ASTÉRISQUE 290

SÉMINAIRE BOURBAKI
VOLUME 2001/2002
EXPOSÉS 894-908

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

Association des collaborateurs de Nicolas Bourbaki, École normale supérieure,
45, rue d'Ulm, F-75230 Paris Cedex 05.
Url : `http://www.bourbaki.ens.fr/`

***Mots clefs et classification mathématique par sujets* (2000)**

Exposé n° 894. — Finite groups, fixed points. — 20E32, 55M35.
Exposé n° 895. — Complexity theory, approximation algorithms, proof verification. — 68W25.
Exposé n° 896. — Iwasawa algebra, p-adic Lie groups. — 11-xx, 16-xx, 22-xx.
Exposé n° 897. — Représentations p-adiques, équations différentielles p-adiques. — 11Sxx, 14Fxx.
Exposé n° 898. — Macdonald polynomials, Hilbert schemes, Cohen-Macaulay, Gorenstein, sheaf cohomology. — 14C05, 05E05, 20C30, 33D45.
Exposé n° 899. — Théorie des champs, théorie quantique des champs, cordes, supercordes, dualité, super-variété, supersymétrie, renormalisation, théorie de jauge, monopôle, équation de Seiberg-Witten, membrane, M-théorie. — 53Cxx, 53Zxx, 57M50, 57R56, 57R57, 58Z05, 81Txx.
Exposé n° 900. — Renormalisation, théories quantiques, problème de Riemann-Hilbert. — 81-xx, 81S40, 58D30.
Exposé n° 901. — Variétés algébriques, cobordisme, anneaux de Chow, formule du degré, anneau de Lazard. — 14A10, 14F35, 14F42, 55N20, 55N22.
Exposé n° 902. — Opérateurs accrétifs, racine carrée accrétive, conjecture de Kato. — 35J30, 35J45, 47F05, 47B44.
Exposé n° 903. — F-isocrystal, Newton stratification, Siegel moduli space, p-divisible group. — 14L05, 14F30.
Exposé n° 904. — Dynamiques génériques, hyperbolicité, transitivité. — 37C20, 37D30, 37C29.
Exposé n° 905. — Rationnellement connexe, variété de Fano, corps C_1, point rationnel, quotient rationnel, variété uniréglée, groupe fondamental, espace de morphismes, application stable, courbe rationnelle libre, monodromie, schéma de Hilbert. — 14J99, 14J45, 14J40, 14J30, 14J26, 14J28, 14G05, 14G15, 14M20.
Exposé n° 906. — Correspondance de Langlands, corps de classes géométrique, faisceaux pervers, fibrés vectoriels. — 11R39, 14F05, 14F20.
Exposé n° 907. — Dessins d'enfants, revêtement. — 14H30, 57M12.
Exposé n° 908. — Invariant tori, Kolmogorov–Arnold–Moser, Hamiltonian systems. — 37K55.

SÉMINAIRE BOURBAKI
VOLUME 2001/2002
EXPOSÉS 894-908

***Résumé.* —** Comme les précédents volumes de ce séminaire, celui-ci contient quinze exposés de synthèse sur des sujets d'actualité : trois exposés de géométrie algébrique, deux sur les systèmes dynamiques, un sur les actions de groupes finis, un de combinatoire et de géométrie algébrique, un d'informatique théorique, un sur la monodromie p-adique, un sur les algèbres d'Iwasawa, un sur la conjecture de Kato, un sur la renormalisation et les diagrammes de Feynman, un sur les dualités en théorie des cordes, un sur la correspondance de Langlands géométrique et un sur les dessins d'enfants.

***Abstract* (Séminaire Bourbaki, volume 2001/2002, exposés 894-908)**

As in the preceding volumes of this seminar, one finds here fifteen survey lectures on topics of current interest: three lectures on algebraic geometry, two on dynamical systems, one on actions of finite groups, one on combinatorics and algebraic geometry, one on theoretical computer sciences, one on p-adic monodromy, one on Iwasawa algebras, one on the Kato conjecture, one on renormalization and Feynman diagrams, one on dualities in string theory, one on the geometric Langlands correspondence and one on "dessins d'enfants".

Alejandro ADEM – *Finite group actions on acyclic* 2*-complexes*

We discuss the recent work of Oliver and Segev, which provides a complete description of the finite groups which can act on a 2–dimensional acyclic complex without fixed points. More precisely they show that there is an essential fixed point free 2–dimensional acyclic G–complex if and only if G is isomorphic to one of the simple groups $\mathrm{PSL}_2(2^k)$ for $k \geqslant 2$, $\mathrm{PSL}_2(q)$ for $q \equiv \pm 3$ $(mod\ 8)$ and $q \geqslant 5$, or $\mathrm{Sz}(2^k)$ for odd $k \geqslant 3$. Furthermore the isotropy subgroups of any such G–complex are all solvable.

Bernard CHAZELLE – *The PCP Theorem* [*after Arora, Lund, Motwani, Safra, Sudan, Szegedy*]

A new characterization of the complexity class NP in terms of probabilistically checkable proofs has had unexpected consequences in combinatorial optimization. The PCP theorem formalizes the idea that statements with short proofs can be checked with a constant number of random probes. This implies that many NP complete problems (such as coloring a graph) cannot be solved approximately with any reasonable level of accuracy.

John COATES – *Iwasawa algebras and arithmetic*

Let p be a prime number, G a compact p-adic Lie group, and $\Lambda(G)$ its Iwsawa algebra. When $G = \mathbb{Z}_p^d$, $d \geqslant 1$, classical commutative algebra gives a beautiful structure theorem for finitely generated $\Lambda(G)$-modules. The aim of the exposé is to generalize this structure theorem to the non-commutative case when G is a compact p-valued p-adic Lie group in the sense of M. Lazard, following recent joint work of the author, P. Schneider, and R. Sujatha; this work is partly based on earlier work of M. Chamarie. We will illustrate the abstract theory with down to earth, but mysterious, examples arising in the arithmetic of elliptic curves without complex multiplication.

Pierre COLMEZ – *Les conjectures de monodromie p-adiques*

Le théorème local de monodromie ℓ-adique de Grothendieck admet, en p-adique, deux généralisations. L'une, conjecturée par Crew, concerne les équations différentielles p-adiques et l'autre, conjecturée par Fontaine, les représentations galoisiennes. L'exposé tente de donner un aperçu des travaux menant aux démonstrations de ces conjectures.

Claudio PROCESI – *On the* $n!$*-conjecture*

We discuss the proof given by M. Haiman of the Macdonald positivity conjecture obtained via the solution to the $n!-$conjecture of Garsia and Haiman. This is obtained from the following remarkable theorem: the Hilbert scheme of $n-$tuples of points in the plane is equal to the $G-$Hilbert scheme of Ito and Nakamura for the action of the symmetric group on the space of such $n-$tuples.

Daniel BENNEQUIN – *Dualités de champs et de cordes* [*d'après 't Hooft, Polyakov, Witten et al.*]

L'étude non perturbative des théories quantiques de champs et de super-cordes a révélé l'existence de *dualités* surprenantes, échangeant électricité et magnétisme, comportements à courte et longue distance, constantes d'interaction faibles et fortes, et provoquant des objets mystérieux (M-théorie). Les implications en mathématiques sont variées : symétries miroirs, formes automorphes, invariants de Seiberg-Witten,... L'exposé veut seulement faire une introduction, pour mathématiciens.

Louis BOUTET de MONVEL – *Algèbre de Hopf des diagrammes de Feynman, renormalisation et factorisation de Wiener-Hopf* [*d'après A. Connes et D. Kreimer*]

La théorie quantique des champs perturbative fournit des séries asymptotiques d'intégrales divergentes. La renormalisation a pour objet d'attribuer à ces intégrales, de façon cohérente, des valeurs numériques précises (parties finies). A. Connes et D. Kreimer ont proposé une méthode mathématiquement limpide et universelle pour accomplir ce programme : les diagrammes de Feynman qui repèrent les termes de la série sont organisés en algèbre de Hopf graduée, et la série renormalisée se déduit d' une factorisation de Wiener-Hopf de la série perturbative, réinterprétée comme fonction (lacet) $\gamma(\mu, \varepsilon)$ à valeur dans le groupe formel associé (dépendant d'une unité de masse μ liée à la graduation, et d'un paramètre ε (dimension) servant au prolongement analytique).

François LOESER – *Cobordisme des variétés algébriques* [*d'après M. Levine et F. Morel*]

M. Levine et F. Morel ont récemment construit un analogue en géométrie algébrique du cobordisme complexe, le cobordisme algébrique. On présentera la construction de cette théorie, ainsi que ses principales propriétés, établies par M. Levine et F. Morel. En particulier, on expliquera comment elle permet d'obtenir une formule du degré très générale qui englobe des résultats antérieurs de Rost et de Voevodsky.

Yves MEYER – *La conjecture de Kato* [*d'après Pascal Auscher, Steve Hofmann, Michael Lacey, John Lewis, Alan McIntosh et Philippe Tchamitchian*]

La conjecture de Kato concerne le domaine de la racine carrée d'opérateurs différentiels accrétifs. Elle vient d'être résolue par P. Auscher et ses collaborateurs. Nous rattacherons cette conjecture aux travaux antérieurs d'A. Calderón, J. Moser et E. de Giorgi. Nous donnerons ensuite un aperçu de la preuve.

Michael RAPOPORT – *On the Newton stratification*

This will be a report on algebraic geometry in characteristic p. Let A/S be a family of abelian varieties over a base scheme S of characteristic p. By associating to each geometric point $\overline{s}$ of S the isogeny class of the p-divisible group of $A_{\overline{s}}$ we obtain a finite disjoint decomposition of S into locally closed subsets, the *Newton stratification of* S associated to A. We will discuss the recent results of de Jong, Oort and others on this stratification in general and in the particular case when S is the solution of a classical moduli problem of abelian varieties.

Christian BONATTI – *Dynamiques génériques : hyperbolicité et transitivité*

Quand les orbites d'un système dynamique passent indéfiniment près de tout point d'un compact K, on dit que K est transitif. *Quels sont les transitifs maximaux d'un système* ? L'exemple des dynamiques hyperboliques de la théorie de Smale montre que la réponse à cette question est l'une des clefs de la description globale de la dynamique. Des idées de R. Mañé ainsi que des théorèmes perturbatifs (lemme de connexion d'Hayashi et ses variantes) ont permis récemment d'identifier, quand ils sont en nombre fini, les maximaux transitifs des systèmes génériques, et de montrer qu'ils sont projectivement hyperboliques. En contraposée, l'absence d'hyperbolicité projective assure l'existence générique d'une infinité de maximaux transitifs.

Olivier DEBARRE – *Variétés rationnellement connexes* [*d'après T. Graber, J. Harris, J. Starr et A.J. de Jong*]

On s'intéresse depuis longtemps aux variétés algébriques *rationnelles* (dont le corps des fonctions rationnelles est une extension transcendante pure du corps de base). En dimension au moins trois, la rationalité n'a pas un bon comportement géométrique, au contraire des variétés *rationnellement connexes*, qui sont telles que par deux points généraux passe une courbe rationnelle. Nous donnerons la démonstration d'un résultat dû à Graber, Harris et Starr en caractéristique nulle et à de Jong et Starr en général, selon lequel toute famille de variétés rationnellement connexes propres paramétrée par une courbe a une section.

Gérard LAUMON – *Travaux de Frenkel, Gaitsgory et Vilonen sur la correspondance de Drinfeld–Langlands*

Langlands a conjecturé l'existence d'une correspondance entre représentations automorphes et représentations de Galois, aussi bien sur les corps de nombres que sur les corps de fonctions. À la suite des travaux de Lang et Rosenlicht sur le corps de classes géométrique, Drinfeld a inventé un analogue géométrique de cette correspondance de Langlands dans le cas des corps de fonctions. La correspondance de Drinfeld-Langlands, appelée aussi correspondance de Langlands géométrique, est une dualité conjecturale entre deux espaces de modules naturellement associés à une courbe algébrique X et à un groupe réductif G. Dans le cas où X est projective et $G = \mathrm{GL}(n)$, une partie essentielle de cette correspondance vient d'être établie par E. Frenkel, D. Gaitsgory et K. Vilonen.

Joseph OESTERLÉ – *Dessins d'enfants*

Les dessins d'enfants, introduits par A. Grothendieck dans *Esquisse d'un programme*, permettent de visualiser les revêtements finis (topologiques, analytiques ou algébriques, cela revient au même) de la droite projective complexe, privée des points 0, 1 et ∞. Un tel revêtement possède un unique modèle sur $\overline{\mathbf{Q}}$ et est déterminé par la donnée d'un ensemble fini muni de deux permutations. Étudier l'action de $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ sur ces objets, et en décrire combinatoirement les modèles entiers sont deux questions centrales du sujet.

Ricardo PÉREZ-MARCO – *KAM techniques in PDE*

Kolmogorov-Arnold-Moser theory of invariant tori in Hamiltonian systems is one of the main achievements of Dynamical Systems. KAM provides the existence of abundant quasi-periodic solutions in non-linear Hamiltonian systems close to integrable ones. More recently, starting with the persistence theory of lower dimensional tori (Melnikov, Eliasson, Kuksin, Pöschel) and the techniques associated to Lyapunov periodic solutions (Craig, Wayne, Bourgain) KAM theory has been extended to the infinite dimensional setting of non-linear PDE. We plan to give an introduction to this new vast field, the most recent progress, and the main unsolved problems.

Séminaire BOURBAKI Novembre 2001
54e année, 2001-2002, no 894, p. 1 à 17

FINITE GROUP ACTIONS ON ACYCLIC 2-COMPLEXES

by **Alejandro ADEM**

1. A BRIEF HISTORY AND MOTIVATION

A simple consequence of the Brouwer fixed point theorem is that any cyclic group acting on a closed disk $\mathbb{D}^n$ must have a *fixed point*. The classical work of P.A. Smith [18] shows that if P is a finite p-group, then any action of P on $\mathbb{D}^n$ must have a fixed point. From this there arises a very evident question: is there a group of *composite order* which can act on some $\mathbb{D}^n$ *without* any fixed points? This was settled in the affirmative by Floyd and Richardson in 1959 (see [7]), when they constructed fixed point free actions of the alternating group A_5 on disks.

These examples stood out as special exceptions for several years — indeed no other such actions were known to exist until Oliver (see [13]) obtained a complete characterization of those finite groups which can act on disks without stationary points. To explain it we first need to introduce some group-theoretic concepts.

DEFINITION 1.1. — *For p and q primes, let $\mathcal{G}_p^q$ be the class of finite groups G with normal subgroups $P \lhd H \lhd G$, such that P is of p-power order, G/H is of q-power order, and H/P is cyclic; and let $\mathcal{G}_p = \cup_q \mathcal{G}_p^q$, $\mathcal{G} = \cup_p \mathcal{G}_p$.*

THEOREM 1.2. — *A finite group G has a fixed point free action on a disk if and only if $G \notin \mathcal{G}$. In particular, any non-solvable group has a fixed point free action on a disk, and an abelian group has such an action if and only if it has three or more non-cyclic Sylow subgroups.*

The smallest group with a fixed point free action on a disk is in fact the alternating group A_5; the smallest *abelian* group with such an action is $C_{30} \times C_{30}$. Oliver also proved that a group G will have a fixed point free action on a finite $\mathbb{Z}_p$-acyclic[(1)]

[(1)] Recall that a complex X is said to be $\mathbb{Z}_p$-*acyclic* if its reduced mod p homology is identically zero; if its reduced integral homology vanishes it is said to be *acyclic*.

complex if and only if $G \notin \mathcal{G}_p$. Note that a group G will act without fixed points on a contractible complex if and only if it acts without fixed points on an *acyclic* complex.

Taking into account Oliver's result, an obvious problem is that of constructing fixed point free actions on contractible or acyclic complexes of small dimension. A well-known theorem by J.-P. Serre states that any finite group acting on a tree must have a fixed point (see [17]). However, the situation for contractible 2-dimensional complexes is much more complicated — in fact it is an open question whether or not it is possible for a finite group to act on such a complex without fixed points. We will restrict our attention from now on to the case of acyclic 2-dimensional complexes.

Our starting point is the classical example of an A_5-action on an acyclic 2-dimensional complex without fixed points, which we now briefly recall. In fact it is an essential ingredient in the construction due to Floyd and Richardson which we discussed above. This example is constructed by considering the left A_5 action on the Poincaré sphere $\Sigma^3 = \mathrm{SO}(3)/\mathrm{A}_5$; as the action has a single fixed point (corresponding to the fact that A_5 is self-normalizing in $\mathrm{SO}(3)$) we may remove an open 3-disk U around it to obtain an acyclic compact 3-manifold $\Sigma^3 - U$ with a fixed point free action of A_5. This in turn can be collapsed to a 2-dimensional subcomplex $X \simeq \Sigma^3 - U$ upon which A_5 still acts without fixed points. Equivalently we could identify Σ^3 with the space obtained by identifying opposite faces of the solid dodecahedron in an appropriate way[(2)] and consider the A_5 action induced by the usual action on the dodecahedron. The fixed point is the center of D and by collapsing to its boundary we obtain an explicit 2-dimensional complex $X = \partial D/\simeq$ with a fixed point free action of A_5 which has 6 pentagonal 2-cells, 10 edges and 5 vertices. Note that if we take the join $A = \mathrm{A}_5 * X$ with the induced diagonal action of A_5, then we obtain a simply connected and acyclic complex, hence a contractible complex with a fixed point free action. From this we can obtain a fixed point free A_5 action on a disk via regular neighborhoods (as explained in [4], p. 57). This is the basic step in the construction of the Floyd-Richardson examples.

Now an obvious question arises from all of this: can we characterize those finite groups which can act without fixed points on acyclic 2-dimensional complexes? Indeed, are there even other examples of such actions? Remarkably it turns out that these actions are only possible for a small class of *simple* groups, and their precise determination and description will require using the classification of finite simple groups.

[(2)] To be precise: identify opposite faces of the dodecahedron by the map which pushes each face through the dodecahedron and twists it by $2\pi/10$ about the axis of the push in the direction of a right hand screw (see [12]).

2. STATEMENT OF RESULTS

In this note we will report on recent work of Oliver and Segev (see [15]) where they provide a complete description of the finite groups which can act on a 2-dimensional acyclic complex without fixed points. Their work builds on previous contributions by Oliver [13], [14], Segev [16] and Aschbacher-Segev [2]. To state their main result we need to introduce a useful technical condition for G-CW complexes. From now on we will use the term G-complex to refer to a G-CW-complex, however these results also hold for simplicial complexes with an admissible G-action[(3)].

DEFINITION 2.1. — *A G-complex X is said to be* essential *if there is no normal subgroup $1 \neq N \lhd G$ with the property that for each $H \subseteq G$, the inclusion $X^{HN} \to X^H$ induces an isomorphism on integral homology.*

If there were such a normal subgroup N, then the G-action on X is 'essentially' the same as the G-action on X^N, which factors through a G/N-action. For 2-dimensional complexes we have:

THEOREM 2.2. — *Let G be any finite group and let X be any 2-dimensional acyclic G-complex. Let N denote the subgroup generated by all normal subgroups $N' \lhd G$ such that $X^{N'} \neq \varnothing$. Then X^N is acyclic, X is essential if and only if $N = 1$, and if $N \neq 1$ then the action of G/N on X^N is essential.*

Based on this we restrict our attention to essential complexes, and we can now state the main result in [15]:

THEOREM 2.3. — *Given a finite group G, there is an essential fixed point free 2-dimensional acyclic G-complex if and only if G is isomorphic to one of the simple groups $\mathrm{PSL}_2(2^k)$ for $k \geqslant 2$, $\mathrm{PSL}_2(q)$ for $q \equiv \pm 3 \pmod 8$ and $q \geqslant 5$ or $\mathrm{Sz}(2^k)$ for odd $k \geqslant 3$. Furthermore the isotropy subgroups of any such G-complex are all solvable.*

Among the groups listed above, only the *Suzuki groups* $\mathrm{Sz}(q)$ are not commonly known; we will provide a precise definition for them as subgroups of $\mathrm{GL}_4(\mathbb{F}_q)$ in § 5. Note that the theorem is stated for arbitrary acyclic 2-dimensional complexes; there is no need to require that the complexes be finite.

Our main goal will be to explain the proof of this result. This naturally breaks up into a number of different steps. We begin in § 3 by explaining how the theorem can be reduced to simple groups, based mostly on a theorem due to Segev [16]. Next in § 4 we describe techniques for constructing the desired actions, using methods derived from Oliver's original work on group actions on acyclic complexes as well as a more detailed analysis of the associated subgroup lattices. This is then applied in § 5 to

(3) A simplicial complex X with a G action is called admissible if the action permutes the simplices linearly and sends a simplex to itself only via the identity.

provide explicit descriptions of fixed point free actions on an acyclic 2-complex for the simple groups listed in the main theorem. In § 6 we sketch conditions which imply the *non-existence* of fixed point free actions on acyclic 2-complexes for most simple groups; this part requires detailed information about the intricate subgroup structure for the finite simple groups. Finally in § 7 we use the classification of finite simple groups and the previous results to outline the proof of the main theorem, which has been previously reduced to verification for simple groups. We also make a few concluding remarks.

Remark 2.4. — The background required to understand these results and their proofs includes: (1) very basic equivariant algebraic topology; (2) familiarity with subgroup complexes and related constructions; and (3) a very detailed knowledge of the subgroup structure of the finite simple groups. As many of the arguments in the proofs depend on the particular properties of these groups, our synopsis cannot hope to contain complete details. However the original paper by Oliver and Segev [15] is written in a clear style accessible to a broad range of mathematicians and hence those interested in a deeper understanding of the results presented here should consult it directly.

3. REDUCTION TO SIMPLE GROUPS

The goal of this section will be to explain how we can restrict our attention to finite simple groups. This is based on the following key result due to Segev [16] :

THEOREM 3.1. — *Let X be any 2-dimensional acyclic G-complex. Then the subcomplex of fixed points X^G is either acyclic or empty. If G is solvable then X^G is acyclic.*

Proof. — Although Segev's original proof uses the Odd Order Theorem, it can be proved more directly. One can show that if X is an acyclic G-complex, then $H_1(X^G, \mathbb{Z}) = H_2(X^G, \mathbb{Z}) = 0$. Hence we are reduced to establishing that there is only one connected component (provided X^G is non-empty). For solvable groups this can be proved directly using induction and Smith Theory. Otherwise we consider a minimal group G for which a counterexample exists. If X^G has k components then in fact it can be shown that X looks roughly like the join of an acyclic fixed point free G-complex Y with a set of k points. However as X is 2-dimensional, Y would have to be 1-dimensional, in other words a tree, and this cannot hold. □

Remark 3.2. — The reader should keep in mind that Theorem 3.1 is a basic tool in many of our subsequent arguments and it will be used explicitly and implicitly on several occasions.

COROLLARY 3.3. — *Let X be any 2-dimensional acyclic G-complex. Assume that $A, B \subset X$ are G-invariant acyclic subcomplexes such that $X^G \subset A \cup B$; then $A \cap B \neq \varnothing$.*

Proof. — Assume that $A \cap B = \varnothing$ and let Z denote the G-complex obtained by identifying A and B each to a point. Then Z is acyclic since A and B both are, and Z^G consists of two points, thus contradicting Theorem 3.1. □

As an immediate consequence of Corollary 3.3 we obtain

LEMMA 3.4. — *Let X be a 2-dimensional acyclic G-complex. Then if $H, K \subset G$ are such that $H \subset N_G(K)$ and X^H, X^K are both non-empty, then $X^{HK} \neq \varnothing$. Moreover, if $H \subset G$ is such that $X^H = \varnothing$, then $X^{\mathrm{C}_G(H)} \neq \varnothing$.*

Proof. — Since H normalizes K, both X^H and X^K are H-invariant acyclic subcomplexes of X. Hence we conclude from Corollary 3.3 that $\varnothing \neq X^H \cap X^K = X^{HK}$. For the second part, it suffices to prove it when H is minimal among subgroups without fixed points. Fix a pair $M, M' \subset H$ of distinct maximal subgroups (note that by Theorem 3.1, H is non-solvable). Then X^M and $X^{M'}$ are non-empty, but $X^M \cap X^{M'} = X^{\langle M, M' \rangle} = X^H = \varnothing$. Hence X^M and $X^{M'}$ are disjoint $\mathrm{C}_G(H)$-invariant acyclic subcomplexes of X, meaning (by Corollary 3.3) that their union cannot contain $X^{\mathrm{C}_G(H)}$, whence it must be non-empty. □

We can now prove one of the main reduction results, which allows us to restrict our attention to essential complexes.

THEOREM 3.5. — *Let G be any finite group, and let X be any 2-dimensional acyclic G-complex. Let N be the subgroup generated by all normal subgroups $N' \lhd G$ such that $X^{N'} \neq \varnothing$. Then X^N is acyclic; X is essential if and only if $N = 1$ and if $N \neq 1$ then the action of G/N on X^N is essential.*

Proof. — If $X^{N_1} \neq \varnothing$ and $X^{N_2} \neq \varnothing$ for $N_1, N_2 \lhd G$, then $X^{\langle N_1, N_2 \rangle} \neq \varnothing$ by Lemma 3.4. So we infer that X^N is non-empty, hence acyclic (by Theorem 3.1). Note that the action of any non-trivial normal subgroup of G/N on X^N has empty fixed point set, hence the action of G/N on X^N is always essential. Finally, assume that $N \neq 1$; by Theorem 3.1 we have that for all $H \subset G$, X^H and X^{NH} are acyclic or empty; and $X^{NH} \neq \varnothing$ if $X^H \neq \varnothing$, by Lemma 3.4. Hence the inclusion $X^{NH} \to X^H$ is always an equivalence of integral homology, and hence X is not essential. □

This result will allow us to focus our attention on actions of simple groups.

THEOREM 3.6. — *If G is a non-trivial finite group for which there exists an essential 2-dimensional acyclic G-complex X, then G is almost simple. In fact there is a normal simple subgroup $L \lhd G$ such that $X^L = \varnothing$ and such that $\mathrm{C}_G(L) = 1$.*

Proof. — We know from Theorem 3.5 that $X^N = \varnothing$ for all normal subgroups $1 \neq N \lhd G$, including the case $N = G$. Now fix a minimal normal subgroup $1 \neq L \lhd G$; we know from Theorem 3.1 that L is not solvable, as $X^L = \varnothing$. Hence L is a direct product of isomorphic non-abelian simple groups (see [8], Thm 2.1.5). Assume that L is not simple; by Lemma 3.4, $X^H \neq \varnothing$ for some simple factor $H \lhd L$. Also, note that $L = \langle gHg^{-1} \mid g \in G \rangle$ since it is a minimal normal subgroup. Now we have that $X^{gHg^{-1}} = gX^H \neq \varnothing$ for all $g \in G$, hence applying the same lemma once again, but now to the L-action on X, we infer that $X^L \neq \varnothing$, a contradiction.

So L is simple; now set $H = \mathrm{C}_G(L)$. Then we have that $H \lhd G$ (this follows from the fact that $L \lhd G$) and so $X^H \neq \varnothing$, by Lemma 3.4. However we have assumed that the action is essential, whence $H = 1$. □

The condition $\mathrm{C}_G(L) = 1$ is equivalent to $G \subseteq \mathrm{Aut}(L)$. Using this proposition we can decide which groups admit essential fixed point free actions on acyclic 2-dimensional complexes by first determining the *simple groups* with such actions and then looking at automorphism groups only for that restricted collection.

As the proof of the main theorem will require explicit knowledge about the finite simple groups, it seems appropriate to briefly recall their classification, we refer to [9] for a detailed explanation. We should point out that it is by now common knowledge that complete details of the proof of the Classification Theorem were not available when it was announced in 1981; crucial work involving the so-called *quasithin groups* was never published and is known to contain gaps. Fortunately this has been resolved thanks to more recent work by Aschbacher and Smith and although a full account has not yet been published, a draft of their manuscript (over 1200 pages long!) is now available on the world wide web (see [3]).

The following theorem encapsulates our understanding of finite simple groups, and its proof requires literally thousands of pages of mathematical arguments by many authors.

THEOREM 3.7. — *Let L denote a non-abelian finite simple group, then it must be isomorphic to one of the following groups:*

- *an alternating group A_n for $n \geqslant 5$*
- *a finite group of Lie type, i.e. a finite Chevalley group or a twisted analogue*[(4)]
- *one of the 26 sporadic simple groups.*

4. TECHNIQUES FOR CONSTRUCTING ACTIONS

One of the main results in [15] is an explicit listing of conditions which imply the existence of fixed point free actions on acyclic complexes. We first introduce

[(4)] We should mention that the Tits group ${}^2\mathrm{F}_4(2)'$ is actually of index 2 in the full Lie type group ${}^2\mathrm{F}_4(2)$.

DEFINITION 4.1. — *A non-empty family*[5] $\mathcal{F}$ *of subgroups of a group* G *is said to be* separating *if it has the following three properties: (a)* $G \notin \mathcal{F}$*; (b) any subgroup of an element in* $\mathcal{F}$ *is in* $\mathcal{F}$*; and (c) for any* $H \lhd K \subseteq G$ *with* K/H *solvable,* $K \in \mathcal{F}$ *if* $H \in \mathcal{F}$*.*

It is not hard to see that any maximal subgroup in a separating family of subgroups of G is self-normalizing. If G is solvable, then it has no separating family of subgroups. For G not solvable we let $\mathcal{SLV}$ denote the family of solvable subgroups, which is the minimal separating family for G.

DEFINITION 4.2. — *Given* G *and a family of subgroups* $\mathcal{F}$*, a* $(G, \mathcal{F})$*-complex is a* G*-complex such that all of its isotropy subgroups lie in* $\mathcal{F}$*. It is said to be* universal *(respectively* H*-universal) if the fixed point set of each* $K \in \mathcal{F}$ *is contractible (respectively acyclic).*

The following proposition relates the two previous concepts in our situation.

PROPOSITION 4.3. — *Let* X *denote a* 2*-dimensional acyclic* G*-complex without fixed points. Let* $\mathcal{F} = \{H \subset G \mid X^H \neq \varnothing\}$*. Then* $\mathcal{F}$ *is a separating family of subgroups of* G*, and* X *is an* H*-universal* $(G, \mathcal{F})$*-complex.*

Given a family of subgroups $\mathcal{F}$, let $N(\mathcal{F})$ denote the nerve of $\mathcal{F}$ (regarded as a poset via inclusion) with a G-action induced by conjugation. Given any set $\mathcal{H}$ of subgroups in G, we let $\mathcal{F}_{\geqslant \mathcal{H}}$ denote the poset of those subgroups in $\mathcal{F}$ which contain some element of $\mathcal{H}$. For a single subgroup H we use the notation $\mathcal{F}_{\geqslant H}$ and $\mathcal{F}_{>H}$ to denote the posets of subgroups containing H or strictly containing H, respectively. We denote $X^{\mathcal{H}} = \cup_{H \in \mathcal{H}} X^H$.

The following are two key technical lemmas which will be required:

LEMMA 4.4. — *If* X *denotes a universal (*H*-universal)* $(G, \mathcal{F})$*-complex then there exists a* G*-map* $X \to N(\mathcal{F})$ *which induces a homotopy equivalence (homology equivalence) between* $X^{\mathcal{H}}$ *and* $N(\mathcal{F}_{\geqslant \mathcal{H}})$*.*

LEMMA 4.5. — *Let* $\mathcal{F}$ *be any family of subgroups of* G*, and let* $\mathcal{F}_0 \subseteq \mathcal{F}$ *be any subfamily such that* $N(\mathcal{F}_{>H}) \simeq *$ *for all* $H \in \mathcal{F} - \mathcal{F}_0$*. Then any (*$H$*-)universal* $(G, \mathcal{F}_0)$*-complex is also an (*H*-)universal* $(G, \mathcal{F})$*-complex; and* $N((\mathcal{F}_0)_{\geqslant \mathcal{H}}) \simeq N(\mathcal{F}_{\geqslant \mathcal{H}})$ *for any set* $\mathcal{H}$ *of subgroups of* G*.*

A complex Y is said to be homologically m-dimensional if $H_n(X, \mathbb{Z}) = 0$ for all $n > m$ and $H_m(X, \mathbb{Z})$ is $\mathbb{Z}$-free. For later use we observe that, for $m \geqslant 1$, if X is an m-dimensional acyclic complex, then any subcomplex of X is homologically $(m-1)$-dimensional and that the intersection of a finite number of homologically $(m-2)$-dimensional complexes is also homologically $(m-2)$-dimensional.

[5] A family is a collection of subgroups of a group G which is closed under conjugation.

The following is a crucial criterion for the constructions we are seeking.

PROPOSITION 4.6. — *Let G be any finite group and let $\mathcal{F}$ be a separating family for G. Then the following are equivalent:*

- *There is a (finite) 2-dimensional H-universal $(G,\mathcal{F})$-complex.*
- *$N(\mathcal{F}_{>H})$ is homologically 1-dimensional for each subgroup $H \in \mathcal{F}$.*
- *$N(\mathcal{F}_{\geqslant\mathcal{H}})$ is homologically 1-dimensional for every set $\mathcal{H}$ of subgroups of G.*

Given a separating family $\mathcal{F}$ of subgroups of G, we say that $H \in \mathcal{F}$ is a *critical* subgroup if $N(\mathcal{F}_{>H})$ is not contractible. Given the above, we can concentrate our attention on the family $\mathcal{SLV}$ and its subfamily of critical subgroups, denoted $\mathcal{SLV}_c$.

First we record conditions which allow one to show that certain subgroups in a family are not critical.

LEMMA 4.7. — *Let $\mathcal{F}$ be any family if subgroups of G which has the property that $H \subseteq H' \subseteq H''$ and $H, H'' \in \mathcal{F}$ imply that $H' \in \mathcal{F}$. Fix a subgroup $H \in \mathcal{F}$; then $N(\mathcal{F}_{>H})$ is contractible if any of the following holds:*

- *H is not an intersection of maximal subgroups in $\mathcal{F}$.*
- *There is a subgroup $\widehat{H} \in \mathcal{F}$ properly containing H and such that $H \subsetneq K \cap \widehat{H}$ for all $H \subsetneq K \in \mathcal{F}_c$.*

We can now state a simple sufficient condition for the existence of a 2-dimensional H-universal $(G,\mathcal{F})$-complex:

PROPOSITION 4.8. — *Let $\mathcal{F}$ be any separating family of subgroups of G. Assume for every non-maximal critical subgroup $1 \neq H \in \mathcal{F}$, that $N_G(H) \in \mathcal{F}$, and that $H \subsetneq K \cap N_G(H)$ for all non-maximal critical subgroups $K \in \mathcal{F}$ properly containing H. Then there exists a 2-dimensional H-universal $(G,\mathcal{F})$-complex.*

We can in fact give a concrete description of the complex. For this we must introduce an integer associated to $H \in \mathcal{F}$.

DEFINITION 4.9. — *If $H \in \mathcal{F}$, a family of subgroups of G, we define*

$$i_{\mathcal{F}}(H) = \frac{1}{[N_G(H) : H]} \cdot (1 - \chi(N(\mathcal{F}_{>H}))).$$

Now let $M_1, \ldots, M_n$ be conjugacy classes representatives for the maximal subgroups of $\mathcal{F}$, and let $H_1, \ldots, H_k$ be conjugacy class representatives for all non-maximal critical subgroups of $\mathcal{F}$. Then there is a 2-dimensional H-universal $(G,\mathcal{F})$-universal complex X which consists of one orbit of vertices of type G/M_i for each $1 \leqslant i \leqslant n$, $[-i_{\mathcal{F}}(H_j)]$-orbits of 1-cells of type G/H_j for each $1 \leqslant j \leqslant k$, and free orbits of 1- and 2-cells. If G is simple[(6)] then X can be constructed to contain exactly $i_{\mathcal{F}}(1)$ free orbits of 2-cells, and no free orbits of 1-cells.

[(6)] In fact G must satisfy an additional technical condition which does not affect the results here.

5. EXPLICIT ACTIONS

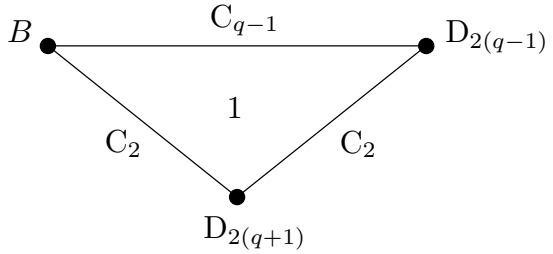
In this section we will outline the construction of fixed point free actions on acyclic 2-dimensional complexes for the simple groups $\mathrm{PSL}_2(2^k)$, for $k \geqslant 2$; $\mathrm{PSL}_2(q)$ for $q \equiv \pm 3 \pmod 8$ and $q \geqslant 5$; and for $\mathrm{Sz}(2^k)$ for odd $k \geqslant 3$.

Example 5.1. — Let $G = \mathrm{PSL}_2(q)$, where $q = 2^k$ and $k \geqslant 2$. Then there is a 2-dimensional acyclic fixed point free G-complex X all of whose isotropy subgroups are solvable. The complex X can be constructed with three orbits of vertices, with isotropy subgroups isomorphic to $B = \mathbb{F}_q \rtimes \mathrm{C}_{q-1}$, $\mathrm{D}_{2(q-1)}$ and $\mathrm{D}_{2(q+1)}$; three orbits of edges with isotropy subgroups isomorphic to C_{q-1}, C_2 and C_2; and one free orbit of 2-cells.

Here B denotes a Borel subgroup, expressed as a semi-direct product isomorphic to $(\mathrm{C}_2)^k \rtimes \mathrm{C}_{q-1}$, identified with the subgroup of projectivized upper triangular matrices. In our notation C_r denotes the cyclic group of order r and D_r denotes the dihedral group of order r. In fact $\mathrm{D}_{2(r-1)}$ can be identified with the subgroup of *monomial matrices*.

This example can be explained from the following analysis. The conjugacy classes of maximal solvable subgroups of G are represented precisely by the groups B, $\mathrm{D}_{2(q-1)}$ and $\mathrm{D}_{2(q+1)}$. The non-maximal critical subgroups must be intersections of maximal subgroups, one can check that up to conjugacy we get C_{q-1}, C_2 and 1. The precise numbers of orbits which appear is determined by calculating the integers $i_{\mathcal{SLV}}(H)$ for the isotropy subgroups.

This example can actually be constructed directly using the 1-skeleton Y_1 of the coset complex Y for the triple of subgroups $(K_1, K_2, K_3) = (B, \mathrm{D}_{2(q-1)}, \mathrm{D}_{2(q+1)})$ in $G = \mathrm{PSL}_2(\mathbb{F}_q)$ given by the maximal solvable subgroups. We can describe Y as the G-complex with vertex set $G/K_1 \sqcup G/K_2 \sqcup G/K_3$, where G acts by left translation, and with a 1-simplex for every pair of cosets with non-empty intersection and a 2-simplex for every triple of cosets with non-empty intersection. The following picture describes the orbit space Y/G:



It is not hard to see that as $G = \langle K_1, K_2, K_3 \rangle$, the complex Y is connected; however (as shown in [2], § 9) it is not acyclic for $k \geqslant 3$, where $q = 2^k$. However, one can show

that the module $H_1(Y_1, \mathbb{Z})$ is stably free — this involves a geometric argument based on the fact that Y_1 is a graph such that the fixed point sets Y_1^H are either contractible or empty for all subgroups $1 \neq H \subset G$ and contractible for all p-subgroups in G. The fact that G is a nonabelian simple group implies that the module must in fact be free (for a proof see [15], Prop. C.4.). Now we can simply attach a single free G-cell to Y_1 to kill its homology, yielding the acyclic complex X.

Carrying out this analysis in the classical case $G = \mathrm{A}_5$ yields an acyclic complex X (which in this case is actually identical to the complex Y) whose cellular chains give a complex of the form[(7)]

$$0 \to \mathbb{Z}[\mathrm{A}_5] \to \begin{array}{c} \mathbb{Z}[\mathrm{A}_5 / \mathrm{C}_2] \\ \oplus \\ \mathbb{Z}[\mathrm{A}_5 / \mathrm{C}_2] \\ \oplus \\ \mathbb{Z}[\mathrm{A}_5 / \mathrm{C}_3] \end{array} \to \begin{array}{c} \mathbb{Z}[\mathrm{A}_5 / \mathrm{A}_4] \\ \oplus \\ \mathbb{Z}[\mathrm{A}_5 / \mathrm{D}_6] \\ \oplus \\ \mathbb{Z}[\mathrm{A}_5 / \mathrm{D}_{10}] \end{array} \to \mathbb{Z} \to 0.$$

Example 5.2. — Let $G = \mathrm{PSL}_2(\mathbb{F}_q)$, where $q = p^k$, $q \geqslant 5$ and $q \equiv \pm 3 \pmod 8$. Then there exists a 2-dimensional acyclic fixed point free G-complex X, all of whose isotropy subgroups are solvable. More precisely, X can be constructed to have four orbits of vertices with isotropy subgroups isomorphic to $\mathbb{F}_q \rtimes \mathrm{C}_{(q-1)/2}$, D_{q-1}, D_{q+1} and A_4; four orbits of edges with isotropy subgroups isomorphic to $\mathrm{C}_{(q-1)/2}$, $\mathrm{C}_2 \times \mathrm{C}_2$, C_3 and C_2; and one free orbit of 2-cells.

These examples are slightly more complicated as the structure of the complex will depend on the value of q modulo 8.

Before explaining the final set of examples, we briefly recall the structure of the Suzuki groups $\mathrm{Sz}(q)$ (see [6], [11], [19] for details). Fix $q = 2^{2k+1}$ and let $\theta \in \mathrm{Aut}(\mathbb{F}_q)$ be the automorphism $x^\theta = x^{2^{k+1}} = x^{\sqrt{2q}}$ (note that $(x^\theta)^\theta = x^2$). For $a, b \in \mathbb{F}_q$ and $\lambda \in (\mathbb{F}_q)^*$, define the elements

$$S(a,b) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & a^\theta & 1 & 0 \\ a^{2+\theta} + ab + b^\theta & a^{1+\theta} + b & a & 1 \end{pmatrix}$$

and

$$M(\lambda) = \begin{pmatrix} \lambda^{1+2^k} & 0 & 0 & 0 \\ 0 & \lambda^{2^k} & 0 & 0 \\ 0 & 0 & \lambda^{-2^k} & 0 \\ 0 & 0 & 0 & \lambda^{-1-2^k} \end{pmatrix}, \quad \tau = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

[(7)] If we consider the original construction discussed in § 1 of an acyclic A_5-complex, then one can obtain the cellular structure below by subdividing each pentagon into a union of ten triangles.

Let $S(q,\theta) = \langle S(a,b) | a,b \in \mathbb{F}_q \rangle$, $T = \langle M(\lambda) | \lambda \in (\mathbb{F}_q)^* \rangle \cong \mathrm{C}_{q-1}$ and

$$B = M(q,\theta) = S(q,\theta) \rtimes T \text{ and } N = \langle T, \tau \rangle \cong \mathrm{D}_{2(q-1)}\,.$$

Then $\mathrm{Sz}(q) \cong \langle M(q,\theta), \tau \rangle$, and under this identification the following hold

(1) $S(q,\theta)$ is the 2-Sylow subgroup of $\mathrm{Sz}(q)$.

(2) There are four conjugacy classes of maximal subgroups in $\mathrm{Sz}(q)$ which are solvable: (B), (N), (M_+) and (M_-), where

$$M_+ \cong \mathrm{C}_{q+\sqrt{2q}+1} \rtimes \mathrm{C}_4 \text{ and } M_- \cong \mathrm{C}_{q-\sqrt{2q}+1} \rtimes \mathrm{C}_4\,.$$

These are all the maximal solvable subgroups in $\mathrm{Sz}(q)$.

(3) $|\mathrm{Sz}(q)| = q^2(q-1)(q^2+1) = q^2(q-1)(q+\sqrt{2q}+1)(q-\sqrt{2q}+1)$; note that the four factors in this expression are relatively prime.

We now describe the third set of examples.

Example 5.3. — Let $q = 2^{2k+1}$, for any $k \geqslant 1$. Then there is a 2-dimensional acyclic fixed point free $\mathrm{Sz}(q)$-complex X, all of whose isotropy subgroups are solvable. X can be constructed to have four orbits of vertices with isotropy subgroups isomorphic to $M(q,\theta)$, $\mathrm{D}_{2(q-1)}$, $\mathrm{C}_{q+\sqrt{2q}+1} \rtimes \mathrm{C}_4$ and $\mathrm{C}_{q-\sqrt{2q}+1} \rtimes \mathrm{C}_4$; four orbits of edges with isotropy subgroups isomorphic to C_{q-1}, C_4, C_4 and C_2; and one free orbit of 2-cells.

6. NON-EXISTENCE OF FIXED POINT FREE ACTIONS

In this section we outline methods for showing that *most* finite simple groups cannot act on an acyclic 2-dimensional complex without fixed points. The first result in this direction is due to Segev [16].

THEOREM 6.1. — *If G is the alternating group A_n, with $n \geqslant 6$, then there is no fixed point free action of G on any acyclic 2-dimensional complex.*

Later this was substantially extended by Aschbacher-Segev [2], who proved:

THEOREM 6.2. — *If G is a finite simple group which acts on an acyclic 2-dimensional complex without fixed points, then G must be isomorphic to either a group of Lie type and Lie rank one, or isomorphic to the sporadic simple group J_1 (the first Janko group).*

We will now sketch the key arguments used to establish these results, which (by the Classification Theorem) rule out most of the finite simple groups. The following lemma will be referred to as the *four subgroup criterion.*

LEMMA 6.3. — *Let G be a finite group and X a 2-dimensional acyclic G-complex. If $H_1, H_2, H_3, H_4 \subset G$ are subgroups such that $X^{\langle H_i, H_j, H_k \rangle} \neq \varnothing$ for any i,j,k then*

$$X^{\langle H_1, H_2, H_3, H_4 \rangle} \neq \varnothing.$$

Proof. — Suppose that in fact $X^{\langle H_1,H_2,H_3,H_4\rangle} = \varnothing$. Let $\mathcal{H} = \{H_1, H_2, H_3, H_4\}$. Now $X^{\mathcal{H}}$ is the union of the acyclic subcomplexes X^{H_i}, which are such that any two or three of them have acyclic intersection, but the four have empty intersection. The homology of $X^{\mathcal{H}}$ is isomorphic (see [5], p. 168) to that of the nerve of the corresponding acyclic covering; yielding $H_2(X^{\mathcal{H}}, \mathbb{Z}) \cong H_2(\mathbb{S}^2, \mathbb{Z}) \cong \mathbb{Z}$. However we know that $X^{\mathcal{H}}$ must be homologically 1-dimensional, which yields a contradiction. □

We now apply this result to multiply transitive groups.

PROPOSITION 6.4. — *Suppose that G acts 4-transitively on a set S with a point stabilizer $H \subset G$. If X is a 2-dimensional acyclic G-complex such that $X^H \neq \varnothing$, then $X^G \neq \varnothing$.*

Proof. — If $|S| = 4$ then G is an extension of the form $1 \to Q \to G \to K \to 1$, where $K \subseteq \Sigma_4$ and $Q \subset H$. By Theorem 3.1, $\varnothing \neq X^Q$ must be acyclic, and as K is solvable its action on X^Q must have a fixed point and we are done. So assume that $|S| \geqslant 5$, and fix four elements $s_1, s_2, s_3, s_4 \in S$. For each $i = 1, 2, 3, 4$, let $H_i \subset G$ be the subgroup of elements which fix s_j for all $j \neq i$. For each $\{i, j, k, r\} = \{1, 2, 3, 4\}$, $\langle H_i, H_j, H_k\rangle$ is the point stabilizer of s_r and therefore fixes a point in X by assumption. Hence by Lemma 6.3 (where $G = \langle H_1, H_2, H_3, H_4\rangle$), $X^G \neq \varnothing$. □

We apply this to show that the alternating groups A_n for $n \geqslant 6$ do not admit fixed point free actions on acyclic 2-complexes. Note that A_n is $(n-2)$-transitive on $\{1, 2, \ldots, n\}$, with point stabilizer A_{n-1} and that m-transitivity implies k-transitivity for $k \leqslant m$; hence A_n is 4-transitive on $\{1, \ldots, n\}$ for all $n \geqslant 6$ (see [1], page 56). If X is a 2-dimensional acyclic A_n-complex, then by our previous proposition, $X^{\mathrm{A}_n} \neq \varnothing$ if $X^{\mathrm{A}_{n-1}} \neq \varnothing$. Hence by induction we are reduced to considering the case when $G = \mathrm{A}_6$; assume that X is a 2-dimensional acyclic G-complex with $X^G = \varnothing$. Using the subgroups $H_i = \langle (i, 5, 6)\rangle$ for $i = 1, 2, 3, 4$ we can show by contradiction that $X^H = \varnothing$ for $H = \mathrm{Alt}\{1, \ldots, 5\}$ (using the covering argument as before). Using an outer automorphism we can thus establish that $X^H = \varnothing$ for any $H \subset G$ with $H \cong \mathrm{A}_5$. Next we consider the collection of subgroups

$$\mathcal{M} = \{\langle (12)(36)\rangle, \langle (12)(45)\rangle, \langle (12)(34)\rangle, \langle (25)(36)\rangle, \langle (26)(35)\rangle\};$$

again applying the covering arguments and comparing with the homology of the nerve of this covering we see that $H_2(X^{\mathcal{M}}, \mathbb{Z}) \neq 0$, a contradiction. We refer to [16], page 39 for details.

This method can also be applied to the Mathieu groups M_n; for $n = 11, 12, 23, 24$ they all act 4-transitively on a set with point stabilizer M_{n-1}. Now M_{10} contains A_6 as a subgroup of index two, hence every action of M_{11} or M_{12} on an acyclic 2-complex must have a fixed point. To obtain the same result for M_{23} and M_{24}, it suffices to establish it for M_{22}, which we will do subsequently.

The case of simple groups of Lie type, and of Lie rank at least equal to 2 can also be handled with these arguments (see [6] for background). We start with a basic lemma about parabolic subgroups.

LEMMA 6.5. — *Let G be a finite simple group of Lie type. Let Σ be the root system associated with G and let Σ_+ and Σ_- be the sets of positive and negative roots. Fix a set J of simple roots which does not contain all of them, and let L_J be the subgroup generated by the diagonal subgroup H together with the root subgroups X_r for all $r \in \langle J \rangle$. Let U_J and V_J be the subgroups generated by all X_r for roots $r \in \Sigma_+$ or $r \in \Sigma_-$, respectively, which are not in $\langle J \rangle$. Then $U_J \lhd P_J = U_J L_J$ and $V_J \lhd P'_J = V_J L_J$, U_J and V_J are nilpotent and $\langle U_J, V_J \rangle = G$.*

In our context we obtain the following fixed point theorem

LEMMA 6.6. — *Let G be a finite simple group of Lie type, and let $P \subsetneq G$ be one of the parabolic subgroups P_J or $P_{J'}$ in the previous lemma. Then for any action of G on an acyclic 2-complex X, $X^P \neq \varnothing$.*

Proof. — Let us assume that $X^G = \varnothing$. Then there are subgroups $U_J \lhd P_J$, $V_J \lhd P'_J$ and $L_J = P_J \cap P'_J$ such that U_J and V_J are nilpotent, $P_J = U_J L_J$, $P'_J = V_J L_J$, and $\langle U_J, V_J \rangle = G$. Note that X^{U_J} and X^{V_J} are acyclic, disjoint and L_J-invariant. Considering the subspaces $A = X^{U_J}$ and $B = X^{V_J}$ and the action of L_J, we see that $X^{L_J} \neq \varnothing$ (Corollary 3.3); similarly we conclude from Lemma 3.4 that X^{P_J} and $X^{P'_J}$ are non-empty. □

Now we can prove

THEOREM 6.7. — *If G is a simple group of Lie type and Lie rank at least two, then every G-action on an acyclic 2-dimensional complex has a fixed point.*

Proof. — Take a root system $\Sigma = \Sigma_+ \sqcup \Sigma_-$ for G and let $J_1 \sqcup J_2$ be a decomposition of the set of simple roots as a disjoint union of non-empty subsets. For each $i = 1, 2$, set $H_i^+ = \langle H, X_s \mid s \in J_i \rangle$, and $H_i^- = \langle H, X_{-s} \mid s \in J_i \rangle$. The subgroup generated by any three of the $H_i^\pm$ is contained in one of the parabolic subgroups P_{J_i} or P'_{J_i} and so has non-empty fixed point set in X. But in fact one can verify that $\langle H_1^\pm, H_2^\pm \rangle = G$, since it contains all subgroups X_s, X_{-s} for simple roots s and hence $X^G \neq \varnothing$ by the four subgroup criterion. □

In [2], Aschbacher and Segev were able to apply the four subgroup criterion to prove that any sporadic simple group other than the Janko group J_1 acting on an acyclic 2-complex has a fixed point. In [15] a different treatment is given, showing that all the sporadics can be handled using a consistent technique which relies on understanding the subgroup structure of these groups in some detail. The essential result is the following.

PROPOSITION 6.8. — *Let $\mathcal{F}$ be a separating family for G and let $K_1, K_2, K_3 \in \mathcal{F}$ be three subgroups such that neither K_2 nor K_3 is conjugate to K_1. Let $K_{ij} = K_i \cap K_j$ and $K = K_1 \cap K_2 \cap K_3$. Let $\mathcal{F}_0 \subseteq \mathcal{F}$ denote the subfamily consisting of $\mathcal{F}_c$ together with all subgroups conjugate to any of the K_i, K_{ij} or K. Assume that the following conditions hold, where $G' = \langle K_1, K_2, K_3 \rangle$:*

- $\frac{1}{[K_{12} : K]} + \frac{1}{[K_{13} : K]} + \frac{1}{[K_{23} : K]} \leqslant 1 + \frac{1}{[K_1 : K]} + \frac{1}{[K_2 : K]} + \frac{1}{[K_3 : K]} - \frac{1}{[G' : K]}$
- *K_1 is maximal in $\mathcal{F}$.*
- *There is no $H \in \mathcal{F}_0$ such that $K \subsetneq H \subsetneq K_{12}$ or $K_{12} \subsetneq H \subsetneq K_1$.*
- $N_G(K_1) \cap N_G(K_{12}) \cap N_G(K) = K$
- *The triples (K_1, K_{12}, K) and (K_1, K_{13}, K) are not G-conjugate.*

Then $H_2(N(\mathcal{F}_{\geqslant(K)}), \mathbb{Z}) \neq 0$ and so there is no 2-dimensional, H-universal $(G, \mathcal{F})$-complex.

This result can be proved as follows: the coset complex Y for the triple (K_1, K_2, K_3) must have $H_2(Y, \mathbb{Z}) \neq 0$ by the first hypothesis (this follows from a counting argument); the other conditions allow one to push a non-zero class non-trivially into

$$H_2(N((\mathcal{F}_0)_{\geqslant(K)}), \mathbb{Z}) \cong H_2(N(\mathcal{F}_{\geqslant(K)}), \mathbb{Z})$$

via the homomorphism induced by the G-equivariant simplicial map $Y^* \to N((\mathcal{F}_0)_{\geqslant(K)})$ sending each vertex in the barycentric subdivision Y^* of Y to its isotropy subgroup. By Proposition 4.6, this implies the stated result. This proposition can be applied systematically to yield

THEOREM 6.9. — *Let G be any of the sporadic simple groups; then there is no 2-dimensional acyclic G-complex without fixed points.*

We illustrate how this may be applied with two examples. Here we assume that we are given a 2-dimensional acyclic X with a fixed point free G-action, and take $\mathcal{F}$ to be the separating family of $H \subset G$ with $X^H \neq \varnothing$.

Example 6.10. — Let $G = M_{22}$, one of the Mathieu groups. We can take $K_3 \cong 2^4 : \mathrm{A}_6$, the subgroup which leaves invariant some hexad in the Steiner system of order 22, and it has an obvious action on this set of order 6 (see [10], Thm 6.8). Then K_1 can be taken to be the stabilizer of a point z in the hexad, and K_2 the stabilizer of some pair of points in the hexad including z. In this case $K_1 \cong \mathrm{L}_3(4)$, $K_2 \cong 2^4 : S_5$, $K_{12} \cong 2^4 : \mathrm{A}_5$, $K_{13} \cong 2^4 : \mathrm{A}_5$, $K_{23} \cong 2^4 : S_4$ and $K \cong 2^4 : \mathrm{A}_4$. Note that the K_{12} and K_{13} are distinct parabolic subgroups in $\mathrm{L}_3(4)$. The conditions in our previous proposition can be checked to hold (note that from our previous results we can see that the K_i all act with fixed points and hence are in $\mathcal{F}$) and so we have completed the verification that the Mathieu groups have no fixed point free actions on acyclic 2-complexes.

Next we deal with the case J_1 which was not covered by [2].

Example 6.11. — Let $G = J_1$, the first Janko group. Take $K_1 \cong (\mathrm{C}_2)^3 \rtimes \mathrm{G}_{21}$, where G_{21} is the Frobenius group of order 21, i.e. $\mathrm{C}_7 \rtimes \mathrm{C}_3$. K_1 is a maximal subgroup in J_1. Let $K_2 \cong \mathrm{C}_7 \times \mathrm{C}_6$ be the normalizer of a subgroup of order 7 in K_{12}, and let $K_3 \cong \mathrm{C}_3 \times \mathrm{D}_{10}$ be the centralizer in G of a subgroup of order 3 in K_2. Then $K_{12} \cong \mathrm{C}_7 \rtimes \mathrm{C}_3$, $K_{13} \cong \mathrm{C}_6 \cong K_{23}$, and $K = K_1 \cap K_2 \cap K_3 \cong \mathrm{C}_3$. Note that all these subgroups are solvable, and so are in $\mathcal{F}$. We can verify that

$$\sum_{i<j} \frac{1}{[K_{ij} : K]} = \frac{1}{7} + \frac{1}{2} + \frac{1}{2} < 1 + \frac{1}{14} + \frac{1}{10} = 1 + \frac{1}{[K_2 : K]} + \frac{1}{[K_3 : K]}$$

while the other conditions are also easy to check, hence showing that J_1 has no fixed point free action on an acyclic 2-complex.

We now consider the finite groups of Lie type which have Lie rank exactly equal to one. There are four families of such groups: the two dimensional projective special linear groups $\mathrm{L}_2(q)$, the three dimensional projective special unitary groups $\mathrm{U}_3(q)$, the Suzuki groups $\mathrm{Sz}(2^{2k+1})$, and the Ree groups $\mathrm{Ree}(3^{2k+1}) = {}^2\mathrm{G}_2(3^{2k+1})$.

The following propositions are used to handle these groups.

PROPOSITION 6.12. — *Let L be one of the simple groups $\mathrm{L}_2(q)$ or $\mathrm{Sz}(q)$, where $q = p^k$ and p is prime ($p = 2$ in the second case). Let $G \subset \mathrm{Aut}(L)$ be any subgroup containing L and $\mathcal{F}$ a separating family for G. Then there exists a 2-dimensional H-universal $(G, \mathcal{F})$-complex if and only if $G = L$, $\mathcal{F} = \mathcal{SLV}$ and q is a power of 2 or $q \equiv \pm 3 \pmod 8$.*

PROPOSITION 6.13. — *Let $G = \mathrm{U}_3(q)$, or ${}^2\mathrm{G}_2(3^{2k+1})$. Then there is no 2-dimensional acyclic G-complex without fixed points.*

These results are proved by combining our previous non-existence techniques with the following additional notion. For any family of subgroups $\mathcal{F}$ and any maximal element $M \in \mathcal{F}$, we set $Lk_{\mathcal{F}>1}(M) = N(\mathcal{F}_{>1}^{<M}) = N(\{H \in \mathcal{F} \mid 1 \neq H \subsetneq M\})$. Then we have

LEMMA 6.14. — *Let $\mathcal{F}$ denote a separating family for G. Let $\mathcal{F}_0 \subset \mathcal{F}$ be any subfamily which contains $\mathcal{F}_c$, and such that each non-maximal subgroup in $\mathcal{F}_0$ is contained in two or more maximal subgroups. Assume that $\mathcal{F}$ satisfies the following two conditions*

(1) *$N(\mathcal{F}_{>1})/G$ is connected and $H_1(N(\mathcal{F}_{>1})/G, \mathbb{Z}) = 0$.*
(2) *There is a maximal subgroup $M \in \mathcal{F}$ such that $Lk_{(\mathcal{F}_0)}(M)$ is not connected.*

Then there is no H-universal 2-dimensional $(G, \mathcal{F})$-complex.

Roughly speaking the proof of this lemma goes as follows: if such a complex did exist, then by (1) the singular set must be acyclic; but the prescribed conditions imply that the links at all vertices must be connected — hence contradicting (2).

In many instances this allows one to prove non-existence of a fixed point free action by contradiction; assuming its existence we can then find a maximal subgroup in the separating family such that the corresponding link is not connected. This of course requires a rather intricate knowledge of the maximal subgroups and more generally the finer structure of the groups under consideration. We refer to [15] § 6 for complete details.

7. PROOF OF THE MAIN THEOREM

We are now prepared to sketch a proof of the main theorem. We recall the statement.

THEOREM 7.1. — *If G is any finite group, then there is an essential fixed point free 2-dimensional finite acyclic G-complex if and only if G is isomorphic to one of the simple groups $\mathrm{PSL}_2(2^k)$ for $k \geqslant 2$; $\mathrm{PSL}_2(q)$ for $q \equiv \pm 3 \pmod 8$ and $q \geqslant 5$; or $\mathrm{Sz}(2^k)$ for odd $k \geqslant 3$. Moreover the isotropy subgroups of any such G-complex are all solvable.*

Proof. — We know that if G has an essential action on an acyclic 2-complex X without fixed points, then there is a non-abelian simple normal subgroup $L \lhd G$ with a fixed point free action and such that $G \subseteq \mathrm{Aut}(L)$. By the Classification Theorem, we know that L must be an alternating group, a group of Lie type, or a sporadic simple group. The results in the previous section rule out all groups on this list[(8)] except possibly the ones in the statement of the theorem. However we have already seen that these groups do in fact act on an acyclic 2-complex without fixed points, and that the isotropy subgroups are all solvable. This completes the proof. □

The work of Oliver and Segev has provided a complete picture for understanding fixed point free group actions on acyclic 2-dimensional complexes. There remains however the problem of considering actions on *contractible* 2-dimensional complexes. In fact Aschbacher and Segev [2] have raised the following

QUESTION 7.2. — *If X is a finite contractible 2-dimensional G-complex, then is $X^G \neq \varnothing$?*

This remains open. The results described here are a basic step towards investigating this question but it will probably require a substantially different approach.

[(8)] In fact the Tits group ${}^2\mathrm{F}_4(2)'$ must be handled separately because it is not the full Lie type group ${}^2\mathrm{F}_4(2)$.

REFERENCES

[1] M. Aschbacher – *Finite Group Theory*, Cambridge Studies in Advanced Mathematics, vol. 10, CUP, 2000.

[2] M. Aschbacher & Y. Segev – A fixed point theorem for groups acting on finite 2-dimensional acyclic simplicial complexes, *Proc. London Math. Soc. (3)* **67** (1993), p. 329–354.

[3] M. Aschbacher & S. Smith – Work in progress: a classification of quasithin groups, available at `http://www.math.uic.edu/~smiths/papers/quasithin/quasithin.dvi`.

[4] G. Bredon – *Introduction to Compact Transformation Groups*, Academic Press, 1972.

[5] K. Brown – *Cohomology of Groups*, GTM, vol. 87, Springer-Verlag, 1982.

[6] R. Carter – *Simple Groups of Lie Type*, Wiley, 1972.

[7] E. Floyd & R. Richardson – An action of a finite group on an n–cell without stationary points, *Bull. Amer. Math. Soc. (N.S.)* **65** (1959), p. 73–76.

[8] D. Gorenstein – *Finite Groups*, Harper and Row, 1969.

[9] ———, *The Classification of Finite Simple Groups*, Plenum Press, New York, 1983.

[10] R. Griess – *Twelve Sporadic Groups*, Springer-Verlag, 1998.

[11] B. Huppert & N. Blackburn – *Finite Groups III*, Springer-Verlag, 1982.

[12] R. Kirby & M. Scharlemann – Eight faces of the Poincaré homology 3-sphere, Geometric Topology, Academic Press, 1979.

[13] R. Oliver – Fixed point sets of group actions on finite acyclic complexes, *Comment. Math. Helv.* **50** (1975), p. 155–177.

[14] ———, Smooth compact group actions on disks, *Math. Z.* **149** (1976), p. 79–96.

[15] R. Oliver & Y. Segev – Fixed point free actions on acyclic 2-complexes, *Acta Mathematica* **189** (2002), p. 203–285.

[16] Y. Segev – Group actions on finite acyclic simplicial complexes, *Israel J. Math.* **82** (1993), p. 381–394.

[17] J.-P. Serre – *Trees*, Springer-Verlag, 1980.

[18] P.A. Smith – Fixed points of periodic transformations, AMS Coll. Pub., vol. XXVII, 1942, p. 350–373.

[19] M. Suzuki – On a class of doubly transitive groups, *Ann. of Math.* **75** (1962), p. 105–145.

Alejandro ADEM

Department of Mathematics
University of Wisconsin
Madison, Wisconsin 53706
U.S.A.
E-mail : `adem@math.wisc.edu`

Séminaire BOURBAKI Novembre 2001
54e année, 2001-2002, no 895, p. 19 à 36

THE PCP THEOREM
[after Arora, Lund, Motwani, Safra, Sudan, Szegedy]

by **Bernard CHAZELLE**

1. INTRODUCTION

The notion of *interactive proof systems* evolved out of cryptography and computational group theory. The cryptographic context is best explained through a little tale (perhaps one day to come true). One fine morning, one of your esteemed colleagues wakes up with, in his head, a crisp, concise, complete proof of Riemann's Hypothesis! Wisdom being one of his many qualities, he is not about to post his proof on the internet. Paranoia being another one, he is not even willing to reveal a single bit of information about the proof; that is, besides its conclusion that the RH is true. Is there any way for your colleague to convince you and the rest of the mathematical community that, indeed, he has a correct proof? Of course, one needs to define what exactly is meant by not "revealing a single bit". That is the subject of *zero-knowledge* cryptography.

The *PCP theorem* addresses a simpler variant: Can your colleague write down his proof in such a way that, were you to peek into it at a constant number of randomly chosen spots, you would leave utterly convinced of its validity? In other words, can he encode the proof as a string of bits so that: (i) a correct proof will never fail to convince you; (ii) an incorrect one will fool you with only a negligible probability? The catch is, you will be allowed to look at only a constant number of bits chosen at random. The PCP theorem asserts the existence of such an encoding. It is striking that the number of lookups can be kept constant regardless of the length of the proof. In fact, if you can put up with a failure rate slightly above $1/2$, i.e., accept a wrong proof half the time, but still never reject a correct one, then the number of bits can be reduced to 3. On the other hand, if you are allowed to read as many bits as are needed to store, say, two lines of this article, the probability of failure drops to 10^{-100}. A key point is that the new proof can be derived from the old one purely syntactically. In other words, one can write a compiler to translate the proof mechanically without any knowledge of mathematics. Furthermore, the new proof is not much longer than the previous one.

A common initial reaction to the PCP theorem is that it must be either wrong or trivial. Why wrong? It seems to imply that any flaw in the proof should spread itself all over the place, so as to be caught immediately in a random peek. But, how can so much information be stored in so few bits? Here is how: If the proof is correct, print it as such; if it is wrong, then intersperse the statement $2+2=3$ at every other step. The problem with that encoding is that a correct proof will not convince anyone. The beauty of the PCP theorem is not that flaws are caught so easily: it is that the mere absence of a flaw is persuasive in and of itself. There is nothing amazing about catching a liar's lie. But it is quite a feat to hear a true story from a congenital liar and end up believing it.

2. THE PCP VIEW OF NP

A *Turing machine* is a computer model whose main feature, for our purposes, is to be universal: in particular, whatever it can compute in time polynomial in the length of the input is believed to constitute what is tractable in *any* (non-quantum) model. The class P consists of the sets for which membership can be decided by a Turing machine in polynomial time. For example, the set of singular square integer matrices is in P, because determinants can be computed in a polynomial number of steps. The class NP includes the sets for which membership can be verified in polynomial time. For instance, the set of polynomials in $\mathbf{Q}[X_1, \ldots, X_n]$ with at least one zero in $\{0,1\}^n$ is in NP. The reason is that, given a polynomial f and a point $x \in \{0,1\}^n$, one can check if $f(x) = 0$ in time polynomial in the number of bits needed to represent f. To find such a zero from scratch seems more difficult (to put it mildly), and it is widely conjectured that P $\neq$ NP. Within computer science, this open question dwarfs all others in importance.

A 3-CNF formula is a conjunction of clauses, each one consisting of three literals; for example, $(v_1 \vee \neg v_2 \vee v_3) \wedge (v_2 \vee v_3 \vee \neg v_4)$. It is satisfiable if some true/false assignment of the v_i's makes the formula true. The one above is, whereas $(v_1 \vee v_1 \vee v_1) \wedge (\neg v_1 \vee \neg v_1 \vee \neg v_1)$ is not. The set of satisfiable 3-CNF formulas is called 3-SAT. A classical result of Cook and Levin says that 3-SAT is *NP-complete*, meaning that, not only it is in NP, but deciding membership in *any* NP set can be reduced to testing the satisfiability of a 3-CNF. The Cook-Levin theorem shows that to understand 3-SAT is to understand all of NP.

Many other sets are known to be NP-complete: for example, the set of 3-colorable graphs. (A graph is 3-colorable if its nodes can be colored red, white, and blue with no edge sharing the same color.) The existence of NP-complete sets brings breathtaking universality into the computing picture. It implies that anyone who can quickly color graphs can also solve algebraic equations over finite fields, factor integers, compute discrete logarithms, find short vectors in lattices, determine the largest clique in a graph, etc.

To formalize what a mathematical proof has to do with NP takes some effort, but the intuition is clear. In any reasonable axiomatic system, this set is in NP:

$$\{\langle T.1^n\rangle \mid T \text{ is a theorem with a proof of size at most } n\},$$

where $\langle T.1^n\rangle$ denotes the 0/1 string formed by writing the theorem T in binary in the axiomatic system and appending n ones at the end. A prover can guess a proof of length at most n, and the verifier can then check it in time polynomial in its length.

The class NP can be described in the language of proofs. If $L \in \mathrm{NP}$ then, given any $x \in L$, there exists a *short proof*, i.e., a polynomial-time computation, that x indeed belongs to L; for example, the solution of an algebraic equation. Conversely, if $x \notin L$, then no proof can convince anyone that x is in L. *Probabilistically checkable proofs* (PCP) add a small twist to this view: randomization. A PCP system for a set L consists of a string of bits (the proof) and a Turing machine with access to random bits (the verifier). Given an input x of n bits, the verifier generates $r(n)$ random bits[(1)]; then it looks up $q(n)$ bits of the proof at locations of its choice. The lookups are done all at once nonadaptively. Finally, after a polynomial amount of (deterministic) computation, the verifier must either accept or reject the proof. The class of sets L that satisfy the two requirements below is denoted by $\mathrm{PCP}[r(n), q(n)]$:

– Given any $x \in L$, there is a proof that causes the verifier to accept x with probability 1.

– Given any $x \notin L$, every proof is rejected with probability at least 1/2.

The functions r and q are called the *random-bit complexity* and *query-bit complexity*, respectively. To alleviate the notation, both of them are understood up to a constant factor. If $r(n) = O(\log n)$, the number of distinct random strings is polynomial and, by running the verifier on all of them, it is immediate that $\mathrm{PCP}[\log n, 1] \subseteq \mathrm{NP}$. Proving the reverse inclusion requires a great deal of ingenuity. The purpose of this article is to explain the proof at a conceptual level, leaving mathematical technicalities aside. The *PCP theorem* states that

$$\mathrm{NP} = \mathrm{PCP}[\log n, 1]. \tag{1}$$

Note that the proof size can be assumed to be polynomial since at most $q(n)2^{r(n)} = n^{O(1)}$ bits of the proof have a chance of ever being read. The PCP theorem can be restated in a way that highlights its "error-spreading" aspect. Given any 3-CNF formula Φ on n variables, there exists another one, denoted by Ψ, which contains $n^{O(1)}$ variables and is satisfiable if and only if Φ is. Furthermore, if Ψ is not satisfiable, then no truth assignment can satisfy more than a fraction $1 - \varepsilon$ of its clauses, for some constant $\varepsilon > 0$. Finally, Ψ can be derived from Φ in polynomial time.

(1) Throughout our discussion, random points or numbers are drawn uniformly, independently from a set that is always clearly understood from the context; in this case the set is $\{0, 1\}$.

It is instructive to see how this follows from (1), because the argument anticipates aspects of the proof of the PCP theorem. Consider a PCP system for Φ. Among the $2^{r(n)}$ possible random strings, some lead to acceptance, others (possibly) to rejection. Given such a string s, let $\Pi_1, \ldots, \Pi_q$ be the bits of the proof read by the verifier. (The locations of these bits depend on s but not on the proof itself.) Let Φ_s be a Boolean formula that evaluates to true if and only if $\Pi_1, \ldots, \Pi_q$ lead to acceptance: Φ_s has $q = O(1)$ variables, each one corresponding to one of the bits read. It is routine to convert Φ_s into a constant size 3-CNF formula Φ'_s by adding a few auxiliary variables if necessary. The formula $\Psi = \bigwedge_s \Phi'_s$ fits the bill. To see why, consider the (only interesting) case: If Φ is not satisfiable, then regardless of the proof, i.e., of the truth assignment of the Φ_s's variables, at least half of these formulas are false and, hence, so is a constant fraction of the clauses in Ψ. □

This characterization of the PCP theorem, which interestingly makes no mention of proofs, verifiers, or even randomization, points to the connection between PCP and inapproximability. Indeed, it implies that it is NP-complete to distinguish between a satisfiable formula and one for which no truth assignment satisfies at least a fraction $1 - \varepsilon$ of the clauses. Another way to look at this result is that if we set out to maximize the number of satisfied clauses in a formula, then we cannot hope to find an approximate solution within a factor $1 - \varepsilon$ of the maximum in polynomial time, unless P = NP. (Other applications are mentioned in the *Historical Notes* section.)

Remark 2.1. — From a mathematician's perspective, the PCP theorem might appear to focus on the "uninteresting" part of mathematics. It is a restatement of NP, not of P; as such, it says nothing about the difficulty of finding proofs. Also, it treats readers as mere fact-checkers. But mathematicians read proofs not so much to find bugs in them but to understand the ideas behind them. This mental picture, so vital to mathematics, is absent from the PCP viewpoint. Within the restrictive framework of verification, the PCP theorem is an impressive statement nevertheless.

Remark 2.2. — The proof of the PCP theorem is a mix of elementary algebra and probability theory; it is long and technical but not particularly difficult. Its originality lies elsewhere: in two places to be precise. One is its use of computational self-reducibility. Instead of keeping the usual separation between proving and verifying, the verifier's work is itself re-encoded as part of the proof: the reader of a proof is made partly its author! The other intriguing aspect of the PCP theorem is its ingenious use of error-correcting codes to express not just signals and bit streams (in typical coding theory fashion) but mathematical proofs, instead.

We close this section presenting a short, archetypical motif of the proof. Given a 3-CNF formula Φ, we wish to design a PCP system to verify its satisfiability. The idea is to construct a large family of multivariate polynomials f_i such that: if Φ is satisfiable, then any satisfying truth assignment corresponds to a common zero to all

the f_i's; otherwise, no more than half of them have a common zero. Suppose that the proof falsely claims that Φ is satisfiable. The verifier asks the prover to present the value of the f_i's at that common zero. If the prover obliges, then half of them will be nonzero and the verifier will easily catch the lie by random sampling. Therefore, the prover must cheat by substituting 0 for the actual values. The verifier's strategy is then to push the prover into the liar's standard pattern of generating new lies to cover up old ones.

Here is one way to do that: Encode any linear combination of the f_i's at the claimed zero as the image of a linear form g at some point x, and ask the prover to present g by values, i.e., provide a table, indexed by x, of all the values of $g(x)$. If the prover lies at one spot $g(x)$, then it must lie all over the place as well, since the verifier can evaluate $g(x)$ as $g(x+y)-g(y)$ for a random y, and hence, quickly spot any inconsistency. Many other such error-detecting mechanisms are needed. They all share the same goal, which is to force the prover to present functions that are very close to some "canonical" functions. Canonical functions are chosen to satisfy certain functional equations. Furthermore, any family of functions that satisfy these equations yield a satisfying assignment for Φ and, hence, a contradiction.

Notation. — $\mathcal{F}^d_{q,m}$ denotes the set of m-variate polynomials of total degree at most d with coefficients in $\mathbf{F}_q$, the finite field of q elements (not to be confused with the query time). We restrict ourselves exclusively to prime fields. We say that a function $f : \mathbf{F}_q^m \mapsto \mathbf{F}_q$ is δ-close to a (finite) family of functions if, for some g in that family, $\mathrm{Prob}[f(x) \neq g(x)] \leqslant \delta$, for random $x \in \mathbf{F}_q^m$. The smallest such δ is called the distance of f to the family. Given any nonempty $H \subseteq \mathbf{F}_q$, we use $\sum_{H^m} f$ as shorthand for $\sum_{x_1,\ldots,x_m \in H} f(x_1,\ldots,x_m)$. All logarithms are to the base 2.

3. TESTING TOOLS

Intuitively, the encoding of a proof in a PCP system must be such that any local deviation from what the verifier expects should be visible nearly everywhere. The relation between a polynomial and the corresponding polynomial map shares this characteristic: Changing a polynomial map at a single point has a rippling effect visible almost everywhere. This analogy suggests a line of attack: encode proofs as polynomials. In this section we pursue this lead and build a number of algebraic tools to be used later when proving the PCP theorem. We specify a polynomial in two different ways. It can be presented, i.e., written down, *by coefficients* (with the obvious meaning) or *by values* (as mentioned earlier). The appeal of the presentation by values is that it is extremely redundant and, hence, provides a built-in error-correcting mechanism.

3.1. The Sumcheck Test

Given $f \in \mathcal{F}^d_{q,m}$, $c \in \mathbf{F}_q$, and some nonempty $H \subseteq \mathbf{F}_q$, how can a proof convince a verifier that $\sum_{H^m} f = c$? Of course, the verifier can compute the sum on its own without a proof, but that requires evaluating f at $O(|H|^m)$ points. Can it be done faster? If the field is big enough, say, $q > 2dm$, it can be done with a single evaluation of f. Write $f_a = f(a, x_2, \ldots, x_m) \in \mathcal{F}^d_{q,m-1}$; if $m = 1$, $\sum_{H^{m-1}} f_a$ denotes $f(a)$. For $m > 0$, the PC proof is defined recursively as follows:

PC Proof that $\sum_{H^m} f = c$

[1] Present $g(x) = \sum_{H^{m-1}} f_x$ by coefficients.

[2] For all $a \in \mathbf{F}_q$, write down the PC proof that $\sum_{H^{m-1}} f_a = g(a)$.

The recursion bottoms out at $m = 0$: no PC proof needed there. The verifier adopts a two-pronged strategy: First, trust the proof and check that it supports the claim; then, test the proof for internal consistency. Accordingly, the verifier begins by checking that $\sum_H g = c$, rejecting the proof if this fails. Next, it verifies that $g(x)$ is, indeed, the polynomial it thinks it is. It picks a random $a \in \mathbf{F}_q$, and uses the PC proof in [2] to verify recursively that, indeed, $\sum_{H^{m-1}} f_a$ is equal to $g(a)$. (If $m = 1$, the verifier does not need to go to [2], since $\sum_{H^{m-1}} f_a$ is available via a single evaluation of the polynomial f.) If this succeeds, the verifier accepts the proof and its claim that $\sum_{H^m} f = c$; else, it rejects it.

To argue that this works, we first observe that if the proof is correct, the verifier always accepts it. On the other hand, if any test fails, rejection ensues. So, the only case worth considering is where the claim is not true but all the tests pass and, therefore, the proof is accepted. Important: all subsequent correctness proofs in this paper will be limited to this case, too, without a need to repeat why.

The case $m = 0$ is error-free. Note that it is the (only) place where the verifier can match the proof against its own knowledge of f. All other tests involve only the internal consistency of the PC proof. Assume now that $m > 0$. We prove by induction on m that the verifier wrongly accepts with probability at most dm/q. Since the first test passes, $g(x)$ cannot agree with the true $\sum_{H^{m-1}} f_x$ everywhere (else their respective sums would be both equal to c). But, being univariate polynomials of degree at most d, they agree at $x = a$ with probability at most d/q. This agreement might lead the verifier to wrongly accept. If there is disagreement at $x = a$, however, the verifier is back to its old task, but now with only $m - 1$ variables. So, by induction, it is fooled with probability at most $d(m-1)/q$. Adding the earlier bound of d/q completes the induction.

LEMMA 3.1. — *Given any $f \in \mathcal{F}^d_{q,m}$, $c \in \mathbf{F}_q$, and nonempty $H \subseteq \mathbf{F}_q$, if $\sum_{x \in H^m} f(x) = c$, then there exists a proof of that fact that the verifier always accepts. Otherwise, no proof is accepted with probability greater than dm/q. The verifier reads $O(dm \log q)$ bits of the proof, needs $O(m \log q)$ random bits, and performs a single evaluation of the polynomial f.*

The most remarkable feature of this test is that only one evaluation of the function f is necessary. Note how increasing the size of the field, by making the polynomial map f increasingly redundant, has the effect of decreasing the error probability. The sumcheck test can be used in conjunction with the simple fact below to test if a function that is close enough to a polynomial vanishes everywhere.

LEMMA 3.2. — *Given $H \subseteq \mathbf{F}_q$ of size $h > 0$, it is possible to build polynomials $R_1, \ldots, R_{q^\ell}$ in $\mathcal{F}^{h\ell}_{q,\ell}$, each one in time $(h+\ell)^{O(1)}$, so that for any nonzero function $g : H^\ell \mapsto \mathbf{F}_q$ and a random index $1 \leqslant i \leqslant q^\ell$, the probability that $\sum_{H^\ell} gR_i = 0$ is at most $h\ell/q$.*

3.2. The Low-Degree Test

We wish to design a PCP system to convince a verifier that a function $f : \mathbf{F}_q^m \mapsto \mathbf{F}_q$ is close to being a polynomial. Restrictions to lines are particularly useful for that purpose. Given $t \in \mathbf{F}_q$, let $f_{a,b}(t)$ denote the univariate function $f(a+tb)$. We define $\Delta_\ell(f)$ as the expected distance of $f_{a,b}$ to $\mathcal{F}^d_{q,1}$ for random $a, b \in \mathbf{F}_q^m$.

LEMMA 3.3. — *If q/md^3 is large enough, given any function $f : \mathbf{F}_q^m \mapsto \mathbf{F}_q$, the distance of f to $\mathcal{F}^d_{q,m}$ is at most $c\Delta_\ell(f)$, for some absolute constant $c > 0$.*

Intuitively, this is saying that if the restriction of a function to a random line is close to a polynomial, so is the function itself. This suggests an obvious PCP system.

> PC PROOF THAT f IS δ-CLOSE TO $\mathcal{F}^d_{q,m}$.
> For every pair $a, b \in \mathbf{F}_q^m$, present by coefficients the polynomial $g_{a,b} \in \mathcal{F}^d_{q,1}$ that is closest to $f_{a,b}$.

The verifier chooses two small constant parameters $\delta, \varepsilon > 0$ and picks $k = \lceil c\delta^{-1} \log \varepsilon^{-1} \rceil$ random (a_i, b_i, t_i), with $a_i, b_i \in \mathbf{F}_q^m$ and $t_i \in \mathbf{F}_q$. Next, it checks that $g_{a_i,b_i}(t_i) = f(a_i + t_i b_i)$, for each $1 \leqslant i \leqslant k$. If all k tests succeed, then the proof is accepted. Any failure means rejection.

Correctness is immediate: If $f \in \mathcal{F}^d_{q,m}$, then $g_{a,b}(t)$ coincides with $f(a+tb)$, for all a, b, and all tests succeed. Suppose now that the distance from f to $\mathcal{F}^d_{q,m}$ exceeds δ. For any fixed a_i, b_i and random t_i, the probability that $f(a_i + t_i b_i) \neq g_{a_i,b_i}(t_i)$ is

precisely the distance from f_{a_i,b_i} to $\mathcal{F}^d_{q,1}$. Averaging over all a_i, b_i, it follows from Lemma 3.3 that, for any fixed i and random a_i, b_i, t_i,

$$\text{Prob}[f(a_i + t_i b_i) \neq g_{a_i,b_i}(t_i)] = \Delta_\ell(f) > \frac{\delta}{c};$$

therefore, all k tests succeed with probability less than $(1 - \delta/c)^k < \varepsilon$. This implies the following lemma.

LEMMA 3.4. — *Let δ, ε be two arbitrarily small positive constants. Given a function $f : \mathbf{F}_q^m \mapsto \mathbf{F}_q$, fix some integer d such that q/md^3 is large enough. If $f \in \mathcal{F}^d_{q,m}$, then there exists a proof that the verifier always accepts. If f is not δ-close to $\mathcal{F}^d_{q,m}$, then no proof is accepted with probability greater than ε. The verifier reads $O(d \log q)$ bits of the proof, needs $O(m \log q)$ random bits, and performs $O(1)$ evaluations of f.*

If f is δ-close to $\mathcal{F}^d_{q,m}$, then its nearest polynomial (unique for small enough δ) can be evaluated at a random point by using f. To evaluate it at an arbitrary point, however, requires a recovery mechanism. The striking feature of the result below is that a single evaluation of f is sufficient to recover $f(x)$ at k (non-necessarily random) points.

LEMMA 3.5. — *For $\delta > 0$ small enough, fix d such that $q\delta/kd$ is large enough. Given a function $f : \mathbf{F}_q^m \mapsto \mathbf{F}_q$, let f_o be a nearest polynomial in $\mathcal{F}^d_{q,m}$. Pick k arbitrary points, $z_1, \ldots, z_k$ in $\mathbf{F}_q^m$. If f is δ-close to $\mathcal{F}^d_{q,m}$, then f_o is unique and there exists a proof that allows the verifier to output $f_o(z_1), \ldots, f_o(z_k)$.*

Otherwise, with probability $1 - O(\sqrt{\delta})$, the verifier either outputs the right values or rejects the proof. The verification reads $O(dk \log q)$ bits of the proof, needs $O(m \log q)$ random bits, takes $(mdk \log q)^{O(1)}$ time, and performs a single evaluation of the function f.

3.3. The Linearity Test

How hard is it to tell whether a function is almost a linear form, i.e., $x \in \mathbf{F}_q^m \mapsto a^T x$, for some $a \in \mathbf{F}_q^m$? We restrict ourselves to the case $q = 2$, the only one of interest for our purposes. Consider a function $f : \mathbf{F}_2^m \mapsto \mathbf{F}_2$ such that, for random $x, y \in \mathbf{F}_2^m$, $\text{Prob}[f(x) + f(y) \neq f(x + y)] \leqslant \delta$, for some small enough $\delta > 0$. A simple argument shows that the function is 2δ-close to some linear form. By now, we trust that the reader can easily write a PC proof for linearity testing.

LEMMA 3.6. — *Given a function $f : \mathbf{F}_2^m \mapsto \mathbf{F}_2$, fix a small enough constant δ. If f is a linear form, then there exists a proof of that fact that the verifier always accepts. On the other hand, if its distance to any linear form exceeds δ, no proof is accepted with probability greater than δ. The verifier reads $O(1)$ bits of the proof, needs $O(m)$ random bits, and performs $O(1)$ evaluations of the function f.*

4. THE PCP THEOREM

The proof consists of three parts: the first two involve the design of suboptimal PCP systems for 3-SAT; the third provides a composition method that allows us to plug the two suboptimal schemes together to produce the desired PCP system.

4.1. Optimal Random Bit Complexity: $\mathrm{NP} \subseteq \mathrm{PCP}[\log n, (\log n)^{O(1)}]$

Let Φ be a 3-CNF formula consisting of m clauses $C_1, \ldots, C_n$ and n variables $v_1, \ldots, v_n$. Since $m = O(n^3)$ and our PCP bounds in this section are all (poly)logarithmic, we might as well assume that $m = n$. We associate a polynomial with each clause in a fairly obvious way: $x_i(1-x_j)x_k$ with $\neg v_i \vee v_j \vee \neg v_k$; plus seven other possibilities. It is clear that Φ is satisfiable if and only if all these n polynomials have a simultaneous zero over $\mathbf{F}_q$. Given a 0/1 assignment f of the x_i's, we define a function $G^f : \{1, \ldots, n\}^4 \mapsto \{0, 1\}$ as follows:

$$G^f(i,j,k,l) = \begin{cases} 0 & \text{if } v_i, v_j, v_k \text{ do not all appear in } C_l \text{ in that order;} \\ f(i)f(j)f(k) & \text{if } C_l \text{ is } \neg v_i \vee \neg v_j \vee \neg v_k\,; \\ f(i)f(j)(1-f(k)) & \text{if } C_l \text{ is } \neg v_i \vee \neg v_j \vee v_k\,; \\ \text{etc.} & \text{(6 other cases).} \end{cases} \tag{2}$$

Φ is satisfiable if and only if there exists an assignment f such that the function G^f vanishes everywhere. To take advantage of the redundancy of polynomial maps mentioned earlier, we encode f and then G^f as polynomial maps.

Let $h = \Theta(\log n)$ and $m = \lceil h/\log h \rceil$, and define q to be a prime sufficiently larger than $h^3 m^4$. Without loss of generality, assume that $n = h^m$. Fix a bijection between $\{1, \ldots, n\}$ and H^m, where $H = \{0, \ldots, h-1\}$; for example, write $i-1$ in base h. From now on, we regard the index i of x_i as an element $(y_1, \ldots, y_m)$ of $H^m \subseteq \mathbf{F}_q^m$. The assignment f maps H^m to $\{0, 1\}$ and, by Lagrange interpolation, can be extended into a map defined by a polynomial of degree $(h-1)m$ (still called f, for simplicity):

$$f(y_1, \ldots, y_m) = \sum_{t_1, \ldots, t_m \in H} f(t_1, \ldots, t_m) \prod_{i=1}^{m} \prod_{u \in H \setminus \{t_i\}} \frac{y_i - u}{t_i - u}\,. \tag{3}$$

Similarly, we regard $G^f(i,j,k,l)$ as a function from H^{4m} to $\mathbf{F}_q$, which we extend into a polynomial G^f_{poly} in $\mathcal{F}^{9hm}_{q,4m}$. This is how we do it. First, we express G^f in a more unified manner:

$$G^f(i_1, i_2, i_3, l) = \prod_{s=1}^{3} a_s(i_s, l)\,(\,b_s(l) - f(i_s)\,),$$

with the obvious meaning of all these functions: s specifies one of the three literals in the clause C_l; i_s is the index (viewed as an element of H^m) of the corresponding

variable x_{i_s} and $f(i_s)$ is its 0/1 assignment; $b_s(l) = 0/1$ indicates if x_{i_s} is negated, etc. The polynomial extension of G^f is defined as

$$G^f_{\text{poly}}(i_1, i_2, i_3, l) = \prod_{s=1}^{3} a'_s(i_s, l)\,(\,b'_s(l) - f(i_s)\,),$$

where a'_s, b'_s are the polynomial extensions of a_s, b_s as defined in (3). It is immediate that G^f_{poly} is of degree less than $9hm$. (Note that we avoid Lagrange interpolation on G^f itself.)

