



HAL
open science

Pavages de l'espace affine

Ilia Smilga

► **To cite this version:**

Ilia Smilga. Pavages de l'espace affine. Théorie des groupes [math.GR]. Université Paris Sud - Paris XI, 2014. Français. NNT : 2014PA112298 . tel-01127044

HAL Id: tel-01127044

<https://theses.hal.science/tel-01127044>

Submitted on 6 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS-SUD

ÉCOLE DOCTORALE 142 :
MATHÉMATIQUES DE LA RÉGION PARIS-SUD

Laboratoire de Mathématiques d'Orsay

THÈSE DE DOCTORAT

MATHÉMATIQUES

par

Ilia SMILGA

Pavages de l'espace affine

Date de soutenance : 12/11/2014

Composition du jury :

Directeur de thèse :
Rapporteurs :
Examineurs :

Yves BENOIST
Thierry BARBOT
Elisha FALBEL
Alessandra IOZZI
Frédéric PAULIN

Professeur à l'Université de Paris-Sud
Professeur à l'Université d'Avignon
Professeur à l'Institut de mathématiques de Jussieu
Professeur à l'ETH Zürich
Professeur à l'Université de Paris-Sud

Rapporteur absent le jour de la soutenance : Anna WIENHARD, professeur à l'Université de Heidelberg

Remerciements

Je tiens avant tout à remercier chaleureusement Yves Benoist pour ces trois années de travail et d'apprentissage. Outre sa compétence mathématique indubitable et sa passion communicative pour l'objet de la recherche, j'ai beaucoup apprécié sa gentillesse, sa patience, son accueil chaleureux notamment à Berkeley et à Jérusalem.

Je remercie également les rapporteurs, Thierry Barbot et Anna Wienhard, pour avoir scrupuleusement relu ce dossier de thèse et rendu un rapport détaillé. Au cours de ma thèse, j'ai eu l'occasion de constater à quel point il pouvait être difficile de lire un texte mathématique. J'ai surtout eu l'occasion de découvrir à quel point il était difficile d'en produire un qui soit lisible ! J'en suis d'autant plus reconnaissant aux rapporteurs.

Merci aux membres du jury (nommés sur la page de couverture) pour être venus jusqu'ici aujourd'hui et m'avoir permis ainsi de soutenir cette thèse, malgré l'éloignement géographique de certains et les plannings chargés de chacun. Un merci particulier à Frédéric Paulin pour ses nombreuses corrections ; il a non seulement éliminé un nombre insoupçonné de coquilles, mais a également fait en sorte que ce mémoire soit rédigé dans un français impeccable, en évitant les anglicismes et en mettant les accents sur les majuscules.

Un grand merci à mes collègues doctorants et jeunes docteurs pour l'intérêt qu'ils ont porté à mon travail, qui a été une grande source de motivation pour moi. Ils sont trop nombreux pour que je puisse dresser une liste exhaustive ; mais je pense en particulier à Paul Mercat (qui m'a permis de rencontrer mon directeur de thèse), Fanny Kassel et François Guéritaud (pour les discussions fructueuses), Bruno Le Floch, Fathi Ben Aribi et Esther Elbaz (pour m'avoir aidé à mettre au point des exposés).

Si j'ai réussi à atteindre ce niveau d'études, c'est en grande partie grâce à ma famille. En particulier, je ne pourrai jamais assez remercier mon père Andreï et feu mon grand-père Voldemar pour m'avoir fait découvrir les mathématiques dès mon plus jeune âge. Bien sûr, je ne serais pas non plus à cette place sans l'amour et l'éducation que j'ai reçu de la part de tous mes proches. Je remercie donc également ma mère Elena, mon frère aîné Boris et tout le reste de la famille.

Mis à part ma famille, de nombreuses autres personnes m'ont motivé et ont contribué à ma passion pour les mathématiques. Je vais tenter d'en faire la liste, bien que je sois sûr d'en oublier. Avant tout, un grand merci aux responsables de la préparation aux Olympiades de Mathématiques : Claude Deschamps, Dimitri Zvonkine, Xavier Caruso, Pierre Bornsztein, Yann Ollivier, Pierre Dehornoy, François Lo Jacomo, Johan Yebbou et les autres. Leur rôle dans ma formation mathématique est difficile à surestimer. Merci également aux camarades qui ont suivi cette préparation avec moi : leur émulation et la bonne ambiance qu'ils créaient m'ont certainement aidé à travailler. Merci aux professeurs que j'ai eus en Russie lors de divers stages et conférences d'été. Merci à mes professeurs et mes colleurs en classe préparatoire (particulièrement Yves Duval et Roger Mansuy), à mes professeurs et mes chargés de TD à l'ENS (là il y avait tellement de cours et de TDs de qualité exceptionnelle que la liste serait trop longue à faire¹ !), ainsi qu'aux brillants camarades qui étaient à mes côtés. Je vais cependant mentionner deux professeurs d'autres disciplines : merci à Jean-Noël Isaïa (professeur de physique en MPSI) et Jean-Christophe Filliâtre (professeur d'informatique à l'ENS) ; suivre leurs cours a été un vrai plaisir, et m'a permis non seulement d'approfondir mais aussi d'élargir ma culture scientifique.

Je tiens à remercier Florent Hivert (et quelques autres) pour des conseils concernant la carrière, la recherche d'emploi et les candidatures, difficiles à obtenir par ailleurs !

Merci à tous mes amis, pour toutes les discussions et les activités de loisir qui m'ont permis de me reposer et redonné de l'énergie pour mon travail. Merci enfin à ma compagne, Albane Mercat, qui m'a soutenu pendant ces années. En m'encourageant en permanence

1. Et puis ne le répétez à personne, mais sachant que certains d'entre eux pourraient bientôt avoir ma carrière entre leurs mains, je ne voudrais pas prendre le risque d'en vexer !

à prendre du recul sur moi-même, elle a une part de mérite dans le fait que j'aie terminé cette thèse à temps : c'est en partie elle qui m'a aidé à mieux m'organiser et à travailler plus efficacement.

Résumé : Pour tout entier naturel impair d , on construit un domaine fondamental pour l'action sur l'espace affine de dimension $2d+1$ de certains groupes de transformations affines libres non abéliens, discrets, agissant proprement et de partie linéaire Zariski-dense dans $SO(d+1, d)$. Pour tout groupe de Lie semisimple réel non compact G , on construit ensuite un groupe de transformations affines de son algèbre de Lie \mathfrak{g} qui est libre non abélien, discret, agit proprement sur \mathfrak{g} et a sa partie linéaire Zariski-dense dans $\text{Ad}G$. Enfin, on donne quelques résultats sur le comportement local des fonctions harmoniques sur le triangle de Sierpinski, plus précisément de leur restriction à un bord du triangle.

Abstract : For every odd positive integer d , we construct a fundamental domain for the action on the $2d+1$ -dimensional space of certain groups of affine transformations which are free, nonabelian, act properly discontinuously and have linear part Zariski-dense in $SO(d+1, d)$. Next for every semisimple noncompact real Lie group G , we construct a group of affine transformations of its Lie algebra \mathfrak{g} which is free, nonabelian, acts properly discontinuously and has linear part Zariski-dense in $\text{Ad}G$. Finally, we give some results about the local behavior of harmonic functions on the Sierpinski triangle restricted to a side of the triangle.

Table des matières

| | |
|----------------------------------------------------------------------------------------|-----------|
| I. Introduction | 9 |
| I.1. Contexte et motivation | 10 |
| I.1.1. Pavages périodiques de l'espace affine | 10 |
| I.1.2. Groupes discrets agissant proprement sur l'espace affine | 11 |
| I.1.3. Fonctions harmoniques sur le triangle de Sierpinski | 12 |
| I.2. Sommaire de la thèse | 12 |
| | |
| II. Domaines fondamentaux pour des groupes affines proprement discontinus | 15 |
| II.1. Introduction | 16 |
| II.2. Conventions, définitions et propriétés basiques | 16 |
| II.2.1. Sous-espaces totalement isotropes maximaux | 16 |
| II.2.2. Applications pseudohyperboliques et cadres | 18 |
| II.2.3. Orientation | 19 |
| II.2.4. Force de contraction et autres considérations métriques | 21 |
| II.3. Pseudohyperbolicité des produits | 23 |
| II.3.1. Cas proximal | 23 |
| II.3.2. Cas pseudohyperbolique | 26 |
| II.4. La « balle de tennis » et les groupes de Schottky généralisés | 31 |
| II.5. Déformations affines | 33 |
| | |
| III. Actions affines propres sur les algèbres de Lie semisimples | 41 |
| III.1. Introduction | 42 |
| III.2. Définitions et propriétés préliminaires | 42 |
| III.2.1. Espace affine étendu et sous-espaces dynamiques | 42 |
| III.2.2. Structure d'algèbre de Lie | 43 |
| III.2.3. Propriétés algébriques de base | 44 |
| III.2.4. Quasi-translations | 46 |
| III.2.5. Identifications canoniques | 47 |
| III.2.6. Propriétés métriques | 48 |
| III.2.7. Comparaison des propriétés métriques dans le cas affine et linéaire | 50 |
| III.3. \mathbb{R} -régularité des produits | 52 |
| III.3.1. Cas proximal | 52 |
| III.3.2. Cas \mathbb{R} -régulier | 55 |
| III.4. Additivité de l'invariant de Margulis | 59 |
| III.5. Invariants de Margulis de mots | 66 |
| III.6. Construction du groupe | 69 |
| | |
| IV. Fonctions harmoniques sur le triangle de Sierpinski | 73 |
| IV.1. Introduction | 74 |
| IV.1.1. Notations et conventions | 74 |

| | |
|-------------------------------------------------------------------------------------|----|
| IV.1.2. Problématique | 74 |
| IV.2. Préliminaires : fonctions harmoniques sur le triangle de Sierpinski | 75 |
| IV.3. Généralités sur u | 79 |
| IV.4. Comportement local de u : existence et direction de la tangente | 82 |
| IV.5. Comportement local de u et de $u_a^c_b$: taux de variation | 85 |
| IV.5.1. Formule pour l'exposant de Hölder | 85 |
| IV.5.2. Cas des points rationnels | 88 |
| IV.5.3. Dérivée de u | 90 |
| IV.5.4. Valeurs numériques et estimations de l'exposant de Hölder | 91 |

Première partie.

Introduction

I.1. Contexte et motivation

I.1.1. Pavages périodiques de l'espace affine

Un *pavage* de \mathbb{R}^n est un ensemble \mathcal{P} de parties de \mathbb{R}^n , appelées *pavés*, qui sont d'intérieurs disjoints, de réunion égale à \mathbb{R}^n tout entier, et vérifient la propriété que tout compact ne rencontre qu'un nombre fini de pavés (on parlera de finitude locale). La figure I.1 représente quelques exemples de pavages familiers au lecteur, en dimension 2.

Ces pavages ont tous la propriété d'être *périodiques*. Cela veut dire qu'il existe un ensemble fini de pavés P_1, \dots, P_k (quitte à rassembler certains pavés, on supposera en fait que $k = 1$) tel que tous les autres pavés sont images de l'un des P_i par une transformation (appartenant à un certain groupe dont on va discuter dans un instant) qui laisse invariante le pavage dans sa globalité.

L'espace dans lequel on vit possède une structure euclidienne. On imagine donc en général que, comme dans ces trois exemples, les transformations en question appartiennent au groupe $\text{Isom}(\mathbb{R}^n)$ des isométries affines de \mathbb{R}^n , qui est égal au produit semidirect du groupe des isométries linéaires (applications linéaires préservant une forme euclidienne) par le groupe des translations :

$$\text{Isom}(\mathbb{R}^n) = \text{O}_n(\mathbb{R}) \ltimes \mathbb{R}^n.$$

De tels pavages ont été étudiés de façon très détaillée, et complètement classés en petite dimension. Ils ont des applications pratiques notamment en cristallographie.

Mais on peut aussi se demander ce qui se passe si on veut paver un espace affine « pur », sans structure euclidienne prédéfinie. Dans ce cas, il faut autoriser les symétries du pavage à être des transformations affines quelconques (pas nécessairement des isométries). Ainsi elles appartiennent au groupe affine tout entier, produit semidirect du groupe linéaire général par le groupe des translations :

$$\text{Aff}(\mathbb{R}^n) = \text{GL}_n(\mathbb{R}) \ltimes \mathbb{R}^n.$$

Un exemple de pavage périodique du plan affine est donné par la figure I.2. Pour une présentation grand public des pavages du plan affine, on renvoie à [7].

On dit qu'un groupe discret Γ agit *proprement* sur un espace topologique (localement compact) si pour tout compact K dans cet espace, l'ensemble $\{\gamma \in \Gamma \mid \gamma K \cap K \neq \emptyset\}$ des images de K qui rencontrent K est fini. Les pavages périodiques de l'espace affine sont fortement liés aux groupes discrets de transformations affines dont l'action est propre. En effet, si on dispose d'un pavage périodique de l'espace affine, son *groupe de symétrie*, c'est-à-dire le groupe des transformations affines laissant le pavage invariant, est un tel groupe, pour peu que le stabilisateur de chaque pavé soit fini. Réciproquement, si on dispose d'un tel groupe Γ , il est toujours possible de construire un *domaine fondamental* pour son action, c'est-à-dire une partie de l'espace affine dont les images par les éléments de Γ pavent l'espace.

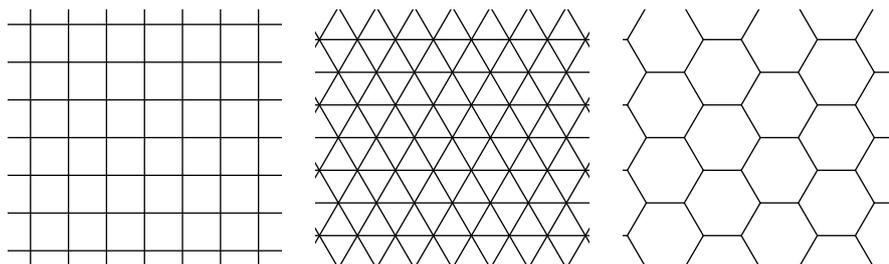


FIGURE I.1.: Quelques pavages familiers.

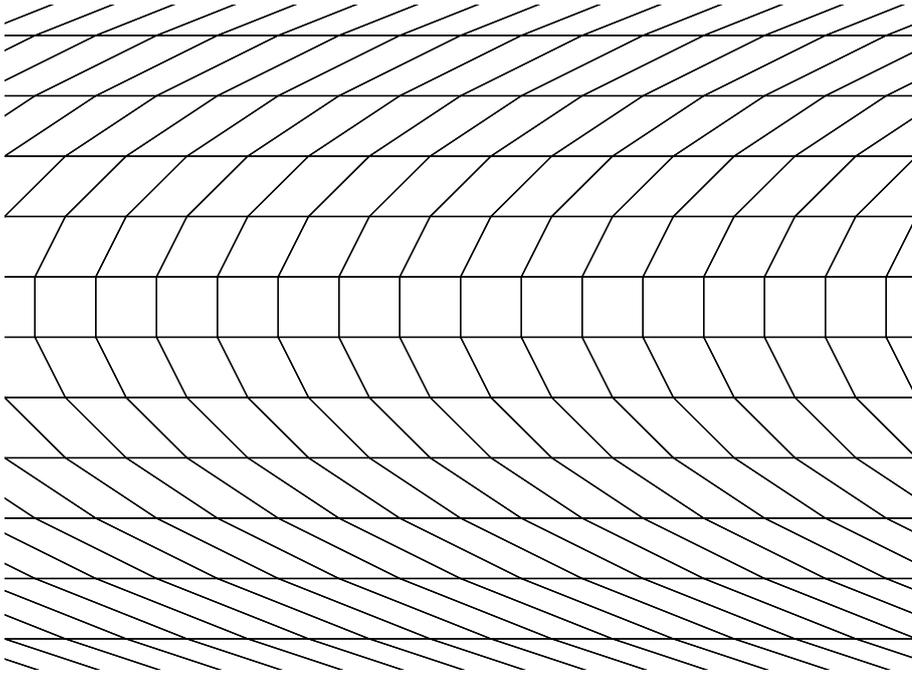


FIGURE I.2.: Un pavage périodique de l'espace affine.

Exemple I.1.1. Le pavage de la figure I.2 correspond au groupe engendré par les applications

$$g_1 : (x, y) \mapsto (x + 1, y) \text{ et } g_2 : (x, y) \mapsto (x + ay, y + 1)$$

(où a est un certain paramètre réel ; en l'occurrence la figure a été faite avec $a = 0.5$.)

On appelle *groupe cristallographique* tout groupe discret Γ de transformations affines agissant proprement sur l'espace affine \mathbb{R}^n et tel que le quotient Γ/\mathbb{R}^n soit compact. Ces groupes correspondent aux pavages par des pavés compacts.

I.1.2. Groupes discrets agissant proprement sur l'espace affine

Les deux premières parties de ce mémoire rentrent dans le cadre d'un travail plus large visant à comprendre les groupes discrets Γ de transformations affines agissant proprement. Dans le cas où Γ est un groupe d'isométries (en d'autres termes, $\Gamma \subset O_n(\mathbb{R}) \times \mathbb{R}^n$), on dispose du résultat suivant :

Théorème I.1.2 (Bieberbach 1911).

- *Tout groupe d'isométries affines agissant proprement sur \mathbb{R}^n est virtuellement abélien.*
- *Tout groupe cristallographique d'isométries affines agissant proprement sur \mathbb{R}^n est virtuellement un réseau de translations.*

(On dit qu'un groupe a *virtuellement* une propriété s'il a un sous-groupe d'indice fini qui a cette propriété). On aimerait bien avoir un résultat analogue dans le cas général. La conjecture suivante a été formulée :

Conjecture I.1.3 (Auslander 1964, [5]). *Tout groupe cristallographique de transformations affines est virtuellement résoluble,*

puis une plus forte :

Conjecture I.1.4 (Milnor 1977, [17]). *Tout groupe discret de transformations affines agissant proprement est virtuellement résoluble.*

Pour mettre ces énoncés en perspective, on doit mentionner le résultat suivant :

Théorème I.1.5 (Alternative de Tits, [22]). *Tout sous-groupe de $GL_n(\mathbb{R})$ est soit virtuellement résoluble, soit contient un groupe libre non abélien.*

Cette alternative s'applique également aux groupes de transformations affines, puisque $GL_n(\mathbb{R}) \ltimes \mathbb{R}^n$ se plonge dans $GL_{n+1}(\mathbb{R})$.

On sait maintenant que la conjecture de Milnor est fautive : Margulis [15, 16] a donné un groupe libre non abélien de transformations affines de partie linéaire Zariski-dense dans $SO(2, 1)$, discret et agissant proprement sur \mathbb{R}^3 . Ce contre-exemple a ensuite été généralisé par Abels, Margulis et Soifer dans [2] à des sous-groupes de $SO(d+1, d)$ pour tout d impair. Jusqu'à aujourd'hui, à ma connaissance, c'étaient les seuls contre-exemples connus à la conjecture de Milnor. Quant à la question d'Auslander, elle reste ouverte depuis cinquante ans. Fried et Goldman [11] l'ont démontrée en dimension 3 (les cas $n = 1$ et 2 étant faciles), et Abels et al. [4] en dimension $n \leq 6$. Quelques autres résultats partiels existent ; voir [1] pour un récapitulatif.

La conjecture d'Auslander est la principale motivation de cette thèse. Notre approche consiste à se focaliser sur les contre-exemples à la conjecture de Milnor. Si on les étudie de façon suffisamment détaillée, on peut peut-être espérer démontrer, un jour, qu'aucun de ces contre-exemples n'a une action avec quotient compact.

I.1.3. Fonctions harmoniques sur le triangle de Sierpinski

Dans la dernière partie de cette thèse, on étudie une question totalement différente : la dérivabilité des applications harmoniques sur le triangle de Sierpinski. Le lien avec les parties précédentes est que cette dérivabilité, comme on le verra au cours de la résolution, dépend de la dynamique d'un certain semigroupe libre (à deux générateurs) d'applications affines. On retrouve ainsi quelques méthodes très similaires aux deux premières parties.

I.2. Sommaire de la thèse

Les trois parties peuvent être lues indépendamment. Chaque partie utilise ses propres notations.

Indiquons également que la partie II a été publiée dans *Geometriae Dedicata* [19], et que la partie III a été soumise aux *Annales de l'Institut Fourier* [20].

Partie II : domaines fondamentaux pour des groupes affines proprement discontinus

Dans cette partie, on construit un domaine fondamental pour certains groupes libres discrets agissant proprement sur l'espace affine \mathbb{R}^{2d+1} (où d est un entier impair quelconque) et préservant une forme quadratique de signature $(d+1, d)$. Le résultat principal de cette partie est le théorème II.5.2.

On commence, dans la section II.2, par donner quelques propriétés géométriques élémentaires de l'espace équipé d'une forme de signature $(d+1, d)$ où d est impair. On décrit, dans la sous-section II.2.1, ses sous-espaces totalement isotropes maximaux ; dans la sous-section II.2.2, ses applications pseudohyperboliques (en gros les applications dont l'espace des points fixes a la plus petite dimension possible) ; dans la sous-section II.2.3, une astuce d'orientation (provenant de [2]) qui permet d'étendre n'importe quelle paire de sous-espaces totalement isotropes maximaux transverses en une paire de demi-espaces de dimension $d+1$ qui ont encore une intersection triviale. Finalement, dans la sous-section II.2.4, on introduit des métriques sur divers espaces (notamment des espaces projectifs)

dans lesquels on devra travailler, et on définit la force de contraction d'une application pseudohyperbolique.

Dans les deux sections qui suivent, on considère des sous-groupes de $\mathrm{SO}(d+1, d)$ engendrés par des applications pseudohyperboliques. Dans la section II.3, on étudie leur action sur $\mathbb{P}(\Lambda^d \mathbb{R}^{d+1, d})$. On démontre que, si on suppose ses générateurs suffisamment contractants, un tel groupe est libre et tout élément est pseudohyperbolique. On contrôle également la géométrie et la force de contraction de tous les mots cycliquement réduits sur les générateurs. Ce résultat est très similaire au lemme 5.24 de [2], et on suit leur preuve de très près. (Pour une preuve plus concise d'un résultat similaire, voir aussi la section 6 de [6].)

Dans la section II.4, on étudie l'action de ces sous-groupes directement sur $\mathbb{P}(\mathbb{R}^{d+1, d})$. On montre que, en supposant à nouveau que les générateurs soient suffisamment contractants, cette action est similaire à l'action d'un groupe de Schottky (ce qui montre à nouveau qu'un tel groupe est libre).

Finalement, dans la section II.5, on étudie les groupes affines Γ dont les parties linéaires satisfont les conditions des deux sections précédentes. On démontre alors le théorème principal. La démonstration suit de près la section 4 de [8]. On commence par décrire un ensemble \mathcal{H}^0 qui est le complémentaire de $2n$ « sources » et « puits » correspondant aux n générateurs de Γ . On démontre (proposition II.5.5) que sous certaines conditions, \mathcal{H}^0 est un domaine fondamental pour Γ . En effet, on voit immédiatement que ses images par les éléments du groupe « s'emboîtent bien ». Pour montrer qu'elles couvrent l'espace tout entier, par l'absurde, on s'intéresse à un point hypothétique qui ne serait couvert par aucun « pavé ». On l'inclut dans une suite emboîtée d'ouverts, puis on montre (par des méthodes adaptées de [8]) que ces ouverts doivent, en un certain sens, s'en aller à l'infini.

Partie III : actions affines propres sur les algèbres de Lie semisimples

Dans cette partie, on construit une nouvelle famille de contre-exemples à la conjecture de Milnor : pour tout groupe de Lie réel semisimple non compact, on donne un sous-groupe libre de $G \ltimes \mathfrak{g}$ agissant proprement (via l'action adjointe) sur \mathfrak{g} et de partie linéaire Zariski-dense dans G . Le résultat principal est le théorème III.1.1.

Dans la section III.2, on donne quelques définitions et quelques propriétés algébriques et métriques de base. Dans la sous-section III.2.1, on remplace l'espace affine par un espace vectoriel $\hat{\mathfrak{g}}$ ayant une dimension de plus, plus commode à manipuler ; et on définit, pour tout élément du groupe $G \ltimes \mathfrak{g}$, une famille de sous-espaces vectoriels et affines « dynamiques ». Dans la sous-section III.2.2, on définit quelques sous-algèbres classiques de \mathfrak{g} . Dans la sous-section III.2.3, on donne quelques propriétés algébriques de base : on fait le lien entre les espaces dynamiques d'une application \mathbb{R} -régulière et les sous-algèbres classiques, et on montre que dans chaque cas, la « géométrie du problème » est essentiellement donnée par un couple d'algèbres paraboliques minimales affines transverses. Dans la sous-section III.2.4, on introduit une classe importante d'automorphismes de l'espace affine parallèle à $\mathfrak{l} = \mathfrak{m} \oplus \mathfrak{a}$, appelés quasi-translations. Dans la sous-section III.2.5, on utilise les deux sous-sections précédentes pour identifier (à quasi-translation près) différents couples d'algèbres paraboliques minimales affines transverses, et montrer que ces identifications sont « naturelles » ; ceci nous permet de définir un invariant de Margulis généralisé (qui est un vecteur). Dans la sous-section III.2.6, on introduit une métrique euclidienne sur l'« espace affine étendu » $\hat{\mathfrak{g}}$, et on l'utilise pour définir deux choses importantes : la notion de couple C -non-dégénéré d'algèbres paraboliques minimales affines transverses (ce qui signifie qu'on peut prétendre qu'elles sont perpendiculaires et se tromper au plus d'un facteur C ou C^2), et la force de contraction d'une application \mathbb{R} -régulière. Dans la sous-section III.2.7, on fait le lien entre ces propriétés métriques d'un élément de $G \ltimes \mathfrak{g}$ et celles de sa partie linéaire.

Dans la section III.3, on démontre que le produit de deux applications \mathbb{R} -régulières est \mathbb{R} -régulier, et on exprime la géométrie et la force de contraction du produit en fonction

de la géométrie et de la force de contraction des facteurs. Pour ce faire, on examine la dynamique de ces applications agissant sur l'espace $\Lambda^q \hat{\mathfrak{g}}$. Cette section est plus ou moins une généralisation de la section II.3, avec des preuves très semblables.

La section III.4 contient le point clé de notre argument. On montre que sous de bonnes hypothèses, l'invariant de Margulis du produit de deux applications \mathbb{R} -régulières est approximativement égal à la somme de leurs invariants de Margulis. On donne également l'expression de l'invariant de Margulis de l'inverse d'une application.

Dans la section III.5, on utilise la récurrence pour démontrer un résultat similaire pour le produit d'un nombre arbitraire d'applications.

Dans la section III.6, on construit un groupe qui satisfait le théorème III.1.1. Comme générateurs, on prend une famille d'applications \mathbb{R} -régulières fortement contractantes, en position générale et avec des invariants de Margulis convenables. En utilisant le résultat de la section précédente, on montre que les éléments du groupe ont des invariants de Margulis qui croissent de façon non bornée, ce qui garantit la propriété de l'action.

Partie IV : fonctions harmoniques sur le triangle de Sierpinski

Dans cette partie, on répond à la question suivante, posée par Kirillov dans son livre [13] (intitulée « Problème 2 », vers la fin de la section 3.2) :

Calculer explicitement la dérivée $u'(t)$ à chaque fois que c'est possible (par exemple en tous les points rationnels),

où u est une fonction harmonique sur le triangle de Sierpinski, restreinte à un de ses côtés. On démontre plusieurs résultats, qui donnent ensemble une réponse à cette question.

On commence, dans la section IV.2, par expliquer le rapport entre les deux définitions de u : celle qu'on vient de donner, et la définition formelle par une équation fonctionnelle. En suivant dans les grandes lignes le livre de Kirillov, on construit les fonctions harmoniques sur le triangle de Sierpinski, et on montre que u est en fait la restriction d'une certaine fonction harmonique « universelle » à un côté du triangle. Le reste de l'article est logiquement indépendant de cette section, sauf pour quelques notations et définitions.

Dans la section IV.3, on établit quelques généralités sur la fonction u .

Dans la suite, on étudie le comportement de $u(t) - u(s)$ quand $t \rightarrow s$. Dans un premier temps (section IV.4), on décrit complètement le comportement de la direction de ce vecteur. Ceci nous donne au passage l'aspect global de la courbe $u(t)$, et permettra un peu plus loin de relier les propriétés des fonctions auxquelles s'intéresse Kirillov à celles de u .

Dans la section IV.5, on étudie la norme de ce vecteur. Le résultat clé de cette partie est la proposition IV.5.4, qui relie le comportement asymptotique de cette norme à celui d'un produit infini de matrices indexé par le développement binaire de s .

Dans la section IV.5.2, on répond à la deuxième partie de la question de Kirillov : on fournit un algorithme qui permet de calculer la dérivée de u en tout point rationnel.

Dans la section IV.5.3, on répond à la question dans le cas général : on montre que la dérivée de u , quand elle existe, ne peut valoir que 0 ou ∞ , et que presque sûrement elle vaut 0.

Enfin, dans la section IV.5.4, on donne quelques valeurs numériques, sur lesquelles on s'appuie pour faire quelques estimations. On établit une condition suffisante simple pour que la dérivée soit égale à 0, et on en conjecture une pour que la dérivée soit égale à ∞ .

Deuxième partie.

Domaines fondamentaux pour des
groupes affines proprement
discontinus

II.1. Introduction

Dans cette partie, on étudie une famille de contre-exemples à la conjecture de Milnor (la conjecture I.1.4), donnée par Abels et al. dans [2]. Il s'agit de groupes libres discrets agissant proprement sur l'espace affine \mathbb{R}^{2d+1} (où d est un entier impair quelconque) et préservant une forme quadratique de signature $(d+1, d)$. Cela généralise le premier contre-exemple [16] de Margulis, qui était le cas particulier où $d = 1$. Les démonstrations qui apparaissent dans [16] et [2] ne sont pas constructives : elles ne font pas explicitement apparaître le domaine fondamental. Cette lacune a été partiellement comblée par Drumm dans sa thèse doctorale et les articles qui ont suivi [9, 10]. Il y a construit des domaines fondamentaux pour les groupes Γ introduits par Margulis (en signature $(2, 1)$), ce qui lui a permis en particulier de déterminer la topologie du quotient \mathbb{R}^3/Γ .

Le but de cette partie est d'adapter la construction de Drumm aux groupes d'Abels-Margulis-Soifer. Comme Drumm, on déterminera ainsi la topologie de l'espace quotient. Le résultat principal de cette partie est le théorème II.5.2. Sa formulation précise s'appuie sur plusieurs définitions introduites progressivement ; en voici un résumé :

Théorème principal (de la partie II). *Soit d un entier positif impair. Alors tout sous-groupe de Schottky généralisé de $\mathrm{SO}(d+1, d)$ avec des générateurs suffisamment contractants a un ensemble ouvert non vide de déformations affines Γ agissant proprement sur $\mathbb{R}^{d+1, d}$, avec le quotient $\mathbb{R}^{d+1, d}/\Gamma$ homéomorphe à un corps à anses plein $(2d+1)$ -dimensionnel.*

Je me suis bien sûr en grande partie inspiré du travail d'Abels et al. [2], ainsi que d'un article de Charette et Goldman [8] exposant les résultats de Drumm.

Le domaine fondamental de Drumm est une partie de \mathbb{R}^3 délimitée par des « crooked planes » — des surfaces composées de deux demi-plans ayant un point commun, et d'une portion de plan qui relie les bords de ces demi-plans. Cette construction ne se généralise pas telle quelle aux dimensions supérieures. Pour pouvoir augmenter la dimension, on va remplacer les « crooked planes » par des « voisinages coniques » de certains demi-plans (à savoir des « ailes positives » définies dans la section II.2.3).

Pour un plan de la partie section par section, on renvoie au sommaire général de la thèse.

II.2. Conventions, définitions et propriétés basiques

Soient p et q deux entiers strictement positifs. On note $\mathbb{R}^{p, q}$ l'espace \mathbb{R}^{p+q} muni d'une forme quadratique Q de signature (p, q) (p coefficients 1 et q coefficients -1). Le groupe des automorphismes de $\mathbb{R}^{p, q}$ (c'est-à-dire des automorphismes de \mathbb{R}^{p+q} qui préservent la forme quadratique) est $\mathrm{O}(p, q)$. Ce groupe a quatre composantes connexes ; on appelle $\mathrm{SO}^+(p, q)$ sa composante neutre.

On introduit un peu plus de structure sur $\mathbb{R}^{p, q}$. On choisit dans $\mathbb{R}^{p, q}$ un sous-espace défini positif maximal S , et on pose $T = S^\perp$ le sous-espace défini négatif maximal correspondant. On peut alors définir les projections orthogonales $\pi_S : \mathbb{R}^{p, q} \rightarrow S$ et $\pi_T : \mathbb{R}^{p, q} \rightarrow T$, et les formes définies positives $N_S := Q|_S$ et $N_T := -Q|_T$, de sorte que

$$\forall x \in \mathbb{R}^{p, q}, \quad Q(x) = N_S(\pi_S(x)) - N_T(\pi_T(x)). \quad (\text{II.2.1})$$

II.2.1. Sous-espaces totalement isotropes maximaux

Désormais l'acronyme SETIM désigne un sous-espace totalement isotrope maximal. Si V est un SETIM de $\mathbb{R}^{p, q}$, alors (en supposant que $p \geq q$) on a $\dim V = q$, $V \subset V^\perp$ et $\dim V^\perp = p$. On note \mathcal{L} l'ensemble de tous les SETIM.

Un outil très pratique pour étudier les SETIM est la bijection suivante entre \mathcal{L} et l'espace $O(T, S)$ des applications linéaires orthogonales entre T et S (vus comme des espaces euclidiens grâce aux formes N_S et N_T) :

$$\begin{array}{ccc} \mathcal{L} & \xleftarrow{\sim} & O(T, S) \\ V & \longmapsto & f_V := \pi_S \circ (\pi_T|_V)^{-1} \\ V_f := \{t + f(t) \mid t \in T\} & \longleftarrow & f \end{array} \quad (\text{II.2.2})$$

Il est immédiat de vérifier que ces deux applications sont bien définies et inverses l'une de l'autre. En effet, pour tout $V \in \mathcal{L}$ et $f \in O(T, S)$, on a :

- $\pi_T|_V$ est bijective. En effet, comme $V \cap T^\perp = V \cap S = \emptyset$, cette application est injective, et les espaces T et V ont même dimension.
- $f_V \in O(T, S)$. En effet, soit $t \in T$; on pose $v := (\pi_T|_V)^{-1}(t)$. Alors $v \in V$, et on a $0 = Q(v) = N_S(\pi_S(v)) - N_T(\pi_T(v)) = N_S(f_V(t)) - N_T(t)$.
- V_f est un SETIM. En effet, cet espace est de dimension q , et pour tout $t \in T$, on a $Q(t + f(t)) = N_S(f(t)) - N_T(t) = 0$.
- $V_{f_V} = V$. En effet, soit $v \in V$; alors on a $v = \pi_T(v) + \pi_S(v) = \pi_T(v) + f_V(\pi_T(v))$, d'où $v \in V_{f_V}$; et on sait que V et V_{f_V} ont la même dimension.
- $f_{V_f} = f$. En effet, soit $t \in T$; alors on a $f_{V_f}(t) = \pi_S(t + f(t)) = f(t)$.

Voici une première utilisation de cette bijection. Plus tard dans cette partie, on démontrera certaines propriétés des familles de $2n$ SETIM deux à deux transverses. Il est prudent de vérifier que l'ensemble de telles familles n'est pas vide. Cela pourrait sembler évident, mais il se trouve que, bien que ce soit le cas pour les valeurs particulières de p et q avec lesquelles on travaillera, c'est faux en général :

Lemme II.2.1. *Soient p et q deux entiers avec $p \geq q \geq 1$. Alors on peut trouver dans $\mathbb{R}^{p,q}$ une infinité de SETIM deux à deux transverses, **sauf** si $p = q$ et p est impair, auquel cas on ne peut pas en trouver plus que deux.*

Démonstration. Soient V_1 et V_2 deux SETIM, et soient $f_i := f_{V_i}$ leurs images par la bijection (II.2.2). Nous affirmons que V_1 et V_2 sont transverses si et seulement si $f_1 - f_2$ est injective. En effet, on a

$$x \in V_1 \cap V_2 \iff \exists t \in T, \quad x = t + f_1(t) = t + f_2(t),$$

d'où $V_1 \cap V_2 = 0 \iff \ker(f_1 - f_2) = 0$.

La question devient alors : combien peut-on trouver d'applications orthogonales de T vers S — ou alors, de façon équivalente, de \mathbb{R}^q vers \mathbb{R}^p — telles que leurs différences deux à deux soient injectives, ou encore telles que tout vecteur non nul ait des images deux à deux distinctes par ces applications ?

Supposons d'abord qu'il existe un entier pair non nul r tel que $q \leq r \leq p$. Soit $f_0 : \mathbb{R}^q \rightarrow \mathbb{R}^p$ n'importe quelle application orthogonale (donc injective), et soit E n'importe quel espace vectoriel de dimension r tel que $f_0(\mathbb{R}^q) \subset E \subset \mathbb{R}^p$. Alors on peut trouver dans $O(E)$ un sous-groupe infini dont les éléments non-triviaux n'ont pas de points fixes : par exemple, le groupe G formé des matrices

$$\begin{pmatrix} R_\theta & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & R_\theta \end{pmatrix}$$

(où $R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$), avec θ qui varie dans \mathbb{R} . Considérons alors l'ensemble de toutes les applications $g \circ f_0$ avec $g \in G$. Soit $x \in \mathbb{R}^q \setminus \{0\}$: alors $f_0(x) \neq 0$, et les images de $f_0(x)$

par les éléments de G sont deux à deux distinctes. On en déduit que toutes les applications $g \circ f_0$ ont effectivement deux à deux des différences injectives.

Sinon, on a $p = q$ avec p est impair. L'identité et l'application $x \mapsto -x$ sont deux applications de $O(p)$ de différence injective. Prenons maintenant trois applications quelconques dans $O(p)$. Alors au moins deux d'entre elles, disons f_1 et f_2 , ont le même déterminant : cela veut dire que $f_1 \circ f_2^{-1} \in SO(p)$. Mais pour p impair, toute application de $SO(p)$ a un point fixe. On en déduit que $f_1 - f_2$ n'est pas injective. \square

II.2.2. Applications pseudohyperboliques et cadres

À partir de maintenant, on fixe un entier $d > 0$ et on pose $(p, q) = (d + 1, d)$. Prenons une application $g \in GL(\mathbb{R}^{d+1, d})$. Alors on peut décomposer $\mathbb{R}^{d+1, d}$ en une somme directe de trois espaces $\mathbb{R}^{d+1, d} = V_{<}(g) \oplus V_{=}(g) \oplus V_{>}(g)$ stables par g et tels que toute valeur propre λ de $g|_{V_{<}(g)}$ (resp. $V_{=}$, $V_{>}$) vérifie $|\lambda| < 1$ (resp. $= 1$, > 1).

Définition II.2.2. On dira que g est *pseudohyperbolique* si $g \in O(d + 1, d)$, $\dim V_{=}(g) = 1$ et la valeur propre de g dans $V_{=}(g)$ est 1 (pas -1). (Comme on le verra bientôt, toutes les applications pseudohyperboliques appartiennent en fait à $SO(d + 1, d)$). Dans ce cas, on appelle *cadre* de g le couple $\mathcal{V}(g) := (V_{<}(g), V_{>}(g))$, et *partie dynamique* de g (par opposition au cadre, qui est la « partie géométrique ») l'application $g_{<} := g|_{V_{<}(g)}$.

Alors une application pseudohyperbolique est uniquement déterminée par son cadre et sa partie dynamique. Cependant, ces derniers doit vérifier certaines conditions. Pour les énoncés, on aura besoin de la notation suivante : pour toute application linéaire g , on note $\rho(g)$ son *rayon spectral*, c'est-à-dire la plus grande valeur absolue d'une valeur propre de g .

Proposition II.2.3. *Les applications pseudohyperboliques sont en correspondance bijective (donnée la définition précédente) avec les triplets $(V_{<}, V_{>}, g_{<})$ où $V_{<}$ et $V_{>}$ sont deux SETIM transverses et $g_{<}$ est un automorphisme de $V_{<}$ avec $\rho(g_{<}) < 1$.*

Démonstration. Commençons par s'assurer que le cadre et la partie dynamique de toute application pseudohyperbolique g vérifient effectivement les conditions requises. En effet :

- Le fait que $g_{<}$ soit un automorphisme de $V_{<}$ et la majoration de son rayon spectral se déduisent immédiatement de la définition de $V_{<}$.
- Par définition également, $V_{<}(g) \cap V_{>}(g) = 0$.
- Soit $x_{<} \in V_{<}(g)$. Alors on a

$$Q(x_{<}) = \lim_{n \rightarrow +\infty} Q(g^n(x_{<})) = Q\left(\lim_{n \rightarrow +\infty} g_{<}^n(x_{<})\right) = 0,$$

puisque $\rho(g_{<}) < 1$. Cela montre que $V_{<}$ est un sous-espace totalement isotrope.

- De même, en utilisant g^{-1} à la place de g , on peut montrer que $V_{>}$ est totalement isotrope. D'autre part comme $\mathbb{R}^{d+1, d} = V_{<} \oplus V_{=} \oplus V_{>}$, on a

$$2d + 1 = \dim V_{<} + \dim V_{=} + \dim V_{>} \leq d + 1 + d = 2d + 1,$$

donc l'inégalité est une égalité, ce qui veut dire que $V_{<}$ et $V_{>}$ ont la dimension maximale.

Soient maintenant $V_{<}$ et $V_{>}$ n'importe quel couple de SETIM transverses et soit g n'importe quel automorphisme de $V_{<}$ avec $\rho(g_{<}) < 1$. Montrons qu'il existe au plus une application pseudohyperbolique de cadre $(V_{<}, V_{>})$ et de partie dynamique $g_{<}$. En effet, soit g une telle application. On peut alors calculer $V_{<}(g)$, $V_{=}(g)$, $V_{>}(g)$ et les restrictions de g à ces sous-espaces, ce qui détermine g de façon unique. En effet :

- Par définition, $V_{<}(g) = V_{<}$ et $V_{>}(g) = V_{>}$.

- Soient $x_< \in V_<$, $x_ = \in V_=(g)$. Alors on a (en notant $\langle \bullet, \bullet \rangle$ la forme bilinéaire correspondant à la forme quadratique Q) :

$$\langle x_<, x_ = \rangle = \lim_{n \rightarrow +\infty} \langle g^n(x_<), g^n(x_ =) \rangle = \left\langle \lim_{n \rightarrow +\infty} g_<^n(x_<), x_ = \right\rangle = 0.$$

Cela montre que $V_=(g) \perp V_<$. De la même façon, on obtient $V_=(g) \perp V_>$; donc $V_=(g) \subset V_<^\perp \cap V_>^\perp$. Mais clairement, l'espace de la partie droite est de dimension au plus 1 ; donc $V_=(g) = V_<^\perp \cap V_>^\perp$.

- Par définition, $g|_{V_<} = g_<$ et $g|_{V_ =}$ est l'identité.
- Pour $x \in \mathbb{R}^{d+1,d}$, on appelle $x_<, x_ =, x_>$ les composantes de x respectivement dans $V_<(g)$, $V_=(g)$, $V_>(g)$ (de sorte que $x = x_< + x_ = + x_>$). Pour tout x , puisque $Q(x_<) = Q(x_>) = \langle x_<, x_ = \rangle = \langle x_>, x_ = \rangle = 0$, on a

$$Q(x) = 2\langle x_<, x_> \rangle + Q(x_ =).$$

Si on applique g , on obtient alors :

$$Q(g(x)) = 2\langle g_<(x_<), g_>(x_>) \rangle + Q(x_ =),$$

donc pour tous $x_< \in V_<$ et $x_> \in V_>$, on a $\langle x_<, x_> \rangle = \langle g_<(x_<), g_>(x_>) \rangle$. On en déduit que $g_>$ est l'application adjointe de $g_<^{-1}$. Plus rigoureusement, on a

$$g_> = \Phi_{\mathcal{V}}^{-1} \circ (g_<^{-1})^* \circ \Phi_{\mathcal{V}}, \quad (\text{II.2.3})$$

où $\Phi_{\mathcal{V}} : V_> \rightarrow V_<^*$ est une restriction et un quotient appropriés de l'isomorphisme canonique $\Phi_Q : \mathbb{R}^{d+1,d} \rightarrow (\mathbb{R}^{d+1,d})^*$ défini par $\Phi_Q(x) \cdot y = \langle x, y \rangle$. Cela détermine $g_>$ de façon unique.

