En particulier

$$(f^n - \mathrm{id})/n \quad et \quad \sum_{i>0} (-1)^{i-1} \frac{T_i}{i}$$

convergent uniformément vers f_* sur $B_R(0)$ quand $|n|_p \to 0$.

La démonstration du Lemme 3.11 est à la fin de cette section.

Corollaire 3.12. — *Si $f : B_R(0) \to B_R(0)$ est un automorphisme, alors $\mathcal{E}(f) = B_R(0)$. Si de plus f est tangent à l'identité en 0, alors pour tout $w \in \mathbb{C}_p$ la série $f^{\circ w}$ est convergente au voisinage de 0.*

Démonstration. — Comme f est un automorphisme de $B_R(0)$ il existe $|\lambda| = 1$ tel que $|f'(z) - \lambda| < 1$ pour tout $z \in B_R(0)$. Par conséquent, quitte à remplacer f par un itéré, on peut supposer $|f'(0) - 1| < 1$. Alors pour tout $r \in (|f(0)|, R)$ il existe $\gamma = \gamma(r) \in (0,1)$ tel que $\|f - \mathrm{id}\,\|_{B_r(0)} < \gamma r$. Donc par (i) du lemme on a $B_r(0) \subset \mathcal{E}(f)$.

Si f est tangent à l'identité, pour tout $w \in \mathbb{C}_p$ on peut trouver $R > 0$ petit tel que $\|f - \mathrm{id}\,\|_{B_R(0)} < \gamma R$ avec $\gamma < \rho(|w|)^{-1}$. Alors on peut appliquer le lemme. □

Notons que la formule du logarithme itératif en (iv) du lemme ressemble celle du logarithme usuel

$$\ln x = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \cdots .$$

Notons aussi l'analogie entre la définition de f_* et l'identité suivante du logarithme p-adique

$$\ln(\lambda) = \lim_{|n|_p \to 0} \frac{\lambda^n - 1}{n}.$$

Les séries dans (ii) sont les *itérés fractionnaires* de f ; voir par exemple [**Ec1**]. De la formule

$$f_* = \lim_{w\to 0} \frac{f^{\circ w} - \mathrm{id}}{w}$$

on peut penser à f_* comme un champ de vecteurs et $f^{\circ w}$ comme le temps w du flot engendré par f_*. Ceci est justifié aussi par la formule de transformation (v) de la Proposition 3.9.

La démonstration du Lemme 3.11 dépend du lemme suivant.

Lemme 3.13. — *Pour tout $w \in \mathbb{C}_p$ et $k \geqslant 1$ on a*

$$\left|\binom{w}{k}\right| \leqslant k|w|\rho(|w|)^k \quad \textit{et} \quad \left|\frac{1}{w}\binom{w}{k} - \frac{(-1)^{k-1}}{k}\right| \leqslant k^2|w|\rho(|w|)^k.$$

Démonstration. — Si $|w| \geqslant 1$ alors il est facile de voir que

$$\left|\binom{w}{k}\right| \leqslant \frac{|w|^k}{|k!|} \leqslant |w|^k (p^{1/(p-1)})^k.$$

Donc on obtient les deux inégalités dans ce cas. Si $p^{-(n+1)} \leqslant |w| \leqslant p^{-n}$ alors notons que $|1 - w/i| \leqslant 1$ si $|i| \geqslant p^{-n}$ et $|1 - w/i| \leqslant |w| \cdot |i|^{-1}$ sinon. Donc on a

$$\begin{aligned}\left|(1-w)\cdot\dots\cdot\left(1 - \frac{w}{k-1}\right)\right| &\leqslant \left|\frac{w}{p^{n+1}}\right|^{k/p^{n+1}} \left|\frac{1}{p}\right|^{k/p^{n+2}+k/p^{n+3}+\dots} \\ &= (|w|p^{n+1})^{k/p^{n+1}} p^{k/p^{n+1}(p-1)} = \rho(|w|)^k,\end{aligned}$$

car il y a au plus k/p^m entiers dans $\{1, \dots, k\}$ que sont divisibles par p^m. Par conséquent

$$\left|\binom{w}{k}\right| = \left|\frac{w}{k}(1-w)\cdot\dots\cdot\left(1-\frac{w}{k-1}\right)\right| \leqslant k|w|\rho(|w|)^k.$$

D'autre part

$$\frac{1}{w}\binom{w}{k} - \frac{(-1)^{k-1}}{k} = \frac{(-1)^k}{k}\sum_{1\leqslant i\leqslant k-1}\frac{w}{i}\prod_{1\leqslant i\leqslant k-1,\, j\neq i}\left(1 - \frac{w}{j}\right).$$

Par le raisonnement précédent on a $\left|\prod_{1\leqslant i\leqslant k-1,\, j\neq i}(1 - w/j)\right| \leqslant \rho(|w|)^k$. Donc

$$\left|\frac{1}{w}\binom{w}{k} - \frac{(-1)^{k-1}}{k}\right| \leqslant k^2|w|\rho(|w|)^k. \qquad \square$$

Démonstration du Lemme 3.11. — On note $\|\cdot\|_R$ au lieu de $\|\cdot\|_{B_R(0)}$.

(ii) Par le Corollaire 1.4 du Lemme de Schwarz, pour tout $n \geqslant 1$ on a

$$\|T_{n+1}\|_R = \|T_n \circ f - T_n\|_R \leqslant \frac{1}{R}\|f - \mathrm{id}\,\|_R \|T_n\|_R \leqslant \gamma\|T_n\|_R.$$

Donc par induction on a $\|T_n\|_R \leqslant R\gamma^n$. Par le Lemme 3.13 $|\binom{w}{n}| \leqslant n|w|\rho(|w|)^n$, par conséquent $\|\binom{w}{n}T_n\|_R \leqslant n|w|(\gamma\rho(|w|))^n \to 0$, quand $n \to \infty$. Donc $f^{\circ w}$ est un automorphisme de $B_R(0)$.

On a l'identité $f^{\circ w_1} \circ f^{\circ w_2} = f^{\circ(w_1+w_2)}$ au sens formel ; voir [**Ec1**]. Donc si les séries sont convergentes on a aussi cette identité au sens analytique.

(iii) Par (ii) la série $\sum_{k\geqslant 1}(-1)^{k+1}T_k/k$ est convergente. Soit $n \geqslant 0$ tel que $p^{-(n+1)} < |w| \leqslant p^{-n}$. Par le lemme précédent on a

$$\left\| \frac{f^w - \mathrm{id}}{w} - \sum_{k=1}^{\infty}(-1)^{k+1}\frac{T_k}{k} \right\|_R = \left\| \sum_{k=2}^{\infty}\left(\frac{1}{w}\binom{w}{k} - \frac{(-1)^{k+1}}{k}\right)T_k \right\|_R$$
$$\leqslant \left(\max_{k\geqslant 1} k(\gamma\rho(|w|))^k\right) R|w| \leqslant CR|w|,$$

où C ne dépend que en $\gamma\rho(|w|)$. Par conséquent $(f^{\circ w} - \mathrm{id})/w$ converge uniformément sur $B_R(0)$, mais par définition

$$f_* = \lim_{|n|_p\to 0}\frac{f^n - \mathrm{id}}{n} = \lim_{|n|_p\to 0}\frac{f^{\circ n} - \mathrm{id}}{n} = \sum_{k\geqslant 1}(-1)^n\frac{T_k}{k}.$$

Donc on obtient la première inégalité du lemme ; la deuxième suit de (ii).

(i) Suit de (iii). □

3.2.2. Propriétés du domaine de quasi-périodicité

Proposition 3.14. — *On a les caractérisations suivantes du domaine de quasi-périodicité.*

(1) $\mathcal{E}(f) = \{z_0 \in U \mid$ *il existe* $n_j \to \infty$ *tel que* $f^{n_j} \to \mathrm{id}$ *sur un voisinage de* $z_0\}$.

(2) $\mathcal{E}(f) = \{z_0 \in U \mid \left\{\frac{f^{nk}-\mathrm{id}}{nk}\right\}_{n\geqslant 0}$ *est localement uniformément bornée*$\}$.

Démonstration

(1) Si $z_0 \in \mathcal{E}(f)$, alors il existe $k \geqslant 1$, $r > 0$ et $C > 0$ tels que $|f^{kp^n}(z) - z| < C1/p^n$, pour $z \in B_r(z_0)$. Alors la suite $\{f^{kp^n}\}_{n>0}$ converge uniformément vers l'identité sur $B_r(z_0)$. Supposons que f^{n_j} converge uniformément vers l'identité sur $B_r(z_0)$. Alors, pour j assez grand on a $f^{n_j}(B_r(z_0)) \subset B_r(z_0)$ et $|(f^{n_j})'(z_0)| = 1$. Donc par le Corollaire 3.12 on a $z_0 \in B_r(z_0) \subset \mathcal{E}(f)$.

(2) Si $z_0 \in \mathcal{E}(f)$, alors la suite $\{\frac{f^nk-\mathrm{id}}{nk}\}_{n>0}$ est uniformément bornée sur un voisinage de z_0. Supposons que $\{\frac{f^{nk}-\mathrm{id}}{nk}\}_{n>0}$ est uniformément bornée sur un voisinage de z_0. Alors $\{f^{kp^n}\}_{n>0}$ converge uniformément vers l'identité sur un voisinage de z_0, et par 1 on a $z_0 \in \mathcal{E}(f)$. □

Corollaire 3.15. — *Tout* $z \in \mathcal{E}(f)$ *est récurrent par* f.

Démonstration. — Par la caractérisation 1. □

Proposition 3.16. — *On a les propriétés suivantes du logarithme itératif.*

(1) *Soit* $z_0 \in \mathcal{E}(f)$. *On a* $f_*(z_0) = 0$ *si et seulement si* z_0 *est un point périodique indifférent de* f. *Si* k *est une période de* z_0, *alors* $f_*'(z_0) = \frac{1}{k}\ln(f^k)'(z_0)$. *En particulier* z_0 *est parabolique si et seulement si* $f_*'(z_0) = 0$.

(2) *Pour tout* $z_0 \in \mathcal{E}(f)$ *tel que* $f_*(z_0) \neq 0$ *il existe* $k \geqslant 1$ *tel que* f^k *est conjugué analytiquement à la translation* $z \mapsto z + k$ *sur un voisinage de* z_0. *Plus précisément pour tout* z *sur un voisinage* U_{z_0} *de* z_0 *on a*

$$f^*(f^k(z)) = f^*(z) + k,$$

où $f^* : U_{z_0} \to B^+_{|k|}(0)$ *est une primitive formelle injective de* $1/f_*$.

Corollaire 3.17. — *Les points périodiques indifférents sont isolés.*

Démonstration. — Par 1 les points périodiques indifférents sont les zéros de la fonction analytique f_*, donc sont isolés. □

Démonstration de la Proposition 3.16

(1) Si z_0 est périodique on a $f_*(z_0) = 0$. Supposons que $f_*(z_0) = 0$. On peut supposer de plus $k(z_0) = 1$, $U_{z_0} = B^+_r(z_0)$, avec $r \in |\mathbb{C}_p|$, et $f(B^+_r(z_0)) \subset B^+_r(z_0)$.

On considère la norme uniforme $\|\cdot\|_r = \|\cdot\|_{B^+_r(z_0)}$ sur $\mathcal{H}(B^+_r(z_0))$; voir Section 1.3.1. On pose $f^{p^n} = \mathrm{id} + g_n$, où $\|g_n\|_r < Cp^{-n}$. Donc

$$f^{kp^n}(z) = z + g_n(f^{p^n}(z)) + \cdots + g_n(f^{(k-1)p^n}(z)) = z + kg_n(z) + r_n(z),$$

où $\|r_n\|_r \leqslant \frac{1}{r}\|g_n\|_r^2$, par le corollaire du lemme de Schwarz. Par conséquent, si n est assez grand on a $\|g_{n+1}\|_r = \frac{1}{p}\|g_n\|_r$.

Alors, soit $g_n \equiv 0$ et tout point de $B_r(z_0)$ est fixé par f^n, soit $f_* \not\equiv 0$ sur $B^+_r(z_0)$ et alors f_* a un nombre fini de zéros sur $B^+_r(0)$. Par (v) de la Proposition 3.9 on a $f_*(f^n(z_0)) = (f^n)'(z_0)f_*(z_0) = 0$ et par conséquent z_0 est prépériodique par f. Comme f est injective sur $B_r(z_0) \subset \mathcal{E}(f)$ on a que z_0 est périodique par f.

(2) Soit $z_0 \in \mathcal{E}(f)$ tel que $f_*(z_0) \neq 0$ et soit f^* une primitive formelle de $1/f_*$ définie sur un voisinage de z_0 et telle que $f^*(z_0) = 0$. Comme $(f^*)'(z_0) = 1/f_*(z_0) \neq 0$, il existe $n > 0$ et une boule U_{z_0} que contient z_0 tel que $f^* : U_{z_0} \to B_{p^{-n}}(0)$ est une bijection.

Soit m tel que $f^m(z_0) \in U_{z_0}$. Par (v) de la Proposition 3.9 on a $f_* \circ f^m \equiv f_*(f^m)'$. Par conséquent $(f^* \circ f^m)' \equiv (f^*)'$ sur U_{z_0}. Donc il existe $w \in B_{p^{-n}}(0)$ tel que $f^* \circ f^m \equiv f^* + w$. Alors

$$f_*(z_0) = \lim_{|n|_p \to 0} \frac{f^{mn}(z_0) - z_0}{mn} = \lim_{|n|_p \to 0} \frac{(f^*)^{-1}(nw) - (f^*)^{-1}(0)}{mn} = \frac{w}{m} f_*(z_0).$$

Comme $f_*(z_0) \neq 0$ on a $w = m$. Par conséquent $f^* \circ f^m \equiv f^* + m$ sur U_{z_0}. □

3.3. Points périodiques indifférents. — Dans cette section on considère la dynamique au voisinage des points périodiques indifférents. Il suffit bien sûr de considérer le cas des points fixes. On considère le groupe (pour la composition) $I(\mathbb{C}_p)$ des séries f à coefficients dans $\mathbb{C}_p$ de la forme

$$f(z) = \lambda z + a_1 z^2 + a_2 z^3 + \cdots, \quad \text{où } |\lambda| = 1$$

qui convergent au voisinage de 0. Comme dans la Section 3.1 on définit, pour $f(z) \neq \lambda z$,

$$r(f) = \Big(\sup_{k \geqslant 1} |a_k|^{1/k}\Big)^{-1} < \infty.$$

Proposition 3.18. — *Soit $f \in I(\mathbb{C}_p)$ une série différente de λz. Alors f est convergente sur $B_{r(f)}(0)$ et on a les propriétés suivantes.*

(i) $B_{r(f)}(0) \subset \mathcal{E}(f)$.
(ii) $r(f) = \sup\{r | \, f$ *est convergente sur* $B_r(0)$ *et* $f(B_r(0)) \subset B_r(0)\}$.

Démonstration. — La preuve de (ii) est similaire à celle de la Proposition 3.2 de la Section 3.1. L'assertion (i) résulte du corollaire du Lemme 3.11 de la section précédente. □

On s'intéresse d'abord aux points périodiques de f dans $B_{r(f)}(0)$. Dans la Section 3.3.1 on considère la structure du groupe $I(\mathbb{C}_p)$. On verra que si $R \in I(\mathbb{C}_p)$ est une fonction rationnelle, alors on a $r(R) \in |\mathbb{C}_p|$; voir Lemme 4.11 dans la Section 4.1.

Considérons dans la suite une série $f \in I(\mathbb{C}_p)$ telle que $r(f) \in |\mathbb{C}_p|$. Après changement de coordonnée on peut supposer que $r(f) = 1$. Alors $f \in \mathcal{O}_K[[z]]$, c'est-à-dire f est à coefficients entiers. Quitte à changer f par un itéré on peut supposer en plus que $|f'(0) - 1| < 1$. On verra que dans ce cas tous les points périodiques de f dans $B_{r(f)}(0)$ ont une période primitive de la forme p^n, où $n \geqslant 0$; voir [**Li1**].

Rappelons que le degré de Weierstrass d'une série $g(z) = a_0 + a_1 z + \cdots \in \mathcal{O}_K[[z]]$, noté wideg$(g)$, est le plus petit entier $d \geqslant 0$ tel que $|a_d| = 1$. Si d est fini alors d est égal au nombre des zéros de g dans $\{|z| < 1\}$, comptés avec multiplicité. De façon équivalente wideg(g) est égal à $ord_{\widetilde{g}}(0)$ où $\widetilde{g}$ dénote la réduction de g. Par conséquent pour $n \geqslant 1$ le nombre

$$\mathrm{wideg}(f^n - \mathrm{id}) = ord_{\widetilde{f}^n - \mathrm{id}}(0) \in \{1, 2, \ldots\} \cup \{\infty\}$$

est égal au nombre de points fixes de f dans $\{|z| < 1\}$. Donc pour connaître le nombre des points périodiques de f dans $\{|z| < 1\}$ d'une période primitive donnée, il suffit d'étudier la suite $\{\mathrm{wideg}(f^n - \mathrm{id})\}_{n \geqslant 1}$.

La condition $|f'(0) - 1| < 1$ est équivalente à ce que $\widetilde{f}$ soit tangente à l'identité en zéro. De plus notons que si $\widetilde{f}(z) = z + a_m z^{m+1} + \cdots$, alors wideg$(\widetilde{f} - \mathrm{id}) = m + 1$ et pour tout entier k,

$$\widetilde{f}^k(z) = z + \widetilde{k} a_m z^{m+1} + \cdots \ .$$

Par conséquent wideg$(\widetilde{f}^k - \mathrm{id}) =$ wideg$(\widetilde{f} - \mathrm{id})$, pour tout entier k qui n'est pas divisible par p. Donc tous les points périodiques de f dans $\{|z| < 1\}$ de période k sont des points fixes. Plus généralement tous les points périodiques de f dans $\{|z| < 1\}$ de période kp^n, avec $n \geqslant 0$ et k qui n'est pas divisible par p, sont aussi de points périodiques de période p^n.

Par conséquent tous les points périodiques de f dans $\{|z| < 1\}$ ont une période primitive de la forme p^n, avec $n \geqslant 0$. Donc il suffit d'étudier la suite

$$i_m = \mathrm{wideg}(f^{p^m} - \mathrm{id}), \text{ pour } m \geqslant 0.$$

***Théorème* (Sen [Se]).** — *Soit $f \in \mathcal{O}_{\mathbb{C}_p}[[x]]$ un série telle que $|f'(0) - 1| < 1$. Si $n > 0$ est tel que $i_n < \infty$, alors $i_n \equiv i_{n-1} \mod (p^n)$.*

Voir aussi [**Lu2**] et [**Li1**]. Le théorème suivant de Keating [**Ke**] décrit des cas où l'on peut connaître explicitement la suite i_m. On utilise ce théorème dans l'exemple 5.15 de la Section 5.2.

***Théorème* (Keating [Ke]).** — *Soit $f \in \mathcal{O}_{\mathbb{C}_p}[[x]]$ une série telle que $i_0 = 2$ et $i_1 = 2 + bp$ avec $0 < b < p - 1$. Alors*

$$i_m = 2 + bp + bp^2 + \cdots + bp^m,$$

pour tout $m \geqslant 1$.

Exemple 3.19. — Considérons la série $g(z) = z + z^2 + az^3 + \cdots \in \mathbb{F}_p[[z]]$. On va montrer que si $p > 2$ alors $g^p(z) = z - (a-1)z^{p+2} + \cdots$. Par conséquent si une série $f \in \mathcal{O}_{\mathbb{C}_p}[[z]]$ a g comme réduction et $a \neq 1$, alors f vérifie les hypothèses du Théorème de Keating avec $b = 1$. Donc $i_m = 1 + (1 + p + \cdots + p^m)$ dans ce cas.

On suppose alors que $p > 2$ et on pose $T(\varphi) = \varphi \circ f - \varphi$, qui est linéaire en φ. Soit $\varphi_0(z) = z$ et $\varphi_{k+1} = T^k(\varphi_0)$ de telle façon que $\varphi_1(z) = z^2$ et $g^p(z) = z + \varphi_p(z)$. Pour $l \geqslant 1$ on a,

$$T(z^l) = (z(1 + z + az^2 + \cdots))^l - z^l = lz^{l+1} + \Big(\frac{l(l-1)}{2} + la\Big)z^{l+2},$$

et en particulier $T(z^p) = z^{2p} + \cdots$. Donc $\varphi_k(z) = k!z^{k+1} + \cdots$, pour $0 \leqslant k < p$, et si on pose $\varphi_{p-2}(z) = (p-2)!z^{p-1}z^{p-1} + \alpha z^p + \cdots$ on a,

$$\begin{aligned}\varphi_{p-1}(z) &= T(\varphi_{p-2}(z)) = (p-2)!\,T(z^{p-1}) + \alpha T(z^p) + \cdots \\ &= (p-1)!z^p + (p-1)!\Big(\frac{p-2}{2} + a\Big)z^{p+1} + \cdots .\end{aligned}$$

Comme $p > 2$ on a $(p-2)/2 = -1$, $2p > p + 2$ et $(p-1)! = 1$. Donc

$$\varphi_p(z) = T(\varphi_{p-1}(z)) = -(a-1)z^{p+2} + \cdots .$$

Exemple 3.20. — Considérons le polynôme $P(z) = z + z^p \in \mathbb{F}_p[z]$. On peut montrer par induction

$$P^n(z) = z + \binom{n}{1}z^p + \cdots + \binom{n}{n}z^{p^n}.$$

Par conséquent $\mathrm{ord}(P^n(z) - z)$ est égal à p^k, la plus grande puissance de p qui divise n. En particulier $\mathrm{ord}(P^{p^m} - z) = p^{p^m}$.

Donc une série $f \in \mathcal{O}_{\mathbb{C}_p}[[x]]$ telle que $\widetilde{f} = P$ vérifie $i_n = \mathrm{wideg}(f^{p^n} - \mathrm{id}) = p^{p^n}$. Si $p = 2$ on a $i_0 = 2$, $i_1 = 4 = 2 + 2b$ avec $b = 1$ et

$$i_m = 2^{2^m} \gg 2^{m+1} = 2 + b(2 + 2^2 + \cdots + 2^m).$$

Si f est à coefficients dans un corps algébrique K et $f(z) \neq \lambda z$ alors f a un nombre fini de points périodiques dans $\{z \in K \mid |z| < 1\}$. En effet, dans ce cas la complétion $\widehat{K}$ de K par $|\cdot|$ est localement compacte, donc $\{z \in \widehat{K} \mid |z| < 1\} \subset \mathcal{E}(f)$ est compact et comme les points périodiques de f dans $\mathcal{E}(f)$ coïncident avec les zéros de $f_* \in \mathcal{H}(\mathbb{D}_p)$ il ne peut avoir qu'un nombre fini de points périodiques de f dans $\{z \in K \mid |z| < 1\}$.

De plus on a les résultats suivants de Lubin ([**Lu1**], Corollaire 4.3.1) et Li ([**Li1**], Théorème 4.3).

***Théorème* (Lubin [Lu1])**. — *Soient $K \subset \mathbb{C}_p$ est un corps à valuation discrète et $f \in \mathcal{O}_K[[x]]$ est une série telle qu'il existe $n > 0$ avec $\mathrm{wideg}(f^n - \mathrm{id}) = \infty$ et $f^n \not\equiv \mathrm{id}$. Alors f a un nombre fini de points périodiques dans $\{z \in K \mid |z| \leqslant 1\}$.*

***Théorème* (Li [Li1])**. — *Soient $K \subset \mathbb{C}_p$ un corps à valuation discrète et $f \in \mathcal{O}_K[[x]]$ une série telle que $|f'(0) - 1| < 1$. Si f a un point périodique de période primitive p^m alors $p^m - p^{m-1} \leqslant e$, où e est le degré de ramification de K sur $\mathbb{Q}_p$.*

Remarque 3.21

(1) En général, si $f \in I(\mathbb{C}_p)$ on peut avoir $r(f^n) > r(f)$. Par exemple la série $f(z) = -z(1 + 2z + 4z^2 + \cdots) = z/(2z - 1)$ est telle que $f^2 = \mathrm{id}$ donc, $r(f^2) = \infty > |1/2|_p = r(f)$.

(2) Comme le montre $f(z) = z + z^2 + z^3 + \cdots = z/(1 - z)$, en général $B_{r(f)}(0)$ ne peut contenir que 0 comme point périodique de f.

3.3.1. Conjugaison locale des points périodiques indifférents. — La classe de conjugaison et les centralisateurs du groupe $I(\mathbb{C}_p)$ ont une structure très simple et bien connue ; voir [**Lu1**]. C'est la même structure que pour le groupe formel correspondant : pour tous f et $g \in I(\mathbb{C}_p)$, f est localement conjugué à g si et seulement si f est formellement conjugué à g. Dans cette section on rappelle brièvement les principaux résultats.

L'étude des classes de conjugaison du groupe $I(\mathbb{C})$ est beaucoup plus subtile ; voir [**Ec1**], [**Vo**] et [**Y**]. Par exemple toute série $f \in I(\mathbb{C}_p)$ telle que $f'(0)$ n'est pas une racine de l'unité est localement linéarisable ; voir par exemple [**HY**] et [**Lu1**], [**TVW**], [**AV**]. Dans le cas complexe ce n'est pas toujours le cas et il y a des problèmes de petit diviseurs assez délicats ; voir [**Y**].

Pour $\lambda \in \mathbb{C}_p$ avec $|\lambda| = 1$ on définit $\mathrm{ord}(\lambda)$ comme le plus petit entier positif tel que $\lambda^n = 1$, si λ est une racine de l'unité et $\mathrm{ord}(\lambda) = \infty$ sinon. Soit $f \in I(\mathbb{C}_p)$ et on pose $\lambda = f'(0)$. Si $\mathrm{ord}(\lambda) = q < \infty$ on définit $\mathrm{valit}(f) = \infty$ si $f^q \equiv \mathrm{id}$ et $\mathrm{valit}(f) = n \geqslant 1$

sinon, où n est l'entier tel que

$$f^q(z) = z + az^{n+1} + \cdots, \text{ avec } a \neq 0.$$

L'entier n est alors un multiple de q. Dans ce dernier cas on a $f_*(z) = (a/q)z^{n+1} + \cdots$ et on note $\mathrm{resit}(f)$ le résidu de $1/f_*$ en 0. Soient f et $g \in I(\mathbb{C}_p)$ avec $\lambda = f'(0)$.

- Si $\mathrm{ord}(\lambda) = \infty$ alors f et g sont conjugués si et seulement si $g'(0) = \lambda$.
- Si $\mathrm{ord}(\lambda) = q < \infty$ et $f^q \equiv \mathrm{id}$, alors f et g sont conjugués si et seulement si $g'(0) = \lambda$.
- Si $\mathrm{ord}(\lambda) = q < \infty$ et $\mathrm{valit}(f) < \infty$ alors f et g sont conjugués si et seulement si $g'(0) = \lambda$, $\mathrm{valit}(f) = \mathrm{valit}(g)$ et $\mathrm{resit}(f) = \mathrm{resit}(g)$.

Dans ce dernier cas f est conjugué a un unique polynôme de la forme $\lambda z(1+z^n+az^{2n})$, avec $a \in \mathbb{C}_p$. Soient $f \in I(\mathbb{C}_p)$, $m \geqslant 1$ et $\zeta \in \mathbb{C}_p$ tels que $\zeta^m = f'(0)$.

- Si $\mathrm{ord}(\zeta) = \infty$ il existe un unique $g \in I(\mathbb{C}_p)$ tel que $g'(0) = \zeta$ et $g^m = f$.
- Si $\mathrm{ord}(\zeta) = q < \infty$ et $f^q \equiv \mathrm{id}$ il existe $g \in I(\mathbb{C}_p)$ tel que $g'(0) = \zeta$ et $g^m = f$.
- Si $\mathrm{ord}(\zeta) = q < \infty$ et $\mathrm{valit}(f) < \infty$ alors il existe $g \in I(\mathbb{C}_p)$ tel que $g'(0) = \zeta$ et $g^m = f$ si et seulement si $q|\mathrm{valit}(f)$. La série g est alors unique.

Soient f, $g \in I(\mathbb{C}_p)$.

- Si f_*, $g_* \not\equiv 0$ alors f et g commutent si et seulement s'il existe $w \in \mathbb{C}_p - \{0\}$ tel que $f_* = wg_*$.
- Si f est linéarisable, son centralisateur se déduit par conjugaison de celui de sa partie linéaire.
- Si $\mathrm{ord}(f'(0)) = q < \infty$ et $\mathrm{valit}(f) = \infty$ posons $\overline{f} = f^q$ et définissons pour $\mathrm{ord}(\zeta)|\mathrm{valit}(f)$ la série $f_\zeta \in I(\mathbb{C}_p)$ par $f'_\zeta(0) = \zeta$ et $f_\zeta^{ord(\zeta)} = \overline{f}$. Alors $g \in I(\mathbb{C}_p)$ commute avec f si et seulement s'il existe ζ (avec $\mathrm{ord}(\zeta)|\mathrm{valit}(f)$) et $w \in \mathbb{C}_p$ tels que

$$g = f_\zeta \circ \overline{f}^{\circ w} = \overline{f}^{\circ w} \circ f_\zeta.$$

4. Dynamique des fonctions rationnelles

Cette partie est dédiée à l'étude de la dynamique globale d'une fonction rationnelle.

Dans la Section 4.1 on étudie les bassins d'attraction et dans la Section 4.2 on étudie les propriétés globales du domaine de quasi-périodicité d'une fonction rationnelle.

On a la partition usuelle en ensembles de Fatou et de Julia comme dans le cas complexe ; voir Section 4.3. Dans la Section 4.4 on montre le Théorème de Classification caractérisant la dynamique de l'ensemble de Fatou. De plus on fait une conjecture de non-errance et une conjecture sur la structure de l'ensemble de Fatou. Dans la Section 4.5 on considère une classe de fonctions rationnelles dites *simples*, qui ont une dynamique particulièrement élémentaire.

On considère d'abord les propriétés de la dynamique d'une fonction rationnelle à coefficients algébriques qui ne dépendent pas d'une valuation donnée. Plus concrètement on s'intéresse aux points périodiques. Il y a une grande littérature à ce sujet et on renvoie le lecteur aux bibliographies de [**MS2**] et [**Be**]. On a la propriété suivante.

***Théorème* (Northcott [No])**. — *Une fonction rationnelle à coefficients sur un corps algébrique a un nombre fini de points prépériodiques.*

Ce théorème vaut en dimension quelconque ; voir aussi Lewis [**Le**]. On donne une nouvelle démonstration de ce théorème (en dimension 1) à l'aide du logarithme itératif ; voir remarque dans la Section 4.5.

Il y a deux types de cycles dont la nature ne dépend pas de d'une valuation : les cycles *super-attractifs* qui sont ceux pour lesquels le multiplicateur est égal à zéro et les cycles *paraboliques*, qui sont ceux pour lesquels le multiplicateur est une racine de l'unité ; voir Section 3.

D'après un théorème de Fatou tout cycle attractif ou parabolique, d'une fonction rationnelle complexe de degré au moins deux, attire au moins un point critique par itération ; voir [**Fa**]. Comme une fonction rationnelle de degré $d > 1$ a au plus $2d - 2$ points critiques on obtient :

***Théorème* (Fatou [Fa])**. — *Soit $R \in \mathbb{C}(z)$ une fonction rationnelle de degré au moins deux. Alors le nombre de cycles attractifs ou paraboliques est majoré par $2\deg(R) - 2$.*

Il y a une amélioration de ce théorème par Shishikura qui dit que le nombre de cycles non-répulsifs d'une fonction rationnelle complexe de degré $d \geqslant 2$ est majoré par $2d - 2$; voir [**Sh**] et une nouvelle démonstration de A. Epstein [**Ep**].

Cette propriété est loin d'être vraie dans le cas p-adique. Par exemple le polynôme $z^p \in \mathbb{C}_p[z]$ a une infinité de cycles attractifs ; voir Section 4.1.1. D'autre part, une fonction rationnelle à coefficients sur $\mathbb{C}_p$ a un cycle indifférent, alors elle en a une infinité ; voir Corollaire 5.17.

Par le principe de Lefschetz on la une propriété analogue au Théorème de Fatou pour les fonctions rationnelles à coefficients dans $\mathbb{C}_p$.

Théorème 1. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré $d \geqslant 2$. Alors le nombre de cycles super-attractifs et paraboliques de R est au plus $2d - 2$.*

Démonstration. — Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré $d \geqslant 2$ et considérons $k \geqslant 0$ cycles super-attractifs ou paraboliques. Considérons l'ensemble T des coefficients de R et des points périodiques super-attractifs et paraboliques en question. Comme T est un ensemble fini le corps $\mathbb{Q}(T)$ a un degré de transcendance fini sur $\mathbb{Q}$ et par conséquent il existe une immersion $i : \mathbb{Q}(T)(z) \to \mathbb{C}(z)$. Alors la fonction rationnelle image $i(R) \in \mathbb{C}(z)$ a au moins k cycles super-attractifs ou paraboliques et donc, par le théorème de Fatou on a $k \leqslant 2d - 2$. $\square$

Si l'on considère seulement des cycles super-attractifs et paraboliques, alors l'énoncé du théorème de Fatou est purement algébrique. Cependant la démonstration de Fatou, ainsi que celle de A. Epstein, utilise des méthodes complexes. Il serait souhaitable d'avoir une démonstration de ce théorème (disons pour les fonctions rationnelles à coefficients algébriques) qui soit purement algébrique.

Remarque 4.1. — Dans le théorème de Fatou ci-dessus (et donc dans le Théorème 1), on peut compter les cycles super-attractifs et paraboliques avec multiplicité comme suit. Si z_0 est un point périodique super-attractif de période primitive n, alors la multiplicité du cycle correspondant est $\deg_{R^n}(z_0) - 1$. Si z_0 est un point périodique parabolique de période primitive n, alors la multiplicité de son cycle est $\mathrm{valit}(R^n, z_0)/\mathrm{ord}(R^n, z_0)$ dans la notation de la Section 3.3.1. C'est-à-dire, si l'on considère une coordonnée telle que $z_0 = 0$ et si $(R^n)'(0)$ est une racine primitive $q^{\text{ième}}$ de l'unité ; alors la multiplicité est égal à l'entier N/q, où N est déterminé par $(R^{nq})(z) = z(1 + az^N + \cdots)$, avec $a \neq 0$.

4.1. Domaines d'attraction. — Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle et $z_0 \in \mathbb{P}(\mathbb{C}_p)$ un point fixe attractif de R. Rappelons que

$$\mathcal{W}_R^s(z_0) = \{z \in \mathbb{P}(\mathbb{C}_p) \mid d(R^n(z), R^n(z_0)) \to 0 \text{ quand } n \to \infty\}$$

est le *bassin d'attraction* de z_0 pour R ; voir Section 3.1. Notons que

$$R^{-1}(\mathcal{W}_R^s(z_0)) = \mathcal{W}_R^s(z_0).$$

Par conséquent R envoie une composante analytique de $\mathcal{W}_R^s(z_0)$ sur une composante analytique de $\mathcal{W}_R^s(z_0)$; voir Proposition 2.6.

Définition 4.2. — Soient $R \in \mathbb{C}_p(z)$ une fonction rationnelle et z_0 un point périodique attractif de R. Alors la composante analytique de $\mathcal{W}_R^s(z_0)$ qui contient z_0 est appelée le *bassin d'attraction immédiat* de z_0 et *composante attractive* de R.

Définition 4.3. — On dit qu'un espace analytique connexe X est de *type Cantor* si l'arbre $\mathcal{A}_X$ de X est non-vide et si $\mathcal{A}_X$ satisfait les propriétés suivantes

(1) Toute arête de $\mathcal{A}_X$ a deux extrémités dans $\mathcal{A}_X$.
(2) Pour tout point $\mathcal{S}$ de $\mathcal{A}_X$ et $\mathcal{P} \in \mathcal{S}$ on a $D_{\mathcal{P}} \cap X \neq \varnothing$.

La condition 2 implique que pour tout sommet $\mathcal{S}$ de $\mathcal{A}_X$ il existe au moins trois arêtes ayant $\mathcal{S}$ comme extrémité ; en particulier $\mathcal{A}_X$ a une infinité de sommets. Dans la notation de la Section 1.2.3, la condition 2 de la définition se traduit par $n_{\mathcal{S}} = 0$ pour tout point $\mathcal{S}$ de $\mathcal{A}_X$.

Par exemple le complémentaire dans $\mathbb{P}(\mathbb{C}_p)$ d'un ensemble de Cantor est un espace analytique connexe de type Cantor ; voir l'exemple 1.3 (vii) où l'on considère l'arbre de $\mathbb{P}(\mathbb{C}_p) - \mathcal{O}_{\mathbb{Q}_p}$ et voir l'exemple 6.1 où on considère un polynôme ayant $\mathbb{P}(\mathbb{C}_p) - \mathcal{O}_{\mathbb{Q}_p}$ comme bassin d'attraction immédiat de l'infini.

On fixe pour le reste de cette section une fonction rationnelle $R \in \mathbb{C}_p(z)$ de degré au moins deux et un point fixe attractif z_0 de R.

Théorème 2. — *Notons C le bassin d'attraction immédiat de z_0 et $D \subset \mathcal{W}_R^s(z_0)$ un disque tel que $z_0 \in R(D) \subset D$. Pour $n \geqslant 0$ notons X_n la composante connexe de l'affinoïde ouvert $R^{-n}(D)$ qui contient z_0. On a alors $C = \cup_{n\geqslant 0} X_n$. De plus il y à deux cas :*

(1) *C est un disque rationnel ouvert.*
(2) *C est un espace analytique de type Cantor.*

La démonstration de ce théorème est à la fin de cette section. Supposons que le point fixe attractif z_0 ait un bassin d'attraction immédiat $\widetilde{D}$ au sens de la Section 3.1. Alors $\widetilde{D}$ est une boule ouverte (Lemme 4.11 ci-dessous). Notons X la composante de $R^{-1}(\widetilde{D})$ qui contient $\widetilde{D}$. On a $X = \widetilde{D}$ sinon il existe $r > r(R)$ tel que le disque ouvert D_1 de centre z_0 et rayon r soit contenu strictement dans X ; alors $R(D_1)$ est contenu strictement dans $R(X) = \widetilde{D} \subset D_1$. Par conséquent $\widetilde{D}$ est la composante analytique de $\mathcal{W}_R^s(z_0)$ qui contient z_0.

Corollaire 4.4. — *Il existe $d > 1$ tel que $R : C \to C$ est de degré d. En particulier $R(C) = C$.*

Démonstration. — Par la Proposition 2.6 il existe $d_n \geqslant 1$ tel que $R : X_n \to X_{n-1}$ est de degré d_n. Il est facile de voir que $d_{n+1} \geqslant d_n$ et par conséquent $d_n = d \geqslant 1$ pour n assez grand. Alors $R : C \to C$ est de degré d.

Si C est un disque alors $d > 1$, car sinon z_0 est indifférent. Si C n'est pas un disque alors il existe n tel que X_n est un disque et X_{n+1} ne l'est pas. Par conséquent $d \geqslant d_n > 1$. $\square$

Corollaire 4.5. — *Si $\lambda = R'(z_0) \neq 0$ il existe $\varphi \in \mathcal{H}(C)$ tel que $\varphi \circ R = \lambda\varphi$ et $\varphi(C) = \mathbb{C}_p$.*

Démonstration. — Par (i) de la Proposition 3.3 il existe $\varphi \in \mathcal{H}(D)$ tel que $\varphi \circ R = \lambda\varphi$ sur D. Comme pour tout $n \geqslant 0$ on $R^n(X_n) = X_0 = D$ on obtient que $\varphi = \lambda^{-n}\varphi \circ R^n \in \mathcal{H}(X_n)$. Comme $C = \cup_{n\geqslant 0} X_n$ et $X_n \subset X_{n+1}$ pour tout $n \geqslant 0$ on a $\varphi \in \mathcal{H}(C)$ et $\varphi \circ R = \lambda\varphi$. De plus notons que $\varphi(X_n) = \lambda^{-n}\varphi(X_0)$ donc $\varphi(C) = \varphi(\cup_{n\geqslant 0} X_n) = \mathbb{C}_p$. $\square$

Corollaire 4.6. — *Si $X \subset \mathcal{W}_R^s(z_0)$ est un affinoïde fermé, alors $\{R^n\}_{n\geqslant 1}$ converge vers z_0 uniformément sur X.*

Démonstration. — Si le disque $D \subset C$ contenant z_0 est suffisamment petit, l'assertion est vraie ; par conséquent il suffit de montrer qu'il existe n tel que $R^n(X) \subset D$. On peut supposer que l'affinoïde X est connexe et, quitte à remplacer R par un itéré on peut supposer $X \cap D \neq \varnothing$, et par conséquent $X \subset C = \cup_{n\geqslant 0} X_n$. Par conséquent il existe n tel que $X \subset X_n$ et $R^n(X) \subset D$. $\square$

4.1.1. Bassins et points critiques. — Dans le cas complexe il y a un théorème du à Fatou qui dit que tout bassin d'attraction immédiat d'un cycle attractif contient un point critique. Ceci n'est pas vrai dans le cadre présent, par exemple tous les cycles de $z^p \in \mathbb{C}_p(z)$ sont attractifs et par conséquent il y a des cycles qui n'attirent aucun point critique ; voir aussi exemple 6.2. Mais on a le corollaire suivant du Théorème 2.

Corollaire 4.7. — *Si le bassin d'attraction immédiat C est un disque et ne contient pas de points critiques, alors R a une infinité de points périodiques attractifs et le degré de $R : C \to C$ est divisible par p.*

Démonstration. — On suppose que $C = \{|z| < 1\}$ et on note $\widetilde{R}$ la réduction de R. Alors $\mathrm{wideg}(R') = \infty$, donc $\widetilde{R}' \equiv 0$ et par conséquent le degré de $R : C \to C$ est divisible par p. Par le Corollaire 4.4 on a $\deg(\widetilde{R}) > 1$ et par (iii) de la Proposition 2.4 il existe une infinité de $\alpha \in \overline{\mathbb{F}}_p$ périodiques par $\widetilde{R}$ tels que $R^n(D_\alpha) = D_{\widetilde{R}^n(\alpha)}$ où $n = n(\alpha)$ est la période de α et D_α dénote $\{z \mid \widetilde{z} = \alpha\}$. Par le lemme de Hensel, chacun de ces D_α contient un point périodique attractif de R et par conséquent il y en a une infinité. Voir aussi la démonstration de (ii) de la Proposition 4.32. □

On conjecture que cette propriété est aussi vraie dans le cas où C est de type Cantor. Ceci impliquerait que *le nombre de cycles attractifs est, soit infini, soit majoré par le nombre de points critiques (et donc au plus* $2\deg(R) - 2$*)*, voir Corollaire 4.9 ci-dessous. On remarque qu'il y a des fonctions rationnelles ayant un bassin d'attraction immédiat de type Cantor (d'un point fixe) qui ne contient pas des points critiques ; voir exemple 6.2.