PC Proof that Φ is Satisfiable

[1] Present $f : \mathbf{F}_q^m \mapsto \mathbf{F}_q$ by values, where f is a polynomial extension of a satisfying assigment for Φ.

[2] Write down the PC proof that f is ε-close to $\mathcal{F}^{hm}_{q,m}$. (*low-degree test*)

[3] Form all R_i's in Lemma 3.2 ($\ell = 4m$) and, for each $1 \leqslant i \leqslant q^\ell$, write down the PC proof that $\sum_{H^{4m}} G^f_{\text{poly}} R_i = 0$. (*sumcheck test*)

The verifier applies the low-degree test on f and rejects the proof if it fails. Otherwise, it picks a random $1 \leqslant i \leqslant q^{4m}$, and sumchecks that $\sum_{H^{4m}} G^f_{\text{poly}} R_i = 0$. It accepts the proof if this test succeeds and rejects it otherwise.

Remark 4.1. — The verifier can compute R_i and a'_s, b'_s on its own and also evaluate G^f_{poly} anywhere by querying f at three places in the proof. Also, we said earlier that the sumcheck test requires that the verifier can trust its evaluations of f. But what if the prover cheated in the presentation of f? This cannot happen: whatever is presented *is* what defines f.

Why does this protocol work? As usual, we do not have false negatives. If Φ is satisfiable, then the prover only has to stick to the scenario above and all tests will succeed. Suppose now that Φ cannot be satisfied. For the usual reasons, we assume that all tests (i.e., low-degree and sumcheck) succeed. We distinguish between three cases:

(1) $f \in \mathcal{F}^{hm}_{q,m}$: Then, G^f_{poly} is a polynomial of degree at most $9hm$, which is nonzero because Φ is not satisfiable. So, by Lemma 3.2, the probability that the sum to check is 0 is at most $4hm/q$. Assume that it is not 0.

The degree of $G^f_{\text{poly}} R_i$ does not exceed $13hm$; therefore, by Lemma 3.1, the probability that the sumcheck test succeeds is $O(hm^2/q)$. This bounds the probability of failure in this case by $O(hm^2/q)$.

(2) f is ε-close (but not 0-close) to $\mathcal{F}^{hm}_{q,m}$: Let f_o be a nearest neighbor in $\mathcal{F}^{hm}_{q,m}$. The sumcheck test requires a single evaluation of G^{f}_{poly} and, hence, evaluations of f at three points. The probability that f and f_o agree at all three points is at least $1-3\varepsilon$. The agreement means that the sumcheck test is, in effect, carried out on $G^{f_o}_{\text{poly}}R_i$. The previous case now shows that the failure probability is at most $3\varepsilon + O(hm^2/q)$.

(3) f is not ε-close to $\mathcal{F}^{hm}_{q,m}$: By Lemma 3.4, the low-degree test will fail to catch that fact with probability at most ε.

To summarize, the verifier might fail to spot an inconsistency with probability $O(\varepsilon+hm^2/q) < 1/2$, for a small enough constant $\varepsilon > 0$. Since $q = O(m^4h^3)$, the number of proof bits read is $O(hm^2 \log q) = (\log n)^{O(1)}$, the running time is $(hm)^{O(1)} = (\log n)^{O(1)}$, the number of random bits needed is $O(m \log q) = O(\log n)$ (the motivation for our choice of h), and the number of places at which the candidate assignment function is evaluated is $O(1)$. This proves that, indeed, NP $\subseteq$ PCP$[\log n, (\log n)^{O(1)}]$.

4.2. Optimal Query Bit Complexity: NP $\subseteq$ PCP$[n^{O(1)}, 1]$

Adding random bits allows the verifier to do with fewer lookups; in fact, a constant number of them. The strategy is roughly the same as before. Since the verifier is given access to only $O(1)$ bits of the proof, however, the ground field must be of constant size, so we set $q = 2$.

Let Φ be a 3-CNF formula consisting of m clauses $C_1, \ldots, C_m$ and n variables $v_1, \ldots, v_n$. As usual, we model each clause, say, $\neg v_i \vee v_j \vee \neg v_k$ as $x_i(1-x_j)x_k$. This defines m cubic polynomials $G_1, \ldots, G_m$, and Φ is satisfiable if and only if all the G_i's have a common zero over $\mathbf{F}_2$. Given $r = (r_1, \ldots, r_m) \in \mathbf{F}_2^m$, let $F_r(x) = \sum r_i G_i(x)$. Our interest in F_r comes from this (trivial) fact:

LEMMA 4.2. — *If Φ is satisfiable, all 2^m polynomials F_r have a common zero; otherwise, given any $a \in \mathbf{F}_2^n$ and a random $r \in \mathbf{F}_2^m$, $F_r(a) = 0$ with probability $1/2$.*

The PC proof of satisfiability is based on this simple test. The prover wants to convince the verifier that it knows a common zero a to all the F_r's (whether that is true or not). To do that, the proof will list the values of $F_r(a)$, for all r, so that the verifier can test that, indeed, $F_r(a) = 0$. The prover must also provide a consistency check that satisfies the verifier that its evaluations of F_r are correct. The cubic polynomial F_r can be written as $F_r(x) = f_r + \sum_i f_r^i x_i + \sum_{i,j} f_r^{ij}\, x_i x_j + \sum_{i,j,k} f_r^{ijk}\, x_i x_j x_k$. The verifier can evaluate $F_r(a)$ by using the three linear forms

$$(4) \quad H_1^a(y) = \sum_i a_i y_i\,; \qquad H_2^a(y) = \sum_{i,j} a_i a_j y_{ij}\,; \qquad H_3^a(y) = \sum_{i,j,k} a_i a_j a_k y_{ijk}\,.$$

Forgive our (abusive) notation y to refer to a set of n, n^2, and n^3 labeled variables, respectively. The PC proof will present each H_i^a by values ($i = 1, 2, 3$). Of course, no guarantee exists that the prover will not corrupt the presentation.

To catch any cheating, the verifier relies on two sets of functional equations: the H_i^a's are (i) linear and (ii) related by the identities

$$H_2^a(y \otimes y') = H_1^a(y) H_1^a(y') \qquad \text{and} \qquad H_3^a(y \otimes z) = H_1^a(y) H_2^a(z)\,, \tag{5}$$

where $y, y' \in \mathbf{F}_2^n$ and $z \in \mathbf{F}_2^{n^2}$. If $y \in \mathbf{F}_2^s$ and $z \in \mathbf{F}_2^t$, then $y \otimes z$ denotes the vector $(y_i z_j) \in \mathbf{F}_2^{st}$.

PC Proof that Φ is Satisfiable

Present the functions H_1^a, H_2^a, H_3^a by values, where a is a common zero to $G_1, \ldots, G_m$.

The verifier performs three basic sets of tests. Fix some small constant $\varepsilon > 0$.

– The first test is to check the linearity of each H_i^a, by using the criterion of Lemma 3.6, with $\delta = \varepsilon^2$. The verifier rejects the proof if any of these 3 tests fail. From now on, *any* evaluation of $H_i^a(y)$ is to be immediately confirmed by the following test: pick a random y' and check that $H_i^a(y) = H_i^a(y + y') - H_i^a(y')$. If this test ever fails, the proof is rejected.

– Next, the verifier checks that the H_i^a's are related by the two identities (5). Each one is tested $O(1/\varepsilon)$ times for random pairs (y, y') and (y, z), where $y, y' \in \mathbf{F}_2^n$ and $z \in \mathbf{F}_2^{n^2}$. Again, any failure implies rejection of the proof.

– Finally, the verifier picks a random r and evaluates $F_r(a)$. To do that, it computes on its own $y_1 = (f_r^i)$, $y_2 = (f_r^{ij})$, $y_3 = (f_r^{ijk})$, as well as f_r, and then looks up the proof at three places to compute the sum $F_r(a) = f_r + H_1^a(y_1) + H_2^a(y_2) + H_3^a(y_3)$. If $F_r(a) = 0$ and none of the previous tests have failed, the verifier accepts the claim that Φ is satisfiable.

Why does this work? If Φ is satisfiable, then the proof needs simply to conform to the directives of the verifier and it will be accepted. Suppose that Φ is not satisfiable and that, by contradiction, the proof is accepted. What is the probability of failure? If the linearity tests passes then, by Lemma 3.6, with probability a least $1 - 3\varepsilon^2$, there exists a linear form $\widehat{H}_i$, for $i = 1, 2, 3$, that disagrees with H_i^a over a fraction at most ε^2 of its domain. This means that, with probability $1 - O(\varepsilon)$, we can assume that all identity tests are performed with respect to the true values of $\widehat{H}_i$. If either identity fails to be satisfied by the $\widehat{H}_i$'s, then by Lemma 4.3 (and its omitted analog for the first identity) a conservative estimate of $O(\varepsilon)$ bounds the probability that the verifier fails to catch that fact.

Therefore, with probability $1 - O(\varepsilon)$, the value $F_r(a)$ computed by the verifier is, indeed, $f_r + \widehat{H}_1(y_1) + \widehat{H}_2(y_2) + \widehat{H}_3(y_3)$, for some linear forms $\widehat{H}_i$ defined by some vector a in accordance with the format specified by (4). Since Φ is not satisfiable, we know by Lemma 4.2 that the value of $F_r(a)$ is zero with probability $1/2$. Therefore,

the verifier will accept a wrong proof with probability $1/2 + O(\varepsilon)$. By setting ε to a small enough constant and repeating the verification, we bring the failure probability below $1/2$. The number of random bits is $O(n^3)$ and the number of bit lookups in the proof is constant. This concludes the proof that $\mathrm{NP} \subseteq \mathrm{PCP}[n^{O(1)}, 1]$.

LEMMA 4.3. — *If $\widehat{H}_3(y \otimes z) - \widehat{H}_1(y)\widehat{H}_2(z)$ is nonzero then, with probability $1/4$, it evaluates to 1 at random y, z.*

4.3. Self-Reduction: $\mathrm{NP} = \mathrm{PCP}[\log n, 1]$

We have built two PCP systems: one, S_1, needs $O(\log n)$ random bits; the other, S_2, uses $O(1)$ queries. We now combine them to extract the best feature from each. The basic idea is simple. We caught a glimpse of it in Section 2. Recall the action of the verifier V_1 for S_1. First, it generates a random bit string s; then, it computes a set of addresses $i_1, \ldots, i_q$ to look up in the proof Π, where $q = (\log n)^{O(1)}$. Upon reading the corresponding bits, $\Pi_{i_1}, \ldots, \Pi_{i_q}$, the verifier evaluates a predicate in time $q^{O(1)}$ to decide whether to accept or not.

Now comes the self-reducibility part. By the Cook-Levin theorem the predicate in question can be expressed as a 3-CNF formula Φ_s of size $q^{O(1)}$. The verifier accepts if and only if there exists a string X_s such that the concatenated string $\Pi_{i_1} \cdots \Pi_{i_q} X_s$ forms a satisfying truth assignment for Φ_s. The key idea is that a PCP system such as S_2 is exactly the sort of thing that can be used to check the satisfiability of Φ_s. So, instead of computing the verification predicate itself, V_1 can hand the problem over to the PC proof system S_2. Its verifier V_2 will then consult its own proof to check whether Φ_s is satisfiable, and will accept or reject accordingly. A minor technical point: Of course, we cannot let *both* V_1 and V_2 err with probability $1/2$. By repeating the verifiers' runs (the standard trick), we can trivially lower the odds of an error to any small constant.

Even though the number of bits read in the proof can be arbitrary, it is important for composition purposes that the number of *entries* be $O(1)$. Polynomial extensions give us a convenient tool for achieving that. A proof Π of length N can be viewed as a function $f : \{1, \ldots, N\} \mapsto \{0, 1\}$, where $f(i) = \Pi_i$. As we did in Section 4.1, we can change our point of view and regard f as a function from H^m to $\{0, 1\}$, where H^m is in bijection with $\{1, \ldots, N\}$ and $H = \{0, \ldots, h-1\}$, for some parameters h, m such that $N = h^m$. (Pad the proof with junk if N is inexpressible in this way.) Let $\widehat{f}$ be a polynomial extension of f in $\mathcal{F}^{hm}_{q,m}$. The proof Π is now rewritten as a presentation of $\widehat{f}$ by values, with all the bells and whistles needed to apply the low-degree test and the recovery mechanism (Lemmas 3.4 and 3.5). The verifier applies the low-degree test to check that the presentation is δ-close to $\mathcal{F}^{hm}_{q,m}$ (for some suitably small $\delta > 0$), and rejects the proof if it is not. Otherwise, it appeals to Lemma 3.5 to evaluate f at $k\,(= q)$ points by a single evaluation of $\widehat{f}$. In this way, the verifier can gain access to $\Pi_{i_1} \cdots \Pi_{i_q}$ in $O(1)$ queries to the new proof. Note that an entry in this new proof is

no longer a single bit but a field element represented as a bit string. For simplicity, we still call the new proof Π: the difference is that now $q = O(1)$.

The benefits of composing proof systems are now obvious. Let us try our hand at composing S_1 with itself, i.e., S_1 with S_2, where S_2 denotes S_1. The verifiers for S_1 and S_2 need $O(\log n)$ and $O(\log(\log n)^{O(1)})$ random bits, respectively, i.e., a total of $O(\log n)$ of them. Obviously, the number of queries remains $O(1)$. All the verification work, being now done by S_2, amounts to $(\log(\log n)^{O(1)})^{O(1)}$, i.e., $(\log\log n)^{O(1)}$. This bound can be further reduced by iterating the composition, but this is not the way to go to make it $O(1)$. For that, we take the previous system, call it S_3, and compose it with S_2. This requires $O(\log n) + (\log\log n)^{O(1)} = O(\log n)$ random bits, $O(1)$ queries and $O(1)$ amount of work. It follows that $O(1)$ bits are read in the proof, and the PCP theorem is proven.

But is it really? The task of V_1 is not *only* to check that Φ_s is satisfiable but that $\Pi_{i_1} \cdots \Pi_{i_q}$ is part of a satisfying assignment.

Here is a simple illustration of the conundrum we face. Say, a prover claims to have a satisfying assignment for $\bigwedge_i C_i$. A verifier might want to check this by picking C_i at random and verifying that the assignment makes C_i true. But suppose that, instead, it chooses to delegate that task to some other PC proof system. A second verifier will then take C_i as input and check that it is satisfiable. But any disjunction of three literals, such as C_i, is always satisfiable. What needs to be checked is not whether C_i is satisfiable on its own, but whether it is by using the assignment *specified* by S_1. Returning to V_2, its job is not to check that Φ_s is satisfiable but that there exists X_s such that $\Pi_{i_1} \cdots \Pi_{i_q} \cdot X_s$ makes Φ_s true. To resolve this *consistency* issue is key to making self-reducibility work. This is easy to do; in fact, by reducing the number of queries to $O(1)$, we have done the hardest part already.

We sketch what remains to be done.

For the verification to be delegated to V_2, of course, it is necessary to encode $\Pi_{i_1} \cdots \Pi_{i_q} X_s$ into the format $\sigma(\Pi_{i_1} \cdots \Pi_{i_q} X_s)$ that V_2 expects. It might be tempting to simply append $\sigma(\Pi_{i_1} \cdots \Pi_{i_q} X_s)$ at the end of Π, but doing so would raise the consistency problem mentioned earlier. Instead, we must effectively *replace* Π (and not just add on to it) with the encoding σ of every possible string $\Pi_{i_1} \cdots \Pi_{i_q} X_s$. But to do so would cause the same Π_{i_j} to appear in different encodings throughout the proof, again raising consistency issues. The solution is to encode each Π_{i_j} separately. Specifically we replace each Π_{i_j} by $\sigma(\Pi_{i_j})$. Likewise, we encode X_s as $\sigma(X_s)$.

This solves one problem, consistency, only to create another one. In this scheme, V_2 does not have access to $\sigma(\Pi_{i_1} \cdots \Pi_{i_q} X_s)$, which is the only encoding it can read, but to $\sigma(\Pi_{i_1}) \cdots \sigma(\Pi_{i_q}) \sigma(X_s)$. Is that good enough? Instead of encoding a whole truth assignment $a_1 \cdots a_n$ via σ, suppose we encode it in chunks: first $\sigma(a_1 \cdots a_{i_1})$, then $\sigma(a_{i_1+1} \cdots a_{i_2})$, etc, and finally $\sigma(a_{i_{q-1}+1} \cdots a_n)$. Can the verifier deal with that

sort of *split form* encoding? The answer is yes. It is, in fact, a rather simple exercise to modify V_2 accordingly.

This completes the proof of the PCP theorem or, at least, of its conceptual outline. The doubting reader can always sample the proof at random and see if that helps...

5. HISTORICAL NOTES

Following the seminal work of Goldwasser, Micali, and Rackoff [20] and Babai [5], which introduced the notion of interactive proofs, an important variant was introduced by Ben-Or et al. [9], in which the verifier interacts with not one but several provers. This framework led to early incarnations of PCP systems by Fortnow, Rompel and Sipser [14]. The idea of putting tight resource restrictions on both the verifier (query time) and the proof (size) originated in the works of Babai et al. [6] and Feige et al. [13]. The algebraic view of Boolean expressions gained currency in a series of papers that highlighted the enormous expressive power of interactive proofs [7, 24]. Turning to co-NP, Lund et al. [24] explained how a prover can convince a verifier that a graph is *not* 3-colorable. (By contrast, to convince someone that a graph is 3-colorable is trivial.) Finally, the ultimate power of interaction was resolved by Shamir [30], who proved that languages with interactive proofs are precisely those that can be decided in polynomial space.

The current notion of PCP itself, with its focus on randomness and query complexity, was formally introduced by Arora and Safra [4]. This development was spurred in large measure by the key insight of Feige et al. [13], which for the first time tied probabilistic proof systems to inapproximability. Babai, Fortnow, and Lund [7] established that PCP$[n^{O(1)}, n^{O(1)}]$ coincides with the class of problems solvable in nondeterministic exponential time. Babai et al. [6] and Feige et al. [13] essentially showed that NP is contained in PCP[polylog,polylog] (the precise bounds being somewhat stronger). Arora and Safra [4] proved that NP = PCP$[\log n, \sqrt{\log n}]$, and introduced the powerful concept of proof composition. The PCP theorem itself, i.e., NP = PCP$[\log n, 1]$, was proven by Arora et al. [3]. Finetuning the constants followed in quick order. Håstad [22] proved the striking result that three queries are sufficient as long as we can tolerate an ε chance of rejecting a correct proof. Building on that result, Guruswami et al. [21] showed that such false-negatives can be avoided provided that the error probability for wrongly accepting is $1/2 + \varepsilon$.

The connection to inapproximability [13] blossomed into a plethora of hardness results, one of the most impressive being Håstad's proof [23] that to approximate the clique number of an n-node graph within a factor of $n^{1-\varepsilon}$ is impossible (unless NP coincides with the randomized version of P, i.e., the class of sets for which membership can be decided in expected polynomial time by a randomized, error-free Turing

machine). At the other end of the spectrum, consider MaxCut, the problem of partitioning the node set of a graph into two subsets with the maximum number of edges joining them. It is possible to find a solution in polynomial time that has a number of edges at least 0.878 times the maximum possible [17]. On the other hand, to push that approximation factor above 0.942 would require that P = NP [22, 33] (building on [8]). A comprehensive 1996 survey of approximation results was compiled by Arora and Lund [2].

Many of the tools for checking the internal consistency of proof systems originated in the area of *program checking* [10, 11, 29]. For example, the low degree test, due to Arora et al. [3], incorporates ideas from [4, 11, 16, 28].

The sumcheck and linearity tests are due respectively to Lund et al. [24] and Blum, Luby, and Rubinfeld [11]. Testing that a polynomial is nonzero (Lemma 3.2) is from [6, 13].

Essential tools in PCP-related work also include the *parallel repetition theorem* by Raz [27], the *long code* by Bellare, Goldreich and Sudan [8], and Fourier transform techniques by Håstad [22]. For background material in complexity theory, the following texts [15, 31, 26, 12], listed in increasing order of technical depth, are excellent entry points. We also mention [1, 2, 25, 32] for in-depth coverage of proof verification and approximation algorithms, and [18, 19, 20] for an introduction to zero-knowledge cryptography.

Acknowledgments. — I wish to thank Lance Fortnow, Oded Goldreich, and Muli Safra for their helpful comments on this manuscript.

REFERENCES

[1] S. Arora – Probabilistic Checking of Proofs and the Hardness of Approximation Problems, Ph.D. Thesis, UC Berkeley, 1994, also available as `http://www.cs.princeton.edu/~arora`.

[2] S. Arora & C. Lund – Hardness of approximations, in *Approximation Algorithms for NP-hard Problems* (D. Hochbaum, ed.), PWS Publishing, 1996.

[3] S. Arora, C. Lund, R. Motwani, M. Sudan & M. Szegedy – Proof verification and the hardness of approximation problems, *J. ACM* **45** (1998), p. 501–555.

[4] S. Arora & M. Safra – Probabilistic checking of proofs: a new characterization of NP, *J. ACM* **45** (1998), p. 70–122.

[5] L. Babai – Trading group theory for randomness, in *Proc. 17th Annual ACM, Symp. Theory Comput.*, 1985, p. 421–429.

[6] L. Babai, L. Fortnow, L. Levin & M. Szegedy – Checking computations in polylogarithmic time, in *Proc. 23rd Annual ACM Symp. Theory Comput.*, 1991, p. 21–31.

[7] L. Babai, L. Fortnow & C. Lund – Non-deterministic exponential time has two-prover interactive protocols, *Computational Complexity* **1** (1991), p. 3–40.

[8] M. Bellare, O. Goldreich & M. Sudan – Free bits, PCPs and non-approximability—towards tight results, *SIAM J. Comput.* **27** (1998), p. 804–915.

[9] M. Ben-Or, S. Goldwasser, J. Kilian & A. Wigderson – Multi-prover interactive proofs: how to remove intractability, in *Proc. 20th Annual ACM Symp. Theory Comput.*, 1988, p. 113–131.

[10] M. Blum & S. Kannan – Designing programs that check their work, in *Proc. 21st Annual ACM Symp. Theory Comput.*, 1989, p. 86–97.

[11] M. Blum, M. Luby & R. Rubinfeld – Self-testing/correcting with applications to numerical problems, *J. Comp. Sys. Sci.* **47** (1993), p. 549–595.

[12] D.-Z. Du & K.-I. Ko – *Theory of Computational Complexity*, Wiley-Interscience, 2000.

[13] U. Feige, S. Goldwasser, L. Lovasz, S. Safra & M. Szegedy – Interactive proofs and the hardness of approximating cliques, *J. ACM* **43** (1996), p. 268–292.

[14] L. Fortnow, J. Rompel & M. Sipser – On the power of multi-prover interactive protocols, *Theoret. Comput. Sci.* **134** (1994), p. 545–557.

[15] M. Garey & D. Johnson – *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, New York, 1979.

[16] P. Gemmell, R. Lipton, R. Rubinfeld, M. Sudan & A. Wigderson – Self-testing/correcting for polynomials and approximate functions, in *Proc. 23rd Annual ACM Symp. Theory Comput.*, 1991, p. 32–42.

[17] M.X. Goemans & D.P. Williamson – Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *J. ACM* **42** (1995), p. 1115–1145.

[18] O. Goldreich – *Modern Cryptography, Probabilistic Proofs and Pseudorandomness*, Algorithms and Combinatorics, vol. 17, Springer, 1999.

[19] O. Goldreich, S. Micali & A. Wigderson – Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems, *J. ACM* **38** (1991), p. 691–729.

[20] S. Goldwasser, S. Micali & C. Rackoff – The knowledge complexity of interactive proof systems, *SIAM J. Comput.* **18** (1989), p. 186–208.

[21] V. Guruswami, D. Lewin, M. Sudan & L. Trevisan – A tight characterization of NP with 3 query PCPs, in *Proc. 39th Annual IEEE Symp. Found. Comput. Sci.*, 1998, also available as ECCC Technical Report TR98-034, p. 8–17.

[22] J. Håstad – Some optimal inapproximability results, in *Proc. 29th Annual ACM Symp. Theory Comput.*, 1997, also available as ECCC Technical Report TR97-037, p. 1–10.

[23] ———, Clique is hard to approximate within $n^{1-\varepsilon}$, *Acta Mathematica* **182** (1999), p. 105–142.

[24] C. Lund, L. Fortnow, H. Karloff & N. Nisan – Algebraic methods for interactive proof systems, *J. ACM* **39** (1992), p. 859–868.

[25] E.W. Mayr, H.J. Promel & A. Steger – *Lectures on Proof Verification and Approximation Algorithms*, LNCS, vol. 1367, Springer Verlag, 1998.

[26] C.H. Papadimitriou – *Computational Complexity*, Addison Wesley, 1994.

[27] R. Raz – A parallel repetition theorem, *SIAM J. Comput.* **27** (1998), p. 763–803.
[28] R. Rubinfeld & M. Sudan – Self-testing polynomial functions efficiently and over rational Domains, in *Proc. 3rd Annual ACM/SIAM Symp. Discrete Algorithms*, 1992, p. 23–32.
[29] ———, Robust characterizations of polynomials with applications to program testing, *SIAM J. Comput.* **25** (1996), p. 252–271.
[30] A. Shamir – IP = PSPACE, *J. ACM* **39** (1992), p. 869–877.
[31] M. Sipser – *Introduction to the Theory of Computation*, PWS Publishing, 1997.
[32] M. Sudan – Efficient Checking of Polynomials and Proofs and the Hardness of Approximation Problems, in *ACM Distinguished Dissertation Series for 1993*, Springer, 1996.
[33] L. Trevisan, G.B. Sorkin, M. Sudan & D.P. Williamson – Gadgets, Approximation, and Linear Programming, *SIAM J. Comput.* **29** (2000), p. 2074–2097.

Bernard CHAZELLE

Princeton University
Department of Computer Science
Princeton, NJ 08544
USA
E-mail : chazelle@cs.princeton.edu

Séminaire BOURBAKI Novembre 2001
54e année, 2001-2002, no 896, p. 37 à 52

IWASAWA ALGEBRAS AND ARITHMETIC

by **John COATES**

1. INTRODUCTION

Let p be a prime number, and G a compact p-adic Lie group. We recall that the Iwasawa algebra of G is defined by

$$\Lambda(G) = \varprojlim_{U} \mathbb{Z}_p\,[G/U]$$

where U runs over the open normal subgroups of G. Any compact $\mathbb{Z}_p$-module on which G acts continuously on the left has a unique structure as a left $\Lambda(G)$-module, extending the G-action. Thanks to this remark, modules over $\Lambda(G)$, where G is usually the image of Galois in a finite dimensional p-adic Galois representation, abound in arithmetic geometry. K. Iwasawa [Iw] was the first to study the structure theory of finitely generated $\Lambda(G)$-modules in the special case when $G = \mathbb{Z}_p$, and deduced from it his celebrated asymptotic formula for the growth of the order of the p-primary subgroup of the ideal class group in a $\mathbb{Z}_p$-extension of a number field. Almost immediately, J-P. Serre [Se1], [Se2] pointed out that, when $G = \mathbb{Z}_p^d$ for any integer $d \geqslant 1$, $\Lambda(G)$ is isomorphic to the local ring $\mathbb{Z}_p[[T_1, \ldots, T_d]]$ of formal power series in d variables with coefficients in $\mathbb{Z}_p$, and that Iwasawa's structure theorem for $\Lambda(G)$-modules could be re-proven for $d = 1$, and generalized to all $d \geqslant 1$, by using classical arguments in commutative algebra about the structure theory of modules up to pseudo-isomorphism ([B-CA], Chap. VII, § 4.4, Theorems 4 and 5).

Intuitively, one might expect that the structure theory for $\Lambda(G)$-modules would be very different in the commutative and non-commutative cases, but the aim of this seminar is to report on joint work of the author, P. Schneider and R. Sujatha [CSS] proving that, surprisingly, the two cases appear to be parallel in many ways. The first step towards elucidating the non-commutative theory was made by O. Venjakob [Ve1], [Ve2], who exploited ideas of J. Björk [Bj] to define in general the notion of a pseudo-null $\Lambda(G)$-module. If G is pro-p and has no element of order p, Venjakob defines

a finitely generated left $\Lambda(G)$-module M to be pseudo-null if it is $\Lambda(G)$-torsion (i.e. each element of M is annihilated by some non-zero element of $\Lambda(G)$), and, in addition, $Ext^1_{\Lambda(G)}(M, \Lambda(G)) = 0$. To establish our structure theory up to pseudo-isomorphism, we need to impose further conditions on G, and we are grateful to B. Totaro for pointing out to us that probably the most natural hypothesis is that G should possess a p-valuation in the sense of M. Lazard [La]. We recall that a p-valuation on G is a function $\omega : G \to (0, \infty]$ satisfying the following axioms for all x and y in G:

(i) $\omega(1) = \infty$, and $\frac{1}{p-1} < \omega(x) < \infty$ for $x \neq 1$;
(ii) $\omega(xy^{-1}) \geqslant \min\{\omega(x), \omega(y)\}$;
(iii) $\omega(x^{-1}y^{-1}xy) \geqslant \omega(x) + \omega(y)$;
(iv) $\omega(x^p) = \omega(x) + 1$.

We say that G is p-valued if it possesses a p-valuation. If G is p-valued, we remark that the compactness of G guarantees that G is complete with respect to the p-valuation ω in the following sense. For each $u > 0$, let G_u denote the subgroup of G consisting of all g such that $\omega(g) \geqslant u$. As J-P. Serre observed to us, G_u is open in G because, choosing $N > u$, G_u contains the subgroup of G generated by the p^N-th powers, and it is well known that this latter subgroup is a neighbourhood of the identity in a p-adic Lie group. Hence the family $\{G_u : u > 0\}$ form an open basis for the topology of G since their intersection is trivial, and the natural map from G to $\varprojlim G/G_u$ is an isomorphism because of the compactness of G. Moreover, Lazard [La] established the following basic facts. Any closed subgroup of a p-valued group is also p-valued. If G is p-valued, then it is pro-p, and has no element of order p. The classic example of a p-valued group is the group of matrices in $GL_n(\mathbb{Z}_p)$ which are congruent to the identity modulo p (resp. mod 4) if p is odd (resp. if $p = 2$). If $p > n + 1$, any pro-p closed subgroup of $GL_n(\mathbb{Z}_p)$ is p-valued.

THEOREM 1.1 ([CSS]). — *Let G be a p-valued compact p-adic Lie group, and let M be a finitely generated torsion $\Lambda(G)$-module. Let M_0 be the maximal pseudo-null submodule of M. Then there exist non-zero left ideals $L_1, \ldots, L_m$, and a $\Lambda(G)$-injection*

$$\phi : \bigoplus_{i=1}^{m} \Lambda(G)/L_i \longrightarrow M/M_0,$$

with Coker(ϕ) *pseudo-null.*

The special case of Theorem 1.1 in which M/M_0 is killed by some power of p was proven earlier by O. Venjakob [Ve1], [Ve2], and S. Howson [Ho]. In § 2, we shall give a sketch of a proof of Theorem 1.1 taken from [CSS], which is remarkably parallel to the classical commutative proof in [B-CA], and which exploits the fact that $\Lambda(G)$ is a filtered ring to which one can apply the techniques of the algebraic theory of microlocalization (see, for example, [LO]). After finding this proof, we also realized that Theorem 1.1 can be derived from the work of M. Chamarie [Ch1], [Ch2], on modules over maximal orders (see [CSS] for the details).

We assume for the rest of this exposé that G is a p-valued compact p-adic Lie group. In particular, it follows that $\Lambda(G)$ is Noetherian, and has no zero divisors. Let $\mathrm{Mod}(G)$ denote the category of all finitely generated left $\Lambda(G)$-modules, and $C^1(G)$ the subcategory whose objects are the pseudo-null modules ($C^1(G)$ is closed under taking subobjects, quotients, and extensions). To discuss questions about the uniqueness of the decomposition in Theorem 1.1, we have to pass to the quotient category

$$\mathfrak{M}(G) = \mathrm{Mod}(G)/C^1(G).$$

We write $Q : \mathrm{Mod}(G) \longrightarrow \mathfrak{M}(G)$ for the canonical functor. If M is an object of $\mathrm{Mod}(G)$, we define its annihilator, which we denote by $\mathrm{ann}_{\Lambda(G)}(M)$, to be the two sided ideal consisting of all r in $\Lambda(G)$ such that $r.M = 0$. We then define the annihilator of the object $Q(M)$ of the quotient category $\mathfrak{M}(G)$, which we denote by $\mathrm{ann}(Q(M))$, to be the sum of all the ideals $\mathrm{ann}_{\Lambda(G)}(N)$, where N runs over all objects of $\mathrm{Mod}(G)$ such that $Q(N)$ is isomorphic to $Q(M)$ in $\mathfrak{M}(G)$. In fact, a lemma of Robson [Ro] shows that

$$\mathrm{ann}(Q(M)) = \mathrm{ann}_{\Lambda(G)}(M/M_0),$$

where, as above, M_0 denotes the maximal pseudo-null submodule of M. Yet another description of $\mathrm{ann}(Q(M))$ can be given in terms of the left ideals $L_1, \ldots, L_m$ appearing in Theorem 1.1. Let J_i be the maximal two-sided ideal of $\Lambda(G)$ which is contained in L_i, and let $J = \cap_{i=1}^m J_i$. Then $J_i = \mathrm{ann}_{\Lambda(G)}(\Lambda(G)/L_i)$, and we have

$$J = \mathrm{ann}(Q(M));$$

in particular, we see that $J \neq 0$ if and only if $J_i \neq 0$ for $i = 1, \ldots, m$.

It is in questions of annihilators that we find a basic difference between the commutative and non-commutative case. R. Greenberg (unpublished) has given an example of a p-valued open subgroup of $GL_2(\mathbb{Z}_p)$, and a finitely generated torsion $\Lambda(G)$-module M such that $\mathrm{ann}(Q(M)) = 0$. Following Chamarie [Ch2], we therefore define $Q(M)$ to be *bounded* (resp. *completely faithful*) if $\mathrm{ann}(Q(M)) \neq 0$ (resp. if $\mathrm{ann}(Q(N)) = 0$ for every torsion $\Lambda(G)$-module N such that $Q(N)$ is a non-zero subquotient of $Q(M)$). It is proven in [Ch2] that, for every finitely generated torsion $\Lambda(G)$-module M, we have a canonical decomposition

$$Q(M) = Q(U) \oplus Q(V),$$

where $Q(U)$ is completely faithful and $Q(V)$ is bounded. Very little is known about completely faithful objects in $\mathfrak{M}(G)$ beyond the fact that Greenberg's example shows that they exist, and also it is shown in [Ch2] that they are cyclic, i.e. isomorphic in $\mathfrak{M}(G)$ to $Q(\Lambda(G)/L)$ where L is a non-zero left ideal. However, Y. Hachimori and O. Venjakob [HV] have recently given examples of completely faithful $\Lambda(G)$-modules which arise naturally in arithmetic geometry, and one suspects that their occurrence in number theory may be rather common.

We write $\mathfrak{M}_b(G)$ for the full subcategory of $\mathfrak{M}(G)$ consisting of the bounded objects, and we define $\mathcal{D}_b(G)$ to be the Grothendieck group of $\mathfrak{M}_b(G)$. In fact, $\mathfrak{M}_b(G)$ is also an abelian category in which every object has finite length, and the Jordan-Hölder theorem shows that $\mathcal{D}_b(G)$ is the free abelian group on the set of isomorphism classes of simple objects in $\mathfrak{M}_b(G)$. It is natural to ask whether we can relate $\mathcal{D}_b(G)$ to a natural group of divisors of the ring $\Lambda(G)$, parallel to the classical theory for commutative, integrally closed, integral domains ([B-CA], Chap. VII, § 4.5, Proposition 11). As we shall now explain, this is indeed the case. Let $K(G)$ denote the skew field of fractions of $\Lambda(G)$, which is well known to exist because $\Lambda(G)$ is Noetherian and has no divisors of zero. Then $\Lambda(G)$ is a *maximal order* (this is the non-commutative analogue of being integrally closed) in the sense that, if B is any intermediate ring with $\Lambda(G) \subset B \subset K(G)$ such that there exist non-zero elements u, v in $K(G)$ with $uBv \subset \Lambda(G)$, then necessarily $B = \Lambda(G)$ (see [CSS], Lemma 2.6). For any left (resp. right) $\Lambda(G)$- module M, we put $M^* = \mathrm{Hom}_{\Lambda(G)}(M, \Lambda(G))$ for the dual right (resp. left) $\Lambda(G)$-module, and we say M is *reflexive* if the natural map from M to M^{**} is an isomorphism. A non-zero left (resp. right) $\Lambda(G)$-submodule L of $K(G)$ is called a *fractional* left (resp. right) ideal if there is a non-zero v in $K(G)$ such that $L \subset \Lambda(G)v$ (resp. $L \subset v\Lambda(G)$). A *fractional ideal* of $\Lambda(G)$ is a subset I of $K(G)$ which is both a fractional left and a fractional right ideal. Finally, we define a fractional c-ideal of $\Lambda(G)$ to be a reflexive fractional ideal of $\Lambda(G)$, and we write $\mathcal{C}(G)$ for the set of fractional c-ideals. As a special case of general results about maximal orders, Asano ([As]) has shown that $\mathcal{C}(G)$ is an abelian group with respect to the product $I.J = (IJ)^{**}$. We recall that a two-sided ideal $\mathfrak{p}$ of $\Lambda(G)$ is said to be *prime* if, whenever x and y are elements of $\Lambda(G)$ such that $x\Lambda(G)y \subset \mathfrak{p}$, we always have x is in $\mathfrak{p}$ or y is in $\mathfrak{p}$. It is then also proven in [As] that $\mathcal{C}(G)$ is the free abelian group on the set $\mathcal{P}$ of all non-zero prime c-ideals, and that every prime c-ideal has height 1 (i.e. is a minimal non-zero prime ideal). There would be great interest in giving an explicit description of this set $\mathcal{P}$ (for example, when G is a p-valued open subgroup of $SL_2(\mathbb{Z}_p)$).

Our aim is to construct a canonical homomorphism

$$\chi : \mathcal{D}_b(G) \to \mathcal{C}(G),$$

and for this we need to localize $\Lambda(G)$ at the prime ideals in $\mathcal{P}$. We recall that a multiplicatively closed subset S of non-zero elements of $\Lambda(G)$ is said to be a right and left Ore set if, for each s in S and a in $\Lambda(G)$ both $aS \cap s\Lambda(G)$ and $Sa \cap \Lambda(G)s$ are non-empty. For each $\mathfrak{p}$ in $\mathcal{P}$, let $S(\mathfrak{p})$ denote the set of all elements of $\Lambda(G)$ whose residue class in $\Lambda(G)/\mathfrak{p}$ is not a zero divisor. Chamarie [Ch1] has proven that $S(\mathfrak{p})$ is a left and right Ore set, and that the localization of $\Lambda(G)$ by $S(\mathfrak{p})$, which we denote by $\Lambda(G)_{\mathfrak{p}}$, is a bounded maximal order, with Jacobson radical $\mathfrak{p}\Lambda(G)_{\mathfrak{p}}$. Moreover, every left and right ideal in $\Lambda(G)_{\mathfrak{p}}$ is principal. Now, up to isomorphism, the objects of finite length in $\mathfrak{M}_b(G)$ are of the form $Q(M)$, where M is a finitely

generated torsion $\Lambda(G)$-module. Hence the localization $M_{\mathfrak{p}} = \Lambda(G)_{\mathfrak{p}} \otimes_{\Lambda(G)} M$ is a finitely generated torsion $\Lambda(G)_{\mathfrak{p}}$-module, which has finite length because $\Lambda(G)_{\mathfrak{p}}$ is a principal ideal domain. Denoting the length of $M_{\mathfrak{p}}$ by $\ell_{\mathfrak{p}}(Q(M))$, we then define

$$\chi(Q(M)) = \prod_{\mathfrak{p} \in \mathcal{P}} \mathfrak{p}^{\ell_{\mathfrak{p}}(Q(M))},$$

and we call $\chi(Q(M))$ the *characteristic ideal* of $Q(M)$. This is well defined, as it is proven in [Ch2], Lemma 4.2.1 that, for any finitely generated torsion $\Lambda(G)$-module M, we have that $M_{\mathfrak{p}} = 0$ for all $\mathfrak{p}$ in $\mathcal{P}$ if and only if $Q(M)$ is completely faithful. Moreover, it is shown in [CSS] that $\mathrm{ann}(Q(M))$ is a c-ideal such that $\chi(Q(M)) \subset \mathrm{ann}(Q(M))$, and such that $\chi(Q(M))$ and $\mathrm{ann}(Q(M))$ have the same prime factors in $\mathcal{P}$. In addition, the exactness of localization and the additivity of the length function show that χ induces a homomorphism from $\mathcal{D}_b(G)$ to $\mathcal{C}(G)$.

THEOREM 1.2 ([CSS]). — *The homomorphism*

$$\chi : \mathcal{D}_b(G) \longrightarrow \mathcal{C}(G)$$

is an isomorphism. In particular, χ induces a bijection between the set of isomorphism classes of simple objects in $\mathfrak{M}_b(G)$ and the set $\mathcal{P}$ of all non-zero prime c-ideals of $\Lambda(G)$.

As far as all finitely generated $\Lambda(G)$-modules are concerned, only the following partial result is proven in [CSS]. If M is an arbitrary finitely generated $\Lambda(G)$-module, we write M_t for its $\Lambda(G)$-torsion submodule. It is shown in [Ve2] that the natural map from $M/M_t \to (M/M_t)^{**}$ is injective and has pseudo-null cokernel.

THEOREM 1.3 ([CSS]). — *Let M be a finitely generated $\Lambda(G)$-module such that $Q(M_t)$ is bounded. Then we have an isomorphism*

$$Q(M) \xrightarrow{\sim} Q(M_t) \oplus Q(M/M_t),$$

where $Q(M/M_t)$ is reflexive in the sense that it is isomorphic to $Q((M/M_t)^{**})$ in the quotient category $\mathfrak{M}(G)$.

The fundamental question left open by [CSS] is whether every prime ideal $\mathfrak{p}$ in $\mathcal{P}$ is principal. This is true when $G = \mathbb{Z}_p^d$, for any integer $d \geqslant 1$, thanks to the Weierstrass Preparation Theorem, and we strongly suspect that it remains true for all compact, p-valued p-adic Lie groups G.

2. SKETCH OF THE PROOF OF THEOREM 1.1

One of the nicest parts of Bourbaki's treatise on commutative algebra is his elegant proof of the analogue of Theorem 1.1 for all finitely generated torsion modules over any Noetherian, integrally closed, integral domain (see [B-CA], Chap. VII, § 4.4, Theorem 5). We now briefly explain how simple ideas from the algebraic theory of

micro-localization allow one to extend these arguments to modules over a wide class of non-commutative filtered rings. Following the spirit of Bourbaki, we proceed axiomatically, and leave to the end of this section the verification that our concrete ring $\Lambda(G)$ satisfies the axioms we impose.

Let A be an associative ring with unit elements, which will not, in general, be commutative. We assume that A is endowed with a filtration $F_\bullet A = \{F_n A : n \in \mathbb{Z}\}$, but we shall be unfaithful to Bourbaki and always assume our filtrations are increasing, i.e. $F_n A \subset F_{n+1} A$ for all n in $\mathbb{Z}$. This filtration will always be assumed to be exhaustive (i.e. $\bigcup_{n\in\mathbb{Z}} F_n A = A$) and separated (i.e. $\bigcap_{n\in\mathbb{Z}} F_n A = 0$). We write

$$\mathrm{gr}_\bullet A = \bigoplus_{n\in\mathbb{Z}} F_n A/F_{n-1}A, \qquad \widehat{A} = \varprojlim A/A_n$$

for the associated graded ring, and the completion of A with respect to the filtration, respectively. We follow the non-commutative literature ([LO], Chap. II, Theorem 2.2), and define A to be a *Zariski ring* if $\mathrm{gr}_\bullet A$ is left and right Noetherian, and $\widehat{A}$ is a faithfully flat left and right A-module.

For the rest of this section, we shall assume that the ring A satisfies the following axioms:

(C1) A is complete with respect to $F_\bullet A$, i.e. the natural injection from A to $\widehat{A}$ is an isomorphism;

(C2) $\mathrm{gr}_\bullet A$ is isomorphic as a graded ring to $k[T_1, \ldots, T_r]$, the ring of polynomials in a finite number of variables with coefficients in a field k, graded by assigning to each of the variables a strictly negative integer as its degree.

Axioms (C1) and (C2) imply that A is left and right Noetherian and has no zero divisors, and that A is a Zariski ring.

For the remainder of the proof, M will denote an arbitrary finitely generated torsion A-module. We endow M with a *good filtration* $F_\bullet M = \{F_n M : n \in \mathbb{Z}\}$. This means that we have

$$F_n M = \sum_{i=1}^{r} F_{n-k_i} A \cdot w_i \qquad (n \in \mathbb{Z}),$$

where $w_1, \ldots, w_r$ is some fixed set of A-generators of M, and $k_1, \ldots, k_r$ are fixed integers. Since $F_\bullet M$ is a good filtration, basic properties of Zariski rings show that not only is $F_\bullet M$ separated, but also any submodule N of M is closed in the filtration topology [LO]). Also, defining the $\mathrm{gr}_\bullet A$ module $\mathrm{gr}_\bullet M$ as usual by

$$\mathrm{gr}_\bullet M = \bigoplus_{n\in\mathbb{Z}} F_n M/F_{n-1}M,$$

we have that $\mathrm{gr}_\bullet M$ is a finitely generated $\mathrm{gr}_\bullet A$-module, which is plainly $\mathrm{gr}_\bullet A$-torsion.

The starting point of our proof is to apply the classical commutative theory to the finitely generated torsion $\mathrm{gr}_\bullet A$-module $\mathrm{gr}_\bullet M$. We write Ass $(\mathrm{gr}_\bullet M)$ for the set of prime ideals $\mathfrak{p}$ in $\mathrm{gr}_\bullet A$ which are the exact annihilator of some non-zero element of

$\mathrm{gr}_\bullet M$. Note that the zero ideal is not in Ass $(\mathrm{gr}_\bullet M)$, because $\mathrm{gr}_\bullet M$ is $\mathrm{gr}_\bullet A$-torsion. As $\mathrm{gr}_\bullet M$ is a graded $\mathrm{gr}_\bullet A$-module, every ideal in Ass $(\mathrm{gr}_\bullet M)$ is graded. We define

$$W(M) = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_m\}$$

to be the set of prime ideals of height 1 in Ass $(\mathrm{gr}_\bullet M)$. It is not difficult to see that $W(M)$ is independent of the very non-canonical choice of the good filtration $F_\bullet M$. We define $S = S(M)$ to be the set of all non-zero homogeneous elements of $\mathrm{gr}_\bullet A$ which do not belong to $\mathfrak{p}_1 \cup \cdots \cup \mathfrak{p}_m$. We can localize $\mathrm{gr}_\bullet A$ with respect to the multiplicative set of S of non-zero homogeneous elements, and obtain in this way the graded ring $S^{-1}\mathrm{gr}_\bullet A$ ([B-CA], Chap. II, § 2.9), which we denote by $(\mathrm{gr}_\bullet A)_S$. The following lemma (see [CSS], Proposition 3.4) is easily established.

LEMMA 2.1. — *The non-zero graded prime ideals of* $(\mathrm{gr}_\bullet A)_S$ *are precisely the* $S^{-1}\mathfrak{p}_i$ $(1 \leqslant i \leqslant m)$*, and these all have height* 1*. Every proper graded ideal of* $(\mathrm{gr}_\bullet A)_S$ *is contained in one of the* $S^{-1}\mathfrak{p}_i$ $(1 \leqslant i \leqslant m)$*.*

COROLLARY 2.2. — *Every graded ideal in* $(\mathrm{gr}_\bullet A)_S$ *is principal.*

To deduce the corollary, we first note that, as A is factorial by axiom (C2), so is its localization $(\mathrm{gr}_\bullet A)_S$. Now, if $\mathfrak{b}$ is any non-zero graded ideal of $(\mathrm{gr}_\bullet A)_S$, every element of Ass $((\mathrm{gr}_\bullet A)_S/\mathfrak{b})$ must be graded prime ideal of $(\mathrm{gr}_\bullet A)_S$, and therefore of height 1 by Lemma 2.1. But then (see [B-CA], Chap. VII, § 1.6, Proposition 10) the ideal $\mathfrak{b}$ is divisorial, and so principal because $(\mathrm{gr}_\bullet A)_S$ is factorial.

The heart of our proof is the following observation on Ore sets, which we understand goes back to Kashiwara [Ka]. As usual, if $x \in F_n A \backslash F_{n-1}A$, we define its principal symbol $\sharp(x)$ by $\sharp(x) = x + F_{n-1}A$. Let us define

$$T = \{t \in A : \sharp(t) \in S\};$$

here S is the multiplicative set of non-zero elements of A defined above.

PROPOSITION 2.3. — *T is a left and right Ore set in A.*

We omit the proof (see [Li] or the last part of [WK]), which uses the fact that A is a Zariski ring. In view of Proposition 2.3, it makes sense to take either the left or the right localization of A with respect to T. As they are isomorphic, we write A_T for either localization. Moreover, as is explained in [Li], A_T is endowed with a natural separated and exhaustive filtration $F_\bullet A_T$, with the property that $\mathrm{gr}_\bullet(A_T) = (\mathrm{gr.A})_S$. Even though we have imposed the axiom that A is complete, it will not in general be true that A_T is complete with respect to the filtration $F_\bullet A_T$. Nevertheless, the filtration $F_\bullet A_T$ makes A_T into a Zariski ring (see [Li]), and this weaker result suffices for our purposes. Finally, if N is any finitely generated A-module endowed with a good filtration $F_\bullet N$, then its localization $N_T = A_T \otimes_A N$ is a finitely generated A_T-module, which (see [Li]) is also endowed with a natural filtration $F_\bullet N_T$ such that $\mathrm{gr}_\bullet(N_T) = (\mathrm{gr}_\bullet N)_S$.

PROPOSITION 2.4. — *Every left and right ideal in A_T is principal.*

To prove this, we can, by symmetry, restrict our attention to left ideals L of A_T. Taking any such left ideal, we endow it with the induced filtration $F_\bullet L$ given by $F_n L = L \cap F_n A_T$. By a basic property of Zariski rings, $F_\bullet L$ is again a good filtration. Plainly $\mathrm{gr}_\bullet L$ is a graded ideal in $\mathrm{gr}_\bullet A_T = (\mathrm{gr}_\bullet A)_S$, and so is principal by Corollary 2.2. Thus we can find a homogeneous z in $\mathrm{gr}_\bullet L$ such that $\mathrm{gr}_\bullet L = (\mathrm{gr}_\bullet A)_S \cdot z$. Now pick w to be any element of L such that $\sharp(w) = z$. But, thanks to a remarkable property of Zariski rings ([LO], Chap. I, § 5, Corollary 5.5), we conclude that $L = A_T w$ (note that we are able to carry out this last step without passing to the completion of A_T).

Once we have established that A_T is a principal ideal domain, we can rapidly complete the proof of Theorem 1.1, following closely the classical commutative argument. As the localized module $M_T = A_T \otimes_A M$ is a finitely generated torsion A_T-module, an old result of Jacobson (see [Ja], Chap. 3, Theorem 10) shows that there exist elements $w_1, \ldots, w_m$ in M_T such that

$$M_T = A_T w_1 \oplus \cdots \oplus A_T w_m.$$

Let $\psi : M \to M_T$ be the canonical A-homomorphism given by $\psi(m) = 1 \otimes m$, and let $M' = \mathrm{Im}(\psi)$, $N = \mathrm{Ker}(\psi)$. Since N is precisely the set of T-torsion elements of M, we have $N_T = 0$, and so N is pseudo-null by Proposition 2.5 below. Now M' is an A-submodule of M_T with $M'_T = M_T$. We are clearly free to multiply any of the elements $w_1, \ldots, w_m$ above by any element of T, and thus we can assume that $w_1, \ldots, w_m$ all belong to M'. We then define the A-submodule M'' of M' by

$$M'' = Aw_1 \oplus \cdots \oplus Aw_m,$$

where the sum is clearly direct because $w_1, \ldots, w_m$ are even linearly independent over A_T. But $N' = M'/M''$ is a quotient of M with $N'_T = 0$, and so N' is also pseudo-null by Proposition 2.5. Now the map $a \mapsto aw_i$ induces an isomorphism of A-modules from A/L'_i to Aw_i, where L'_i is the annihilator of w_i in A. The composed map

$$\rho : \bigoplus_{i=1}^{m} A/L'_i \xrightarrow{\sim} \bigoplus_{i=1}^{m} Aw_i = M'' \subset M' = M/N$$

is an injective A-homorphism with pseudo-null cokernel. Let L_i be the unique left ideal such that L_i/L'_i is the maximal pseudo-null submodule of A/L_i. We deduce easily that ρ induces an injection of A-modules

$$\varphi : \bigoplus_{i=1}^{m} A/L_i \longrightarrow M/M_0,$$

where M_0 is the maximal pseudo-null submodule of M, and $\mathrm{Coker}(\varphi)$ is pseudo-null. Thus we have established Theorem 1.1 with $\Lambda(G)$ replaced by our ring A satisfying axioms (C1) and (C2).

PROPOSITION 2.5. — *Let U be any A-subquotient of the A-torsion module M. Then $U_T = 0$ implies that U is pseudo-null. The converse statement holds if A is assumed to be Auslander regular.*

In fact, the notion of a module over an arbitrary ring A with identity element being pseudo-null is defined in [CSS]. We refer the reader to [CSS] for the somewhat delicate proof of Proposition 2.5, as well as a discussion of the notion of Auslander regularity.

The fact that $\Lambda(G)$ satisfies axioms (C1) and (C2) when G is a compact p-valued p-adic Lie group is a consequence of the following result, which is essentially contained in the important but difficult paper of M. Lazard [La]. The explanation given in [CSS] of how to derive this result from Lazard's work was given to us by B. Totaro. The last assertion of Proposition 2.6 is due to O. Venjakob [Ve1], [Ve2].

PROPOSITION 2.6. — *Assume that G is a compact p-adic Lie group, which is p-valued. Then $\Lambda(G)$ possesses a complete, separated and exhaustive filtration $F_\bullet\Lambda(G)$ such that $\mathrm{gr}_\bullet\Lambda(G)$ is isomorphic as a graded ring to the polynomial ring $\mathbb{F}_p[X_0, \ldots, X_d]$ in $d+1$ variables, where d is the dimension of G; here the grading on $\mathbb{F}_p[X_0, \ldots, X_d]$ is given by assigning to each of the variables X_i a strictly negative integer degree. In particular, $\Lambda(G)$ satisfies axioms (C1) and (C2). In addition, $\Lambda(G)$ is Auslander regular.*

3. ARITHMETIC EXAMPLES

Concrete examples of finitely generated torsion $\Lambda(G)$-modules, which are of great arithmetic interest, abound in arithmetic geometry. Because of lack of space, we only discuss two classes of examples. In both cases, G is non-commutative, and is the image of Galois in a 2-dimensional p-adic Galois representation; thus, by Lazard [La], G is automatically p-valued provided G is pro-p and $p \geqslant 5$.

Example 1. — Let $p \geqslant 5$, and let $\mu_{p^n} (1 \leqslant n \leqslant \infty)$ denote the group of all p^n-th roots of unity. We write F for any finite extension of $\mathbb{Q}$ containing μ_p, and define

$$F^{\mathrm{cyc}} = F(\mu_{p^\infty}), \quad \Gamma = G(F^{\mathrm{cyc}}/F).$$

Now fix a non-zero element α of F, which is not a root of unity, and define

$$K_\infty = F^{\mathrm{cyc}}(\alpha^{1/p^n} : n = 1, 2, \ldots), \quad G = G(K_\infty/F).$$

If we define H to be $G(K_\infty/F_{\mathrm{cyc}})$, then both H and Γ are isomorphic to $\mathbb{Z}_p$, so that G is a p-adic Lie group of dimension 2, which is p-valued. Moreover, G is not commutative. Let $\psi : \Gamma \to \mathbb{Z}_p^\times$ bethe character giving the action of Γ on μ_{p^∞}. Then, as α belongs to F, Kummer theory shows that the natural action of Γ on H via inner automorphism is given by the character ψ. One can study many left $\Lambda(G)$-modules

which are of arithmetic interest, but the simplest is probably the following. Let L_∞ denote the maximal unramified abelian p-extension of K_∞, and put $X = G(L_\infty/K_\infty)$. As usual, there is a continuous left action of G on X via inner automorphism (if σ is in G and x in X, we define $\sigma \cdot x = \widetilde{\sigma} x \widetilde{\sigma}^{-1}$, where $\widetilde{\sigma}$ denotes any lifting of σ to the Galois group of L_∞ over F). Y. Ochi [Oc] has proven that X is a finitely generated torsion $\Lambda(G)$-module. At present, very little else is known about the module $\Lambda(G)$; in particular, it seems that at present no example is known in which we can prove that X is not pseudo-null as a $\Lambda(G)$-module. In the special case $F = \mathbb{Q}(\mu_p)$ and $\alpha = p$, one can easily show that $X \neq 0$ if and only if p is an irregular prime. Moreover, in this case, O. Venjakob [Ve3] has shown that if X is pseudo-null, then the p-primary subgroup of the ideal class group of K_∞ is zero.

Example 2. — Let F be a finite extension of $\mathbb{Q}$, and E an elliptic curve defined over F, with $\mathrm{End}_{\overline{\mathbb{Q}}}(E) = \mathbb{Z}$. Let $p \geqslant 5$, and let E_{p^n} $(1 \leqslant n \leqslant \infty)$ denote the group of p^n-division points on E. We define

$$F_\infty = F(E_{p^\infty}), \quad G = G(F_\infty/F).$$

The action of G on E_{p^∞} defines an injection of G into $\mathrm{Aut}(E_{p^\infty}) \xrightarrow{\sim} GL_2(\mathbb{Z}_p)$, and, by a theorem of Serre [Se3], the image of G is open in $GL_2(\mathbb{Z}_p)$. By the Weil pairing, $F^{\mathrm{cyc}} = F(\mu_{p^\infty})$ is contained in F_∞, and we put

$$H = G(F_\infty/F^{\mathrm{cyc}}), \quad \Gamma = G(F^{\mathrm{cyc}}/F).$$

We shall assume from now on that G is pro-p (this can always be achieved, if necessary, by replacing F by a finite extension, e.g. by $F(E_p)$). Hence Γ is pro-p, and so is isomorphic to $\mathbb{Z}_p$.

For each intermediate field L with $F \subseteq L \subseteq F_\infty$, we recall that the *Selmer group* of E over L is defined by

$$S(E/L) = \mathrm{Ker}(H^1(G(\overline{\mathbb{Q}}/L), E_{p^\infty}) \longrightarrow \prod_v H^1(G(\overline{L_v}/L_v), E(\overline{L_v})),$$

where v runs over all finite places of L, and L_v denotes the union of the completions at v of all finite extensions of F contained in L. As usual, we have the exact sequence

$$0 \longrightarrow E(L) \otimes \mathbb{Q}_p/\mathbb{Z}_p \longrightarrow S(E/L) \longrightarrow Ш(E/L)(p) \longrightarrow 0,$$

where $Ш(E/L)(p)$ denotes the p-primary subgroup of the Tate-Shafarevich group of E over L. We write

$$X(E/L) = \mathrm{Hom}(S(E/L), \mathbb{Q}_p/\mathbb{Z}_p)$$

for the compact Pontrjagin dual of the discrete p-primary module $S(E/L)$. If L is Galois over F, then the Galois group $G(L/F)$ of L over F has a natural action on both $S(E/L)$ and $X(E/L)$, and it is easily seen that $X(E/L)$ is always a finitely generated $\Lambda(G(E/L))$-module. We shall be primarily interested in the $\Lambda(G)$-module $X(E/F_\infty)$. If E has good ordinary reduction at all places v of F dividing p, old conjectures due

to B. Mazur [Ma] and M. Harris [Ha] affirm, respectively, that $X(E/F^{\mathrm{cyc}})$ is $\Lambda(\Gamma)$-torsion and $X(E/F_\infty)$ is $\Lambda(G)$-torsion. It is easy to see that the validity of Mazur's conjecture for all finite extensions L of F contained in F_∞ implies the validity of Harris' conjecture for F_∞, but this is of little use in practice since the number of cases in which we can prove Mazur's conjecture remains very limited (the best result to date is K. Kato's [Kat] theorem that $X(E/F^{\mathrm{cyc}})$ is $\Lambda(\Gamma)$-torsion when E is an elliptic curve defined over $\mathbb{Q}$ with good ordinary reduction at p, and F is an abelian extension of $\mathbb{Q}$). In [CH], an alternative approach is given which does enable one to give the first proven examples of Harris' conjecture, and to deduce new examples of Mazur's conjecture.

THEOREM 3.1. — *Assume that* (i) $p \geqslant 5$, (ii) G *is pro-p,* (iii) E *has good ordinary reduction at all places* v *of* F *dividing* p, *and* (iv) $X(E/F^{\mathrm{cyc}})$ *is a finitely generated* $\mathbb{Z}_p$*-module. Then* $X(E/F_\infty)$ *is a finitely generated* $\Lambda(H)$*-module, where* $H = G(F_\infty/F^{\mathrm{cyc}})$. *In particular,* $X(E/F_\infty)$ *is a torsion* $\Lambda(G)$*-module.*

Remark 1. — Every $\Lambda(G)$-module, which is finitely generated over $\Lambda(H)$, is automatically $\Lambda(G)$-torsion. This is because $\Lambda(G)$ is not finitely generated over $\Lambda(H)$, since G/H is infinite.

Remark 2. — In the general framework and notation of Theorem 1.1, we say that the $\Lambda(G)$-module M has μ-invariant zero if none of the left ideals $L_1, \ldots, L_m$ appearing in Theorem 1.1 is of the form $\Lambda(G)p^k$ for some integer $k \geqslant 1$. If Γ is a group isomorphic to $\mathbb{Z}_p$, it follows from Theorem 1.1 and the Weierstrass preparation theorem that a $\Lambda(\Gamma)$-module Y is a finitely generated $\mathbb{Z}_p$-module if and only if Y is $\Lambda(\Gamma)$-torsion and has μ-invariant zero. Note also that if M is a finitely generated $\Lambda(H)$-module, then it must have μ-invariant zero, because $\Lambda(G)/\Lambda(G)p^k$ is not a finitely generated $\Lambda(H)$-module when $k \geqslant 1$.

A second important result about $X(E/F_\infty)$ is due to Y. Ochi and O. Venjakob [OV].

THEOREM 3.2. — *Assume that hypotheses* (i), (ii), (ii), *and* (iv) *of Theorem 3.1 are valid. Then* $X(E/F_\infty)$ *contains no non-zero pseudo-null submodule, the* $\Lambda(H)$*-torsion submodule of* $X(E/F_\infty)$ *is zero, and* $X(E/F_\infty)$ *has strictly positive* $\Lambda(H)$*-rank.*

To prove the last assertion of Theorem 3.2, we must use the fact that always $X(E/F_\infty) \neq 0$ (this was first remarked by R. Greenberg, and a proof is given in the Appendix of [CH]). Assume now that E over F satisfies hypotheses (i), (ii), (ii), and (iv) of Theorem 3.1. We conclude from the above results and Theorem 1.1 that there exist non-zero left ideals $L_1, \ldots, L_m$ of $\Lambda(G)$ such that we have an exact sequence of $\Lambda(G)$-modules

$$0 \longrightarrow \bigoplus_{i=1}^{m} \Lambda(G)/L_i \longrightarrow X(E/F_\infty) \longrightarrow D \to 0,$$

where D is pseudo-null. We stress that there is great arithmetic interest in studying the left ideals $L_1, \ldots, L_m$, even in particular numerical examples. Of course, one imagines that these left ideals must be related to the values at $s = 1$ of the twists of the complex L-function of E over F by Artin characters of G (i.e. those which factor through finite quotients of G). On a much more elementary level, one can ask whether or not $\mathrm{ann}_{\Lambda(G)}(X(E/F_\infty)) \neq 0$. This still has not been settled in a single numerical example when $F_\infty = F(E_{p^\infty})$. Surprisingly, when K_∞ and G are as defined in Example 1 above, Hachimori and Venjakob [HV] have recently given many examples of elliptic curves E over K_∞ such that the dual of the Selmer group of E over K_∞ is a finitely generated torsion $\Lambda(G)$-module, which is not pseudo-null, but which is completely faithful. For example, they prove that this is the case for the elliptic curve $E = X_1(11)$ given below, when $p = 5$, and K_∞ is the field obtained by adjoining to $\mathbb{Q}$ all 5-power roots of unity and all 5-power roots of 11.

We end by discussing two specific numerical examples of elliptic curves over their field of p-power division points.

Numerical example 1. — I am grateful to T. Fisher for first pointing out this example to me. Let E be the elliptic curve over $\mathbb{Q}$

$$E : y^2 + xy = x^3 - x - 1.$$

This is the curve $B1$ of conductor 294 in [Cr1]. Take $F = \mathbb{Q}(\mu_7)$ and $p = 7$. Although E has bad reduction at 7 over $\mathbb{Q}$, it is easily seen that E has good ordinary reduction at the unique prime of F above 7. Moreover, μ_7 is a Galois submodule of E_7, whence we see easily that $F_\infty = \mathbb{Q}(E_{7^\infty})$ is a pro-7 extension of F. Fisher [Fi1] has shown that $S(E/F) = 0$. One can then use arguments from Iwasawa theory ([CS], p. 83) to deduce that we also have $S(E/F^{\mathrm{cyc}}) = 0$. Hence hypotheses (i), (ii), (iii), and (iv) of Theorem 3.1 are valid for E over F and $p = 7$. We conclude that $X(E/F_\infty)$ is a torsion $\Lambda(G)$-module, with μ-invariant equal to zero, and with no non-zero pseudo-null submodule. Moreover, $X(E/F_\infty)$ is finitely generated over $\Lambda(H)$, its $\Lambda(H)$-torsion submodule is zero, and it has positive $\Lambda(H)$-rank. Here $G = G(\mathbb{Q}(E_{7^\infty})/\mathbb{Q}(\mu_7))$ and $H = G(\mathbb{Q}(E_{7^\infty}/\mathbb{Q}(\mu_{7^\infty})))$.

Numerical example 2. — Let E be the elliptic curve $X_1(11)$ over $\mathbb{Q}$, namely

$$E : y^2 + y = x^3 - x^2.$$

Then E has good ordinary reduction at 5. Take $F = \mathbb{Q}(\mu_5)$ and $p = 5$. As $(0, 0)$ is a rational point of order 5 on E, $F_\infty = \mathbb{Q}(E_{5^\infty})$ is a pro-5 extension of F. Indeed, it is well known and easy to see that the image of G in $\mathrm{Aut}(E_{5^\infty})$ can be identified with the subgroup of all matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $GL_2(\mathbb{Z}_5)$ with $a \equiv d \equiv 1 \mod 5$, and $c \equiv 0 \mod 5^2$, and this group in turn is isomorphic to the group of all matrices in $GL_2(\mathbb{Z}_5)$, which are congruent to the identity modulo 5. Finally, it is well known that, in this

case, $S(E/F^{\mathrm{cyc}}) = 0$ (see [CS], Chap. 5). Hence hypotheses (i), (ii), (iii), and (iv) of Theorem 3.1 hold for E over F, and we conclude that $X(E/F_\infty)$ is a torsion $\Lambda(G)$-module, with μ-invariant equal to zero, and with no non-zero pseudo-null submodule. In fact, the following stronger result is true.

PROPOSITION 3.3. — *Take $E = X_1(11)$, $p = 5$, $F = \mathbb{Q}(\mu_5)$ and $H = G(F_\infty/\mathbb{Q}(\mu_{5^\infty}))$. Then $X(E/F_\infty)$ is a finitely generated $\Lambda(H)$-module of rank 4, its $\Lambda(H)$-torsion submodule is zero, but it is not a free $\Lambda(H)$-module.*

For each finite Galois extension L of F which is contained in F_∞, we write

$$G_L = G(F_\infty/L), \quad H_L = G(F_\infty/L^{\mathrm{cyc}}),$$

so that G_L ranges over the open normal subgroups of G, and H_L ranges over the open normal subgroups of H. The proof of Proposition 3.3 hinges on the remarkable fact that one can use ideas of Y. Hachimori and K. Matsuno [HM] to determine the exact $\mathbb{Z}_5$-rank of the H_L-coinvariants of $X(E/F_\infty)$ for every finite Galois extension L of F contained in F_∞. In particular, the following result is proven in § 7 of [CH].

PROPOSITION 3.4. — *For each finite Galois extension L of F contained in F_∞, $X(E/L^{\mathrm{cyc}})$ is a free $\mathbb{Z}_5$-module of rank 4. $[L^{\mathrm{cyc}} : F^{\mathrm{cyc}}] - \tau_L$, where τ_L denotes the number of prime of L^{cyc} above 11. In particular, Mazur's conjecture is true for E over L^{cyc}, and $E(L^{\mathrm{cyc}})$ is a finitely generated abelian group of rank at most $4 \cdot [L^{\mathrm{cyc}} : F^{\mathrm{cyc}}] - \tau_L$.*

One deduces easily from this proposition that $X(E/F_\infty)_{H_L}$ has $\mathbb{Z}_5$-rank equal to $4 \cdot [L^{\mathrm{cyc}} : F^{\mathrm{cyc}}]$ for all L. At the same time, one shows that the $\mathbb{Z}_5$-torsion subgroup of $X(E/F_\infty)_{H_L}$ is never zero, which shows that $X(E/F_\infty)$ is not a free $\Lambda(H)$-module.

By contrast, it is a deep arithmetic problem to determine the exact $\mathbb{Z}_5$-rank of the G_L-coinvariants of $X(E/F_\infty)$ for all L. In particular, the following lemma is not difficult to prove using the methods of [CH].

LEMMA 3.5. — *Let L be any finite Galois extension of F contained in F_∞. Then $X(E/F_\infty)_{G_L}$ is finite if and only if both $E(L)$ and the 5-primary subgroup of the Tate-Shafarevich group of E over L are finite.*

We have already used the classical fact that $X(E/F_\infty)_G$ is finite. By some remarkable explicit descent calculations, T. Fisher [Fi2] has recently shown that $X(E/F_\infty)_{G_L}$ is finite for L ranging over the following three cyclic extensions of F of degree 5 which are contained in F_∞:

$$L_1 = \mathbb{Q}(X_1(11)_5), \quad L_2 = \mathbb{Q}(X_2(11)_5), \quad L_3 = FQ(\mu_{11})^+;$$

here $X_2(11)$ denotes the unique elliptic curve of conductor 11 over $\mathbb{Q}$ which has no non-zero rational point, and $\mathbb{Q}(\mu_{11})^+$ denotes the maximal real subfield of $\mathbb{Q}(\mu_{11})$. The field L_1 is the splitting field of the polynomial

$$x^5 + 2x^4 + 6x^3 - 2x^2 + 4x - 1.$$

Then T. Fisher's descent calculations [Fi2] show that $E(L_1)$ is finite, and

$$\text{Ш}(E/L_1)(5) = (\mathbb{Z}/5\mathbb{Z})^2.$$

However, it does not seem possible to extend his explicit calculations even to the field $K_2 = L_1(\mu_{5^2})$, which has degree 100 over $\mathbb{Q}$. Nevertheless, if

$$K_n = L_1(\mu_{5^{n+1}}) \qquad (n = 0, 1, \dots),$$

T. Fisher, R. Greenberg and myself have shown that simple theoretical arguments from the Iwasawa theory of elliptic curves do enable one to prove that, for all integers $n = 0, 1, \dots, E(K_n)$ is finite and $\text{Ш}(E/K_n)(5)$ is finite of order 5^{16n+2}.

When this exposé was given in November 2001, I naively imagined that perhaps $X_1(11)$ had no points of infinite order in the whole tower F_∞ of 5-division points. I am very grateful to K. Matsuno (unpublished) for producing overwhelming numerical evidence that this is not true. Take the field L_3 above, and let $H_2 = L_3(\mu_{5^2})$, so that H_2 is an abelian extension of $\mathbb{Q}$ of degree 100. K. Matsuno calculated the complex L-function of $X_1(11)$ over H_2, and proved that it has a zero at $s = 1$ of order 4. Thus, unless the conjecture of Birch and Swinnerton-Dyer is false (which I do not for one moment believe), $X_1(11)$ should have rank 4 over H_2. So far, no point of infinite order has been found, nor has Fisher been able to extend his explicit descent calculations to $X_1(11)$ over H_2. Thus, in conclusion, two important questions remain unanswered about the arithmetic of $X_1(11)$ over its field F_∞ of 5-power division points. Is the dual of the Selmer group of $X_1(11)$ over F_∞ a completely faithful $\Lambda(G)$-module? What is the $\mathbb{Z}$-rank of the group $X_1(11)(F_\infty)$ of F_∞-rational points modulo its torsion subgroup?

REFERENCES

[As] K. Asano – Zur Arithmetik in Schiefringen, *Osaka J. Math.* **2** (1949), p. 98–134.

[Bj] J.-E. Björk – Filtered Noetherian Rings, in *Noetherian rings and their applications*, Math. Survey Monographs, vol. 24, AMS, 1987, p. 59–97.

[B-CA] N. Bourbaki – *Algèbre Commutative*, Paris, Hermann, 1972.

[Ch1] M. Chamarie – Anneaux de Krull non commutatifs, *J. Algebra* **72** (1981), p. 210–222.

[Ch2] ———, Modules sur les anneaux de Krull non commutatifs, in *Sém. d'Algèbre P. Dubreil et M.-P. Malliavin 1982*, Springer Lecture Notes, vol. 1029, Springer, 1983, p. 283–310.

[CH] J. Coates & S. Howson – Euler characteristics and elliptic curves II, *J. Math. Soc. Japan* **53** (2001), p. 175–235.

[CSS] J. Coates, P. Schneider & R. Sujatha – Modules over Iwasawa algebras, *J. Inst. Math. Jussieu* **2** (2003), p. 73–108.

[CS] J. Coates & R. Sujatha – *Galois cohomology of elliptic curves*, Lecture Notes, vol. 88, TIFR-AMS, 2000.

[Cr1] J. Cremona – *Algorithms for modular elliptic curves*, 2nd ed., Cambridge University Press, 1997.

[Fi1] T. Fisher – On 5 and 7 descents for elliptic curves, Ph.D. Thesis, Cambridge University, 2000.

[Fi2] ———, Descent calculations for the elliptic curves of conductor 11, *Proc. London Math. Soc.* **86** (2003), p. 583–606.

[HM] Y. Hachimori & K. Matsuno – An analogue of Kida's formula for the Selmer groups of elliptic curves, *J. Alg. Geometry* **8** (1999), p. 581–601.

[HV] Y. Hachimori & O. Venjakob – Completely faithful Selmer groups over Kummer extensions, to appear.

[Ha] M. Harris – p-adic representations arising from descent on abelian varieties, *Compositio Math.* **30** (1979), p. 177–245.

[Ho] S. Howson – Structure of central torsion Iwasawa modules, *Bull. Soc. math. France* **130** (2002), no. 4, p. 507–535.

[Iw] K. Iwasawa – On Γ-extensions of number fields, *Bull. AMS* **65** (1959), p. 183–226.

[Ja] H. Jacobson – *Theory of rings*, Math. Surveys, vol. 2, AMS, Providence, 1943.

[Ka] M. Kashiwara – Algebraic study of systems of partial differential equations, Master thesis, Tokyo University, 1971, English translation in Mem. Soc. math. France (N.S.), vol. 63 (1995).

[Kat] K. Kato – p-adic Hodge theory and values of zeta functions of modular forms, to appear.

[La] M. Lazard – Groupes analytiques p-adiques, *Publ. Math. IHÉS* **26** (1965), p. 380–603.

[Li] Li Huishi – Lifting Ore sets of Noetherian filtered rings and applications, *J. Algebra* **179** (1996), p. 686–703.

[LO] Li Huishi & F. van Oystaeyen – *Zariskian Filtrations*, Kluwer, Dordrecht, 1996.

[Ma] B. Mazur – Rational points of abelian varieties in towers of number fields, *Invent. Math.* **18** (1992), p. 183–266.

[Oc] O. Ochi – Iwasawa modules via homotopy theory, Ph.D. Thesis, Cambridge University, 1999.

[OV] Y. Ochi & O. Venjakob – On the structure of Selmer groups over p-adic Lie extensions, *J. Alg. Geom* **11** (2002), p. 547–576.

[Ro] J.C. Robson – Cyclic and faithful objects in quotient categories with applications to Noetherian simple or Asano rings, in *Noncommutative Ring Theory*, Springer Lecture Notes, vol. 545, Springer, 1976, p. 151–172.

[Se2] J.-P. Serre – Letter to K. Iwasawa, dated August 27th 1958.

[Se3] ———, Propriétés galoisiennes des points d'ordre fini des courbes elliptiques, *Invent. Math.* **15** (1972), p. 259–331.

[Se1] ———, Classes des corps cyclotomiques (d'après K. Iwasawa), in *Sém. Bourbaki*, Collection hors série, vol. 5, Société mathématique de France, 1995, exp. n° 174, décembre 1958, p. 83–93.

[Ve1] O. Venjakob – Iwasawa Theory of p-adic Lie Extensions, Thesis, Heidelberg University, 2000.

[Ve2] ———, On the structure of the Iwasawa algebra of a p-adic Lie group, *J. European Math. Soc* **4** (2002), p. 271–311.

[Ve3] ———, A non-commutative Weierstrass preparation theorem and its applications to Iwasawa theory, *J. reine angew. Math.* **559** (2003), p. 153–191.

[WK] E. Wexler-Kreindler – Microlocalisation, platitude et théorie de torsion, *Comm. Algebra* **16** (1988), p. 1813–1852.

John COATES

Cambridge University
DPMMS
Wilberforce Road
GB-Cambridge, CB3 0WB
U.K.
E-mail : J.H.Coates@dpmms.cam.ac.uk

Séminaire BOURBAKI Novembre 2001
54e année, 2001-2002, no 897, p. 53 à 101

LES CONJECTURES DE MONODROMIE p-ADIQUES

par **Pierre COLMEZ**

0. INTRODUCTION

0.1. Survol des résultats

Le théorème local de monodromie ℓ-adique de Grothendieck peut s'énoncer de la manière suivante.

THÉORÈME 0.1. — *Soit* K *un corps complet pour une valuation discrète de corps résiduel k fini et de caractéristique p. Soient $\ell \neq p$ un nombre premier et* V *un $\mathbf{Q}_\ell$-espace vectoriel de dimension finie sur lequel le groupe de Galois absolu $\mathscr{G}_K$ de* K *agit continûment ; alors l'action de l'inertie est quasi-unipotente : il existe un sous-groupe ouvert du sous-groupe d'inertie I_K de $\mathscr{G}_K$ dont les éléments agissent de manière unipotente.*

Ce théorème [45], bien que « quasi-trivial » (c'est une conséquence facile de l'existence d'une « structure de Frobenius » : le pro-ℓ-quotient de I_K est isomorphe à $\mathbf{Z}_\ell$ et, si $q = \operatorname{card} k$, un Frobenius $\sigma \in \mathscr{G}_K$ opérant par conjugaison intérieure, agit sur ce pro-ℓ-quotient par multiplication par q), a des implications globales intéressantes concernant la cohomologie étale ℓ-adique des variétés algébriques (voir [50] pour une discussion et des variations sur ce thème).

L'énoncé correspondant étant complètement faux pour $\ell = p$ (le pro-p-quotient de I_K étant nettement plus compliqué que le pro-ℓ-quotient), il n'est pas très facile d'imaginer ce que peut en être un analogue p-adique et, de fait, il a fallu un certain temps, ne serait-ce que pour définir les objets à considérer. Il y a deux cas suivant que le corps K est de caractéristique 0 ou de caractéristique p. Dans le cas où K est de caractéristique 0, la difficulté pour arriver à formuler cette conjecture vient de ce que $\mathscr{G}_K$ possède beaucoup trop de $\mathbf{Q}_p$-représentations et qu'il faut commencer par comprendre lesquelles on veut garder (représentations « de de Rham »). Le théorème de monodromie locale prend alors la forme suivante :

THÉORÈME 0.2. — *Toute $\mathbf{Q}_p$-représentation de de Rham de $\mathscr{G}_K$ est potentiellement semi-stable.*

Si K est de caractéristique p, le groupe $\mathscr{G}_K$ a « encore plus » de $\mathbf{Q}_p$-représentations et il faut remplacer ces dernières par des modules différentiels sur une couronne « infiniment fine » de rayon 1. Dans ce cadre, l'analogue du théorème de Grothendieck est l'énoncé suivant :

THÉORÈME 0.3. — *Tout module différentiel sur l'anneau de Robba, muni d'une structure de Frobenius, est quasi-unipotent.*

L'intérêt de l'énoncé du théorème 0.3 « conjecture de monodromie p-adique de Crew[(1)] » a été mis en évidence par Crew [31]. Cet énoncé permet, en particulier, d'obtenir des résultats de finitude [31] pour la cohomologie rigide. Ces résultats de finitude ont entre-temps été démontrés par Mebkhout [62] et Berthelot [7] (dans le cas des coefficients constants), le premier en utilisant ses résultats obtenus en collaboration avec Christol (cf. [17], [18], [19], [20] et [21]) et le second en utilisant les altérations de de Jong [51]. Pour d'autres applications, voir [53, 54, 75].

La conjecture de Crew vient d'être démontrée de manière quasi-simultanée par André [1], par Mebkhout [63] et par Kedlaya [57], par des méthodes très différentes. D'autre part, Berger [6] avait, peu auparavant, montré comment déduire le théorème 0.2 « conjecture de monodromie p-adique de Fontaine [40][(2)] » de la conjecture de Crew.

0.2. Modules différentiels sur l'anneau de Robba

0.2.1. La conjecture de Crew. — Soit L un corps complet pour une valuation p-adique de corps résiduel k. Soit $\mathscr{R}_L$ (resp. $\mathscr{E}_L^\dagger$) l'anneau des fonctions à coefficients dans L qui sont analytiques (resp. analytiques bornées) sur une couronne $0 < v_p(T) \leqslant r$, où $r > 0$ dépend de la fonction que l'on considère. L'anneau $\mathscr{R}_L$ est souvent appelé *anneau de Robba*. L'anneau $\mathscr{E}_L^\dagger$ est un anneau valué non complet d'anneau résiduel $k((T))$.

Un Frobenius φ sur ces anneaux est un morphisme d'anneaux de $\mathscr{E}_L^\dagger$ respectant la valuation et induisant le morphisme $x \mapsto x^p$ sur $k((T))$ (ou, plus généralement $x \mapsto x^q$, où q est une puissance de p) ; un tel morphisme s'étend par continuité à $\mathscr{R}_L$. Des exemples agréables de Frobenius sont $T \mapsto T^p$ et $T \mapsto (1+T)^p - 1$. Un φ-module $\mathscr{D}$ sur un de ces anneaux est un module libre de rang fini muni d'un endomorphisme φ-semi-linéaire $\varphi_{\mathscr{D}}$ dont la matrice dans une base est *inversible*.

(1) Ou de Crew-Tsuzuki. Crew [31] est très prudent, et se cantonne aux modules différentiels provenant de la géométrie, et Tsuzuki [75] est nettement plus optimiste, mais ni Crew, ni Tsuzuki n'ont, formellement, présenté cet énoncé comme une conjecture.
(2) Cet énoncé n'est, lui non plus, pas formellement présenté comme une conjecture dans [40].

Choisissons une dérivation continue ∂ de l'anneau $\mathscr{E}_{\mathrm{K}}^{\dagger}$ (comme $d/d\mathrm{T}$, $\mathrm{T}d/d\mathrm{T}$ ou $(1+\mathrm{T})d/d\mathrm{T}$). Un ∂-module sur un de ces anneaux est un module libre de rang fini muni d'une connexion $\partial_{\mathscr{D}}$ au-dessus de ∂, c'est-à-dire vérifiant

$$\partial_{\mathscr{D}}(f\cdot x)=\partial f\cdot x+f\cdot\partial_{\mathscr{D}}x \quad \text{si } f\in\mathscr{E}_{\mathrm{K}}^{\dagger} \text{ (resp. } \mathscr{R}_{\mathrm{K}}) \text{ et } x\in\mathscr{D}.$$

Un (φ,∂)-module $\mathscr{D}$ est un module qui est à la fois un φ-module et un ∂-module de telle sorte que $\varphi_{\mathscr{D}}$ et $\partial_{\mathscr{D}}$ vérifient la relation de commutation

$$\partial_{\mathscr{D}}\circ\varphi_{\mathscr{D}}=\frac{\partial(\varphi(\mathrm{T}))}{\varphi(\partial\mathrm{T})}\cdot\varphi_{\mathscr{D}}\circ\partial_{\mathscr{D}}.$$

Dans le cas particulier $\partial=(1+\mathrm{T})d/d\mathrm{T}$ et $\varphi(\mathrm{T})=(1+\mathrm{T})^p-1$, on a $\partial(\varphi(\mathrm{T}))/\varphi(\partial\mathrm{T})=p$. Si on fixe une base de $\mathscr{D}$ et si A (resp. B) est la matrice de $\varphi_{\mathscr{D}}$ (resp. $\partial_{\mathscr{D}}$) dans cette base, cette relation de commutation se traduit par

$$\mathrm{BA}+\partial\mathrm{A}=\frac{\partial(\varphi(\mathrm{T}))}{\varphi(\partial\mathrm{T})}\mathrm{A}\varphi(\mathrm{B}).$$

Si E est une extension finie séparable de $k((\mathrm{T}))$, il lui correspond des extensions $\mathscr{E}_{\mathrm{L}}^{\dagger}(\mathrm{E})$ et $\mathscr{R}_{\mathrm{L}}(\mathrm{E})$ de $\mathscr{E}_{\mathrm{L}}^{\dagger}$ et $\mathscr{R}_{\mathrm{L}}$ respectivement auxquelles les actions de φ et ∂ s'étendent canoniquement. La conjecture de Crew peut alors s'énoncer de façon précise sous la forme :

Conjecture 0.4. — *Si $\mathscr{D}$ est un (φ,∂)-module de rang d sur $\mathscr{R}_{\mathrm{L}}$, alors il existe une extension finie séparable* E *de $k((\mathrm{T}))$ telle que l'équation différentielle $\partial_{\mathscr{D}}\mathrm{X}=0$ admette d solutions linéairement indépendantes dans $\mathscr{R}_{\mathrm{L}}(\mathrm{E})[\log\mathrm{T}]\otimes\mathscr{D}$.*

Cette conjecture a une « reformulation filtrée », faisant disparaître $\log\mathrm{T}$, sous la forme : « si $\mathscr{D}$ est un (φ,∂)-module sur $\mathscr{R}_{\mathrm{L}}$, alors il existe une extension finie séparable E de $k((\mathrm{T}))$ et une filtration croissante de $\mathscr{D}$ par des sous-∂-modules $\mathscr{D}_i$ telles que, pour tout i, l'équation différentielle $\partial_{\mathscr{D}_i/\mathscr{D}_{i-1}}\mathrm{X}=0$ admette $d_i=\mathrm{rang}(\mathscr{D}_i/\mathscr{D}_{i-1})$ solutions linéairement indépendantes dans $\mathscr{R}_{\mathrm{L}}(\mathrm{E})\otimes(\mathscr{D}_i/\mathscr{D}_{i-1})$ ».

Remarque 0.5. — Il y a une subtilité cachée dans cette conjecture : la géométrie algébrique fournit naturellement des (φ,∂)-modules sur $\mathscr{E}_{\mathrm{L}}^{\dagger}$, mais, en rang $\geqslant 2$, il faut vraiment étendre les coefficients à $\mathscr{R}_{\mathrm{L}}(\mathrm{E})$ [et pas seulement à $\mathscr{E}_{\mathrm{L}}^{\dagger}(\mathrm{E})$]. Christol et Mebkhout se sont retrouvés confrontés au même problème (rem. 1.6). (On trouvera une explication de ce phénomène au n° 2.3 : les ensembles des $\mathscr{E}$-pentes et $\mathscr{R}$-pentes d'un (φ,∂)-module sur $\mathscr{E}_{\mathrm{L}}^{\dagger}$ ne sont pas toujours égaux.)

0.2.2. Les démonstrations. — Crew [30] avait établi la conjecture en rang 1, et avant cet été, les seuls résultats un peu généraux concernant cette conjecture étaient, d'une part, un théorème de Tsuzuki [74] montrant qu'un (φ,∂)-module sur $\mathscr{E}^\dagger$ isocline (i.e. n'ayant qu'une φ-pente [(3)]) devient trivial après une extension finie séparable de $k((\mathrm{T}))$ et, d'autre part, un théorème de Christol et Mebkhout [18] montrant qu'un (φ,∂)-module de ∂-pente 0 devient unipotent sur une extension modérée de $k((\mathrm{T}))$ (i.e. une extension de la forme $k((\mathrm{T}^{1/d}))$ avec $(d,p)=1$).

Malgré ces résultats encourageants, les experts étaient plutôt sceptiques en ce qui concerne la conjecture de Crew. L'exemple [20, Ex. 3.0-11] du ∂-module associé, en 2-adique, à l'opérateur différentiel $9\mathrm{T}^3d^2/d\mathrm{T}^2+9\mathrm{T}^2d/d\mathrm{T}+4/3-\mathrm{T}$ constituant un contre-exemple potentiel sérieux. La situation a radicalement changé avec l'article [2] dans lequel André montre que ce contre-exemple n'en est pas un. Sa démonstration dans le cas général [1] est une généralisation de celle qu'il a utilisée dans [2] ; elle est purement existentielle et ne fournit aucun renseignement sur l'extension à faire[(4)] pour rendre le module unipotent. Elle s'appuie de manière essentielle sur les résultats de Christol-Mebkhout et mêle adroitement la théorie de Galois différentielle et les contraintes combinatoires résultant de la théorie de Christol-Mebkhout.

La démonstration de Mebkhout [63] utilise aussi à fond les résultats de Christol-Mebkhout mais, contrairement à celle d'André, est constructive : elle fournit un algorithme (pour courageux) explicitant pas à pas les extensions séparables de $k((\mathrm{T}))$ qu'il faut faire pour aboutir à la ∂-pente 0.