Finalement, soit $V_ = := V_<^\perp \cap V_>^\perp$. Alors $\dim V_ = = 1$ et $\mathbb{R}^{d+1,d} = V_< \oplus V_ = \oplus V_>$. Considérons l'application $g := g_< \oplus \text{Id}_{V_ =} \oplus g_>$, avec $g_>$ défini par (II.2.3). Alors il est immédiat de vérifier que g est une application pseudohyperbolique de cadre $(V_<, V_>)$ et de partie dynamique $g_<$. (Notez que de la formule (II.2.3), on déduit que les valeurs propres de $g_>$ sont les inverses des valeurs propres de $g_<$). \square

Au passage, on peut maintenant montrer — comme on l'avait annoncé — que toutes les applications pseudohyperboliques g appartiennent à $\text{SO}(d+1, d)$. En effet, pour tout tel g , on a $\det g_> = (\det g_<)^{-1}$, d'où $\det g = (\det g_<)(\det \text{Id})(\det g_>) = 1$.

Définition II.2.4. On appelle *cadre* (de façon abstraite) un couple de SETIM transverses. Si \mathcal{V} est un cadre, on note :

- $V_<$ sa première composante et $V_>$ sa deuxième composante ;
- $V_ =$ la droite $V_<^\perp \cap V_>^\perp$;
- $V_ \leq := V_<^\perp = V_< \oplus V_ =$ et $V_ \geq := V_>^\perp = V_> \oplus V_ =$.

II.2.3. Orientation

Proposition II.2.5. *On peut choisir une orientation sur tous les SETIM V (resp. sur leurs orthogonaux V^\perp) telle que toute application $f \in \text{SO}^+(d+1, d)$ induise un isomorphisme direct de V sur $f(V)$ (resp. de V^\perp sur $f(V^\perp) = f(V)^\perp$).*

Démonstration. On commence par traiter le cas des orthogonaux des SETIM. On fixe une certaine orientation sur S et une certaine orientation sur T (on rappelle que ce sont deux sous-espaces orthogonaux, l'un défini positif maximal et l'autre défini négatif maximal). Alors pour tout SETIM V , π_S induit une bijection linéaire de V^\perp sur S . En effet, les deux espaces ont dimension $d+1$, et on a

$$\ker \pi_S|_{V^\perp} = V^\perp \cap \ker \pi_S = V^\perp \cap T = \{0\},$$

puisque V^\perp est un sous-espace positif et T un sous-espace défini négatif. On choisit alors sur V^\perp l'orientation qui fait de $\pi_S|_{V^\perp}$ un isomorphisme direct.

Considérons maintenant l'application de S dans lui-même donnée par la composition de

$$S \xrightarrow{\pi_S^{-1}} V^\perp \xrightarrow{f} f(V^\perp) \xrightarrow{\pi_S} S.$$

Il est facile de voir que son déterminant dépend continûment de f et ne s'annule jamais pour $f \in \text{SO}^+(d+1, d)$. Puisque $\text{SO}^+(d+1, d)$ est connexe, ce déterminant doit avoir un signe constant, d'où le résultat.

En remplaçant S par T , le même argument s'adapte aux SETIM eux-mêmes. \square

On fixe désormais une telle famille d'orientations.

Définition II.2.6. L'aile positive construite sur un SETIM V est le demi-espace

$$V^{\mathbf{L}} := \{v + xe \mid v \in V, x \geq 0\},$$

où $e \in V^\perp$ est n'importe quel vecteur qui complète toute base directe (e_1, \dots, e_d) de V en une base directe (e_1, \dots, e_d, e) de V^\perp . (Le symbole \mathbf{L} se lit « demi-orthogonal » ; il est censé représenter la moitié du symbole \perp .)

Proposition II.2.7. Si d est impair, les ailes positives construites sur deux SETIM transverses $V_<$ et $V_>$ ont toujours une intersection triviale :

$$V_<^{\mathbf{L}} \cap V_>^{\mathbf{L}} = \{0\}.$$

Démonstration. Soit $\mathcal{B}_> = (e_>^1, \dots, e_>^d)$ une base directe de $V_>$. On pose

$$\mathcal{B}_< = (e_<^1, \dots, e_<^d) := \Phi_V^{-1}(-\mathcal{B}_>^*)$$

la base de $V_>$ duale à la base $-\mathcal{B}_> = (-e_>^1, \dots, -e_>^d)$ (voir la preuve de la proposition II.2.3 pour la définition de Φ_V). On pose également e_- le vecteur de norme unité se trouvant dans V_- tel que $(\mathcal{B}_>, e_-)$ soit une base directe de $V_>^\perp = V_> \oplus V_-$. On définit alors une base de $\mathbb{R}^{d+1, d}$ en concaténant ces bases de $V_<$, $V_>$ et V_- :

$$\mathcal{B} := (e_<^1, \dots, e_<^d, e_>^1, \dots, e_>^d, e_-).$$

Dans cette base, la forme quadratique Q est alors donnée par la matrice

$$\begin{pmatrix} 0 & -I_d & 0 \\ -I_d & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Considérons maintenant l'automorphisme f donné, en base \mathcal{B} , par la matrice

$$\begin{pmatrix} 0 & I_d & 0 \\ I_d & 0 & 0 \\ 0 & 0 & (-1)^d \end{pmatrix}.$$

Il est facile de voir que $f \in \text{SO}^+(d+1, d)$ (pour les détails, voir dans [2] la démonstration du lemme 3.1 ; ils appellent cette application h_π). Mais f envoie $\mathcal{B}_>$ sur $\mathcal{B}_<$ et $(\mathcal{B}_>, e_-)$ sur $(\mathcal{B}_<, (-1)^d e_-)$. Par la proposition II.2.5, on en déduit que ces dernières sont des bases directes de $V_>$ et $V_>^\perp$. Ceci implique que

$$\begin{cases} V_<^{\mathbf{L}} = V_< + (-1)^d \mathbb{R}^{\geq 0} e_- \\ V_>^{\mathbf{L}} = V_> + \mathbb{R}^{\geq 0} e_- \end{cases}$$

Donc

$$V_<^{\mathbf{L}} \cap V_>^{\mathbf{L}} = \left((-1)^d \mathbb{R}^{\geq 0} \cap \mathbb{R}^{\geq 0} \right) e_-;$$

puisque d est impair, on peut conclure. \square

Remarque II.2.8. Supposons maintenant que d soit pair. Alors le même argument montre que deux ailes positives ont *toujours* une intersection non triviale. Avec nos méthodes, il n'y a donc aucun espoir de construire un sous-groupe libre non abélien proprement discontinu dans $\mathrm{SO}(d+1, d) \times \mathbb{R}^{2d+1}$ pour d pair, puisqu'un point crucial de la construction est l'existence de $2n$ ailes positives deux à deux disjointes (avec $n > 1$). Et effectivement, le théorème A de [2] affirme que de tels sous-groupes n'existent pas; sa preuve utilise un argument d'orientation très similaire.

Définition II.2.9. Pour tout cadre \mathcal{V} (ou \mathcal{V}' , \mathcal{V}_i et ainsi de suite), on note $e_=-$ (resp. $e'_=$, $e_{i,=}$ et ainsi de suite) le vecteur de norme unité contenu dans la demi-droite $V_-= \cap V_>^L$. Si d est impair, on a alors :

$$\begin{cases} V_<^L = V_< - \mathbb{R}^{\geq 0}e_=- \\ V_>^L = V_> + \mathbb{R}^{\geq 0}e_=-. \end{cases}$$

II.2.4. Force de contraction et autres considérations métriques

On suppose dorénavant que d est impair. On introduit sur $\mathbb{R}^{d+1, d}$, en plus de sa forme quadratique structurelle Q , plusieurs formes quadratiques définies positives. Toute telle forme N donne un produit scalaire (noté $\langle x, y \rangle_N$), une norme euclidienne (notée $\|x\|_N := N(x)^{\frac{1}{2}}$; donc aussi une métrique sur $\mathbb{R}^{d+1, d}$), et une norme d'opérateur (notée également $\|g\|_N := \sup \frac{\|g(x)\|_N}{\|x\|_N}$).

On a d'abord besoin d'une norme « globale », qu'on utilisera la plupart du temps : elle nous permettra de faire des mesures qui ne dépendent pas d'un cadre particulier. Dans la mesure où toutes les normes sur un espace de dimension finie sont équivalentes, son choix n'a pas vraiment d'importance; cependant, l'expression particulière suivante simplifiera certaines preuves. On définit la forme N_0 par

$$\forall x \in \mathbb{R}^{d+1, d}, \quad N_0(x) = N_S(\pi_S(x)) + N_T(\pi_T(x)) \quad (\text{II.2.4})$$

(comparer ceci avec (II.2.1)).

D'un autre côté, pour tout cadre \mathcal{V} , on a aussi besoin d'une norme « locale », qui simplifiera les calculs impliquant ce cadre particulier. On pose $N_{\mathcal{V}}$ la forme quadratique (définie positive) de $\mathbb{R}^{d+1, d}$ qui rend les espaces $V_<$, $V_=-$ et $V_>$ deux à deux orthogonaux, mais dont la restriction à chacun de ces espaces coïncide avec N_0 .

Considérons un espace vectoriel E (pour le moment, le lecteur peut supposer que $E = \mathbb{R}^{d+1, d}$; plus tard, on aura également besoin du cas où $E = \Lambda^d \mathbb{R}^{d+1, d}$). On définit

$$\begin{cases} \pi_{\mathbb{S}} : E \setminus \{0\} \rightarrow \mathbb{S}(E) \\ \pi_{\mathbb{P}} : E \setminus \{0\} \rightarrow \mathbb{P}(E) \end{cases}$$

comme étant, respectivement, les projections canoniques sur la sphère $\mathbb{S}(E) := (E \setminus \{0\})/\mathbb{R}^{>0}$ et sur l'espace projectif $\mathbb{P}(E) := (E \setminus \{0\})/\mathbb{R}^*$. (Envisager la sphère comme une partie de E risque d'engendrer des confusions lorsqu'on va changer de norme; c'est pour cette raison qu'on la définit comme un espace quotient abstrait.) Pour toute application linéaire $g : E \rightarrow E$, on définit les applications correspondantes $g_{\mathbb{S}} : \mathbb{S}(E) \rightarrow \mathbb{S}(E)$ et $g_{\mathbb{P}} : \mathbb{P}(E) \rightarrow \mathbb{P}(E)$ (notées simplement g lorsqu'il n'y a pas de confusion possible).

Prenons un espace métrique (\mathcal{M}, δ) ; soient X et Y deux parties de \mathcal{M} . La distance ordinaire, ou distance « minimum », entre X et Y sera notée

$$\delta(X, Y) := \inf_{x \in X} \inf_{y \in Y} \delta(x, y),$$

par opposition à la distance de Hausdorff, qui sera notée

$$\delta^{\text{Haus}}(X, Y) := \max \left(\sup_{x \in X} \delta(x, Y), \sup_{y \in Y} \delta(y, X) \right).$$

Pour toute forme quadratique définie positive N sur E , pour tous $\bar{x}, \bar{y} \in \mathbb{S}(E)$, on définit la distance

$$\alpha_N(\bar{x}, \bar{y}) := \arccos \frac{\langle x, y \rangle_N}{\|x\|_N \|y\|_N},$$

où x et y sont n'importe quels représentants de \bar{x} et \bar{y} (clairement, la valeur ci-dessus ne dépend pas du choix des représentants). Elle mesure l'angle entre les demi-droites \bar{x} et \bar{y} . Par souci de concision, on se contentera généralement de noter simplement $\alpha_N(x, y)$ avec $x, y \in E \setminus \{0\}$, pour désigner $\alpha_N(\pi_{\mathbb{S}}(x), \pi_{\mathbb{S}}(y))$.

De même, on munit $\mathbb{P}(E)$ de la distance

$$\alpha_N^{\text{Proj}}(x, y) := \alpha_N(\mathbb{R}x, \mathbb{R}y) = \min(\alpha_N(x, y), \alpha_N(x, -y)).$$

Notez que si X et Y sont des parties symétriques par rapport à l'origine (notamment des espaces vectoriels), on a $\alpha_N^{\text{Proj}}(X, Y) = \alpha_N(X, Y)$: dans cette situation, on peut ignorer la distinction entre le cas sphérique et le cas projectif.

Pour toute partie $X \subset \mathbb{S}(E)$ et tout rayon $\varepsilon > 0$, on notera

$$B_N(X, \varepsilon) := \{x \in \mathbb{S}(E) \mid \alpha_N(x, X) < \varepsilon\}.$$

l' ε -voisinage de X selon la distance α_N . Lorsque X est symétrique, on traitera parfois $B_N(X, \varepsilon)$ comme une partie de $\mathbb{P}(E)$.

Pour éviter des notations trop lourdes, on va souvent spécifier une forme quadratique « par défaut » au début de certaines sections ou certaines sous-sections. Dans le reste de cette section ou cette sous-section, il est sous-entendu que *toute* mention de *n'importe laquelle* des valeurs ou fonctions dépendant d'une métrique définies ci-dessus sans mention explicite de la métrique elle-même (comme $\langle x, y \rangle$, $\alpha(x, y)$, $B(X, \varepsilon)$ et ainsi de suite) fait référence à la métrique « par défaut » du moment.

On introduit finalement la notation suivante. Soient A et B deux quantités strictement positives, et p_1, \dots, p_k des paramètres. Lorsqu'on écrit

$$A \lesssim_{p_1, \dots, p_k} B,$$

on veut dire qu'il existe une constante C , qui dépend de p_1, \dots, p_k et de rien d'autre, telle que $A \leq CB$. (Si aucun paramètre n'est écrit, cela signifie bien sûr que C est une constante absolue.) Lorsqu'on écrit

$$A \gtrsim_{p_1, \dots, p_k} B,$$

on veut dire que $A \lesssim_{p_1, \dots, p_k} B$ et $B \lesssim_{p_1, \dots, p_k} A$ à la fois.

Définition II.2.10. Soit g une application pseudohyperbolique, \mathcal{V} son cadre, $g_{<} = g|_{V_{<}(g)}$ sa partie dynamique, $g_{>} = g|_{V_{>}(g)}$. Puisque $V_{<}$ et $V_{>}$ sont transverses, d'après la proposition II.2.7, $V_{<}^{\perp}$ et $V_{>}^{\perp}$ ont une intersection nulle. Leurs projections sur la sphère sont alors disjointes ; étant compactes, elles sont toujours séparées par une distance non nulle. On appelle *séparation* de \mathcal{V} (ou, par abus de terminologie, de g) le nombre

$$\varepsilon(g) = \varepsilon(\mathcal{V}) := \alpha_{N_0}(V_{<}^{\perp}, V_{>}^{\perp}),$$

à savoir la distance entre les projections mentionnées ci-dessus dans la métrique globale. (La distance dans la métrique locale, $\alpha_{N_{\mathcal{V}}}(V_{<}^{\perp}, V_{>}^{\perp})$, est par définition toujours égale à $\frac{\pi}{2}$). Pour toute constante $\varepsilon > 0$, on dit que \mathcal{V} (ou g) est ε -séparée si $\varepsilon(\mathcal{V}) \geq \varepsilon$.

La *force de contraction* de g est la quantité

$$s(g) := \max(\|g_{<}\|, \|g_{>}^{-1}\|)$$

(dans la métrique donnée indifféremment par N_0 ou $N_{\mathcal{V}(g)}$: les deux coïncident sur $V_<(g)$ et $V_>(g)$.) Pour $s > 0$, on dit que g est s -contractante si $s(g) \leq s$. Dans ce cas, pour tous $x_< \in V_<(g)$ et $x_> \in V_>(g)$, on a

$$\frac{\|g(x_<)\|}{\|x_<\|} \leq s \text{ et } \frac{\|g(x_>)\|}{\|x_>\|} \geq s^{-1}.$$

Notez que si $d > 1$, il n'y a pas de constante C telle que toute application pseudohyperbolique soit C -contractante, car la norme peut être beaucoup plus grande que le rayon spectral. En revanche, pour toute application pseudohyperbolique g , on a

$$s(g^n) = \underset{n \rightarrow \infty}{\text{O}} (\rho(g_<)^n) \underset{n \rightarrow \infty}{\rightarrow} 0.$$

Nous allons maintenant formuler une propriété essentielles des métriques définies au-dessus, qu'on utilisera très souvent par la suite. Toutes les normes $\|\bullet\|_{N_{\mathcal{V}}}$ et les distances associées $\alpha_{N_{\mathcal{V}}}$ sont Lipschitz-équivalentes, avec une constante de Lipschitz commune qui ne dépend que de la séparation de \mathcal{V} . Plus précisément :

Lemme II.2.11. *Pour tout $\varepsilon > 0$ et pour tout cadre ε -séparé \mathcal{V} , on a :*

$$\begin{aligned} \forall x \in \mathbb{R}^{d+1,d}, \quad \|x\|_{N_{\mathcal{V}}} \asymp_{\varepsilon} \|x\|_{N_0}; \\ \forall x, y \in \mathbb{S}(\mathbb{R}^{d+1,d}), \quad \alpha_{N_{\mathcal{V}}}(x, y) \asymp_{\varepsilon} \alpha_{N_0}(x, y). \end{aligned}$$

Démonstration. Pour tout cadre \mathcal{V} , soit $C(\mathcal{V})$ la constante de Lipschitz entre les normes données par N_0 et $N_{\mathcal{V}}$, à savoir la plus petite constante qui satisfait la première inégalité ci-dessus. Alors $C(\mathcal{V})$ est toujours finie, et peut s'exprimer comme la norme d'opérateur de l'application identité subordonnée aux normes donnée par N_0 et $N_{\mathcal{V}}$: elle dépend donc continûment de \mathcal{V} . Sachant que pour tout $\varepsilon > 0$ fixé, l'ensemble des cadres ε -séparés est compact, la première affirmation en découle.

Maintenant si deux normes données par N et N' sont C -Lipschitz-équivalentes, alors les distances correspondantes α_N et $\alpha_{N'}$ sont toujours C^2 -Lipschitz-équivalentes. En effet, en dimension 2, cela découle d'un calcul immédiat ; dans le cas général, on peut simplement fixer deux vecteurs x et y et faire le calcul dans l'espace vectoriel qu'ils engendrent. La deuxième estimation découle alors de la première. \square

II.3. Pseudohyperbolicité des produits

Le but de cette section est de démontrer la proposition II.3.8, qui dit essentiellement que sous certaines conditions, le produit de plusieurs applications pseudohyperboliques est encore pseudohyperbolique.

II.3.1. Cas proximal

Soit E un espace vectoriel. On fixe sur E une forme quadratique par défaut \hat{N}_0 . (En pratique, on appliquera les résultats de cette sous-section à $E = \Lambda^d \mathbb{R}^{d+1,d}$.)

Notre premier objectif est de démontrer le lemme II.3.4, qui est analogue à la proposition II.3.8 (et sera utilisé pour la démontrer), mais avec des applications proximales à la place d'applications pseudohyperboliques. On commence par quelques définitions.

Définition II.3.1. Soit $f \in \text{GL}(E)$. Soit λ une valeur propre de f de valeur absolue maximale. On dit que f est *proximal* si λ est unique et a multiplicité 1. On peut alors décomposer E en somme directe d'une droite $V_s(f)$, appelée son *espace attractif*, et d'un hyperplan $V_u(f)$, appelé son *espace répulsif*, tous deux stables par f et tels que :

$$\begin{cases} f|_{V_s} = \pm \lambda \text{Id} \\ \text{pour toute valeur propre } \mu \text{ de } f|_{V_u}, |\mu| < |\lambda|. \end{cases}$$

On appelle *separation* de f le nombre $\eta(f) := \alpha(V_s(f), V_u(f))$. Pour toute constante $\eta > 0$, on dit que f est η -séparée si $\eta(f) \geq \eta$. Pour toute forme quadratique N sur E , on définit la *force de contraction* de f par rapport à N par

$$\hat{s}_N(f) := \frac{\|f|_{V_u}\|_N}{|\lambda|}$$

(on rappelle que \hat{s} tout court signifie $\hat{s}_{\hat{N}_0}$.) Notez que ces définitions diffèrent de celles qu'on a utilisées dans le contexte des applications pseudohyperboliques (d'où la nouvelle notation \hat{s}).

Définition II.3.2. Un *système proximal indépendant* est un n -uplet $F = (f_1, \dots, f_n)$ d'applications $f_i \in \text{GL}(E)$ telles que :

- (i) chaque f_i et chaque f_i^{-1} est proximal ;
- (ii) pour tous les indices i, i' et pour tous les signes σ, σ' tels que $(i', \sigma') \neq (i, -\sigma)$, on a

$$\alpha(V_s(f_i^\sigma), V_u(f_{i'}^{\sigma'})) > 0.$$

Dans ce cas, on appelle *séparation* de F le nombre

$$\eta(F) := \min_{(i', \sigma') \neq (i, -\sigma)} \alpha(V_s(f_i^\sigma), V_u(f_{i'}^{\sigma'})),$$

et *force de contraction* de F le nombre

$$\hat{s}(F) := \max_{i, \sigma} \hat{s}(f_i^\sigma).$$

Définition II.3.3. Prenons un entier naturel k , et prenons k couples $(i_1, \sigma_1), \dots, (i_k, \sigma_k)$ tels que pour tout l , $1 \leq i_l \leq n$ et $\sigma_l = \pm 1$. Considérons le mot $f = f_{i_1}^{\sigma_1} \dots f_{i_k}^{\sigma_k}$.

On dit que f est *réduit* si pour tout l tel que $1 \leq l \leq k-1$, on a $(i_{l+1}, \sigma_{l+1}) \neq (i_l, -\sigma_l)$. On dit que f est *cycliquement réduit* s'il est réduit et si, en plus, on a $(i_1, \sigma_1) \neq (i_k, -\sigma_k)$.

On démontre maintenant un équivalent de la proposition II.3.8 dans le cas proximal :

Lemme II.3.4. *Pour tout $\eta > 0$, il existe une constante $\hat{s}(\eta) > 0$ avec la propriété suivante. Soit $F = (f_1, \dots, f_n)$ un système proximal indépendant η -séparé et $\hat{s}(\eta)$ -contractant. Soit $f = f_{i_1}^{\sigma_1} \dots f_{i_k}^{\sigma_k}$ (avec $\sigma_l = \pm 1$) n'importe quel mot cycliquement réduit.*

Alors f est proximal, $\hat{s}(f) \lesssim_\eta \hat{s}(F)$ et

$$\alpha(V_s(f), V_s(f_{i_1}^{\sigma_1})) \lesssim_\eta \hat{s}(F).$$

Avant de poursuivre, on a besoin d'un lemme technique qui met en relation la force de contraction abstraite $\hat{s}(f)$ et certaines constantes de Lipschitz de f agissant sur l'espace projectif $\mathbb{P}(E)$. Pour toute partie $X \subset \mathbb{P}(E)$, on introduit la notation suivante pour la constance de Lipschitz de f restreinte à X en métrique donnée par N :

$$\mathcal{L}_N(f, X) := \sup_{\substack{(x,y) \in X^2 \\ x \neq y}} \frac{\alpha_N^{\text{proj}}(f(x), f(y))}{\alpha_N^{\text{proj}}(x, y)}$$

Lemme II.3.5. *Pour tous $\eta > 0$, $\zeta > 0$, pour toute application f proximale et η -séparée, on a :*

$$\mathcal{L}(f, \mathbb{P}(E) \setminus B(V_u(f), \zeta)) \lesssim_{\eta, \zeta} \hat{s}(f) \tag{II.3.1a}$$

$$\hat{s}(f) \lesssim_{\eta, \zeta} \mathcal{L}(f, B(V_s(f), \zeta)) \tag{II.3.1b}$$

(en utilisant bien entendu la métrique donnée par \hat{N}_0 .)

Démonstration. Soient $\eta > 0, \zeta > 0$. Pour toute f proximale, on définit sur E une forme quadratique \hat{N}_f qui rend $V_s(f)$ et $V_u(f)$ orthogonaux mais qui coïncide avec \hat{N}_0 sur ces espaces. Par une généralisation évidente du lemme II.2.11, pour toute application f proximale et η -séparée, on a

$$\alpha_{\hat{N}_f}^{\text{proj}} \asymp_{\eta} \alpha_{\hat{N}_0}^{\text{proj}}. \quad (\text{II.3.2})$$

(le lemme II.2.11 mentionnait α plutôt que α^{proj} , mais comme ces deux distances sont localement égales, cette différence a peu d'importance.) Considérons maintenant une application proximale f , et remarquons les choses suivantes :

– De (II.3.2), on déduit que

$$\begin{aligned} \mathbb{P}(E) \setminus B(V_u(f), \zeta) &\subset \mathbb{P}(E) \setminus B_{\hat{N}_f}(V_s(f), \zeta'), \\ B(V_s(f), \zeta) &\supset B_{\hat{N}_f}(V_s(f), \zeta'), \end{aligned}$$

où $\zeta' = C\zeta$ pour une certaine constante C qui ne dépend que de η . De plus, il est clair que $X \subset Y$ implique $\mathcal{L}(f, X) \leq \mathcal{L}(f, Y)$.

– Pour toute partie X , on a

$$\mathcal{L}_{\hat{N}_0}(f, X) \asymp_{\eta} \mathcal{L}_{\hat{N}_f}(f, X).$$

– Pour tout $\zeta' > 0$, on a

$$\mathcal{L}_{\hat{N}_f}(f, B_{\hat{N}_f}(V_s(f), \zeta')) \asymp_{\zeta'} \hat{s}_{\hat{N}_f}(f).$$

En effet, considérons la projection naturelle $\pi_u : \mathbb{P}(E) \setminus V_u(f) \rightarrow V_u(f)$ définie par $\pi_u(x_u : 1) = x_u$ (avec les notations évidentes). Elle induit un homéomorphisme de $B_{\hat{N}_f}(V_s(f), \zeta')$ sur la boule $\left\{x \in V_u(f) \mid \|x\|_{\hat{N}_f} \leq \frac{1}{\tan \zeta'}\right\}$. Un calcul immédiat montre que cet homéomorphisme est bilipschitzien (par rapport aux métriques $\alpha_{\hat{N}_f}^{\text{proj}}$ et $\|\bullet\|_{\hat{N}_f}$), avec une constante de Lipschitz $C(\zeta')$ qui ne dépend pas du tout de f ni de η . D'autre part, la constante de Lipschitz de la fonction conjuguée $\pi_u \circ f \circ \pi_u^{-1}$ n'est autre que $\hat{s}_{\hat{N}_f}(f)$. Donc f est lipschitzienne avec la constante $C(\zeta')^2 \hat{s}_{\hat{N}_f}(f)$, d'où la conclusion.

– Puisque \hat{N}_f et \hat{N}_0 coïncident sur $V_u(f)$ et $V_s(f)$, on a $\hat{s}_{\hat{N}_f}(f) = \hat{s}_{\hat{N}_0}(f)$.

Pour montrer (II.3.1a), il ne reste plus qu'à appliquer successivement toutes ces étapes, en tenant compte du fait que

$$\mathbb{P}(E) \setminus B_{\hat{N}_f}(V_u(f), \zeta') = B_{\hat{N}_f}(V_s(f), \frac{\pi}{2} - \zeta').$$

Pour montrer (II.3.1b), il suffit d'appliquer les mêmes étapes dans l'ordre inverse. \square

Démonstration du lemme II.3.4. Soit $\eta > 0$, et soit $F = (f_1, \dots, f_n)$ un système proximal indépendant η -séparé et $\hat{s}(\eta)$ -contractant (pour une certaine valeur $\hat{s}(\eta)$ à préciser plus tard).

Un corollaire immédiat du lemme II.3.5 est que pour toute application ϕ proximale et η -séparée et pour tout $\zeta \leq \eta$, on a

$$\phi(\mathbb{P}(E) \setminus B(V_u(\phi), \zeta)) \subset B(V_s(\phi), C(\eta, \zeta) \hat{s}(\phi)) \quad (\text{II.3.3})$$

pour une certaine constante $C(\eta, \zeta)$. En effet, $V_s(\phi) \in \mathbb{P}(E) \setminus B(V_u(\phi), \zeta)$ est un point fixe de ϕ et $\text{diam}(\mathbb{P}(E) \setminus B(V_u(\phi), \zeta)) \leq \frac{\pi}{2} \lesssim 1$.

Soit $\eta' = C(\eta, \frac{\eta}{3}) \hat{s}(F)$. Pour tout l compris entre 1 et k , on pose

$$\begin{cases} X_l^- := B(V_u(f_i^{\sigma_l}), \frac{\eta}{3}) \\ X_l^+ := B(V_s(f_i^{\sigma_l}), \eta'). \end{cases}$$

Alors d'après (II.3.3), pour tout l on a $f_{i_l}^{\sigma_l}(\mathbb{P}(E) \setminus X_l^-) \subset X_l^+$. Comme $\hat{s}(F) \leq \hat{s}(\eta)$, en choisissant $\hat{s}(\eta)$ suffisamment petit, on peut supposer que $\eta' \leq \frac{\eta}{3}$. Alors pour tout l on a aussi $X_l^+ \subset \mathbb{P}(E) \setminus X_{l-1}^-$ (puisque le mot f est réduit). Par récurrence, on en déduit que

$$f(\mathbb{P}(E) \setminus X_k^-) \subset X_1^+.$$

D'autre part on sait, d'après (II.3.1a), que pour tout l on a

$$\mathcal{L}\left(f_{i_l}^{\sigma_l}, \mathbb{P}(E) \setminus X_l^-\right) \lesssim_{\eta} \hat{s}(F) \leq \hat{s}(\eta). \quad (\text{II.3.4a})$$

Encore une fois, en choisissant $\hat{s}(\eta)$ suffisamment petit, on peut en fait supposer que

$$\mathcal{L}\left(f_{i_l}^{\sigma_l}, \mathbb{P}(E) \setminus X_l^-\right) < 1. \quad (\text{II.3.4b})$$

Comme le mot f est cycliquement réduit, on a $X_1^+ \subset \mathbb{P}(E) \setminus X_k^-$; donc X_1^+ est stable par f et, par récurrence, on obtient

$$\mathcal{L}(f, X_1^+) < 1.$$

On en déduit que f est proximale et $V_s(f) \in X_1^+$ (voir le lemme 3.8 dans [22] pour une preuve), ce qui montre le premier et le troisième énoncés de la conclusion. D'autre part, on voit facilement que $V_u(f) \subset X_k^-$ (en effet, considérons n'importe quel point $x \in \mathbb{P}(E)$ appartenant à $V_u(f)$ mais pas à X_k^- : alors on aurait $\lim_{n \rightarrow \infty} f^n(x) = V_s(f)$, ce qui contredit le fait que $V_u(f)$ est un sous-espace stable). Or on sait que

$$\begin{aligned} \alpha(X_1^+, X_k^-) &\geq \alpha(V_s(f_{i_1}^{\sigma_1}), V_u(f_{i_k}^{\sigma_k})) - \eta' - \frac{\eta}{3} \\ &\geq \eta - \frac{\eta}{3} - \frac{\eta}{3} \\ &= \frac{\eta}{3}, \end{aligned}$$

donc f est $\frac{\eta}{3}$ -séparée.

Cela nous permet d'appliquer (II.3.1b) à f :

$$\hat{s}(f) \lesssim_{\eta} \mathcal{L}\left(f, B(V_s(f), \frac{\eta}{3})\right).$$

On sait que $B(V_s(f), \frac{\eta}{3}) \subset B(V_s(f_{i_1}^{\sigma_1}), \frac{2\eta}{3}) \subset \mathbb{P}(E) \setminus X_k^-$, d'où

$$\mathcal{L}\left(f, B(V_s(f), \frac{\eta}{3})\right) \leq \mathcal{L}(f, \mathbb{P}(E) \setminus X_k^-).$$

D'autre part, en utilisant (II.3.4a) combinée avec (II.3.4b), on obtient que

$$\mathcal{L}(f, \mathbb{P}(E) \setminus X_k^-) \lesssim_{\eta} \hat{s}(F).$$

En rassemblant ces inégalités, on obtient

$$\hat{s}(f) \lesssim_{\eta} \hat{s}(F),$$

ce qui montre le deuxième énoncé de la conclusion. \square

II.3.2. Cas pseudohyperbolique

Tout au long de cette section, on travaille par défaut dans la métrique donnée par N_0 .

Définition II.3.6. On définit un *multicadre* \mathcal{W} comme une famille de n cadres $\mathcal{V}_1, \dots, \mathcal{V}_n$ dont les $2n$ composantes $V_{1,<}, V_{1,>}, \dots, V_{n,<}, V_{n,>}$ sont deux à deux transverses. On définit la *séparation* $\varepsilon(\mathcal{W})$ du multicadre comme la plus petite séparation d'une paire de SETIM choisie parmi tous ceux qui composent le cadre.

Soit $\mathcal{W} = (\mathcal{V}_1, \dots, \mathcal{V}_n)$ un multicadre. On appelle *groupe basé sur \mathcal{W}* tout groupe G engendré par des applications pseudohyperboliques g_1, \dots, g_n de cadres respectifs $\mathcal{V}_1, \dots, \mathcal{V}_n$. Pour $s > 0$, on dit que G est *s -contractant* si tous ces générateurs sont s -contractants ; on appelle *force de contraction* de G le nombre

$$s(G) := \max_i s(g_i).$$

Remarque II.3.7.

- Attention : ici, la transversalité et la séparation doivent vraiment être vérifiées sur les $\binom{2n}{2}$ paires de SETIM possibles, pas seulement sur les cadres $\mathcal{V}_1, \dots, \mathcal{V}_n$.
- Le lemme II.2.1 garantit l'existence de multicadres contenant un nombre de cadres arbitrairement grand.

Proposition II.3.8. *Pour tout $\varepsilon > 0$, il existe une constante $s_1(\varepsilon) > 0$ ayant la propriété suivante. Soit \mathcal{W} un multicadre ε -séparé, $G = \langle g_1, \dots, g_n \rangle$ n'importe quel groupe $s_1(\varepsilon)$ -contractant basé sur \mathcal{W} , $g = g_{i_1}^{\sigma_1} \dots g_{i_k}^{\sigma_k}$ (avec $\sigma_l = \pm 1$) n'importe quel mot cycliquement réduit non vide.*

Alors l'application g est pseudohyperbolique, $\frac{\varepsilon}{3}$ -séparée, 1-contractante, et on a

$$\alpha_{N_0}^{\text{Haus}}(V_{>}(g), V_{>}(g_{i_1}^{\sigma_1})) \lesssim_{\varepsilon} s(G).$$

Définition II.3.9. Un groupe vérifiant ces propriétés sera appelé un *groupe pseudohyperbolique*.

Remarque II.3.10. Un groupe pseudohyperbolique est toujours libre. En effet, considérons n'importe quel mot réduit en ses générateurs. Il existe alors un conjugué de ce mot qui est cycliquement réduit, et qui est donc pseudohyperbolique. En particulier ce n'est pas l'identité.

La proposition découle du lemme II.3.4 appliqué à l'espace $E := \Lambda^d \mathbb{R}^{d+1,d}$. En effet, il existe une correspondance entre les applications pseudohyperboliques dans $\mathbb{R}^{d+1,d}$ et les applications proximales dans E , comme on le montrera bientôt.

Pour toute application $g \in L(\mathbb{R}^{d+1,d})$, on définit l'application correspondante $\Lambda^d g \in L(E)$, et pour toute forme quadratique N sur $\mathbb{R}^{d+1,d}$, on définit la forme quadratique correspondante $\Lambda^d N$ sur E par la formule

$$\langle x_1 \wedge \dots \wedge x_d, y_1 \wedge \dots \wedge y_d \rangle_{\Lambda^d N} := \sum_{\sigma \in \mathcal{S}_d} \epsilon_{\sigma} \prod_{i=1}^d \langle x_i, y_{\sigma(i)} \rangle_N$$

(où \mathcal{S}_d est l'ensemble des permutations de $\{1, \dots, d\}$ et ϵ_{σ} désigne la signature de σ). La forme par défaut sur E sera $\hat{N}_0 := \Lambda^d N_0$. On peut maintenant énoncer la correspondance souhaitée :

Lemme II.3.11.

- (i) *Pour $g \in \text{SO}(d+1, d)$, $\Lambda^d g$ est proximale si et seulement si g est pseudohyperbolique. De plus, l'espace attractif (resp. répulsif) de $\Lambda^d g$ ne dépend de rien d'autre que de $V_{>}(g)$ (resp. $V_{<}(g)$) :*

$$\begin{cases} V_s(\Lambda^d g) = \Lambda^d V_{>}(g) \\ V_u(\Lambda^d g) = \{x \in E \mid x \wedge \Lambda^{d+1} V_{<}(g) = 0\}. \end{cases} \quad (\text{II.3.5})$$

(ii) Pour tout $\varepsilon > 0$, il existe une constante $\eta(\varepsilon) > 0$ telle que pour tout cadre ε -séparé \mathcal{V} , on a

$$\alpha(V_s, V_u) \geq \eta(\varepsilon)$$

(avec V_s et V_u définis par les formules (II.3.5)).

(iii) Pour tout $\varepsilon > 0$, pour toute application $g \in \text{SO}(d+1, d)$ pseudohyperbolique et ε -séparée, on a

$$s(g) \lesssim_\varepsilon \hat{s}(\Lambda^d g).$$

Si de plus on suppose que $s(g) < 1$, on a

$$s(g) \asymp_\varepsilon \hat{s}(\Lambda^d g).$$

(iv) Si V_1 et V_2 sont deux sous-espaces de dimension d dans $\mathbb{R}^{d+1, d}$, on a

$$\alpha_{N_0}^{\text{Haus}}(V_1, V_2) \asymp \alpha_{\Lambda^d N_0}(\Lambda^d V_1, \Lambda^d V_2).$$

Démonstration.

(i) Soit $g \in \text{SO}(d+1, d)$. Soient $\lambda_1, \dots, \lambda_{2d+1}$ les valeurs propres de g , comptées avec multiplicité et ordonnées par valeur absolue croissante. Alors on sait que les valeurs propres de $\Lambda^d g$, comptées avec multiplicité, sont exactement les produits de la forme $\lambda_{i_1} \dots \lambda_{i_d}$, avec $1 \leq i_1 < \dots < i_d \leq 2d+1$. Comme les deux plus grandes sont $\lambda_{d+2} \dots \lambda_{2d+1}$ et $\lambda_{d+1} \lambda_{d+3} \dots \lambda_{2d+1}$, il s'ensuit que $\Lambda^d g$ est proximal si et seulement si $|\lambda_{d+1}| < |\lambda_{d+2}|$.

Supposons que ce soit le cas. Étant des espaces isotropes, $V_<(g)$ et $V_>(g)$ sont de dimension au plus d ; on en déduit que $|\lambda_{d+1}| = 1$. On a alors $|\lambda_{d+2}| > 1$, d'où $\dim V_>(g) = d$. Comme $V_=(g) \subset V_>(g)^\perp$ et que $V_=(g)$ est transverse à $V_>(g)$, on obtient que $\dim V_=(g) = 1$. En connaissant tout cela, il est facile de montrer que l'identité (II.2.3) est vérifiée, donc $\lambda_{d+1} = \frac{\det g}{(\det g_<)(\det g_>)} = 1$. On conclut que g est pseudohyperbolique. La réciproque est évidente.

Quant à l'expression de V_s et de V_u , elle découle immédiatement en considérant une base qui trigonalise g .

(ii) Soit $\varepsilon > 0$. Clairement, $\alpha(V_s, V_u)$ dépend continûment de $V_<$ et de $V_>$, et ne s'annule jamais lorsque $V_<$ et $V_>$ sont transverses. Comme l'ensemble de tous les cadres ε -séparés est compact, cette expression doit avoir une borne inférieure non nulle.

(iii) Soit $\varepsilon > 0$; soit $g \in \text{SO}(d+1, d)$ une application pseudohyperbolique et ε -séparée de cadre \mathcal{V} . On procède en trois étapes.

– Premièrement, on remarque que, d'après (II.2.3), on a $\|g_>^{-1}\| \asymp_\varepsilon \|g_<\|$, d'où

$$s(g) \asymp_\varepsilon \|g_>^{-1}\|. \quad (\text{II.3.6})$$

– Deuxièmement, montrons que pour toute application proximale f , on a

$$\hat{s}_{\Lambda^d N_0}(f) \asymp_\varepsilon \hat{s}_{\Lambda^d N_{\mathcal{V}}}(f). \quad (\text{II.3.7})$$

(Attention : la forme $\Lambda^d N_{\mathcal{V}}$ n'est en général pas la même que \hat{N}_f .) En effet, notez que si deux normes donnée par N et N' sont C -Lipschitz-équivalentes, alors les normes données par $\Lambda^d N$ et $\Lambda^d N'$ sont C^d -Lipschitz-équivalentes. Les inégalités ci-dessus découlent alors du lemme II.2.11.

– La dernière étape consiste à démontrer le résultat dans la métrique donnée par $N_{\mathcal{V}(g)}$. Soient $s_1 \leq \dots \leq s_d$ (resp. $s'_1 \geq \dots \geq s'_d$) les valeurs singulières de $g_>$ (resp. $g_<$), de sorte que $\|g_<\|_{N_{\mathcal{V}}} = \|g_<\|_{N_0} = s'_1$ et $\|g_>^{-1}\| = s_1^{-1}$. Comme les espaces $V_<$,

V_{\leq} et $V_{>}$ sont stables par g et deux à deux $N_{\mathcal{V}}$ -orthogonaux, on obtient que les valeurs singulières de g dans la métrique donnée par $N_{\mathcal{V}}$ sont

$$s'_d, \dots, s'_1, 1, s_1, \dots, s_d$$

(remarquez cependant que si on ne suppose pas que $s(g) < 1$, cette liste n'est pas forcément triée par ordre croissant.) D'autre part, on sait que les valeurs singulières de $\Lambda^d g$ dans la métrique donnée par $\Lambda^d N_{\mathcal{V}}$ sont les produits de d valeurs singulières distinctes de g dans la métrique donnée par $N_{\mathcal{V}}$. Puisque $V_s(\Lambda^d g)$ est $\Lambda^d N_{\mathcal{V}}$ -orthogonal à $V_u(\Lambda^d g)$, on peut encore une fois analyser les valeurs singulières séparément dans chaque sous-espace. On sait que la valeur singulière correspondant à V_s est égale à $s_1 \dots s_d$; on en déduit que $\|\Lambda^d g|_{V_u}\|_{\Lambda^d N_{\mathcal{V}}}$ est égale à la plus grande des valeurs singulières restantes. En particulier ce nombre est supérieur ou égal à $1 \cdot s_2 \dots s_d$. D'autre part, si λ est la plus grande valeur propre de $\Lambda^d g$, on a

$$|\lambda| = |\lambda_1 \dots \lambda_d| = |\det g_{>}| = s_1 \dots s_d$$

(où $\lambda_1, \dots, \lambda_d$ sont les valeurs propres de $g_{>}$). On en déduit que :

$$\hat{s}_{\Lambda^d N_{\mathcal{V}}}(\Lambda^d g) = \frac{\|\Lambda^d g|_{V_u}\|_{\Lambda^d N_{\mathcal{V}}}}{|\lambda|} \geq \frac{1 \cdot s_2 \dots s_d}{s_1 \dots s_d} = s_1^{-1} = \|g_{>}^{-1}\|. \quad (\text{II.3.8})$$

En rassemblant (II.3.6), (II.3.7) et (II.3.8), on obtient la première estimation. Supposons maintenant que $s(g) < 1$. Alors on a $s'_1 \leq s(g) < 1$ et $1 < (s(g))^{-1} \leq s_1$, ce qui signifie que les valeurs singulières de $\Lambda^d g$ sont effectivement rangées dans le « bon ordre ». Donc $1 \cdot s_2 \dots s_d$ est effectivement la plus grande valeur propre de $\Lambda^d g|_{V_u}$, et l'inégalité devient une égalité : $\hat{s}_{\Lambda^d N_{\mathcal{V}}}(\Lambda^d g) = \|g_{>}^{-1}\|$. La deuxième estimation en découle.

(iv) Soient V_1 et V_2 deux espaces de dimension d . On introduit les notations :

$$\alpha_1 := \alpha_{N_0}^{\text{Haus}}(V_1, V_2);$$

$$\alpha_2 := \alpha_{\Lambda^d N_0}(\Lambda^d V_1, \Lambda^d V_2).$$

On peut trouver une base N_0 -orthonormée (e_1, \dots, e_{2d+1}) de $\mathbb{R}^{d+1, d}$ telle que V_1 a pour base (e_1, \dots, e_d) et V_2 a pour base

$$((\cos \theta_i)e_i + (\sin \theta_i)e_{d+i})_{1 \leq i \leq d},$$

pour certains angles $0 \leq \theta_1 \leq \dots \leq \theta_d \leq \frac{\pi}{2}$. Dans ce cas, on a :

$$\alpha_1 = \theta_d$$

et

$$\cos \alpha_2 = \prod_{i=1}^d \cos \theta_i,$$

d'où

$$(\cos \alpha_1)^d \leq \cos \alpha_2 \leq \cos \alpha_1.$$

D'autre part, la concavité de la fonction $y \mapsto (\arccos \exp y)^2$ implique que pour tout $\theta \in [0, \frac{\pi}{2}]$, on a

$$\arccos((\cos \theta)^d) \leq \sqrt{d}\theta.$$

On obtient finalement

$$\alpha_1 \leq \alpha_2 \leq \sqrt{d}\alpha_1,$$

CQFD. □

On peut maintenant démontrer la proposition principale.