Proposition 4.8. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Alors R a au plus* $\deg(R) - 1$ *cycles de bassins d'attraction immédiat de type Cantor.*

Corollaire 4.9. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins deux. Si R a plus de* $3\deg(R) - 3$ *cycles attractifs, alors R en a une infinité.*

Démonstration. — Supposons que R a un nombre fini de cycles attractifs. Par le Corollaire 4.7 chaque cycle ayant un disque comme bassin d'attraction immédiat attirés au moins un point critique de R, et par conséquent il y a au plus $2\deg(R) - 2$ tels cycles. Par la proposition il y a au plus $\deg(R) - 1$ cycles attractifs avec bassin d'attraction immédiat qui n'est pas un disque. □

La démonstration de la Proposition 4.8 se base sur le lemme suivant ; voir [**Be**].

Lemme 4.10

(1) *Soient X un affinoïde fermé connexe non-vide et $V \subset X$ un affinoïde ouvert. Alors le nombre de composantes connexes de l'affinoïde fermé $X - V$ est égal au nombre de bouts de V.*

(2) *Soient* $V_1, \ldots, V_n$ *des affinoïdes ouverts disjoints ayant chacun au moins deux bouts. Alors l'affinoïde fermé* $\mathbb{P}(\mathbb{C}_p) - V_1 \cup \cdots \cup V_n$ *a au moins* $n+1$ *composantes connexes.*

(3) *Soit* $R \in \mathbb{C}(z)$ *une fonction rationnelle de degré* $d \geqslant 1$ *et* $D_1, \ldots, D_n$ *disques disjoints deux à deux tels que* $R^{-1}(D_i)$ *n'est pas une union de disques, pour* $1 \leqslant i \leqslant n$. *Alors* $n \leqslant d-1$.

Démonstration

(1) Pour chaque bout $\mathcal{P}$ de V l'ensemble $B_{\mathcal{P}} \cap X$ est un affinoïde fermé connexe non-vide. Par conséquent le nombre de composantes connexes de

$$X - V = \sqcup_{\mathcal{P} \text{ bout de } V} B_{\mathcal{P}} \cap X$$

est égal au nombre de bouts de V.

(2) Par la partie 1 et par induction en n.

(3) On pose $X = \mathbb{P}(\mathbb{C}_p) - (D_1 \sqcup \cdots \sqcup D_n)$. Par hypothèse, pour chaque $1 \leqslant i \leqslant n$ il existe une composante connexe V_i de $R^{-1}(D_i)$ qui n'est pas un disque. Par la partie 2, $R^{-1}(X)$ a au moins $n+1$ composantes connexes. Par conséquent $n+1 \leqslant \deg(R) = d$, d'où $n \leqslant d-1$. ☐

Démonstration de la Proposition 4.8. — Soit z_0 un point périodique de R de période primitive k, soit D un disque qui contient z_0 et contenu dans le bassin d'attraction immédiat de z_0. De plus soit X_l la composante connexe de $R^{-l}(D)$ qui intersecte le cycle de z_0. Alors les bassins d'attraction immédiats des points du cycle de z_0 sont de type Cantor si et seulement si il existe l tel que X_l est un disque et X_{l+1} ne l'est pas. Donc si'l y a n cycles attractifs avec bassin d'attraction immédiat de type Cantor, on peut trouver des disques $D_1, \ldots, D_n$ deux à deux disjoints tels que $R(D_i)^{-1}$ n'est pas une union de disques. Donc par le point 3 du lemme précédent on a $n \leqslant d-1$. ☐

4.1.2. Preuve du Théorème 2. — La preuve du Théorème 2 dépend des lemmes suivants.

Lemme 4.11. — *Soit* $R \in \mathbb{C}_p(z)$ *une fonction rationnelle telle que* $R(0) = 0$ *et* $|R'(0)| \leqslant 1$ *et soit* $r(R)$ *comme dans la Section 3.1. Alors* $r(R) \in |\mathbb{C}_p|$.

Démonstration. — On pose $R(z) = a_0 + a_1 z + \cdots$. Soit il existe i tel que $|a_i|(r(R))^i = 1$ et le résultat est clair ; soit $r(R)$ est le rayon de convergence de la série $a_0 + a_1 z + \cdots$ qui est donc égal à la plus petite norme d'un pôle et par conséquent $r(R) \in |\mathbb{C}_p|$. ☐

Lemme 4.12. — *Soit* $\{X_n\}_{n \geqslant 0}$ *une suite croissante d'affinoïdes ouverts connexes et soit* $C = \cup_{n \geqslant 0} X_n$, *qui est un espace analytique connexe. Supposons que pour tout bout* $\mathcal{P}$ *de* X_n *on a* $B_{\mathcal{P}} \not\subset C$ *et il existe de bouts* $\mathcal{P}_0$ *et* $\mathcal{P}_1$ *de* X_{n+1} *tels que* $B_{\mathcal{P}_0} \subset B_{\mathcal{P}}$ *et* $B_{\mathcal{P}_1} \subset B_{\mathcal{P}}$. *Alors* C *est de type Cantor.*

Démonstration. — Si $\mathcal{S}$ est un système projectif tel que $\mathcal{S} \prec C$ alors il existe $n \geqslant 0$ tel que $\mathcal{S} \prec X_n$. De plus si $\mathcal{S} \prec X_n$ n'est pas un point de $\mathcal{A}_{X_n}$, alors $\mathcal{S}$ n'est pas un point de $\mathcal{A}_C$, donc $\mathcal{A}_C \subset \cup_{n\geqslant 0}\mathcal{A}_{X_n}$. Par définition pour tout point (resp. sommet) $\mathcal{S}$ de $\mathcal{A}_{X_n}$, il existe des bouts $\mathcal{P}_i \in \mathcal{S}$ distincts, tels que $D_{\mathcal{P}_i} \not\subset X_n$, $i = 0, 1$ (resp. $i = 0, 1, 2$) ; donc pour chaque $\mathcal{P}_i$ il existe un bout $\mathcal{Q}_i$ de X_n tel que $B_{\mathcal{Q}_i} \subset D_{\mathcal{P}_i}$. Par hypothèse $B_{\mathcal{Q}_i} \not\subset C$, donc $\mathcal{S}$ est un point (resp. sommet) de $\mathcal{A}_C$ et par conséquent $\mathcal{A}_C = \cup_{n\geqslant 0}\mathcal{A}_{X_n}$. Alors notons que C satisfait la propriété 2 car chaque X_n, étant un affinoïde ouvert connexe, la satisfait.

Pour montrer la propriété 1 il suffit de montrer que pour toute arête I de $\mathcal{A}_C$, chaque composante B de $\mathbb{P}(\mathbb{C}_p) - C_I$ intersecte C. Soit $\mathcal{S} \in I$ et n tel que $\mathcal{S} \in \mathcal{A}_{X_n}$. Si $B \cap X_n = \varnothing$ alors il existe un bout $\mathcal{P}$ de X_n tel que $B \subset B_{\mathcal{P}}$. Par hypothèse le nombre de bouts $\mathcal{Q}$ de X_{n+1} tels que $B_{\mathcal{Q}} \subset B_{\mathcal{P}}$ est au moins deux, donc $B_{\mathcal{Q}} \subset B$ et $B \cap X_{n+1} \neq \varnothing$. □

Lemme 4.13. — *Soit $Q \in \mathbb{C}_p(z)$ une fonction rationnelle et $\mathcal{P}$ un bout tel que $B_{\mathcal{P}} \subset B_{Q(\mathcal{P})}$ et $\mathcal{P} \neq Q(\mathcal{P})$. Alors $B_{\mathcal{P}}$ contient un point fixe de Q.*

Démonstration. — Après changement de coordonnée on suppose que $B_{\mathcal{P}} = \{|z| \leqslant 1\}$ et $B_{Q(\mathcal{P})} = \{|z| \leqslant r\}$ où $r \in |\mathbb{C}_p|$ est tel que $r > 1$. Par le Lemme 2.3, $(Q - \mathrm{id})(\mathcal{P}) = \mathcal{P}$ et par le Corollaire 2.2 on a $\{|z| \leqslant r\} \subset (Q - \mathrm{id})(B_{\mathcal{P}})$. Par conséquent il existe $z_0 \in B_{\mathcal{P}}$ tel que $Q(z_0) - z_0 = 0$. □

Démonstration du Théorème 2. — Soit $\widetilde{C}$ le bassin d'attraction immédiat de z_0 ; on a donc $C \subset \widetilde{C}$. Il y a deux cas.

Cas 1. — X_n est un disque pour tout $n \geqslant 0$. Alors $C = \cup_{n\geqslant 0}X_n$ est un disque. Considérons $r > 0$ et une coordonnée telle que $z_0 = 0$ et $C = \{|z| < r\}$. Comme $R(C) = C$ le développement en série de R est convergent dans C. Considérons $r(R) > 0$ comme dans la Section 3.1 ; on a $r \leqslant r(R)$ par la Proposition 3.2. D'autre part pour tout $s < r(R)$ il existe n tel que $R^n(B_s(0)) \subset X_0$ et donc $B_s(0) \subset X_n \subset B_{r(R)}(0)$; on conclut que $r = r(R)$. Par le Lemme 4.11 on a $r(R) \in |\mathbb{C}_p|$, donc après changement de coordonnée on peut supposer $C = \{|z| < 1\}$.

Considérons la réduction $\widetilde{R}$ de R et $\mathcal{S} = \{\mathcal{P}(\xi)\}_{\xi\in\mathbb{P}(\overline{\mathbb{F}}_p)}$ le système projectif canonique. On va montrer que $\widetilde{C} = \{|z| < 1\}$. Il suffit de montrer qu'il existe une infinité de $\xi \in \overline{\mathbb{F}}_p$ tels que $D_{\mathcal{P}(\xi)} \cap \mathcal{W}^s_R(0) = \varnothing$. Par (iii) de la Proposition 2.4 il suffit de montrer il existe une infinité de points périodiques de $\widetilde{R}$ dans $\overline{\mathbb{F}}_p$. Ceci est immédiat car $\deg(\widetilde{R}) > 1$.

Cas 2. — Il existe n_0 tel que X_{n_0} n'est pas un disque. On suppose sans perte de généralité que $n_0 = 1$, de telle façon que X_0 est un disque. Soit $\mathcal{P}_0$ le bout de X_0 ; notons que X_1 a au moins deux bouts. De plus pour chaque bout $\mathcal{Q}$ de X_1 on a $B_{\mathcal{Q}} \subset B_{\mathcal{P}_0}$ et $\mathcal{Q} \neq \mathcal{P}_0$.

Comme pour tout bout $\mathcal{P}$ de X_n on a $R^n(\mathcal{P}) = \mathcal{P}_0$, il existe au moins deux bouts $\mathcal{Q}_0$ et $\mathcal{Q}_1$ de X_{n+1} tels que $B_{\mathcal{Q}_i} \subset B_{\mathcal{P}}$, pour $i = 0, 1$. De plus $\mathcal{P} \neq \mathcal{P}_0$ et $R^n(\mathcal{P}) = \mathcal{P}_0$ donc $B_{\mathcal{P}} \subset B_{R^n(\mathcal{P})}$ et par le Lemme 4.13 la fonction rationnelle R^n a un point fixe dans $B_{\mathcal{P}}$. Par conséquent $B_{\mathcal{P}} \not\subset \mathcal{W}^s_R(z_0)$ et en particulier $B_{\mathcal{P}} \not\subset \widetilde{C}$ et $B_{\mathcal{P}} \not\subset C$. Donc C satisfait les hypothèses du Lemme 4.12 et par conséquent C est un espace analytique de type Cantor.

Il reste à montrer que $C = \widetilde{C}$. Soit X un affinoïde fermé connexe qui intersecte $X_0 \subset C$ et $\mathbb{P}(\mathbb{C}_p) - C$. Soit $\mathcal{P}_n$ un bout de X_n défini par induction ($\mathcal{P}_0$ est le bout de X_0) tel que $B_{\mathcal{P}_{n+1}} \subset B_{\mathcal{P}_n}$ et tel que $B_{\mathcal{P}_n} \cap X \neq \varnothing$. Par le raisonnement précédent on a $B_{\mathcal{P}_n} \not\subset X$.

Soit D_0 une composante connexe de $\mathbb{P}(\mathbb{C}_p) - X$. Alors, soit $D_0 \cap (B_{\mathcal{P}_n} - B_{\mathcal{P}_{n+1}}) = \varnothing$, soit $D_0 \subset B_{\mathcal{P}_n} - B_{\mathcal{P}_{n+1}}$. Comme le nombre de composantes connexes de $\mathbb{P}(\mathbb{C}_p) - X$ est fini il existe n tel que $B_{\mathcal{P}_n} - B_{\mathcal{P}_{n+1}} \subset X$. Par le raisonnement précédent il existe un bout $\mathcal{Q}$ de X_{n+1} tel que $B_{\mathcal{Q}} \subset B_{\mathcal{P}_n} - B_{\mathcal{P}_{n+1}} \subset X$. Mais on a montré que $B_{\mathcal{Q}} \not\subset \mathcal{W}^s_R(z_0)$. Donc $X \not\subset \mathcal{W}^s_R(z_0)$. Par conséquent $C = \widetilde{C}$. □

4.2. Domaine de quasi-périodicité. — Fixons une fonction rationnelle $R \in \mathbb{C}_p(z)$. On note par $\mathcal{E}(R)$ l'ensemble de points $z_0 \in \mathbb{P}(\mathbb{C}_p)$ tels qu'il existe une suite $n_j \to \infty$, quand $j \to \infty$, telle que R^{n_j} converge uniformément vers l'identité pour la distance chordale, sur un voisinage de z_0. On appelle $\mathcal{E}(R)$ *domaine de quasi-périodicité de* R.

Par définition $\mathcal{E}(R)$ est ouvert, $R(\mathcal{E}(R)) = \mathcal{E}(R)$, R est injective sur $\mathcal{E}(R)$ et tout point de $\mathcal{E}(R)$ est récurrent par R. Dans la Section 4.4 on verra que $\mathcal{E}(R)$ est égal à l'intérieur de l'ensemble des points récurrents par R (Corollaire 4.27).

Notons que la définition de $\mathcal{E}(R)$ coïncide localement avec la définition donnée dans la Section 3.2. En effet pour chaque $z_0 \in \mathcal{E}(R)$ il existe un disque $D \subset \mathcal{E}(R)$ contenant z_0 et $n \geqslant 1$ tel que $R^n(D) = D$ et par conséquent $R^n : D \to D$ est de degré 1. Par le Corollaire 3.12, D est contenue dans le domaine de quasipériodicité de R, au sens de la Section 3.2. En particulier $\mathcal{E}(R^n) = \mathcal{E}(R)$ pour $n \geqslant 1$.

On voudrait avoir un analogue du Lemme 3.11 pour les affinoïdes fermés contenues dans $\mathcal{E}(R)$. Dans $\mathbb{C}_p$ la convergence locale n'implique pas la convergence globale, comme dans le cas complexe. Donc on ne peut pas appliquer le Lemme 3.11 directement. A cause de cela on considère la Proposition 4.14, ci-dessous.

Étant donné un affinoïde fermé connexe $X \subset \mathbb{C}_p$ et $x \in X$ on dénote par $D_x \subset X$ le plus grand disque contenu dans X qui contient x. Alors $\pi_X(x) \in \mathcal{A}_X$ est le système projectif associé à D_x ; voir Section 1.2.3.

Proposition 4.14. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle et $X \subset \mathcal{E}(R)$ un affinoïde fermé connexe tel que $X \subset \mathbb{C}_p$. Alors il existe un entier $n \geqslant 1$ tel que $R^n : X \to X$ est de degré 1 et tel que R^n induit l'identité sur $\mathcal{A}_X$ et sur chaque système projectif $\mathcal{S} \in \mathcal{A}_X$. (En particulier pour chaque $x \in X$ on a $R^n(D_x) = D_x$.) De plus, il existe*

$\gamma \in (0,1)$ *tel que pour tout* $x \in X$ *on a*

$$\|R^n - \mathrm{id}\,\|_{D_x} \leqslant \gamma \cdot \mathrm{diam}(D_x).$$

Corollaire 4.15. — *Soit* $R \in \mathbb{C}_p(z)$ *une fonction rationnelle. Alors* R *permute les composantes analytiques du domaine de quasi-périodicité et chaque composante est périodique.*

Démonstration. — Soit C une composante analytique de $\mathcal{E}(R)$. Alors il est clair que $R(C)$ est contenu dans une composante analytique C' de $\mathcal{E}(R)$. Soit $x \in C$ et soit $Y \subset C'$ un affinoïde fermé connexe tel que $R(x) \in Y$. Par la Proposition 4.14 il existe $n \geqslant 1$ tel que $R^n(Y) = Y$. Alors $X = R^{n-1}(Y)$ est un affinoïde fermé connexe qui contient x car $R(X) = R^n(Y) = Y$ et R est injective sur $\mathcal{E}(R)$. Donc $X \subset C$ et par conséquent $Y \subset R(C)$. Alors $C' \subset R(C)$ et $R^n(C) = C$. □

D'après la Proposition 4.14 on peut appliquer le Lemme 3.11 à chaque disque D_x, pour $x \in X$, avec une constante $\gamma \in (0,1)$ uniforme. Donc le corollaire suivant est immédiat.

Corollaire 4.16. — *Soient* $R \in \mathbb{C}_p(z)$ *une fonction rationnelle et* $X \subset \mathcal{E}(R) \cap \mathbb{C}_p$ *un affinoïde fermé connexe. Supposons* R *vérifie avec* $n = 1$*, les propriétés de la Proposition 4.14. Considérons la fonction* $\rho : (0,\infty) \to (0,\infty)$ *définie dans la Section 3.2. Alors on a les propriétés suivantes.*

(1) *Soit* $T_0 = \mathrm{id}$ *et pour* $n \geqslant 0$ *posons* $T_{n+1} = T_n \circ R - T_n$*. Pour* $w \in \mathbb{C}_p$ *tel que* $\rho(|w|) < \gamma^{-1}$ *on a* $\|T_n\|_X \leqslant R\gamma^n$*. De plus, la série*

$$R^{\circ w} = \sum_{i=0}^{m} \binom{w}{i} T_i,$$

converge uniformément dans X *vers un automorphisme de* X*. De plus si* $\rho(|w_i|) < \gamma^{-1}$*, pour* $i = 1, 2$*, alors* $R^{\circ w_1} \circ R^{\circ w_2} = R^{\circ (w_1 + w_2)}$*.*

(2) *Pour tout* $w \in \mathbb{C}_p$ *avec* $\rho(|w|) < \gamma^{-1}$ *on a*

$$\left\| R_* - \frac{R^{\circ w} - \mathrm{id}}{w} \right\|_X \leqslant C \mathrm{diam}(X)|w| \quad \textit{et} \quad \left\| R_* - \sum_{1 \leqslant i \leqslant k} (-1)^{i-1} \frac{T_i}{i} \right\|_X \leqslant k\gamma^k \mathrm{diam}(X),$$

où la constante C *ne dépend que de* $\gamma\rho(|w|)$*. En particulier* $(R^n - \mathrm{id})/n$ *et* $\sum_{i>0} (-1)^{i-1} T_i/i$ *convergent uniformément vers* R_* *sur* X *quand* $|n|_p \to 0$*.*

Corollaire 4.17. — *Soit* $C \subset \mathbb{C}_p$ *une composante analytique de* $\mathcal{E}(R)$*. Alors* $R_* \in \mathcal{H}(C)$*.*

Démonstration. — Par 2 du corollaire précédent on a $R_* \in \mathcal{H}(X)$ pour tout affinoïde fermé $X \subset C$. Donc on a $R_* \in \mathcal{H}(C)$ par définition de $\mathcal{H}(C)$. □

Corollaire 4.18. — *Un affinoïde fermé* $X \subset \mathcal{E}(R)$ *contient au plus un nombre fini de points périodiques de* R*.*

Démonstration. — Après changement de coordonnée on suppose $X \subset \mathbb{C}_p$. Par 1 de la Proposition 3.16 les points périodiques dans $\mathcal{E}(R)$ coïncident avec les zéros de R_*. Donc la condition $R_* \equiv 0$ implique qu'il existe $n \geqslant 1$ tel que $R^n = \mathrm{id}$ sur un ouvert, mais ceci n'est pas possible car $\deg(R) > 1$. Par conséquent $R_* \not\equiv 0$. Par le corollaire précédent $R_* \in \mathcal{H}(X)$ et donc R_* a un nombre fini de zéros dans X. □

Corollaire 4.19. — *Pour tout affinoïde fermé $X \subset \mathcal{E}(R)$ tel que $X \subset \mathbb{C}_p$, il existe une suite d'entiers $n_j \to \infty$ telle que R^{n_j} converge uniformément vers l'identité sur X.*

Démonstration. — Par le point 2 du Corollaire 3.11 ; voir aussi démonstration de (i) de la Proposition 3.14. □

Corollaire 4.20. — *Soit $X \subset \mathcal{E}(R)$ un affinoïde fermé. Alors tout bout $\mathcal{P} \prec X$ (resp. système projectif $\mathcal{S} \prec X$) est périodique par R.*

Démonstration. — Comme $\mathcal{S} \prec X$ implique $\mathcal{P} \prec X$ pour une infinité de bouts $\mathcal{P} \in \mathcal{S}$, il suffit de montrer l'affirmation sur les bouts.

Considérons une coordonnée telle que $X \subset \mathbb{C}_p$. Alors il suffit de noter que par le corollaire précédent il existe un entier n tel que

$$\|R^n - \mathrm{id}\,\|_X < \min\{\mathrm{diam}(B_{\mathcal{P}}), \mathrm{diam}(D_{\mathcal{P}})\}.$$

Donc on a $R^n(\mathcal{P}) = \mathcal{P}$ par le Lemme 2.3. □

4.2.1. Preuve de la Proposition 4.14. — La démonstration de la Proposition 4.14 est divisée en plusieurs lemmes. Fixons une fonction rationnelle $R \in \mathbb{C}_p(z)$ de degré au moins deux et un affinoïde fermé $X \subset \mathcal{E}(R)$ tel que $X \subset \mathbb{C}_p$.

Lemme 4.21. — *Soit $\mathcal{P}$ un bout tel que $D_{\mathcal{P}} \subset \mathcal{E}(R)$. Alors $R : D_{\mathcal{P}} \to D_{R(\mathcal{P})}$ est de degré 1 et il existe un entier $n \geqslant 1$ tel que $R^n(D_{\mathcal{P}}) = D_{\mathcal{P}}$.*

Démonstration. — Par (ii) de la Proposition 3.9 R est injective sur $D_{\mathcal{P}} \subset \mathcal{E}(R)$ et donc $\deg_R(\mathcal{P}) = 1$. Donc par (iv) du Lemme 2.1 $R : D_{\mathcal{P}} \to D_{R(\mathcal{P})}$ est de degré 1.

Comme tout point de $\mathcal{E}(R)$ est récurrent par R (Corollaire 3.15) il existe un entier $n \geqslant 1$ tel que $R^n(D_{\mathcal{P}}) \cap D_{\mathcal{P}} \neq \varnothing$. Notons que $D_{\mathcal{P}} \cup D_{R^n(\mathcal{P})} \neq \mathbb{P}(\mathbb{C}_p)$, sinon $\mathcal{E}(R) = \mathbb{P}(\mathbb{C}_p)$ et alors R est injective ; ceci n'est pas possible car $\deg(R) > 1$. Donc $D_{R^n(\mathcal{P})} \subset D_{\mathcal{P}}$ ou $D_{\mathcal{P}} \subset D_{R^n(\mathcal{P})}$. Si on a pas l'égalité alors $D_{\mathcal{P}} \subset \mathcal{E}(R)$ contient un point périodique attractif ou répulsif, mais ceci n'est pas possible. Donc $R^n(D_{\mathcal{P}}) = D_{\mathcal{P}}$. □

Comme par hypothèse l'affinoïde fermé X est contenu dans $\mathbb{C}_p$, l'arbre $\mathcal{A}_X$ de X est non-vide et on a la partition canonique

$$X = \left(\sqcup_{\mathcal{S} \text{ sommet de } \mathcal{A}_X} X_{\mathcal{S}}\right) \sqcup \left(\sqcup_{I \text{ arête de } \mathcal{A}_X} X_I\right).$$

La Proposition 4.14 est une conséquence immédiate des Lemmes 4.22 et 4.23 ci-dessous.

Lemme 4.22. — *Il existe $n \geqslant 1$ tel que pour tout sommet $\mathcal{S}$ de $\mathcal{A}_X$ on a $R^n(\mathcal{S}) = \mathcal{S}$ et R^n induit l'identité sur le système projectif $\mathcal{S}$. Dans ce cas il existe $\gamma_{\mathcal{S}} \in (0,1)$ tel que pour tout $x \in X_{\mathcal{S}}$ on a*

$$\|R^n - \mathrm{id}\,\|_{D_x} \leqslant \gamma_{\mathcal{S}} \cdot \mathrm{diam}(D_x).$$

Démonstration. — Soit $\mathcal{P} \in \mathcal{S}$ tel que $D_{\mathcal{P}} \subset \mathcal{E}(R)$. Alors par le lemme précédent il existe k tel que $R^k(\mathcal{P}) = \mathcal{P}$. Comme R est injective sur $\mathcal{E}(R)$ on a $\deg_{R^k}(\mathcal{S}) = 1$ (Lemme 2.11). Comme tout automorphisme de $\mathbb{P}(\overline{\mathbb{F}}_p)$ est d'ordre fini, il existe m tel que R^{km} induit l'identité sur $\mathcal{S}$. Comme $\mathcal{A}_X$ n'a que un nombre fini de sommets, l'assertion du lemme en résulte.

Notons que pour tout $x \in X_{\mathcal{S}}$ on a $\mathrm{diam}(D_x) = \mathrm{diam}(X_{\mathcal{S}})$. De plus $R^n(x) \in D_x$, donc on a $|R^n(x) - x| < \mathrm{diam}(X_{\mathcal{S}})$. Comme $X_{\mathcal{S}}$ est un affinoïde fermé son image par $R^n - \mathrm{id}$ est aussi un affinoïde fermé (Proposition 2.6) et par conséquent il existe $\gamma_{\mathcal{S}} \in (0,1)$ tel que $\|R^n - \mathrm{id}\,\|_{X_{\mathcal{S}}} \leqslant \gamma_{\mathcal{S}} \cdot \mathrm{diam}(X_{\mathcal{S}})$. □

Lemme 4.23. — *Soit I une arête de $\mathcal{A}_X$ et supposons que R induit l'identité sur les extrémités de I. Alors il existe $\gamma_I \in (0,1)$ tel que pour tout $x \in X_I$ on ait*

$$\|R - \mathrm{id}\,\|_{D_x} \leqslant \gamma_I \cdot \mathrm{diam}(D_x).$$

En particulier $R(D_x) = D_x$ et donc R induit l'identité sur chaque point dans I.

Démonstration. — Après changement de coordonnée on suppose $X_I = \{r < |z| < 1\}$, avec $r \in (0,1)$. Alors pour $x \in X_I$ on a $D_x = \{|z - x| < |x|\}$ et en particulier $\mathrm{diam}(D_x) = |x|$.

Notons que $|R(z)| = |z|$ pour $r < |z| < 1$. Comme R est l'identité sur les systèmes projectifs associés aux bouts de X_I, la fonction $|R|_0$ coïncide avec l'identité sur un voisinage (r_0, r_1) de $[r, 1]$. On choisit r_0 et r_1 tels que tous les zéros et pôles de R dans $\{r_0 < |z| < r_1\}$ sont de norme égale à r ou 1. Comme R induit l'identité sur les systèmes projectifs associés aux bouts de X_I, à tout zéro a de R sur $\{|z| = 1\}$ correspond un pôle $b(a)$ tel que $|a - b(a)| < 1$. Comme R fixe le bout associé à $\{|z| < r\}$, R a au moins un zéro dans $\{|z| \leqslant r_0\}$.

De la même façon, à tout zéro a de R sur $\{|z| = r\}$ correspond un pôle $b(a)$ tel que $|a - b(a)| < r$. Par conséquent R est de la forme

$$R(z) = (\lambda z - \mu) \cdot \prod_{a \text{ zéro sur } \{|z|=1, |z|=r\}} \left(\frac{z-a}{z-b(a)}\right)$$
$$\cdot \frac{\prod_{a \text{ zéro sur } \{|z|>r_1\}} (1 - a^{-1}z)}{\prod_{b \text{ pôle sur } \{|z|>r_1\}} (1 - b^{-1}z)} \cdot \frac{\prod_{a \text{ zéro sur } \{|z|<r_0\}} (1 - az^{-1})}{\prod_{b \text{ pôle sur } \{|z|<r_0\}} (1 - bz^{-1})}$$

où $|\lambda - 1| < 1$ et $|\mu| \leqslant r_0$. Notons que pour $z \in X_I$ on a

$$\frac{z-a}{z-b(a)} = 1 + \frac{a-b(a)}{z-b(a)} \quad \text{et} \quad \left|\frac{a-b(a)}{z-b(a)}\right| \leqslant \left|\frac{a-b(a)}{b(a)}\right| < 1.$$

Par conséquent

$$\begin{aligned}|R(z)-z| &\leqslant \max\{|R(z)-(\lambda z-\mu)|, |(\lambda-1)z+\mu|\}\\ &\leqslant |z| \max\{|\lambda-1|, |\mu/z|, |(a-b(a))/b(a)|, r_1^{-1}, r_0/r\}.\end{aligned}$$

Donc $\gamma_I = \max\{|\lambda-1|, |(a-b(a))/b(a)|, r_1^{-1}, r_0/r\} < 1$ satisfait l'assertion du lemme. □

4.3. Ensembles de Fatou et de Julia. — Dans cette section on considère la décomposition en ensembles de Fatou et de Julia, comme dans le cas complexe. On considère d'abord le théorème suivant de Hsia qui est l'analogue p-adique du Théorème de Montel ; voir [**Hs**] et voir [**CG**] pour le Théorème de Montel.

***Théorème* (Hsia [Hs])**. — *Soit $r > 0$ et considérons une famille $F \subset \mathcal{H}(B_r(0))$. S'il existe $z_0 \in \mathbb{C}_p$ tel que $f(z) \neq z_0$ pour tout $f \in F$ et tout $z \in B_r(0)$ alors F est uniformément lipschitzienne pour la distance chordale.*

La démonstration de ce théorème est simple : il suffit de noter que l'image d'un disque par une série f convergente sur $B_r(0)$ est aussi un disque. Par conséquent, si les éléments d'une famille $F \subset \mathcal{H}(B_r(0))$ ne prennent pas la valeur z_0 (que on suppose de norme égale à 1), les images de $B_r(0)$ par les $f \in F$ ont un diamètre chordal plus petit que 1 ; alors le théorème suit du Corollaire 2.2 du Lemme de Schwarz (Section 1.3.1).

Définition 4.24. — Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. L'*ensemble de Fatou* de R, noté $F(R)$, est l'ensemble de tous les points $z_0 \in \mathbb{P}(\mathbb{C}_p)$ tels qu'il existe un voisinage U de z_0 où la famille $\{R^n|_U\}_{n\geqslant 1}$ est uniformément lipschitzienne pour la distance chordale. De plus $J(R) = \mathbb{P}(\mathbb{C}_p) - F(R)$ est appelé l'*ensemble de Julia* de R.

Il est facile de voir que l'on a les propriétés usuelles ; $R^{-1}(F(R)) = F(R)$, $R^{-1}(J(R)) = J(R)$ et pour tout $n \geqslant 1$, $F(R^n) = F(R)$ et $J(R^n) = J(R)$. De plus, si w est un automorphisme de $\mathbb{P}(\mathbb{C}_p)$ alors $F(w \circ R \circ w^{-1}) = w(F(R))$ et $J(w \circ R \circ w^{-1}) = w(J(R))$.

Par définition $F(R)$ est ouvert et par conséquent $J(R)$ est fermé, mais en général $J(R)$ n'est pas compact. De plus $J(R)$ n'a pas des points isolés ; voir [**Hs**].

Proposition 4.25. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Alors $\mathcal{E}(R) \subset F(R)$ et $F(R)$ contient tous les bassins d'attraction.*

Démonstration. — Par le Lemme 4.21 pour tout disque $D \subset \mathcal{E}(R)$ il existe $n \geqslant 1$ tel que $R^n(D) = D$. Donc $\mathcal{E}(R) \subset F(R)$. Le cas des bassins d'attraction est clair. □

En particulier $F(R)$ contient tous les points périodiques non-répulsifs de R et il n'est pas difficile de voir que tous les points périodiques répulsifs de R appartiennent à $J(R)$. De plus on a le théorème suivant de Benedetto.

***Théorème* (Benedetto [Be]).** — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Alors R a un point fixe non-répulsif, en particulier $F(R) \neq \varnothing$.*

Voici une différence avec le cas complexe. Dans le cas complexe l'ensemble de Julia d'une fonction rationnelle est non-vide et l'ensemble de Fatou peut être vide, mais dans le cas p-adique c'est le contraire. Par le théorème précédent, l'ensemble de Fatou est non-vide et l'ensemble de Julia peut être vide. Par exemple il n'est pas difficile de voir que $J(z^p) = \varnothing$. Plus généralement $J(R) = \varnothing$ pour toutes les fonctions rationnelles simples ; voir Section 4.5.

***Corollaire 4.26*.** — *L'ensemble de Julia est d'intérieur vide.*

Démonstration. — Si l'ensemble de Julia d'une fonction rationnelle $R \in \mathbb{C}_p(z)$ est d'intérieur non-vide, alors par le Théorème de Hsia $\cup_{n \geqslant 0} R^n(J(R)) \subset J(R)$ est égal à $\mathbb{P}(\mathbb{C}_p)$ ou égal à $\mathbb{P}(\mathbb{C}_p)$ moins un point. Comme $J(R)$ est fermé on obtient $J(R) = \mathbb{P}(\mathbb{C}_p)$, ce qui n'est pas possible par le théorème de Benedetto. □

4.4. Structure de l'ensemble de Fatou et disques errants. — Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. On dit qu'un disque $D \subset \mathbb{P}(\mathbb{C}_p)$ est *errant* si pour tous $k > l \geqslant 0$ on a $R^k(D) \cap R^l(D) = \varnothing$.

***Théorème de Classification*.** — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Alors l'ensemble de Fatou de R se décompose dans les ensembles disjoints suivants.*

(1) *Bassins d'attraction.*
(2) $\mathcal{E}'(R) = \cup_{n \geqslant 0} R^{-n}(\mathcal{E}(R))$.
(3) *L'union des disques errants qui ne sont pas attirés par un cycle attractif.*

***Corollaire 4.27*.** — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle. Alors $\mathcal{E}(R)$ est égal à l'intérieur de l'ensemble des points récurrents par R.*

Démonstration. — Par définition $\mathcal{E}(R)$ est ouvert et tous les points de $\mathcal{E}(R)$ sont récurrents par R ; voir Section 4.2. D'autre part considérons un point $z_0 \in \mathbb{P}(\mathbb{C}_p)$ récurrent par R.

Par [**Hs**] on peut trouver un point périodique $w_0 \in \mathbb{P}(\mathbb{C}_p)$ tel que tout point de $J(R)$ soit accumulé par des préimages de w_0, par des itérés de R. Donc z_0 appartient à l'ensemble de Fatou de R. Clairement z_0 n'appartient pas à un disque errant et n'est pas attiré par un cycle attractif. Donc le théorème implique que $z_0 \in \mathcal{E}(R)$. □

Dans [**Hs**] Hsia a montré que l'ensemble de Julia est contenu dans l'ensemble de points de accumulation des points périodiques. Mais en effet on à l'égalité.

***Corollaire 4.28*.** — *L'ensemble de Julia est égal à l'ensemble des points d'accumulation des points périodiques.*

Démonstration. — Par [**Hs**] il suffit de montrer que si $w \in \mathbb{P}(\mathbb{C}_p)$ est accumulé par points périodiques, alors $w \in J(R)$. Clairement w n'appartient pas a un bassin d'attraction et n'est pas contenue dans un disque errant. De plus $w \notin \mathcal{E}'(R)$ car les points périodiques indifférents sont isolés (Corollaire 3.17). Donc par le théorème on a $w \in J(R)$. □

Il n'est pas clair si $J(R)$ est égal à la fermeture de points périodiques répulsifs ; voir aussi [**Hs**]. Par analogie au cas complexe on fait la conjecture suivante.

Conjecture de Non-Errance. — *Tout disque errant est attiré par un cycle attractif.*

On peut comparer à Sullivan [**Sul**] et Guckenheimer [**Gu**]. Benedetto a fait une conjecture de non-errance qui n'est pas tout à fait équivalente à celle-ci. Par exemple la conjecture de Benedetto ne dit rien dans le cas où l'ensemble de Fatou est $\mathbb{P}(\mathbb{C}_p)$ tout entier. Néanmoins on peut avoir une dynamique non-triviale dans ce cas ; voir exemples 6.3 et 6.6.

D'après le théorème précédent la Conjecture de Non-Errance est équivalente à la conjecture suivante.

Structure Conjecturale de l'Ensemble de Fatou. — *Tout point de l'ensemble de Fatou appartient à $\mathcal{E}'(R)$ ou est attiré par un cycle attractif.*

La démonstration du Théorème de Classification dépend du lemme suivant.

Lemme 4.29. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle et supposons que D est un disque errant. Alors les affirmations suivantes sont vraies.*

(1) *$D \subset F(R)$.*

(2) *$D \cap \mathcal{E}'(R) = \varnothing$.*

(3) *Pour tout $n \geqslant 1$, $R^n(D)$ est un disque.*

(4) *Si* diam *dénote le diamètre chordal, on a*

$$\liminf_{j\to\infty} \operatorname{diam}(R^j(D)) = 0.$$

(5) *Si D intersecte un bassin d'attraction, alors D est attiré par le cycle correspondant.*

Démonstration

(1) Comme $R^n(D) \cap D = \varnothing$, pour $n \geqslant 1$ la famille $\{R^n|_D\}_{n\geqslant 1}$ est uniformément lipschitzienne (Théorème de Hsia, Section 4.3). Donc $D \subset F(R)$.

(2) Par le Corollaire 3.15 tout point de $\mathcal{E}(R)$ est récurrent par R. Donc l'orbite de D est disjointe de $\mathcal{E}(R)$; c'est-à-dire $D \cap \mathcal{E}'(R) = \varnothing$.

(3) Par le Corollaire 2.2 on a que $R^n(D)$ est un disque ou $R^n(D) = \mathbb{P}(\mathbb{C}_p)$. Mais comme D est un disque errant on a $R^n(D) \subset \mathbb{P}(\mathbb{C}_p) - D$ et par conséquent $R^n(D)$ est un disque.

(4) Supposons $\liminf_{j\to\infty} \operatorname{diam}(R^j(D)) > 0$. Par le Lemme 2.3 on a que si D_0 et $R(D_0)$ sont des disques et Q est une fonction rationnelle telle que $d(R(z), Q(z)) < \operatorname{diam}(R(D_0)) < 1$, alors $Q(D_0) = R(D_0)$. Donc quitte à faire une petite perturbation dans les coefficients de R, on peut supposer qu'il existe une extension finie K de $\mathbb{Q}_p$, tel que $R \in K(z)$ et tel que $D \cap \mathbb{P}(K) \neq \varnothing$.

Alors $\mathbb{P}(K)$ est invariant par R et comme $\mathbb{P}(K)$ est compact, pour tout $w \in D$ on peut trouver des entiers $m > n$ tels que $|R^n(w) - R^m(w)| < \liminf_{j\to\infty} \operatorname{diam}(R^j(D))$. Mais ceci implique $R^n(D) \cap R^m(D) \neq \varnothing$. On obtient une contradiction.

(5) Supposons qu'un point $z_0 \in D$ est attiré par un cycle attractif. Par la partie 3 on a $\liminf_{j\to\infty} \operatorname{diam}(R^j(D)) = 0$. Donc il existe n tel que $R^n(z_0)$ soit suffisamment proche d'un point périodique attractif w_0 et tel que le diamètre de $R^n(D)$ soit suffisamment petit, de telle façon que $R^n(D)$ soit contenu dans le bassin d'attraction de w_0. Alors D est attiré par le cycle (attractif) de w_0. □

Démonstration du Théorème de Classification. — Soit $z_0 \in F(R)$. Par définition de $F(R)$ il existe un disque ouvert D qui contient z_0 et tel que $\operatorname{diam}(R^n(D)) < 1 = \operatorname{diam}(\mathbb{P}(\mathbb{C}_p))$ pour tout $n \geqslant 1$. D'après le Corollaire 2.2 $R^n(D)$ est un disque pour $n \geqslant 1$.

Si D est un disque errant alors $D \cap \mathcal{E}'(R) = \varnothing$ par le lemme précédent. De plus soit D est disjoint des bassins d'attraction, soit D est attiré par un cycle attractif. Donc le théorème est vérifié dans ce cas.