Ces deux démonstrations n'utilisent la structure de Frobenius que de manière anecdotique et permettent de démontrer un résultat plus fort que la conjecture de Crew. La démonstration de Kedlaya [57] est totalement orthogonale. Il fait une étude poussée des φ-modules sur l'anneau de Robba (et ses généralisations) obtenant en particulier un analogue du théorème de Dieudonné-Manin. Ceci lui permet de munir un tel module d'une filtration stable par Frobenius telle que chaque morceau du gradué soit isocline. D'autre part, si on est parti d'un (φ,∂)-module, la filtration est stable par ∂, ce qui permet d'utiliser le théorème de Tsuzuki pour conclure.

0.2.3. Compléments. — On peut réinterpréter la conjecture de Crew « à la Fontaine ». Supposons k algébriquement clos (cela évitera d'avoir à faire des extensions du corps des coefficients dans ce qui suit). Si E est une extension finie galoisienne de $k((\mathrm{T}))$, l'anneau $\mathscr{R}_{\mathrm{L}}(\mathrm{E})$ est muni d'actions de $\mathscr{G}_{k((\mathrm{T}))}=\mathrm{Gal}(k((\mathrm{T}))^{\mathrm{sep}}/k((\mathrm{T})))$ et de la dérivation $\mathscr{R}_{\mathrm{L}}(\mathrm{E})$-linéaire N normalisée par $\mathrm{N}(\log \mathrm{T})=1$; ces deux actions commutent entre elles et commutent aussi à ∂.

[(3)] Les φ-modules et les ∂-modules ont chacun une notion de pente. Ces deux notions n'ont rien à voir ; en particulier elles se comportent de manières très différentes par produit tensoriel.
[(4)] Dans [2], André, après avoir prouvé l'existence d'une extension trivialisant le module, détermine explicitement cette extension par une méthode qu'il semble difficile de généraliser.

Soit $\overline{\mathscr{R}}_{\log} = \bigcup_{\mathrm{E}\subset k((\mathrm{T}))^{\mathrm{sep}}} \mathscr{R}_{\mathrm{L}}(\mathrm{E})[\log \mathrm{T}]$; c'est un anneau muni des dérivations N et ∂ et d'une action discrète de $\mathscr{G}_{k((\mathrm{T}))}$, ces actions commutant deux à deux. Si $\mathscr{D}$ est un (φ,∂)-module de rang d, l'espace $\mathrm{V}(\mathscr{D})$ des solutions de l'équation différentielle $\partial_{\mathscr{D}}\mathrm{X} = 0$ dans $\overline{\mathscr{R}}_{\log}\otimes\mathscr{D}$ est un L-espace vectoriel de dimension d (d'après la conjecture de Crew) muni d'actions de N et $\mathscr{G}_{k((\mathrm{T}))}$ commutant entre elles, l'action de $\mathscr{G}_{k((\mathrm{T}))}$ se faisant à travers un quotient fini. Un tel objet sera appelé *un* $(\mathscr{G}_{k((\mathrm{T}))},\mathrm{N})$*-module sur* L.

Réciproquement, si V est un $(\mathscr{G}_{k((\mathrm{T}))},\mathrm{N})$-module sur L, le $\mathscr{R}_{\mathrm{L}}$-module $\mathscr{D}(\mathrm{V})$ des points fixes de $\overline{\mathscr{R}}_{\log}\otimes_{\mathrm{L}}\mathrm{V}$ sous les actions de N et $\mathscr{G}_{k((\mathrm{T}))}$ est un (φ,∂)-module de rang d sur $\mathscr{R}_{\mathrm{L}}$.

Il n'est pas difficile de montrer que les foncteurs $\mathrm{V}\mapsto\mathscr{D}(\mathrm{V})$ et $\mathscr{D}\mapsto\mathrm{V}(\mathscr{D})$ que l'on vient d'introduire sont inverses l'un de l'autre et donc que la catégorie des (φ,∂)-modules sur $\mathscr{R}_{\mathrm{L}}$ est équivalente à celle des $(\mathscr{G}_{k((\mathrm{T}))},\mathrm{N})$-modules sur L, ce qui en fournit une description particulièrement simple. (Pour une description conjecturale de cette catégorie en termes purement différentiels, voir [20, p. 671-672].)

D'autre part, à une représentation V de $\mathscr{G}_{k((\mathrm{T}))}$ d'image finie (resp. à un (φ,∂)-module $\mathscr{D}$, plus généralement à un ∂-module *soluble*), on sait associer un entier, à savoir son conducteur de Swan $\mathrm{Sw}(\mathrm{V})$ (resp. son irrégularité $\mathrm{Irr}(\mathscr{D})$) et Tsuzuki a montré [76] (voir [32, 61] pour d'autres démonstrations) que l'on a

$$\mathrm{Irr}(\mathscr{D}(\mathrm{V})) = \mathrm{Sw}(\mathrm{V}).$$

0.3. Anneaux de Fontaine et représentations p-adiques

0.3.1. Notations. — On se fixe une clôture algébrique $\overline{\mathbf{Q}}_p$ de $\mathbf{Q}_p$ et un système $(\varepsilon^{(n)})_{n\in\mathbf{N}}$ d'éléments de $\overline{\mathbf{Q}}_p$ vérifiant $\varepsilon^{(0)}=1$, $\varepsilon^{(1)}\neq 1$ et $(\varepsilon^{(n+1)})^p=\varepsilon^{(n)}$ si $n\in\mathbf{N}$. Ceci fait de $\varepsilon^{(n)}$ une racine primitive p^n-ième de l'unité et on note F_n le corps $\mathbf{Q}_p(\varepsilon^{(n)})$ et F_∞ l'extension cyclotomique de $\mathbf{Q}_p$ réunion des F_n. Soit $\mathbf{C}_p$ le complété de $\overline{\mathbf{Q}}_p$ pour la valuation v_p.

On note $\mathscr{G}_{\mathbf{Q}_p}$ le groupe de Galois $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ et $\chi : \mathscr{G}_{\mathbf{Q}_p}\to\mathbf{Z}_p^*$ le caractère cyclotomique. Soit aussi $\mathscr{H}_{\mathbf{Q}_p}$ le noyau de la restriction de χ à $\mathscr{G}_{\mathbf{Q}_p}$ de telle sorte que $\mathscr{H}_{\mathbf{Q}_p}=\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathrm{F}_\infty)$ et soit $\Gamma_{\mathbf{Q}_p}=\mathscr{G}_{\mathbf{Q}_p}/\mathscr{H}_{\mathbf{Q}_p}=\mathrm{Gal}(\mathrm{F}_\infty/\mathbf{Q}_p)$.

(On peut remplacer $\mathbf{Q}_p$ par un corps complet pour la valuation v_p, de corps résiduel parfait, et c'est ce qui est fait dans le texte principal, mais travailler avec $\mathbf{Q}_p$ a l'avantage de simplifier certaines formules.)

0.3.2. Le programme de Fontaine. — Soit G un groupe topologique (comme $\mathscr{G}_{\mathbf{Q}_p}$ ou $\mathscr{H}_{\mathbf{Q}_p}$). Pour mettre un peu d'ordre dans les $\mathbf{Q}_p$-représentations de G, la stratégie de Fontaine est de construire des anneaux topologiques munis d'une action de G et de structures additionnelles respectées par l'action de G. Chacun de ces anneaux permet de découper dans l'ensemble des $\mathbf{Q}_p$-représentations de G celles qui sont B-*admissibles* (une représentation V de dimension d est dite B-admissible si $\mathrm{B}\otimes_{\mathbf{Q}_p}\mathrm{V}$ est isomorphe à B^d en tant que G-module). Si V est une représentation B-admissible de $\mathscr{G}_{\mathrm{K}}$, le

B^{G}-module $\mathbf{D}_{\mathrm{B}}(\mathrm{V}) = (\mathrm{B}\otimes\mathrm{V})^{\mathrm{G}}$ est libre de rang $\dim_{\mathbf{Q}_p}\mathrm{V}$ et est muni de toutes les structures additionnelles de B respectées par l'action de G. Ceci permet d'associer aux représentations de G des invariants plus maniables (en général des objets provenant de l'algèbre linéaire) et, si l'anneau B est assez fin (i.e. si on peut retrouver $\mathbf{Q}_p$ à l'intérieur de B en utilisant les structures respectées par G), de classifier les représentations B-admissibles en termes de ces invariants. Cette approche a l'avantage de ramener l'étude de toutes les représentations B-admissibles à celle de l'anneau B ; tout l'art résidant dans la construction d'anneaux intéressants. Si V est B-admissible, la matrice permettant de passer d'une base de $\mathbf{D}_{\mathrm{B}}(\mathrm{G})$ sur B^{G} à une base de V sur $\mathbf{Q}_p$ appartient à $\mathrm{GL}_d(\mathrm{B})$; ses coefficients sont *les* B-*périodes* de la représentation V.

0.3.3. Quelques anneaux de Fontaine. — La théorie du corps des normes de Fontaine et Wintenberger [44] et [79] permet d'associer à toute extension finie K de $\mathbf{Q}_p$ un corps $\mathbf{E}_{\mathrm{K}}$ de caractéristique p, complet pour une valuation discrète. Elle permet de munir la clôture séparable $\mathbf{F}_p((\mathrm{T}))^{\mathrm{sep}}$ de $\mathbf{F}_p((\mathrm{T}))$ d'une action naturelle de $\mathscr{G}_{\mathbf{Q}_p}$ identifiant $\mathscr{H}_{\mathbf{Q}_p}$ à $\mathscr{G}_{\mathbf{F}_p((\mathrm{T}))}$; le groupe $\mathscr{G}_{\mathbf{Q}_p}$ agit alors sur $\mathbf{E}_{\mathbf{Q}_p} = \mathbf{F}_p((\mathrm{T}))$ à travers $\Gamma_{\mathbf{Q}_p}$ par $\sigma(\mathrm{T}) = (1+\mathrm{T})^{\chi(\sigma)} - 1$.

On peut relever cette action en caractéristique 0 à $\mathbf{Z}_p[[\mathrm{T}]]$ par la même formule et, si on fait agir le Frobenius φ par $\mathrm{T}\mapsto(1+\mathrm{T})^p-1$, les actions de φ et $\mathscr{G}_{\mathbf{Q}_p}$ sur $\mathbf{Z}_p[[\mathrm{T}]]$ commutent. Ces actions s'étendent par continuité à $\mathscr{E}^{\dagger}_{\mathbf{Q}_p}$, $\mathscr{R}_{\mathbf{Q}_p}$ et $\mathscr{R}_{\mathbf{Q}_p}[\log\mathrm{T}]$. On dispose d'autre part de « flèches de localisation en $\varepsilon^{(n)}-1$ » commutant à l'action de Galois, $\mathscr{R}_{\mathbf{Q}_p}[\log\mathrm{T}] \dashrightarrow \mathrm{F}_n[[\mathrm{T}]]$; cette flèche est en pointillés car elle n'est définie, par la formule

$$f\mapsto\varphi^{-n}(f) = f\big(\varepsilon^{(n)}(1+\mathrm{T})^{p^{-n}}-1\big) = f\Big(\varepsilon^{(n)}-1+\varepsilon^{(n)}\Big(\sum_{k=1}^{+\infty}\binom{p^{-n}}{k}\mathrm{T}^k\Big)\Big),$$

que sur le sous-anneau des $f\in\mathscr{R}_{\mathbf{Q}_p}[\log\mathrm{T}]$ qui convergent en $\varepsilon^{(n)}-1$. On obtient le diagramme d'anneaux

$$\mathscr{E}^{\dagger}_{\mathbf{Q}_p}\longrightarrow\mathscr{R}_{\mathbf{Q}_p}\longrightarrow\mathscr{R}_{\mathbf{Q}_p}[\log\mathrm{T}]\dashrightarrow\mathrm{F}_n[[\mathrm{T}]]\longrightarrow\mathrm{F}_n\ .$$

La procédure[(5)] permettant de construire $\mathscr{E}^{\dagger}_{\mathbf{Q}_p}$, $\mathscr{R}_{\mathbf{Q}_p}$, ... à partir de $\mathbf{F}_p((\mathrm{T}))$, permet, en partant de $\mathbf{F}_p((\mathrm{T}))^{\mathrm{sep}}$ d'obtenir le diagramme d'anneaux

$$\mathbf{B}^{\dagger}\longrightarrow\widetilde{\mathbf{B}}^{\dagger}_{\mathrm{rig}}\longrightarrow\widetilde{\mathbf{B}}^{\dagger}_{\log}\dashrightarrow\mathbf{B}^{+}_{\mathrm{dR}}\longrightarrow\mathbf{C}_p\ .$$

(5) Pour construire $\mathscr{E}^{\dagger}_{\mathbf{Q}_p}$ à partir de $\mathbf{F}_p((\mathrm{T}))$, il y a un choix : le choix d'un relèvement de T ou, ce qui revient au même, d'un Frobenius sur $\mathscr{E}^{\dagger}_{\mathbf{Q}_p}$; une fois ce choix fait, la construction s'étend canoniquement à $\mathbf{F}_p((\mathrm{T}))^{\mathrm{sep}}$. Notre choix de T en fait un analogue p-adique de $e^{2i\pi}-1$ et $\log(1+\mathrm{T})$ devient un analogue p-adique de $2i\pi$.

et comme on a muni $\mathbf{F}_p((\mathrm{T}))^{\mathrm{sep}}$ d'une action de $\mathscr{G}_{\mathbf{Q}_p}$ grâce à la théorie du corps des normes, tous ces anneaux sont munis d'une action continue de $\mathscr{G}_{\mathbf{Q}_p}$ commutant à celle de φ sur $\mathbf{B}^\dagger$, $\widetilde{\mathbf{B}}^\dagger_{\mathrm{rig}}$ et $\widetilde{\mathbf{B}}^\dagger_{\log}$.

0.3.4. La hiérarchie [40] *des représentations galoisiennes.* — Une représentation de $\mathscr{G}_{\mathrm{K}}$ qui est $\mathbf{B}^+_{\mathrm{dR}}[\frac{1}{\log(1+\mathrm{T})}]$-admissible est *de de Rham*; elle est *cristalline* si elle est $\widetilde{\mathbf{B}}^\dagger_{\mathrm{rig}}[\frac{1}{\log(1+\mathrm{T})}]$-admissible, *semi-stable* si elle est $\widetilde{\mathbf{B}}^\dagger_{\log}[\frac{1}{\log(1+\mathrm{T})}]$-admissible[(6)], et potentiellement semi-stable si ses périodes de de Rham appartiennent au sous-anneau de $\mathbf{B}^+_{\mathrm{dR}}[\frac{1}{\log(1+\mathrm{T})}]$ engendré par $\widetilde{\mathbf{B}}^\dagger_{\log}[\frac{1}{\log(1+\mathrm{T})}]$ et $\overline{\mathbf{Q}}_p$.

Les implications « cristalline ⇒ semi-stable ⇒ potentiellement semi-stable ⇒ de Rham » sont immédiates; les deux premières sont strictes et la conjecture de monodromie p-adique de Fontaine est que la dernière implication est en fait une équivalence. Cette conjecture était connue en dimension 1 depuis longtemps; Fontaine [41] l'avait récemment démontrée en dimension 2 et Hyodo [49] l'avait vérifiée pour les extensions de représentations semi-stables. D'autre part, Tsuji [72, 73] et Faltings [35] avaient démontré (à la suite de travaux de nombreuses personnes dont Tate, Bloch, Kato, Fontaine, Messing, Faltings, Hyodo etc.) la conjecture C_{st} de Fontaine, prouvant ainsi que les représentations provenant de la géométrie (du moins une bonne partie d'entre elles) sont potentiellement semi-stables (il faut en outre utiliser les altérations de de Jong).

0.3.5. La théorie des (φ,Γ)*-modules.* — La théorie des (φ,Γ)-modules de Fontaine ([38]+[12]) établit une équivalence de catégories entre, d'une part, les $\mathbf{Q}_p$-représentations de $\mathscr{G}_{\mathbf{Q}_p}$ (i.e. les $\mathbf{Q}_p$-espaces vectoriels de dimension finie munis d'une action continue de $\mathscr{G}_{\mathbf{Q}_p}$) et, d'autre part, les (φ,Γ)-modules étales sur $\mathscr{E}^\dagger_{\mathbf{Q}_p}$ (i.e. les $\mathscr{E}^\dagger_{\mathbf{Q}_p}$-espaces vectoriels de dimension finie munis d'actions semi-linéaires de φ et $\Gamma_{\mathbf{Q}_p}$ commutant entre elles tels que φ soit étale, c'est-à-dire de φ-pente 0). Cette équivalence de catégorie est obtenue par la procédure décrite plus haut pour le groupe $\mathscr{H}_{\mathbf{Q}_p}$, en utilisant l'anneau $\mathbf{B}^\dagger$: elle associe à V le $\mathscr{E}^\dagger_{\mathbf{Q}_p}=(\mathbf{B}^\dagger)^{\mathscr{H}_{\mathbf{Q}_p}}$-espace vectoriel $\mathbf{D}^\dagger(\mathrm{V})=(\mathbf{B}^\dagger\otimes\mathrm{V})^{\mathscr{H}_{\mathbf{Q}_p}}$.

Comme $\Gamma_{\mathbf{Q}_p}\cong\mathbf{Z}_p^*$ est essentiellement procyclique, on peut retraduire l'équivalence de catégories ci-dessus en disant qu'une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathbf{Q}_p}$ de dimension d peut se décrire entièrement à l'aide de deux matrices A et B de $\mathrm{GL}_d(\mathscr{O}^\dagger_{\mathscr{E},\mathbf{Q}_p})$ (où $\mathscr{O}^\dagger_{\mathscr{E},\mathbf{Q}_p}$ désigne l'anneau des entiers de $\mathscr{E}^\dagger_{\mathbf{Q}_p}$) vérifiant la relation de commutation $\mathrm{A}\varphi(\mathrm{B})=\mathrm{B}\gamma(\mathrm{A})$, ce qui semble plus simple *a priori* que de décrire la représentation directement.

(6) La définition habituelle fait intervenir des anneaux $\mathbf{B}^+_{\mathrm{cris}}$ et $\mathbf{B}^+_{\mathrm{st}}$ qui sont étroitement reliés à $\widetilde{\mathbf{B}}^\dagger_{\mathrm{rig}}$ et $\widetilde{\mathbf{B}}^\dagger_{\log}$; en particulier les seuls éléments intéressants pour ces histoires de classification des représentations vivent dans des sous-F-espaces vectoriels de dimension finie stables par φ, et ceux-ci sont dans l'intersection $\widetilde{\mathbf{B}}^+_{\mathrm{rig}}=\widetilde{\mathbf{B}}^\dagger_{\mathrm{rig}}\cap\mathbf{B}^+_{\mathrm{cris}}$ (resp. $\widetilde{\mathbf{B}}^+_{\log}=\widetilde{\mathbf{B}}^\dagger_{\log}\cap\mathbf{B}^+_{\mathrm{st}}$). Passer de $\mathbf{B}^+_{\mathrm{cris}}$ à $\widetilde{\mathbf{B}}^+_{\mathrm{rig}}$ revient à considérer la cohomologie rigide au lieu de la cohomologie cristalline : $\mathbf{B}^+_{\mathrm{cris}}$ peut s'interpréter [37] comme le H^0 cristallin de l'anneau des entiers de $\overline{\mathrm{F}}$ et $\widetilde{\mathbf{B}}^+_{\mathrm{rig}}$ comme son H^0 rigide.

0.3.6. Le (φ,∇)-module attaché à une représentation galoisienne. — On peut transformer un (φ,Γ)-module en un φ-module avec connexion en considérant l'action ∇_{V} de l'algèbre de Lie de $\Gamma_{\mathbf{Q}_p}$ (i.e. l'action infinitésimale de $\Gamma_{\mathbf{Q}_p}$). Le passage à l'algèbre de Lie introduit des dénominateurs (la série définissant le logarithme n'est pas à coefficients entiers), et on doit étendre les coefficients de $\mathscr{E}^{\dagger}_{\mathbf{Q}_p}$ à $\mathscr{R}_{\mathbf{Q}_p}$. L'étude du φ-module à connexion ainsi obtenu a été menée à bien par Berger [6], ce qui lui permet d'obtenir un certain nombre de résultats concernant la classification des représentations p-adiques.

Ce module différentiel ne rentre pas tout à fait dans le cadre du paragraphe précédent car, sur les coefficients, la dérivation est $\nabla = (1+\mathrm{T})\log(1+\mathrm{T})\frac{d}{d\mathrm{T}}$ et $\log(1+\mathrm{T}) \notin \mathscr{E}^{\dagger}_{\mathbf{Q}_p}$, ce qui fait que la connexion a une infinité de singularités régulières en les zéros de $\log(1+\mathrm{T})$, c'est-à-dire les $\zeta-1$, $\zeta \in \boldsymbol{\mu}_{p^\infty}$. Le résidu en chacune de ces singularités est l'opérateur de Sen (cf. [69, 23, 42]) dont les valeurs propres sont les « poids de Hodge-Tate généralisés » de V ; en particulier, ce résidu est nul si et seulement si V est $\mathbf{C}_p$-admissible.

D'autre part, utilisant la flèche $\mathscr{E}^{\dagger}_{\mathbf{Q}_p} \dashrightarrow \mathrm{F}_n[[t]]$, on montre [42] que la représentation V est de de Rham si et seulement si les singularités ne sont qu'apparentes. Pour ramener la conjecture de Fontaine à celle de Crew, Berger résout simultanément toutes ces singularités en construisant un $\mathscr{R}_{\mathbf{Q}_p}$-réseau de $\mathscr{R}_{\mathbf{Q}_p}[\frac{1}{\log(1+\mathrm{T})}] \otimes_{\mathscr{E}^{\dagger}_{\mathbf{Q}_p}} \mathbf{D}^{\dagger}(\mathrm{V})$ stable par $\partial_{\mathrm{V}} = \frac{1}{\log(1+\mathrm{T})}\nabla_{\mathrm{V}}$ et par φ, et montre que le (φ,∂)-module ainsi obtenu est quasi-unipotent si et seulement si V est potentiellement semi-stable.

1. ÉQUATIONS DIFFÉRENTIELLES p-ADIQUES

1.1. Anneaux de séries de Laurent

Soit L un corps de caractéristique 0 complet pour la valuation p-adique v_p supposée discrète (on pourrait remplacer cette hypothèse par maximalement complet). On suppose, pour se simplifier la vie, que le corps résiduel k de L est algébriquement clos (ça évitera d'avoir à étendre les scalaires à certains endroits). Remarquons que cette restriction n'est pas très sérieuse car, pour résoudre une équation différentielle, on peut toujours étendre les scalaires puis utiliser le théorème d'Ax-Sen-Tate (prop. 2.1) pour revenir au corps de base.

Soit $\mathscr{E}_{\mathrm{L}}$ l'ensemble des séries de Laurent $\sum_{k\in\mathbf{Z}} a_k \mathrm{T}^k$ telle que la suite $(v_p(a_k))_{k\in\mathbf{Z}}$ soit minorée et vérifie $\lim_{k\to-\infty} v_p(a_k) = +\infty$.

Si $r>0$, soit $\mathscr{E}_{\mathrm{L}}^{]0,r]}$ (resp. $\mathscr{E}_{\mathrm{L}}^{(0,r]}$) l'anneau des fonctions analytiques (resp. analytiques bornées) sur la couronne $0 < v_p(\mathrm{T}) \leqslant r$. C'est aussi l'ensemble des séries de Laurent $\sum_{k\in\mathbf{Z}} a_k \mathrm{T}^k$ vérifiant $\lim_{k\to-\infty} v_p(a_k) + kr = +\infty$ et $\lim_{k\to+\infty} v_p(a_k) + ks = +\infty$ quel que soit $0 < s \leqslant r$ (resp. la suite $(v_p(a_k))_{k\in\mathbf{Z}}$ est minorée). On peut aussi

obtenir $\mathscr{E}_{\mathrm{L}}^{]0,r]}$ en complétant $\mathscr{E}_{\mathrm{L}}^{(0,r]}$ pour la topologie de Fréchet induite par la famille de valuations w_s, $0 < s \leqslant r$ définies par $w_s(\sum_{k\in\mathbf{Z}} a_k \mathrm{T}^k) = \inf_{k\in\mathbf{Z}}(v_p(a_k) + sr)$.

On obtient alors $\mathscr{E}_{\mathrm{L}}^{\dagger}$ (resp. $\mathscr{R}_{\mathrm{L}}$) comme limite inductive des espaces vectoriels topologiques $\mathscr{E}_{\mathrm{L}}^{(0,r]}$ (resp. $\mathscr{E}_{\mathrm{L}}^{]0,r]}$). Finalement, soient $\mathscr{E}_{\mathrm{L}}^{+}$ et $\mathscr{R}_{\mathrm{L}}^{+}$ les intersections respectives de $\mathscr{E}_{\mathrm{L}}$ et $\mathscr{R}_{\mathrm{L}}$ avec L[[T]] ; ceci fait de $\mathscr{R}_{\mathrm{L}}^{+}$ (resp. $\mathscr{E}_{\mathrm{L}}^{+}$) l'anneau des fonctions analytiques (resp. analytiques bornées) sur le disque $0 < v_p(\mathrm{T})$. Tout élément de $\mathscr{R}_{\mathrm{L}}$ peut alors s'écrire comme la somme d'un élément de $\mathscr{E}_{\mathrm{L}}^{\dagger}$ et d'un élément de $\mathscr{R}_{\mathrm{L}}^{+}$ et on a $\mathscr{R}_{\mathrm{L}}^{+} \cap \mathscr{E}_{\mathrm{L}}^{\dagger} = \mathscr{E}_{\mathrm{L}}^{+}$.

De manière imagée, $\mathscr{E}_{\mathrm{L}}$ est l'anneau des fonctions analytiques sur une couronne vide de rayon 1 et $\mathscr{R}_{\mathrm{L}}$ (resp. $\mathscr{E}_{\mathrm{L}}^{\dagger}$) est l'anneau des fonctions analytiques (resp. analytiques bornées) sur une couronne infiniment fine de rayon 1.

Proposition 1.1. — (i) *$\mathscr{E}_{\mathrm{L}}^{+}$ et $\mathscr{E}^{(0,r]}$, $r > 0$, sont des anneaux principaux* ;

(ii) *$\mathscr{E}_{\mathrm{L}}$ et $\mathscr{E}_{\mathrm{L}}^{\dagger}$ sont des corps* ;

(iii) *$\mathscr{R}_{\mathrm{L}}$, $\mathscr{R}_{\mathrm{L}}^{+}$ et $\mathscr{E}_{\mathrm{L}}^{]0,r]}$, $r > 0$, sont des anneaux de Bézout* ;

(iv) *tout sous-module fermé d'un module libre* M *de rang fini sur $\mathscr{R}_{\mathrm{L}}$, $\mathscr{R}_{\mathrm{L}}^{+}$ ou $\mathscr{E}_{\mathrm{L}}^{(0,r]}$, $r > 0$, est libre de rang* $\leqslant \mathrm{rang}(\mathrm{M})$.

Démonstration. — Les (i) et (ii) sont des conséquences de la théorie des polygones de Newton (cf. [34], par exemple) pour les séries de Laurent (et du fait que l'on a supposé L de valuation discrète). Le (iii) est un résultat de Lazard [58] ; rappelons qu'un anneau de Bézout est un anneau dans lequel tout idéal de type fini est principal. Pour le (iv), voir [6] ; ce résultat joue un grand rôle dans la réduction de la conjecture de Fontaine à celle de Crew.

On se fixe une dérivation ∂ et un Frobenius φ de $\mathscr{E}_{\mathrm{L}}^{\dagger}$. Si E est une extension finie séparable de $k((\mathrm{T}))$, il lui correspond des extensions $\mathscr{E}_{\mathrm{L}}^{\dagger}(\mathrm{E})$ et $\mathscr{R}_{\mathrm{L}}(\mathrm{E})$ de $\mathscr{E}_{\mathrm{L}}^{\dagger}$ et $\mathscr{R}_{\mathrm{L}}$ respectivement auxquelles les actions de φ et ∂ s'étendent canoniquement. Voir le n° 2.2 pour un point de vue plus général. Une propriété que nous aurons à utiliser est l'existence (cf. prop 2.3) d'un isomorphisme d'anneaux topologiques de $\mathscr{R}_{\mathrm{L}}$ sur $\mathscr{R}_{\mathrm{L}}(\mathrm{E})$ [E est topologiquement isomorphe à $k((\mathrm{T}))$] ; en particulier, après extension séparable finie de $k((\mathrm{T}))$, on se retrouve avec la même théorie.

1.2. Les résultats de Christol et Mebkhout

Les travaux de Christol et Mebkhout ont déjà fait l'objet d'un exposé à ce séminaire [59] ; nous nous contenterons donc d'un résumé rapide. Pour une présentation détaillée de la théorie, nous renvoyons à [21].

1.2.1. Valuation de convergence d'un module différentiel. — Rappelons que, si $0 < s \leqslant r$, on définit sur $\mathscr{E}_{\mathrm{L}}^{]0,r]}$ une valuation w_s par la formule $w_s(\sum_{k\in\mathbf{Z}} a_k \mathrm{T}^k) = \inf_{k\in\mathbf{Z}} v_p(a_k) + ks$. Si $f \in \mathscr{R}_{\mathrm{L}}$, la fonction $s \mapsto w_s(f)$ est définie sur un intervalle de la forme $]0, r]$, $r > 0$ et est une fonction concave de s, affine par morceaux. Une manière commode de voir w_s est d'introduire « le »*point générique t_s de valuation s* ; il vit dans une extension transcendante algébriquement close L_s de L, complète pour une valuation étendant v_p (et encore notée v_p), et est caractérisé (à automorphisme de corps valué près) par $v_p(t_s - a) = s$ quel que soit $a \in \mathrm{L}$ de valuation s. On a alors $w_s(f) = v_p(f(t_s))$ si $f \in \mathscr{E}_{\mathrm{L}}^{]0,r]}$ et $r \geqslant s$. (Pour construire L_s, on munit $\mathrm{L}[\mathrm{X}, \mathrm{X}^{-1}]$ de la valuation $v_p(\sum_{k=k_0}^{k_1} a_k \mathrm{X}^k) = \inf_{k\in\mathbf{Z}} v_p(a_k) + ks$; cette valuation est multiplicative et donc passe au corps des fractions ; on peut alors compléter ce dernier, prendre une clôture algébrique et recompléter (maximalement si besoin est) pour obtenir L_s et prendre $t_s = \mathrm{X}$.)

Si $\mathrm{G} = (g_{i,j}) \in \mathrm{M}_d(\mathscr{R}_{\mathrm{L}})$, on définit $w_s(\mathrm{G})$ comme le minimum des $w_s(g_{i,j})$, $1 \leqslant i, j \leqslant d$.

Soient $\mathscr{D}$ un ∂-module de rang d sur $\mathscr{R}_{\mathrm{L}}$, $e_1, \dots, e_d$ une base de $\mathscr{D}$ sur $\mathscr{R}_{\mathrm{L}}$ et, si $n \in \mathbf{N}$, soit $\mathrm{G}_n = (g_{i,j})$ la matrice définie par $\partial_{\mathscr{D}}^n e_j = \sum_{i=1}^d g_{i,j} e_i$. Si $s > 0$, la suite de terme général $\frac{1}{n} w_s(\mathrm{G}_n)$ a une limite (quand elle est définie) quand n tend vers $+\infty$ (norme spectrale d'un opérateur). Ceci permet d'introduire la fonction

$$\mathrm{Val}(\mathscr{D}, s) = \sup\Big(s, \lim_{n\to+\infty} \frac{v_p(n!) - w_s(\mathrm{G}_n)}{n}\Big) = \sup\Big(s, \frac{1}{p-1} - \lim_{n\to+\infty} \frac{w_s(\mathrm{G}_n)}{n}\Big).$$

La quantité $\mathrm{Val}(\mathscr{D}, s)$ s'interprète comme la valuation de convergence des solutions de l'équation différentielle $\partial_{\mathscr{D}} \mathrm{X} = 0$ autour en t_s. (Une base des solutions autour d'un point x_0 est donnée par les séries $\sum_{n=0}^{+\infty} \partial_{\mathscr{D}}^n e_i \cdot (x_0 - \mathrm{T})^n / n!$, $1 \leqslant i \leqslant d$.)

Pour calculer cette fonction, on peut, dans les cas favorables, utiliser un résultat de Young [80] :

Proposition 1.2. — *Soient $\mathscr{L} = \partial^d + a_{d-1}(\mathrm{T})\partial^{d-1} + \cdots + a_0(\mathrm{T})$, où $a_0(\mathrm{T}), \dots, a_{d-1}(\mathrm{T}) \in \mathscr{R}_{\mathrm{L}}$ (et $\partial = \frac{d}{d\mathrm{T}}$), $\mathrm{P} = \mathrm{X}^d + a_{d-1}(t_s)\mathrm{X}^{d-1} + \cdots + a_0(t_s)$ et $u_1, \dots, u_d$ les valuations des racines de* P (ces valuations se lisent sur le polygone de Newton de P)*, alors $\mathscr{L}\mathrm{X} = 0$ admet une base de solutions $f_1, \dots, f_d$ autour de t_s dont les valuations de convergence vérifient :*

(i) *si $u_i < s$, la valuation de convergence de f_i est exactement $\frac{1}{p-1} + u_i$;*

(ii) *si $u_i \geqslant s$, la valuation de convergence de f_i est $\leqslant \frac{1}{p-1} + s$.*

Passer d'un module différentiel à un opérateur différentiel demande de trouver un *vecteur cyclique* (i.e. un vecteur X tel que $\mathrm{X}, \partial_{\mathscr{D}}\mathrm{X}, \dots, \partial_{\mathscr{D}}^{d-1}\mathrm{X}$ forment une base de $\mathscr{D}$ sur $\mathscr{R}_{\mathrm{L}}$). Cela ne peut, en général, se faire sans passer au corps des fractions de $\mathscr{R}_{\mathrm{L}}$, ce qui introduit quelques petits problèmes techniques pour se débarrasser des singularités apparentes que l'on récupère ainsi.

Un ingrédient technique très utile est *l'antécédent de Frobenius*. Soit φ un Frobenius sur $\mathscr{E}_{\mathrm{L}}^{\dagger}$. Si $\mathscr{D}$ est un ∂-module sur $\mathscr{R}_{\mathrm{L}}$, on note $\varphi^*\mathscr{D}$ le transformé de $\mathscr{D}$ par φ : si la matrice de $\partial_{\mathscr{D}}$ dans une base est G, la matrice de $\partial_{\varphi^*\mathscr{D}}$ est $\partial(\varphi(\mathrm{T}))\varphi(\mathrm{G})$. On appelle antécédent de Frobenius de $\mathscr{D}$ un ∂-module $\mathscr{N}$ tel que l'on ait $\varphi^*\mathscr{N} = \mathscr{D}$. Il ne faut pas confondre cette notion avec celle de structure de Frobenius ; un ∂-module admet une structure de Frobenius (i.e. de (φ, ∂)-module) si $\varphi^*\mathscr{D} \cong \mathscr{D}$. L'existence d'antécédents de Frobenius est garantie sous des conditions assez larges par un théorème de Christol et Dwork [16] :

PROPOSITION 1.3. — *Soit $\mathscr{D}$ un ∂-module sur l'anneau $\mathscr{E}_{\mathrm{L}}^{]r_1,r_2[}$ des fonctions analytiques sur la couronne $0 < r_1 < v_p(\mathrm{T}) < r_2$ tel que $\mathrm{Val}(\mathscr{D}, s) < \frac{1}{p} + s$ si $s \in]r_1, r_2[$. Il existe alors un* (unique) *φ-module $\mathscr{N}$ sur $\mathscr{E}_{\mathrm{L}}^{]pr_1,pr_2[}$ tel que $\mathscr{D} = \varphi^*\mathscr{N}$. De plus, on a $\mathrm{Val}(\mathscr{N}, ps) = p\mathrm{Val}(\mathscr{D}, s)$.*

On remarquera que, si la valuation de convergence de $\mathscr{D}$ est assez petite (ce qui sera le cas pour les modules solubles, cf. ci-dessous), on peut réitérer le procédé pour se ramener en un point où le théorème de Young peut être utilisé de manière efficace. Une application de ces techniques est la très utile proposition :

PROPOSITION 1.4 ([19]). — *La fonction $s \mapsto \mathrm{Val}(\mathscr{D}, s)$ est convexe, affine par morceaux, et sa dérivée ne prend pour valeurs que des nombres rationnels de dénominateurs $\leqslant d$.*

1.2.2. La théorie des ∂-pentes. — Par convexité, la limite $\lim_{s\to 0^+} \mathrm{Val}(\mathscr{D}, s)$ existe (elle peut être infinie) et le module $\mathscr{D}$ est dit *soluble* si $\lim_{s\to 0^+} \mathrm{Val}(\mathscr{D}, s) = 0$. Il existe alors un nombre rationnel $\beta \geqslant 0$, de dénominateur $\leqslant d$, tel que $\mathrm{Val}(\mathscr{D}, s) = (1+\beta)s$, si s est assez petit. Ce nombre rationnel β est appelé *la plus grande pente* de $\mathscr{D}$. On dit que $\mathscr{D}$ est *purement de pente β* si *toutes* les solutions de l'équation $\partial_{\mathscr{D}}\mathrm{X} = 0$ (et pas seulement une solution générale) ont, en t_s (s proche de 0), une valuation de convergence égale à $(1+\beta)s$.

THÉORÈME 1.5 (« de décomposition » [19, 20]). — *Un ∂-module soluble $\mathscr{D}$ sur $\mathscr{R}_{\mathrm{L}}$ possède une décomposition canonique $\mathscr{D} = \oplus_{\gamma\geqslant 0}\mathscr{D}_\gamma$, où $\mathscr{D}_\gamma$ est purement de pente γ.*

Remarque 1.6. — Si on part d'un ∂-module sur $\mathscr{E}_{\mathrm{L}}^{\dagger}$, on pourrait espérer obtenir une décomposition sur $\mathscr{E}_{\mathrm{L}}^{\dagger}$, mais l'exemple du ∂-module associé à l'opérateur $\mathscr{L} = \mathrm{T}^2 d/d\mathrm{T} + (3\mathrm{T} - \pi)d/d\mathrm{T} + 1$ (où $\pi^{p-1} = -p$) montre qu'il n'en est rien ; un calcul montre qu'il possède deux pentes distinctes sur $\mathscr{R}_{\mathrm{L}}$ et que la série $\sum_{n=0}^{+\infty} \pi^{-n} n! \mathrm{T}^n$ est une solution de $\mathscr{L}f = 0$ appartenant à $\mathscr{R}_{\mathrm{L}}$ mais pas à $\mathscr{E}_{\mathrm{L}}^{\dagger}$, ce qui exclut une décomposition sur $\mathscr{E}_{\mathrm{L}}^{\dagger}$.

THÉORÈME 1.7 (« de Hasse-Arf » [19]). — *Si $\mathscr{D}$ est un ∂-module soluble, alors la quantité $\mathrm{Irr}(\mathscr{D}) = \sum_{\gamma\geqslant 0} \gamma \cdot \mathrm{rang}\, \mathscr{D}_\gamma$ est un entier.*

Remarque 1.8. — Ce théorème est une étape dans la démonstration du *théorème de l'indice* qui était une des motivations de Christol et Mebkhout (généraliser en rang quelconque les résultats de Robba [64, 65, 66, 67, 22] en rang 1). L'irrégularité $\mathrm{Irr}(\mathscr{D})$ de $\mathscr{D}$ est définie comme l'indice d'un certain opérateur ; c'est donc un entier par définition et le théorème ci-dessus donne une formule permettant de la calculer.

Corollaire 1.9. — *Un ∂-module soluble irréductible, de rang fini sur $\mathscr{R}_{\mathrm{L}}$, n'a qu'une seule pente, et le dénominateur de cette pente divise le rang du module.*

Ce qui précède s'applique en particulier aux (φ,∂)-modules : l'existence d'un Frobenius permet de montrer que l'on a $\mathrm{Val}(\mathscr{D},s)\leqslant\frac{1}{p}\mathrm{Val}(\mathscr{D},ps)$ si s est assez petit, avec égalité si $\mathrm{Val}(\mathscr{D},ps)<\frac{p}{p-1}+ps$.

1.2.3. ∂-modules de pente 0. — Si $\mathscr{D}$ est un module différentiel de rang d sur une couronne non vide A du corps des nombres complexes et si $x_0\in\mathrm{A}$, les solutions locales de l'équation $\partial_{\mathscr{D}}\mathrm{X}=0$ autour de x_0 forment un $\mathbf{C}$-espace vectoriel de dimension d. On peut prolonger analytiquement une solution le long d'un chemin et, au bout d'un tour, on récupère une nouvelle solution autour de x_0 ; d'où un opérateur « de monodromie » N. Les valeurs propres de $\frac{1}{2i\pi}\log\mathrm{N}$ s'appellent les *exposants* de $\mathscr{D}$; ce sont des éléments de $\mathbf{C}/\mathbf{Z}$ bien définis à l'ordre près.

En p-adique, on ne dispose pas du prolongement analytique, mais Frobenius sous toutes ses coutures constitue un substitut efficace. En utilisant les antécédents de Frobenius, Christol et Mebkhout [18] (voir aussi [33]) ont réussi à définir les exposants p-adiques pour un ∂-module $\mathscr{D}$ sur $\mathscr{R}_{\mathrm{L}}$ purement de ∂-pente 0 (on dit aussi parfois que $\mathscr{D}$ vérifie *la condition de Robba*). Ces exposants appartiennent à $(\mathbf{Z}_p/\mathbf{Z})^d$ modulo une certaine relation d'équivalence assez compliquée, mais si les différences de ces exposants sont toutes non Liouville (un nombre p-adique α est non Liouville si $\lim_{n\to\pm\infty}\frac{1}{n}v_p(\alpha-n)=0$, i.e. si α est entier ou pas trop proche de $\mathbf{Z}$), alors ces exposants sont des éléments de $\mathbf{Z}_p/\mathbf{Z}$ bien définis à l'ordre près.

Théorème 1.10 (« de la monodromie p-adique » [18]). — *Si $\mathscr{D}$ est un ∂-module de rang d sur $\mathscr{R}_{\mathrm{L}}$ de ∂-pente 0, et si les différences de ses exposants sont non Liouville, alors l'équation différentielle $\partial_{\mathscr{D}}\mathrm{X}=0$ admet d solutions linéairement indépendantes dans $\mathscr{R}_{\mathrm{L}}[\mathrm{T}^{\alpha_1},\dots,\mathrm{T}^{\alpha_d},\log\mathrm{T}]\otimes_{\mathscr{R}_{\mathrm{L}}}\mathscr{D}$, où $\alpha_1,\dots,\alpha_d$ sont les exposants de $\mathscr{D}$.*

Corollaire 1.11. — *Un (φ,∂)-module sur $\mathscr{R}_{\mathrm{L}}$ de ∂-pente 0 devient unipotent sur une extension de $\mathscr{R}_{\mathrm{L}}$ associée à une extension modérément ramifiée de $k((\mathrm{T}))$.*

Démonstration. — Comme $\mathscr{D}$ est muni d'une structure de Frobenius, l'ensemble des exposants de $\mathscr{D}$ est stable par $\alpha\mapsto p\alpha$ et les exposants de $\mathscr{D}$ sont des nombres rationnels de dénominateur premier à p.

1.3. Théorie de Galois différentielle

Ce n° est consacré à la démonstration d'André [1]. Celui-ci se place dans un cadre abstrait s'appliquant à d'autres situations que celles des ∂-modules solubles sur l'anneau de Robba, et nous avons choisi de nous restreindre à ce cadre pour nous concentrer sur la construction qui est au cœur de la démonstration (il s'agit du module $\mathscr{M}_\ell^\otimes$ ci-dessous, obtenu par induction tensorielle) et son utilisation.

1.3.1. Filtration des groupes de Galois différentiels. — On peut associer [8] à un ∂-module $\mathscr{D}$ de rang d sur $\mathscr{R}_\mathrm{L}$ un groupe de Galois différentiel $\mathrm{G}_\mathscr{D}$; c'est un sous-groupe algébrique de GL_d défini sur une extension finie L' de L ; ses représentations algébriques sont en correspondance avec les ∂-modules de rang fini sur $\mathscr{R}_{\mathrm{L}'}$ que l'on obtient à partir de $\mathscr{D}$ et de son dual par produit tensoriel, sous-objet et quotient. Réciproquement, si $\mathscr{D}'$ est un tel ∂-module, son groupe de Galois différentiel est un quotient de $\mathrm{G}_\mathscr{D}$. Nous utiliserons cette correspondance pour passer du langage des représentations à celui des ∂-modules sans plus de commentaire.

Si $\gamma \geqslant 0$, on note $\mathrm{G}_\mathscr{D}^\gamma$ le plus grand sous-groupe distingué de $\mathrm{G}_\mathscr{D}$ tel que les représentations de $\mathrm{G}_\mathscr{D}/\mathrm{G}_\mathscr{D}^\gamma$ n'aient que des ∂-pentes $\leqslant \gamma$. Comme $\mathrm{G}_\mathscr{D}$ est algébrique, toute suite strictement décroissante de sous-groupes algébriques est de longueur finie. En particulier, la filtration de $\mathrm{G}_\mathscr{D}$ par les ∂-pentes n'a qu'un nombre fini de sauts (un saut est un réel $\gamma > 0$ tel $\mathrm{G}_\mathscr{D}^{\gamma-\varepsilon} \neq \mathrm{G}_\mathscr{D}^\gamma$ si $\varepsilon > 0$) $\gamma_1, \ldots, \gamma_r$ et on appelle $\gamma_0 = 0$, $\gamma_1, \ldots, \gamma_r$ les *∂-pentes* de $\mathrm{G}_\mathscr{D}$.

Remarque 1.12. — (i) Si $\mathscr{D}'$ est un sous-∂-module ou un quotient de $\mathscr{D}$, le théorème de décomposition montre que les ∂-pentes de $\mathscr{D}'$ sont incluses dans celles de $\mathscr{D}$; en particulier, elles sont inférieures ou égales à la plus grande pente de $\mathscr{D}$.

(ii) Si $\mathscr{D}_1$ et $\mathscr{D}_2$ sont deux ∂-modules sur $\mathscr{R}_\mathrm{L}$, les pentes de $\mathscr{D}_1 \otimes \mathscr{D}_2$ sont inférieures ou égales à la plus grande des plus grandes ∂-pentes de $\mathscr{D}_1$ et $\mathscr{D}_2$: la valuation de convergence du produit de deux séries entières est inférieure ou égale au maximum des valuations de convergence de ces séries.

On déduit de cette remarque que les ∂-pentes d'une représentation de $\mathrm{G}_\mathscr{D}$ sont incluses dans celles de $\mathrm{G}_\mathscr{D}$ et qu'elles sont $\leqslant \gamma$ si et seulement si $\mathrm{G}_\mathscr{D}$ agit à travers $\mathrm{G}_\mathscr{D}/\mathrm{G}_\mathscr{D}^\gamma$.

1.3.2. La conjecture de Crew. — Si $\mathscr{D}$ est un ∂-module sur $\mathscr{R}_\mathrm{L}$ et E est une extension finie séparable de $k((\mathrm{T}))$, le $\mathscr{R}_\mathrm{L}(\mathrm{E})$-module $\mathscr{R}_\mathrm{L}(\mathrm{E}) \otimes_{\mathscr{R}_\mathrm{L}} \mathscr{D}$ est un ∂-module sur $\mathscr{R}_\mathrm{L}(\mathrm{E})$ dont le groupe de Galois différentiel $\mathrm{G}_\mathscr{D}(\mathrm{E})$ est un sous-groupe de $\mathrm{G}_\mathscr{D}$. La conjecture de Crew (sous sa forme filtrée) peut se reformuler sous la forme : « si $\mathscr{D}$ est un (φ, ∂)-module *irréductible* sur $\mathscr{R}_\mathrm{L}$, il existe une extension finie séparable E de $k((\mathrm{T}))$ telle que $\mathrm{G}_\mathscr{D}(\mathrm{E}) = \{1\}$ ». Le cas de rang 1 permet de démontrer que cet énoncé est vrai si $\mathrm{G}_\mathscr{D}$ est résoluble et, plus généralement, il permet de prouver qu'il existe une extension finie séparable E de $k((\mathrm{T}))$ telle que $\mathrm{G}_\mathscr{D}(\mathrm{E})$ n'ait pas de quotient abélien non trivial.

Quitte à faire une extension séparable finie de $k((\mathrm{T}))$, on est donc ramené à vérifier le résultat suivant.

PROPOSITION 1.13. — *Si $\mathscr{D}$ est un (φ,∂)-module irréductible sur $\mathscr{R}_{\mathrm{L}}$, alors tout quotient simple de $\mathrm{G}_{\mathscr{D}}$ est abélien.*

Démonstration. — La démonstration se fait par l'absurde. Le principe est de faire apparaître des dénominateurs dans les pentes en induisant à partir d'une extension modérée de $k((\mathrm{T}))$ pour obtenir une contradiction avec le théorème de Hasse-Arf.

Soit H un quotient simple non abélien de $\mathrm{G}_{\mathscr{D}}$. Comme H est simple, il a exactement une ∂-pente γ, et celle-ci est non nulle d'après le corollaire 1.9. Soit $\mathscr{M}$ une représentation irréductible de H de dimension $d>1$; alors $\mathscr{M}$ est de pente γ. Soit $\ell\neq p$ un nombre premier ne divisant ni d, ni le numérateur de γ, ni l'ordre du groupe Out(H), et soit ζ une racine primitive ℓ-ième de l'unité. Si $0\leqslant i\leqslant \ell-1$, soit $\mathscr{M}^{(i)}$ le $\mathscr{R}_{\mathrm{L}}(\mathrm{T}^{1/\ell})$-$\partial$-module obtenu à partir de $\mathscr{M}$ via le changement de variable $\mathrm{T}\mapsto\zeta^i\mathrm{T}^{1/\ell}$; si $x\in\mathscr{M}^{(0)}$, on note $x^{(i)}$ son image dans $\mathscr{M}^{(i)}$ via le changement de variable $\mathrm{T}^{1/\ell}\mapsto\zeta^i\mathrm{T}^{1/\ell}$ (et donc $x^{(0)}=x$). Considérons les ∂-modules $\mathscr{M}_\ell$ et $\mathscr{M}_\ell^{\otimes}$ définis par :

$$\mathscr{M}_\ell=\{(x^{(0)},\dots,x^{(\ell-1)}),\ x\in\mathscr{M}^{(0)}\}\ ,$$

et $\mathscr{M}_\ell^{\otimes}$ est le sous-$\mathscr{R}_{\mathrm{L}}$-module de

$$\mathscr{M}^{(0)}\otimes_{\mathscr{R}_{\mathrm{L}}(\mathrm{T}^{1/\ell})}\cdots\otimes_{\mathscr{R}_{\mathrm{L}}(\mathrm{T}^{1/\ell})}\mathscr{M}^{(\ell-1)}$$

engendré par les $x^{(0)}\otimes\cdots\otimes x^{(\ell-1)}$, $x\in\mathscr{M}^{(0)}$. L'intérêt du passage de $\mathscr{M}$ à $\mathscr{M}_\ell$ est que la pente a été divisée par ℓ (comme la dimension a été multipliée par ℓ, cela ne fournit pas de contradiction au théorème de Hasse-Arf). Quand on étend les scalaires de $\mathscr{R}_{\mathrm{L}}(\mathrm{T})$ à $\mathscr{R}_{\mathrm{L}}(\mathrm{T}^{1/\ell})$, le module $\mathscr{M}_\ell$ devient isomorphe à la somme des $\mathscr{M}^{(i)}$; comme le groupe de Galois différentiel de chacun des $\mathscr{M}^{(i)}$ est isomorphe à H, le groupe de Galois différentiel H_ℓ de $\mathscr{M}_\ell$ est un sous-groupe du produit en couronne de H par le groupe cyclique C_ℓ (ce produit en couronne U est un produit semi-direct $1\to\mathrm{H}^\ell\to\mathrm{U}\to\mathrm{C}_\ell\to1$, où C_ℓ agit par permutation circulaire des facteurs). Comme H est simple, il y a *a priori* 3 cas :

(I) H_ℓ est le produit direct de C_ℓ et H ;

(II) H_ℓ est un produit semi-direct de C_ℓ et H ;

(III) H_ℓ est égal au produit en couronne.

Dans tous les cas, une suite de sous-groupes distingués de H_ℓ est au plus de longueur 3 et H_ℓ a au plus 3 pentes. Deux de ces pentes sont nulles [une par définition et l'autre car le quotient C_ℓ de H_ℓ correspond à l'extension modérée $\mathscr{R}_{\mathrm{L}}(\mathrm{T}^{1/\ell})$ de $\mathscr{R}_{\mathrm{L}}(\mathrm{T})$], et la troisième est la pente de $\mathscr{M}_\ell$, c'est-à-dire γ/ℓ. Pour conclure, il suffit de montrer que ces trois cas sont exclus par notre choix de ℓ.

Dans le premier cas, l'action de C_ℓ découpe $\mathscr{M}_\ell$ en ℓ morceaux de rang d et ∂-pente γ/ℓ, et comme $d\gamma/\ell$ n'est pas entier, cela contredit le théorème de Hasse-Arf. Le second cas est exclu par la condition « ℓ ne divise pas l'ordre de Out(H) ».

Dans le troisième cas, le module $\mathscr{M}_\ell^{\otimes}$ est une représentation irréductible de H_ℓ de dimension d^ℓ ; comme l'action de H_ℓ ne se factorise pas à travers C_ℓ (car $\mathscr{M}$ est irréductible, de dimension >1), la ∂-pente de ce module est γ/ℓ, ce qui contredit le théorème de Hasse-Arf puisque $d^\ell\gamma/\ell$ n'est pas entier.

1.4. Le théorème de Turrittin p-adique

Dans ce n$^{\mathrm{o}}$, nous esquissons la démonstration du « théorème de Turritin p-adique » de Mebkhout [63], la conjecture de Crew en étant un cas particulier. Pour simplifier les calculs, on utilise la dérivation $\partial = \mathrm{T}d/d\mathrm{T}$ et on note π le « π de Dwork » (i.e. une racine $(p-1)$-ième de $-p$).

1.4.1. Construction de ∂-modules solubles de rang 1. — Si P est un polynôme sans terme constant, on note $\chi(\mathrm{P})$ le module de rang 1 dont $\exp(\pi\mathrm{P}(\mathrm{T}^{-1}))$ est une base des sections horizontales.

Proposition 1.14. — *Soient $a \in \mathrm{L}$ vérifiant $v_p(a) = 0$ et $\beta \in \mathbf{N}-\{0\}$. Si $v_p(\beta) = n$, il existe un polynôme*

$$\mathrm{P}_{\beta,a}(\mathrm{X}) = \frac{a}{\beta}\mathrm{X}^\beta + \frac{pb_1}{\beta}\mathrm{X}^{\beta/p} + \cdots + \frac{p^n b_n}{\beta}\mathrm{X}^{\beta/p^n},$$

où les b_i sont des entiers d'une extension convenable de K*, tel que le module $\chi(\mathrm{P}_{\beta,a})$ soit soluble. De plus, ce module*

(i) *est de ∂-pente β ;*

(ii) *peut être muni d'une structure de Frobenius.*

L'existence de $\mathrm{P}_{\beta,a}$ est due à Robba [65]. Pour Robba, l'extension convenable dont il est question dans la proposition est assez imposante (c'est un corps algébriquement clos maximalement complet), mais Matsuda [60] a montré que l'on peut prendre les b_i dans une extension finie de L, du moins si $p \neq 2$. Le (i) se démontre alors par un calcul direct et le (ii) est dû à Chiarellotto et Christol [14].

1.4.2. Modules de ∂-pente entière. — Pour rester dans le cadre « valuation discrète », nous supposons désormais $p \neq 2$. La démonstration du théorème de Turritin p-adique repose sur l'énoncé suivant

Proposition 1.15 ([63]). — (i) *Si $\mathscr{D}$ est un ∂-module irréductible soluble sur $\mathscr{R}_\mathrm{L}$ de ∂-pente entière $\beta > 0$, alors il existe $h \in \mathbf{N}$, une extension finie L' de* L *et $a \in \mathrm{L}'$ de valuation* 0 *tel que le module $\mathscr{D} \otimes \chi(\mathrm{P}_{\beta,a})$ soit de pente $< \beta$.*

(ii) *Si $\mathscr{D}$ est un ∂-module soluble de rang premier à p n'ayant qu'une seule ∂-pente $\beta \in \mathbf{N}$, alors $\mathscr{D}$ peut se casser en deux.*

La démonstration de cette proposition est assez technique ; ce qui suit n'en est qu'une esquisse. Soit $\mathrm{I} =]\frac{1}{(p-1)\beta}, \frac{p}{(p-1)\beta}[$. On commence par prendre un antécédent de Frobenius $\mathscr{D}_h$ d'ordre h suffisamment grand pour que l'on soit dans les conditions d'utilisation du théorème de Young (prop. 1.2) et que l'on ait $\mathrm{Val}(\mathscr{D}_h, s) = (1+\beta)s$

si $s \in \mathrm{I}$; puis on choisit un vecteur cyclique de telle sorte que $\mathscr{D}_h$ soit associé à un opérateur différentiel $\mathrm{Q}(\partial) = \partial^d + a_{d-1}(\mathrm{T})\partial^{d-1} + \cdots + a_0(\mathrm{T})$, où les a_i appartiennent au corps des fractions des fonctions holomorphes sur la couronne $v_p(\mathrm{T}) \in \mathrm{I}$.

Tordre $\mathscr{D}_h$ par $\chi(a\mathrm{X}^\beta)$ revient à remplacer $\mathrm{Q}(\partial)$ par $\mathrm{Q}(\partial - \pi\beta a/\mathrm{T}^\beta) = \partial^d + \cdots + b_0(\mathrm{T})$. Comme $\mathscr{D}_h$ est purement de ∂-pente β, on a

$$w_s(a_0(\mathrm{T})) = d\Big(\frac{1}{p-1} - \beta s\Big) \text{ et } w_s(a_i(\mathrm{T})) \geqslant (d-i)\Big(\frac{1}{p-1} - \beta s\Big) \quad \text{si } s \in \mathrm{I},$$

et on montre que l'on peut trouver a dans une extension finie de L de telle sorte qu'il existe $s_0 \in \mathrm{I}$ tel que l'on ait $w_{s_0}(b_0(\mathrm{T})) > d(\frac{1}{p-1} - \beta s_0)$. Un petit calcul montre alors que, si

$$\mathrm{Q}_1(\partial) = \mathrm{Q}(\partial - \pi\partial(\mathrm{P}_{\beta,a}(\mathrm{T}^{-1}))) = \partial^d + c_{d-1}(\mathrm{T})\partial^{d-1} + \cdots + c_0(\mathrm{T})$$

est l'opérateur différentiel correspondant au module tordu $\mathscr{D}_h \otimes \chi(\mathrm{P}_{\beta,a})$, alors

(i) $w_s(c_i(\mathrm{T})) \geqslant (d-i)(\frac{1}{p-1} - \beta s)$ si $0 \leqslant i \leqslant d-1$ et $s \in \mathrm{I}$;

(ii) $w_{s_0}(c_0(\mathrm{T})) > d(\frac{1}{p-1} - \beta s)$;

(iii) $w_s(c_{d-1}(\mathrm{T})) = \frac{1}{p-1} - \beta s$ si $(d,p) = 1$ et $s \in \mathrm{I}$.

Le (i) assure que $\mathscr{D}_h \otimes \chi(\mathrm{P}_{\beta,a})$ est de plus grande ∂-pente $\leqslant \beta$ et le (iii) implique que l'on a égalité si $(d,p) = 1$. Le (ii), quant à lui, montre que $\mathscr{D}_h \otimes \chi(\mathrm{P}_{\beta,a})$ n'est pas purement de ∂-pente β (cf. th. 1.2). Il n'y a plus qu'à utiliser le théorème de décomposition et appliquer h fois Frobenius pour revenir au module initial et conclure.

Remarque 1.16. — La proposition 1.15 admet (cf. [46, prop. 3.3]) un analogue pour les représentations galoisiennes d'un corps local (à corps résiduel fini).

1.4.3. Énoncé et démonstration du théorème. — Si $\mathscr{D}$ est un ∂-module soluble de rang d sur $\mathscr{R}_\mathrm{L}$, nous dirons de façon informelle « si on ne rencontre pas de nombre de Liouville en cours de route » pour signifier « si la catégorie tannakienne engendrée par $\mathscr{D}$ et les $\mathscr{R}_\mathrm{L}(\mathrm{E})$, où E parcourt l'ensemble des extensions séparables de $k((\mathrm{T}))$, n'admet que des objets dont les exposants et leurs différences sont non Liouville ».

Théorème 1.17 ([63]). — *Si $\mathscr{D}$ est un ∂-module soluble de rang d sur $\mathscr{R}_\mathrm{L}$, et si on ne rencontre pas de nombre de Liouville en cours de route, alors il existe une extension séparable finie* E *de $k((\mathrm{T}))$ et des entiers p-adiques non Liouville $\alpha_1, \ldots, \alpha_d$ tels que l'équation différentielle $\partial_\mathscr{D}\mathrm{X} = 0$ admette d solutions linéairement indépendantes dans $\mathscr{R}_\mathrm{L}(\mathrm{E})[\mathrm{T}^{\alpha_1}, \ldots, \mathrm{T}^{\alpha_d}, \log \mathrm{T}]$.*

Remarque 1.18. — On ne rencontre pas de nombre de Liouville en cours de route, si on part d'un (φ,∂)-module, et le théorème ci-dessus admet comme conséquence la conjecture de Crew. D'un autre coté, bien qu'a priori il faille le faire exprès pour tomber sur un nombre de Liouville, on n'a aucun critère satisfaisant permettant de garantir que cela ne va pas arriver. En particulier, on ne sait pas démontrer que les exposants d'un module différentiel défini sur $\overline{\mathbf{Q}}(\mathrm{T})$ sont non Liouville.

Remarque 1.19. — Le théorème ci-dessus donne une description tout à fait satisfaisante des ∂-modules solubles ; le cas non soluble reste, quant à lui, totalement mystérieux.

Pour démontrer le théorème (sous sa forme filtrée faisant disparaître $\log \mathrm{T}$), il suffit de traiter le cas d'un module irréductible.

Si $\mathscr{D}$ est de rang 1, sa ∂-pente est entière et les propositions 1.14 et 1.15 montrent que l'on peut faire baisser cette ∂-pente en tordant par un module de rang 1 avec structure de Frobenius. Une récurrence immédiate montre que $\mathscr{D}$ est de la forme $\mathscr{D}_0 \otimes \mathscr{D}_f$, où $\mathscr{D}_0$ est de ∂-pente 0 et $\mathscr{D}_f$ admet une structure de Frobenius ; il n'y a plus qu'à utiliser le théorème de Crew (i.e. la conjecture de Crew en rang 1) pour trivialiser $\mathscr{D}_f$ et le théorème 1.10 pour conclure.

Si $\mathscr{D}$ est de rang premier à p, sa ∂-pente a un dénominateur premier à p et donc devient entière après une extension modérée de $k((\mathrm{T}))$. Une récurrence immédiate, utilisant le (ii) de la proposition 1.15 (et le fait que si $d = d_1 + d_2$ est premier à p, alors d_1 ou d_2 est premier à p), permet de montrer que $\mathscr{D}$ acquiert un constituant de rang 1 après une suite d'extensions modérées de $k((\mathrm{T}))$.

Si $\mathscr{D}$ est de rang divisible par p, le module $\mathrm{End}^0(\mathscr{D})$ des éléments de trace nulle est de rang premier à p. D'après ce qui précède, après une extension séparable finie E de $k((\mathrm{T}))$, le module $\mathrm{End}^0(\mathscr{D})$ acquiert un constituant de rang 1 et de ∂-pente 0 ; il existe alors $\alpha \in \mathbf{Z}_p$ et un morphisme non trivial de $\mathscr{D}$ dans $\mathscr{D} \otimes \mathrm{T}^\alpha$. Quitte à faire une extension modérée, on se retrouve dans l'un des deux cas suivants :

(i) le module $\mathrm{End}^0(\mathscr{D})$ a une section horizontale, et la dimension des endomorphismes de $\mathscr{D}$ commutant à la connexion est $\geqslant 2$; $\mathscr{D}$ n'est donc pas irréductible ;

(ii) il existe $\alpha \in \mathbf{Z}_p$ irrationnel et un morphisme non trivial de $\mathscr{D}$ dans $\mathscr{D} \otimes \mathrm{T}^\alpha$, et comme $\mathscr{D}$ et $\mathscr{D} \otimes \mathrm{T}^\alpha$ ne sont pas isomorphes (considérer les déterminants), cela implique que $\mathscr{D}$ n'est pas irréductible.

En conclusion, après une extension finie séparable, $\mathscr{D}$ acquiert une composante irréductible de rang strictement inférieur à celui de $\mathscr{D}$. Le théorème s'en déduit par récurrence sur le rang de $\mathscr{D}$.

2. φ-MODULES

Soit L un corps complet pour une valuation discrète étendant v_p, et dont le corps résiduel est algébriquement clos.

2.1. La stratégie de Kedlaya

Si E est une extension finie galoisienne de $k((\mathrm{T}))$ et $\mathrm{G} = \mathrm{Gal}(\mathrm{E}/k((\mathrm{T})))$, l'anneau $\mathscr{R}_{\mathrm{L}}(\mathrm{E})[\log \mathrm{T}]$ est muni d'actions de φ, ∂, G et N (où $\mathrm{N} = d/d\log \mathrm{T}$ est la dérivation $\mathscr{R}(\mathrm{E})$-linéaire normalisée par $\mathrm{N}(\log \mathrm{T}) = 1$). L'action de G commute aux autres ; celle de N commute à ∂, et on a $\mathrm{N}\varphi = p\,\varphi\mathrm{N}$.

Maintenant, si $\mathscr{D}$ est un (φ, ∂)-module sur $\mathscr{R}_{\mathrm{L}}$, alors (modulo la conjecture de Crew), il existe une extension finie galoisienne de $k((\mathrm{T}))$ telle que le L-espace vectoriel $\mathrm{V}(\mathscr{D}) = (\mathscr{R}_{\mathrm{L}}(\mathrm{E})[\log \mathrm{T}] \otimes_{\mathscr{R}_{\mathrm{L}}} \mathscr{D})^{\partial=0}$ soit de dimension d (cf. n° 0.2.3). Le L-espace vectoriel $\mathrm{V}(\mathscr{D})$ est muni d'actions de G, φ et N, et G commute à φ et N, tandis que φ et N vérifient la relation $\mathrm{N}\varphi = p\,\varphi\mathrm{N}$.

D'après le théorème de Dieudonné-Manin, $\mathrm{V}(\mathscr{D})$ admet une décomposition suivant les pentes de φ (i.e. les valuations de valeurs propres de φ) ; à cette décomposition est associée une (unique) filtration $0 = \mathrm{V}_0 \subset \mathrm{V}_1 \subset \cdots \subset \mathrm{V}_\ell = \mathrm{V}(\mathscr{D})$ telle que V_i soit stable par φ, que φ n'ait qu'une seule pente r_i sur $\mathrm{V}_i/\mathrm{V}_{i-1}$, et que l'on ait $r_1 < \cdots < r_\ell$. La relation $\mathrm{N}\varphi = p\,\varphi\mathrm{N}$ montre que $\mathrm{N}(\mathrm{V}_i) \subset \mathrm{V}_{i-1}$ et donc que cette filtration est stable par N, et que N agit trivialement sur le gradué associé.

On en déduit l'existence d'une filtration $0 = \mathscr{D}_0 \subset \mathscr{D}_1 \subset \cdots \subset \mathscr{D}_\ell = \mathscr{D}$ de $\mathscr{D}$ par des sous-φ-modules sur $\mathscr{R}_{\mathrm{L}}$, vérifiant les conditions suivantes :

(i) il existe un sous-φ-module Δ^i sur $\mathscr{E}^\dagger_{\mathrm{L}}$ de $\mathscr{D}_i/\mathscr{D}_{i-1}$ tel que $\mathscr{D}_i/\mathscr{D}_{i-1} = \mathscr{R}_{\mathrm{L}} \otimes_{\mathscr{E}^\dagger_{\mathrm{L}}} \Delta^i$;

(ii) il existe un φ-module W_i sur L tel que

$$\mathscr{R}_{\mathrm{L}}(\mathrm{E}) \otimes_{\mathrm{L}} \mathrm{W}_i = \mathscr{R}_{\mathrm{L}}(\mathrm{E}) \otimes_{\mathscr{R}_{\mathrm{L}}} (\mathscr{D}_i/\mathscr{D}_{i-1}) \quad \text{et} \quad \mathscr{E}^\dagger_{\mathrm{L}}(\mathrm{E}) \otimes_{\mathrm{L}} \mathrm{W}_i = \mathscr{E}^\dagger_{\mathrm{L}}(\mathrm{E}) \otimes_{\mathscr{E}^\dagger_{\mathrm{L}}} \Delta^i.$$

(On a $\mathrm{W}_i = \mathrm{V}_i/\mathrm{V}_{i-1}$ et $\Delta^i = \Delta_i/\Delta_{i-1}$ avec $\Delta_i = (\mathscr{E}^\dagger_{\mathrm{L}}(\mathrm{E})[\log \mathrm{T}] \otimes_{\mathrm{L}} \mathrm{V}_i)^{\mathrm{N}=0,\mathrm{G}=1}$; pour vérifier que tout marche bien, il faut utiliser le fait que $\mathscr{E}^\dagger_{\mathrm{L}}(\mathrm{E})[\log \mathrm{T}]$ est un sous-anneau de $\mathscr{R}_{\mathrm{L}}(\mathrm{E})[\log \mathrm{T}]$ stable par ∂, φ, G et N.)

Maintenant, si on part d'un φ-module sur $\mathscr{R}_{\mathrm{L}}$ (sans ∂-structure), on peut se demander ce qui reste vrai. Le (ii) est manifestement trop fort (il est plus ou moins équivalent à l'existence d'une ∂-structure pour laquelle W_i est le L-espace vectoriel des sections horizontales). Ceci amène Kedlaya à introduire deux notions de φ-pentes, les $\mathscr{E}$-pentes et les $\mathscr{R}$-pentes (« pentes génériques » et « pentes spéciales » chez Kedlaya) ; il remplace alors (ii) par « (ii′) les $\mathscr{E}$-pentes et les $\mathscr{R}$-pentes de Δ^i sont les mêmes » (c'est trivialement le cas si W_i existe), et montre qu'*un φ-module quelconque sur $\mathscr{R}_{\mathrm{L}}$ a une filtration vérifiant* (i) et (ii′). D'autre part, il montre que, si on est parti d'un (φ, ∂)-module, alors la filtration est stable par ∂, ce qui permet de ramener la conjecture de Crew à un cas traité par Tsuzuki [74].

Si on essaie de faire en sens inverse le chemin ci-dessus, on tombe sur un os : on ne sait pas quelle extension galoisienne E de $k((\mathrm{T}))$ va marcher. Cela oblige à les considérer toutes et donc à construire un anneau $\mathscr{R}(k((\mathrm{T}))^{\mathrm{sep}})$ (cf. n° 2.2) ; malheureusement cet anneau est beaucoup trop gros et il faut ensuite « décompléter » pour redescendre à une extension finie de $k((\mathrm{T}))$. Toutes ces étapes sont assez techniques ...

2.2. Les « foncteurs » $\mathscr{E}$, $\mathscr{E}^\dagger$ et $\mathscr{R}$

Soit E_0 un corps de caractéristique p complet pour une valuation discrète v_E de corps résiduel k parfait. Notre but dans ce n° est d'associer[7] à une extension algébrique E de E_0 ou au complété d'une telle extension, des anneaux $\mathscr{E}(\mathrm{E})$, $\mathscr{E}^\dagger(\mathrm{E})$ et $\mathscr{R}(\mathrm{E})$. Si E est parfait, ces anneaux sont canoniques, mais, dans le cas général, la construction dépend du choix d'un Frobenius sur $\mathscr{E}^\dagger(\mathrm{T})$ et d'un isomorphisme de $k((\mathrm{T}))$ sur E_0 ou, ce qui revient au même, d'une uniformisante de E_0. Si $\mathrm{E}_0 = k((\mathrm{T}))$, on retombe sur les anneaux $\mathscr{E} = \mathscr{E}(\mathrm{T})$, $\mathscr{E}^\dagger = \mathscr{E}^\dagger(\mathrm{T})$ et $\mathscr{R} = \mathscr{R}(\mathrm{T})$ définis précédemment.

2.2.1. Généralités sur les corps valués complets

Proposition 2.1. — (i) *Si* K *est un corps complet pour une valuation* $v : \mathrm{K} \to \mathbf{R} \cup \{+\infty\}$*, et si* $\overline{\mathrm{K}}$ *est une clôture algébrique de* K*, alors* v *s'étend de manière unique à* $\overline{\mathrm{K}}$*.*

(ii) *Le complété* $\widehat{\overline{\mathrm{K}}}$ *de* $\overline{\mathrm{K}}$ *pour cette valuation est un corps algébriquement clos.*

(iii) *La clôture séparable* $\mathrm{K}^{\mathrm{sep}}$ *de* K *dans* $\overline{\mathrm{K}}$ *est dense dans* $\widehat{\overline{\mathrm{K}}}$*.*

(iv) $\mathrm{Gal}(\mathrm{K}^{\mathrm{sep}}/\mathrm{K})$ *s'identifie au groupe des automorphismes continus de* $\widehat{\overline{\mathrm{K}}}$ *laissant* K *fixe et, si* H *est un sous-groupe de* $\mathrm{Gal}(\mathrm{K}^{\mathrm{sep}}/\mathrm{K})$*, le sous-corps de* $\widehat{\overline{\mathrm{K}}}$ *fixé par* H *est le complété de la clôture radicielle de* $(\mathrm{K}^{\mathrm{sep}})^{\mathrm{H}}$*.*

Remarque 2.2. — Le (iii) n'a, bien évidemment, d'intérêt qu'en caractéristique p et le (iv) est le théorème d'Ax-Sen-Tate ([3], [71]) ; contrairement aux apparences, ce n'est pas une conséquence formelle de la théorie de Galois.

2.2.2. Le cas E *parfait.* — Soit $\mathrm{F} = \mathrm{W}(k)[\frac{1}{p}]$ le corps des fractions de l'anneau des vecteurs de Witt à coefficients dans k, ce qui fait de F un corps complet pour la valuation v_p, d'anneau des entiers $\mathscr{O}_\mathrm{F} = \mathrm{W}(k)$ et de corps résiduel k_F.

Si E est un corps valué de caractéristique p, on note E^+ l'anneau de ses entiers.

Soit maintenant $\widetilde{\mathrm{E}}$ un corps parfait de caractéristique p muni d'une valuation v_E. Soient $\mathscr{O}_\mathscr{E}(\widetilde{\mathrm{E}}) = \mathrm{W}(\widetilde{\mathrm{E}})$ et $\mathscr{O}^+_\mathscr{E}(\widetilde{\mathrm{E}}) = \mathrm{W}(\widetilde{\mathrm{E}}^+)$. Soient $\mathscr{E}(\widetilde{\mathrm{E}}) = \mathscr{O}_\mathscr{E}(\widetilde{\mathrm{E}})[\frac{1}{p}]$ et $\mathscr{E}^+(\widetilde{\mathrm{E}}) = \mathscr{O}^+_\mathscr{E}(\widetilde{\mathrm{E}})[\frac{1}{p}]$. Les anneaux $\mathscr{O}^+_\mathscr{E}(\widetilde{\mathrm{E}}) \subset \mathscr{O}_\mathscr{E}(\widetilde{\mathrm{E}})$ sont, par construction, séparés et complets pour la topologie p-adique (*topologie forte*) et on a $\mathscr{O}_\mathscr{E}(\widetilde{\mathrm{E}})/p\mathscr{O}_\mathscr{E}(\widetilde{\mathrm{E}}) = \widetilde{\mathrm{E}}$ et $\mathscr{O}^+_\mathscr{E}(\widetilde{\mathrm{E}})/p\mathscr{O}^+_\mathscr{E}(\widetilde{\mathrm{E}}) = \widetilde{\mathrm{E}}^+$. Par ailleurs, $\mathscr{E}(\widetilde{\mathrm{E}})$ est un corps complet pour la valuation p-adique (la topologie associée est la *topologie forte*), d'anneau de valuation $\mathscr{O}_\mathscr{E}(\widetilde{\mathrm{E}})$ et de corps résiduel $\widetilde{\mathrm{E}}$.

Si $x \in \widetilde{\mathrm{E}}$, on note $[x]$ son représentant de Teichmüller dans $\mathscr{O}_\mathscr{E}(\widetilde{\mathrm{E}})$; rappelons que c'est l'unique élément de $\mathscr{O}_\mathscr{E}(\widetilde{\mathrm{E}})$ ayant x pour réduction modulo p et possédant une racine p^n-ième dans $\mathscr{O}_\mathscr{E}(\widetilde{\mathrm{E}})$ quel que soit $n \in \mathbf{N}$. Tout élément x de $\mathscr{O}_\mathscr{E}(\widetilde{\mathrm{E}})$

[7] C'est une extension de la théorie des anneaux de Cohen [9] développée dans [57] ; les anneaux $\mathscr{E}(\mathrm{E})$, $\mathscr{E}^\dagger(\mathrm{E})$ et $\mathscr{R}(\mathrm{E})$ correspondent respectivement aux anneaux $\Gamma^\mathrm{E}[\frac{1}{p}]$, $\Gamma^\mathrm{E}_{\mathrm{con}}[\frac{1}{p}]$ et $\Gamma^\mathrm{E}_{\mathrm{an,con}}$ de Kedlaya.

(resp. $\mathscr{O}^+_{\mathscr{E}}(\widetilde{\mathrm{E}})$) s'écrit donc de manière unique sous la forme $\sum_{k=0}^{+\infty} p^k[x_k]$, où $(x_k)_{k\in\mathbf{N}}$ est une suite d'éléments de $\widetilde{\mathrm{E}}$ (resp. $\widetilde{\mathrm{E}}^+$) et tout élément de $\mathscr{E}(\widetilde{\mathrm{E}})$ ou $\mathscr{E}^+(\widetilde{\mathrm{E}})$ sous la forme $\sum_{k\gg -\infty}^{+\infty} p^k[x_k]$.

La *topologie faible* sur $\mathscr{O}_{\mathscr{E}}(\widetilde{\mathrm{E}})$ (resp. $\mathscr{O}^+_{\mathscr{E}}(\widetilde{\mathrm{E}})$) est, par définition, la topologie qui fait de l'application $x = \sum_{k=0}^{+\infty} p^k[x_k] \to (x_k)_{k\in\mathbf{N}}$ un homéomorphisme de $\mathscr{O}_{\mathscr{E}}(\widetilde{\mathrm{E}})$ [resp. $\mathscr{O}^+_{\mathscr{E}}(\widetilde{\mathrm{E}})$] sur $\widetilde{\mathrm{E}}^{\mathbf{N}}$ [resp. $(\widetilde{\mathrm{E}}^+)^{\mathbf{N}}$] muni de la topologie produit ($\widetilde{\mathrm{E}}$ et $\widetilde{\mathrm{E}}^+$ étant muni de la topologie définie par la valuation $v_{\mathbf{E}}$) ; si $\overline{\pi} \in \widetilde{\mathrm{E}}$ vérifie $v_{\mathrm{E}}(\overline{\pi}) > 0$, la topologie faible est aussi obtenue en prenant les $[\overline{\pi}]^k\mathscr{O}^+_{\mathscr{E}}(\widetilde{\mathrm{E}}) + p^{n+1}\mathscr{O}_{\mathscr{E}}(\widetilde{\mathrm{E}})$ [resp. les $[\overline{\pi}]^k\mathscr{O}^+_{\mathscr{E}}(\widetilde{\mathrm{E}}) + p^{n+1}\mathscr{O}^+_{\mathscr{E}}(\widetilde{\mathrm{E}})$], pour $k, n \in \mathbf{N}$, comme base de voisinages de 0. On munit $\mathscr{E}(\widetilde{\mathrm{E}}) = \bigcup_{n\in\mathbf{N}} p^{-n}\mathscr{O}_{\mathscr{E}}(\widetilde{\mathrm{E}})$ et $\mathscr{E}^+(\widetilde{\mathrm{E}})$ de la topologie de la limite inductive.

Si $r > 0$, soit $\mathscr{O}^{(0,r]}_{\mathscr{E}}(\widetilde{\mathrm{E}})$ le sous-anneau des $x = \sum_{k=0}^{+\infty} p^k[x_k] \in \mathscr{O}_{\mathscr{E}}(\widetilde{\mathrm{E}})$ tels que $rv_{\mathrm{E}}(x_k) + k$ tend vers $+\infty$ quand k tend vers $+\infty$. On munit $\mathscr{O}^{(0,r]}_{\mathscr{E}}(\widetilde{\mathrm{E}})$ de la topologie définie par la valuation w_r, avec $w_r(x) = \inf_{k\in\mathbf{Z}} rv_{\mathrm{E}}(x_k) + k$, topologie pour laquelle il est complet. On munit $\mathscr{E}^{(0,r]}(\widetilde{\mathrm{E}}) = \mathscr{O}^{(0,r]}_{\mathscr{E}}(\widetilde{\mathrm{E}})[\frac{1}{p}]$ et $\mathscr{E}^\dagger(\widetilde{\mathrm{E}}) = \bigcup_{r>0}\mathscr{E}^{(0,r]}(\widetilde{\mathrm{E}})$ de la topologie de la limite inductive.

Les w_r, $r > 0$, forment une famille de valuations sur $\mathscr{E}^+(\widetilde{\mathrm{E}})$ et on définit $\mathscr{R}^+(\widetilde{\mathrm{E}})$ comme le complété de $\mathscr{E}^+(\widetilde{\mathrm{E}})$ pour la topologie de Fréchet définie par cette famille (une suite u_n tend vers u dans $\mathscr{R}^+(\widetilde{\mathrm{E}})$ si et seulement si, quel que soit $r > 0$, $w_r(u_n - u) \to +\infty$ quand $n \to +\infty$).

De même, si $r > 0$, soit $\mathscr{E}^{]0,r]}(\widetilde{\mathrm{E}})$ le complété de $\mathscr{E}^{(0,r]}(\widetilde{\mathrm{E}})$ pour la topologie de Fréchet définie par la famille de valuations w_s, $r \geqslant s > 0$. Finalement, soit $\mathscr{R}(\widetilde{\mathrm{E}}) = \bigcup_{r>0}\mathscr{E}^{]0,r]}(\widetilde{\mathrm{E}})$ muni de la topologie de la limite inductive.

L'anneau $\mathscr{E}^\dagger(\widetilde{\mathrm{E}})$ est un sous-corps de $\mathscr{E}(\widetilde{\mathrm{E}})$; ses éléments sont dits *surconvergents*. Les anneaux $\mathscr{E}^\dagger(\widetilde{\mathrm{E}})$ et $\mathscr{R}^+(\widetilde{\mathrm{E}})$ sont des sous-anneaux de $\mathscr{R}(\widetilde{\mathrm{E}})$, tout élément de $\mathscr{R}(\widetilde{\mathrm{E}})$ peut s'écrire comme la somme d'un élément de $\mathscr{E}^\dagger(\widetilde{\mathrm{E}})$ et d'un élément de $\mathscr{R}^+(\widetilde{\mathrm{E}})$ et $\mathscr{R}^+(\widetilde{\mathrm{E}}) \cap \mathscr{E}^\dagger(\widetilde{\mathrm{E}}) = \mathscr{E}^+(\widetilde{\mathrm{E}})$.

L'action de φ sur $\mathscr{O}_{\mathscr{E}}(\widetilde{\mathrm{E}})$ s'étend par continuité aux anneaux $\mathscr{E}^\dagger(\widetilde{\mathrm{E}})$, $\mathscr{R}(\widetilde{\mathrm{E}})$ et $\mathscr{R}^+(\widetilde{\mathrm{E}})$. L'action de φ est est bijective sur $\mathscr{E}^\dagger(\widetilde{\mathrm{E}})$, $\mathscr{R}(\widetilde{\mathrm{E}})$ et $\mathscr{R}^+(\widetilde{\mathrm{E}})$ et induit un isomorphisme de $\mathscr{E}^{(0,r]}(\widetilde{\mathrm{E}})$ sur $\mathscr{E}(\widetilde{\mathrm{E}})^{(0,\frac{r}{p}]}$ et de $\mathscr{E}^{]0,r]}(\widetilde{\mathrm{E}})$ sur $\mathscr{E}(\widetilde{\mathrm{E}})^{]0,\frac{r}{p}]}$.

2.2.3. Le cas général. — Si E est une extension algébrique de E_0 ou le complété d'une telle extension, notons $\widetilde{\mathrm{E}}$ le complété de sa clôture radicielle.

Choisissons un Frobenius φ sur $\mathscr{E}^\dagger_{\mathrm{L}}$ et une uniformisante $\overline{\pi}$ de E_0. On a alors $\varphi(\mathrm{T}) = \mathrm{T}^p + p\sum_{k\in\mathbf{Z}} a_k\mathrm{T}^k$, où les a_k sont des éléments de $\mathscr{O}_{\mathrm{L}}$ vérifiant une condition de décroissance convenable, et il existe dans $\mathscr{O}^\dagger_{\mathscr{E}}(\widetilde{\mathrm{E}})$ un unique élément π dont la réduction modulo p est $\overline{\pi}$ et tel que l'on ait $\varphi(\pi) = \pi^p + p\sum_{k\in\mathbf{Z}} a_k\pi^k$. On note $\mathscr{E}(\mathrm{E}_0)$ l'image de $\mathscr{E}$ dans $\mathscr{E}(\widetilde{\mathrm{E}})$ par l'application $f \mapsto f(\pi)$. Par construction, $\mathscr{E}(\mathrm{E}_0)$ est un sous-corps de $\mathscr{E}(\widetilde{\mathrm{E}})$ stable par φ, complet pour la topologie forte, de corps résiduel E_0.

Maintenant, si E est une extension finie de E_0, alors $\mathscr{E}(\widetilde{E})$ contient une unique extension algébrique $\mathscr{E}(E)$ de $\mathscr{E}(E_0)$ dont le corps résiduel est E.

Si E est une extension algébrique de E_0 (resp. le complété d'une extension algébrique de E_0), on note $\mathscr{E}(E)$ l'adhérence dans $\mathscr{E}(\widetilde{E})$ de $\bigcup_{E'\subset E}\mathscr{E}(E')$ pour la topologie forte (resp. faible), où E' parcourt les extensions finies de E_0 contenues dans E. On note $\mathscr{O}_{\mathscr{E}}(E)=\mathscr{O}_{\mathscr{E}}(\widetilde{E})\cap\mathscr{E}(E)$ l'anneau des entiers de $\mathscr{E}(E)$ et on a, dans tous les cas, $\mathscr{O}_{\mathscr{E}}(E)/p\mathscr{O}_{\mathscr{E}}(E)=E$.

Soit $\mathscr{E}^{\dagger}(E)=\mathscr{E}^{\dagger}(\widetilde{E})\cap\mathscr{E}(E)$ le sous-corps des éléments surconvergents de $\mathscr{E}(E)$. Plus généralement, si $r>0$, soit $\mathscr{E}^{(0,r]}(E)=\mathscr{E}^{(0,r]}(\widetilde{E})\cap\mathscr{E}(E)$. On note $\mathscr{E}^{]0,r]}(E)$ l'adhérence de $\mathscr{E}^{(0,r]}(E)$ dans $\mathscr{E}^{]0,r]}(\widetilde{E})$ et $\mathscr{R}(E)$ l'adhérence de $\mathscr{E}^{\dagger}(E)$ dans $\mathscr{R}(\widetilde{E})$. On a $\mathscr{R}(\widehat{E})=\mathscr{R}(E)$, si $\widehat{E}$ est le complété de E pour la valuation v_E.

Si H est un sous-groupe fermé du groupe $\mathrm{Aut}_{E_0}E$ des automorphismes continus de E laissant E_0 fixe, et X est l'un des foncteurs $\mathscr{E}$, $\mathscr{E}^{\dagger}$, $\mathscr{E}^{]0,r]}$ ... ci-dessus, alors H opère continûment sur X(E), mais il faut faire un peu attention quand on prend les points fixes sous l'action de H. Par exemple, on a $\mathscr{E}(E)^H=\mathscr{E}(E^H)$ et $\mathscr{E}^{\dagger}(E)^H=\mathscr{E}^{\dagger}(E^H)$, mais $\mathscr{R}(E)^H=\mathscr{R}(\widehat{E})^H=\mathscr{R}(\widehat{E}^H)$ n'est pas, en général, égal à $\mathscr{R}(E^H)$.

Plus généralement, si L est une extension finie totalement ramifiée de F, munie d'une extension de φ, on peut tensoriser tous les anneaux précédents par $\mathscr{O}_L$ au-dessus de $\mathscr{O}_F=W(k)$; on dénote cette opération en rajoutant un L en indice.

Si E est une extension finie E_0 de corps résiduel k', soient $E'_0=k'\cdot E_0$ et $L'=W(k')\otimes_{W(k)}L$. Si $\overline{\pi}_E$ est une uniformisante de E, si $\overline{P}(X)=X^d+\overline{a}_{d-1}X^{d-1}+\cdots+\overline{a}_0\in(E'_0)^+[X]$ est le polynôme minimal de $\overline{\pi}_E$ et si $a_i\in\mathscr{O}^{\dagger}_{\mathscr{E},L}(E'_0)$ est un relèvement surconvergent de $\overline{a}_i$, alors $P=X^d+a_{d-1}X^{d-1}+\cdots+a_0$ (resp. $\varphi(P)$) a une unique racine π_E (resp. $\varphi(\pi_E)$) dans $\mathscr{O}_{\mathscr{E}}(\widetilde{E})$ dont l'image dans E est $\overline{\pi}_E$ (resp. $\overline{\pi}_E^p$) et on a $\pi_E\in\mathscr{E}^{\dagger}(E)$. (Lemme de Hensel si E/E_0 est séparable ; le cas général s'en déduit en appliquant φ le nombre de fois qu'il faut.) Ceci permet de montrer que φ s'étend de manière unique à tous les anneaux construits ci-dessus ; son action commute à celle de $\mathrm{Aut}_{E_0}E$.

La dérivation ∂ s'étend de manière unique à $\mathscr{E}^{\dagger}_L(E)$ et $\mathscr{R}_L(E)$; on a par exemple

$$\partial\pi_E=-\frac{\partial a_{d-1}\pi_E^{d-1}+\cdots+\partial a_0}{d\pi_E^{d-1}+(d-1)a_{d-1}+\cdots+a_1}.$$

Elle s'étend par continuité à $\mathscr{E}_L(E)$ si E est une extension *séparable* quelconque de E_0. Par contre, *elle ne s'étend pas par continuité à* $\mathscr{E}^{\dagger}_L(E)$ *ou* $\mathscr{R}_L(E)$ si E est une extension infinie « trop ramifiée » de E_0.