Démonstration de la proposition II.3.8. Soit $\varepsilon > 0$; soit $\mathcal{W} = (\mathcal{V}_1, \dots, \mathcal{V}_n)$ un multicaire ε -séparé et $G = \langle g_1, \dots, g_n \rangle$ un groupe $s_1(\varepsilon)$ -contractant basé sur \mathcal{W} , pour une certaine constante $s_1(\varepsilon)$ à préciser au cours de la preuve. Soit $g = g_{i_1}^{\sigma_1} \dots g_{i_k}^{\sigma_k}$ un mot cycliquement réduit non vide.

Pour tout i , on prend $f_i = \Lambda^d g_i$. Vérifions que le lemme II.3.4 s'applique. En effet :

- D'après le lemme II.3.11 (i), $F = (f_1, \dots, f_n)$ est un système proximal indépendant. (Les conditions (i) et (ii) de la définition découlent respectivement de la première et deuxième partie du lemme II.3.11 (i).)
- D'après le lemme II.3.11 (ii), on a $\eta(F) \leq \eta(\varepsilon)$; en d'autres termes, F est $\eta(\varepsilon)$ -séparé. On pose $\eta = \eta(\varepsilon)$: alors « \lesssim_η » implique toujours « \lesssim_ε ».
- Sans perte de généralité, on peut supposer que $s(G) < 1$. Alors d'après le lemme II.3.11 (iii), on a $\hat{s}(F) \lesssim_\varepsilon s(G)$, qui à son tour est majoré par $s_1(\varepsilon)$. Si on prend $s_1(\varepsilon)$ suffisamment petit (puisque η est entièrement déterminé par ε), on a alors

$$\hat{s}(F) \leq \hat{s}(\eta).$$

Déduisons maintenant les conclusions de la proposition II.3.8 à partir de celles du lemme II.3.4, appliqué au mot $\Lambda^d g = f_{i_1}^{\sigma_1} \dots f_{i_k}^{\sigma_k}$:

- Le fait que g soit pseudohyperbolique découle du lemme II.3.11 (i).
- Montrons que

$$\alpha^{\text{Haus}}(V_{>}(g), V_{>}(g_{i_1}^{\sigma_1})) \lesssim_\varepsilon s(G).$$

En effet, on a :

$$\begin{aligned} \alpha^{\text{Haus}}(V_{>}(g), V_{>}(g_{i_1}^{\sigma_1})) &\lesssim \alpha(\Lambda^d V_{>}(g), \Lambda^d V_{>}(g_{i_1}^{\sigma_1})) && \text{par le lemme II.3.11 (iv)} \\ &= \alpha(V_s(\Lambda^d g), V_s(f_{i_1}^{\sigma_1})) && \text{par le lemme II.3.11 (i)} \\ &\lesssim_\eta \hat{s}(F) && \text{par le lemme II.3.4} \\ &\lesssim_\varepsilon s(G) && \text{par le lemme II.3.11 (iii);} \end{aligned}$$

et on sait que « \lesssim_η » implique « \lesssim_ε ».

- Montrons que g est $\frac{\varepsilon}{3}$ -séparée. Comme $s(G) \leq s_1(\varepsilon)$, on peut choisir $s_1(\varepsilon)$ suffisamment petit pour déduire du point précédent l'inégalité suivante :

$$\alpha^{\text{Haus}}(V_{>}(g), V_{>}(g_{i_1}^{\sigma_1})) \leq \frac{\varepsilon}{3}.$$

En remplaçant g par g^{-1} , on obtient de même

$$\alpha^{\text{Haus}}(V_{<}(g), V_{<}(g_{i_k}^{\sigma_k})) \leq \frac{\varepsilon}{3}.$$

Finalement, comme g est cycliquement réduit et que \mathcal{W} est ε -séparé, on sait que

$$\alpha(V_{>}(g_{i_1}^{\sigma_1}), V_{<}(g_{i_k}^{\sigma_k})) \geq \varepsilon.$$

De ces trois inégalités, on déduit que

$$\alpha(V_{<}(g), V_{>}(g)) \geq \frac{\varepsilon}{3}. \quad (\text{II.3.9})$$

- Montrons que g est 1-contractante. En utilisant le lemme II.3.4 et le lemme II.3.11 (iii), on obtient

$$\begin{aligned} s(g) &\lesssim_{\varepsilon(g)} \hat{s}(\Lambda^d g) && \text{par le lemme II.3.11 (iii)} \\ &\lesssim_\eta \hat{s}(F) && \text{par le lemme II.3.4} \\ &\lesssim_\varepsilon s(G) && \text{par le lemme II.3.11 (iii) (car } s(G) < 1. \end{aligned}$$

Comme $\varepsilon(g) \geq \frac{\varepsilon}{3}$ et $\eta = \eta(\varepsilon)$, on obtient $s(g) \lesssim_\varepsilon s(G) \leq s_1(\varepsilon)$. En prenant $s_1(\varepsilon)$ suffisamment petit, on en déduit que

$$s(g) < 1. \quad \square$$

II.4. La « balle de tennis » et les groupes de Schottky généralisés

Soit $\varepsilon > 0$, et soit \mathcal{V} un cadre.

Définition II.4.1. On définit, sur la sphère $\mathbb{S}(\mathbb{R}^{d+1,d})$ (désormée simplement notée \mathbb{S}), les régions suivantes :

$$\begin{cases} \mathcal{H}_\mathbb{S}^- := B_{N_\mathcal{V}}(\pi_\mathbb{S}(V_<^\mathbb{L}), \varepsilon) \\ \mathcal{H}_\mathbb{S}^+ := B_{N_\mathcal{V}}(\pi_\mathbb{S}(V_>^\mathbb{L}), \varepsilon). \end{cases}$$

(Bien sûr, elles dépendent de \mathcal{V} et de ε , mais pour alléger les notations, on laissera cette dépendance implicite.) On les appelle *régions de balle de tennis* : voir la figure II.1 pour comprendre pourquoi.

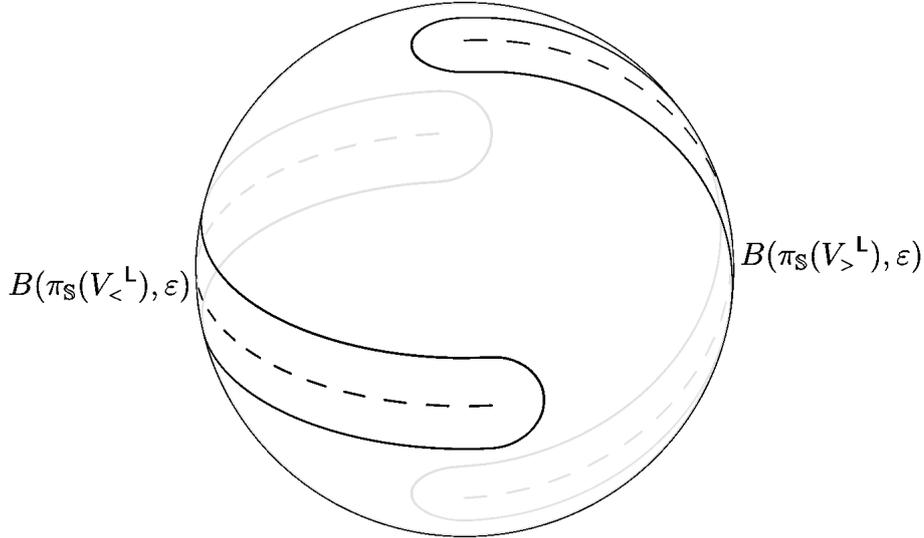


FIGURE II.1.: Les « régions de balle de tennis » pour $d = 1$. Dans ce cas $V_>^\mathbb{L}$ et $V_<^\mathbb{L}$ sont des demi-plans ; leurs projections sur la sphère $\pi_\mathbb{S}(V_>^\mathbb{L})$ et $\pi_\mathbb{S}(V_<^\mathbb{L})$ sont les demi-cercles en pointillé ; et les régions elles-mêmes, soit les ε -voisinages de ces projections, sont tracées en trait plein.

Dans la proposition qui suit et dans sa démonstration, on travaille dans la métrique donnée par $N_{\mathcal{V}(g)}$.

Proposition II.4.2. *Pour tout $\varepsilon > 0$, il existe une constante $s_2(\varepsilon)$ telle que pour toute application g pseudohyperbolique et $s_2(\varepsilon)$ -contractante (de cadre \mathcal{V}), on a*

$$g_\mathbb{S} \left(\mathbb{S} \setminus \overline{\mathcal{H}_\mathbb{S}^-} \right) \subset \mathcal{H}_\mathbb{S}^+ \quad (\text{II.4.1a})$$

$$g_\mathbb{S}^{-1} \left(\mathbb{S} \setminus \overline{\mathcal{H}_\mathbb{S}^+} \right) \subset \mathcal{H}_\mathbb{S}^-. \quad (\text{II.4.1b})$$

Remarque II.4.3. Comme on travaille ici dans la métrique donnée par $N_\mathcal{V}$, la séparation de g n'a aucune importance et le nombre ε n'a rien à voir avec celle-ci. Ce nombre définit seulement l'« épaisseur » des régions de balle de tennis $\mathcal{H}_\mathbb{S}^\pm$.

Remarque II.4.4. Comme $g_{\mathbb{S}}$ est un homéomorphisme et que les régions qu'on considère sont des ouverts réguliers, les deux relations sont en fait équivalentes. De plus, comme $V_{<}(g^{-1}) = V_{>}$, $V_{>}(g^{-1}) = V_{<}$ et $s(g^{-1}) = s(g)$, (II.4.1b) n'est rien d'autre que (II.4.1a) appliquée à g^{-1} .

Soit $\varepsilon > 0$ et soit g une application pseudohyperbolique de cadre \mathcal{V} . Comme on l'avait fait auparavant, pour $x \in \mathbb{R}^{d+1,d}$, on définit le triplet $(x_{<}, x_{=}, x_{>}) \in V_{<} \times V_{=} \times V_{>}$ tel que $x_{<} + x_{=} + x_{>} = x$; ce sont les projections $N_{\mathcal{V}}$ -orthogonales de x sur les espaces correspondants. Le vecteur $e_{=}$ (voir la définition II.2.9) donne une orientation sur $V_{=}$, ce qui nous permet de définir un ordre sur cet espace de dimension 1 : on dira que $x_{=} \geq y_{=}$ lorsque $\langle x_{=}, e_{=} \rangle \geq \langle y_{=}, e_{=} \rangle$.

Lemme II.4.5. *Soit $x \in \mathbb{R}^{d+1,d} \setminus \{0\}$. Alors on a :*

$$\pi_{\mathbb{S}}(x) \in B(\pi_{\mathbb{S}}(V_{<}^{\perp}), \varepsilon) \iff \begin{cases} x_{=} \leq 0 \text{ et } \frac{\|x_{>}\|}{\|x_{<}+x_{=}\|} < \tan \varepsilon \\ \text{or} \\ x_{=} \geq 0 \text{ et } \frac{\|x_{>}+x_{=}\|}{\|x_{<}\|} < \tan \varepsilon \end{cases} \quad (\text{II.4.2a})$$

et

$$\pi_{\mathbb{S}}(x) \in B(\pi_{\mathbb{S}}(V_{>}^{\perp}), \varepsilon) \iff \begin{cases} x_{=} \leq 0 \text{ et } \frac{\|x_{<}+x_{=}\|}{\|x_{>}\|} < \tan \varepsilon \\ \text{or} \\ x_{=} \geq 0 \text{ et } \frac{\|x_{<}\|}{\|x_{>}+x_{=}\|} < \tan \varepsilon \end{cases} \quad (\text{II.4.2b})$$

(et en remplaçant partout « $< \tan \varepsilon$ » par « $\leq \tan \varepsilon$ », on peut caractériser de la même façon les adhérences de ces régions.)

Démonstration. Sans perte de généralité, on se concentre sur (II.4.2b) (l'autre formule s'en déduit simplement en échangeant $V_{<}$ et $V_{>}$ et en renversant l'orientation de $V_{=}$.) On rappelle que $x \in V_{>}^{\perp}$ si et seulement si $x_{<} = 0$ et $x_{=} > 0$.

- Supposons que $x_{=} \geq 0$. Comme $V_{>}^{\perp} \subset V_{\geq}$, on a $\alpha(x, V_{>}^{\perp}) \geq \alpha(x, V_{\geq})$. D'autre part, on a $\alpha(x, V_{\geq}) = \alpha(x, x_{>} + x_{=})$ et $x_{>} + x_{=} \in V_{>}^{\perp}$, ce qui montre l'inégalité inverse. Donc $\alpha(x, V_{>}^{\perp}) = \alpha(x, V_{\geq})$.
- Supposons que $x_{=} \leq 0$; sans perte de généralité, on peut supposer que $\|x\| = 1$. Comme $V_{>} \subset V_{>}^{\perp}$, clairement on a $\alpha(x, V_{>}^{\perp}) \leq \alpha(x, V_{>})$. Maintenant soit $y \in V_{>}^{\perp}$; alors on a :

$$\begin{aligned} \cos \alpha(x, y) &= \frac{\langle x, y \rangle}{\|y\|} \\ &= \frac{\langle x_{<} + x_{>} + x_{=}, y_{>} + y_{=} \rangle}{\|y_{>} + y_{=}\|} \\ &= \frac{\langle x_{>}, y_{>} \rangle + \langle x_{=}, y_{=} \rangle}{\|y_{>} + y_{=}\|} \\ &\leq \frac{\langle x_{>}, y_{>} \rangle}{\|y_{>}\|} \\ &= \cos \alpha(x, y_{>}), \end{aligned}$$

puisque $\langle x_{=}, y_{=} \rangle \leq 0$ et $\|y_{>} + y_{=}\| \geq \|y_{>}\|$. Donc $\alpha(x, y) \geq \alpha(x, y_{>})$, avec $y_{>} \in V_{>}$. Cela montre l'inégalité inverse. Donc $\alpha(x, V_{>}^{\perp}) = \alpha(x, V_{>})$.

Le résultat découle maintenant du fait que pour tout sous-espace vectoriel $E \subset \mathbb{R}^{d+1,d}$, on a

$$\alpha(x, E) = \alpha(x, x_E) = \arccos \frac{\|x_E\|}{\|x\|} = \arctan \frac{\|x - x_E\|}{\|x_E\|},$$

où x_E est la projection $N_{\mathcal{V}}$ -orthogonale de x sur E . □

Démonstration de la proposition II.4.2. En vertu de la remarque II.4.4, il suffit de montrer (II.4.1a). Soit $x \in \mathbb{R}^{d+1,d} \setminus \{0\}$ tel que $\alpha(x, V_{<}^L) > \varepsilon$; il suffit de montrer que si $s(g) \leq s_2(\varepsilon)$ (pour une valeur de $s_2(\varepsilon)$ à préciser au cours de la preuve), alors on a $\alpha(g(x), V_{>}^L) < \varepsilon$.

Supposons que $x_{=} \leq 0$. Alors on a, d'après le lemme II.4.5,

$$\frac{\|x_{>}\|}{\|x_{<} + x_{=}\|} > \tan \varepsilon.$$

On en déduit que

$$\begin{aligned} \frac{\|g(x)_{<} + g(x)_{=}\|^2}{\|g(x)_{>}\|^2} &= \frac{\|g(x_{<})\|^2 + \|g(x_{=})\|^2}{\|g(x_{>})\|^2} \\ &\leq \frac{s(g)^2 \|x_{<}\|^2 + \|x_{=}\|^2}{s(g)^{-2} \|x_{>}\|^2} \\ &\leq s(g)^2 \frac{\|x_{<}\|^2 + \|x_{=}\|^2}{\|x_{>}\|^2} \\ &\leq s(g)^2 (\tan \varepsilon)^{-2} \\ &< (\tan \varepsilon)^2, \end{aligned}$$

pourvu que $s(g) < (\tan \varepsilon)^4$, ce qui est vrai si on prend $s_2(\varepsilon)$ plus petit que cette valeur. Donc $\alpha(g(x), V_{>}^L) < \varepsilon$. D'autre part, on a $g(x)_{=} = g(x_{=}) = x_{=} \leq 0$. On en déduit que $\alpha(g(x), V_{>}^L) < \varepsilon$. Dans le cas où $x_{=} \geq 0$, un calcul tout à fait analogue donne le même résultat. \square

Considérons maintenant un multicaire $\mathcal{W} = (\mathcal{V}_1, \dots, \mathcal{V}_n)$ et une famille de rayons $\varepsilon_1, \dots, \varepsilon_n$.

Définition II.4.6. Comme dans la définition II.4.1, on définit pour chaque indice i les régions

$$\begin{cases} \mathcal{H}_{\mathbb{S},i}^- := B_{N_{\mathcal{V}_i}}(\pi_{\mathbb{S}}(V_{i,<}^L), \varepsilon_i) \\ \mathcal{H}_{\mathbb{S},i}^+ := B_{N_{\mathcal{V}_i}}(\pi_{\mathbb{S}}(V_{i,>}^L), \varepsilon_i). \end{cases}$$

(Encore une fois, elles dépendent de \mathcal{W} et de ε_i , mais pour alléger les notations, on laissera cette dépendance implicite.) Soit $G = \langle g_1, \dots, g_n \rangle$ n'importe quel groupe basé sur \mathcal{W} . Si les fermés $\overline{\mathcal{H}_{\mathbb{S},i}^{\pm}}$ sont deux à deux disjoints et si pour chaque i , $s(g_i)$ est suffisamment petit pour appliquer la proposition II.4.2, on dit que G est $(\varepsilon_1, \dots, \varepsilon_n)$ -Schottky.

Dans ce cas, il découle de la proposition II.4.2 que le groupe G est libre. En effet, on a pour tout i :

$$\begin{cases} g_i \left(\mathbb{S} \setminus \overline{\mathcal{H}_{\mathbb{S},i}^-} \right) \subset \mathcal{H}_{\mathbb{S},i}^+ \\ g_i^{-1} \left(\mathbb{S} \setminus \overline{\mathcal{H}_{\mathbb{S},i}^+} \right) \subset \mathcal{H}_{\mathbb{S},i}^-, \end{cases} \quad (\text{II.4.3})$$

et on peut appliquer le lemme du ping-pong (voir par exemple la proposition 1.1 de [22]).

II.5. Déformations affines

Définition II.5.1. Soit $G \subset \text{SO}(d+1, d)$ un groupe de transformations linéaires. On appelle *déformation affine* de G tout groupe $\Gamma \subset \mathbb{R}^{d+1,d} \rtimes \text{SO}(d+1, d)$ tel que la projection canonique $\ell : \mathbb{R}^{d+1,d} \rtimes \text{SO}(d+1, d) \rightarrow \text{SO}(d+1, d)$ induise un isomorphisme de Γ sur G . En d'autres termes, c'est un groupe de transformations affines qui ne contient pas de translations pures et dont les parties linéaires forment le groupe G .

Supposons maintenant que $G = \langle g_1, \dots, g_n \rangle$ est un groupe libre ; soit Γ une déformation affine de G . Alors celle-ci est engendrée par les éléments $\gamma_1, \dots, \gamma_n$ dont les parties linéaires valent respectivement g_1, \dots, g_n . Cela veut dire que, G étant fixé, Γ est entièrement déterminé par les parties de translation de ses générateurs, à savoir les vecteurs $\gamma_1(0), \dots, \gamma_n(0)$. Réciproquement, pour toute famille de vecteurs $\mathbf{t} = (t_1, \dots, t_n) \in (\mathbb{R}^{d+1,d})^n$, on peut définir $\gamma_1, \dots, \gamma_n$ par $\gamma_i(x) = g_i(x) + t_i$ pour tout i . Comme G est libre, le groupe engendré par ces éléments est alors une déformation affine de G , qu'on appellera $G(\mathbf{t})$. Cela définit une bijection entre l'ensemble des déformations affines de G et $(\mathbb{R}^{d+1,d})^n$.

On peut alors énoncer le théorème principal de façon plus précise :

Théorème II.5.2. *Pour tout $\varepsilon > 0$, il existe une constante $s_3(\varepsilon)$ ayant la propriété suivante. Soit \mathcal{W} un multicadre ε -séparé, G un groupe $s_3(\varepsilon)$ -contractant basé sur \mathcal{W} . Alors on peut affirmer que :*

- (i) *Le groupe G est libre ;*
- (ii) *Il existe un cône ouvert non vide $\mathbf{T} \subset (\mathbb{R}^{d+1,d})^n$ (dépendant de G) tel que pour tout $\mathbf{t} \in \mathbf{T}$, la déformation affine $G(\mathbf{t})$ soit discrète et agisse proprement sur $\mathbb{R}^{d+1,d}$;*
- (iii) *Pour tout $\mathbf{t} \in \mathbf{T}$, l'espace quotient $\mathbb{R}^{d+1,d}/G(\mathbf{t})$ est homéomorphe à un corps plein à n anses, de dimension $(2d + 1)$.*

Démonstration. On commence par donner quelques définitions et notations, fixées jusqu'à la fin de cette section.

- On fixe $\varepsilon > 0$, \mathcal{W} un multicadre ε -séparé, G un groupe $s_3(\varepsilon)$ -contractant basé sur \mathcal{W} . On déterminera la valeur de $s_3(\varepsilon)$ au cours de la preuve.
- Pour tout i , on choisit une constante $\varepsilon_i > 0$ telle que pour toute partie $X \subset \mathbb{S}$, on ait

$$B_{N_{\nu_i}}(X, \varepsilon_i) \subset B_{N_0}(X, \frac{\varepsilon}{3}). \quad (\text{II.5.1})$$

On définit les régions de balle de tennis $\mathcal{H}_{\mathbb{S},i}^\sigma$ en utilisant ces constantes (voir la définition II.4.6).

- Pour tout $\mathbf{t} \in (\mathbb{R}^{d+1,d})^n$, pour chaque i compris entre 1 et n et pour chaque $\sigma = \pm 1$, on définit l'ouvert suivant. C'est une partie de $\mathbb{R}^{d+1,d}$ en forme de cône, dont le sommet dépend du vecteur de translation et dont la base est la région de balle de tennis correspondante (voir la figure II.2) :

$$\mathcal{H}_i^\sigma(\mathbf{t}) := \pi_{\mathbb{S}}^{-1}(\mathcal{H}_{\mathbb{S},i}^\sigma) + \sigma u_i,$$

où u_i est la solution de l'équation $u_i + g_i(u_i) = t_i$. (Comme l'application g_i est pseudohyperbolique, elle n'a pas -1 pour valeur propre, et donc cette équation a bien une unique solution.)

- Pour tout tel \mathbf{t} , on définit aussi, pour tous les indices i , les ouverts

$$\begin{cases} \tilde{\mathcal{H}}_i^-(\mathbf{t}) := \mathcal{H}_i^-(\mathbf{t}) \\ \tilde{\mathcal{H}}_i^+(\mathbf{t}) := \gamma_i \left(\mathbb{R}^{d+1,d} \setminus \overline{\mathcal{H}_i^-(\mathbf{t})} \right) \end{cases}$$

(où $\gamma_i : x \mapsto g_i(x) + t_i$ est le i -ième générateur de la déformation affine $G(\mathbf{t})$: il dépend implicitement de \mathbf{t} .) On introduit également l'ouvert

$$\mathcal{H}^0 := \mathbb{R}^{d+1,d} \setminus \bigcup_{i=1}^n \bigcup_{\sigma=\pm} \overline{\tilde{\mathcal{H}}_i^\sigma}.$$

- On définit \mathbf{T} comme l'ensemble de tous les $\mathbf{t} \in (\mathbb{R}^{d+1,d})^n$ tels que les $2n$ fermés $\overline{\mathcal{H}_i^\pm(\mathbf{t})}$ soient deux à deux disjoints.

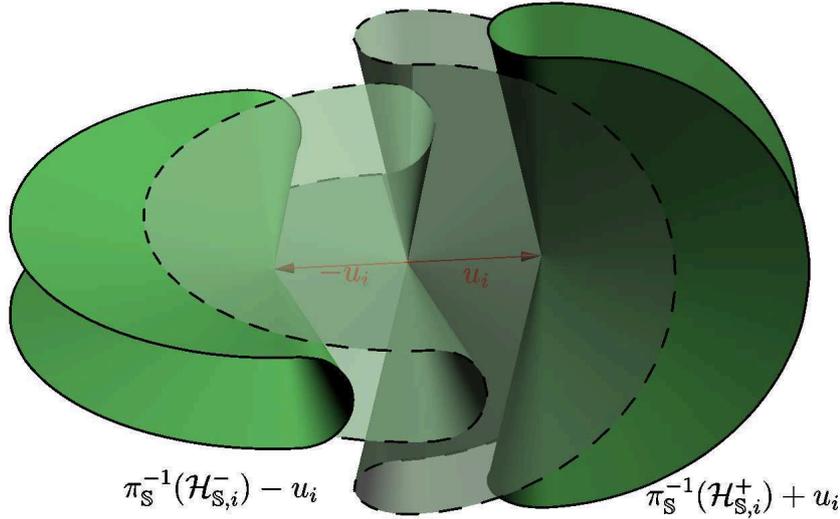


FIGURE II.2.: Les ouverts $\mathcal{H}_i^\pm(\mathbf{t})$, pour $d = 1$ et i et \mathbf{t} fixés arbitrairement. Pour les construire, on part des régions de balle de tennis (voir la figure II.1), on prend ensuite l'image réciproque par π_S (ce qui revient à former un cône), puis on translate de $\pm u_i$.

Alors (i) découle immédiatement ou bien de la proposition II.3.8, ou bien de la proposition II.4.2 combinée avec (II.5.1). L'énoncé (ii) découle du lemme II.5.3, du lemme II.5.4 et de la proposition II.5.5 ci-dessous (l'existence d'un domaine fondamental implique en effet que le groupe est discret et agit proprement). Cette dernière proposition est intéressante en soi, dans la mesure où elle décrit la forme précise du domaine fondamental. Elle nous permet également de démontrer (iii). En effet, si le domaine fondamental est \mathcal{H}^0 , alors l'espace quotient $\mathbb{R}^{d+1,d}/G(\mathbf{t})$ est homéomorphe à l'espace obtenu à partir du fermé $\overline{\mathcal{H}^0}$ en identifiant, pour chaque i , la frontière de $\tilde{\mathcal{H}}_i^-$ avec la frontière de $\tilde{\mathcal{H}}_i^+$. Mais clairement, toutes ces frontières sont homéomorphes à \mathbb{R}^{2d} (ou bien, si on préfère, à des « disques » ouverts de dimension $2d$), et \mathcal{H}^0 est homéomorphe à \mathbb{R}^{2d+1} (ou bien à une boule ouverte de dimension $2d + 1$). \square

Lemme II.5.3. *L'ensemble \mathbf{T} est non vide.*

Démonstration. On rappelle que $e_{i,=}$ est le vecteur de norme N_0 unité qui est fixé par g_i et dont le signe est choisi convenablement (voir la définition II.2.9). On pose

$$\mathbf{t}_0 := (2e_{1,=}, \dots, 2e_{n,=}).$$

Alors on a, pour tout i , $u_i = e_{i,=}$, puisque par définition $g_i(e_{i,=}) = e_{i,=}$.

Montrons que $\mathbf{t}_0 \in \mathbf{T}$, c'est-à-dire que les fermés $\overline{\mathcal{H}_i^\pm(\mathbf{t}_0)}$ sont deux à deux disjoints. Pour faire cela, on va montrer qu'ils sont inclus dans les ouverts $\pi_S^{-1}(\mathcal{H}_{S,i}^\pm)$, dont on sait déjà qu'ils sont deux à deux disjoints.

En effet, fixons un indice i et un signe σ . Dans cette preuve, on travaille dans la métrique donnée par N_{V_i} . On doit montrer que :

$$\overline{\mathcal{H}_i^\sigma(\mathbf{t}_0)} := \overline{\pi_S^{-1}(\mathcal{H}_{S,i}^\sigma)} + \sigma e_{i,=} \subset \pi_S^{-1}(\mathcal{H}_{S,i}^\sigma).$$

Supposons, sans perte de généralité, que $\sigma = +1$; soit $x \in \overline{\pi_S^{-1}(\mathcal{H}_{S,i}^+)}$. Si $x = 0$, clairement, on a

$$x + e_{i,=} = e_{i,=} \in V_{i,>}^L \subset \pi_S^{-1}(\mathcal{H}_{S,i}^+).$$

Sinon, il est facile de voir que $\pi_{\mathbb{S}}(x) \in \overline{\mathcal{H}_{\mathbb{S},i}^+}$. Clairement, on peut appliquer le lemme II.4.5, pourvu qu'on remplace les inégalités strictes par des inégalités larges. On reprend la notation $x = x_{<} + x_{=} + x_{>}$ de ce lemme. On distingue trois cas :

– Si $x_{=} \geq 0$, alors on a

$$\frac{\|x_{<}\|}{\|x_{>} + x_{=}\|} \leq \tan \varepsilon_i.$$

On a toujours $x_{=} + e_{=} \geq 0$ et $\|x_{=} + e_{=}\| > \|x_{=}\|$, d'où

$$\frac{\|x_{<}\|}{\|x_{>} + x_{=} + e_{=}\|} < \frac{\|x_{<}\|}{\|x_{>} + x_{=}\|} \leq \tan \varepsilon_i,$$

et on conclut que $x + e_{=} \in \pi_{\mathbb{S}}^{-1}(B_N(\pi_{\mathbb{S}}(V_{>}^L), \varepsilon_i))$.

– Si $x_{=} \leq -e_{=}$, alors de même

$$\frac{\|x_{<} + x_{=} + e_{=}\|}{\|x_{>}\|} < \frac{\|x_{<} + x_{=}\|}{\|x_{>}\|} \leq \tan \varepsilon_i,$$

et on arrive à la même conclusion.

– Si $-e_{=} < x_{=} < 0$, alors on a

$$\frac{\|x_{<} + x_{=}\|}{\|x_{>}\|} \leq \tan \varepsilon_i,$$

d'où on déduit

$$\begin{aligned} \|x_{<}\| &< \|x_{<} + x_{=}\| \\ &\leq (\tan \varepsilon_i) \|x_{>}\| \\ &< (\tan \varepsilon_i) \|x_{>} + x_{=} + e_{=}\|; \end{aligned}$$

puisque $x_{=} + e_{=} \geq 0$, on arrive encore à la même conclusion. \square

Lemme II.5.4. *L'ensemble \mathbf{T} est un cône ouvert.*

Démonstration. La stabilité de \mathbf{T} par les homothéties de rapport strictement positif est évidente. En effet, pour tout réel $\lambda > 0$, les ensembles $\mathcal{H}_i^{\pm}(\lambda \mathbf{t})$ sont les images des ensembles $\mathcal{H}_i^{\pm}(\mathbf{t})$ par l'homothétie de rapport λ ; donc ces premiers sont disjoints si et seulement si ces derniers sont disjoints.

Montrons qu'il est ouvert. Soit $\mathbf{t}_0 = (t_{0,1}, \dots, t_{0,n})$ un élément quelconque de \mathbf{T} . On sait que deux quelconques parmi les fermés $\overline{\mathcal{H}_i^{\pm}(\mathbf{t}_0)}$ sont disjoints; on prétend que, de plus, ils sont séparés par une distance non nulle. En effet, prenons une boule B dont le rayon est très grand comparé à \mathbf{t}_0 . Alors les parties de ces fermés qui se trouvent à l'intérieur de B sont compactes et disjointes, donc séparées par une distance non nulle. Quant aux parties qui se trouvent à l'extérieur de B , elles sont séparées car asymptotiquement, leurs projections sur \mathbb{S} — à savoir $\overline{\mathcal{H}_{\mathbb{S},i}^{\pm}}$ — sont aussi compactes et disjointes.

Soit d_{\min} la plus petite de ces distances. Considérons l'ensemble de tous les $\mathbf{t} = (t_1, \dots, t_n)$ tels que pour tout indice i , $\|u_i - u_{0,i}\| < \frac{d_{\min}}{2}$. Alors clairement cet ensemble est un voisinage de \mathbf{t}_0 , et est inclus dans \mathbf{T} . \square

Proposition II.5.5. *Pour tout $\mathbf{t} \in \mathbf{T}$, l'action de la déformation affine $\Gamma := G(\mathbf{t})$ sur l'espace affine $\mathbb{R}^{d+1,d}$ a pour domaine fondamental \mathcal{H}^0 . Plus précisément :*

- (i) *Les images de \mathcal{H}^0 par les éléments de Γ sont deux à deux disjointes;*
- (ii) *Les images de son adhérence recouvrent l'espace tout entier :*

$$\bigcup_{\gamma \in \Gamma} \overline{\gamma(\mathcal{H}^0)} = \mathbb{R}^{d+1,d}.$$

Démonstration. On fixe une valeur $\mathbf{t} = (t_1, \dots, t_n) \in \mathbf{T}$; on appelle $\Gamma = \langle \gamma_1, \dots, \gamma_n \rangle := G(\mathbf{t})$ la déformation affine correspondante. La première chose à comprendre est que les régions \mathcal{H}_i^\pm (désormais on ne mentionnera plus la dépendance en \mathbf{t}) satisfont des « identités de ping-pong » similaires à (II.4.3). Plus précisément, on déduit de (II.4.3) que

$$\begin{aligned} \gamma_i \left(\mathbb{R}^{d+1,d} \setminus \overline{\mathcal{H}_i^-} \right) &= \gamma_i \left(\mathbb{R}^{d+1,d} \setminus \overline{\pi_{\mathbb{S}}^{-1}(\mathcal{H}_{\mathbb{S},i}^-) - u_i} \right) \\ &= g_i \left(\pi_{\mathbb{S}}^{-1} \left(\mathbb{S} \setminus \overline{\mathcal{H}_{\mathbb{S},i}^-} \right) \right) - g_i(u_i) + t_i \\ &\subset \pi_{\mathbb{S}}^{-1}(\mathcal{H}_{\mathbb{S},i}^+) + u_i \\ &= \mathcal{H}_i^+, \end{aligned}$$

et la même chose s'applique à γ_i^{-1} . Ainsi on obtient

$$\begin{cases} \gamma_i \left(\mathbb{R}^{d+1,d} \setminus \overline{\mathcal{H}_i^-} \right) \subset \mathcal{H}_i^+ \\ \gamma_i^{-1} \left(\mathbb{R}^{d+1,d} \setminus \overline{\mathcal{H}_i^+} \right) \subset \mathcal{H}_i^- \end{cases} \quad (\text{II.5.2})$$

En remplaçant les \mathcal{H}_i^\pm par les $\tilde{\mathcal{H}}_i^\pm$, les inclusions deviennent des égalités exactes :

$$\gamma_i \left(\mathbb{R}^{d+1,d} \setminus \overline{\tilde{\mathcal{H}}_i^-} \right) = \tilde{\mathcal{H}}_i^+; \quad (\text{II.5.3a})$$

$$\gamma_i^{-1} \left(\mathbb{R}^{d+1,d} \setminus \overline{\tilde{\mathcal{H}}_i^+} \right) = \tilde{\mathcal{H}}_i^-. \quad (\text{II.5.3b})$$

En effet (II.5.3a) est vrai par la définition de $\tilde{\mathcal{H}}_i^+$, et (II.5.3b) découle de (II.5.3a) car γ_i est continue et \mathcal{H}_i^- est un ouvert régulier.

Pour avoir une vraie « configuration de ping-pong », on a aussi besoin de vérifier que les ensembles $\tilde{\mathcal{H}}_i^\pm$ sont deux à deux disjoints. Mais en utilisant leur définition et (II.5.2), on sait qu'ils sont inclus dans les ensembles $\tilde{\mathcal{H}}_i^\pm$, qui sont deux à deux disjoints par hypothèse ($\mathbf{t} \in \mathbf{T}$).

- (i) est alors immédiat. En effet, on voit par récurrence que pour tout $\gamma = \gamma_{i_1}^{\sigma_1} \dots \gamma_{i_k}^{\sigma_k} \in \Gamma$, l'image $\gamma(\mathcal{H}^0)$ se trouve dans $\tilde{\mathcal{H}}_{i_1}^{\sigma_1}$, qui est par définition disjointe de \mathcal{H}^0 .
 - (ii) est la partie difficile. Le cas particulier $d = 1$ a été fait par Drumm dans [10] (preuve du théorème 4); voir aussi [8], section 4. Notre preuve est assez proche.
- Dorénavant, on travaille dans la métrique donnée par N_0 .

Avant de poursuivre, nous avons besoin d'un petit lemme géométrique. On rappelle que S est un certain sous-espace défini positif maximal de $\mathbb{R}^{d+1,d}$ fixé une fois pour toutes.

Lemme II.5.6.

(i) Soient V et W deux SETIM. On a alors

$$\alpha(V_Q^\perp, W_Q^\perp) = \alpha(V, W)$$

(où V_Q^\perp désigne le Q -orthogonal de V).

(ii) Soit V' le Q -orthogonal d'un certain SETIM V et soit $x \in S$. Alors on a

$$\sin \alpha(x, V' \cap S) = \sqrt{2} \sin \alpha(x, V').$$

Démonstration.

- (i) Il est évident que $\alpha(V_{N_0}^\perp, W_{N_0}^\perp) = \alpha(V, W)$ (où $V_{N_0}^\perp$ désigne, de même, le N_0 -orthogonal de V). D'autre part, pour tout SETIM V , on a $V_Q^\perp = \varsigma(V_{N_0}^\perp)$, où l'application $\varsigma := \text{Id}_S \oplus (-\text{Id}_T)$ est N_0 -isométrique. L'identité recherchée en découle.

- (ii) Remarquons d'abord que tout plan contenu dans V' forme avec S un angle de $\frac{\pi}{4}$. En effet, un tel plan contient des vecteurs x tels que $Q(x, x) = 0$ (car il a une intersection non triviale avec V), mais aucun vecteur x tel que $Q(x, x) < 0$. D'autre part, il découle des définitions de N_0 et Q que $\alpha(x, S) < (\text{resp. } =, >) \frac{\pi}{4}$ si et seulement si $Q(x, x) > (\text{resp. } =, <) 0$.

Maintenant soit X l'espace de dimension 3 engendré par x , $\pi_{V'}(x)$ et la droite $V' \cap S$ (où $\pi_{V'}$ désigne la projection N_0 -orthogonale sur V'). On sait que les plans $V' \cap X$ et $S \cap X$ forment un angle de $\frac{\pi}{4}$, et que le plan engendré par x et $\pi_{V'}(x)$ est perpendiculaire à $V' \cap X$. La version sphérique de la loi des sinus entraîne alors l'identité $\sin \alpha(x, V' \cap S) = \sqrt{2} \sin \alpha(x, \pi_{V'}(x))$, CQFD. \square

Démonstration de la proposition II.5.5, suite. On procède par l'absurde : soit $x_0 \in \mathbb{R}^{d+1, d}$ tel que

$$\forall \gamma \in \Gamma, \quad \gamma(x_0) \notin \overline{\mathcal{H}^0}. \quad (\text{II.5.4})$$

Il existe alors une (unique) suite (i_k, σ_k) (indexée par $k \geq 1$) d'éléments de $\{1, \dots, n\} \times \{-, +\}$, telle que pour tout $k \geq 0$, on a

$$\gamma^{-[k]}(x_0) \in \tilde{\mathcal{H}}_{i_{k+1}}^{\sigma_{k+1}},$$

où $\gamma^{[k]} := \gamma_{i_1}^{\sigma_1} \dots \gamma_{i_k}^{\sigma_k}$, et $\gamma^{-[k]}$ est une notation abrégée pour $(\gamma^{[k]})^{-1}$. En effet, par récurrence, supposons qu'on ait construit les k premiers termes (pour un certain $k \geq 0$) ; alors on a, par hypothèse,

$$\gamma^{-[k]}(x_0) \in \mathbb{R}^{d+1, d} \setminus \overline{\mathcal{H}^0} = \bigcup_{i=1}^n \bigcup_{\sigma=\pm} \tilde{\mathcal{H}}_i^{\sigma},$$

ce qui nous permet de choisir un couple (i_{k+1}, σ_{k+1}) approprié (qui est en fait unique car les $\tilde{\mathcal{H}}_i^{\pm}$ sont disjoints).

Notez aussi que le mot $\gamma^{[k]}$ est toujours réduit, c'est-à-dire que pour tout $k \geq 1$, on a $(i_{k+1}, \sigma_{k+1}) \neq (i_k, -\sigma_k)$. En effet, on a $\gamma^{-[k-1]}(x_0) \in \tilde{\mathcal{H}}_{i_k}^{\sigma_k}$; comme $\gamma_{i_k}^{\sigma_k}$ est bijective, en appliquant (II.5.3a) (si $\sigma_k = +1$; sinon (II.5.3b)), on obtient

$$\begin{aligned} \gamma^{-[k]}(x_0) &= \gamma_{i_k}^{-\sigma_k} \left(\gamma^{-[k-1]}(x_0) \right) \\ &\in \mathbb{R}^{d+1, d} \setminus \overline{\tilde{\mathcal{H}}_{i_k}^{-\sigma_k}}. \end{aligned}$$

On peut aussi supposer qu'il y a une infinité de valeurs de k pour lesquelles le mot $\gamma^{[k]}$ est cycliquement réduit (en d'autres termes $(i_k, \sigma_k) \neq (i_1, -\sigma_1)$.) En effet, sinon, on peut remplacer x_0 par $\gamma_i^{\sigma}(x_0)$, où (i, σ) est un couple tel que l'ensemble des indices k vérifiant $(i_k, \sigma_k) \neq (i, -\sigma)$ soit infini et contienne aussi 1 (un tel couple existe toujours). Alors la nouvelle valeur de x_0 vérifie encore (II.5.4), et la suite (i_k, σ_k) est modifiée en ajoutant (i, σ) au début.

Sans perte de généralité, on suppose que $(i_1, \sigma_1) = (1, +)$.

Une récurrence facile montre alors que les ouverts suivants forment une suite décroissante et se resserrent autour de x_0 :

$$\tilde{\mathcal{H}}_1^+ \supset \gamma^{[1]}(\tilde{\mathcal{H}}_{i_2}^{\sigma_2}) \supset \gamma^{[2]}(\tilde{\mathcal{H}}_{i_3}^{\sigma_3}) \supset \dots \ni x_0. \quad (\text{II.5.5})$$

Ensuite, on pose

$$\Delta := S \cap V_{1, >}^L.$$

On sait que S est un espace de dimension $d+1$, que $V_{1, >}^L$ est un demi-espace de dimension $d+1$ et que $S \cap V_{1, >} = 0$ (car S est défini positif et $V_{1, >}$ est isotrope). Ainsi Δ est

une demi-droite. On pose également P l'hyperplan (de dimension d) de S qui est N_S -orthogonal (ou de façon équivalente Q -orthogonal, ou encore N_0 -orthogonal) à Δ . Pour éviter des périphrases alambiquées, dans la suite, on utilisera souvent des termes comme « au-dessus » ou « en dessous » ; il est sous-entendu que le « haut » est la direction indiquée par Δ .

Considérons maintenant l'ensemble $\tilde{\mathcal{H}}_1^+ \cap (x_0 + S)$ (ici $(x_0 + S)$ désigne l'espace affine passant par x_0 et parallèle à S). Il est contenu dans un demi-espace affine de $(x_0 + S)$ situé au-dessus de l'hyperplan parallèle à P . En effet, de (II.5.1) on déduit que

$$\mathcal{H}_{S,1}^+ \subset B_{N_0} \left(\pi_S(V_{1,>}^L), \frac{\varepsilon}{3} \right).$$

Sans perte de généralité, on peut supposer que $\varepsilon \leq \text{diam } \mathbb{P}(\mathbb{R}^{d+1,d}) = \frac{\pi}{2}$ (en effet, aucun cadre ni aucun multcadre ne peut avoir une séparation qui excède cette valeur). Alors le rayon du voisinage dans la partie droite est inférieur ou égal à $\frac{\pi}{6}$. On déduit du lemme II.5.6 que

$$B_{N_0} \left(\pi_S(V_{1,>}^L), \frac{\pi}{6} \right) \cap S \subset B_{N_0} \left(\pi_S(V_{1,>}^L \cap S), \frac{\pi}{4} \right)$$

(l'angle $\frac{\pi}{4}$ est la solution de $\sin x = \sqrt{2} \sin \frac{\pi}{6}$). On peut alors en déduire la propriété recherchée.

En appliquant (II.5.5), on voit que pour tout $k \geq 0$, l'ouvert

$$\gamma^{[k]}(\tilde{\mathcal{H}}_{i_{k+1}}^{\sigma_{k+1}}) \cap (x_0 + S)$$

est inclus dans le demi-espace de $(x_0 + S)$ situé au-dessus d'un hyperplan parallèle à P . Parmi tous ces hyperplans, on appelle P_k celui qui se trouve le plus haut ; on appelle a_k l'intersection de P_k avec la droite contenant $(x_0 + \Delta)$, et on pose, pour $k \geq 1$, $\delta_k := a_k - a_{k-1}$.