Supposons que D n'est pas errant. Soient n et $k \geqslant 1$ tels que $R^n(D) \cap R^{n+k}(D) \neq \varnothing$. Si $R^{n+k}(D)$ est strictement contenu dans $R^n(D)$ alors D contient un point périodique attractif et D est attiré par le cycle correspondant ; cf. Corollaire 1.4. Sinon $R^n(D) \subset R^{n+k}(D)$ et on considère le disque $D_0 = \cup_{i\geqslant 0} R^{n+ik}(D)$. On a $R^k(D_0) = D_0$. Si $R^k : D_0 \to D_0$ est de degré 1 alors $D_0 \subset \mathcal{E}(R)$ par le Corollaire 3.11. Sinon D_0 est contenu dans le bassin d'attraction d'un point fixe de R^k ; voir Proposition 3.2. □

4.5. Bonne réduction et fonctions rationnelles simples. — En coordonnées homogènes, une fonction rationnelle R est de la forme

$$R([x, y]) = [P_1(x, y), P_2(x, y)],$$

où P_1, $P_2 \in \mathbb{C}_p[x, y]$ sont des polynômes homogènes de degré $\deg(R)$. Comme pour $\lambda \in \mathbb{C}_p - \{0\}$, $[\lambda P_1, \lambda P_2]$ représente la même fonction rationnelle, on peut supposer P_1 et P_2 à coefficients entiers et qu'au moins un des coefficients de P_1 ou de P_2 est de norme égale à 1. On considère la définition suivante, due à Morton et Silverman [**MS2**].

Définition 4.30. — On dit que $R \in \mathbb{C}_p(z)$ a une *bonne réduction* si $\widetilde{P}_1$ et $\widetilde{P}_2$ n'ont pas de racine commune sur $\overline{\mathbb{F}}_p \times \overline{\mathbb{F}}_p$, autre que $(0, 0)$.

Donc R a une bonne réduction si et seulement si $|\mathrm{disc}(P_1, P_2)| = 1$. Autrement dit $R \in \mathbb{C}_p(z)$ a une bonne réduction si et seulement si le système projectif canonique $\mathcal{S}$ est invariant par R et la réduction $\widetilde{R}$ de R a un degré maximal, $\deg(\widetilde{R}) = \deg(R)$.

La notion de bonne réduction dépend du choix de la coordonnée. Par exemple le polynôme pz^2 n'a pas une bonne réduction sur $\mathbb{C}_p$, mais il est conjugué à z^2 qui en a une. Pour $n > 1$ la fonction rationnelle R^n a une bonne réduction si et seulement si R a une bonne réduction ; voir [**Be**].

Notons que, si $R \in \mathbb{Q}(z)$, alors R a une bonne réduction sur $\mathbb{C}_p$ pour presque tout nombre premier p.

***Théorème* (Morton, Silverman [MS2])**. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle ayant une bonne réduction. Alors R n'augmente pas la distance chordale et en particulier $J(R) = \varnothing$.*

Définition 4.31. — On dit qu'une fonction rationnelle $R \in \mathbb{C}_p(z)$ est *simple* s'il existe un choix de coordonnée pour laquelle R a une bonne réduction.

La dynamique des fonctions rationnelles ayant une bonne réduction, ou plus généralement la dynamique des fonctions rationnelles simples, est intéressante car on peut comprendre une grande partie de la dynamique si l'on comprend la dynamique de la réduction ; voir la proposition suivante.

Proposition 4.32. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle, ayant une bonne réduction $\widetilde{R}$. Pour $\alpha \in \mathbb{P}(\overline{\mathbb{F}}_p)$ on pose $D_\alpha = \{\widetilde{z} = \alpha\}$. Alors on a les propriétés suivantes.*

(i) *$R : D_\alpha \to D_{\widetilde{R}(\alpha)}$ est de degré $\deg_{\widetilde{R}}(\alpha)$. En particulier $R(D_\alpha) = D_{\widetilde{R}(\alpha)}$.*

(ii) *D_α contient un point périodique attractif si et seulement si α est périodique par $\widetilde{R}$ et $(\widetilde{R}^k)'(\alpha) = 0$, où k est la période de α. Dans ce cas D_α contient un unique point périodique et D_α est son bassin d'attraction immédiat.*

(iii) *$D_\alpha \cap \mathcal{E}(R) \neq \varnothing$ si et seulement si α est périodique par $\widetilde{R}$ et $(\widetilde{R}^k)'(\alpha) \neq 0$, où k est la période de α. Dans ce cas $D_\alpha \subset \mathcal{E}(R)$. De plus, si $\deg(R) > 1$ alors D_α est une composante analytique de $\mathcal{E}(R)$ qui contient un point périodique de R de même période que α.*

(iv) *Soit $\widetilde{R}' \equiv 0$ et dans ce cas $p | \deg(R)$ et tout point périodique de R est attractif ; soit $\widetilde{R}' \not\equiv 0$ et alors tout cycle attractif de R attire un point critique (et donc le nombre de cycles attractifs est majoré par $2\deg(R) - 2$) et $\mathcal{E}(R)$ est non-vide.*

Voir aussi Proposition 2.4 dans la Section 2. La démonstration de cette proposition est à la fin de cette section.

Corollaire 4.33. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle ayant une bonne réduction. Alors tout point de $\mathbb{P}(\mathbb{C}_p)$ appartient à $\mathcal{E}'(R) = \cup_{n \geqslant 0} R^{-n}(\mathcal{E}(R))$ ou à un bassin d'attraction.*

Démonstration. — Par la proposition précédente et en considérant que tout élément de $\mathbb{P}(\overline{\mathbb{F}}_p)$ est prépériodique pour une fonction rationnelle quelconque. $\square$

Corollaire 4.34. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins deux ayant une bonne réduction et telle que $\mathcal{E}(R) \neq \varnothing$. Alors $\mathcal{E}(R)$ est de la forme $\cup_A D_\alpha$ où $A \subset \mathbb{P}(\overline{\mathbb{F}}_p)$ est tel que $\#A$ et $\#(\mathbb{P}(\overline{\mathbb{F}}_p) - A)$ sont infini.*

Démonstration. — Comme $\mathcal{E}(R) \neq \varnothing$ la réduction $\widetilde{R}$ a un point périodique avec multiplicateur différent de zéro. Par conséquent $\widetilde{R}$ en a une infinité. Comme le degré de R est au moins deux il y une infinité de $\alpha \in \overline{\mathbb{F}}_p$ strictement prépériodiques. Alors le corollaire suit de la proposition. $\square$

Proposition 4.35. — *Soit L une extension finie de $\mathbb{Q}_p$ et $R \in L(z)$ une fonction rationnelle de degré au moins deux, avec une bonne réduction. Alors R a un nombre fini des points prépériodiques dans $\mathbb{P}(L)$.*

Démonstration. — Comme $\mathbb{P}(\widetilde{L}) \subset \mathbb{P}(\overline{\mathbb{F}}_p)$ est fini il suffit de montrer que pour chaque $\alpha \in \mathbb{P}(\widetilde{L})$ périodique par $\widetilde{R}$, le disque $\{\widetilde{z} = \alpha\}$ rencontre au plus un nombre fini de points prépériodiques par R. On se ramène au cas où $\alpha = 0$ et le disque $\{|z| < 1\}$ est fixé par R.

Supposons d'abord que R est injective sur $\{|z| < 1\}$. Dans ce cas, $\{|z| < 1\} \subset \mathcal{E}(R)$ et les points prépériodiques dans ce disque sont périodiques. La fonction R_* est holomorphe dans $\{|z| < 1\}$ et comme R est de degré au moins deux on a $R_* \not\equiv 0$. Comme les points périodiques indifférents sont les zéros de R_* et $\mathbb{P}(L)$ est compact, $\{|z| < 1\}$ rencontre au plus un nombre fini de points périodiques.

Supposons maintenant que R n'est pas injective sur $\{|z| < 1\}$. Alors $\{|z| < 1\}$ est une composante attractive de R. De plus R n'a qu'un seul point périodique dans $\{|z| < 1\}$, qui est alors fixe et qu'on peut supposer égal à 0. Soit $\varepsilon > 0$ tel que R n'a pas de préimage de 0 sur $\{|z| < \varepsilon\}$ différente de 0. Comme $|L|$ est discret il existe $n \geqslant 1$ tel que $|R^n(z)| < \varepsilon$ pour tout $z \in L$ tel que $|z| < 1$. Donc tout point prépériodique contenu dans $\{z \in L \mid |z| < 1\}$ appartient à l'ensemble fini $R^{-n}(0)$. $\square$

Voici une démonstration du théorème de Northcott unidimensionnelle. Si K est un corps de nombres, alors une fonction rationnelle $R \in K(z)$ a une bonne réduction pour toute valeur absolue non-archimédienne, sauf pour un nombre fini. Si l'on considère la complétion L de K pour une telle valeur absolue la proposition précédente dit que R a un nombre fini de points prépériodiques dans L et par conséquent dans $\mathbb{P}(K)$.

Les exemples canoniques de fonctions rationnelles ayant une bonne réduction, sont les polynômes moniques entiers.

Exemple 4.36. — Considérons les fonctions rationnelles *bicritiques*, qui sont les fonctions rationnelles, de degré au moins deux qui ont seulement deux points critiques, le minimum possible ; voir [**Mi**]. Notons que cette classe comprend les fonctions rationnelles de degré deux. Si l'on choisit une coordonnée telle que les points critiques soient

0 et ∞, alors les fonctions rationnelles bicritiques sont de la forme $(az^n+b)/(cz^n+d)$, avec $ad-bc=1$.

Il n'est pas difficile de voir que, si a, b, c et d sont entiers, alors la fonction rationnelle correspondante est simple. De plus si $p \not| n$ cette condition est aussi nécessaire. Si $p|n$ alors cette condition n'est pas nécessaire. Par exemple le polynôme $z^2+\frac{1}{4}$ est simple sur $\mathbb{C}_2$, car il est conjugué à $z+z^2$.

Remarque 4.37

(1) Le théorème de Morton et Silverman suit du (i) de la Proposition 4.32 et du Lemme de Schwarz.

(2) Si $R \in \mathbb{C}_p(z)$ a une bonne réduction et $D_\alpha \subset \mathcal{E}(R)$, alors D_α contient une infinité de point périodiques, tous indifférents ; voir Corollaire 5.13.

Démonstration de la Proposition 4.32. — Soit $\mathcal{S}=\{\mathcal{P}(\xi)\}_{\mathbb{P}(\overline{\mathbb{F}}_p)}$ le système projectif canonique, de telle façon que $D_{\mathcal{P}(\alpha)}=D_\alpha$.

(i) Par la partie (iv) du Lemme 2.1 un point $x \in D_\beta$, où $\beta \in \mathbb{P}(\overline{\mathbb{F}}_p)$, a au moins $\deg_R(\mathcal{P}(\xi))$ préimages dans D_ξ, pour chaque $\xi \in \mathbb{P}(\overline{\mathbb{F}}_p)$ tel que $\widetilde{R}(\xi)=\beta$.

Comme R a bonne réduction $\deg(\widetilde{R})=\deg(R)$ et par conséquent

$$\sum_{\widetilde{R}(\xi)=\beta} \deg_R(\mathcal{P}(\xi)) = \sum_{\widetilde{R}(\xi)=\beta} \deg_{\widetilde{R}}(\xi) = \deg(\widetilde{R}) = \deg(R).$$

Donc $R : D_\xi \to D_\beta$ est de degré $\deg_{\widetilde{R}}(\xi)$.

(ii) Clairement $\widetilde{R}^k(\alpha)=\alpha$ et $(\widetilde{R}^k)'(\alpha)=0$ est équivalent à ce que $R^k : D_\alpha \to D_\alpha$ soit de degré plus grand que 1. Par le Lemme de Schwarz ceci est équivalent à ce que R^k ait un point fixe attractif sur D_α.

Supposons que $R : D_0 \to D_0$ est de degré plus grand que 1. Alors

$$\widetilde{R}(z) = a_0 + a_1 z + a_2 z^2 + \cdots, \text{ avec } |a_0|<1,\ |a_1|<1 \text{ et } \max|a_i|=1.$$

Donc $r(R)=1$, où r est comme dans la Section 3.1. Par conséquent D_0 est un bassin d'attraction immédiat ; voir Section 3.1.

(iii) Supposons $D_\alpha \cap \mathcal{E}(R) \neq \varnothing$. Comme tout point de $\mathcal{E}(R)$ est récurrent par R, α est périodique par $\widetilde{R}$; voir Corollaire 3.15. Par conséquent $R^k(D_\alpha)=D_\alpha$, où $k \geqslant 1$ est la période de α. De plus $(\widetilde{R}^k)'(\alpha) \neq 0$, car sinon on a par (ii) que D_α est contenu dans le bassin d'attraction d'un cycle attractif.

D'autre part supposons que $\alpha \in \mathbb{P}(\overline{\mathbb{F}}_p)$ est périodique par $\widetilde{R}$, de période k, et $(\widetilde{R}^k)'(\alpha) \neq 0$. Alors par (i), $R^k : D_\alpha \to D_\alpha$ est de degré 1. Donc $D_\alpha \subset \mathcal{E}(R)$ par le Corollaire 3.12 de la Section 3.2.

Notons que l'on a montré que si $D_\alpha \cap \mathcal{E}(R) \neq \varnothing$, alors $D_\alpha \subset \mathcal{E}(R)$. Pour montrer que D_α est une composante analytique de $\mathcal{E}(R)$ il suffit de montrer qu'il existe une infinité de $\beta \in \mathbb{P}(\overline{\mathbb{F}}_p)$ tels que $D_\beta \cap \mathcal{E}(R) = \varnothing$.

Si $\deg(R)>1$ alors il existe $\beta \in \overline{\mathbb{F}}_p$ tel que $\widetilde{R}(\beta)=\alpha$ et tel que β n'appartient pas au cycle de α. Ceci implique que β a une infinité de préimages distinctes, donc il suffit de montrer que $D_\beta \cap \mathcal{E}(R) = \varnothing$. Comme $R(D_\beta) = D_\alpha \subset \mathcal{E}(R)$ et β n'appartient pas au cycle de α on a $D_\beta \cap \mathcal{E}(R) = \varnothing$, car R est injective sur $\mathcal{E}(R)$; voir le (ii) de la Proposition 3.9.

Il reste à montrer que, si α est de période k, alors D_α contient un point périodique de R de période k. Après changement de coordonnée on peut supposer $D_\alpha = \{|z| < 1\}$. Donc on a

$$R^k(z) = a_0 + a_1 z + a_2 z^2 + \cdots, |a_0| < 1.$$

Comme $\deg(\widetilde{R}) > 1$ et $\max_{i \geqslant 2} |a_i| = 1$ on a $\mathrm{wideg}(R^k - \mathrm{id}) > 1$. Par conséquent il existe $x \in D_\alpha = \{|z| < 1\}$ tel que $R^k(x) = x$.

(iv) Si $\widetilde{R}' \equiv 0$, alors $\widetilde{R}$ est de la forme $\widetilde{Q}(z^p)$ et donc $p|\deg(\widetilde{R}) = \deg(R)$. De plus par la partie (i) tout point périodique de R est attractif.

Supposons que $\widetilde{R}' \not\equiv 0$ et soit x un point périodique attractif de R de période k. Après un changement de coordonnée on suppose $x = 0$. Alors

$$R^k(z) = a_1 z + a_2 z^2 + \cdots, \text{ avec } |a_1| < 1, |a_i| \leqslant 1$$

et il existe un entier $i \geqslant 1$ tel que $|ia_i| = 1$ (car $\widetilde{R}' \not\equiv 0$). Donc $0 < \mathrm{wideg}((R^k)') < \infty$ et par conséquent il existe une solution de l'équation $(R^k)'(z) = 0$ dans $\{|z| < 1\}$. Donc le bassin d'attraction immédiat du cycle de 0 contient un point critique de R.

En particulier le nombre de cycles attractifs est au plus $2\deg(R) - 2$ et donc R a une infinité des points périodiques indifférents. Par conséquent $\mathcal{E}(R) \neq \varnothing$. $\square$

5. Composantes analytiques du domaine de quasi-périodicité

Rappelons que la composante analytique d'un ensemble $U \subset \mathbb{P}(\mathbb{C}_p)$ qui contient un point $x \in U$ est l'union de tous les affinoïdes fermés connexes contenus dans U qui contiennent x.

Théorème 3. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins deux et C une composante analytique de $\mathcal{E}(R)$. Alors C est un affinoïde ouvert connexe, c'est-à-dire ;*

$$C = \mathbb{P}(\mathbb{C}_p) - B_0 \cup \cdots \cup B_n,$$

où $n \geqslant 0$ et $B_0, \ldots, B_n$ sont des boules fermées. De plus, chaque bout de C (dont la boule fermée associée est l'un des B_i) est périodique et le degré de l'action induite par l'itéré correspondant sur le système projectif associé est plus grand que 1.

On peut voir les composantes analytiques du domaine de quasi-périodicité comme l'analogue p-adique des disques de Siegel et des anneaux de Herman en dynamique complexe. Mais il y a des différences : sur $\mathbb{C}_p$ on peut avoir des composantes avec un nombre fini arbitraire de bouts ; voir les exemples dans la Section 6. Par analogie avec le cas complexe on considère la définition suivante.

Définition 5.1. — Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins deux, C une composante analytique de $\mathcal{E}(R)$ et $n+1$ le nombre de bouts de C. Alors on dit que C est un *n-anneau de Herman* si $n > 0$ et on dit que C est un *disque de Siegel* si $n = 0$.

La proposition suivante ne dépend pas du théorème.

Proposition 5.2. — *Soit X un affinoïde ouvert connexe et $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins 2 telle que $R : X \to X$ est de degré 1. Alors $X \subset \mathcal{E}(R)$. Si de plus pour tout système projectif $\mathcal{S}$ il existe $i \geqslant 1$ tel que $\deg_{R^i}(\mathcal{S}) > 1$, alors X est une composante analytique de $\mathcal{E}(R)$.*

Démonstration. — On montre d'abord que $X \subset \mathcal{E}(R)$. Si X est un disque ceci est une conséquence du Corollaire 3.12. Comme $\deg(R) > 1$ on a $X \neq \mathbb{P}(\mathbb{C}_p)$, donc si X n'est pas un disque on a $\mathcal{A}_X \neq \varnothing$.

D'après le Corollaire 2.16 il existe n tel que R^n induit l'identité sur $\mathcal{A}_X$. Si $\mathcal{S} \in \mathcal{A}_X$ alors $\mathcal{S} \prec X$, donc on a $\deg_R(\mathcal{S}) = 1$ par le Lemme 2.11. Par conséquent il existe m tel que R^{nm} induit l'identité sur $\mathcal{S}$. Donc si $\mathcal{P} \in \mathcal{S}$ est tel que $D_{\mathcal{P}} \subset X$ alors $R^{mn} : D_{\mathcal{P}} \to D_{\mathcal{P}}$ est de degré 1. Par le Corollaire 3.12 on a $D_{\mathcal{P}} \subset \mathcal{E}(R)$. Par conséquent $X_{\mathcal{S}} \subset \mathcal{E}(R)$ et donc $X \subset \mathcal{E}(R)$.

Soit C la composante analytique de $\mathcal{E}(R)$ qui contient X. Si $\mathcal{S}$ est un système projectif associé à un bout $\mathcal{P}$ de X tel que $\deg_{R^i}(\mathcal{S}) > 1$, alors on a $\mathcal{S} \not\prec C$ par le Lemme 2.11, car R est injective sur $C \subset \mathcal{E}(R)$. C'est-à-dire $C \subset D_{\mathcal{P}}$. Par conséquent on a

$$X \subset C \subset \cap_{\text{bout de } C} D_{\mathcal{P}} = X. \qquad \square$$

On a le corollaire suivant du Théorème 3.

Corollaire 5.3. — *Soit $R \in \mathbb{C}_p(z)$. Alors tout bout $\mathcal{P} \prec \mathcal{E}(R)$ est périodique par R.*

Démonstration. — Si $\mathcal{P}$ est un bout d'une composante analytique de $\mathcal{E}(R)$ alors il est périodique par le théorème précédent. Sinon il existe un affinoïde fermé $X \prec \mathcal{E}(R)$ tel que $\mathcal{P} \prec X$ et par conséquent $\mathcal{P}$ est périodique par le Corollaire 4.20. $\square$

La démonstration du Théorème 3 occupe le reste de cette section.

Lemme d'Approximation. — *Soit R une fonction rationnelle, fixant le bout $\mathcal{P}_0$ associé à $\mathbb{D}_p = \{|z| < 1\}$ et injective dans la couronne $\{r < |z| < 1\}$.*

(1) *Pour tout bout $\mathcal{P}$ tel que la boule fermée associée $B_{\mathcal{P}}$ vérifie $B_{\mathcal{P}} \subset \mathbb{D}_p$ et $\operatorname{diam}(B_{\mathcal{P}}) \geqslant r$, l'image $R(\mathcal{P})$ vérifie $B_{R(\mathcal{P})} \subset \mathbb{D}_p$ et $\operatorname{diam}(B_{R(\mathcal{P})}) = \operatorname{diam}(B_{\mathcal{P}})$.*

(2) *Il existe un fonction rationnelle R_0 injective dans $\mathbb{D}_p$, fixant $\mathcal{P}_0$, telle que pour tout bout $\mathcal{P}$ vérifiant les hypothèses précédentes on ait $R_0(\mathcal{P}) = R(\mathcal{P})$.*

Démonstration. — Soit h un automorphisme de $\mathbb{P}(\mathbb{C}_p)$ préservant $\mathbb{D}_p$ tel que $h \circ R(\{r < |z| < 1\}) = \{r < |z| < 1\}$. Si $\widehat{R}_0$ vérifie les conclusions du lemme par rapport à $h \circ R$, alors $R_0 = h^{-1} \circ \widehat{R}_0$ les vérifiera par rapport à R. On peut donc se ramener à $h = \mathrm{id}$; par conséquent, pour tout $r < |z| < 1$ on a $|R(z)| = |z|$.

Comme $\deg_R(\mathcal{P}_0) = 1$ et par (iv) du Lemme 2.1, si n est le nombre des pôles de R sur $\mathbb{D}_p$, alors R a $n+1$ zéros sur $\mathbb{D}_p$. Soient $a_1, \ldots, a_{n+1}$ et $b_1, \ldots, b_n$ les zéros et les pôles de R sur $\mathbb{D}_p$ respectivement. On pose

$$Q(z) = \prod_{i=1}^{n} \Big(\frac{z - b_i}{z - a_i}\Big) = \prod_{i=1}^{n} \Big(1 + \frac{a_i - b_i}{z - a_i}\Big).$$

Donc $R_0 = R \cdot Q : \mathbb{D}_p \to \mathbb{D}_p$ est de degré 1 et par conséquent, si le bout $\mathcal{P}$ vérifie les conditions dans 1 on a $\mathrm{diam}(B_{R_0(\mathcal{P})}) = \mathrm{diam}(B_{\mathcal{P}}) \geqslant r$. D'autre part on a

$$\Big|\frac{a_i - b_i}{z - a_i}\Big| \leqslant r|z|^{-1} \quad \text{et} \quad |R(z)| = |z|$$

pour tout $z \in \{r < |z| < 1\}$. Donc

$$|R_0(z) - R(z)| = |R(z)| \Big|\prod_{i=1}^{n} \Big(1 + \frac{a_i - b_j}{z - a_i}\Big) - 1\Big| \leqslant r.$$

On a alors $R_0(\mathcal{P}) = R(\mathcal{P})$ par le Lemme 2.3. Donc $B_{\mathcal{P}} \subset \mathbb{D}_p$ et $\mathrm{diam}(B_{R(\mathcal{P})}) = \mathrm{diam}(B_{R_0(\mathcal{P})}) = \mathrm{diam}(B_{\mathcal{P}})$. □

Corollaire 5.4. — *Soient $R \in \mathbb{C}_p(z)$ une fonction rationnelle et $D_i = \{|z - z_i| < r_i\}$ des disques, pour $0 \leqslant i \leqslant k$, avec $D_0 = D_k$. On suppose que pour $0 \leqslant i < k$ l'image par R du bout associé à D_i est le bout associé à D_{i+1} et que R est injective sur $\{rr_i < |z - z_i| < r_i\}$. Alors tout bout $\mathcal{P}$ dont la boule fermée associée $B_{\mathcal{P}}$ vérifie $B_{\mathcal{P}} \subset D_0$ et $\mathrm{diam}(B_{\mathcal{P}}) \geqslant rr_0$ est périodique sous l'action de R.*

Démonstration. — Par le lemme il existe des fonctions rationnelles R_i, pour $0 \leqslant i < k$ telles que $R : D_i \to D_{i+1}$ est de degré 1 et tel que pour tout bout $\mathcal{Q}$ qui vérifie $B_{\mathcal{Q}} \subset D_0$ et $\mathrm{diam}(B_{\mathcal{Q}}) \geqslant rr_0$ le bout $R^k(\mathcal{P}) = R_{k-1} \circ \cdots \circ R_0(\mathcal{P})$, vérifie les mêmes propriétés. Donc il suffit de vérifier que $\mathcal{P}$ est périodique par $Q = R_{k-1} \circ \cdots \circ R_0$. Comme $Q : D_0 \to D_0$ est de degré 1, on a par le Corollaire 3.12 que $D_0 \subset \mathcal{E}(Q)$. Par le Corollaire 4.20 de la Section 4.2, $\mathcal{P}$ est périodique par Q. □

Lemme 5.5. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins deux et C une composante analytique de $\mathcal{E}(R)$. Pour tout $x \notin C$ il existe un bout $\mathcal{P}_x \prec C$ périodique par R tel que $x \in B_{\mathcal{P}_x}$ et $B_{\mathcal{P}_x} \cap C = \varnothing$. De plus si $\mathcal{S}_x$ est le système projectif associé à $\mathcal{P}_x$ et n_x est la période de $\mathcal{P}_x$, alors $\deg_{R^{n_x}}(\mathcal{S}_x) > 1$.*

Démonstration. — On suppose que C est fixe par R. Soit $\mathcal{P}_0$ un bout tel qu'il existe un affinoïde fermé $X \subset C$ tel que $D_{\mathcal{P}_0} \subset X$. Par le Lemme 4.21 il existe $n \geqslant 1$ tel que $R^n : D_{\mathcal{P}_0} \to D_{\mathcal{P}_0}$ est de degré 1. Donc quitte à remplacer R par un itéré on peut supposer que $R : D_{\mathcal{P}_0} \to D_{\mathcal{P}_0}$ est de degré 1. Après changement de coordonnée on peut supposer que $D_{\mathcal{P}_0} = \{|z| > 1\} \cup \{\infty\}$.

Soit I la composante d'injectivité de R qui contient C. Rappelons que I est un affinoïde ouvert connexe ; voir Proposition 2.9 . Notons qu'on a $D_{\mathcal{P}_0} \subset I$.

(1) Supposons que $\mathcal{P}$ est un bout qui satisfait les propriétés suivantes.

(i) $x \in B_{\mathcal{P}}$.

(ii) $\mathcal{P}$ est périodique par R de période N et $R^i(\mathcal{P}) \prec C$ pour tout $0 \leqslant i < N$.

(iii) $B_{R^i(\mathcal{P})} \subset \{|z| \leqslant 1\}$ et il existe $r \in |\mathbb{C}_p|$ tel que $\operatorname{diam}(B_{R^i(\mathcal{P})}) = r$, pour $0 \leqslant i < N$.

Notons que $\mathcal{P}_0$ satisfait ces propriétés avec $r = 1$.

Soit $\mathcal{S}$ le système projectif associé à $\mathcal{P}$; on suppose que $\deg_{R^N}(\mathcal{S}) = 1$. Alors on va trouver un bout $\mathcal{P}'$ qui satisfait aussi (i)-(iii) ci-dessus, avec $r' < r$ appartenant à un ensemble fini.

Donc on conclut qu'il existe un bout $\mathcal{P}$ qui satisfait (i)-(iii) et que de plus $\deg_{R^N}(\mathcal{S}) > 1$; alors on continue en 3.

(2) Quitte à remplacer N par un multiple, on peut supposer que R^N induit l'identité sur $\mathcal{S}$. Par (iii) de la Proposition 2.4, sauf pour un nombre fini de $\mathcal{Q} \in \mathcal{S}$, $R : D_{\mathcal{Q}} \to D_{\mathcal{Q}}$ est de degré 1 et par le Corollaire 3.12 on a $D_{\mathcal{Q}} \subset \mathcal{E}(R)$. Par conséquent $\mathcal{S} \prec C$ et $R^i(\mathcal{S}) \prec C$, pour $0 \leqslant i < N$. En particulier $B_{R^i(\mathcal{P})} \cap I \neq \varnothing$.

D'autre part, $\cup B_{R^i(\mathcal{P})} \not\subset I$, sinon $R^N : B_{\mathcal{P}} \to B_{\mathcal{P}}$ est de degré 1 et par conséquent $B_{\mathcal{P}} \subset \mathcal{E}(R)$ ce qui implique $x \in B_{\mathcal{P}} \subset C$, mais $x \notin C$ par hypothèse. Par conséquent on a

$$0 < r' = \max_{0 \leqslant i < N} \operatorname{diam}(B_{R^i(\mathcal{P})} - I) < r.$$

(2.1) Soit $\mathcal{P}'$ le bout associé à $\{|z - x| \leqslant r'\} \subset \{|z| \leqslant 1\}$, donc $\mathcal{P}'$ satisfait la propriété (i) par définition. Soit $\mathcal{Q}' \in \mathcal{S}$ le bout tel que $x \in D_{\mathcal{Q}'}$. Notons que $R^N(\mathcal{Q}') = \mathcal{Q}'$ et $\deg_{R^N}(\mathcal{Q}') = 1$ car R^N induit l'identité sur $\mathcal{S}$. Donc il existe une couronne ouverte A ayant $\mathcal{Q}'$ comme bout telle que $R : A \to A$ est de degré 1. Par la Proposition 5.2 on a $A \subset \mathcal{E}(R)$ et comme $\mathcal{S} \prec C$ on a $\mathcal{Q}' \prec C$.

Par définition de r' on a $D_{\mathcal{Q}'} - B_{\mathcal{P}'} \subset I$, donc R est injective sur la couronne $D_{\mathcal{Q}'} - B_{\mathcal{P}'}$. Donc

$$B_{R(\mathcal{P}')} \subset D_{R(\mathcal{Q}')} \subset B_{\mathcal{P}} \subset \{|z| \leqslant 1\},$$

et par le Lemme d'Approximation on a $\operatorname{diam}(R(\mathcal{P}')) = \operatorname{diam}(B_{\mathcal{P}'}) = r'$. Alors par définition de r' on a $D_{R(\mathcal{Q}')} - B_{R(\mathcal{P}')} \subset I$.

Alors on peut montrer par induction (en utilisant le Lemme d'Approximation) que pour tout $i \geqslant 0$ on a $B_{R^i(\mathcal{P}')} \subset \{|z| \leqslant 1\}$ et $\operatorname{diam}(B_{R^i(\mathcal{P}')}) = r'$; c'est-à-dire (iii) du 1. On applique le Corollaire 5.4 à les disques $D_{R^i(\mathcal{Q}')}$ et $r = r'$ et on obtient que $\mathcal{P}'$ est périodique par R.

(2.2) Il reste à montrer que $\mathcal{P}' \prec C$. Comme dans 2.1 on peut montrer que tout bout $\widehat{\mathcal{P}}$ tel que $B_{\widehat{\mathcal{P}}} \subset B_{\mathcal{P}}$ et $\operatorname{diam}(B_{\widehat{\mathcal{P}}}) = r'$ est périodique par R ; soit $n(\widehat{\mathcal{P}})$ la période

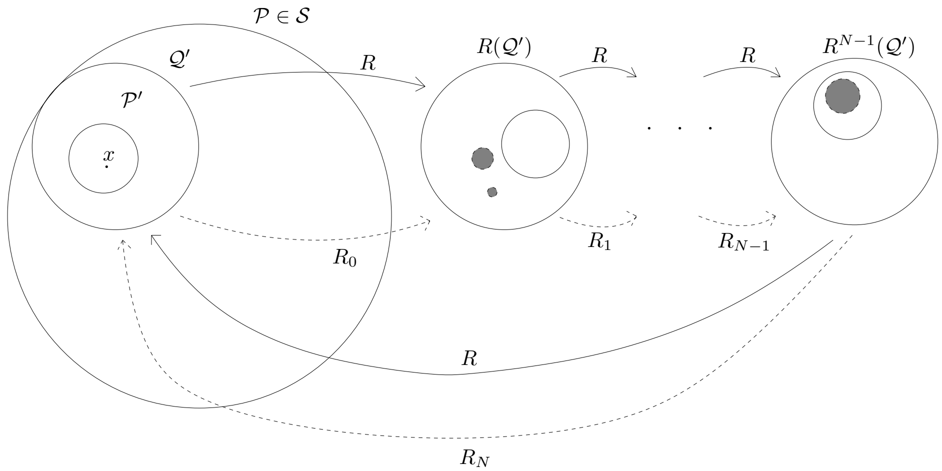

FIGURE 3. La partie grise représente le complémentaire de I.

de $\widehat{\mathcal{P}}$. Si de plus

$$\cup B_{R^i(\widehat{\mathcal{P}})} \subset I, \text{ pour } 0 \leqslant i < n(\widehat{\mathcal{P}}),$$

on a que $R^{n(\widehat{\mathcal{P}})} : B_{\widehat{\mathcal{P}}} \to B_{\widehat{\mathcal{P}}}$ est de degré 1 et on a $B_{\widehat{\mathcal{P}}} \subset \mathcal{E}(R)$ par le Corollaire 3.12.

Notons que l'ensemble $\mathcal{T}$ des bouts $\widehat{\mathcal{P}}$ tels que $B_{\widehat{\mathcal{P}}} \subset B_{\mathcal{P}}$, $\operatorname{diam}(B_{\widehat{\mathcal{P}}}) = r'$ et

$$\cup B_{R^i(\widehat{\mathcal{P}})} \not\subset I, \text{ pour } 0 \leqslant i < n(\widehat{\mathcal{P}}),$$

est fini. Par conséquent l'ensemble $D_{\mathcal{Q}'} - \cup_{\mathcal{T}} B_{\widehat{\mathcal{P}}} \subset \mathcal{E}(R)$ est un affinoïde ouvert connexe. Comme $\mathcal{Q}' \prec C$ cet affinoïde ouvert est contenu dans C. Donc $\mathcal{P}' \prec C$.

(3) Le bout $\mathcal{P}_0$ satisfait (i)-(iii) du 1. Dans 2 on a montré que si $\mathcal{P}$ est un bout qui satisfait ces propriétés et $\deg_{R^N}(\mathcal{S}) = 1$ (où $\mathcal{S}$ est le système projectif associé à $\mathcal{P}$) alors il existe un bout $\mathcal{P}'$ dont $r' < r$ appartient à l'ensemble fini,

$$\{\operatorname{diam}(B - I) \mid B \subset \mathbb{C}_p \text{ boule fermée telle que } B \cap I \neq \varnothing\}.$$

Par conséquent il existe un bout $\mathcal{P}$ qui satisfait les propriétés (i)-(iii) du 1 et de plus $\deg_{R^N}(\mathcal{S}) > 1$.

(3.1) On va montrer qu'un tel bout $\mathcal{P}_x = \mathcal{P}$ satisfait les assertions du lemme. Comme $\mathcal{P}_x$ satisfait (i)-(iii) il reste à montrer que $B_{\mathcal{P}_x} \cap C = \varnothing$. Si $B_{\mathcal{P}_x} \cap C \neq \varnothing$ alors $\mathcal{S} \prec C$ car $\mathcal{P}_x \prec C$. Par conséquent $R^i(\mathcal{S}) \prec C$ et $\deg_R(R^i(\mathcal{S})) = 1$, pour $0 \leqslant i < N$, par le Lemme 2.11. Mais ceci contredit $\deg_{R^N}(\mathcal{S}) > 1$. Donc $B_{\mathcal{P}_x} \cap C = \varnothing$. □

Démonstration du Théorème 3. — Sans perte de généralité on suppose que C est fixé par R. Par le lemme précédent, pour chaque $x \notin \mathcal{E}(R)$ il existe un bout $\mathcal{P}_x \prec C$ périodique par R de période n_x tel que $x \in B_{\mathcal{P}_x}$, $B_{\mathcal{P}_x} \cap C = \varnothing$ et tel que si $\mathcal{S}_x$ est le système projectif associé à $\mathcal{P}_x$ alors $\deg_{R^{n_x}}(\mathcal{S}_x) > 1$.

Notons que $R^i(\mathcal{P}_x) \prec C$ et $B_{R^i(\mathcal{P}_x)} \cap C = \varnothing$, donc $A_x = \cup_{i \geqslant 0} B_{R^i(\mathcal{P}_x)}$ est disjoint de C. De plus, si $A_x \cap A_y \neq \varnothing$ alors $A_x = A_y$. Comme $\deg_{R^{n_x}}(S_x) > 1$, on a $R^{n_x}(B_{\mathcal{P}_x}) = \mathbb{P}(\mathbb{C}_p)$, sinon $R^{n_x} : B_{\mathcal{P}_x} \to B_{\mathcal{P}_x}$ est de degré 1 par le Corollaire 2.2 et par conséquent $x \in B_{\mathcal{P}_x} \subset C$ par le Corollaire 3.12. Donc on a $R(A_x) = \mathbb{P}(\mathbb{C}_p)$ par le Corollaire 2.2.

Fixons $z \in C$. Alors C et chaque A_x contient un point de $R^{-1}(z)$, donc il y a au plus $\deg(R) - 1$ ensembles A_x disjoints deux à deux. Par conséquent $C = \mathbb{P}(\mathbb{C}_p) - \cup_{x \notin C} A_x$ est un affinoïde ouvert et les $\mathcal{P}_x$ sont ses bouts. La dernier assertion du théorème suit du fait que $\deg_{R^{n_x}}(\mathcal{S}_x) > 1$. □

5.1. Dynamique sur une composante analytique. — Considérons une fonction rationnelle complexe $R \in \mathbb{C}(z)$ ayant une disque de Siegel ou un anneau de Herman C fixé par R. Alors R est conjuguée sur C à une rotation irrationnelle par une application holomorphe ; voir [**CG**].

Soit $\alpha \in \mathbb{R} - \mathbb{Q}$ le nombre de rotation correspondant. Alors pour tout $w \in \mathbb{T} = \mathbb{R}/\mathbb{Z}$ on peut définir une automorphisme (holomorphe) $R^{\circ w}$ de C de nombre de rotation w. Donc on a une action de $\mathbb{T}$ sur C définie par $(w, z) \to R^{\circ w}(z)$. De plus pour toute suite $\{n_j\}_{j \geqslant 1}$ d'entiers positifs telle que $n_j \alpha \to w$ dans $\mathbb{T}$, on a que R^{n_j} converge uniformément à $R^{\circ w}$ sur chaque partie compacte de C.

Dans cette section on montre des propriétés analogues sur $\mathbb{C}_p$, avec l'anneau $\mathbb{Z}_p = \{w \in \mathbb{Q}_p \mid |w| \leqslant 1\}$ au lieu de $\mathbb{T}$. L'anneau $\mathbb{Z}_p$ est compacte et $\mathbb{Z} \subset \mathbb{Z}_p$ est dense sur $\mathbb{Z}_p$.

Une différence avec le cas complexe est qu'il ne suffit pas que la composante soit fixée. Il faut que les bouts de la composante soient fixés et que la fonction rationnelle soit tangente à l'identité en chaque bout de C. On dit qu'une fonction rationnelle $R \in \mathbb{C}_p(z)$ est *tangente à l'identité en un bout* $\mathcal{P}$ si R fixe $\mathcal{P}$ et si la multiplicité de $\mathcal{P}$ comme point fixe de l'action de R dans le système projectif associé à $\mathcal{P}$ est strictement plus grande que 1.

Proposition 5.6. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins deux et soit C une composante analytique de $\mathcal{E}(R)$ de telle façon que R fixe C et les bouts de C. De plus supposons que R est tangente à l'identité en chaque bout de C. Alors on les propriétés suivantes.*

(1) *Pour chaque $w \in \mathbb{Z}_p$ il existe un automorphisme $R^{\circ w}$ de C tel que pour toute suite $\{n_j\}_{j \geqslant 1}$ d'entiers positifs telle que $n_j \to w$ au sens p-adique, R^{n_j} converge uniformément (pour la métrique chordale) à $R^{\circ w}$ sur chaque affinoïde fermé contenu dans C.*

(2) *L'application qui à $(w, z) \in \mathbb{Z}_p \times C$ associe $R^{\circ w}(z) \in C$ définit une action. En particulier $R^{\circ 0}$ est l'identité et on a $R^{\circ(w_1 + w_2)} = R^{\circ w_1} \circ R^{\circ w_2}$ pour tous w_1, $w_2 \in \mathbb{Z}_p$.*

Comme $\mathbb{Z}_p$ est compact le corollaire suivant est immédiat.

Corollaire 5.7. — *La suite* $\{R^n\}_{n\geqslant 1}$ *est normale sur chaque composante analytique* C *de* $\mathcal{E}(R)$ *; c'est-à-dire que pour toute suite d'entiers positifs il existe une sous-suite* $\{m_j\}_{j\geqslant 1}$ *telle que* R^{m_j} *est uniformément convergent sur chaque affinoïde fermé contenu dans* C.

La démonstration de la proposition dépend de deux lemmes.

Lemme 5.8. — *Soit* $R \in \mathbb{C}_p(z)$ *une fonction rationnelle qui soit tangente à l'identité au bout associé à* $\{|z| < 1\}$. *Alors il existe* $r_0 \in (0,1)$ *tel que* R *induit l'identité sur le système projectif associé à* $\{|z| < r\}$, *pour chaque* $r \in (r_0, 1)$.

Démonstration. — Comme R est tangente à l'identité au bout associé à $\{|z| < 1\}$ on a $R(z) = z + z^2 P(z)/Q(z)$ où $P, Q \in \mathbb{C}_p[z]$ sont des polynômes à coefficients entiers tels que $\widetilde{Q}(0) \neq 0$. Par conséquent il existe $r_0 \in (0,1)$ tel que pour chaque λ qui satisfait $r_0 < |\lambda| < 1$ on a $|P(\lambda)| \leqslant 1$ et $|Q(\lambda)| = 1$; voir Lemme 2.1. Donc la réduction de

$$\lambda^{-1}R(\lambda z) = z + \lambda z^2 \frac{P(\lambda z)}{Q(\lambda z)}$$

est l'identité. $\square$

Lemme 5.9. — *Soit* C *un affinoïde ouvert et* $R \in \mathbb{C}_p(z)$ *une fonction rationnelle telle que* $R : C \to C$ *est de degré* 1 *et tel que* R *fixe les bouts de* C. *De plus supposons que* R *est tangente à l'identité en chaque bout de* C. *Alors* R *induit l'identité sur* $\mathcal{A}_C$ *et sur chaque système projectif* $\mathcal{S} \in \mathcal{A}_C$.