PROPOSITION 2.3. — *L'application* $f(T)\mapsto f(\pi_E)$ *induit un isomorphisme de*

(i) $\mathscr{E}_{L'}(T)$ *sur* $\mathscr{E}_L(E)$ *;*

(ii) $\mathscr{E}^{\dagger}_{L'}(T)$ *sur* $\mathscr{E}^{\dagger}_L(E)$ *;*

(iii) $\mathscr{R}_{L'}(T)$ *sur* $\mathscr{R}_L(E)$.

D'autre part, on a le résultat suivant généralisant le théorème de Lazard.

PROPOSITION 2.4 ([57]). — *Si* E *est une extension algébrique de* E_0 *ou le complété d'une telle extension, et si* L *est une extension finie de* F*, alors* $\mathscr{R}_{\mathrm{L}}(\mathrm{E})$ *est un anneau de Bézout.*

2.3. φ-Modules sur l'anneau de Robba

Si A est un anneau muni d'un endomorphisme φ, *un φ-module sur* A est un A-module libre de rang fini muni d'une action semi-linéaire de φ telle que la matrice de φ dans une base soit inversible.

2.3.1. Le théorème de Dieudonné-Manin. — Soit κ un corps algébriquement clos de caractéristique p et M une extension finie du corps des fractions $\mathrm{W}(\kappa)[\frac{1}{p}]$ de l'anneau des vecteurs de Witt à coefficients dans κ. On suppose que l'action de φ sur $\mathrm{W}(\kappa)$ s'étend à M. Le sous-corps de M fixe par φ est une extension finie totalement ramifiée de $\mathbf{Q}_p$. ayant même groupe de valuation que M ; il contient donc une uniformisante π de M.

Soit D un φ-module de rang d sur M. Soit v un vecteur cyclique (tel que $v, \varphi(v), \dots, \varphi^{d-1}(v)$ forment une base de D sur M) et soit $\mathrm{P}(\mathrm{X}) = \mathrm{X}^d + a_{d-1}\mathrm{X}^{d-1} + \cdots + a_0$ le polynôme défini par $\varphi^d(v)+a_{d-1}\varphi^{d-1}(v)+\cdots+a_0 v = 0$. Soient $r_1, \dots, r_d$ les valuations des racines de p dans une extension de M (ces valuations se lisent directement sur le polygone de Newton de P).

PROPOSITION 2.5. — *Les r_i ne dépendent, à l'ordre près, que de* D *et pas du choix de v et si, pour $1 \leqslant i \leqslant d$, il existe $\alpha_i \in \mathrm{M}$ avec $v_p(\alpha_i) = r_i$, alors* D *admet une base $(e_1, \dots, e_d)$ telle que l'on ait $\varphi(e_i) = \alpha_i e_i$ pour $1 \leqslant i \leqslant d$.*

La proposition ci-dessus est une version du théorème de Dieudonné-Manin. Remarquons que l'on peut assurer l'existence des α_i en adjoignant à M une racine d'ordre convenable de π. Les r_i s'appellent les *φ-pentes* de D.

Notons, pour le reste de cette partie, $\widetilde{\mathrm{E}}$ le complété de la clôture algébrique de $k((\mathrm{T}))$, où k est algébriquement clos de caractéristique p. On note $\widetilde{\mathscr{R}}$, $\widetilde{\mathscr{E}}^{\dagger}$ et $\widetilde{\mathscr{E}}$ respectivement les anneaux $\mathscr{R}_{\mathrm{L}}(\widetilde{\mathrm{E}})$, $\mathscr{E}^{\dagger}_{\mathrm{L}}(\widetilde{\mathrm{E}})$ et $\mathscr{E}_{\mathrm{L}}(\widetilde{\mathrm{E}})$.

La proposition 2.5 ci-dessus s'applique en particulier à un φ-module sur $\widetilde{\mathscr{E}}$. Si $\mathrm{E} \subset \widetilde{\mathrm{E}}$ et si D est un φ-module sur $\mathscr{E}^{\dagger}(\mathrm{E})$, on appelle *$\mathscr{E}$-pentes de* D les φ-pentes de $\widetilde{\mathscr{E}} \otimes_{\mathscr{E}^{\dagger}_{\mathrm{L}}(\mathrm{E})} \mathrm{D}$. On dit que D est *isocline* s'il n'a qu'une $\mathscr{E}$-pente et *étale* si cette pente est nulle.

2.3.2. Un analogue du théorème de Dieudonné-Manin sur l'anneau de Robba

THÉORÈME 2.6 ([57]). — *Si* D *est un φ-module de rang d sur $\widetilde{\mathscr{R}}$, alors* D *possède une base de vecteurs propres sur $\mathscr{R}_{\mathrm{L}'}(\widetilde{\mathrm{E}})$, où* L' *est une extension finie de* L *; de plus, les valuations des valeurs propres correspondantes ne dépendent, à l'ordre près, que de* D *; ces valuations s'appellent* les $\mathscr{R}$-pentes *de* D.

Remarque 2.7. — En rang 1, un φ-module sur $\widetilde{\mathscr{R}}$ est défini sur $\widetilde{\mathscr{E}}^{\dagger}$, et les $\mathscr{E}$-pente et $\mathscr{R}$-pente coïncident.

La démonstration de ce théorème s'apparente à un numéro de funambulisme. On commence par montrer que, si $v_p(\lambda)$ est assez grand, il existe $v \in \mathrm{D}$ vérifiant $\varphi(v) = \lambda v$ (c'est loin d'être trivial). On utilise alors le théorème de Lazard généralisé (prop. 2.4) pour montrer que ce vecteur est multiple d'un vecteur propre w [pour une valeur propre de valuation $\leqslant v_p(\lambda)$] primitif (i.e. que l'on peut compléter en une base de D sur $\widetilde{\mathscr{R}}$). Appliquant ceci au module $\mathrm{D}/\langle w\rangle$, on fabrique une base de D dans laquelle la matrice de φ est triangulaire. La somme des valuations des termes diagonaux est égale à la $\mathscr{E}$-pente de $\det \mathrm{D}$; en particulier, elle ne dépend pas de la base choisie.

Passer d'une matrice triangulaire à une matrice diagonale n'est pas une mince affaire. Par exemple, en dimension 2, si on part d'une base v_1, v_2 dans laquelle l'action de φ est donnée par $\varphi(v_1) = \lambda_1 v_1$ et $\varphi(v_2) = \lambda_2 v_2 + a v_1$, il y a deux cas suivant que $v_{\mathrm{L}}(\lambda_2) - v_{\mathrm{L}}(\lambda_1) \geqslant 0$ ou $v_{\mathrm{L}}(\lambda_2) - v_{\mathrm{L}}(\lambda_1) < 0$ (où v_{L} désigne la valuation normalisée de L). Dans le premier cas, l'équation $\varphi(b) - \frac{\lambda_2}{\lambda_1} b = \frac{a}{\lambda_1}$ a une solution dans $\widetilde{\mathscr{R}}$ et $(v_1, v_2 + b v_1)$ est une base constituée de vecteurs propres ; dans le second cas, l'équation précédente n'a, en général, pas de solution et il faut procéder autrement. Kedlaya montre (et c'est là le point le plus délicat) que, si $v_{\mathrm{L}}(\lambda_2) - v_{\mathrm{L}}(\lambda_1) \leqslant -2$, alors D contient un vecteur propre primitif pour une valeur propre λ_1' avec $v_{\mathrm{L}}(\lambda_1') < v_{\mathrm{L}}(\lambda_1)$. Si $v_{\mathrm{L}}(\lambda_2) - v_{\mathrm{L}}(\lambda_1)$ est pair (et on peut toujours se ramener à ce cas en remplaçant L par une extension quadratique), après un nombre fini d'étapes, on se ramène au cas $v_{\mathrm{L}}(\lambda_2) - v_{\mathrm{L}}(\lambda_1) \geqslant 0$.

Le cas général ne peut pas se ramener au cas de la dimension 2 car la construction précédente ne fera jamais apparaître que des pentes dont le dénominateur est une puissance de 2, et la démonstration dans le cas général est franchement technique.

2.3.3. Filtration par les pentes

THÉORÈME 2.8 ([57]). — *Si* D *est un φ-module sur $\mathscr{R}_{\mathrm{L}}$, alors* D *admet une unique filtration $0 = \mathrm{D}_0 \subset \mathrm{D}_1 \subset \cdots \subset \mathrm{D}_\ell = \mathrm{D}$ par des sous-φ-modules sur $\mathscr{R}_{\mathrm{L}}$ vérifiant les conditions suivantes* :

(i) $\mathrm{D}_i/\mathrm{D}_{i-1}$ *est un φ-module sur $\mathscr{R}_{\mathrm{L}}$ n'ayant qu'une seule $\mathscr{R}$-pente s_i* ;

(ii) $s_1 < s_2 < \cdots < s_\ell$;

(iii) $\mathrm{D}_i/\mathrm{D}_{i-1}$ *contient un sous-φ-module Δ_i sur $\mathscr{E}^{\dagger}$ isocline de $\mathscr{E}$-pente s_i tel que l'on ait $\mathrm{D}_i/\mathrm{D}_{i-1} = \mathscr{R}_{\mathrm{L}} \otimes_{\mathscr{E}_{\mathrm{L}}^{\dagger}} \Delta_i$.*

La démonstration de ce théorème est presque aussi acrobatique que celle du précédent. Soient $e_1, \dots, e_d$ une base de D sur $\mathscr{R}_{\mathrm{L}}$ et A la matrice de φ dans cette base. D'après le théorème 2.6, il existe une matrice $\mathrm{M} \in \mathrm{GL}_d(\widetilde{\mathscr{R}})$ telle que $\mathrm{M}^{-1}\mathrm{A}\varphi(\mathrm{M})$ soit diagonale. Approximant M par une matrice à coefficients dans $\mathscr{R}_{\mathrm{L}}(\mathrm{E})$, où E est une extension finie de $k((\mathrm{T}))$, on construit un sous-φ-module Δ sur $\mathscr{E}^{\dagger}_{\mathrm{L}}(\mathrm{E})$ tel que $\mathscr{R}_{\mathrm{L}}(\mathrm{E}) \otimes_{\mathscr{E}^{\dagger}_{\mathrm{L}}(\mathrm{E})} \Delta = \mathscr{R}_{\mathrm{L}}(\mathrm{E}) \otimes_{\mathscr{R}_{\mathrm{L}}} \mathrm{D}$ et ayant les mêmes ensembles de $\mathscr{R}$-pentes et $\mathscr{E}$-pentes. C'est cette dernière condition qui demande le plus de travail ; en effet, si D est un φ-module de rang d sur $\mathscr{E}^{\dagger}_{\mathrm{L}}$, et si $r_1 \leqslant \cdots \leqslant r_d$ (resp. $s_1 \geqslant \cdots \geqslant s_d$) sont les $\mathscr{E}$-pentes (resp. les $\mathscr{R}$-pentes) de D, les seules relations que l'on ait en général entre ces ensembles sont l'égalité $s_1 + \cdots + s_d = r_1 + \cdots + r_d$ et les inégalité $s_1 + \cdots + s_i \geqslant r_1 + \cdots + r_i$, si $i < d$. Cette condition permet de montrer que la filtration croissante de Δ par les $\mathscr{E}$-pentes, qui n'est a priori définie que sur $\mathscr{E}_{\mathrm{L}}(\mathrm{E}) \otimes_{\mathscr{E}^{\dagger}_{\mathrm{L}}(\mathrm{E})} \Delta$, se descend en une filtration sur Δ. Pour terminer, on étend cette filtration à $\mathscr{R}_{\mathrm{L}}(\mathrm{E}) \otimes_{\mathscr{R}_{\mathrm{L}}} \mathrm{D}$ et un peu de descente galoisienne permet de conclure.

2.3.4. Application aux (φ, ∂)-modules. — Soit ∂ une dérivation sur $\mathscr{E}^{\dagger}_{\mathrm{L}}$. Tsuzuki [74] (voir aussi [15]) a démontré le résultat suivant.

THÉORÈME 2.9 ([74]). — *Si* D *est un (φ, ∂)-module isocline sur $\mathscr{E}^{\dagger}_{\mathrm{L}}$, il existe une extension finie séparable* E *de $k((\mathrm{T}))$ telle que $\mathscr{E}^{\dagger}_{\mathrm{L}}(\mathrm{E}) \otimes_{\mathscr{E}^{\dagger}_{\mathrm{L}}} \mathrm{D}$ possède une base $e_1, \dots, e_d$ sur $\mathscr{E}^{\dagger}_{\mathrm{L}}(\mathrm{E})$ vérifiant $\varphi(e_i) = e_i$ et $\partial e_i = 0$ si $1 \leqslant i \leqslant d$.*

Les deux conditions $\varphi(x) = x$ et $\partial x = 0$ ne sont pas loin d'être équivalentes : les L-espaces vectoriels engendrés par les solutions sont les mêmes dans les deux cas. Notons que, en rang 1, cet énoncé est équivalent à la conjecture de Crew car un élément inversible de $\mathscr{R}_{\mathrm{L}}$ appartient à $\mathscr{E}^{\dagger}_{\mathrm{L}}$ et que, par force, un φ-module de rang 1 est isocline. Ceci nous fournit donc une autre démonstration du théorème de Crew.

Le théorème 2.8 permet de déduire la conjecture de Crew de cet énoncé. En effet, une récurrence immédiate montre qu'il suffit de prouver que le $\mathscr{E}^{\dagger}$-module isocline D_1 est stable par ∂, ce qui suit, avec quelque effort, de ce que ∂ diminue les φ-pentes.

2.3.5. Compléments. — Le résultat de Tsuzuki peut se réinterpréter en termes galoisiens en anticipant un peu sur les résultats des §§ suivants (cf. remarque 5.2). La dérivation ∂ s'étend par continuité à $\mathscr{E}$ et à $\mathscr{E}(k((\mathrm{T}))^{\mathrm{sep}})$. D'autre part, les foncteurs

$$\mathrm{V} \mapsto \mathrm{D}(\mathrm{V}) = \Big(\mathscr{E}\big(k((\mathrm{T}))^{\mathrm{sep}}\big) \otimes_{\mathbf{Q}_p} \mathrm{V}\Big)^{\mathscr{G}_{k((\mathrm{T}))}} \quad \text{et} \quad \mathrm{D} \mapsto \mathrm{V}(\mathrm{D}) = \Big(\mathscr{E}\big(k((\mathrm{T}))^{\mathrm{sep}}\big) \otimes_{\mathscr{E}} \mathrm{D}\Big)^{\varphi=1}$$

sont inverses l'un de l'autre, et établissent une équivalence de catégories entre la catégorie des $\mathbf{Q}_p$-représentations de $\mathscr{G}_{k((\mathrm{T}))}$ et celle des φ-modules étales sur $\mathscr{E}$. On en déduit l'existence sur tout φ-module étale D sur $\mathscr{E}$ d'une unique connexion ∂_{D} au-dessus de ∂ vérifiant la relation $\partial_{\mathrm{D}} \circ \varphi = \frac{\partial(\varphi(\mathrm{T}))}{\varphi(\partial(\mathrm{T}))}\, \varphi \circ \partial_{\mathrm{D}}$: en effet, une telle connexion s'étend de manière unique à $\mathscr{E}(k((\mathrm{T}))^{\mathrm{sep}}) \otimes_{\mathscr{E}} \mathrm{D} = \mathscr{E}(k((\mathrm{T}))^{\mathrm{sep}}) \otimes_{\mathbf{Q}_p} \mathrm{V}(\mathrm{D})$ et le fait que φ est étale et vérifie la relation de commutation ci-dessus implique que ∂_{D} est

nul sur V(D). En résumé, *tout φ-module étale* D *sur $\mathscr{E}$ peut être muni d'une unique structure de (φ,∂)-module.*

COROLLAIRE 2.10 ([74]). — *Si* D *est un (φ,∂)-module étale sur $\mathscr{E}$ les conditions suivantes sont équivalentes* :

(i) *∂ est surconvergente* (il existe une base de D sur $\mathscr{E}$ dans laquelle la matrice de ∂ est à coefficients dans $\mathscr{E}^\dagger$) ;

(ii) (l'inertie de) $\mathscr{G}_{k((\mathrm{T}))}$ *agit à travers un quotient fini sur* V(D).

3. REPRÉSENTATIONS p-ADIQUES ET COHOMOLOGIE GALOISIENNE

3.1. Généralités

Dans toute cette partie, G est un groupe profini ; en particulier, G est compact. Un G-*anneau* est un anneau topologique muni d'une action continue de G respectant sa structure d'anneau (on demande que $(\sigma,x)\to\sigma(x)$ soit continue).

Si B est un G-anneau, une B-*représentation* de G est un B-module libre W de rang fini muni d'une action semi-linéaire continue de G.

Si on choisit une base $e_1,\dots,e_d$ de W sur B et que l'on note $\mathrm{U}_\sigma=(u_{i,j}^\sigma)$ la matrice des vecteurs $\sigma(e_1),\dots,\sigma(e_d)$ dans la base $e_1,\dots,e_d$ (i.e. $\sigma(e_j)=\sum_{i=1}^d u_{i,j}^\sigma e_i$), la semi-linéarité de l'action de G se traduit par la relation de cocycle $\mathrm{U}_{\sigma\tau}=\mathrm{U}_\sigma\sigma(\mathrm{U}_\tau)$ quels que soient $\sigma,\tau\in\mathrm{G}$. En particulier, en prenant $\tau=\sigma^{-1}$, on en déduit le fait que U_σ est inversible et que $\sigma\to\mathrm{U}_\sigma$ est un cocycle continu à valeurs dans $\mathrm{GL}_d(\mathrm{B})$. D'autre part, si on choisit une autre base $f_1,\dots,f_d$ de W sur B et que l'on note $\mathrm{M}=(m_{i,j})$ la matrice de passage et U'_σ la matrice de $\sigma(f_1),\dots,\sigma(f_d)$ dans la base $f_1,\dots,f_d$, on a $\mathrm{U}'_\sigma=\mathrm{M}^{-1}\mathrm{U}_\sigma\sigma(\mathrm{M})$, ce qui montre que les cocycles associés à deux bases différentes sont cohomologues. Ceci permet d'associer à toute B-représentation de G de rang d un élément de l'ensemble de cohomologie continue $\mathrm{H}^1(\mathrm{G},\mathrm{GL}_d(\mathrm{B}))$.

Si V est une $\mathbf{Q}_p$-représentation de G, alors $\mathrm{B}\otimes_{\mathbf{Q}_p}\mathrm{V}$, muni de l'action diagonale de G, est une B-représentation de G ; on dit que V est B-*admissible* si la classe de cohomologie de $\mathrm{B}\otimes_{\mathbf{Q}_p}\mathrm{V}$ dans $\mathrm{H}^1(\mathrm{G},\mathrm{GL}_d(\mathrm{B}))$ qui lui est associée est triviale ou, autrement dit, si $\mathrm{B}\otimes_{\mathbf{Q}_p}\mathrm{V}\cong\mathrm{B}^d$ en tant que G-module. On remarquera que le groupe de cohomologie continue $\mathrm{H}^1(\mathrm{G},\mathrm{GL}_d(\mathrm{B}))$ dépend fortement de la topologie que l'on a mise sur B mais que la B-admissibilité d'une $\mathbf{Q}_p$-représentation n'en dépend pas (tant que l'injection de $\mathbf{Q}_p$ dans B est continue).

Remarque 3.1. — Si V est une $\mathbf{Q}_p$-représentation B-admissible, $\mathbf{D}_\mathrm{B}(\mathrm{V})=(\mathrm{B}\otimes_{\mathbf{Q}_p}\mathrm{V})^\mathrm{G}$ est un B^G-module libre de rang $d=\dim_{\mathbf{Q}_p}\mathrm{V}$ et on dispose d'un isomorphisme naturel $\alpha_\mathrm{B}:\mathrm{B}\otimes_{\mathrm{B}^\mathrm{G}}\mathbf{D}_\mathrm{B}(\mathrm{V})\to\mathrm{B}\otimes_{\mathbf{Q}_p}\mathrm{V}$ commutant à l'action de G. En particulier, si

$B^{\text{structure}} = \mathbf{Q}_p$, on peut retrouver V en tant que $\mathbf{Q}_p$-représentation de G à partir de $\mathbf{D}_B(V)$ en prenant les éléments de $B \otimes_{B^G} \mathbf{D}_B(V)$ fixés par les structures.

3.2. Le théorème de Hilbert 90 et ses variantes

La discussion précédente montre que l'on a intérêt à étudier les ensembles $H^1(G, GL_d(B))$. En particulier, si cet ensemble est trivial, toutes les $\mathbf{Q}_p$-représentations de G sont B-admissibles et le module $\mathbf{D}_B(V)$ est un invariant non trivial de V pour tout V.

Pour ce faire, on dispose de deux outils classiques. D'une part la suite exacte « d'inflation-restriction » : si Λ est un G-anneau et si H est un sous-groupe fermé distingué de G, alors la suite d'ensembles pointés

$$1 \longrightarrow H^1(G/H, GL_d(\Lambda^H)) \xrightarrow{\inf} H^1(G, GL_d(\Lambda)) \xrightarrow{\text{res}} H^1(H, GL_d(\Lambda))$$

est exacte. D'autre part, du théorème de Hilbert 90 (et des techniques entrant dans sa démonstration : « séries de Poincaré », cf. [70] par exemple).

Proposition 3.2 (Hilbert 90). — *Si* L/F *est une extension galoisienne de groupe de Galois* G *et si on munit* L *de la topologie discrète, alors*

(i) $H^1(G, GL_d(L)) = \{1\}$ *si* $d \geqslant 1$;

(ii) $H^1(G, L) = \{0\}$.

Le résultat suivant se déduit du théorème de Hilbert 90 par approximations successives.

Proposition 3.3. — *Soit* Λ *un* G*-anneau et soit* $\pi \in \Lambda$ *tel que l'idéal engendré par* π *soit stable par* G *et* Λ *soit séparé et complet pour la topologie* π*-adique supposée plus forte que la topologie de* Λ. *Si* $H^1(G, GL_d(\Lambda/\pi\Lambda)) = \{1\}$ *quel que soit* $d \geqslant 1$ ($\Lambda/\pi\Lambda$ étant muni de la topologie quotient)*, alors*

(i) $H^1(G, GL_d(\Lambda)) = \{1\}$ *si* $d \geqslant 1$;

(ii) $H^1(G, \Lambda) = \{0\}$.

Proposition 3.4. — *Soient* Λ *un* G*-anneau principal et* $\omega \in \Lambda$ *tel que l'idéal* (ω) *soit stable par* G. *Si* $\Lambda' = \Lambda[\frac{1}{\omega}] = \bigcup_{n\in\mathbf{N}} \omega^{-n}\Lambda$ *est muni de la topologie de la limite inductive alors l'application naturelle* $H^1(G, GL_d(\Lambda)) \to H^1(G, GL_d(\Lambda'))$ *est surjective.*

Démonstration. — A un cocycle à valeurs dans $GL_d(\Lambda')$ correspond une Λ'-représentation de G et la compacité de G implique l'existence d'un sous-Λ-réseau stable.

3.3. La méthode de Sen

La méthode de Sen [69] permet, sous certaines *conditions de Tate-Sen*, de réduire beaucoup la complexité apparente des ensembles de cohomologie que l'on considère.

Soit G_0 un groupe profini muni d'un caractère continu $\chi : G_0 \to \mathbf{Z}_p^*$ dont l'image est ouverte. Si $g \in G_0$, on note $n(g)$ l'entier défini par $n(g) = v_p(\chi(g)^{p-1} - 1)$.

Soit $\widetilde{\Lambda}$ une $\mathbf{Z}_p$-algèbre munie de $v : \widetilde{\Lambda} \to \mathbf{R} \cup \{+\infty\}$ vérifiant les conditions :

(i) $v(x) = +\infty \Leftrightarrow x = 0$;

(ii) $v(xy) \geqslant v(x) + v(y)$;

(iii) $v(x + y) \geqslant \inf(v(x), v(y))$;

(iv) $v(p) > 0$ et $v(px) = v(p) + v(x)$ si $x \in \widetilde{\Lambda}$.

La condition (iii) permet d'utiliser v pour munir $\widetilde{\Lambda}$ d'une topologie et la condition (i) montre que cette topologie est séparée. On suppose de plus que $\widetilde{\Lambda}$ est complet pour cette topologie et que $\widetilde{\Lambda}$ est muni d'une action continue de G_0 telle que l'on ait $v(g(x)) = v(x)$ si $g \in G_0$ et $x \in \widetilde{\Lambda}$.

Considérons les propriétés suivantes :

(TS1) Il existe $c_1 > 0$ tel que, quels que soient les sous-groupes ouverts $H_1 \subset H_2$ du noyau H_0 de χ, il existe $\alpha \in \widetilde{\Lambda}^{H_1}$ vérifiant $v(\alpha) > -c_1$ et $\sum_{\tau \in H_2/H_1} \tau(\alpha) = 1$.

(TS2) Il existe $c_2 > 0$ et, pour tout sous-groupe ouvert H de H_0, un entier $n(H) \in \mathbf{N}$, une suite croissante $(\Lambda_{H,n})_{n \in \mathbf{N}}$ de sous-$\mathbf{Z}_p$-algèbres fermées de $\widetilde{\Lambda}^H$ et, pour $n \geqslant n(H)$, une application $\mathbf{Z}_p$-linéaire $R_{H,n} : \widetilde{\Lambda}^H \to \Lambda_{H,n}$ vérifiant :

a) si $H_1 \subset H_2$, alors $\Lambda_{H_2,n} = (\Lambda_{H_1,n})^{H_2}$ et $R_{H_1,n} = R_{H_2,n}$ sur $\widetilde{\Lambda}^{H_2}$;

b) $g(\Lambda_{H,n}) = \Lambda_{gHg^{-1},n}$ et $g(R_{H,n}(x)) = R_{gHg^{-1},n}(gx)$ si $g \in G_0$;

c) $R_{H,n}$ est $\Lambda_{H,n}$-linéaire et $R_{H,n}(x) = x$ si $x \in \Lambda_{H,n}$;

d) si $n \geqslant n(H)$ et si $x \in \widetilde{\Lambda}^H$, alors $v(R_{H,n}(x)) \geqslant v(x) - c_2$;

e) si $x \in \widetilde{\Lambda}^H$, alors $\lim_{n \to +\infty} R_{H,n}(x) = x$.

(TS3) Il existe $c_3 > 0$ et, pour tout sous-groupe ouvert G de G_0, un entier $n(G) \geqslant n(H)$, où $H = G \cap H_0$, tel que, si $n \geqslant n(G)$, si $\gamma \in G/H$ vérifie $n(\gamma) \leqslant n$, alors $\gamma - 1$ est inversible sur $X_{H,n} = (1 - R_{H,n})(\widetilde{\Lambda}^H)$ et on a $v((\gamma - 1)^{-1}(x)) \geqslant v(x) - c_3$ si $x \in X_{H,n}$.

Les propriétés (TS1) et (TS3) sont les plus délicates à vérifier. Dans les situations arithmétiques, on peut en général prendre $c_1 > 0$ quelconque, et la théorie de la ramification supérieure est un ingrédient essentiel (on est dans le cadre de la théorie des extensions « presque étales » dans la terminologie de Faltings). Les applications $R_{H,n}$ sont souvent appelées des *traces de Tate normalisées*.

PROPOSITION 3.5. — *Soit $\widetilde{\Lambda}$ vérifiant les conditions de Tate-Sen* (TS1), (TS2) *et* (TS3) *et soit $\sigma \mapsto U_\sigma$ un cocycle continu sur G_0 à valeurs dans $\mathbf{GL}_d(\widetilde{\Lambda})$. Si G est un sous-groupe ouvert distingué de G_0 tel que $v(U_\sigma - 1) > c_1 + 2c_2 + 2c_3$ quel que soit $\sigma \in G$ et si $H = G \cap H_0$, alors il existe $M \in \mathbf{GL}_d(\widetilde{\Lambda})$ vérifiant $v(M-1) > c_2 + c_3$ tel que le cocycle $\sigma \mapsto V_\sigma = M^{-1} U_\sigma \sigma(M)$ soit trivial sur H et à valeurs dans $\mathbf{GL}_d(\Lambda_{H,n(G)})$.*

On peut retraduire ce résultat en l'existence d'invariants attachés aux $\mathbf{Q}_p$-représentations de G_0. Si V est une telle représentation de dimension d, la compacité de G_0 assure l'existence d'un réseau T de V stable par G_0, et la continuité de l'action de G_0 se traduit par l'existence, pour tout $k \in \mathbf{N}$, d'un sous-groupe ouvert distingué de G_0 agissant trivialement sur T/p^kT. On a alors le résultat suivant :

PROPOSITION 3.6. — *Si k est un entier tel que $v(p^k) > c_1 + 2c_2 + 2c_3$ et si G est un sous-groupe distingué de G_0 agissant trivialement sur T/p^kT, soit $H = G \cap H_0$ et soit $n \geqslant n(H)$. Alors $\widetilde{\Lambda} \otimes_{\mathbf{Z}_p} T$ contient un unique sous-$\Lambda_{H,n}$-module $D_{H,n}(T)$ libre de rang d vérifiant les propriétés suivantes :*

(i) $D_{H,n}(T)$ *est fixe par* H *et stable par* G_0 *;*

(ii) *l'application naturelle* $\widetilde{\Lambda} \otimes_{\Lambda_{H,n}} D_{H,n}(T) \to \widetilde{\Lambda} \otimes_{\mathbf{Z}_p} T$ *est un isomorphisme ;*

(iii) $D_{H,n}(T)$ *possède une base sur $\Lambda_{H,n}$ dans laquelle la matrice U_γ de $\gamma \in G/H$, vérifie $v(U_\gamma - 1) > c_3$.*

Remarque 3.7. — Le groupe G_0 agit sur $D_{H,n}(T)$ à travers G_0/H. On peut encore diminuer le quotient agissant (sans tuer l'invariant que l'on vient de construire) en prenant les points fixes de $D_{H,n}(T)$ sous H_0 (qui agit à travers le groupe fini H_0/H), mais le $\Lambda_{H_0,n}$-module ainsi obtenu n'est plus forcément libre (cela dépend beaucoup des propriétés algébriques des $\Lambda_{H,n}$).

4. LES ANNEAUX DE FONTAINE

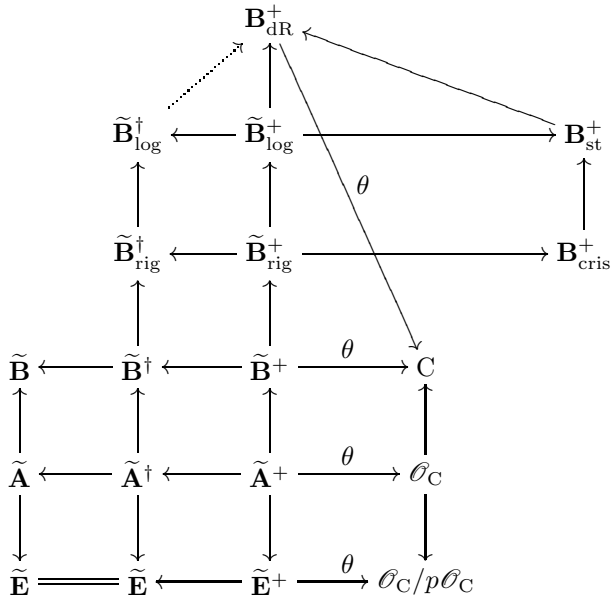
4.1. Cartographie

4.1.1. Notations. — Soit k_F un corps parfait de caractéristique p, $\mathscr{O}_F = W(F)$ l'anneau des vecteurs de Witt à coefficients dans k_F et $F = \mathscr{O}_F[\frac{1}{p}]$ ce qui fait de F un corps complet pour la valuation p-adique v_p, de corps résiduel k_F.

Soit $\overline{F}$ une clôture algébrique de F ; la valuation v_p s'étend de manière unique à $\overline{F}$. Choisissons un système $\varepsilon = (1, \varepsilon^{(1)}, \dots, \varepsilon^{(n)}, \dots)$, où les $\varepsilon^{(n)} \in \overline{F}$ vérifient $\varepsilon^{(1)} \neq 1$ et $(\varepsilon^{(n+1)})^p = \varepsilon^{(n)}$; en particulier, $\varepsilon^{(n)}$ est une racine primitive p^n-ième de l'unité.

Si $K \subset \overline{F}$ est une extension finie totalement ramifiée de F, soient $K_n = K(\varepsilon^{(n)})$ si $n \in \mathbf{N}$ et $K_\infty = \cup_{n \in \mathbf{N}} K_n$ l'extension cyclotomique de K ; on note F′ l'extension maximale non ramifiée de F contenue dans K_∞. Soient $\mathscr{G}_K = \mathrm{Gal}(\overline{F}/K)$, $\mathscr{H}_K = \mathrm{Gal}(\overline{F}/K_\infty)$ et $\Gamma_K = \mathscr{G}_K/\mathscr{H}_K = \mathrm{Gal}(K_\infty/K)$. Soit aussi $\chi : \mathscr{G}_K \to \mathbf{Z}_p^*$ le caractère cyclotomique ; il induit un isomorphisme de Γ_K sur un sous-groupe ouvert de $\mathbf{Z}_p^*$.

4.1.2. Navigation dans le monde des anneaux. — La plupart des anneaux construits par Fontaine s'obtiennent à partir d'un anneau[8] $\widetilde{\mathbf{A}}^+$ en localisant et en complétant. Comme l'anneau $\widetilde{\mathbf{A}}^+$ est de dimension 2, les anneaux que l'on obtient de cette manière forment naturellement un tableau en deux dimensions (et même en quatre si on rajoute les actions de Frobenius et Galois) ; il est donc quasi-impossible de les présenter de manière satisfaisante dans un texte qui, par nature, est de dimension 1. Le lecteur aura intérêt à se reporter au tableau ci-dessous pour s'orienter.



(i) Plus on monte et plus p est inversible : la ligne du bas vit en caractéristique p ; c'est la réduction modulo p de la ligne du dessus et on passe de l'avant-dernière ligne à la précédente en rendant p-inversible. Plus on va vers la gauche (en ne tenant pas compte de la dernière colonne qui n'est là que pour mémoire), et plus $\pi = [\varepsilon] - 1$ est inversible : π est nul dans la quatrième colonne, et l'application $\theta : \mathbf{B}_{\mathrm{dR}}^+ \to \mathrm{C}$ est la réduction modulo π.

(ii) Toutes les flèches de ce diagramme sont injectives à l'exception de la réduction modulo p et de θ.

(iii) $\mathscr{G}_{\mathrm{F}}$ opère continûment sur tout le diagramme.

(iv) φ opère sur les trois premières colonnes de manière bijective (et sur la dernière, mais pas bijectivement), et son action commute à celle de $\mathscr{G}_{\mathrm{F}}$.

(v) Les anneaux des trois premières colonnes (sauf $\widetilde{\mathbf{B}}_{\mathrm{log}}^+$) ont un petit frère sans $\widetilde{\ }$ sur lequel l'action de φ n'est qu'injective. On passe d'un anneau sans $\widetilde{\ }$ à un anneau

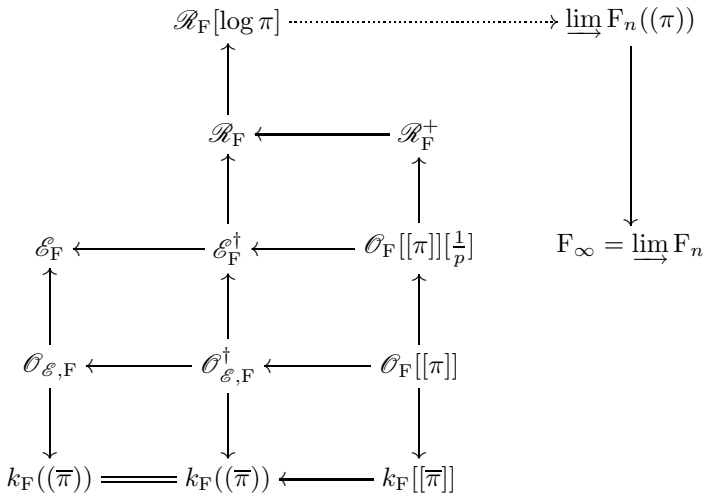
[8] L'anneau $\widetilde{\mathbf{A}}^+$ en question est souvent noté W(R) et quelquefois $\mathrm{A}_{\mathrm{inf}}$ dans la littérature.

avec en rendant φ inversible et en complétant. Par exemple, $\widetilde{\mathbf{E}}$ est le complété de la clôture radicielle de $\mathbf{E}$.

(vi) La flèche en pointillés de $\widetilde{\mathbf{B}}^{\dagger}_{\log}$ vers $\mathbf{B}^{+}_{\mathrm{dR}}$ est une famille de morphismes $(\iota_n)_{n\in\mathbf{N}}$, où ι_n est défini sur le sous-anneau $\widetilde{\mathbf{B}}^{]0,p^{-n}]}[\log[\tilde{p}]]$ de $\widetilde{\mathbf{B}}^{\dagger}_{\log}$ des éléments « convergeant sur la couronne $0 < v_p(\mathrm{T}) \leqslant p^{-n}$ »; la réunion de ces anneaux est $\widetilde{\mathbf{B}}^{\dagger}_{\log}$ et on a $\iota_n(x) = \iota_{n-1}(\varphi^{-1}(x))$ si tout est défini.

(vii) La théorie des (φ,Γ)-modules utilise les anneaux $\mathbf{E}$, $\mathbf{A}$ et $\mathbf{B}$ de la première colonne, alors que la théorie de Hodge p-adique utilise les anneaux $\widetilde{\mathbf{B}}^{+}_{\mathrm{rig}}$, $\widetilde{\mathbf{B}}^{+}_{\log}$, $\mathbf{B}^{+}_{\mathrm{dR}}$ et C. Le lecteur remarquera que la plupart des flèches allant de la théorie des (φ,Γ)-modules à la théorie de Hodge p-adique sont dans le mauvais sens; il faut donc à chaque fois démontrer que l'on peut effectivement remonter.

(viii) Si K est une extension finie de F, un K en indice indique le sous-anneau fixé par $\mathscr{H}_{\mathrm{K}}$. Les anneaux sans $\widetilde{\ }$ avec un K en indice sont des anneaux de séries en une variable π_{K}. Dans le cas K = F, on a $\pi_{\mathrm{F}} = \pi = [\varepsilon] - 1$; les actions de $\mathscr{G}_{\mathrm{F}}$ et φ sont données par les formules $g(\pi) = (1+\pi)^{\chi(g)} - 1$ et $\varphi(\pi) = (1+\pi)^p - 1$, et on obtient le diagramme suivant :



(ix) $[\varepsilon] = 1+\pi$ est un analogue p-adique de $e^{2i\pi}$, et l'application

$$\mu \longmapsto \int_{\mathbf{Z}_p} (1+\pi)^x \mu(x)$$

qui, à une distribution (resp. une mesure) μ sur $\mathbf{Z}_p$ à valeurs dans F (resp. $\mathscr{O}_{\mathrm{F}}$) associe sa transformée de Fourier, induit un isomorphisme de l'algèbre des distributions (resp. des mesures) sur $\mathbf{B}^{+}_{\mathrm{rig,F}} \cong \mathscr{R}^{+}_{\mathrm{F}}$ (resp. sur $\mathbf{A}^{+}_{\mathrm{F}} = \mathscr{O}_{\mathrm{F}}[[\pi]]$). (On rappelle qu'une mesure est une application linéaire continue sur l'espace des fonctions continues et

qu'une distribution est une application linéaire continue sur l'espace des fonctions localement analytiques.) Cette remarque n'est pas utilisée dans ce texte, mais on pourra consulter [26] pour un survol des questions dans lesquelles elle joue un rôle.

4.2. C et ses sous-corps

4.2.1. Le corps C *et l'action de* $\mathscr{G}_{\mathrm{F}}$. — Soit C le complété de $\overline{\mathrm{F}}$ pour la valuation v_p. C'est un corps algébriquement clos muni d'une action continue de $\mathscr{G}_{\mathrm{F}}$. Si L est une extension finie de K, alors $\mathrm{C}^{\mathscr{H}_{\mathrm{L}}}$ est le complété $\widehat{\mathrm{L}}_\infty$ de L_∞ pour v_p. Par ailleurs, si $n \in \mathbf{N}$ et si $x \in \mathrm{L}_\infty$, alors $\frac{1}{[\mathrm{L}_{n+k}:\mathrm{L}_n]}\mathrm{Tr}_{\mathrm{L}_{n+k}/\mathrm{L}_n}(x)$ ne dépend pas du choix de l'entier k tel que $x \in \mathrm{L}_{n+k}$. L'application de L_∞ dans L_n ainsi définie se prolonge par continuité en une application $\mathrm{R}_{\mathrm{L},n} : \widehat{\mathrm{L}}_\infty \to \mathrm{L}_n$.

Proposition 4.1 ([71]). — *L'anneau* $\widetilde{\Lambda} = \mathrm{C}$ *vérifie les conditions* (TS1), (TS2) *et* (TS3)*, avec* $v = v_p$, $\widetilde{\Lambda}^{\mathscr{H}_{\mathrm{L}}} = \widehat{\mathrm{L}}_\infty$, $\Lambda_{\mathscr{H}_{\mathrm{L}},n} = \mathrm{L}_n$, $\mathrm{R}_{\mathscr{H}_{\mathrm{L}},n} = \mathrm{R}_{\mathrm{L},n}$, *les constantes* $c_1 > 0$, $c_2 > 0$ *et* $c_3 > \frac{1}{p-1}$ *pouvant être choisies arbitrairement.*

La démonstration de Tate repose sur la théorie de la ramification qui permet en particulier de montrer que $v_p(\mathfrak{d}_{\mathrm{K}_n/\mathrm{F}_n})$ tend vers 0 quand n tend vers $+\infty$ (si $\mathrm{K} \subset \mathrm{L}$ sont deux extensions finies de F, on note $\mathfrak{d}_{\mathrm{L}/\mathrm{K}}$ la différente de l'extension K/L) ; dans la terminologie de Faltings cela se traduit par : « l'extension $\mathrm{K}_\infty/\mathrm{F}_\infty$ est presque étale ».

On déduit de la proposition 4.1 [en fait on a juste besoin de $\mathrm{C}^{\mathscr{H}_{\mathrm{K}}} = \widehat{\mathrm{K}}_\infty$ et de la propriété (TS2)] et de ce que Γ_{K} agit sur K_n à travers un quotient fini, le fait que C ne contient ni d'analogue p-adique de $\log 2i\pi$ ni d'analogue p-adique de $2i\pi$. De manière précise, on a

Proposition 4.2 ([71]). — (i) C *ne contient aucun élément* y *tel que l'on ait* $g(y) = y + \log\chi(g)$ *quel que soit* $g \in \mathscr{G}_{\mathrm{K}}$.

(ii) *Si* $k \in \mathbf{Z} - \{0\}$ *et si* $x \in \mathrm{C}$ *vérifie* $g(x) = \chi(g)^k x$ *pour tout* $g \in \mathscr{G}_{\mathrm{K}}$, *alors* $x = 0$.

4.2.2. L'anneau $\mathbf{B}_{\mathrm{HT}}$. — Soit $\mathbf{B}_{\mathrm{HT}} = \mathrm{C}[t, t^{-1}]$. On munit $\mathbf{B}_{\mathrm{HT}}$ d'une action continue de $\mathscr{G}_{\mathrm{F}}$ en faisant agir $g \in \mathscr{G}_{\mathrm{F}}$ sur t par la formule $g(t) = \chi(g)t$. On a alors $\mathbf{B}_{\mathrm{HT}}^{\mathscr{G}_{\mathrm{K}}} = \mathrm{K}$ d'après la proposition 4.2 et le théorème d'Ax-Sen-Tate. On munit $\mathbf{B}_{\mathrm{HT}}$ d'une graduation indexée par les entiers en posant $\mathrm{Gr}^i\mathbf{B}_{\mathrm{HT}} = \mathrm{C} \cdot t^i$ si $i \in \mathbf{Z}$. Cette graduation est stable sous l'action de $\mathscr{G}_{\mathrm{F}}$.

4.3. Les anneaux de caractéristique p

Soit $\mathfrak{a} = \{x,\ v_p(x) \geqslant 1/p\}$ et soit

$$\widetilde{\mathbf{E}}^+ = \{(x_n)_{n\in\mathbf{N}}, x_n \in \mathscr{O}_{\mathrm{C}}/\mathfrak{a} \text{ et } x_{n+1}^p = x_n \text{ si } n \in \mathbf{N}\}.$$

L'anneau $\mathscr{O}_{\mathrm{C}}/\mathfrak{a}$ étant de caractéristique p, l'application $x \mapsto x^p$ en est un morphisme et $\widetilde{\mathbf{E}}^+$ est un anneau[9] de caractéristique p sur lequel $\mathscr{G}_{\mathrm{F}}$ agit naturellement (composante par composante). D'autre part, si $(x_n)_{n\in\mathbf{N}} \in \widetilde{\mathbf{E}}^+$ et si $\hat{x}_n$ est un relèvement quelconque de x_n dans $\mathscr{O}_{\mathrm{C}}$, la suite de terme général $\hat{x}_{n+k}^{p^k}$ converge dans $\mathscr{O}_{\mathrm{C}}$ vers une limite $x^{(n)}$ qui ne dépend pas du choix des $\hat{x}_n$. Ceci permet de décrire $\widetilde{\mathbf{E}}^+$ comme l'ensemble des suites $x = (x^{(0)}, \dots, x^{(n)}, \dots)$ d'éléments de $\mathscr{O}_{\mathrm{C}}$ vérifiant $(x^{(n+1)})^p = x^{(n)}$. Soit $v_{\mathrm{E}} : \widetilde{\mathbf{E}} \to \mathbf{R}$ l'application définie par $v_{\mathrm{E}}(x) = v_p(x^{(0)})$.

On peut voir $\varepsilon = (1, \varepsilon^{(1)}, \dots, \varepsilon^{(n)}, \dots)$ comme un élément de $\widetilde{\mathbf{E}}^+$ et, si on pose $\overline{\pi} = \varepsilon - 1$, on a $v_{\mathbf{E}}(\overline{\pi}) = \frac{p}{p-1}$.

Si K est une extension finie de F, soient

$$\widetilde{\mathbf{E}}_{\mathrm{K}}^+ = \{(x_n)_{n\in\mathbf{N}} \in \widetilde{\mathbf{E}}^+,\ x_n \in \mathscr{O}_{\mathrm{K}_\infty}/\mathfrak{a} \text{ si } n \in \mathbf{N}\},$$
$$\mathbf{E}_{\mathrm{K}}^+ = \{(x_n)_{n\in\mathbf{N}} \in \widetilde{\mathbf{E}}^+,\ x_n \in \mathscr{O}_{\mathrm{K}_n}/\mathfrak{a} \text{ si } n \text{ est assez grand}\}.$$

$\mathbf{E}_{\mathrm{K}}^+$ contient ε et $\overline{\pi}$, ce qui nous permet de poser $\widetilde{\mathbf{E}} = \widetilde{\mathbf{E}}^+[\overline{\pi}^{-1}]$, $\widetilde{\mathbf{E}}_{\mathrm{K}} = \widetilde{\mathbf{E}}_{\mathrm{K}}^+[\overline{\pi}^{-1}]$ et $\mathbf{E}_{\mathrm{K}} = \mathbf{E}_{\mathrm{K}}^+[\overline{\pi}^{-1}]$ si K est une extension finie de F.

THÉORÈME 4.3 ([44, 79]). — (i) *$\widetilde{\mathbf{E}}$ est un corps dont $v_{\mathbf{E}}$ est une valuation pour laquelle il est complet et dont le corps résiduel est $\overline{k}_{\mathrm{F}}$. De plus l'action naturelle de $\mathscr{G}_{\mathrm{F}}$ sur $\widetilde{\mathbf{E}}$ est continue.*

(ii) *$\mathbf{E}_{\mathrm{F}} = k_{\mathrm{F}}((\overline{\pi}))$ et, plus généralement, si* K *est une extension finie de* F*, alors $\mathbf{E}_{\mathrm{K}}$ muni de $v_{\mathbf{E}}$ est un corps complet pour une valuation discrète de corps résiduel $k_{\mathrm{F}'}$.*

(iii) *Le sous-corps $\mathbf{E} = \cup_{\mathrm{K}\subset\overline{\mathrm{F}}}\mathbf{E}_{\mathrm{K}}$ de $\widetilde{\mathbf{E}}$ est une clôture séparable de $\mathbf{E}_{\mathrm{F}}$ stable par $\mathscr{G}_{\mathrm{F}}$ et, si* K *est une extension finie de* F*, alors $\mathrm{Gal}(\mathbf{E}/\mathbf{E}_{\mathrm{K}}) = \mathscr{H}_{\mathrm{K}}$; en particulier, $\mathscr{H}_{\mathrm{K}}$ agit continûment sur $\mathbf{E}$ muni de la topologie discrète.*

(iv) *$\widetilde{\mathbf{E}}$ (resp. $\widetilde{\mathbf{E}}_{\mathrm{K}}$) est le complété de $\mathbf{E}$ (resp. de la clôture radicielle de $\mathbf{E}_{\mathrm{K}}$) pour la valuation $v_{\mathbf{E}}$; en particulier, $\widetilde{\mathbf{E}}$ est algébriquement clos et $\widetilde{\mathbf{E}}_{\mathrm{K}} = \widetilde{\mathbf{E}}^{\mathscr{H}_{\mathrm{K}}}$.*

Remarque 4.4. — (i) Ce théorème permet de relier la théorie de Galois des corps locaux de caractéristique 0 à celle des corps locaux de caractéristique p ; c'est le point de départ de beaucoup des constructions de Fontaine.

(ii) Le corps $\mathbf{E}_{\mathrm{K}}$ peut naturellement être mis en bijection avec la limite projective des K_n relativement aux applications $\mathrm{N}_{\mathrm{K}_{n+1}/\mathrm{K}_n}$; c'est ce qui lui vaut l'appellation de « corps des normes ».

(iii) La théorie du corps des normes a été développée par Fontaine et Wintenberger [44, 79], dans un cadre beaucoup plus général que celui de l'extension cyclotomique (celui des extensions « arithmétiquement profinies »). En particulier, le cas où on remplace l'extension cyclotomique par l'extension obtenue en rajoutant un système compatible de racines p^n-ièmes de p a l'air prometteur (cf. [10]).

[9] L'anneau $\widetilde{\mathbf{E}}^+$ et le corps $\widetilde{\mathbf{E}}$ sont très souvent notés R et F*r* R respectivement.

(iv) Les (ii) et (iii) du théorème sont loin d'être évidents, mais les ingrédients utilisés pour leur démonstration se résument à

— le lemme de Hensel ;

— l'extension $\mathrm{K}_\infty/\mathrm{F}_\infty$ est presque étale ;

— si $x \in \mathscr{O}_{\mathrm{F}_{n+1}}$ et $\sigma \in \mathscr{G}_{\mathrm{F}_n}$, alors $v_p(\sigma(x)-x) \geqslant \frac{1}{p-1}$;

les deux derniers points permettant, en particulier, de montrer que l'on a $x^p \equiv \mathrm{N}_{\mathrm{K}_{n+1}/\mathrm{K}_n}(x)$ mod. $\mathfrak{a}$, si n est assez grand et $x \in \mathscr{O}_{\mathrm{K}_{n+1}}$.

4.4. De la caractéristique p à la caractéristique 0

4.4.1. Notations. — On peut utiliser les foncteurs $\mathscr{E}$, $\mathscr{E}^\dagger$, $\mathscr{R}$, ... du n° 2.2 pour remonter les anneaux précédents en caractéristique 0. Pour cela, on doit choisir une uniformisante de $\mathbf{E}_\mathrm{F}$ et un Frobenius sur $\mathscr{E}(\mathrm{T})$ et un choix judicieux, pour des raisons qui apparaîtront bientôt, consiste à prendre $\overline{\pi}$ comme uniformisante de $\mathbf{E}_\mathrm{F}$ et $\mathrm{T} \mapsto (1+\mathrm{T})^p - 1$ pour φ. La grande nouveauté par rapport au n° 2.2 est que tous ces anneaux se retrouvent munis d'une action de $\mathscr{G}_\mathrm{F}$ plutôt que d'une action de $\mathscr{G}_{k((\mathrm{T}))}$. On introduit donc de nouvelles notations mettant l'accent sur les extensions de F plutôt que sur celles de $k((\mathrm{T}))$. On pose

$$\widetilde{\mathbf{B}} = \mathscr{E}(\widetilde{\mathbf{E}}), \qquad \widetilde{\mathbf{B}}_\mathrm{K} = \mathscr{E}(\widetilde{\mathbf{E}}_\mathrm{K}), \qquad \mathbf{B} = \mathscr{E}(\mathbf{E}), \qquad \mathbf{B}_\mathrm{K} = \mathscr{E}(\mathbf{E}_\mathrm{K})$$

$$\widetilde{\mathbf{B}}^\dagger = \mathscr{E}^\dagger(\widetilde{\mathbf{E}}), \qquad \widetilde{\mathbf{B}}^\dagger_\mathrm{K} = \mathscr{E}^\dagger(\widetilde{\mathbf{E}}_\mathrm{K}), \qquad \mathbf{B}^\dagger = \mathscr{E}^\dagger(\mathbf{E}), \qquad \mathbf{B}^\dagger_\mathrm{K} = \mathscr{E}^\dagger(\mathbf{E}_\mathrm{K})$$

$$\widetilde{\mathbf{B}}^\dagger_\mathrm{rig} = \mathscr{R}(\widetilde{\mathbf{E}}), \qquad \widetilde{\mathbf{B}}^\dagger_{\mathrm{rig},\mathrm{K}} = \mathscr{R}(\widetilde{\mathbf{E}}_\mathrm{K}), \qquad \mathbf{B}^\dagger_{\mathrm{rig},\mathrm{K}} = \mathscr{R}(\mathbf{E}_\mathrm{K})$$

Les anneaux ci-dessus et $\widetilde{\mathbf{A}}^+ = \mathscr{O}^+_{\mathscr{E}}(\widetilde{\mathbf{E}}) = \mathrm{W}(\widetilde{\mathbf{E}}^+)$, $\widetilde{\mathbf{B}}^+ = \mathscr{E}^+(\widetilde{\mathbf{E}})$, $\widetilde{\mathbf{B}}^+_\mathrm{rig} = \mathscr{R}^+(\widetilde{\mathbf{E}})$ sont plus ou moins suffisants pour les énoncés, mais pour travailler, on a aussi besoin des anneaux intermédiaires

$$\widetilde{\mathbf{A}} = \mathscr{O}_\mathscr{E}(\widetilde{\mathbf{E}}) = \mathrm{W}(\widetilde{\mathbf{E}}), \qquad \widetilde{\mathbf{A}}_\mathrm{K} = \mathscr{O}_\mathscr{E}(\widetilde{\mathbf{E}}_\mathrm{K}), \qquad \mathbf{A} = \mathscr{O}_\mathscr{E}(\mathbf{E}), \qquad \mathbf{A}_\mathrm{K} = \mathscr{O}_\mathscr{E}(\mathbf{E}_\mathrm{K})$$

$$\widetilde{\mathbf{A}}^{(0,r]} = \mathscr{O}^{(0,r]}_\mathscr{E}(\widetilde{\mathbf{E}}), \qquad \widetilde{\mathbf{A}}^{(0,r]}_\mathrm{K} = \mathscr{O}^{(0,r]}_\mathscr{E}(\widetilde{\mathbf{E}}_\mathrm{K}), \qquad \mathbf{A}^{(0,r]} = \mathscr{O}^{(0,r]}_\mathscr{E}(\mathbf{E}), \qquad \mathbf{A}^{(0,r]}_\mathrm{K} = \mathscr{O}^{(0,r]}_\mathscr{E}(\mathbf{E}_\mathrm{K})$$

$$\widetilde{\mathbf{B}}^{(0,r]} = \mathscr{E}^{(0,r]}(\widetilde{\mathbf{E}}), \qquad \widetilde{\mathbf{B}}^{(0,r]}_\mathrm{K} = \mathscr{E}^{(0,r]}(\widetilde{\mathbf{E}}_\mathrm{K}), \qquad \mathbf{B}^{(0,r]} = \mathscr{E}^{(0,r]}(\mathbf{E}), \qquad \mathbf{B}^{(0,r]}_\mathrm{K} = \mathscr{E}^{(0,r]}(\mathbf{E}_\mathrm{K})$$

$$\widetilde{\mathbf{B}}^{]0,r]} = \mathscr{E}^{]0,r]}(\widetilde{\mathbf{E}}), \qquad \widetilde{\mathbf{B}}^{]0,r]}_\mathrm{K} = \mathscr{E}^{]0,r]}(\widetilde{\mathbf{E}}_\mathrm{K}), \qquad \mathbf{B}^{]0,r]}_\mathrm{K} = \mathscr{E}^{]0,r]}(\mathbf{E}_\mathrm{K})$$

Notons que les anneaux $\mathbf{B}^\dagger_\mathrm{rig}$ et $\mathbf{B}^{]0,r]}$ sont absents des listes ci-dessus (comme $\mathbf{E}$ est dense dans $\widetilde{\mathbf{E}}$, on a $\mathscr{R}(\mathbf{E}) = \mathscr{R}(\widetilde{\mathbf{E}})$ et $\mathbf{B}^{]0,r]}(\mathbf{E}) = \mathbf{B}^{]0,r]}(\widetilde{\mathbf{E}})$ ce qui nous empêche de prendre la définition évidente). On les définit comme étant les sous-anneaux $\mathbf{B}^\dagger_{\mathrm{rig},\mathrm{K}} \otimes_{\mathbf{B}^\dagger_\mathrm{K}} \mathbf{B}^\dagger$ (ceci ne dépend pas du choix de K) et $\mathbf{B}^{]0,r]}_\mathrm{F} \otimes_{\mathbf{B}^{(0,r]}_\mathrm{F}} \mathbf{B}^{(0,r]}$ de $\widetilde{\mathbf{B}}^\dagger_\mathrm{rig}$ et $\widetilde{\mathbf{B}}^{]0,r]}$ respectivement.

On peut aussi décrire les anneaux ci-dessus de manière plus algébrique (cette description étant d'ailleurs fort utile [6] pour étudier leur lien avec $\mathbf{B}_{\mathrm{dR}}^{+}$). Par exemple, si $\widetilde{\mathbf{A}}^{+}\{\mathrm{X}\}$ désigne l'ensemble des séries $\sum_{k=0}^{+\infty} a_k \mathrm{X}^k$, où $a_k \to 0$ quand $k \to +\infty$, et si $\omega \in \widetilde{\mathbf{A}}^{+}$ n'est pas une unité de $\widetilde{\mathbf{A}}^{+}$ (i.e. si $v_{\mathbf{E}}(\overline{\omega}) > 0$), alors :

$$\widetilde{\mathbf{B}}^{\dagger} = \bigcup_{k\geqslant 1}\Big(\widetilde{\mathbf{A}}^{+}\Big\{\frac{p}{\omega^k}\Big\}\Big[\frac{1}{p}\Big]\Big)$$
$$\widetilde{\mathbf{B}}_{\mathrm{rig}}^{+} = \bigcap_{n\geqslant 1}\Big(\widetilde{\mathbf{A}}^{+}\Big\{\frac{\omega^n}{p}\Big\}\Big[\frac{1}{p}\Big]\Big)$$
$$\widetilde{\mathbf{B}}_{\mathrm{rig}}^{\dagger} = \bigcup_{k\geqslant 1}\Big(\bigcap_{n\geqslant k}\Big(\widetilde{\mathbf{A}}^{+}\Big\{\frac{p}{\omega^k},\frac{\omega^n}{p}\Big\}\Big[\frac{1}{p}\Big]\Big)\Big)$$

4.4.2. Action de $\mathscr{G}_{\mathrm{F}}$ et de φ. — L'action de $\mathscr{G}_{\mathrm{F}}$ sur $\widetilde{\mathbf{E}}$ se prolonge en une action continue (pour la topologie faible) sur $\widetilde{\mathbf{A}}$ et $\widetilde{\mathbf{B}}$ qui commute à celle du morphisme de Frobenius φ. De manière explicite, on a

$$\varphi\Big(\sum_{k\gg -\infty}^{+\infty} p^k[x_k]\Big) = \sum_{k\gg -\infty}^{+\infty} p^k[x_k^p] \quad \text{et} \quad \sigma\Big(\sum_{k\gg -\infty}^{+\infty} p^k[x_k]\Big) = \sum_{k\gg -\infty}^{+\infty} p^k[\sigma(x_k)] \text{ si } \sigma\in\mathscr{G}_{\mathrm{F}}.$$

Ces actions s'étendent par continuité aux anneaux $\widetilde{\mathbf{B}}^{\dagger}$, $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{\dagger}$, $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}$, $\widetilde{\mathbf{B}}^{(0,r]}$ et $\widetilde{\mathbf{B}}^{]0,r]}$. L'action de φ est bijective sur $\widetilde{\mathbf{B}}^{\dagger}$, $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{\dagger}$ et $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}$ et induit un isomorphisme de $\widetilde{\mathbf{B}}^{(0,r]}$ sur $\widetilde{\mathbf{B}}^{(0,p^{-1}r]}$ et de $\widetilde{\mathbf{B}}^{]0,r]}$ sur $\widetilde{\mathbf{B}}^{]0,p^{-1}r]}$

Soit $\pi = [\varepsilon]-1 \in \widetilde{\mathbf{A}}^{+}$. Les choix que l'on a faits identifient $\mathbf{B}_{\mathrm{F}}$ au sous-corps $\mathscr{E}_{\mathrm{F}}(\pi)$ de $\widetilde{\mathbf{B}}$. La formule $g(\pi) = (1+\pi)^{\chi(g)}-1$ montre que $\mathbf{B}_{\mathrm{F}}$ est stable sous l'action de $\mathscr{G}_{\mathrm{F}}$ et que l'action de $\mathscr{G}_{\mathrm{F}}$ commute à celle de φ (c'est ce qui justifie les choix en question). Cette stabilité entraine la stabilité de $\mathbf{A}$, $\mathbf{B}$ et par suite, celle de $\mathbf{B}^{\dagger}$, $\mathbf{B}_{\mathrm{rig}}^{\dagger}$, $\mathbf{B}_{\mathrm{rig}}^{+}$, $\mathbf{B}^{(0,r]}$ et $\mathbf{B}^{]0,r]}$ sous l'action de $\mathscr{G}_{\mathrm{F}}$.

Comme $\mathbf{E}^{\mathscr{H}_{\mathrm{K}}} = \mathbf{E}_{\mathrm{K}}$ et $\widetilde{\mathbf{E}}^{\mathscr{H}_{\mathrm{K}}} = \widetilde{\mathbf{E}}_{\mathrm{K}}$, les points fixes de l'un des anneaux ci-dessus sous l'action de $\mathscr{H}_{\mathrm{K}}$ est l'anneau ayant le même nom mais avec un K en indice ; par exemple $\mathbf{B}^{\mathscr{H}_{\mathrm{K}}} = \mathbf{B}_{\mathrm{K}}$, $(\mathbf{B}^{\dagger})^{\mathscr{H}_{\mathrm{K}}} = \mathbf{B}_{\mathrm{K}}^{\dagger}$ ou $(\mathbf{B}_{\mathrm{rig}}^{\dagger})^{\mathscr{H}_{\mathrm{K}}} = \mathbf{B}_{\mathrm{rig,K}}^{\dagger}$. D'autre part, l'action de $\mathscr{H}_{\mathrm{K}}$ sur $\mathbf{E}$ étant continue pour la topologie discrète, $\mathscr{H}_{\mathrm{K}}$ agit continûment sur $\mathbf{A}$ et $\mathbf{B}$ pour la topologie forte.

4.4.3. Éléments surconvergents et propriétés de Tate-Sen. — L'action de φ sur les anneaux avec $\widetilde{\ }$ est bijective, mais pas sur les anneaux sans $\widetilde{\ }$, et on rajoute un n en indice pour indiquer que l'on a appliqué φ^{-n}. Par exemple,

$$\mathbf{B}_{\mathrm{K},n}^{\dagger} = \varphi^{-n}(\mathbf{B}_{\mathrm{K}}^{\dagger}) \subset \widetilde{\mathbf{B}}_{\mathrm{K}}^{\dagger}, \quad \mathbf{A}_{\mathrm{K},n}^{(0,r]} = \varphi^{-n}(\mathbf{A}_{\mathrm{K}}^{(0,p^{-n}r]}) \subset \widetilde{\mathbf{A}}_{\mathrm{K}}^{(0,r]}.$$

Le corps $\mathbf{B}$ est une extension de degré p de $\varphi(\mathbf{B})$, et on définit un opérateur $\psi : \mathbf{B} \to \mathbf{B}$ par la formule $\psi(x) = p^{-1}\varphi^{-1}(\mathrm{Tr}_{\mathbf{B}/\varphi(\mathbf{B})}(x))$. L'opérateur ψ est un inverse à gauche de φ qui commute à l'action de $\mathscr{G}_{\mathrm{K}}$. Par ailleurs, si L est une extension finie

de F, et r est assez petit, alors $\psi(\mathbf{A}_{\mathrm{L}}^{(0,r]}) \subset \mathbf{A}_{\mathrm{L}}^{(0,r]}$. Ceci permet, si $r > 0$ et si n est assez grand ($n \geqslant n(r, \mathrm{L})$) et $k \in \mathbf{N}$, de définir une application $\mathrm{R}_{\mathrm{L},n} = \varphi^{-n} \circ \psi^k \circ \varphi^{n+k}$: $\mathbf{A}_{\mathrm{L},n+k}^{(0,r]} \to \mathbf{A}_{\mathrm{L},n}^{(0,r]}$, et on montre que la réunion des $\mathbf{A}_{\mathrm{L},n}^{(0,r]}$, $n \in \mathbf{N}$ est dense dans $\widetilde{\mathbf{A}}_{\mathrm{L}}^{(0,r]}$ et que $\mathrm{R}_{\mathrm{L},n}$ se prolonge par continuité à $\widetilde{\mathbf{A}}_{\mathrm{L}}^{(0,r]}$.

Proposition 4.5 ([12]). — *L'anneau $\widetilde{\Lambda} = \widetilde{\mathbf{A}}^{(0,r]}$ vérifie les conditions* (TS1), (TS2) *et* (TS3), *avec $v = w_r$, $\widetilde{\Lambda}^{\mathscr{H}_{\mathrm{L}}} = \widetilde{\mathbf{A}}_{\mathrm{L}}^{(0,r]}$, $\Lambda_{\mathscr{H}_{\mathrm{L}},n} = \mathbf{A}_{\mathrm{L},n}^{(0,r]}$, $\mathrm{R}_{\mathscr{H}_{\mathrm{L}},n} = \mathrm{R}_{\mathrm{L},n}$, les constantes $c_1 > 0$, $c_2 > 0$ et $c_3 > \frac{r}{p-1}$ pouvant être choisies arbitrairement.*

4.5. Le logarithme et l'anneau $\widetilde{\mathbf{B}}_{\log}^{\dagger}$

Si $x \in \widetilde{\mathbf{A}}^{\dagger}$ vérifie $v_{\mathbf{E}}(\overline{x} - 1) > 0$, la série $\log x = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}(x-1)^n$ converge dans $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{\dagger}$. Par exemple, $t = \log[\varepsilon]$ est un élément de $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}$ sur lequel $\mathscr{G}_{\mathrm{F}}$ agit par multiplication par le caractère cyclotomique et qui peut être vu comme un analogue p-adique de $2i\pi$. On aimerait bien étendre cette application à $(\widetilde{\mathbf{B}}^{\dagger})^*$, mais pour cela, on est forcé d'étendre un peu l'anneau $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{\dagger}$ en rajoutant un analogue p-adique u de $\log p$.

Soit $\widetilde{p} = (p, p^{1/p}, \dots) \in \widetilde{\mathbf{E}}^{+}$. Si $\sigma \in \mathscr{G}_{\mathrm{F}}$, il existe $c(\sigma) \in \mathbf{Z}_p$ tel que l'on ait $\sigma(\widetilde{p}) = \widetilde{p}\varepsilon^{c(\sigma)}$ ($\sigma \mapsto c(\sigma)$ est le cocycle à valeurs dans $\mathbf{Z}_p(1)$ associé à p par la théorie de Kummer).

Soit $\widetilde{\mathbf{B}}_{\log}^{\dagger} = \widetilde{\mathbf{B}}_{\mathrm{rig}}^{\dagger}[u]$. On munit $\widetilde{\mathbf{B}}_{\log}^{\dagger}$ d'un opérateur « de monodromie » $\mathrm{N} = -d/du$ et d'une action de φ (resp. $\mathscr{G}_{\mathrm{F}}$), compatible avec celle existant sur $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{\dagger}$, en posant $\varphi(u) = pu$ (resp. $\sigma(u) = u + c(\tau)t$). Les actions de φ et N commutent à celle de $\mathscr{G}_{\mathrm{F}}$ et on a $\mathrm{N}\varphi = p\varphi\mathrm{N}$.

Proposition 4.6. — *Il existe une unique application* $\log : (\widetilde{\mathbf{B}}^{\dagger})^* \to \widetilde{\mathbf{B}}_{\log}^{\dagger}$ *vérifiant les propriétés suivantes* :

(i) $\log xy = \log x + \log y$;

(ii) $\log x = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}(x-1)^n$ *si la série converge* ;

(iii) $\log[a] = 0$ *si* $a \in \overline{k}_{\mathrm{F}}$;

(iv) $\log p = 0$ *et* $\log[\widetilde{p}] = u$.

De plus, on a $\varphi(\log x) = \log \varphi(x)$ *et* $\sigma(\log x) = \log \sigma(x)$ *si* $\sigma \in \mathscr{G}_{\mathrm{F}}$.

On note $\widetilde{\mathbf{B}}_{\log}^{+}$ le sous-anneau $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}[u]$ de $\widetilde{\mathbf{B}}_{\log}^{\dagger}$; il est stable par N, φ et $\mathscr{G}_{\mathrm{F}}$. Si K est une extension finie de F, on note $\mathbf{B}_{\log,\mathrm{K}}^{\dagger}$ le sous-anneau $\mathbf{B}_{\mathrm{rig},\mathrm{K}}^{\dagger}[\log \pi_{\mathrm{K}}]$ de $\widetilde{\mathbf{B}}_{\log}^{\dagger}$; il est stable par N, φ et $\mathscr{G}_{\mathrm{K}}$ (qui agit à travers Γ_{K}).

4.6. L'anneau $\mathbf{B}_{\mathrm{dR}}^{+}$

4.6.1. Construction de $\mathbf{B}_{\mathrm{dR}}^{+}$

PROPOSITION 4.7 ([36]). — *L'application* $\theta : \widetilde{\mathbf{A}}^{+} \to \mathscr{O}_{\mathrm{C}}$, *donnée par* $\sum_{n=0}^{+\infty} p^n[x_n] \mapsto \sum_{n=0}^{+\infty} p^n x_n^{(0)}$, *est un morphisme surjectif d'anneaux dont le noyau est un idéal principal engendré par* $\omega = \pi/\varphi^{-1}(\pi) = 1 + [\varepsilon^{1/p}] + \cdots + [\varepsilon^{1/p}]^{p-1}$ *ou par* $\xi = [\widetilde{p}] - p$.

Remarque 4.8. — $(\widetilde{\mathbf{A}}^{+}, \theta)$ est l'épaississement p-adique universel[(10)] de $\mathscr{O}_{\mathrm{C}}$ (cf. [39]).

On prolonge θ en un morphisme de $\widetilde{\mathbf{B}}^{+}$ sur C, on note $\mathbf{B}_{\mathrm{dR}}^{+}$ l'anneau $\varprojlim \widetilde{\mathbf{B}}^{+}/(\ker\theta)^n$. Ceci fait de $\mathbf{B}_{\mathrm{dR}}^{+}$ un anneau de valuation discrète d'idéal maximal $\ker\theta$ et de corps résiduel C. En particulier, la clôture séparable de F dans $\mathbf{B}_{\mathrm{dR}}^{+}$ s'identifie à $\overline{\mathrm{F}}$.

On munit $\mathbf{B}_{\mathrm{dR}}^{+}$ de la topologie pour laquelle les $p^k\widetilde{\mathbf{A}}^{+} + (\ker\theta)^n$, avec $n, k \in \mathbf{N}$, forment une base de voisinages de 0. L'action de $\mathscr{G}_{\mathrm{F}}$ sur $\widetilde{\mathbf{B}}^{+}$ s'etend par continuité en une action continue de $\mathscr{G}_{\mathrm{F}}$ sur $\mathbf{B}_{\mathrm{dR}}^{+}$. La série définissant $t = \log[\varepsilon] = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}\pi^n$ converge dans $\mathbf{B}_{\mathrm{dR}}^{+}$ et t est un générateur de $\ker\theta$ (ce qui explique pourquoi on a mis autant de temps à comprendre ce qu'était l'analogue p-adique de $2i\pi$).

On pose $\mathbf{B}_{\mathrm{dR}} = \mathbf{B}_{\mathrm{dR}}^{+}[t^{-1}]$, ce qui fait de $\mathbf{B}_{\mathrm{dR}}$ un corps et on munit $\mathbf{B}_{\mathrm{dR}}$ de la filtration décroissante définie par $\mathbf{B}_{\mathrm{dR}}^{i} = t^i\mathbf{B}_{\mathrm{dR}}^{+}$. Cette filtration est stable par l'action de $\mathscr{G}_{\mathrm{K}}$.

4.6.2. $\mathbf{B}_{\mathrm{dR}}^{+}$ *et le reste du monde.* — Il n'existe pas de morphisme continu de $\widetilde{\mathbf{B}}_{\log}^{\dagger}$ dans $\mathbf{B}_{\mathrm{dR}}$ car, bien que ces anneaux soient obtenus en localisant et complétant $\widetilde{\mathbf{A}}^{+}$, les topologies sont trop différentes. Toutefois, $\widetilde{\mathbf{B}}_{\log}^{\dagger}$ est une limite inductive d'anneaux qui se plongent dans $\mathbf{B}_{\mathrm{dR}}^{+}$, mais il faut changer de plongement pour chacun de ces sous-anneaux. De manière précise,

(i) $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}$ et $\mathbf{B}_{\mathrm{dR}}^{+}$ sont deux complétés de $\widetilde{\mathbf{B}}^{+}$ et l'identité $\widetilde{\mathbf{B}}^{+} \to \widetilde{\mathbf{B}}^{+}$ se prolonge en une injection continue de $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}$ dans $\mathbf{B}_{\mathrm{dR}}^{+}$.

(ii) La série $\log\frac{[\widetilde{p}]}{p} = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}\big(\frac{[\widetilde{p}]}{p} - 1\big)^n$ converge dans $\mathbf{B}_{\mathrm{dR}}^{+}$, ce qui nous fournit une injection naturelle de $\widetilde{\mathbf{B}}_{\log}^{+}$ dans $\mathbf{B}_{\mathrm{dR}}^{+}$ qui commute à l'action de $\mathscr{G}_{\mathrm{F}}$. Les anneaux $\widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}$, $\widetilde{\mathbf{B}}_{\log}^{+}$ et $\mathbf{B}_{\mathrm{dR}}$ sont alors reliés par les *suites exactes fondamentales* [39, 28]

$$0 \longrightarrow \widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}[\tfrac{1}{t}] \longrightarrow \widetilde{\mathbf{B}}_{\log}^{+}[\tfrac{1}{t}] \xrightarrow{\mathrm{N}} \widetilde{\mathbf{B}}_{\log}^{+}[\tfrac{1}{t}] \longrightarrow 0$$
$$0 \longrightarrow \big(\widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}[\tfrac{1}{t}]\big)^{\varphi=1} \longrightarrow \widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}[\tfrac{1}{t}] \xrightarrow{1-\varphi} \widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}[\tfrac{1}{t}] \longrightarrow 0$$
$$0 \longrightarrow \mathbf{Q}_p \longrightarrow \big(\widetilde{\mathbf{B}}_{\mathrm{rig}}^{+}[\tfrac{1}{t}]\big)^{\varphi=1} \longrightarrow \mathbf{B}_{\mathrm{dR}}/\mathbf{B}_{\mathrm{dR}}^{0} \longrightarrow 0$$

D'autre part, on a $\mathbf{B}_{\mathrm{dR}}^{\mathscr{G}_{\mathrm{K}}} = \mathrm{K}$ et $(\widetilde{\mathbf{B}}_{\log}^{+}[\frac{1}{t}])^{\mathscr{G}_{\mathrm{K}}} = \mathrm{F}$, et l'application naturelle de $\mathrm{K} \otimes_{\mathrm{F}} \widetilde{\mathbf{B}}_{\log}^{+}[\frac{1}{t}]$ dans $\mathbf{B}_{\mathrm{dR}}$ est injective [36, 39, 28].

(10) L'anneau $\widetilde{\mathbf{A}}^{+}$ est très souvent noté W(R) et parfois $\mathrm{A}_{\mathrm{inf}}$.

(iii) Si $x = \sum_{k=0}^{+\infty} p^k[x_k] \in \widetilde{\mathbf{A}}$, la série converge dans $\mathbf{B}_{\mathrm{dR}}^+$ si et seulement si la série $\theta(x) = \sum_{k=0}^{+\infty} p^k x_k^{(0)}$ converge dans C, c'est-à-dire si et seulement si $k+v_{\mathbf{E}}(x_k)$ tend vers $+\infty$ quand k tend vers $+\infty$. On en déduit une application naturelle $\iota_0 : \widetilde{\mathbf{B}}^{(0,1]} \to \mathbf{B}_{\mathrm{dR}}^+$ qui s'avère être injective.

(iv) Comme $\widetilde{\mathbf{B}}^{]0,1]} = \widetilde{\mathbf{B}}_{\mathrm{rig}}^+ + \widetilde{\mathbf{B}}^{(0,1]}$ et comme φ^{-n} induit une bijection continue de $\widetilde{\mathbf{B}}^{]0,p^{-n}]}[u]$ sur $\widetilde{\mathbf{B}}^{]0,1]}[u]$, on obtient [6], quel que soit $n \in \mathbf{N}$, une injection $\iota_n = \iota_0 \circ \varphi^{-n}$ de $\widetilde{\mathbf{B}}^{]0,p^{-n}]}[u]$ dans $\mathbf{B}_{\mathrm{dR}}^+$. Cette injection peut être vue comme la « localisation en $\varepsilon^{(n)} - 1$ ». Si K est une extension finie de F, il existe $n(\mathrm{K})$ tel que, pour tout $n \geqslant n(\mathrm{K})$, on ait

$$\iota_n\Big(\mathbf{B}_{\mathrm{K}}^{]0,p^{-n}]}[\log \pi_{\mathrm{K}}]\Big) \subset \mathrm{K}_n[[t]].$$

5. CLASSIFICATION DES REPRÉSENTATIONS p-ADIQUES

5.1. Le (φ,Γ)-module associé à une représentation de $\mathscr{G}_{\mathrm{K}}$

Si V est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$, le $\mathbf{B}_{\mathrm{K}}$-espace vectoriel $\mathbf{D}(\mathrm{V}) = (\mathbf{B} \otimes_{\mathbf{Q}_p} \mathrm{V})^{\mathscr{H}_{\mathrm{K}}}$ est muni d'une action semi-linéaire de φ provenant de l'action de φ sur $\mathbf{B}$ (et qui est donc *étale*, c'est-à-dire de φ-pente 0) et d'une action résiduelle de $\Gamma_{\mathrm{K}} = \mathscr{G}_{\mathrm{K}}/\mathscr{H}_{\mathrm{K}}$. Pour des raisons évidentes, un tel objet s'appelle *un (φ,Γ)-module étale sur $\mathbf{B}_{\mathrm{K}}$*.

Si D est un (φ,Γ)-module sur $\mathbf{B}_{\mathrm{K}}$, alors $\mathbf{V}(\mathrm{D}) = (\mathbf{B} \otimes_{\mathbf{B}_{\mathrm{K}}} \mathrm{D})^{\varphi=1}$ est un $\mathbf{Q}_p$-espace vectoriel muni d'une action continue de $\mathscr{G}_{\mathrm{K}}$.

Théorème 5.1 ([38]). — *La catégorie des $\mathbf{Q}_p$-représentations de $\mathscr{G}_{\mathrm{K}}$ est équivalente à celle des (φ,Γ)-modules étales sur $\mathbf{B}_{\mathrm{K}}$. Plus précisément, si* V *est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$, alors $\mathbf{D}(\mathrm{V})$ est un (φ,Γ)-module étale sur $\mathbf{B}_{\mathrm{K}}$ et $\mathbf{V}(\mathbf{D}(\mathrm{V})) = \mathrm{V}$ et, réciproquement, si* D *est un (φ,Γ)-module étale sur $\mathbf{B}_{\mathrm{K}}$, alors $\mathbf{V}(\mathrm{D})$ est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$ et $\mathbf{D}(\mathbf{V}(\mathrm{D})) = \mathrm{D}$.*

La démonstration de ce théorème repose sur les faits suivants :

— si on munit $\mathbf{E}$ de la topologie discrète, alors $\mathrm{H}^1(\mathscr{H}_{\mathrm{K}}, \mathrm{GL}_d(\mathbf{E})) = 1$ d'après le théorème de Hilbert 90 et donc, si on munit $\mathbf{B}$ de la topologie forte, alors $\mathrm{H}^1(\mathscr{H}_{\mathrm{K}}, \mathrm{GL}_d(\mathbf{B})) = 1$ (cf. prop. 3.3 et 3.4). Autrement dit, toute représentation de $\mathscr{H}_{\mathrm{K}}$ est $\mathbf{B}$-admissible et $\mathrm{V} \mapsto \mathbf{D}(\mathrm{V})$ a de bonnes propriétés.

— $\mathbf{B}^{\varphi=1} = \mathbf{Q}_p$, et donc $\mathbf{V}(\mathbf{D}(\mathrm{V})) = \mathrm{V}$ si V est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$.

— si $(a_{i,j})_{1\leqslant i,j\leqslant d} \in \mathrm{GL}_d(\mathbf{E})$, alors le système d'équations $x_i^p = \sum_{j=1}^d a_{i,j}x_j$, $1 \leqslant j \leqslant d$ admet p^d solutions dans $\mathbf{E}^d$ et ces solutions forment un $\mathbf{F}_p$-espace vectoriel de dimension d engendrant $\mathbf{E}^d$. Un (φ,Γ)-module est étale si et seulement si il existe une base dans laquelle la matrice de φ appartient à $\mathrm{GL}_d(\mathbf{A}_{\mathrm{K}})$ et le résultat précédent permet de montrer que, si Δ est le $\mathbf{A}_{\mathrm{K}}$ réseau de D engendré par cette base, alors $(\mathbf{E} \otimes_{\mathbf{A}_{\mathrm{K}}} \Delta)^{\varphi=1}$ est un $\mathbf{F}_p$-espace vectoriel de dimension d engendrant $\mathbf{E} \otimes_{\mathbf{A}_{\mathrm{K}}} \Delta$;

autrement dit, « $\mathbf{V}(\mathrm{D})$ mod p » a la bonne dimension. Le résultat cherché s'en déduit « par dévissage et passage à la limite ».

Remarque 5.2. — (i) Nous n'avons pas utilisé Γ_{K} pour montrer que les foncteurs $\mathrm{V} \mapsto \mathbf{D}(\mathrm{V})$ et $\mathrm{D} \mapsto \mathbf{V}(\mathrm{D})$ sont inverses l'un de l'autre. Ceci permet de montrer que la catégorie des $\mathbf{Q}_p$-représentations de $\mathscr{H}_{\mathrm{K}} \cong \mathscr{G}_{k_{\mathrm{F}}((\mathrm{T}))}$ est équivalente à celle des φ-modules étales sur $\mathbf{B}_{\mathrm{K}} \cong \mathscr{E}_{\mathrm{F}}$. Ce résultat est une version locale de résultats généraux de Katz [52, chap. 4].

(ii) Nous avons privilégié dans ce texte les $\mathbf{Q}_p$-représentations, mais la théorie des (φ, Γ)-modules donne d'aussi bons résultats, en tensorisant par $\mathbf{A}$ au lieu de $\mathbf{B}$, si on considère des $\mathbf{Z}_p$-modules de type fini (pas nécessairement libres) munis d'une action continue de $\mathscr{G}_{\mathrm{K}}$ (ou $\mathscr{H}_{\mathrm{K}}$).

Comme application, on a par exemple le résultat suivant qui entre dans la démonstration du théorème 2.9 : si D est un φ-module étale sur $\mathbf{B}_{\mathrm{K}}$, il existe une extension finie L de K et une base de $\mathbf{B}_{\mathrm{L}} \otimes_{\mathbf{B}_{\mathrm{K}}} \mathrm{D}$ sur $\mathbf{B}_{\mathrm{L}}$ dans laquelle la matrice de φ appartient à $1 + p^n \mathrm{M}_d(\mathbf{A}_{\mathrm{L}})$; en effet, ce n'est qu'une traduction de ce que, si V est une $\mathbf{Q}_p$-représentation de $\mathscr{H}_{\mathrm{K}}$, il existe un $\mathbf{Z}_p$-réseau T de V stable par $\mathscr{H}_{\mathrm{K}}$ et une extension finie L de K telle que $\mathscr{H}_{\mathrm{L}}$ agisse trivialement sur $\mathrm{T}/p^n\mathrm{T}$.

(iii) On peut utiliser l'équivalence de catégories ci-dessus pour étudier la cohomologie galoisienne et la théorie d'Iwasawa des représentations de $\mathscr{G}_{\mathrm{K}}$; nous renvoyons à [24] pour un résumé des résultats et à [47, 48, 4, 13] pour les détails.

(iv) Le théorème ci-dessus n'est pas assez fin pour vraiment étudier les problèmes de classification des représentations p-adiques ; le problème est que le corps $\mathbf{B}_{\mathrm{K}}$ est un objet un peu grossier et aller plus loin demande de descendre les coefficients à $\mathbf{B}_{\mathrm{K}}^{\dagger}$.

La méthode de Sen permet, utilisant la proposition 4.5, de prouver le résultat suivant :

Proposition 5.3 ([12]). — *Si* K *est une extension finie de* F *et* d *un entier* $\geqslant 1$*, les applications naturelles*

$$\varinjlim \mathrm{H}^1(\Gamma_{\mathrm{K}}, \mathrm{GL}_d(\mathbf{B}_{\mathrm{K},n}^{\dagger})) \longrightarrow \mathrm{H}^1(\Gamma_{\mathrm{K}}, \mathrm{GL}_d(\widetilde{\mathbf{B}}_{\mathrm{K}}^{\dagger})) \longrightarrow \mathrm{H}^1(\mathscr{G}_{\mathrm{K}}, \mathrm{GL}_d(\widetilde{\mathbf{B}}^{\dagger})),$$

induites par l'inflation de Γ_{K} *à* $\mathscr{G}_{\mathrm{K}}$ *et les inclusions* $\mathbf{B}_{\mathrm{K},n}^{\dagger} \subset \widetilde{\mathbf{B}}_{\mathrm{K}}^{\dagger} \subset \widetilde{\mathbf{B}}^{\dagger}$*, sont des bijections.*

En particulier, comme le cocycle correspondant à une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$ est fixe par φ^n pour tout n, il est dans l'image de $\mathrm{H}^1(\Gamma_{\mathrm{K}}, \mathrm{GL}_d(\mathbf{B}_{\mathrm{K}}^{\dagger}))$; on en déduit la $\mathbf{B}^{\dagger}$-admissibilité des représentations de $\mathscr{H}_{\mathrm{K}}$ obtenues par restriction d'une représentation de $\mathscr{G}_{\mathrm{K}}$. Ceci permet, en définissant $\mathbf{D}^{\dagger}(\mathrm{V}) = (\mathbf{B}^{\dagger} \otimes_{\mathbf{Q}_p} \mathrm{V})^{\mathscr{H}_{\mathrm{K}}}$ pour une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$ et $\mathbf{V}^{\dagger}(\mathrm{D}) = (\mathbf{B}^{\dagger} \otimes_{\mathbf{B}_{\mathrm{K}}^{\dagger}} \mathrm{D})^{\varphi=1}$ pour un (φ, Γ)-module sur $\mathbf{B}_{\mathrm{K}}^{\dagger}$, de raffiner le théorème 5.1 sous la forme

THÉORÈME 5.4 ([12]). — *La catégorie des $\mathbf{Q}_p$-représentations de $\mathscr{G}_{\mathrm{K}}$ est équivalente à celle des (φ,Γ)-modules étales sur $\mathbf{B}_{\mathrm{K}}^{\dagger}$. Plus précisément, si* V *est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$, alors $\mathbf{D}^{\dagger}(\mathrm{V})$ est un (φ,Γ)-module étale sur $\mathbf{B}_{\mathrm{K}}$ et $\mathbf{V}^{\dagger}(\mathbf{D}^{\dagger}(\mathrm{V}))=\mathrm{V}$ et, réciproquement, si* D *est un (φ,Γ)-module étale sur $\mathbf{B}_{\mathrm{K}}^{\dagger}$, alors $\mathbf{V}^{\dagger}(\mathrm{D})$ est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$ et $\mathbf{D}^{\dagger}(\mathbf{V}^{\dagger}(\mathrm{D}))=\mathrm{D}$.*

D'autre part, on montre [11] que tout sous-$\mathbf{B}_{\mathrm{K}}^{\dagger}$-espace vectoriel de dimension finie de $\mathbf{D}(\mathrm{V})$ stable par φ est inclus dans $\mathbf{D}^{\dagger}(\mathrm{V})$; ceci nous fournit le « résultat de descente » suivant dans lequel les représentations galoisiennes ont disparu.

COROLLAIRE 5.5. — *Si* D *est un (φ,Γ)-module étale sur $\mathbf{B}_{\mathrm{K}}$, alors l'ensemble des sous-$\mathbf{B}_{\mathrm{K}}^{\dagger}$-espaces vectoriels de dimension finie stables par φ admet un plus grand élément $\mathrm{D}^{\dagger}$ et on a $\mathrm{D}=\mathbf{B}_{\mathrm{K}}\otimes_{\mathbf{B}_{\mathrm{K}}^{\dagger}}\mathrm{D}^{\dagger}$.* [On a $\mathrm{D}^{\dagger}=\mathbf{D}^{\dagger}(\mathbf{V}(\mathrm{D}))$.]

5.2. La hiérarchie des représentations galoisiennes [40]

5.2.1. Représentations de Hodge-Tate. — Si V est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$, le K-espace vectoriel $\mathbf{D}_{\mathrm{HT}}(\mathrm{V})=(\mathbf{B}_{\mathrm{HT}}\otimes\mathrm{V})^{\mathscr{G}_{\mathrm{K}}}$ est muni d'une graduation indexée par les entiers et un entier i tel que $\mathrm{Gr}^{-i}\mathbf{D}_{\mathrm{HT}}(\mathrm{V})\neq 0$ est un *poids de Hodge-Tate* de V. Une représentation $\mathbf{B}_{\mathrm{HT}}$-admissible est dite *de Hodge-Tate*. Une représentation C-admissible est clairement de Hodge-Tate et, réciproquement, une représentation de Hodge-Tate est C-admissible si 0 est son seul poids de Hodge-Tate.