Le résultat est alors une conséquence du résultat suivant :

Lemme II.5.7. *Il existe une constante $\delta_{\min} > 0$ telle que pour tout $k \geq 1$, dès que $(i_k, \sigma_k) \neq (1, -)$, on a $\|\delta_k\|_{N_0} \geq \delta_{\min}$.*

En effet, de (II.5.5), on déduit que la suite (a_k) est croissante et majorée par x_0 . Or on a choisi x_0 de telle sorte que la condition du lemme II.5.7 soit vérifiée une infinité de fois. On en déduit que (a_k) est non bornée, ce qui est une contradiction. \square

Démonstration du lemme II.5.7. On travaille toujours dans la métrique donnée par N_0 .

Soit $k \geq 1$ un indice tel que $(i_k, \sigma_k) \neq (1, -)$, de sorte que $g^{[k]}$ soit cycliquement réduit. On sait que le groupe G est pseudohyperbolique ; par la proposition II.3.8, on a

$$\alpha(V_{>}(g^{[k]}), V_{1,>}) \lesssim_{\varepsilon} s(G).$$

Comme $s(G) \leq s_3(\varepsilon)$, en choisissant $s_3(\varepsilon)$ suffisamment petit, on peut supposer que cet angle ne dépasse pas $\frac{\pi}{6}$. Le lemme II.5.6 implique alors que

$$\alpha(V_{\geq}(g^{[k]}) \cap S, V_{1,\geq} \cap S) \leq \frac{\pi}{4}$$

(on rappelle que $V_{1,\geq} \cap S$ est la droite contenant Δ).

Maintenant soit η_k le projeté de δ_k sur $V_{\geq}(g^{[k]}) \cap S$ parallèlement à P . Alors on a

$$\|\delta_k\| \geq \left(\cos \frac{\pi}{4} \right) \|\eta_k\| = \frac{\sqrt{2}}{2} \|\eta_k\|. \quad (\text{II.5.6})$$

Ensuite, toujours par la proposition II.3.8, on sait que $g^{[k]}$ est pseudohyperbolique, $\frac{\varepsilon}{3}$ -séparée et 1-contractante ; soit $\mathcal{V}^{[k]}$ son cadre. Par définition, la norme de $g^{[k]}$ restreinte à

$V_{>}(g^{[k]})$ (resp. $V_{=}(g^{[k]})$) est égale à $s(g^{[k]})$ (resp. 1). On en déduit que

$$\begin{aligned} \left\| g^{-[k]} \Big|_{V_{\geq}(g^{[k]})} \right\|_{N_{\mathcal{V}^{[k]}}} &= \max \left(\left\| g^{-[k]} \Big|_{V_{>}(g^{[k]})} \right\|, \left\| g^{-[k]} \Big|_{V_{=}(g^{[k]})} \right\| \right) \\ &= \max \left(\left\| (g_{>}^{[k]})^{-1} \right\|, 1 \right) \\ &= 1. \end{aligned}$$

Du lemme II.2.11, on déduit que

$$\left\| g^{-[k]} \Big|_{V_{\geq}(g^{[k]})} \right\|_{N_0} \lesssim_{\varepsilon} 1;$$

étant donné que, par construction, $\eta_k \in V_{\geq}(g^{[k]})$, on obtient

$$\|\eta_k\| \gtrsim_{\varepsilon} \|g^{-[k]}(\eta_k)\|. \quad (\text{II.5.7})$$

Finalement, soit x_k n'importe quel point situé à la fois dans P_k et dans $\overline{\gamma^{[k]}(\tilde{\mathcal{H}}_{i_{k+1}}^{\sigma_{k+1}})}$ (l'intersection est non vide par définition de P_k). On pose $y_{k-1} := x_k - \eta_k$. Comme le projeté orthogonal de η_k sur Δ est égal à δ_k , on en déduit que $y_{k-1} \in P_{k-1}$, et en particulier $y_{k-1} \notin \gamma^{[k-1]}(\tilde{\mathcal{H}}_{i_k}^{\sigma_k})$. En appliquant $\gamma^{-[k]}$, on obtient

$$\begin{cases} \gamma^{-[k]}(x_k) \in \overline{\tilde{\mathcal{H}}_{i_{k+1}}^{\sigma_{k+1}}} \\ \gamma^{-[k]}(y_{k-1}) \in \tilde{\mathcal{H}}_{i_k}^{-\sigma_k}. \end{cases}$$

Comme $(i_{k+1}, \sigma_{k+1}) \neq (i_k, -\sigma_k)$, on a

$$\|g^{-[k]}(\eta_k)\| = \|\gamma^{-[k]}(x_k) - \gamma^{-[k]}(y_{k-1})\| \geq d_{\min}, \quad (\text{II.5.8})$$

où d_{\min} est la plus petite distance entre deux quelconques des ensembles $\tilde{\mathcal{H}}_i^{\sigma}$ (non nulle comme on l'a vu dans preuve du lemme II.5.4).

En rassemblant (II.5.6), (II.5.7) et (II.5.8), on trouve comme on le voulait une minoration de $\|\delta_k\|$ qui ne dépend pas de k . \square

Troisième partie.

Actions affines propres sur les algèbres de Lie semisimples

III.1. Introduction

Le but de cette partie est de construire une nouvelle famille de contre-exemples à la conjecture de Milnor (la conjecture I.1.4). Elle est entièrement consacrée à la démonstration du théorème suivant :

Théorème III.1.1. *Soit G un groupe de Lie linéaire réel, semisimple et non compact. Considérons le « groupe affine » $G \ltimes \mathfrak{g}$, pour l'action adjointe de G sur son algèbre de Lie \mathfrak{g} . Alors il existe un sous-groupe $\Gamma \subset G \ltimes \mathfrak{g}$ dont la partie linéaire est Zariski-dense dans G et qui est libre, non abélien, discret et agit proprement sur l'espace affine correspondant à \mathfrak{g} .*

La stratégie générale de la preuve est la même que celle de Margulis [16] (puisque la composante neutre de $\mathrm{SO}(2,1)$ agissant sur \mathbb{R}^3 est isomorphe à $\mathrm{PSL}_2(\mathbb{R})$ agissant sur $\mathfrak{sl}_2(\mathbb{R})$, le premier exemple de Margulis est effectivement un cas particulier de notre théorème); quelques idées ont également été inspirées par [3]. On va introduire, pour certaines applications affines g , un invariant qui mesure la partie de translation de g le long de son espace neutre A_g^- (défini plus loin). Le point clé du raisonnement est alors de démontrer que sous de bonnes hypothèses, l'invariant de la composée de deux applications est proche de la somme de leurs invariants. Il y a cependant deux difficultés qui n'étaient pas présentes dans [16].

Premièrement, alors que l'invariant de Margulis originel était un scalaire, notre invariant est un vecteur. Pour le définir proprement, on a besoin d'introduire des identifications canoniques entre les différents espaces A_g^- , puis suivre les transformations que subit le représentant canonique d'un certain vecteur se trouvant dans un de ces espaces en le projetant sur d'autres espaces de cette sorte.

Deuxièmement, il se trouve que dans le cas général, la restriction de g à A_g^- n'est pas toujours une translation pure. Elle a parfois une partie de rotation, mais celle-ci est toujours confinée à un sous-espace vectoriel propre de A_g^- . L'argument s'adapte à cette situation, mais devient plus compliqué.

Une autre nouveauté que nous avons apportée est la notion de couple d'espaces C -non-dégénéré, qui, dans le cas des espaces affines, englobe à la fois une mesure quantitative de la transversalité des espaces et une majoration de leur distance à l'origine. Cette notion rend les preuves plus claires et plus simples.

III.2. Définitions et propriétés préliminaires

On fixe un groupe de Lie linéaire réel G , semisimple et non compact. Sans perte de généralité, on peut supposer que G est connexe et a un centre trivial. On voit le groupe G comme un groupe d'automorphismes de \mathfrak{g} , via la représentation adjointe; en d'autres termes, on identifie le groupe abstrait G avec le groupe linéaire $\mathrm{Ad} G \subset \mathrm{GL}(\mathfrak{g})$. Soit $\mathfrak{g}_{\mathrm{Aff}}$ l'espace affine correspondant à \mathfrak{g} . Le groupe des transformations affines de $\mathfrak{g}_{\mathrm{Aff}}$ dont la partie linéaire se trouve dans G peut alors s'écrire $G \ltimes \mathfrak{g}$ (où \mathfrak{g} représente le groupe des translations).

III.2.1. Espace affine étendu et sous-espaces dynamiques

On commence par quelques définitions.

On choisit une fois pour toutes un point de $\mathfrak{g}_{\mathrm{Aff}}$ qu'on prend comme origine; on appelle \mathbb{R}_0 l'espace vectoriel à une dimension formellement engendré par ce point, et on pose $\hat{\mathfrak{g}} := \mathfrak{g} \oplus \mathbb{R}_0$ l'espace affine étendu correspondant à \mathfrak{g} . Alors $\mathfrak{g}_{\mathrm{Aff}}$ est l'hyperplan affine « à hauteur un » de cet espace, et \mathfrak{g} est l'hyperplan vectoriel correspondant :

$$\mathfrak{g} = \mathfrak{g} \times \{0\} \subset \mathfrak{g} \times \mathbb{R}_0; \quad \mathfrak{g}_{\mathrm{Aff}} = \mathfrak{g} \times \{1\} \subset \mathfrak{g} \times \mathbb{R}_0.$$

Toute application affine g de partie linéaire $\ell(g)$ et de vecteur de translation v , définie sur $\mathfrak{g}_{\text{Aff}}$ par

$$g : x \mapsto \ell(g)(x) + v,$$

s'étend de façon unique en une application linéaire définie sur $\hat{\mathfrak{g}}$, donnée par la matrice

$$\begin{pmatrix} \ell(g) & v \\ 0 & 1 \end{pmatrix}.$$

Ainsi le groupe affine $G \ltimes \mathfrak{g}$ agit naturellement sur $\hat{\mathfrak{g}}$.

On appelle *sous-espace affine étendu* de $\hat{\mathfrak{g}}$ tout sous-espace vectoriel de $\hat{\mathfrak{g}}$ non contenu dans \mathfrak{g} . Il y a une correspondance bijective entre les sous-espaces affines étendus de $\hat{\mathfrak{g}}$ et les sous-espaces affines de $\mathfrak{g}_{\text{Aff}}$. On adoptera la convention d'utiliser la lettre A pour les sous-espaces affines étendus de $\hat{\mathfrak{g}}$, et la lettre V pour les sous-espaces vectoriels de \mathfrak{g} .

Par abus de terminologie, les éléments du sous-groupe distingué $\mathfrak{g} \triangleleft G \ltimes \mathfrak{g}$ seront toujours appelés *translations*, même si on les considérera la plupart du temps comme des endomorphismes de $\hat{\mathfrak{g}}$ (ce sont donc formellement des transvections). Pour tout vecteur $v \in \mathfrak{g}$, on note τ_v la translation correspondante.

Pour tout $g \in G \ltimes \mathfrak{g}$, on décompose $\hat{\mathfrak{g}}$ en une somme directe de trois sous-espaces

$$\hat{\mathfrak{g}} = V_g^> \oplus A_g^= \oplus V_g^< ,$$

appelés *sous-espaces dynamiques* de g , qui sont tous stables par g et tels que toute valeur propre λ de la restriction de g à $V_g^>$ (resp. $A_g^=$, $V_g^<$) vérifie $|\lambda| > 1$ (resp. $|\lambda| = 1$, $|\lambda| < 1$). On pose également $A_g^{\geq} := V_g^> \oplus A_g^=$ et $A_g^{\leq} := V_g^< \oplus A_g^=$.

Dans ce cas, on a bien entendu $V_g^> \subset \mathfrak{g}$ et $V_g^< \subset \mathfrak{g}$ mais $A_g^= \not\subset \mathfrak{g}$ (ce qui justifie le choix des lettres A et V). On a par conséquent

$$\mathfrak{g} = V_g^> \oplus V_g^= \oplus V_g^< ,$$

où on note $V_g^= := A_g^= \cap \mathfrak{g}$ la partie linéaire de $A_g^=$ (ou plus précisément de l'espace affine correspondant). On pose également

$$\begin{aligned} V_g^{\geq} &:= A_g^{\geq} \cap \mathfrak{g} = V_g^> \oplus V_g^= \\ \text{et } V_g^{\leq} &:= A_g^{\leq} \cap \mathfrak{g} = V_g^< \oplus V_g^= . \end{aligned}$$

On dira qu'une application $g \in G \ltimes \mathfrak{g}$ est \mathbb{R} -régulière si sa partie linéaire est \mathbb{R} -régulière, c'est-à-dire si la dimension de l'espace $A_g^=$ (ou de sa partie linéaire $V_g^=$) est la plus petite possible. (La valeur minimale de la dimension sera donnée dans la preuve de l'affirmation III.2.2.)

III.2.2. Structure d'algèbre de Lie

On introduit maintenant quelques sous-algèbres classiques de \mathfrak{g} . L'intérêt est que si un élément $g \in G \ltimes \mathfrak{g}$ est \mathbb{R} -régulier, alors ses sous-espaces dynamiques sont égaux, à conjugaison près, à certaines de ces sous-algèbres (voir le corollaire III.2.3).

On choisit dans \mathfrak{g} :

- une involution de Cartan θ . Alors on a la décomposition de Cartan correspondante $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{q}$, où on appelle \mathfrak{k} l'espace des points fixes de θ et \mathfrak{q} l'espace des points fixes de $-\theta$. On appelle K le sous-groupe compact maximal de G d'algèbre de Lie \mathfrak{k} .
- un *sous-espace de Cartan* \mathfrak{a} compatible avec θ (c'est-à-dire une sous-algèbre abélienne de \mathfrak{g} maximale parmi celles qui sont contenues dans \mathfrak{q}). On pose $A := \exp \mathfrak{a}$.

- un système Σ^+ de racines restreintes positives dans \mathfrak{a}^* . On rappelle qu'une *racine restreinte* est un élément non nul $\alpha \in \mathfrak{a}^*$ tel que l'espace de racine

$$\mathfrak{g}_\alpha := \{Y \in \mathfrak{g} \mid \forall X \in \mathfrak{a}, [X, Y] = \alpha(X)Y\}$$

est non trivial. Elles forment un système de racines Σ ; un système de racines positives Σ^+ est une partie de Σ contenue dans un demi-espace et telle que $\Sigma = \Sigma^+ \sqcup -\Sigma^+$. On pose

$$\mathfrak{a}^+ := \{X \in \mathfrak{a} \mid \forall \alpha \in \Sigma^+, \alpha(X) > 0\}$$

la chambre de Weyl (ouverte) correspondante de \mathfrak{a} .

On pose alors :

- M le centralisateur de \mathfrak{a} dans K , \mathfrak{m} son algèbre de Lie.
- L le centralisateur de \mathfrak{a} dans G , \mathfrak{l} son algèbre de Lie. Il est clair que $\mathfrak{l} = \mathfrak{a} \oplus \mathfrak{m}$, et il est bien connu (voir par exemple [14], Proposition 7.82a) que $L = MA$.
- \mathfrak{n}^+ (resp. \mathfrak{n}^-) la somme des espaces de racine restreinte de Σ^+ (resp. de $-\Sigma^+$).
- $\mathfrak{p}^+ := \mathfrak{l} \oplus \mathfrak{n}^+$ et $\mathfrak{p}^- := \mathfrak{l} \oplus \mathfrak{n}^-$ les algèbres paraboliques minimales correspondantes.
- $\hat{\mathfrak{l}}$, $\hat{\mathfrak{p}}^+$ et $\hat{\mathfrak{p}}^-$ les extensions vectorielles des sous-espaces affines de $\mathfrak{g}_{\text{Aff}}$ parallèles respectivement à \mathfrak{l} , \mathfrak{p}^+ et \mathfrak{p}^- et passant par l'origine. En d'autres termes :

$$\hat{\mathfrak{l}} := \mathfrak{l} \oplus \mathbb{R}_0 \text{ et } \hat{\mathfrak{p}}^\pm := \mathfrak{p}^\pm \oplus \mathbb{R}_0.$$

Il est commode pour nous de définir une *algèbre parabolique minimale* (abrégée a.p.m. dans la suite) dans \mathfrak{g} comme l'image de \mathfrak{p}^+ par n'importe quel élément de G . De même, on définit une *a.p.m. affine* dans $\hat{\mathfrak{g}}$ comme l'image de $\hat{\mathfrak{p}}^+$ par n'importe quel élément de $G \times \mathfrak{g}$. De façon équivalente, un sous-espace $\hat{\mathfrak{p}}_1 \subset \hat{\mathfrak{g}}$ est une a.p.m. affine si et seulement s'il n'est pas contenu dans \mathfrak{g} et que sa partie linéaire $\hat{\mathfrak{p}}_1 \cap \mathfrak{g}$ est une a.p.m.

On dit que deux a.p.m. (ou deux a.p.m. affines) sont *transverses* si leur intersection a la plus petite dimension possible.

Exemple III.2.1. Un cas particulier important est $G = \text{PSL}_n(\mathbb{R})$. Dans ce cas, on peut prendre pour θ l'involution $X \mapsto -X^t$, pour \mathfrak{a} l'ensemble des matrices diagonales (de trace nulle), et pour Σ^+ l'ensemble de toutes les racines $e_i^* - e_j^*$ telles que $i < j$ (où e_i^* est la forme linéaire définie par $e_i^*(\text{Diag}(a_1, \dots, a_n)) = a_i$). Dans ce cas :

- \mathfrak{k} (resp. \mathfrak{q}) est l'ensemble des matrices antisymétriques (resp. symétriques) de trace nulle, et $K = \text{SO}_n(\mathbb{R})$;
- \mathfrak{a}^+ est l'ensemble des matrices de la forme $\text{Diag}(\lambda_1, \dots, \lambda_n)$ de trace nulle et avec $\lambda_1 > \dots > \lambda_n$;
- \mathfrak{m} est réduit à 0; \mathfrak{l} est égal à \mathfrak{a} ; A (resp. M , L) est le groupe des matrices diagonales de déterminant 1 de coefficients positifs (resp. égaux à ± 1 , quelconques);
- \mathfrak{n}^+ (resp. \mathfrak{n}^-) est l'ensemble des matrices triangulaires supérieures (resp. inférieures) avec des zéros sur la diagonale;
- \mathfrak{p}^+ (resp. \mathfrak{p}^-) est l'ensemble de toutes les matrices triangulaires supérieures (resp. inférieures) de trace nulle.

III.2.3. Propriétés algébriques de base

On a les propriétés algébriques suivantes.

Affirmation III.2.2. Soit $g \in G \times \mathfrak{g}$.

- L'application g est \mathbb{R} -régulière si et seulement si elle est conjuguée (par un élément de $G \times \mathfrak{g}$) à un produit $\tau_v m \exp(a)$ avec $v \in \mathfrak{l}$, $m \in M$ et $a \in \mathfrak{a}^+$ (on identifie ici le sous-groupe du groupe affine $G \times \mathfrak{g}$ fixant l'« origine » \mathbb{R}_0 avec le groupe linéaire G).*
- Dans ce cas, A_g^{\geq} et A_g^{\leq} sont des a.p.m. affines transverses.*

(iii) De plus, dans ce cas $V_g^>$ (resp. $V_g^<$) est déterminé de façon unique par A_g^{\geq} (resp. by A_g^{\leq}).

Démonstration.

(i) Montrons que pour tout $g \in G \times \mathfrak{g}$, on a $\dim A_g^- \geq \dim \hat{\mathfrak{l}}$, avec égalité si et seulement si g a la forme requise.

On peut décomposer g en un produit $g = \tau_v g_h g_e g_u$, où v est un certain vecteur dans \mathfrak{g} , $g_h \in G$ est hyperbolique (semisimple à valeurs propres réelles positives), $g_e \in G$ est elliptique (semisimple à valeurs propres de module 1), $g_u \in G$ est unipotent, et les trois dernières applications commutent deux à deux. À conjugaison près, on peut supposer que A_g^- passe par l'origine \mathbb{R}_0 , ce qui signifie que $v \in V_g^-$.

Comme g_e et g_u ont toutes leurs valeurs propres de module 1 et commutent avec g_h , on a $A_{g_h}^- = A_{g_h}^-$. À conjugaison près, on peut supposer que $g_h = \exp a$ où a est un élément de \mathfrak{a} , et même, plus précisément, de l'adhérence de \mathfrak{a}^+ . Alors clairement l'espace $A_{g_h}^-$ est la somme de $\hat{\mathfrak{l}}$ et de tous les espaces de racine restreinte \mathfrak{g}_α tels que la valeur $\alpha(a)$ se trouve être nulle. Ceci montre que $A_{g_h}^-$ contient $\hat{\mathfrak{l}}$, et que l'égalité se produit si et seulement si $a \in \mathfrak{a}^+$.

Clairement si g a la forme requise, alors $g_h = \exp a$, $g_e = m$ et $g_u = 1$, de sorte que $a \in \mathfrak{a}^+$. Réciproquement, supposons que $a \in \mathfrak{a}^+$; montrons que g a la forme requise. On commence par observer que deux chambres de Weyl distinctes sont toujours disjointes; par conséquent le seul conjugué de \mathfrak{a}^+ contenant a est \mathfrak{a}^+ lui-même. On en déduit que $Z_G(a) = Z_G(\mathfrak{a}) = L$. Alors g_e est un élément elliptique de L , donc un élément de M ; g_u est un élément unipotent de L , donc égal à 1; et $v \in V_g^- = \mathfrak{l}$. On en déduit que $g = \tau_v g_e \exp a$ avec $v \in \mathfrak{l}$, $g_e \in M$ et $a \in \mathfrak{a}^+$, comme demandé.

(ii) D'après le point précédent, à conjugaison près, on peut supposer que $g = \tau_v m \exp a$ avec $v \in \mathfrak{l}$, $m \in M$ et $a \in \mathfrak{a}^+$. Mais alors clairement on a $A_g^{\geq} = A_{\exp a}^{\geq} = \hat{\mathfrak{p}}^+$ et de même $A_g^{\leq} = \hat{\mathfrak{p}}^-$.

(iii) Si g est de la forme $\tau_v m \exp a$, alors, de même, on a $V_g^> = \mathfrak{n}^+$; et on sait que \mathfrak{n}^+ est le nilradical (plus grand idéal nilpotent) de \mathfrak{p}^+ . Donc pour toute application g \mathbb{R} -régulière, $V_g^>$ est le nilradical de V_g^{\geq} , qui est la partie linéaire de A_g^{\geq} (c'est-à-dire que $V_g^{\geq} = A_g^{\geq} \cap \mathfrak{g}$). De même, $V_g^<$ est le nilradical de la partie linéaire de A_g^{\leq} . \square

Corollaire III.2.3. *Pour toute application \mathbb{R} -régulière $g \in G \times \mathfrak{g}$, il existe une application « canonisante » $\phi \in G \times \mathfrak{g}$ telle que :*

$$\begin{aligned} \phi(V_g^>) &= \mathfrak{n}^+ & \phi(A_g^-) &= \hat{\mathfrak{l}} \\ \phi(V_g^=) &= \mathfrak{l} & \phi(A_g^{\geq}) &= \hat{\mathfrak{p}}^+ \\ \phi(V_g^<) &= \mathfrak{n}^- & \phi(A_g^{\leq}) &= \hat{\mathfrak{p}}^- \\ \phi(V_g^{\geq}) &= \mathfrak{p}^+ & & \\ \phi(V_g^{\leq}) &= \mathfrak{p}^- & & \end{aligned}$$

Exemple III.2.4. Pour $G = \mathrm{PSL}_n(\mathbb{R})$, une application $g \in G \times \mathfrak{g}$ est \mathbb{R} -régulière si et seulement si sa partie linéaire, vue comme un automorphisme de \mathbb{R}^n (et non pas de \mathfrak{g} selon notre convention usuelle), a des valeurs propres réelles de valeurs absolues deux à deux distinctes.

Affirmation III.2.5. *Tout couple d'a.p.m. transverses (resp. d'a.p.m. affines transverses) peut être envoyé sur $(\mathfrak{p}^+, \mathfrak{p}^-)$ (resp. $(\hat{\mathfrak{p}}^+, \hat{\mathfrak{p}}^-)$) par un certain élément de G (resp. de $G \times \mathfrak{g}$).*

Démonstration. On va démontrer la version linéaire ; la version affine s'en déduit immédiatement. Soit $(\mathfrak{p}_1, \mathfrak{p}_2)$ un tel couple. Par définition, pour $i = 1, 2$, on peut écrire $\mathfrak{p}_i = \phi_i(\mathfrak{p}^+)$ pour un certain $\phi_i \in G$. On va appliquer la décomposition de Bruhat à l'application $\phi_1^{-1}\phi_2$: on peut écrire

$$\phi_1^{-1}\phi_2 = p_1 w p_2,$$

où p_1, p_2 appartiennent au sous-groupe parabolique minimal $P^+ := N_G(\mathfrak{p}^+)$, et w est un élément du groupe de Weyl $W := N_G(\mathfrak{a})/Z_G(\mathfrak{a})$ (voir par exemple le théorème 7.40 dans [14]). Soit $\phi := \phi_1 p_1 = \phi_2 p_2^{-1} w^{-1}$; alors on a

$$\mathfrak{p}_1 = \phi(\mathfrak{p}^+) \text{ et } \mathfrak{p}_2 = \phi(w\mathfrak{p}^+).$$

On en déduit que $w\mathfrak{p}^+$ est transverse à \mathfrak{p}^+ . Ceci se produit si et seulement si w est égal à w_0 , le mot le plus long du groupe de Weyl ; or $w_0\mathfrak{p}^+ = \mathfrak{p}^-$. Ainsi $\mathfrak{p}_1 = \phi(\mathfrak{p}^+)$ et $\mathfrak{p}_2 = \phi(\mathfrak{p}^-)$ comme demandé. \square

Affirmation III.2.6. *Toute application $\phi \in G \times \mathfrak{g}$ qui laisse invariants les deux espaces $\hat{\mathfrak{p}}^+$ et $\hat{\mathfrak{p}}^-$ appartient au groupe $L \times \mathfrak{l}$.*

Démonstration. Il est bien connu (cela découle par exemple du lemme 7.64 de [14]) que $N_G(\mathfrak{p}^+) \cap N_G(\mathfrak{p}^-) = Z_G(\mathfrak{a}) = L$; donc la partie linéaire d'une telle application ϕ doit se trouver dans L . Puisque ϕ laisse invariant l'espace $\hat{\mathfrak{p}}^+ \cap \hat{\mathfrak{p}}^- = \hat{\mathfrak{l}}$, sa partie de translation doit se trouver dans \mathfrak{l} . \square

III.2.4. Quasi-translations

Dans cette sous-section, on va compléter l'affirmation III.2.6 en étudiant l'action des éléments de $L \times \mathfrak{l}$ sur l'espace $\hat{\mathfrak{l}}$.

Définition III.2.7. On appelle *quasi-translation* tout automorphisme affine de $\hat{\mathfrak{l}}$ induit par un élément du groupe $L \times \mathfrak{l}$.

On va maintenant expliquer et justifier cette terminologie. On appelle $Z := Z(L)$ le centre de L , $D := [L, L]$ son sous-groupe dérivé, et \mathfrak{z} et \mathfrak{d} les algèbres de Lie correspondantes. Il est bien connu que le groupe L est réductif, ce qui permet d'écrire

$$\mathfrak{l} = \mathfrak{z} \oplus \mathfrak{d}.$$

On peut en fait montrer (voir le théorème 7.53 b et c de [14]) que Z rencontre chaque composante connexe de L . On peut donc également écrire

$$L = ZD.$$

Par définition, L agit trivialement sur \mathfrak{z} et Z agit trivialement sur \mathfrak{l} ; la seule action non triviale est celle de D sur \mathfrak{d} . De plus, D préserve la forme de Killing, qui est définie négative sur \mathfrak{d} (comme \mathfrak{a} est abélienne, on a $\mathfrak{d} \subset \mathfrak{m} \subset \mathfrak{k}$). Pour résumer tout cela :

Proposition III.2.8. *Toute quasi-translation est un élément de $(O(\mathfrak{d}) \times \mathfrak{d}) \times \mathfrak{z}$.*

En d'autres termes, les quasi-translations sont des isométries affines de \mathfrak{l} qui préservent les directions de \mathfrak{d} et \mathfrak{z} et agissent seulement par translation sur la composante \mathfrak{z} .

Remarque III.2.9. Sachant que \mathfrak{l} peut aussi se décomposer comme $\mathfrak{a} \oplus \mathfrak{m}$, on a $\mathfrak{z} = \mathfrak{a} \oplus \mathfrak{z}(\mathfrak{m})$ (somme directe de \mathfrak{a} et du centre de l'algèbre de Lie \mathfrak{m}) et $\mathfrak{d} = [\mathfrak{m}, \mathfrak{m}]$. Par conséquent la dernière proposition, et en fait l'ensemble des affirmations du reste de la partie III, resteraient valables si on substituait, respectivement, \mathfrak{a} et \mathfrak{m} à la place de \mathfrak{z} et \mathfrak{d} . L'avantage d'introduire \mathfrak{z} et \mathfrak{d} est que les invariants de Margulis (voir plus loin) vivent dans un espace plus grand, et sont donc des invariants plus fins. Cela pourrait servir pour des travaux ultérieurs.

Exemple III.2.10. Pour $G = \mathrm{PSL}_n(\mathbb{R})$, comme l'algèbre $\mathfrak{l} = \mathfrak{a}$ est abélienne, \mathfrak{z} coïncide avec \mathfrak{l} ou de façon équivalente \mathfrak{d} est réduit à 0, donc une quasi-translation est simplement une translation. Le cas le plus simple où $\mathfrak{d} \neq 0$ est $G = \mathrm{SO}^+(4,1)$ (dans ce cas D , ainsi que M , est isomorphe à $\mathrm{SO}(3)$).

III.2.5. Identifications canoniques

Ici on introduit des identifications canoniques (à quasi-translation près) entre les différents espaces A_g^- (dans le corollaire III.2.11), et on les utilise pour définir l'invariant de Margulis d'une application \mathbb{R} -régulière. On vérifie également que ces identifications commutent avec certaines projections « naturelles » (dans le lemme III.2.14).

Les deux propriétés suivantes sont des conséquences immédiates des affirmations III.2.2, III.2.5 et III.2.6 :

Corollaire III.2.11. *Soit $(\hat{\mathfrak{p}}_1, \hat{\mathfrak{p}}_2)$ un couple d'a.p.m. affines transverses. Alors toute application $\phi \in G \times \mathfrak{g}$ telle que $\phi(\hat{\mathfrak{p}}_1, \hat{\mathfrak{p}}_2) = (\hat{\mathfrak{p}}^+, \hat{\mathfrak{p}}^-)$ donne, par restriction, une identification de l'intersection $\hat{\mathfrak{p}}_1 \cap \hat{\mathfrak{p}}_2$ avec $\hat{\mathfrak{l}}$, unique à quasi-translation près.*

Ici la notation $\phi(\hat{\mathfrak{p}}_1, \hat{\mathfrak{p}}_2)$ désigne le couple $(\phi(\hat{\mathfrak{p}}_1), \phi(\hat{\mathfrak{p}}_2))$. Notez que si le même espace est obtenu d'une autre façon comme intersection de deux a.p.m. affines, son identification avec $\hat{\mathfrak{l}}$ ne sera pas la même (elle diffèrera par un élément du groupe de Weyl).

Corollaire III.2.12. *Soit $g \in G \times \mathfrak{g}$ une application \mathbb{R} -régulière. Soit $\phi \in G \times \mathfrak{g}$ n'importe quelle application vérifiant $\phi(A_g^{\geq}, A_g^{\leq}) = (\hat{\mathfrak{p}}^+, \hat{\mathfrak{p}}^-)$. Alors la restriction du conjugué $\phi g \phi^{-1}$ à $\hat{\mathfrak{l}}$ est une quasi-translation.*

Cela nous permet de donner la définition suivante. On appelle $\pi_{\mathfrak{z}}$ la projection depuis \mathfrak{l} sur \mathfrak{z} parallèlement à \mathfrak{d} .

Définition III.2.13. Soit $g \in G \times \mathfrak{g}$ une application \mathbb{R} -régulière. Prenons n'importe quel point x dans l'espace affine $A_g^- \cap \mathfrak{g}_{\mathrm{Aff}}$ et n'importe quelle application $\phi \in G$ telle que $\phi(V_g^{\geq}, V_g^{\leq}) = (\mathfrak{p}^+, \mathfrak{p}^-)$. Alors le vecteur

$$M(g) := \pi_{\mathfrak{z}}(\phi(g(x) - x)) \in \mathfrak{z}$$

ne dépend pas du choix de x ni de ϕ . On l'appelle l'*invariant de Margulis* de g . (Il s'agit d'un invariant de conjugaison.)

Lemme III.2.14. *Soit $\hat{\mathfrak{p}}_1$ une a.p.m. affine quelconque. Soit \mathfrak{n}_1 le nilradical de sa partie linéaire, et $\hat{\mathfrak{p}}_2$ et $\hat{\mathfrak{p}}'_2$ deux a.p.m. affines quelconques toutes deux transverses à $\hat{\mathfrak{p}}_1$. Soit ϕ (resp. ϕ') un élément de $G \times \mathfrak{g}$ qui envoie le couple $(\hat{\mathfrak{p}}_1, \hat{\mathfrak{p}}_2)$ (resp. $(\hat{\mathfrak{p}}_1, \hat{\mathfrak{p}}'_2)$) sur $(\hat{\mathfrak{p}}^+, \hat{\mathfrak{p}}^-)$ (ces deux applications existent d'après l'affirmation III.2.5). Soit*

$$\psi : \hat{\mathfrak{p}}_1 \cap \hat{\mathfrak{p}}_2 \longrightarrow \hat{\mathfrak{p}}_1 \cap \hat{\mathfrak{p}}'_2$$

la projection parallèle à \mathfrak{n}_1 . Alors l'application $\bar{\psi}$ définie par le diagramme commutatif

$$\begin{array}{ccc} \hat{\mathfrak{l}} & \xrightarrow{\bar{\psi}} & \hat{\mathfrak{l}} \\ \uparrow \phi & & \uparrow \phi' \\ \hat{\mathfrak{p}}_1 \cap \hat{\mathfrak{p}}_2 & \xrightarrow{\psi} & \hat{\mathfrak{p}}_1 \cap \hat{\mathfrak{p}}'_2 \end{array}$$

est une quasi-translation.

La projection ψ est bien définie car $\hat{\mathfrak{p}}^+ = \mathfrak{n}^+ \oplus \hat{\mathfrak{l}} = \mathfrak{n}^+ \oplus (\hat{\mathfrak{p}}^+ \cap \hat{\mathfrak{p}}^-)$, et donc $\hat{\mathfrak{p}}_1 = \phi'^{-1}(\hat{\mathfrak{p}}^+) = \mathfrak{n}_1 \oplus (\hat{\mathfrak{p}}_1 \cap \hat{\mathfrak{p}}'_2)$.

Démonstration. Sans perte de généralité, on peut supposer que $\phi = \text{Id}$ (sinon on peut simplement remplacer les trois a.p.m. affines par leurs images par ϕ^{-1} .) Alors on a $\hat{\mathfrak{p}}_1 = \hat{\mathfrak{p}}^+$, $\hat{\mathfrak{p}}_2 = \hat{\mathfrak{p}}^-$ et $\hat{\mathfrak{p}}'_2 = \phi'^{-1}(\hat{\mathfrak{p}}^-)$, où ϕ' peut être n'importe quelle application qui stabilise l'espace $\hat{\mathfrak{p}}^+$. On veut montrer que l'application $\phi' \circ \psi$ est une quasi-translation.

On sait que ϕ' se trouve dans le stabilisateur $N_{G \times \mathfrak{g}}(\hat{\mathfrak{p}}^+)$, qui est égal à $P^+ \ltimes \mathfrak{p}^+$, où $P^+ := N_G(\mathfrak{p}^+)$ est le sous-groupe parabolique minimal d'algèbre de Lie \mathfrak{p}^+ . On utilisera la décomposition de Langlands

$$P^+ = MAN^+ = LN^+,$$

où N^+ est le groupe connexe d'algèbre de Lie \mathfrak{n}^+ (voir par exemple la proposition 7.83 de [14]). Puisque L normalise \mathfrak{n}^+ et que $\mathfrak{l} \oplus \mathfrak{n}^+ = \mathfrak{p}^+$, cela se généralise à la « décomposition de Langlands affine »

$$P^+ \ltimes \mathfrak{p}^+ = (L \ltimes \mathfrak{l})(N^+ \ltimes \mathfrak{n}^+).$$

On peut donc écrire $\phi' = l \circ n$ avec $l \in L \ltimes \mathfrak{l}$ et $n \in N^+ \ltimes \mathfrak{n}^+$.

On utilisera la propriété suivante : tout élément n du groupe $N^+ \ltimes \mathfrak{n}^+$ stabilise l'espace \mathfrak{n}^+ et induit l'identité sur l'espace quotient $\hat{\mathfrak{p}}^+/\mathfrak{n}^+$. En effet, lorsque n est dans le groupe « linéaire » N^+ , ce groupe étant connexe, cela découle du fait que \mathfrak{n}^+ est un idéal de \mathfrak{p}^+ . Lorsque n est une translation pure par un vecteur de \mathfrak{n}^+ , c'est évident.

Par définition, la même propriété est vraie pour ψ ; par conséquent l'application $n \circ \psi$ stabilise encore \mathfrak{n}^+ et induit encore l'identité sur $\hat{\mathfrak{p}}^+/\mathfrak{n}^+$. Mais on sait d'autre part que $n \circ \psi$ est définie sur $\hat{\mathfrak{p}}_1 \cap \hat{\mathfrak{p}}_2 = \hat{\mathfrak{l}}$, et l'envoie sur

$$n \circ \psi(\hat{\mathfrak{p}}_1 \cap \hat{\mathfrak{p}}_2) = n(\hat{\mathfrak{p}}_1 \cap \hat{\mathfrak{p}}'_2) = l^{-1}(\hat{\mathfrak{l}}) = \hat{\mathfrak{l}}.$$

Donc l'application $n \circ \psi$ est égale à l'identité sur $\hat{\mathfrak{l}}$. On en déduit que $\bar{\psi} = \phi' \circ \psi = l \circ n \circ \psi = l$ est une quasi-translation (sur $\hat{\mathfrak{l}}$) comme demandé. \square

III.2.6. Propriétés métriques

Ici on introduit quelques conventions et on définit deux propriétés métriques importantes des applications \mathbb{R} -régulières : la C -non-dégénérescence (qui signifie que la géométrie de l'application n'est pas trop proche d'un cas dégénéré) et la force de contraction.

On introduit sur $\hat{\mathfrak{g}}$ une norme euclidienne qui rend les sous-espaces \mathfrak{n}^+ , \mathfrak{n}^- , \mathfrak{d} , \mathfrak{z} et \mathbb{R}_0 deux à deux orthogonaux, et dont la restriction à \mathfrak{d} coïncide avec la forme de Killing au signe près (en effet cette dernière est définie négative sur $\mathfrak{d} \subset \mathfrak{m}$). Pour toute application linéaire g agissant sur $\hat{\mathfrak{g}}$, on note $\|g\| := \sup_{x \neq 0} \frac{\|g(x)\|}{\|x\|}$ sa norme d'opérateur.

Considérons un espace euclidien E (pour le moment, le lecteur peut supposer que $E = \hat{\mathfrak{g}}$; plus tard on aura aussi besoin du cas où $E = \Lambda^q \hat{\mathfrak{g}}$ pour un certain entier q). On définit sur l'espace projectif $\mathbb{P}(E)$ une distance en posant, pour tous $\bar{x}, \bar{y} \in \mathbb{P}(E)$,

$$\alpha(\bar{x}, \bar{y}) := \arccos \frac{|\langle x, y \rangle|}{\|x\| \|y\|},$$

où x et y sont n'importe quels vecteurs représentant respectivement \bar{x} et \bar{y} (bien entendu la valeur ne dépend pas du choix de x et y). Celle-ci mesure l'angle entre les droites \bar{x} et \bar{y} . Par souci de concision, on se contentera généralement d'écrire $\alpha(x, y)$ avec x et y des vecteurs de $E \setminus \{0\}$.

Pour tout sous-espace vectoriel non nul $F \subset E$ et pour tout rayon $\varepsilon > 0$, on notera

$$B_{\mathbb{P}}(F, \varepsilon) := \{x \in \mathbb{P}(E) \mid \alpha(x, \mathbb{P}(F)) < \varepsilon\}$$

le ε -voisinage de F dans $\mathbb{P}(E)$. (On peut le voir comme une sorte de « voisinage conique ».)

Considérons un espace métrique (\mathcal{M}, δ) ; soient X et Y deux parties non vides de \mathcal{M} . On notera

$$\delta(X, Y) := \inf_{x \in X} \inf_{y \in Y} \delta(x, y)$$

la distance ordinaire (distance minimale) entre X et Y , et

$$\delta^{\text{Haus}}(X, Y) := \max \left(\sup_{x \in X} \delta(x, Y), \sup_{y \in Y} \delta(y, X) \right)$$

la distance de Hausdorff entre X et Y .

Finalement, on introduit la notation suivante. Soient A et B deux quantités strictement positives, et p_1, \dots, p_k des paramètres. En écrivant

$$A \lesssim_{p_1, \dots, p_k} B,$$

on voudra dire qu'il existe une constante K , qui ne dépend de rien d'autre que p_1, \dots, p_k , telle que $A \leq KB$. (Si on n'écrit aucun paramètre, cela signifie bien sûr que K est une constante absolue.) En écrivant

$$A \gtrsim_{p_1, \dots, p_k} B,$$

on voudra dire que $A \lesssim_{p_1, \dots, p_k} B$ et $B \lesssim_{p_1, \dots, p_k} A$ à la fois.

Définition III.2.15. Soient $(\hat{\mathfrak{p}}_1, \hat{\mathfrak{p}}_2)$ un couple d'a.p.m. affines. Une *application canonisante optimale* pour ce couple est une application $\phi \in G \times \mathfrak{g}$ vérifiant

$$\phi(\hat{\mathfrak{p}}_1, \hat{\mathfrak{p}}_2) = (\hat{\mathfrak{p}}^+, \hat{\mathfrak{p}}^-)$$

et minimisant la quantité $\max(\|\phi\|, \|\phi^{-1}\|)$. D'après l'affirmation III.2.5, une telle application existe si et seulement si $\hat{\mathfrak{p}}_1$ et $\hat{\mathfrak{p}}_2$ sont transverses.

On appelle *application canonisante optimale* d'une application \mathbb{R} -régulière $g \in G \times \mathfrak{g}$ toute application canonisante optimale du couple (A_g^{\geq}, A_g^{\leq}) .

Soit $C \geq 1$. On dit qu'un couple d'a.p.m. affines $(\hat{\mathfrak{p}}_1, \hat{\mathfrak{p}}_2)$ (resp. une application \mathbb{R} -régulière g) est *C-non-dégénéré* s'il a une application canonisante optimale ϕ telle que $\|\phi^{\pm 1}\| \leq C$.

Prenons maintenant g_1 et g_2 deux applications \mathbb{R} -régulières dans $G \times \mathfrak{g}$. On dit que le couple (g_1, g_2) est *C-non-dégénéré* si chacun des quatre couples possibles $(A_{g_i}^{\geq}, A_{g_j}^{\leq})$ est *C-non-dégénéré*.

L'intérêt de cette définition est qu'il y a beaucoup de calculs où, en traitant un couple d'espaces *C-non-dégénéré* comme s'ils étaient perpendiculaires, on se trompe au plus d'une constante (multiplicative) qui dépend de C . Le résultat suivant sera souvent utile :

Lemme III.2.16. *Soit $C \geq 1$. Alors toute application $\phi \in \text{GL}(E)$ telle que $\|\phi^{\pm 1}\| \leq C$ induit une application lipschitzienne sur $\mathbb{P}(E)$, de constante de Lipschitz égale à C^2 .*

Démonstration. Il suffit de vérifier ceci pour toute restriction de ϕ à un sous-espace de dimension 2 de E . Mais dans un plan, en utilisant la décomposition en valeurs singulières (voir la preuve du lemme III.3.8 (iii) pour une définition), cette identité est évidente. \square

Remarque III.2.17. L'ensemble des couples transverses de sous-espaces affines étendus est caractérisé par deux conditions ouvertes : il y a bien sûr la transversalité des espaces, mais il y a aussi la condition, pour chaque espace, de ne pas être inclus dans \mathfrak{g} . Le mot « dégénérescence » signifie ici la mise en défaut d'une des deux conditions. Ainsi la propriété qu'une paire $(\hat{\mathfrak{p}}_1, \hat{\mathfrak{p}}_2)$ soit *C-non-dégénérée* englobe en fait deux propriétés différentes.

Premièrement, elle implique que les espaces $\hat{\mathfrak{p}}_1$ et $\hat{\mathfrak{p}}_2$ sont transverses, de façon quantitative. Plus précisément, cela veut dire qu'une certaine fonction continue qui s'annulerait si les espaces étaient transverses est minorée. Un exemple d'une telle fonction est le plus

petit terme non identiquement nul dans la suite des « angles principaux » définie dans la preuve du lemme III.3.8 (iv).