Démonstration. — Comme $R : C \to C$ est de degré 1, R induit une isométrie sur $\mathcal{A}_C$; voir Section 2.5. Comme R fixe les bouts de C, R induit l'identité sur $\mathcal{A}_C$.

Comme C est un affinoïde ouvert, chaque sommet $\mathcal{S}$ de $\mathcal{A}_C$ est l'extrémité d'au moins trois arêtes de $\mathcal{A}_C$. Par conséquent R fixe au moins trois éléments d'un sommet. Donc R induit l'identité sur chaque sommet de $\mathcal{A}_C$.

Donc pour toute arête $I = (\mathcal{S}_0, \mathcal{S}_1) \subset \mathbb{H}_p$ de $\mathcal{A}_C$, R est tangente à l'identité en chaque bout de C_I. Par le lemme précédent pour chaque $i \in \{0,1\}$ il existe un système projectif $\mathcal{S}'_i$ proche de $\mathcal{S}_i$ tel que R induit l'identité sur chaque système projectif dans $(\mathcal{S}_i, \mathcal{S}'_i]$. En particulier R induit l'identité sur $\mathcal{S}'_0$ et $\mathcal{S}'_1$. Alors R induit l'identité sur chaque système projectif dans $[\mathcal{S}'_0, \mathcal{S}'_1]$, par le Lemme 4.23. $\square$

Démonstration de la Proposition 5.6. — Comme le degré de R est au moins deux $C \neq \mathbb{P}(\mathbb{C}_p)$ et on peut supposer que $C \subset \mathbb{C}_p$.

Considérons une suite croissante d'affinoïdes fermés $X_i \subset C$ de telle façon que $\cup X_i = C$. Dans le cas où C n'est pas un disque R satisfait (avec $n = 1$) les propriétés de la Proposition 4.14 pour X_i, par le lemme précédent. On notera $\gamma_i \in (0,1)$ une constante telle que $\|R - \mathrm{id}\,\|_{X_i} \leqslant \gamma_i \cdot \mathrm{diam}(D_x)$ pour $x \in X_i$.

Si C est un disque, on peut supposer que $C = \{|z| < 1\}$ et $X_i = \{|z| \leqslant r_i\}$, avec $r_i \in |\mathbb{C}_p| \cap (0,1)$ tel que $r_i \to 1$. Alors par le Lemme 5.8 on peut supposer que R

induit l'identité sur le système projectif $\mathcal{S}_i$ associé à X_i, pour chaque i. Comme dans ce cas $\mathcal{A}_{X_i} = \{\mathcal{S}_i\}$, R satisfait les propriétés de la Proposition 4.14 et alors on peut continuer comme dans le cas où C n'est pas un disque.

Étant donné $w \in \mathbb{Z}_p$ considérons une suite croissante $\{m_i\}_{i\geqslant 1}$ d'entiers positifs de façon que $\rho(|w - m_i|_p) < \gamma_j^{-1}$, pour tout $j \geqslant i$. Alors on a $R^{\circ(w-m_i)} \in \mathcal{H}(X_i)$ d'après le Corollaire 4.16. Notons que $R^{\circ(w-m_{i+1})}$ définie sur X_i coïncide avec la restriction de la même application définie sur X_{i+1}.

De plus on a $\rho(|m_{i+i} - m_i|_p) \leqslant \gamma_i^{-1}$, donc $R^{\circ(m_{i+1}-m_i)}$ est définie sur X_i et coïncide avec $R^{m_{i+1}-m_i}$. Par conséquent $R^{\circ(w-m_{i+1})}$ définie sur X_{i+1} et $R^{\circ(w-m_i)}$ définie sur X_i, satisfont l'identité suivante sur X_i,

$$R^{m_{i+1}} \circ R^{\circ(w-m_{i+1})} = R^{m_i} \circ R^{\circ(m_{i+1}-m_i)} \circ R^{\circ(w-m_{i+1})} = R^{m_i} \circ R^{\circ(w-m_i)}.$$

Donc la fonction $R^{\circ w}$ définie par $R^{m_i} \circ R^{\circ(w-m_i)}$ sur X_i appartient à $\mathcal{H}(C)$. Notons qu'on a $R^{\circ 0} = \mathrm{id}$; voir Corollaire 4.16.

Soit $\{n_j\}_{j\geqslant 1}$ une suite d'entiers positifs tels que $n_j \to w$ au sens p-adique et fixons $i \geqslant 0$. Alors il existe $J \geqslant 0$ tel que $\rho(|n_j - m_i|_p) < \gamma_i^{-1}$ pour tout $j \geqslant J$. Donc $R^{n_j - m_i} = R^{\circ(n_j-m_i)}$ sur X_i et par conséquent $R^{n_j} = R^{m_i} \circ R^{\circ(n_j-m_i)}$ converge uniformément à $R^{\circ w} = R^{m_i} \circ R^{\circ(w-m_i)}$ sur X_i ; voir Corollaire 4.16.

Soient $w_1, w_2 \in \mathbb{Z}_p$; considérons des suites $\{n_j\}_{j\geqslant 1}$ et $\{k_j\}_{j\geqslant 1}$ d'entiers positifs tels que $n_j \to w_1$ et $k_j \to w_2$ au sens p-adique. Comme R^{n_j} et R^{m_j} convergent uniformément à $R^{\circ w_1}$ et $R^{\circ w_2}$ sur chaque X_i on a que $R^{n_j+k_j}$ converge uniformément à $R^{\circ(w_1+w_2)}$ sur chaque X_i. Par conséquent $R^{\circ w_1} \circ R^{\circ w_2} = R^{\circ(w_1+w_2)}$.

Comme pour chaque $w \in \mathbb{Z}_p$ on a $R^{\circ w} \circ R^{\circ(-w)} = R^{\circ 0} = \mathrm{id}$, on a que $R^{\circ(-w)} \in \mathcal{H}(C)$ est l'inverse de $R^{\circ w}$ et par conséquent $R^{\circ w}$ est un automorphisme de C. $\square$

5.2. Composantes analytiques et points périodiques. — Dans cette section on s'intéresse aux points périodiques contenus dans une composante analytique donnée du domaine de quasi-périodicité.

Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins deux et soit C une composante analytique de $\mathcal{E}(R)$ fixée par R. Les bouts de C sont permutés par R. Notons $\mathcal{P}_0, \ldots, \mathcal{P}_n$, les bouts *fixés* par R. Pour chaque $0 \leqslant i \leqslant n$, R induit une action projective (de degré au moins 2 par le Théorème 3) sur le système projectif $\mathcal{S}_i$ associé à $\mathcal{P}_i$; on note $n_R(\mathcal{P}_i)$ la multiplicité du point fixe $\mathcal{P}_i$ pour cette action. Rappelons que quand $n_R(\mathcal{P}_i) > 1$ on dit que R est tangent à l'identité en $\mathcal{P}_i$.

Proposition 5.10. — *Le nombre de points fixes de R dans C comptés avec multiplicité est égal à*

$$2 + \sum_{0\leqslant i\leqslant n} (n_R(\mathcal{P}_i) - 2).$$

La démonstration de cette proposition est à la fin de cette section. On considère d'abord des corollaires et un exemple.

Corollaire 5.11. — *Si C est un disque de Siegel fixe, alors C contient au moins un point fixe.*

Corollaire 5.12. — *Soit $R \in \mathbb{C}_p(z)$ de degré au moins deux et C une composante analytique de $\mathcal{E}(R)$ fixée par R. Si aucun des bouts de C n'est fixé par R, alors C contient exactement deux points fixes de R, comptés avec multiplicités.*

Corollaire 5.13. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins deux. Alors chaque composante analytique de $\mathcal{E}(R)$ continent une infinité de points périodiques de R.*

Démonstration. — Quitte à remplacer R par un itéré on suppose que tous les bouts $\mathcal{P}_1, \ldots, \mathcal{P}_n$ de C sont fixes et que R est tangente à l'identité en $\mathcal{P}_i$, pour $1 \leqslant i \leqslant n$. Notons que pour $k \geqslant 1$ on a $n_{R^{pk}}(\mathcal{P}_i) > n_{R^k}(\mathcal{P}_i)$. Par conséquent le nombre de points fixes de R^{p^m} dans C (qui est égal à $2 + \sum(n_{R^{p^m}}(\mathcal{P}_i) - 2)$ par la proposition) tend vers l'infini quand $m \to \infty$. Chaque point fixe de R^k a une multiplicité comme point fixe de R^{lk} qui est bornée indépendamment de l. Donc R a une infinité de points périodiques dans C ; voir Section 3.3. $\square$

Corollaire 5.14. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle et C une composante analytique de $\mathcal{E}(R)$ fixée par R. Supposons que R fixe les bouts de C et que R est tangente à l'identité en chaque bout de C. Alors tous les points périodiques de R dans C ont une période primitive de la forme p^m, $m \geqslant 0$.*

Démonstration. — Simplement note que pour tout $m \geqslant 0$ et tout $k \geqslant 1$ qui n'est pas divisible par p on a $n_{R^{kp^m}}(\mathcal{P}) = n_{R^{p^m}}(\mathcal{P})$ pour tout bout $\mathcal{P}$ de C ; voir Section 3.3. $\square$

Voici d'autres différences entre le cas complexe et le cas p-adique. Dans le cas complexe les disques de Siegel et les anneaux de Herman, coïncident avec le domaine de linéarisation, mais sur $\mathbb{C}_p$ tous les composantes analytiques du domaine de quasi-périodicité contiennent une infinité de points périodiques et par conséquent les domaines de linéarisation sont strictement plus petits que les composantes analytiques correspondantes.

Une autre différence est que dans le cas complexe les disques de Siegel et les anneaux de Herman sont instables (voir [**Ma**]) ; c'est-à-dire que pour chaque fonction rationnelle à coefficients complexes on peut trouver des fonctions rationnelles arbitrairement proches qui n'ont pas de disque de Siegel ni de anneaux d'Herman. Dans le cas p-adique, si $R \in \mathbb{C}_p(z)$ est une fonction rationnelle ayant l'affinoïde ouvert connexe C comme composante analytique de $\mathcal{E}(R)$, alors pour toute fonction rationnelle $Q \in \mathbb{C}_p(z)$ proche de R, C est une composante analytique de $\mathcal{E}(Q)$.

Exemple 5.15. — Étant donnés a, b, $\lambda \in \mathbb{C}_p$ tels que $|a| = |b| < 1$ et $|\lambda| = 1$ considérons la fonction rationnelle

$$R(z) = \lambda z^2 \frac{1 - bz}{z - a}$$

et la couronne $C = \{z \in \mathbb{C}_p \mid |a| < |z| < |b|^{-1}\}$. Notons qu'on a

$$\frac{1}{R(1/z)} = \lambda^{-1} z^2 \frac{1 - az}{z - b}.$$

(1) On va montrer que C est une composante analytique de $\mathcal{E}(R)$. Notons que $R : C \to C$ est de degré 1 et R fixe les bouts de C. Soit $\mathcal{P}_a$ le bout associé à $\{|z| > |a|\} \cup \{\infty\}$ et $\mathcal{P}_b$ le bout associé à $\{|z| < |b|^{-1}\}$; ce sont les bouts de C. La coordonnée $w_a = -a/z$ (resp. $w_b = bz$) est telle que $w_a(\mathcal{P}_a)$ (resp. $w_b(\mathcal{P}_b)$) est le bout associé à $\{|z| < 1\}$. De plus

$$R_a(z) = w_a \circ R \circ w_a^{-1}(z) = \lambda^{-1} z^2 \frac{z+1}{z+ab} \quad \text{et} \quad R_b(z) = w_b \circ R \circ w_b^{-1}(z) = \lambda z^2 \frac{1-z}{z+ab}.$$

Donc $\widetilde{R}_a(z) = \lambda^{-1}(z+z^2)$ et $\widetilde{R}_b(z) = \lambda(z-z^2)$ sont de degré 2. Par la Proposition 5.2, C est une composante analytique de $\mathcal{E}(R)$.

(2) Notons que $n_R(\mathcal{P}_a) = n_R(\mathcal{P}_b) = 1$ si et seulement si $|\lambda - 1| = 1$. Donc par la Proposition 5.10, R a un point fixe dans C si et seulement si $|\lambda - 1| < 1$.

(2.1) On suppose que $|\lambda - 1| < 1$ de telle façon que $\widetilde{R}_a(z) = z + z^2$ et $\widetilde{R}_b(z) = z - z^2 = -\widetilde{R}_a(-z)$. Donc $n_{R^k}(\mathcal{P}_a) = n_{R^k}(\mathcal{P}_b)$, pour $k \geqslant 1$. D'après l'exemple 3.19 on a $n_{R^{p^m}}(\mathcal{P}_a) = 1 + (1 + p + \cdots + p^m)$. Par la Proposition 5.10 la fonction rationnelle R^{p^m} a

$$2 + (n_{R^{p^m}}(\mathcal{P}_a) - 2) + (n_{R^{p^m}}(\mathcal{P}_b) - 2) = 2(1 + \cdots + p^m)$$

points fixes dans C, comptés avec multiplicité. Par conséquent, si R n'a pas de cycles paraboliques dans C, alors R a exactement 2 cycles de période primitive p^m. Par le Corollaire 5.14 ce sont tous les cycles de R contenus dans C.

Si $p = 2$ on obtient (par l'exemple 3.20 et par un raisonnement similaire) que si R n'a pas de cycles paraboliques dans C, alors R a exactement $2^{-m}(2^{2^m} - 2^{2^{m-1}})$ cycles de période primitive 2^m. Par le Corollaire 5.14 ce sont tous les cycles de R contenus dans C.

(3) On va montrer que si $\lambda = 1$ et $a = b = p$ alors R n'a pas de cycles paraboliques. Un théorème du à Fatou dit que tout cycle parabolique attire, dans le sens complexe, au moins un point critique ; voir [**Fa**]. Donc il suffit de vérifier que tous les points critiques de R sont attirés (dans le sens complexe) par les points fixes super-attractifs 0 et ∞.

Les points 0 et ∞ sont des points critiques. Notons que p est un pôle de R et pour tout autre réel $x \in \mathbb{R}$, $R(x)$ est réel. De plus R est négative dans l'intervalle $(p, +\infty)$ et par conséquent R a un point critique c dans $(p, +\infty)$. Comme $R(z) = -1/R(1/z)$, $-1/c$ est aussi un point critique de R. Alors 0, ∞, c et $-1/c$ sont tous les points critiques de R.

Donc il suffit de vérifier que $R^n(c) \to \infty$ quand $n \to \infty$. Notons que -1 est un point fixe répulsif de R et pour tout $x \in (-\infty, -1)$ on a $R(x) < x$. Par conséquent $R^n(x) \to \infty$ quand $n \to \infty$ car R n'a pas d'autre point fixe dans $(-\infty, -1)$. Finalement il n'est pas difficile de voir que pour tout $x \in (p, +\infty)$ on a $R(x) < -1$, donc $R(c) < -1$ et par conséquent $R^n(c) \to \infty$ quand $n \to \infty$.

Démonstration de la Proposition 5.10. — On va appliquer le Théorème des Résidus à $R - \mathrm{id}$ sur $X = C$; voir Section 2.1.1. Considérons une coordonnée telle que $\infty \in C$ et telle que $R(\infty) \neq \infty$. Alors $R - \mathrm{id}$ a deux pôles simples dans C : ∞ qui est un pôle simple de id et la préimage w de ∞ par R dans C, qui est un pôle simple car R est injective sur $C \subset \mathcal{E}(R)$.

Le nombre des zéros de $R - \mathrm{id}$ dans C est évidement égal au nombre F de points fixes de R dans C (comptés avec multiplicité). Par conséquent,

$$\sum_C \mathrm{ord}_{R-\mathrm{id}}(z) = F + \mathrm{ord}_{R-\mathrm{id}}(\infty) + \mathrm{ord}_{R-\mathrm{id}}(w) = F - 2.$$

Soit $\mathcal{P}$ un bout de C. Notons que pour toute coordonnée affine w on a $\mathrm{ord}_{w\circ R\circ w^{-1}-\mathrm{id}}(\mathcal{P}) = \mathrm{ord}_{R-\mathrm{id}}(\mathcal{P})$. Donc après changement de coordonnée affine on peut supposer $D_{\mathcal{P}} = \{|z| > 1\} \cup \{\infty\}$.

Si $\mathcal{P}$ n'est pas fixé par R on a par le Lemme 2.3 $\mathcal{P}' = (R - \mathrm{id})(\mathcal{P}) = R(\mathcal{P})$, qui est un bout de C. Par conséquent $B_{\mathcal{P}} \subset \mathbb{C}_p - \{0\}$. Donc $\mathrm{ord}_{R-\mathrm{id}}(\mathcal{P}) = 0$ dans ce cas.

Supposons que $\mathcal{P}$ est fixé par R. Soit $\widetilde{R} \in \widetilde{\mathbb{C}}_p(z)$ la réduction de R. Par définition de $n_R(\mathcal{P}) \geqslant 1$ on peut écrire

$$\frac{1}{\widetilde{R}(1/z)} = z + az^{n_R(\mathcal{P})}\frac{P(z)}{Q(z)},$$

où $a \in \widetilde{\mathbb{C}}_p - \{0\}$ et les polynômes $P, Q \in \widetilde{\mathbb{C}}_p[z]$ sont tels que $P(0) \neq 0$ et $Q(0) \neq 0$. Par conséquent

$$\widetilde{R}(z) - z = \frac{1}{z^{n_R(\mathcal{P})-2}}\frac{P(1/z)}{aQ(1/z) + z^{1-n_R(\mathcal{P})}P(1/z)}.$$

Supposons $n_R(\mathcal{P}) = 1$. Comme $\deg_{\widetilde{R}}(\infty) = \deg_R(\mathcal{P}) = 1$ on a $1 + aP(0)/Q(0) \neq 0$. Donc $\deg_{R-\mathrm{id}}(\mathcal{P}) = \deg_{\widetilde{R}-\mathrm{id}}(\infty) = 1$ et $(\widetilde{R} - \mathrm{id})(\infty) = \infty$. Par conséquent $(R - \mathrm{id})(\mathcal{P}) = \mathcal{P}$ et

$$\mathrm{ord}_{R-\mathrm{id}}(\mathcal{P}) = -\deg_{R-\mathrm{id}}(\mathcal{P}) = -1 = n_R(\mathcal{P}) - 2.$$

Si $n_R(\mathcal{P}) = 2$ alors $\widetilde{R}(\infty) \in \widetilde{\mathbb{C}}_p - \{0\}$. Donc $D_{(R-\mathrm{id})(\mathcal{P})} \subset \mathbb{C}_p - \{0\}$ et par conséquent $\mathrm{ord}_{R-\mathrm{id}}(\mathcal{P}) = 0 = n_R(\mathcal{P}) - 2$.

Si $n_R(\mathcal{P}) > 2$ alors notons que $\deg_{R-\mathrm{id}}(\mathcal{P}) = \deg_{\widetilde{R}-\mathrm{id}}(\infty) = n_R(\mathcal{P}) - 2$ et $\widetilde{R}(\infty) = 0$. Par conséquent $\mathrm{ord}_{R-\mathrm{id}}(\mathcal{P}) = \deg_{R-\mathrm{id}}(\mathcal{P}) = n_R(\mathcal{P}) - 2$.

Par le Théorème des Résidus on a

$$F - 2 = \sum_{\text{bouts fixes de } C} (n_R(\mathcal{P}) - 2). \qquad \square$$

5.3. Dynamique au bouts d'une composante analytique

Proposition 5.16. — *Soit R une fonction rationnelle de degré au moins deux, C une composante analytique de $\mathcal{E}(R)$ et $\mathcal{P}$ un bout de C. Soit $\mathcal{S}$ le système projective associé à $\mathcal{P}$. Alors il existe une infinité de $\mathcal{P}_0 \in \mathcal{S}$ tels que $D_{\mathcal{P}_0}$ est un disque de Siegel.*

Voir figure 4.

Démonstration. — Quitté à remplacer R par un itéré on peut supposer que $\mathcal{P}$, et par conséquent $\mathcal{S}$, est fixé par R. Considérons un paramétrage $\mathcal{S} = \{\mathcal{P}(\xi)\}_{\mathbb{P}(\overline{\mathbb{F}}_p)}$. Soit $\mathcal{T} \subset \mathcal{S}$ comme dans (iii) de la Proposition 2.4 et soit $T \subset \mathbb{P}(\overline{\mathbb{F}}_p)$ tel que $\mathcal{T} = \cup_T \mathcal{P}(\xi)$.

Comme $\mathcal{P}$ est un bout de C on a $\deg_R(\mathcal{P}) = 1$ et donc $\widetilde{R}' \not\equiv 0$. Par conséquent il existe une infinité des $\xi \in \mathbb{P}(\overline{\mathbb{F}}_p))$ périodiques par $\widetilde{R}$ tels que $(\widetilde{R}^k)'(\xi) \neq 0$, où k est la période de ξ. On peut supposer de plus que $R^j(\xi) \notin \mathbb{P}(\overline{\mathbb{F}}_p) - T$, pour $i \leqslant j \leqslant k$. Donc par (iii) de la Proposition 2.4 et le Corollaire 3.12 on a $D_{\mathcal{P}(\xi)} \subset \mathcal{E}(R)$. Par le Théorème 3 on a $\deg(\widetilde{R}) > 1$, donc D_α est un disque de Siegel par la Proposition 5.2. $\square$

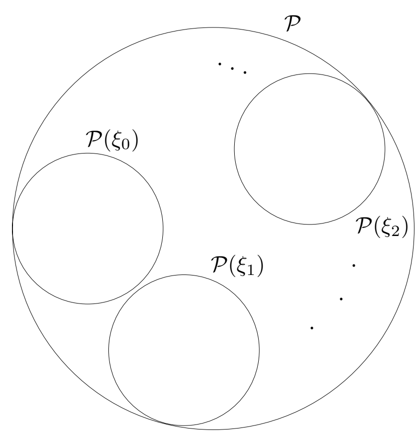

FIGURE 4. Chaque bout $\mathcal{P}$ d'une composante analytique de $\mathcal{E}(R)$ a une infinité de disques de Siegel attachés.

Corollaire 5.17. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle de degré au moins deux. Alors les affirmations suivantes sont équivalentes.*

(i) *R a un point périodique indifférent.*
(ii) *R a une infinité de points périodiques indifférents.*
(iii) *$\mathcal{E}(R) \neq \varnothing$.*
(iv) *$\mathcal{E}(R)$ à une infinité de disques de Siegel.*

Démonstration

(ii) $\Rightarrow$ (i) est trivial ; (i) $\Rightarrow$ (iii) résulte du point 2 de la Proposition 3.16.

(iv) $\Rightarrow$ (ii) suit du Corollaire 5.13 et (iii) $\Rightarrow$ (iv) suit de la Proposition 5.16 et de (i) de la Proposition 5.10. $\square$

6. Exemples

Dans cette section on considère des exemples. Dans la Section 6.1 on montre que pour tout affinoïde ouvert il existe une fonction rationnelle ayant chaque composante de l'affinoïde comme composante analytique de son domaine de quasi-périodicité (Proposition 6.4) ; voir aussi Proposition 6.7.

Exemple 6.1. — On va montrer que le bassin d'attraction de l'infini du polynôme $R(z) = p^{-1}(z - z^p)$ est égal à $\mathbb{P}(\mathbb{C}_p) - \mathcal{O}_{\mathbb{Q}_p}$. Dans la Section 1.2.3 il y a une description de l'arbre de cet espace analytique.

Notons que pour $|w| > 1$ on a $|R(w)| \geqslant p|w|^p$, donc $\{|z| > 1\} \subset \mathcal{W}_r^s(\infty)$. De plus pour tout w tel que $|w - i| = 1$ pour tout $i \in \{0, 1, \ldots, p-1\}$, on a $|w - w^p| = 1$. Donc $|R(w)| = p > 1$ et par conséquent $w \in \mathcal{W}_R^s(\infty)$.

D'autre part $\mathcal{O}_{\mathbb{Q}_p}$ est invariant par R donc $\mathcal{O}_{\mathbb{Q}_p} \cap \mathcal{W}_R^s(\infty) = \varnothing$. Pour chaque $i \in \{0, \ldots, p-1\}$ l'application

$$R : D_i = \{|z - i| < 1\} \longrightarrow \{|z| < p\},$$

est de degré 1. Soit $R_i : \{|z| < p\} \to D_i$ la branche inverse correspondante. Pour tout $k \geqslant 0$ fixé on a

$$\mathbb{C}_p - \cup_{i_0,\ldots,i_k \in \{0,\ldots,p-1\}} R_{i_0} \circ \cdots \circ R_{i_k}(\{|z| < p\}) \subset \mathcal{W}_R^s(\infty).$$

Alors l'égalité $\mathcal{W}_R^s(\infty) = \mathbb{P}(\mathbb{C}_p) - \mathcal{O}_{\mathbb{Q}_p}$ suit du fait que

$$R_{i_0} \circ \cdots \circ R_{i_k}(\{|z| < p\}) = \{z \mid |z - (i_0 + \cdots + i_k p^k)| < p^{-k}\}.$$

Exemple 6.2. — Considérons la fonction rationnelle $R(z) = \lambda(\frac{z}{z-1} + z^p) \in \mathbb{C}_p(z)$, où $\lambda \in \mathbb{C}_p$ est tel que $p^{-(p-1)/(p+1)} < |\lambda| < 1$. On montrera que le bassin d'attraction du point fixe attractif 0 est de type Cantor et ne contient pas des points critiques.

Notons que pour chaque $z \in \mathbb{C}_p$ tel que $|z| < 1$ on a $|R(z)| = |\lambda z|$. De plus $R(1) = R(\infty) = \infty$ et par conséquent $D = \{|z| < 1\}$ est le plus grand disque contenu dans le bassin d'attraction de 0. Comme $R(D) \neq D$, le bassin d'attraction immédiat de 0 est non-trivial.

D'autre part, si $z \in \mathbb{C}_p$ est tel que $|z| > 1$ alors $|R(z)| = |\lambda||z|^p$; donc, si de plus $|z| > |\lambda|^{-1/(p-1)}$, on a que $|R^n(z)| \to \infty$ quand $n \to \infty$. Comme $R'(z) = \lambda(-\frac{1}{(z-1)^2} + pz^{p-1})$, les points critiques finis de R sont de norme $p^{1/(p+1)}$. Donc la condition $|\lambda| > p^{-(p-1)/(p+1)}$ impliques que tous les points critiques de R tend vers l'infini par itération, et par conséquent le bassin d'attraction de 0 ne contient pas des points critiques.

Exemple 6.3. — Considérons le polynôme $P(z) = \frac{1}{p}(z^p - z^{p^2})$. Notons que $0, 1 \in K(P) \subset \{|z| \leqslant 1\}$ et pour $|w| \leqslant 1$ on a,

$$P(z+w) - P(w) = \frac{1}{p}(z^p - z^{p^2}) + Q_w(z),$$

où $Q_w \in \mathbb{C}_p[z]$ est entier et $Q_w(0) = 0$. Par conséquent $|P(z+w) - P(w)| = p|z|^p$ pour tout $p^{-1/(p-1)} < |z| < 1$. En particulier pour toute boule fermée $B \subset \{|z| \leqslant 1\}$ de diamètre au moins $p^{-1/(p-1)}$, $P^{-1}(B)$ est une réunion de p boules $B_0, \ldots, B_{p-1}$ disjointes deux à deux avec

$$\mathrm{diam}(B_i) = \left(\frac{1}{p}\mathrm{diam}(B)\right)^{1/p} > p^{-1/(p-1)},$$

et $P : B_i \to B$ est de degré p. On pose $P^{-1}(\{|z| \leqslant 1\}) = B_0 \sqcup \cdots \sqcup B_{p-1}$ où $\mathrm{diam}(B_i) = p^{-1/p}$. En général,

$$P^{-n}(\{|z| \leqslant 1\}) = \sqcup_{a_i \in \{0,1,\ldots,p-1\}} B_{a_0 a_1 \ldots a_{n-1}},$$

où les boules $B_{a_0 \ldots a_{n-1}}$ sont de diamètre

$$p^{-(1/p + 1/p^2 + \cdots + 1/p^n)} = p^{-(1-p^{-n})/(p-1)} \longrightarrow p^{-1/(p-1)} \text{ quand } n \to \infty,$$

et les indices sont tels que $B_{a_0 \ldots a_{n-1} a_n} \subset B_{a_0 \ldots a_{n-1}}$ et $P(B_{a_0 a_1 \ldots a_n}) = B_{a_1 \ldots a_n}$. Considère l'espace de codage Σ_p avec alphabet $\{0, \ldots, p-1\}$,

$$\Sigma_p = \{.a_0 a_1 \ldots \mid a_i \in \{0, \ldots, p-1\}\},$$

et à chaque point $w \in K(P)$ on associe la suite $\pi(w) = .a_0 a_1 \ldots$ de telle façon que $P^n(w) \in B_{a_n}$. Par conséquent l'application π respecte la dynamique de P sur $K(P)$ et la dynamique du décalage $\sigma : \Sigma_p \to \Sigma_p$ défini par $\sigma(.a_0 a_1 \ldots) = .a_1 \ldots$. Alors notons que $w \in \cap_{n \geqslant 0} B_{a_0 a_1 \ldots a_n}$ et par conséquent w appartient à une composante errante si et seulement si $\pi(w)$ n'est pas prépériodique pour le décalage σ.

Jusque à maintenant on est dans une situation standard en dynamique. Mais notons que à chaque $\underline{a} = .a_0 a_1 \ldots \in \Sigma_p$ on peut associer l'intersection $B_{\underline{a}} = \cap_{n \geqslant 0} B_{a_0 \ldots a_n}$ et comme Σ_p n'est pas dénombrable il y a beaucoup d'*intersections vides*, car $\mathbb{C}_p$ est séparable. (Ceci montre que $\mathbb{C}_p$ n'est pas maximalement complet, c'est-à-dire que il existe des suites emboîtées de boules B_n avec intersection vide ; voir e.g. [**Ro**]).

Si l'intersection $B_{\underline{a}}$ est non-vide alors c'est une boule de diamètre $p^{-1/(p-1)}$. Par le Lemme 4.29 si D est un disque errant alors

$$\liminf_{n \to \infty} \mathrm{diam}(P^n(D)) = 0.$$

Par conséquent tous les intersections $B_{\underline{a}}$, avec $\underline{a} \in \Sigma_p$ qui n'est pas prépériodique par σ, sont vides. De plus par le Lemme 4.13 pour toute suite périodique $\underline{a} \in \Sigma_p$ il existe un point périodique w de P, de même période, tel que $\pi(w) = \underline{a}$. Par conséquent $\pi(K(P)) \subset \Sigma_p$ est égal à l'ensemble des suites prépériodiques pour le décalage σ.

6.1. Exemples de composantes du domaine de quasi-périodicité

Proposition 6.4. — *Soit X un affinoïde ouvert. Alors il existe une fonction rationnelle R telle que chaque composante connexe de X est une composante analytique du domaine de quasi-périodicité de R.*

La démonstration de cette proposition se base sur le lemme suivant.

Lemme 6.5. — *Soit X un affinoïde ouvert et $\mathcal{S}$ un système projectif tel que pour toute composante analytique Y de X il existe $\mathcal{P} \in \mathcal{S}$ tel que $Y \subset D_{\mathcal{P}}$. Alors il existe une fonction rationnelle R telle que $\deg_R(\mathcal{S}) > 1$, $R : X \to X$ est de degré 1 et telle que R fixe les bouts de X.*

Démonstration. — Après changement de coordonnée on peut supposer que $\mathcal{S}$ est le système projectif canonique et que $X \subset \{|z| \leqslant 1\}$. Alors il existe un ensemble fini $A \subset \{|z| \leqslant 1\}$ tel que $X \subset \cup_A D_a$, ou $D_a = \{|z - a| < 1\}$. Le polynôme $Q(z) = z + \prod_A (z-a)^2$ est tel que $Q : D_a \to D_a$ est de degré 1 pour tout $a \in A$. Donc $D_a \subset \mathcal{E}(Q)$ par le Corollaire 3.12 et les bouts $\mathcal{P} \prec D_a$ de X sont périodiques par Q, par le Corollaire 4.20. Donc il existe n tel que $R = Q^n$ fixe les bouts de X. Par conséquent $R : X \to X$ est de degré 1 et $\deg_R(\mathcal{S}) = \deg(R) = n \cdot \deg(Q) \geqslant 2$. □

Démonstration de la Proposition 6.4. — Soient $\mathcal{S}_0, \dots, \mathcal{S}_k$ les systèmes projectifs associés aux bouts de X et soit $R_{\mathcal{S}_i}$ comme dans le lemme précédent. Alors $R = R_{\mathcal{S}_0} \circ \dots \circ R_{\mathcal{S}_k} : X \to X$ est de degré 1 et $\deg_R(\mathcal{S}_i) > 1$. Par la Proposition 5.2 chaque composante connexe de X est une composante analytique de $\mathcal{E}(R)$. □

Exemple 6.6. — On va montrer que la fonction rationnelle,

$$R(w) = \frac{1 + w^{p+1}}{w^p (1 + (pw)^{p+1})} \in \mathbb{C}_p(w),$$

a $C = \{1 < |z| < p\}$ comme composante analytique de $\mathcal{E}(R)$ et que l'ensemble de Julia de R est vide.

Notons que la coordonnée $1/pw$ est une symétrie de R, c'est-à-dire $R(w) = 1/pR(1/pw)$. Soit $\mathcal{S}_0$ le système projectif canonique et $\mathcal{P}_0 \in \mathcal{S}_0$ le bout associé à $\{|w| \leqslant 1\}$, qui est un bout de C. On pose $D_0 = \{|w| < 1\}$ et $W_0 = \{|z| = 1\}$. Notons que la réduction de R est $(1 + w^{p+1})/w^p$ et pour tout $\mathcal{P} \in \mathcal{S}_0 - \{\mathcal{P}_0\}$, $R : D_{\mathcal{P}} \to D_{R(\mathcal{P})}$ est de degré $\deg_R(\mathcal{P})$. Par conséquent $R(W_0) = W_0 \sqcup D_0$ et R n'augmente pas la distance chordale $d_{\mathcal{S}_0}$ dans $W_0 \sqcup D_0$.

On pose $\mathcal{S}_1 = (\mathcal{S}_0)_{1/pw}$, $D_1 = (D_0)_{1/pw}, \dots$ de telle façon que $C = D_{\mathcal{P}_0} \cap D_{\mathcal{P}_1}$ et on a les propriétés analogues ; en particulier R n'augmente pas la distance chordale $d_{\mathcal{S}_1}$ (correspondante à la coordonnée $1/pw$) dans $W_1 \sqcup D_1 = \{|w| \geqslant p\} \cup \{\infty\}$; voir figure 5. De plus la réduction de R dans la coordonnée $1/pw$ est aussi $(1 + w^{p+1})/w^p$. Il est facile de voir que $R : C \to C$ est de degré 1, et comme $\deg_R(\mathcal{S}_0) = \deg_R(\mathcal{S}_1) = p + 1 > 1$, C est une composante analytique de $\mathcal{E}(R)$ (Proposition 5.2).

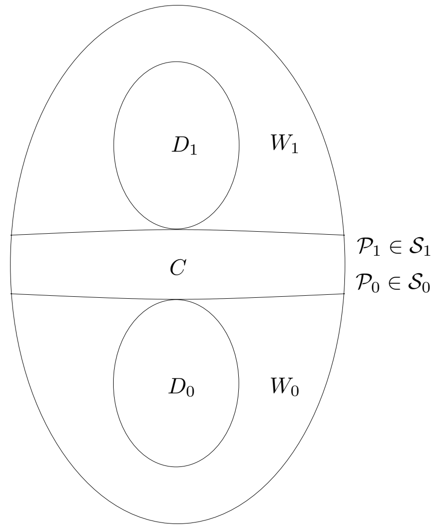

FIGURE 5. Partition de $\mathbb{P}(\mathbb{C}_p)$.

(1) On va montrer que, si $z, w \in D_0$ sont tels que

$$d_{\mathcal{S}_0}(z,w) = |z-w| \leqslant p^{-1/(p-1)} \text{ et } R(w) \in \{|z| \geqslant p\} \cup \{\infty\},$$

alors $d_{\mathcal{S}_1}(R(z), R(w)) \leqslant p^{-1/(p-1)}$. Ceci est équivalent à montrer que si $|1/pR(w)| \leqslant 1$ alors $|1/pR(z) - 1/pR(w)| \leqslant p^{-1/(p-1)}$. On a

$$\begin{aligned}\frac{1}{pR(z)} - \frac{1}{pR(w)} &= \frac{z^p(1+(pz)^{p+1})}{p(1+z^{p+1})} - \frac{w^p(1+(pw)^{p+1})}{p(1+w^{p+1})} \\ &= \frac{z^p - w^p - (zw)^p(z-w) + (pz)^{p+1}(1+w^{p+1}) - (pw)^{p+1}(1+z^{p+1})}{p(1-z^{p+1})(1+w^{p+1})}.\end{aligned}$$

Notons que $|1+z^{p+1}| = |1+w^{p+1}| = 1$ et

$$|(pz)^{p+1}(1+w^{p+1}) - (pw)^{p+1}(1+z^{p+1})| \leqslant p^{-(p+1)} < \frac{1}{p}\, p^{-1/(p-1)}.$$

D'autre part, si $\varepsilon = z - w$, alors,

$$|z^p - w^p| = |p\varepsilon w^{p-1} + \cdots + p\varepsilon^{p-1}w + \varepsilon^p| \leqslant \frac{1}{p}\, p^{-1/(p-1)},$$

car $|\varepsilon| \leqslant p^{1/(p-1)}$ et $|\varepsilon^p| \leqslant \frac{1}{p}\, p^{-1/(p-1)}$. De plus, comme

$$1 \geqslant \Big|\frac{1}{pR(w)}\Big| = \Big|\frac{w^p(1+(pw)^{p+1})}{p(1+w^{p+1})}\Big| = p|w|^p,$$

on a $p|z|^p \leqslant 1$ et donc $|(zw)^p(z-w)| \leqslant \frac{1}{p^2}\, p^{-1/(p-1)}$. Par conséquent

$$\Big|\frac{1}{pR(z)} - \frac{1}{pR(w)}\Big| \leqslant p\Big(\frac{1}{p}\, p^{-1/(p-1)}\Big) = p^{-1/(p-1)}.$$

(2) Montrons maintenant que $F(R) = \mathbb{P}(\mathbb{C}_p)$. Notons que $C \subset \mathcal{E}(R) \subset F(R)$. Fixons $z_0 \in \mathbb{P}(\mathbb{C}_p) - C$; par symétrie on peut supposer $|z_0| \leqslant 1$. Comme $R^{-1}(F(R)) = F(R)$, s'il existe n tel que $R^n(z_0) \in C$ alors $z_0 \in F(R)$; donc on peut supposer que $R^n(z_0) \notin C$, pour $n \geqslant 0$.

Par la partie 1, pour tout $w_0 \in B = \{|w - z_0| < p^{-1/(p-1)}\}$ et tout $n \geqslant 1$, on a $d_{\mathcal{S}_i}(R^n(z_0), R^n(w_0))) \leqslant p^{-1/(p-1)}$, où $i = 0$ si $R^n(z_0) \in \{|z| \leqslant 1\}$ et $i = 1$ sinon. Par conséquent la famille $\{R^n\}_{n \geqslant 0}$ est uniformément lipschitzienne sur B, et donc $z_0 \in B \subset F(R)$.

La Proposition 6.4 ne nous donne aucun information sur le degré de la fonction rationnelle. Avec la proposition suivante, dans un cadre plus restrictif, on aura un contrôle sur le degré.

L'idée est de modifier une fonction rationnelle $R \in \mathbb{C}_p(z)$ ayant une bonne réduction ; voir Section 4.5. Par la Proposition 4.32, si $\deg(R) > 1$, les composantes analytiques de $\mathcal{E}(R)$ sont des disques de la forme $\{\widetilde{z} = \xi\}$. Si $\deg(R) = 1$ alors $\mathcal{E}(R) = \mathbb{P}(\mathbb{C}_p)$; voir exemples dans la Section 3.2.

Proposition 6.7. — *Soit $R \in \mathbb{C}_p(z)$ une fonction rationnelle ayant une bonne réduction. Soit $r \in |\mathbb{C}_p|$ tel que $r < 1$ et soient des $\mathcal{P}_1, \ldots, \mathcal{P}_n$ bouts tels que $B_{\mathcal{P}_i} \subset \mathcal{E}(R)$ et* $\operatorname{diam}(B_{\mathcal{P}_i}) = r$. *Alors il existe une fonction rationnelle Q de degré* $\deg(Q) = \deg(R) + n$ *telle que pour toute composante analytique C de $\mathcal{E}(R)$, l'affinoïde ouvert connexe*

$$C - \cup_{\mathcal{T}} B_{\mathcal{P}}, \ \text{où } \mathcal{T} = \{R^j(\mathcal{P}_i) \mid \text{pour } 1 \leqslant i \leqslant n \text{ et } j \geqslant 1\},$$

est une composante analytique de $\mathcal{E}(Q)$.

Notons que par le Corollaire 5.3, l'ensemble $\mathcal{T}$ de la proposition est fini. La démonstration de cette proposition est à la fin de cette section.

Exemple 6.8. — Considérons des bouts $\mathcal{P}_1, \ldots, \mathcal{P}_n$ tels que les boules $B_{\mathcal{P}_1}, \ldots, B_{\mathcal{P}_n}$ sont disjointes deux à deux et de même diamètre. Alors par la proposition (en prenant $R = \mathrm{id}$) il existe une fonction rationnelle Q de degré $n+1$ telle que

$$\mathbb{P}(\mathbb{C}_p) - B_1 \cup \cdots \cup B_n,$$

est une composante analytique de $\mathcal{E}(Q)$.

Exemple 6.9. — Soit $n \geqslant 1$ et ζ racine primitive $n^{\text{ième}}$ de l'unité et soit $r \in |\mathbb{C}_p|$ tel que 1 est la seul racine $n^{\text{ième}}$ de l'unité dans la boule $\{|z - 1| \leqslant r\}$. Soit $\mathcal{P}$ le bout associé à cette boule. Posons $R(z) = \zeta z$ et notons que $R^n(B_{\mathcal{P}}) = B_{\mathcal{P}}$ et que les boules $B_{\mathcal{P}}, \ldots, R^{n-1}(B_{\mathcal{P}})$ sont disjointes deux à deux. Par la proposition il existe une fonction rationnelle Q de degré 2 telle que

$$\mathbb{P}(\mathbb{C}_p) - B_{\mathcal{P}} \cup \cdots \cup R^{n-1}(B_{\mathcal{P}}),$$

est une composante analytique de $\mathcal{E}(Q)$ (un $(n-1)$-anneau de Herman).