5.2.2. Représentations de de Rham. — Si V est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$, le K-espace vectoriel $\mathbf{D}_{\mathrm{dR}}(\mathrm{V})=(\mathbf{B}_{\mathrm{dR}}\otimes\mathrm{V})^{\mathscr{G}_{\mathrm{K}}}$ est muni d'une filtration par des sous-K-espaces vectoriels $\mathbf{D}_{\mathrm{dR}}^{i}(\mathrm{V})$ pour $i\in\mathbf{Z}$ qui est décroissante (i.e. $\mathbf{D}_{\mathrm{dR}}^{i+1}(\mathrm{V})\subset\mathbf{D}_{\mathrm{dR}}^{i}(\mathrm{V})$) exhaustive (i.e. $\mathbf{D}_{\mathrm{dR}}^{i}(\mathrm{V})=\mathbf{D}_{\mathrm{dR}}(\mathrm{V})$ si $i\ll 0$) et séparée (i.e $\mathbf{D}_{\mathrm{dR}}^{i}(\mathrm{V})=\{0\}$ si $i\gg 0$). Une représentation $\mathbf{B}_{\mathrm{dR}}$-admissible est dite *de de Rham*. Si V est de de Rham, alors V est de Hodge-Tate et $\mathbf{D}_{\mathrm{HT}}(\mathrm{V})$ est le gradué associé à $\mathbf{D}_{\mathrm{dR}}(\mathrm{V})$.

Utilisant l'injection $\iota_n:\widetilde{\mathbf{B}}^{(0,p^{-n}]}\to\mathbf{B}_{\mathrm{dR}}^{+}$, on peut retrouver $\mathbf{D}_{\mathrm{dR}}(\mathrm{V})$ et sa filtration à partir de $\mathbf{D}^{\dagger}(\mathrm{V})$. (Si $r>0$, on note $\mathbf{D}^{(0,r]}(\mathrm{V})$ le $\mathbf{B}_{\mathrm{K}}^{(0,r]}$-module $(\mathbf{B}^{(0,r]}\otimes\mathrm{V})^{\mathscr{H}_{\mathrm{K}}}$; c'est [11] le plus grand sous-$\mathbf{B}_{\mathrm{K}}^{(0,r]}$-module D de type fini de $\mathbf{D}(\mathrm{V})$ tel que l'on ait $\varphi(\mathrm{D})\subset\mathbf{B}_{\mathrm{K}}^{(0,p^{-1}r]}\otimes_{\mathbf{B}_{\mathrm{K}}^{(0,r]}}\mathrm{D}$.)

PROPOSITION 5.6 ([6]+[42]). — *Si* V *est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$, il existe $n(\mathrm{V})$ tel que si $n\geqslant n(\mathrm{V})$ et $i\in\mathbf{Z}$, alors*

$$\mathbf{D}_{\mathrm{dR}}^{i}(\mathrm{V})=\Bigl(t^{i}\mathrm{K}_n[[t]]\otimes_{\mathbf{B}_{\mathrm{K}}^{(0,p^{-n}]}}\mathbf{D}^{(0,p^{-n}]}(\mathrm{V})\Bigr)^{\Gamma_{\mathrm{K}}},$$

où $t^{i}\mathrm{K}_n[[t]]$ est considéré comme un $\mathbf{B}_{\mathrm{K}}^{(0,p^{-n}]}$-module via $\iota_n:\mathbf{B}_{\mathrm{K}}^{(0,p^{-n}]}\to\mathrm{K}_n[[t]]$.

5.2.3. Représentations semi-stables. — Une $\mathbf{Q}_p$-représentation de $\mathscr{G}_K$ qui est $\widetilde{\mathbf{B}}^+_{\log}[\frac{1}{t}]$-admissible (resp. $\widetilde{\mathbf{B}}^+_{\mathrm{rig}}[\frac{1}{t}]$-admissible) est dite *semi-stable* (resp. *cristalline*). Une $\mathbf{Q}_p$-représentation V de $\mathscr{G}_K$ est dite *potentiellement semi-stable* s'il existe une extension finie L de K telle que la restriction de V à $\mathscr{G}_L$ soit semi-stable (en tant que représentation de $\mathscr{G}_L$). On a les implications suivantes :

(i) « V cristalline » ⇒ « V semi-stable » ;

(ii) « V semi-stable » ⇒ « V potentiellement semi-stable » ;

(iii) « V potentiellement semi-stable » ⇒ « V de de Rham » ;

(iv) « V de de Rham » ⇒ « V de Hodge-Tate ».

La première implication vient de l'inclusion $\widetilde{\mathbf{B}}^\dagger_{\mathrm{rig}}[\frac{1}{t}] \subset \widetilde{\mathbf{B}}^\dagger_{\log}[\frac{1}{t}]$, la seconde est une évidence, la troisième vient de ce que $\mathbf{B}_{\mathrm{dR}}$ contient $\widetilde{\mathbf{B}}^\dagger_{\log}[\frac{1}{t}]$ et $\overline{\mathrm{F}}$, et la dernière de ce que $\mathbf{B}_{\mathrm{HT}}$ est l'algèbre graduée associée à $\mathbf{B}_{\mathrm{dR}}$.

En ce qui concerne les implications en sens inverse, on a

(o) il existe des représentations cristallines non triviales ; par exemple, si $k \in \mathbf{Z}$ le $\mathbf{Q}_p$-espace vectoriel $\mathbf{Q}_p(k)$ de dimension 1 sur lequel $\sigma \in \mathscr{G}_K$ agit par multiplication par la puissance k-ième $\chi(\sigma)^k$ du caractère cyclotomique ; plus généralement, et ça a été la motivation [36] pour l'introduction de toutes ces notions, si X est une variété propre et lisse sur K possédant un modèle sur $\mathscr{O}_K$ ayant bonne réduction, alors les $\mathbf{Q}_p$-représentations $\mathrm{H}^i_{\mathrm{et}}(\mathrm{X}_{\overline{K}}, \mathbf{Q}_p)$ de $\mathscr{G}_K$ fournies par la cohomologie étale de X, sont cristallines ;

(i) la représentation $\sigma \mapsto \begin{pmatrix} \chi(\sigma) & c(\sigma) \\ 0 & 1 \end{pmatrix}$ associée à $\log p$ est semi-stable mais pas cristalline ;

(ii) une représentation sur laquelle l'inertie de $\mathscr{G}_K$ agit à travers un quotient fini non trivial est potentiellement semi-stable (et même potentiellement cristalline), mais pas semi-stable ;

(iii) l'implication « V de Rham » ⇒ « V potentiellement semi-stable » est la conjecture de monodromie p-adique de Fontaine ;

(iv) une extension non triviale $0 \to \mathbf{Q}_p \to \mathrm{V} \to \mathbf{Q}_p(k) \to 0$ est de Hodge-Tate mais pas de de Rham si $k > 0$;

(v) la représentation $\sigma \mapsto \begin{pmatrix} 1 & \log \chi(\sigma) \\ 0 & 1 \end{pmatrix}$ associée à $\log 2i\pi$ n'est pas de Hodge-Tate.

5.2.4. (φ, N)-modules filtrés. — Si V est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_K$, le F-espace vectoriel $\mathbf{D}_{\mathrm{st}}(\mathrm{V}) = (\widetilde{\mathbf{B}}^+_{\log}[\frac{1}{t}] \otimes \mathrm{V})^{\mathscr{G}_K}$ est muni d'une action semi-linéaire de φ et d'une action linéaire de N vérifiant la relation $\mathrm{N}\varphi = p\varphi\mathrm{N}$. D'autre part, $\mathrm{K} \otimes_{\mathrm{F}} \mathbf{D}_{\mathrm{st}}(\mathrm{V})$ s'injecte dans $\mathbf{D}_{\mathrm{dR}}(\mathrm{V})$ et est muni d'une filtration par des sous-K-espaces vectoriels qui est décroissante, exhaustive et séparée. Un tel objet est appelé *un (φ, N)-module filtré sur* K. Le sous-F-espace vectoriel $\mathbf{D}_{\mathrm{cris}}(\mathrm{V}) = \mathbf{D}_{\mathrm{st}}(\mathrm{V})^{\mathrm{N}=0}$ de $\mathbf{D}_{\mathrm{st}}(\mathrm{V})$ s'identifie à $(\widetilde{\mathbf{B}}^+_{\mathrm{rig}}[\frac{1}{t}] \otimes \mathrm{V})^{\mathscr{G}_K}$; c'est *un φ-module filtré sur* K.

Si D est un (φ, N)-module filtré sur K et $\mathrm{D}_{\mathrm{K}} = \mathrm{K} \otimes_{\mathrm{F}} \mathrm{D}$, l'injection de $\widetilde{\mathbf{B}}^{+}_{\log}[\frac{1}{t}]$ dans $\mathbf{B}_{\mathrm{dR}}$ induit une application

$$\Big(\widetilde{\mathbf{B}}^{+}_{\log}[\frac{1}{t}] \otimes_{\mathrm{F}} \mathrm{D}\Big)^{\mathrm{N}=0,\varphi=1} \longrightarrow \Big(\mathbf{B}_{\mathrm{dR}} \otimes_{\mathrm{K}} \mathrm{D}_{\mathrm{K}}\Big)/\Big(\mathbf{B}_{\mathrm{dR}} \otimes_{\mathrm{K}} \mathrm{D}_{\mathrm{K}}\Big)^{0}$$

dont le noyau $\mathbf{V}_{\mathrm{st}}(\mathrm{D})$ est un $\mathbf{Q}_p$-espace vectoriel (pas forcément de dimension finie) muni d'une action de $\mathscr{G}_{\mathrm{K}}$. (Les actions de φ, N et la filtration sur un produit tensoriel étant données par les formules naturelles $\mathrm{N}(a \otimes d) = \mathrm{N}(a) \otimes d + a \otimes \mathrm{N}(d)$, $\varphi(a \otimes d) = \varphi(a) \otimes \varphi(d)$ et $(\mathbf{B}_{\mathrm{dR}} \otimes_{\mathrm{K}} \mathrm{D}_{\mathrm{K}})^{i} = \sum_{j \in \mathbf{Z}} \mathbf{B}_{\mathrm{dR}}^{i-j} \otimes \mathrm{D}_{\mathrm{K}}^{j}$.)

Si D est un (φ, N)-module filtré sur K, on peut lui associer deux invariants numériques $t_{\mathrm{N}}(\mathrm{D})$ et $t_{\mathrm{H}}(\mathrm{D})$ définis par

$$t_{\mathrm{N}}(\mathrm{D}) = v_p(\det \varphi) \quad \text{et} \quad t_{\mathrm{H}}(\mathrm{D}) = \sum_{i \in \mathbf{Z}} i \cdot \dim_{\mathrm{K}}(\mathrm{D}_{\mathrm{K}}^{i}/\mathrm{D}_{\mathrm{K}}^{i+1}).$$

(Comme φ n'est que semi-linéaire, son déterminant dépend de la base dans laquelle il est calculé, mais sa valuation p-adique n'en dépend pas.) On dit que D est *admissible* si on a $t_{\mathrm{H}}(\mathrm{D}) = t_{\mathrm{N}}(\mathrm{D})$ et $t_{\mathrm{H}}(\mathrm{D}') \leqslant t_{\mathrm{N}}(\mathrm{D}')$ pour tout sous-F-espace vectoriel D' de D stable par φ et N (D'_{K} étant muni de la filtration induite).

5.2.5. Construction des représentations semi-stables. — Le théorème suivant fournit une description concrète des représentations semi-stables et donc (modulo la conjecture de monodromie de Fontaine) des représentations de de Rham.

THÉORÈME 5.7 ([29, 28]). — *La catégorie des représentations semi-stables de $\mathscr{G}_{\mathrm{K}}$ est équivalente à celle des (φ, N)-modules filtrés admissibles sur* K. *Plus précisément, si* V *est une représentation semi-stable, alors $\mathbf{D}_{\mathrm{st}}(\mathrm{V})$ est un (φ, N)-module filtré admissible sur* K *et $\mathbf{V}_{\mathrm{st}}(\mathbf{D}_{\mathrm{st}}(\mathrm{V})) = \mathrm{V}$ et, réciproquement, si* D *est un (φ, N)-module filtré admissible sur* K, *alors $\mathbf{V}_{\mathrm{st}}(\mathrm{D})$ est une représentation semi-stable de $\mathscr{G}_{\mathrm{K}}$ et $\mathbf{D}_{\mathrm{st}}(\mathbf{V}_{\mathrm{st}}(\mathrm{D})) = \mathrm{D}$.*

De plus, cette équivalence de catégories induit une équivalence entre la catégorie des représentations cristallines et celle des φ-modules filtrés admissibles sur K.

On dispose de deux descriptions des représentations semi-stables : l'une, donnée par le théorème ci-dessus, en termes de (φ, N)-modules filtrés admissibles sur K et l'autre en termes de (φ, Γ)-modules étales sur $\mathbf{B}^{\dagger}_{\mathrm{K}}$. Une question naturelle qui se pose est : « comment passe-t-on de l'une à l'autre ? ». Dans un sens, en utilisant le fait que $(\widetilde{\mathbf{B}}^{\dagger}_{\mathrm{rig,K}}, (\mathbf{B}^{\dagger}_{\mathrm{rig,K},n})_{n \in \mathbf{N}})$ vérifie la propriété de Tate-Sen (TS2), on obtient le résultat suivant qui montre comment retrouver les actions de φ et N sur $\mathbf{D}_{\mathrm{st}}(\mathrm{V})$ à partir de $\mathbf{D}^{\dagger}(\mathrm{V})$; couplé avec la proposition 5.6, il fournit une description du (φ, N)-module filtré $\mathbf{D}_{\mathrm{st}}(\mathbf{V}^{\dagger}(\mathrm{D}))$ si D est un (φ, Γ)-module étale sur $\mathbf{B}^{\dagger}_{\mathrm{K}}$.

PROPOSITION 5.8 ([6]). — *Si* V *est une $\mathbf{Q}_p$-représentation de $\mathscr{G}_{\mathrm{K}}$, alors*

$$\mathbf{D}_{\mathrm{cris}}(\mathrm{V}) = \Big(\mathbf{B}^{\dagger}_{\mathrm{rig,K}}[\frac{1}{t}] \otimes_{\mathbf{B}^{\dagger}_{\mathrm{K}}} \mathbf{D}^{\dagger}(\mathrm{V})\Big)^{\Gamma_{\mathrm{K}}} \quad \text{et} \quad \mathbf{D}_{\mathrm{st}}(\mathrm{V}) = \Big(\mathbf{B}^{\dagger}_{\log,\mathrm{K}}[\frac{1}{t}] \otimes_{\mathbf{B}^{\dagger}_{\mathrm{K}}} \mathbf{D}^{\dagger}(\mathrm{V})\Big)^{\Gamma_{\mathrm{K}}}.$$

Voir [77] pour des résultats de nature similaire. Le vrai problème est d'aller dans l'autre sens car il est nettement plus facile de construire à la main un (φ, N)-module filtré admissible qu'un (φ, Γ)-module. En d'autre termes, comment décrire les matrices de φ et γ sur $\mathbf{D}^{\dagger}(\mathbf{V}_{\mathrm{st}}(\mathrm{D}))$ à partir de D, si D est un (φ, N)-module filtré admissible sur K. Ce n'est pas une question complètement gratuite car on peut lire sur un (φ, Γ)-module des propriétés de la représentation modulo p qui sont fort mystérieuses sur le (φ, N)-module filtré. Une réponse « raisonnable » à cette question aurait des applications à la théorie des déformations des représentations galoisiennes.

Le seul cas où l'on ait une réponse satisfaisante est le cas [78] où $\mathrm{K} = \mathrm{F}$ est non ramifié, et où $\mathbf{V}_{\mathrm{st}}(\mathrm{D})$ est une représentation cristalline à poids de Hodge-Tate compris entre 0 et $p-1$. (Il s'agit d'une démonstration des résultats de Fontaine et Laffaille [43] via la théorie des (φ, Γ)-modules.) D'autre part, dans le cas où $\mathrm{K} = \mathrm{F}$ est non ramifié, et où $\mathbf{V}_{\mathrm{st}}(\mathrm{D})$ est une représentation cristalline (sans hypothèse sur les poids de Hodge-Tate), on sait [25] que l'on peut se débrouiller pour que les matrices de φ et γ soient à coefficients dans $\mathscr{O}_{\mathrm{F}}[[\pi]]$.

5.3. Représentations galoisiennes et équations différentielles

5.3.1. L'opérateur de Sen. — Si on utilise la méthode de Sen et la proposition 4.1, on obtient le résultat suivant qui est précisément le cas qu'avait considéré Sen [69].

THÉORÈME 5.9 ([69]). — *Si $d \geqslant 1$, les applications naturelles*

$$\varinjlim \mathrm{H}^1(\Gamma_{\mathrm{K}}, \mathrm{GL}_d(\mathrm{K}_n)) \longrightarrow \mathrm{H}^1(\Gamma_{\mathrm{K}}, \mathrm{GL}_d(\widehat{\mathrm{K}}_{\infty})) \longrightarrow \mathrm{H}^1(\mathscr{G}_{\mathrm{K}}, \mathrm{GL}_d(\mathrm{C})),$$

induites par l'inflation de Γ_{K} à $\mathscr{G}_{\mathrm{K}}$ et les inclusions $\mathrm{K}_n \subset \widehat{\mathrm{K}}_{\infty} \subset \mathrm{C}$, sont des bijections.

Ce résultat permet de montrer que, si V est une représentation p-adique de $\mathscr{G}_{\mathrm{K}}$ et n est assez grand, alors $\mathrm{C} \otimes_{\mathbf{Q}_p} \mathrm{V}$ possède un sous-K_n-espace vectoriel fixe par $\mathscr{H}_{\mathrm{K}}$, stable par $\mathscr{G}_{\mathrm{K}}$, de dimension $\dim_{\mathbf{Q}_p} \mathrm{V}$ et engendrant $\mathrm{C} \otimes_{\mathbf{Q}_p} \mathrm{V}$. Il peut exister plusieurs sous-K_n-espaces vectoriels de $\mathrm{C} \otimes_{\mathbf{Q}_p} \mathrm{V}$ vérifiant ces propriétés, mais la proposition 3.6 et la remarque 3.7 nous en fournissent un privilégié noté $\mathbf{D}_{\mathrm{Sen},n}(\mathrm{V})$; de plus, il existe k indépendant de n tel que, si D' est un autre sous-K_n-espace vectoriel de $\mathrm{C} \otimes_{\mathbf{Q}_p} \mathrm{V}$ vérifiant ces propriétés, alors $\mathrm{K}_{n+k} \otimes_{\mathrm{K}_n} \mathrm{D}' = \mathrm{K}_{n+k} \otimes_{\mathrm{K}_n} \mathbf{D}_{\mathrm{Sen},n}(\mathrm{V})$. On montre alors facilement que $\frac{\gamma-1}{\chi(\gamma)-1}$ vu comme opérateur $\mathbf{Q}_p$-linéaire de $\mathbf{D}_{\mathrm{Sen},n}(\mathrm{V})$ tend vers une limite quand $\gamma \in \Gamma_{\mathrm{K}}$ tend vers 1. On obtient de la sorte [69] un opérateur Θ_{V} qui est K_n-linéaire et peut être vu comme un « poids de Hodge-Tate généralisé » :

PROPOSITION 5.10. — *Si* V *est une représentation de Hodge-Tate de $\mathscr{G}_{\mathrm{K}}$, alors Θ_{V} est diagonalisable et ses valeurs propres sont les poids de Hodge-Tate de* V (avec multiplicité). *Réciproquement, si Θ_{V} est diagonalisable et ses valeurs propres sont des entiers, alors* V *est de Hodge-Tate. En particulier,* V *est* C*-admissible si et seulement si $\Theta_{\mathrm{V}} = 0$.*

D'autre part, Θ_{V} peut être vu comme un élément de $\mathrm{C} \otimes_{\mathbf{Q}_p} \mathrm{End}(\mathrm{V})$ et Sen a démontré le résultat suivant :

THÉORÈME 5.11 ([68, 69]). — *La sous-algèbre de Lie de* $\mathrm{End}(\mathrm{V})$ *engendrée par les logarithmes des éléments du sous-groupe d'inertie de* $\mathscr{G}_{\mathrm{K}}$ *est la plus petite sous-algèbre de Lie de* $\mathrm{End}(\mathrm{V})$ *définie sur* $\mathbf{Q}_p$ *dont les* C*-points contiennent* Θ_{V}. *En particulier,* $\Theta_{\mathrm{V}} = 0$ (i.e. V est C-admissible) *si et seulement si le sous-groupe d'inertie de* $\mathscr{G}_{\mathrm{K}}$ *agit à travers un quotient fini.*

5.3.2. Le (φ, ∇)*-module associé à une représentation galoisienne.* — On note ∂ la dérivation $(1+\pi)d/d\pi$ de $\mathbf{B}_{\mathrm{F}}^{\dagger}$. Cette dérivation s'étend de manière unique à $\mathbf{B}_{\mathrm{K}}^{\dagger}$ pour toute extension finie K de F et on a $\partial \circ \varphi = p\,\varphi \circ \partial$.

PROPOSITION 5.12 ([6]). — *Si* $x \in \mathbf{B}_{\mathrm{rig,K}}^{\dagger}$, *alors* $\lim_{\gamma\to 1,\ \gamma\in\Gamma_{\mathrm{K}}} \frac{\gamma-1}{\chi(\gamma)-1}x = \nabla x$, *où* $\nabla = t\partial = td/dt$ *est une dérivation de* $\mathbf{B}_{\mathrm{rig,K}}^{\dagger}$.

Démonstration. — L'existence de la limite résulte d'une étude directe de l'action de Γ_{K} ; le reste est un calcul immédiat.

Ce qui précède correspond au cas de la représentation triviale ; dans le cas général, on a le résultat suivant :

PROPOSITION 5.13 ([6]). — *Si* V *est une* $\mathbf{Q}_p$*-représentation de* $\mathscr{G}_{\mathrm{K}}$, *la famille d'opérateurs* $\frac{\gamma-1}{\chi(\gamma)-1}$ *tend, quand* $\gamma \in \Gamma_{\mathrm{K}}$ *tend vers* 1, *vers une connexion* ∇_{V} *au-dessus de* ∇ *sur* $\mathbf{D}_{\mathrm{rig}}^{\dagger}(\mathrm{V}) = \mathbf{B}_{\mathrm{rig,K}}^{\dagger} \otimes_{\mathbf{B}_{\mathrm{K}}^{\dagger}} \mathbf{D}^{\dagger}(\mathrm{V})$. *De plus, si* n *est assez grand, cette connexion laisse stable le sous-*$\mathbf{B}_{\mathrm{K}}^{]0,p^{-n}]}$*-module* $\mathbf{D}^{]0,p^{-n}]}(\mathrm{V}) = \mathbf{B}_{\mathrm{K}}^{]0,p^{-n}]} \otimes_{\mathbf{B}_{\mathrm{K}}^{(0,p^{-n}]}} \mathbf{D}^{(0,p^{-n}]}(\mathrm{V})$ *de* $\mathbf{D}_{\mathrm{rig}}^{\dagger}(\mathrm{V})$.

On a donc associé à toute $\mathbf{Q}_p$-représentation V de $\mathscr{G}_{\mathrm{K}}$ un module avec connexion et structure de Frobenius sur l'anneau de Robba $\mathbf{B}_{\mathrm{rig,K}}^{\dagger}$. Ce n'est malheureusement pas suffisant pour pouvoir utiliser les résultats généraux sur les équations différentielles p-adiques car t n'est pas inversible dans $\mathbf{B}_{\mathrm{rig,K}}^{\dagger}$, ce qui se traduit par l'existence potentielle d'une infinité de singularités dans toute couronne du type $0 < v_p(\mathrm{T}) \leqslant r$. Comme t n'a que des pôles simples, ces singularités sont régulières et on peut « localiser en $\varepsilon^{(n)} - 1$ » pour les étudier.

5.3.3. Localisation. — La proposition 3.3 permet de déduire du théorème 5.9 le résultat suivant :

PROPOSITION 5.14. — *Si* $d \geqslant 1$, *les applications naturelles*

$$\varinjlim \mathrm{H}^1(\Gamma_{\mathrm{K}}, \mathrm{GL}_d(\mathrm{K}_n[[t]])) \longrightarrow \mathrm{H}^1(\Gamma_{\mathrm{K}}, \mathrm{GL}_d((\mathbf{B}_{\mathrm{dR}}^{+})^{\mathscr{H}_{\mathrm{K}}})) \longrightarrow \mathrm{H}^1(\mathscr{G}_{\mathrm{K}}, \mathrm{GL}_d(\mathbf{B}_{\mathrm{dR}}^{+})),$$

induites par l'inflation de Γ_{K} *à* $\mathscr{G}_{\mathrm{K}}$ *et les inclusions* $\mathrm{K}_n[[t]] \subset (\mathbf{B}_{\mathrm{dR}}^{+})^{\mathscr{H}_{\mathrm{K}}} \subset \mathbf{B}_{\mathrm{dR}}^{+}$, *sont des bijections.*

Comme précédemment, on peut utiliser cette proposition pour définir, si n est assez grand, un élément privilégié $\mathbf{D}_{\mathrm{Dif},n}^{+}(\mathrm{V})$ parmi les sous-$\mathrm{K}_n[[t]]$-modules libres de rang d de $\mathbf{B}_{\mathrm{dR}}^{+} \otimes_{\mathbf{Q}_p} \mathrm{V}$, fixes par $\mathscr{H}_{\mathrm{K}}$, stables par $\mathscr{G}_{\mathrm{K}}$ et engendrant $\mathbf{B}_{\mathrm{dR}}^{+} \otimes_{\mathbf{Q}_p} \mathrm{V}$.

PROPOSITION 5.15. — (i) *L'image de* $\mathbf{D}^+_{\mathrm{Dif},n}(\mathrm{V})$ *par* $\theta : \mathbf{B}^+_{\mathrm{dR}} \otimes_{\mathbf{Q}_p} \mathrm{V} \to \mathrm{C} \otimes_{\mathbf{Q}_p} \mathrm{V}$ *est* $\mathbf{D}_{\mathrm{Sen},n}(\mathrm{V})$.

(ii) *La famille d'opérateurs* $\frac{\gamma-1}{\chi(\gamma)-1}$ *tend, quand* $\gamma \in \Gamma_{\mathrm{K}}$ *tend vers* 1, *vers une connexion* ∇_{V} *au-dessus de* $\nabla = td/dt$ *sur* $\mathbf{D}^+_{\mathrm{Dif},n}(\mathrm{V})$ *et on a* $\theta \circ \nabla_{\mathrm{V}} = \Theta_{\mathrm{V}}$.

(iii) *De plus,* $\mathbf{D}^+_{\mathrm{Dif},n}(\mathrm{V}) = \mathrm{K}_n[[t]] \otimes_{\mathbf{B}_{\mathrm{K}}^{]0,p^{-n}]}} \mathbf{D}^{]0,p^{-n}]}(\mathrm{V})$ *et cette identification commute à* ∇_{V}.

Le (iii) montre que $\mathbf{D}^+_{\mathrm{Dif},n}(\mathrm{V})$ est le localisé du (φ,∇)-module $\mathbf{D}^{]0,p^{-n}]}(\mathrm{V})$ en $\varepsilon^{(n)}-1$ et le (ii) montre que le résidu de la connexion ∇_{V} en ce point est Θ_{V}. Pour pouvoir se débarrasser de ces singularités, il faut au moins que Θ_{V} soit diagonalisable à valeurs propres entières (i.e. que V soit de Hodge-Tate), mais ce n'est pas suffisant (voir ci-dessous) ; en tout cas, ce résultat montre que le (φ,∇)-module associé à une représentation galoisienne admet en général une infinité de singularités « non enlevables ».

PROPOSITION 5.16 ([6, 42]). — (i) $\mathbf{D}^+_{\mathrm{Dif},n}(\mathrm{V})$ *est sans singularité si et seulement si* V *est* C*-admissible.*

(ii) $\mathbf{D}^+_{\mathrm{Dif},n}(\mathrm{V})$ *est à singularités apparentes* (la connexion a une base de sections horizontales dans $\mathbf{D}_{\mathrm{Dif},n}(\mathrm{V}) = \mathbf{D}^+_{\mathrm{Dif},n}(\mathrm{V})[\frac{1}{t}]$) *si et seulement si* V *est de de Rham.*

Le (i) est une évidence et le (ii) se démontre en utilisant la proposition 5.6. Remarquons que si V est de de Rham, alors $\mathbf{D}_{\mathrm{Dif},n}(\mathrm{V}) = \mathrm{K}_n((t)) \otimes_{\mathrm{K}} \mathbf{D}_{\mathrm{dR}}(\mathrm{V})$ et donc $\mathbf{D}_{\mathrm{Dif},n}(\mathrm{V})$ possède un sous-$\mathrm{K}_n[[t]]$-réseau sur lequel on peut diviser ∇_{V} par t, à savoir $\mathrm{K}_n[[t]] \otimes_{\mathrm{K}} \mathbf{D}_{\mathrm{dR}}(\mathrm{V})$.

5.3.4. Le (φ,∂)*-module attaché à une représentation de de Rham.* — D'après le paragraphe précédent, si V est de de Rham, les singularités de ∇_{V} ne sont qu'apparentes d'un point de vue local ; le problème est donc d'arriver à les supprimer simultanément ou, ce qui revient au même, de trouver un réseau sur lequel on puisse diviser ∇_{V} par t. Le lemme suivant fournit la clé de ce que l'on doit faire.

LEMME 5.17. — x *est divisible par* t *dans* $\mathbf{B}^{\dagger}_{\mathrm{rig,K}}$ *si et seulement si* $\iota_n(x) \in t\mathbf{B}^+_{\mathrm{dR}}$, *pour tout* n *assez grand.*

Soit alors $\mathbf{D}^{\dagger}_{\mathrm{rig}}(\mathrm{V}) = \mathbf{B}^{\dagger}_{\mathrm{rig,K}} \otimes_{\mathbf{B}^{\dagger}_{\mathrm{K}}} \mathbf{D}^{\dagger}(\mathrm{V})$, et soient $\mathbf{N}^+_{\mathrm{dR}}(\mathrm{V}) = \mathbf{B}^+_{\mathrm{dR}} \otimes_{\mathrm{K}} \mathbf{D}_{\mathrm{dR}}(\mathrm{V})$ et $\mathbf{N}^{\dagger}_{\mathrm{rig}}(\mathrm{V})$ l'ensemble des $x \in \mathbf{D}^{\dagger}_{\mathrm{rig}}(\mathrm{V})[\frac{1}{t}]$, tels que $\iota_n(x) \in \mathbf{N}^+_{\mathrm{dR}}(\mathrm{V})$ si n est assez grand (dépendant de x). Ce $\mathbf{B}^{\dagger}_{\mathrm{rig,K}}$-module est fermé et, si a et b sont respectivement le plus grand et le plus petit poids de Hodge-Tate de V, alors $\mathbf{N}^{\dagger}_{\mathrm{rig}}(\mathrm{V})$ contient $t^{-b}\mathbf{D}^{\dagger}_{\mathrm{rig}}(\mathrm{V})$ et est contenu dans $t^{-a}\mathbf{D}^{\dagger}_{\mathrm{rig}}(\mathrm{V})$; c'est donc un $\mathbf{B}^{\dagger}_{\mathrm{rig,K}}$-module libre de rang d (cf. prop. 1.1). Un peu plus de travail permet d'obtenir :

THÉORÈME 5.18 ([6]). — *Si* V *est une* $\mathbf{Q}_p$*-représentation de de Rham de* $\mathscr{G}_{\mathrm{K}}$ *de dimension* d, *alors* $\mathbf{D}^{\dagger}_{\mathrm{rig}}(\mathrm{V})[\frac{1}{t}]$ *contient un unique sous-*$\mathbf{B}^{\dagger}_{\mathrm{rig,K}}$*-module* $\mathbf{N}^{\dagger}_{\mathrm{rig}}(\mathrm{V})$ *libre de rang* d *stable par* $\partial_{\mathrm{V}} = t^{-1}\nabla_{\mathrm{V}}$.

De plus, ce module est stable par φ et Γ_K et on a

$$\mathbf{B}^{\dagger}_{\mathrm{rig,K}}[\frac{1}{t}] \otimes_{\mathbf{B}^{\dagger}_{\mathrm{rig,K}}} \mathbf{N}^{\dagger}_{\mathrm{rig}}(V) = \mathbf{B}^{\dagger}_{\mathrm{rig,K}}[\frac{1}{t}] \otimes_{\mathbf{B}^{\dagger}_{K}} \mathbf{D}^{\dagger}(V).$$

Ce théorème admet comme corollaire la conjecture de monodromie de Fontaine (modulo la conjecture de monodromie de Crew) : si V est de de Rham, $\mathbf{N}^{\dagger}_{\mathrm{rig}}(V)$ est un (φ, ∂)-module sur l'anneau de Robba $\mathbf{B}^{\dagger}_{\mathrm{rig,K}}$. Un tel module étant quasi-unipotent d'après la conjecture de Crew, il existe une extension finie L de K telle que $\mathbf{B}^{\dagger}_{\mathrm{log,L}}[\frac{1}{t}] \otimes_{\mathbf{B}^{\dagger}_{K}} \mathbf{D}^{\dagger}(V)$ admette une base constituée de sections horizontales pour ∇_V. Vu le lien entre Γ_L et ∇_V, cela implique que Γ_L agit à travers un quotient fini sur l'espace vectoriel des sections horizontales, et la proposition 5.8 permet de montrer que V est semi-stable en tant que représentation de $\mathscr{G}_{L_n}$ si n est assez grand.

Remarque 5.19 ([6]). — Si V est C-admissible, alors $\mathbf{N}^{\dagger}_{\mathrm{rig}}(V) = \mathbf{D}^{\dagger}(V)$. Ceci permet de déduire du théorème de Tsuzuki (cor. 2.10) l'équivalence (cf. th. 5.11)

« V est C-admissible » $\Leftrightarrow$ « l'inertie de $\mathscr{G}_K$ agit à travers un quotient fini ».

POSTFACE (JANVIER 2003)

Depuis la rédaction de cet exposé, le sujet a connu quelques prolongements :

— Kedlaya a annoncé, comme conséquences de la conjecture de Crew, des résultats généraux de finitude [55] pour la cohomologie rigide avec coefficients, et une preuve [56] des conjectures de Weil par voie p-adique (la possibilité d'obtenir une telle preuve avait été suggérée par Mebkhout [62]).

— En faisant du mécano avec certains des arguments de Berger [5, 6] et Kedlaya [57], on peut obtenir [27] une démonstration de la conjecture de Fontaine évitant complètement la théorie des équations différentielles p-adiques ; les détails seront publiés ailleurs.

Remerciements. — La rédaction de cet exposé a bénéficié des remarquables conditions de travail offertes par le K.I.A.S. de Séoul, et je voudrais remercier cette institution et Minhyong Kim de leur hospitalité.

RÉFÉRENCES

[1] Y. André – « Filtrations de type Hasse-Arf et monodromie p-adique », *Invent. Math.* **148** (2002), p. 285–317.

[2] ———, « Représentations galoisiennes et opérateurs de Bessel p-adiques », *Ann. Inst. Fourier* (*Grenoble*) **52** (2002), p. 779–808.

[3] J. Ax – « Zeroes of polynomials over local fields. The Galois action », *J. Algebra* **15** (1970), p. 417–428.

[4] D. Benois – « On Iwasawa theory of crystalline representations », *Duke Math. J.* **104** (2000), p. 211–267.

[5] L. Berger – « Représentations p-adiques et équations différentielles », Thèse, Université Paris VI, 2001.

[6] ———, « Représentations p-adiques et équations différentielles », *Invent. Math.* **148** (2002), p. 219–284.

[7] P. Berthelot – « Finitude et pureté cohomologique en cohomologie rigide », *Invent. Math.* **128** (1997), p. 329–377.

[8] D. Bertrand – « Groupes algébriques et équations différentielles linéaires », in *Séminaire Bourbaki 1991/92*, Astérisque, vol. 206, Soc. Math. France, 1992, exp. 750, p. 183–204.

[9] N. Bourbaki – *Algèbre commutative,* IX-X, Masson, Paris, 1983.

[10] C. Breuil – « Une application du corps des normes », *Compositio Math.* **117** (1999), p. 189–203.

[11] F. Cherbonnier – « Représentations p-adiques surconvergentes », Thèse, Université d'Orsay, 1996.

[12] F. Cherbonnier & P. Colmez – « Représentations p-adiques surconvergentes », *Invent. Math.* **133** (1998), p. 581–611.

[13] ———, « Théorie d'Iwasawa des représentations p-adiques d'un corps local », *J. Amer. Math. Soc.* **12** (1999), p. 241–268.

[14] B. Chiarellotto & G. Christol – « Overconvergent isocrystals and F-isocrystals », *Compositio Math.* **100** (1996), p. 77–99.

[15] G. Christol – « About a Tsuzuki theorem », in *p-adic functional analysis* (*Ioannina, 2000*), Lecture Notes in Pure and Appl. Math., vol. 222, Dekker, New York, 2001, p. 63–74.

[16] G. Christol & B. Dwork – « Modules différentiels sur des couronnes », *Ann. Inst. Fourier* (*Grenoble*) **44** (1994), p. 689–720.

[17] G. Christol & Z. Mebkhout – « Sur le théorème de l'indice des équations différentielles p-adiques I », *Ann. Inst. Fourier* (*Grenoble*) **43** (1993), p. 1545–1574.

[18] ———, « Sur le théorème de l'indice des équations différentielles p-adiques II », *Ann. of Math.* **146** (1997), p. 345–410.

[19] ———, « Sur le théorème de l'indice des équations différentielles p-adiques III », *Ann. of Math.* **151** (2000), p. 385–457.

[20] ———, « Sur le théorème de l'indice des équations différentielles p-adiques IV », *Invent. Math.* **143** (2001), p. 629–672.

[21] ———, « Équations différentielles p-adiques et coefficients p-adiques sur les courbes », in *Périodes p-adiques II*, Astérisque, vol. 279, Soc. Math. France, 2002, p. 125–183.

[22] G. Christol & P. Robba – *Équations différentielles p-adiques. Applications aux sommes exponentielles*, Actualités Mathématiques, Hermann, Paris, 1994.

[23] P. Colmez – « Sur un résultat de Shankar Sen », *C. R. Acad. Sci. Paris Sér. I Math.* **318** (1994), p. 983–985.

[24] ———, « Représentations p-adiques d'un corps local », in *Proceedings of the International Congress of Mathematicians II* (*Berlin 1998*), Doc. Mat. Extra, vol. II, Deutsche Math. Verein., 1998, p. 153–162.

[25] ———, « Représentations cristallines et représentations de hauteur finie », *J. reine Angew. Math.* **514** (1999), p. 119–143.

[26] ———, « Fonctions L p-adiques », in *Séminaire Bourbaki 1998/99*, Astérisque, vol. 266, Soc. Math. France, 2000, exp. 851, p. 21–58.

[27] ———, exposé à Orsay, novembre 2001.

[28] ———, « Espaces de Banach de dimension finie », *J. Inst. Math. Jussieu* **1** (2002), p. 331–439.

[29] P. Colmez & J.-M. Fontaine – « Construction des représentations p-adiques semi-stables », *Invent. Math.* **140** (2000), p. 1–43.

[30] R. Crew – « F-isocrystals and p-adic representations », in *Algebraic Geometry* (*Bowdoin 1985*), Proc. Symp. Pure Math., vol. XLVI (2), Amer. Math. Soc., 1987, p. 111–138.

[31] ———, « Finiteness theorems for the cohomology of an overconvergent isocrystal on a curve », *Ann. scient. Éc. Norm. Sup.* 4^{e} *série* **31** (1998), p. 717–763.

[32] ———, « Canonical extensions, irregularities, and the Swan conductor », *Math. Ann.* **316** (2000), p. 19–37.

[33] B. Dwork – « On exponents of p-adic differential modules », *J. reine Angew. math.* **484** (1997), p. 85–126.

[34] B. Dwork, N. Gerotto & F. Sullivan – *An introduction to* G*-functions*, Ann. of Math. Studies, vol. 133, Princeton Univ. Press, 1994.

[35] G. Faltings – « Almost étale extensions », in *Périodes p-adiques II*, Astérisque, vol. 279, Soc. Math. France, 2002, p. 185–270.

[36] J.-M. Fontaine – « Sur certains types de représentations p-adiques du groupe de Galois d'un corps local ; construction d'un anneau de Barsotti-Tate », *Ann. of Math.* **115** (1982), p. 529–577.

[37] ———, « Cohomologie de de Rham, cohomologie cristalline et représentations p-adiques », in *Algebraic geometry* (*Tokyo/Kyoto,1982*), SLN, vol. 1016, Springer, 1983, p. 86–108.

[38] ———, « Représentations p-adiques des corps locaux », in *The Grothendieck Festschrift*, vol. II, Birkhäuser, Boston, 1991, p. 249–309.

[39] ———, « Le corps des périodes p-adiques », in *Périodes p-adiques*, Astérisque, vol. 223, Soc. Math. France, 1994, exp. II, p. 59–102.

[40] ———, « Représentations p-adiques semi-stables », in *Périodes p-adiques*, Astérisque, vol. 223, Soc. Math. France, 1994, exp. III, p. 113–184.

[41] ———, exposé à Orsay, 1998.

[42] ———, « Arithmétique des représentations galoisiennes p-adiques », Astérisque, à paraître.

[43] J.-M. Fontaine & G. Laffaille – « Construction de représentations p-adiques », *Ann. scient. Éc. Norm. Sup.* 4^{e} *série* **15** (1982), p. 547–608.

[44] J.-M. Fontaine & J.-P. Wintenberger – « Le corps des normes de certaines extensions algébriques de corps locaux », *C. R. Acad. Sci. Paris Sér. I Math.* **288** (1979), p. 367–370.

[45] A. Grothendieck – Lettres des 24/9/64 p. 183 et 3-5/10/64 p. 204, in *Correspondance Grothendieck-Serre*, Documents Mathématiques, vol. 2, Soc. Math. France, 2001.

[46] G. Henniart – « La conjecture de Langlands locale numérique pour GL(n) », *Ann. scient. Éc. Norm. Sup.* 4$^{\text{e}}$ *série* **21** (1988), p. 497–544.

[47] L. Herr – « Sur la cohomologie galoisienne des corps p-adiques », *Bull. Soc. math. France* **126** (1998), p. 563–600.

[48] ———, « Une approche nouvelle de la dualité de Tate », *Math. Ann.* **320** (2001), p. 307–337.

[49] O. Hyodo – « $\mathrm{H}^1_g(\mathrm{K},\mathrm{V}) = \mathrm{H}^1_{\mathrm{st}}(\mathrm{K},\mathrm{V})$ », in *Proceedings of a symposium on arithmetic geometry* (K. Kato, M. Kurihara & T. Saito, éds.), Univ. Tokyo, 1991.

[50] L. Illusie – « Autour du théorème de monodromie locale », in *Périodes p-adiques*, Astérisque, vol. 223, Soc. Math. France, 1994, exp. I, p. 9–57.

[51] A.J. de Jong – « Smoothness, semi-stability and alterations », *Publ. Math. Inst. Hautes Études Sci.* **83** (1996), p. 51–93.

[52] N. Katz – « p-adic properties of modular schemes and modular forms », in *Modular functions of one variable III*, SLN, vol. 350, Springer, 1973, p. 69–190.

[53] K. Kedlaya – « Full faithfulness for overconvergent F-isocrystals », preprint, 2001.

[54] ———, « Semi-stable reduction for overconvergent F-isocrystals on a curve », preprint, 2001.

[55] ———, « Finiteness of rigid cohomology with coefficients », preprint, 2002.

[56] ———, « Fourier transforms and p-adic Weil II », preprint, 2002.

[57] ———, « A p-adic local monodromy theorem », Ann. of Maths, à paraître.

[58] M. Lazard – « Les zéros des fonctions analytiques d'une variable sur un corps valué complet », *Publ. Math. Inst. Hautes Études Sci.* **14** (1962), p. 47–75.

[59] F. Loeser – « Exposants p-adiques et théorèmes d'indice pour les équations différentielles p-adiques (d'après G. Christol et Z. Mebkhout) », in *Séminaire Bourbaki 1996/97*, Astérisque, vol. 245, Soc. Math. France, 1997, exp. 822, p. 57–81.

[60] S. Matsuda – « Local indices of p-adic differential operators corresponding to Artin-Schreier-Witt coverings », *Duke Math. J.* **77** (1995), p. 607–625.

[61] ———, « Katz correspondence for quasi-unipotent overconvergent isocrystals », *Compositio Math.* **134** (2002), p. 1–34.

[62] Z. Mebkhout – « Sur le théorème de finitude de la cohomologie p-adique d'une variété affine non singulière », *Amer. J. Math.* **119** (1997), p. 1027–1081.

[63] ———, « Analogue p-adique du théorème de Turrittin et le théorème de la monodromie p-adique », *Invent. Math.* **148** (2002), p. 319–351.

[64] P. Robba – « On the index of p-adic differential operators I », *Ann. of Math.* **101** (1975), p. 280–316.

[65] ———, « On the index of p-adic differential operators II », *Duke Math. J.* **43** (1976), p. 19–31.

[66] ———, « On the index of p-adic differential operators III, applications to twisted exponential sums », in *Cohomologie p-adique*, Astérisque, vol. 119-120, Soc. Math. France, 1984, p. 191–266.

[67] ———, « Indice d'un opérateur différentiel p-adique IV. Cas des systèmes. Mesure de l'irrégularité dans un disque », *Ann. Inst. Fourier* (*Grenoble*) **35** (1985), p. 13–55.

[68] S. Sen – « Lie algebras of Galois groups arising from Hodge-Tate modules », *Ann. of Math.* **97** (1973), p. 160–170.

[69] ———, « Continuous cohomology and p-adic Galois representations », *Invent. Math.* **62** (1980/81), p. 89–116.

[70] J.-P. Serre – *Corps locaux*, 2ème éd., Publications de l'Université de Nancago, vol. VIII, Hermann, Paris, 1968.

[71] J. Tate – « p-divisible groups », in *Proc. Conf. Local Fields* (*Driebergen, 1966*), Springer, Berlin, 1967, p. 158–183.

[72] T. Tsuji – « p-adic étale cohomology and crystalline cohomology in the semi-stable reduction case », *Invent. Math.* **137** (1999), p. 233–411.

[73] ———, « Semi-stable conjecture of Fontaine-Jannsen : a survey », in *Périodes p-adiques II*, Astérisque, vol. 279, Soc. Math. France, 2002, p. 323–370.

[74] N. Tsuzuki – « Finite local monodromy of overconvergent unit-root F-isocrystal on a curve », *Amer. J. Math.* **120** (1998), p. 1165–1190.

[75] ———, « Slope filtration of quasi-unipotent overconvergent F-isocrystals », *Ann. Inst. Fourier* (*Grenoble*) **48** (1998), p. 379–412.

[76] ———, « The local index and the Swan conductor », *Compositio Math.* **111** (1998), p. 245–288.

[77] N. Wach – « Représentations p-adiques potentiellement cristallines », *Bull. Soc. math. France* **124** (1996), p. 375–400.

[78] ———, « Représentations cristallines de torsion », *Compositio Math.* **108** (1997), p. 185–240.

[79] J.-P. Wintenberger – « Le corps des normes de certaines extensions infinies des corps locaux ; applications », *Ann. scient. Éc. Norm. Sup.* 4^{e} *série* **16** (1983), p. 59–89.

[80] P. Young – « Radii of convergence and index for p-adic differential operators », *Trans. Amer. Math. Soc.* **333** (1992), p. 769–785.

Pierre COLMEZ

Institut de mathématiques de Jussieu
4 place Jussieu
F-75005 PARIS
E-mail : colmez@math.jussieu.fr

Séminaire BOURBAKI Novembre 2001
54e année, 2001-2002, no 898, p. 103 à 115

ON THE $n!$-CONJECTURE

by **Claudio PROCESI**

1. MACDONALD POLYNOMIALS, THE POSITIVITY CONJECTURE

In 1988 Macdonald defined 2-parameter symmetric functions unifying the theory of Hall-Littlewood and Jack polynomials (*cf.* [M], [M1]).

We recall a variant of his construction (*cf.* [H1]). For a given positive integer n we want to define symmetric functions $\widetilde{H}_\lambda(X)$ indexed by partitions of n and with coefficients in $\mathbb{Q}(q,t)$. We use the fact that symmetric functions (in infinitely many variables) are polynomials in the Newton functions $\psi_k := \sum_i x_i^k$.

The $\widetilde{H}_\lambda(X)$ are implicitly defined using dominance order and plethystic transformations.

For a symmetric function $f(X)$ the plethystic transformation $f(X) \to f(X[1-q])$ is the unique morphism $\mathbb{Q}[\psi_1, \dots, \psi_m, \dots]$ to $\mathbb{Q}(q)[\psi_1, \dots, \psi_m, \dots]$ sending $\psi_k \to \psi_k(1-q^k)$.

The dominance order of partitions, is

$$(p_1, p_2, \dots, p_n) \geqslant (q_1, q_2, \dots, q_n), \iff \forall k \; p_1 + p_2 + \cdots + p_k \geqslant q_1 + q_2 + \cdots + q_k.$$

Finally for a partition λ its dual λ' is obtained exchanging rows and columns in its Young diagram.

1.1. Theorem (Macdonald[(1)]). — *There exist unique symmetric functions $\widetilde{H}_\lambda(X)$ satisfying:*

(1) *$\widetilde{H}_\lambda(X[1-q])$ lies in the vector space over $\mathbb{Q}(q,t)$ generated by the Schur functions $S_\mu(X)$, $\mu \geqslant \lambda$.*

(2) *$\widetilde{H}_\lambda(X[1-t])$ lies in the vector space over $\mathbb{Q}(q,t)$ generated by the Schur functions $S_\mu(X)$, $\mu \geqslant \lambda'$.*

(3) *In the expansion of $\widetilde{H}_\lambda(X)$ through Schur functions, the coefficient of S_n is* 1.

(1) I follow here Haiman's approach [H1].

POSITIVITY CONJECTURE. — *The coefficients of the expansion $\widetilde{H}_\lambda(X)$ through Schur functions are polynomials in q,t with coefficients positive integers.*

For further discussion and a guide through the literature we refer to [H1].

2. $n!$-CONJECTURE

Since the work of Frobenius, the connection between symmetric functions and representations of the symmetric group has been well understood. In particular it is useful to associate to the irreducible representation indexed by a partition λ the Schur function $S_\lambda(X)$. Extending this by linearity one has a linear isomorphism $\chi \to F(\chi)$ (called Frobenius character) between the space of characters of the symmetric group on n letters and the space of symmetric functions of degree n.

With this convention suppose we have a bigraded representation of the symmetric group $V_{i,j}$ and let $\chi_{i,j}$ be the corresponding bigraded character. Then we construct the 2-parameter symmetric function, called its bigraded Frobenius character:

$$\sum_{i,j} q^i t^j F(\chi_{i,j}).$$

So, to prove the positivity conjecture one should construct, for each partition λ, a bigraded representation whose bigraded Frobenius character is the Macdonald polynomial $\widetilde{H}_\lambda(X)$.

In 1991 Adriano Garsia and Mark Haiman, inspired by similar constructions for the simpler case of $q-$Kostka polynomials (*cf.* [GP]), proposed such a construction.

Let $R := \mathbb{C}[x_1, \dots, x_n; y_1, \dots, y_n]$ be the polynomial ring in $2n$ variables.

A partition λ of n will be always identified to a set of n points in the integral lattice.

The n pairs $\lambda := \{(i_h, j_h)\}$, $h = 1, \dots, n$ of numbers give (up to sign) the polynomial:

$$D_\lambda := \det(x_k^{i_h} y_k^{j_h}) \in R, \quad (h,k = 1, \dots, n).$$

Consider the R module structure on R setting:

$$p \cdot f := p\Big(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}; \frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_n}\Big) f$$

set:

$$I_\lambda := \{p \in R \mid p \cdot D_\lambda = 0\}, \quad V_\lambda := \{p \cdot D_\lambda \mid p \in R\},$$

$V_\lambda \cong R/I_\lambda$ is the space spanned by all the derivatives of the polynomial D_λ.

$n!$-CONJECTURE ([GH]). — $\dim R/I_\lambda = \dim V_\lambda = n!$

We let the symmetric group S_n act on $R := \mathbb{C}[x_1, \dots, x_n; y_1, \dots, y_n]$ by the diagonal action (or simultaneously on the x and y). Since D_λ is bihomogeneous and skew symmetric, it is clear that $V_\lambda = R/I_\lambda$ is a bigraded representation of the symmetric group so:

SECOND CONJECTURE ([GH]). — *The bigraded Frobenius character of V_λ is the Macdonald polynomial $\tilde{H}_\lambda(X)$.*

Both conjectures have now been proved by Mark Haiman, his final results are in [H2]. It has turned out that the most difficult part of the project has been to establish the $n!$-conjecture.

The proof of the $n!$-conjecture is based on a deep property of the Hilbert scheme of n-tuples of points in the plane. This connection has also allowed Haiman to solve other conjectures on diagonal harmonics as we shall explain at the end (*cf.* [H], [H3]).

In order to see how the Hilbert scheme enters, let us first make an elementary remark.

Define a linear form T_λ on R as:

$$T_\lambda(p) := (p \cdot D_\lambda)(0). \tag{2.1}$$

2.2. LEMMA. — $I_\lambda := \{p \in R \mid T_\lambda(pq) = 0,\ \forall q \in R\}$.

Proof. — A polynomial $p \in R$ is 0 if and only if $(q \cdot p)(0) = 0$ for every $q \in R$. Thus $p \cdot D_\lambda = 0$ if and only if $q \cdot (p \cdot D_\lambda)(0) = 0$, for every q. Since $q \cdot (p \cdot D_\lambda) = (qp) \cdot D_\lambda$ we have the claim. □

Remark that, if h, k is a pair of natural integers, external to the partition λ we have, for every i:

$$\frac{\partial}{\partial x_i}^h \frac{\partial}{\partial y_i}^k D_\lambda = 0$$

It follows that I_λ contains the monomials $x_i^h y_i^k$ for every such pair. In other words, set J_λ to be the ideal of $\mathbb{C}[x, y]$ generated by the monomials $x^h y^k$, $(h, k) \notin \lambda$ and $A_\lambda := \mathbb{C}[x, y]/J_\lambda$. A_λ has dimension n and, as basis, the monomials $x^i y^j$, $(i, j) \in \lambda$.

Finally, identifying $R = \mathbb{C}[x, y]^{\otimes n}$ we have that R/I_λ is a quotient of $A_\lambda^{\otimes n}$.

The linear form T_λ factors through $A_\lambda^{\otimes n}$ and it is antisymmetric. Any antisymmetric linear form factors through antisymmetrization $A_\lambda^{\otimes n} \to \wedge^n A_\lambda$. We have $\dim \wedge^n A_\lambda = 1$, since $\dim A_\lambda = n$, hence such a form up to scalar is unique.

Thus the $n!$-conjecture is equivalent to:

RANK CONJECTURE. — *The form $T_\lambda(pq)$ on $A_\lambda^{\otimes n}$ has rank $n!$.*

It has turned out to be too difficult to analyze directly the ideal J_λ, but rather one must work more globally on H_n, the Hilbert scheme of all ideals I of codimension n in $\mathbb{C}[x,y]$. H_n comes together with the universal family

$$\mathscr{F} := \{(I,p) \in H_n \times \mathbb{C}^2 \mid I(p) = 0\}. \tag{2.3}$$

The projection map $p : \mathscr{F} \to H_n$ is flat and $F := p_*\mathscr{O}_{\mathscr{F}}$ is the tautological vector bundle $F = \{(I, \mathbb{C}[x,y]/I)\}$, a bundle of algebras of dimension n over H_n.

A fundamental theorem of Fogarty [F] states that the Hilbert scheme is smooth of dimension $2n$, from which it follows easily that it gives a (crepant) resolution of singularities:

$$H_n \xrightarrow{\rho} \mathbb{C}^{2n}/S_n.$$

In the Hilbert scheme the ideals J_λ play a special role. In fact on $\mathbb{C}^2$ and hence on H_n, acts a two dimensional torus $\mathscr{T} := \{(\alpha,\beta)\}$ through $(\alpha,\beta)(x,y) := (\alpha x, \beta y)$. A fixed point is just a bihomogeneous ideal and one easily sees that these are exactly the ideals J_λ (indexed by partitions of n).

One uses always the fact that any $\mathscr{T}$ stable closed subset of H_n contains a fixed point.

Now we can globalize the rank conjecture. Consider the antisymmetrization $T : F^{\otimes n} \to \wedge^n F$ which induces a *form* $T(ab)$ on the bundle $F^{\otimes n}$.

2.4. Lemma. — *T has generically rank $n!$. The $n!$-conjecture is equivalent to the statement that the form T has always rank $n!$. In this case $F^{\otimes n}/Ker(T)$ is a bundle of algebras each carrying the regular representation of S_n.*

Proof. — In a generic point of the Hilbert scheme, the ideal I defines $n-$distinct points and the algebra $\mathbb{C}[x,y]/I = \oplus_{i=1}^n \mathbb{C}e_i$ with the e_i orthogonal idempotents.

It is easily seen that the kernel of T has, as basis, the elements $e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_n}$ where the indices are not all distinct. Its complement is the regular representation with basis $e_{\sigma 1} \otimes e_{\sigma 2} \otimes \cdots \otimes e_{\sigma n}$, $\sigma \in S_n$.

If the rank is constant we must have a bundle of regular representations. The set of points with rank $n!$ is open dense and $\mathscr{T}$ stable, hence if the complement, where rank $< n!$, is non empty it must contain a fixed point J_λ contradicting the $n!$ conjecture. □

Let now $\mathscr{B}_n$ be the sheaf of sections of $F^{\otimes n}$ (a bundle of algebras) and $\mathscr{I}_n$ the sheaf of ideals kernel of T.

We have:

$$\mathrm{Spec}(\mathscr{B}_n) := \{(I, p_1, p_2, \ldots, p_n) \in H_n \times (\mathbb{C}^2)^n \mid I(p_i) = 0,\ \forall i = 1, \ldots, n\}. \tag{2.5}$$

Now we can define the subvariety X_n of $\mathrm{Spec}(\mathscr{B}_n)$ which is the closure of its open subset where all p_i are distinct.

2.6. LEMMA. — *X_n is defined by the sheaf of ideals $\mathscr{I}_n$. The commutative diagram:*

$$\begin{array}{ccc} X_n & \longrightarrow & \mathbb{C}^{2n} \\ {\scriptstyle p}\downarrow & & \downarrow \\ H_n & \longrightarrow & \mathbb{C}^{2n}/S_n \end{array} \tag{2.7}$$

identifies X_n to the reduced fiber product.

Sketch of proof. — The sheaf of ideals $\mathscr{I}_n$ restricted to the part of H_n consisting of reduced subschemes, defines the set where all the p_i are distinct. One obtains immediately the first statement. In the commutative diagram, by construction, X_n is a subvariety of the reduced fiber product $\overline{X}_n$. On the regular part the fiber product is reduced, so it is enough to show that $\overline{X}_n$ is irreducible. One sees this by induction on n using the fact that the preimage under p, of a subscheme supported in a unique point, is a point. □

2.8. THEOREM. — *The kernel of T has constant rank if and only if X_n is Cohen-Macaulay and Gorenstein. In this case $\mathscr{B}_n/\mathscr{I}_n = p_*(\mathscr{O}(X_n))$.*

X_n is Cohen-Macaulay if and only if the morphism $p : X_n \to H_n$ is flat. In this case $\mathscr{B}_n/\mathscr{I}_n = p_*(\mathscr{O}(X_n))$.

Proof. — Some parts are fairly straightforward. Let us see how the Gorenstein property plays a role. Assume X_n is Cohen-Macaulay and Gorenstein. Take a point $I_\lambda \in H_n$ we have seen that there is a unique point in $p^{-1}(I_\lambda)$. The coordinate ring B_λ of the sheme theoretic fiber $p^{-1}(I_\lambda)$ is the local ring in this point modulo a regular sequence, hence by the Gorenstein assumption it has a 1-dimensional socle (a unique minimal ideal). It must necessarily be S_n stable and hence carry the sign representation, then the kernel of T on B_λ is an ideal which, if non 0, must contain the sign representation. This is clearly absurd so the form $T(ab)$ is non degenerate on the $n!$ dimensional algebra B_λ.

The converse follows a similar line. □

3. THE G-HILBERT SCHEME

It is quite interesting (and useful) to reinterpret the previous discussion as follows.

If we have that the map $p : X_n \to H_n$ is flat, we also have that each fiber of p is a subscheme of length $n!$ in $\mathbb{C}^{2n}$. From the theory of the Hilbert scheme we have then a classifying map $i : H_n \to H_{n!,2n}$ where $H_{n!,2n}$ is the Hilbert scheme parameterizing subschemes of length $n!$ in $\mathbb{C}^{2n}$. On the other hand, the open part of H_n corresponding to subschemes with $n-$distinct points parametrizes the generic orbits of S_n in $\mathbb{C}^{2n}$.

In general, given a finite group G acting faithfully on an irreducible quasi-projective variety X we have the following construction of Ito and Nakamura [IN]. Consider the

open set X^0 (the union of the generic orbits) over which G acts freely. The set of such orbits X^0/G can be identified with a locally closed subset of the Hilbert scheme $H_{|G|,X}$ of finite subschemes of length $|G|$ in X. One sees easily that this is in fact open in the subscheme of G fixed points of $H_{|G|,X}$.

The closure $\widetilde{H}_{G,X}$ of X^0/G in $H_{|G|,X}$ is an irreducible component of the subscheme of G-stable finite schemes for which the coordinate ring carries the regular representation[(2)].

By continuity the universal family restricted to $\widetilde{H}_{G,X}$ is a flat family of G-stable finite schemes for which the coordinate ring carries the regular representation.

For a G-variety X this closure of X^0/G will be called the *G-Hilbert scheme* and denoted $\widetilde{H}_{G,X}$. It comes equipped with a proper birational morphism $\rho : \widetilde{H}_{G,X} \to X/G$. It seems to be quite interesting to determine when $\widetilde{H}_{G,X}$ is smooth, and hence ρ a canonical resolution of singularities (*cf.* [BKR]). It is not hard to prove that, in our setting, from the flatness of $p : X_n \to H_n$ follows that H_n is identified to the G-equivariant Hilbert scheme of $\mathbb{C}^{2n}$ and X_n to its universal family.

4. X_n IS COHEN-MACAULAY AND GORENSTEIN

The main theorem proved by Haiman in [H2], from which he deduces both the $n!$ and the Macdonald positivity conjectures, is:

4.1. THEOREM. — *X_n is Cohen-Macaulay and Gorenstein.*

The idea of the proof is to use induction on n and the fact that, on the points H_n^0 of H_n which do not define n coincident points, we have a local structure, analytically a product

$$\begin{array}{ccc} X_h \times X_k & \xrightarrow{\ g\ } & \mathbb{C}^{2h} \times \mathbb{C}^{2k} \\ {\scriptstyle p_h \times p_k}\downarrow & & \downarrow \\ H_h \times H_k & \longrightarrow & \mathbb{C}^{2h} \times \mathbb{C}^{2k}/S_h \times S_k \end{array}$$

By induction if X_n^0 is the open set of X_n lying over H_n^0, then X_n^0 is Cohen-Macaulay and Gorenstein, moreover $Rg_*^i\mathscr{O}(X_n^0) = 0$, $\forall i > 0$.

More precisely one has to use the *flag* Hilbert scheme $H_{n,n-1}$ and the corresponding variety $X_{n,n-1}$ and exploit the birational map $X_{n,n-1} \to X_n$.

Set theoretically $H_{n,n-1}$ is made of pairs of ideals

$$I \subset J \subset \mathbb{C}[x,y] \mid \dim \mathbb{C}[x,y]/I = n,\ \dim \mathbb{C}[x,y]/J = n-1.$$

(2) These subschemes are called $G-$clusters in [BKR].

We have 3 natural morphisms:

$$H_{n,n-1} \xrightarrow{q} H_n, \quad H_{n,n-1} \xrightarrow{r} H_{n-1}, \quad H_{n,n-1} \xrightarrow{f} \mathbb{C}^2.$$

From a result of Tikhomirov and also Cheah (*cf.* [Ch]) we have:

4.2. THEOREM. — *$H_{n,n-1}$ is smooth and $H_{n,n-1} \xrightarrow{q\times f} H_n \times \mathbb{C}^2$ has as image the universal family $\mathscr{F}$ and it is a resolution of the singularities of $\mathscr{F}$.*

Composing r with the morphism $H_{n-1} \xrightarrow{s} \mathbb{C}^{2(n-1)}/S_{n-1}$ we obtain

$$H_{n,n-1} \xrightarrow{sr \times f} \mathbb{C}^{2(n-1)}/S_{n-1} \times \mathbb{C}^2 = \mathbb{C}^{2n}/S_{n-1}.$$

At this point define $X_{n,n-1}$ as reduced fiber product:

$$\begin{array}{ccc} X_{n,n-1} & \longrightarrow & X_{n-1} \\ \downarrow & & \downarrow p \\ H_{n,n-1} & \longrightarrow & H_{n-1} \end{array} \tag{4.3}$$

On the varieties $X_{n,n-1}, X_n, H_{n,n-1}, H_n$ we have standard line bundles. On H_n define $\mathscr{O}_{H_n}(1)$ as the maximal exterior power of the tautological bundle and $\mathscr{O}_{H_n}(k) = \mathscr{O}_{H_n}(1)^k$. On $H_{n,n-1}$ we have the line bundles obtained by pull-back through the 2 projections, $H_{n,n-1} \xrightarrow{q} H_n$, $H_{n,n-1} \xrightarrow{r} H_{n-1}$ we set:

$$\mathscr{O}(h,k) := r^*\mathscr{O}_{H_{n-1}}(h) \otimes q^*\mathscr{O}_{H_n}(k).$$

On $X_{n,n-1}$, X_n define also sheaves $\mathscr{O}(h,k)$, $\mathscr{O}(k)$ by pull-back from $H_{n,n-1}$, H_n.

We can now reformulate Theorem 4.1 in a more precise form:

4.4. THEOREM

$T(n)$ *X_n is Cohen-Macaulay, Gorenstein with dualizing sheaf $\mathscr{O}(-1)$.*

$U(n)$ *$X_{n,n-1}$ is Cohen-Macaulay and Gorenstein with dualizing sheaf $\mathscr{O}(0,-1)$.*

The technique of the proof will be to follow the sequence of implications

$$T(n-1) \implies U(n) \implies T(n).$$

To proceed we must compute some canonical sheaves (*cf.* [H2], § 3.6):

4.5. THEOREM. — *For $H_{n,n-1}, H_n$ we have as canonical sheaves:*

$$\omega_{H_{n,n-1}} = \mathscr{O}(1,-1), \qquad \omega_{H_n} = \mathscr{O}(0) \quad \textit{structure sheaf.}$$

It is necessary first of all to prove:

4.6. Lemma. — *Suppose by induction that $X_{n-1} \to H_{n-1}$ is flat, the diagram*

$$\begin{array}{ccc} X_{n,n-1} & \longrightarrow & X_{n-1} \\ \downarrow & & \downarrow p \\ H_{n,n-1} & \longrightarrow & H_{n-1} \end{array}$$

is a fiber product.

Proof. — The statement claims that, forming the fiber product

$$\begin{array}{ccc} Y & \longrightarrow & X_{n-1} \times \mathbb{C}^2 \\ p' \downarrow & & \downarrow p \\ H_{n,n-1} & \longrightarrow & H_{n-1} \times \mathbb{C}^2 \end{array}$$

we have that Y is reduced and thus it coincides with $X_{n,n-1}$.

In order to prove it we use the fact that p is flat and finite and so also p' is flat and finite, thus Y is Cohen-Macaulay and it suffices to prove that it is reduced in codimension 0, which one obtains restricting to the regular part of the diagram. □

Now prove that $T(n-1) \implies U(n)$.

From the previous Lemma it follows that $X_{n,n-1}$ is Gorenstein, moreover the dualizing sheaf relative to the morphism $X_{n,n-1} \to X_{n-1}$ is the pull-back of the dualizing sheaf relative to the morphism $H_{n,n-1} \to H_{n-1}$ which is $\mathscr{O}(1,-1)$ while the dualizing sheaf of X_{n-1} is by induction $\mathscr{O}(-1,0)$ thus the tensor product is $\mathscr{O}(0,-1)$.

Now the implication $U(n) \implies T(n)$.

One has to analyze the morphism $g : X_{n,n-1} \to X_n$. The main Theorem follows from general principles from the proposition:

4.7. Proposition

$$R^i g_* \mathscr{O}(X_{n,n-1}) = 0, \ \forall i > 0, \quad g_* \mathscr{O}(X_{n,n-1}) = \mathscr{O}(X_n). \tag{4.8}$$

This proposition is based on a basic Lemma and some geometric considerations.

4.9. Lemma. — *Given a proper morphism $g : Y \to X$ between algebraic varieties over $\mathbb{C}$. Suppose we have given m global functions $z_1, \dots, z_m$ on X and let Z be the subvariety of X where they vanish and $U := X - Z$ the complement.*

Assume the following conditions:

(1) *The z_i form a regular sequence in every local ring $\mathscr{O}_{X,P}$, $P \in Z$.*
(2) *The z_i form a regular sequence in every local ring $\mathscr{O}_{Y,Q}$, $Q \in g^{-1}(Z)$.*
(3) *Every fiber of g has dimension $< m-1$.*
(4) *On the open set U the canonical morphism $\mathscr{O}_X \to Rg_* \mathscr{O}_Y$ is an isomorphism,*

then the canonical morphism $\mathscr{O}_X \to Rg_* \mathscr{O}_Y$ is an isomorphism (everywhere).

This Lemma has a fairly simple cohomological proof ([H2], Lemma 3.8.5). The hard point is to apply this Lemma to the morphism g. We choose as sequence z_i the $n-1$ functions $y_1 - y_2, \ldots, y_1 - y_n$ and we need to verify the hypotheses (1)–(4).

The most difficult is (1).

In any case both for (1) and (2) one proves the stronger statement that $y_1, y_2, \ldots, y_n$ is a regular sequence.

One observes:

For (2) by induction all the local rings are Cohen-Macaulay, one must verify that the codimension of the variety given by the equations $y_1 = y_2 = \cdots = y_n = 0$ is n.

It is enough to do it for $H_{n,n-1}$, since $X_{n,n-1} \to H_{n,n-1}$ is finite.

For (3) it is enough to analyze the morphism $H_{n,n-1} \to H_n$.

The fiber has maximal dimension on the fixed points, i.e. the ideals J_λ, and consists of the ideals of dimension 1 of $\mathbb{C}[x,y]/J_\lambda$. By direct inspection one sees that this is a projective space of dimension $d-1$ where d is the number boundary cases of the diagram. For $n > 3$ the inequality is easy while for $n \leqslant 3$ one must prove the Lemma directly.

To prove (4) we see that on U the morphism is locally isomorphic to a product of two morphisms $g : X_{k,k-1} \times X_{h,h-1} \to X_k \times X_h$. Thus we can proceed by induction.

(1) is the difficult part.

First a reduction: Presenting H_n and X_n as blow-ups.

First H_n.

Let $R := \mathbb{C}[x_1, \ldots, x_n; y_1, \ldots, y_n]$ be the polynomial ring in $2n$ variables:

$$S := \mathbb{C}[x_1, \ldots, x_n; y_1, \ldots, y_n]^{S_n}, \quad A := \{f \in R \mid \sigma(f) = \varepsilon_\sigma f,\ \sigma \in S_n\}$$

the invariants and the alternating elements under the diagonal action of S_n (ε_σ denotes the sign of the permutation). Finally let $I := AR$ be the ideal of R generated by A. We add an indeterminate t and consider the two graded algebras:

$$U_n := S \bigoplus_{i=1}^{\infty} A^i t^i, \quad V_n := R \bigoplus_{i=1}^{\infty} I^i t^i.$$

One acts with S_n on $R[t]$ by the diagonal action on R and acting on t with the sign representation.

4.10. THEOREM. — a) $U = V^{S_n}$.

b) $H_n = \mathrm{Proj}(U_n), \quad X_n = \mathrm{Proj}(V_n)$.

Sketch of proof. — It is well known that H_n can be described as follows (*cf.* [N]).

Consider the variety Z of triples (X, Y, v) where X, Y are two $n \times n$ matrices with $XY = YX$, $v \in \mathbb{C}^n$ a vector. Define Z_0 to be the open set of Z where the vectors $X^i Y^j v$ generate the space $\mathbb{C}^n$.

Z_0 is smooth, the group $\mathrm{GL}(n, \mathbb{C})$ acts freely on Z_0 and finally $H_n = Z_0/\mathrm{GL}(n, \mathbb{C})$.

On Z many conjectures are open but we know that it is irreducible and that pairs of diagonalizable matrices are dense in Z.

Construct the quotient in two steps: $\widetilde{H}_n = Z_0/\operatorname{SL}(n,\mathbb{C})$, $H_n = \widetilde{H}_n/\mathbb{C}^*$ we see that $Z_0/\operatorname{SL}(n,\mathbb{C})$ is an open set of the variety $Z/\!/\operatorname{SL}(n,\mathbb{C})$.

By definition $Z/\!/\operatorname{SL}(n,\mathbb{C})$ is the spectrum of the ring of invariants and, by classical invariant theory such invariants are generated by:

$$\operatorname{tr}(M),\ [M_1v, M_2v, \dots, M_nv]$$

where M, M_i denote monomials in the matrices X, Y and $[M_1v, M_2v, \dots, M_nv]$ denotes the determinant of these vectors.

By the previous remarks we can compute this ring by restricting it to pairs of diagonal matrices and then it is easy to see that it is identified to U_n. Moreover the open set $Z_0/\operatorname{SL}(n,\mathbb{C})$ is the part where at least one of the determinants $[M_1v, M_2v, \dots, M_nv]$ is non 0 from which the statement easily follows.

As for X_n one has clearly a commutative diagram:

$$\begin{array}{ccc} \operatorname{Proj}(V_n) & \longrightarrow & \mathbb{C}^{2n} \\ {\scriptstyle p}\downarrow & & \downarrow \\ \operatorname{Proj}(U_n) & \longrightarrow & \mathbb{C}^{2n}/S_n \end{array} \tag{4.11}$$

comparing it to 2.7 one has the claim. □

At this point one has the next reduction:

In order to prove condition (1) it suffices to prove that, for every d the ideal I^d is a free module over the polynomial ring $\mathbb{C}[y_1, \dots, y_n]$.

This last statement will be further reduced to a more combinatorial statement. In order to explain it we must introduce some new objects, the *polygraphs*.

Thus, given a positive integer ℓ let us denote by $[\ell]$ the segment $[1, 2, \dots, \ell]$.

Given a function $f : [\ell] \to [n]$ consider the induced linear map $\pi_f : (\mathbb{C}^2)^n \to (\mathbb{C}^2)^\ell$, and its graph $W_f \subset (\mathbb{C}^2)^n \times (\mathbb{C}^2)^\ell$.

The union $Z(n,\ell) := \cup_f W_f$ as f varies on the set of all functions $f : [\ell] \to [n]$ is a *polygraph*. Clearly on this polygraph operate various groups, in particular we will use the group S_ℓ, and its subgroups, which permutes functions and graphs.

Call x_i, y_i the coordinates on $(\mathbb{C}^2)^n$ and a_i, b_i those on $(\mathbb{C}^2)^\ell$.

The coordinate ring of $Z(n,\ell)$ is a quotient $R(n,\ell) := \mathbb{C}[x_i, y_i, a_j, b_j]/I(n,\ell)$.

We consider next $\ell = nd$ and decompose

$$(\mathbb{C}^2)^n \times (\mathbb{C}^2)^n \times \cdots \times (\mathbb{C}^2)^n = (\mathbb{C}^2)^\ell$$

the group S_n^d operates permuting separately the coordinates of the factors and on the ring $R(n,\ell)$. Consider the subspace $R(n,\ell)^\varepsilon$ of antisymmetric elements (with respect to all factors S_n).

4.12. Lemma. — *There is a canonical isomorphism as R modules of $R(n,\ell)^\varepsilon$ and I^d.*

Proof. — Let $f_0 : [\ell] \to [n]$ be defined as $f_0(kn+i) = i$, $\forall 1 \leqslant i \leqslant n, 0 \leqslant k < d$ call $W_{f_0} = (\mathbb{C}^2)^n$ the associated graph. The restriction of functions of $R(n,\ell)$ to such a space is a morphism of R modules and sends $a_{kn+i} \to x_i$, $b_{kn+i} \to y_i$; it is easily seen that it maps $R(n,\ell)^\varepsilon$ surjectively to J^d.

It is enough to prove that it is also injective. In fact an antisymmetric function vanishes on W_f if $f(kn+i) = f(kn+j)$, $1 \leqslant i \leqslant n$, $0 \leqslant k < d$, $i \neq j$. Any other function is in the orbit of f_0 relatively to S_n^d. By antisymmetry a function in $R(n,\ell)^\varepsilon$ is determined completely by its values on W_{f_0}. □

So finally let us show that everything will follow from the final key statement:

4.13. THEOREM. — *$R(n,\ell)$ is a free module on the polynomial ring $\mathbb{C}[y_1,\dots,y_n]$.*

Assume the last statement. Decomposing $R(n,\ell)$ into isotypic components with respect to S_n^d we see that $R(n,\ell)^\varepsilon$ is a direct summand as module and so from the freeness of $R(n,\ell)$ follows also that of I^d isomorphic to $R(n,\ell)^\varepsilon$.

At the end of this long sequence of reductions we have to face Theorem 4.13. This is proved by Haiman with a very long and complex induction only using a very careful bookkeeping and commutative algebra which occupies more than 30 pages of his paper. Rather than try to discuss this highly technical point we prefer to discuss some further developments. Nevertheless one should point out that, from the considerations that we will see in the next paragraph (Theorem 5.1), the appearence of polygraphs and their properties are geometrically natural.

5. DIAGONAL HARMONICS

The space of diagonal harmonics D_n can be defined as the subspace of polynomials in $\mathbb{C}[x_1,\dots,x_n;y_1,\dots,y_n]$ solutions of the system of differential equations $P(\partial/\partial x_i, \partial/\partial y_i)f = 0$ where P runs over all polynomials without constant term which are symmetric with respect to the diagonal action of S_n.

D_n can also be identified with the polynomial ring $\mathbb{C}[x_1,\dots,x_n;y_1,\dots,y_n]$ modulo the ideal generated by all S_n invariant polynomials without constant term.

Garsia and Haiman discovered a series of interesting conjectures mixing algebra, combinatorics and geometry on the space D_n, which is a bigraded, finite dimensional representation of S_n. The simplest of which describe its dimension as a vector space and its structure as representation.

(1) $\dim D_n = (n+1)^{n-1}$.

(2) As a representation D_n is isomorphic to the permutation representation on the set of *parking functions* tensored by the sign representation[(3)].

More precise conjectures on the bigraded character can be found in [GH1], in particular a rather remarkable expression for its bigraded Frobenius character.

One can attack these conjectures using the Lefschetz fixed point formula of Atiyah-Bott (*cf.* [AB]).

The idea is to prove that the space of diagonal harmonics can be identified with the global sections of the vector bundle $p_*\mathscr{O}_{X_n}$ restricted to the subvariety of H_n consisting of subschemes supported at 0 and then compute its character by localization principles.

This vector bundle is acted upon by the torus T and the Lefschetz fixed point formula of Atiyah-Bott can be applied provided one knows the vanishing of suitable cohomology groups. Finally these vanishing theorems can be deduced by applying the theory of Bridgeland King Reid to X_n (this can be done because of the solution of the $n!-$conjecture). Their theory in our case establishes an equivalence of derived categories *the BKR correspondence* (generalized McKay correspondence) between the derived category of coherent sheaves on H_n and that of S_n-equivariant modules over R.

The announced geometric interpretation of polygraphs and their property is:

5.1. Theorem. — *Under the BKR correspondence the polygraph $R(n,\ell)$ corresponds to the bundle F^ℓ.*

The locus $V(y_1,\dots,y_n)$ in X_n is a complete intersection.

This provides the requested vanishing theorems.

Once all of this is done one obtains finally an explicit formula, in term of Macdonald polynomials of the bigraded character of D_n (*cf.* [H3], 3.10 and 3.11). In particular the two previous conjectures follow from this formula.

Final comments

The theory we have sketched applies to the sum of two copies of the standard reflection representation of S_n. It thus suggests possible generalizations to other reflection groups.

At the moment it is not very clear what can be generalized and in which form, as the question of understanding which G-Hilbert schemes are smooth is completely mysterious. Haiman has made some computations for type B_n.

(3) $n-$people choose a spot out of n linearly ordered parking spots. They must reach their chosen spot and, if free park in it, otherwise reach the first free spot and park. The choice made is a *parking function* if, no matter in which order the people arrive they always park.

REFERENCES

[AB] M. Atiyah & R. Bott – A Lefschetz fixed point formula for elliptic differential operators, *Bull. Amer. Math. Soc.* **72** (1966), p. 245–250.

[BKR] T. Bridgeland, A. King & M. Reid – Mukai implies McKay: the McKay correspondence as an equivalence of derived categories, Electronic preprint, `arXiv:math.AG/9908027`, 1999.

[Ch] J. Cheah – Cellular decompositions for nested Hilbert schemes of points, *Pacific J. Math.* **183** (1998), p. 39–90.

[F] J. Fogarty – Algebraic families on an algebraic surface, *Amer. J. Math.* **90** (1968), p. 511–521.

[GH] A. Garsia & M. Haiman – A graded representation model for Macdonald's polynomials, *Proc. Nat. Acad. Sci. U.S.A.* **90** (1993), no. 8, p. 3607–3610.

[GH1] ———, A remarkable $q-t$-Catalan sequence and q-Lagrange inversion, *J. Algebraic Comb.* **5** (1996), no. 3, p. 191–244.

[GP] A. Garsia & C. Procesi – On certain graded S_n-modules and the q-Kostka polynomials, *Adv. Math.* **94** (1992), no. 1, p. 82–138.

[H] M. Haiman – Conjectures on the quotient ring by diagonal invariants, *J. Algebraic Combin.* **5** (1994), no. 1, p. 17–76.

[H1] ———, Macdonald polynomials and geometry, in *New perspectives in geometric combinatorics* (Billera, Björner, Greene, Simion & Stanley, eds.), vol. 38, M.S.R.I. Publications, 1999, p. 207–254.

[H2] ———, Hilbert schemes, polygraphs, and the Macdonald positivity conjecture, *Journal of the A.M.S.* (to appear), 2001.

[H3] ———, Vanishing theorems and character formulas for the Hilbert scheme of points in the plane, preprint, 2001.

[IN] Y. Ito & I. Nakamura – McKay correspondence and Hilbert schemes, *Proc. Japan Acad. Ser. A Math. Sci.* **72** (1996), no. 7, p. 135–138.

[M] I.G. Macdonald – A new class of symmetric functions, in *Actes du 20ème séminaire lotharingien*, vol. 372/S-20, Publ. I.R.M.A. Strasbourg, 1988, p. 131–171.

[M1] ———, *Symmetric functions and Hall polynomials*, 2nd ed., The Clarendon Press, Oxford University Press, New York, 1995.

[N] H. Nakajima – *Lectures on Hilbert schemes of points on surfaces*, American Math. Society, Providence RI, 1999.

Claudio PROCESI

Dipartimento di Matematica G. Castelnuovo
Università di Roma La Sapienza
Piazzale A. Moro
I-00185 Roma
Italie
E-mail : `claudio@mat.uniroma1.it`

Séminaire BOURBAKI Mars 2002
54e année, 2001-2002, no 899, p. 117 à 148

DUALITÉS DE CHAMPS ET DE CORDES
[d'après 't Hooft, Polyakov, Witten et al.]

par **Daniel BENNEQUIN**

L'auteur souhaite seulement présenter, pour des mathématiciens, le cadre géométrique des dualités en physique. En particulier, faute de temps et de place, l'exposé écrit ne développe pas les principales applications en mathématique : symétrie miroir et théorie de Seiberg-Witten.

1. DUALITÉS CLASSIQUES

1.1. Une dualité s'installe lorsque deux points de vue opposés sont nécessaires pour décrire un même ensemble de phénomènes. La dualité est l'échange de points de vue complémentaires. Elle est partout présente en physique.

Une première manifestation de dualité est expliquée dans le cours de Feynman : la description d'un champ doit passer par la description du mouvement d'une particule test ou d'un appareil amplificateur (par exemple de la limaille de fer pour un champ magnétique ou des antennes pour les ondes gravitationnelles). À un niveau plus fin, la particule crée des champs ; plus fondamentalement, particule et appareil sont eux-mêmes des champs ; comment décrire cette interaction entre champs sans choisir un point de vue extérieur aux champs ?

La perception visuelle fait aussi intervenir une dualité, que B. Julia m'avait signalée : les photons qui renseignent sur une position de la matière ne communiquent que des impulsions ; à partir d'une fonction des impulsions et des fréquences (p, ω), notre système nerveux extrait une fonction des positions et du temps (q, t). La transformation de Fourier est une expression mathématique de cette dualité. Elle traduit en Analyse la plus ancienne dualité développée en mathématique : la dualité projective de Poncelet.

En mécanique quantique, la transformation de Fourier devient un changement de repère (W. Heisenberg, [38]), exprimant la *dualité onde-corpuscule*. Le principe d'incertitude limite la compatibilité des points de vue mis en dualité, mais la mécanique

quantique rétablit aussi un équilibre entre la matière et les champs d'interaction, en introduisant la fonction d'onde de Schrödinger.

Le domaine classique impose des équations différentielles à ses champs, par exemple l'équation de Maxwell pour le champ électromagnétique, l'équation d'Einstein pour le champ gravitationnel. La Mécanique quantique relativiste ajoute les équations de Dirac pour les fonctions d'ondes. Dualement, le champ de Maxwell peut être vu comme fonction d'onde du photon.

Dans le vide sur $\mathbb{R}^4$, les équations de Maxwell pour les champs électrique et magnétique E et B s'écrivent

$$\nabla \times B = \frac{\partial E}{\partial t}, \qquad \nabla \times E = -\frac{\partial B}{\partial t}$$
$$\nabla \cdot B = 0, \qquad \nabla \cdot E = 0$$

elles sont symétriques pour la transformation

$$E \longmapsto B, \qquad B \longmapsto -E.$$

C'est surtout cette sorte de symétrie qui s'appelle *dualité* dans les développements récents (Misner et Wheeler, *cf.* [57]).

Pour que cette symétrie reste valable en présence de charges et de courants, Dirac a proposé, en 1931, d'ajouter des charges et des courants magnétiques aux côtés des charges et des courants électriques usuels ([14], [15]). Ce fut l'invention du *monopôle magnétique*, activement recherché mais toujours non détecté.

Cependant, le potentiel magnétique est un vecteur et le potentiel électrique, lui, est un scalaire, ce qui fait penser qu'à un niveau plus profond, notamment en mécanique quantique, des difficultés s'opposent à la dualité électrique–magnétique (*cf.* Witten, 1997, [97]).

Dans la théorie des interactions faibles (principales responsables des effets radioactifs), on rencontre une généralisation des équations de Maxwell, les équations de Yang-Mills-Higgs (*cf.* § 1.2). Pour des équations analogues, en 1974, 't Hooft et Polyakov ont découvert des *monopôles magnétiques non abéliens*, et établi une forme de dualité échangeant magnétisme et électricité (*cf.* § 2.3). Pourtant, là aussi, les degrés de liberté des potentiels sont différents.

C. Montonen et D. Olive (1977) ont conjecturé une vraie dualité électrique–magnétique, mais dans une *théorie de jauge* super-Yang-Mills-Higgs, avec 4 *supersymétries*, en dimension 4. Après des progrès notables de D. Olive, E. Witten (1978) et N. Seiberg (1988), la conjecture a été établie en 1994 par A. Sen (*cf.* § 2.4).

Une propriété remarquable de cette dualité est qu'elle échange deux *échelles* différentes de la théorie : les propriétés à courte distance avec celles aux grandes distances.

La *dualité* se révèle être un principe fondamental en théorie quantique des champs. Elle transforme le principe de correspondance de Bohr : une théorie quantique est

associée à une théorie de champ classique, mais, une fois la théorie quantique développée, il apparaît d'autres théories classiques représentant certaines limites, certains points de vue différents, sur la même théorie quantique. Comme une variété peut avoir besoin de plusieurs cartes.

De plus, la plupart des dualités connues proviennent de dualités en théorie des cordes. Le creuset des dualités semble être la M-théorie (§ 3.3).

La fin du § 1 expose brièvement le cadre des champs classiques, avec des fermions à partir de 1.3 et des super-champs à partir de 1.4. Le § 2 donnera une idée de la renormalisation, ce qui est nécessaire pour préciser la définition d'une dualité et il exposera quelques exemples de dualité de champs. Le § 3 fera une introduction aux dualités de cordes qui semblent détenir le secret des dualités de champs.

1.2. En théorie de Yang-Mills (*cf.* [40], [56]), la variable est une connection ∇ sur un fibré principal P en groupe de Lie G au-dessus d'une variété W de dimension D. En choisissant une trivialisation locale de P, on écrit $\nabla = d + A$, avec une 1-forme différentielle A à valeurs dans l'algèbre de Lie $\mathfrak{g}$ de G. La courbure F, ou F_A, de ∇ est la 2-forme définie par $F_A = dA + A \wedge A$; à l'origine classique, c'est F qui s'appelait le champ. Soit d_A la dérivée extérieure covariante, étendant ∇ en une dérivation graduée du module des formes différentielles à valeurs dans $\mathfrak{g}$ (sur l'algèbre des formes différentielles scalaires), l'identité de Bianchi, toujours satisfaite, est $d_A F_A = 0$. Pour écrire l'équation de Yang-Mills, on choisit une métrique sur W, lorentzienne, de signature $(D-1,1)$ si on souhaite un temps réel, riemannienne si on préfère le temps imaginaire pur, et l'on note $*$ l'opérateur de Hodge-de Rham associé, échangeant les formes extérieures de degré p et $D-p$. Dans le vide, l'équation est

$$d_A * F_A = 0.$$

S'il y a des courants j, l'équation devient

$$d_A * F_A = *j.$$

Les équations de Maxwell sont obtenues dans le cas particulier $D = 4$, $G = U_1$; alors F est une 2-forme différentielle ordinaire. En signature lorentzienne, et en choisissant une coordonnée t pour le temps, la courbure F se décompose en un champ de (co)vecteurs E (les composantes en $dt \wedge dx^i$) et un champ de (bi)vecteurs B (les composantes en $dx^i \wedge dx^j$). Alors l'action de l'opérateur $*$ se traduit par $E \longmapsto B$, $B \longmapsto -E$.