Deuxièmement, elle implique que chacun des espaces $\hat{\mathfrak{p}}_1$ et $\hat{\mathfrak{p}}_2$ n'est « pas trop proche » de l'espace \mathfrak{g} (dans le même sens). En termes purement affines, cela signifie que les espaces affines $\hat{\mathfrak{p}}_1 \cap \mathfrak{g}_{\text{Aff}}$ et $\hat{\mathfrak{p}}_2 \cap \mathfrak{g}_{\text{Aff}}$ ne passent pas trop loin de l'origine.

Les deux conditions sont nécessaires, et apparaissent dans la littérature précédente (comme [16] et [2]) ; mais pour le moment, elles ont toujours été traitées séparément.

Définition III.2.18. Soit $s > 0$. On dit qu'une application \mathbb{R} -régulière $g \in G \times \mathfrak{g}$ est *s-contractante* si on a :

$$\forall (x, y) \in V_g^< \times A_g^{\geq}, \quad \frac{\|g(x)\|}{\|x\|} \leq s \frac{\|g(y)\|}{\|y\|}.$$

On appelle *force de contraction* de g le plus petit nombre $s(g)$ tel que g est $s(g)$ -contractante. En d'autres termes, on a

$$s(g) = \left\| g|_{V_g^<} \right\| \left\| g^{-1}|_{A_g^{\geq}} \right\|.$$

Encore en d'autres termes, $s(g)$ est l'*inverse* du « trou de valeurs singulières » entre $V_g^<$ et A_g^{\geq} (voir la preuve du lemme III.3.8 (iii) pour la définition des valeurs singulières). Dans la convention que nous avons choisie, une application « fortement contractante » a une *petite* valeur de s .

Remarque III.2.19. Notez que pour toute application \mathbb{R} -régulière $g \in G \times \mathfrak{g}$, on a, lorsque $N \rightarrow +\infty$,

$$s(g^N) = \underset{N \rightarrow +\infty}{\mathbf{O}} (\rho^{-N}),$$

où ρ est le trou spectral de g entre $V_g^<$ et A_g^{\geq} (quotient de la plus petite valeur absolue d'une valeur propre de $g|_{A_g^{\geq}}$ — en l'occurrence 1 — par la plus grande valeur absolue d'une valeur propre de $g|_{V_g^<}$). Par définition, $\rho > 0$; on en déduit que

$$s(g^N) \underset{N \rightarrow +\infty}{\rightarrow} 0.$$

III.2.7. Comparaison des propriétés métriques dans le cas affine et linéaire

Pour toute application $f \in G \times \mathfrak{g}$, on note $\ell(f)$ la partie linéaire de f , vue comme un élément de $G \times \mathfrak{g}$ en identifiant G avec le stabilisateur de l'« origine » \mathbb{R}_0 . En d'autres termes, pour tous $(x, t) \in \mathfrak{g} \oplus \mathbb{R}_0 = \hat{\mathfrak{g}}$, on pose

$$\ell(f)(x, t) = f(x, 0) + (0, t).$$

(Voir G comme un sous-groupe de $G \times \mathfrak{g}$ nous permet d'éviter de devoir introduire de nouvelles définitions de la C -non-dégénérescence et de la force de contraction pour les éléments de G .)

Lemme III.2.20. Soit $C \geq 1$, et prenons dans $G \times \mathfrak{g}$ n'importe quelle application g (ou couple d'applications (g, h)) \mathbb{R} -régulière(s) et C -non-dégénérée(s). Supposons que $s(g^{-1}) \leq 1$. Dans ce cas :

- (i) L'application $\ell(g)$ (resp. le couple $(\ell(g), \ell(h))$) est encore C -non-dégénéré(e) ;
- (ii) On a $s(\ell(g)) \leq s(g)$;
- (iii) Plus précisément, on a $s(g) \asymp_C s(\ell(g)) \left\| g|_{A_g^{\geq}} \right\|$.

La démonstration des deux premiers points est une simple vérification formelle, et ne contient aucune surprise.

Démonstration.

- (i) On va démontrer le résultat seulement pour une application g ; pour un couple d'applications le raisonnement est analogue.

Soit ϕ une application canonisante optimale pour g . Alors clairement $A_{\ell(g)}^{\geq} = V_g^{\geq} \oplus \mathbb{R}_0$, et

$$\begin{aligned} \ell(\phi)(A_{\ell(g)}^{\geq}) &= \ell(\phi)(V_g^{\geq} \oplus \mathbb{R}_0) \\ &= \phi(V_g^{\geq}) \oplus \mathbb{R}_0 \\ &= \hat{\mathfrak{p}}^+; \end{aligned}$$

de même, $\ell(\phi)(A_{\ell(g)}^{\leq}) = \hat{\mathfrak{p}}^-$. Donc $\ell(\phi)$ est une application canonisante pour $\ell(g)$. D'autre part, on a

$$\begin{aligned} \|\ell(\phi)\| &= \max\left(\|\phi|_{\mathfrak{g}}\|, 1\right) \\ &\leq \max(\|\phi\|, 1) \\ &\leq \max(C, 1), \end{aligned}$$

et de même pour ϕ^{-1} . Comme $C \geq 1$, on obtient que $\ell(g)$ est C -non-dégénérée.

- (ii) On a :

$$\begin{aligned} s(\ell(g)) &= \left\| \ell(g)|_{V_{\ell(g)}^{\leq}} \right\| \left\| \ell(g)^{-1}|_{A_{\ell(g)}^{\geq}} \right\| \\ &= \left\| g|_{V_g^{\leq}} \right\| \max\left(\left\| g^{-1}|_{V_g^{\geq}} \right\|, 1\right) \\ &\leq \left\| g|_{V_g^{\leq}} \right\| \max\left(\left\| g^{-1}|_{A_g^{\geq}} \right\|, 1\right) \\ &= s(g). \end{aligned}$$

(Pour justifier la dernière égalité, remarquer que $V_g^{\leq} \subset A_g^{\geq}$ contient des points fixes de g et donc $\left\| g^{-1}|_{A_g^{\geq}} \right\| \geq 1$.)

- (iii) On a, par définition :

$$s(g) = \left\| g|_{V_g^{\leq}} \right\| \left\| g^{-1}|_{A_g^{\geq}} \right\|.$$

Soit ϕ une application canonisante optimale pour g . Comme g est C -non-dégénérée (et $\phi(A_g^{\leq})$ est orthogonal à $\phi(V_g^{\geq})$), on obtient que

$$s(g) \asymp_C \left\| g|_{V_g^{\leq}} \right\| \max\left(\left\| g^{-1}|_{A_g^{\leq}} \right\|, \left\| g^{-1}|_{V_g^{\geq}} \right\|\right).$$

Or, comme A_g^{\leq} contient des points fixes de g , clairement on a $\left\| g^{-1}|_{A_g^{\leq}} \right\| \geq 1$. D'autre part, puisque $s(g^{-1}) \leq 1$, on a $\left\| g^{-1}|_{V_g^{\geq}} \right\| \leq 1$. On en déduit que

$$s(g) \asymp_C \left\| g|_{V_g^{\leq}} \right\| \left\| g^{-1}|_{A_g^{\leq}} \right\|.$$

D'après le corollaire III.2.12, le conjugué de $g|_{A_g^{\leq}}$ par ϕ est une quasi-translation. D'après la proposition III.2.8 caractérisant les quasi-translations, on peut écrire

$$\phi g|_{A_g^{\leq}} \phi^{-1} = \tau_v \rho,$$

où ρ est un automorphisme orthogonal du sous-espace \mathfrak{d} , et τ_v est la translation par un certain vecteur $v \in \mathfrak{l}$. Comme ρ préserve la norme euclidienne (il préserve la forme

de Killing, et par convention elles coïncident sur \mathfrak{d}), elle n'a aucune influence sur la norme d'opérateur ; et clairement $\|\tau_v\| = \|\tau_{-v}\|$. On en déduit que $\|\rho^{-1}\tau_v^{-1}\| = \|\tau_{-v}\| = \|\tau_v\| = \|\tau_v\rho\|$, d'où $\left\|g^{-1}|_{A_g^{\bar{=}}}\right\| \asymp_C \left\|g|_{A_g^{\bar{=}}}\right\|$. Ainsi on obtient

$$s(g) \asymp_C \left\|g|_{V_g^<}\right\| \left\|g|_{A_g^{\bar{=}}}\right\|.$$

Une estimation similaire est valable pour $\ell(g)$; mais comme la restriction de $\ell(g)$ à $A_{\ell(g)}^{\bar{=}}$ n'a pas de partie de translation, le deuxième facteur disparaît :

$$s(\ell(g)) \asymp_C \left\|\ell(g)|_{V_g^<}\right\|.$$

Comme g et $\ell(g)$ coïncident sur $V_g^<$, on conclut que

$$s(g) \asymp_C s(\ell(g)) \left\|g|_{A_g^{\bar{=}}}\right\|$$

comme demandé. \square

III.3. \mathbb{R} -régularité des produits

Le but de cette section est de démontrer la proposition III.3.6, qui dit essentiellement d'une façon quantitative que sous certaines conditions, le produit de deux applications \mathbb{R} -régulières est encore \mathbb{R} -régulier. Cette section ressemble très fortement à la section II.3. Les preuves sont toutes très similaires, mais les énoncés sont différents. Un changement important est l'introduction de la notion de C -non-dégénérescence, qui permet entre autres de traiter directement le cas affine. En revanche, contrairement à la section II.3, nous avons choisi dans cette partie de nous intéresser uniquement au produit de deux applications, et de reléguer la récurrence qui permet de généraliser à un produit quelconque à la section III.5.

III.3.1. Cas proximal

Soit E un espace euclidien. (En pratique, on appliquera les résultats de cette sous-section à $E = \Lambda^q \hat{\mathfrak{g}}$ pour un certain entier q .)

Notre premier objectif est de démontrer la proposition III.3.4, qui est analogue à la proposition III.3.6 (et sera utilisée pour la démontrer), mais avec des applications proximales à la place d'applications \mathbb{R} -régulières. On commence par quelques définitions.

Définition III.3.1. Soit $\gamma \in \text{GL}(E)$. Soit λ une valeur propre de γ de valeur absolue maximale. On dira que γ est *proximale* si λ est unique et a multiplicité 1. On peut alors décomposer E en une somme directe d'une droite E_γ^s , appelée son *espace attractif*, et d'un hyperplan E_γ^u , appelé son *espace répulsif*, tous deux stables par γ et tels que :

$$\begin{cases} \gamma|_{E_\gamma^s} = \lambda \text{Id} \\ \text{pour toute valeur propre } \mu \text{ de } \gamma|_{E_\gamma^u}, |\mu| < |\lambda|. \end{cases}$$

Définition III.3.2. Considérons une droite E^s et un hyperplan E^u de E , transverses l'un à l'autre. Une *application canonisante optimale* pour le couple (E^s, E^u) est une application $\phi \in \text{GL}(E)$ vérifiant

$$\phi(E^s) \perp \phi(E^u)$$

et minimisant la quantité $\max(\|\phi\|, \|\phi^{-1}\|)$.

On appelle *application canonisante optimale* pour une application proximale $\gamma \in \text{GL}(E)$ toute application canonisante optimale pour le couple (E_γ^s, E_γ^u) .

Soit $C \geq 1$. On dit qu'un couple d'une droite et d'un hyperplan (E^s, E^u) (resp. une application proximale γ) est C -non-dégénéré(e) s'il ou elle a une application canonisante optimale ϕ telle que $\|\phi^{\pm 1}\| \leq C$.

Soient maintenant γ_1, γ_2 deux applications proximales dans $\text{GL}(E)$. On dit que le couple (γ_1, γ_2) est C -non-dégénéré si chacun des quatre couples possibles $(E_{\gamma_i}^s, E_{\gamma_j}^u)$ est C -non-dégénéré.

Définition III.3.3. Soit $\gamma \in \text{GL}(E)$ une application proximale. On définit la *force de contraction* de γ par

$$\tilde{s}(\gamma) := \frac{\|\gamma|_{E_\gamma^u}\|}{|\lambda|}.$$

Notez que cette définition est différente de celle qu'on avait utilisée dans le contexte des applications \mathbb{R} -régulières (d'où la nouvelle notation \tilde{s}).

Proposition III.3.4. *Pour tout $C \geq 1$, il existe une constante strictement positive $\tilde{s}_1(C)$ ayant la propriété suivante. Prenons un couple C -non-dégénéré d'applications proximales γ_1, γ_2 dans $\text{GL}(E)$, et supposons que γ_1 et γ_2 sont toutes les deux $\tilde{s}_1(C)$ -contractantes. Alors $\gamma_1\gamma_2$ est proximale, et on a :*

$$(i) \quad \alpha(E_{\gamma_1\gamma_2}^s, E_{\gamma_1}^s) \lesssim_C \tilde{s}(\gamma_1);$$

$$(ii) \quad \tilde{s}(\gamma_1\gamma_2) \lesssim_C \tilde{s}(\gamma_1)\tilde{s}(\gamma_2).$$

Avant de poursuivre, on a besoin d'un lemme technique qui fait le lien entre la force de contraction abstraite $\tilde{s}(\gamma)$ et quelques vraies constantes de Lipschitz de γ agissant sur l'espace projectif $\mathbb{P}(E)$. Pour toute partie non vide $X \subset \mathbb{P}(E)$, on introduit la notation suivante pour la constante de Lipschitz de γ restreinte à X :

$$\mathcal{L}(\gamma, X) := \sup_{\substack{(x,y) \in X^2 \\ x \neq y}} \frac{\alpha(\gamma(x), \gamma(y))}{\alpha(x, y)}$$

Lemme III.3.5. *Pour tous $C \geq 1$, $\zeta \in]0, \frac{\pi}{2}[$, pour toute application proximale et C -non-dégénérée γ , on a :*

$$\tilde{s}(\gamma) \lesssim_{C,\zeta} \mathcal{L}(\gamma, B_{\mathbb{P}}(E_\gamma^s, \zeta)). \quad (\text{III.3.1a})$$

$$\mathcal{L}(\gamma, \mathbb{P}(E) \setminus B_{\mathbb{P}}(E_\gamma^u, \zeta)) \lesssim_{C,\zeta} \tilde{s}(\gamma) \quad (\text{III.3.1b})$$

Démonstration. Soient $C \geq 1$, $\zeta \in]0, \frac{\pi}{2}[$. Considérons une application proximale et C -non-dégénérée γ ; soit ϕ une application canonisante optimale pour γ . Alors sans perte de généralité, on peut remplacer γ par $\gamma' := \phi\gamma\phi^{-1}$. En effet l'estimation $\tilde{s}(\gamma) \asymp_C \tilde{s}(\gamma')$ est évidente. Quant à l'autre côté, d'après le lemme III.2.16, on a $\mathcal{L}(\phi, \mathbb{P}(E)) \asymp_C 1$, d'où $\mathcal{L}(\gamma, X) \asymp_C \mathcal{L}(\gamma', \phi(X))$ pour tout ensemble X . On a aussi

$$\begin{aligned} \phi(B_{\mathbb{P}}(E_\gamma^s, \zeta)) &\supset B_{\mathbb{P}}(E_{\gamma'}^s, C^{-2}\zeta), \\ \phi(\mathbb{P}(E) \setminus B_{\mathbb{P}}(E_\gamma^u, \zeta)) &\subset \mathbb{P}(E) \setminus B_{\mathbb{P}}(E_{\gamma'}^u, C^{-2}\zeta); \end{aligned}$$

et clairement $X \subset Y$ implique $\mathcal{L}(\gamma, X) \leq \mathcal{L}(\gamma, Y)$.

Il reste à montrer que pour tout $\zeta' > 0$, on a

$$\mathcal{L}(\gamma', B_{\mathbb{P}}(E_{\gamma'}^s, \zeta')) \asymp_{\zeta'} \tilde{s}(\gamma')$$

(ceci implique (III.3.1a) en prenant $\zeta' = C^{-2}\zeta$, et (III.3.1b) en prenant $\zeta' > \frac{\pi}{2} - C^{-2}\zeta$). En effet, considérons la projection

$$\begin{aligned} \pi_u : \mathbb{P}(E) \setminus \mathbb{P}(E_{\gamma'}^u) &\longrightarrow E_{\gamma'}^u \\ \bar{x} &\longmapsto \frac{x_u}{x_s}, \end{aligned}$$

où x_u et x_s représentent les composantes de x dans la décomposition $E = E_{\gamma'}^u \oplus E_{\gamma'}^s$ (et pour donner un sens à la division par x_s , on choisit une identification isométrique de $E_{\gamma'}^s$ avec \mathbb{R}). Comme $E_{\gamma'}^s$ et $E_{\gamma'}^u$ sont, par construction, orthogonaux, elle induit un homéomorphisme de $B_{\mathbb{P}}(E_{\gamma'}^s, \zeta')$ sur la boule $\left\{ x \in E_{\gamma'}^u \mid \|x\| \leq \frac{1}{\tan \zeta'} \right\}$. Un calcul immédiat montre que cet homéomorphisme est bilipschitzien, avec une constante de Lipschitz $K(\zeta')$ qui ne dépend pas du tout de γ ni de C . D'un autre côté, la constante de Lipschitz de l'application conjuguée $\pi_u \gamma' \pi_u^{-1}$ n'est rien d'autre que $\tilde{s}(\gamma')$. Donc γ' est lipschitzienne avec la constante $K(\zeta')^2 \tilde{s}(\gamma')$, ce qui permet de conclure. \square

Démonstration de la proposition III.3.4. Soit $C \geq 1$, et soit (γ_1, γ_2) un couple C -non-dégénéré d'applications proximales $\tilde{s}_1(C)$ -contractantes (pour une valeur $\tilde{s}_1(C)$ qui sera précisée au cours de la preuve). Alors d'après le lemme III.2.16, pour tous i et j on a $\alpha(E_{\gamma_i}^s, E_{\gamma_j}^u) \geq \eta$ où on pose $\eta := \frac{\pi}{2C^2}$.

Un corollaire immédiat du lemme III.3.5 est que pour toute application proximale C -non-dégénérée γ et pour tout $\zeta \leq \eta$, on a

$$\gamma(\mathbb{P}(E) \setminus B_{\mathbb{P}}(E_{\gamma}^u, \zeta)) \subset B_{\mathbb{P}}(E_{\gamma}^s, K(C, \zeta) \tilde{s}(\gamma)) \quad (\text{III.3.2})$$

pour une certaine constante $K(C, \zeta)$. En effet, $E_{\gamma}^s \in \mathbb{P}(E) \setminus B_{\mathbb{P}}(E_{\gamma}^u, \zeta)$ est un point fixe de γ et $\text{diam}(\mathbb{P}(E) \setminus B_{\mathbb{P}}(E_{\gamma}^u, \zeta)) \leq \frac{\pi}{2} \lesssim 1$.

Pour $i = 1, 2$, on introduit les nombres $\eta_i := K(C, \frac{\eta}{3}) \tilde{s}(\gamma_i)$ et les ensembles

$$\begin{cases} X_i^+ := B_{\mathbb{P}}(E_{\gamma_i}^s, \eta_i) \\ X_i^- := B_{\mathbb{P}}(E_{\gamma_i}^u, \frac{\eta}{3}). \end{cases}$$

Alors d'après (III.3.2), pour tout i on a $\gamma_i(\mathbb{P}(E) \setminus X_i^-) \subset X_i^+$. Comme $\tilde{s}(\gamma_i) \leq \tilde{s}_1(C)$, en choisissant $\tilde{s}_1(C)$ suffisamment petit, on peut supposer que $\eta_i \leq \frac{\eta}{3}$. Alors ces quatre ensembles sont deux à deux disjoints : pour tous i et j , on a $X_i^+ \subset \mathbb{P}(E) \setminus X_j^-$. En particulier, on en déduit que

$$\gamma_1 \gamma_2(\mathbb{P}(E) \setminus X_2^-) \subset X_1^+.$$

Or d'après (III.3.1b), on sait que pour tout i

$$\mathcal{L}(\gamma_i, \mathbb{P}(E) \setminus X_i^-) \lesssim_C \tilde{s}(\gamma_i) \leq \tilde{s}_1(C). \quad (\text{III.3.3})$$

Encore une fois, en choisissant $\tilde{s}_1(C)$ suffisamment petit, on peut en fait supposer que

$$\mathcal{L}(\gamma_i, \mathbb{P}(E) \setminus X_i^-) < 1.$$

Puisque $X_1^+ \subset \mathbb{P}(E) \setminus X_2^-$, on en déduit que X_1^+ est stable par $\gamma_1 \gamma_2$ et que

$$\mathcal{L}(\gamma_1 \gamma_2, X_1^+) < 1.$$

On déduit de tout cela que $\gamma_1 \gamma_2$ est proximale et que $E_{\gamma_1 \gamma_2}^s \in X_1^+$ (voir le lemme 3.8 de [22] pour une démonstration), ce qui montre l'inégalité (i). D'autre part, il est facile de voir que $E_{\gamma_1 \gamma_2}^u \subset X_2^-$ (en effet, considérons un point $x \in \mathbb{P}(E)$ qui appartiendrait à $E_{\gamma_1 \gamma_2}^u$ mais pas à X_2^- : alors on aurait $\lim_{n \rightarrow \infty} (\gamma_1 \gamma_2)^n(x) = E_{\gamma_1 \gamma_2}^s$, ce qui contredit le fait que $E_{\gamma_1 \gamma_2}^u$ est un sous-espace stable). On en déduit que

$$\begin{aligned} \alpha(E_{\gamma_1 \gamma_2}^s, E_{\gamma_1 \gamma_2}^u) &\geq \alpha(E_{\gamma_1}^s, E_{\gamma_2}^u) - \eta_1 - \frac{\eta}{3} \\ &\geq \eta - \frac{\eta}{3} - \frac{\eta}{3} \\ &= \frac{\eta}{3}. \end{aligned}$$

Clairement, cela implique que $\gamma_1\gamma_2$ est C' -non-dégénérée pour une certaine constante C' qui dépend seulement de η , donc seulement de C .

Cela nous permet d'appliquer (III.3.1a) à $\gamma_1\gamma_2$:

$$\tilde{s}(\gamma_1\gamma_2) \lesssim_C \mathcal{L}(\gamma_1\gamma_2, B_{\mathbb{P}}(E_{\gamma_1\gamma_2}^s, \frac{\eta}{3})).$$

On sait que $B_{\mathbb{P}}(E_{\gamma_1\gamma_2}^s, \frac{\eta}{3}) \subset B_{\mathbb{P}}(E_{\gamma_1}^s, \frac{2\eta}{3}) \subset \mathbb{P}(E) \setminus X_2^-$, d'où

$$\mathcal{L}(\gamma_1\gamma_2, B_{\mathbb{P}}(E_{\gamma_1\gamma_2}^s, \frac{\eta}{3})) \leq \mathcal{L}(\gamma_1\gamma_2, \mathbb{P}(E) \setminus X_2^-).$$

D'autre part, de (III.3.3), on déduit que

$$\mathcal{L}(\gamma_1\gamma_2, \mathbb{P}(E) \setminus X_2^-) \lesssim_C \tilde{s}(\gamma_1)\tilde{s}(\gamma_2).$$

En rassemblant ces inégalités, on obtient

$$\tilde{s}(\gamma_1\gamma_2) \lesssim_C \tilde{s}(\gamma_1)\tilde{s}(\gamma_2);$$

ainsi nous avons également démontré (ii). □

III.3.2. Cas \mathbb{R} -régulier

La proposition suivante estime la position des espaces dynamiques et la force de contraction pour un produit de deux applications \mathbb{R} -régulières suffisamment contractantes qui forment un couple non dégénéré.

Proposition III.3.6. *Pour tout $C \geq 1$, il existe une constante strictement positive $s_1(C) \leq 1$ ayant la propriété suivante. Prenons n'importe quel couple C -non-dégénéré (g, h) d'applications \mathbb{R} -régulières dans $G \times \mathfrak{g}$; supposons que les applications $g^{\pm 1}$ et $h^{\pm 1}$ sont toutes $s_1(C)$ -contractantes. Alors gh est \mathbb{R} -régulière, $2C$ -non-dégénérée, et on a :*

$$(i) \quad \begin{cases} \alpha^{\text{Haus}}(A_{gh}^{\geq}, A_g^{\geq}) \lesssim_C s(g) \\ \alpha^{\text{Haus}}(A_{gh}^{\leq}, A_h^{\leq}) \lesssim_C s(h^{-1}) \end{cases};$$

$$(ii) \quad s(gh) \lesssim_C s(g)s(h).$$

Notez que si g est C -non-dégénérée avec $s(g) \leq 1$, on peut montrer qu'en fait on a $s(g^{-1}) \asymp_C s(g)$ (ce qui simplifierait la deuxième inégalité). Mais on n'a pas besoin de ce résultat.

Avant de donner la démonstration, commençons par formuler un cas particulier :

Corollaire III.3.7. *Sous les mêmes hypothèses, on a*

$$\begin{cases} \alpha^{\text{Haus}}(V_{gh}^{\geq}, V_g^{\geq}) \lesssim_C s(\ell(g)) \\ \alpha^{\text{Haus}}(V_{gh}^{\leq}, V_h^{\leq}) \lesssim_C s(\ell(h)^{-1}). \end{cases}$$

Démonstration. Si un couple (g, h) vérifie les hypothèses de la proposition III.3.6, alors le lemme III.2.20 montre que le couple $(\ell(g), \ell(h))$ les vérifie encore. Or pour toute application \mathbb{R} -régulière f , puisque $\ell(f)$ et f ont la même action sur \mathfrak{g} , il est évident que $V_{\ell(f)}^{\geq} = V_f^{\geq}$ et $V_{\ell(f)}^{\leq} = V_f^{\leq}$. □

Pour démontrer la proposition III.3.6, on utilise le résultat de la sous-section précédente, en établissant une correspondance entre la \mathbb{R} -régularité et la proximalité dans une puissance extérieure convenable.

On introduit les entiers :

$$\begin{aligned} q &:= \dim \hat{\mathfrak{p}}^+; \\ p &:= \dim \mathfrak{n}^-; \\ d &:= \dim \hat{\mathfrak{g}} = p + q. \end{aligned}$$

Pour tout $g \in G \ltimes \mathfrak{g}$, on peut définir sa puissance extérieure $\Lambda^q g : \Lambda^q \hat{\mathfrak{g}} \rightarrow \Lambda^q \hat{\mathfrak{g}}$. La structure euclidienne de $\hat{\mathfrak{g}}$ induit de façon canonique une structure euclidienne sur $\Lambda^q \hat{\mathfrak{g}}$.

Lemme III.3.8.

(i) Pour $g \in G \ltimes \mathfrak{g}$, $\Lambda^q g$ est proximale si et seulement si g est \mathbb{R} -régulière. De plus, l'espace attractif (resp. répulsif) de $\Lambda^q g$ dépend uniquement de A_g^\geq (resp. de $V_g^<$) :

$$\begin{cases} E_{\Lambda^q g}^s = \Lambda^q A_g^\geq \\ E_{\Lambda^q g}^u = \{x \in \Lambda^q \hat{\mathfrak{g}} \mid x \wedge \Lambda^p V_g^< = 0\}. \end{cases} \quad (\text{III.3.4})$$

(ii) Pour tout $C \geq 1$, dès que (g_1, g_2) est un couple C -non-dégénéré d'applications \mathbb{R} -régulières, $(\Lambda^q g_1, \Lambda^q g_2)$ est un couple C^q -non-dégénéré d'applications proximales.

(iii) Pour tout $C \geq 1$, pour toute application \mathbb{R} -régulière C -non-dégénérée $g \in G \ltimes \mathfrak{g}$, on a

$$s(g) \lesssim_C \tilde{s}(\Lambda^q g).$$

Si de plus $s(g) \leq 1$, on a

$$s(g) \succsim_C \tilde{s}(\Lambda^q g).$$

(iv) Pour tout couple de sous-espaces A_1 et A_2 de dimension q dans $\hat{\mathfrak{g}}$, on a

$$\alpha^{\text{Haus}}(A_1, A_2) \asymp \alpha(\Lambda^q A_1, \Lambda^q A_2).$$

Démonstration.

(i) Soit $g \in G \ltimes \mathfrak{g}$. Soient $\lambda_1, \dots, \lambda_d$ les valeurs propres de g (agissant sur $\hat{\mathfrak{g}}$) comptées avec multiplicité et ordonnées par valeur absolue croissante. Alors on sait que les valeurs propres de $\Lambda^q g$ comptées avec multiplicité sont exactement les produits de la forme $\lambda_{i_1} \dots \lambda_{i_q}$, où $1 \leq i_1 < \dots < i_q \leq d$. Comme les deux plus grandes d'entre elles (en valeur absolue) sont $\lambda_{p+1} \dots \lambda_d$ et $\lambda_p \lambda_{p+2} \dots \lambda_d$, on en déduit que $\Lambda^q g$ est proximale si et seulement si $|\lambda_p| < |\lambda_{p+1}|$.

D'autre part, d'après l'affirmation III.2.2 (i), on sait que $\dim A_g^- \geq \dim \hat{\mathfrak{l}} = d - 2p$, avec égalité si et seulement si g est \mathbb{R} -régulière. Or, comme la partie linéaire de g préserve la forme de Killing, l'espace $V_g^<$ est Killing-orthogonal à V_g^\leq , qui est supplémentaire (dans \mathfrak{g}) à $V_g^>$; donc $\dim V_g^< \leq \dim V_g^>$. Par symétrie, on obtient $\dim V_g^< = \dim V_g^>$. On en déduit que $\dim V_g^< = \frac{1}{2}(d - \dim A_g^-) \leq p$, avec égalité si et seulement si g est \mathbb{R} -régulière.

En particulier, on a toujours $|\lambda_{p+1}| = 1$. En rassemblant le tout, on conclut que

$$\Lambda^q g \text{ est proximal} \iff |\lambda_p| < 1 \iff \dim V_g^< = p \iff g \text{ est } \mathbb{R}\text{-régulière.}$$

Quant à l'expression de E^s et E^u , elle se déduit immédiatement en considérant une base qui trigonalise g .

(ii) Prenons n'importe quel couple $(i, j) \in \{1, 2\}^2$. Soit ϕ une application canonisante optimale pour le couple $(A_{g_i}^\geq, A_{g_j}^\leq)$. Alors on a $\phi(A_{g_i}^\geq) = \hat{\mathfrak{p}}^+$ et (d'après l'affirmation III.2.2 (iii)) $\phi(V_{g_j}^\leq) = \mathfrak{n}^-$. Avec la structure euclidienne que nous avons choisie, $\hat{\mathfrak{p}}^+$ est orthogonal à \mathfrak{n}^- ; donc $\Lambda^q \hat{\mathfrak{p}}^+$ est orthogonal à $\{x \in \Lambda^q \hat{\mathfrak{g}} \mid x \wedge \Lambda^p \mathfrak{n}^- = 0\}$. Du point précédent, on déduit que $\Lambda^q \phi$ est une application canonisante pour le couple $(E_{\Lambda^q g_i}^s, E_{\Lambda^q g_j}^u)$. Comme $\|\Lambda^q \phi\| \leq \|\phi\|^q$ et de même pour ϕ^{-1} , on a la conclusion.

(iii) Soit $C \geq 1$, et soit $g \in G \rtimes \mathfrak{g}$ une application \mathbb{R} -régulière C -non-dégénérée. Remarquons d'abord la chose suivante : soit ϕ une application canonisante optimale pour g , et soit $g' = \phi g \phi^{-1}$. Alors il est clair que $s(g') \asymp_C s(g)$ et que $\tilde{s}(\Lambda^q g') \asymp_C \tilde{s}(\Lambda^q g)$. On peut donc supposer que $V_g^<, A_g^=$ et $V_g^>$ sont deux à deux orthogonaux.

On appelle *valeurs singulières* de g les racines carrées des valeurs propres de l'application g^*g (où g^* est l'application adjointe, par rapport à la norme euclidienne). Soient $s_1 \leq \dots \leq s_q$ (resp. $s'_1 \leq \dots \leq s'_p$) les valeurs singulières de g restreinte à A_g^{\geq} (resp. $V_g^<$), de sorte que $\|g^{-1}|_{A_g^{\geq}}\| = s_1^{-1}$ et $\|g|_{V_g^<}\| = s'_p$. Puisque les espaces A_g^{\geq} et $V_g^<$ sont stables par g et orthogonaux, on obtient que les valeurs singulières de g dans $\hat{\mathfrak{g}}$ tout entier sont

$$s'_1, \dots, s'_p, s_1, \dots, s_q$$

(notez cependant que si on ne suppose pas $s(g) \leq 1$, cette liste pourrait ne pas être rangée par ordre croissant.) D'autre part, on sait que les valeurs singulières de $\Lambda^q g$ sont les produits de q valeurs singulières distinctes de g . Puisque $E_{\Lambda^q g}^s$ est orthogonal à $E_{\Lambda^q g}^u$, on peut encore une fois analyser les valeurs singulières séparément dans chaque sous-espace. On sait que la valeur singulière correspondant à E^s est égale à $s_1 \dots s_q$; on en déduit que $\|\Lambda^q g|_{E^u}\|$ est égal à la plus grande des valeurs singulières restantes. En particulier c'est supérieur ou égal à $s'_p \cdot s_2 \dots s_q$. D'autre part, si λ est la plus grande valeur propre de $\Lambda^q g$, on a

$$|\lambda| = |\lambda_{p+1} \dots \lambda_d| = \left| \det(g|_{A_g^{\geq}}) \right| = s_1 \dots s_q$$

(où $\lambda_1, \dots, \lambda_d$ sont les valeurs propres de g rangées par valeur absolue croissante). On en déduit que :

$$\tilde{s}(\Lambda^q g) = \frac{\|\Lambda^q g|_{E_{\Lambda^q g}^u}\|}{|\lambda|} \geq \frac{s'_p \cdot s_2 \dots s_q}{s_1 \dots s_q} = s'_p s_1^{-1} = s(g), \quad (\text{III.3.5})$$

ce qui est la première estimation qu'on cherchait.

Supposons maintenant que $s(g) \leq 1$. Alors on a $s'_p \leq s_1$, ce qui signifie que les valeurs singulières de $\Lambda^q g$ sont effectivement rangées dans le « bon » ordre. Donc $s'_p \cdot s_2 \dots s_q$ est en fait la plus grande valeur singulière de $\Lambda^q g|_{E^u}$, et l'inégalité devient une égalité : $\tilde{s}(\Lambda^q g) = s(g)$. La deuxième estimation en découle.

(iv) Soient A_1 et A_2 deux sous-espaces de dimension q dans $\hat{\mathfrak{g}}$. On pose

$$\begin{aligned} \alpha_1 &:= \alpha^{\text{Haus}}(A_1, A_2); \\ \alpha_2 &:= \alpha(\Lambda^q A_1, \Lambda^q A_2). \end{aligned}$$

On peut trouver une base orthonormée (e_1, \dots, e_d) de $\hat{\mathfrak{g}}$ telle que A_1 a pour base (e_1, \dots, e_q) et A_2 a pour base

$$((\cos \theta_i) e_i + (\sin \theta_i) e_{q+i})_{1 \leq i \leq q},$$

pour une certaine famille d'angles $\frac{\pi}{2} \geq \theta_1 \geq \dots \geq \theta_p \geq \theta_{p+1} = \dots = \theta_q = 0$ (bien entendu e_j n'est pas défini pour $j > d$, mais dans cette formule tous ces vecteurs ont un coefficient nul). Dans ce cas, on a $\alpha_1 = \theta_1$ et $\cos \alpha_2 = \prod_{i=1}^q \cos \theta_i$, d'où

$$(\cos \alpha_1)^q \leq \cos \alpha_2 \leq \cos \alpha_1.$$

D'autre part, de la concavité de la fonction $t \mapsto (\arccos \exp t)^2$, on déduit que pour tout $\theta \in [0, \frac{\pi}{2}]$, on a $\arccos((\cos \theta)^q) \leq \sqrt{q} \theta$. Finalement on obtient

$$\alpha_1 \leq \alpha_2 \leq \sqrt{q} \alpha_1,$$

CQFD. □

On aura également besoin du lemme technique suivant :

Lemme III.3.9. *Il existe une constante $\varepsilon > 0$ ayant la propriété suivante. Soient $\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2$ deux a.p.m. affines telles que*

$$\begin{cases} \alpha^{\text{Haus}}(\hat{\mathbf{p}}_1, \hat{\mathbf{p}}^+) \leq \varepsilon \\ \alpha^{\text{Haus}}(\hat{\mathbf{p}}_2, \hat{\mathbf{p}}^-) \leq \varepsilon. \end{cases}$$

Alors elles forment un couple 2-non-dégénéré.

(Évidemment la constante 2 est arbitraire ; on pourrait la remplacer par n'importe quel nombre strictement supérieur à 1.)

Démonstration. Soit \mathcal{P} l'ensemble de tous les couples d'a.p.m. affines, $\mathcal{P}' \subset \mathcal{P}$ le sous-ensemble des couples transverses. Comme \mathcal{P}' est une partie ouverte de \mathcal{P} , pour ε assez petit $\hat{\mathbf{p}}_1$ et $\hat{\mathbf{p}}_2$ seront transverses. De plus, \mathcal{P}' est un espace homogène sous l'action de $G \times \mathfrak{g}$ (d'après l'affirmation III.2.5), donc l'application orbitale qui à un élément $\phi \in G \times \mathfrak{g}$ associe le couple $\phi(\hat{\mathbf{p}}^+, \hat{\mathbf{p}}^-)$ est ouverte. On en déduit que pour tout C , l'ensemble des couples « strictement C -non-dégénérés » (c'est-à-dire C' -non-dégénérés pour un certain $C' < C$) est ouvert. \square

Démonstration de la proposition III.3.6. Soit $C \geq 1$, et soit (g, h) un couple C -non-dégénéré d'applications \mathbb{R} -régulières dans $G \times \mathfrak{g}$. Supposons que $g^{\pm 1}$ et $h^{\pm 1}$ sont $s_1(C)$ -contractantes, pour une certaine constante $s_1(C)$ à préciser au cours de la preuve.

On prend $\gamma_1 = \Lambda^q g$ et $\gamma_2 = \Lambda^q h$. Vérifions les hypothèses de la proposition III.3.4. En effet :

- D'après le lemme III.3.8 (i), γ_1 et γ_2 sont proximales.
- D'après le lemme III.3.8 (ii), le couple (γ_1, γ_2) est C^q -non-dégénéré.
- Comme on a supposé que $s_1(C) \leq 1$, on déduit du lemme III.3.8 (iii) que $\tilde{s}(\gamma_1) \lesssim_C s(g)$ et que $\tilde{s}(\gamma_2) \lesssim_C s(h)$. Si on choisit $s_1(C)$ suffisamment petit, alors γ_1 et γ_2 sont $\tilde{s}_1(C^q)$ -contractantes.

On applique maintenant la proposition III.3.4 à l'application $\Lambda^q(gh) = \gamma_1\gamma_2$. Il reste à déduire les conclusions de la proposition III.3.6.

- Que gh est \mathbb{R} -régulière découle du lemme III.3.8 (i).
- De la proposition III.3.4 (i), en utilisant le lemme III.3.8 (i), (iii) et (iv), on obtient

$$\alpha^{\text{Haus}} \left(A_{gh}^{\geq}, A_g^{\geq} \right) \lesssim_C s(g),$$

ce qui démontre la première ligne de la proposition III.3.6 (i).

- En appliquant la proposition III.3.4 à $\gamma_2^{-1}\gamma_1^{-1}$ à la place de $\gamma_1\gamma_2$, on obtient de même la deuxième ligne de la proposition III.3.6 (i).
- Soit ϕ une application canonisante optimale pour le couple (A_g^{\geq}, A_h^{\leq}) . Par hypothèse, $\|\phi^{\pm 1}\| \leq C$. Mais si on prend $s_1(C)$ suffisamment petit, les deux inégalités qu'on vient de démontrer, combinées au lemme III.3.9, nous permettent de trouver une application ϕ' avec $\|\phi'\| \leq 2$, $\|\phi'^{-1}\| \leq 2$ et

$$\phi' \circ \phi(A_{gh}^{\geq}, A_{gh}^{\leq}) = (\hat{\mathbf{p}}^+, \hat{\mathbf{p}}^-).$$

On en déduit que l'application composée gh est $2C$ -non-dégénérée.

- La dernière inégalité, à savoir la proposition III.3.6 (ii), découle maintenant de la proposition III.3.4 (ii) en utilisant le lemme III.3.8 (iii). \square

III.4. Additivité de l'invariant de Margulis

La proposition III.4.1 est l'ingrédient clé de toute la partie III. Elle explique l'effet des opérations de groupe (l'inverse et la composition) sur l'invariant de Margulis. Le premier point est trivial à démontrer, mais néanmoins important pour la suite. La démonstration du deuxième point occupe la totalité de cette section.

On appelle w_0 n'importe quel élément de G tel que $w_0(\mathfrak{p}^+, \mathfrak{p}^-) = (\mathfrak{p}^-, \mathfrak{p}^+)$. (Par l'affirmation III.2.6, le résultat ci-dessous ne dépend pas du choix de w_0 .)

Proposition III.4.1.

(i) Pour toute application \mathbb{R} -régulière $g \in G \times \mathfrak{g}$, on a

$$M(g^{-1}) = -w_0(M(g)).$$

(ii) Pour tout $C \geq 1$, il existe des constantes positives $s_2(C) \leq 1$ et $\mu(C)$ ayant la propriété suivante. Soient $g, h \in G \times \mathfrak{g}$ des applications \mathbb{R} -régulières formant un couple C -non-dégénéré, avec $g^{\pm 1}$ et $h^{\pm 1}$ toutes $s_2(C)$ -contractantes. Alors gh est \mathbb{R} -régulière, et on a :

$$\|M(gh) - M(g) - M(h)\| \leq \mu(C).$$

Soit $C \geq 1$. On choisit une constante $s_2(C) \leq 1$, suffisamment petite pour satisfaire toutes les contraintes qui apparaîtront au cours de la preuve. Pour le reste de cette section, on fixe $g, h \in G \times \mathfrak{g}$ un couple C -non-dégénéré d'applications \mathbb{R} -régulières telles que $g^{\pm 1}$ et $h^{\pm 1}$ sont $s_2(C)$ -contractantes.

La remarque suivante sera utilisée tout au long de la section.

Remarque III.4.2. On peut supposer que les couples $(A_{gh}^{\geq}, A_g^{\leq})$, $(A_{hg}^{\geq}, A_h^{\leq})$, $(A_g^{\geq}, A_{gh}^{\leq})$ et $(A_{hg}^{\geq}, A_h^{\leq})$ sont tous $2C$ -non-dégénérés. En effet, on rappelle que (d'après la proposition III.3.6), on a

$$\begin{cases} \alpha^{\text{Haus}} \left(A_{gh}^{\geq}, A_g^{\leq} \right) \lesssim_C s(g) \\ \alpha^{\text{Haus}} \left(A_{hg}^{\geq}, A_h^{\leq} \right) \lesssim_C s(h^{-1}) \end{cases}$$

et des majorations similaires avec g et h interchangeés. D'autre part, par hypothèse, le couple (A_g^{\geq}, A_h^{\leq}) est C -non-dégénéré. Si on choisit $s_2(C)$ suffisamment petit, ces quatre propriétés découlent alors du lemme III.3.9.

Démonstration de la proposition III.4.1.

(i) On sait que w_0 commute avec $\pi_{\mathfrak{z}}$. En effet, w_0 stabilise $\mathfrak{l} = \mathfrak{p}^+ \cap \mathfrak{p}^-$; en tant qu'élément de G (identifié à $\text{Ad } G$), c'est un morphisme d'algèbres de Lie; il stabilise donc \mathfrak{z} et \mathfrak{d} , respectivement le centre et l'algèbre dérivée de \mathfrak{l} . Compte tenu du fait que $V_{g^{-1}}^{\geq} = V_g^{\leq}$ et vice-versa, l'identité à démontrer découle immédiatement de la définition de l'invariant de Margulis.

(ii) Si on prend $s_2(C) \leq s_1(C)$, alors la proposition III.3.6 garantit que gh est \mathbb{R} -régulière. Pour estimer $M(gh)$, on décompose $gh : A_{gh}^{\geq} \rightarrow A_{gh}^{\leq}$ en un produit de plusieurs applications plus simples.

– On commence par séparer les deux facteurs du produit gh . On a le diagramme commutatif

$$\begin{array}{ccccc} & & gh & & \\ & & \curvearrowright & & \\ A_{gh}^{\geq} & \longleftarrow & A_{hg}^{\geq} & \longleftarrow & A_{gh}^{\leq} \\ & \longleftarrow g & & \longleftarrow h & \end{array}$$

En effet, comme hg est le conjugué de gh par h et vice-versa, on a $h(A_{gh}^{\geq}) = A_{hg}^{\geq}$ et $g(A_{hg}^{\geq}) = A_{gh}^{\geq}$.