Démonstration de la Proposition 6.7. — Soit $\mathcal{S}$ le système projectif canonique et $\mathcal{I} \subset \mathcal{S}$ l'ensemble des bouts $\mathcal{P} \in \mathcal{S}$ tels que $\deg_R(\mathcal{P}) = 1$. Comme $\mathcal{E}(R) \neq \varnothing$ l'ensemble $\mathcal{I}$ est infini et on peut trouver $\mathcal{Q} \in \mathcal{I}$ tel que $D_{\mathcal{Q}} \cap (\cup_{\mathcal{T}} B_{\mathcal{P}}) = \varnothing$. Après changement de coordonnée on suppose que $\infty \in D_{\mathcal{Q}}$, de telle façon que $D_{\mathcal{Q}} = \{|z| > 1\} \cup \{\infty\}$. Soit $\mathcal{S}_i$ le système projectif associé à $\mathcal{P}_i$, pour $1 \leqslant i \leqslant n$.

1.– On va définir par induction des fonctions rationnelles $R_i \in \mathbb{C}_p(z)$ qui satisfont les propriétés suivantes.

(1) $\deg(R_i) = \deg(R) + i$, pour $1 \leqslant i \leqslant n$.
(2) $\deg_{R_i}(\mathcal{S}_j) > 1$, pour $1 \leqslant j \leqslant i$.
(3) R_i est injective sur $D_{\mathcal{P}} - (B_{\mathcal{P}_1} \cup \cdots \cup B_{\mathcal{P}_i})$, pour tout $\mathcal{P} \in \mathcal{I}$.
(4) Pour tout bout $\mathcal{P} \notin \{\mathcal{P}_1, \ldots, \mathcal{P}_i\}$ tel que $B_{\mathcal{P}} \subset \{|z| \leqslant 1\}$ et $\operatorname{diam}(B_{\mathcal{P}}) = r$ on a $R_i(\mathcal{P}) = R(\mathcal{P})$ et $R_i : B_{\mathcal{P}} \to B_{R(\mathcal{P})}$ est de degré 1.

Notons que 4 implique 3. Posons $R_0 = R$ et supposons que R_{i-1} est déjà définie. Après changement affine de coordonnée au départ et à l'arrivée, on suppose que $0 \in B_{\mathcal{P}_i}$ et $R_{i-1}(0) = 0$ de telle façon que $R_{i-1}(B_{\mathcal{P}_i}) = \{|z| \leqslant r\} = B_{\mathcal{P}_i}$.

On pose

$$R_i(z) = \frac{z}{z-\mu} R_{i-1}(z) = R_{i-1}(z) + \frac{\mu}{z-\mu} R_{i-1}(z),$$

où $\mu \in \mathbb{C}_p$ est de norme égale à r. Clairement $\deg(R_i) = \deg(R_{i-1}) + 1 = \deg(R) + i$, donc R vérifie 1. De plus $\deg_{R_i}(\mathcal{S}_i) > 1$.

Comme R_{i-1} vérifie 4 on a

$$|R_{i-1}(z_0)| = |z_0| \text{ pour tout } z_0 \in \{|z| \leqslant 1\} - B_{\mathcal{P}_1} \cup \cdots \cup B_{\mathcal{P}_{i-1}}.$$

Par conséquent, si de plus $z_0 \notin B_{\mathcal{P}_i}$, on a $|R_i(z) - R_{i-1}(z)| \leqslant r$. Donc si $\mathcal{P}$ est comme dans 4 on a $R_i(\mathcal{P}) = R_{i-1}(\mathcal{P}) = R(\mathcal{P})$, $\deg_{R_i}(\mathcal{P}) = 1$ (Lemme 2.3) et $R_i : B_{\mathcal{P}} \to B_{R(\mathcal{P})}$ est de degré 1. C'est à dire R_i vérifie 4, et par conséquent 3.

Fixons $1 \leqslant j < i$ et considérons $\zeta_0 \in B_{\mathcal{P}_j}$ et $\zeta_1 \in B_{R(\mathcal{P}_j)}$. On a $|\zeta_0| = |\zeta_1| > r$. On pose $Q_0(z) = R_{i-1}(z + \zeta_0) - \zeta_1$ et $Q_1(z) = R_i(z + \zeta_0) - \zeta_1$. Comme R_{i-1} satisfait la propriété 4, on a $Q_0(\mathcal{P}) = \mathcal{P}$ et $\deg_{Q_0}(\mathcal{S}) > 1$, où $\mathcal{P}$ est le bout de $\{|z| \leqslant r\}$ et $\mathcal{S}$ est le système projectif contenant $\mathcal{P}$. Par définition de R_i et Q_1 on a

$$Q_1(z) = \frac{z+\zeta_0}{z+\zeta_0-\mu} R_{i-1}(z+\zeta_0) - \zeta_1 = \frac{z+\zeta_0}{z+\zeta_0-\mu} Q_0(z) + \zeta_1 \frac{\mu}{z+\zeta_0-\mu}.$$

Comme $|\zeta_0| = |\zeta_1| > r = |\mu|$ on conclut que $Q_1(\mathcal{P}) = Q_0(\mathcal{P}) = \mathcal{P}$ et $\deg_{Q_1}(\mathcal{S}) = \deg_{Q_0}(\mathcal{S}) > 1$ (Lemme 2.3). Donc $R_i(\mathcal{P}_j) = R(\mathcal{P}_j)$ et $\deg_{R_i}(\mathcal{S}_j) > 1$.

2.– Posons $Q = R_n$ et considérons un cycle $C_1, \ldots, C_k$ de composantes analytiques de $\mathcal{E}(R)$ de telle façon que $R(C_i) = C_{i+1}$, pour $1 \leqslant i \leqslant k$, où $C_{k+1} = C_0$. Par les propriétés 2 et 3, Q^k vérifie les hypothèses de la Proposition 5.2 avec $X = C_i - \cup_{\mathcal{T}} B_{\mathcal{P}}$ pour chaque $1 \leqslant i \leqslant k$. Par conséquent les affinoïdes ouverts $C_i - \cup_{\mathcal{T}} B_{\mathcal{P}}$, pour $1 \leqslant i \leqslant k$, forment un cycle de composantes analytiques de $\mathcal{E}(Q)$. $\square$

Références

[AV] D. Arrowsmith, F. Vivaldi. *Geometry of p-adic Siegel discs*. Phys. D **71** #1-2 (1994), 222–236.

[Be] R. Benedetto. *Fatou components in p-adic dynamics*. Thesis, Brown University 1998.

[CG] L. Carleson, T. Gamelin. *Complex dynamics*. Springer-Verlag 1991.

[Ca] J.W.S. Cassels. *Local fields*. LMS student texts **3**. Cambridge Univ. Press (1986).

[DH] A. Douady, J. Hubbard. *On the dynamics of polynomial-like mappings*. Ann. Scient. Éc. Norm. Sup **18** (1985), 287–343.

[Ec1] J. Ecalle. *Théorie des invariants holomorphes*. Thèse d'état, Orsay, Mars 1974. Voir aussi *Théorie itérative : introduction à la théorie des invariants holomorphes*. J. math. pures et appl. **54** (1975), 183 –258.

[Ep] A. Epstein. *Infinitesimal Thurston rigidity and the Fatou-Shishikura inequality*. IMS preprint #1999 / 1.

[Es] A. Escassut. *Algèbres de Banach ultramétriques et algèbres de Krasner-Tate*. Astérisque **10** (1973).

[Fa] P. Fatou. *Sur les équations fonctionelles*. Bull. Soc. Math. France **47** (1919), 161–217.

[FvP] J. Fresnel, M. van der Put. *Géométrie Analytique rigide et applications*. PM 18, Birkhäuser 1981.

[GvP] L. Gerritzen, M. van der Put. *Schottky groups and Mumford curves*. LMN 817, Springer-Verlag Berlin-Heidelberg-New York, 1980.

[GH] Sur les groupes hyperboliques d'après Mikhael Gromov. Edité par E. Ghys et P. de la Harpe. Progress in Mathematics, 83. Birkhauser Boston, Inc., Boston, MA, 1990.

[Gu] J. Gukenheimer. *Sensitive dependence on initial conditions for one dimensional maps*. Comm. Math. Phys. **70** (1979), 133–160.

[HY] M. Herman, J.-C. Yoccoz. *Generalizations of some therorems of small divisors to non-Archimedean fields*. Geometric Dynamics (Rio de Janeiro, 1981), LMN 1007, Springer-Verlag Berlin-New York, (1983), 408–447.

[Hs] L.C. Hsia. *Closure of periodic points over a non-Archimidean field*. J. London Math. Soc. **62** (2000), 685–700.

[Ke] K. Keating. *Automorphisms and extensions of $k((t))$*. J. Number Th. **41** (1992), 314–321.

[Ku] M. Kuczma. *Functional equations in a simple variable*. Monografie Mat. **46** (1968).

[La] P. Lavaurs. *Systèmes dynamiques holomorphes : explosion de points périodiques paraboliques*. Thèse, Université de Paris-Sud, 1989.

[Le] J. Lewis. *Invariant sets of morphisms on projective and affine number spaces*. J. Algebra **20** (1972) 419–434.

[Li1] H.-C. Li. *p-adic periodic points and Sen's theorem*. J. Number Th. **56** (1996), 309–318.

[Li2] H.-C. Li. *p-adic power series which commute under composition*. Tran. of A.M.S. **349** (1997), 1437–1446.

[Li3] H.-C. Li. *Counting periodic points of p-adic power series*. Compositio Math. **100** (1996), 351–364.

[Lu1] J. Lubin. *Nonarchimedean dynamical systems*. Compositio Math. **94** (1994), 321–346.

[Lu2] J. Lubin. *Another proof of Sen's theorem on iteration of power series*. Proc. Amer. Math. Soc. **123** (1995), 63–66.

[Ma] R. Mañé. *On the instability of Herman rings*. Invent. Math. **81** (1985), 459–471.

[Mi] J. Milnor. *On bicritical rational maps*. IMS preprint #1997 / 10.

[MS1] P. Morton, J. Silverman. *Rational periodic points of rational fonctions*. Inter. Math. Res. Notices. **1** (1994), 97–110.

[MS2] P. Morton, J. Silverman. *Periodic points, multiplicities, and dynamical units*. J. Reine Agnew. Math. **461** (1995), 81–122.

[Mo] E. Motzkin. *L'arbre d'un quasi connexe : un invariant conforme p-adique*. Groupe d'étude d'analyse ultramétrique, 9ème année, **3** (1981/82), 18p.

[No] D. Northcott. *Periodic points of an algebraic variety*. Ann. Math. **51** (1950), 167–177.

[R-L] J. Rivera-Letelier. *Dynamique des fonctions rationnelles sur des corps locaux.* Thèse, Orsay 2000.

[Ro] P. Robba. *Fonctions analytiques sur les corps valués ultramétriques complets*. Astérisque **10** (1973).

[Ru1] R. Rumely. *Capacity theory on algebraic curves and canonical heights*. Groupe d'étude d'analyse ultamétrique, 12ème année, **22** (1984/85), 17p.

[Ru2] R. Rumely. *Capacity theory on algebraic curves*. LNM 1378, Springer-Verlag, Berlin-New York, 1989.

[Se] S. Sen. *On automorphisms of local fields*. Ann. of Math. **90** (1969), 33–46.

[Sh] M. Shishikura. *On the Quasiconformal surgery of rational functions*. Ann. Scient. Éc. Norm. Sup. **20** (1987), 1–29.

[Sul] D. Sullivan. *Quasiconformal homeomorphisms and dynamics I : Solution of the Fatou-Julia problem on wandering domains*. Ann. of Math. **122** (1985), 401–418.

[TVW] E. Thiran, D. Verstegen, J. Weyers. *p-adic dynamics*. J. Stat. Phys. **54#3/4** (1989), 839–913.

[Vo] S.M. Voronin. *Classification analytique des germes d'applications conformes* $(\mathbb{C}, 0) \to (\mathbb{C}, 0)$ *tangents à l'identité* (en russe), Functional Analysis **15/1** (1981), 1–17.

[Y] J.C. Yoccoz. *Petits diviseurs en dimension* 1. Astérisque **231** (1995).

J. Rivera-Letelier, Depto de Matemáticas, Univ. Católica del Norte, Casilla 1280, Antofagasta, Chile • *E-mail :* `rivera-letelier@ucn.cl`

Astérisque
287, 2003, p. 231–249

ON THE DIVERGENCE OF GEODESIC RAYS IN MANIFOLDS WITHOUT CONJUGATE POINTS, DYNAMICS OF THE GEODESIC FLOW AND GLOBAL GEOMETRY

by

Rafael Oswaldo Ruggiero

Dedicated to J. Palis, on his 60th. birthday

Abstract. — Let (M, g) be a compact Riemannian manifold without conjugate points. Suppose that the horospheres in $(\widetilde{M}, g)$ depend continuously on their normal directions. Then we show that geodesic rays diverge uniformly in the universal covering $(\widetilde{M}, g)$. We give some applications of this result to the study of the dynamics of the geodesic flow and the global geometry of manifolds without conjugate points.

Introduction

The problem of the divergence of geodesic rays in manifolds without conjugate points is one of the most natural, yet unsolved, questions of the theory. Recall that a C^∞ Riemannian, n-dimensional manifold (M, g) has no conjugate points if the exponential map is nonsingular at every point. The universal covering $\widetilde{M}$ of a manifold (M, g) is diffeomorphic to R^n, and the metric spheres in $(\widetilde{M}, g)$ — the universal covering endowed with the pullback of g — are diffeomorphic to the standard sphere in R^n. Given a point $p \in \widetilde{M}$, and two geodesics γ, β in $(\widetilde{M}, g)$ parametrized by arclength such that $p = \gamma(0) = \beta(0)$, we say that these geodesics diverge if $\lim_{t\to+\infty} d(\gamma(t), \beta(t)) = \infty$. Although two different geodesic rays starting from a point in $\widetilde{M}$ diverge in all well-known examples of manifolds without conjugate points (e.g., nonpositive curvature, no focal points, bounded asymptote), there is no general proof of this fact so far. The problem has been already considered by L. Green [**11**] in the late 50's, where Green deals with the divergence of radial Jacobi fields. Later, P. Eberlein [**6**] proves that radial Jacobi fields diverge along any geodesic in

2000 ***Mathematics Subject Classification.*** — 37D40, 53C23, 53C70.
Key words and phrases. — Conjugate points, horospheres, divergence of geodesic rays, local product structure, Gromov hyperbolic space, expansive flow.

Partially supported by CNPq, FAPERJ, and TWAS.

(M, g), but observes that the divergence might not be uniform, it could depend on the geodesic (in the same work [**6**], Eberlein points out a gap in Green's paper). The divergence of rays and Jacobi fields is related with many important geometric properties of manifolds without conjugate points, like the continuity of the horospherical foliations and Green bundles, and the existence of good compactifications of $\widetilde{M}$. This motivated somehow the introduction of some categories of manifolds without conjugate points in the literature (see for instance [**5**], [**6**], [**8**], for the so-called bounded asymptote condition, [**16**] for the Axiom of asymptoticity, [**14**] for the proof of the superlinear divergence of radial Jacobi fields in manifolds with bounded asymptote). The usual approach to the proofs of the continuity of horospheres, Green bundles, and divergence of rays, relies on strong assumptions on the asymptotic behaviour of geodesics and Jacobi fields (e.g., convexity in the case of nonpositive curvature; uniformly bounded asymptotic behaviour of Green Jacobi fields in the case of manifolds without focal points and manifolds with bounded asymptote). We shall present in this paper a more topological approach to the problem of the divergence of rays, based on simple variational properties of horospheres. Given $\theta = (p, v)$ in the unit tangent bundle $T_1\widetilde{M}$ of $\widetilde{M}$, we shall denote by $\gamma_\theta(t)$ the geodesic parametrized by arclength whose initial conditions are $\gamma_\theta(0) = p$, $\gamma'_\theta(0) = v$. We shall denote by $H_\theta(t)$ the horosphere of the geodesic γ_θ containing the point $\gamma_\theta(t)$. We say that the map $\theta \mapsto H_\theta(0)$ is continuous (in the compact open topology) if given a compact ball $B_r(q) \subset \widetilde{M}$ of radius r, and $\varepsilon > 0$, there exists $\delta = \delta(r, q, \varepsilon)$ such that if $\| \theta - \alpha \| \leqslant \delta$ then the Hausdorff distance d_H between the sets

$$d_H(H_\theta(0) \cap B_r(q), H_\alpha(0) \cap B_r(q)) \leqslant \varepsilon.$$

The introduction of this notion is motivated by the works of Pesin [**16**], Eschenburg [**8**], and Ballmann, Brin, and Burns [**1**]. Observe that, if M is compact, the number δ above does not depend on the point q, since every horosphere has an isometric image that meets a compact fundamental domain of $\widetilde{M}$ (horospheres are preserved by isometries of $(\widetilde{M}, g)$). In all known examples of manifolds without conjugate points the map $\theta \mapsto H_\theta(0)$ is continuous. Moreover, the assumption of the continuity of horospheres does not carry (a priori) any restrictions on either the convexity of the metric or the behaviour of Jacobi fields. The main result of the paper is the following:

Theorem 1. — *Let (M, g) be a compact, C^∞ Riemannian manifold without conjugate points. Assume that the map $\theta \mapsto H_\theta(0)$ is continuous in $T_1\widetilde{M}$. Then, every two different geodesics $\gamma(t)$, $\beta(t)$ with $\gamma(0) = \beta(0)$ in $\widetilde{M}$ diverge.*

The proof of Theorem 1 is done in Sections 1 and 2, where we also study some general problems concerning asymptoticity properties of geodesics which were introduced by Croke and Schroeder in [**4**]. Namely, consider the relation $\boldsymbol{R}$ between geodesics in $\widetilde{M}$ defined by: $\gamma \;\boldsymbol{R}\; \beta$ if and only if γ is a Busemann asymptote of β. We show in

Section 1 that, under our continuity hypothesis, this relation is an equivalence relation. In the remaining sections we give some applications of Theorem 1. The results in Section 3 are inspired in the following classical result of Eberlein: Let (M, g) be a compact, C^∞ Riemannian manifold without conjugate points. Assume that the Green subbundles $E^s(\theta)$, $E^u(\theta)$ are linearly independent at every point $\theta \in T_1M$. Then the geodesic flow of (M, g) is Anosov. Recall that the geodesic flow $\phi_t : T_1M \to T_1M$ is defined by $\phi_t(\theta) = (\gamma_\theta(t), \gamma'_\theta(t))$. We obtain in Section 3 a sort of topological version of Eberlein's result. Recall that $(\widetilde{M}, g)$ is a *Gromov hyperbolic space* if there exists $\delta > 0$ such that every geodesic triangle formed by the union of three geodesic segments $[x_0, x_1]$, $[x_1, x_2]$, $[x_2, x_0]$ satisfies the following property: the distance from any $p \in [x_i, x_{i+1}]$ to $[x_{i+1}, x_{i+2}] \cup [x_{i+2}, x_i]$ is bounded above by δ (the indices are taken mod. 3). The main Theorem of Section 3 is the following.

Theorem 2. — *Let (M, g) be a compact Riemannian manifold without conjugate points. Suppose that the map $\theta \mapsto H_\theta(0)$ is continuous in the compact open topology in $\widetilde{M}$. Then, if $H_{(p,v)}(0) \cap H_{(p,-v)}(0) = \{p\}$ for every $(p, v) \in T_1\widetilde{M}$, the universal covering $(\widetilde{M}, g)$ is a Gromov hyperbolic space.*

Using some results in [**18**] we shall show that Theorem 2 is equivalent to the following result:

Theorem 3. — *Let (M, g) be a compact Riemannian manifold without conjugate points. Suppose that the canonical liftings in T_1M of the submanifolds $H_{(p,v)}(0)$, $H_{(p,-v)}(0)$ give rise to continuous foliations H^s, H^u having a local product structure. Then $(\widetilde{M}, g)$ is a Gromov hyperbolic space.*

For the definition of the canonical liftings of the horospheres we refer to Section 3. A pair of ϕ_t-invariant foliations F_1, F_2 in T_1M has a *local product structure* if there exists an atlas $\{\Phi_i : U_i \subset T_1M \to R^{2n-1}\}$ of T_1M such that

(1) Every Φ_i is continuous.

(2) Each local chart is of the form $\Phi_i = (x^i, y^i, t)$, $t \in (-\varepsilon, \varepsilon)$, where the level sets $x^i = \mathit{constant}$, $y^i = \mathit{constant}$ are connected components of the foliations F_1, F_2 respectively.

In virtue of Theorems 2 and 3, we can say that the topological transversality (meaning local product structure) of the horospherical foliations in T_1M implies that $\widetilde{M}$ is a Gromov hyperbolic space. Notice that Theorem 1 is true for manifolds of nonpositive curvature, because the hypotheses in the Theorem imply that there are no flat planes in $\widetilde{M}$ ([**5**]). It also holds for manifolds without focal points, but if we allow focal points many key facts of the theory (convexity, bounded asymptotic behaviour of Jacobi fields and geodesics, etc.) might not hold.

In Section 4 we get some results concerning the boundary of a Gromov hyperbolic group that covers a compact manifold without conjugate points. Suppose that the map $\theta \mapsto H_\theta(0)$ is continuous. Then we show that, if the fundamental group of M is Gromov hyperbolic, its ideal boundary is a sphere. This fact is well known for compact manifolds of nonpositive curvature whose fundamental group is Gromov hyperbolic. However, if we drop the assumption on the curvature it is not clear whether the ideal boundary of the fundamental group is a sphere.

Finally, in Section 5, we apply the results of Sections 1, 2 to manifolds satisfying the so-called Axiom of Asymptoticity, introduced by Pesin [**16**]. This notion is perhaps the first one in the literature of the research about continuity of horospheres which does not consider any assumptions on the C^2 features of the metric (convexity, Jacobi fields).

1. Horospheres and Busemann flows in $\widetilde{M}$

Throughout the paper, (M,g) will be a C^∞, compact Riemannian manifold without conjugate points. All the geodesics will be parametrized by arc length. We shall often call by $[p,q]$ the geodesic segment joining two points in $\widetilde{M}$. A very special property of manifolds with no conjugate points is the existence of the so-called *Busemann functions*: given $\theta = (p,v) \in T_1\widetilde{M}$ the *Busemann function* $b^\theta : \widetilde{M} \to R$ associated to θ is defined by

$$b^\theta(x) = \lim_{t\to+\infty}(d(x,\gamma_\theta(t)) - t)$$

The level sets of b^θ are the *horospheres* $H_\theta(t)$ where the parameter t means that $\gamma_\theta(t) \in H_\theta(t)$. Notice that $\gamma_\theta(t)$ intersects each level set of b^θ perpendicularly at only one point in $H_\theta(t)$, and that $b^\theta(H_\theta(t)) = -t$ for every $t \in R$. Next, we list some basic properties of horospheres and Busemann functions that will be needed in the forthcoming sections (see [**16**], [**4**] for instance, for details).

Lemma 1.1

(1) *b^θ is a C^1 function for every θ.*

(2) *The gradient ∇b^θ has norm equal to one at every point.*

(3) *Every horosphere is a C^{1+K}, embedded submanifold of dimension $n-1$ (C^{1+K} means K-Lipschitz normal vector field), where K is a constant depending on curvature bounds.*

(4) *The orbits of the integral flow of $-\nabla b^\theta$, $\psi_t^\theta : \widetilde{M} \to \widetilde{M}$, are geodesics which are everywhere perpendicular to the horospheres H_θ. In particular, the geodesic γ_θ is an orbit of this flow and we have that*

$$\psi_t^\theta(H_\theta(s)) = H_\theta(s+t)$$

for every $t, s \in R$.

A geodesic β is *asymptotic* to a geodesic γ in $\widetilde{M}$ if there exists a constant $C > 0$ such that $d(\beta(t),\gamma(t)) \leqslant C$ for every $t \geqslant 0$. We shall denote by *Busemann asymptotes* of γ_θ

the orbits of the flow ψ_t^θ. Busemann asymptotes of γ_θ might not be asymptotic to γ_θ, so the relation between geodesics $\gamma \boldsymbol{R} \beta$ if and only if "γ is a Busemann asymptote of β" might not be an equivalence relation. Observe that in all known examples of manifolds without conjugate points (nonpositive curvature, no focal points, metrics on surfaces without conjugate points [**1**]), the relation $\boldsymbol{R}$ is an equivalence relation. Lemma 1.1, item 4, implies that the horospheres $H_\theta(t)$ are equidistant, i.e., given any point $p \in H_\theta(s)$, then the distance $d(p, H_\theta(t))$ is equal to $|t-s|$. The canonical lifting in $T_1\widetilde{M}$ of $H_\theta(t)$ is the set $\widetilde{H}_\theta(t) = \{(p, -\nabla_p b^\theta), p \in H_\theta(t)\}$. Another way of defining the horosphere $H_\theta(0)$ is

$$H_\theta(0) = \lim_{r\to+\infty} S_r(\gamma_\theta(r)),$$

where $S_r(p)$ is the sphere of radius r centered at p, and the limit is uniform on compact subsets of $\widetilde{M}$. In other words, given $D > 0$, $q \in \widetilde{M}$, and $\varepsilon > 0$, there exists $T > 0$ such that the Haussdorf distance between the restrictions of $H_\theta(0)$ and $S_r(\gamma_\theta(r))$ to the ball $B_D(q)$ of radius D centered at q is less than ε for every $r \geqslant T$. We shall denote by d_H the Haussdorf distance between subsets in $\widetilde{M}$. Actually, given $\theta \in T_1\widetilde{M}$, the spheres $S_r(\gamma_\theta(r))$ converge to $H_\theta(0)$ in the C^1 topology uniformly of compact subsets (see [**16**]) as a consequence of the bounded geometry of $\widetilde{M}$. The notion of continuity of $\theta \mapsto H_\theta(0)$ given in the introduction is equivalent to the following: let θ_n converge to θ, then $H_{\theta_n}(0)$ converges to $H_\theta(0)$ uniformly on compact subsets of $\widetilde{M}$. Although it is clear that $H_\theta(t)$ depends continuously on $t \in R$, it is not known whether $H_\theta(0)$ depends continuously on θ. The continuity of $\theta \mapsto H_\theta$ is equivalent to the continuity in the C^1 topology of the map $\theta \mapsto b^\theta$ uniformly on compact subsets of $\widetilde{M}$. For the purposes of this section, it will be more convenient to state our results in terms of Busemann functions instead of using horospheres, in general the notation becomes simpler. Let $b^{\theta,t}(p) = d(p, \gamma_\theta(t)) - t$, so $b^\theta(p) = \lim_{t\to+\infty} b^{\theta,t}(p)$, for every $p \in \widetilde{M}$. We shall denote by $d_{T_1M}(\cdot,\cdot)$ the Sasaki metric in T_1M. The following result tells us that the continuity of $\theta \mapsto H_\theta(0)$ implies that horospheres can be uniformly approached by large spheres.

Lemma 1.2. — *Let (M,g) be a compact manifold without conjugate points, such that the map $\theta \mapsto b^\theta$ is continuous. Then, given $D > 0$, $\varepsilon > 0$, there exists $T > 0$ such that for every $\theta = (p_0, v_0) \in T_1\widetilde{M}$, and every ball B_D of radius D containing p_0, we have*

$$|b^\theta(p) - b^{\theta,t}(p)| \leqslant \varepsilon,$$

for every $p \in B_D$ and $t \geqslant T$.

Proof. — Let us first recall that the family of functions $b^{\theta,t}$ converges monotonically to b^θ, i.e., the difference $b^\theta(p) - b^{\theta,t}(p)$ decreases with respect to t for every p. This is due to the fact that the spheres $S_t(\gamma_\theta(t))$ converge monotonically to $H_\theta(0)$, i.e., the Hausdorff distance between the restrictions of $S_t(\gamma_\theta(t))$ and $H_\theta(0)$ to compact sets decreases to 0 as t goes to $+\infty$. Since $b^\theta(p) = \lim_{t\to+\infty} b^{\theta,t}(p)$, there exists $T_{\theta,p,\varepsilon} > 0$ such that

$$|b^\theta(p) - b^{\theta,t}(p)| \leqslant \varepsilon$$

for every $t \geqslant T_{\theta,p,\varepsilon}$.

Claim 1. — *The number $T_{\theta,p,\varepsilon}$ can be made independent of p in B_D. Moreover, it depends on θ, D, ε.*

Indeed, this is a straighforward application of Dini's Lemma about uniform convergence of monotone sequences of functions: the family of functions

$$f_t : B_D \longrightarrow R, \quad f_t(p) = |b^\theta(p) - b^{\theta,t}(p)|$$

converges monotonically to zero at each point $p \in B_D$; so by Dini's Lemma the family converges uniformly to zero in the compact set B_D.

Next, we would like to relate $T_{\theta,D,\varepsilon}$ and $T_{\alpha,D,\varepsilon}$ for α close to θ. At this point we need the continuity of horospheres, we have not used this hypothesis so far. Recall that

$$|b^{\theta,t}(p) - b^{\alpha,t}(p)| = |d(p,\gamma_\theta(t)) - d(p,\gamma_\alpha(t))| \leqslant d(\gamma_\theta(t),\gamma_\alpha(t)).$$

By hypothesis, $\theta \mapsto b^\theta$ depends continuously on θ uniformly on compact subsets, so there exists $\delta > 0$ such that if $d_{T_1\widetilde{M}}(\theta,\alpha) < \delta$ then $|b^\theta(p) - b^\alpha(p)| \leqslant \varepsilon$ for every $p \in B_D$. On the other hand, by continuity of geodesics with respect to initial conditions, given $\varepsilon > 0$ there exists $\delta' = \delta'(\theta,\varepsilon) > 0$ such that if $d_{T_1\widetilde{M}}(\theta,\alpha) \leqslant \delta'$ then

$$|b^{\theta,t}(p) - b^{\alpha,t}(p)| \leqslant d(\gamma_\theta(t),\gamma_\alpha(t)) \leqslant \varepsilon,$$

for every $0 < t \leqslant T_{\theta,D,\varepsilon}$, and every $p \in B_D$. This implies that

$$\begin{aligned}|b^\alpha(p) - b^{\alpha,T}(p)| &\leqslant |b^\alpha(p) - b^\theta(p)| + |b^\theta(p) - b^{\theta,T}(p)| + |b^{\theta,T}(p) - b^{\alpha,T}(p)| \\ &\leqslant \varepsilon + \varepsilon + \varepsilon = 3\varepsilon,\end{aligned}$$

for every $d_{T_1\widetilde{M}}(\theta,\alpha) \leqslant \min\{\delta,\delta'\}$, for $T = T_{\theta,D,\varepsilon}$, and every $p \in B_D$. Therefore, by the monotonicity of the limit $b^\alpha(p) = \lim_{t\to+\infty} b^{\alpha,t}(p)$ we have that $|b^\alpha(p) - b^{\alpha,t}(p)| \leqslant 3\varepsilon$ for every $t \geqslant T = T_{\theta,D,\delta}$ and $d_{T_1\widetilde{M}}(\theta,\alpha) < \varepsilon_\theta = \min\{\varepsilon,\varepsilon'\}$. This means that

$$T_{\alpha,D,3\varepsilon} \leqslant T_{\theta,D,\varepsilon},$$

for every $d_{T_1\widetilde{M}}(\theta,\alpha) < \varepsilon_\theta$. Now, take a compact fundamental domain M_0 of the manifold M, let $K = \{\theta = (x,w), x \in M_0, w \in T_1\widetilde{M}\}$, and let us cover K by open neighborhoods $V_{\varepsilon_\theta}(\theta)$ where $\theta \in K$. Take a finite covering $\cup_{i=1}^m V_{\varepsilon_{\theta_i}}(\theta_i)$ of K by these neighborhoods. Let $T = \max_{i=1,2,..,m}\{T_{\theta_i,D,\delta}\}$, and assume without loss of generality that the ball B_D contains the fundamental domain M_0. Then we get that $|b^\alpha(p) - b^{\alpha,t}(p)| \leqslant 3\delta$ for every $p \in B_D$, $\alpha \in K$, and every $t \geqslant T$. Since balls, horospheres, and Busemann functions are preserved by isometries in $\widetilde{M}$, we deduce that $|b^\alpha(p) - b^{\alpha,t}(p)| \leqslant 3\varepsilon$ for every $\alpha = (q,w) \in T_1\widetilde{M}$, $t \geqslant T$, and p in every ball of radius D containing q, as we wished to show. $\square$

Corollary 1.1. — *Let M be a compact Riemannian manifold without conjugate points, and assume that the map $\theta \mapsto H_\theta(0)$ is continuous. Then, if $q \in H_\theta(0)$, and $\alpha = (q, -\nabla_q b^\theta)$, we have that $b^\alpha = b^\theta$. In particular, $H_\alpha(0) = H_\theta(0)$ and the relation $\boldsymbol{R}$ is an equivalence relation between geodesics in $T_1\widetilde{M}$.*

Proof. — Let γ_α be the Busemann asymptote of θ through α. Let us denote by $[x,y]$ the geodesic segment joining the points x, y in $\widetilde{M}$. By definition, the geodesic γ_α is the limit of the geodesic segments $[q, \gamma_\theta(t)]$, where $t \to +\infty$. Let γ_{α_t} be the geodesic containing the segment $[q, \gamma_\theta(t)]$, with $\gamma_{\alpha_t}(0) = q$ and $\gamma_{\alpha_t}(r_t) = \gamma_\theta(t)$ for some positive r_t. Then, $\lim_{t\to+\infty} \alpha_t = \alpha$, and by the continuity hypothesis, $\lim_{t\to+\infty} b^{\alpha_t} = b^\alpha$, this limit being uniform on compact subsets of $\widetilde{M}$. Let $\theta = (p_0, v_0)$, and consider a compact ball B containing the foot points p_0 and q of the vectors θ and α respectively. By Lemma 1.2, the functions b^θ, b^α, b^{α_t}, $t > 0$, can be uniformly approached by radial functions $b^{\theta,T}$, $b^{\alpha,T}$, $b^{\alpha_t,T}$ in the compact ball B. Namely, given $\varepsilon > 0$ there exists $T_\varepsilon > 0$ such that for every $T > T_\varepsilon$, every $p \in B$, we have that

$$|b^\theta(p) - b^{\theta,T}(p)| \leqslant \varepsilon, \quad |b^\alpha(p) - b^{\alpha,T}(p)| \leqslant \varepsilon, \quad |b^{\alpha_t}(p) - b^{\alpha_t,T}(p)| \leqslant \varepsilon,$$

for every $t > 0$. Since the functions b^{α_t} converge uniformly in B to the function b^α, there exists $S_\varepsilon > 0$ such that

$$|b^\alpha(p) - b^{\alpha,t}(p)| \leqslant \varepsilon,$$

for every $p \in B$ and $t \geqslant S_\varepsilon$. Recall that the number r_t is defined by $\gamma_{\alpha_t}(r_t) = \gamma_\theta(t)$.

Claim 1. — $\lim_{t\to+\infty} |r_t - t| = 0$.

This follows easily by definition: $r_t = d(\gamma_{\alpha_t}(0), \gamma_\theta(t)) = d(q, \gamma_\theta(t))$; which implies that

$$r_t - t = d(q, \gamma_\theta(t)) - t = b^{\theta,t}(q).$$

And, since $q \in H_\theta(0) = (b^\theta)^{-1}(0)$, we get

$$0 = b^\theta(q) = \lim_{t\to+\infty} b^{\theta,t}(q) = \lim_{t\to+\infty} |r_t - t|.$$

Claim 2. — $b^{\alpha_t,r_t}(p) = b^{\theta,t}(p) + t - r_t$ *for every* $p \in \widetilde{M}$.

Just check the definitions:

$$\begin{aligned} b^{\alpha_t,r_t}(p) &= d(p, \gamma_{\alpha_t}(r_t)) - r_t \\ &= d(p, \gamma_\theta(t)) - r_t \\ &= b^{\theta,t}(p) + t - r_t. \end{aligned}$$

Hence, if $t, r_t \geqslant \sup\{T_\varepsilon, S_\varepsilon\}$ we obtain,

$$\begin{aligned} |b^\alpha(p) - b^\theta(p)| &\leqslant |b^\alpha(p) - b^{\alpha_t}(p)| + |b^{\alpha_t}(p) - b^{\alpha_t,r_t}(p)| \\ &\quad + |b^{\alpha_t,r_t}(p) - b^{\theta,t}(p)| + |b^{\theta,t}(p) - b^\theta(p)| \\ &\leqslant 3\varepsilon + |r_t - t|, \end{aligned}$$

for every $p \in B$. Since ε and B are arbitrary, and by Claim 2, $\lim_{t\to+\infty} |r_t - t| = 0$, we deduce that $b^\alpha(p) = b^\theta(p)$ for every $p \in \widetilde{M}$. Since the level sets of Busemann functions are horospheres, then the horospheres of γ_θ are the same horospheres of γ_α. It is clear that the relation $\gamma\ \boldsymbol{R}\ \beta$ if and only if γ is a Busemann asymptote of β is an equivalence relation: their horospherical foliations will be the same, and therefore, their Busemann flows will coincide. □

We would like to point out that Croke and Schroeder in [**4**] posed the question of whether the relation $\boldsymbol{R}$ is an equivalence relation in the universal covering of a compact manifold without conjugate points. They show in fact that, if M is analytic, then $b^\alpha(p) - b^\theta(p)$ is constant if γ_θ, γ_α are axes of the same deck transformation in $\pi_1(M)$. Corollary 1.1 asserts that, under the assumption of continuity of horospheres, $b^\alpha(p) - b^\theta(p)$ is constant whenever γ_α is a Busemann asymptote of γ_θ.

2. The divergence of geodesic rays

The main result of the section is the following:

Theorem 2.1. — *Let (M,g) be a compact manifold without conjugate points such that the map $\theta \mapsto H_\theta(0)$ is continuous. Then the geodesic rays diverge uniformly in $\widetilde{M}$, namely, given $\varepsilon > 0$, $L > 0$, there exists $s_{\varepsilon,L} > 0$ such that any two geodesic rays $\gamma_{(p,v)}$, $\gamma_{(p,w)}$, where $p \in \widetilde{M}$ and v, w form an angle $\angle(v,w) \geqslant \varepsilon$, satisfy*

$$d(\gamma_{(p,v)}(t), \gamma_{(p,w)}(t)) \geqslant L$$

for every $t \geqslant s_{\varepsilon,L}$.

We begin by recalling the first variation formula.

Lemma 2.1. — *Let $\gamma(t)$ be a geodesic of (M,g) parametrized by arclength. Let $f : [a,b] \times (-\varepsilon,\varepsilon) \to M$ be a differentiable variation of $\gamma[a,b]$, i.e., $f(t,0) = \gamma(t)$. Then, the length $L(x)$ of the curve $f_x(t) = f(t,x)$ satisfies*

$$L'(0) = \left\langle \frac{\partial f}{\partial x}(t,0), \gamma'(t) \right\rangle \Big|_a^b .$$

Now, let γ_θ, γ_α be two geodesics in $\widetilde{M}$ with $\gamma_\theta(0) = \gamma_\alpha(0) = p$. Assume that $\theta \neq \alpha$. Take $s < 0$, and let $f : [s,0] \times (-a,a) \to \widetilde{M}$ be the variation of $\gamma_\theta[s,0]$ defined by

– $f(t,0) = \gamma_\theta(t)$ for every $t \in [s,0]$.

– $f_x(t) = f(t,x)$ is the geodesic segment joining $\gamma_\theta(s)$ and $\gamma_\alpha(x)$ for each $|x| < a$ (observe that t might not be the arclength parameter of the geodesic $f_x(t)$).

If $|s|$ and a are small enough, the geodesic segments $f_x(t)$ are unique and minimizing. Hence, the variation f is differentiable because of the smooth dependence of small geodesic segments with respect to initial conditions. Notice that $\frac{\partial f}{\partial x}(s,x) = 0$ for every $x \in (-a,a)$. Also, we have that $\frac{\partial f}{\partial x}(0,x) = \gamma'_\alpha(x)$, and $\frac{\partial f}{\partial t}(t,0) = \gamma'_\theta(t)$.

Corollary 2.1. — *Let $\varepsilon = \angle(\gamma_\theta'(0), \gamma_\alpha'(0)) = \| \theta - \alpha \|_{T_1\widetilde{M}}$. Then, there exist $s_0 > 0$, $K_0 = K_0(\| K \|_\infty)$, such that for every variation $f : [s, 0] \times (-a, a) \to M$ as above, with $|s| \leqslant s_0$, $|a| \leqslant s$; there exists $\overline{\delta} = \overline{\delta}(\varepsilon, s, \| K \|_\infty)$, such that for every $x \in (-\overline{\delta}, \overline{\delta})$ we have*

$$|L(x) - (L(0) + L'(0)x)| \leqslant K_0 x^2,$$

where $\| K \|_\infty$ is the supremum of the sectional curvatures of (M, g).

The proof of Corollary 2.1 is straightforward from the first and second variation formulas and the compactness of M. Corollary 2.1 can be viewed as a shortcut lemma, and we shall use it to give a sort of lower estimate of the distance between the horospheres $H_\theta(t)$, $H_\alpha(t)$ when γ_θ, γ_α are two different geodesic rays starting at the same point.