Dans ce formalisme, avec $W = M \times \mathbb{R}$, M de dimension 3, pour toute surface fermée $\Sigma \subset M$, on définit les charges électriques et magnétiques par

$$Q_e(\Sigma) = \int_\Sigma \frac{*F}{2\pi}, \qquad Q_m(\Sigma) = \int_\Sigma \frac{F}{2\pi}.$$

Le fait que F est la courbure d'un fibré impose à $Q_m(\Sigma)$ d'être un nombre entier. En l'absence de mouvement des courants, $Q_e(\Sigma)$ est indépendant de t en vertu des

équations de Maxwell. On dit que Q_e est une charge de Noether et que Q_m est une charge topologique.

Les équations de Yang-Mills sont les équations d'Euler-Lagrange de l'action dans le vide

$$S = \frac{1}{4g^2}\int_W \mathrm{Tr}(F_A \wedge *F_A).$$

On dit que $\frac{1}{4g^2}\,\mathrm{Tr}(F_A \wedge *F_A)$ est le lagrangien de la théorie.

La théorie est invariante par *changement de jauge*, c'est-à-dire par automorphisme de P induisant l'identité de W.

La variable de la gravitation d'Einstein est une métrique $\mathbf{g}$ (lorentzienne pour un temps réel) sur W et le lagrangien est celui de Hilbert, c'est-à-dire $R \cdot \mathrm{Vol}_{\mathbf{g}}$, où $\mathrm{Vol}_{\mathbf{g}}$ est l'élément de volume de $\mathbf{g}$ et R la courbure scalaire de $\mathbf{g}$. Einstein a aussi pensé au lagrangien plus général $(R - 2\Lambda)\cdot \mathrm{Vol}_{\mathbf{g}}$, pour une constante Λ, la *constante cosmologique* (*cf.* [37], [93]). La théorie est invariante par difféomorphisme de W. Afin d'obtenir les équations du mouvement en présence d'autres champs, comme ceux de Maxwell ou de Yang et Mills, on ajoute les actions entre elles et on prend les équations variationnelles d'Euler-Lagrange. Les théories d'Einstein et de Maxwell ou Yang-Mills ensemble deviennent invariantes par tout automorphisme de P. (*Cf.* [9], [37], [42].)

Un des exemples les plus étonnants de dualité classique est celui d'Ehlers et Geroch pour la gravitation en dimension 4 ([23], [28]) :

Lorsqu'existe un champ de vecteurs K préservant $\mathbf{g}$, c'est-à-dire un champ de Killing, on peut voir la théorie d'Einstein comme une théorie de dimension 3 sur le quotient par K ; les champs étant la métrique quotient $\mathbf{h}$, un champ scalaire λ, venant de la norme de K, et une 1-forme $\mathbf{A}$ exprimant le produit scalaire par K des vecteurs transverses. Cette forme $\mathbf{A}$ hérite d'une invariance de jauge U_1, et l'équation d'Einstein en dimension 4 impose l'équation de Maxwell : $d*d\mathbf{A} = 0$. J. Ehlers a donc pu introduire le *champ scalaire* dual, au sens de Poincaré-Hodge : $d\omega = *d\mathbf{A}$. Pour écrire les équations différentielles sur $\mathbf{h}, \lambda, \omega$, le mieux selon R. Geroch est de poser $\tau = \omega + i\lambda$ et $\widetilde{\mathbf{h}} = \lambda\mathbf{h}$; dès lors, en notant $\widetilde{\nabla}$ la dérivée covariante associée à $\widetilde{\mathbf{h}}$, on a

$$\widetilde{R} = \frac{1}{2}\,(\mathrm{Im}\,\tau)^{-2}(\widetilde{\nabla}\tau)\circ(\widetilde{\nabla}\bar{\tau}), \qquad d^{\widetilde{\nabla}} * d^{\widetilde{\nabla}}\,\tau = (\mathrm{Im}\,\tau)^{-1}|\widetilde{\nabla}\tau|^2.$$

Le miracle de dualité est l'invariance par $SL_2(\mathbb{R})$ de ces équations : en posant $\widetilde{\mathbf{h}}' = \widetilde{\mathbf{h}}$, et

$$\tau' = \frac{a\tau + b}{c\tau + d}\,,$$

pour tout ensemble de nombres réels a, b, c, d, tels que $ad - bc = 1$, on déduit d'une solution $\mathbf{h}, \tau$ une autre solution $\mathbf{h}', \tau'$.

C'était le rêve d'Einstein de trouver une théorie unitaire mettant tous les champs sur le même pied que la métrique $\mathbf{g}$. Une proposition dans ce sens fut celle de Th. Kaluza (1921), précisée par O. Klein (1926). L'idée est présente dans l'exemple précédent : la réduction dimensionnelle (*cf.* [8], [9], [21]).

Partant du lagrangien d'Einstein en dimension D, on impose à W d'être le produit d'une variété X par une variété M de dimension k, et on impose à la métrique le long de M de varier dans une famille de dimension finie, par exemple $M = G/H$ et $\mathbf{g}_{|M}$ G-invariante, ou encore $(M, \mathbf{g})$, variété riemannienne d'Einstein-Kähler. Alors le champ $\mathbf{g}$ sur W se réinterprète comme un ensemble de champs sur X. Dans le cas le plus simple, $M = S^1$, $X = \mathbb{R}^4$, on trouve le couplage sur X des équations d'Einstein et de Maxwell, mais accompagné d'un champ scalaire supplémentaire qui ne peut pas être pris constant dans les solutions. Lorsque $M = G$, groupe de Lie compact, on trouve à côté de $\mathbf{g}$ sur X un champ de Yang-Mills ∇, et comme champ supplémentaire, une métrique invariante (des deux côtés) sur G. Lorsque $M = G/H$, on trouve une métrique, une théorie de jauge (pour le normalisateur N de H dans G) et des champs de Higgs φ_1, encore accompagnés de champs scalaires φ_2 (*cf.* [9]). Ces *champs scalaires ajoutés* φ, et leurs généralisations en présence d'autres champs que $\mathbf{g}$ sur W, par exemple en présence d'un champ anti-symétrique $\mathbf{b}$ sur W vont fournir le principal réservoir de dualité, par un mécanisme analogue à celui de l'exemple d'Ehlers ([10], [11], [47], [49]).

1.3. En face des tenseurs comme $\mathbf{g}$, F, φ véhiculant des interactions entre particules et se manifestant quantiquement comme des bosons, il y a, pour la matière ordinaire, des champs de spineurs ψ qui ont le comportement quantique des fermions.

Lorsque $\mathbf{g}$ est une métrique définie positive, ou de signature lorentzienne $(D-1, 1)$, sur un espace vectoriel de dimension $D \geq 3$, la composante connexe de l'identité dans le groupe des rotations, $SO(\mathbf{g})$, possède un groupe fondamental à deux éléments. Le groupe $Spin(\mathbf{g})$ est le revêtement universel de $SO(\mathbf{g})$; c'est donc un revêtement à deux feuillets. Une *représentation spin* est une représentation linéaire complexe de $Spin(\mathbf{g})$ ayant la plus petite dimension possible. En dimension D impaire, il existe une représentation spin unique à isomorphisme près S de dimension $2^{(D-1)/2}$ sur $\mathbb{C}$; en dimension D paire, il y en a deux S^+, S^- de dimension $2^{D/2}$, dites de Weyl ([13], [33]). La somme $S = S^+ \oplus S^-$ s'appelle espace des spineurs de Dirac. Quand S est la complexifiée d'une représentation réelle $S_{\mathbb{R}}$, on dit que $S_{\mathbb{R}}$ est un espace de spineurs de Majorana. S'il existe $S^+_{\mathbb{R}}$ et $S^-_{\mathbb{R}}$, on parle de Majorana-Weyl.

Une structure spin sur une variété riemannienne orientée W est un revêtement double du fibré des repères orthonormés directs qui induit le revêtement $Spin(\mathbf{g})$ de $SO(\mathbf{g})$ en chaque point de W. Lorsqu'une structure spin existe, on peut définir des fibrés de spineurs encore notés S et un opérateur de Dirac $D : S \to S$; l'équation de Dirac sans masse $D\psi = 0$ dérive du lagrangien $\widetilde{\psi} \cdot D\psi$ ([13]).

En présence d'un fibré principal $P \to W$ en groupe G muni d'une connexion ∇ et en choisissant un fibré vectoriel E associé à P, on a un opérateur de Dirac tordu $D_A : S \otimes E \to S \otimes E,\ D_A = D \otimes 1 + 1 \otimes \nabla$; le lagrangien associé est $\widetilde{\psi} \cdot D_A \psi$.

Si l'on ajoute à ∇ et ψ les *champs de Higgs*, qui sont des sections de certains fibrés associés à P, on obtient tous les champs du « modèle standard ».

La théorie des représentations unitaires du groupe de Poincaré-Lorentz selon E. Wigner associe à toute particule une composante irréductible de la représentation du groupe des spineurs sur une puissance tensorielle $S \otimes S \otimes \cdots \otimes S$ avec n facteurs ; dans ce cas, on parle d'*hélicité* $\frac{n}{2}$, ou encore de *spin* $\frac{n}{2}$. Par exemple, l'électron est de spin $\frac{1}{2}$; le photon, un champ de vecteurs, un champ de covecteurs, une connexion ∇ est de spin 1 ; une métrique est de spin 2, la particule quantique associée est le *graviton*. Les particules de spin demi-entier sont les *fermions*, celle de spin entier, les *bosons*.

Élie Cartan a montré que le couplage des équations d'Einstein avec celles de Dirac exige un changement de point de vue : il ne suffit plus de considérer $\mathbf{g}$, il faut ajouter une torsion des connexions métriques, car le tenseur énergie-moment d'un spineur n'est pas symétrique. La méthode retenue pour résoudre le problème (Sciama-Kibble, 1961) consiste à introduire un sous-fibré réel V de $S \otimes S^*$, isomorphe au fibré tangent $T(W)$, et à faire varier l'isomorphisme $\varepsilon : T(W) \to V$, nommé *vielbein* et noté e^α_μ. Un choix de ε détermine une métrique. Les variables de la théorie sont alors ε, ψ et une connexion lorentzienne auxiliaire ω sur V (*cf.* [33]).

1.4. Toutes ces interactions, tous ces lagrangiens qui s'ajoutent font intervenir des échelles de grandeurs extrêmement différentes (des rapports de 10^{20}). La *supersymétrie*, introduite au début des années 70 (pour les besoins des cordes), se propose d'échanger les bosons avec les fermions ; si elle n'était qu'approximativement réalisée, cela pourrait expliquer les différences d'échelles.

Salam et Strathdee ont introduit la notion de *super-espace* dans ce contexte ; elle a depuis envahi toute la théorie des champs et des cordes (*cf.* [85], [56], [33], [13]).

Remarque. — Il y a là un point qui gêne souvent le dialogue entre mathématiciens et physiciens, car les définitions adoptées couramment en mathématique et en physique sont différentes. Mais A.S. Schwartz [74] a montré qu'elles sont équivalentes :

Pour $d \in \mathbb{N} \cup \{\infty\}$, on note L_d l'algèbre extérieure $\Lambda(\mathbb{R}^d)$, $L_d^0 = \mathbb{R}$ l'ensemble de ses éléments de degré 0, L_d^+ et L_d^- les facteurs de degré pair > 0 et de degré impair respectivement.

Le modèle local d'une *super-variété* au sens de De Witt de dimension $(m|n)$ et de dimension cachée $d \in \mathbb{N} \cup \{\infty\}$ est un produit $U \times (\mathbb{R}^m \otimes L_d^+) \times (\mathbb{R}^n \otimes L_d^-)$, pour U ouvert de $\mathbb{R}^m$, que l'on note $U^{m|n}$. (*Cf.* [73], [85].)

Pour toute fonction C^∞, $f : U \to \mathbb{R}$, on note $Z[f]$ la fonction de $U \times (\mathbb{R}^m \otimes L_d^+)$ dans L_d définie par la formule

$$Z[f](u_1, \ldots, u_m) = \sum_{\alpha \in \mathbb{N}^m} \frac{D^\alpha f(a)}{\alpha!} (u^+)^\alpha \,,$$

en décomposant les coordonnées : $u_i = a_i + u_i^+$ avec $a_i \in L_d^0$ et $u_i^+ \in L_d^+$.

Une *super-fonction* Φ de $U^{m|n}$ dans L_d est une fonction de $U^{m|n}$ dans L_d de la forme suivante, pour $u \in U \times (\mathbb{R}^m \otimes L_d^+)$ et $v \in \mathbb{R}^n \otimes L_d^-$:

$$\Phi(u, v) = \sum_{\beta \in \mathbb{N}^n} Z[f_\beta](u) \cdot v^\beta \,,$$

où les f_β sont des fonctions C^∞ de U dans $\mathbb{R}$.

On dit que Φ est paire (resp. impaire) si son image appartient à $L_d^+ \oplus L_d^0$ (resp. L_d^-).

À partir de là, on définit les *morphismes* $F = (\Phi_1, \ldots, \Phi_p, \Psi_1, \ldots, \Psi_q)$ de $U^{m|n}$ dans $V^{p|q}$ par leurs composantes super-fonctions, les Φ_i étant paires et les Ψ_j impaires. Ces morphismes s'appellent applications de classe H^∞ dans [73].

En recollant les modèles, on obtient une catégorie de super-variétés. Cette catégorie pour $d = \infty$ est équivalente à la catégorie des super-variétés de Berezin, Kostant, Manin ([13], [56], [73], [74]), c'est-à-dire celle des espaces annelés avec le modèle local de faisceaux $\Lambda^{m|n}(U) = C^\infty(U, L_n)$, pour U ouvert dans $\mathbb{R}^m$, les morphismes étant les morphismes pairs d'algèbres graduées (*cf.* [13], p. 66, lemme de Leites et Manin).

Le super-espace-temps est le produit de l'espace-temps usuel W de dimension D avec le produit $(\mathbb{R}^D \otimes L_d^+) \times (S^N \otimes L_d^-)$, pour d assez grand. Le nombre N se nomme *nombre de super-symétrie*. L'espace qui sert pour la dimension impaire est l'espace S d'une représentation spinorielle. La super-algèbre de Lie des *super-translations* est construite à partir d'une application bilinéaire symétrique $\Gamma : S \otimes S \to W$ envoyant la diagonale dans le cône du futur. L'espace total des super-translations est le super-espace temps et le super-crochet est donné par $\{s_1, s_2\} = -2\Gamma(s_1, s_2)$ sur chaque composante S. La super-algèbre de Lie dite de super-Poincaré est l'extension de l'algèbre de Lie $so(\mathbf{g})$ du groupe de Lorentz, ou plutôt de $so(\mathbf{g}) \otimes (L_d^0 \oplus L_d^+)$, par l'algèbre des super-translations. Elle agit naturellement sur le super-espace-temps.

Les théories de jauge s'étendent en théories de « super-Yang-Mills-Higgs », mais aucune de ces théories n'est évidente à construire. Il était encore plus difficile de construire des théories de *supergravité*. C'est pourtant ce qu'ont fait, pour $D = 4$, $N = 1$, d'abord, Freedman et Van Nieuwenhuizen, Deser et Zumino en 1976 (*cf.* [24]). Puis la « théorie maximale » $D = 4$, $N = 8$, fut construite par E. Cremmer et B. Julia (1979, [10]) en compactifiant, selon une recette Kaluza-Klein super, l'extraordinaire théorie $D = 11$, $N = 1$ de Cremmer, Julia, Scherk (1978, [12]).

Un argument de W. Nahm, joint au parathéorème affirmant que les champs de spin > 2 n'ont pas d'évolution causale, entraîne que, pour la physique, $N = 8$ est un

maximum en $D = 4$ et $N = 1$ est un maximum en $D = 11$. D'où l'importance de la supergravité en dimension 11.

En plus de la métrique $\mathbf{g}$ et de son super-partenaire, qui est un vecteur de spineurs appelé *gravitino*, il faut faire appel à une 3-forme $A_{\mu\nu\rho}$ pour construire cette théorie.

Les théories de supergravité contiennent souvent des champs antisymétriques $F_{\mu\nu}, A_{\mu\nu\rho}, \dots$. Les compactifications à la Kaluza-Klein font apparaître des champs scalaires φ couplés à des formes. Une conséquence des symétries de la fibre compacte et de la dualité de Hodge sur les espaces de solutions est que l'ensemble de ces champs $F, A, \varphi, \dots$ forme des sections de fibrés en *espaces riemanniens symétriques* de Cartan G/H de type non compact (voir [8], [21]).

Par exemple, en compactifiant sur un tore $\mathbb{T}^k$ la supergravité $D = 11$, Cremmer et Julia ont détecté comme espace G/H : pour $k = 1$, $\mathbb{R}_+^\times$; $k = 2$, $(SL_2(\mathbb{R}) \times \mathbb{R}_+^\times)/U_1$; $k = 3$, $(SL_3(\mathbb{R}) \times SL_2(\mathbb{R}))/SO_3 \times U_1$; $k = 4$, SL_5/SO_5 ; $k = 5$, $SO_{5,5}/SO_5 \times SO_5$; $k = 6$, $E_{6(6)}/Sp_4$; $k = 7$, $E_{7(7)}/SU_8$; $k = 8$, $E_{8(8)}/SO_{16}$, à l'aide d'Ehlers pour le dernier cas.

Pour la supergravité $D = 4$, $N = 4$, couplée avec 22 champs de vecteurs A_μ, on trouve $G = SL_2(\mathbb{R}) \times O(6, 22)$, $H = U_1 \times O_6 \times O_{22}$. Pour $D = 4$, $N = 8$, on a vu $G = E_{7(7)}$, $H = SU_8$. Pour $D = 4$, $N = 2$, il y a plusieurs théories, l'une d'entre elles donne $G = E_{7(-25)}$, $H = E_6 \times U_1$. On rencontre les espaces AI, BDI, EI, EV, EVII, EVIII en formes réelles normales (Helgason).

On trouvera les dualités quantiques de champs et de cordes à l'intérieur d'un sous-groupe discret $G(\mathbb{Z})$ du groupe de symétrie G (*cf.* [11], [19], [30], [47], [59], [92]).

2. DUALITÉS QUANTIQUES

2.1. En vertu du principe de correspondance de Bohr, une théorie quantique de champs, abrégée en TQC, est associée à un ensemble $\mathcal{A}$ de champs classiques $A_1, A_2, \dots$, comme des connexions, des métriques ou des spineurs, sur une variété (ou une super-variété) W de dimension D, et à un lagrangien $\mathcal{L}(A_1, A_2, \dots)$ qui est une fonction sur W, dont la valeur en $x \in W$ ne dépend des A_i qu'à travers leurs jets en x.

En mécanique quantique, la densité de probabilité de présence d'une particule est le carré du module d'une fonction d'onde ; en TQC, on suppose que la probabilité de trouver les particules correspondant aux champs $A_1, A_2, \dots$ avec des impulsions-énergies données $p_1, p_2, \dots$ à l'entrée ou à la sortie d'une collision s'obtient à partir de nombres complexes notés $\langle A_1(p_1)A_2(p_2)\cdots\rangle$, appelés *amplitudes*. Il est pratique de rassembler les amplitudes pour toutes les valeurs $p_1, p_2, \dots$ possibles en prenant les transformées de Fourier inverses de ces fonctions des p_i : d'où une distribution notée (abusivement) $\langle A_1(x_1)A_2(x_2)\cdots\rangle$, et interprétée comme amplitude de probabilité de présence. (*Cf.* [31], [48], [63], [77].)

On se donne en plus une famille de fonctions des champs, certaines locales, d'autres non $F(A_1, A_2, \dots)$, les quantités observables, et l'on souhaite calculer leurs amplitudes aussi $\langle F(A)\rangle$. Pour simplifier, nous noterons momentanément A la collection $A_1, A_2, \dots$. Richard Feynman a proposé de faire comme s'il existait une mesure de volume $\mathcal{D}A$ pour laquelle les amplitudes s'écriraient

$$\langle F(A)\rangle = \frac{1}{Z}\int \mathcal{D}A e^{\frac{i}{\hbar}\int_W \mathcal{L}(A)} F(A),$$

où $Z = \int \mathcal{D}(A)\, \exp\Big(\frac{i}{\hbar}\int_W \mathcal{L}(A)\Big)$ s'appelle la *fonction de partition*. L'intégrale sur W de $\mathcal{L}(A)$ est l'*action classique* $S(A)$. Les intégrales suivant $\mathcal{D}A$ portent le nom d'*intégrales de Feynman*.

A priori, cette écriture n'est qu'un guide formel, un programme à réaliser. La théorie de la *renormalisation* a pour but de réaliser ce programme quand c'est possible (*cf.* [5], [43, 44], [48], [63], [72]).

La plupart des résultats sont *perturbatifs* : c'est-à-dire que les physiciens ont dû choisir des paramètres, considérés comme petits, comme la valeur du g dans le $\frac{1}{g^2}$ devant le $\mathcal{L}$ de Yang-Mills, ou $\hbar$, et procéder ordre par ordre en les puissances de g ou de $\hbar$ (voir l'exposé de L. Boutet de Monvel à ce séminaire sur les travaux de Connes et Kreimer). Mais quelques constructions rigoureuses, comme celles des champs conformes, où les calculs numériques impressionnants de E. Wilson sur la discrétisation de la théorie de Yang-Mills justifient l'espoir d'une théorie non-perturbative.

Remarque. — G. Segal ([77]) donne une définition plus précise des TQC, inspirée de son axiomatisation (1987) des théories de champs conformes et de celle donnée par Atiyah (1988) pour les théories topologiques des champs. Dans ces derniers cas, la théorie toute renormalisée peut être définie directement. Le point de vue de Segal fait le lien entre le point de vue, dit constructif, d'une vraie mesure euclidienne $\mathcal{D}A\, \exp(-S(A))$ et la quantification géométrique à la Souriau-Kostant-Weinstein.

Bien qu'on ne puisse pas entrer ici dans les détails, il me semble indispensable d'indiquer en peu de mots l'idée géométrique de la renormalisation ou plutôt comment elle devrait se passer en suivant l'approche de Wilson au début des années 70. En effet, la définition même d'une dualité en TQC et, comme on le verra, en théorie des cordes repose sur les concepts de la renormalisation. Et ne rien savoir de la renormalisation en TQC est à peu près comme ne rien savoir de la notion de limite quand on parle de vitesse et d'équation différentielle ordinaire.

D'abord les pionniers Feynman, Schrödinger, Tomonoga ont découvert qu'il n'y a pas de définition univoque des amplitudes, leur valeur dépend de l'*échelle d'observation*. (Aujourd'hui, l'un des principaux tests de la théorie des interactions fortes (QCD) est l'évolution de la charge renormalisée en fonction de l'échelle de l'énergie mise en jeu.)

Pour les théories de jauge, les principales recettes sont la « troncature », élimination des A de trop grandes et trop petites impulsions, la « régularisation dimensionnelle », déplacement en dimension non entière par prolongement analytique ou la « discrétisation sur réseaux ». Mais, dans un deuxième temps, il faut s'affranchir de ces régularisations arbitraires : des années 40 aux années 70 furent élaborés des procédés pour définir la « partie finie » des amplitudes en adaptant arbitrairement la limite d'un nombre fini de quantités (*cf.* [48], [63], [13], [72]). Au début des années 70, l'approche de Wilson par le groupe de renormalisation a bien éclairci la situation ([86], [87], [88], [5], [64]). Voici la démarche :

Soit $\mathcal{L}$ un lagrangien de champ classique ; par exemple

$$\mathcal{L}(\varphi) = \frac{a}{2}|\nabla\varphi|^2 - \frac{m^2}{2}\varphi^2 + \frac{\lambda}{24}\varphi^4$$

pour la théorie dite φ^4 d'un champ scalaire (juste une fonction à valeur réelle sur l'espace-temps, modèle simplifié du magnétisme si l'on rend le temps et la masse imaginaires purs). Et soit U_0 la variété des paramètres dont $\mathcal{L}$ dépend ; dans l'exemple, U_0 est l'espace de coordonnées (a, m^2, λ), a se nomme l'« intensité de champ », m la « masse » et λ la « constante de couplage » ; dans un $\mathcal{L}$ plus général, on trouve des coordonnées naturelles pour U_0, avec des coefficients de monômes de dérivées des champs ; on les appelle toutes « constantes de couplages » pour simplifier.

L'échelle d'observation des particules (que $A_1, A_2, \dots$ sont censés décrire) est donnée par l'inverse d'une distance ; elle est mesurée par un nombre Λ réel strictement positif, i.e. $\Lambda \in \mathbb{R}_+^\times$. Dans une théorie relativiste quantique, avec des unités où $c = \hbar = 1$, ce nombre Λ fixe aussi l'ordre de grandeur d'une fréquence, d'une énergie, d'une masse ou d'une impulsion.

La théorie achevée de $\mathcal{L}(A)$ doit prévoir des amplitudes $\langle A_1(x_1)A_2(x_2)\cdots A_n(x_n)\rangle$ bien définies à une échelle donnée, pour des valeurs données des constantes de couplages :

une théorie quantique relativiste est décrite par une variété V munie d'un isomorphisme π sur un ouvert de $U_0 \times \mathbb{R}_+^\times$ et d'un feuilletage ρ de dimension 1.

Un point de V représente une collection cohérente d'amplitudes et ρ représente l'évolution des champs quantiques en fonction de Λ. L'application π correspond aux *conditions de renormalisation*, interprétation des constantes de couplages physiques par certaines amplitudes bien choisies.

Par exemple dans la théorie φ^4, la masse (fois i) donne le pôle des fonctions à deux points $\langle\varphi(p_1)\varphi(p_2)\rangle$ en fonction de l'impulsion $p = p_1 - p_2$, et la constante λ donne la valeur d'une amplitude (convenablement normalisée) à quatre points $\langle\varphi(p_1)\varphi(p_2)\varphi(p_3)\varphi(p_4)\rangle$.

Comme le choix d'une unité d'échelle est indifférent, les feuilles de ρ se projettent sur les trajectoires d'un champ de vecteurs β sur U_0 ; si bien que β sur un ouvert de $U_0 \times \mathbb{R}_+^\times$ peut servir à paramétrer les amplitudes de V. Si l'on maintient (V, π, ρ), c'est

pour indiquer qu'un point de V n'est pas un lagrangien de U_0, mais un lagrangien en général bien plus compliqué, non polynomial.

Le champ β, ainsi que son action (simple) sur les amplitudes (Gell-Mann et Low), porte le nom de *flot de renormalisation* ou *groupe de renormalisation*. (Selon la convention la plus répandue, le champ β est orienté dans le sens des énergies décroissantes.)

Donnons une idée de la construction de V, ρ et π, ou plutôt racontons l'histoire comme elle devrait se passer en général. Pour fixer les idées, travaillons avec la régularisation tronquant les grandes impulsions ([64]) ; les autres cas se laissent décrire de manière analogue ([43, 44], [87]). On commence par plonger U_0 dans un espace U assez grand de lagrangiens (en général U est de dimension infinie), en supposant que chaque point de U donne une valeur finie de l'intégrale fonctionnelle tronquée au-delà d'une certaine fréquence. En outre, pour tout couple (Λ_0, Λ_1) dans $\mathbb{R}_+^\times \times \mathbb{R}_+^\times$ tel que $\Lambda_0 < \Lambda_1$, on se donne une fibration $\pi_{\Lambda_1,\Lambda_0}$ de U sur U_0 qui impose la valeur de certaines amplitudes tronquées par Λ_0 à l'échelle d'observation Λ_1. (Dans l'exemple φ^4, les valeurs de a, m^2, λ sont données par le développement en impulsions de l'intégrale du lagrangien jusqu'aux monômes de degré ≤ 4 en φ.) On suppose en plus que, partant de $\mathcal{L}$ dans U, tronquant l'intégrale de Feyman au-dessus de Λ_0 et calculant la moyenne sur les Λ dans l'intervalle $[\Lambda_1, \Lambda_0]$, on obtient les mêmes amplitudes qu'en tronquant simplement à Λ_1 un autre lagrangien $\mathcal{L}' = R_{\Lambda_1,\Lambda_0}(\mathcal{L})$, quitte à faire un changement de variables $x' = f(x)$ et normaliser autrement les champs $A_i' = Z_i A_i$. Ce $\mathcal{L}'$ est le *potentiel effectif* selon Wilson, associé à $\mathcal{L}$, à l'échelle Λ_1 et à la troncature au-dessus de Λ_0.

Le *(semi-)groupe de renormalisation* est l'action multiplicative de $]0,1[$ sur $U \times \mathbb{R}_+^\times$ définie par la formule suivante : $R_\lambda(\mathcal{L}, \Lambda) = (R_{\lambda\Lambda,\Lambda}(\mathcal{L}), \lambda\Lambda)$, pour $\lambda \in]0,1[$, $\mathcal{L} \in U$ et $\Lambda \in \mathbb{R}_+^\times$. On a $R_{\lambda_2,\lambda_1} \circ R_{\lambda_1,\lambda_0} = R_{\lambda_2,\lambda_0}$ donc $R_{\lambda\mu} = R_\lambda \circ R_\mu$, $\lambda < 1$, $\mu < 1$. La variété $U_{\Lambda_1,\Lambda_0} = R_{\Lambda_1,\Lambda_0}(U_0)$ est censée couper transversalement les fibres de $\pi_{\Lambda_1,\Lambda_0}$. Tout cela est espéré pour que la théorie soit renormalisable mais c'est surtout le point suivant qui constitue un « théorème de renormalisation » : lorsque Λ_0 tend vers $+\infty$ la variété U_{Λ_1,Λ_0} converge vers une sous-variété U_{Λ_1} de U et la restriction de $\pi_{\Lambda_1,\Lambda_0}$ à U_{Λ_1,Λ_0} converge vers une application π_{Λ_1} de U_{Λ_1} sur U_0. Pour tout couple $\Lambda_1 < \Lambda_2$, on a $R_{\Lambda_1,\Lambda_2}(U_{\Lambda_2}) = U_{\Lambda_1}$. Les $U_\Lambda \times \{\Lambda\}$ forment une sous-variété V de $U \times \mathbb{R}_+^\times$ que la collection des π_Λ envoie dans $U_0 \times \mathbb{R}_+^\times$; le semi-groupe R induit un flot sur V ; c'est l'invariance d'échelle qui assure que ce flot provient d'un champ de vecteurs β sur U_0. On peut voir V et son feuilletage ρ comme la « vraie théorie » (ou l'ensemble des vraies théories), l'identification à $U_0 \times \mathbb{R}_+^\times$ et le champ β sont des souvenirs du lagrangien classique ; ce sont eux qui expriment qu'une théorie classique a été quantifiée.

Attention : Très peu de théories sont renormalisables au-dessus de $U_0 \times \mathbb{R}_+^\times$ tout entier ; en particulier, les rôles des Λ petits et des Λ grands sont en général dissymétriques dès le début de la construction. Dire que π projette bien V sur un ouvert où Λ tend vers ∞ signifie que la théorie s'analyse bien en ultraviolet, c'est-à-dire pour des distances

de plus en plus petites ; au contraire, au-dessus d'un ouvert où Λ tend vers 0, on décrit le comportement infrarouge, pour les grandes distances.

L'adjectif « renormalisable » est souvent réservé aux théories qui sont renormalisables perturbativement en ultraviolet, cependant en mécanique statistique, qui est une version de la théorie euclidienne des champs quantiques ([69]), la renormalisation se fait surtout en infrarouge comme l'a expliqué K.G. Wilson ([87], [88], [5]). Qu'il existe aussi une théorie perturbative infrarouge en TQC est signalé entre autres par E. Witten (*cf.* [13], p. 1141).

La renormalisabilité perturbative UV dépend des poids des constantes de couplage g vis-à-vis des dilatations d'espace $x \mapsto t^{-1}x$: la théorie est formellement renormalisable si tous les poids sont positifs ou nuls. Exemple : φ^4 pour $D \leq 4$, mais pas $D > 4$; φ^3 pour $D = 6$; φ^6 pour $D = 3$; φ^k quel que soit k pour $D = 2$; Einstein, $D = 2$; Yang-Mills-Higgs, $D = 4$.

La dynamique β sur U_0 a des conséquences importantes pour la dynamique des champs dans l'espace-temps. *A priori* β pourrait engendrer un flot compliqué, cycles limites, attracteurs étranges, *etc.*, mais une conjecture qui tient bon prétend que β est le gradient d'une fonction. D'ailleurs ce sont surtout les points d'équilibre qui ont retenu l'attention.

Les points de U_0 qui sont des attracteurs de β lorsque $\Lambda \to \infty$ sont dits *stables dans l'ultraviolet* (ou UV-stables), ceux qui sont attracteurs quand $\Lambda \to 0$ sont dits *stables dans l'infrarouge* (ou IR-stables). Les limites 0 et ∞ sont les deux racines de l'arc-en-ciel. Lorsqu'un point correspondant à une théorie découplée, sans interaction, est UV-stable, la théorie est dite *asymptotiquement libre* ; c'est alors que la théorie des perturbations est pleinement justifiée. D'après 't Hooft, c'est le cas de la théorie de Yang-Mills non abélienne qui décrit les interactions fortes (QCD), interactions entre quarks, au contraire de l'électrodynamique quantique (QED) où l'absence d'interaction est plutôt IR-stable (le propre des théories de jauge abéliennes). Donc, aux grands moments de transfert (c'est-à-dire Λ grand), le couplage entre quarks et gluons tend vers zéro ; la théorie perturbative peut faire des prédictions. Le problème est à grande échelle (c'est-à-dire Λ petit). Là, logiquement, la théorie des perturbations ne peut plus dire grand chose.

C'est la volonté de comprendre le comportement infrarouge de ces théories qui mène à la dualité.

Il est temps d'indiquer ce qui s'appelle *dualité* dans une théorie quantique des champs :

Soit $\mathcal{C}, \mathcal{C}'$ deux théories de champs renormalisées avec des algèbres d'observables $\mathcal{A}, \mathcal{A}'$ et des ensembles U_0, U_0' de lagrangiens classiques (ou de constantes de couplages) ; une *dualité* entre $\mathcal{C}$ et $\mathcal{C}'$ est une application α d'une partie de $\mathcal{A}$ dans une

partie de $\mathcal{A}'$ et une application φ d'une partie de $U_0 \times \mathbb{R}_+^\times$ dans une partie de $U_0' \times \mathbb{R}_+^\times$ envoyant le flot de β sur celui de β', de telle sorte que les amplitudes se correspondent :

$$\langle \alpha(F(A)) \rangle_{\varphi(u,\Lambda)} = \langle F(A) \rangle_{u,\Lambda}.$$

(Pour une parfaite dualité, on considère la collection complète des amplitudes, mais il peut arriver que la collection se réduise à une fonction de partition.)

Une dualité est un changement de coordonnées ; le plus proche analogue en Analyse classique est la transformation de Fourier $f \mapsto \widehat{f}$. Un exemple d'identité d'amplitude est la formule de Plancherel $\|f\|_2 = C\|\widehat{f}\|_2$. On peut aussi penser aux généralisations en Analyse harmonique non commutative.

Les exemples les plus intéressants échangent une théorie descriptible en UV avec une théorie descriptible en IR.

Les principales dualités sont des « dualités faible/fort », c'est-à-dire qu'elles échangent les couplages faibles de $\mathcal{C}$ avec les couplages forts de $\mathcal{C}'$ et réciproquement.

Par exemple, pour la théorie φ^4 en dimension $D = 2$, le groupe de renormalisation β agissant sur le cadran $\{\lambda \in \mathbb{R}_+, m^2 \in \mathbb{R}_+\}$ possède deux sources $(0,0)$, $(\infty, 0)$ et un col $(\lambda_c, 0)$. Il existe une dualité de la théorie φ^4 avec elle-même qui préserve le col et échange les sources.

Une discrétisation remarquable de φ^4 en dimension 2 (euclidienne) est le système d'Ising décrivant un ensemble de spins classiques $\sigma_{i,j} \in \{\pm 1\}$ aux sommets du réseau $\Gamma = \mathbb{Z}^2$, avec une statistique de Boltzmann $\exp\bigl(-\frac{1}{kT} H(\sigma)\bigr)$ (*cf.* [27], [69]). La théorie duale provient du réseau dual Γ^* dont les sommets sont les centres des faces de Γ, avec les variables de *désordre* $\mu_{i^*,j^*} \in \mathbb{R}$, qu'on peut définir par des formules (non locales) à partir des corrélations des spins (*cf.* [69]). À la température exactement critique $T = T_c$, les corrélations des σ sur Γ coïncident avec les corrélations des μ sur Γ^*. On peut prendre la *limite continue critique* des systèmes d'Ising au voisinage de la température critique : en partant d'une boîte finie de côté n, on fait tendre T vers T_c et n vers ∞ de sorte que $n(T - T_c)$ tende vers une limite finie $\lambda_c^1 - \lambda^1$, et le plan continu est rapporté aux coordonnées $(x, y) = \lim \frac{1}{n}(i, j)$. L'ensemble des corrélations tend vers l'ensemble des amplitudes d'une théorie quantique des champs qui peut être identifiée à la théorie φ^4 (*cf.* Polyakov, Zamolodchikov).

Les formules de la théorie des champs holonomes de Jimbo, Miwa, Sato offrent ainsi l'un des rares cas où il est possible d'expliciter les applications $\alpha : \sigma \mapsto \mu$ et $\varphi : \frac{\lambda}{\lambda_c} \mapsto \frac{\lambda_c}{\lambda}$ de la définition d'une dualité.

Un autre exemple de dualité en dimension 2 a l'intérêt d'échanger magnétisme et électricité : le modèle de Thirring est dual de la théorie de Sine-Gordon. Comme dans le cas d'Ising, il s'agit de théories *complètement intégrables*, c'est-à-dire qu'on connaît des formules fermées pour des amplitudes renormalisées.

2.2. Dualités abéliennes

Quelques exemples simples de dualités quantiques faibles/fortes en petites dimensions exploitent la dualité de Poincaré-Hodge (*cf.* Witten *in* [13]). On peut traiter ces exemples rigoureusement à partir d'intégrales de Feynman gaussiennes, qui se ramènent à des mesures de Wiener ; leur exposé se fait donc dans le cadre des variétés riemanniennes. Les équations de champs classiques associées sont linéaires.

a) Le plus simple est en dimension $D = 2$, sur une surface de Riemann Σ. (Il est connu comme « équivalence $R \longleftrightarrow 1/R$ ».) On considère un champ scalaire unitaire, c'est-à-dire une fonction $\varphi : \Sigma \to S^1 = \mathbb{R}/2\pi\mathbb{Z}$. (On note aussi φ le relèvement aux revêtements universels $\widetilde{\Sigma} \to \mathbb{R}$.) Le lagrangien est $\mathcal{L}(\varphi) = \frac{R^2}{4\pi}\, d\varphi \wedge *d\varphi$. L'équation classique est $\Delta\varphi = 0$, dont les solutions sont les fonctions harmoniques. Localement, si φ est harmonique, il existe une fonction σ à valeurs réelles telle que $*d\sigma = d\varphi$; on a aussi $\Delta\sigma = 0$. Et, à une constante additive près, σ et φ se déterminent l'une l'autre, $*d\varphi = -d\sigma$.

Afin de réaliser au mieux la dualité sous la forme d'une transformation de Fourier, on introduit un espace de champs φ plus grand ; celui de toutes les sections de tous les fibrés en cercle triviaux L au-dessus de Σ. Et à côté de φ et σ, on considère aussi les champs ∇, connexions unitaires sur L ; leur courbure est notée F. Le lagrangien complet est

$$\mathcal{L}(\varphi, \nabla, \sigma) = \frac{R^2}{4\pi}\, \nabla\varphi \wedge *\nabla\varphi - \frac{i}{2\pi}\, \sigma F.$$

Son équation d'Euler-Lagrange donne $F = 0$. Localement l'annulation de F permet d'utiliser ∇ pour trivialiser L et on retrouve dans cette trivialisation l'équation $\Delta\varphi = 0$. L'intégrale fonctionnelle de Feynman justifie cette trivialisation globalement : en intégrant d'abord sur σ, on trouve $F = 0$, mais en intégrant ensuite sur les classes d'équivalence de jauge de ∇, on trouve que ∇ ne peut avoir d'holonomie non triviale.

Surtout, en intégrant d'abord sur φ et ∇ sans toucher à σ (ce qui se fait avec la méthode des fantômes de Faddeev-Popov, *cf.* [33]), on obtient la formule suivante :

$$\int \mathcal{D}\varphi e^{-\frac{R^2}{4\pi}\int d\varphi\wedge *d\varphi} = \Big(\frac{R}{\pi}\Big)^{\chi(\Sigma)} \int \mathcal{D}\sigma e^{-\frac{1}{4\pi R^2}\int d\sigma\wedge *d\sigma},$$

qui généralise le calcul de la transformée de Fourier des gaussiennes en dimension finie.

Cette formule recouvre la propriété modulaire des fonctions thêta, *cf.* le cours de K. Gawędski dans [13].

Mais les calculs ne s'arrêtent pas aux fonctions de partitions : par exemple, (mêmes références) l'observable $e^{i(\varphi(P)-\varphi(Q))}$ pour deux points P et Q de Σ, correspond par dualité à la fonction de partition des champs σ, tels que la forme $d\sigma$ s'étend à $\Sigma\backslash(P \cup Q)$ en une forme fermée avec résidus 1 en P et -1 en Q. On voit que le

dual d'un observable local peut ne pas être local. Il peut arriver aussi qu'il le soit : par exemple le dual de $d\varphi$ est $\frac{1}{iR^2} \star d\sigma$.

b) En dimension 3, la dualité échange les champs scalaires $\varphi : \Sigma \to S^1$ comme ci-dessus, avec les 1-formes A modulo équivalence de jauge $A \sim A + df$. La théorie duale de celle des fonctions harmoniques est la théorie magnétostatique de Maxwell. Le champ qui couple les deux, comme ∇ au a), est une connexion ∇_B sur un fibré L_B dont φ est une section :

$$\mathcal{L}(\varphi, \nabla_B, A) = \frac{R^2}{4\pi} \nabla_B\varphi \wedge *\nabla_B\varphi - \frac{i}{2\pi} A \wedge F_B.$$

Pour les fonctions de partition, on a

$$\int \mathcal{D}\varphi e^{-\frac{R^2}{4\pi}\int d\varphi\wedge *d\varphi} = C \sum_{L_A} \int \mathcal{D}A e^{-\frac{1}{4\pi R^2}\int F_A\wedge *F_A}.$$

L'observable « boucle de Wilson » $e^{i\mu \oint_C A}$, donnée par l'holonomie de A sur un lacet C dans M^3, correspond aux sections φ d'un fibré L_B singulières le long de C, avec une monodromie prescrite $e^{2\pi i\mu}$.

c) Sur une variété W de dimension 4, la dualité abélienne, ou gaussienne, échange une connexion ∇_A sur un fibré en cercle L et une connexion ∇_C sur un fibré N, toutes les deux en théorie électromagnétique de Maxwell. C'est l'étoile de Hodge du § 1, $F_C = *F_A$, qui fournit la dualité. Le champ intermédiaire est une 2-forme ordinaire G.

't Hooft et Polyakov ont montré qu'on gagne beaucoup en ajoutant le terme topologique $\frac{i\theta}{4\pi^2} F_A \wedge F_A$ au lagrangien, qui devient :

$$\mathcal{L}(A) = \frac{1}{2e^2} F_A \wedge *F_A + \frac{i\theta}{4\pi^2} F_A \wedge F_A.$$

L'action associée, si W est compacte, est modifiée par un élément de $2\pi i\mathbb{Z}$. En posant $\tau = \frac{\theta}{\pi} + i\frac{2\pi}{e^2}$, et en introduisant les composantes autoduales et antiautoduales $F_\pm = \frac{1}{2}(F_A \pm *F_A)$, on a

$$\mathcal{L}(A) = \Big(\frac{i\bar\tau}{4\pi}\|F_+\|^2 - \frac{i\tau}{4\pi}\|F_-\|^2\Big) \operatorname{Vol}_{\mathbf{g}}.$$

Mais pour établir une dualité, on introduit à côté de A une 2-forme G et une connexion C sur un fibré en droite N, et on considère le lagrangien

$$\mathcal{L}(A, G, C) = \Big(\frac{i\bar\tau}{4\pi}\|\mathcal{F}_+\|^2 - \frac{i\tau}{4\pi}\|\mathcal{F}_-\|^2\Big) \operatorname{Vol}_{\mathbf{g}} - \frac{i}{2\pi} F_C \wedge G,$$

où $\mathcal{F} = F_A - G$. La dualité s'énonce ainsi : la fonction de partition du côté A avec la constante τ est égale (à une constante multiplicative près) à celle du côté C, mais avec la constante $-1/\tau$. D'après E. Witten :

$$Z(\tau) = \tau^{-\frac{(\chi+\sigma)}{4}} \bar\tau^{-\frac{(\chi-\sigma)}{4}} Z\Big(-\frac{1}{\tau}\Big),$$

où χ et σ sont les nombres d'Euler et la signature de W.

Comme tout est manifestement invariant (sur W compacte) par le changement $\theta \mapsto \theta + 2\pi$ (et invariant par $\theta \mapsto \theta + \pi$ si W est spinorielle), rien ne change par $\tau \mapsto \tau + 2$ (resp. $\tau \mapsto \tau + 1$), et Z est une forme modulaire.

Lorsque $W = M \times \mathbb{R}$, on peut écrire la théorie sous forme hamiltonienne, le moment conjugué π_A de A est l'étoile de Hodge de

$$F_A^\vee = 2\pi i \frac{\partial \mathcal{L}(A)}{\partial F_A} = \frac{2\pi i}{e^2} * F_A - \frac{\theta}{\pi} F_A.$$

La dualité est la transformation de Legendre $F_A^\vee = F_C$, $F_C^\vee = -F_A$. Elle échange les charges électriques et magnétiques ; pour $\Sigma \subset M^3$:

$$Q_e(\Sigma) = \int_\Sigma \frac{F_A^\vee}{2\pi} = \int_\Sigma \frac{F_C}{2\pi}, \quad Q_m(\Sigma) = \int_\Sigma \frac{F_A}{2\pi} = -\int_\Sigma \frac{F_C^\vee}{2\pi}.$$

Le groupe $SL_2(\mathbb{Z})$ agit alors sur l'espace de Hilbert des états associé à M.

2.3. Dualités non abéliennes

Le premier exemple de dualité non linéaire est dû à A.M. Polyakov (1975) : il s'agissait d'une théorie de jauge en dimension 3, avec une connexion ∇ sur un fibré en groupe SO_3 et un champ vectoriel $\vec{s}$ à valeur dans la représentation standard $\mathbb{R}^3$. Un mécanisme de « brisure de symétrie » (*cf.* ci-dessous) fait apparaître un champ de jauge abélien A ; le champ principal de la théorie duale φ provient de solutions particulières $(\vec{s}, \nabla)$, les « monopôles » de la théorie initiale, à travers l'équation $*d\varphi = F_A$. Ce champ acquiert un potentiel transcendant $\cos(2\varphi)$, renormalisable en infrarouge mais pas en ultraviolet (*cf.* [13], tome 2).

Les monopôles de 't Hooft et Polyakov sont des solutions particulières des équations de Yang-Mills-Higgs, version classique du « modèle standard » pour les particules en « interaction faible » ([42]). La théorie standard n'est pas purement celle de Yang et Mills : il y a bien un groupe de jauge G (ici $SU_2 \times U_1$ ou $SU_3 \times SU_2 \times U_1$ si l'on veut tenir compte des quarks et des gluons) et une connexion ∇, mais il y a aussi au moins un champ φ, appelé *champ de Higgs*, section du fibré adjoint en algèbre de Lie $\mathfrak{g}$. Le lagrangien de Yang-Mills-Higgs est

$$\mathcal{L}(A, \varphi) = \frac{1}{4g^2}(\mathrm{Tr}(F_A \wedge *F_A) + |d_A\varphi|^2 + \lambda(|\varphi|^2 - a^2)^2).$$

Cependant, les monopôles les plus étudiés (nous ne parlerons que de ceux-là) sont des solutions statiques des équations de Bogomolny en dimension 3

$$F_A = *d_A\varphi.$$

Ce sont les équations dérivées de $\mathcal{L}$ dans la limite $\lambda \to 0$, introduite par Prasad et Sommerfield (1975). Pour que ces monopôles correspondent à des solutions de Yang-Mills-Higgs dans $\mathbb{R}^4$ d'énergie finie, on impose qu'à l'infini dans $\mathbb{R}^3$ (à temps fixé), on ait dans chaque direction $|\varphi| = a$; si bien que la configuration à l'infini du champ

de Higgs φ fournit un invariant topologique, une classe caractéristique. Par exemple, pour le groupe de jauge $G = SU_2$, φ définit une application de la sphère à l'infini de $\mathbb{R}^3$ dans la sphère de rayon a de l'algèbre de Lie so_3 de SO_3 ; son degré est un nombre entier k qui s'interprète comme une charge magnétique, $b = 2\pi k g^{-1}$. Autrement dit, les plans perpendiculaires aux directions de φ définissent, en dehors d'un compact de $\mathbb{R}^3$, un fibré en plans, de groupe structural SO_2 ; le nombre k est sa classe d'Euler.

On dit alors que φ a *spontanément brisé* le groupe de jauge SU_2 sur le sous-groupe $H = SO_2$. (Attention : à proprement parler aucune symétrie n'est brisée, la symétrie de jauge est intacte.) Suivant B. Julia et A. Zee (1974, [51]) et M.K. Prasad et C.M. Sommerfield (1975, [70]), il existe aussi des *dyons*, devinés par J. Schwinger, possédant à la fois une charge magnétique et une charge électrique ; alors la condition qui remplace $bg \in 2\pi\mathbb{Z}$ s'écrit $e_n b_m - e_m b_n \in 2\pi\mathbb{Z}$, pour chaque paire de particules n et m de charges électriques e_n, e_m et magnétiques b_n, b_m (J. Schwinger, [76], D. Zwanziger, [99], 1968). C'est la condition de quantification généralisant celle que Dirac avait trouvée pour le monopôle magnétique.

Polyakov et 't Hooft (1975, 1976, *cf.* [42], [69]) évaluèrent l'effet des instantons et des monopôles dans la théorie des perturbations : si g est la constante de couplage de jauge, les contributions aux amplitudes calculées par intégrales de Feynman viennent avec un facteur de l'ordre de $\exp(-8\pi^2/g^2)$. 't Hooft et Polyakov remarquèrent aussi l'importance du facteur $\exp(i\theta)$ qui intervient devant les amplitudes. Cet angle θ n'apparaît que dans la combinaison complexe $(\theta/2\pi) + i(4\pi/g^2)$; il vient de la possibilité d'ajouter un terme $\frac{g^2\theta}{32\pi^2}\mathrm{Tr}(F_A^2)$ au lagrangien de Yang-Mills-Higgs. Ce terme est de nature topologique, il ne modifie pas la théorie classique, mais s'avère fondamental pour la théorie quantique.

Dans une théorie de jauge en dimension 4, pour $G = SU_n$ avec des champs de Higgs, mais sans brisure de symétrie, après un fixage de jauge (c'est-à-dire une trivialisation unitaire des fibrés), 't Hooft montre que la théorie renferme des champs vectoriels avec ou sans masse, des champs scalaires et des SU_2-monopôles. En couplage fort, on espère que la théorie se réexprime, dualement, en séries de $1/g^2$ à partir de monopôles. Les physiciens attendent un comportement statistique analogue à celui des supraconducteurs, mais avec les rôles de l'électricité et du magnétisme échangé : plus précisément, ils suspectent un comportement en troupeau de moutons des paires de monopôles. Or il est établi que les champs magnétiques, qui ne sont pas piégés dans les supraconducteurs, en sont rejetés ; c'est ce qui s'appelle l'« effet Meissner ». Dès lors un effet Meissner dual exclurait, ou confinerait, les courants électriques. C'est le scénario de confinement des quarks que 't Hooft proposa.

Justifier ce scénario est un des principaux problèmes de la physique des particules.

't Hooft détailla même la description d'un portrait de plusieurs phases pour la chromodynamique quantique renormalisée ; au moins trois phases : le mode du confinement, dual du mode de Higgs et une « phase de Coulomb » autoduale où l'électricité

et le magnétisme sont symétriques, qui permet des forces de longue portée. La théorie duale pour une théorie de jauge pour SU_n a pour groupe de jauge le groupe adjoint PSU_n.

Plus généralement, on peut tenter l'analyse des théories de Yang-Mills avec tout groupe de jauge compact G et des champs de Higgs ramenant spontanément la symétrie à un sous-groupe H. Goddard, Nuyts et Olive ([32]) ont démontré que la « charge magnétique » des monopôles de 't Hooft et Polyakov se généralise naturellement en une représentation d'un groupe de Lie compact $H^\vee$ qu'on peut construire explicitement à partir de H, le « groupe dual ». On a $(H^\vee)^\vee = H$. Par exemple, $U_1^\vee = U_1$, alors la charge est un entier. La conjecture est qu'il existe une vraie dualité faible/fort échangeant H et $H^\vee$. Les nombres quantiques de Noether (électriques) et topologiques (magnétiques) s'échangeraient. Les monopôles pour H formeraient une théorie de jauge pour $H^\vee$.

Il est très remarquable, mais aussi rarement signalé (A.J. Schwartz l'a pourtant fait), que le groupe dual $H^\vee$ est exactement celui que R.P. Langlands avait défini à la fin des années 60 pour les besoins de l'arithmétique et de la théorie des groupes ([54]). Par exemple, $SU_n^\vee = PSU_n = SU_n/(\mathbb{Z}/n\mathbb{Z})$. La dualité induit la transposition des matrices de Cartan, ainsi les séries A, D, E, F, G sont auto-duales mais B et C s'échangent. Le groupe dual de Langlands intervient dans un vaste ensemble de conjectures reliant l'analyse harmonique, les formes automorphes et la théorie des nombres, pour donner des lois de réciprocité nouvelles. (Langlands parle de groupes réductifs complexes ; ce sont les complexifications des groupes compacts. Il travaille aussi avec des adèles et pas seulement sur $\mathbb{R}$ ou $\mathbb{C}$.) Il serait merveilleux qu'une même théorie mathématique à venir éclaire à la fois la vraie nature des quarks et la correspondance de Langlands (dont un cas particulier traité par A. Wiles a entraîné la démonstration de la conjecture de Fermat).

2.4. En considérant des théories de jauge avec suffisamment de super-symétrie, le scénario du confinement a pu être établi presque entièrement.

Dans les théories de super-Yang-Mills-Higgs, les variables sont une super-connexion $\mathcal{A}$ sur un G-fibré et un super-champ vectoriel Φ dont on note Φ_i, $i \in I$, les super-fonctions coordonnées (*cf.* § 1.4). On décompose $\Phi = \varphi + \psi + F$ suivant L_d^0, L_d^-, L_d^+. Le champ φ est à valeurs dans un multiple de la représentation adjointe $\mathfrak{g}$ de G. Le lagrangien est somme de deux termes $\mathcal{F}(\mathcal{A}, \Phi)$ et $\mathcal{K}(\Phi, \overline{\Phi})$. Le premier, $\mathcal{F}$, est une fonction holomorphe des Φ_i appelée super-potentiel, alors que le second, $\mathcal{K}$, est la fonction génératrice d'une métrique kählérienne $g_{i\bar{j}} = \partial^2 K/\partial\Phi_i \partial\overline{\Phi}_j$ (*cf.* [85], [33]).

Le principal paramètre dont dépend la théorie est issu du coefficient u du terme quadratique en Φ dans $\mathcal{F}$; après renormalisation, $u = C\langle \mathrm{Tr}\varphi^2 \rangle$. Il s'appelle « vide » et il est responsable des brisures de symétrie par un mécanique de Higgs.

Pour $G = SU_2$, spontanément brisé sur U_1, on a $|\varphi| = a$ à l'infini dans $\mathbb{R}^3$, alors $u = a^2$.

Dans $\mathcal{F}$, en facteur du terme quadratique en les super-courbures, on trouve

$$\tau(\Phi) = \frac{\theta}{2\pi} + \frac{4\pi i}{g^2},$$

avec $\theta, g \in \mathbb{R}$. Lorsque $N = 2$, il existe des relations entre τ et $\mathcal{K}$, par exemple la métrique kählérienne sur la droite complexe engendrée par a s'écrit $ds^2 = \frac{1}{2\pi}\Im(\tau\, da d\overline{a})$.

Soit $\widetilde{\mathcal{F}}$ le super-potentiel renormalisé ; on introduit la « variable duale » de a, en posant $a_D = \partial\widetilde{\mathcal{F}}/\partial a$. On a donc $\tau = da_D/da$.

Le groupe de renormalisation β s'exprime par la dépendance de a, a_D, o, en fonction de u et de la fréquence Λ. Seiberg et Witten ([78], [79]) démontrent l'existence d'une série

$$\tau(u, \Lambda) = \frac{i}{\pi}\, \ell n \frac{u}{\Lambda^2} + \sum_0^\infty F_n \frac{\Lambda^n}{u^{2n}}.$$

Donc, quand $\Lambda \to \infty$, on a $a_D \sim 2i(a/\pi)\, ln\,(a/\Lambda) + i\, a/\pi$.

Ainsi la monodromie à l'infini en u de a_D est non triviale, il doit donc exister une singularité au moins dans le plan des u. Seiberg et Witten montrent que la positivité de ds^2 force l'existence d'un groupe de monodromie global non abélien, d'où l'existence d'au moins deux singularités $\pm u_c$. Une bonne normalisation des singularités pour rendre compte de la limite $\Lambda \to \infty$ est $u_c = \Lambda^2$.

Toujours selon Seiberg et Witten, a et a_D sont des périodes de la forme différentielle holomorphe $\frac{1}{2\pi}(x-u)\frac{dx}{y}$ sur la courbe elliptique d'équation $y^2 = (x^2 - \Lambda^4)(x-u)$.

Plus généralement, pour un groupe de jauge compact simple G de rang r, la théorie de champs renormalisée est décrite par une courbe complexe Σ de genre r variant au-dessus d'un tore de dimension r (la « branche de Coulomb »). Le groupe de renormalisation fait évoluer les couplages, interprétés comme des périodes d'une forme holomorphe, en fonction de Λ et des coordonnées sur le tore, suivant le système complètement intégrable associé à G par Hitchin (*cf.* [13]).

La limite de basse énergie en un point singulier u_c est la théorie duale infrarouge de la théorie de jauge $N = 2$. Elle s'appelle « théorie de Seiberg-Witten ». Ses champs fondamentaux, un champ de spineurs et une connexion, sont des dégénérescences de masse 0 d'états quantiques massifs très particuliers qui existent pour tout $u \neq \pm u_c$: les *états BPS*.

En effet, on trouve dans la théorie super-YMH, $N = 2$, des super-monopôles magnétiques et des super-dyons de charge électrique Q_e et de charge magnétique Q_m, qui sont invariants par la moitié des super-symétries ; on les appelle états BPS pour Bogomolny, Prasad, Sommerfield. Leur masse satisfait à $M^2 = C^2(Q_e^2 + Q_m^2)$. Dans cette théorie, on n'a qu'une action projective de l'algèbre de super-Poincaré $N = 2$ ([61], [98]), c'est-à-dire une action linéaire d'une extension centrale de la super-algèbre : pour un super-dyon, la charge centrale est égale à $Z = aN_e + a_D N_m$, où $N_e = Q_e - \frac{\theta}{2\pi}Q_m$ et $N_m = Q_m$ sont des nombres entiers en vertu de l'effet θ de Witten ([91]). Ces

nombres Z forment un réseau dans $\mathbb{C}$ lorsque $u \neq \pm u_c$. Le groupe de monodromie de (a, a_D) agit sur Z ; il est d'indice fini dans $SL_2(\mathbb{Z})$.

C'est seulement pour $N = 4$, avec deux super-symétries de plus, que A. Sen [80] a pu démontrer l'existence d'états BPS pour N_e, N_m premiers entre eux quelconques, établissant ainsi la conjecture de dualité $N = 4$ de Montonen et Olive [58].

Notons $N_m = r$ et $N_e = s$. L'espace des modules de monopôles de charge magnétique r pour les équations de Bogomolny sur $\mathbb{R}^3$ s'écrit

$$\mathbb{R}^3 \times (S^1 \times \mathcal{M}_r^0)/(\mathbb{Z}/r\mathbb{Z}),$$

où $\mathcal{M}_r^0$ est une variété hyperkählérienne de dimension $4r - 4$, l'espace des monopôles réduits, sur laquelle le sous-groupe $\mathbb{Z}/r\mathbb{Z}$ de S^1 agit librement ([3]). D'après Witten, Manton et Sen ([80]), les états quantiques de la théorie $N = 4$ sont des formes différentielles sur $\mathcal{M}_r^0$; la conjecture de Montonen et Olive se ramenait à prouver l'existence d'une unique forme harmonique L^2 (de degré $2r - 2$) sur laquelle l'action de $t \in \mathbb{Z}/r\mathbb{Z}$ est la multiplication par $\exp(-2\pi i\, ts/r)$.

A. Sen a d'abord traité le cas de $r = 2$, les autres cas ont été résolus peu après par Segal, Selby, Parrati. Pour G simple compact quelconque, *cf.* [26], [29], [55].

3. DUALITÉS DE CORDES

3.1. Depuis plus de 20 ans, la théorie des cordes, et surtout celle des supercordes, était considérée comme la meilleure chance pour unifier toutes les interactions avec la gravitation. Elle avait éclipsé les tentatives de champs unitaires, en particulier la supergravité $D = 11$. Mais depuis 6 ou 7 ans, à la suite des découvertes de dualités étranges dans le monde des cordes (Schwarz, Sen, Duff, Hull-Townsend, Witten,...), le paysage a changé ; une mystérieuse théorie M en dimension 11 (ou peut-être F en dimension 12), rassemblant des cordes qui se propagent, mais aussi des membranes (surfaces dans l'espace, variétés de dimension 3 dans l'espace-temps), des trous noirs et des variétés spéciales de toutes dimensions, s'est dévoilée. Ses limites les plus intéressantes sont des supercordes, mais aussi une de ses limites est la supergravité en dimension 11. Ainsi c'est difficile de dire qui a mangé l'autre, de la corde ou de la supergravité.

L'idée des cordes est d'attribuer aux particules très énergétiques la forme d'une ficelle d'une longueur de l'ordre de 10^{-32} cm, et de « quantifier deux fois » les mouvements et les accidents de cette ficelle relativiste dans l'espace-temps de Minkowski pour retrouver, sans contradiction, toute la physique connue. D'où vient cette idée ? Quel est son rapport avec la gravitation ? Pour le comprendre, reportons-nous à la plus naïve des questions qui se pose en théorie quantique des champs : comment un champ devient-il une particule ? En effet, dans quelle approximation passe-t-on des amplitudes des connexions et des sections de fibrés spin à des particules ayant

des trajectoires comme des projectiles ? La relation se fait en deux temps, en « deux quantifications » : une première quantification mène de l'action du point matériel (sans masse) dans l'espace de Minkowski à l'équation de Klein-Gordon $d * d\varphi = 0$; c'est l'équation de Schrödinger de la particule. La deuxième quantification fait passer à une assemblée d'oscillateurs, pour tenir compte des créations et annihilations.

Afin de rendre compte des interactions, il est possible d'introduire un potentiel, par exemple φ^4, ce qui donne une théorie de champ scalaire. Pour les particules chargées, les photons ou les super-points, les interactions peuvent s'expliquer en couplant les équations de Dirac, de Maxwell et d'Einstein. De même les quarks, les bosons W et Z, les gluons sont soumis au couplage de Dirac et de Yang-Mills. L'exigence de la renormalisation sélectionne sévèrement le choix des interactions mais laisse encore un goût d'arbitraire.

Avec les cordes, la vision des interactions est changée : dans une limite où $\hbar$ serait négligeable, mais pas la longueur typique des cordes, mesurée par la constante α', inverse de la *tension de la corde*, on aurait affaire à des objets possédant une dimension d'espace et une de temps. Les trajectoires sont remplacées par des cylindres. Et, pour décrire les *interactions* de cordes, il suffit de remplacer le cylindre par des surfaces dont le bord a plus de deux composantes.

Le nombre n de trous compte le nombre de cordes à l'entrée et à la sortie de l'accident. On montre que le genre k de la surface (son nombre d'anses) mesure l'ordre des *corrections quantiques* dans les amplitudes de cordes quantiques.

La deuxième bonne nouvelle apportée par les cordes est la présence naturelle de particules véhiculant des interactions gravitationnelles : les *gravitons*. La première quantification des théories des cordes est une théorie de champ bidimensionnelle qui incorpore comme condition de renormalisation, au premier ordre en α',la théorie de la gravitation d'Einstein classique à côté de la théorie de Yang-Mills classique dans l'espace ambiant. D'ailleurs la constante α' vaut $\hbar\frac{G}{c^3}$, où G est la constante de gravitation universelle de Newton.

Mais aux ordres plus élevés en α' des corrections arrivent qui pourraient bien régulariser la gravitation aux grandes énergies et guider sa quantification. Notons que les théories de cordes conservent un sens, au moins perturbatif, si l'on remplace l'espace-temps plat de Minkowski par des variétés lorentziennes munies de métriques d'Einstein, i.e. solutions des équations d'Einstein classiques.

Une autre « prédiction » est la *super-symétrie* : sans elle, c'est-à-dire sans passer aux *supercordes*, des particules aberrantes violant les principes de la relativité, se propageant plus vite que les photons, s'imposeraient subrepticement.

Enfin, et surtout, déjà la première quantification des cordes et supercordes est une théorie quantique de champ très exigeante : si l'on veut y maintenir, au niveau quantique, les symétries de la corde, ou supercorde classique (à savoir le groupe de (super) Poincaré ambiant, les difféomorphismes de la surface source, les dilatations locales de

sa métrique (transformations de Weyl), les super-symétries de super-surface), il n'y a que 5 possibilités, toutes en dimension 10 d'espace-temps. Celles-ci portent les noms suivants : Type I, Type IIA, Type IIB, Hétérotique SO_{32}, Hétérotique $E_8 \times E_8$.

3.2. Suivant Polyakov (*cf.* [69], [2], [6], [33]), une théorie de cordes à valeurs dans la variété W, de dimension D, équipée d'une métrique lorentzienne $(D-1,1)$, est composée de deux champs en dimension 2 : une métrique sur Σ (de signature $(1,1)$) et une application X de Σ dans W ; le lagrangien est l'*énergie* $-\frac{1}{2\alpha'}\|\nabla X\|^2$.

Lorsque W est l'espace plat de Minkowski, un choix convenable des coordonnées ramène X à $D-2$ solutions de l'équation des ondes (les coordonnées transverses à $\Sigma = S^1 \times \mathbb{R}$ ou $[0,1] \times \mathbb{R}$), et deux constantes (*cf.* [33], [13]). La première quantification donne une algèbre de Weyl et sa représentation de Fock, et les symétries, la reparamétrisation des branches du cône de lumière fournissent une représentation ρ de l'algèbre de Virasoro $\mathrm{Vir} \otimes \overline{\mathrm{Vir}}$ dans l'espace de Fock.

Si N est l'opérateur de nombre, la masse d'un état pur est donnée par une formule $\alpha' M^2 = C_0 + N$, où C_0 est une constante à déterminer. L'invariance lorentzienne externe (de W) force $C_0 = -1$ et l'invariance conforme (de Σ) force $D = 26$. Alors, en décomposant la représentation ρ en facteurs irréductibles (les champs primaires) (*cf.* [27]), on peut identifier les vecteurs de masse 0 ; ils forment diverses représentations irréductibles de $SO(D-1,1)$: on y trouve les espaces de sections des fibrés en $S^2, \Lambda^2, \mathbb{R}$, c'est-à-dire les 2-formes symétriques, $G_{\mu\nu}$, les 2-formes antisymétriques $B_{\mu\nu}$ et un champ scalaire Φ.

Les conditions de renormalisation du premier ordre en α' s'expriment comme des équations aux dérivées partielles sur $G_{\mu\nu}$, $B_{\mu\nu}$, Φ ; c'est la surprise : on tombe sur une pertubation des équations d'Einstein, celle qui dérive du lagrangien :

$$\mathcal{L}(G,B,\Phi) = \frac{1}{2\kappa^2} e^{-\Phi}(R_G + 4|\nabla\Phi|^2 - \frac{1}{12}|dB|^2).$$

On voit que $\alpha' \sim \kappa^2 e^{2\langle\Phi\rangle}$, donc Φ, nommé le *dilaton*, fixe la « constante de couplage » α'. C'est la morale des cordes : toutes les constantes y deviennent des champs.

La présence d'une particule de masse imaginaire (le « vertex » ordinaire) rend la théorie peu physique, mais cela ne l'empêche pas d'être riche de propositions mathématiques.

Avec les *supercordes*, les physiciens semblent avoir eu plus de chance, et ils ont fait un cadeau encore plus beau aux mathématiciens.

Il y a *a priori* deux définitions différentes : celle dite de Green et Schwarz, où une super-surface de Riemann s'envoie dans un espace de super-Minkowski modelé sur une puissance du fibré des spineurs S^N, et celle dite de Neveu-Schwarz-Ramond, où la super-surface s'envoie dans le super-espace modelé sur le fibré tangent $T(W)$. Mais, compte tenu des symétries, les deux approches s'avèrent équivalentes.

Pour maintenir les invariances super-Poincaré et super-conformes, il faut $D = 10$; en cette dimension, l'espace des spineurs de Dirac est de dimension 32. À côté des vecteurs de l'espace transverse à la corde, on trouve les spineurs de l'espace de dimension 8, S^+, S^- tous deux de dimension 8 aussi. (La *trialité* est l'équivalence par automorphisme extérieur des trois représentations.)

Pour décrire les solutions classiques, on doit tenir compte de plusieurs possibilités pour les ondes de spineurs $\psi(z, \bar{z}, \theta, \bar{\theta}) \in S^N$

1) elles se propagent soit à droite soit à gauche,

2) elles appartiennent à l'un ou l'autre des deux fibrés spin sur S^1.

On met un indice + pour les ondes se déplaçant vers la droite, c'est-à-dire les fonctions de z et un indice − pour celles qui se déplacent vers la gauche, c'est-à-dire les fonctions de $\bar{z}$. La seconde distinction sépare l'espace des multiples de $(zdz)^{\frac{1}{2}}$, noté R (comme Ramond) de l'espace des multiples de $(dz)^{\frac{1}{2}}$, noté NS (comme Neveu-Schwarz).

Toutes les possibilités sans tachyons ont été classifiées (c'est le contenu du Théorème de projection GSO, Goddard, Sent, Olive).

Le type I, avec une super-symétrie $N = 1$ sur W (mais toujours 32 pour Σ). Ses états de masse 0, déterminés par la représentation des algèbres super-Virasoro, appelés *champs effectifs*, sont une métrique $G_{\mu\nu}$ (NS $\otimes$ NS), un dilaton Φ (NS $\otimes$ NS), une 2-forme anti-symétrique $B_{\mu\nu}$ ($R \otimes R$), un champ de covecteur-spineur ξ_μ^α, le *gravitino*, et un champ de spineurs λ_α, le *dilatino*. C'est tout pour les cordes fermées. Mais, en type I, on permet des cordes ouvertes, il s'ajoute alors un champ de jauge A_μ pour le groupe SO_{32} et un champ de gaugino ψ^σ.

Le type IIA, $N = 2$, champs effectifs : G, B, Φ, tous trois $\mathrm{NS}_+ \otimes \mathrm{NS}_+$, ξ_1, λ_1 ($\mathrm{R}_+ \otimes \mathrm{NS}_+$), ξ_2, λ_2 ($\mathrm{NS}_+ \otimes \mathrm{R}_-$), plus une 1-forme A_μ et une 3-forme $A_{\mu\nu\rho}$, de type $\mathrm{R}_+ \otimes \mathrm{R}_-$.

Le type IIB (dit *chiral*), $N = 2$, champs effectifs : G, B, Φ ($\mathrm{NS}_+ \otimes \mathrm{NS}_+$), ξ_1, λ_1 ($\mathrm{NS}_+ \otimes \mathrm{R}_-$), ξ_2, λ_2 ($\mathrm{NS}_+ \otimes \mathrm{R}_+$), mais trois formes paires (toutes de type $\mathrm{R}_+ \otimes \mathrm{R}_+$) : A^0 l'*axion*, $A^2_{\mu\nu}$ une 2-forme, et une 4-forme $A_{\mu\nu\rho\sigma}$, avec la restriction $dA = *dA$.

En plus de ces trois possibilités, fut inventée la théorie des *cordes hétérotiques* ; une mise en forme géométrique fait appel à une notion de super-feuilletage dans une super-variété de dimension $4|n$ (*cf.* [52]) ; elle correspond à une corde (bosonique) pour les champs allant à droite, et à une supercorde (fermionique) pour les champs allant à gauche. Elle vient sous deux formes, avec des champs de Yang-Mills supplémentaires, soit SO_{32}, soit $E_8 \times E_8$. Ses champs effectifs sont G,B,Φ,A (de jauge), uniquement NS, et χ,λ,ψ, de type R, comme pour le type I.

Lorsqu'on pratique une réduction à la Kaluza-Klein pour descendre en dimension < 10, les cinq types n'en donnent plus que trois : I, II, Het.

Les lagrangiens effectifs, qui sont les limites effectives de basses énergies des supercordes, ont été construits : ils donnent les théories de supergravité à 10 dimensions.

Par exemple, pour le type I, la partie bosonique de l'action est

$$S_{\mathrm{I}} = \frac{1}{16\pi\alpha'^4}\int \mathrm{Vol}_{10}\left(e^{-\Phi}(R_G + |\nabla\Phi|^2) - \frac{1}{12}|H_3|^2 - \frac{1}{4}e^{-\frac{\Phi}{2}}|F_2|^2\right),$$

où F_2 est la courbure du champ de Yang-Mills SO_{32} et $H_3 = dB_2$.
Pour les théories hétérotiques, avec mes mêmes conventions,

$$S_{\mathrm{Het}} = \frac{1}{16\pi\alpha'^4}\int \mathrm{Vol}_{10}\, e^{-\Phi}(R_G + |\nabla\Phi|^2 - \frac{1}{12}|H_3|^2 - \frac{1}{4}|F_2|^2).$$

Pour la théorie de type IIA,

$$S_{\mathrm{IIA}} = \frac{1}{16\pi\alpha'^4}\int \mathrm{Vol}_{10}\left(e^{-\Phi}(R_G+|\nabla\Phi|^2-\frac{1}{12}|H_3|^2)-\frac{1}{4}|F_2|^2-\frac{1}{48}|F_4|^2\right)+\frac{1}{2}B_2\wedge F_4\wedge F_4\,,$$

où $H_3 = dB_2$, $F_2 = dA_1$, $F_4 = dA_3$ et $F_4' = F_4 + A_1 \wedge H_3$.
Pour la théorie de type IIB,

$$S_{\mathrm{IIB}} = \frac{1}{16\pi\alpha'^4}\int \mathrm{Vol}_{10}\left(e^{-\Phi}(R_G + |\nabla\Phi|^2 - \frac{1}{12}|H_3|^2) - \frac{1}{2}|\nabla A^0|^2 - \frac{1}{12}|H_3' + A^0 H_3|^2 - \frac{1}{240}|F_5|^2\right) + A_4 \wedge H_3 \wedge H_3'\,,$$

où $H_3 = dB_2, H_3' = dB_2'$ et $F_5 = dA_4 + B_2' \wedge H_3$.

3.3. Si l'on ajoute aux cinq théories toutes les théories de cordes en dimension plus petite obtenues par compactification de Kaluza-Klein, on a une grande famille qui a beaucoup embarrassé les chevaliers de la théorie unique de tout. L'espoir est revenu lorsque Witten, Hull, Townsend, Duff, Sen et al., entre 1995 et 1996, ont observé que, par dualité, toutes ces théories avaient l'air d'être équivalentes ([47], [92], [66]).

Afin de préserver une partie des super-symétries, les compactifications ont surtout été faites le long de tores $\mathbb{T}^k$, de produits $M \times \mathbb{T}^{k-4}$ avec M une surface $K3$ ou de produits $Y \times \mathbb{T}^{k-6}$ avec une variété de Calabi-Yau Y de dimension complexe 3. Ces variétés possèdent des espaces de modules (ne serait-ce que le rayon d'un cercle, ou le volume) qui forment pour chaque théorie $\mathcal{T}$ un espace $U_{\mathcal{T}}$. Cet espace se projette sur des axes de constantes de couplages, en particulier celui de α' ou $\langle\Phi\rangle$.

Une *dualité* entre deux théories de cordes $\mathcal{T}_1$ et $\mathcal{T}_2$ est une bijection f d'une partie $\mathcal{O}_1$ des observables de $\mathcal{T}_1$ sur une partie $\mathcal{O}_2$ des observables de $\mathcal{T}_2$, et une bijection f^* d'une partie V_1 de $U_{\mathcal{T}_1}$, sur V_2 dans $U_{\mathcal{T}_2}$ qui échangent les amplitudes, c'est-à-dire $\langle f(F_1)\rangle_{f^*(u_1)} = \langle F_1\rangle_{u_1}$.

À proprement parler, il n'y a que des *conjectures de dualité* mais elles ont franchi des tests qui portent, soit sur les théories de supergravités effectives, soit sur les solutions classiques invariantes par une partie des super-symétries, encore appelées solutions BPS (*cf.* [81]).

Certaines dualités échangent les couplages forts et faibles des cordes ; ce sont les *S-dualités* ([75], [19], [20], [47], [59], [60]).

Exemples. — Pour $D = 10$, type I $\longleftrightarrow$ Het_{32}, pour $D = 6$, IIA/$K3$ $\longleftrightarrow$ Het/$\mathbb{T}^4$ (i.e. la théorie de type IIA compactifiée sur une surface K3 est duale de la théorie hétérotique compactifiée sur un tore $\mathbb{T}^4$), pour $D = 10$, IIB $\longleftrightarrow$ IIB.

D'autres dualités respectent les couplages faibles et peuvent être détectées sur la théorie perturbative ; ce sont les *T-dualités*, T comme target ([25], [30], [39]).

Exemples. — IIA $\longleftrightarrow$ IIB si on compactifie sur des cercles de rayons inverses (i.e. IIA/S^1_R $\longleftrightarrow$ IIB/$S^1_{R^{-1}}$). Dans les mêmes conditions $\mathrm{Het}_{E_8\times E_8}$ $\longleftrightarrow$ $\mathrm{Het}_{SO_{32}}$.

Les T-dualités constituent des sous-groupes discrets des symétries signalées à la fin du § 1. Par exemple Het/$\mathbb{T}^6$ a un $O(6,22;\mathbb{Z})$ de T-dualités dans le $O(6,22;\mathbb{R})$. Mais dans ce cas, il y a aussi un $SL_2(\mathbb{R})$ mieux caché de S-dualités agissant sur $\tau = \langle A + ie^{-\Phi}\rangle$, A l'axion, Φ le dilaton. Dedans, un $SL_2(\mathbb{Z})$ passe aux cordes quantiques. On retrouve la dualité de Sen (§ 2.4) comme une des conséquences de basse énergie.

Pour la théorie de type II compactifiée sur $\mathbb{T}^6$, le groupe E_7 des symétries de la super-gravité $N = 8$ contient un sous-groupe discret $E_7(\mathbb{Z})$ dont les éléments sont appelés U-dualités, formé à partir d'un $O(6,6;\mathbb{Z})$ de T-dualités et du $SL_2(\mathbb{Z})$ de S-dualités de la théorie IIB en dimension 10.

La difficulté à chaque fois est de construire les états BPS, analogues des dyons, avec des charges quantifiées.

C'est là que les *branes* apparaissent :

Dans toutes les théories de supercordes, l'examen non perturbatif (mais spéculatif) a révélé l'existence de solutions (semi-)classiques particulières électriquement et/ou magnétiquement chargées, jouant le rôle des instantons, des solitons et des monopôles en théorie des champs : les *p-branes*. Ainsi nommées car elles définissent soit comme lieu singulier, soit comme centre, des sous-variétés de dimension $p + 1$ de l'espace-temps, donc des sous-variétés dimension p dans l'espace (au moins pour les solitons et les monopôles). Les 2-branes s'appellent simplement des membranes.

Sur ces variétés, on peut intégrer les $(p+1)$-formes différentielles du lagrangien de supergravité effective ou leurs formes duales.

La dualité de Poincaré accompagne les dualités de cordes et associe des q-branes duales aux p-branes. La duale d'une p-brane est donc une $(D - p - 4)$-brane.

Par exemple, avec $D = 10$, en théorie IIA, la corde est une 1-brane et donne une 5-brane. En théorie IIB, on trouve une 3-brane autoduale.

Les théories de supercordes induisent sur les p-branes des théories de jauge dont les constantes de couplage proviennent de la géométrie en dimension D.

Les branes associées comme courants aux formes différentielles de type NS $\otimes$ NS furent assez faciles à construire ; on trouve par exemple les 0-branes associées aux dilatons Φ de toutes les théories, une 1-brane pour Het, courant de la jauge A, duale d'une 5-brane, des membranes associées aux formes B_2 des théories de type II et Het.

La construction de courants pour les formes $\mathrm{R} \otimes \mathrm{R}$ fut plus longue à venir. J. Polchinski ([65], [67]) les trouva sous formes de $\mathcal{D}$-branes, $\mathcal{D}$ comme Dirichlet ; ce sont des solutions singulières, comme des déchirures et des trous noirs dans l'espace-temps, le long desquelles les cordes, *a priori* toutes fermées de IIA et IIB peuvent s'ouvrir. Elles ont permis de compléter la liste BPS conjecturée pour la S-dualité de IIB/S^1 ou la T-dualité du type I. En théorie de type IIA (resp. IIB), il y a une p-brane qui est $\mathcal{D}$-brane pour tout entier pair entre 0 et 8 (resp. impair entre -1 et 9). La T-dualité échange les $\mathcal{D}$-branes de IIA avec celles de IIB. En théorie de type I, il y a des $\mathcal{D}$-branes pour $p = 1, 5, 9$ (*cf.* [4], [8], [18], [30]).

La dernière révolution fut la découverte de la M-théorie (Townsend, Witten, 1996, [47], [92]) : en cherchant à comprendre le couplage fort de IIA, on trouve quelque chose de dual en dimension 11 dont la limite de basse énergie est la supergravité $D = 11$ de Cremmer, Julia, Scherk !

Ce « quelque chose » fut baptisé M-théorie, comme membrane, merveille ou mystère. Cette théorie est encore largement conjecturale, sa fonction principale est de compléter le diagramme suivant :

$$\begin{array}{ccc} \text{M-théorie} & \longrightarrow & \text{Supercorde IIA} \\ \downarrow & & \downarrow \\ \text{Supergravité, } D = 11 & \longrightarrow & \text{Supergravité IIA, } D = 10 \end{array}$$

où les flèches horizontales sont des compactifications sur un cercle S^1 et les flèches verticales sont les limites de basse énergie ([16], [17], [46], [59], [81]).

Par compactification sur un intervalle compact, on trouve la duale de $\mathrm{Het}_{E_8 \times E_8}$ en dimension 10. La M-théorie possède une membrane et une 5-brane duale.

En compactifiant une théorie de supercorde ou la M-théorie sur une variété X de dimension k, on peut enrouler une p-brane sur un d-cycle de X et obtenir dans $\mathbb{R}^{D-k}$ une $(p-d)$-brane ([1], [16], [19], [20], [82], [95]).

Ainsi la corde IIA est la membrane de M enroulée sur S^1, la 2-brane de IIA est la projection de la membrane, sa 4-brane est la 5-brane de M enroulée et sa 5-brane non singulière est la projection de la 5-brane de M. Il y a aussi en M-théorie des branes noires analogues de trous noirs de dimensions 2, 4 et 6 ([16], [34]).

Les ressources mathématiques de la M-théorie semblent inépuisables :

1) Selon Strominger, Yau, Zaslow, toute variété de Calabi-Yau X de dimension 3 sur $\mathbb{C}$ contient une famille à trois paramètres réels de 0-branes pour IIA compactifiée sur X. Si on applique la dualité T, on trouve une famille de 3-branes pour IIB formant la variété miroir Y de X. *Cf.* [36], [39], [41], [45], [53], [83].

2) Selon Witten, les états BPS de la théorie super-Yang-Mills $N = 2$, $D = 4$ qui permettent d'établir la dualité électrique-magnétique quantique et de construire la

théorie de Seiberg-Witten viennent de configurations de branes de IIA qui correspondent par dualité à des 2-branes, de la forme $D^2 \times \mathbb{R}$ en M-théorie accrochées sur des 5-branes de la forme $\Sigma \times \mathbb{R}^4$ dans $S^1 \times \mathbb{R}^{10}$ ([97], [96]).

Par un glissement de sens, l'ultime théorie dont les limites variées sont IIA, IIB, I, supergravité $D = 11$, $\mathrm{Het}_{E_8 \times E_8}$, $\mathrm{Het}_{SO_{32}}$, s'est aussi appelée M. Schwarz et Sen ont bien proposé la lettre U comme unité, mais notre époque est plus mystique.

3.4. Paracelse avait deviné la théorie M en 1525 au cours de ses recherches sur l'Astrologie. Il n'a pas dit M comme quoi. Sans doute comme Mysterium. Il a même parlé du M.m. en précisant plus loin M.magnum.

Écoutons d'abord Lucien Braun dans son commentaire récent ([7]) à la traduction du Volumen Paramirum ([62]) :

« ...La recherche sur l'*ens astrale* le conduit à s'interroger, à partir de l'Astre, sur les conditions d'existence de la vie humaine. L'homme est incapable de penser sa propre origine, estime Paracelse. Dès 1525, il se heurte, dans sa réflexion têtue, au « toujours déjà-là », au primordial, au primexistant : avant tout surgissement d'un existant quelconque, quelque chose est toujours déjà donné d'où il procède. Pour le faire entendre, Paracelse se sert de l'air comme comparaison (ailleurs aussi appelé *chaos*) et comme métaphore (qui doit nous *conduire plus loin*). Nous respirons l'air, écrit-il, sans lui nous ne pouvons exister. En filant la métaphore suffisamment loin, l'air lui-même finit par procéder de quelque chose que décidément on ne peut plus nommer – même par métaphore ! Paracelse désigne cet « abyssal-toujours-premier » par la lettre M (qui est mystère). Toujours déjà il y a quelque chose, avant tout donné, qui est la condition d'apparition de tout donné possible ; et qui est réfractaire à toute analyse subséquente. »

Écoutons Paracelse lui-même avant que, quelques lignes plus loin, il ne parle du poison contenu dans le M.m. qui risque d'infecter le corps de l'homme :

« ... Vous dites que si l'air n'existait pas, toutes choses dépériraient, disparaîtraient ; tout ce qui vit étoufferait et mourrait. On peut dire de même : il existe quelque chose qui maintient et conserve le corps (qui contient la vie) ; la perte de ce « quelque chose » serait aussi dommageable pour le corps que serait pour le vivant le défaut de l'air. Il convient donc de le préserver. L'air lui-même est conservé dans et par ce « quelque chose », et si cela venait à manquer, l'air aussi disparaîtrait. Le firmament en vit également, et si cela n'était, il disparaîtrait. Ce « quelque chose », je l'appelle M. Il n'existe rien de plus éminent dans tout le domaine de la création, et pour un médecin il n'y a rien de plus utile à méditer.

Je m'efforce de vous l'expliquer : le M. ne provient pas du firmament ; il n'y est pas né ; ce n'est pas le firmament qui nous l'envoie, etc. Rien de tout cela ne vaut. Mais le M. maintient toutes les créatures en leur être, que ce soit dans le ciel ou sur la terre ; tous les éléments vivent en lui et par lui... ».

Remerciements. — L'auteur est seul responsable des fautes dans le texte, mais il tient à remercier ceux qui l'ont aidé à entrer dans le sujet : C. Bachas, A. Bahraini, J. Dunaud, B. Julia, O. Maspfuhl, S. Paycha, M. Slupinski et J.-B. Zuber.

RÉFÉRENCES

[1] O. Aharony – « String theory dualities from M-theory », *Nuclear Physics B* **476** (1996), p. 470–483.

[2] I. Antoniadis, E. Cremmer & K.S. Stelle – « Les supercordes », in *Gazette des Mathématiciens*, Soc. Math. France, 2001, 1ère partie, n° 87, janvier ; 2ème partie, n° 88, avril.

[3] M.F. Atiyah & N.J. Hitchin – *The geometry and dynamics of magnetic monopoles*, Princeton U.P., 1988.

[4] C. Bachas – « Lectures on D-branes », in *Duality and supersymmetric theories* [60].

[5] G. Benfatto & G. Galavotti – *Renormalization group*, Physics Notes, vol. 1, Princeton University Press, 1995.

[6] J.-B. Bost – « Fibrés déterminants, déterminants régularisés et mesures sur les espaces de modules des courbes complexes », in *Sém. Bourbaki*, Astérisque, vol. 152-153, Soc. Math. France, 1987, exp. n° 676, p. 113–149.

[7] L. Braun – *Paracelse, de l'Astrologie*, Presses Universitaires de Strasbourg, 2002.

[8] P. Breitenlohner, D. Maison & G. Gibbons – « 4-dimensional black holes from Kaluza-Klein theories », *Comm. Math. Physics* (1988), p. 295–333.

[9] R. Coquereaux & A. Jadczyk – « Symmetries of Einstein-Yang-Mills fields and dimensional reductions », *Comm. Math. Physics* **98** (1985), no. 1, p. 79–104.

[10] E. Cremmer & B. Julia – « The $SO(8)$ supergravity », *Nuclear Physics B* **159** (1979), p. 141–212.

[11] E. Cremmer, B. Julia, H. Lü & C. Pope – « Dualisation of dualities », *Nuclear Physics B* **523** (1998), p. 73–144.

[12] E. Cremmer, B. Julia & J. Scherk – « Supergravity theory in 11 dimensions », *Physics Letters B* **76** (1978), p. 409–412.

[13] P. Deligne, P. Etingof & al. (éds.) – *Quantum fields and strings : A course for mathematicians*, AMS, IAS, 1999.

[14] P.A.M. Dirac – « Quantised Singularities in the Electromagnetic Field », *Proc. Roy. Soc. A* **133** (1931), p. 60–72.

[15] ———, « The Theory of Magnetic Poles », *Phys. Review* **74** (1948), p. 817–830.

[16] M.J. Duff – « M-theory (the theory formerly known as strings) », *International Journal of Modern Physics A* **11** (1996), no. 32, p. 5623–5641.

[17] ———, « The world in eleven dimensions : a tribute to Oskar Klein », `arXiv: hep-th/0111237`, 38 pp, 2001.

[18] M.J. Duff, R.R. Khuri & J.X. Lu – « String solitons », *Physics Reports* **259** (1995), p. 213–326.