- Ensuite on factorise l'application $g : A_{hg}^- \rightarrow A_{gh}^-$ à travers l'application $g : A_g^- \rightarrow A_g^-$, qu'on connaît mieux. On a le diagramme commutatif

$$\begin{array}{ccc}
 A_{gh}^- & \xleftarrow{g} & A_{hg}^- \\
 \searrow \pi_g & & \swarrow \pi_g \\
 A_g^- & \xleftarrow{g} & A_g^-
 \end{array}$$

où π_g est la projection sur A_g^- parallèlement à $V_g^> \oplus V_g^<$. (Elle commute avec g car ces trois espaces sont tous invariants par g .)

- Finalement, on décompose encore chaque flèche diagonale du dernier diagramme en deux facteurs. On pose

$$\begin{cases}
 A_{g,gh}^- := A_g^{\geq} \cap A_{gh}^{\leq} \\
 A_{hg,g}^- := A_{hg}^{\geq} \cap A_g^{\leq},
 \end{cases}$$

et P_1 (resp. P_2) la projection sur $A_{g,gh}^-$ (resp. $A_{hg,g}^-$) parallèlement à $V_g^> \oplus V_g^<$. Pour justifier cette définition, on doit vérifier que $A_{g,gh}^-$ (et de même $A_{hg,g}^-$) est supplémentaire à $V_g^> \oplus V_g^<$. En effet, d'après la remarque III.4.2, A_{gh}^{\leq} est transverse à A_g^{\geq} , donc (d'après l'affirmation III.2.2 (iii)) supplémentaire à $V_g^>$; par conséquent $A_g^{\geq} = V_g^> \oplus A_{g,gh}^-$ et $\hat{\mathfrak{g}} = V_g^< \oplus A_g^{\geq} = V_g^< \oplus V_g^> \oplus A_{g,gh}^-$. On a alors les diagrammes commutatifs

$$\begin{array}{ccccc}
 A_{gh}^- & \xrightarrow{P_1} & A_{g,gh}^- & \xrightarrow{\pi_g} & A_g^- \\
 & \searrow & & \searrow & \\
 & & & & A_g^-
 \end{array}$$

et

$$\begin{array}{ccccc}
 A_{hg}^- & \xrightarrow{P_2} & A_{hg,g}^- & \xrightarrow{\pi_g} & A_g^- \\
 & \searrow & & \searrow & \\
 & & & & A_g^-
 \end{array}$$

Les deuxième et troisième étapes peuvent être répétées avec h à la place de g . La façon d'adapter la deuxième étape est évidente; pour la troisième étape, on factorise $\pi_h : A_{hg}^- \rightarrow A_h^-$ à travers $A_{h,hg}^-$ et $\pi_h : A_{gh}^- \rightarrow A_h^-$ à travers $A_{gh,h}^-$ (définis de même). En combinant ces trois décompositions, on obtient la moitié inférieure du diagramme III.1. (Nous avons omis le développement de h ; nous laissons la construction du diagramme complet aux lecteurs particulièrement motivés.) On va maintenant interpréter toutes ces applications comme des endomorphismes de $\hat{\mathfrak{l}}$. Pour le faire, on choisit des applications canonisantes optimales

$$\phi_g, \phi_{gh}, \phi_{hg}, \phi_{g,gh}, \phi_{hg,g}$$

respectivement de g , de gh , de hg , du couple $(A_g^{\geq}, A_{gh}^{\leq})$ et du couple $(A_{hg}^{\geq}, A_g^{\leq})$. Ceci nous permet de définir \overline{ggh} , \overline{hgh} , $\overline{g_{g,gh}}$, $\overline{g_-}$, $\overline{P_1}$, $\overline{P_2}$, $\overline{\psi_1}$, $\overline{\psi_2}$ comme les applications qui rendent commutatif la totalité du diagramme III.1.

Maintenant on définit

$$\begin{cases}
 M_{gh}(g) := \pi_3(\overline{ggh}(x) - x) \\
 M_{gh}(h) := \pi_3(\overline{hgh}(x) - x)
 \end{cases}$$

pour tout $x \in \mathfrak{l}_{\text{Aff}}$, où $\mathfrak{l}_{\text{Aff}} := \hat{\mathfrak{l}} \cap \mathfrak{g}_{\text{Aff}}$ est l'espace affine parallèle à \mathfrak{l} et passant par l'origine. Justifions cette définition : comme gh est le conjugué de hg par g et

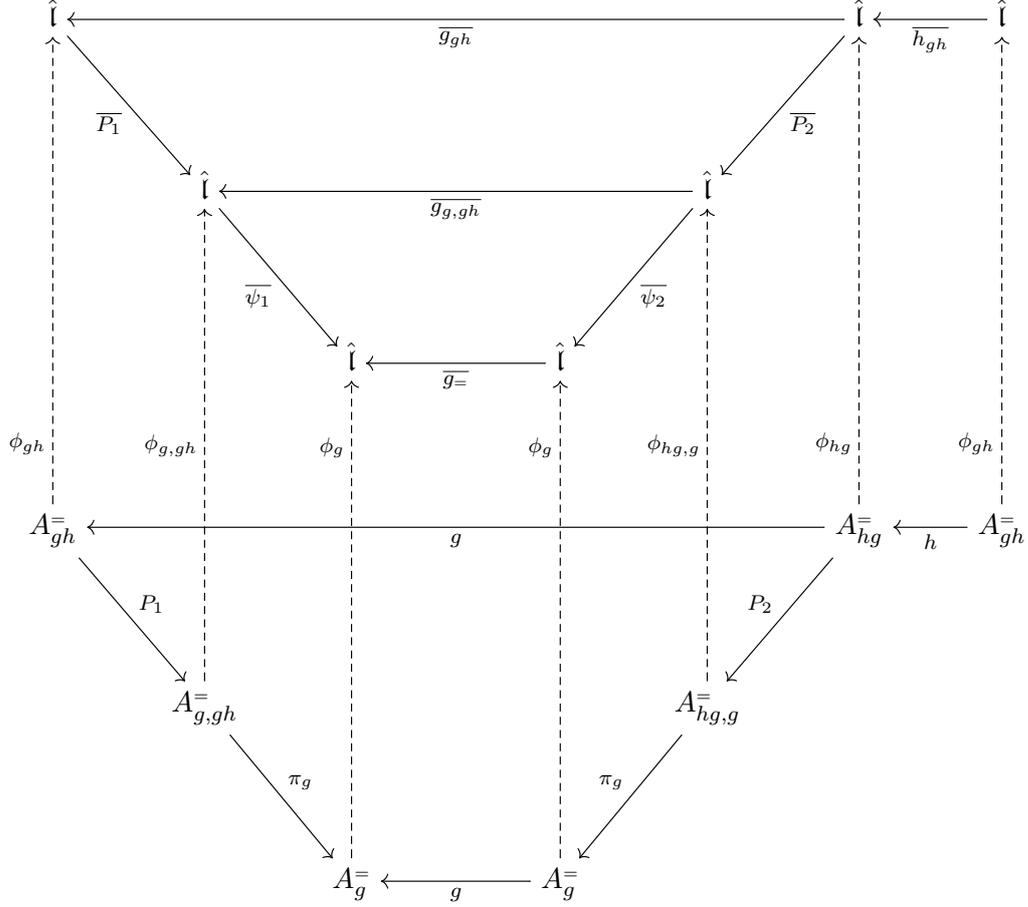


Diagramme III.1

vice-versa, on a :

$$\begin{aligned}
 h(A_{gh}^{\geq}) &= A_{hg}^{\geq} & g(A_{hg}^{\geq}) &= A_{gh}^{\geq} \\
 h(A_{gh}^{\leq}) &= A_{hg}^{\leq} & g(A_{hg}^{\leq}) &= A_{gh}^{\leq},
 \end{aligned}$$

et donc les applications $\overline{g_{gh}}$ et $\overline{h_{gh}}$ stabilisent les espaces $\hat{\mathfrak{p}}^+$ et $\hat{\mathfrak{p}}^-$. Par l'affirmation III.2.6, elles sont donc des quasi-translations ; on en déduit que les valeurs définies ci-dessus ne dépendent pas du choix de x .

Cette définition est à comparer avec celle de l'invariant de Margulis (la définition III.2.13) : on a $M(gh) = \pi_3(\overline{g_{gh}} \circ \overline{h_{gh}}(x) - x)$ pour tout $x \in \mathfrak{L}_{\text{Aff}}$. Soit un tel x , et posons $x' := \overline{h_{gh}}(x)$. On a alors

$$\begin{aligned}
 M(gh) &= \pi_3(\overline{g_{gh}} \circ \overline{h_{gh}}(x) - x) \\
 &= \pi_3(\overline{g_{gh}}(x') - x') + \pi_3(\overline{h_{gh}}(x) - x) \\
 &= M_{gh}(g) + M_{gh}(h).
 \end{aligned}$$

En effet, comme $\overline{h_{gh}}$ est une quasi-translation, x' est encore un élément de $\mathfrak{L}_{\text{Aff}}$.

Il suffit donc de montrer que $\|M_{gh}(g) - M(g)\| \lesssim_C 1$ et $\|M_{gh}(h) - M(h)\| \lesssim_C 1$. C'est une conséquence immédiate du lemme III.4.4 ci-dessous. (Notez que, alors que les vecteurs $M_{gh}(g)$ et $M_{gh}(h)$ sont des éléments de \mathfrak{z} , les applications $\overline{g_{gh}}$ et $\overline{h_{gh}}$ sont des isométries agissant sur le sous-espace $\hat{\mathfrak{I}}$ tout entier.) \square

Remarque III.4.3. Contrairement aux vrais invariants de Margulis, les valeurs $M_{gh}(g)$ et $M_{gh}(h)$ *dépendent* du choix des applications canonisantes. Un choix différent d'applications canonisantes nous obligerait à soustraire une constante de la première et à l'ajouter à la seconde.

On dira qu'une bijection linéaire f entre deux sous-espaces de $\hat{\mathfrak{g}}$ est $O(C)$ -bornée si elle est bornée par une constante qui ne dépend que de C , c'est-à-dire que $\|f\| \lesssim_C 1$ et $\|f^{-1}\| \lesssim_C 1$. On dira que deux autormorphismes f_1, f_2 de $\hat{\mathfrak{I}}$ (qui dépendent d'une certaine manière de g et de h) sont $O(C)$ -presque-équivalents, et on écrira $f_1 \approx_C f_2$, s'ils vérifient la condition

$$\|f_1 - \xi \circ f_2 \circ \xi'\| \lesssim_C 1$$

pour certaines quasi-translations $O(C)$ -bornées ξ, ξ' . C'est effectivement une relation d'équivalence.

Lemme III.4.4. *Les applications $\overline{g_{gh}}$ et $\overline{h_{gh}}$ sont $O(C)$ -presque-équivalentes à $\overline{g_-}$ et $\overline{h_-}$ respectivement.*

Pour démontrer ceci, on utilise la propriété suivante :

Lemme III.4.5. *Toutes les flèches non horizontales du diagramme III.1 représentent des applications bijectives $O(C)$ -bornées.*

Démonstration.

- Pour les flèches verticales, c'est une conséquence immédiate de la remarque III.4.2.
- Montrons que $P_1 : A_{gh}^- \rightarrow A_{g,gh}^-$, la projection parallèlement à $V_g^> \oplus V_g^<$, est bijective et $O(C)$ -bornée. Il suffit de démontrer que les distances $\alpha(A_{gh}^-, V_g^> \oplus V_g^<)$ et $\alpha(A_{g,gh}^-, V_g^> \oplus V_g^<)$ sont minorées par une constante strictement positive qui ne dépend que de C .
- Estimons la deuxième quantité. On sait que $\phi_{g,gh}$ envoie, respectivement, A_{gh}^{\leq} et $V_g^>$ sur $\hat{\mathfrak{p}}^-$ et \mathfrak{n}^+ (en utilisant l'affirmation III.2.2 (iii) à propos de l'unicité de $V^>$). Par convention, les deux derniers espaces mentionnés sont orthogonaux ; comme le couple $(A_g^{\geq}, A_{gh}^{\leq})$ est $2C$ -non-dégénéré (remarque III.4.2), on en déduit par le lemme III.2.16 à propos de la constante de Lipschitz des applications bornées que

$$\alpha(A_{gh}^{\leq}, V_g^>) \gtrsim_C 1$$

(en fait la partie de gauche est précisément minorée par $\frac{1}{4C^2} \frac{\pi}{2}$), et en particulier

$$\alpha(A_{g,gh}^-, V_g^>) \gtrsim_C 1.$$

Ensuite on applique ϕ_g ; puisque $\|\phi_g\| \leq C$ et $\|\phi_g^{-1}\| \leq C$, la distance est, encore une fois, divisée par au plus C^2 . Mais après cette transformation, l'espace $\phi_g(A_{g,gh}^-) = \hat{\mathfrak{p}}^+$, qui contient à la fois $\phi_g(A_{g,gh}^-)$ et $\phi_g(V_g^>)$, est orthogonal à $\phi_g(V_g^<) = \mathfrak{n}^+$; on a donc

$$\alpha(\phi_g(A_{g,gh}^-), \phi_g(V_g^>)) = \alpha(\phi_g(A_{g,gh}^-), \phi_g(V_g^> \oplus V_g^<)).$$

En appliquant ϕ_g^{-1} pour revenir aux espaces d'origine, on introduit encore une fois un facteur qui n'est pas plus petit que $\frac{1}{C^2}$. On conclut que

$$\alpha(A_{g,gh}^-, V_g^> \oplus V_g^<) \gtrsim_C 1$$

comme demandé.

- On va maintenant estimer la première quantité. Pour montrer que l'espace A_{gh}^- est « loin » de $V_g^> \oplus V_g^<$, on va montrer qu'il est « proche » de l'espace $A_{g,gh}^-$, dont on vient de montrer qu'il est « loin » de $V_g^> \oplus V_g^<$.

On utilisera la propriété suivante : si les sous-espaces vectoriels F et G sont perpendiculaires (c'est-à-dire que les supplémentaires orthogonaux de $F \cap G$ respectivement

dans F et dans G sont orthogonaux l'un à l'autre), alors pour tout sous-espace F' , on a

$$\alpha^{\text{Haus}}(F \cap G, F' \cap G) \leq \alpha^{\text{Haus}}(F, F'),$$

pourvu que $F' \cap G$ ait la même dimension que $F \cap G$.

En prenant pour F, G et F' les images par l'application $O(C)$ -bornée ϕ_{gh} des espaces $A_{gh}^{\geq}, A_{gh}^{\leq}$ et A_g^{\geq} respectivement, on déduit que

$$\alpha^{\text{Haus}}(A_{gh}^{\leq}, A_{g,gh}^{\leq}) \lesssim_C \alpha^{\text{Haus}}(A_{gh}^{\geq}, A_g^{\geq}).$$

D'autre part, la proposition III.3.6 nous dit que

$$\alpha^{\text{Haus}}(A_{gh}^{\geq}, A_g^{\geq}) \lesssim_C s(g).$$

Puisque $s(g) \leq s_2(C)$, en prenant $s_2(C)$ assez petit, on peut supposer que la distance de Hausdorff entre A_{gh}^{\leq} et $A_{g,gh}^{\leq}$ est inférieure à la moitié de notre minoration de la distance minimale entre $A_{g,gh}^{\leq}$ et $V_g^> \oplus V_g^<$. On conclut que

$$\alpha(A_{gh}^{\leq}, V_g^> \oplus V_g^<) \gtrsim_C 1.$$

- On montre de même que les applications $P_2 : A_{hg}^{\leq} \rightarrow A_{hg,g}^{\leq}$, $\pi_g : A_{g,gh}^{\leq} \rightarrow A_g^{\leq}$ et $\pi_g : A_{hg,g}^{\leq} \rightarrow A_g^{\leq}$ sont bijectives et $O(C)$ -bornées.
- Les applications $\overline{P}_1, \overline{P}_2, \overline{\psi}_1$ et $\overline{\psi}_2$ de la moitié supérieure du diagramme s'obtiennent alors en composant plusieurs bijections $O(C)$ -bornées, donc sont elles-mêmes bijectives et $O(C)$ bornées. \square

Démonstration du lemme III.4.4. On va se concentrer sur la première estimation ; la démonstration de la deuxième estimation est analogue.

D'après le lemme III.2.14, les applications $\overline{\psi}_1$ et $\overline{\psi}_2$ sont des quasi-translations. Donc $\overline{g_{g,gh}}$ est aussi une quasi-translation.

On voudrait pouvoir prétendre que $\overline{g_{gh}}$ et $\overline{g_{g,gh}}$ sont en fait des translations. Pour faire ceci, on modifie légèrement le coin supérieur droit du diagramme III.1. On pose

$$\begin{cases} \phi'_{hg} := \ell(\overline{g_{gh}}) \circ \phi_{hg} \\ \phi'_{hg,g} := \ell(\overline{g_{g,gh}}) \circ \phi_{hg,g}, \end{cases}$$

où ℓ désigne la partie linéaire comme définie dans la section III.2.7, et on définit $\overline{P}'_2, \overline{\psi}'_2, \overline{g}'_{gh}, \overline{g}'_{g,gh}$ de sorte à rendre le nouveau diagramme (le diagramme III.2) commutatif. Les facteurs que nous avons introduits (les flèches horizontales courtes du diagramme III.2) sont de norme 1 : en effet, en tant que quasi-translations de $\hat{\mathbb{I}}$ fixant \mathbb{R}_0 , elles sont des transformations linéaires orthogonales (d'après la proposition III.2.8). Donc le lemme III.4.5 reste valable pour le diagramme III.2 ; mais maintenant, les applications modifiées \overline{g}'_{gh} et $\overline{g}'_{g,gh}$ sont des translations par construction.

On peut écrire :

$$\overline{g}'_{gh} = (\overline{P}_1^{-1} \circ \overline{g}'_{g,gh} \circ \overline{P}_1) \circ (\overline{P}_1^{-1} \circ \overline{P}'_2).$$

Alors, comme \overline{g}'_{gh} et $\overline{g}'_{g,gh}$ sont des translations, \overline{P}_1 et \overline{P}'_2 ont la même partie linéaire, de sorte que $\overline{P}_1^{-1} \circ \overline{P}'_2$ est aussi une translation. Étant la composée de deux applications $O(C)$ -bornées, elle est aussi $O(C)$ -bornée. On a donc

$$\overline{g}'_{gh} \approx_C \overline{P}_1^{-1} \circ \overline{g}'_{g,gh} \circ \overline{P}_1.$$

Il est clair que $\overline{g_{gh}}$ est $O(C)$ -presque-équivalente à \overline{g}'_{gh} et que $\overline{g_{g,gh}}$ est $O(C)$ -presque-équivalente à $\overline{g}'_{g,gh}$. Il reste à vérifier que l'application $\overline{g}'_{g,gh}$ est $O(C)$ -presque-équivalente à son conjugué $\overline{P}_1^{-1} \circ \overline{g}'_{g,gh} \circ \overline{P}_1$.

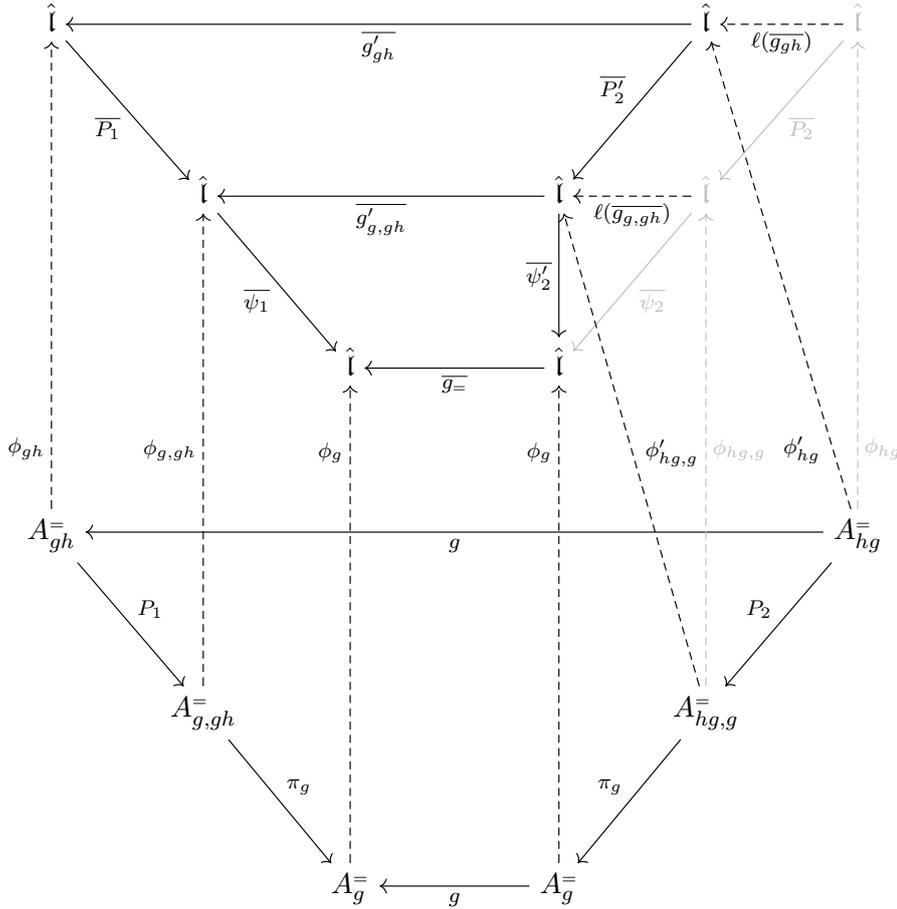


Diagramme III.2

Cela peut se déduire du lemme III.4.6 ci-dessous. En effet, soit $\overline{P_1''}$ la quasi-translation construite dans le lemme III.4.6. Soit $v \in \mathfrak{I}$ le vecteur de translation de $\overline{g'_{g,gh}}$, de sorte que

$$\overline{g'_{g,gh}} = \tau_v.$$

Alors on a

$$\left\| \overline{P_1}^{-1} \circ \overline{g'_{g,gh}} \circ \overline{P_1} - \overline{P_1''}^{-1} \circ \overline{g'_{g,gh}} \circ \overline{P_1''} \right\| = \left\| \tau_{\overline{P_1}^{-1}(v)} - \tau_{\overline{P_1''}^{-1}(v)} \right\|.$$

Il ne faut pas oublier que, pour tout vecteur u , bien qu'on appelle l'application τ_u une « translation », elle est en toute rigueur une transvection ; sa norme $\|\tau_u\|$ est donc égale à la norme de la matrice $\begin{pmatrix} 1 & \|u\| \\ 0 & 1 \end{pmatrix}$. En particulier on a $\|u\| \leq \|\tau_u\| \leq \|u\| + 1$. On en déduit que

$$\begin{aligned} \left\| \tau_{\overline{P_1}^{-1}(v)} - \tau_{\overline{P_1''}^{-1}(v)} \right\| &\leq \left\| \overline{P_1}^{-1}(v) - \overline{P_1''}^{-1}(v) \right\| + 1 \\ &\leq \left\| (\overline{P_1}^{-1} - \overline{P_1''}^{-1}) \right\|_{\mathfrak{I}} \|v\| + 1 \end{aligned}$$

(puisque $v \in \mathfrak{I}$).

Maintenant d'après la proposition III.2.8, on sait que la quasi-translation $\overline{P_1''}$ restreinte à \mathfrak{I} est un élément du groupe D ; comme celui-ci est compact, l'application $\rho \mapsto \rho^{-1}$ est lipschitzienne sur ce groupe. On peut alors déduire du lemme III.4.6 que

$$\left\| (\overline{P_1}^{-1} - \overline{P_1''}^{-1}) \right\|_{\mathfrak{I}} \lesssim_C s(\ell(g)).$$

D'autre part, on a $\|v\| \leq \|\tau_v\| = \left\| \overline{g'_{g,gh}} \right\| \lesssim_C \left\| g|_{A_g^=} \right\|$, car $\overline{g'_{g,gh}}$ est la composée de $g|_{A_g^=}$ avec plusieurs applications $O(C)$ -bornées. On en déduit que

$$\left\| \overline{P_1}^{-1} \circ \overline{g'_{g,gh}} \circ \overline{P_1} - \overline{P_1}''^{-1} \circ \overline{g'_{g,gh}} \circ \overline{P_1}'' \right\| \lesssim_C s(\ell(g)) \left\| g|_{A_g^=} \right\| + 1.$$

D'après le lemme III.2.20 (iii), on a $s(\ell(g)) \left\| g|_{A_g^=} \right\| \lesssim_C s(g)$; et on sait que $s(g) \leq 1$. On obtient finalement

$$\left\| \overline{P_1}^{-1} \circ \overline{g'_{g,gh}} \circ \overline{P_1} - \overline{P_1}''^{-1} \circ \overline{g'_{g,gh}} \circ \overline{P_1}'' \right\| \lesssim_C 1,$$

ce qui conclut. \square

Lemme III.4.6. *La partie linéaire de l'application $\overline{P_1}$ est « presque » une quasi-translation. Plus précisément, il existe une quasi-translation $\overline{P_1}''$ telle que*

$$\left\| (\overline{P_1} - \overline{P_1}'') \Big|_{\mathfrak{l}} \right\| \lesssim_C s(\ell(g)).$$

On rappelle que $\ell(g)$ est l'application qui a la même partie linéaire que g , mais sans partie de translation : voir la sous-section III.2.7. On utilise deux primes parce que le rapport entre $\overline{P_1}''$ et $\overline{P_1}$ n'a rien à voir avec le rapport entre $\overline{P_2}'$ et $\overline{P_2}$.

Démonstration. Soit $P_1'' : A_{gh}^= \rightarrow A_{g,gh}^=$ la projection parallèlement à $V_{gh}^<$; on pose $\overline{P_1}'' := \phi_{g,gh} \circ P_1'' \circ \phi_{gh}^{-1}$ l'endomorphisme correspondant de $\hat{\mathfrak{l}}$. Alors par le lemme III.2.14, $\overline{P_1}''$ est une quasi-translation.

On doit montrer que pour tout $x \in \mathfrak{l}$, on a

$$\|\overline{P_1}(x) - \overline{P_1}''(x)\| \lesssim_C s(\ell(g))\|x\|.$$

D'après la remarque III.4.2, c'est vrai si et seulement si pour tout $x \in V_{gh}^=$, on a

$$\|P_1(x) - P_1''(x)\| \lesssim_C s(\ell(g))\|x\|.$$

Prenons n'importe quel vecteur $x \in V_{gh}^=$. Décomposons-le de deux façons différentes :

$$\begin{aligned} x &=: \underbrace{x_1}_{\in V_{g,gh}^=} + \underbrace{x_2}_{\in V_g^<} + \underbrace{x_3}_{\in V_g^>} \\ &=: \underbrace{x'_1}_{\in V_{g,gh}^=} + \underbrace{x'_2}_{\in V_{gh}^<} \end{aligned}$$

(où on pose naturellement $V_{g,gh}^= := A_{g,gh}^= \cap \mathfrak{g}$), de sorte que $x_1 = P_1(x)$ et $x'_1 = P_1''(x)$. Notre premier but est d'établir l'estimation (III.4.1). En gros, l'idée est que comme $x \in V_{gh}^= \subset V_{gh}^{\geq}$, et comme ce dernier sous-espace est « proche » de $V_g^{\geq} = V_{g,gh}^= \oplus V_g^>$, la composante x_2 est « petite ».

Plus précisément, x_2 est l'image de x par la projection sur $V_g^<$ parallèlement à V_g^{\geq} ; donc $\phi_g(x_2)$ est l'image de $\phi_g(x)$ par la projection sur \mathfrak{n}^- parallèlement à \mathfrak{p}^+ , qui est une projection orthogonale. On en déduit que

$$\begin{aligned} \frac{\|\phi_g(x_2)\|}{\|\phi_g(x)\|} &= \sin \alpha(\phi_g(x), \mathfrak{p}^+) \\ &\leq \alpha(\phi_g(x), \mathfrak{p}^+) \\ &\leq \alpha^{\text{Haus}}(\phi_g(V_{gh}^{\geq}), \phi_g(V_g^{\geq})). \end{aligned}$$

Comme g est C -non-dégénérée, en utilisant le lemme III.2.16 on obtient

$$\|x_2\| \lesssim_C \alpha^{\text{Haus}}(V_{gh}^{\geq}, V_g^{\geq}) \|x\|.$$

Du corollaire III.3.7, on déduit que

$$\|x_2\| \lesssim_C s(\ell(g)) \|x\|. \quad (\text{III.4.1})$$

D'autre part, on a :

$$x_2 = \underbrace{(x'_1 - x_1)}_{\in V_{g,gh}^{\leq}} - \underbrace{x_3}_{\in V_g^{\geq}} + \underbrace{x'_2}_{\in V_{gh}^{\leq}},$$

d'où

$$\|\phi_{g,gh}(x_2)\|^2 = \|\phi_{g,gh}(x'_1 - x_1)\|^2 + \|\phi_{g,gh}(x_3)\|^2 + \|\phi_{g,gh}(x'_2)\|^2$$

et en particulier

$$\|\phi_{g,gh}(x'_1 - x_1)\| \leq \|\phi_{g,gh}(x_2)\|.$$

Puisque $(A_g^{\geq}, A_{gh}^{\leq})$ est $2C$ -non-dégénérée (remarque III.4.2), il découle que

$$\|x'_1 - x_1\| \lesssim_C \|x_2\|. \quad (\text{III.4.2})$$

En combinant cela avec (III.4.1), on obtient

$$\|x'_1 - x_1\| \lesssim_C s(\ell(g)) \|x\|,$$

CQFD. □

III.5. Invariants de Margulis de mots

Nous avons déjà étudié comment les forces de contraction (proposition III.3.6) et les invariants de Margulis (proposition III.4.1) se comportent quand on prend le produit de deux applications \mathbb{R} -régulières, C -non-dégénérées, suffisamment contractantes. Le but de cette section est de généraliser ces résultats à des mots de longueur arbitraire sur un ensemble de générateurs donné.

Définition III.5.1. Prenons k générateurs g_1, \dots, g_k . Considérons un mot $g = g_{i_1}^{\sigma_1} \dots g_{i_l}^{\sigma_l}$ de longueur $l \geq 1$ en ces générateurs et leurs inverses (avec, pour tout m , $1 \leq i_m \leq k$ et $\sigma_m = \pm 1$). On dit que g est *réduit* si pour tout m tel que $1 \leq m \leq l-1$, on a $(i_{m+1}, \sigma_{m+1}) \neq (i_m, -\sigma_m)$. On dit que g est *cycliquement réduit* s'il est réduit et s'il vérifie aussi $(i_1, \sigma_1) \neq (i_l, -\sigma_l)$.

Proposition III.5.2. *Pour tout $C \geq 1$, il existe une constante strictement positive $s_3(C) \leq 1$ ayant la propriété suivante. Prenons n'importe quelle famille d'applications $g_1, \dots, g_k \in G \rtimes \mathfrak{g}$ vérifiant les hypothèses suivantes :*

(H1) *Chaque g_i est \mathbb{R} -régulière.*

(H2) *N'importe quel couple pris parmi les applications $\{g_1, \dots, g_k, g_1^{-1}, \dots, g_k^{-1}\}$ est C -non-dégénéré, sauf évidemment s'il est de la forme (g_i, g_i^{-1}) pour un certain i .*

(H3) *Pour tout i , on a $s(g_i) \leq s_3(C)$ et $s(g_i^{-1}) \leq s_3(C)$.*

Prenons n'importe quel mot non vide cycliquement réduit $g = g_{i_1}^{\sigma_1} \dots g_{i_l}^{\sigma_l}$ (avec $1 \leq i_m \leq k$, $\sigma_m = \pm 1$ pour tout m). Alors l'application g est \mathbb{R} -régulière, $2C$ -non-dégénérée, et on a

$$\left\| M(g) - \sum_{m=1}^l M(g_{i_m}^{\sigma_m}) \right\| \leq l\mu(2C)$$

(où $\mu(2C)$ est la constante introduite dans la proposition III.4.1).

La démonstration se fait bien sûr par récurrence, en utilisant la proposition III.3.6 et la proposition III.4.1 ; cependant il y a une subtilité. En supposant que le couple (g, h) est C -non-dégénéré, on peut seulement conclure que gh est $2C$ -non-dégénérée ; cela casserait la récurrence si on utilisait une approche directe. Pour garantir la $2C$ -non-dégénérescence de tous les mots, on doit utiliser le fait que la force de contraction de g croît exponentiellement avec sa longueur, de sorte que la distance (de Hausdorff) entre A_g^\geq et $A_{g_{i_1}^\geq}$ est en fait la somme d'une série dont les termes décroissent exponentiellement et reste donc bornée. Pour le prendre en compte, on va démontrer par récurrence une série d'affirmations un peu plus compliquées.

Démonstration. On fixe $C \geq 1$, une constante strictement positive $s_3(C) \leq 1$ à déterminer au cours de la preuve et une famille g_1, \dots, g_k vérifiant les hypothèses (H1), (H2) et (H3). On va montrer par récurrence sur $\max(l, l')$ que dès qu'on prend un mot non vide cycliquement réduit $g = g_{i_1}^{\sigma_1} \dots g_{i_l}^{\sigma_l}$, on a les propriétés suivantes :

- (i) L'application g est \mathbb{R} -régulière.
- (ii)
$$\begin{cases} \alpha^{\text{Haus}} \left(A_g^\geq, A_{g_{i_1}^\geq}^\geq \right) \lesssim_C 2(1 - 2^{-(l-1)}) s_3(C) \\ \alpha^{\text{Haus}} \left(A_g^\leq, A_{g_{i_l}^\leq}^\leq \right) \lesssim_C 2(1 - 2^{-(l-1)}) s_3(C). \end{cases}$$
- (iii) $s(g) \leq 2^{-(l-1)} s_3(C)$.
- (iv)
$$\left\| M(g) - \sum_{m=1}^l M(g_{i_m}^{\sigma_m}) \right\| \leq (l-1)\mu(2C).$$
- (v) Si $h = g_{i'_1}^{\sigma'_1} \dots g_{i'_{l'}}^{\sigma'_{l'}}$ est un autre mot non vide cycliquement réduit tel que gh est aussi cycliquement réduit, le couple (g, h) est $2C$ -non-dégénéré.

En particulier, les propriétés (i), (iv) et (v) impliquent la proposition.

En effet, les cinq affirmations sont vraies pour $l = 1$ (et $l' = 1$). Soit maintenant $l \geq 2$, et supposons que les affirmations (i) à (v) sont vraies pour tous les mots cycliquement réduits de longueur m avec $1 \leq m \leq l-1$. Prenons n'importe quel mot cycliquement réduit $g = g_{i_1}^{\sigma_1} \dots g_{i_l}^{\sigma_l}$. Alors on affirme qu'il est possible de décomposer g en deux sous-mots cycliquement réduits

$$g' := g_{i_1}^{\sigma_1} \dots g_{i_m}^{\sigma_m} \quad \text{et} \quad g'' := g_{i_{m+1}}^{\sigma_{m+1}} \dots g_{i_l}^{\sigma_l},$$

tous deux non vides (c'est-à-dire $0 < m < l$).

En effet, supposons le contraire : supposons que pour tout tel m , on ait

$$(i_m, \sigma_m) = (i_1, -\sigma_1) \quad \text{ou} \quad (i_{m+1}, \sigma_{m+1}) = (i_l, -\sigma_l).$$

Montrons, par récurrence sur m , que la première condition est toujours fautive, donc que la deuxième est toujours vraie. Pour $m = 1$, c'est évident. Supposons qu'on le sait pour $m-1$; alors on a

$$(i_m, \sigma_m) = (i_l, -\sigma_l) = (i_2, \sigma_2) \neq (i_1, -\sigma_1)$$

car le mot est réduit. En prenant $m = l-1$, on obtient alors une contradiction.

Par les hypothèses de récurrence (i) et (v), g' et g'' sont \mathbb{R} -régulières et forment un couple $2C$ -non-dégénéré ; par l'hypothèse de récurrence (iii), on a $s(g') \leq 2^{-(m-1)} s_3(C) \leq s_3(C)$ et on peut supposer que $s_3(C) \leq s_1(2C)$ (de même pour g'^{-1} , g'' , g''^{-1}). Ainsi le couple (g', g'') vérifie la proposition III.3.6 (avec la constante $2C$). Montrons que g vérifie les propriétés (i) à (v).

- La propriété (i) (que g est \mathbb{R} -régulière) est une conséquence directe de la proposition III.3.6.

- Vérifions la propriété (iii). De la proposition III.3.6 (ii), on déduit que $s(g) \lesssim_C s(g')s(g'')$; on a alors, par l'hypothèse de récurrence (iii) :

$$\begin{aligned} s(g) &\lesssim_C \left(2^{-(m-1)}s_3(C)\right) \left(2^{-(l-m-1)}s_3(C)\right) \\ &= s_3(C) \left(2^{-(l-2)}s_3(C)\right). \end{aligned}$$

Si on prend $s_3(C)$ suffisamment petit, on obtient

$$s(g) \leq 2^{-(l-1)}s_3(C).$$

- Vérifions maintenant (ii); il suffit de vérifier la première inégalité (la deuxième en découle en substituant g^{-1} .) On rappelle que α^{Haus} est une distance sur l'ensemble de tous les sous-espaces vectoriels de $\hat{\mathfrak{g}}$, d'où

$$\alpha^{\text{Haus}} \left(A_g^{\geq}, A_{g_{i_1}^{\sigma_1}}^{\geq} \right) \leq \alpha^{\text{Haus}} \left(A_g^{\geq}, A_{g'}^{\geq} \right) + \alpha^{\text{Haus}} \left(A_{g'}^{\geq}, A_{g_{i_1}^{\sigma_1}}^{\geq} \right).$$

En estimant le premier terme par la proposition III.3.6 (i) et le deuxième terme par l'hypothèse de récurrence (ii), on obtient :

$$\alpha^{\text{Haus}} \left(A_g^{\geq}, A_{g_{i_1}^{\sigma_1}}^{\geq} \right) \lesssim_C s(g') + 2 \left(1 - 2^{-(m-1)}\right) s_3(C).$$

Or par l'hypothèse de récurrence (iii) on a $s(g') \leq 2^{-(m-1)}s_3(C)$, d'où

$$\begin{aligned} \alpha^{\text{Haus}} \left(A_g^{\geq}, A_{g_{i_1}^{\sigma_1}}^{\geq} \right) &\lesssim_C 2^{-(m-1)}s_3(C) + 2 \left(1 - 2^{-(m-1)}\right) s_3(C) \\ &= 2 \left(1 - 2^{-m}\right) s_3(C) \\ &\leq 2 \left(1 - 2^{-(l-1)}\right) s_3(C), \end{aligned}$$

puisque $m \leq l - 1$. (Ici la constante multiplicative implicite est la même que dans la proposition III.3.6 (i), et ne change pas d'un pas de récurrence à l'autre.)

- Ensuite on vérifie (iv). Par l'hypothèse de récurrence (iv), on a

$$\left\{ \left\| \begin{aligned} M(g') - \sum_{p=1}^m M(g_{i_p}^{\sigma_p}) \\ M(g'') - \sum_{p=m+1}^l M(g_{i_p}^{\sigma_p}) \end{aligned} \right\| \leq \begin{aligned} (m-1)\mu(2C) \\ (l-m-1)\mu(2C). \end{aligned} \right.$$

Si on prend $s_3(C) \leq s_2(2C)$, g' et g'' vérifient la proposition III.4.1, d'où

$$\|M(g) - M(g') - M(g'')\| \leq \mu(2C).$$

En ajoutant les trois inégalités, on obtient la conclusion souhaitée.

- Il reste à vérifier (v) : soit $h = g_{i'_1}^{\sigma'_1} \dots g_{i'_l}^{\sigma'_l}$ un autre mot cycliquement réduit (avec $1 \leq l' \leq l$) tel que gh est aussi cycliquement réduit. On doit vérifier que les quatre couples (A_g^{\geq}, A_g^{\leq}) , (A_g^{\geq}, A_h^{\leq}) et ainsi de suite sont $2C$ -non-dégénérés. Ceci découle par le lemme III.3.9 de la propriété (ii) (appliquée à g et à h) et de l'hypothèse (H2), pourvu qu'on ait pris $s_3(C)$ suffisamment petit. \square

III.6. Construction du groupe

On démontre maintenant (lemme III.6.1) que si on prend un groupe engendré par une famille finie d'applications \mathbb{R} -régulières, C -non-dégénérées, suffisamment contractantes et ayant des invariants de Margulis bien choisis, il satisfait toutes les conclusions du théorème III.1.1, sauf la Zariski-densité. On exhibe ensuite un tel groupe qui est en plus Zariski-dense (ce qui démontre le théorème III.1.1).

On rappelle (voir la section III.4) que w_0 est un certain élément de G tel que $w_0(\mathfrak{p}^+, \mathfrak{p}^-) = (\mathfrak{p}^-, \mathfrak{p}^+)$. Notez qu'il existe un vecteur non nul $v \in \mathfrak{z}$ fixe par $-w_0$. En effet, w_0 normalise l'espace $\mathfrak{p}^+ \cap \mathfrak{p}^- = \mathfrak{l}$; donc il normalise son centre déployé \mathfrak{a} . Comme w_0 échange les racines positives et négatives, la chambre de Weyl ouverte \mathfrak{a}^+ est stable par l'involution $-w_0$. Comme \mathfrak{a}^+ est un ensemble convexe, on peut prendre $v = v' - w_0(v')$ pour n'importe quel v' dans \mathfrak{a}^+ . Alors on a en effet $v \neq 0$, $v \in \mathfrak{a} \subset \mathfrak{z}$ et $-w_0(v) = v$.

On fixe dorénavant un vecteur M_0 colinéaire à v et tel que $\|M_0\| = 2\mu(2C)$.

Lemme III.6.1. *Considérons une famille $g_1, \dots, g_k \in G \rtimes \mathfrak{g}$ vérifiant les hypothèses (H1), (H2) et (H3) de la proposition III.5.2, et aussi la condition supplémentaire*

(H4) Pour tout i , $M(g_i) = M_0$.

Alors ces applications engendrent un groupe libre discret agissant proprement sur l'espace affine $\mathfrak{g}_{\text{Aff}}$.

Démonstration. Pour montrer que le groupe est libre, il suffit de remarquer que tout mot réduit non vide en les $g_i^{\pm 1}$ est conjugué à un certain mot non vide cycliquement réduit, lequel, par la proposition III.5.2, est \mathbb{R} -régulier et en particulier différent de l'identité.

Pour montrer que le groupe est discret et agit proprement, la première étape est de montrer l'inégalité (III.6.1) ci-dessous, qui dit que les éléments cycliquement réduits du groupe ont des invariants de Margulis qui croissent de façon non bornée. Prenons n'importe quel mot cycliquement réduit $g = g_{i_1}^{\sigma_1} \dots g_{i_l}^{\sigma_l}$. Alors de la proposition III.5.2, il découle que

$$\|M(g)\| \geq \left\| \sum_{m=1}^l M(g_{i_m}^{\sigma_m}) \right\| - l\mu(2C).$$

D'autre part, on sait que pour tous i et σ , on a

$$M(g_i^\sigma) = M_0.$$

En effet pour $\sigma = +1$, c'est vrai par hypothèse; pour $\sigma = -1$, on a

$$M(g_i^{-1}) = -w_0(M(g_i)) = -w_0(M_0) = M_0,$$

d'après la proposition III.4.1 (i) et par définition de M_0 . On conclut que

$$\begin{aligned} \|M(g)\| &\geq \|lM_0\| - l\mu(2C) \\ &= 2l\mu(2C) - l\mu(2C) \\ &= l\mu(2C). \end{aligned} \tag{III.6.1}$$

Maintenant soit X n'importe quelle partie compacte de l'espace affine $\mathfrak{g}_{\text{Aff}}$, et supposons que g est un mot réduit en les $g_i^{\pm 1}$. On doit montrer que dès que g est suffisamment long, on a $g(X) \cap X = \emptyset$.

On commence par remarquer qu'il est toujours possible de trouver un indice i et un signe σ tels que $g_i^\sigma g$ est cycliquement réduit. On a alors :

$$g(X) \cap X = \emptyset \iff g_i^\sigma g(X) \cap g_i^\sigma(X) = \emptyset.$$

En posant $X' = \bigcup_{i,\sigma} g_i^\sigma(X)$ (qui est bien sûr encore compact), il suffit de montrer que dès que g est *cycliquement* réduit et suffisamment long, on a

$$g(X) \cap X' = \emptyset. \quad (\text{III.6.2})$$

Soit ϕ_g une application canonisante optimale pour g , $\hat{\pi}_3$ la projection (orthogonale) sur $\mathfrak{z} \oplus \mathbb{R}_0$ parallèlement à \mathfrak{d} , \mathfrak{n}^+ et \mathfrak{n}^- (qu'on peut voir comme une application affine agissant sur $\mathfrak{g}_{\text{Aff}}$). Alors par la définition de l'invariant de Margulis, on a

$$\begin{aligned} \hat{\pi}_3 \circ \phi_g(g(X)) &= \tau_{M(g)} \circ \hat{\pi}_3 \circ \phi_g(X) \\ &= \hat{\pi}_3 \circ \phi_g(X) + M(g). \end{aligned}$$

Maintenant on remarque que, d'une part, l'application g est $2C$ -non-dégénérée d'après la proposition III.5.2, d'où

$$\|\hat{\pi}_3 \circ \phi_g(x - y)\| \leq \|\phi_g(x - y)\| \leq 2C\|x - y\|$$

pour tous $x, y \in \mathfrak{g}_{\text{Aff}}$. D'autre part, on rappelle l'inégalité (III.6.1) :

$$\|M(g)\| \geq l\mu(2C),$$

où l est la longueur de g . On en déduit que dès qu'on a

$$l > \frac{2C}{\mu(2C)} \max_{x \in X, y \in X'} \|x - y\|,$$

les images $\hat{\pi}_3 \circ \phi_g(g(X))$ et $\hat{\pi}_3 \circ \phi_g(X')$ sont disjointes. Ceci implique (III.6.2), ce qui montre que le groupe est discret et agit proprement. \square

Démonstration du théorème III.1.1. La stratégie est désormais claire : on va trouver une constante positive $C \geq 1$ et une famille d'applications $g_1, \dots, g_k \in G \ltimes \mathfrak{g}$ (avec $k \geq 2$) qui vérifient les conditions (H1) à (H4) et dont les parties linéaires engendrent un sous-groupe Zariski-dense de G , puis on va appliquer le lemme III.6.1. On procède en plusieurs étapes.