Lemma 2.2. — *Let γ_θ, γ_α be two geodesic rays with $\gamma_\theta(0) = \gamma_\alpha(0) = p$. Let $\theta = (p, v)$, $\rho = \| \theta - \alpha \|$. Then, for every $\rho > 0$ there exists $\delta_1 = \delta_1(\rho, \| K \|_\infty) > 0$, such that for every t, t', satisfying $\gamma_\alpha(t') = \gamma_\alpha \cap H_{(p,-v)}(-t)$ we have that*

$$|t' - t| \geqslant \delta_1(1 - \cos(\rho)).$$

Proof. — Let s_0, $\delta_0 = \overline{\delta}(\varepsilon, s_0, \| K \|_\infty)$, K_0 be the numbers defined in Corollary 2.1. We know that $t' \geqslant t$, because

$$\begin{aligned} t' = d(\gamma_\alpha(0), \gamma_\alpha(t')) &= d(\gamma_\theta(0), \gamma_\alpha(t')) \\ &\geqslant d(\gamma_{(p,-v)}(0), H_{(p,-v)}(-t)) \\ &= d(\gamma_\theta(0), \gamma_\theta(t)) = t, \end{aligned}$$

where we used that $\gamma_\theta(t) = \gamma_{(p,-v)}(-t)$ is the geodesic that realizes the distance between $p = \gamma_{(p,-v)}(0)$ and $H_{(p,-v)}(-t)$. Thus, $s_0 + t' \geqslant s_0 + t$. By Corollary 2.1,

$$\begin{aligned} d(\gamma_\theta(-s_0), \gamma_\alpha(\delta)) + d(\gamma_\alpha(\delta), \gamma_\alpha(t')) &= d(\gamma_\theta(-s_0), \gamma_\theta(0)) + \delta\cos(\rho) + O(\delta^2) + t' - \delta \\ &= s_0 + \delta\cos(\rho) + t' - \delta + O(\delta^2), \end{aligned}$$

for every $|\delta| \leqslant \delta_0$; which implies that

$$s_0 + t = d(\gamma_\theta(-s_0), \gamma_\theta(t)) < s_0 + \delta\cos(\rho) + t' - \delta + O(\delta^2).$$

for every $|\delta| < \delta_0$ (see figure 1).

So we get, by Corollary 2.1,

$$t' - t > \delta - \delta\cos(\rho) + O(\delta^2) = \delta(1 - \cos(\rho)) + O(\delta^2),$$

for every $|\sigma| < \delta_0$. On the other hand, there exists $0 < \delta_1 = \delta_1(\rho, K_0) \leqslant \delta_0$ such that $\delta_1(1 - \cos(\rho)) + O(\delta_1^2) > \frac{\delta_1}{2}(1 - \cos(\rho))$, therefore $t' - t > \frac{\delta_1}{2}(1 - \cos(\rho))$ as we wished to show. □

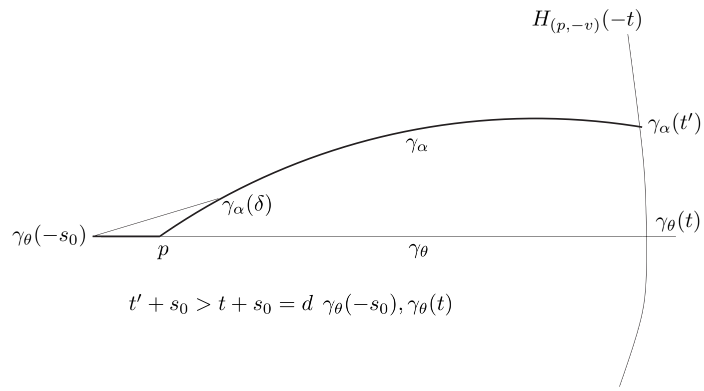

FIGURE 1

Proof of Theorem 2.1. — Let (M, g) be a compact manifold without conjugate points such that the map $\theta \mapsto H_\theta(0)$ is continuous. Let $C > 0$, and consider two geodesic rays γ_θ, γ_α with $\gamma_\theta(0) = \gamma_\alpha(0) = p$. Let us define

$$T_{\alpha,\theta,C} = \sup\{t > 0, d(\gamma_\theta(t), \gamma_\alpha(t)) \leqslant C\}.$$

We are going to show that there exists $T(\| \theta - \alpha \|, C) > 0$ such that $T_{\alpha,\theta,C} \leqslant T(\| \theta - \alpha \|, C)$ for every α, θ in $T_1\widetilde{M}$ with the same foot point p. In particular, $T(\| \theta - \alpha \|, C)$ depends on C and on the angle between θ and α. So let $t > 0$ be such that $d(\gamma_\theta(t), \gamma_\alpha(t)) \leqslant C$.

Claim 1. — *There exists a number t' such that*

$$\gamma_\alpha(t') = \gamma_\alpha \cap H_{(p,-v)}(-t + C).$$

In fact, recalling that $b^{(p,-v)}$ is the Busemann function of $\gamma_{(p,-v)}$, the hypotheses on $\gamma_\alpha(t)$ implies that

$$t - C \leqslant b^{(p,-v)}(\gamma_\alpha(t)) \leqslant t + C,$$

which is easy to check by the definition of the Busemann function (recall that $\gamma_\theta(t) = \gamma_{(p,-v)}(-t)$). On the other hand, observe that $H_{(p,-v)}(-t) = (b^{(p,-v)})^{-1}(t)$, and since $\gamma_\alpha(0) = \gamma_\theta(0) \in H_{(p,-v)}(0)$ we have that $b^{(p,-v)}(\gamma_\alpha(0)) = 0$, $b^{(p,-v)}(\gamma_\alpha(t)) \geqslant t - C$. Since the function $g(t) = b^{(p,-v)}(\gamma_\alpha(t))$ is continuous, it assumes all the values in the interval $[0, t - C]$. Hence, there exists t' such that $g(t') = t - C$. Or equivalently, $\gamma_\alpha(t') \in (b^{(p,-v)})^{-1}(t - C) = H_{(p,-v)}(-t + C)$. It is clear that $t' > t - C$, because $t - C$ is the distance between $p = \gamma_{(p,-v)}(0)$ and $H_{(p,-v)}(-t + C)$.

We can assume without loss of generality, by changing t by $t - C$, that γ_α intersects $H_{\gamma_\theta}(t)$. In this case we have $t' > t$, and we can assume that

$$t' = \inf\{s > t, \gamma_\alpha(s) \in H_{(p,-v)}(-t)\}.$$

Claim 2. — *Given $\varepsilon > 0$, there exists $T = T(\varepsilon, C) > 0$ such that $|t - t'| \leqslant \varepsilon$ for every $t \geqslant T$.*

This follows from Lemma 1.2. Indeed, since $\gamma_\theta(0) = \gamma_\alpha(0) = p$, the points $\gamma_\theta(t), \gamma_\alpha(t)$ belong to the sphere $S_t(p)$ for every t. Let $\phi_t : T_1\widetilde{M} \to T_1\widetilde{M}$ be the geodesic flow of $\widetilde{M}$. Let $\beta = \phi_{-t}(p, -v)$. The geodesic γ_β is a reparametrization of γ_θ satisfying $\gamma_\beta(0) = \gamma_\theta(t)$, and $\gamma_\beta(t) = \gamma_\theta(0) = p$. By the choice of t', we have that the ball $B_{3C}(\gamma_\beta(0))$ contains $\gamma_\alpha(t)$ and $\gamma_\alpha(t')$. Moreover, the points $\gamma_\theta(t) = \gamma_\beta(0)$ and $\gamma_\alpha(t)$ belong to the sphere $S_t(p) = S_t(\gamma_\beta(t))$. Thus, by Lemma 1.2, for $D = 3C$ and $\varepsilon > 0$, there exists $T > 0$ such that

$$d_H(H_\beta(0) \cap B_D(\gamma_\theta(t)), S_t(p) \cap B_D(\gamma_\theta(t))) \leqslant \varepsilon,$$

for every $t \geqslant T$. In particular, the point $\gamma_\alpha(t)$, that belongs to $S_t(p) \cap B_D(\gamma_\theta(t))$, is within a distance ε from $H_\beta(0)$ (see figure 2).

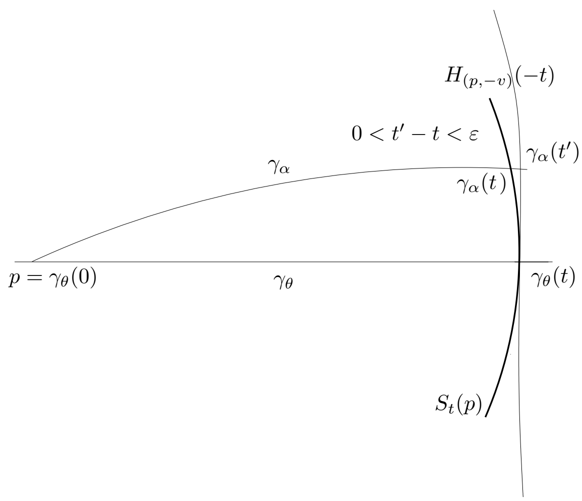

FIGURE 2

Hence, choosing ε small enough, the first time $s = t' > t$ where occurs a transversal intersection between $\gamma_\alpha(s)$ and $H_\beta(0) = H_{(p,-v)}(-t)$ satisfies

$$d(\gamma_\alpha(t'), \gamma_\alpha(t)) = |t' - t| \leqslant \varepsilon,$$

for every $t \geqslant T$. This finishes the proof of the Claim.

Applying Lemma 2.2 and Claim 2 to the geodesics γ_θ, γ_α, we have that for every $t \geqslant T(\varepsilon, C)$,

$$\varepsilon \geqslant |t' - t| \geqslant \delta_1(1 - \cos(\rho)),$$

where ρ is the angle between θ and α. Therefore, if $\rho > 0$, the number $\varepsilon_\rho = \delta_1(1 - \cos(\rho))$ is strictly greater than 0. Hence, the above inequality holds only if

$\varepsilon \geqslant \varepsilon_\rho$, and therefore, t has to be smaller than $T(\varepsilon_\rho/2, C)$ for instance. This implies that $T_{\theta,\alpha,C} \leqslant T(\varepsilon_\rho/2, C)$, as we wished to prove.

3. Topological transversality of horospheres and expansiveness are equivalent

The goal of this section is to combine the divergence of geodesic rays with some ideas connecting topological dynamics and global geometry. We shall obtain a sort of topological version of Eberlein's theorem about the characterization of Anosov geodesic flows by the transversality of the Green bundles. We begin by showing one of the main consequences of the divergence of geodesic rays.

Lemma 3.1. — *Let M be a compact manifold without conjugate points, such that the map $\theta \mapsto H_\theta(0)$ is continuous. If a geodesic β is asymptotic to a geodesic γ_θ in $\widetilde{M}$, then β is a Busemann asymptote of γ_θ.*

Proof. — Let us assume that $d(\gamma_\theta(t), \beta(t)) \leqslant C$ for every $t \geqslant 0$. Consider the geodesic segments $[\beta(0), \gamma_\theta(t)]$ (using the notation of Corollary 1.1), where $t \geqslant 0$. Let β_t be the geodesic defined by $\beta_t(0) = \beta(0)$, $\beta_t(T_t) = \gamma_\theta(t)$ for some positive T_t. Since $d(\beta_t(T_t), \beta(t)) \leqslant C$ for every $t \geqslant 0$, letting $t \to +\infty$ we have, by the uniform divergence of geodesic rays, that
$$\lim_{t \to +\infty} \beta_t = \beta,$$
uniformly on compact sets. This implies that β is a Busemann asymptote of γ_θ by definition. □

Lemma 3.2. — *Let M be a compact manifold without conjugate points, such that the map $\theta \mapsto H_\theta(0)$ is continuous. Then, if $\gamma_{(p,v)}$ and γ_α are bi-asymptotic, we have that, up to a reparametrization of γ_α, $\gamma_\alpha(0) \in H_{(p,v)}(0) \cap H_{(p,-v)}(0)$.*

Proof. — Let $\alpha = (q, w)$. Since γ_α is bi-asymptotic to $\gamma_{(p,v)}$, then, according to Lemma 3.1 and Corollary 1.1, γ_α is Busemann asymptotic to $\gamma_{(p,v)}$, and $\gamma_{(q,-w)}$ is Busemann asymptotic to $\gamma_{(p,-v)}$. Let us assume that $q = \gamma_\alpha(0) = \gamma_\alpha \cap H_{(p,v)}(0)$. In this case, we have $b^\alpha(\gamma_\alpha(0)) = 0$, $b^{(p,v)}(\gamma_\alpha(0)) = 0$, and by Corollary 1.1, $b^\alpha = b^{(p,v)}$. If we also had that $b^{(q,-w)} = b^{(p,-v)}$, then Lemma 3.2 would hold.

Claim. — $\gamma_\alpha(0) = q \in H_{(p,-v)}(0)$.

In fact, let t_0 be such that $\gamma_\alpha(t_0) = \gamma_\alpha \cap H_{(p,-v)}(0)$. Since $-t_0 = b^{(p,v)}(\gamma_\alpha(t_0))$, and $H_{(p,-v)}(0)$ is in the region $(b^{(p,v)})^{-1}[0, +\infty)$, we have that $t_0 \leqslant 0$. Let ϕ_t be the geodesic flow of $\widetilde{M}$. Notice that
$$\phi_{-t_0}(q, -w) = (\gamma_\alpha(t_0), -\gamma'_\alpha(t_0)),$$
so we have, by Corollary 1.1, that $b^{\phi_{-t_0}(q,-w)} = b^{(p,-v)}$ as long as $\gamma_\alpha(t_0) \in H_{(p,-v)}(0)$. Hence, $\psi_t^{(p,v)} = \psi_t^{(q,w)}$, $\psi_t^{(p,-v)} = \psi_t^{(q,-w)}$ for every $t \in R$ — where ψ^θ is the Busemann

flow of θ — and the foliations $\{H_{(q,-w)}(t), t \in R\}$, $\{H_{(p,-v)}(t), t \in R\}$, coincide. We get, by the choice of t_0, that

$$H_{(q,-w)}(-t_0) = H_{\phi_{-t_0}(q,-w)}(0) = H_{(p,-v)}(0).$$

But

$$H_{(q,-w)}(-t_0) = \psi_{-t_0}^{(q,-w)}(H_{(q,-w)}(0)),$$

(see figure 3) which implies that

$$\psi_{-t_0}^{(q,-w)}(H_{(q,-w)}(0)) = \psi_{-t_0}^{(p,-v)}(H_{(q,-w)}(0)) = H_{(p,-v)}(0).$$

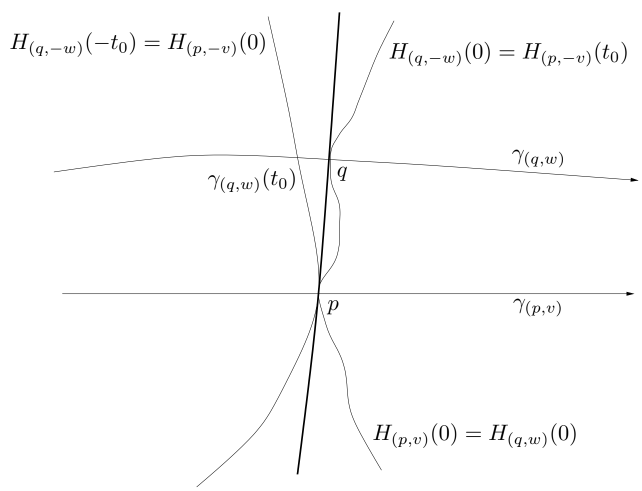

FIGURE 3

Therefore, we get

$$H_{(q,-w)}(0) = \psi_{t_0}^{(p,-v)}(H_{(p,-v)}(0)) = H_{(p,-v)}(t_0).$$

Since $H_{(q,-w)}(0)$ is in the region $(b^{(q,w)})^{-1}[0,+\infty) = (b^{(p,v)})^{-1}[0,+\infty)$, and $\gamma_{(p,v)}(-t_0) = \gamma_{(p,-v)}(t_0) \in H_{(p,-v)}(t_0)$, we have that

$$b^{(p,v)}(\gamma_{(p,v)}(-t_0)) = t_0 \geqslant 0.$$

Since t_0 was already nonpositive, we conclude that $t_0 = 0$, thus proving the Claim and the Lemma. $\square$

Lemmas 3.1, 3.2 are concerned with one basic question of the theory of manifolds without conjugate points: Are the geodesics asymptotic to γ Busemann asymptotic to γ? In all known examples of manifolds without conjugate points the answer to this question is affirmative. However, there is no proof of this fact, as far as I know, without strong assumptions on either the curvature, the convexity of the metric, or on the asymptotic behaviour of Jacobi fields.

Next, let us recall a notion that appears very often in topological dynamics. Given a C^∞ Riemannian manifold (N,g), a differentiable flow $f_t : N \to N$ without singularities is said to be *expansive* if there exists $\varepsilon > 0$ such that the following holds: let $p \in N$, and suppose that there exist $q \in N$, and a continuous, surjective reparametrization $\rho : R \to R$, with $\rho(0) = 0$, of the orbit of q such that $d(f_t(p), f_{\rho(t)}(q)) \leqslant \varepsilon$ for every $t \in R$; then q belongs to the orbit of p. The following results are proved in [**17**], [**18**].

Theorem 3.1. — *Let M be a compact manifold without conjugate points. If the geodesic flow is expansive, the universal covering of M endowed with the pullback of the metric of M is a Gromov hyperbolic space.*

Lemma 3.3. — *Let (M,g) be a compact Riemannian manifold without conjugate points. Then the geodesic flow is expansive if and only if for every pair of geodesics γ, β in $(\widetilde{M}, g)$ with $d(\gamma,\beta) \leqslant D$ we have that $\gamma = \beta$.*

So expansiveness of the geodesic flow is equivalent to the nonexistence of bi-asymptotic geodesics. Hence, to show Theorem 2 it is enough to show the following:

Lemma 3.4. — *Let (M,g) be a compact manifold without conjugate points such that $\theta \mapsto H_\theta(0)$ is continuous. If $H_{(p,v)}(0) \cap H_{(p,-v)}(0) = \{p\}$ for every $(p,v) \in T_1\widetilde{M}$, then there are no bi-asymptotic geodesics in $\widetilde{M}$.*

The proof of Lemma 3.4 is immediate from Lemma 3.2: bi-asymptotic geodesics are Busemann bi-asymptotic, and there is an injection between geodesics which are bi-asymptotic to $\gamma_{(p,v)}$ and the set $H_{(p,v)}(0) \cap H_{(p,-v)}(0)$.

For the proof of Theorem 3, we shall need some other definitions. The stable horosphere $H^s(p,v)$ of (p,v) in T_1M is defined by the following canonical lift:

$$H^s(p,v) = \Pi\{(q,w), q \in H_{(\widetilde{p},\widetilde{v})}(0), w = -\nabla_q b^{(\widetilde{p},\widetilde{v})}\},$$

where $\Pi : T_1\widetilde{M} \to T_1M$ is the canonical projection, and $\Pi(\widetilde{p},\widetilde{v}) = (p,v)$. The unstable horosphere of (p,v) is defined by

$$H^u(p,v) = \Pi\{(q,w), q \in H_{(\widetilde{p},-\widetilde{v})}(0), w = \nabla_q b^{(\widetilde{p},-\widetilde{v})}\}.$$

When the geodesic flow is Anosov, these sets coincide with the dynamical stable and unstable sets of (p,v). The collection of the stable horospheres is denoted by H^s, and the collection of unstable horospheres is denoted by H^u.

Lemma 3.5. — *Let M be a compact manifold without conjugate points such that the collections H^s, H^u are continuous foliations of T_1M having a local product structure (as defined in the introduction). Then the geodesic flow is expansive.*

Proof. — The definition of local product structure includes the continuity of the map $\theta \mapsto H_\theta(0)$, so our hypothesis implies that geodesic rays diverge in $\widetilde{M}$. Observe also that the local product structure provides a number $r > 0$, such that for each (p, v) in T_1M there exists an open ball $V(p, v)$ of radius r where

$$V(p, v) \cap H^s(p, v) \cap H^u(p, -v) = \{(p, v)\}.$$

The number $r > 0$ is uniform in T_1M by the compactness of M and the continuity of the invariant foliations. Thus, the natural projection of the neighborhood $V(p, v)$ in M, lifted to $\widetilde{M}$, gives us an open neighborhood $W(\widetilde{p})$ of a lift $\widetilde{p}$ of p in $\widetilde{M}$ where

$$W(\widetilde{p}) \cap H_{(\widetilde{p},v)}(0) \cap H_{(\widetilde{p},-v)}(0) = \{\widetilde{p}\}.$$

Moreover, $W(\widetilde{p})$ contains an open ball $B_\delta(\widetilde{p})$ centered at $\widetilde{p}$, and δ does not depend on $\widetilde{p}$. Let us prove that this local transversality implies expansiveness. Arguing by contradiction, suppose that the geodesic flow is not expansive. Then, given any $\varepsilon > 0$ there would exist a pair of different geodesics $\gamma_{\theta_\varepsilon}$, $\gamma_{\beta_\varepsilon}$ in T_1M such that $d(\gamma_{\theta_\varepsilon}(t), \gamma_{\beta_\varepsilon}(\rho(t))) \leqslant \varepsilon$, for every $t \in R$, where $\rho : R \to R$ is a reparametrization of $\gamma_{\beta_\varepsilon}$ satisfying the conditions in the definition of expansiveness. It is clear that for ε small enough, we can lift the above geodesics to $\widetilde{M}$ and get a pair of geodesics $\widetilde{\gamma}_{\overline{\theta}_\varepsilon}$, $\widetilde{\gamma}_{\overline{\beta}_\varepsilon}$ such that $d_H(\widetilde{\gamma}_{\overline{\theta}_\varepsilon}, \widetilde{\gamma}_{\overline{\beta}_\varepsilon}) \leqslant \varepsilon$. Denote $(p_\varepsilon, v_\varepsilon) = \overline{\theta}_\varepsilon$. By Lemma 3.2 and Corollary 1.1, $\widetilde{\gamma}_{\overline{\theta}_\varepsilon}$ and $\widetilde{\gamma}_{\overline{\beta}_\varepsilon}$ would be Busemann bi-asymptotic to each other and hence, the set

$$H_{(p_\varepsilon, v_\varepsilon)}(0) \cap H_{(p_\varepsilon, -v_\varepsilon)}(0)$$

would contain a point within a very small distance from $(p_\varepsilon, v_\varepsilon) = \overline{\theta}_\varepsilon$. This clearly contradicts the existence of the neighborhood $B_\delta(p_\varepsilon)$. □

4. Visibility and the ideal boundary of $\pi_1(M)$

Recall that $\widetilde{M}$ is a visibility manifold if given $p \in \widetilde{M}$, $\varepsilon > 0$, there exists $L = L(p, \varepsilon) > 0$ such that if the distance from every point of a geodesic segment $[x, y]$ in $\widetilde{M}$ to p is greater than L, then the angle formed by the geodesic segments $[p, x]$ and $[p, y]$ at the point p is less than ε. When L does not depend on p, $\widetilde{M}$ is said to be a uniform visibility manifold. Visibility manifolds were introduced by Eberlein [**6**], and their geometric properties were extensively studied by Eberlein and O'Neill in the 70's (see [**7**] for instance). Visibility manifolds of nonpositive curvature enjoy many properties of negatively curved manifolds. In fact, if M is compact and has nonpositive curvature, then $\widetilde{M}$ is a visibility manifold if and only if $\widetilde{M}$ is a Gromov hyperbolic space. In [**17**] is stated that if M is compact and has no conjugate points, then $\widetilde{M}$ is a visibility manifold if and only if $\widetilde{M}$ is a Gromov hyperbolic space. It is first shown that the visibility property implies Gromov hyperbolicity, however, the proof of the converse statement has a gap based precisely in the (implicit) assumption of the divergence of geodesic rays. In the light of the results in Section 2, what we have is the following:

Lemma 4.1. — *Let M be a compact manifold without conjugate points. Assume that the map $\theta \mapsto H_\theta(0)$ is continuous. If $\widetilde{M}$ is Gromov hyperbolic then $\widetilde{M}$ is a visibility manifold.*

Proof. — We just make a sketch of the proof pointing out the role of the divergence of geodesic rays in the argument. We want to show that the Gromov hyperbolicity of $\widetilde{M}$ implies visibility. Let $\delta > 0$ be such that every geodesic triangle in $\widetilde{M}$ is δ-thin. It is easy to see that there exists $D = D(\delta)$ such that in every geodesic triangle $[x_0, x_1] \cup [x_1, x_2] \cup [x_2, x_0]$ there exists three points $y_i \in [x_i, x_{i+1}]$ (indices taken mod. 3) with $d(y_i, y_{i+1}) \leqslant D$ (see figure 4).

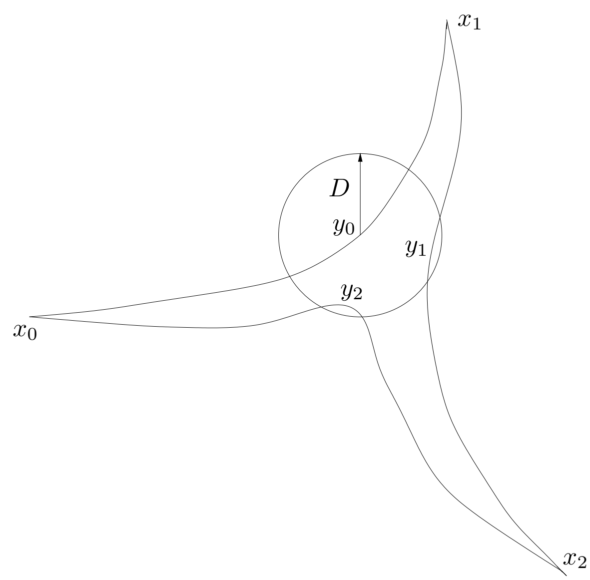

Figure 4

Let us suppose that the distance between x_0 and every point in $[x_1, x_2]$ is greater than $L > 0$. By the triangle inequality, we have that

$$\inf\{d(x_0, y_0), d(x_0, y_2)\} \geqslant L - D.$$

So we have two geodesic rays γ_0, γ_2 starting at $(x_0 = \gamma_0(0) = \gamma_2(0)$, namely, the geodesic rays containing the geodesic segments $[x_0, x_1]$, $[x_0, x_2]$ respectively, having points $y_0 \in \gamma_0$, $y_2 \in \gamma_2$, such that

(1) $d(\gamma_i(0), y_i) \geqslant L - D$, for $i = 0, 2$.
(2) $d(y_0, y_2) \leqslant D$.

The continuity of $\theta \mapsto H_\theta(0)$ implies the uniform divergence of rays in $\widetilde{M}$. So given $\varepsilon > 0$, there exists $T >$ such that if $L - D \geqslant T$, the angle formed by γ_0 and γ_2 at x_0 is less than ε. This clearly implies Lemma 4.1. $\square$

Once we have that $\widetilde{M}$ is a visibility manifold, the theory of Eberlein and O'Neil [**7**] grants the existence of a well defined compactification of $\widetilde{M}$, similar to the compactification of the universal covering of negatively curved manifolds. The boundary of the compactification, called the ideal boundary, is homeomorphic to a sphere of dimension $n-1$, if n is the dimension of M. The action of $\pi_1(M)$ extends to the boundary and the complexity of the dynamics of the action is comparable with the complexity of the actions of Kleinian groups in the sphere. A similar theory is made for Gromov hyperbolic groups. For the definitions and proofs of statements we refer to [**12**], [**10**]. Indeed, the group $\pi_1(M)$ has a compactification as a metric space, via its Cayley graph endowed with the word metric on a finite set of generators. There is an ideal boundary for this compactification, and action of $\pi_1(M)$ induces an action in this ideal boundary. The point is that these two boundaries, the first one of geometric nature and the second one of algebraic nature, are homeomorphic. This is a straightforward consequence of the following fact: the Cayley graph endowed with the word metric and $\widetilde{M}$ are quasi-isometric spaces. Therefore, we have proven the following:

Lemma 4.2. — *Let M be a compact manifold without conjugate points such that the map $\theta \mapsto H_\theta(0)$ is continuous. If $\pi_1(M)$ is Gromov hyperbolic, then its ideal boundary is homeomorphic to a sphere.*

We would expect that Lemma 4.2 holds for every compact manifold without conjugate points, with no extra assumptions on the manifold. It is not known if the $n-1$ sphere is the ideal boundary of a Gromov hyperbolic group covering a compact, n-dimensional manifold, for $n \geqslant 3$ (a good survey of results and conjectures can be found in [**13**]).

5. Is the divergence of geodesic rays equivalent to the continuity of horospheres?

We would like to finish with some remarks about a class of manifolds without conjugate points introduced by Pesin in [**16**]. We say that a manifold M without conjugate points satisfies the so-called Axiom of Asymptoticity, if given any $\theta = (p, v)$ in $T_1\widetilde{M}$, a point $q \in \widetilde{M}$, sequences $\theta_n \to \theta$, $q_n \to q$, and $t_n \to +\infty$, then the sequence of geodesics $[q_n, \gamma_{\theta_n}(t_n)]$ converges to a geodesic β that is asymptotic to γ_θ. Pesin in [**16**] claimed that if M is a compact manifold without conjugate points satisfying the Axiom of Asymptoticity, then the map $\theta \mapsto H_\theta(0)$ is continuous in the sense defined in Section 1. However, the argument is based in Green's result [**11**] about the divergence of geodesic rays in manifolds without conjugate points which has a gap in its proof, as we already mentioned in the introduction. Therefore, what is actually proved in [**16**] is the following:

Lemma 5.1. — *Let M be a compact manifold without conjugate points such that geodesic rays diverge in $\widetilde{M}$. If $\widetilde{M}$ satisfies the Axiom of Asymptoticity, then the map $\theta \mapsto H_\theta(0)$ is continuous.*

According to the results in Section 2 we have,

Corollary 5.1. — *Let M be a compact manifold without conjugate points such that $\widetilde{M}$ satisfies the Axiom of Asymptoticity. Then, geodesic rays diverge in $\widetilde{M}$ if and only if the map $\theta \mapsto H_\theta(0)$ is continuous.*

It is natural to expect that the divergence of geodesic rays implies the continuity of the horospheres. Straighforward generalizations of the proofs of Lemmas 3.1 and 3.2 apply to manifolds satisfying the Axiom of Asymptoticity.

Corollary 5.2. — *Let M be a compact manifold without conjugate points such that the map $\theta \mapsto H_\theta(0)$ is continuous. If $\widetilde{M}$ satisfies the Axiom of Asymptoticity then*

(1) *The relation between geodesics defined by $\gamma\, \boldsymbol{R}\, \beta$ if and only if β is a Busemann asymptote of γ, is an equivalence relation.*

(2) *If γ_α is asymptotic to γ_β, then $b^\theta - b^\alpha$ is constant.*

(3) *A geodesic β is bi-asymptotic to γ_θ if and only if, up to a reparametrization of β, $\beta(0) \in H_{(p,v)}(0) \cap H_{(p,-v)}(0)$.*

References

[1] Ballmann, W., Brin, M., Burns, K.: *On surfaces with no conjugate points.* J. Diff. Geom. 25 (1987), 249-273.

[2] Ballman, W., Gromov, M., Schroeder, V.: *Manifolds of Non-positive curvature.* Boston, Birkhausser 1985.

[3] Busemann, H.: The geometry of geodesics. New York, Academic Press. 1955.

[4] Croke, C., Schroeder, V.: *The fundamental group of manifolds without conjugate points.* Comm. Math. Helv. 61 (1986) 161-175.

[5] Eberlein, P.: *Geodesic flows in certain manifolds with no conjugate points.* Trans. Amer. Math. Soc. 167 (1972) 151-170.

[6] Eberlein, P.: *When is a geodesic flow of Anosov type I.* J. Differ. Geom. 8 (1973), 437-463.

[7] Eberlein, P., O'Neill, B.: *Visibility manifolds.* Pacific Journal of Math. 46, 1 (1973), 45-109.

[8] Eschenburg, J.-H.: Horospheres and the stable part of the geodesic flow. Math. Z. 153 (1977), 237-251.

[9] Freire, A., Mañé, R.: *On the entropy of the geodesic flow in manifolds without conjugate points.* Inv. Math. 69 (1982), 375-392.

[10] Ghys, E., De La Harpe, P.: *Sur les groupes hyperboliques d'apres M. Gromov.* Progress in Math. 83 Birkhausser.

[11] Green, L.: *Geodesic instability.*. Proc. Amer. Math. Soc. 7 (1956), 438-448.

[12] Gromov, M.: *Hyperbolic Groups.* Essays in Group Theory 8, 75-264 S. M. Gersten Editor, Springer-Verlag.

[13] Gromov, M.: *Geometric group theory.* Volume 2: Asymptotic invariants of infinite groups. London Math. Society Lecture Notes Series 182, G. Niblo and M. Roller editors, Cambridge University Press 1993.

[14] Knieper, G.: *Mannifaltigkeiten ohne konjugierte Punkte.* Bonner Mathematische Schriften 168 (1986).

[15] Morse, J. M.: *A fundamental class of geodesics on any closed surface of genus greater than one.* Trans. Amer. Math. Soc. 26 (1924), 25-60.

[16] Pesin, Ja. B.: *Geodesic flows on closed Riemannian manifolds without focal points.* Math. USSR Izvestija, Vol 11, 6 (1977).

[17] Ruggiero R.: *Expansive Dynamics and Hyperbolic Geometry.* Bol. Soc. Brasil. Mat. 25 (1994), 139-172.

[18] Ruggiero, R.: *Expansive geodesic flows in manifolds without conjugate points.* Ergod. Th. Dynam. Sys. 17 (1997), 211-225.

[19] Ruggiero, R.: *Weak stability of the geodesic flow and Preissmann's theorem.* Ergod. Th. Dynam. Sys. 20 (2000), 1231-1251.

R.O. Ruggiero, Pontificia Universidade Católica do Rio de Janeiro, PUC-Rio, Dep. de Matemática, Rua Marqués de São Vicente 225, Gávea, Rio de Janeiro, Brasil • *E-mail :* `rorr@mat.puc-rio.br`

Astérisque
287, 2003, p. 251–272

COMPLEX SCHOTTKY GROUPS

by

José Seade & Alberto Verjovsky

Abstract. — In this work we study a certain type of discrete groups acting on higher dimensional complex projective spaces. These generalize the classical Schottky groups acting on the Riemann sphere. We study the limit sets of these actions, which turn out to be solenoids. We also look at the compact complex manifols obtained as quotient of the region of discontinuity, divided by the action. We determine their topology and the dimension of the space of their infinitesimal deformations. We show that every such deformation arises from a deformation of the embedding of the group in question into the group of automorphisms of the corresponding complex projective space, which is a reminiscent of the classical Teichmüller theory.

Introduction

The theory of Kleinian groups introduced by Poincaré [**Po**] in the 1880's played a major role in many parts of mathematics throughout the 20th century, as for example in Riemann surfaces and Teichmüller theory, automorphic forms, holomorphic dynamics, conformal and hyperbolic geometry, 3-manifolds theory, etc. These groups are, by definition, discrete groups of holomorphic automorphisms of the complex projective line $P^1_{\mathbb{C}}$, whose limit set is not the whole $P^1_{\mathbb{C}}$. Equivalently, these can be regarded as groups of isometries of the hyperbolic 3-space, or as groups of conformal automorphisms of the sphere S^2. Much of the theory of Kleinian groups has been generalised to conformal Kleinian groups in higher dimensions (also called

2000 ***Mathematics Subject Classification.*** — Primary: 32H, 58D, 58E20, 58F23, 53C; Secondary: 20H10, 22E40, 30D05, 30F40, 30F35, 32L25, 32J18, 32G05, 32M05, 57S20, 57S25.
Key words and phrases. — Kleinian groups, Schottky groups, complex manifolds, complex solenoids or laminations, minimal sets, ergodicity, Hausdorff dimension.

Research partially supported by CONACYT (Mexico), projects E-27992 and E-27958, PAPIIT (UNAM, Mexico) IN 104 596 IM, Academia Mexicana de Ciencias and the London Royal Society.

Möbius or *hyperbolic* Kleinian groups), *i.e.*, to discrete groups of conformal automorphisms of the sphere S^n whose limit set is not the whole sphere (see, for instance, [**Ku1, Ku2, Ma1, Su1, Su2, Su3, Su4**]).

Many interesting results about the dynamics of rational maps on $P^1_{\mathbb{C}}$ in the last decades have been motivated by the dynamics of Kleinian groups, and there is an interesting "dictionary" between these two theories (see, for instance, [**Su1, Su2, Su3, Su4, Mc1, Mc2**]). The theory of rational maps has also been generalised to automorphisms of $P^2_{\mathbb{C}}$, and recently many results are being obtained about the dynamics of automorphisms and rational endomorphisms of $P^n_{\mathbb{C}}$ in general. This led us to define in [**SV**] the concept of a *higher dimensional complex Kleinian groups*. By this we meant (infinite) discrete subgroups of $\mathrm{PSL}(n+1,\mathbb{C})$, the group of *holomorphic* automorphisms of $P^n_{\mathbb{C}}$, $n > 1$, acting with a non-empty region of discontinuity.

One of the most interesting families of (conformal) Kleinian groups is provided by the Schottky groups, and the aim of this article is to study the analogous construction for groups acting by holomorphic transformations on complex projective spaces. We call these *Complex Schottky Groups*.

We consider an arbitrary configuration $\{(L_1, M_1), \ldots, (L_r, M_r)\}$ of pairs of projective n-spaces in $P^{2n+1}_{\mathbb{C}}$, which are all of them pairwise disjoint. Given arbitrary neighbourhoods $U_1, \ldots, U_r$ of the L_i's, pairwise disjoint, we show that there exists, for each $i = 1, \ldots, r$, projective transformations T_i of $P^{2n+1}_{\mathbb{C}}$, which interchange the interior with the exterior of a compact tubular neighbourhood N_i of L_i contained in U_i, leaving invariant the boundary $E_i = \partial(N_i)$. The E_i's are *mirrors*, they play the same role in $P^{2n+1}_{\mathbb{C}}$ as circles play in S^2 to define the classical Schottky groups. Each mirror E_i is a $(2n+1)$-sphere bundle over $P^n_{\mathbb{C}}$. The group of automorphisms of $P^{2n+1}_{\mathbb{C}}$ generated by the T_i's is a complex Kleinian group Γ. The region of discontinuity $\Omega(\Gamma)$ is a fibre bundle over $P^n_{\mathbb{C}}$ with fibre S^{2n+2} minus a Cantor set $\mathcal{C}$. The limit set Λ is the complement of $\Omega(\Gamma)$ in $P^{2n+1}_{\mathbb{C}}$; it is the set of accumulation points of the Γ-orbit of the $L_i's$, and it is a product $\mathcal{C} \times P^n_{\mathbb{C}}$. The action of Γ on this set of projective lines is minimal in the sense that the Γ-orbit of every point x_o in $P^{2n+1}_{\mathbb{C}}$ accumulates to (at least a point in) each one of the projective lines in Λ. This set is transversally *projectively self-similar*, i.e., Λ corresponds to a Cantor set in the Grassmannian $G_{2n+1,n}$, which is dynamically-defined. Hence Λ is a *solenoid* (or *lamination*) by projective spaces, which is transversally Cantor and projectively self-similar. Each of these groups Γ contains a subgroup $\check{\Gamma}$ of index two, which is a free group of rank $r-1$ and acts *freely* on $\Omega(\Gamma)$. The quotient $\Omega(\Gamma)/\check{\Gamma}$ is a compact complex manifold, which is a fibre bundle over $P^n_{\mathbb{C}}$ with fibre the connected sum of $(r-1)$ copies of $S^{2n+1} \times S^1$. As mentioned above, these manifolds have a canonical projective structure [**Gu**], *i.e.*, they have an atlas $\{(\mathcal{U}_i, \phi_i)\}$ whose changes of coordinates are restrictions of complex projective transformations. However, these manifolds are never Kähler, due to cohomological reasons. When $n = 1$, the manifolds that we obtain are *Pretzel twistor spaces* in the sense of [**Pe**]; and if the configuration $\{(L_1, M_1), \ldots, (L_r, M_r)\}$ consists of twistor lines of the fibration $p\colon P^3_{\mathbb{C}} \to S^4$, then Γ and $\check{\Gamma}$ descend to conformal Schottky groups on S^4. In this case $\Omega(\Gamma)/\check{\Gamma}$ is the twistor space of the conformally flat manifold $S^4/p(\check{\Gamma})$, which is a Schottky manifold [**Ku2**]; $\Omega(\Gamma)/\check{\Gamma}$ is a *flat twistor space* [**Si**]. We also generalise our construction of Schottky groups to $P^{\infty}_{\mathbb{C}}$, the projectivization of a separable complex infinite dimensional Hilbert space.

We then compare the deformations of our Schottky groups with the deformations of the complex manifolds that one gets as quotients of the action of the group on its region of discontinuity. For this we estimate an upper bound for the Hausdorff dimension of the limit set of the complex Schottky groups. We use this to show that, with the appropriate conditions for the Schottky group $\check{\Gamma}$, the Kuranishi space $\mathfrak{K}$ of versal deformations of the complex manifold $M_{\check{\Gamma}} := \Omega(\check{\Gamma})/\check{\Gamma}$, is smooth near the reference point determined by $M_{\check{\Gamma}}$. Furthermore, we estimate the dimension of $\mathfrak{K}$ and we prove that every infinitesimal deformation of $M_{\check{\Gamma}}$ actually corresponds to an infinitesimal deformation of the group $\check{\Gamma}$ in the projective group $\mathrm{PSL}(2n+2, \mathbb{C})$, in analogy with the classical Teichmüller and moduli theory for Riemann surfaces.

While preparing this article we had very useful conversations with a number of people, and we wish to thank all of them for their support. We are most grateful to François Lescure for very useful comments regarding the algebraic aspects of this paper. Special thanks are also due to Misha Kapovich and Etienne Ghys.

1. Complex Schottky groups

We recall that (in the classical case) the Schottky groups are obtained by considering pairwise disjoint $(n-1)$-spheres $\mathbb{S}_1, \ldots, \mathbb{S}_r$ in S^n, see [**Ma2**]. Each sphere $\mathbb{S}_i$ plays the role of a mirror: it divides S^n in two diffeomorphic components, and one has an involution T_i of S^n interchanging these components, the inversion on $\mathbb{S}_i$. The Schottky group is defined to be the group of conformal transformations generated by these involutions. We are going to make a similar construction on $P^{2n+1}_{\mathbb{C}}$, $n > 0$. (For $n = 0$, if we take $P^0_{\mathbb{C}}$ to be a point, this construction gives the classical Schottky groups on $P^1_{\mathbb{C}}$.)