[19] M.J. Duff, J.T. LIU & R. Minasian – « Eleven-dimensional origin of string/string duality : a one-loop test », *Nuclear Physics B* **452** (1995), p. 261–282.
[20] M.J. Duff, J.T. Liu & J. Rahmfeld – « Four-dimensional string/string/string triality », *Nuclear Physics B* **459** (1996), p. 125–159.
[21] M.J. Duff, B.E.W. Nilsson & C.N. Pope – « Kaluza-Klein supergravity », *Physics Reports* **130** (1986), no. 1 & 2, p. 1–142.
[22] M.J. Duff & K.S. Stelle – « Multi-membrane solutions of $D = 11$ supergravity », *Physics Letters B* **253** (1991), p. 113–118.
[23] J. Ehlers – « Exterior solutions of Einstein's gravitational field equations admitting a two-dimensional abelian group of isometric correspondences », in *Colloque sur la théorie de la relativité 1959*, Centre Belge Rech. Math., 1959, p. 49–57.
[24] S. Ferrara, J. Scherk & B. Zumino – « Algebraic properties of extended supergravity theories », *Nuclear Physics B* **121** (1977), p. 393–402.
[25] A. Font, L.E. I. nez, D. Lüst & F. Quevedo – « Strong-weak coupling duality and non-perturbative effects in string theory », *Physics Letters B* **249** (1990), p. 35–43.
[26] J. Gaunlett – « Supersymmetric monopoles and duality », in *Duality and supersymmetric theories* [60].
[27] K. Gawędski – « Conformal field theory », in *Sém. Bourbaki*, Astérisque, vol. 177-178, Soc. Math. France, 1989, exp. n° 704, p. 95–126.
[28] R. Geroch – « A method for generating solutions of Einstein's equations », *Journal of Math. Physics* **12** (1971), no. 6, p. 918–924.
[29] G.W. Gibbons – « The Sen conjecture for fundamental monopoles of distinct types », *Physics Letters B* **382** (1996), p. 53–59.
[30] A. Giveon, M. Porrati & E. Rabinovici – « Target space duality in string theory », *Physics Reports* **244** (1994), p. 77–202.
[31] J. Glimm & A. Jaffe – *Quantum physics, a functional integral point of view*, Springer-Verlag, 1981.
[32] P. Goddard, J. Nuyts & D. Olive – « Gauge theory and magnetic charge », *Nuclear Physics B* **125** (1977), p. 1–28.
[33] M.B. Green, J.H. Schwarz & E. Witten – *Superstring theory*, vol. 1 & 2, Cambridge University Press, 1987.
[34] R. Güven – « Black p-brane solutions of $D = 11$ supergravity theory », *Physics Letters B* **276** (1992), p. 49–55.
[35] A. Hanany & E. Witten – « Type IIB superstrings, PBS monopoles, and three-dimensional gauge dynamics », *Nuclear Physics B* **492** (1997), p. 152–190.
[36] J.A. Harvey & G. Moore – « Algebras, BPS states, and strings », *Nuclear Physics B* **463** (1996), p. 315–368.
[37] S. Hawking & G. Ellis – *The large scale structure of space-time*, Cambridge University Press, 1973.
[38] W. Heisenberg – *Les principes de la théorie de quanta*, Gauthier-Villars, 1972.
[39] M. Henningson & G. Moore – « Counting curves with modular forms », *Nuclear Physics B* **472** (1999), p. 518–528.

[40] N. Hitchin – « The Yang-Mills equation and the topology of 4-manifolds (after Simon K. Donaldson) », in *Sém. Bourbaki*, Astérisque, vol. 105-106, Soc. Math. France, 1983, exp. n° 606, p. 167–178.

[41] ———, « Lectures on special Lagrangian submanifolds », `arXiv:math.DG/9907034` v1, 6/7/99.

[42] G. 't Hooft – *Under the spell of the gauge principle*, Advanced Series in Mathematical Physics, vol. 19, World Scientific, 1994.

[43] ———, « The renormalization group in QFT », in *Under the spell of the gauge principle* [42].

[44] G. 't Hooft & M. Veltman – « Diagrammar », in *Under the spell of the gauge principle* [42].

[45] K. Hori & C. Vafa – « Mirror Symmetry », `arXiv:hep-th/00022V3`, mars 2000.

[46] P. Hořava & E. Witten – « Heterotic and Type I string dynamics from eleven dimensions », *Nuclear Physics B* **460** (1996), p. 506–524.

[47] C.M. Hull & P.K. Townsend – « Unity of superstring dualities », *Nuclear Physics B* **438** (1995), p. 109–137.

[48] C. Itzykson & J.B. Zuber – *Quantum field theory*, McGraw-Hill, 1980.

[49] B. Julia – « Dualities in the classical supergravity limits », `arXiv:hep-th/9805083`.

[50] ———, « Magics of M-gravity », `arXiv:hep-th/0105031` v1.

[51] B. Julia & A. Zee – « Poles with both magnetic and electric charges in non-Abelian gauge theory », *Phys. Rev. D* **11** (1975), p. 22–27 à 22–32.

[52] A. Konechny & A.S. Schwartz – « On $(k \oplus \ell|q)$-Dimensional Supermanifolds », in *Supersymmetry and Quantum Field Theory, Proc. Kharkov, Ukraine 1997* (J. Weiss & V.P. Akulov, éds.), L. N. in Physics, vol. 509, Springer, 1998, p. 201–206.

[53] M. Kontsevich – « Mirror symmetry in dimension 3 », in *Sém. Bourbaki*, Astérisque, vol. 237, Soc. Math. France, 1996, exp. n° 801, p. 275–293.

[54] R.P. Langlands – « Problems in the theory of automorphic forms », in *Lectures in Modern Analysis and Applications III*, Lect. Notes in Math., vol. 170, Springer, 1970.

[55] K. Lee, E. Weinberg & P. Yi – « Moduli space of many BPS monopoles for arbitrary gauge groups », *Phys. Rev D* **54** (1996), p. 1633–1643.

[56] Y.I. Manin – *Gauge field theory and complex geometry*, Springer-Verlag, 1984.

[57] C.W. Misner, K.S. Thorne & J.A. Wheeler – *Gravitation*, W.H. Freeman and Co., San Francisco, Calif., 1973.

[58] C. Montonen & D. Olive – « Magnetic monopoles as gauge particles », *Physics Letters B* **72** (1977), p. 117–120.

[59] N.A. Obers & B. Pioline – « U-duality and M-theory », *Physics Reports* **318** (1999), p. 113–225.

[60] D. Olive & P. West (éds.) – *Duality and supersymmetric theories*, Publications of the Newton Institute, Cambridge University Press, 1999.

[61] H. Osborn – « Topological charges for $N = 4$ supersymmetric gauge theories and monopoles of spin 1 », *Physics Letters B* **83** (1979), p. 321–326.

[62] T. Paracelse – *Volumen (Medicinae) Paramirum*, Sudhoff, 1525.

[63] M.E. Peskin & D.V. Schroeder – *An introduction to Quantum Field Theory*, Addison-Wesley Publishing Company, 1995.

[64] J. Polchinski – « Renormalization and effective lagrangians », *Nuclear Physics B* **231** (1984), p. 261–295.

[65] ———, « Combinatorics of boundaries in string theory », *Physical Review D* **50** (1994), p. 6041–6045.

[66] ———, « String duality », *Reviews of Modern Physics* **68** (1996), p. 1245–1258.

[67] ———, « Dirichlet-branes and Ramond-Ramond charges », `arXiv:hep-ph/9510017` V3, novembre 1995.

[68] J. Polchinski & E. Witten – « Evidence for heterotic-Type I string duality », *Nuclear Physics B* **460** (1996), p. 525–540.

[69] A.M. Polyakov – *Gauge fields and strings*, Contemporary concepts in Physics, vol. 3, Harwood Academic Publishers, 1987.

[70] M.K. Prasad & C.M. Sommerfield – « Exact Solution for the 't Hooft Monopole and the Julia-Zee Dyon », *Phys. Rev. Letters* **35** (1975), p. 760–762.

[71] S.-J. Rey – « Confining phase of superstrings and axionic strings », *Physical Review D* **43** (1991), p. 526–538.

[72] V. Rivasseau – *From perturbative to constructive renormalization*, Princeton Series in Physics, Princeton University Press, 1991.

[73] A. Rogers – « A global theory of supermanifolds », *J. Math. Phys.* **21** (1980), no. 6, p. 1352–1365.

[74] A.S. Schwartz – « On the definition of superspace », *Theoret. and Math. Phys.* **60** (1984), no. 1, p. 657–660.

[75] J.H. Schwarz & A. Sen – « Duality symmetries of $4D$-heterotic strings », *Physics Letters B* **312** (1993), p. 105–114.

[76] J. Schwinger – « Magnetic Charge in Quantum Field Theory », *Phys. Rev.* **144** (1966), p. 1087–1093.

[77] G. Segal – « Notes on quantum field theory », `http://online.itp.ucsb.edu/online/geom99/segal1_n/`.

[78] N. Seiberg & E. Witten – « Electric magnetic duality, monopole condensation and confinement in $N = 2$ supersymmetric Yang-Mills theory », *Nuclear Physics B* **426** (1994), p. 19–52.

[79] ———, « Monopoles, duality and chiral symmetric breaking in $N = 2$ supersymmetric QCD », *Nuclear Physics B* **431** (1994), p. 484–550.

[80] A. Sen – « Dyon monopole bound states, self-dual harmonic forms on the multi-monopole moduli space, and $SL_2(\mathbb{Z})$ invariance in string theory », *Physics Letters B* **329** (1994), p. 217–221.

[81] ———, « An introduction to non-perturbative string theory », in *Duality and supersymmetric theories* [60], p. 297–413.

[82] D. Sorokin – « Superbranes and superembeddings », *Physics Reports* **329** (2000), p. 1–101.
[83] A. Strominger, S.-T. Yau & E. Zaslow – « Mirror symmetry is T-duality », *Nuclear Physics B* **479** (1996), p. 243–259.
[84] P.K. Townsend – « The eleven-dimensional supermembrane revisited », *Physics Letters B* **350** (1995), p. 184–188.
[85] J. Wess & J. Bagger – *Supersymmetry and supergravity*, 2ème éd., Princeton University Press, 1992.
[86] K.G. Wilson – « Renormalization of a scalar field theory in strong coupling », *Physical Review D* **6** (1972), no. 2, p. 419–426.
[87] ———, « The renormalization group : Critical phenomena and the Kondo problem », *Review of Modern Physics* **47** (1975), p. 773–840.
[88] ———, « The renormalization group and critical phenomena », *Reviews of Modern Physics* **55** (1983), no. 3, p. 583–600.
[89] E. Witten – « Overview of K-theory applied to strings », `arXiv:hep-th/0007175` v1.
[90] ———, « Deconstruction, G_2 holonomy, and Doublet-Triplet splitting », `arXiv:Hep-ph/0201018` V2, 15.01.2002.
[91] ———, « Dyons of charge $e\theta/2\pi$ », *Physics Letters B* **86** (1979), p. 283–287.
[92] ———, « String theory dynamics in various dimensions », *Nuclear Physics B* **443** (1995), p. 85–126.
[93] ———, « Strong coupling and the cosmological constant », *Modern Physics Letters A* **10** (1995), p. 2153–2155.
[94] ———, « Bound states of strings and p-branes », *Nuclear Physics B* **460** (1996), p. 335–350.
[95] ———, « Five-branes and M-theory on an orbifold », *Nuclear Physics B* **463** (1996), p. 383–397.
[96] ———, « Solutions of four-dimensional field theories via M-theory », *Nuclear Physics B* **500** (1997), p. 3–42.
[97] ———, « Duality, spacetime and quantum mechanics », *Physics today* (May 1997), p. 28–33.
[98] E. Witten & D. Olive – « Supersymmetry algebras that include topological charges », *Physics Letters B* **78** (1978), p. 97–101.
[99] D. Zwanziger – « Exactly Soluble Non-relativistic Model of Particles with Both Electric and Magnetic Charges », *Phys. Rev.* **176** (1968), no. 5, p. 1480–1495.

Daniel BENNEQUIN

Université Denis-Diderot (Paris VII)
Équipe de Géométrie et Dynamique
Institut de Mathématique de Jussieu
UMR 9994 du CNRS
2 place Jussieu
F-75251 Paris Cedex 05

Séminaire BOURBAKI Mars 2002
54e année, 2001-2002, nO 900, p. 149 à 165

ALGÈBRE DE HOPF DES DIAGRAMMES DE FEYNMAN, RENORMALISATION ET FACTORISATION DE WIENER-HOPF
[d'après A. Connes et D. Kreimer]

par **Louis BOUTET de MONVEL**

1. INTRODUCTION

1.1. Renormalisation

Cet exposé a pour objet de décrire le point de vue de A. Connes et D. Kreimer sur la renormalisation en théorie quantique des champs (théorie perturbative), expliquées dans les articles [CK1, CK2]. La plupart des démonstrations sont omises, et on en trouvera les détails dans *loc. cit.*. Je remercie chaleureusement A. Connes, J. Zinn-Justin, A. Arabia pour leur aide et leurs conseils. Je remercie particulièrement P. Cartier qui a relu et corrigé ces notes dans la meilleure tradition de N. Bourbaki.

Un des effets frappants de la théorie de la renormalisation est de faire prévoir que les constantes de structure qui gouvernent un grand système physique dépendent de l'échelle d'observation. Ce fait se constate aussi dans d'autres phénomènes étendus où tous les ordres de grandeur contribuent, en particulier dans la description des transitions de phase pour lesquelles aussi l'idée de la renormalisation est pertinente (*cf.* [ZJ]). Il s'observe déjà dans des phénomènes plus classiques et connus depuis le 19ème siècle : ainsi une balle de ping-pong de masse $m = 2$gr, immergée dans l'eau (masse d'eau déplacée : $M = 32$gr environ), subit une poussée vers le haut correspondant au poids $M - m = 30$gr (force d'Archimède moins son poids). Mais, si on la lâche, elle ne remonte pas avec une accélération de $\frac{M-m}{m}g = 15g$: il y a une interaction avec l'eau environnante et le résultat final est que, tant que la vitesse reste petite, la balle remonte avec une accélération de $\frac{M-m}{m+\frac{M}{2}}\ g < 2g$. Vu de l'extérieur, tout se passe comme si la balle était affectée de la masse apparente $m' = m + \frac{M}{2}$, et si l'on n'avait pas le moyen de disséquer en détail l'interaction qui a lieu près de la balle, m' serait la seule « masse effective » qu'on puisse lui attribuer à cette échelle.

La théorie quantique des champs veut décrire et expliquer les phénomènes fondamentaux qui gouvernent la physique à l'échelle nucléaire ou particulaire. Dès le départ, elle se heurte à des difficultés considérables car les quantités « naturelles » qu'on veut

calculer sont décrites par des analogues d'intégrales qui n'ont pour l'instant pas beaucoup de sens mathématique, ou dans le cas plus simple de la théorie perturbative, par des séries asymptotiques d'intégrales divergentes[(1)].

Un des objets de la renormalisation est d'attribuer à ces intégrales une valeur (« partie finie »). Les parties finies d'intégrales interviennent dans beaucoup de questions d'E.D.P. et Hadamard en a donné une description systématique par troncature ou par prolongenment analytique. Ici, elles arrivent en série et il faut les organiser de façon cohérente. Les théoriciens du champ ont énoncé pour cela un certain nombre de règles qui, pour être précises et efficaces, n'en sont pas moins très compliquées.

A. Connes et D. Kreimer montrent que ces règles peuvent être résumées très élégamment par une factorisation à la Wiener-Hopf (ou Riemann-Hilbert) d'un lacet à valeurs dans un groupe associé aux diagrammes de Feynman, et que cela donne lieu à des formules « universelles » qui vivent dans le groupe « universel » des difféomorphismes formels tangents à l'identité.

1.2. Champs

Dans la théorie classique, un champ[(2)] est une fonction A sur $\mathbf{R}^d$ qui satisfait une équation

$$\partial^2 A - F'(A) = 0$$

où ∂^2 désigne le d'Alembertien $-\partial_t^2 + \sum \partial_{x_i^2}$, et F' est la dérivée d'une fonction donnée F. Il revient au même de dire que A extrémise l'intégrale d'action $\int d^d x\ \mathcal{L}(A)$ où le lagrangien est $\mathcal{L}(A) = \frac{1}{2}(\partial A)^2 + F(A)$; d est la dimension de l'espace-temps ; dans les formules finales c'est un entier D, $D = 4$ dans notre monde, $D = 6$ ci-dessous, mais de toute façon on devra « le faire varier continûment » pour la régularisation dimensionnelle.

Un champ quantique ϕ est une fonction (généralisée) à valeurs opérateurs, satisfaisant à un certain nombre d'axiomes pour lesquels je renvoie à la bibliographie. Il est déterminé par un état vide (d'énergie minimale) noté $\langle 0\rangle$, et les « fonctions de Green »

$$G(x_1, \dots, x_N) = \langle 0 | T\phi(x_1) \dots \phi(x_N) | 0\rangle\,,$$

(le signe T signifie qu'on réordonne les facteurs x_i par ordre de temps $t = x^0$ décroissant).

---

(1) Il ne s'agit en aucune façon d'une théorie floue ou imprécise puisque les résultats théoriques qu'elle fournit concordent aussi bien que possible avec les résultats expérimentaux — jusqu'à 14 chiffres significatifs en électrodynamique quantique — ce qui est un exploit inégalé par la physique antérieure. C'est plutôt un exemple de plus du fait que les résultats les plus spectaculaires dans la description du monde qui nous entoure précèdent souvent les mathématiques qui les expliquent le plus élégamment.

(2) Pour simplifier, nous ne considérerons ici qu'un champ scalaire, i.e. A est une fonction numérique ; la physique utilise aussi des champs vectoriels, tensoriels, spinoriels, etc.

Pour quantifier une théorie classique, les physiciens s'inspirent de la formulation variationnelle et de la mécanique statistique, et affirment que la fonction de Green doit être décrite par une intégrale fonctionnelle sur l'espace de tous les champs classiques :

$$G(x_1,\dots,x_N) = \mathcal{N}^{-1}\int [dA]\ e^{iS(A)} A(x_1)\dots A(x_N)$$

où $S(A) = \int d^dx\ \mathcal{L}(A))$ est l'intégrale d'action, et $\mathcal{N} = \int [dA] e^{iS(A)}$ est un facteur de normalisation assurant $\langle 0|0\rangle = 1$. Dans cette intégrale aucun des termes $\mathcal{N}, [dA], \int$ n'a de sens, du moins pris séparément ; le tout évoque une intégrale gaussienne comme celles du mouvement brownien, mais la notation des physiciens est plus suggestive.

Dans la théorie perturbative, l'action est $\mathcal{L}(A) = \mathcal{L}_0(A) + \mathcal{L}_{\rm int}(A)$ où

$$\mathcal{L}_0 = \frac{1}{2}(\partial A)^2 - \frac{m^2}{2}A^2 \tag{1}$$

est le lagrangien du champ libre de masse m, et $\mathcal{L}_{\rm int}(A)$ est le terme perturbatif. Le développement en série (formelle) de l'exponentielle suggère

$$G(x_1,\dots,x_N) \sim \mathcal{N}^{-1}\sum \frac{i^n}{n!}\int [dA] e^{iS_0(A)}\,(S_{\rm int}(A))^n\,A(x_1)\dots A(x_N)$$

avec

$$\mathcal{N} \sim \sum \frac{i^n}{n!}\int [dA] e^{iS_0(A)} (S_{\rm int}(A)^n.$$

Dans ces expressions interviennent une foison d'intégrales « gaussiennes » de la forme

$$\sum \int [dA] e^{iS_0(A)}\ S_{\rm int}(A)^n\ A(x_1)\dots A(x_N) \tag{2}$$

qui se ramènent à des intégrales en dimension finie.

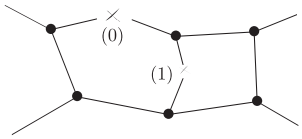
Dans [CK1, CK2] et dans beaucoup d'autres travaux, les auteurs ont choisi comme exemple le lagrangien correctif $\mathcal{L}_{\rm int}(A) = \frac{g}{3!}A^3$ en dimension $d = 6$ (« théorie ϕ^3 »), parce que cela illustre déjà de façon simple, mais significative, le cas général, même si cela ne correspond pas à une théorie physique vraisemblable. Nous ferons de même ici.

2. INTÉGRALES ET DIAGRAMMES DE FEYNMAN

2.1. Diagrammes de Feynman

Les intégrales de Feynman (2) et les problèmes que pose leur divergence sont connus depuis le début de la théorie (Dirac). Le génie de Feynman l'a conduit à les indexer par les diagrammes qui portent son nom, qui en plus d'indiquer clairement l'intégrand et l'espace où l'on intègre, évoquent de façon imagée des interactions ou collisions virtuelles.

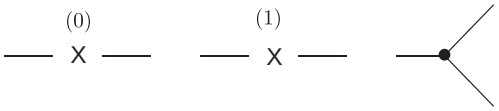
Un diagramme de Feynman est un graphe coloré (non orienté)[3] : il a un ensemble de sommets colorés, reliés par des arêtes.



Les arêtes extérieures, ou pattes, n'ont qu'un seul sommet ; les arêtes intérieures ont deux sommets distincts. Les types de sommets se distinguent par le nombre d'arêtes, et une couleur (pour les sommets à deux arêtes).

Dans la théorie ϕ^3, il y a 3 types de sommets :

(3)



correspondant respectivement aux termes quadratiques A^2, $(\partial A)^2$ et au terme cubique A^3 de $\mathcal{L}(A)$.

2.2. Intégrales de Feynman

Un tel diagramme Γ code une distribution U_Γ de N variables vectorielles $(x_1, \dots, x_N)$, N le nombre de pattes, qu'on repère plutôt par sa transformée de Fourier $U_\Gamma(p_1, \dots, p_N)$ des moments extérieurs p_e, qui vérifient des relations de conservation $\sum p_e = 0$ (une pour chaque composante connexe). Celle-ci est définie comme suit : à chaque arête i est associée une variable de moment p_i à d dimensions, libre si i est une patte extérieure, muette (variable d'intégration, notée k_i dans [CK1]) sinon. U_Γ est alors définie par une intégrale sur l'espace des moments internes (ou plutôt sur un sous-espace, car l'intégrand comporte des facteurs δ) :

$$U_\Gamma(p_1, \dots, p_N) = \int \prod d^d p_i \;\; I_\Gamma(p_i). \tag{4}$$

L'intégrand I_Γ est un produit de facteurs auxquels contribuent séparément les arêtes intérieures et les sommets, par des règles explicites suivantes (pour un autre $\mathcal{L}_{\text{int}}$ il faudrait en général plus de types de sommets (ou d'arrêtes), et compléter les règles qui décrivent les facteurs associés aux sommets ou aux arêtes).

I_1 Chaque arête i produit un facteur $\frac{1}{m^2+p_i^2}$.

I_2 Chaque sommet v produit un facteur $\delta(\sum_{i \ni v} p_i^v)$ (cela fait autant de relations linéaires que de sommets, exprimant la conservation des moments)[4].

[3] on est aussi amené à colorer les arêtes, et à orienter le graphe, s'il y a plusieurs espèces de particules et qu'on doit distinguer entre particules et anti-particules

[4] p_i^v désigne le flux de moment entrant dans v indexé par l'arête i : si v, v' sont les deux sommets d'une arête i, on a $p_i^{v'} = -p_i^v$.

I_3 Chaque 2-sommet de type 0 produit un facteur m^2.

I_4 Chaque 2-sommet de type 1 produit un facteur p^2 (les deux moments associés sont opposés et ont même p^2).

I_5 Chaque 3-sommet produit un facteur $\mu^{3-d/2}g$.

μ est l'unité de masse — pour l'instant une variable formelle — introduite pour que les constantes qui figurent dans les intégrales à renormaliser soient des nombres purs, « sans dimension ». (Il s'agit cette fois de la dimension des physiciens : les mesures sont éléments des puissances tensorielles d'un espace vectoriel (ou sections d'un fibré vectoriel) de base. Dans une théorie scalaire relativiste comme ici, celui-ci est de rang 1, et la dimension d'une quantité physique est l'exposant de la puissance tensorielle où on la mesure ; de façon équivalente, toutes les quantités (masse, moment,...) de la théorie sont commensurables à une puissance de l'unité de masse μ ; la « dimension » d'une quantité P est l'exposant s tel que $\mu^{-s}P$ soit un nombre pur).

On peut résumer l'effet des facteurs δ en disant que les moments extérieurs vérifient la relation de conservation $\sum p_e = 0$ et que l'intégrale est effectuée sur l'espace E^i_Γ de dimension $(I - V + C)d$ des moments intérieurs indépendants (I est le nombres d'arêtes internes, V le nombre de sommets, C le nombre de composantes connexes).

Ces intégrales sont le plus souvent divergentes — doublement : à distance finie elles ont des singularités car la norme utilisée est celle de la relativité : $p^2 = -p_0^2 + \sum p_i^2$; les physiciens s'empressent de la remplacer par la norme euclidienne en faisant une « rotation de Wick » (i.e. en remplaçant le temps $t = x^0$ par un temps imaginaire pur it), ce qui élimine ces singularités (sauf pour le cas limite $m = 0$) ; le cas relativiste s'en déduit en principe par prolongement analytique, et cela ne change de toute façon pas grand'chose au principal problème de la renormalisation, qui vient des divergences à l'infini de ces intégrales (« catastrophe ultra-violette »).

2.3. Parties finies

Hadamard nous a montré comment on peut attribuer une valeur à une intégrale divergente. Par exemple, pour une singularité à l'origine, on tronque et on enlève une partie infinie « évidente » dans un développement asymptotique :

$$\mathrm{pf}\int_0^1 dt\ t^{s-1}\varphi(t) = \lim_{\varepsilon\to 0}\int_\varepsilon^1 dt\ t^{s-1}\varphi(t) + \sum \frac{\varphi^{(j)}(0)}{j!}\ \frac{\varepsilon^{s+j}}{s+j}.$$

Il y a un problème supplémentaire dans le cas — le plus utile — où l'exposant s est entier.

Un autre moyen « canonique » est de remarquer que le membre de gauche est une fonction méromorphe de s, avec ici des pôles simples aux entiers négatifs. Ce prolongement redonne la partie finie pour s régulier. Aux pôles, on retient le terme constant du développement de Laurent ; c'est un peu moins canonique, car il dépend du choix de la borne supérieure d'intégration, ou d'une unité de longueur.

Une des méthodes utilisées par les physiciens, qui est celle qu'utilisent Connes et Kreimer, est la méthode de régularisation dimensionnelle, qui est une méthode de prolongement analytique : pour une fonction $f(p)$ invariante par rotation sur $\mathbf{R}^d$, cela revient à intégrer en coordonnées polaires :

$$\int d^d p\ f(p) = v_d \int_0^\infty dr\ r^{d-1} f(r),$$

où $v_d = 2\pi^{d/2}/\Gamma(d/2)$ est le $(d-1)$-volume de la sphère unité, de sorte qu'on ait $\int d^d x\ e^{-x^2} = \pi^{d/2}$. Le résultat est une fonction méromorphe de d.

La régularisation dimensionnelle peut être effectuée pour les intégrales de Feynman. Elle conduit pour chacune à une distribution $U_\Gamma = U_\Gamma(p, \mu, d)$ des moments extérieurs, qui dépend du paramètre μ et, de façon méromorphe, de la dimension d, avec un pôle pour $d = D$ (en général multiple, mais d'ordre fini ; ici $D = 6$; il peut y avoir d'autres pôles en dehors de D).

Remarque 2.1. — Les valeurs (renormalisées) des intégrales de Feynman sont intéressantes pour elles-mêmes : en dehors de puissances de π, on trouve (entre autres) des combinaisons rationnelles de $\zeta(3), \zeta(5), \ldots$ et plus généralement de nombres multizêtas dont le degré de complexité (poids, profondeur) est lié à la topologie du diagramme ; mais ceci est une autre histoire.

3. ALGÈBRE DE HOPF DES DIAGRAMMES DE FEYNMAN

3.1. Diagrammes habillés et coproduit

Dans la théorie ϕ^3, les diagrammes utiles (nous dirons « admissibles ») sont réunion disjointe de diagrammes « simples », i.e. connexes, et qui restent connexes si on leur enlève une arête (le sigle des physiciens est 1PI = irréductibles à une particule) ; comme pour les groupes, on compte que les objets de la fig. (3), les plus « simples », ne le sont pas.

Pour fabriquer des nombres à partir des diagrammes et des intégrales de Feynman, et surtout pour décrire la règle de réinsertion de diagrammes pour le calcul de ces intégrales par itération d'intégrations partielles, il convient de les « habiller ». Un diagramme « habillé » est une combinaison de diagrammes admissibles :

$$\sum \sigma_\Gamma \otimes \Gamma,$$

où chaque coefficient σ_Γ est une distribution sur l'espace E_Γ des moments extérieurs p_e (indexés par les N arêtes extérieures, et liés par la relation de conservation $\sum p_e = 0$, de sorte que E_Γ est de dimension $(N-1)d$).[(5)]

[(5)] L'espace des coefficients distribution n'est pas précisé et cela n'a guère d'importance ; on peut se limiter à des distributions invariantes par le groupe de Lorentz, et il faut en outre qu'on puisse

Les diagrammes admissibles, habillés ou non, forment une algèbre commutative graduée $\mathcal{H}$: la loi de produit est la somme disjointe, et $\otimes$ pour les coefficients, l'unité est le diagramme vide (affecté du coefficient 1). La graduation est définie par le nombre de boucles $L = I + V - C = \dim H^1(\Gamma)$, voir ci-dessous § 3.2.

Connes et Kreimer définissent sur $\mathcal{H}$ une structure d'algèbre de Hopf au moyen du coproduit

$$\Delta(\Gamma) = \Gamma \otimes 1 + 1 \otimes \Gamma + \sum_{\gamma \subset \Gamma, \iota} \gamma_{(\iota)} \otimes \Gamma/\gamma_{(\iota)}. \tag{5}$$

Dans cette somme, γ désigne un diagramme admissible non vide, inséré dans Γ, dont les composantes ont deux ou trois pattes : les ensembles de sommets des composantes de γ dans Γ sont deux à deux disjoints, seules des arêtes extérieures peuvent être communes.

ι est une collection de signes 0 ou 1 indexée par les composantes à deux pattes.

Le deuxième facteur, $\Gamma/\gamma_{(\iota)}$ s'obtient en contractant en un point chaque composante α, et en l'affectant de la couleur $\iota(\alpha)$ (0 ou 1) s'il s'agit d'une composante à deux pattes.

Dans le premier facteur les composantes de $\gamma_{(\iota)}$ ont deux ou trois pattes, et un habit fixé par ι : $\sigma_0 \otimes \gamma_\alpha$ si γ_α est une composante à trois pattes, ou deux pattes et la couleur est $\iota(\alpha) = 0$, $\sigma_1 \otimes \gamma_\alpha$ si γ_α a deux pattes et la couleur correspondante est 1. Les distributions $\sigma_{0,1}$ sont choisies de sorte que $\sigma_0(am^2 + bp^2) = a$, $\sigma_1(am^2 + bp^2) = b$ [6].

THÉORÈME 3.1. — *Muni du coproduit Δ ci-dessus, $\mathcal{H}$ est une algèbre de Hopf.*

Ce théorème est démontré en détail dans [CK1] et nous ne reproduisons pas ici la démonstration. En fait, il correspond à l'idée naturelle qu'on peut dans les intégrales (4) faire des intégrations partielles par paquets sur des sous-diagrammes (pourvu que ceux-ci restent admissibles), et itérer cette opération autant de fois qu'on veut. Il y a une complication mineure venant du fait que, pour les diagrammes à deux pattes, la règle de réinsertion dépend de la couleur du diagramme qu'on réinsert.

3.2. Graduation

Un poids multiplicatif sur $\mathcal{H}$ (valuation) est une fonction w à valeurs entières telle que

$$w(fg) = w(f) + w(g).$$

appliquer le procédé de régularisation dimensionnelle, pour pouvoir faire varier d continûment autour de $D = 6$ en contournant le pôle.

[6] L'article [CK1] dit plutôt que $\gamma_{(0)}$ est le diagramme γ affecté de moments extérieurs nuls, et $\gamma_{(1)} = \frac{\partial}{\partial p^2}\gamma|_{p=0}$: j'espère avoir interprété correctement. Il faut modifier cette règle dans le cas limite $m = 0$.

Il est complètement défini par ses valeurs sur les diagrammes simples puisque $\mathcal{H}$ est l'algèbre de polynômes engendrée par ceux-ci. Nous nous intéressons plus particulièrement à ceux qui respectent le coproduit : $w(\Delta f) = w(f)$: les plus évidents sont

$$I(\Gamma) = \quad I \quad = \text{ le nombre d'arêtes intérieures de } \Gamma$$
$$v(\Gamma) = \quad V - C \quad = \text{ le nombre de sommets } - \text{ le nombre de composantes.}$$

Définition 3.2. — *La graduation canonique de $\mathcal{H}$ est définie par $L(\Gamma) = I - V + C = \dim H^1(\Gamma)$. Pour cette graduation $\mathcal{H}$ est connexe, i.e. $\mathcal{H}^0$ est le corps des constantes (**C** ou n'importe quel corps de caractéristique* 0*).*

L'assertion résulte de ce que, par définition, on a $\dim H^1(\Gamma) > 0$ si Γ est non vide.

On note $\mathcal{H}_+$ l'idéal des éléments de degré > 0.Il est immédiat qu'on a

$$\Delta(X) - X \otimes 1 - 1 \otimes X \in \mathcal{H}_+ \otimes \mathcal{H}_+ \quad \text{si} \quad X \in \mathcal{H}_+.$$

Un autre poids de bigèbre utile est

$$\omega(\Gamma) = V_3 - N + 2, \tag{6}$$

où V_3 est le nombre de sommets à trois pattes, N le nombre d'arêtes extérieures.

3.3. Groupe dual

Le dual gradué $\mathcal{A}$ de $\mathcal{H}$ (ou plutôt un prédual) est l'espace des $\sum \varphi_\Gamma \Gamma$ où pour chaque Γ le coefficient φ_Γ parcourt un espace de fonctions test $\mathcal{S}(E_\Gamma)$. C'est une algèbre de Hopf graduée, cocommutative, mais non commutative. On notera $\widehat{\mathcal{A}}$ le complété.

Un élément $x \in \widehat{\mathcal{A}}$ est de type groupe, resp. de type Lie (primitif) si

$$\Delta_{\mathcal{A}} x = x \otimes x \quad \text{resp.} \quad \Delta_{\mathcal{A}} x = x \otimes 1 + 1 \otimes x$$

(rappelons que le coproduit $\Delta_{\mathcal{A}}$ est dual du produit de $\mathcal{H}$). Les éléments de type groupe sont les caractères de $\mathcal{H}$; ils forment un groupe G_F gradué pronilpotent, dont l'algèbre de Lie est constituée des éléments de type Lie.

4. RENORMALISATION

4.1. Lacet fondamental

À chaque diagramme de Feynman Γ est donc associée une distribution $U_\Gamma(\mu, d)$. Dans cette théorie, il est essentiel de ne renormaliser que des nombres purs (quantités « sans dimension »).

La dimension $\dim \Gamma$ de la distribution U_Γ résulte des règles $I_1, \ldots, I_5$ ci-dessus. Pour un diagramme simple (1PI), compte tenu des relations de conservation de moment (on intègre sur un espace de dimension Ld), de la définition $L = I - V + 1$ et de la relation de comptage élémentaire $3V_3 + 2V_2 = 2I + N$ obtenue en énumérant par

sommets ou par arêtes les couples ($\bullet$ —) constitués d'une arête et un de ses sommets (demi-arête), on obtient :

$$\begin{aligned}\dim\Gamma &= \dim\prod\frac{d^dp_i}{p_i^2+m^2}\prod_{V_3}g\mu^{3-d/2}\prod_{V_2}(m^2 \text{ ou } p^2)\\ &= Ld-2I+V_3(3-\frac{d}{2})+2V_2=(1-\frac{N}{2})d+N.\end{aligned}\tag{7}$$

Noter que l'exposant de g est $V_3=\omega(\Gamma)+N-2$.

À chaque diagramme habillé $\varphi\otimes\Gamma$ on associe alors un nombre sans dimension, dépendant de μ et de d (d voisin de 6, $d\neq 6$) :

$$\langle\gamma_\mu(d),\varphi\otimes\Gamma\rangle=g^{2-N}\mu^{-B}U(\Gamma)=\langle\varphi,U_\Gamma\rangle\tag{8}$$

avec $B=(1-\frac{N}{2})d+N$.[7]

On a ainsi défini un lacet $\gamma=\gamma_\mu(d)$ ou $\gamma(\mu,d)$ (méromorphe en d, dépendant de μ) autour de $d=D=6$, à valeurs dans $G_F\subset\mathcal{A}$ (il s'agit évidemment d'un caractère de $\mathcal{H}$).

On pose $d=D-\varepsilon$ et on note $\mathcal{B}$ le complété, pour la graduation L, de $\mathcal{A}[1/\varepsilon,\varepsilon]]$, dont γ est un élément.

Il est clair que tout élément $\varphi=\sum\varphi_j\varepsilon^j\in\mathcal{B}$ se décompose additivement $\varphi=\varphi_+-\varphi_-$ avec $\varphi_+=\sum_{j\geqslant 0}\varphi_j\varepsilon^j$, $\varphi_-=-\sum_{j<0}\varphi_j\varepsilon^j$ ($\varphi_\pm$ est holomorphe en ε resp. $1/\varepsilon$; ce sont les parties positive et négative d'une série de Laurent).

De même tout élément inversible $f\in 1+\mathcal{B}_+$ (i.e. le terme de plus bas degré pour la graduation est $f^0=1$) admet une unique factorisation

$$f=f_-^{-1}f_+\ ,$$

où de nouveau $f_\pm$ est holomorphe en ε, resp. $1/\varepsilon$, le terme de plus bas degré est 1, et $f_-=1$ pour $1/\varepsilon=0$ (factorisation de Wiener-Hopf, qui est liée à la classification des fibrés sur la sphère de Riemann, et au problème de Riemann-Hilbert, mais immédiate ici dans une algèbre filtrée complète). L'unicité de la factorisation assure que f_+ et f_- sont de type groupe si f l'est, de même que, dans la version infinitésimale additive, φ_+ et φ_- sont de type Lie si φ l'est.

Le résultat principal de [CK1] est le suivant :

THÉORÈME 4.1. — *Le lacet renormalisé γ_{renorm} est la valeur pour $d=D$ ($\varepsilon=0$) du facteur γ_+ dans la factorisation de Wiener-Hopf $\gamma=\gamma_-^{-1}\gamma_+$.*

Une quantité observable se réinterprète comme série à coefficients dans $\mathcal{H}$ ou, ce qui revient au même, comme fonction $F(\gamma)$ sur le groupe G_F :

[7] Il est parfois utile de prolonger cette définition au cas où l'habit φ est dimensionné : on remplace alors par $B=(1-\frac{N}{2})d+N+\dim\varphi$.

SCHOLIE. — *La quantité effective renormalisée (resp. effective non renormalisée, nue) s'obtient en substituant $\gamma_{\mu+}(\varepsilon)$ resp. $\gamma_\mu(\varepsilon), \gamma_{\mu-}$ dans son expression en fonction de γ. La valeur renormalisée est la valeur en $\gamma_{\mu+}$ pour $\varepsilon = 0$.*

(Cette dernière dépend encore de l'unité de masse μ, i.e. de l'échelle d'observation).

Ce théorème regroupe de façon concise et élégante toutes les règles de la renormalisation par régularisation dimensionnelle. Il est démontré dans [CK1]. Il n'est pas question de le démontrer ici, ne serait-ce que parce qu'il faudrait d'abord décrire les règles en question.

Rappelons que la régularisation dimensionnelle est essentiellement équivalente aux autres procédés de régularisation, par exemple par troncature, du moins pour les intégrales de dimension finie ; elle n'est pas définie pour les intégrales de dimension infinie (qui ne le sont guère non plus), et ne s'applique donc qu'à la théorie perturbative.

4.2. Groupe de renormalisation

La L-graduation et l'unité de masse sont liées de façon étroite en vertu des règles ci-dessus. De façon plus précise, notons (comme dans [CK2]) Y l'opérateur de la graduation L : ($Y \sum f_n = \sum n f_n$ si f_n est de degré n). C'est une dérivation (automorphisme infinitésimal) de $\mathcal{H}$ ou de $\mathcal{A}$, i.e. à la fois pour le produit et le coproduit.

PROPOSITION 4.2. — *On a $e^{t\varepsilon Y}\gamma(\varepsilon, \mu) = \gamma(\varepsilon, e^t\mu)$.*

La graduation et les exposants de μ ont été choisis de sorte que cela soit évident. Le groupe de renormalisation est le groupe à 1 paramètre $e^{t\varepsilon Y}$.

4.3. Contreterme, fonction β

Le résultat suivant résulte simplement des règles $I_1 - I_5$ pour les intégrales de Feynman et est démontré dans [CK2]. En fait, il s'agit plutôt d'un axiome pour la renormalisation, et il faut s'assurer qu'on a fait les bons choix pour μ :

THÉORÈME 4.3. — *Dans la factorisation $\gamma = \gamma_-^{-1}\gamma_+$ le facteur γ_- est indépendant de μ.*

Il est alors important d'analyser quels sont les contretermes γ_- possibles. L'inverse $\varphi(1/\varepsilon) = \gamma_-^{-1}$ a la propriété suivante : son terme de degré 0 est 1, il est de type groupe, et on a

$$e^{t\varepsilon Y}(\varphi) = \varphi F_t(\varepsilon), \tag{9}$$

où $F_t \in \mathcal{A}[[\varepsilon]]$ est holomorphe en ε. F_t vérifie évidemment la condition de cocycle

$$F_{t+s} = F_t \; e^{t\varepsilon Y}(F_s) \; .$$

En particulier, pour $\varepsilon = 0$, il définit un groupe à un paramètre

$$F_t|_{\varepsilon=0} = e^{\beta t} \quad \text{avec} \quad \beta \in \mathcal{A}_+. \tag{10}$$

L'élément primitif $\beta \in \mathcal{A}_+$ est le « résidu » de la théorie.

Dérivant (9) pour $t = 0$ on obtient $\varepsilon\varphi^{-1} \cdot Y\varphi = \frac{d}{dt}F_t(\varepsilon)|_{t=0}$. Le premier membre de cette égalité est holomorphe en $1/\varepsilon$, le second en ε, donc les deux sont constants et égaux à β :

$$Y\varphi = \varphi\frac{\beta}{\varepsilon}. \tag{11}$$

PROPOSITION 4.4. — *Soient $\mathcal{C}$ le complété pour la graduation L de $\mathcal{A}[1/\varepsilon]$, et $\alpha \in \mathcal{C}_+$; alors l'équation $Y\varphi = \varphi\alpha$ admet pour unique solution $\varphi \in 1 + \mathcal{C}_+$ la série*

$$\varphi = (1 - Y^{-1}R_\alpha)^{-1}(1) \tag{12}$$

où R_α désigne la multiplication à droite par α. Alors φ est de type groupe ssi α est de type Lie.

Preuve : Y est inversible sur $\mathcal{C}_+$; l'équation équivaut donc à

$$Y(1 - Y^{-1}R_\alpha)\varphi = 0, \quad \text{i.e.} \quad (1 - Y^{-1}R_\alpha)\varphi = \text{constante} = 1$$

puisque le terme de degré 0 de φ est 1. On a alors aussi

$$Y(\varphi \otimes \varphi) = (\varphi \otimes \varphi)(\alpha \otimes 1 + 1 \otimes \alpha), \quad Y(\Delta\varphi) = \Delta\varphi.\Delta\alpha,$$

donc $\Delta\varphi = \varphi \otimes \varphi$ ssi $\Delta\alpha = \alpha \otimes 1 + 1 \otimes \alpha$.

Ainsi, les contretermes possibles dans une théorie de renormalisation comme ci-dessus sont de la forme $(1 - Y^{-1}R_\alpha)^{-1}(1)$, avec $\alpha = \beta/\varepsilon$, $\beta \in \mathcal{A}_+$ de type Lie.

Dans [CK2] on trouve la description suivante du contreterme : adjoignons à $\mathcal{A}$ un nouvel élément Z_0 avec la relation $\operatorname{ad} Z_0 = Y$. On a alors $Z_0\varphi = \varphi(Z_0 + \beta/\varepsilon)$ d'où $e^{tZ_0}\varphi = \varphi e^{t(Z_0+\beta/\varepsilon)}$. Compte tenu que $e^{-tZ_0}\varphi e^{tZ_0} = e^{-tY}(\varphi) = \sum e^{-nt}\varphi_n \to 1$ pour $t \to \infty$, on obtient :

$$\varphi = e^{-tY}(\varphi)\ e^{-tZ_0}e^{t(Z_0+\beta/\varepsilon)} = \lim_{t\to\infty}\ e^{-tZ_0}e^{t(Z_0+\beta/\varepsilon)}. \tag{13}$$

L'élément $\beta \in \mathcal{A}$ qui apparaît dans la théorie est étroitement lié à la « fonction β » des physiciens.

5. CONSTANTE DE STRUCTURE ET FORMULES UNIVERSELLES.

5.1. Diagrammes réduits (cas $m = 0$)

Dans ce dernier numéro, on se limite à des diagrammes réduits correspondant au cas $m = 0$. Dans ce cas limite, on peut effacer tous les sommets de type 2 : ceux de couleur 0 produisent un facteur $m^2 = 0$, donc ne contribuent plus du tout, et ceux de couleur 1 produisent un facteur p^2 qui neutralise le facteur $\frac{1}{p^2+m^2}$ d'une des arêtes adjacentes, donc on peut aussi les effacer. La règle de réinsertion définissant le coproduit doit être modifiée : elle reste la même pour les diagrammes à trois pattes,

mais les diagrammes à deux pattes sont réinsérés au milieu de n'importe quelle arête intérieure. On peut en outre oublier les habits puisqu'il ne reste qu'une couleur et un habit pour la réinsertion des diagrammes à deux pattes.

Notons $\mathcal{H}_r$ et G_r l'algèbre de Hopf et le groupe correspondants (ce sont respectivement une sous-algèbre de $\mathcal{H}$ et un quotient de G_F).

On repère les éléments primitifs de $\mathcal{H}_r^*$ comme suit : à tout diagramme simple Γ on associe l'élément

$$\underline{\Gamma} = S(\Gamma)\ p\Gamma \in \mathcal{H}^*, \tag{14}$$

où $S(\Gamma)$ est le nombre d'automorphismes de Γ et p désigne le projecteur canonique sur les éléments primitifs (si $x \in \mathcal{H}_+$, px est l'unique élément primitif égal à x mod. $\mathcal{H}_+^{*2}$). Suivant F. Patras [Pat1, Pat2] p se calcule ainsi : si $a, b \in L(\mathcal{H}^*)$, on pose $a * b = M \circ (a \otimes b) \circ \Delta$ (où $M : \mathcal{H}^* \otimes \mathcal{H}^* \to \mathcal{H}^*$ est la multiplication) ; il s'agit d'un produit associatif, qui prolonge celui de l'algèbre de Hopf $\mathcal{H}^* \otimes \mathcal{H}$, et pour lequel l'unité est le projecteur p_0 sur le corps de base (la counité) ; on a $\mathrm{Id} = p_0 + p_+$ où p_+ est le projecteur sur l'idéal des éléments de degré > 0 de $\mathcal{H}^*$. On a alors $p = \mathrm{Log}^* \mathrm{Id} = \sum \frac{(-1)^{k-1}}{k} p_+^{*k}$.

Les $\underline{\Gamma}$ forment une base de l'algèbre de Lie $\mathcal{G}_r \subset \mathcal{H}_r^*$, et la règle de réinsertion définissant le coproduit donne

$$[\underline{\Gamma}, \underline{\Gamma'}] = \sum_{v'} \underline{\Gamma' \circ^{v'} \Gamma} - \sum_{v} \underline{\Gamma \circ^{v} \Gamma'}\ , \tag{15}$$

où v' resp. v parcourt l'ensemble des sommets (resp. lignes pour les diagrammes à deux pattes) et $\circ^v$ signifie qu'on greffe le diagramme de gauche à la v-ième place du diagramme de droite.

Pour un diagramme simple et réduit à N pattes, on a $V = 2(L-1)+N$, $\quad I = 3(L-1)+N$. Observons qu'il y a (à un isomorphisme près) un seul diagramme simple $\Gamma_{2,L}$ à deux pattes de degré L, resp. $\Gamma_{3,L}$ à trois pattes, et pour ceux-ci on a $I = 3L-1, V = 2L$, resp. $I = 3L, V = 2L+1$; en outre, le degré du diagramme greffé $\Gamma \circ^v \Gamma'$ est toujours $L + L'$.

On obtient donc pour les crochets de diagrammes simples à 2 ou 3 pattes :

$$\begin{aligned}
[\underline{\Gamma}_{2,L}, \underline{\Gamma}_{2,L'}] &= 3(L'-L)\ \underline{\Gamma}_{2,L+L'} \\
[\underline{\Gamma}_{3,L}, \underline{\Gamma}_{3,L'}] &= 2(L'-L)\ \underline{\Gamma}_{3,L+L'} \\
[\underline{\Gamma}_{3,L}, \underline{\Gamma}_{2,L'}] &= 3L'\ \underline{\Gamma}_{3,L+L'} - 2L\ \underline{\Gamma}_{2,L+L'}.
\end{aligned} \tag{16}$$

5.2. Groupe des difféomorphismes formels de la droite

Notons G_2 le groupe des difféomorphismes formels tangents à l'identité sur la droite :

$$\varphi = X + \sum_{2}^{\infty} a_n X^n.$$

Son algèbre de Hopf $\mathcal{H}_2$ est l'algèbre des polynômes en les $a_n = a_n(\varphi)$; si T désigne la série de Taylor $X + \sum a_n X^n$ ($T_\varphi = \langle T, \varphi\rangle$) $= \varphi(X)$, le coproduit est défini par

$$\langle \Delta T, \varphi \otimes \psi\rangle = T_\varphi \circ T_\psi.$$

C'est un groupe gradué, la graduation correspondant à l'action du groupe des homothéties (a_n est de degré $n-1$). Son algèbre de Lie admet pour base graduée les champs de vecteurs

$$Z'_k = X^{k+1}\frac{\partial}{\partial X}.$$

Z_k est de degré k et on a

$$[Z'_p, Z'_q] = (p-q)Z'_{p+q}. \tag{17}$$

Les formules de crochet (16) impliquent :

PROPOSITION 5.1. — *Posons $\rho_\Gamma = 1$ si Γ a trois pattes, $\rho_\Gamma = 3/2$ s'il en a deux ; il existe un unique homomorphisme ρ d'algèbres de Lie ou de groupes formels $G_r \to G_2$ tel que*

$$\rho(\underline{\Gamma}) = \rho_\Gamma Z'_{2L} \quad (L = L(\Gamma)). \tag{18}$$

5.3. Constante de couplage et difféomorphismes

La renormalisation a pour effet de remplacer les quantités observables par des séries de diagrammes (dépendant de ε et μ) ; en particulier, à la constante de couplage g est associée la série génératrice

$$g_0 = X + \sum_2^\infty \alpha_n X^n = (gZ_1)(Z_3)^{-3/2} \quad \text{(série impaire)} \tag{19}$$

avec

$$gZ_1 = \Big(X + \sum_{\Gamma\in F_3} X^{2L+1}\frac{\Gamma}{S(\Gamma)}\Big) \quad Z_3 = \Big((1 - \sum_{\Gamma\in F_2} X^{2L}\frac{\Gamma}{S(\Gamma)}\Big) \tag{20}$$

où F_2 resp. F_3 désigne l'ensemble des diagrammes simples et réduits à deux resp. trois pattes ($2L$, resp. $2L+1$ est l'exposant de g dans ces diagrammes) ; $S(\Gamma)$ est le nombre d'automorphismes de Γ.

La série génératrice g_0 définit une fonction $\Phi^* : G_r \to G_2$: si $m \in G_r$, $\Phi^*(m) = \varphi(X)$ est l'évaluation en m de la série $g_0(X)$ (de façon duale : un homomorphisme d'algèbres $\Phi : \mathcal{H}_2 \to \mathcal{H}_r$ qui envoie a_n sur le coefficient α_n).

THÉORÈME 5.2. — *La fonction $\Phi^* : G_F \to G_2$ définie par la série génératrice g_0 est un homomorphisme de groupes (de façon duale : Φ est un homomorphisme d'algèbres de Hopf).*

En fait, Connes et Kreimer montrent qu'on a $\Phi^* = \rho$, en prouvant que les deux transportent de la même façon le champ de vecteurs invariant à droite ∂_Z associé à un élément primitif $Z \in \mathcal{H}_r$ ($\partial_Z a = \langle Z \otimes \mathrm{Id}, \Delta a \rangle$; il suffit de vérifier ceci quand Z parcourt l'ensemble des éléments basiques $\underline{\Gamma}$).

5.4. Remarques finales

L'homomorphisme ρ défini précédemment est compatible avec les graduations (plus exactement, il est homogène de degré 2) puisque a_{n+1} et $X^{n+1}\partial/\partial X$ sont de degré n). Le paramètre g est essentiellement reflété dans la graduation, car on a, élémentairement, en remettant dans la notation le paramètre g qui n'y figurait pas : $\gamma_{\mu,g}(\varepsilon) = e^{2tY}\gamma_{\mu,1}(\varepsilon)$.

On montre que l'image $\rho(\beta)$ de l'élément générateur β du n° 4.3 est exactement la « fonction » β des physiciens, d'où la notation.

Le théorème 5.2 implique que la renormalisation pour g_{eff} se lit directement dans la factorisation $g_{\mathrm{eff}} = g_{\mathrm{eff}-}^{-1} g_{\mathrm{eff}+}$ dans le groupe G_2, et G_r n'intervient plus. En ce sens G_2 apparaît comme groupe universel pour ces calculs, et on peut imaginer qu'un groupe des difféomorphismes à plusieurs variables intervient de même pour une théorie de champ vectorielle.

Le lacet universel $\gamma_\mu(d)$ définit, par recollement via g_{eff} un fibré P sur la droite projective P_1, de fibre la droite formelle X. Ici, comme probablement dans toute théorie perturbative, le lacet se factorise et ce fibré est trivialisable ; on peut imaginer que dans une théorie non perturbative, comme dans le problème de Riemann-Hilbert non nilpotent, interviendront de tels fibrés non triviaux.

RÉFÉRENCES

[CK1] A. Connes & D. Kreimer – « Renormalization in quantum field theory and the Riemann-Hilbert problem I : The Hopf algebra structure of graphs and the main theorem », *Comm. Math. Phys.* **210** (2000), no. 1, p. 249–273.

[CK2] ———, « Renormalization in quantum field theory and the Riemann-Hilbert problem. II : the β-function, diffeomorphisms and the renormalization group », *Comm. Math. Phys.* **216** (2001), no. 1, p. 215–241.

[BK1] D.J. Broadhurst & D. Kreimer – « Knots and numbers in ϕ^4 theory to 7 loops and beyond », *Internat. J. Modern Phys. C* **6** (1995), no. 4, p. 519–524.

[BK2] ———, « Association of multiple zeta values with positive knots via Feynman diagrams up to 9 loops », *Phys. Lett. B* **393** (1997), no. 3-4, p. 403–412.

[BK3] D.J. Broadhurst, J.A. Gracey & D. Kreimer – « Beyond the triangle and uniqueness relations : non-zeta counterterms at large N from positive knots », *Z. Phys. C* **75** (1997), no. 3, p. 559–574.

[BK4] D.J. Broadhurst & D. Kreimer – « Feynman diagrams as a weight system : four-loop test of a four-term relation », *Phys. Lett. B* **426** (1998), no. 3-4, p. 339–346.

[BK5] ———, « Renormalization automated by Hopf algebra », *J. Symbolic Comput.* **27** (1999), no. 6, p. 581–600, and `hep-th/9810087`.

[BK6] ———, « Combinatoric explosion of renormalization tamed by Hopf algebra : 30-loop Padé-Borel resummation », *Phys. Lett. B* **475** (2000), no. 1-2, p. 63–70, and `hep-th/9912093`.

[BK7] ———, « Towards cohomology of renormalization : bigrading the combinatorial Hopf algebra of rooted trees », *Comm. Math. Phys.* **215** (2000), no. 1, p. 217–236.

[CK3] A. Connes & D. Kreimer – « Hopf algebras, renormalization and noncommutative geometry », *Comm. Math. Phys.* **199** (1998), no. 1, p. 203–242.

[CK4] ———, « Hopf algebras, renormalization and noncommutative geometry », in *Quantum field theory : perspective and prospective (Les Houches 1998)*, NATO Sci. Ser. C Math. Phys. Sci., vol. 530, Kluwer Acad. Publ., Dordrecht, 1999, p. 59–108.

[CK5] ———, « Renormalization in quantum field theory and the Riemann-Hilbert problem », *J. High Energy Phys.* (1999), no. 9, p. Paper 24, 8 pp., (electronic) and `hep-th/9909126`.

[CK6] ———, « Lessons from quantum field theory : Hopf algebras and spacetime geometries, Moshé Flato (1937-1998) », *Lett. Math. Phys.* **48** (1999), no. 1, p. 85–96, and `hep-th/9904044`.

[CK7] ———, « From local perturbation theory to Hopf- and Lie-algebras of Feynman graphs », in *Mathematical physics in mathematics and physics (Siena, 2000)*, Fields Inst. Commun., vol. 30, Amer. Math. Soc., 2001, p. 105–114.

[DK] R. Delbourgo & D. Kreimer – « Using the Hopf algebra structure of QFT in calculations », *Phys. Rev. D (3)* **60** (1999), no. 10, and `hep-th/9903249`.

[K1] D. Kreimer – « Renormalization and knot theory », *J. Knot Theory and Ramifications* **6** (1997), no. 4, p. 479–581.

[K2] ———, « On the Hopf algebra structure of perturbative quantum field theories », *Adv. Theor. Math. Phys.* **2** (1998), no. 2, p. 303–334.

[K3] ———, « On overlapping divergences », *Comm. Math. Phys.* **204** (1999), no. 3, p. 669–689, and `hep-th/9810022`.

[K4] ———, « Chen's iterated integral represents the operator product expansion », *Adv. Theor. Math. Phys.* **3** (1999), no. 3, and `hep-th/9901099`.

[K5] ———, *Knots and Feynman diagrams*, Cambridge Lecture Notes in Physics, vol. 13, Cambridge University Press, Cambridge, 2000.

[K6] ———, « Shuffling quantum field theory », *Lett. Math. Phys.* **51** (2000), no. 3, p. 179–191.

Renormalisation

[Coll] J. Collins – *Renormalization,*, Cambridge monographs in math. phys., Cambridge University Press, Cambridge, 1984.

[Dres] M. Dresden – « Renormalization in historical perspective – The first stage », in *Renormalization*, Springer-Verlag, New York, Berlin, Heidelberg, 1994.

[Drou] J.-M. Drouffe & C. Itzykson – *Théorie statistique des champs*, Savoirs actuels, InterEditions/Editions du C.N.R.S., 1989, 2 volumes.

[EG] H. Epstein & V. Glaser – « The role of locality in perturbation theory », *Ann. Inst. H. Poincaré A* **19** (1973), p. 211–295.

[FMRS] J. Feldman, J. Magnen, V. Rivasseau & R. Seneor – « Massive Gross-Neveu model : a rigorous perturbative construction », *Phys. Rev. Lett.* **54** (1985).

[GK] K. Gawedski & A. Kupianen – « Exact renormalization of the Gross-Neveu model of quantum fields », *Phys. Rev. Lett* **54** (1985).

[GJ] J. Glimm & A. Jaffe – *Quantum Physics*, Springer-Verlag, New York, Berlin, Heidelberg, 1987.

[LeBe] M. Le Bellac – *Des phénomènes critiques aux champs de jauge*, Savoirs actuels, InterEditions/Editions du C.N.R.S., 1988.

[ZJ] J. Zinn-Justin – *Quantum Field Theory and Critical Phenomena*, International series of monographies on physics, vol. 92, Oxford science publications, 1996.

Wiener-Hopf, Riemann-Hilbert, algèbres de Hopf

[Beau] A. Beauville – « Monodromie des systèmes différentiels linéaires à pôles simples sur la sphère de Riemann », in *Sém. Bourbaki*, Astérisque, vol. 216, Soc. Math. France, Paris, 1993, exp. n° 765 (mars 1993), p. 103–119.

[Boli] A. Bolibruch – « Fuchsian systems with reducible monodromy and the Riemann-Hilbert problem », Lecture Notes in Math., vol. 1250, Springer, 1992, p. 139–155.

[BKI] N. Bourbaki – *Éléments de mathématique. Algèbre. Chapitres* 1 *à* 3, Masson, Paris, 1982.

[ENS] L. Boutet de Monvel, A. Douady & J.-L. Verdier (éds.) – *Mathématique et Physique, Séminaire de l'E.N.S. 1979-82*, Progress in Math., vol. 35, Birkhäuser, 1983.

[Drin] V.G. Drinfel'd – « Almost cocommutative Hopf algebras », *Algebra i Analiz* **1** (1989), no. 2, p. 30–46, and translation in *Leningrad Math. J.* **1** (1990), no. 2, p. 321-342.

[WH] I. Gohberg & M.A. Kaashoek (éds.) – *Constructive methods of Wiener-Hopf factorization*, Operator Theory : Advances and Applications, vol. 21, Birkhäuser Verlag, Basel, 1986.

[LP] I. Lappo-Danilevskii – *Mémoire sur la théorie des systèmes d'équations différentielles linéaires*, Chelsea, New York, 1953.

[Pat1] F. Patras & C. Reutenauer – « Higher Lie idempotents », *J. Algebra* **222** (1999), no. 1, p. 51–64.

[Pat2] F. Patras – « La décomposition en poids des algèbres de Hopf », *Ann. Inst. Fourier* **43** (1993), no. 4, p. 1067–1087.

[Spe] F.-O. Speck – *General Wiener-Hopf factorization methods*, vol. 119, Pitman (Advanced Publishing Program), Boston, MA, 1985, With a foreword by E. Meister.

Général

[BCE] J. Brodzki, A. Connes & D. Ellwood – « Polarized modules and Fredholm modules », *Mat. Fiz. Anal. Geom.* **2** (1995), no. 1, p. 15–24.

[C1] A. Connes – *Noncommutative geometry*, Academic Press, Inc., San Diego, CA, 1994, 661 pp.

[C2] ———, « Non-commutative geometry and physics », in *Gravitation et quantifications (Les Houches, 1992)*, North-Holland, Amsterdam, 1995, p. 805–950.

[C3] ———, « Géométrie non commutative et physique quantique », in *Mathématiques quantiques*, SMF Journ. Annu., Soc. Math. France, Paris, 1992, 20 pp.

[C4] ———, « The action functional in noncommutative geometry », *Comm. Math. Phys.* **117** (1988), no. 4, p. 673–683.

[CM] A. Connes & H. Moscovici – « Hopf algebras, cyclic cohomology and the transverse index theorem », *Comm. Math. Phys.* **198** (1998), no. 1, p. 199–246.

[CS] A. Connes & E. Størmer – « A connection between the classical and the quantum mechanical entropies », in *Operator algebras and group representations (Neptun, 1980),* vol. I, Monographs Stud. Math., vol. 17, Pitman, Boston, Mass.-London, 1984, p. 113–123.

Louis BOUTET de MONVEL

Université de Paris VI, analyse algébrique
Institut de Math. de Jussieu
UMR 7586 du CNRS
Case 82, 4 place Jussieu
F-75252 Paris Cedex 05
E-mail : boutet@math.jussieu.fr

Séminaire BOURBAKI Mars 2002
54e année, 2001-2002, nO 901, p. 167 à 192

COBORDISME DES VARIÉTÉS ALGÉBRIQUES
[d'après M. Levine et F. Morel]

par **François LOESER**

INTRODUCTION

Le but de cet exposé est de décrire les travaux de Levine et Morel concernant une nouvelle théorie (co)homologique des variétés algébriques, le cobordisme algébrique, ainsi que ses applications à la formule du degré de Rost. Traditionnellement, deux variétés $\mathcal{C}^\infty$ compactes sans bord de dimension n sont dites cobordantes si leur somme disjointe est le bord d'une variété $\mathcal{C}^\infty$ compacte de dimension $n+1$. On définit ainsi une relation d'équivalence et l'ensemble des classes d'équivalence peut être naturellement muni d'une structure d'anneau gradué (la graduation provenant de la dimension), l'anneau de cobordisme réel N_*. Cet anneau a été introduit par Thom dans [21], qui a démontré que les N_n sont isomorphes aux groupes d'homotopie stable des « espaces de Thom » $\mathrm{MO}(n)$ et que N_* est isomorphe à une algèbre de polynômes sur $\mathbf{Z}/2\mathbf{Z}$. En considérant de façon similaire des variétés $\mathcal{C}^\infty$ stablement presque complexes, on définit l'anneau de cobordisme complexe U_* dont la construction est rappelée dans la section 1. Milnor a démontré dans [17] que U_* est isomorphe à une algèbre de polynômes sur $\mathbf{Z}$ et qu'elle est engendrée comme algèbre par les classes de certaines variétés projectives complexes, en fait les espaces projectifs complexes et les hypersurfaces de bidegré $(1,1)$ dans le produit de deux espaces projectifs complexes (*cf.* 2.3.9).

Dans un travail fondamental [20], Quillen démontre plus précisément que U_* est canoniquement isomorphe à l'anneau de Lazard $\mathbf{L}_*$ classifiant les lois de groupe formel commutatif de rang 1. Plus généralement, il considère les théories homologiques et cohomologiques $X \mapsto U_*(X)$ et $X \mapsto U^*(X)$ représentées par le spectre de Thom complexe. Que U_* soit canoniquement isomorphe à l'anneau de Lazard $\mathbf{L}_*$ peut être compris comme le fait que, dans la théorie U_*, la classe de Chern du produit tensoriel de deux fibrés en droite est « la plus compliquée possible » contrairement à ce qui se passe en homologie singulière où $c_1(L\otimes M) = c_1(L)+c_1(M)$. Ceci reflète une propriété

plus générale de U_* qui est d'être universelle parmi « les théories homologiques avec orientation complexe ».

En géométrie algébrique, pour les variétés sur un corps k, il existe un analogue naturel du cobordisme qui consiste à considérer un morphisme projectif $W \to \mathbf{A}^1$, avec, disons, W lisse et irréductible, qui est transverse à $\{0\}$ et à $\{1\}$. Les fibres W_0 et W_1 correspondantes sont alors lisses sur k et on aurait envie de dire qu'elles sont algébriquement cobordantes. La théorie que l'on obtient (*cf.* 2.3.3) de cette façon est malheureusement bien trop naïve : par exemple deux courbes cobordantes en ce sens ont toujours même genre, ce qui ne se produit pas avec la théorie Ω_* de Levine et Morel dont une des propriétés fondamentales est que sa valeur $\Omega_*(\operatorname{Spec} k)$ sur le point est canoniquement isomorphe à l'anneau de Lazard $\mathbf{L}_*$. De plus, comme on le verra au cours de l'exposé, la théorie homologique Ω_* est également universelle parmi les théories homologiques « orientées » pour les variétés algébriques sur un corps k.

Pour conclure remarquons que, comme il l'exprime lui-même dans l'introduction de [20], le travail de Quillen est déjà fortement influencé par les travaux de Grothendieck en géométrie algébrique : « I have been strongly influenced by Grothendieck's theory of motives in algebraic geometry ... and like to think of a cobordism theory as a universal contravariant functor on the category of $\mathcal{C}^\infty$ manifolds endowed with Gysin homomorphisms for a class of proper « oriented maps »... ».

J'ai bénéficié de l'aide et des conseils de J.-B. Bost, A. Chambert-Loir, M. Levine et F. Morel pendant la préparation de cet exposé. Qu'ils en soient amicalement remerciés ici.

Table des matières

1. PRÉLIMINAIRES SUR LE COBORDISME COMPLEXE

1.1. Cobordisme complexe

La notion d'orientation complexe pour un morphisme $f : Z \to X$ de variétés $\mathcal{C}^\infty$ généralise celle de structure faiblement complexe sur Z lorsque X est un point. Soit f

un tel morphisme. On suppose que les variétés Z et X sont équidimensionnelles, et on définit la dimension relative de f comme $\dim f := \dim Z - \dim X$. Commençons par supposer que f est de dimension relative paire. Dans ce cas, une orientation complexe de f consiste en la donnée d'une classe d'équivalence de factorisations de f en le composé d'un morphisme structural de fibré complexe $p : E \to X$ et d'un plongement $i : Z \to E$ dont le fibré normal est muni d'une structure complexe. Deux factorisations $f = p \circ i$ et $f = p' \circ i'$ sont considérées comme équivalentes si les fibrés complexes E et E' peuvent être plongés dans le même fibré complexe E'' de façon que, dans E'', i et i' soient isotopes de façon compatible avec la structure complexe dont sont munis leurs fibrés normaux respectifs. Quand f est de dimension relative impaire, on remplace E par $E \times \mathbf{R}$ dans la définition précédente. On dira morphisme orienté pour morphisme muni d'une orientation complexe.

Deux morphismes propres orientés $f_1 : Z_1 \to X$ et $f_2 : Z_2 \to X$ sont dits cobordants s'il existe un morphisme propre orienté $b : W \to X \times \mathbf{R}$ tel que les morphismes $\epsilon_i : X \to X \times \mathbf{R}$ qui à x associent (x, i) soient transverses à b et que le pull-back de b par les ϵ_i soient isomorphes comme morphismes orientés aux f_i.

Le cobordisme est une relation d'équivalence et on note $U^q(X)$ l'ensemble des classes de cobordisme de morphismes propres orientés de dimension relative $-q$. Il sera aussi commode de passer en notation homologique en posant $U_q(X) := U^{\dim X - q}(X)$.

1.1.1. Le cobordisme U^* est un foncteur contravariant : si $g : Y \to X$ est un morphisme de variétés $\mathcal{C}^\infty$ et $f : Z \to X$ un morphisme propre orienté, quitte à bouger g par une homotopie, on peut supposer, par le théorème d'isotopie de Thom que le morphisme g est transverse à f. La classe de cobordisme du morphisme $Y \times_X Z \to Y$ ne dépend que de celle de f, et on obtient ainsi un morphisme

$$g^* : U^q(X) \to U^q(Y).$$

1.1.2. Tout morphisme propre orienté $g : X \to Y$ de dimension relative d définit un morphisme de Gysin

$$g_* : U^q(X) \to U^{q-d}(Y),$$

par composition.

1.1.3. Le produit cartésien induit un produit externe naturel

$$\otimes : U^*(X_1) \times U^*(X_2) \to U^*(X_1 \times X_2).$$

On en tire une structure d'anneau gradué sur $U^*(X)$ en posant $x_1 \cdot x_2 := \Delta^*(x_1 \otimes x_2)$, Δ désignant le morphisme diagonal. L'unité est la classe de cobordisme $1 = 1_X$ du morphisme identité $X \to X$.

1.2. L'anneau de Lazard

Rappelons qu'une loi de groupe formel commutatif de rang 1 à coefficients dans un anneau commutatif A est une série formelle

$$F(u,v) = \sum_{(i,j)\in\mathbf{N}^2} a_{i,j}u^i v^j \tag{1.2.1}$$

appartenant à $A[[u,v]]$ et satisfaisant les relations

$$F(u,0) = F(0,u) = u \tag{1.2.2}$$

$$F(u,v) = F(v,u) \tag{1.2.3}$$

et

$$F(u,F(v,w)) = F(F(u,v),w). \tag{1.2.4}$$

D'après Lazard [9], il existe une loi de groupe formel commutatif de rang 1 universelle $(\mathbf{L}, F_{\mathbf{L}})$. La construction de $(\mathbf{L}, F_{\mathbf{L}})$ est très simple. On considère l'anneau $\widetilde{\mathbf{L}}$ des polynômes à coefficients entiers en les variables $A_{i,j}$, pour i, j dans $\mathbf{N}$ et la série formelle universelle $\widetilde{F} = \sum_{i,j} A_{i,j}u^i v^j$ dans $\widetilde{\mathbf{L}}[[u,v]]$. Il suffit alors de définir $\mathbf{L}$ comme le quotient de l'anneau $\widetilde{\mathbf{L}}$ par les relations obtenues en imposant les relations (1.2.2), (1.2.3) et (1.2.4) à $\widetilde{F}$ et de définir la série $F_{\mathbf{L}} := \sum_{i,j} a_{i,j}u^i v^j$ comme l'image de $\widetilde{F}$ dans $\mathbf{L}[[u,v]]$. On munit l'anneau $\mathbf{L}$ d'une structure d'anneau gradué $\mathbf{L}_*$ en donnant le degré $i+j-1$ à $a_{i,j}$. On aura aussi à considérer l'anneau gradué $\mathbf{L}^*$ défini par $\mathbf{L}^n := \mathbf{L}_{-n}$. Un fait remarquable, démontré par Lazard [9], est que l'anneau de Lazard $\mathbf{L}$ est un anneau de polynômes à coefficients entiers en un nombre dénombrable de variables.

1.3. Le théorème de Quillen

Soit E un fibré vectoriel complexe de rang n sur une variété X. On note $i : X \to E$ le morphisme donné par la section nulle. La classe d'Euler de E est l'élément

$$e(E)(= c_n(E)) := i^* i_* 1 \tag{1.3.1}$$

de $U^{2n}(X)$.

Notons $q : \mathbf{P}(E) \to X$ le fibré projectif représentant les quotients de rang 1 de E et γ le fibré en droites sur $\mathbf{P}(E)$ quotient canonique de $q^*(E)$. Dans les sections suivantes ce fibré en droites sera parfois également noté $O(1)$ (c'est le fibré qui a des sections globales et non son dual). On note ξ la classe d'Euler du fibré γ. On vérifie alors par un calcul classique, *cf.* [8], que $U^*(\mathbf{P}(E))$ est, via q^*, un $U^*(X)$-module libre de base $1, \xi, \ldots, \xi^{n-1}$. De plus si E est somme directe de fibrés en droites $L_1, \ldots, L_n$, on a

$$\prod_{i=1}^{n} (\xi - e(L_i)) = 0. \tag{1.3.2}$$

On en tire en particulier que

$$U^*(\mathbf{P}^n) \simeq U^*(\text{point})[u]/u^{n+1} \tag{1.3.3}$$

avec $u = e(\gamma)$ et que

$$U^*(\mathbf{P}^n \times \mathbf{P}^n) \simeq U^*(\text{point})[u,v]/(u^{n+1}, v^{n+1}), \tag{1.3.4}$$

avec $u = e(\mathrm{pr}_1^*(\gamma))$ et $u = e(\mathrm{pr}_2^*(\gamma))$. On peut alors construire une loi de groupe formel commutative sur $U^*(\text{point})$ de la façon suivante. D'après (1.3.4), on peut écrire

$$e(\mathrm{pr}_1^*(\gamma) \otimes \mathrm{pr}_2^*(\gamma)) = \sum_{i,j \leqslant n} c_{i,j}^n u^i v^j. \tag{1.3.5}$$

De plus $c_{i,j}^n$ ne change pas quand on augmente n. En faisant tendre n vers l'infini on en tire une série formelle $F(u,v)$ dans $U^*(\text{point})[[u,v]]$ de la forme

$$F(u,v) = \sum_{i,j} c_{i,j} u^i v^j \tag{1.3.6}$$

avec $c_{i,j}$ dans $U^{2-2i-2j}(\text{point})$. On vérifie facilement que c'est une loi de groupe formel sur $U^*(\text{point})$. Par exemple, on démontre l'associativité (1.2.4) en évaluant de deux façons différentes la classe d'Euler de $\mathrm{pr}_1^*(\gamma) \otimes \mathrm{pr}_2^*(\gamma) \otimes \mathrm{pr}_3^*(\gamma)$ sur $\mathbf{P}^n \otimes \mathbf{P}^n \otimes \mathbf{P}^n$.

Comme tout fibré en droites complexes sur une variété X est image inverse de γ pour un morphisme $X \to \mathbf{P}^n$ convenable, on en déduit que pour toute paire de fibrés en droites complexes E_1 et E_2 sur la variété X on a

$$e(L_1 \otimes L_2) = F(e(L_1), e(L_2)). \tag{1.3.7}$$

Par la propriété universelle de l'anneau de Lazard, on dispose d'un morphisme canonique $\delta : \mathbf{L}^* \to U^*(\text{point})$.

1.3.1. Théorème (Quillen [20]). — *Le morphisme canonique $\delta : \mathbf{L}^* \to U^*(\text{point})$ est un isomorphisme d'anneaux gradués.*

Dans cet isomorphisme, les coefficients $a_{i,j}$ de la série de Lazard admettent une interprétation géométrique en terme des classes des espaces projectifs et de celles des hypersurfaces de Milnor $H_{i,j}$ (*cf.* 2.3.9, la proposition 2.3.10 restant bien sûr vraie en remplaçant Ω_* par U_* ; voir aussi [2] Proposition 10.6). La preuve de Quillen a été exposée dans ce séminaire par M. Karoubi [8] (voir aussi [2]). Elle utilise de façon essentielle que $U^q(X)$ est de type fini lorsque X a le type d'homotopie d'un CW-complexe fini, ce qui résulte de l'interprétation homotopique à la Thom (Théorème 7.1.1) et des énoncés de finitude usuels en homotopie, le lien entre opérations de Steenrod en cobordisme complexe et opérations de Landweber-Novikov ainsi qu'un calcul explicite faisant intervenir les espaces lenticulaires.

2. CONSTRUCTION DU COBORDISME ALGÉBRIQUE

2.1. Notations

On fixe un corps k. On note $\mathcal{L}_k$ la catégorie des variétés lisses quasi-projectives sur k et $\mathcal{V}_k$ celle des schémas de type fini sur k. On note $\mathcal{L}'_k$ et $\mathcal{V}'_k$ les catégories formées des mêmes objets mais où les morphismes sont les morphismes projectifs. On dira parfois variété lisse pour variété quasi-projective lisse.

2.2. Cycles de cobordisme

2.2.1. Soit X dans $\mathcal{V}_k$. On note $\mathcal{M}_*^+(X)$ le groupe abélien libre sur les classes d'isomorphisme de cycles $f : Y \to X$ avec f projectif et Y lisse et intègre sur k, gradué par la dimension de Y.

On désire définir le groupe de cobordisme algébrique $\Omega_*(X)$ comme quotient de $\mathcal{M}_*^+(X)$ par une relation d'équivalence convenable, le « cobordisme algébrique ».

Pour des raisons de commodité technique, en particulier pour disposer dès le départ de classes de Chern de fibrés en droites, il est préférable de considérer la donnée supplémentaire de familles de fibrés en droites sur Y. Ainsi on définit un cycle de cobordisme sur X comme la donnée d'un morphisme $f : Y \to X$ avec f projectif et Y lisse et intègre sur k ainsi que de r fibrés en droites $L_1, \ldots L_r$ sur Y, avec r éventuellement nul. On a une notion évidente d'isomorphisme pour les cycles de cobordisme sur X (on tolère le renumérotage des L_i). La dimension d'un tel cycle est l'entier $\dim_k(Y) - r$. On notera 1_X le cycle $\mathrm{Id} : X \to X$ et $1 = 1_{\mathrm{Spec}\, k}$.

2.2.2. On note $\mathcal{Z}(X)$ le groupe abélien libre sur les classes d'isomorphisme de cycles de cobordisme sur X. En considérant la dimension des cycles de cobordisme, on le munit d'une structure de groupe abélien gradué $\mathcal{Z}_*(X)$.

Si $g : X \to X'$ est un morphisme projectif, on dispose d'un morphisme d'image directe

$$g_* : \mathcal{Z}_*(X) \longrightarrow \mathcal{Z}_*(X')$$

qui à $[f : Y \to X, (L_1, \ldots, L_r)]$ associe $[g \circ f : Y \to X', (L_1, \ldots, L_r)]$. On obtient ainsi un foncteur

$$\mathcal{Z}_* : \mathcal{V}'_k \longrightarrow \mathbf{Ab}_*$$

qui est additif, autrement dit, le morphisme canonique

$$\bigoplus_{1 \leqslant i \leqslant r} \mathcal{Z}_*(X_i) \longrightarrow \mathcal{Z}_*\Big(\coprod_{1 \leqslant i \leqslant r} X_i \Big)$$

est un isomorphisme.

Si $g : X' \to X$ est un morphisme lisse équidimensionnel de dimension relative d on dispose d'un morphisme d'image inverse

$$g^* : \mathcal{Z}_*(X) \longrightarrow \mathcal{Z}_{*+d}(X')$$

qui à $[f : Y \to X, (L_1, \dots, L_r)]$ associe $[p_2 : Y \times_X X' \to X', (p_1^* L_1, \dots, p_1^* L_r)]$. Enfin, pour tout fibré en droites L sur la variété X, on dispose d'un morphisme tautologique, la première classe de Chern,

$$\widetilde{c}_1(L) : \mathcal{Z}_*(X) \longrightarrow \mathcal{Z}_{*-1}(X)$$

qui à $[f : Y \to X, (L_1, \dots, L_r)]$ associe $[f : Y \to X, (L_1, \dots, L_r, f^* L)]$.

Ces données vérifient les conditions suivantes :

(A1) On a $\mathrm{Id}^* = \mathrm{Id}$ et, chaque fois que cela a un sens, $(f \circ g)^* = g^* \circ f^*$.

(A2) Si $f : X \to Z$ est projectif et $g : Y \to Z$ est lisse équidimensionnel, on a $g^* f_* = f'_* g'^*$, pour tout diagramme cartésien

$$\begin{array}{ccc} W & \xrightarrow{g'} & X \\ {\scriptstyle f'}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow{g} & Z. \end{array}$$

(A3) La première classe de Chern est compatible à l'image directe :

$$f_* \circ \widetilde{c}_1(f^* L) = \widetilde{c}_1(L) \circ f_*$$

et à l'image inverse :

$$\widetilde{c}_1(f^* L) \circ f^* = f^* \circ \widetilde{c}_1(L).$$

De plus on a toujours

$$\widetilde{c}_1(L) \circ \widetilde{c}_1(M) = \widetilde{c}_1(M) \circ \widetilde{c}_1(L)$$

et $\widetilde{c}_1(L) = \widetilde{c}_1(M)$ si L et M sont isomorphes.

On pose $\mathcal{Z}_*(k) := \mathcal{Z}_*(\mathrm{Spec}\, k)$. On a un produit externe évident

$$\mathcal{Z}_*(X) \times \mathcal{Z}_*(Y) \longrightarrow \mathcal{Z}_*(X \times_k Y)$$

qui est associatif, commutatif, et pour lequel l'élément 1 de $\mathcal{Z}_0(k)$ est une unité. Ainsi $\mathcal{Z}_*(k)$ est muni d'une structure d'anneau gradué et $\mathcal{Z}_*(X)$ est naturellement muni d'une structure de $\mathcal{Z}_*(k)$-module gradué.

Ce produit externe est compatible en un sens évident aux images directes et inverses, ainsi qu'aux classes de Chern $\widetilde{c}_1$.

2.2.3. Un foncteur additif

$$H_* : \mathcal{V}'_k \longrightarrow \mathbf{Ab}_*$$

muni de morphismes d'image inverse $g^* : H_*(X) \to H_{*+d}(X')$ pour les morphismes $g : X' \to X$ lisses équidimensionnels de dimension d et de classes de Chern $\widetilde{c}_1(L) : H_*(X) \to H_{*-1}(X)$ pour les fibrés en droites, vérifiant l'analogue des conditions (A1)-(A3) est appelé foncteur de Borel-Moore orienté sur $\mathcal{V}_k$. On écrira $H_*(k)$ pour $H_*(\mathrm{Spec}\, k)$. Si de plus, $H_0(k)$ contient un élément 1 et pour toute paire (X, Y) d'objets de $\mathcal{V}_k$, on dispose d'un produit externe, donné par une forme bilinéaire graduée

$$H_*(X) \times H_*(Y) \longrightarrow H_*(X \times Y),$$

qui est strictement commutatif, associatif et pour lequel 1 est une unité et que ce produit externe est compatible aux images directes et inverses, ainsi qu'aux classes de Chern $\widetilde{c}_1$, on dira que H_* est un foncteur de Borel-Moore orienté avec produit.

Remarquons que le foncteur $\mathcal{Z}_*$ est universel parmi les foncteurs de Borel-Moore orientés avec produit : en effet, si H_* est un tel foncteur, il existe un unique morphisme de foncteurs de Borel-Moore orientés $\theta : \mathcal{Z}_* \to \mathcal{H}_*$ tel que $\theta(1) = 1$, car nécessairement $\theta([f : Y \to X, (L_1, \ldots, L_r)]) = f_* \circ \widetilde{c}_1(L_1) \circ \cdots \circ \widetilde{c}_1(L_r)(1)$.

2.3. De $\mathcal{Z}_*$ à Ω_*

2.3.1. Condition dimensionnelle. — Un premier axiome naturel que doit satisfaire la théorie Ω_* que nous cherchons à construire est la condition de dimensionalité

(Dim) Pour toute variété lisse Y sur k et toute famille $(L_1, \ldots, L_n)$ de fibrés en droites sur Y, on a

$$\widetilde{c}_1(L_1) \circ \cdots \circ \widetilde{c}_1(L_n)(1_Y) = 0$$

si $n > \dim(Y)$.

En quotientant les $\mathcal{Z}_*(X)$ par les sous-groupes engendrés par les éléments de la forme $[Y \to X, (\pi^* L_1, \ldots, \pi^* L_n, M_1, \ldots M_m)]$ pour $\pi : Y \to Z$ un morphisme lisse équidimensionnel entre variétés lisses et $n > \dim(Z)$, on obtient un foncteur de Borel-Moore orienté $\underline{\mathcal{Z}}_*(X)$ universel parmi les foncteurs de Borel-Moore orientés avec produit satisfaisant (Dim). Remarquons que, par construction, $\underline{\mathcal{Z}}_n(X) = 0$ si $n < 0$.

2.3.2. Zéros d'une section et $\widetilde{c}_1$. — Une autre condition naturelle est que la première classe de Chern corresponde aux zéros d'une section générale. Plus précisément

(Sec) Pour toute variété lisse Y sur k et tout fibré en droites L sur Y, pour toute section s de L qui est transverse à la section nulle,

$$\widetilde{c}_1(L)(1_Y) = [i : Z \to Y]$$

avec $i : Z \to Y$ l'inclusion du lieu des zéros de s.

En quotientant les $\underline{\mathcal{Z}}_*(X)$ par les sous-groupes engendrés par les éléments de la forme

$$[Y \to X, (L, L_1, \ldots, L_n)] - [Z \to X, (i^* L_1, \ldots, i^* L_n)]$$

on obtient un foncteur $\underline{\Omega}_*(X)$ universel parmi les foncteurs de Borel-Moore orientés avec produit satisfaisant (Dim) et (Sec).

2.3.3. Remarque. — Soit $W \to X \times \mathbf{A}^1$ un morphisme projectif avec W lisse et irréductible. On suppose que le morphisme composé $W \to \mathbf{A}^1$ est transverse à $\{0\}$ et à $\{1\}$ de telle sorte que les fibres W_0 et W_1 correspondantes soient lisses sur k. Par analogie avec le cas complexe, on définit le groupe de cobordisme naïf $\Omega_*^{\mathrm{naif}}(X)$ comme le quotient de $\mathcal{M}_*^+(X)$ par le sous-groupe engendré par les différences $[W_0 \to X] - [W_1 \to X]$ (cobordisme naïf). En appliquant (Sec) au fibré trivial sur W, on vérifie qu'on a un morphisme $\Omega_*^{\mathrm{naif}}(X) \to \underline{\Omega}_*(X)$. En général, ce morphisme n'est pas surjectif.

2.3.4. R_*-foncteurs de Borel-Moore orientés. — Si R_* est un anneau gradué (commutatif avec unité), un R_*-foncteur de Borel-Moore orienté est un foncteur de Borel-Moore orienté avec produit A_* muni d'un morphisme d'anneaux gradués $R_* \to A_*(k)$. Dans ce cas, tous les groupes $A_*(X)$ sont naturellement munis d'une structure de R_*-module pour laquelle toutes les opérations précédemment considérées sont R_* linéaires.

2.3.5. Construction de Ω_*, enfin !— On désire que Ω_* soit un $\mathbf{L}_*$-foncteur de Borel-Moore orienté et que la classe de Chern dans Ω_* du produit de deux fibrés en droites soit donnée par la loi de Lazard universelle $F_{\mathbf{L}}$.

Soit A_* un $\mathbf{L}_*$-foncteur de Borel-Moore orienté vérifiant (Dim). On note F_A l'image de la série de Lazard $F_{\mathbf{L}}$ dans $A_*(k)[[u,v]]$. On considère la condition

(LGF) Pour toute variété lisse Y sur k et toute paire de fibrés en droites (L,M) sur Y, on a

$$F_A(\widetilde{c}_1(L),\widetilde{c}_1(M))(1_Y) = \widetilde{c}_1(L\otimes M)(1_Y).$$

Remarquons que l'évaluation de F_A a un sens grâce à l'axiome (Dim).

On peut maintenant définir Ω_*. Pour cela on considère comme précédemment le quotient des $\mathbf{L}_* \otimes_{\mathbf{Z}} \underline{\Omega}_*(X)$ par des relations convenables (*cf.* [11]) pour obtenir le foncteur Ω_* universel parmi les $\mathbf{L}_*$-foncteurs de Borel-Moore orientés avec produit satisfaisant (Dim), (Sec) et (LGF).

2.3.6. Considérons le foncteur $X \mapsto \mathrm{CH}_*(X)$ qui à une variété X sur k associe son groupe de Chow. C'est un foncteur de Borel-Moore orienté avec produit. De plus on a la relation

$$\widetilde{c}_1(L\otimes M) = \widetilde{c}_1(L) + \widetilde{c}_1(M)$$

pour deux fibrés en droites sur une même base. On en tire une structure de $\mathbf{L}_*$-foncteur sur CH_* correspondant à la loi additive $F_a(u,v) = u+v$ qui munit CH_* d'une structure de $\mathbf{L}_*$-foncteur de Borel-Moore orienté avec produit satisfaisant (Dim), (Sec) et (LGF). On dispose donc d'un morphisme canonique de foncteurs $\Omega_* \to \mathrm{CH}_*$.