- D'après le lemme 7.2 dans [6], on peut trouver une famille d'applications $\gamma_1, \dots, \gamma_k \in G$ (qu'on verra comme des éléments de $G \ltimes \mathfrak{g}$ en identifiant G au stabilisateur de \mathbb{R}_0), telle que :
 - (i) Chaque γ_i est \mathbb{R} -régulière (c'est (H1)).
 - (ii) Pour tout couple d'indices i, i' et de signes σ, σ' tels que $(i', \sigma') \neq (i, -\sigma)$, les espaces $V_{\gamma_i^\sigma}^{\geq}$ et $V_{\gamma_{i'}^{\sigma'}}^{\leq}$ sont transverses.
 - (iii) Chaque γ_i engendre, à lui tout seul, un groupe Zariski-connexe.
 - (iv) Tous les γ_i ensemble engendrent un sous-groupe Zariski-dense de G .

Notez que la Zariski-densité n'est possible que pour $k \geq 2$.

- Clairement, tout couple d'espaces transverses est C -non-dégénéré pour un certain réel fini C ; et nous avons ici un nombre fini de tels couples. Donc en choisissant une bonne valeur de C (qu'on fixera pour le reste de la preuve), l'hypothèse (H2) devient une conséquence directe de la condition (ii) ci-dessus.
- La condition de Zariski-connexité nous permet de remplacer chaque γ_i par une certaine puissance γ_i^N sans sacrifier la Zariski-densité. Remplacer une application par sa puissance ne change pas sa géométrie, donc les conditions (i) et (ii) sont aussi préservées. En choisissant N suffisamment grand, on peut supposer que les nombres $s(\gamma_i^{\pm 1})$ sont aussi petits qu'on veut : cela nous donne (H3). En fait, on va même supposer que pour tout i , on a $s(\gamma_i^{\pm 1}) \leq s_4(C)$ pour une constante $s_4(C)$ encore plus petite que $s_3(C)$, qu'on spécifiera bientôt.

– Pour satisfaire (H4), on remplace les applications γ_i par les applications

$$g_i := \tau_{\phi_i^{-1}(M_0)} \circ \gamma_i$$

(pour $1 \leq i \leq k$), où ϕ_i est une application canonisante pour γ_i .

On doit vérifier que cela ne brise pas les trois premières conditions. En effet, pour tout i , on a $\gamma_i = \ell(g_i)$; même mieux, puisque le vecteur de translation $\phi_i^{-1}(M_0)$ se trouve dans le sous-espace $V_{\gamma_i}^=$ stable par γ_i , clairement la translation commute avec γ_i , donc g_i a la même géométrie que γ_i (dans le sens que $A_{g_i}^{\geq} = A_{\gamma_i}^{\geq} = V_{\gamma_i}^{\geq} \oplus \mathbb{R}_0$ et $A_{g_i}^{\leq} = A_{\gamma_i}^{\leq} = V_{\gamma_i}^{\leq} \oplus \mathbb{R}_0$). Donc les g_i vérifient encore les hypothèses (H1) et (H2), mais maintenant on a $M(g_i) = M_0$ (c'est (H4)). Quant à la force de contraction, on a, d'après le lemme III.2.20 :

$$s(g_i) \lesssim_C s(\gamma_i) \|\tau_{M_0}\| \leq s_4(C) \|\tau_{M_0}\|,$$

et de même pour g_i^{-1} . Étant donné que $\|M_0\| = 2\mu(2C)$, $\|\tau_{M_0}\|$ ne dépend que de C . On en déduit qu'en choisissant $s_4(C)$ suffisamment petit, l'hypothèse (H3) est vérifiée. On conclut que le groupe engendré par les éléments g_1, \dots, g_k est discret et agit proprement (d'après le lemme III.6.1), est libre (selon le même résultat), non abélien (puisque $k \geq 2$), et de partie linéaire Zariski-dense dans G , CQFD. \square

Quatrième partie.

Fonctions harmoniques sur le triangle de Sierpinski

Le but de cette partie est de répondre à la question de Kirillov énoncée en introduction. On appelle u une certaine fonction harmonique sur le triangle de Sierpinski restreinte à un de ses côtés. Les résultats principaux sont le théorème IV.5.4, qui donne la méthode générale pour déterminer la dérivée de u , le théorème IV.5.7, qui donne un algorithme effectif pour la calculer en tout point rationnel, et le théorème IV.5.12, qui dit que la dérivée de u est toujours soit égale à 0, soit à ∞ , soit n'est pas définie.

IV.1. Introduction

IV.1.1. Notations et conventions

Dans toute cette partie, ω désigne la racine cubique de -1 de partie imaginaire positive. Pour $z \in \mathbb{C}$, on note $h_z : \mathbb{C} \rightarrow \mathbb{C}$ l'homothétie de centre z et de rapport $\frac{1}{2}$: $h_z(w) = \frac{w+z}{2}$.

En parlant d'une quantité strictement positive, on dira qu'elle est *positivement bornée* si elle est majorée par une constante finie et minorée par une constante strictement positive. Ceci équivaut à dire que son logarithme est borné.

Lorsqu'on dira qu'une dérivée existe ou qu'elle est définie, on voudra dire que le taux de variation a une limite dans \mathbb{R} , i. e. on admettra implicitement aussi les dérivées infinies.

On aura souvent besoin de faire des estimations à un facteur positivement borné près. Comme toutes les normes sur un espace vectoriel de dimension finie sont équivalentes, dans un tel contexte, on n'a pas besoin de les distinguer. On va donc généralement se contenter d'écrire « $\|\bullet\|$ », en sous-entendant que l'affirmation qu'on fait est vraie quelle que soit la norme. Cependant, pour faire certains calculs, on aura besoin d'utiliser des normes spécifiques : on rajoutera alors un indice pour préciser laquelle.

On devra aussi souvent manipuler des produits de la forme $F_{a_1} \dots F_{a_n}$, où F est un certain opérateur qui peut prendre une valeur parmi F_0 et F_1 (par exemple h , M , \vec{M} , \tilde{M} et ainsi de suite). Par abus de notation, on écrira alors simplement $F_{a_1 \dots a_n}$, ou même F_w (si on a déjà convenu que $w = a_1 \dots a_n$).

Tout ce qui est écrit en notation positionnelle avec une virgule est en binaire (par exemple $\frac{1}{2} = 0,1$), sauf si le contexte indique clairement le contraire (par exemple $\pi \approx 3,14\dots$). De même, à tout mot fini $w = a_1 \dots a_n$ sur l'alphabet $\{0,1\}$, on associe l'entier $\sum_{i=1}^n a_i 2^{n-i}$ qu'il représente en binaire.

Il est bien connu que les points rationnels dyadiques ont deux développements binaires. Pour $s \in [0,1[$ (resp. $s \in]0,1]$), on appelle son *développement binaire supérieur* (resp. *inférieur*) l'unique développement binaire de s qui ne se termine pas par une suite infinie de uns (resp. de zéros). Pour les problèmes qu'on va étudier, les deux développements jouent des rôles parfaitement symétriques ; on parlera donc souvent de « développement binaire » sans préciser, et il faudra comprendre « développement binaire inférieur ou supérieur ».

IV.1.2. Problématique

On appelle u l'unique fonction continue de $[0,1]$ dans \mathbb{R}^3 qui vérifie :

$$\begin{cases} u(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, & u(1) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ u \circ h_0 = M_0 \circ u \\ u \circ h_1 = M_1 \circ u, \end{cases} \quad (\text{IV.1.1})$$

où on pose

$$M_0 := \frac{1}{5} \begin{pmatrix} 5 & 2 & 2 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad M_1 := \frac{1}{5} \begin{pmatrix} 2 & 0 & 1 \\ 2 & 5 & 2 \\ 1 & 0 & 2 \end{pmatrix}$$

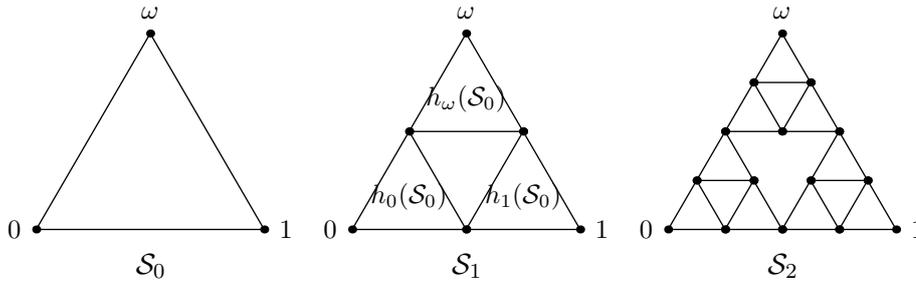


FIGURE IV.1.: Approximations discrètes du triangle de Sierpinski.

(voir aussi la figure IV.5). On justifiera plus loin l'existence et l'unicité de cette fonction.

Rappelons la question de Kirillov ([13], « Problème 2 », vers la fin de la Section 3.2) :

Calculer explicitement la dérivée $u'(t)$ à chaque fois que c'est possible (par exemple en tous les points rationnels).

Le but de cette partie est de répondre à cette question. Il faut cependant préciser que nos notations diffèrent légèrement des siennes. Kirillov utilise la lettre u pour désigner des fonctions harmoniques particulières à valeurs réelles, notamment quatre d'entre elles : $\phi = u_0^0_1$, $\psi = u_0^1_1$, $\chi = u_0^{-1}_1$ et $\xi = u_0^2_1$. Mais comme ces fonctions s'obtiennent par projection à partir de celle que nous appelons u , la question telle qu'on la formule ici est une généralisation de celle posée par Kirillov.

Pour une introduction plus complète aux applications harmoniques sur le triangle de Sierpinski et sur d'autres fractales, voir le livre de Strichartz [21].

IV.2. Préliminaires : fonctions harmoniques sur le triangle de Sierpinski

Définition IV.2.1. On définit le triangle de Sierpinski par approximations successives. On pose d'abord

$$\mathcal{S}_0 := \{0, 1, \omega\},$$

puis on définit récursivement, pour tout $n \in \mathbb{N}$,

$$\mathcal{S}_{n+1} := h_0(\mathcal{S}_n) \cup h_1(\mathcal{S}_n) \cup h_\omega(\mathcal{S}_n).$$

Donnons un nom à la réunion de toutes ces approximations :

$$\mathcal{S}_\infty := \bigcup_{n=0}^{+\infty} \mathcal{S}_n,$$

et le *triangle de Sierpinski* est alors défini comme l'adhérence de cette réunion :

$$\mathcal{S} := \overline{\mathcal{S}_\infty}.$$

Notez qu'on peut naturellement voir \mathcal{S}_n comme un graphe : \mathcal{S}_0 est un graphe complet sur trois sommets, et comme arêtes de \mathcal{S}_{n+1} , on prend les images des arêtes de \mathcal{S}_n par les trois homothéties. Il est alors facile de voir que chaque sommet sauf 0, 1 et ω a exactement quatre voisins. (Voir la figure IV.1.) Dans la suite, E désigne un certain espace vectoriel réel, et l'écriture « $s \sim t$ » signifie « s est un voisin de t ».

Définition IV.2.2. Soit $n \in \mathbb{N}$. Une *fonction harmonique* sur \mathcal{S}_n est une fonction $f : \mathcal{S}_n \rightarrow E$ telle que :

$$\forall s \in \mathcal{S}_n \setminus \{0, 1, \omega\}, \quad f(s) = \frac{1}{4} \sum_{t \sim s} f(t).$$

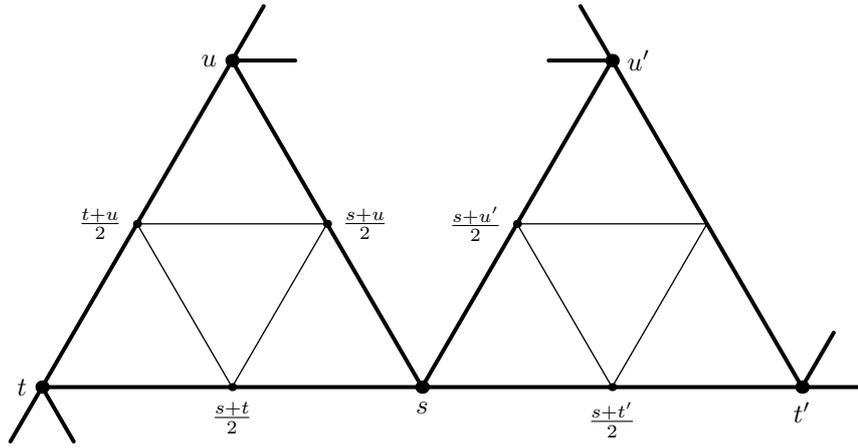


FIGURE IV.2.: Subdivision de \mathcal{S}_n .

On n'impose aucune condition sur les valeurs $f(0)$, $f(1)$ et $f(\omega)$: on considère que ce sont les conditions aux limites.

Définition IV.2.3. Une *fonction harmonique* sur \mathcal{S} est une fonction continue $f : \mathcal{S} \rightarrow E$ dont la restriction à chaque \mathcal{S}_n est harmonique.

Le résultat suivant, ainsi que sa démonstration, viennent du livre de Kirillov [13]. La première fois qu'il est mentionné dans la littérature semble être dans [12].

Proposition IV.2.4. Soient $a, b, c \in E$. Alors il existe une unique fonction f qui est harmonique sur \mathcal{S} et qui satisfait les conditions aux limites :

$$f(0) = a, \quad f(1) = b, \quad f(\omega) = c.$$

(Cela signifie notamment que l'ensemble de toutes les fonctions harmoniques sur \mathcal{S} à valeurs réelles est un espace vectoriel à trois dimensions.)

Cette fonction sera notée $f_a^c b$.

Démonstration. Pour vérifier cela, on va construire explicitement, par récurrence, la famille de fonctions qui vérifient ces conditions sur les approximations successives de \mathcal{S} .

D'abord on définit la fonction harmonique sur $\mathcal{S}_0 = \{0, 1, \omega\}$ qui satisfait les conditions aux limites (qui est évidemment unique) :

$$f_0 : \quad 0 \mapsto a, \quad 1 \mapsto b, \quad \omega \mapsto c.$$

Supposons maintenant qu'on a une fonction harmonique $f_n : \mathcal{S}_n \rightarrow E$ qui satisfait les conditions aux limites. On va maintenant définir une fonction f_{n+1} qui étend f_n à \mathcal{S}_{n+1} . Soient s, t, u n'importe quels trois sommets adjacents de \mathcal{S}_n qui forment un triangle ayant la même orientation que \mathcal{S}_0 . Alors \mathcal{S}_{n+1} contient encore ces sommets, mais aussi les sommets $\frac{s+t}{2}$, $\frac{t+u}{2}$, $\frac{s+u}{2}$ (voir la figure IV.2). Réciproquement, il est facile de voir que tous les sommets de \mathcal{S}_{n+1} peuvent être pris en compte de cette manière. On pose alors

$$\begin{aligned} f_{n+1} \left(\frac{s+t}{2} \right) &= \frac{2f_n(s) + 2f_n(t) + f_n(u)}{5} \\ f_{n+1} \left(\frac{s+u}{2} \right) &= \frac{2f_n(s) + f_n(t) + 2f_n(u)}{5} \\ f_{n+1} \left(\frac{t+u}{2} \right) &= \frac{f_n(s) + 2f_n(t) + 2f_n(u)}{5} \end{aligned}$$

Étant un prolongement de f_n , cette fonction satisfait encore les conditions aux limites. Vérifier qu'elle est harmonique sur \mathcal{S}_{n+1} est un calcul immédiat ; il suffit de le faire pour s et pour $\frac{s+t}{2}$, puisque n'importe quel sommet est analogue à un de ces deux-là.

- Les voisins de s dans \mathcal{S}_{n+1} sont $\frac{s+t}{2}$, $\frac{s+u}{2}$, $\frac{s+t'}{2}$ et $\frac{s+u'}{2}$, où t' et u' sont les deux autres voisins de s dans \mathcal{S}_n (voir la figure IV.2). On a :

$$\begin{aligned}
& f_{n+1}\left(\frac{s+t}{2}\right) + f_{n+1}\left(\frac{s+u}{2}\right) + f_{n+1}\left(\frac{s+t'}{2}\right) + f_{n+1}\left(\frac{s+u'}{2}\right) = \\
&= \frac{2f_n(s) + 2f_n(t) + f_n(u)}{5} + \frac{2f_n(s) + f_n(t) + 2f_n(u)}{5} + \\
&\quad + \frac{2f_n(s) + 2f_n(t') + f_n(u')}{5} + \frac{2f_n(s) + f_n(t') + 2f_n(u')}{5} \\
&= \frac{8f_n(s) + 3(f_n(t) + f_n(u) + f_n(t') + f_n(u'))}{5} \\
&= \frac{8 + 12}{5} f_n(s) \\
&= 4f_{n+1}(s)
\end{aligned}$$

- Les voisins de $\frac{s+t}{2}$ sont s , $\frac{s+u}{2}$, $\frac{t+u}{2}$ et t . On a :

$$\begin{aligned}
& f_{n+1}(s) + f_{n+1}\left(\frac{s+u}{2}\right) + f_{n+1}\left(\frac{t+u}{2}\right) + f_{n+1}(t) = \\
&= f_n(s) + \frac{2f_n(s) + f_n(t) + 2f_n(u)}{5} + \frac{f_n(s) + 2f_n(t) + 2f_n(u)}{5} + f_n(t) \\
&= \frac{8f_n(s) + 8f_n(t) + 4f_n(u)}{5} \\
&= 4f_{n+1}\left(\frac{s+t}{2}\right)
\end{aligned}$$

Montrons maintenant que ces fonctions sont uniques. En effet, considérons deux fonctions harmoniques sur \mathcal{S}_n qui vérifient les mêmes conditions aux limites. Alors leur différence vaut 0 en 0, 1 et ω , et est encore harmonique. Le principe du maximum garantit alors qu'elle est identiquement nulle.

Cela nous permet de définir $f_a^c_b$ sur \mathcal{S}_∞ , et le calcul qu'on vient de faire nous donne une relation de récurrence qui permet de la calculer en un point quelconque. Réécrivons-la sous une forme plus commode :

$$\begin{cases} f_a^c_b \circ h_0 = f_a^{\frac{2a+b+2c}{5}}_{\frac{2a+2b+c}{5}} \\ f_a^c_b \circ h_1 = f_{\frac{2a+2b+c}{5}}^{\frac{a+2b+2c}{5}}_b \\ f_a^c_b \circ h_\omega = f_{\frac{2a+b+2c}{5}}^c_{\frac{a+2b+2c}{5}}. \end{cases} \quad (\text{IV.2.1})$$

Pour pouvoir étendre continument cette fonction à tout le triangle de Sierpinski, on doit vérifier qu'elle est uniformément continue. On pose

$$D(a, b, c) := \max(\|a - b\|, \|b - c\|, \|a - c\|).$$

Par le principe du maximum, on a, pour toutes les valeurs de a, b, c et pour tous $s, t \in \mathcal{S}_\infty$,

$$\|f_a^c_b(s) - f_a^c_b(t)\| \leq D(a, b, c).$$

D'un autre côté, on vérifie facilement qu'on a

$$D\left(a, \frac{2a+b+2c}{5}, \frac{2a+2b+c}{5}\right) \leq \frac{3}{5}D(a, b, c).$$

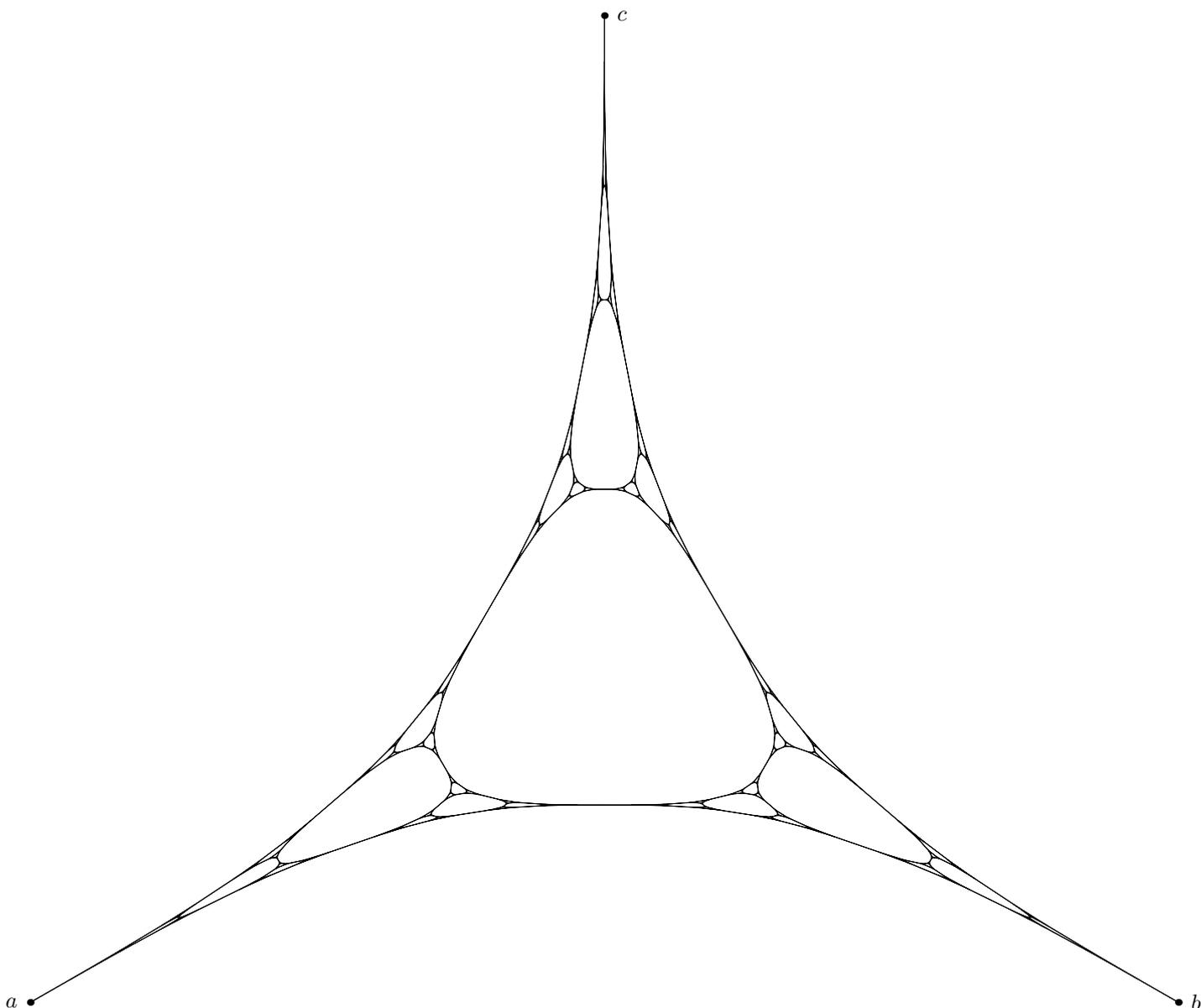


FIGURE IV.3.: L'image harmonique du triangle de Sierpinski, $f_a^c_b(\mathcal{S})$.

En utilisant (IV.2.1), par récurrence, il suit que si T est un « triangle élémentaire » de niveau n , on a

$$\forall s, t \in T, \quad \|f_a^c_b(s) - f_a^c_b(t)\| \leq \left(\frac{3}{5}\right)^n C,$$

où C ne dépend que de a , b et c . Maintenant considérons deux points de \mathcal{S}_∞ dont la distance est au plus égale à $(\frac{1}{2})^n$; alors il est facile de se convaincre que les deux triangles élémentaires de niveau $n - 1$ qui les contiennent sont soit confondus, soit adjacents. On en déduit

$$\forall s, t \in \mathcal{S}_\infty, \quad |s - t| \leq \left(\frac{1}{2}\right)^n \implies \|f_a^c_b(s) - f_a^c_b(t)\| \leq \left(\frac{3}{5}\right)^n C',$$

où C' est encore une constante réelle strictement positive qui ne dépend que de a , b et c — d'où la conclusion. Cela montre que $f_a^c_b$ est effectivement bien définie sur \mathcal{S} . \square

De façon très informelle, $f_a^c_b$ décrit la forme que prendrait un triangle de Sierpinski en caoutchouc si on le tendait sur trois clous fixés en a , b et c (voir la figure IV.3).

Notez que par continuité, les relations (IV.2.1) sont encore valables dans \mathcal{S} . Elles sont très importantes : elles fournissent une définition intrinsèque des fonctions $f_a^c_b$.

À partir de maintenant, en suivant Kirillov, on se restreint à un côté du triangle de Sierpinski : on pose $u_a^c b = f_a^c b|_{[0,1]}$. On peut alors abandonner la troisième ligne dans (IV.2.1). De plus, Kirillov s'intéresse principalement aux fonctions à valeurs réelles, notamment quatre d'entre elles : $\phi = u_0^0 1$, $\psi = u_0^1 1$, $\chi = u_0^{-1} 1$ et $\xi = u_0^2 1$ (chez lui, u désigne en fait une fonction quelconque parmi ces quatre). Elles ont cependant un inconvénient : il est difficile d'étudier une de ces fonctions indépendamment des autres, car (IV.2.1) nous oblige à changer les valeurs de a , b et c .

Ceci nous amène à introduire la fonction $u_{e_0 e_1 e_\omega}$, où (e_0, e_1, e_ω) est la base canonique de \mathbb{R}^3 . Son intérêt est qu'elle contient toute l'information à propos de toutes les fonctions $u_a^c b$. En effet, quels que soient a , b , et c , on a $u_a^c b = \Phi \circ u_{e_0 e_1 e_\omega}$, où $\Phi : \mathbb{R}^3 \rightarrow E$ est l'application linéaire qui envoie la base canonique sur (a, b, c) . (Si a , b et c sont réels, il s'agit simplement de la forme linéaire de matrice $\begin{pmatrix} a & b & c \end{pmatrix}$.) Les relations (IV.2.1) fournissent alors une équation fonctionnelle qui exprime $u_{e_0 e_1 e_\omega}$ en termes d'elle-même — qui n'est autre que (IV.1.1). On voit donc que la fonction u qu'on a définie au début n'est qu'une notation abrégée (à laquelle on se tiendra désormais) pour $u_{e_0 e_1 e_\omega}$. Son existence et son unicité découlent de la preuve de la proposition IV.2.4.

IV.3. Généralités sur u

Notez que tout ce qu'on dira dans cette section peut très naturellement se généraliser à \mathcal{S} tout entier, c'est-à-dire à $f_{e_0 e_1 e_\omega}$. Il suffit pour cela d'introduire la matrice M_ω , par analogie avec M_0 et M_1 , et de comprendre qu'un point générique de \mathcal{S} peut être décrit par une suite infinie de symboles 0, 1 ou ω , par analogie avec le développement binaire.

On aura d'abord besoin de diagonaliser les M_i . Chacune d'entre elles a pour valeurs propres 1, $\frac{3}{5}$ et $\frac{1}{5}$, avec les vecteurs propres respectifs e_i , \vec{v}_i et \vec{w}_i , où :

$$e_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_0 = \begin{pmatrix} -1 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \vec{w}_0 = \begin{pmatrix} 0 \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix},$$

$$e_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_1 = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{pmatrix}, \quad \vec{w}_1 = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix}.$$

Notez que

$$f_1^0 0 + f_0^0 1 + f_0^1 0 = f_1^1 1 \equiv 1,$$

ce qui montre que tout se passe en fait en deux dimensions : toutes les valeurs de $f_{e_0 e_1 e_\omega}$, donc aussi de u , sont comprises dans le plan affine

$$\mathcal{H} := \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 1\}.$$

Bien entendu, les applications M_i le laissent stable, et induisent des applications affines sur celui-ci (voir la figure IV.4). C'est d'ailleurs pour cela qu'on a choisi d'écrire e_i plutôt que \vec{e}_i : les e_i sont des éléments du plan affine \mathcal{H} , à savoir — avec cette interprétation — les points fixes des applications affines M_i .

Essayons maintenant d'écrire une formule explicite pour $u(s)$. On doit d'abord exprimer s à partir de 0, de 1 et des h_i , ce qui se fait en écrivant :

$$s = \lim_{n \rightarrow \infty} h_{a_1 \dots a_n}(0) = \lim_{n \rightarrow \infty} h_{a_1 \dots a_n}(1),$$

où $s = 0, a_1 a_2 \dots$ est un développement binaire (inférieur ou supérieur) de s . En appliquant les relations ci-dessus, on obtient le résultat suivant :

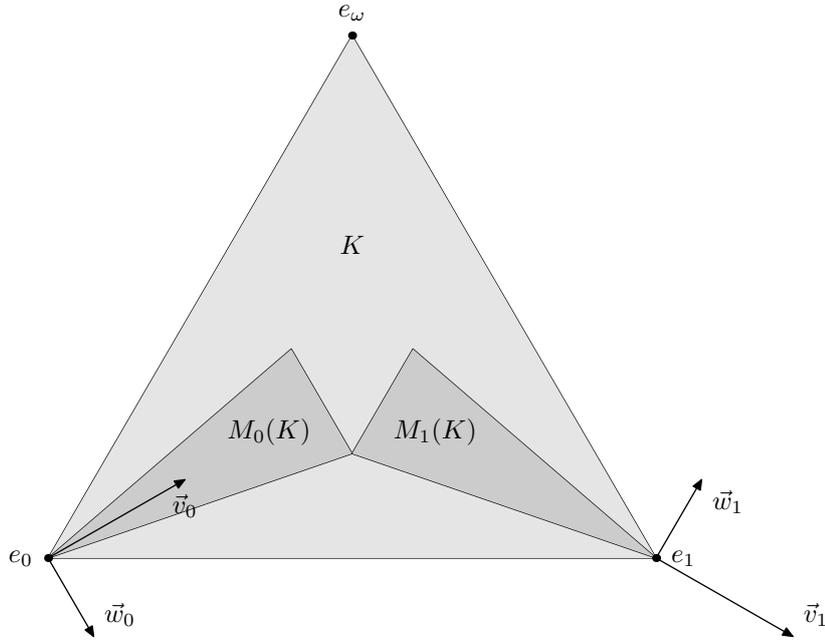


FIGURE IV.4.: Vecteurs propres des opérateurs M_i ; le plan de la feuille correspond au plan affine \mathcal{H} . L'origine de \mathbb{R}^3 se trouve donc sous le centre de la figure. Le triangle K est défini un peu plus loin ; il est rajouté pour mieux illustrer l'action des M_i sur \mathcal{H} .

Proposition IV.3.1. Soit $K := \mathcal{H} \cap \mathbb{R}_{\geq 0}^3$ le triangle (plein) avec les vecteurs de la base canonique pour sommets. Alors pour tout $s \in [0, 1]$ et pour toute valeur initiale $u_0 \in K$, on a

$$u(s) = \lim_{n \rightarrow \infty} M_{a_1 \dots a_n}(u_0), \quad (\text{IV.3.1})$$

où $s = 0, a_1 a_2 \dots$ est un développement binaire.

(Voir la figure IV.5.)

Démonstration. Réécrivons d'abord cette formule de façon à mettre en évidence le fait que le résultat ne dépend pas de la valeur initiale. Il est facile de vérifier que K est stable aussi bien par M_0 que par M_1 : par conséquent, les $M_{a_1 \dots a_n}(K)$ forment une suite décroissante pour l'inclusion. Dans la preuve de la proposition IV.2.4, on a déjà vu que les diamètres de ces ensembles tendaient vers 0. Ils sont clairement fermés. Leur intersection est donc réduite à un point :

$$\forall u_0 \in K, \quad \left\{ \lim_{n \rightarrow \infty} M_{a_1 \dots a_n}(u_0) \right\} = \bigcap_{n=0}^{\infty} M_{a_1 \dots a_n}(K).$$

D'autre part, une récurrence immédiate montre que pour tout mot fini w sur l'alphabet $\{0, 1\}$, $h_w(0) = 0, w$, où $0, w$ est vu comme une fraction binaire, et que $u \circ h_w = M_w \circ u$. Par conséquent, par les relations (IV.1.1) et par la continuité de u , on a

$$u \left(\sum_{i=1}^{\infty} a_i 2^{-i} \right) = \lim_{n \rightarrow \infty} M_{a_1 \dots a_n} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

d'où la formule. □

Corollaire IV.3.2. u est injective.

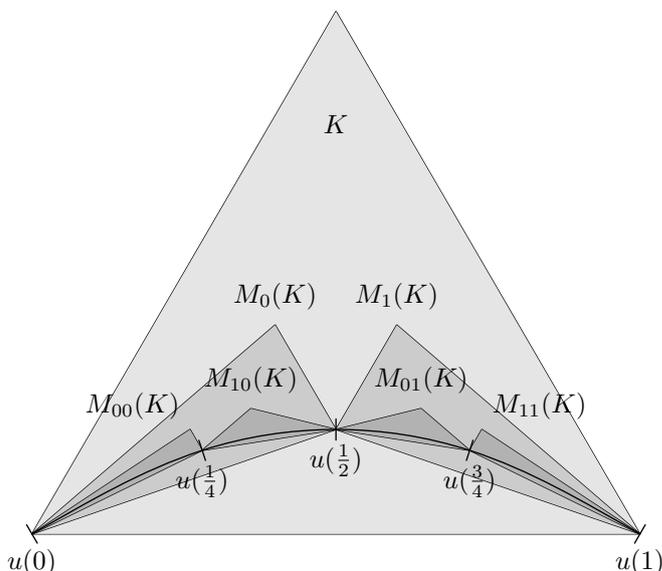


FIGURE IV.5.: Illustration de la proposition IV.3.1. La courbe qui traverse le dessin est la courbe paramétrée $u(t)$.

Démonstration. Soient $(a_i)_{i \geq 1}$ et $(a'_i)_{i \geq 1}$ deux suites à valeurs dans $\{0, 1\}$. Soit k le premier indice où ces suites diffèrent, i. e. supposons $a_1 = a'_1, \dots, a_{k-1} = a'_{k-1}, a_k \neq a'_k$. Sans perte de généralité, on peut supposer qu'en fait $a_k = 0$ et $a'_k = 1$. Choisissons une valeur arbitraire $u_0 \in K$. Alors les affirmations suivantes sont équivalentes (n'oublions pas que les M_i sont bijectives) :

$$\begin{aligned} \lim_{n \rightarrow \infty} M_{a_1 \dots a_n}(u_0) &= \lim_{n \rightarrow \infty} M_{a'_1 \dots a'_n}(u_0) \\ \bigcap_{n=0}^{\infty} M_{a_1 \dots a_n}(K) &= \bigcap_{n=0}^{\infty} M_{a'_1 \dots a'_n}(K) \\ \bigcap_{n=k}^{\infty} M_{0a_{k+1} \dots a_n}(K) &= \bigcap_{n=k}^{\infty} M_{1a'_{k+1} \dots a'_n}(K) \end{aligned}$$

Dans la dernière ligne, chaque côté est inclus dans le $M_i(K)$ respectif. Mais un rapide calcul (ou un rapide coup d'œil à la figure IV.5) montre que les deux images du triangle ont un et un seul point commun :

$$M_0(K) \cap M_1(K) = \left\{ \frac{1}{5} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \right\} = \{M_0(e_1)\} = \{M_1(e_0)\}.$$

L'égalité précédente a donc lieu si et seulement si

$$\begin{cases} \bigcap_{n=k}^{\infty} M_{a_{k+1} \dots a_n}(K) = \{e_1\} \\ \bigcap_{n=k}^{\infty} M_{a'_{k+1} \dots a'_n}(K) = \{e_0\}, \end{cases}$$

ce qui est vrai si et seulement si pour tout $i > k$, $a_i = 1$ et $a'_i = 0$. En d'autres termes, les deux suites donnent la même valeur si et seulement si on a $\sum_{i=1}^{\infty} a_i 2^{-i} = \sum_{i=1}^{\infty} a'_i 2^{-i}$. Le sens « seulement si » démontre la proposition, et le sens « si » confirme l'équivalence des différents développements binaires. \square

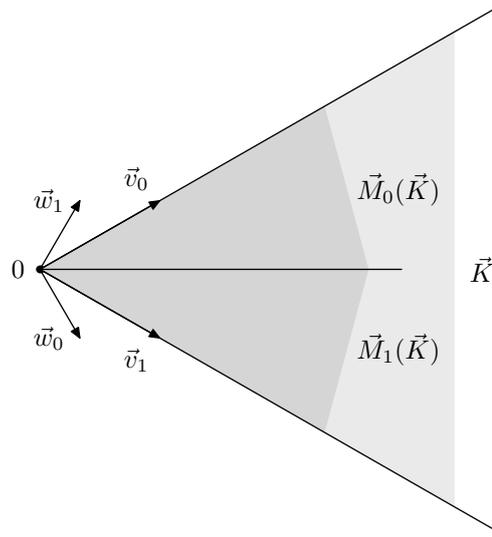


FIGURE IV.6.: Principaux objets du plan vectoriel $\vec{\mathcal{H}}$.

IV.4. Comportement local de u : existence et direction de la tangente

On va maintenant étudier le comportement du vecteur $u(t) - u(s)$ quand $t \rightarrow s$. Dans cette section, on va décrire complètement le comportement de la norme de ce vecteur : de façon quantitative par la formule (IV.4.1), et de façon qualitative dans le corollaire IV.4.3.

La différence $u(t) - u(s)$ vit dans $\vec{\mathcal{H}} := \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$, le plan vectoriel parallèle au plan affine \mathcal{H} . Ce plan vectoriel est bien entendu stable par M_0 et M_1 ; il est donc naturel d'introduire les restrictions de ces opérateurs à $\vec{\mathcal{H}}$, qu'on appellera \vec{M}_0 et \vec{M}_1 . Chaque \vec{M}_i a pour valeurs propres $\frac{3}{5}$ et $\frac{1}{5}$, avec les vecteurs propres respectifs \vec{v}_i et \vec{w}_i . Notez que dans les deux cas, \vec{v}_i et \vec{w}_i forment une base orthogonale de $\vec{\mathcal{H}}$.

On introduit maintenant l'ensemble

$$\vec{K} := \{a\vec{v}_0 + b\vec{v}_1 \mid a, b \geq 0 \text{ et } (a, b) \neq (0, 0)\}$$

(voir la figure IV.6.) Cette notation ne doit pas faire croire au lecteur que cet ensemble serait, d'une façon ou d'une autre, dérivé de K : il ne l'est pas, en tout cas pas directement. Mais il joue bel et bien un rôle analogue, et le lemme suivant va montrer précisément dans quel sens :

Lemme IV.4.1. *L'ensemble \vec{K} est stable par les \vec{M}_i . Pour tous $s, t \in [0, 1]$ tels que $s < t$, $u(t) - u(s) \in \vec{K}$.*

Démonstration. La première partie se vérifie par un calcul immédiat (voir aussi la figure IV.6).

Quant à la deuxième partie, montrons d'abord le résultat pour les valeurs rationnelles dyadiques de s et t (telles que $0 \leq s < t \leq 1$), par récurrence sur leurs valuations dyadiques. On pose $I_n := \mathcal{S}_n \cap [0, 1] = \{\frac{k}{2^n} \mid 0 \leq k \leq 2^n\}$, de sorte que $I_0 = \{0, 1\}$ et $I_{n+1} = h_0(I_n) \cup h_1(I_n)$. Si $s, t \in I_0$, alors $s = 0$, $t = 1$ et le résultat se vérifie par un calcul immédiat. Supposons maintenant que le résultat est vrai pour tous $s, t \in I_n$; soient $s < t$ deux points dans I_{n+1} . On doit distinguer trois cas : $s < t \leq \frac{1}{2}$, $s < \frac{1}{2} < t$ et $\frac{1}{2} \leq s < t$. On va traiter seulement le cas du milieu, qui est le plus difficile ; les deux autres sont analogues. On peut

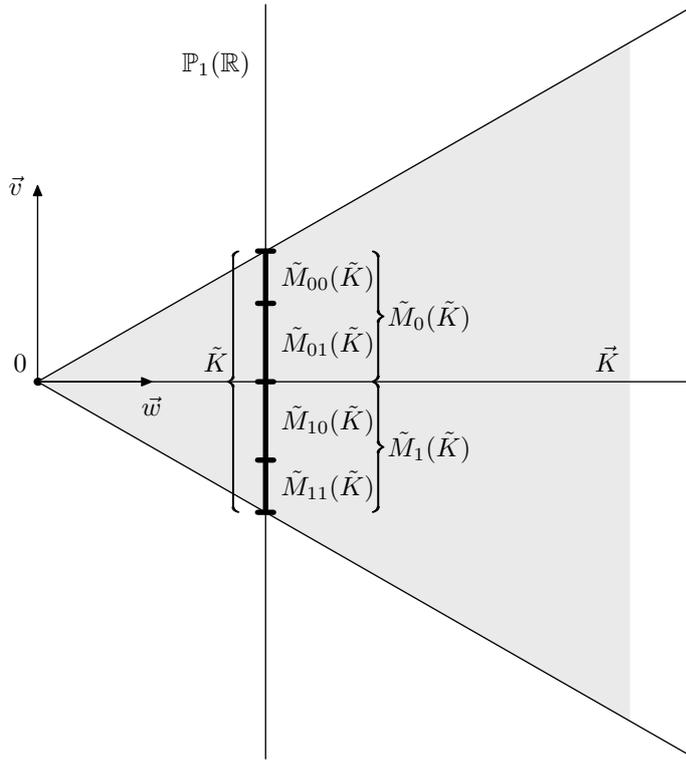


FIGURE IV.7.: Illustration de la proposition IV.4.2

écrire $s = h_0(S)$ et $t = h_1(T)$, de sorte qu'on a :

$$\begin{aligned}
 u(t) - u(s) &= u(t) - u(\frac{1}{2}) + u(\frac{1}{2}) - u(s) \\
 &= u(h_1(T)) - u(h_1(0)) + u(h_0(1)) - u(h_0(S)) \\
 &= \vec{M}_1(u(T) - u(0)) + \vec{M}_0(u(1) - u(S)).
 \end{aligned}$$

En appliquant, dans l'ordre, l'hypothèse de récurrence, le fait que \vec{K} est stable par les \vec{M}_i , et le fait que \vec{K} est stable par combinaison linéaire à coefficients strictement positifs (qui découle immédiatement de la formule qui le définit), on obtient le résultat.

Soient maintenant $0 \leq s < t \leq 1$ des valeurs quelconques. Comme u est continue et que les rationnels dyadiques sont denses, $u(t) - u(s)$ se trouve dans l'adhérence de \vec{K} , qui est égale à $\vec{K} \cup \{0\}$. Mais étant donné que, par le corollaire IV.3.2, u est injective, la différence ne peut pas valoir 0. \square

On a à présent besoin d'introduire les équivalents projectifs des objets pertinents. Soit $\pi : \vec{\mathcal{H}} \setminus \{0\} \rightarrow \mathbb{P}^1(\mathbb{R})$ la projection canonique ; on pose $\tilde{K} := \pi(\vec{K})$ et

$$\begin{aligned}
 \tilde{M}_i : \quad \mathbb{P}^1(\mathbb{R}) &\rightarrow \mathbb{P}^1(\mathbb{R}) \\
 \pi(\vec{x}) &\mapsto \pi(M_i(\vec{x}))
 \end{aligned}$$

(voir la figure IV.7). Clairement, \tilde{K} est stable par les \tilde{M}_i .

Proposition IV.4.2. *Soit $s \in [0, 1]$; soit $s = 0, a_1 a_2 \dots$ son développement binaire supérieur. Alors pour tout $\tilde{u}_0 \in \tilde{K}$, on a*

$$\lim_{\substack{t \rightarrow s \\ t > s}} \pi(u(t) - u(s)) = \lim_{n \rightarrow \infty} \tilde{M}_{a_1 \dots a_n}(\tilde{u}_0), \tag{IV.4.1}$$

c'est-à-dire que les deux limites existent et sont égales.