Consider the subspaces of $\mathbb{C}^{2n+2} = \mathbb{C}^{n+1} \times \mathbb{C}^{n+1}$ defined by $\widehat{L}_0 := \{(a, 0) \in \mathbb{C}^{2n+2}\}$ and $\widehat{M}_0 := \{(0, b) \in \mathbb{C}^{2n+2}\}$. Let $\widehat{S}$ be the involution of $\mathbb{C}^{2n+2}$ defined by $\widehat{S}(a, b) = (b, a)$. This interchanges $\widehat{L}_0$ and $\widehat{M}_0$.

1.1. Lemma. — *Let* $\Phi\colon \mathbb{C}^{2n+2} \to \mathbb{R}$ *be given by* $\Phi(a, b) = |a|^2 - |b|^2$. *Then:*

i) $\widehat{E}_{\widehat{S}} := \Phi^{-1}(0)$ *is a real algebraic hypersurface in* $\mathbb{C}^{2n+2}$ *with an isolated singularity at the origin* 0. *It is embedded in* $\mathbb{C}^{2n+2}$ *as a (real) cone over* $S^{2n+1} \times S^{2n+1}$, *with vertex at* $0 \in \mathbb{C}^{2n+2}$.

ii) $\widehat{E}_{\widehat{S}}$ *is invariant under multiplication by* $\lambda \in \mathbb{C}$*, so it is in fact a complex cone.* $\widehat{E}_{\widehat{S}}$ *separates* $\mathbb{C}^{2n+2} - \{(0,0)\}$ *in two diffeomorphic connected components* U *and* V*, which contain respectively* $\widehat{L_0} - \{(0,0)\}$ *and* $\widehat{M_0} - \{(0,0)\}$*. These two components are interchanged by the involution* $\widehat{S}$*, for which* $\widehat{E}_{\widehat{S}}$ *is an invariant set.*

iii) *Every linear subspace* $\widehat{K}$ *of* $\mathbb{C}^{2n+2}$ *of dimension* $n+2$ *containing* $\widehat{L_0}$ *meets transversally* $\widehat{E}_{\widehat{S}}$ *and* $\widehat{M_0}$*. Therefore a tubular neighbourhood* V *of* $\widehat{M_0} - \{(0,0)\}$ *in* $P^{2n+1}_{\mathbb{C}}$ *is obtained, whose normal disc fibres are of the form* $\widehat{K} \cap V$*, with* $\widehat{K}$ *as above.*

Proof. — The first statement is clear because Φ is a quadratic form with $0 \in \mathbb{C}^{2n+2}$ as unique critical point. Clearly $\widehat{E}_{\widehat{S}}$ is invariant under multiplication by complex numbers, so it is a complex cone. That $\widehat{E}_{\widehat{S}} \cap S^{4n+3} = S^{2n+1} \times S^{2n+1} \subset \mathbb{C}^{2n+2}$, is because this intersection consists of all pairs (x,y) so that $|x| = |y| = 1/\sqrt{2}$. That $\widehat{S}$ leaves $\widehat{E}_{\widehat{S}}$ invariant is obvious, and so is that $\widehat{S}$ interchanges the two components of $\mathbb{C}^{2n+2} - \{(0,0)\}$ determined by $\widehat{E}_{\widehat{S}}$, which must be diffeomorphic because $\widehat{S}$ is an automorphism. Finally, if $\widehat{K}$ is a subspace as in the statement (iii), then $\widehat{K}$ meets transversally $\widehat{E}_{\widehat{S}}$, because through every point in $\widehat{E}_{\widehat{S}}$ there exists an affine line in $\widehat{K}$ which is transverse to $\widehat{E}_{\widehat{S}}$. □

Let S be the linear projective involution of $P^{2n+1}_{\mathbb{C}}$ defined by $\widehat{S}$. Since $\widehat{E}_{\widehat{S}}$ is a complex cone, it projects to a codimension 1 real submanifold of $P^{2n+1}_{\mathbb{C}}$, that we denote by E_S.

1.2. Corollary

i) E_S *is an invariant set of* S*.*

ii) E_S *is a* S^{2n+1}*-bundle over* $P^n_{\mathbb{C}}$*, in fact* E_S *is the sphere bundle associated to the holomorphic bundle* $(n+1)\mathcal{O}_{P^n_{\mathbb{C}}}$*, which is the normal bundle of* $P^n_{\mathbb{C}}$ *in* $P^{2n+1}_{\mathbb{C}}$*.*

iii) E_S *separates* $P^{2n+1}_{\mathbb{C}}$ *in two connected components which are interchanged by* S *and each one is diffeomorphic to a tubular neighbourhood of the canonical* $P^n_{\mathbb{C}}$ *in* $P^{2n+1}_{\mathbb{C}}$*.*

Definition. — We call E_S *the canonical mirror* and S *the canonical involution*.

It is an exercise to show that (1.1) holds in the following more generally setting. Of course one has the equivalent of (1.2) too.

1.3. Lemma. — *Let* λ *be a positive real number and consider the involution*

$$\widehat{S}_\lambda : \mathbb{C}^{n+1} \times \mathbb{C}^{n+1} \longrightarrow \mathbb{C}^{n+1} \times \mathbb{C}^{n+1},$$

given by $\widehat{S}_\lambda(a,b) = (\lambda b, \lambda^{-1} a)$*. Then* $\widehat{S}_\lambda$ *also interchanges* $\widehat{L}_0$ *and* $\widehat{M}_0$*, and the set*

$$\widehat{E}_\lambda = \{(a,b) : |a|^2 = \lambda^2 |b|^2\}$$

satisfies, with respect to $\widehat{S}_\lambda$*, the analogous properties* (i)-(iii) *of (1.1) above.*

We notice that as λ tends to ∞, the manifold E_λ gets thiner and approaches the L_0-axes. Consider now two arbitrary disjoint projective subspaces L and M of dimension n in $P^{2n+1}_{\mathbb{C}}$, and the corresponding linear subspaces $\widehat{L}, \widehat{M}$ of $\mathbb{C}^{2n+2}$. It is clear that $\mathbb{C}^{2n+2} = \widehat{L} \oplus \widehat{M}$ and there is a linear automorphism $\widehat{H}$ of $\mathbb{C}^{2n+2}$ taking $\widehat{L}$ to $\widehat{L}_0$ and $\widehat{M}$ to $\widehat{M}_0$. For every $\lambda \in \mathbb{R}_+$, the automorphism $\widehat{H}^{-1} \circ \widehat{S}_\lambda \circ \widehat{H}$, is an involution that descends to an involution $H^{-1} \circ S_\lambda \circ H$ of $P^{2n+1}_{\mathbb{C}}$ that interchanges L and M. It is clear that one has results analogous to (1.1) and to (1.2). One also has:

1.4. Lemma. — *Let T be a linear projective involution of $P^{2n+1}_{\mathbb{C}}$ that interchanges L and M. Then T is conjugate in* $\mathrm{PSL}(2n+2, \mathbb{C})$ *to the canonical involution S.*

Proof. — Let $\widehat{L}$ and $\widehat{M}$ be linear subspaces of $\mathbb{C}^{2n+2}$ as above. Let $\{l_1, \dots, l_{n+1}\}$ be a basis of $\widehat{L}$. Then $\{l_1, \dots, l_{n+1}, \widehat{T}(l_1), \dots, \widehat{T}(l_{n+1})\}$ is a basis of $\mathbb{C}^{2n+2}$. The linear transformation that sends the canonical basis of $\mathbb{C}^{2n+2} = \mathbb{C}^{n+1} \oplus \mathbb{C}^{n+1}$ to this basis induces a projective transformation which realizes the required conjugation. □

In this paper, *mirrors* in $P^{2n+1}_{\mathbb{C}}$ are, by definition, the images of E_S under the action of $\mathrm{PSL}(2n+2, \mathbb{C})$. A mirror is the boundary of a tubular neighbourhood of a $P^n_{\mathbb{C}}$ in $P^{2n+1}_{\mathbb{C}}$, so it is an S^{2n+1}-bundle over $P^n_{\mathbb{C}}$.

We summarise the previous discussion in the following result.

1.5. Proposition. — *Let $L \cong M \cong P^n_{\mathbb{C}}$ be disjoint projective subspaces of $P^{2n+1}_{\mathbb{C}}$. Then:*

i) *There exist involutions of $P^{2n+1}_{\mathbb{C}}$ that interchange L and M.*

ii) *Each of these involutions has a mirror, i.e., an invariant set $E = E_T \subset P^{2n+1}_{\mathbb{C}}$ which separates $P^{2n+1}_{\mathbb{C}}$ in two connected components which are interchanged by T. Each component is diffeomorphic to a tubular neighbourhood of the canonical $P^n_{\mathbb{C}} \subset P^{2n+1}_{\mathbb{C}}$.*

iii) *Given an arbitrary tubular neighbourhood U of L, we can choose T so that the corresponding mirror E_T is contained in the interior of U.*

In fact one can obviously make stronger the last statement of (1.5):

1.6. Lemma. — *Let L and M be as above. Given an arbitrary constant λ, $0 < \lambda < 1$, we can find an involution T interchanging L and M, with a mirror E such that if U^* is the open component of $P^{2n+1}_{\mathbb{C}} - E$ which contains M and $x \in U^*$, then $d(T(x), L) < \lambda d(x, M)$, where the distance d is induced by the Fubini-Study metric.*

Proof. — The involution $T_\lambda := H^{-1} \circ S_\lambda \circ H$, with H and S_λ as above, satisfies (1.6). □

We notice that the parameter λ in (1.6) gives control upon the degree of expansion and contraction of the generators of the groups, so one can estimate bounds on the Hausdorff dimension of the limit set (see section 2 below).

The previous discussion can be summarized in the following theorem (cf. [**No**]):

1.7. Theorem. — *Let $\mathcal{L} := \{(L_1, M_1), \ldots, (L_r, M_r)\}$, $r > 1$, be a set of r pairs of projective subspaces of dimension n of $P^{2n+1}_{\mathbb{C}}$, all of them pairwise disjoint. Then:*

i) *There exist involutions $T_1, \ldots, T_r$ of $P^{2n+1}_{\mathbb{C}}$, such that each T_i, $i = 1, \ldots, r$, interchanges L_i and M_i, and the corresponding mirrors E_{T_i} are all pairwise disjoint.*

ii) *If we choose the T_i's in this way, then the subgroup of* $\mathrm{PSL}(2n+2, \mathbb{C})$ *that they generate is complex Kleinian.*

iii) *Moreover, given a constant $C > 0$, we can choose the T_i's so that if $T := T_{j_1} \cdots T_{j_k}$ is a* reduced word of length $k > 0$ *(i.e., $j_1 \neq j_2 \neq \cdots \neq j_{k-1} \neq j_k$), then $T(N_i)$ is a tubular neighbourhood of the projective subspace $T(L_i)$ which becomes very thin as k increases: $d(x, T(L_i)) < C\lambda^k$ for all $x \in T(N_i)$, where N_i is the connected component of $P^{2n+1}_{\mathbb{C}} - E_{T_i}$ that contains L_i, for all $i = 1, \ldots, r$.*

1.7.1. Definition. — A Complex Kleinian group constructed as above will be called a *Complex Schottky Group.*

1.7.2. Definition. — Given a Complex Schottky group Γ, we define its *limit set* $\Lambda := \Lambda(\Gamma)$ to be the set of accumulation points of the Γ-orbit of the union $L_1 \cup \cdots \cup L_r$. Its complement $\Omega = \Omega(\Gamma) := P^{2n+1}_{\mathbb{C}} - \Lambda$ is the *region of discontinuity.*

1.7.3. Remark. — We notice that this definition is not standard but it is suitable for Schottky groups.

1.8. Theorem. — *Let Γ be a complex Schottky group in $P^{2n+1}_{\mathbb{C}}$, generated by involutions $\{T_1, \ldots, T_r\}$, $n \geqslant 1$, $r > 1$, as in (1.7) above. Let $\Omega(\Gamma)$ be the region of discontinuity of Γ and let $\Lambda(\Gamma) = P^{2n+1}_{\mathbb{C}} - \Omega(\Gamma)$ be the limit set. Then, one has:*

i) *Let $W = P^{2n+1}_{\mathbb{C}} - \cup_{i=1}^{r} \mathring{N}_i$, where $\mathring{N}_i$ is the interior of the tubular neighbourhood N_i as in (1.7). Then W is a compact fundamental domain for the action of Γ on $\Omega(\Gamma)$. One has: $\Omega(\Gamma) = \bigcup_{\gamma \in \Gamma} \gamma(W)$, and the action on Ω is properly discontinuous.*

ii) *$\Lambda(\Gamma)$ is an intersection of nested sets: $\Lambda(\Gamma) = \cap_{i=1}^{\infty} \gamma_i(N_{j(i)})$, where $\{\gamma_i\}_{i=1}^{\infty}$ is a sequence of distinct elements of Γ and $j \colon \mathbb{N} \to \{1, \ldots, r\}$ is a function such that $\gamma_{i+1}(N_{j(i+1)}) \subset \gamma_i(N_{j(i)})$.*

iii) *If $r = 2$, then $\Gamma \cong \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$, the infinite dihedral group, and $\Lambda(\Gamma)$ is the union of two disjoint projective subspaces L and M of dimension n. In this case we say that Γ is* elementary*, in analogy with Kleinian groups acting on $P^1_{\mathbb{C}}$.*

iv) *If $r > 2$, then $\Lambda(\Gamma)$ is a complex solenoid (lamination), homeomorphic to $P^n_{\mathbb{C}} \times \mathcal{C}$, where $\mathcal{C}$ is a Cantor set. Γ acts minimally on the set of projective subspaces in $\Lambda(\Gamma)$ considered as a closed subset of the Grassmannian $G_{2n+1,n}$.*

v) *If $r > 2$, let $\check{\Gamma} \subset \Gamma$ be the index 2 subgroup consisting of the elements which are reduced words of even length. Then $\check{\Gamma}$ is free of rank $r-1$ and acts freely on $\Omega(\Gamma)$. The compact manifold with boundary $\check{W} = W \cup T_1(W)$ is a fundamental domain for the action of $\check{\Gamma}$ on $\Omega(\Gamma)$. We also call $\check{\Gamma}$ a* complex Schottky group.

vi) *Each element $\gamma \in \check{\Gamma}$ leaves invariant two copies, P_1 and P_2, of $P^n_{\mathbb{C}}$ in $\Lambda(\Gamma)$. For every $L \subset \Lambda(\Gamma)$, $\gamma^i(L)$ converges to P_1 (or to P_2) as $i \to \infty$ (or $i \to -\infty$).*

In fact we prove that if $r > 2$, then Γ acts on a graph whose vertices have all valence either 2 or r. This graph is actually a tree, which can be compactified by adding its "ends". These form a Cantor set and the action of Γ can be extended to this compactification. The limit set $\Lambda(\Gamma)$ corresponds to the uncountable set of ends of this tree. We use this to prove statement v) above.

Proof of i). — Let ∂W be the boundary of $W = P^{2n+1}_{\mathbb{C}} - \cup \mathring{N}_i$, i.e., the union $E_1 \cup \cdots \cup E_r$ of the mirrors. Set $W_0 := W$. Now define $W_1 = \bigcup_{i=0}^r T_i(W)$, where T_0 is the identity, by definition. Then W_1 is a manifold whose boundary consists of $r(r-1)$ components $E_{ij} := T_i(E_j)$, $i \neq j$, $i, j = 1, \ldots, r$, each one being a mirror. Define, by induction on $k > 1$, $W_k = \bigcup_{i=o}^r T_i(W_{k-1})$. Then W_k is a manifold whose boundary consists of $r(r-1)^k$ components, $E_{j_1,\ldots,j_k} := T_{j_1} \cdots Tj_{k-1}(E_{j_k})$, where $j_1, j_2, \ldots, j_k \in \{1, \ldots, r\}$ and $j_1 \neq j_2, \ldots, j_{k-1} \neq j_k$. Thus W_k is contained in the interior of W_{k+1}: $W_k \subset \mathring{W}_{k+1}$.

Let $U = \bigcup_{k=0}^\infty W_k$, so U is Γ-invariant, since $T_j(W_k) \subset W_{k+1}$ for every $j \in \{1, \ldots, r\}$. It is clear that U is open, since any $x \in U$ is contained in the interior of some W_k. Let $\gamma = T_{j_1} \cdots T_{j_k}$ be any element of Γ represented as a reduced word of length $k > 1$. Then $\gamma(W) \subset W_k - \mathring{W}_{k-1}$. Thus, for any $\gamma \neq \beta$, $\gamma(\mathring{W}) \bigcap \beta(\mathring{W}) = \varnothing$. Since $U = \bigcup_{\gamma \in \Gamma} \gamma(W)$, then U is obtained from translates of W, glued along some boundary components. Thus U is open, connected, with a properly discontinuous action of Γ. Therefore $U \subset \Omega(\Gamma)$. To finish the proof of i) we must prove $P^{2n+1}_{\mathbb{C}} - U = \Lambda(\Gamma)$. For this we consider, for each $k \geqslant 0$, the set $F_k := P^{2n+1}_{\mathbb{C}} - \mathring{W}_k$. Then $F_{k+1} \subset F_k$, hence $\bigcap_{k=0}^\infty F_k = P^{2n+1}_{\mathbb{C}} - U$ is a nonempty closed invariant set. For each $k \geqslant 0$, F_k is a disjoint union of closed tubular neighbourhoods of projective subspaces of dimension n of $P^{2n+1}_{\mathbb{C}}$. These are of the form $\gamma(N_i) = T_{j_1} \cdots T_{j_k}(N_i)$, for a $\gamma \in \Gamma$ which is represented in terms of the generators as the reduced word $T_{j_1} \cdots T_{j_k}$. They are closed tubular neighbourhoods of the projective subspace $T_{j_1} \cdots T_{j_k}(L_i)$. For each sequence $\{\gamma_j\}_{j=1}^\infty$ in Γ, such that the length of γ_{j+1} is bigger than the length of γ_j and $\gamma_{j+1}(N_i) \subset \gamma_j(N_i)$, the tubular neighbourhood becomes thinner. By (1.7), the sequence $\{\gamma_j(L_i)\}_{j=1}^\infty$ converges, in the Hausdorff metric, to a linear subspace of dimension n. Hence, also by (1.7), $P^{2n+1}_{\mathbb{C}} - U$ is a nowhere dense closed subset of $P^{2n+1}_{\mathbb{C}}$, which is a disjoint union of projective subspaces of dimension n. Thus U is open and dense in $P^{2n+1}_{\mathbb{C}}$; since $U \subset \Omega(\Gamma)$, it follows that $\Omega(\Gamma)$ is also connected. We have that U/Γ is compact and it is obtained from the compact fundamental domain W after identifications in each component of its boundary. If $\Omega(\Gamma) \neq U$ we arrive to a contradiction, because Ω/Γ is connected and U/Γ is open, compact and properly contained in Ω/Γ. Therefore, $\Omega(\Gamma) = U$ and $\Lambda(\Gamma) = \bigcap_{i=0}^\infty F_i$. This proves i).

Proof of ii). — If $x \in \Lambda(\Gamma)$ then, as shown above, $x \in \bigcap_{i=0}^{\infty} F_i$. To prove ii) it is sufficient to choose, for each i, the component of F_i which contains x. Such component is of the form $\gamma(N_j)$ for a *unique* $\gamma \in \Gamma$ (we set $\gamma = \gamma_i$) and a *unique* $j \in \{1, \dots, r\}$. We set $j = j(i)$. This proves ii). This also shows that $\bigcap_{i=0}^{\infty} F_i$ is indeed the limit set according to Kulkarni's definition in [**Ku1**].

Proof of iii). — We have two involutions, T and S, and two neighbourhoods, N_T and N_S, whose boundaries are the mirrors of T and S, respectively. The limit set is the disjoint union $A \cup B$, where $A := \bigcap_{\gamma \in \Gamma'} \gamma(N_S)$, $B := \bigcap_{\gamma \in \Gamma''} \gamma(N_T)$, Γ' is the set of elements in Γ which are words ending in T and Γ'' is the set of elements which are words ending in S. By (1.7), A and B are each the intersection of a nested sequence of tubular neighbourhoods of projective subspaces of dimension n, whose intersection is a projective subspace of dimension n. Hence A and B are both projective subspaces of dimension n, and they are disjoint. Two reduced words ending in T and S, act differently on N_T (or N_S). Hence Γ is the free product of the groups generated T and S, proving iii).

Proof of iv). — Let $L \subset P_{\mathbb{C}}^{2n+1}$ be a subspace of dimension n and let N be a closed tubular neighbourhood of L as above. Let D be a closed disc which is an intersection of the form $\widehat{L} \cap N$, where $\widehat{L}$ is a subspace of complex dimension $n+1$, transversal to L. If M is a subspace of dimension n contained in the interior of N, then M is transverse to D, otherwise the intersection of M with $\widehat{L}$ would contain a complex line and M would not be contained in N. From the proofs of i) and ii) we know that $\Lambda(\Gamma)$ is the disjoint union of uncountable subspaces of dimension n. Let $x \in \Lambda(\Gamma)$ and let $L \subset \Lambda(\Gamma)$ be a projective subspace with $x \in L$. Let N be a tubular neighbourhood of L and D a transverse disc as above. Then $\Lambda(\Gamma) \cap D$ is obtained as the intersection of families of discs of decreasing diameters, exactly as in the construction of Cantor sets. Therefore $\Lambda(\Gamma) \cap D$ is a Cantor set and $\Lambda(\Gamma)$ is a solenoid (or lamination) by projective subspaces which is transversally Cantor. It follows that $\Lambda(\Gamma)$ is a fibre bundle over $P_{\mathbb{C}}^n$, with fibre a Cantor set $\mathcal{C}$. Since $P_{\mathbb{C}}^n$ is simply connected and $\mathcal{C}$ is totally disconnected, this fibre bundle must be trivial, hence the limit set is a product $P_{\mathbb{C}}^n \times \mathcal{C}$, as stated.

There is another way to describe the above construction: Γ acts, via the differential, on the Grassmannian $G_{2n+1,n}$ of projective subspaces of dimension n of $P_{\mathbb{C}}^{2n+1}$. This action also has a region of discontinuity and contains a Cantor set which is invariant. This Cantor set corresponds to the closed family of disjoint projective subspaces in $\Lambda(\Gamma)$. It is clear that the action on the Grassmannian is minimal on this Cantor set.

Proof of v). — Choose a point x_0 in the interior of W. Let Γ_{x_0} be the Γ-orbit of x_0. We construct a graph $\check{\mathcal{G}}$ as follows: to each $\gamma(x_0) \in \Gamma_{x_0}$ we assign a vertex v_γ. Two vertices $v_\gamma, v_{\gamma'}$ are joined by an edge if $\gamma(W)$ and $\gamma'(W)$ have a common boundary component, which corresponds to a mirror E_i. This means that γ' is γ followed by an involutions T_i or vice-versa. This graph can be realized geometrically by joining the corresponding points $\gamma(x_0), \gamma'(x_0) \in \Omega(\Gamma)$ by an arc $\alpha_{\gamma,\gamma'}$ in $\Omega(\Gamma)$, which is chosen to be transversal to the corresponding boundary component of $\gamma(W)$; we also choose these arcs so that no two of them intersect but at the extreme points. Clearly $\check{\mathcal{G}}$ is a tree and each vertex has valence r. To construct a graph $\mathcal{G}$ with an appropriate Γ-action we introduce more vertices in $\check{\mathcal{G}}$: we put one vertex at the middle point of each edge in $\check{\mathcal{G}}$; these new vertices correspond to the points where the above arcs intersect the boundary components of $\gamma(W)$. Then we have an obvious simplicial action of Γ on $\mathcal{G}$. Let $\check{\Gamma}$ be the index-two subgroup of Γ consisting of elements which can be written as reduced words of even length in terms of $T_1, \dots, T_r$. A fundamental domain for $\check{\Gamma}$ in $\Omega(\Gamma)$ is $\check{W} = W \cup T_1(W)$, so this group acts freely on the vertices of $\check{\mathcal{G}}$. Hence $\check{\Gamma}$ is a free group of rank $r-1$. The tree $\check{\mathcal{G}}$ can be compactified by its ends by adding a Cantor set on which $\check{\Gamma}$ acts minimally; this corresponds to the fact that Γ acts minimally on the set of projective subspaces which constitute $\Lambda(\Gamma)$.

Proof of vi). — By (1.7), if $\gamma \in \check{\Gamma}$, then either $\gamma(N_1)$ is contained in N_1 or $\gamma^{-1}(N_1)$ is contained in N_1; say $\gamma(N_1)$ is contained in N_1. Thus $\{\gamma^i(N_1)\}$, $i > 0$, is a nested sequence of tubular neighbourhoods of projective subspaces whose intersection is a projective subspace P_1 of dimension n; $\{\gamma^i(N_1)\}$, $i < 0$, is also nested sequence of tubular neighbourhoods of projective subspaces whose intersection is a projective subspace P_2 of dimension n. For every $L \subset \Lambda(\Gamma)$, $\gamma^i(L)$ converges to P_1 and P_2 as $i \to \infty$ or $i \to -\infty$, respectively, and both P_1 and P_2 are invariant under γ, as claimed. $\square$

1.9. Remarks

i) The action of $\check{\Gamma}$ in the Cantor set of projective subspaces is analogous to the action of a classical Fuchsian group of the second kind on its Cantor limit set. We also observe that, since each involution T_i is conjugate to the canonical involution defined in lemma 1.1, the laminations obtained in theorem 1.8 are transversally *projectively self-similar*. Hence one could try to apply results analogous to the results for (conformally) self-similar sets (for instance Bowen's formula [**Bo**]) to estimate the transverse Hausdorff dimension of the laminations obtained. Here by *transverse Hausdorff dimension* we mean the Hausdorff dimension of the Cantor set $\mathcal{C}$ of projective subspaces of $G_{2n+1,n}$ which conform the limit set. If $\widetilde{T}_i$, $i = 1, \dots, r$, denote the maps induced in the Grassmannian $G_{2n+1,n}$ by the linear projective transformations T_i, then $\mathcal{C}$ is dynamically-defined by the group generated by the set $\{\widetilde{T}_i\}$.

ii) The construction of Kleinian groups given in 1.8 actually provides families of Kleinian groups, obtained by varying the size of the mirrors that bound tubular neighbourhoods around the $L_i' s$. In Section 3 below we will look at these families.

iii) The above construction of complex Kleinian groups, using involutions and mirrors, can be adapted to produce discrete groups of automorphisms of quaternionic projective spaces of odd (quaternionic) dimension. Every "quaternionic Kleinian group" on $P_{\mathcal{H}}^{2n+1}$ lifts canonically to a complex Kleinian group on $P_{\mathbb{C}}^{4n+3}$.

2. Quotient Spaces of the region of discontinuity

We now discuss the nature of the quotients $\Omega(\Gamma)/\Gamma$ and $\Omega(\Gamma)/\check{\Gamma}$, for the groups of section 1. The proof of proposition (2.1) is straightforward and is left to the reader.

2.1. Proposition. — *Let L be a copy of the projective space $P^n_{\mathbb{C}}$ in $P^{2n+1}_{\mathbb{C}}$ and let x be a point in $P^{2n+1}_{\mathbb{C}} - L$. Let $K_x \subset P^{2n+1}_{\mathbb{C}}$ be the unique copy of the projective space $P^{n+1}_{\mathbb{C}}$ in $P^{2n+1}_{\mathbb{C}}$ that contains L and x. Then K_x intersects transversally every other copy of $P^n_{\mathbb{C}}$ embedded in $P^{2n+1}_{\mathbb{C}} - L$, and this intersection consists of one single point. Thus, given two disjoint copies L and M of $P^n_{\mathbb{C}}$ in $P^{2n+1}_{\mathbb{C}}$, there is a canonical projection map*

$$\pi := \pi_L \colon P^{2n+1}_{\mathbb{C}} - L \longrightarrow M,$$

which is a (holomorphic) submersion. Each fibre $\pi^{-1}(x)$ is diffeomorphic to $\mathbb{R}^{2n+2}$.

2.2. Theorem. — *Let Γ be a complex Schottky group as in theorem 1.8, with $r > 2$. Then:*

i) *The fundamental domain W of Γ is (the total space of) a locally trivial differentiable fibre bundle over $P^n_{\mathbb{C}}$ with fibre $S^{2n+2} - \mathring{D}_1 \cup \cdots \cup \mathring{D}_r$, where each $\mathring{D}_i$ is the interior of a smooth closed $(2n+2)$-disc D_i in S^{2n+2} and the D_i's are pairwise disjoint.*

ii) $\Omega(\Gamma)$ *fibres differentiably over $P^n_{\mathbb{C}}$ with fibre S^{2n+2} minus a Cantor set.*

iii) *If $\check{\Gamma}$ is the subgroup of index two as in theorem 1.8, which acts freely on $\Omega(\Gamma)$, then $\Omega(\Gamma)/\check{\Gamma}$ is a compact complex manifold that fibres differentiably over $P^n_{\mathbb{C}}$ with fibre $(S^{2n+1} \times S^1)\# \cdots \#(S^{2n+1} \times S^1)$, the connected sum of $r-1$ copies of $S^{2n+1} \times S^1$.*

Proof of i). — Let $P_1, P_2 \subset \Lambda(\Gamma)$ be two disjoint projective subspaces of dimension n contained in $\Lambda(\Gamma) \subset P^{2n+1}_{\mathbb{C}}$. Since $\Omega(\Gamma)$ is open in P^{2n+1}, the restriction to $\Omega(\Gamma)$ of the map π given by 2.1, using P_1 as L and P_2 as M, is a holomorphic submersion. We know, by theorem 1.8.iv, that $\Lambda(\Gamma)$ is a compact set which is a *disjoint* union of projective subspaces of dimension n and which is a transversally Cantor lamination. By 2.1, for each $y \in P_2$, K_y meets transversally each of these projective subspaces (in other words, K_y is transverse to the lamination $\Lambda(\Gamma)$, outside P_1). Hence, by theorem 1.8, for each $y \in P_2$, K_y intersects $\Lambda(\Gamma) - P_1$ in a Cantor set minus one point (this point corresponds to P_1). The family of subspaces K_y of dimension $n+1$ are all transverse to P_2.

Let us now choose P_1 and P_2 as in 1.8.vi, so they are invariant sets for some $\gamma \in \check{\Gamma}$, and $\gamma^j(L)$ converges to P_2 as $j \to \infty$ for every projective n-subspace $L \subset \Lambda(\Gamma) - P_1$. We see that *every* mirror E_i, $i \in \{1, \ldots, r\}$ is transverse to all K_y. Hence the restriction

$$\pi_1 := \pi_{P_1}|_W \colon W \longrightarrow P_2 \cong P^n_{\mathbb{C}},$$

of π to W, is a submersion which restricted to each component of the boundary is also a submersion. For each $y \in P_2$ one has $\pi_1^{-1}(\{y\}) = K_y \cap W$, so $\pi_1^{-1}(\{y\})$ is compact. Thus π_1 is the projection of a locally trivial fibre bundle with fibres $K_y \cap W$, $y \in P_2$, by Ehresmann's lemma [**Eh**]. On the other hand for a fixed $y_0 \in P_2$, $K_{y_0} \cap W$ is a closed $(2n+2)$-disc with $r-1$ smooth closed $(2n+2)$-discs removed from its interior. This is true because P_1 is contained in exactly one of the $N_i's$, say N_1, the tubular neighbourhood of P_1, and K_{y_0} intersects each N_j, $j \neq 1$, in a smooth closed $(2n+2)$-disc. This proves i).

Proof of ii). — The above arguments show that for each $\overline{\gamma} \in \Gamma$, the image $\overline{\gamma}(E_i)$ of a mirror E_i is transverse to K_y for all $y \in P_2$ and $i \in \{1, \ldots, r\}$. Hence the restriction $\pi_1^k := \pi_{P_1}|_{W_k}$, where W_k is as above, is a submersion whose restriction to each boundary component of W_k is also a submersion. Thus π_1^k is a locally trivial fibration. Since $\Omega(\Gamma) = \bigcup_{k \geqslant 0} W_k$, we finish the proof of the first part of ii) by applying the slight generalisation below of Ehresmann fibration lemma [**Eh**]; we leave the proof to the reader.

Lemma. — *Let $\mathcal{M} = \bigcup_{i=1}^{\infty} \mathcal{N}_i$ be a smooth manifold which is the union of compact manifolds with boundary $\mathcal{N}_i$, so that each $\mathcal{N}_i$ is contained in the interior of $\mathcal{N}_{i+1}$. Let $\mathcal{L}$ be a smooth manifold and $f : \mathcal{M} \to \mathcal{L}$ a submersion whose restriction to each boundary component of $\mathcal{N}_i$, for every i, is also a submersion. Then f is a locally trivial fibration.*

Thus $\pi_{P_1} : \Omega(\Gamma) \to P_2 \cong P^n_{\mathbb{C}}$ is a holomorphic submersion which is a locally trivial differentiable fibration. To finish the proof of ii) we only need to show that the fibres of π_{P_1} are S^{2n+2} minus a Cantor set. Just as above, one shows that $K_y \cap W_k$ is diffeomorphic to the sphere S^{2n+2} minus the interior of $r(r-1)^k$ disjoint $(2n+2)$-discs. Therefore the fibre of π_{P_1} at y, which is $K_y \cap \Omega(\Gamma)$, is the intersection of S^{2n+2} minus a nested union of discs, which gives a Cantor set as claimed in ii).

Proof of iii). — We recall that by theorem 1.8.v, the fundamental domain of $\check{\Gamma}$ is the manifold $\check{W} = W \cup T_1(W)$. Then, as above, the restriction of π to $\check{W}$ is a submersion which is also a submersion in each connected component of the boundary:

$$\partial \check{W} = \Big(\bigcup_{j \neq 1} T_1(E_j)\Big) \bigcup_{j \neq 1} E_j,$$

which is the disjoint union of the $r-1$ mirrors E_j, $j \neq 1$, together with the mirrors $E_{1j} := T_1(E_j)$, $j \neq 1$. The mirror E_j is identified with E_{1j}, $j \neq 1$, by T_1, and $\Omega(\Gamma)/\check{\Gamma}$ is obtained through these identifications. Let $\check{\pi} : \check{W} \to P_2$ be the restriction of π to $\check{W}$. By the proof of i), $\check{\pi}^{-1}(y) = K_y \cap \check{W}$, $y \in P_2$, is diffeomorphic to S^{2n+2} minus the interior of $2(r-1)$ disjoint $(2n+2)$-discs. The restriction of $\check{\pi}$ to each E_j and E_{1j} determines fibrations $\check{\pi}_j : E_j \to P_2$ and $\check{\pi}_{1j} : E_{1j} \to P_2$, respectively, whose fibres are S^{2n+1}. Set $\widehat{\pi}_j := \check{\pi}_{1j} \circ (T_1|_{E_j})$. If we had that $\widehat{\pi}_j = \check{\pi}_j$ for all $j = 2, \ldots, r$, then we would have a fibration from $\check{W}/\check{\Gamma}$ to P_2, because we would have compatibility of the projections on the boundary. In fact we only need that $\widehat{\pi}_j$ and $\check{\pi}_j$

be homotopic through a smooth family of fibrations $\pi_t : E_{1j} \to P_2$, $\pi_1 = \widehat{\pi}_j$, $\pi_0 = \check{\pi}_j$, $t \in [0,1]$. Actually, to be able to glue well the fibrations at the boundary we need that $\pi_t = \check{\pi}_j$ for t in a neighbourhood of 0 and $\pi_t = \widehat{\pi}_j$ for t in a neighbourhood of 1. But this is almost trivial: $\check{\pi}_j : E_{1j} \to P_2$ is the projection of E_{1j} onto P_2 from P_1 and $\widehat{\pi} - j$ is the projection of E_{1j} from $T(P_1)$ onto P_2. The n-dimensional subspaces P_1 and $T(P_1)$ are disjoint from P_2, so there exists a smooth family of n-dimensional subspaces P_t, $t \in [0,1]$, such that the family is disjoint from P_2 and $P_t = P_1$ for t in a neighbourhood of 0 and $P_t = T(P_1)$ for t in a neighbourhood of 1. We can choose the family so that for each $t \in [0,1]$, the set of $n+1$ dimensional subspaces which contain P_t meet transversally E_{1j}. To achieve this we only need to take an appropriate curve in the Grassmannian of projective n-planes in $P_{\mathbb{C}}^{2n+1}$, consisting of a family P_t which is transverse to all K_y; this is possible by (2.1) and the fact that the set of n-dimensional subspaces which are *not* transverse to the $K_y' s$, is a proper algebraic variety of $P_{\mathbb{C}}^{2n+1}$. In this way we obtain the desired homotopy. Hence $\check{W}$ fibres over $P_2 \cong P_{\mathbb{C}}^n$; the fibre is obtained from S^{2n+2} minus the interior of $2(r-1)$ disjoint $(2n+2)$-discs whose boundaries are diffeomorphic to S^{2n+1} and are identified by pairs by diffeomorphisms which are *isotopic* to the identity (using a fixed diffeomorphism to S^{2n+1}). Hence the fibre is *diffeomorphic* to $(S^{2n+1} \times S^1)\# \cdots \#(S^{2n+1} \times S^1)$, the connected sum of $r-1$ copies of $S^{2n+1} \times S^1$. This proves iii). ☐

2.3. Theorem. — *Let M_Γ be the compact complex orbifold $M_\Gamma := \Omega(\Gamma)/\Gamma$, which has complex dimension $(2n+1)$. Then:*

i) *The singular set of M_Γ, $\mathrm{Sing}(M_\Gamma)$, is the disjoint union of r submanifolds analytically equivalent to $P_{\mathbb{C}}^n$, one contained in (the image in M_Γ of) each mirror E_i of Γ.*

ii) *Each component of $\mathrm{Sing}(M_\Gamma)$ has a neighbourhood homeomorphic to the normal bundle of $P_{\mathbb{C}}^n$ in $P_{\mathbb{C}}^{2n+1}$ modulo the involution $v \mapsto -v$, for v a normal vector.*

iii) *M_Γ fibres over $P_{\mathbb{C}}^n$ with fibre a real analytic orbifold with r singular points, each having a neighbourhood (in the fibre) homeomorphic to the cone over the real projective space $P_{\mathbb{R}}^{2n+1}$.*

Proof. — We notice that M_Γ is obtained from the fundamental domain W after an identification on the boundary E_j by the action of T_j. The singular set of M_Γ is the union of the images, under the canonical projection $p : \Omega(\Gamma) \to \Omega(\Gamma)/\Gamma$, of the fixed point sets of the r involutions T_j. Now, T_j is conjugate to the canonical involution S of (1.2). The lifting of S to $\mathbb{C}^{2n+2} = \mathbb{C}^{n+1} \times \mathbb{C}^{n+1}$ has as fixed point set the $(n+1)$-subspace $\{(a,a) : a \in \mathbb{C}^{n+1}\}$. This projectivizes to a n-dimensional projective subspace. Since we can assume, *for a fixed j*, that T_j is an isometry, we obtain the local structure of a neighbourhood of each component of the singular set. The same arguments as in theorem 2.2.iii prove that $\Omega(\Gamma)/\Gamma$ fibres over $P_{\mathbb{C}}^n$ and that the fibre has r singular points, corresponding to the r components of $\mathrm{Sing}(M_\Gamma)$, and each of these r points has a neighbourhood (in the fibre) homeomorphic to the cone over $P^{2n+1}_{\mathbb{R}}$. □

2.4. Remarks

i) The map π in (2.2.ii) is holomorphic, but the fibration is not holomorphically locally trivial, because the complex structure on the fibres may change.

ii) The Kleinian groups of 2.2 provide a method for constructing complex manifolds which is likely to produce interesting examples (cf. [**No, Ka1, Ka2, Ka3, Ka4, Pe, Si**]). These are never Kähler, because the fibration $\pi : \Omega(\Gamma)/\check{\Gamma} \to P^n_{\mathbb{C}}$ has a section, by dimensional reasons, so there can not exist a 2-cocycle with a power which is the fundamental class of $\Omega(\Gamma)/\check{\Gamma}$. The bundle $(n+1)\mathcal{O}_{P^n_{\mathbb{C}}}$ is nontrivial as a real bundle, because it has non-vanishing Pontryagin classes (except for $n=1$), hence π is a nontrivial fibration. We notice that the fundamental group of a compact Riemann surface of genus greater than zero is never a free group; similarly, by Kodaira's classification, the only compact complex surface with non trivial free fundamental group is the Hopf surface $S^3 \times S^1$. Our examples above give compact complex manifolds with free fundamental groups (of arbitrarly high rank) in all odd dimensions greater than one. Multiplying these examples by $P^1_{\mathbb{C}}$, one obtains similar examples in all even dimensions. As pointed out by the referee, it would be interesting to know if there are other examples which are minimal, *i.e.*, they are not obtained by blowing up along a smooth subvariety of the examples above. It is natural to conjecture that our examples in odd dimensions are the only ones which have a projective structure and free fundamental group of rank greater than one.

iii) The manifolds obtained by resolving the singularities of the orbifolds in (2.3) have very interesting topology. We recall that the orbifold M_Γ is singular along r disjoint copies of $P^n_{\mathbb{C}}$: $S_1, \ldots, S_r$. The resolution $\widetilde{M}_\Gamma$ of M_Γ is obtained by a monoidal transformation along each S_i, and it replaces each point $x \in S_i$, $1 \leqslant i \leqslant r$ by a projective space $P^n_{\mathbb{C}}$. Hence, if $\mathcal{P} : \widetilde{M} \to M$ denotes the resolution map, then $\mathcal{P}^{-1}(S_i)$ is a non-singular divisor in $\widetilde{M}$, which fibres holomorphically over $P^n_{\mathbb{C}}$ with fibre $P^n_{\mathbb{C}}$, $1 \leqslant i \leqslant r$.