Quant aux développements inférieurs, la formule devient :

$$\lim_{\substack{t \rightarrow s \\ t < s}} \pi(u(s) - u(t)) = \lim_{n \rightarrow \infty} \tilde{M}_{a_1 \dots a_n}(\tilde{u}_0).$$

Démonstration. La preuve est assez similaire à la preuve de la proposition IV.3.1. Comme avant, les ensembles $\tilde{M}_{a_1 \dots a_n}(\tilde{K})$ sont tous fermés et forment une suite décroissante. Pour s'assurer que leur intersection est réduite à un point, il suffit de vérifier que leurs diamètres tendent vers 0; pour le faire, il suffit de montrer que les \tilde{M}_i sont contractantes. Dans la métrique usuelle (où la distance entre deux points est égale à l'angle entre les droites correspondantes), ce n'est pas le cas : la constante de Lipschitz vaut alors 1. On va donc utiliser une autre métrique — celle qui est définie par $d(\pi(\vec{u}_1), \pi(\vec{u}_2)) = |\frac{a_1}{b_1} - \frac{a_2}{b_2}|$, où on pose $\vec{u}_i = a_i \vec{v} + b_i \vec{w}$ avec $\vec{v} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$ et $\vec{w} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}$. Moralement, cela revient à projeter sur une droite et non pas sur un cercle : voir la figure IV.7. Comme il existe un voisinage du point à l'infini qui ne rencontre pas \tilde{K} , cette métrique est bien définie sur \tilde{K} et y induit la bonne topologie. Quelques lignes de calcul montrent alors que dans cette métrique, \tilde{M}_0 et \tilde{M}_1 ont tous les deux leur constante de Lipschitz égale à $\frac{3}{4}$. On peut donc écrire

$$\forall \tilde{u}_0 \in \tilde{K}, \quad \left\{ \lim_{n \rightarrow \infty} \tilde{M}_{a_1 \dots a_n}(\tilde{u}_0) \right\} = \bigcap_{n=0}^{\infty} \tilde{M}_{a_1 \dots a_n}(\tilde{K}).$$

Soit maintenant $n \in \mathbb{N}$. Alors pour tout t suffisamment proche, mais strictement supérieur, à s , les n premiers bits des développements binaires de s et t coïncident (c'est là qu'on utilise l'hypothèse que le développement binaire est supérieur), de sorte qu'on a $s = h_{a_1 \dots a_n}(S)$ et $t = h_{a_1 \dots a_n}(T)$ pour certains réels $S, T \in [0, 1]$. Par le lemme IV.4.1, on a

$$\pi(u(t) - u(s)) = \tilde{M}_{a_1 \dots a_n}(\pi(u(T) - u(S))) \in \tilde{M}_{a_1 \dots a_n}(\tilde{K}),$$

d'où le résultat.

Quant au cas symétrique, la preuve est complètement analogue. □

Corollaire IV.4.3. *La quantité*

$$\vec{D}u(s) := \lim_{\substack{t \rightarrow s \\ t \in [0, 1]}} \pi(u(t) - u(s)) \operatorname{sgn}(t - s)$$

est une fonction de s bien définie, injective et continue.

En termes géométriques, $\vec{D}u(s)$ est le vecteur directeur de la tangente à la courbe que constitue l'image de u . On peut interpréter ce résultat comme le fait que cette courbe est C^1 et convexe — ce qui, après tout, n'est pas si surprenant si on pense à cette courbe comme à la forme que prend le bord d'un triangle de Sierpinski élastique tendu (voir la figure IV.3). Notez que la *paramétrisation* de la courbe est bien moins régulière — comme on le verra, elle n'est même pas toujours dérivable.

Démonstration. Étant donné deux suites $(a_i)_{i \geq 1}$ et $(a'_i)_{i \geq 1}$ à valeurs dans $\{0, 1\}$, on a

$$\bigcap_{n=0}^{\infty} \tilde{M}_{a_1 \dots a_n}(\tilde{K}) = \bigcap_{n=0}^{\infty} \tilde{M}_{a'_1 \dots a'_n}(\tilde{K})$$

si et seulement si les deux suites sont des développements binaires d'un même nombre. La preuve n'en fait que suivre, *mutatis mutandis*, celle du corollaire IV.3.2; on a besoin du fait que les \tilde{M}_i sont injectives et du fait que $\tilde{M}_0(\tilde{K})$ et $\tilde{M}_1(\tilde{K})$ se coupent en un et un

seul point, qui se vérifient facilement. Combiné à la proposition précédente, le sens « si » montre que la fonction est bien définie, et le sens « seulement si » — qu'elle est injective.

Quant à la continuité, soit $\varepsilon > 0$. Comme \tilde{M}_0 et \tilde{M}_1 sont contractantes (dans la bonne métrique), on peut trouver un nombre n tel que le diamètre de n'importe quel $\tilde{M}_{a_1 \dots a_n}(\tilde{K})$ soit inférieur à $\frac{\varepsilon}{2}$. Soient maintenant $s, s' \in [0, 1]$ tels que $|s - s'| \leq 2^{-n}$. Soient $s = 0, a_1 \dots a_n \dots$ et $s' = 0, a'_1 \dots a'_n \dots$ leurs développements binaires; soit $w = a_1 \dots a_n$, de sorte que $\vec{D}u(s) \in \tilde{M}_w(\tilde{K})$, et de même w' . Alors deux cas sont possibles : ou bien $w = w'$, ou bien w et w' sont les représentations binaires de deux entiers consécutifs; en d'autres termes

$$\left| \sum_{i=1}^n a_i 2^{n-i} - \sum_{i=1}^n a'_i 2^{n-i} \right| \leq 1.$$

Cependant, sachant que $\tilde{M}_0(\pi(\vec{w}_0)) = \pi(\vec{w}_0)$, $\tilde{M}_1(\pi(\vec{w}_1)) = \pi(\vec{w}_1)$ et $\tilde{M}_1(\pi(\vec{w}_0)) = \tilde{M}_0(\pi(\vec{w}_1))$, une récurrence immédiate montre que dans le deuxième cas, $\tilde{M}_w(\tilde{K})$ et $\tilde{M}_{w'}(\tilde{K})$ ont un point commun. Ainsi, la distance entre $\vec{D}u(s)$ et $\vec{D}u(s')$ est inférieure ou égale à ε , ce qui conclut. \square

IV.5. Comportement local de u et de $u_a^c b$: taux de variation

IV.5.1. Formule pour l'exposant de Hölder

Dans la section précédente, on a vu que quand t tend vers s , la direction de $u(t) - u(s)$ se comporte plutôt bien. On va voir maintenant que sa norme a un comportement beaucoup plus chaotique. Une approche naïve consisterait à essayer simplement de calculer la dérivée $u'(s)$; cependant, on verra dans la proposition IV.5.12 que quand cette dérivée existe, elle vaut soit 0 soit ∞ : ce n'est donc pas le paramètre le plus pertinent ici.

On introduit les paramètres suivants, qui vont nous fournir une description plus précise du comportement local de u :

Définition IV.5.1. Soit $s \in [0, 1]$.

– L'*exposant de Hölder supérieur* de u en s est

$$\alpha_{\text{sup}}(u, s) := \limsup_{\substack{t \rightarrow s \\ t \in [0, 1]}} \frac{\ln \|u(t) - u(s)\|}{\ln |t - s|}.$$

– L'*exposant de Hölder inférieur* de u en s est

$$\alpha_{\text{inf}}(u, s) := \liminf_{\substack{t \rightarrow s \\ t \in [0, 1]}} \frac{\ln \|u(t) - u(s)\|}{\ln |t - s|}.$$

– Quand les deux sont égaux, on les appellera simplement l'*exposant de Hölder* de u en s , et on écrira

$$\alpha(u, s) := \lim_{\substack{t \rightarrow s \\ t \in [0, 1]}} \frac{\ln \|u(t) - u(s)\|}{\ln |t - s|}.$$

Une fois qu'on connaît l'exposant de Hölder de u , on peut en général calculer la dérivée de u , puis celle de $u_a^c b$. En effet :

Lemme IV.5.2. Soit $s \in [0, 1]$. Alors on a :

- $\alpha_{\text{inf}}(u, s) > 1 \implies u'(s) = 0 \implies \alpha_{\text{inf}}(u, s) \geq 1$;
- $\alpha_{\text{sup}}(u, s) < 1 \implies u'(s) = \infty \implies \alpha_{\text{sup}}(u, s) \leq 1$,

où la dernière égalité est un abus de notation pour dire « $\lim_{t \rightarrow s} \frac{\|u(t) - u(s)\|}{|t - s|} = \infty$ ».

Démonstration. C'est un exercice élémentaire d'analyse réelle. \square

Lemme IV.5.3. Soit $\Phi = (a \ b \ c)$ une forme linéaire, et soit $s \in [0, 1]$ tel que $\vec{D}u(s) \notin \ker \Phi$. On a alors :

- $(u_a^c)_b'(s) = 0$ si et seulement si $u'(s) = 0$;
- $(u_a^c)_b'(s) = \infty$ si et seulement si $u'(s) = \infty$;
- $(u_a^c)_b'(s)$ n'est pas définie si et seulement si $u'(s)$ n'est pas définie.

Démonstration. C'est essentiellement une application de la formule de la dérivée d'une fonction composée. \square

Notez que d'après le corollaire IV.4.3, pour une valeur donnée de Φ , ce lemme s'applique à toutes les valeurs de s sauf au plus une. En particulier, il est facile de vérifier que pour les fonctions ϕ , ψ , χ et ξ , ce lemme s'applique avec seulement deux exceptions : $\chi'(0)$ et $\xi'(1)$ (et Kirillov montre dans [13] que les deux valent 0.)

Voici la formule qui permet de calculer l'exposant de Hölder :

Proposition IV.5.4. Soit $s \in [0, 1]$, $s = 0, a_1 a_2 \dots$ son développement binaire. Alors on a

$$\begin{aligned} \alpha_{\text{inf}}(u, s) &= \liminf_{n \rightarrow \infty} \frac{\ln \|\vec{M}_{a_1 \dots a_n}\|}{n \ln \left(\frac{1}{2}\right)}; \\ \alpha_{\text{sup}}(u, s) &= \limsup_{n \rightarrow \infty} \frac{\ln \|\vec{M}_{a_1 \dots a_n}\|}{n \ln \left(\frac{1}{2}\right)} \end{aligned} \tag{IV.5.1}$$

(comparer cette formule à (IV.3.1) et (IV.4.1)). Pour la démontrer, on aura besoin du lemme suivant :

Lemme IV.5.5. Soit f un automorphisme linéaire de \mathbb{R}^2 , soit \mathcal{C} un cône fermé saillant qui est stable par f , soit enfin $\mathcal{C}' \subset \text{int } \mathcal{C}$ tel que $\mathcal{C}' \cup \{0\}$ est aussi un cône fermé. Alors pour tout $\vec{x} \in \mathcal{C}'$, le rapport

$$\frac{\|f(\vec{x})\|}{\|f\| \|\vec{x}\|}$$

est positivement borné par des constantes qui ne dépendent que de \mathcal{C} et de \mathcal{C}' .

Démonstration. Choisissons une base dans laquelle \mathcal{C} devient le quadrant supérieur droit. On peut alors écrire :

- $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\}$;
- $f = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, où les coefficients a, b, c et d sont tous positifs ;
- $\mathcal{C}' = \{(x, y) \in \mathcal{C} \setminus \{0\} \mid m \leq \frac{x}{y} \leq M\}$, où m et M sont des bornes finies et strictement positives.

Soit $\vec{x} = (x, y) \in \mathcal{C}'$. Utilisons la norme infinie : on a $\|\vec{x}\|_\infty = \max(x, y)$, $\|f\|_\infty = \max(a, b, c, d)$ et $\|f(\vec{x})\|_\infty = \max(ax + by, cx + dy)$. Mais on a, d'un côté

$$\max(ax + by, cx + dy) \leq 2 \max(a, b, c, d) \max(x, y),$$

et d'un autre côté

$$\begin{aligned} \max(ax + by, cx + dy) &\geq \max(a, b, c, d) \min(x, y) \\ &\geq \mu \max(a, b, c, d) \max(x, y), \end{aligned}$$

où $\mu = \min(m, M^{-1})$ ne dépend que de \mathcal{C} et \mathcal{C}' , conformément à ce qui avait été annoncé. \square

Soit $\vec{K}' = \{a\vec{w}_0 + b\vec{w}_1 \mid a, b \geq 0\}$ (sur la figure IV.6, c'est le cône qui est délimité par \vec{w}_0 et \vec{w}_1 , non représenté pour alléger la figure). Alors il est facile de vérifier qu'il est stable par \vec{M}_0 et \vec{M}_1 , et donc aussi par tous leurs produits. En appliquant le lemme à $\mathcal{C} = \vec{K}'$ et à $\mathcal{C}' = \vec{K}$ (attention aux notations légèrement incohérentes), on obtient l'estimation suivante :

Corollaire IV.5.6. *Pour tout mot fini w sur l'alphabet $\{0, 1\}$ et pour tout $\vec{u}_0 \in \vec{K}$, le rapport*

$$\frac{\|\vec{M}_w(\vec{u}_0)\|}{\|\vec{M}_w\| \|\vec{u}_0\|}$$

est positivement borné.

Démonstration de la proposition IV.5.4. Soit maintenant $s \in [0, 1]$, et soit $s' \in [0, 1]$ un point proche, mais distinct, de s . Soit $n = \left\lfloor \frac{\ln|s'-s|}{\ln(\frac{1}{2})} \right\rfloor$, de sorte que $\frac{1}{2} \leq 2^n|s' - s| \leq 1$. Soit $s = 0, a_1 \dots a_n \dots$ et $s' = 0, a'_1 \dots a'_n \dots$ leurs développements binaires respectifs. On pose $w = a_1 \dots a_n$ et $w' = a'_1 \dots a'_n$; on peut alors écrire $s = h_w(S)$ et $s' = h_{w'}(S')$, pour certaines valeurs $S, S' \in [0, 1]$. On en déduit :

$$\begin{aligned} \frac{\ln \|u(s') - u(s)\|}{\ln |s' - s|} &= \frac{\ln \|u(h_{w'}(S')) - u(h_w(S))\|}{n \ln(\frac{1}{2}) + O(1)} \\ &= \frac{\ln \|M_{w'}(u(S')) - M_w(u(S))\|}{n \ln(\frac{1}{2}) + O(1)}. \end{aligned} \quad (\text{IV.5.2})$$

Deux cas sont alors possibles :

- Supposons $w' = w$. Alors d'après le corollaire précédent, on a

$$\begin{aligned} \ln \|M_{w'}(u(S')) - M_w(u(S))\| &= \ln \|\vec{M}_w(u(S') - u(S))\| \\ &= \ln \|\vec{M}_w\| + \ln \|u(S') - u(S)\| + O(1). \end{aligned} \quad (\text{IV.5.3a})$$

- Dans l'autre cas, par la définition de n , w et w' sont forcément les écritures binaires de deux entiers consécutifs. Sans perte de généralité, on peut supposer que $s < s'$: on peut alors écrire $w = a_0 \dots a_k 0 1 \dots 1$ et $w' = a_0 \dots a_k 1 0 \dots 0$ (où les deux groupes de bits finaux ont la même longueur, éventuellement nulle). On a alors $h_{w'}(0) = h_w(1)$, d'où

$$\ln \|M_{w'}(u(S')) - M_w(u(S))\| = \ln \|\vec{M}_{w'}(u(S') - u(0)) + \vec{M}_w(u(1) - u(S))\|.$$

Or il est facile de voir que, étant donné deux vecteurs \vec{u}, \vec{v} qui varient dans un cône saillant, le rapport entre $\|\vec{u} + \vec{v}\|$ et $\|\vec{u}\| + \|\vec{v}\|$ est positivement borné. Comme les deux termes de la somme ci-dessus appartiennent à \vec{K} , on peut écrire, en utilisant une fois de plus le corollaire précédent :

$$\begin{aligned} \ln \|M_{w'}(u(S')) - M_w(u(S))\| &= \\ &= \ln \left(\|\vec{M}_{w'}(u(S') - u(0))\| + \|\vec{M}_w(u(1) - u(S))\| \right) + O(1) \\ &= \max \left(\ln \|\vec{M}_{w'}(u(S') - u(0))\|, \ln \|\vec{M}_w(u(1) - u(S))\| \right) + O(1) \\ &= \max \left(\ln \|\vec{M}_{w'}\| + \ln \|u(S') - u(0)\|, \ln \|\vec{M}_w\| + \ln \|u(1) - u(S)\| \right) + O(1). \end{aligned}$$

Remarquez maintenant qu'on a

$$\vec{M}_w(\vec{v}_0) = \vec{M}_{w'}(\vec{v}_1).$$

En effet, comme $w = a_0 \dots a_k 0 1 \dots 1$ et $w' = a_0 \dots a_k 1 0 \dots 0$, l'identité découle des égalités $\vec{M}_0(\vec{v}_0) = \frac{3}{5}\vec{v}_0$, $\vec{M}_1(\vec{v}_1) = \frac{3}{5}\vec{v}_1$ et $\vec{M}_0(\vec{v}_1) = \vec{M}_1(\vec{v}_0)$, dont la vérification est immédiate (voir aussi la figure IV.6). Comme $\vec{v}_{0,1} \in \vec{K}$, on obtient

$$\ln \|\vec{M}_w\| + \ln \|\vec{v}_0\| + O(1) = \ln \|\vec{M}_{w'}\| + \ln \|\vec{v}_1\| + O(1),$$

ou, en d'autres termes, puisque $\|\vec{v}_0\| = \|\vec{v}_1\|$,

$$\ln \|\vec{M}_{w'}\| = \ln \|\vec{M}_w\| + O(1).$$

Cela nous permet d'écrire

$$\begin{aligned} \ln \|M_{w'}(u(S')) - M_w(u(S))\| &= & \text{(IV.5.3b)} \\ &= \ln \|\vec{M}_w\| + \max \left(\ln \|u(S') - u(0)\|, \ln \|u(1) - u(S)\| \right) + O(1). \end{aligned}$$

Vérifions maintenant que dans les deux cas, le terme du milieu est borné. Comme $[0, 1]$ est compact et que u est continue, il est clairement majoré. Pour l'autre inégalité, notez que, dans le premier cas, on a $(S' - S) = 2^n(s' - s) \geq \frac{1}{2} > \frac{1}{4}$; et dans le second cas, on a $(S' - 0) + (1 - S) = 2^n(s' - s) \geq \frac{1}{2}$, donc ou bien $(S' - 0) \geq \frac{1}{4}$ ou bien $(1 - S) \geq \frac{1}{4}$. Considérons maintenant l'ensemble

$$X := \left\{ (s_1, s_2) \in [0, 1]^2 \mid |s_1 - s_2| \geq \frac{1}{4} \right\}$$

et la fonction

$$\begin{aligned} X &\rightarrow \mathbb{R}_{\geq 0}, \\ (s_1, s_2) &\mapsto \|u(s_1) - u(s_2)\|. \end{aligned}$$

Comme elle est continue et que X est compact, elle atteint son minimum. Mais par le corollaire IV.3.2, elle ne peut pas s'annuler; elle est donc positivement bornée. On en déduit que $\|u(S') - u(S)\|$ dans le premier cas, et $\max(\|u(S') - u(0)\|, \|u(1) - u(S)\|)$ dans le second cas, est positivement borné.

On peut donc rassembler (IV.5.3a) et (IV.5.3b), et écrire :

$$\ln \|M_{w'}(u(S')) - M_w(u(S))\| = \ln \|\vec{M}_w\| + O(1),$$

et en injectant ceci dans (IV.5.2), on obtient le résultat. \square

IV.5.2. Cas des points rationnels

Théorème IV.5.7. *Soit $s \in [0, 1] \cap \mathbb{Q}$. On sait alors que son développement binaire $s = 0, a_1 a_2 \dots$ est ultimement périodique; soit n la longueur de sa période, $p = a_k \dots a_{k+n-1}$ sa période. Alors \vec{M}_p a des valeurs propres réelles, strictement positives et distinctes, et on a*

$$\alpha(u, s) = \frac{\ln \lambda_p}{n \ln \frac{1}{2}},$$

où λ_p est la plus grande valeur propre de \vec{M}_p .

Concrètement, comme on connaît déjà le déterminant de \vec{M}_p , on peut calculer λ_p par la formule

$$\lambda_p = \frac{1}{2} \left(\text{tr } \vec{M}_p + \sqrt{(\text{tr } \vec{M}_p)^2 - \left(\frac{3}{25}\right)^n} \right).$$

Notez que la proposition IV.5.2 permet d'en déduire $u'(s)$. En effet, on verra dans la section suivante que le cas $\alpha(u, s) = 1$ ne se produit jamais.

Pour la démonstration, on aura besoin du lemme suivant :

Lemme IV.5.8. Soit $s \in [0, 1]$, $w = a_1 \dots a_n$ un mot sur l'alphabet $\{0, 1\}$. Alors on a :

$$\alpha_{\text{inf}}(u, h_w(s)) = \alpha_{\text{inf}}(u, s);$$

$$\alpha_{\text{sup}}(u, h_w(s)) = \alpha_{\text{sup}}(u, s).$$

Démonstration. Soit $s = 0, b_1 b_2 \dots$ un développement binaire ; soit $\vec{u}_0 \in \vec{K}$. Il est bien connu que $\frac{\|\vec{M}_w(\vec{u})\|}{\|\vec{u}\|}$ est borné quand \vec{u} varie ; sachant que \vec{M}_w est inversible, l'inverse de cette quantité est également borné ; donc la quantité

$$\ln \|\vec{M}_{a_1 \dots a_k b_1 \dots b_n}(\vec{u}_0)\| - \ln \|\vec{M}_{b_1 \dots b_n}(\vec{u}_0)\|$$

est bornée. Le résultat découle immédiatement de la proposition IV.5.4. \square

Démonstration du théorème IV.5.7. Grâce au lemme qu'on vient de démontrer, on sait qu'on peut se débarrasser de la préperiode du développement de s ; on peut donc supposer que celui-ci est purement périodique. De la formule (IV.5.1) il découle :

$$\alpha_{\text{inf}}(u, s) = \alpha_{\text{sup}}(u, s) = \lim_{k \rightarrow \infty} \frac{\ln \|\vec{M}_p^k\|}{kn \ln \left(\frac{1}{2}\right)}$$

Comme \vec{K} est stable par \vec{M}_p , on sait aussi que \vec{K} est stable par \vec{M}_p . Puisque \vec{M}_p est contractante, elle a donc un et un seul point fixe dans \vec{K} , ce qui veut dire que \vec{M}_p a un unique vecteur propre \vec{v}_p qui se trouve dans \vec{K} et qui soit associé à une valeur propre strictement positive λ_p (on justifiera dans un instant que c'est effectivement la plus grande). Soit μ_p l'autre valeur propre ; comme $\det(\vec{M}_p) = \left(\frac{3}{5}\right)^n > 0$, elle est également strictement positive, donc le vecteur propre correspondant \vec{u}_p ne peut pas se trouver dans \vec{K} . En utilisant à nouveau le fait que $\vec{M}_p(\vec{K})$ est strictement inclus dans \vec{K} et en considérant la disposition des vecteurs concernés les uns par rapport aux autres et par rapport à \vec{K} , il est facile de voir que $\lambda_p > \mu_p$. Il en découle que $\|\vec{M}_p^k\| \sim \lambda_p^k$, d'où le résultat. \square

Exemple IV.5.9. Il découle du théorème IV.5.7 que si s est rationnel dyadique, alors $\alpha(u, s) = \frac{\ln \frac{3}{5}}{\ln \frac{1}{2}} \approx 0,737$ (en effet, dans ce cas, λ_p est simplement la plus grande valeur propre de \vec{M}_0 ou \vec{M}_1 , à savoir $\frac{3}{5}$). En particulier, on a dans ce cas $u'(s) = \infty$.

Notez que le lemme IV.5.8 est intéressant en soi, même en dehors du cas $s \in \mathbb{Q}$: il montre que l'exposant de Hölder ne change pas quand on change un nombre fini de bits dans le développement binaire. Par souci de complétude, mentionnons aussi au passage la symétrie suivante de u (bien qu'elle soit assez évidente). Soit

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(dans la figure IV.5, cela correspond simplement à l'axe de symétrie vertical). Alors

Proposition IV.5.10. Pour tout $s \in [0, 1]$, on a $u(1 - s) = Pu(s)$.

Corollaire IV.5.11. Pour tout $s \in [0, 1]$, $\alpha_{\text{inf}}(u, 1 - s) = \alpha_{\text{inf}}(u, s)$ et $\alpha_{\text{sup}}(u, 1 - s) = \alpha_{\text{sup}}(u, s)$.

Démonstration. Clairement, $P^2 = 1$, $M_1 = PM_0P^{-1} = PM_0P$, $M_0 = PM_1P$, et K est stable par P . Soit $s = 0, a_1a_2 \dots$ un développement binaire ; alors, en utilisant simplement (IV.3.1), on obtient :

$$\begin{aligned} u(1-s) &= u(0, \bar{a}_1\bar{a}_2 \dots) \\ &= \lim_{n \rightarrow \infty} M_{\bar{a}_1 \dots \bar{a}_n} u_0 \\ &= \lim_{n \rightarrow \infty} PM_{a_1 \dots a_n} Pu_0 \\ &= Pu(s), \end{aligned}$$

où \bar{a} est un raccourci pour dire $1-a$ et $u_0 \in K$ est un point de départ arbitraire. \square

IV.5.3. Dérivée de u

Théorème IV.5.12. *Pour tout $s \in [0, 1]$, $u'(s)$ est soit égale à 0, soit égale à ∞ , soit n'est pas définie. Si $s \in \mathbb{Q}$, seuls les deux premiers cas sont possibles.*

Démonstration. Par l'absurde : soit s un nombre dans $[0, 1]$ tel que $u'(s)$ existe, est finie et non nulle. On a alors en particulier

$$\lim_{t \rightarrow s} \frac{\|u(t) - u(s)\|}{|t - s|} = C$$

pour une certaine constante C , où $0 < C < \infty$.

Soit $s = 0, a_1a_2 \dots$ un développement binaire. Soit t_n une suite à valeurs dans $[0, 1]$ et telle que pour tout $n \in \mathbb{N}$, le développement binaire de t_n (disons au moins un des deux) coïncide avec celui de s dans les n premiers bits. Une telle suite converge évidemment vers s . On a alors, pour tout $n \in \mathbb{N}$:

$$\frac{\|u(t_n) - u(s)\|}{|t_n - s|} = \frac{\|\vec{M}_{a_1 \dots a_n}(u(T_n) - u(S_n))\|}{2^{-n}|T_n - S_n|},$$

où $S_n = h_{a_1 \dots a_n}^{-1}(s) = 0, a_{n+1}a_{n+2} \dots$ et $T_n = h_{a_1 \dots a_n}^{-1}(t_n)$. Comme $[0, 1]$ est compact, on peut extraire une suite $\phi(n)$ telle que $S_{\phi(n)}$ converge vers une certaine valeur S_∞ .

Supposons que $S_\infty < 1$ (par symétrie, on peut traiter de même le cas $S_\infty > 0$). Désormais, on impose que $t_n > s$. Par le corollaire IV.5.6, on sait que

$$\frac{\|\vec{M}_{a_1 \dots a_n}\| \|u(T_n) - u(S_n)\|}{\|\vec{M}_{a_1 \dots a_n}(u(T_n) - u(S_n))\|}$$

est positivement borné. On sait également que

$$\frac{\|\vec{M}_{a_1 \dots a_n}(u(T_n) - u(S_n))\|}{2^{-n}|T_n - S_n|}$$

tend vers C : c'est donc aussi positivement borné. Supposons un instant que t_n est choisi de telle sorte que T_n est constant égal à 1. Alors pour n suffisamment grand, $S_{\phi(n)}$ varie dans un intervalle fermé qui ne contient pas 1, et puisque u est continue et injective, $\frac{\|u(T_{\phi(n)}) - u(S_{\phi(n)})\|}{|T_{\phi(n)} - S_{\phi(n)}|}$ est positivement borné.

Cela signifie que $\frac{\vec{M}_{a_1 \dots a_{\phi(n)}}}{2^{-\phi(n)}}$ est positivement borné en norme, et que donc cette expression varie dans un espace compact. On peut donc extraire de $\phi(n)$ encore une sous-suite, qu'on appellera $\psi(n)$, telle que $\frac{\vec{M}_{a_1 \dots a_{\psi(n)}}}{2^{-\psi(n)}}$ converge ; appelons \vec{M}_∞ sa limite. Notez que, bien qu'on

ait utilisé une valeur particulière de t_n pour montrer que cette construction était possible, les objets eux-mêmes ne dépendent pas de t_n .

Soit maintenant T un point quelconque dans $]S_\infty, 1]$; on pose $t_n = h_{a_1 \dots a_n}(T)$, de sorte que $T_n = T$ pour tout n . On a alors :

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\|u(t_{\psi(n)}) - u(s)\|}{|t_{\psi(n)} - s|} &= \lim_{n \rightarrow \infty} \frac{\|\vec{M}_{a_1 \dots a_{\psi(n)}}(u(T) - u(S_{\psi(n)}))\|}{2^{-\psi(n)}|T - S_{\psi(n)}|} \\ &= \frac{\|\vec{M}_\infty(u(T) - u(S_\infty))\|}{|T - S_\infty|}, \end{aligned}$$

ce qui signifie que ce dernier rapport vaut toujours C indépendamment de la valeur de T .

On sent bien que cette affirmation semble trop forte pour être vraie, et on va bientôt montrer qu'elle est absurde. Notez que

$$\det \left(\frac{\vec{M}_{a_1 \dots a_{\psi(n)}}}{2^{-\psi(n)}} \right) = \det \left(2\vec{M}_i \right)^{\psi(n)} = \left(\frac{12}{25} \right)^{\psi(n)},$$

ce qui tend vers 0 quand n augmente. Donc $\det(\vec{M}_\infty) = 0$, en d'autres termes, \vec{M}_∞ est de rang 1. L'affirmation peut alors se reformuler de la façon suivante : il existe une forme linéaire Φ telle que le rapport

$$\frac{\Phi(u(T) - u(S_\infty))}{T - S_\infty}$$

soit constant. L'intervalle $]S_\infty, T[$ contient une infinité de points rationnels dyadiques; d'après le corollaire IV.4.3, on peut trouver dans cet intervalle une valeur T' rationnelle dyadique telle que $\Phi(\vec{D}u(T')) \neq 0$. Mais l'exemple IV.5.9 combiné avec le lemme IV.5.3 montrent que dans ce cas, $\Phi(u(T) - u(S_\infty))$, en tant que fonction de T , doit avoir une dérivée infinie au point T' . Le rapport tout entier doit donc également avoir une dérivée infinie en ce point, ce qui est absurde.

Soit maintenant, par l'absurde, $s \in [0, 1] \cap \mathbb{Q}$ tel que u n'est pas dérivable en s . Appliquons le théorème IV.5.7 : ce n'est possible que si $\alpha(u, s) = 1$. Soit p la période du développement binaire de s , n la longueur de la période; on a alors $\lambda_p = (\frac{1}{2})^n$. On sait que M_p a pour déterminant $(\frac{3}{25})^n$ et que l'une de ses valeurs propres vaut 1. On a donc $\text{tr } M_p = 1 + (\frac{1}{2})^n + (\frac{6}{25})^n$, et $\text{tr } 25^n M_p = 25^n + 6^n + (\frac{25}{2})^n$. D'autre part, on sait que dans la base canonique, $25M_0$ et $25M_1$ sont données par des matrices à coefficients entiers, il en est donc de même pour $25^n M_p$, et donc $\text{tr } 25^n M_p$ doit être entier. Contradiction. \square

Corollaire IV.5.13. *Pour Lebesgue-presque tout $s \in [0, 1]$, $u'(s) = 0$.*

Démonstration. Soit Φ une forme linéaire dont le noyau ne rencontre pas \vec{K} ; il est alors facile de déduire du corollaire IV.4.3 que la fonction $\Phi \circ u$ est monotone. Mais c'est un résultat bien connu (voir par exemple [18], Theorem 8.19) qu'une fonction monotone est dérivable (au sens usuel de « a une dérivée finie ») presque partout. Comme, d'après le lemme IV.5.3, $\Phi \circ u$ se comporte de la même façon que u , on peut conclure. \square

IV.5.4. Valeurs numériques et estimations de l'exposant de Hölder

Voici la table de toutes les périodes possibles de longueur au plus 7 (à permutation cyclique près), ordonnées par leurs exposants de Hölder :

| s | Période p | Longueur n | $5^n \text{tr } \vec{M}_p$ | Valeur approchée de α |
|------------------|-------------|--------------|----------------------------|------------------------------|
| $\frac{1}{3}$ | 01 | 2 | 7 | 1,119 |
| $\frac{21}{127}$ | 0010101 | 7 | 388 | 1,096 |
| $\frac{11}{63}$ | 001011 | 6 | 175 | 1,086 |
| $\frac{5}{31}$ | 00101 | 5 | 76 | 1,085 |
| $\frac{1}{5}$ | 0011 | 4 | 34 | 1,078 |
| $\frac{19}{127}$ | 0010011 | 7 | 436 | 1,072 |
| $\frac{11}{127}$ | 0001011 | 7 | 472 | 1,055 |
| $\frac{13}{127}$ | 0001101 | 7 | 472 | 1,055 |
| $\frac{1}{7}$ | 001 | 3 | 16 | 1,050 |
| $\frac{3}{31}$ | 00011 | 5 | 88 | 1,040 |
| $\frac{5}{63}$ | 000101 | 6 | 211 | 1,039 |
| $\frac{1}{9}$ | 000111 | 6 | 223 | 1,025 |
| $\frac{9}{127}$ | 0001001 | 7 | 580 | 1,012 |
| $\frac{5}{127}$ | 0000101 | 7 | 616 | 0,999 |
| $\frac{1}{21}$ | 000011 | 6 | 250 | 0,997 |
| $\frac{7}{127}$ | 0000111 | 7 | 628 | 0,995 |
| $\frac{1}{15}$ | 0001 | 4 | 43 | 0,982 |
| $\frac{3}{127}$ | 0000011 | 7 | 736 | 0,962 |
| $\frac{1}{31}$ | 00001 | 5 | 124 | 0,936 |
| $\frac{1}{63}$ | 000001 | 6 | 367 | 0,903 |
| $\frac{1}{127}$ | 0000001 | 7 | 1096 | 0,880 |
| 0 | 0 | 1 | 4 | 0,737 |

Quand on regarde cette table, il devient clair que globalement, les nombres dont les développements binaires contiennent de longs blocs de bits consécutifs égaux ont tendance à avoir des exposants de Hölder plus petits, tandis que ceux dont les bits sont plus « mélangés » ont tendance à avoir des exposants plus grands. Les plus petits exposants apparaissent quand le développement binaire est constant à partir d'un certain rang, c'est-à-dire pour les rationnels dyadiques — voir plus bas. Les plus grands exposants semblent apparaître quand les zéros et les uns sont strictement en alternance, c'est-à-dire pour $\frac{1}{3}$ et ses multiples par les rationnels dyadiques ; l'auteur ne sait pas le démontrer. L'auteur a réussi à démontrer l'affirmation suivante, qui confirme en partie cette intuition :

Proposition IV.5.14. *Soit $s \in [0, 1]$, $s = 0, a_1 a_2 \dots$ son développement binaire. On pose*

$$d_{\text{inf}, \text{sup}}(s) := \liminf_{n \rightarrow \infty} \sup \frac{1}{n} \sum_{i=1}^n |a_i - a_{i+1}|.$$

On a alors :

$$\alpha_{\text{inf}}(u, s) \leq \frac{\ln \frac{3}{5}}{\ln \frac{1}{2}} + d_{\text{inf}}(s);$$

$$\alpha_{\text{sup}}(u, s) \leq \frac{\ln \frac{3}{5}}{\ln \frac{1}{2}} + d_{\text{sup}}(s).$$

Moralement, $d_{\text{inf}}(s)$ et $d_{\text{sup}}(s)$ encadrent la proportion asymptotique de positions où le bit change de valeur ; cette dernière peut aussi se comprendre comme l'inverse de la longueur moyenne asymptotique des blocs de bits consécutifs égaux.

Corollaire IV.5.15. *Si $d_{\text{sup}}(s) < 1 - \frac{\ln \frac{3}{5}}{\ln \frac{1}{2}} \approx 0,263 \approx \frac{1}{3,802}$, alors $u'(s) = \infty$.*

Avant de démontrer la proposition, on aura besoin d'un petit lemme technique :

Lemme IV.5.16. *Quels que soient $\vec{u} \in \vec{K}$, $n \in \mathbb{N}$ et $i \in \{0, 1\}$, on a*

$$\frac{\|\vec{M}_i^n(\vec{u})\|_2}{\|\vec{u}\|_2} \geq \frac{1}{2} \left(\frac{3}{5}\right)^n.$$

Démonstration. Commençons par exprimer \vec{K} dans les bases qui diagonalisent \vec{M}_i . Un calcul rapide montre qu'on a :

$$\vec{K} = \left\{ a\vec{v}_0 + b\vec{w}_0 \mid 0 \leq \frac{b}{a} \leq 3 \right\} = \left\{ a\vec{v}_1 + b\vec{w}_1 \mid 0 \geq \frac{b}{a} \geq -3 \right\}$$

Soit $i \in \{0, 1\}$ et $\vec{u} = a\vec{v}_i + b\vec{w}_i \in \vec{K}$. On a alors $\vec{M}_i(\vec{u}) = \frac{3}{5}a\vec{v}_i + \frac{1}{5}b\vec{w}_i$. Soit maintenant $n \in \mathbb{N}$; on a (puisque \vec{v}_i et \vec{w}_i sont orthogonaux) :

$$\begin{aligned} \left(\frac{\|\vec{M}_i^n(\vec{u})\|_2}{\|\vec{u}\|_2} \right)^2 &= \frac{\|(\frac{3}{5})^n a\vec{v}_i + (\frac{1}{5})^n b\vec{w}_i\|_2^2}{\|a\vec{v}_i + b\vec{w}_i\|_2^2} \\ &= \frac{(\frac{3}{5})^{2n} \|a\vec{v}_i\|_2^2 + (\frac{1}{5})^{2n} \|b\vec{w}_i\|_2^2}{\|a\vec{v}_i\|_2^2 + \|b\vec{w}_i\|_2^2} \\ &\geq (\frac{3}{5})^{2n} \frac{\|a\vec{v}_i\|_2^2}{\|a\vec{v}_i\|_2^2 + \|b\vec{w}_i\|_2^2} \\ &= (\frac{3}{5})^{2n} \frac{1}{1 + (\frac{b}{a})^2 \frac{\|\vec{w}_i\|_2^2}{\|\vec{v}_i\|_2^2}} \\ &= (\frac{3}{5})^{2n} \frac{1}{1 + \frac{1}{3} (\frac{b}{a})^2} \\ &\geq (\frac{3}{5})^{2n} \frac{1}{1 + \frac{1}{3} 3^2} \\ &= \left(\frac{1}{2} \left(\frac{3}{5}\right)^n\right)^2. \quad \square \end{aligned}$$

Démonstration de la proposition IV.5.14. Soit $\vec{u}_0 \in \vec{K}$; d'après la proposition IV.5.4 on peut alors écrire

$$\alpha_{\inf}(u, s) = \liminf_{n \rightarrow \infty} \frac{\ln \|\vec{M}_{a_1 \dots a_n}\|}{n \ln(\frac{1}{2})} = \liminf_{n \rightarrow \infty} \frac{\ln \|\vec{M}_{a_1 \dots a_n}(\vec{u}_0)\|_2}{n \ln(\frac{1}{2})}.$$

On peut maintenant décomposer le développement binaire de s en blocs de bits consécutifs égaux. Soit $n \in \mathbb{N}$; on a :

$$\vec{M}_{a_1 \dots a_n} = \vec{M}_{a_1} \dots \vec{M}_{a_n} = \vec{M}_{\alpha_1}^{l_1} \vec{M}_{\alpha_2}^{l_2} \dots \vec{M}_{\alpha_{k(n)-1}}^{l_{k(n)-1}} \vec{M}_{\alpha_{k(n)}}^{r(n)},$$

où les valeurs des α_i alternent entre 0 et 1 (de sorte que $\alpha_1 = a_1$, $\alpha_2 = 1 - a_1$, $\alpha_3 = a_1$, \dots , $\alpha_{k(n)} = a_n$) et tous les l_i sont des entiers naturels non nuls. Du lemme IV.5.16, il découle que

$$\|\vec{M}_{a_1 \dots a_n}(\vec{u}_0)\|_2 \geq \left(\frac{1}{2}\right)^{k(n)} \left(\frac{3}{5}\right)^n \|\vec{u}_0\|_2$$

d'où (ne pas oublier que $\ln(\frac{1}{2}) < 0$) :

$$\frac{\ln \|\vec{M}_{a_1 \dots a_n}(\vec{u}_0)\|_2}{n \ln(\frac{1}{2})} \leq \frac{k(n)}{n} + \frac{\ln(\frac{3}{5})}{\ln(\frac{1}{2})} + \frac{\ln \|\vec{u}_0\|_2}{n}.$$

Quand n augmente, le troisième terme tend vers 0, et la limite inférieure du premier terme est par définition égale à $d_{\inf}(x)$, d'où la conclusion. \square

Minorer l'exposant de Hölder semble beaucoup plus difficile que de le majorer. Voici une borne triviale : on a clairement

$$\forall \vec{u} \in \vec{\mathcal{H}}, \forall i \in \{0, 1\}, \quad \|M_i(\vec{u})\| \leq \frac{3}{5} \|\vec{u}\|,$$

donc l'exposant de Hölder est toujours supérieur ou égal à $\frac{\ln(\frac{3}{5})}{\ln(\frac{1}{2})} \approx 0,737$. Mais cela ne suffit pas pour trouver des points où la dérivée est nulle. L'auteur soupçonne que si le développement binaire de s ne contient jamais trois bits consécutifs égaux, alors $u'(s) = 0$ — mais cela reste à vérifier.

Bibliographie

- [1] H. Abels. Properly discontinuous groups of affine transformations, a survey. *Geom. Dedicata*, 87:309–333, 2001.
- [2] H. Abels, G.A. Margulis, and G.A. Soifer. On the Zariski closure of the linear part of a properly discontinuous group of affine transformations. *J. Differential Geometry*, 60:315–344, 2002.
- [3] H. Abels, G.A. Margulis, and G.A. Soifer. The linear part of an affine group acting properly discontinuously and leaving a quadratic form invariant. *Geom. Dedicata*, 153:1–46, 2011.
- [4] H. Abels, G.A. Margulis, and G.A. Soifer. The Auslander conjecture for dimension less than 7. Préppublication, [arXiv:1211.2525](https://arxiv.org/abs/1211.2525) [math.GR], 2013.
- [5] L. Auslander. The structure of compact locally affine manifolds. *Topology*, 3:131–139, 1964.
- [6] Y. Benoist. Actions propres sur les espaces homogènes réductifs. *Annals of Mathematics*, 144:315–347, 1996.
- [7] Y. Benoist. Pavages du plan. Disponible sur <http://www.math.u-psud.fr/~benoist/prepubli/01pavage.pdf>, 2001.
- [8] V. Charette and W.M. Goldman. Affine Schottky groups and crooked tilings. *Contemporary Mathematics*, 262:69–90, 2000.
- [9] T. Drumm. Fundamental polyhedra for Margulis space-times. *Topology*, 31(4):677–683, 1992.
- [10] T. Drumm. Linear holonomy of Margulis space-times. *J. Differential Geometry*, 38:679–691, 1993.
- [11] D. Fried and W.M. Goldman. Three-dimensional affine crystallographic groups. *Adv. in Math.*, 47:1–49, 1983.
- [12] J. Kigami. A harmonic calculus on the Sierpinski spaces. *Japan J. Appl. Math.*, 6:2:259–290, 1989.
- [13] A.A. Kirillov. *A tale of two fractals*. Birkhäuser, 2013.
- [14] A.W. Knap. *Lie groups beyond an introduction*. Birkhäuser, 1996.
- [15] G.A. Margulis. Free properly discontinuous groups of affine transformations. *Dokl. Akad. Nauk SSSR*, 272:937–940, 1983.
- [16] G.A. Margulis. Complete affine locally flat manifolds with a free fundamental group. *J. Soviet Math.*, 134:129–134, 1987.
- [17] J. Milnor. On fundamental groups of complete affinely flat manifolds. *Adv. in Math.*, 25:178–187, 1977.
- [18] W. Rudin. *Real and complex analysis*. Mc Graw-Hill, 1966.
- [19] I. Smilga. Fundamental domains for properly discontinuous affine groups. *Geom. Dedicata*, 171:203–229, 2014.
- [20] I. Smilga. Proper affine actions on semisimple Lie algebras. Soumis; disponible sur [arXiv:1406.5906](https://arxiv.org/abs/1406.5906) [math.GR], 2014.
- [21] R.S. Strichartz. *Differential equations on fractals: A tutorial*. Princeton University Press, 2006.
- [22] J. Tits. Free subgroups in linear groups. *J. of Algebra*, 20:250–270, 1972.