2.5. Symmetric products of classical Kleinian groups. — Let $\Gamma \subset \mathrm{PSL}(2,\mathbb{C})$ be a classical Kleinian group acting on $P^1_{\mathbb{C}}$. Let $\Lambda(\Gamma)$ and $\Omega(\Gamma) := P^1_{\mathbb{C}} - \Lambda(\Gamma)$ be, respectively, the limit set and the region of discontinuity of Γ. Since $P^n_{\mathbb{C}}$ is the n^{th} symmetric product of $P^1_{\mathbb{C}}$, $P^n_{\mathbb{C}} \cong S^n(P^1_{\mathbb{C}})$, there is a canonical diagonal action of Γ on $P^n_{\mathbb{C}}$, for all $n > 1$. The group Γ acts properly and discontinuously on $\Omega^n := P^n_{\mathbb{C}} - S^n(\Lambda(\Gamma))$. In particular, if Γ is a Schottky group of the second kind acting in $P^1_{\mathbb{C}}$ whose limit set $\Lambda(\Gamma)$ is a Cantor set, then $S^n(\Lambda(\Gamma))$ is again a Cantor set, and the action of Γ on its complement is discontinuous. Every point in $S^n(\Lambda(\Gamma))$ is an accumulation point of orbits of Γ. This provides examples of complex Klenian groups acting on $P^n_{\mathbb{C}}$ whose limit sets are Cantor sets. If in addition, the quotient of the action of Γ in $P^1_{\mathbb{C}}$ in the region of discontinuity is compact, then Ω^n/Γ is also compact.

3. Hausdorff dimension and moduli spaces

Let $\mathcal{L} := \{(L_1, M_1), \dots, (L_r, M_r)\}$ be a configuration of $P^n_{\mathbb{C}}$'s in $P^{2n+1}_{\mathbb{C}}$ as before, $r > 2$. Let Γ and Γ' be complex Schottky groups obtained from this same configuration, i.e., they are generated by sets $\{T_1, \dots, T_r\}$ and $\{T'_1, \dots, T'_r\}$ of holomorphic involutions of $P^{2n+1}_{\mathbb{C}}$ that interchange the L_i's with the M_i's. For each $i = 1, \dots, r$, the composition $T'_i \circ T_i^{-1}$ preserves the subspaces L_i, M_i. It is an exercise to see that the subgroup of $\mathrm{PSL}(n+2, \mathbb{C})$ of transformations that preserve these subspaces is the projectivization of a copy of $\mathrm{GL}(n+1, \mathbb{C}) \times \mathrm{GL}(n+1, \mathbb{C}) \subset \mathrm{GL}(2n+2, \mathbb{C})$. Therefore, we can always find an analytic family $\{\Gamma_t\}$, $0 \leqslant t \leqslant 1$, of complex Schottky groups, with configuration $\mathcal{L}$, such that $\{\Gamma_0\} = \Gamma$ and $\{\Gamma_1\} = \Gamma'$. Furthermore, let $\mathcal{L} := \{(L_1, M_1), \dots, (L_r, M_r)\}$ and $\mathcal{L}' := \{(L'_1, M'_1), \dots, (L'_r, M'_r)\}$ be two configurations of $P^n_{\mathbb{C}}$'s in $P^{2n+1}_{\mathbb{C}}$ as before. Due to dimensional reasons, we can always move these configurations to obtain a differentiable family of pairs of disjoint n-dimensional subspaces $\{(L_{1,t}, M_{1,t}), \dots, (L_{r,t}, M_{r,t})\}$, with $0 \leqslant t \leqslant 1$, providing an isotopy between $\mathcal{L}$ and $\mathcal{L}'$. Thus one has a differentiable family Γ_t of complex Kleinian groups, where $\Gamma_0 = \Gamma$ and $\Gamma_1 = \Gamma'$. The same statements hold if we replace Γ and Γ' by their subgroups $\check{\Gamma}$ and $\check{\Gamma}'$, consisting of words of even length. So one has a differentiable family $\check{\Gamma}_t$ of Kleinian groups, where $\check{\Gamma}_0 = \check{\Gamma}$ and $\check{\Gamma}_1 = \check{\Gamma}'$. Hence the manifolds $\Omega(\Gamma_t)/\check{\Gamma}_t$ are all diffeomorphic. By section 2, these manifolds are (in general nontrivial) fibre bundles over $P^n_{\mathbb{C}}$ with fibre $\#^{(r-1)}(S^{2n+1} \times S^1)$, a connected sum of $(r-1)$-copies of $S^{2n+1} \times S^1$. If $n = 1$, given any configuration of r pairwise disjoint lines in $P^3_{\mathbb{C}}$, there exist an isotopy of $P^3_{\mathbb{C}}$ which carries the configuration into a family of r twistor lines. Hence $P^3_{\mathbb{C}}$ minus this configuration is diffeomorphic to the Cartesian product of S^4 minus r points with $P^1_{\mathbb{C}}$. Moreover, the attaching functions that we use to glue the boundary components of W, the fundamental domain of Γ, are all isotopic to the identity, because they live in $\mathrm{PSL}(4, \mathbb{C})$, which is connected. Thus, if $n = 1$, then $\Omega(\Gamma_t)/\check{\Gamma}_t$ is diffeomorphic to a product $P^1_{\mathbb{C}} \times \#^{(r-1)}(S^3 \times S^1)$. Hence we have:

3.1. Proposition. — *The differentiable type of the compact (complex) manifold $\Omega(\Gamma_t)/\check{\Gamma}_t$ is independent of the choice of configuration. It is a manifold of real dimension $(4n+2)$, which is a fibre bundle over $P^n_{\mathbb{C}}$ with fibre $\#^{(r-1)}(S^{2n+1} \times S^1)$; moreover, this bundle is trivial if $n = 1$. We denote the corresponding manifold by M^n_r.*

The fact that the bundle is trivial when $n = 1$ is interesting because, as pointed out in the introduction, when the configuration $\mathcal{L}$ consists of twistor lines in $P^3_{\mathbb{C}}$, the quotient $\Omega(\Gamma)/\check{\Gamma}$ is the twistor space of the conformally flat manifold $p(\Omega(\Gamma))/p(\check{\Gamma})$, which is a connected sum of the form $\#^{(r-1)}(S^3 \times S^1)$. Hence, in this case the natural fibration goes the other way round, i.e., it is a fibre bundle over $\#^{(r-1)}(S^3 \times S^1)$ with fibre $P^1_{\mathbb{C}}$.

Given a configuration $\mathcal{L}$ as above, let us denote by $[\mathcal{L}]_G$ its orbit under the action of the group $G = \mathrm{PSL}(2n+2, \mathbb{C})$. These orbits are equivalence classes of such configurations. Let us denote by $\mathcal{C}^n_r$ the set of equivalence classes of configurations consisting of r pairs of $P^n_{\mathbb{C}}$'s as above. Then $\mathcal{C}^n_r$ is a Zariski open set of the moduli space $\mathfrak{M}^n_r$, of configurations of r unordered couples of projective subspaces of dimension n in $P^{2n+1}_{\mathbb{C}}$, which is obtained as the Mumford quotient [**MFK**] of the action of G on such configurations. By [**MFK**], $\mathcal{C}^n_r$ is a complex algebraic variety: the *moduli space of configurations* of r pairs of n-planes $P^n_{\mathbb{C}}$ in $P^{2n+1}_{\mathbb{C}}$. Similarly, we denote by $\mathfrak{G}^n_r$ the equivalence classes, or moduli space, of the corresponding Schottky groups, where two such groups are equivalent if they are conjugate by an element in $\mathrm{PSL}(n+2, \mathbb{C})$. Given $\mathcal{L} := \{(L_1, M_1), \ldots, (L_r, M_r)\}$, and r-tuples of involutions $(T_1, \ldots, T_r)$ and $(S_1, \ldots, S_r)$ as above, i.e., interchanging L_i with M_i for all $i = 1, \ldots, r$ and having pairwise disjoint mirrors, we say that these r-tuples are *equivalent* if there exists $h \in G$ such that $hT_ih^{-1} = S_i$ for all i. Let $\mathfrak{T}_{\mathcal{L}}$ denote the set of equivalence classes of such r-tuples of involutions. It is clear that a conjugation h as above must leave $\mathcal{L}$ invariant. Hence, if r is big enough with respect to n, then h must be actually the identity, so the equivalence classes in fact consist of a single element.

3.2. Theorem. — *There exists a holomorphic surjective map $\pi\colon \mathfrak{G}^n_r \to \mathcal{C}^n_r$ which is a C^∞ locally trivial fibration with fibre $\mathfrak{T}_{\mathcal{L}}$. Furthermore, let Γ, Γ' be complex Schottky groups as above and let $\Omega(\Gamma), \Omega(\Gamma')$ be their regions of discontinuity. Then the complex orbifolds $M_\Gamma := \Omega(\Gamma)/\Gamma$ and $M_{\Gamma'} := \Omega(\Gamma')/\Gamma'$ are biholomorphically equivalent if and only if Γ and Γ' are projectively conjugate, i.e., they represent the same element in $\mathfrak{G}^n_r$. Similarly, if $\check{\Gamma}, \check{\Gamma}'$ are the corresponding index 2 subgroups, consisting of the elements which are words of even length, then the manifolds $M_{\check{\Gamma}} := \Omega(\Gamma)/\check{\Gamma}$ and $M_{\check{\Gamma}'} := \Omega(\Gamma')/\check{\Gamma}'$, are biholomorphically equivalent if and only if $\check{\Gamma}$ and $\check{\Gamma}'$ are projectively conjugate.*

Proof. — The first statement in (3.2) is obvious, i.e., that we have a holomorphic surjection $\pi\colon \mathfrak{G}^n_r \to \mathcal{C}^n_r$ with kernel $\mathfrak{T}_{\mathcal{L}}$. The other statements are immediate consequences of the following lemma (3.3), proved for us by Sergei Ivashkovich. Our proof below is a variation of Ivashkovich's proof.

3.3. Lemma. — *Let U be a connected open set in $P^{2n+1}_{\mathbb{C}}$ that contains a subspace $L \subset P^{2n+1}_{\mathbb{C}}$ of dimension n, and let $h\colon U \to V$ be a biholomorphism onto an open set $V \subset P^{2n+1}_{\mathbb{C}}$. Suppose that V also contains a subspace M of dimension n. Then h extends uniquely to an element in* $\mathrm{PSL}(2n+2, \mathbb{C})$.

Proof. — Let $f : U \to P^n_{\mathbb{C}}$ be a holomorphic map. Then f is defined by n meromorphic functions $f_1, \ldots, f_n$ from U to $P^1_{\mathbb{C}}$ (see [**Iva**]), *i.e.*, holomorphic functions which are defined outside of an analytic subset of U (the indeterminacy set).

Consider the set of all subspaces of $P^{2n+1}_{\mathbb{C}}$ of dimension $n+1$ which contain L. Then, if N is such subspace, one has a neighbourhood U_N of L in N which is the complement of a round ball in the affine part, $\mathbb{C}^{n+1}$, of N. Since the boundary of such a ball is a round sphere S_N and, hence, it is pseudo-convex, it follows from E. Levi extension theorem, applied to each f_i, that the restriction, f_N, of f to $U \cap N$ extends to all of N as a meromorphic function. The union of all subspaces N is $P^{2n+1}_{\mathbb{C}}$ and they all meet in L. The functions f_N depend holomorphically on N as is shown in [**Iva**]. One direct way to prove this is by considering the Henkin-Ramirez reproducing kernel defined on each round sphere S_N, [**He, Ram**]. One can choose the spheres S_N in such a way that the kernel depends holomorphically on N by considering a tubular neighbourhood of L in N whose radius is independent of N. Hence the extended functions to all $N's$ define a meromorphic function in all of $P^{2n+1}_{\mathbb{C}}$, which extends f. Now let h be as in the statement lemma 3.3 and let $\widetilde{h}$ be its meromorphic extension. Then, since by hypothesis h is a biholomorphism from the open set $U \subset P^n_{\mathbb{C}}$ onto the the open set $V := h(U) \subset P^n_{\mathbb{C}}$, one can apply the above arguments to $h^{-1} : V \to U$. Let $g : P^n_{\mathbb{C}} \to P^n_{\mathbb{C}}$ be the meromorphic extension of h^{-1}. Then, outside of their sets of indeterminacy, one has $\widetilde{h}g = g\widetilde{h} = Id$. Hence the indeterminacy sets are empty and both $\widetilde{h}$ and g are biholomorphisms of $P^n_{\mathbb{C}}$. In fact, in [**Iva**] it is shown that if f is as in the statement of lemma 3.3 and if f is required only to be locally injective, then f extends as a holomorphic function. ◻

Notice that if $n = 1$, then (3.3) becomes Lemma 3.2 in [**Ka1**].

3.4. Corollary. — *For $r > 2$ sufficiently large, the manifold $\Omega(\Gamma)/\check{\Gamma}$ has non-trivial moduli.*

In fact, if the manifolds $\Omega(\Gamma)/\check{\Gamma}$ and $\Omega(\Gamma')/\check{\Gamma}'$ are complex analytically equivalent, then $\check{\Gamma}$ is conjugate to $\check{\Gamma}'$ in $\mathrm{PSL}(2n+2, \mathbb{C})$, by (3.2), and the corresponding configurations $\mathcal{L}$ and $\mathcal{L}'$ are projectively equivalent. Now it is sufficient to choose r big enough to have two such configurations which are not projectively equivalent. This is possible because the action induced by the projective linear group G on the Grassmannian $G_{2n+1,n}$ is obtained from the projectivization of the action of $\mathrm{SL}(2n+2, \mathbb{C})$ acting on the Grassmann algebra Λ^{n+1}, of $(n+1)$-vectors of $\mathbb{C}^{2n+2}$, restricted to the set of decomposable $(n+1)$-vectors $\mathcal{D}^{n+1}$. The set $\mathcal{D}^{n+1}$ generates the Grassmann algebra and $G_{2n+1,n} = (\mathcal{D}^{n+1} - \{0\})/\sim$, where $\sim$ is the equivalence relation of projectivization.

If r is small with respect to n, then $\mathcal{C}^n_r$ consists of one point, because any two such configurations are in the same $\mathrm{PSL}(2n+2, \mathbb{C})$-orbit. Therefore, in this case $\mathfrak{T}_{\mathcal{L}}$ coincides with $\mathfrak{G}^n_r$. That is, to change the complex structure of M^n_r we need to change the corresponding involutions into a family of involutions, with the same configuration (up to conjugation), which is not conjugate to the given one.

The following result is a generalization of Theorem 1.2 in [**Ka1**]. This can be regarded as a restriction for a complex orbifold (or manifold) to be of the form $\Omega(\Gamma)/\Gamma$ (or $\Omega(\Gamma)/\check{\Gamma}$).

3.5. Proposition. — *If $r > 2$, then the compact complex manifolds and orbifolds $\Omega(\Gamma)/\check{\Gamma}$ and $\Omega(\Gamma)/\Gamma$, obtained in theorem 2.2, have no non-constant meromorphic functions.*

Proof. — Let f be a meromorphic function on one of these manifolds (or orbifolds). Then f lifts to a meromorphic function $\widetilde{f}$ on $\Omega(\Gamma) \subset P^{2n+1}_{\mathbb{C}}$, which is $\check{\Gamma}$-invariant. By lemma (3.6) below, f extends to a meromorphic function on all of $P^{2n+1}_{\mathbb{C}}$. Hence $\widetilde{f}$ must be constant, because $\check{\Gamma}$ is an infinite group. □

***3.6. Lemma* ([Iva])**. — *Let $U \subset P^{2n+1}_{\mathbb{C}}$, $n \geqslant 1$, be an open set that contains a projective subspace $P^n_{\mathbb{C}}$. Let $f : U \to P^1_{\mathbb{C}}$ be a meromorphic function. Then f can be extended to a meromorphic function $\widetilde{f} : U \to P^1_{\mathbb{C}}$.*

We refer to [**Iva**] for the proof of (3.6). In the following proposition we estimate an upper bound for the Hausdorff dimension of the limit set of some Schottky groups.

3.7. Proposition. — *Let $r > 2$, $0 < \lambda < (r-1)^{-1}$ and let Γ and $\check{\Gamma}$ be as in (1.7). Then, for every $\delta > 0$, the Hausdorff dimension of $\Lambda(\Gamma) = \Lambda(\check{\Gamma})$ is less than $2n+1+\delta$, i.e., the* transverse *Hausdorff dimension of $\Lambda(\Gamma) = \Lambda(\check{\Gamma})$ is less than $1+\delta$.*

Proof. — We recall that $\Lambda(\Gamma) = \cap_{k=0}^{\infty} F_k$, by the proof of theorem 1.8.i), where F_k is the disjoint union of the $r(r-1)^k$ closed tubular neighbourhoods $\gamma(N_i)$, $i \in \{1, \ldots, r\}$, where $\gamma \in \Gamma$ is an element which can be represented as a reduced word of length k in terms of the generators. $\gamma(N_i)$ is a closed tubular neighbourhood of $\gamma(L_i)$, as in theorem 1.7, and the "width" of each $\gamma(N_i)$, $w_{(\gamma,i)} := d(\gamma(E_i), L_i)$, satisfies $w_{(\gamma,i)} \leqslant C\lambda^k$, as was shown in lemma 1.6 and corollary 1.7. Hence,

$$w(k) := \sum_{\substack{l(\gamma)=k \\ i\in\{1,\ldots,r\}}} w_{(\gamma,i)}^{1+\delta} \leqslant Cr(r-1)^k \lambda^{k(1+\delta)} < Cr(r-1)^{-\delta k}.$$

Thus, $\lim_{k\to\infty} w(k) = 0$. Hence, just as in the proof of the theorem of Marstrand [**Mr**], the Hausdorff dimension of $\Lambda(\Gamma)$ can not exceed $2n+1+\delta$. □

Next we will apply the previous estimates to compute the versal deformations of manifolds obtained from complex Schottky groups as in (3.7), whose limit sets have small Hausdorff dimension.

We first recall [**Kod**] that given a compact complex manifold X, a deformation of X consists of a triple $(\mathcal{X}, \mathcal{B}, \omega)$, where $\mathcal{X}$ and $\mathcal{B}$ are complex analytic spaces and $\omega : \mathcal{X} \to \mathcal{B}$ is a surjective holomorphic map such that $\omega^{-1}(t)$ is a complex manifold for all $t \in \mathcal{B}$ and $\omega^{-1}(t_0) = X$ for some t_0, which is called the reference point. It is known [**Kur**] that given X, there is always a deformation $(\mathcal{X}, \mathfrak{K}_X, \omega)$ which is *universal*, in the sense that every other deformation is induced from it (see also [**KNS, Kod**]). The space $\mathfrak{K}_X$ is the Kuranishi space of *versal* deformations of X [**Kur**]. If we let $\Theta := \Theta_X$ be the sheaf of germs of local holomorphic vector fields on X, then every deformation of X determines, via differentiation, an element in $H^1(X, \Theta)$, so $H^1(X, \Theta)$ is called the *space of infinitesimal deformations* of X ([**Kod**], Ch. 4). Furthermore ([**KNS**] or [**Kod**, Th. 5.6]), if $H^2(X, \Theta) = 0$, then the Kuranishi space $\mathfrak{K}_X$ is smooth at the reference point t_0 and its tangent space at t_0 is canonically identified with $H^1(X, \Theta)$. In particular, in this case *every infinitesimal deformation* of X comes from an actual deformation, and vice-versa, every deformation of the complex structure on X, which is near the original complex structure, comes from an infinitesimal deformation.

The following lemma is an immediate application of (3.7) and Harvey's Theorem 1 in [**Ha**], which generalises the results of Scheja [**Schj**].

3.8. Lemma. — *Let $r > 2$, $0 < \lambda < (r-1)^{-1}$, let $\check{\Gamma}$ be as in proposition 3.7 and let $\Omega := \Omega(\Gamma) \subset P_{\mathbb{C}}^{2n+1}$ be its region of discontinuity. Then one has:*

$$H^j(\Omega, i^*(\Theta_{P_{\mathbb{C}}^{2n+1}})) \cong H^j(P_{\mathbb{C}}^{2n+1}, \Theta_{P_{\mathbb{C}}^{2n+1}}), \quad \textit{for } 0 \leqslant j < n,$$

where i is the inclusion of Ω in $P_{\mathbb{C}}^{2n+1}$. Hence, if $n > 1$, then one has:

$$H^0(\Omega, i^*(\Theta_{P_{\mathbb{C}}^{2n+1}})) \cong \mathfrak{sl}(2n+2, \mathbb{C}) \quad \textit{and} \quad H^j(\Omega, i^*(\Theta_{P_{\mathbb{C}}^{2n+1}})) \cong 0,$$

for all $0 < j < n$, where $\mathfrak{sl}(2n+2, \mathbb{C})$ is the Lie algebra of $\mathrm{PSL}(2n+2, \mathbb{C})$, and it is being considered throughout this section as an additive group.

Proof. — By (3.7) we have that the Hausdorff dimension d of the limit set $\Lambda(\check{\Gamma})$ satisfies $d < 2n + 1 + \delta$ for every $\delta > 0$. Therefore the Hausdorff measure of $\Lambda(\Gamma)$ of dimension s, $\mathcal{H}_s(\Lambda(\Gamma))$, is zero for every $s > 2n + 1$. Hence the first isomorphism in (3.8) follows from Theorem 1.ii) in [**Ha**], because the sheaf Θ is locally free. The second statement in (3.8) is now immediate, because

$$H^0(P_{\mathbb{C}}^{2n+1}, \Theta_{P_{\mathbb{C}}^{2n+1}}) \cong \mathfrak{sl}(2n+2, \mathbb{C}) \quad \text{and} \quad H^j(P_{\mathbb{C}}^{2n+1}, \Theta_{P_{\mathbb{C}}^{2n+1}}) \cong 0 \text{ for } j > 0,$$

a fact which follows immediately by applying the long exact sequence in cohomology derived from the short exact sequence:

$$0 \longrightarrow \mathcal{O} \longrightarrow [\mathcal{O}(1)]^{n+1} \longrightarrow \Theta_{P_{\mathbb{C}}^{2n+1}} \longrightarrow 0,$$

where $\mathcal{O}$ is the structural sheaf of $P_{\mathbb{C}}^{2n+1}$ and $[\mathcal{O}(1)]^{n+1}$ is the direct sum of $n+1$ copies of $\mathcal{O}_{P_{\mathbb{C}}^{2n+1}}(1)$, the sheaf of germs of holomorphic sections of the holomorphic line bundle over $P_{\mathbb{C}}^{2n+1}$ with Chern class 1. See Hartshorne [**Ht**], Example 8.20.1, page 182. □

We let $M := \Omega/\check{\Gamma}$, where $\check{\Gamma}$ is as above. We notice that Ω is simply connected when $n > 0$, so that Ω is the universal covering $\widetilde{M}$ of M. Let $p : \widetilde{M} \to M$ be the covering projection; since $\check{\Gamma}$ acts freely on Ω, this projection is actually given by the group action. Let Θ_M be the sheaf of germs of local holomorphic vector fields on M and let $\widetilde{\Theta}$ be the pull-back of Θ to $\widetilde{M}$ under the covering p; $\widetilde{\Theta}$ is the sheaf $i^*(\Theta_{P_{\mathbb{C}}^{2n+1}})$ on $\widetilde{M} = \Omega$.

3.9. Lemma. — *If $n > 2$, then for $0 \leqslant j \leqslant 2$ we have:*

$$H^j(M, \Theta_M) \cong H^j_\rho(\check{\Gamma}, \mathfrak{sl}(2n+2, \mathbb{C})),$$

where $\mathfrak{sl}(2n+2, \mathbb{C}))$ is considered as a $\check{\Gamma}$-left module via the representation

$$\rho : \check{\Gamma} \longrightarrow \mathrm{Aut}(\mathfrak{sl}(2n+2, \mathbb{C})))$$

given by:

$$\rho(\gamma)(v) = dT_g \circ v \circ {T_g}^{-1}, \quad v \in \mathfrak{sl}(2n+2, \mathbb{C}),$$

where T_g is the action of $g \in \check{\Gamma}$ on $P_{\mathbb{C}}^{2n+1}$.

Proof. — If $n > 2$, then (3.8) and Mumford's formula (c) in [**Mu**], pag 23, (see also Grothendieck [**Gr**], Chapter V) imply that there exists an isomorphism

$$\phi : H^j_\rho(\check{\Gamma}, H^0(\Omega, \widetilde{\Theta})) \longrightarrow H^j(M, \Theta_M),$$

for $0 \leqslant j \leqslant 2$, where $H^0(\Omega, \widetilde{\Theta})$ is the vector space of holomorphic vector fields on the universal covering $\widetilde{M} = \Omega \subset P_{\mathbb{C}}^{2n+1}$ of M.

Now, by [**Ha**], Theorem 1.i), every holomorphic vector field in $\Omega(\Gamma)$, extends to a holomorphic vector field defined in all of $P_{\mathbb{C}}^{2n+1}$. Therefore,

$$H^0(\Omega, \widetilde{\Theta}) = H^0(P_{\mathbb{C}}^{2n+1}, \Theta_{P_{\mathbb{C}}^{2n+1}}) = \mathfrak{sl}(2n+2, \mathbb{C}). \quad \square$$

We recall that $\check{\Gamma}$ is a free group of rank $r-1$; let $g_1, \ldots, g_{r-1}$ be generators of $\check{\Gamma}$. By [**HS**], page 195 Corollary 5.2, applied to $\check{\Gamma}$, we obtain:

$$H^1_\rho(\check{\Gamma}, \mathfrak{sl}(2n+2, \mathbb{C})) \cong \mathfrak{sl}(2n+2, \mathbb{C}) \times \cdots \times \mathfrak{sl}(2n+2, \mathbb{C})/Im(\psi),$$

where

$$\psi : \mathfrak{sl}(2n+2, \mathbb{C}) \longrightarrow \mathfrak{sl}(2n+2, \mathbb{C}) \times \cdots \times \mathfrak{sl}(2n+2, \mathbb{C})$$

is given by $\psi(v) = (g_1(v) - v, \ldots, g_{r-1}(v) - v)$. We claim that ψ is injective. Indeed, if v is a linear vector field in $P_{\mathbb{C}}^{2n+1}$ which is invariant by $g_1, \ldots, g_{r-1}$, then, by Jordan's theorem, this vector field is tangent to a hyperplane Π which is $\check{\Gamma}$-invariant. This can not happen. In fact, if L is a n dimensional projective subspace contained in $\Lambda(\check{\Gamma})$, then L must intersect Π transversally in a subspace of dimension $n-2$, for otherwise Π would contain the whole limit set $\Lambda(\check{\Gamma})$, which is a disjoint union of projective subspaces of dimension n. Hence, there exists $L \subset \Pi$, a projective n-subspace such that $L \cap \Lambda(\check{\Gamma}) = \varnothing$. Then, as we have shown in section 1, there exists a sequence $\{\gamma_i\}_{i \in \mathbb{N}}$ such that $\lim_{i \to \infty}(\gamma_i(L)) = L_1$, where $L_1 \subset \Lambda(\check{\Gamma})$, where L_1 is not contained in Π. This is a contradiction to the invariance of Π. $\square$

Therefore,

$$\dim_{\mathbb{C}} H^1(\Omega, \widetilde{\Theta}) = \dim_{\mathbb{C}}[\mathfrak{sl}(2n+2,\mathbb{C})^{r-2}] = (r-2)\left((2n+2)^2-1\right).$$

By [**HS**], page 197, Corollary 5.6 we have $H^2_\rho(\check{\Gamma}, \mathfrak{sl}(2n+2,\mathbb{C})) = 0$. Hence, by 3.9 above, one obtains,

$$H^2(M, \Theta_M) \cong H^2_\rho(\check{\Gamma}, \mathfrak{sl}(2n+2,\mathbb{C})) = 0.$$

Thus we arrive to the following theorem:

3.10. Theorem. — *Let $n, r > 2$ and let λ be an arbitrary scalar such that $0 < \lambda < (r-1)^{-1}$. Let Γ be a Schottky group as in (1.7.iii), so that the (Fubini-Study) distance from $\gamma(x)$ to the limit set Λ decreases faster that $C\lambda^k$ for every point $x \in P^{2n+1}_{\mathbb{C}}$ and any $\gamma \in \Gamma$ of word-length k (where C is some positive constant). Let $\check{\Gamma}$ be the index-two subgroup of Γ consisting of words of even length. Let Ω be the region of discontinuity of Γ, $M := \Omega/\check{\Gamma}$, and let $\mathfrak{K}^n_r$ denote the Kuranishi space of* versal *deformations of M, with reference point $t_0 \in \mathfrak{K}^n_r$ corresponding to M. Then, we have:*

$$H^1(M, \Theta_M) \cong H^1_\rho(\check{\Gamma}, \mathfrak{sl}(2n+2,\mathbb{C})) \cong \mathbb{C}^{(r-2)((2n+2)^2-1)},$$

and

$$H^2(M, \Theta_M) = 0.$$

Hence $\mathfrak{K}^n_r$ is non-singular at t_0, of complex dimension $(r-2)((2n+2)^2-1)$, and every small deformation of M is obtained by a small deformation of $\check{\Gamma}$ as a subgroup of $\mathrm{PSL}(2n+2,\mathbb{C})$*, unique up to conjugation.*

Although we only considered $n > 2$ above, the last theorem remains valid for $n = 0, 1$. In fact, if $n = 0$ and $r > 2$, we have the classical Schottky groups. The manifold $\Omega/\check{\Gamma}$ is a compact Riemann surface of genus $r-1$. It is well known that in this case the moduli space has dimension $3(r-1)-3 = 3(r-2)$, which, of course, coincides with the formula above. When $n = 1$ and $r > 2$ the manifolds $\Omega/\check{\Gamma}$ are Pretzel twistor spaces of genus $g = r-1$, in the sense of Penrose [**Pe**]. The theorem above gives that the dimension of the moduli space of this manifold is $15g-15$, which coincides with Penrose's calculations in page 251 of [**Pe**].

References

[Bo] R. Bowen – Hausdorff dimension of quasi-circles, *Publ. Math. IHES* **50** (1979), p. 11–25.

[Eh] C. Ehresmann – Sur les espaces fibrés différentiables, *C. R. Acad. Sci. Paris Sér. I Math.* **224** (1947), p. 1611–1612.

[Gr] A. Grothendieck – Sur quelques points d'algèbre homologique, *Tohoku Math. J., II. Ser.* **9** (1957), p. 119–221.

[Gu] R. Gunning – *On uniformization of complex manifolds; the role of connections*, Math. Notes, vol. 22, Princeton University Press, 1978.

[Ht] R. Hartshorne – *Algebraic geometry*, Graduate Texts in Mathematics, vol. 52, Springer-Verlag, New York - Heidelberg - Berlin, 1984.

[Ha] R. Harvey – Removable singularities of cohomology classes, *Amer. Journal of Math.* **96** (1974), no. 3, p. 498–504.

[He] G.M. Henkin – Integral representation of functions in strictly pseudoconvex domains and applications to the $\overline{\partial}$-problem, *Math. USSR, Sb.* **11** (1970), p. 273–281.

[HS] P.J. Hilton & U. Stammbach – *A Course in Homological Algebra*, Graduate Texts in Mathematics, vol. 4, Springer-Verlag, New York - Heidelberg - Berlin, 1971.

[Iva] S.M. Ivashkovich – Extension of locally biholomorphic mappings of domains into complex projective space, *Math. USSR, Izv.* (1984), p. 181–189, Translation from Izv. Akad. Nauk SSSR, Ser. Mat. **47** (1983), p. 197-206 (Russian).

[Ka1] M. Kato – On compact complex 3-folds with lines, *Japanese J. Math.* **11** (1985), p. 1–58.

[Ka3] ———, Compact Complex 3-folds with Projective Structures; The infinite Cyclic Fundamental Group Case, *Saitama Math. J.* **4** (1986), p. 35–49.

[Ka2] ———, Factorization of compact complex 3-folds which admit certain projective structures, *Tôhoku Math. J.* **41** (1989), p. 359–397.

[Ka4] ———, Compact Quotient Manifolds of Domains in a Complex 3-Dimensional Projective Space and the Lebesgue Measure of Limit Sets, *Tokyo J. Math.* **19** (1996), p. 99–119.

[Kod] K. Kodaira – *Complex Manifolds and Deformation of Complex Structures*, Springer Verlag, 1985.

[KNS] K. Kodaira, L. Nirenberg & D.C. Spencer – On the existence of deformations of complex analytic structures, *Ann. of Math.* **68** (1958), p. 450–459.

[Ku1] R.S. Kulkarni – Groups with domains of discontinuity, *Math. Ann.* **237** (1978), p. 253–272.

[Ku2] ———, Conformal structures and Möbius structures, in *Aspects of Mathematics* (R.S. Kulkarni & U. Pinkhall, eds.), Vieweg, 1988, p. 1–39.

[Kur] M. Kuranishi – On the locally complete families of complex analytic structures, *Ann. of Math.* **75** (1962), p. 536–577.

[Mr] J.M. Marstrand – The dimension of cartesian product sets, *Proc. Cambridge Phil. Soc.* **50** (1954), p. 198–202.

[Ma2] B. Maskit – A characterization of Schottky groups, *J. d'Analyse Math.* **19** (1967), p. 227–230.

[Ma1] ———, *Kleinian groups*, Grundelehren math. Wissenschaften, vol. 287, Springer Verlag, 1980.

[Mc2] C.T. McMullen – Rational maps and Kleinian groups, in *Proceedings International Congress of Mathematicians (Kyoto, 1990)*, Springer Verlag, 1991, p. 889–900.

[Mc1] ———, *Complex Dynamics and Renormalization*, Annals of Mathematics Studies, vol. 135, Princeton University Press. NJ, 1994.

[Mu] D. Mumford – *Abelian Varieties*, Tata Lectures, Oxford University Press, 1974.

[MFK] D. Mumford, J. Fogarty & F. Kirwan – *Geometric invariant theory*, 3rd enl. ed., Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 34, Springer-Verlag, Berlin, 1993.

[No] M.V. Nori – The Schottky groups in higher dimensions, in *Proceedings of Lefschetz Centennial Conference, Mexico City*, AMS Contemporary Math., vol. 58, part I, 1986, p. 195–197.

[Pe] R. Penrose – Pretzel twistor spaces, in *Further advances in twistor theory* (L.J. Mason & L.P. Hughston, eds.), Research Notes in Math., vol. 231, Pitman, 1990, p. 246–253.

[Po] H. Poincaré – *Papers on Fuchsian Functions*, Springer Verlag, 1985, translated by J. Stillwell.

[Ram] E. Ramirez de Arellano – Ein Divisionsproblem und Randintegraldarstellungen in der Komplexen Analysis, *Math. Ann.* **184** (1970), p. 172–187.

[Schj] G. Scheja – Riemannsche Hebbarkeitsätze für Cohomoligieklassen, *Math. Ann.* **144** (1961), p. 345–360.

[SV] J. Seade & A. Verjovsky – Higher Dimensional complex Kleinian groups, preprint.

[Si] M.A. Singer – Flat twistor spaces, conformally flat manifolds and four-dimensional field theory, *Commun. Math. Phys.* **133** (1990), p. 75–90.

[Su1] D.P. Sullivan – Seminar on Conformal and Hyperbolic geometry, Preliminary Publ. IHES, notes by M. Baker and J. Seade, 1982.

[Su2] ———, Conformal Dynamical Systems, in *Geometric dynamics (Rio de Janeiro, 1981)*, Lect. Notes in Math., vol. 1007, Springer Verlag, 1983, p. 725–752.

[Su3] ———, Quasiconformal homeomorphisms and dynamics I, *Ann. of Math.* **122** (1985), p. 401–418.

[Su4] ———, Quasiconformal homeomorphisms and dynamics II, *Acta Math.* **155** (1985), p. 243–260.

J. Seade, Instituto de Matemáticas, Universidad Nacional Autónoma de México, Unidad Cuernavaca, Apartado Postal 273-3, C.P. 62210, Cuernavaca, Morelos, México
E-mail : jseade@matcuer.unam.mx

A. Verjovsky, Instituto de Matemáticas, Universidad Nacional Autónoma de México, Unidad Cuernavaca, Apartado Postal 273-3, C.P. 62210, Cuernavaca, Morelos, México
E-mail : alberto@matcuer.unam.mx

ASTÉRISQUE

2003

288. M. REES – *Views of Parameter Space : Topographer and Resident*
287. Geometric Methods in Dynamics (II), volume in honor of Jacob Palis, W. DE MELO, M. VIANA, J.-C. YOCCOZ, editors
286. Geometric Methods in Dynamics (I), volume in honor of Jacob Palis, W. DE MELO, M. VIANA, J.-C. YOCCOZ, editors
285. PO HU – *Duality for Smooth Families in Equivariant Stable Homotopy Theory*
284. Autour de l'analyse microlocale, volume en l'honneur de Jean-Michel Bony, G. LEBEAU, éditeur

2002

283. L. ROBBIANO, C. ZUILY – *Analytic theory for the quadratic scattering wave front set and application to the Schrödinger equation*
282. SÉMINAIRE BOURBAKI, volume 2000/2001, exposés 880-893
281. T. DUQUESNE, J.-F. LE GALL – *Random Trees, Lévy Processes and Spatial Branching Processes*
280. A. MOKRANE, P. POLO, J. TILOUINE – *Cohomology of Siegel varieties*
279. Cohomologies p-adiques et applications arithmétiques (II), P. BERTHELOT, J.-M. FONTAINE, L. ILLUSIE, K. KATO, M. RAPOPORT, éditeurs
278. Cohomologies p-adiques et applications arithmétiques (I), P. BERTHELOT, J.-M. FONTAINE, L. ILLUSIE, K. KATO, M. RAPOPORT, éditeurs
277. B. RÉMY – *Groupes de Kac-Moody déployés et presque déployés*
276. SÉMINAIRE BOURBAKI, volume 1999/2000, exposés 865-879

2001

275. J.-M. BISMUT, S. GOETTE – *Families torsion and Morse functions*
274. A. BONNET, G. DAVID – *Cracktip is a global Mumford-Shah minimizer*
273. K. NISHIYAMA, H. OCHIAI, K. TANIGUCHI, H. YAMASHITA, S. KATO – *Nilpotent orbits, associated cycles and Whittaker models for highest weight representations*
272. M. BOILEAU, J. PORTI – *Geometrization of 3-orbifolds of cyclic type (avec la collaboration de M. HEUSENER)*
271. M. KASHIWARA, P. SCHAPIRA – *Ind-Sheaves*
270. M. BONK, J. HEINONEN, P. KOSKELA – *Uniformizing Gromov hyperbolic spaces*
269. J.-L. WALDSPURGER – *Intégrales orbitales nilpotentes et endoscopie pour les groupes classiques non ramifiés*

2000

268. J. FRANCHETEAU, G. MÉTIVIER – *Existence de chocs faibles pour des systèmes quasi-linéaires hyperboliques multidimensionnels*
267. R. CERF – *Large Deviations for three Dimensional Supercritical Percolation*
266. SÉMINAIRE BOURBAKI, volume 1998/1999, exposés 850-864
265. O. BIQUARD – *Métriques d'Einstein asymptotiquement symétriques*
264. L. LIPSHITZ, Z. ROBINSON – *Rings of Separated Power Series and Quasi-Affinoid Geometry*
263. C. SABBAH – *Équations différentielles à points singuliers irréguliers et phénomène de Stokes en dimension 2*
262. A. VASY – *Propagation of singularities in three-body scattering*
261. Géométrie complexe et systèmes dynamiques, colloque en l'honneur d'Adrien Douady, Orsay 1995, M. FLEXOR, P. SENTENAC et J.-C. YOCCOZ, éditeurs

1999

260. S. D. CUTKOSKY – *Local monomialization and factorization of morphisms*
259. R. KRIKORIAN – *Réductibilité des systèmes produits-croisés à valeurs dans des groupes compacts*
258. Structure Theory of Set Addition, J.-M. DESHOUILLERS, B. LANDREAU and A.A. YUDIN, editors
257. J.-P. LABESSE – *Cohomologie, stabilisation et changement de base (avec la collaboration de L. BREEN et L. CLOZEL)*
256. F. MOREL – *Théorie homotopique des schémas*
255. R.E. KOTTWITZ, D. SHELSTAD – *Foundations of twisted endoscopy*
254. C.J. BUSHNELL, G. HENNIART – *Local tame lifting for* $\mathrm{GL}(n)$ *II : wildly ramified supercuspidals*
253. B. MAGNERON – *Involutions complexes et vecteurs sphériques associés pour les groupes de Lie nilpotents réels*

1998

252. SÉMINAIRE BOURBAKI, volume 1997/1998, exposés 835-849
251. Nombre et répartition de points de hauteur bornée, E. PEYRE, éditeur
250. C. BONATTI, R. LANGEVIN – *Difféomorphismes de Smale des surfaces (avec la collaboration de E. JEANDENANS)*
249. P. AUSCHER, P. TCHAMITCHIAN – *Square root problem for divergence operators and related topics*
248. P. COLMEZ – *Intégration sur les variétés* p*-adiques*
247. G. PISIER – *Non-commutative vector valued* L_p*-spaces and completely* p*-summing maps*

1997

246. V. TARASOV, A. VARCHENKO – *Geometry of* q*-hypergeometric functions, quantum affine algebras and elliptic quantum groups*
245. SÉMINAIRE BOURBAKI, volume 1996/1997, exposés 820-834
244. J.-M. BISMUT – *Holomorphic families of immersions and higher analytic torsion forms*
243. L. LAFFORGUE – *Chtoucas de Drinfeld et conjecture de Ramanujan-Petersson*
242. N. BURQ – *Pôles de diffusion engendrés par un coin*
241. SÉMINAIRE BOURBAKI, volume 1995/1996, exposés 805-819

1996

240. A. SÀ BARETTO, R. B. MELROSE, M. ZWORSKI – *Semilinear diffraction of conormal waves*
239. J.-L. VERDIER – *Des catégories dérivées des catégories abéliennes*
238. A. BROISE, F. DAL'BO, M. PEIGNÉ – *Méthodes des opérateurs de transfert : transformations dilatantes de l'intervalle et dénombrement de géodésiques fermées*
237. SÉMINAIRE BOURBAKI, volume 1994/1995, exposés 790-804
236. Hommage à P. A. MEYER et J. NEVEU
235. J.-P. OTAL – *Le théorème d'hyperbolisation pour les variétés fibrées de dimension 3*
234. A. GENESTIER – *Espaces symétriques de Drinfeld*

1995

233. I. KRIZ, J.P. MAY – *Operads, algebras modules and motives*
232. Recent advances in operator algebras (Orléans, 1992)
231. J.-C. YOCCOZ – *Petits diviseurs en dimension 1*
230. J.-Y. CHEMIN – *Fluides parfaits incompressibles*
229. B. PERRIN-RIOU – *Fonctions* L p*-adiques des représentations* p*-adiques*
228. Columbia University number theory seminar (New-York, 1992)
227. SÉMINAIRE BOURBAKI, volume 1993/1994, exposés 775-789