2.3.7. Un autre exemple est fourni par le foncteur de G_0-théorie qui à une variété X sur k associe le groupe de Grothendieck $G_0(X)$ de la catégorie des $\mathcal{O}_X$-modules cohérents. On pose $G_0(X)[\beta,\beta^{-1}] := G_0(X)\otimes_{\mathbf{Z}} \mathbf{Z}[\beta,\beta^{-1}]$. On munit ce groupe d'une graduation en décrétant que β est de degré 1. On définit $f_* : G_0(Y)[\beta,\beta^{-1}] \to G_0(X)[\beta,\beta^{-1}]$ pour $f : Y \to X$ un morphisme projectif par extension des scalaires à partir de $f_* : G_0(Y) \to G_0(X)$. Par contre, pour l'image inverse par un morphisme $f : Y \to X$ lisse équidimensionnel de dimension d, on multiplie de plus par β^d. Enfin, si L est un fibré en droites sur X, on définit $\widetilde{c}_1(L)$ comme la multiplication par $(1-[\mathcal{L}^\vee])\beta$, $\mathcal{L}$ désignant le faisceau des sections de L. On a la relation

$$\widetilde{c}_1(L\otimes M) = \widetilde{c}_1(L) + \widetilde{c}_1(M) - \beta\, \widetilde{c}_1(L)\circ\widetilde{c}_1(M)$$

pour deux fibrés en droites sur une même base. On en tire une structure de $\mathbf{L}_*$-foncteur sur $G_0[\beta,\beta^{-1}]$ correspondant à la loi multiplicative $F_m(u,v) = u + v - \beta uv$. On vérifie qu'on a ainsi une structure de $\mathbf{L}_*$-foncteur de Borel-Moore orienté avec produit satisfaisant (Dim), (Sec) et (LGF) et donc un morphisme canonique de foncteurs $\Omega_* \to G_0[\beta,\beta^{-1}]$.

2.3.8. Classe d'un diviseur. — À tout diviseur de Cartier D dans X on peut associer la classe $[D \to X] := \widetilde{c}_1(O_X(D))(1_X)$, $O_X(D)$ désignant le fibré des sections de $\mathcal{O}_X(D)$. Si D et D' sont linéairement équivalents on a $[D \to X] = [D' \to X]$.

2.3.9. Hypersurfaces de Milnor. — On considère des entiers $n > 0$ et $m > 0$. On note γ_n le fibré sur $\mathbf{P}^n$ dont le faisceau des sections est $\mathcal{O}(1)$ et on considère le fibré $\gamma_{n,m} := p_1^*\gamma_n \otimes p_2^*\gamma_m$ sur $\mathbf{P}^n \times \mathbf{P}^m$. On considère l'hypersurface $H_{n,m}$ de $\mathbf{P}^n \times \mathbf{P}^m$ lieu des zéros d'une section de $\gamma_{n,m}$ transverse à la section nulle. Si $m \leqslant n$, on peut choisir $H_{n,m}$ définie par l'équation bihomogène $\sum_{0\leqslant i\leqslant m} X_iY_i = 0$.

L'énoncé suivant résulte aisément des constructions précédentes.

2.3.10. Proposition. — *On a les relations*

$$[H_{n,m} \to \mathbf{P}^n \times \mathbf{P}^m] = \sum_{i\geqslant 0}^{n}\sum_{j\geqslant 0}^{m} a_{i,j}[\mathbf{P}^{n-i} \times \mathbf{P}^{m-j} \to \mathbf{P}^n \times \mathbf{P}^m]$$

dans $\Omega_*(\mathbf{P}^n \times \mathbf{P}^m)$, *et*

$$[H_{n,m}] = [\mathbf{P}^n][\mathbf{P}^{m-1}] + [\mathbf{P}^{n-1}][\mathbf{P}^m] + \sum_{i\geqslant 1}^{n}\sum_{j\geqslant 1}^{m} a_{i,j}[\mathbf{P}^{n-i}][\mathbf{P}^{m-j}]$$

dans $\Omega_*(k)$, *les* $a_{i,j}$ *étant les coefficients de la série de Lazard* $F_{\mathbf{L}}$.

Ceci donne une interprétation géométrique aux coefficients $a_{i,j}$. En particulier, pour $i,j > 1$, $a_{i,j}$ est égal à la classe de $H_{i,j}$ dans $\Omega_*(k)$ modulo des éléments décomposables. Aussi l'image de $\mathbf{L}_*$ dans $\Omega_*(k)$ est engendrée par les images des espaces projectifs et des hypersurfaces de Milnor $H_{i,j}$. On en déduit que pour toute variété X sur k le morphisme canonique $\mathcal{Z}_*(X) \to \Omega_*(X)$ est surjectif : les $a_{i,j}$ étaient en fait déjà là avant qu'on ne les y mette de force !

En fait le morphisme canonique $\mathcal{M}_*^+(X) \to \Omega_*(X)$ est aussi surjectif, autrement dit $\Omega_*(X)$ est engendré comme groupe abélien par des cycles naïfs de la forme $[Y \to X]$. Pour se débarrasser des fibrés L_i dans les générateurs, on commence par remarquer que sur une variété quasi-projective, pour tout fibré en droites L, il existe un fibré en droites M tel que M et $M \otimes L$ soient tous les deux très amples. Comme $\widetilde{c}_1(L)$ s'exprime en fonction de $\widetilde{c}_1(M)$ et $\widetilde{c}_1(L \otimes M)$ on peut supposer que les L_i sont très amples. Quand k est infini, on a alors des sections transverses à la section nulle par Bertini et on conclut par (SEC). Quand k est fini, un petit argument supplémentaire est nécessaire ([11] Lemma 4.17).

2.4. Premières propriétés

On regroupe ici quelques propriétés fondamentales dont la démonstration n'est pas très difficile à partir de ce qui précède.

2.4.1. Formule pour un éclatement. — Soit $i : Z \to X$ une immersion fermée entre variétés lisses. On note X_Z l'éclatement de Z dans X et η le fibré conormal à i. La déformation au cône normal de i est l'éclatement $\pi : Y \to X \times \mathbf{P}^1$ de $X \times \mathbf{P}^1$ le long de $Z \times \{0\}$. Le diviseur exceptionnel P de π s'identifie au projectifié du fibré conormal à $Z \times \{0\}$ et X_Z au transformé strict de $X \times \{0\}$. On note $O_Y(P)$ le fibré en droites associé au diviseur P sur Y et $O_P(-1)$ le dual du fibré canonique quotient sur P. La série $\chi(u)$ donnant l'inverse pour la loi de groupe formel $F_{\mathbf{L}}$ (i.e. vérifiant $F_{\mathbf{L}}(u, \chi(u)) = u$) étant divisible par u, on peut l'écrire $\chi(u) = ug(u)$.

2.4.2. Proposition. — *Avec les notations précédentes, on a les égalités*

$$[X_Z \to Y] = [X \to Y] + \chi([O_Y(P)])$$

dans $\Omega_*(Y)$ *et*

$$[X_Z \to X] = [X \to X] + i_* q_*([g(O_P(-1))])$$

dans $\Omega_*(X)$, q *désignant la projection* $P \to X$.

2.4.3. Trace et dimension zéro. — Le lemme suivant se démontre en considérant un cobordisme naïf (*cf.* remarque 2.3.3) convenable.

2.4.4. Lemme. — *Soit* L *une* k*-algèbre séparable finie. Alors* $[\operatorname{Spec} L] = ([L : k])1$ *dans* $\Omega_0(k)$.

On en déduit aisément :

2.4.5. Proposition. — *Soit* $L|k$ *une extension séparable finie de corps. Alors la composition des morphismes canoniques*

$$\Omega_*(X) \longrightarrow \Omega_*(X \otimes L) \longrightarrow \Omega_*(X)$$

est la multiplication par $[L : k]$.

D'où finalement :

2.4.6. Théorème. — *Pour tout corps* k *le morphisme canonique* $\mathbf{L}_0 \to \Omega_0(k)$ *est un isomorphisme. Pour toute variété lisse de dimension 0,* $X = \operatorname{Spec} A$ *sur* k, $[X] = (\dim_k A)\, 1$ *dans* $\Omega_0(k)$.

3. PROPRIÉTÉS FONDAMENTALES

On aura besoin dans cette section de supposer que k admet la résolution des singularités, ce qui est vérifié pour k de caractéristique 0. Par souci de simplicité, on supposera dans la suite de cet exposé que k est toujours de caractéristique 0, même si ce n'est pas explicite et/ou utile.

3.1. Le théorème de localisation

L'énoncé suivant est vraiment le résultat technique essentiel qui permet la démonstration des propriétés fondamentales de Ω_*.

3.1.1. Théorème. — *On suppose que le corps k admet la résolution des singularités. Soient X une variété sur k, $i : Z \to X$ une sous-variété fermée et $j : U \to X$ l'ouvert complémentaire. Alors on a une suite exacte*

$$\Omega_*(Z) \longrightarrow \Omega_*(X) \longrightarrow \Omega_*(U) \longrightarrow 0.$$

Expliquons pourquoi l'hypothèse que le corps k admette la résolution des singularités permet de démontrer la surjectivité de $j^* : \Omega_*(X) \to \Omega_*(U)$. On part de $Y \to U$ un morphisme projectif avec Y lisse quasi-projective. En appliquant la résolution des singularités et le lemme de Chow à une adhérence convenable de Y on parvient à construire un morphisme projectif $\widetilde{Y} \to X$ prolongeant $Y \to U$ avec $\widetilde{Y}$ lisse quasi-projective, ce qui permet de démontrer la surjectivité au niveau de $\mathcal{M}_*^+$. La suite de la démonstration est bien plus délicate. Elle procède en plusieurs étapes successives, pour lesquelles nous renvoyons à [11], et consiste à prolonger, de façon plus ou moins analogue, certaines relations de U à X. Elle utilise en particulier de façon fondamentale l'énoncé suivant, qui est une conséquence (d'une variante de) la proposition 2.4.2 et de la résolution des singularités.

3.1.2. Proposition. — *Soit $W' \to W$ un morphisme birationnel projectif entre variétés quasi-projectives lisses. On note F et E les lieux exceptionnels de W et W', respectivement. On suppose que E est un diviseur à croisements normaux. Alors il existe η dans $\Omega_*(F)$ tel que*

$$[W' \to W] = [W \to W] + i_*(\eta),$$

i désignant le morphisme d'inclusion de F dans W.

3.2. Invariance par homotopie

3.2.1. Théorème. — *Soit X une variété sur k. Pour tout entier n le morphisme d'image inverse $p^* : \Omega_*(X) \to \Omega_{*+n}(X \times \mathbf{A}^n)$ associé à la projection $p : X \times \mathbf{A}^n \to X$ est un isomorphisme.*

On se ramène au cas $n = 1$ et on commence par démontrer :

3.2.2. Lemme. — *Avec les notations précédentes, pour tout a dans k, le morphisme $T_a : X \times \mathbf{A}^1 \to X \times \mathbf{A}^1$ donné par la translation par a sur le deuxième facteur induit l'identité sur $\Omega_*(X \times \mathbf{A}^1)$*

Démonstration. — Soit $f : Y \to X \times \mathbf{A}^1$. On considère $F : X \times \mathbf{A}^1 \times \mathbf{A}^1 \to X \times \mathbf{A}^1$ donné par $(x, t_1, t_2) \mapsto (x, t_1 + at_2)$. Le produit fibré F^*Y donne un cobordisme naïf (au sens de la remarque 2.3.3) entre f qui est isomorphe à la fibre en 0 et son translaté par a qui est isomorphe à la fibre en 1. □

Suite de la preuve. — Démontrons la surjectivité de p^*. Soit $f : Y \to X \times \mathbf{A}^1$. D'après le lemme précédent on peut supposer que le composé $Y \to \mathbf{A}^1$ est lisse en 0. On considère le produit fibré $Y_m := Y \times_{\mathbf{A}^1} (\mathbf{A}^1 \times \mathbf{A}^1)$, le morphisme $\mathbf{A}^1 \times \mathbf{A}^1 \to \mathbf{A}^1$ étant celui de multiplication, et le cycle $g : Y_m \to X \times \mathbf{A}^1 \times \mathbf{A}^1$ qui s'en déduit en composant avec la projection. On vérifie que ce cycle donne un cobordisme naïf entre f qui est isomorphe à la préimage de $X \times \mathbf{A}^1 \times \{1\}$ et la préimage de $X \times \mathbf{A}^1 \times \{0\}$ qui s'identifie à $p^*(f^{-1}(X \times \{0\}) \times X)$.

On va démontrer l'injectivité de p^* en exhibant un inverse à gauche. Soit $f : Y \to X \times \mathbf{P}^1$ un cycle sur $X \times \mathbf{P}^1$ de classe η dans $\Omega_*(X \times \mathbf{P}^1)$. La classe $\psi(\eta)$ du cycle $f^{-1}(X \times \{a\}) \to X)$ dans $\Omega_{*-1}(X)$ est bien définie et indépendante de a, pour a assez général. En effet $\psi = q_* \circ \widetilde{c}_1(\gamma_1)$ avec $q : X \times \mathbf{P}^1 \to X$ la projection. En particulier ψ s'annule sur l'image de $\Omega_*(X \times \{\infty\})$ dans $\Omega_*(X \times \mathbf{P}^1)$. Par le théorème de localisation 3.1.1, il s'ensuit que ψ se factorise par un morphisme $\Omega_*(X \times \mathbf{A}^1) \to \Omega_{*-1}(X)$ qui est l'inverse à gauche de p^* désiré. □

3.3. Fibrés projectifs

Soit X une k-variété et soit E un fibré vectoriel de rang n sur X. On considère le fibré projectif $q : \mathbf{P}(E) \to X$ représentant les quotients de rang 1 de E. On note $O(1) \to \mathbf{P}(E)$ le fibré en droites quotient canonique de $q^*(E)$. Pour tout entier i dans $\{0, \ldots, n\}$ on considère le morphisme

$$\xi^{(i)} : \Omega_{*-n+i}(X) \longrightarrow \Omega_*(\mathbf{P}(E))$$

défini en posant $\xi^{(i)} := \widetilde{c}_1(O(1))^i \circ q^*$.

Par des méthodes essentiellement classiques, Levine et Morel obtiennent à partir des théorèmes 3.1.1 et 3.2.1, la formule suivante pour les fibrés projectifs.

3.3.1. Théorème. — *Soit X une k-variété et soit E un fibré vectoriel de rang n sur X. Le morphisme canonique*

$$\oplus_{i=0}^{n-1} \xi^{(i)} : \oplus_{i=0}^{n-1} \Omega_{*-n+i}(X) \longrightarrow \Omega_*(\mathbf{P}(E))$$

est un isomorphisme.

En utilisant ce résultat joint au théorème 3.1.1, ils obtiennent la forme plus générale suivante de l'invariance par homotopie.

3.3.2. Théorème. — *Soit X une k-variété, soit E un fibré vectoriel de rang n sur X et soit $p : V \to X$ un E-torseur. Alors le morphisme canonique*

$$p^* : \Omega_*(X) \longrightarrow \Omega_{*+n}(V)$$

est un isomorphisme.

3.4. Classes de Chern

Soit X une k-variété et soit E un fibré vectoriel de rang n sur X. On déduit du théorème 3.3.1 que $\Omega_*(\mathbf{P}(E))$ est muni d'une structure canonique de $\operatorname{End}\Omega_*(X)$-module gradué à gauche. On en tire l'existence de morphismes

$$\widetilde{c}_i(E) : \Omega_*(X) \longrightarrow \Omega_*(X)$$

pour $0 \leqslant i \leqslant n$, caractérisés par les relations

$$\sum_{i=0}^{n} \widetilde{c}_i(E)\widetilde{c}_1(O(1))^{n-i} = 0$$

et $\widetilde{c}_0(E) = \mathrm{id}$, les classes de Chern de E, qui vérifient les propriétés « usuelles » des classes de Chern.

Il est aussi possible de définir ces classes de Chern à partir des classes de Chern des fibrés en droites en utilisant le principe de scindage. Plus généralement considérons l'anneau gradué $\mathbf{Z}[t] = \mathbf{Z}[t_1, \ldots, t_n, \ldots]$ des polynômes à coefficients entiers en une infinité de variables t_i de degré 1 et posons

$$\Omega_*(X)[t] := \Omega_*(X) \otimes_{\mathbf{Z}} \mathbf{Z}[t].$$

Pour tout fibré en droites L sur X on considère l'automorphisme $\widetilde{c}_t(L) := \sum_{i=0}^{\infty} \widetilde{c}_1(L)^i t_i$ de $\Omega_*(X)[t]$. Par le principe de scindage, cette construction s'étend à tout fibré vectoriel E de rang n sur X. On développe alors $\widetilde{c}_t(E)$ en

$$\widetilde{c}_t(E) = \sum_I \widetilde{c}_I(E) t^I,$$

I parcourant $\mathbf{N}^{\mathbf{N}}$. Les coefficients $\widetilde{c}_I(E)$ sont les classes de Chern de Conner et Floyd attachées à E. On retrouve les classes de Chern usuelles comme coefficients des t_i.

4. THÉORÈMES DE COMPARAISON

4.1. Loi de groupe formel multiplicative et invariance birationnelle

On s'intéresse dans cette section à des foncteurs de Borel-Moore orientés avec produit satisfaisant (Dim), (Sec) et (LGF) avec loi de groupe formel multiplicative, c'est-à-dire de la forme

$$u + v - \beta uv,$$

avec β non nécessairement inversible. Le foncteur $\Omega_*^\beta := \Omega_* \otimes_{\mathbf{L}_*} \mathbf{Z}[\beta]$ est universel parmi de tels foncteurs. On notera avec un indice α_β l'image d'un élément α de $\Omega_*(X)$ dans $\Omega_*^\beta(X)$. Il résulte de la proposition 2.3.10 que $[\mathbf{P}^1]_\beta = \beta$. De plus on peut montrer, à l'aide de la proposition 2.4.2, du théorème 3.3.2 et de l'astuce de Jouanolou, que, pour tout fibré vectoriel E sur X lisse quasi-projective de rang n, on a

$$[\mathbf{P}(E) \to X]_\beta = \beta^{n-1}[X \to X]_\beta. \tag{4.1.1}$$

En particulier on a $[\mathbf{P}^n]_\beta = \beta^n$. La formule pour un éclatement 2.4.2 se simplifie dramatiquement alors pour donner :

4.1.1. Proposition. — *Soit X une variété quasi-projective lisse quasi-projective sur k et soit Z une sous-variété lisse fermée. Alors on a*

$$[X_Z \to X]_\beta = [X \to X]_\beta$$

dans $\Omega_*^\beta(X)$.

En utilisant le théorème de factorisation faible de Abramovich, Karu, Matsuki, et Włodarczyk ([1], [3]), on en déduit :

4.1.2. Corollaire. — *Soient Z et Z' deux variétés projectives lisses birationnellement équivalentes. Alors $[Z]_\beta = [Z']_\beta$ dans $\Omega_*^\beta(k)$.*

Plus précisément, on a le résultat suivant :

4.1.3. Proposition. — *Le noyau du morphisme canonique*

$$\Omega_*(k) \longrightarrow \Omega_*^\beta(k)$$

est l'idéal I engendré par les différences $[Z]-[Z']$ avec Z et Z' des variétés projectives lisses birationnellement équivalentes.

Démonstration. — Au vu du corollaire 4.1.2, il suffit de démontrer que $\Omega_*(k) = I + \mathbf{Z}[\mathbf{P}^1]$, ce qui résulte du fait que les hypersurfaces de Milnor $H_{n,m}$ sont birationnellement équivalentes à $\mathbf{P}^{n+m-1}$ et de ce que $\Omega_*(k)$ est engendré par les classes des hypersurfaces de Milnor et des espaces projectifs (ce dernier point est conséquence de l'isomorphisme $\mathbf{L}_* \to \Omega_*(k)$ démontré en 4.3.2). □

On en déduit le théorème :

4.1.4. Théorème. — *Le foncteur $X \mapsto \Omega_*(X) \otimes_{\mathbf{L}_*} \mathbf{Z}[\beta]$ est universel parmi les $\mathbf{L}_*$-foncteurs de Borel-Moore orientés avec produit satisfaisant* (Dim), (Sec) *et* (LGF) *satisfaisant la condition d'invariance birationnelle suivante : pour tout morphisme projectif birationnel $f : Y \to X$ entre variétés lisses irréductibles, on a $f_* 1_Y = 1_X$.*

Démonstration. — Il résulte de la proposition 4.1.3 que si A_* est un $\mathbf{L}_*$-foncteur de Borel-Moore orienté avec produit satisfaisant (Dim), (Sec) et (LGF) satisfaisant la condition d'invariance birationnelle, alors on a un morphisme de foncteurs $\Omega_* \otimes_{\mathbf{L}_*} \mathbf{Z}[\beta] \to A_*$. Il reste à vérifier que le foncteur $\Omega_* \otimes_{\mathbf{L}_*} \mathbf{Z}[\beta]$ satisfait la condition d'invariance birationnelle, ce qui résulte de la proposition 4.1.1 et du théorème de factorisation faible. □

4.2. Comparaison avec le K^0

Dans ce numéro on restreindra des foncteurs de Borel-Moore orientés avec produit A_* à la catégorie $\mathcal{L}'_k$. On passera en numérotation cohomologique en posant $A^*(X) = A_{\dim X - *}(X)$.

Pour X dans $\mathcal{L}'_k$ le morphisme canonique entre groupes de Grothendieck $K^0(X) \to G_0(X)$ est un isomorphisme (on rappelle que $K^0(X)$ désigne le groupe de Grothendieck des faisceaux localement libres de type fini sur X). On identifiera donc la restriction du foncteur $G_0[\beta, \beta^{-1}]$ à $\mathcal{L}'_k$ au foncteur $K^0[\beta, \beta^{-1}] := K^0 \otimes_{\mathbf{Z}} \mathbf{Z}[\beta, \beta^{-1}]$, en convenant maintenant que β est de degré -1.

4.2.1. Théorème. — *Soit k un corps de caractéristique zéro. Le morphisme canonique de foncteurs de Borel-Moore orientés avec produit sur $\mathcal{L}'_k$*

$$\Omega^* \longrightarrow K^0[\beta, \beta^{-1}]$$

induit un isomorphisme

$$\Omega^* \otimes_{\mathbf{L}_*} \mathbf{Z}[\beta, \beta^{-1}] \simeq K^0[\beta, \beta^{-1}].$$

Ébauche de preuve. — Compte tenu du caractère universel de Ω_*, il suffit de montrer qu'il existe un unique morphisme de foncteurs de Borel-Moore orientés avec produit sur $\mathcal{L}'_k$:

$$\lambda : K^0[\beta, \beta^{-1}] \longrightarrow \Omega^* \otimes_{\mathbf{L}_*} \mathbf{Z}[\beta, \beta^{-1}].$$

Remarquons que nécessairement λ doit envoyer β sur la classe de $[\mathbf{P}^1]$, autrement dit β. De plus, pour X dans $\mathcal{L}_k$, on a, pour tout fibré en droites L de faisceau des sections $\mathcal{L}$, la relation $[\mathcal{L}] = 1 - c_1(L^\vee)\beta$ dans $K^0[\beta, \beta^{-1}]$, et donc nécessairement λ doit envoyer $[\mathcal{L}]$ sur $1 - c_1(L^\vee)\beta$ dans $\Omega^* \otimes_{\mathbf{L}_*} \mathbf{Z}[\beta, \beta^{-1}]$, ce qui garantit l'unicité de λ par le principe de scindage. Pour définir λ, on pose $\lambda([\mathcal{E}]\beta^n) = (r - c_1(E^\vee)\beta)\beta^n$, pour E un fibré vectoriel de rang r sur X de faisceau des sections $\mathcal{E}$. Par le principe de scindage, λ est un morphisme d'anneaux gradués. Le fait qu'il commute à l'image inverse par les morphismes lisses étant clair, il reste à vérifier qu'il commute à l'image directe par les morphismes projectifs. Ceci se démontre de façon tout à fait analogue[(1)] à la preuve du théorème de Riemann-Roch-Grothendieck donnée dans [5]. Il suffit de vérifier l'énoncé pour une projection $\mathbf{P}^n_k \times X \to X$ sur un schéma lisse et pour une immersion fermée

[(1)] En fait, d'après [11] remarque 8 page 93, le théorème est essentiellement équivalent au théorème de Riemann-Roch-Grothendieck.

$Y \to X$ entre schémas lisses. Pour le premier cas on se ramène au cas où $X = \operatorname{Spec} k$, par compatibilité avec le produit externe, qui est de vérification directe. Le cas de l'immersion fermée se démontre lui par déformation au cône normal. □

4.3. Calcul de $\Omega_*(k)$

Commençons par déterminer $\Omega_*(k) \otimes_{\mathbf{L}_*} \mathbf{Z}$.

4.3.1. Proposition. — *Soit k un corps de caractéristique zéro. On a*

$$\Omega_*(k) \otimes_{\mathbf{L}_*} \mathbf{Z} \simeq \mathbf{Z}.$$

Démonstration. — Posons $\Omega^{\mathrm{ad}}_*(k) = \Omega_*(k) \otimes_{\mathbf{L}_*} \mathbf{Z}$. D'après le théorème 2.4.6, $\Omega^{\mathrm{ad}}_0(k) \simeq \mathbf{Z}$. Il suffit donc de démontrer que $\Omega^{\mathrm{ad}}_n(k) = 0$, pour $n > 0$, autrement dit que la classe de X dans $\Omega^{\mathrm{ad}}_*(k)$ est nulle, pour toute variété projective lisse X de dimension > 0. D'après 4.1, la classe de $\mathbf{P}^n$ est nulle pour $n > 0$. Si $n > 0$, la classe de toute hypersurface lisse Y dans $\mathbf{P}^{n+1}$ est nulle ; en effet une telle hypersurface est linéairement équivalente au diviseur $d\mathbf{P}^n$ et $[d\mathbf{P}^n] = d[\mathbf{P}^n] = 0$ dans $\Omega^{\mathrm{ad}}_*(k)$. On utilise ici l'additivité de la loi de groupe formel associée à $\Omega^{\mathrm{ad}}_*(k)$ ainsi que les notations de 2.3.8. Considérons maintenant une variété projective lisse irréductible Y. Elle est birationnellement équivalente à une hypersurface réduite (peut-être singulière) $\overline{Y}$. Par le théorème d'Hironaka, il existe un morphisme $\mu : S \to \mathbf{P}^{n+1}$ composé d'éclatements de centres lisses tel que le diviseur $\mu^{-1}(\overline{Y})$ soit à croisements normaux. En particulier la transformée stricte $\widetilde{Y}$ de $\overline{Y}$ est lisse et birationnellement équivalente à Y. On a donc $[Y] = [\widetilde{Y}]$ dans $\Omega^{\mathrm{ad}}_*(k)$. Il résulte de (4.1.1) appliqué aux centres des éclatements successifs que $[\widetilde{Y}] = [\mu^{-1}(\overline{Y})]$ dans $\Omega^{\mathrm{ad}}_*(k)$. Pour conclure, on remarque que la classe de $\mu^{-1}(\overline{Y})$ est égale à celle de $\mu^{-1}(D)$ pour tout diviseur D de même degré que $\overline{Y}$, et que $\mu^{-1}(D)$ est lisse et birationnellement équivalent à D pour D assez général. □

4.3.2. Théorème. — *Soit k un corps de caractéristique zéro. Le morphisme canonique*

$$\Psi : \mathbf{L}_* \longrightarrow \Omega_*(k)$$

est un isomorphisme.

Ébauche de preuve. — Démontrons l'injectivité. Lorsque k est un sous-corps de $\mathbf{C}$, on peut considérer le foncteur cohomologique $X \mapsto U^{2*}(X(\mathbf{C}))$ sur les variétés lisses sur k. Par le caractère universel du cobordisme algébrique, on en tire un morphisme canonique $\Omega^*(X) \to U^{2*}(X(\mathbf{C}))$, et en particulier un morphisme canonique d'anneaux $\theta : \Omega_*(k) \to U_{2*}(\text{point})$. Mais le composé $\theta \circ \Psi$ n'est autre que l'isomorphisme

de Quillen $\delta : \mathbf{L}^* \to U^*(\text{point})$, ce qui prouve l'injectivité de Ψ. Pour démontrer l'injectivité sans hypothèse sur k, Levine et Morel considèrent le morphisme canonique[(2)]

$$\Omega_*(k) \longrightarrow \mathbf{Z}[t_1, \dots, t_n, \dots]$$

qui associe à une variété projective lisse sur k « l'ensemble de ses nombres de Chern ». En composant avec le morphisme canonique $\mathbf{L}_* \to \Omega_*(k)$, on en tire un morphisme

$$\mathbf{L}_* \longrightarrow \mathbf{Z}[t_1, \dots, t_n, \dots].$$

Pour conclure, Levine et Morel démontrent que ce morphisme correspond à la loi de groupe formel $\lambda^{-1}(\lambda(u) + \lambda(v))$, avec $\lambda(u)$ la série $\sum_{i \geqslant 0} t_i u^i$ et $\lambda^{-1}(u)$ son inverse pour la composition, dont il est connu qu'il est injectif [20]. Cette preuve n'utilise pas le théorème de Quillen.

Enfin, il résulte de la proposition 4.3.1 que, pour $i > 0$, $\Omega_i(k) = \sum_{1 \leqslant j \leqslant i} \mathbf{L}_j \Omega_{i-j}(k)$, et la surjectivité suit. □

4.3.3. Remarque. — La preuve du théorème précédent est toute différente (et indépendante) de celle du théorème de Quillen. En effet, comme nous l'avons déjà signalé, celle-ci repose de façon essentielle sur les propriétés de finitude du cobordisme complexe, dont l'analogue strict en cobordisme algébrique n'est vraisemblablement pas vérifié en général. Par contre on a utilisé des arguments géométriques qui n'ont pas d'analogue en topologie : la suite exacte de localisation, le théorème d'Hironaka, le fait que toute variété soit birationnellement une hypersurface, etc.

4.3.4. Corollaire. — *Soit $k \subset K$ une extension de corps de caractéristique zéro. Le morphisme canonique $\Omega_*(k) \to \Omega_*(K)$ est un isomorphisme.*

4.4. Comparaison avec les groupes de Chow

4.4.1. Théorème. — *Soit k un corps de caractéristique zéro. Le morphisme canonique de foncteurs de Borel-Moore orientés*

$$\Omega_* \longrightarrow \mathrm{CH}_*$$

induit un isomorphisme

$$\Omega_* \otimes_{\mathbf{L}_*} \mathbf{Z} \longrightarrow \mathrm{CH}_*.$$

Démonstration. — Commençons par remarquer que si $\widetilde{Z} \to Z$ est un morphisme projectif birationnel entre variétés lisses, alors l'image de $[\widetilde{Z} \to Z]$ dans $\Omega_*(Z) \otimes_{\mathbf{L}_*} \mathbf{Z}$ est égale à 1_Z. En effet cela résulte du théorème 5.2.1 car les termes du second membre de (5.2.1) sont nuls dans $\Omega_*(Z) \otimes_{\mathbf{L}_*} \mathbf{Z}$ pour des raisons de degré.

Soit Z un sous-schéma fermé intègre d'un schéma X de type fini sur k. Il résulte de la remarque précédente que la classe de $[\widetilde{Z} \to X]$ dans $\Omega_*(X) \otimes_{\mathbf{L}_*} \mathbf{Z}$, pour $\widetilde{Z} \to Z$ un morphisme projectif birationnel avec $\widetilde{Z}$ lisse, ne dépend que de Z. On la note $[Z \hookrightarrow X]$.

(2) Merkurjev a démontré, sans hypothèse sur k, que ce morphisme se factorise en une rétraction $\Omega_*(k) \to \mathbf{L}_*$ du morphisme canonique $\mathbf{L}_* \to \Omega_*(k)$.

On définit ainsi un morphisme de groupes abéliens $\Phi : Z_*(X) \to \Omega_*(X) \otimes_{\mathbf{L}_*} \mathbf{Z}$ dont le composé avec le morphisme $\Omega_*(X) \otimes_{\mathbf{L}_*} \mathbf{Z} \to \mathrm{CH}_*(X)$ est le morphisme canonique $Z_*(X) \to \mathrm{CH}_*(X)$. Comme il résulte du théorème 5.2.1 que le morphisme Φ est surjectif, il reste à savoir que Φ se factorise par $\mathrm{CH}_*(X)$, ce qui est donné par le lemme suivant. □

4.4.2. Lemme. — *Soit W un sous-schéma fermé intègre d'un schéma X de type fini sur k et soit f dans $k(W)^*$ une fonction rationnelle non nulle de diviseur* $\mathrm{div}(f)$ *dans $Z_*(X)$. Alors $\Phi(\mathrm{div}(f)) = 0$ dans $\Omega_*(X) \otimes_{\mathbf{L}_*} \mathbf{Z}$.*

Ébauche de preuve. — Grâce à la résolution des singularités on peut considérer un morphisme $\pi : \widetilde{W} \to W$ avec $\widetilde{W}$ quasi-projective lisse tel que f induise un morphisme $\widetilde{f} : \widetilde{W} \to \mathbf{P}^1$. On peut de plus supposer que le diviseur de $\widetilde{f}$ soit de la forme $D_0 - D_\infty$ avec D_0 et D_∞ des diviseurs à croisements normaux ne s'intersectant pas. Considérons les classes $[D_0 \to \widetilde{W}]$ et $[D_\infty \to \widetilde{W}]$ définies dans 2.3.8. Comme on a $[D_0 \to \widetilde{W}] = [D_\infty \to \widetilde{W}]$, il suffit de démontrer que $\Phi(D_0)$ est égale à l'image de $[D_0 \to \widetilde{W}]$ dans $\Omega_*(X) \otimes_{\mathbf{L}_*} \mathbf{Z}$, et de même pour D_∞, ce qui résulte des formules explicites pour la classe d'un diviseur à croisements normaux de [11] et de la proposition 4.3.1. □

En particulier, pour X une variété algébrique complexe lisse et projective, on déduit du théorème 4.4.1 l'existence d'un morphisme canonique

$$\mathrm{CH}^*(X) \longrightarrow U^{2*}(X(\mathbf{C})) \otimes_{\mathbf{L}^*} \mathbf{Z}.$$

Ce morphisme avait été introduit par B. Totaro dans son travail sur les cycles algébriques de torsion [22].

5. LA FORMULE DU DEGRÉ

On rappelle que l'on fait partout l'hypothèse que k est de caractéristique zéro. On renvoie à [11], [15], [16], [14] pour une discussion de ce qui reste connu en caractéristique positive.

5.1. Définition du degré

Soit X une variété irréductible sur k de corps des fractions $k(X)$. Pour tout ouvert U non vide de X on dispose d'un morphisme canonique $\Omega_*(U) \to \Omega_*(k(X))$ et on vérifie que le morphisme canonique

$$\varinjlim \Omega_*(U) \to \Omega_*(k(X))$$

est un isomorphisme. En particulier on dispose d'un morphisme $\Omega_*(X) \to \Omega_*(k(X))$, qui composé avec l'inverse de l'isomorphisme du corollaire 4.3.4 fournit un morphisme canonique $\deg : \Omega_*(X) \to \Omega_*(k)$. Plus généralement, si $X_1, \ldots, X_n$ sont les

composantes irréductibles de X, on définit de façon similaire des morphismes degré $\deg_i : \Omega_*(X) \to \Omega_*(k)$, pour $1 \leqslant i \leqslant n$.

5.2. Formule générale du degré

Nous sommes maintenant en mesure d'énoncer la formule générale du degré de Levine et Morel.

5.2.1. Théorème (Formule générale du degré). — *Soit X un schéma de type fini sur k de composantes irréductibles $X_1, \ldots, X_n$. Pour tout sous-schéma fermé Z de X, on fixe un morphisme projectif birationnel $\widetilde{Z} \to Z$ avec $\widetilde{Z}$ dans $\mathcal{L}_k$. Alors, pour tout élément α de $\Omega_*(X)$, on peut écrire*

$$\alpha - \sum_{1 \leqslant i \leqslant n} \deg_i(\alpha)[\widetilde{X}_i \to X_i] = \sum_{Z,\, \mathrm{codim} Z > 0} \omega_Z[\widetilde{Z} \to X] \tag{5.2.1}$$

pour un choix convenable d'éléments ω_Z dans $\Omega_*(k)$.

Ébauche de preuve. — On raisonne par récurrence sur la dimension de X, en utilisant le théorème de localisation 3.1.1. □

5.2.2. Corollaire. — *Pour tout schéma de type fini X, $\Omega_*(X)$ est engendré comme $\Omega_*(k)$-module par les classes de degré $\leqslant \dim X$. Si, de plus, X est irréductible, fixons un morphisme projectif birationnel $\widetilde{X} \to X$ avec X dans $\mathcal{L}_k$. Alors $\Omega_*(X)$ est engendré comme $\Omega_*(k)$-module par les classes de degré $\leqslant \dim X - 1$ et $[\widetilde{X} \to X]$.*

5.2.3. Remarque. — Le corollaire précédent est l'analogue algébrique d'un résultat de Quillen [20] selon lequel l'anneau de cobordisme complexe $U_*(X)$ est engendré comme $\mathbf{L}_*$-module par les classes de degré $\leqslant 2 \dim X$, si X est un complexe fini.

Soit X une variété projective irréductible et lisse sur k. Considérons le noyau $\widetilde{\Omega}_*(X)$ du morphisme $\deg : \Omega_*(X) \to \Omega_*(k)$. Remarquons que le morphisme composé $\Omega_*(k) \to \Omega_*(X) \to \Omega_*(k)$ étant l'identité, on a un scindage canonique $\Omega_*(X) = \Omega_*(k) \oplus \widetilde{\Omega}_*(X)$.

Il résulte du théorème 5.2.1 que l'image de $\widetilde{\Omega}_*(X)$ par le morphisme d'image directe $\Omega_*(X) \to \Omega_*(k)$ associé au morphisme structural $X \to \operatorname{Spec} k$ coïncide avec l'idéal $M(X)$ de $\Omega_*(k)$ engendré par les classes $[Y]$, pour Y projective lisse de dimension $< \dim X$ admettant un k-morphisme $Y \to X$.

L'énoncé suivant qui est également une conséquence du théorème 5.2.1, a été conjecturé par Rost (*cf.* [15]).

5.2.4. Théorème. — *Soit $f : Y \to X$ un morphisme entre variétés projectives lisses irréductibles. Alors $[Y] - \deg(f)[X]$ appartient à l'idéal $M(X)$ de $\Omega_*(k)$.*

5.3. La formule du degré de Rost

Soit $f : Y \to X$ un morphisme entre variétés projectives lisses irréductibles de même dimension d. Remarquons que dans ce cas $\deg(f)$ appartient à $\Omega_0(k)$, on peut donc le voir comme un entier, égal au degré usuel d'après le lemme 2.4.4. On peut ainsi réécrire (5.2.1) en

$$[Y] - \deg(f)[X] = \sum_{Z,\, \mathrm{codim Z}>0} \omega_Z[\widetilde{Z} \to X] \tag{5.3.1}$$

pour un choix convenable d'éléments ω_Z dans $\Omega_{d-\dim Z}(k)$.

Pour toute variété projective lisse X de dimension d, on considère N_d, le d-ième polynôme de Newton en les classes de Chern du fibré tangent dans $\mathrm{CH}_0(X)$ et on note $s_d(X)$ l'entier $-\deg N_d$ (il s'agit ici du degré usuel $\mathrm{CH}_0(X) \to \mathbf{Z}$). Les faits suivants sont bien connus des topologues (*cf.* [2]).

(S1) Si d est de la forme $d = p^n - 1$ avec p un nombre premier, alors l'entier $s_d(X)$ est divisible par p.

(S2) Pour X et X' de dimension d et d' tous deux > 0, on a $s_{d+d'}(X \times X') = 0$.

On va maintenant déduire de la formule générale du degré l'énoncé suivant, dû à Rost (*cf.* [15], [16]), qui généralise des résultats antérieurs de Voevodsky.

5.3.1. Théorème (Formule du degré de Rost). — *Soit $f : Y \to X$ un morphisme entre variétés projectives lisses irréductibles de même dimension $d > 0$. On suppose que d est de la forme $d = p^n - 1$ avec p un nombre premier. Alors il existe un zéro-cycle z dans $\mathrm{CH}_0(X)$ de degré égal à*

$$\frac{s_d(Y)}{p} - \deg(f)\frac{s_d(X)}{p}.$$

Démonstration. — Si on applique s_d à l'égalité (5.3.1), on obtient

$$s_d(Y) - \deg(f)s_d(X) = \sum_{Z,\, \mathrm{dim Z}=0} s_d(\omega_Z[Z \to X])$$

pour un choix convenable d'éléments ω_Z dans $\Omega_d(k)$, car les Z de dimension > 0 disparaissent grâce à (S2). Le zéro-cycle $z = \sum \frac{s_d(\omega_Z)}{p} Z$ vérifie alors la propriété requise. □

5.3.2. Remarque. — Il est important de remarquer qu'il est vraiment nécessaire de travailler avec Ω_* pour pouvoir utiliser (S2) dans la preuve précédente. En particulier celle-ci ne fonctionnerait plus si on se contentait de travailler directement dans l'anneau de Chow.

5.3.3. Remarque. — La version plus générale du théorème 5.3.1 due à Borghesi [4] peut également être déduite du théorème 5.2.1.

La formule du degré de Rost admet un certain nombre de conséquences frappantes (*cf.* [16]). Elle permet en particulier de retrouver les résultats suivants de la théorie des formes quadratiques (*cf.* [16]).

5.3.4. Théorème (Hoffmann [6]). — *Soient Q_1 et Q_2 deux quadriques anisotropes sur k. Si Q_1 est de dimension $\geqslant 2^r-1$ et si Q_2 est anisotrope sur le corps des fonctions de Q_1, alors Q_2 est de dimension $\geqslant 2^r-1$.*

5.3.5. Théorème (Izhboldin [7]). — *Soient Q_1 et Q_2 deux quadriques anisotropes sur k. Si Q_1 est de dimension 2^r-1 et si Q_2 est isotrope sur le corps des fonctions de Q_1, alors Q_1 est isotrope sur le corps des fonctions de Q_2.*

6. IMAGES INVERSES EN COBORDISME ALGÉBRIQUE

On présente dans cette section des résultats contenus dans [10].

6.1. Théories homologiques de Borel-Moore orientées

Pour les besoins de la construction de Ω_*, nous avons été minimalistes pour ce qui est des propriétés demandées a priori. En particulier nous n'avons considéré que des images inverses par des morphismes lisses. Il est souhaitable d'avoir des images inverses en plus grande généralité, comme celle fournie par le formalisme des théories homologiques de Borel-Moore orientées.

Une théorie de Borel-Moore orientée sur $\mathcal{V}_k$ consiste en les données suivantes [10] :

(D1) Un foncteur additif

$$H_* : \mathcal{V}_k' \longrightarrow \mathbf{Ab}_*.$$

(D2) Pour tout morphisme $f : Y \to X$ dans $\mathcal{V}_k$ localement d'intersection complète[(3)] de dimension relative d, un morphisme

$$f^* : H_*(X) \longrightarrow H_{*+d}(Y).$$

(D3) Un produit externe bilinéaire

$$H_*(X) \times H_*(Y) \longrightarrow H_*(X \times_k Y)$$

qui est associatif, commutatif, et admettant une unité 1 de $H_0(k)$.

On demande de plus que ces données vérifient les propriétés suivantes :

(BM1) Fonctorialité de f^*.

[(3)] i.e. qui admet une factorisation $f = p \circ i$ avec i une immersion régulière fermée et p un morphisme lisse.

(BM2) Si $f : X \to Z$ est un morphisme projectif et $g : Y \to Z$ est un morphisme localement d'intersection complète et que g est transverse à f, on a $g^* f_* = f'_* g'^*$, pour tout diagramme cartésien

$$\begin{array}{ccc} W & \xrightarrow{g'} & X \\ {\scriptstyle f'}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow{g} & Z. \end{array}$$

(BM3) Compatibilité de f_* et f^* au produit externe.

(PB) Soit E un fibré vectoriel de rang n sur X dans $\mathcal{V}_k$, le morphisme canonique

$$\oplus_{i=0}^{n-1} \xi^{(i)} : \oplus_{i=0}^{n-1} H_{*-n+i}(X) \longrightarrow H_*(\mathbf{P}(E))$$

est un isomorphisme, le morphisme $\xi^{(i)} : H_{*-n+i}(X) \to H_*(\mathbf{P}(E))$ étant défini comme en (3.3) par $\xi^{(i)} := \widetilde{c}_1(O(1))^i \circ q^*$.

(H) Soit X dans $\mathcal{V}_k$, soit E un fibré vectoriel de rang n sur X et soit $p : V \to X$ un E-torseur. Alors $p^* : H_*(X) \to H_{*+n}(V)$ est un isomorphisme.

Bien sûr un tel foncteur est un foncteur de Borel-Moore orienté avec produits au sens de (2.2.3), la première classe de Chern d'un fibré en droites $\mathcal{L}$ sur une variété X étant définie par $\widetilde{c}_1(\mathcal{L}) := i^* \circ i_*$, pour $i : X \to \mathcal{L}$ la section nulle (ce qui permet de définir $\xi^{(i)}$ dans (PB)).

6.2. Images inverses en cobordisme algébrique

La théorie de l'intersection de Fulton [5] permet de munir le foncteur $X \mapsto \mathrm{CH}_*(X)$ d'une structure de théorie de Borel-Moore orientée. On peut également munir le foncteur $G_0[\beta, \beta^{-1}]$ d'une telle structure.

6.2.1. Théorème. — *Si k est de caractéristique zéro, le cobordisme algébrique Ω_* admet une unique structure de théorie de Borel-Moore orientée compatible avec les données déjà définies. De plus le cobordisme algébrique est universel parmi les théories de Borel-Moore orientées.*

Le point-clé dans cet énoncé est la construction d'images inverses fonctorielles $f^* : \Omega_*(X) \to \Omega_{*+d}(Y)$ pour $f : Y \to X$ une immersion régulière fermée de dimension relative d vérifiant les propriétés demandées. Ce travail est mené à bien dans [10], et nécessite de reprendre en les étendant au cadre du cobordisme algébrique l'essentiel des constructions de la théorie de l'intersection de [5].

7. RELATIONS AVEC LA THÉORIE HOMOTOPIQUE DES SCHÉMAS

7.1. Le spectre du cobordisme complexe

Commençons par rappeler l'interprétation homotopique du cobordisme complexe.

On note $G_{n,m}$ la grassmannienne des n-plans dans $\mathbf{C}^{n+m}$ et $\xi_{n,m}$ le fibré complexe tautologique de rang n sur $G_{n,m}$. Le classifiant $\mathrm{BU}(n)$ du groupe $U(n)$ (et des fibrés complexes de rang n) est la colimite des espaces $G_{n,m}$. Il est muni du fibré ξ_n, colimite des fibrés $\xi_{n,m}$. On note $\mathrm{MU}(n)$ l'espace de Thom du fibré ξ_n et on définit $\mathbf{MU}$, le spectre du cobordisme complexe, par $\mathbf{MU}_{2n} = \mathrm{MU}(n)$ et $\mathbf{MU}_{2n+1} = \Sigma\mathrm{MU}(n)$, Σ désignant le foncteur de suspension.

L'énoncé suivant est une reformulation du théorème de Thom sur le cobordisme (dans le cas complexe).

7.1.1. Théorème. — *Pour toute variété X, $U^q(X)$ est canoniquement isomorphe à* $\mathrm{colim}\,[\Sigma^{2k-q}X, \mathrm{MU}(k)]$. *Autrement dit, le spectre* $\mathbf{MU}$ *représente*[(4)] *la théorie cohomologique U^* dans la catégorie homotopique stable.*

7.2. Le spectre du cobordisme algébrique

Rappelons la construction de la catégorie homotopique stable en géométrie algébrique d'après [23], [18]. Morel et Voevodsky ont défini dans [19] la catégorie homotopique $\mathcal{H}(k)$ des schémas lisses sur un corps k. C'est une catégorie de modèles fermée dont les objets sont des faisceaux simpliciaux pour la topologie de Nisnevich. En particulier, tout fibré vectoriel $\mathcal{E}$ sur une k-variété admet un espace de Thom $\mathrm{Th}(\mathcal{E})$ dans $\mathcal{H}(k)$. Ainsi on peut définir dans la catégorie $\mathcal{H}(k)$ l'analogue noté $\mathrm{MGL}(n)$ de l'espace $\mathrm{MU}(n)$ considéré en (7.1). Remarquons que dans la catégorie homotopique $\mathcal{H}(k)$ on dispose de deux « cercles » différents, le cercle simplicial S^1_s et le cercle de Tate $S^1_t := \mathbf{A}^1 \setminus \{0\}$. On vérifie que leur smash-produit T est faiblement équivalent à la droite projective. La catégorie homotopique stable $\mathcal{S}H(k)$ est définie[(5)] en localisant par rapport au smash-produit $\Sigma_T : X \mapsto T \wedge X$. On définit alors les T-spectres de façon usuelle, comme une suite d'objets E_i munis de morphismes $\Sigma_T(E_i) \to E_{i+1}$. À tout T-spectre $\mathbf{E}$ est associée une théorie cohomologique bigraduée

$$E^{p,q}(X) := \mathrm{Hom}_{\mathcal{S}H(k)}(\Sigma^\infty, S^{p,q} \wedge \mathbf{E})$$

sur $\mathcal{H}(k)$, en posant $S^{p,q} := (S^1_s)^{p-q} \wedge (S^1_t)^q$.

En particulier, on peut associer naturellement à la suite d'espaces $\mathrm{MGL}(n)$ un T-spectre noté $\mathbf{MGL}$ et une théorie cohomologique bigraduée $\mathrm{MGL}^{*,*}$. Ceci illustre

(4) Au grain de sel près que les objets de la catégorie homotopique stable ne sont pas tous des variétés différentiables.

(5) Pour être correct il faudrait considérer ici des espaces pointés, mais nous négligeons délibérément ce point.

le principe général selon lequel en géométrie algébrique les théories cohomologiques sur la catégorie des variétés lisses sur un corps sont naturellement bigraduées. Ainsi la cohomologie singulière $H^*(X, \mathbf{Z})$ est-elle remplacée par la cohomologie motivique $H^{*,*}(X, \mathbf{Z})$, la K-théorie complexe par la K-théorie algébrique $K^{*,*}(X)$ (avec $K^{n,i}(X) = K_{2i-n}^{\text{Quillen}}(X)$), ..., le premier degré correspondant au degré cohomologique et le second au poids. En particulier, on obtient, pour X lisse, un morphisme

$$\Omega^*(X) \longrightarrow \mathrm{MGL}^{2*,*}(X)$$

dont Levine et Morel conjecturent qu'il est toujours un isomorphisme.

Pour conclure, signalons que Hopkins et Morel ont annoncé, sans imposer d'hypothèse sur k, que pour toute variété lisse (simpliciale) X sur k on a un isomorphisme

$$\mathrm{MGL}^{*,*}(X) \otimes_{\mathbf{L}^*} \mathbf{Z}[\beta, \beta^{-1}] \simeq K^{*,*}(X) \otimes_{\mathbf{Z}} \mathbf{Z}[\beta, \beta^{-1}],$$

β étant maintenant de bidegré $(-2, -1)$. Ils ont également annoncé, sous l'hypothèse que le corps k est de caractéristique zéro, l'existence d'une suite spectrale de type Atiyah-Hirzebruch reliant la cohomologie motivique à $\mathrm{MGL}^{*,*}$.

RÉFÉRENCES

[1] D. Abramovich, K. Karu, K. Matsuki & J. Włodarczyk – « Torification and factorization of birational morphisms », preprint 2000, `AG/9904135`.

[2] J. Adams – *Stable homotopy and generalised homology*, University of Chicago Press, 1974.

[3] L. Bonavero – « Factorisation faible des applications birationnelles », in *Séminaire Bourbaki*, Astérisque, vol. 282, Soc. Math. France, 2002, exp. n° 880, Novembre 2000, p. 1–37.

[4] S. Borghesi – « Algebraic Morava K-theories and the higher degree formula », preprint 2000, disponible à l'adresse `http://www.math.uiuc.edu/K-theory/0412/index.html`.

[5] W. Fulton – *Intersection theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 2, Springer-Verlag, Berlin, 1984.

[6] D. Hoffmann – « Isotropy of quadratic forms over the function field of a quadric », *Math. Z.* **220** (1995), p. 461–476.

[7] O. Izhbolin – « Motivic equivalence of quadratic forms II », *Man. Math.* **102** (2000), p. 41–52.

[8] M. Karoubi – « Cobordisme et groupes formels (d'après D. Quillen et T. tom Dieck) », in *Séminaire Bourbaki, 24ème année (1971/1972)*, Lecture Notes in Math., vol. 317, Springer, Berlin, 1973, exp. n° 408, p. 141–165.

[9] M. Lazard – « Sur les groupes de Lie formels à un paramètre », *Bull. Soc. Math. France* **83** (1955), p. 251–274.

[10] M. Levine – « Algebraic cobordism II », en préparation.

[11] M. Levine & F. Morel – « Algebraic cobordism I », preprint 2002, disponible à l'adresse `http://www.math.uiuc.edu/K-theory/0547/index.html`.

[12] ———, « Cobordisme algébrique I », *C. R. Acad. Sci. Paris Sér. I Math.* **332** (2001), p. 723–728.

[13] ———, « Cobordisme algébrique II », *C. R. Acad. Sci. Paris Sér. I Math.* **332** (2001), p. 815–820.

[14] A. Merkurjev – « Algebraic oriented cohomology theories », preprint 2002, disponible à l'adresse `http://www.math.uiuc.edu/K-theory/0535/index.html`.

[15] ———, « Degree Formula », preprint 2000, disponible à l'adresse `http://www.math.ohio-state.edu/~rost/chain-lemma.html`.

[16] ———, « Rost's degree formula (Notes of mini-course in Lens, June 2001) », preprint 2001, disponible à l'adresse `http://www.math.ucla.edu/~merkurev/publicat.htm`.

[17] J. Milnor – « On the cobordism ring Ω^* and a complex analogue. I », *Amer. J. Math.* **82** (1960), p. 505–521.

[18] F. Morel – « Voevodsky's proof of Milnor's conjecture », *Bull. Amer. Math. Soc.* **35** (1998), p. 123–143.

[19] F. Morel & V. Voevodsky – « $\mathbf{A}^1$-homotopy theory of schemes », *Inst. Hautes Études Sci. Publ. Math.* **90** (1999), p. 45–143.

[20] D. Quillen – « Elementary proofs of some results of cobordism theory using Steenrod operations », *Advances in Math.* **7** (1971), p. 29–56.

[21] R. Thom – « Quelques propriétés globales des variétés différentiables », *Comment. Math. Helv.* **28** (1954), p. 17–86.

[22] B. Totaro – « Torsion algebraic cycles and complex cobordism », *J. Amer. Math. Soc.* **10** (1997), p. 467–493.

[23] V. Voevodsky – « The Milnor Conjecture », preprint 1996, disponible à l'adresse `http://www.math.uiuc.edu/K-theory/0170/index.html`.

François LOESER

École normale supérieure
Département de mathématiques et applications
UMR 8553 du CNRS
45 rue d'Ulm
75230 Paris Cedex 05
France
E-mail : `Francois.Loeser@ens.fr`
URL : `http://www.dma.ens.fr/~loeser/`

Séminaire BOURBAKI Mars 2002
54e année, 2001-2002, no 902, p. 193 à 206

LA CONJECTURE DE KATO
[d'après Pascal Auscher, Steve Hofmann, Michael Lacey, John Lewis, Alan McIntosh et Philippe Tchamitchian]

par **Yves MEYER**

1. INTRODUCTION

La conjecture de Kato concerne les racines carrées d'opérateurs accrétifs. Cette conjecture est magique au sens suivant : son énoncé est innocent et ne fait appel qu'à des notions très naturelles concernant les opérateurs agissant sur un espace de Hilbert. Il s'agit de démontrer que le domaine de la racine carrée d'un opérateur accrétif coïncide avec celui de la forme ayant servi à construire l'opérateur. Mais la conjecture de Kato implique d'autres conjectures importantes. Par exemple, un cas particulier, en dimension 1, de la conjecture de Kato entraîne une conjecture de Calderón portant sur la continuité du noyau de Cauchy pour les graphes lipschitziens. Quand Alan McIntosh observa ce fait en 1980, il traça du même coup le chemin conduisant à la preuve de la conjecture de Calderón. Il ne restait plus qu'à terminer la démonstration en ramenant le problème à une estimation portant sur une fonctionnelle quadratique et en utilisant alors la technique des « mesures de Carleson » (section 6). Les mesures de Carleson ont été introduites par Lennart Carleson pour résoudre un problème portant sur les idéaux de fonctions holomorphes et bornées dans le disque unité. Leur usage en théorie des opérateurs a été systématisé par R. Coifman et ses collaborateurs. Ces mêmes estimations quadratiques et ces mêmes mesures de Carleson joueront un rôle essentiel dans la résolution de la conjecture de Kato.

2. LA PREMIÈRE VERSION DE LA CONJECTURE DE KATO

Commençons par une définition introduite par Ralph Phillips en 1957. La partie réelle du nombre complexe z est notée $\mathrm{Re}\, z$.

DÉFINITION 2.1. — *Soit H un espace de Hilbert sur $\mathbb{C}$ et désignons par $\langle x, y\rangle$ et $\|x\|$ la forme sesquilinéaire et la norme correspondantes. Soit V un sous-espace vectoriel de H et $T : V \to H$ un opérateur linéaire. On dira que T est dissipatif si*

$$\mathrm{Re}\langle T(x), x\rangle \leqslant 0, \quad x \in V. \tag{2.1}$$

L'opérateur T est dissipatif maximal s'il est dissipatif et s'il ne peut être prolongé en un opérateur dissipatif $\widetilde{T} : \widetilde{V} \to H$ où $\widetilde{V}$ contienne strictement V. Nous dirons alors que V est le domaine de T.

Rappelons également la définition d'un semi-groupe de contractions.

DÉFINITION 2.2. — *Un semi-groupe de contractions est une famille $S(t)$, $t \geqslant 0$, d'opérateurs linéaires continus sur l'espace de Hilbert H vérifiant, pour tout $x \in H$,*

$$S(t+\tau) = S(t)S(\tau), \quad t, \tau \geqslant 0 \tag{2.2}$$

$$\|S(t)(x) - x\| \to 0, \quad t \to 0 \tag{2.3}$$

$$\|S(t)(x)\| \leqslant \|x\|, \quad t \geqslant 0. \tag{2.4}$$

Avec ces notations, le théorème de Phillips est l'énoncé suivant

THÉORÈME 2.3. — *Un opérateur T est le générateur infinitésimal d'un semi-groupe de contractions $S(t) : H \to H$ si et seulement si T est dissipatif maximal et si son domaine V est dense dans H.*

Selon une suggestion de Friedrichs, un opérateur T est appelé accrétif si $-T$ est dissipatif.

Voici un exemple des situations que nous venons d'étudier. L'opérateur $\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$ est dissipatif quand $H = L^2(\mathbb{R}^n)$ et le semi-groupe associé $\exp(t\Delta)$ est le semi-groupe de la chaleur. Notre propos sera de définir la racine carrée d'un opérateur accrétif maximal. Dans le cas particulier du laplacien, cette racine carrée est l'opérateur de Calderón $\Lambda = \sqrt{-\Delta}$. Son domaine est celui de la forme de Dirichlet $B(u, v) = \int \nabla u \cdot \overline{\nabla v}\, dx$. Le semi-groupe $S(t) = \exp(-t\Lambda)$, $t \geqslant 0$, décrit l'évolution $u(x, 0) \to u(x, t)$ des fonctions harmoniques $u(x, t)$ dans $\Omega = \{(x, t) \in \mathbb{R}^n \times (0, \infty)\}$ qui sont nulles à l'infini.

Si la construction d'opérateurs accrétifs est aisée, celle d'opérateurs accrétifs maximaux est un peu plus délicate. On commence par la définition d'une forme sesquilinéaire accrétive.

DÉFINITION 2.4. — *Soit H un espace de Hilbert pour lequel le produit scalaire et la norme seront notés $\langle x, y\rangle$ et $\|x\|$. Soit H_1 un second espace de Hilbert. On désigne par $\|x\|_1$ la norme correspondante et l'on suppose que H_1 est un sous-espace vectoriel dense dans H, cette injection étant continue.*
Soit $B(x, y)$ une forme sesquilinéaire définie et bicontinue sur $H_1 \times H_1$:

$$|B(x, y)| \leqslant C\|x\|_1\|y\|_1. \tag{2.5}$$

Nous dirons que B est accrétive si, pour tout x dans H_1, on a

$$\operatorname{Re} B(x,x) \geqslant 0. \tag{2.6}$$

Si β est un nombre réel positif, nous dirons que $B(x,y)$ est β-accrétive si, pour tout $x \in H_1$, on a

$$\operatorname{Re} B(x,x) + \|x\|^2 \geqslant \beta \|x\|_1^2. \tag{2.7}$$

Dans ce cas, nous dirons que H_1 est le domaine de la forme $B(x,y)$.

Ceci étant, voici comment on construit des opérateurs accrétifs maximaux.

LEMME 2.5. — *Soit $B : H_1 \times H_1 \to \mathbb{C}$ une forme sesquilinéaire β-accrétive. On désigne par V l'ensemble des $x \in H$ pour lesquels il existe une constante $C = C(x)$ de sorte que, pour tout $y \in H$, on ait*

$$|B(x,y)| \leqslant C\|y\|. \tag{2.8}$$

Alors l'opérateur $T : V \to H$ défini par

$$\langle T(x), y\rangle = B(x,y), \quad x \in V,\ y \in H \tag{2.9}$$

est accrétif maximal.

En restant toujours dans un cadre hilbertien général, voici la définition, puis la construction de la racine carrée accrétive d'un opérateur accrétif maximal.

DÉFINITION 2.6. — *Soit H un espace de Hilbert et $T : V \to H$ un opérateur accrétif maximal de domaine $V \subset H$. Alors il existe un et un seul opérateur accrétif maximal S tel que $S^2 = T$. On écrit $S = \sqrt{T}$. En outre, sur le domaine V de T, on a*

$$S = \frac{2}{\pi}\int_0^\infty (1+\lambda^2 T)^{-1} T\, d\lambda = \frac{8}{\pi}\int_0^\infty (1+\lambda^2 T)^{-3}\lambda^2 T^2\, d\lambda. \tag{2.10}$$

Le problème posé par Kato dans [7] est la question suivante :

Supposons que l'opérateur maximal accrétif T soit défini à l'aide d'une forme β-accrétive B au sens du lemme 2.5. Le domaine de la racine carrée accrétive maximale $\sqrt{T}$ est-il alors le même que le domaine H_1 de la forme qui a servi à construire T ?

Ceci est évidemment vrai si l'opérateur T est auto-adjoint, c'est-à-dire si $B(u,u)\geqslant 0$ pour tout $u \in V$.

J.-L. Lions a démontré dans [7] que la conjecture de Kato est équivalente à l'énoncé suivant : il existe une constante C telle que, pour tout $x \in H_1$, on ait, en désignant par T^* l'adjoint de T

$$\|\sqrt{T}x\| \leqslant C\|x\|_1, \quad \|\sqrt{T^*}x\| \leqslant C\|x\|_1, \quad x \in H_1. \tag{2.11}$$

Alan McIntosh a trouvé un contre-exemple à cette version hilbertienne de la conjecture de Kato. Il restait à savoir si la conjecture de Kato est exacte dans le cas particulier qui a motivé le travail de Tosio Kato, celui où T est un opérateur différentiel accrétif du second ordre, écrit sous forme de divergence.

3. LA CONJECTURE DE KATO PRÉCISÉE

On pose $H = L^2(\mathbb{R}^n)$ et $H_1 \subset H$ désignera l'espace de Sobolev $H^1(\mathbb{R}^n)$ qui se compose des fonctions de carré intégrable dont le gradient, pris au sens des distributions, est aussi de carré intégrable. Ensuite on pose $\partial_j = \partial/\partial x_j$ et l'on considère n^2 fonctions $a_{j,k}(x)$, $1 \leqslant j,k \leqslant n$, à valeurs réelles ou complexes, mesurables et bornées (appartenant à $L^\infty(\mathbb{R}^n)$). On suppose qu'il existe deux constantes $C \geqslant \beta > 0$ telles que l'on ait, pour tout $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{C}^n$ et pour presque tout $x \in \mathbb{R}^n$,

$$\beta|\xi|^2 \leqslant \mathrm{Re}\Big(\sum_{1\leqslant j,k\leqslant n} a_{j,k}(x)\xi_k\overline{\xi_j}\Big) \leqslant C|\xi|^2. \tag{3.1}$$

On désigne par $\Omega(\beta, C)$ l'ensemble des matrices $A(x) = (a_{j,k}(x))_{1\leqslant j,k\leqslant n}$ qui vérifient (3.1) et $\|a_{j,k}\|_\infty \leqslant C$, $1 \leqslant j,k \leqslant n$. Ensuite Ω est la réunion des $\Omega(\beta, C)$, $0 < \beta \leqslant C < \infty$. Alors Ω est un ensemble ouvert dans $(L^\infty(\mathbb{R}^n))^{n^2}$ de matrices $n \times n$.

Lemme 3.1. — *Si* (3.1) *est vérifiée,*

$$B(f,g) = \sum_{1\leqslant j,k\leqslant n} \int a_{j,k}(x)\partial_k f(x)\overline{\partial_j g(x)}\, dx \tag{3.2}$$

est une forme sesquilinéaire $\min(1,\beta)$*-accrétive dont le domaine est l'espace de Sobolev* $H^1(\mathbb{R}^n)$.

L'opérateur accrétif maximal défini par cette forme est alors

$$T(f) = -\sum_{1\leqslant j,k\leqslant n} \partial_j\big(a_{j,k}(x)\partial_k f(x)\big) \tag{3.3}$$

et son domaine V est le sous-espace vectoriel de $H^1(\mathbb{R}^n)$ caractérisé par la condition suivante : $f \in V$ si et seulement si les fonctions $g_j \in L^2(\mathbb{R}^n)$ définies par $g_j(x) = \sum_{1\leqslant k\leqslant n} a_{j,k}(x)\partial_k f(x)$ vérifient $\sum_{1\leqslant j\leqslant n} \partial_j g_j \in L^2(\mathbb{R}^n)$. Ce domaine dépend non linéairement des fonctions mesurables $a_{j,k}$ et n'est pas un espace fonctionnel classique.

Pascal Auscher et ses collaborateurs (S. Hofmann, M. Lacey, J. Lewis, A. McIntosh et Ph. Tchamitchian) ont démontré le résultat suivant :

THÉORÈME 3.2. — *Soit T l'opérateur maximal accrétif défini par* (3.3) *lorsque la matrice $A(x)$ vérifie les conditions* (3.1)*. Alors le domaine de l'opérateur maximal accrétif $\sqrt{T}$ est l'espace de Sobolev $H^1(\mathbb{R}^n)$ et l'application $\Phi : A \to \sqrt{T}$ est holomorphe sur l'ouvert Ω.*

La seconde assertion du théorème 3.2 conduit à étudier $\sqrt{T}$ en utilisant son développement en série entière. Lorsque ce théorème était encore une conjecture, plusieurs équipes concurrentes établissaient la continuité des opérateurs multilinéaires $T_k(A, A, \ldots, A)(f)$, $k \in \mathbb{N}$, apparaissant dans cette série entière, mais nous (R. Coifman, G. David, J-L. Journé, C. Kenig, Y.M., etc.) n'étions capables de démontrer la convergence de cette série que si la matrice $A(x)$ est suffisamment proche de l'identité : $\|A(x) - I\|_\infty \leqslant \varepsilon_n$ et le résultat le plus précis dans cette direction (fournissant la valeur la moins faible de la constante ε_n) avait été obtenu par J-L. Journé [6].

Si les fonctions $a_{j,k}$ sont suffisamment régulières, $\sqrt{T}$ est alors un opérateur pseudo-différentiel classique d'ordre 1 et, à ce titre, est continu sur l'espace de Sobolev H^1. La même remarque s'applique à T^* et le théorème de J-L. Lions implique alors la conjecture de Kato.

Dans le cas général où les fonctions $a_{j,k}(x)$ sont mesurables et bornées, on ne peut plus utiliser les ressources du calcul pseudo-différentiel. Les méthodes hilbertiennes abstraites sont également insuffisantes. Des outils plus puissants sont nécessaires. Ces outils ont été forgés par Alberto Calderón et « l'École de Chicago », mais il faut aussi mentionner les travaux de De Giorgi, Moser, Carleson, etc. Le programme de Calderón a consisté à s'affranchir des hypothèses superflues de régularité qui limitent la portée du calcul pseudo-différentiel classique et à construire des algèbres incluant à la fois les opérateurs de multiplication ponctuelle par des fonctions mesurables et bornées, comme les $a_{j,k}(x)$, et les opérateurs de dérivation, comme les ∂_j.

Après avoir attaqué la conjecture de Kato dans la perspective donnée par le programme de Calderón (voir le théorème 5.3), Pascal Auscher et Philippe Tchamitchian ont finalement décidé d'aborder cette conjecture par une voie un peu différente, en s'inspirant des travaux de De Giorgi et de Moser que nous rappelons maintenant.

4. ENNIO DE GIORGI ET JÜRGEN MOSER

En 1956, E. De Giorgi étudia la régularité des solutions faibles u de l'équation aux dérivées partielles $T(u) = 0$ où T est défini par (3.3) et où les coefficients $a_{j,k}(x)$ sont réels, mesurables et bornés. Il démontra que toute solution faible d'une telle équation aux dérivées partielles est höldérienne d'exposant $\mu > 0$ et que μ ne dépend que des normes L^∞ des coefficients, de la constante d'accrétivité et de la dimension.

Plus précisément on a, en posant $B_R = \{x; |x - x_0| < R\}$ et $B_\rho = \{x; |x - x_0| < \rho\}$

THÉORÈME 4.1. — *Soit T l'opérateur accrétif maximal défini par* (3.3). *Supposons que les fonctions $a_{j,k}(x)$, $1 \leqslant j,k \leqslant n$ soient à valeurs réelles. Alors il existe un exposant positif μ et une constante positive C tels que, pour tout $\rho \in]0,R[$ et pour toute solution $u \in H^1(B_R)$ de $T(u)=0$, on ait*

$$\int_{B_\rho} |\nabla u|^2\,dx \leqslant C\big(\frac{\rho}{R}\big)^{n-2+2\mu} \int_{B_R} |\nabla u|^2\,dx. \tag{4.1}$$

En outre μ et C ne dépendent que des normes L^∞ des fonctions $a_{j,k}(x)$, de la constante β d'ellipticité et de la dimension n.

La propriété (4.1) implique la régularité höldérienne. Si les fonctions $a_{j,k}(x)$ sont à valeurs complexes, le théorème 4.1 cesse d'être vrai, comme l'ont établi Maz'ya, Nazarov et Plamenevskii [9] en dimension $n \geqslant 5$. Nous dirons que l'opérateur T a la propriété de Dirichlet si (4.1) est vérifiée.

Une autre définition utile concerne le « noyau de la chaleur » $K_t(x,y)$ défini par

$$\exp(-tT)[f(x)] = \int K_t(x,y) f(y)\,dy. \tag{4.2}$$

On peut penser que sous l'hypothèse (3.1) le noyau de la chaleur $K_t(x,y)$ possède des propriétés analogues à celles du noyau de la chaleur usuel (correspondant à $A(x)=I$). Cette remarque conduit à la définition suivante

DÉFINITION 4.2. — *L'opérateur accrétif maximal T est de type gaussien s'il existe un exposant $\mu \in (0,1]$ et deux constantes positives $\beta>0, c>0$ tels que l'on ait d'une part*

$$|K_t(x,y)| \leqslant ct^{-n/2} \exp\Big(-\beta\frac{|x-y|^2}{t}\Big) \tag{4.3}$$

$$|K_t(x,y) - K_t(x+h,y)| \leqslant c\Big(\frac{|h|}{t^{1/2}+|x-y|}\Big)^{\mu} t^{-n/2} \exp\Big(-\beta\frac{|x-y|^2}{t}\Big) \tag{4.4}$$

pour $t>0, x,y,h \in \mathbb{R}^n$ tels que $|h| \leqslant t^{1/2}+|x-y|$. On impose, d'autre part, la propriété (4.4) *en échangeant les rôles de x et y.*

J. Nash, D. Aronson, E. Fabes et D. Stroock ont établi le théorème suivant

THÉORÈME 4.3. — *Si les n^2 fonctions $a_{j,k}(x)$ sont toutes à valeurs réelles, alors l'opérateur accrétif maximal T défini par* (3.3) *est du type gaussien.*

Ce théorème a été complété par un résultat remarquable dû à P. Auscher. En effet on a

THÉORÈME 4.4. — *Revenons au cas où les $a_{j,k}(x)$ sont à valeurs réelles ou complexes. Alors T défini par* (3.3) *est de type gaussien si et seulement si T et T^* ont la propriété de Dirichlet.*

Ceci a toujours lieu si $n = 1$ et $n = 2$, mais il existe des contre-exemples en dimension $n \geqslant 5$ [9].

Venons-en aux contributions de Jürgen Moser. On considère un opérateur accrétif maximal T défini par (3.3) et l'on suppose encore que les fonctions $a_{j,k}(x)$ soient réelles. On a alors

$$c_0|\xi|^2 \leqslant \sum_{j,k} a_{j,k}(x)\xi_j\xi_k \leqslant \frac{1}{c_0}|\xi|^2 \tag{4.5}$$

uniformément en x. En 1961, J. Moser utilisa les ressources de l'espace BMO qui venait d'être créé pour démontrer l'inégalité de Harnack qui s'énonce ainsi

THÉORÈME 4.5. — *Soit Ω un ouvert de $\mathbb{R}^n$ et $K \subset \Omega$ un ensemble compact. Il existe alors une constante $C = C(c_0, K, \Omega)$ telle que, pour toute fonction positive $f \in H^1(\Omega)$ vérifiant $T(f) = 0$ au sens des distributions, ont ait*

$$\sup\{f(x); x \in K\} \leqslant C \inf\{f(x); x \in K\}. \tag{4.6}$$

5. ALBERTO CALDERÓN

Intrigué par le théorème de De Giorgi, A. Calderón se proposait de le redémontrer en construisant de nouvelles algèbres d'opérateurs incluant les opérateurs accrétifs donnés par (3.3). Calderón s'est d'abord posé le problème d'établir l'estimation L^2 de base en utilisant des conditions minimales de régularité sur noyau-distribution $K(x, y)$ d'un opérateur T. On veut savoir si, pour toute fonction $f \in L^2(\mathbb{R}^n)$, on a $\|T(f)\|_2 \leqslant C\|f\|_2$ quand $T(f)(x) = \text{v.p.} \int K(x, y)f(y)\,dy$. Lars Hörmander a proposé la condition suivante

$$\int_{\{|x-y| \geqslant 2|x'-x|\}} |K(x', y) - K(x, y)|\,dy \leqslant C. \tag{5.1}$$

Cette propriété, combinée à l'estimation $L^2(\mathbb{R}^n)$ qui s'écrit

$$\|T(f)\|_2 \leqslant C\|f\|_2 \tag{5.2}$$

entraîne les estimations $L^p(\mathbb{R}^n)$ pour $2 \leqslant p < \infty$, ainsi que l'estimation limite (L^∞, BMO) où, là encore, BMO est l'espace de John et Nirenberg. C'est-à-dire que $f \in BMO$ si et seulement s'il existe une constante C telle que l'on ait

$$\sup_Q \Big\{ \int_Q |f(x) - m_Q(f)|^2 \frac{dx}{|Q|} \Big\} \leqslant C. \tag{5.3}$$

On a désigné par $|Q|$ le volume du cube Q, la borne supérieure est calculée sur tous les cubes $Q \subset \mathbb{R}^n$ et $m_Q(f)$ désigne la moyenne de f sur Q.

DÉFINITION 5.1. — *Si à la fois T et son adjoint T^* vérifient* (5.1) *et* (5.2), *nous dirons que T est un opérateur de Calderón-Zygmund.*

Les opérateurs pseudo-différentiels classiques d'ordre 0 sont des opérateurs de Calderón-Zygmund, mais la réciproque n'est pas vraie. Les opérateurs de Calderón-Zygmund jouent un rôle fondamental dans la preuve, par Guy David, de la conjecture de Vitushkin (caractérisation des ensembles compacts du plan complexe de capacité analytique nulle et de mesure de Hausdorff unidimensionnelle finie comme étant les ensembles totalement non rectifiables).

Le problème essentiel dans l'étude des opérateurs de Calderón-Zygmund est celui de la continuité de l'opérateur T sur $L^2(\mathbb{R}^n)$. Ce programme de travail a débouché sur le « théorème $T(1)$ » de David et Journé que nous énonçons sous une forme simplifiée

THÉORÈME 5.2. — *Supposons que l'on ait*

$$K(x,y) = -K(y,x), \quad |K(x,y)| \leqslant |x-y|^{-n}, \quad |\nabla K(x,y)| \leqslant |x-y|^{-(n+1)}. \tag{5.4}$$

Alors une condition nécessaire et suffisante de continuité de T sur $L^2(\mathbb{R}^n)$ est $T(1) \in BMO$.

Les conditions imposées à $K(x,y)$ peuvent être considérablement allégées ; une régularité höldérienne suffit. Mais l'on ne sait pas démontrer le « théorème $T(1)$ » sous la condition (5.1). L'action de T sur la fonction identiquement égale à 1 doit être définie en donnant un sens à l'intégrale $g(x) = \int K(x,y)dy$. Cette fonction $g(x)$ est définie, modulo une fonction constante. En principe le théorème $T(1)$ pourraît s'appliquer à notre propos, grâce à la relation entre la conjecture de Kato et les opérateurs de Calderón-Zygmund. On a, en effet

THÉORÈME 5.3. — *En conservant les notations du théorème* 4.4, *supposons que T soit de type gaussien et que la conjecture de Kato soit vraie, c'est-à-dire que* (2.11) *soit vérifiée lorsque H_1 est l'espace de Sobolev $H^1(\mathbb{R}^n)$. Alors il existe n opérateurs de Calderón-Zygmund $U_1, \ldots, U_n$, tels que l'on ait*

$$\sqrt{T} = U_1\partial_1 + \cdots + U_n\partial_n. \tag{5.5}$$

Ce théorème a été établi par Auscher et Tchamitchian avant que la preuve complète de la conjecture de Kato ne soit obtenue. Ce théorème nous incite à démontrer la conjecture de Kato en utilisant le « théorème $T(1)$ » de David et Journé. La preuve générale doit beaucoup à cette remarque. Voici cependant des exemples où le théorème $T(1)$ ne permet pas de conclure. C'est, par exemple, le cas si T est l'opérateur de Cauchy sur une courbe lipschitzienne Γ qui est le graphe d'une fonction lipschitzienne $a(x)$. Dans ce cas, l'opérateur T est défini par le noyau singulier vp $\frac{1}{\pi}(z(x)-z(y)^{-1}$ où $z(x) = x + ia(x)$. Le calcul de $T(1)$ ne peut se faire et c'est pourquoi un nouveau théorème a été recherché, puis démontré. Il s'agit du « théorème $T(b)$ » dont on trouvera la preuve dans [10]. On part d'une fonction $b(x)$, appartenant à L^∞, qui est accrétive au sens où sa partie réelle $\operatorname{Re} b(x)$ vérifie $\operatorname{Re} b(x) \geqslant \beta > 0$. Dans ces conditions, le théorème $T(1)$ se généralise et l'on a

THÉORÈME 5.4. — *En conservant les hypothèses du théorème* 5.2*, supposons qu'il existe une fonction accrétive* $b(x) \in L^\infty(\mathbb{R}^n)$ *telle que* $T(b)$ *appartienne à BMO. Alors* T *est borné sur* $L^2(\mathbb{R}^n)$*.*

En revenant à l'opérateur de Cauchy, il suffit, pour appliquer le théorème 5.4, de choisir $b(x) = 1 + ia'(x)$. En effet, cette fonction satisfait évidemment la condition d'accrétivité et l'on a alors $T(b) = 0$, comme le montre immédiatement la formule de Cauchy. La force du « théorème $T(b)$ » réside dans la possibilité de choisir la fonction $b(x)$ « après coup », en prenant en compte les propriétés de l'opérateur que l'on étudie.

Une autre version du théorème $T(b)$ peut être trouvée dans un remarquable travail de M. Christ et J.L. Journé [4].

Comme nous le verrons dans la huitième section, la preuve de la conjecture de Kato dépend d'une troisième version (adaptée) du « théorème $T(b)$ ». Cette nouvelle version, qui étend celle de M. Christ et J.L. Journé, s'applique à des situations beaucoup plus générales que celles décrites dans le théorème 5.4. Il n'est plus nécessaire de supposer que T soit de type gaussien, ce qui, comme nous l'avons vu, n'est pas toujours vérifié.

6. LES MESURES DE CARLESON

La définition des mesures de Carleson apparaît pour la première fois dans la preuve par Lennart Carleson de la « conjecture de la couronne ». Il s'agit de l'énoncé suivant : si $f_1(z), \cdots, f_n(z)$ sont n fonctions holomorphes et bornées dans le disque unité D et s'il existe une constante C telle que l'on ait, pour tout $z \in D$,

$$1 \leqslant \sum_{1 \leqslant k \leqslant n} |f_k(z)| \leqslant C \tag{6.1}$$

alors on peut trouver n autres fonctions holomorphes et bornées dans D, notées $g_1(z), \cdots, g_n(z)$, telles que

$$\sum_{1 \leqslant k \leqslant n} f_k(z) g_k(z) = 1. \tag{6.2}$$

Pour démontrer ce théorème, Carleson définit ce qui est aujourd'hui connu sous le nom de mesures de Carleson et démontra le théorème 6.2 qui suit.

DÉFINITION 6.1. — *Soit* $\Omega \subset \mathbb{R}^n$ *un ensemble ouvert. On désigne alors par* $\widehat{\Omega}$ *l'ensemble des couples* $(x,t) \in \mathbb{R}^n \times]0,\infty[$ *tels que la boule* $B(x,t)$ *de centre* x *et de rayon* t *soit incluse dans* Ω*. Soit* $d\mu(x,t)$ *une mesure de Radon positive sur* $\mathbb{R}^n \times]0,\infty[$*. On dit que* $d\mu$ *est une mesure de Carleson s'il existe une constante* C *telle que pour toute partie ouverte* $\Omega \subset \mathbb{R}^n$*, on ait*

$$\mu(\widehat{\Omega}) \leqslant C|\Omega|, \tag{6.3}$$

où $|E|$ *est la mesure de Lebesgue de* $E \subset \mathbb{R}^n$*.*

La plus petite constante C pouvant figurer dans (6.3) est la norme de la mesure de Carleson μ et sera notée $\|\mu\|_{\mathcal{C}}$. Pour établir (6.3), il suffit de le faire si Ω est une boule arbitraire de $\mathbb{R}^n$. Le passage au cas général utilise un recouvrement de Vitali de Ω par une suite de boules B_j, $j \in \mathbb{N}$. Ce faisant, la constante figurant dans le second membre de (6.3) aura changé.

Pour vérifier qu'une mesure de Radon positive $d\mu(x,t)$ est une mesure de Carleson, on utilise le plus souvent une variante de la condition (6.3). Si Q est une cube arbitraire de $\mathbb{R}^n$ de côté l_Q, on désignera par $\widetilde{Q}$ le cube $Q \times [0, l_Q]$ de $\mathbb{R}^{n+1}$ et la condition de Carleson s'écrit simplement $\mu(\widetilde{Q}) \leqslant C|Q|$.

Nous arrivons maintenant au théorème annoncé. Il concerne le calcul des intégrales $I = \int_0^\infty \int_{\mathbb{R}^n} f(x,t) d\mu(x,t)$ où $f(x,t)$ est une fonction borélienne positive ou nulle et où $d\mu$ est une mesure de Carleson de norme $\|\mu\|_{\mathcal{C}}$.

THÉORÈME 6.2. — *Posons*

$$f^*(x) = \sup_{|y-x| \leqslant t} f(y,t). \tag{6.4}$$

Alors on a

$$\int_0^\infty \int_{\mathbb{R}^n} f(x,t) d\mu(x,t) \leqslant \|\mu\|_{\mathcal{C}} \int_{\mathbb{R}^n} f^*(x)\, dx. \tag{6.5}$$

7. LA PREUVE DE LA CONJECTURE DE KATO (PREMIÈRE PARTIE)

La preuve de la conjecture de Kato utilise deux idées maintenant classiques en théorie des opérateurs. La première qui a été développée par Elias Stein, mais qui peut être rattachée aux travaux de Littlewood et Paley, concerne l'utilisation de fonctionnelles quadratiques. Supposons qu'un opérateur T soit naturellement décomposé en une série $\sum_{j \in J} T_j$. Pour démontrer que cet opérateur T est borné sur $L^2(\mathbb{R}^n)$, c'est-à-dire pour établir l'estimation fondamentale $\|T(f)\|_2 \leqslant C\|f\|_2$, il suffit, dans certains cas, de démontrer que

$$\Big(\sum_{j \in J} \|T_j(f)\|_2^2 \Big)^{1/2} \leqslant C\|f\|_2. \tag{7.1}$$

On dira alors que les morceaux T_j, $j \in J$, sont presque orthogonaux.

Le second ingrédient est l'utilisation des mesures de Carleson pour établir l'estimation quadratique (7.1). La preuve du théorème $T(1)$ de David et Journé suit ce programme.

Commençons par la réduction à une fonctionnelle quadratique. On part de T défini par (3.3) et l'on considère sa racine carrée accrétive, définie par (2.10). Nous devons démontrer que l'on a

$$\|\sqrt{T}(f)\|_2 \leqslant C\|\nabla f\|_2 \tag{7.2}$$

et cela nous amène à calculer

$$\sup\{|\langle\sqrt{T}(f), g\rangle|;\ \|g\|_2 \leqslant 1\}. \tag{7.3}$$

Mais ce produit scalaire s'écrit, grâce à (2.10),

$$\langle\sqrt{T}(f), g\rangle = \frac{8}{\pi}\int_0^\infty \langle W_\lambda(f), V_\lambda(g)\rangle \frac{d\lambda}{\lambda}, \tag{7.4}$$

où

$$W_\lambda = (1+\lambda^2 T)^{-1}\lambda T, \quad V_\lambda = (1+\lambda^2 T^*)^{-2}\lambda^2 T^*. \tag{7.5}$$

Arrivés à ce point, on applique les estimations quadratiques « abstraites » de McIntosh et Yagi ([3] est la meilleure référence) qui fournissent

$$\int_0^\infty \|V_\lambda(g)\|_2^2 \frac{d\lambda}{\lambda} \leqslant C\|g\|_2^2. \tag{7.6}$$

La conjecture de Kato résultera donc de l'estimation quadratique

$$\int_0^\infty \|W_\lambda(f))\|_2^2 \frac{d\lambda}{\lambda} \leqslant C\|\nabla f\|_2^2. \tag{7.7}$$

L'opérateur Θ_λ est défini sur les n-uples de fonctions de $L^2(\mathbb{R}^n)$ par

$$\Theta_\lambda(f_1, \ldots, f_n)(x) = \lambda(1+\lambda^2 T)^{-1}\sum_{j,k}\partial_j[(a_{j,k}(x))(f_k(x))]. \tag{7.8}$$

Alors (7.7) s'écrit

$$\int_0^\infty \|\Theta_\lambda \nabla f\|_2^2 \frac{d\lambda}{\lambda} \leqslant C\|\nabla f\|_2^2. \tag{7.9}$$

Nous désignons ensuite par $\varphi(x)$ une fonction de la classe de Schwartz, portée par la boule unité, d'intégrale égale à 1 et par Φ_λ l'opérateur de convolution avec $\varphi_\lambda(x) = \lambda^{-n}\varphi(x/\lambda)$. On pose ensuite

$$\gamma_\lambda(x) = [(1+\lambda^2 T)^{-1}\lambda\partial_j a_{j,k}(x)]_{1\leqslant k\leqslant n}. \tag{7.10}$$

En utilisant des estimations sur l'action « à longue portée » de l'opérateur $(I+\lambda^2 T)^{-1}$, on obtient

$$\Theta_\lambda(\nabla f)(x) = \gamma_\lambda(x)\cdot\Phi_\lambda\nabla f(x) + \rho(x,\lambda), \tag{7.11}$$

où

$$\int_0^\infty\int_{\mathbb{R}^n} |\rho(x,\lambda)|^2\, dx\frac{d\lambda}{\lambda} \leqslant C\|\nabla f\|_2^2. \tag{7.12}$$

La conjecture de Kato découlera alors de

$$\int_0^\infty \int_{\mathbb{R}^n} |\gamma_\lambda(x)|^2 |\Phi_\lambda \nabla f(x)|^2 \, dx \frac{d\lambda}{\lambda} \leqslant C \|\nabla f\|_2^2. \tag{7.13}$$

Le théorème de Hardy et Littlewood sur la fonction maximale implique

$$\| \sup_{|y-x| \leqslant \lambda} |\Phi_\lambda \nabla f(y)| \|_2 \leqslant C \|\nabla f\|_2. \tag{7.14}$$

La conjecture de Kato sera donc établie si nous montrons que la mesure $d\mu(x,\lambda) = |\gamma_\lambda|^2 \, dx \frac{d\lambda}{\lambda}$ est une mesure de Carleson. La preuve de ce fait est classique. Partant d'une boule arbitraire B de $\mathbb{R}^n$, on découpe la fonction mesurable et bornée $a_{j,k}(x)$ en deux morceaux. Le premier (noté u) est porté par la boule double $\widetilde{B}$, de même centre que B et de rayon double et le second (noté v) est nul sur $\widetilde{B}$. Pour traiter la contribution de u dans la mesure $d\mu$, on observe que $\|u\|_2 \leqslant C \|a_{j,k}\|_\infty |B|^{1/2}$ et l'estimation recherchée proviendra de

$$\int_0^\infty \|(1+\lambda^2 T)^{-1} \lambda \nabla f\|_2^2 \frac{d\lambda}{\lambda} \leqslant C \|f\|_2^2. \tag{7.15}$$

Mais cette dernière estimation n'est autre que l'estimation (7.7) que nous nous proposions de démontrer. Nous sommes donc dans une situation de cercle vicieux !

8. UN THÉORÈME T(b) POUR LES MESURES DE CARLESON

Le paragraphe précédent nous a appris que la preuve de l'estimation (H^1, L^2) pour l'opérateur $\sqrt{T}$ repose sur le fait que la mesure $d\mu(x,\lambda) = |\gamma_\lambda|^2 \, dx \frac{d\lambda}{\lambda}$ est une mesure de Carleson, ce qui semble essentiellement dépendre de l'estimation cherchée. P. Auscher et Ph. Tchamitchian ont réussi à sortir de cette impasse en démontrant un « théorème $T(b)$ pour les mesures de Carleson ». La fonction accrétive b est ici remplacée par une collection F_Q de fonctions de test adaptées à l'opérateur T. Soyons plus précis. Nous désignons par $\mathcal{Q}$ l'ensemble de tous les cubes $Q \subset \mathbb{R}^n$ et, pour tout cube $Q \in \mathcal{Q}$, de côté l_Q, nous supposons avoir pu construire une fonction F_Q ayant les cinq propriétés suivantes

$$F_Q \in H^1(5Q) \tag{8.1}$$

$$\int_{5Q} |\nabla F_Q|^2 \, dx \leqslant C|Q| \tag{8.2}$$

$$T(F_Q) = -\mathrm{div}[A(x) \nabla F_Q] \in L^2(5Q) \tag{8.3}$$

$$\int_{5Q} |T(F_Q)|^2 \, dx \leqslant C l_Q^{-2} |Q| \tag{8.4}$$

$$\sup_{Q \in \mathcal{Q}} \int_0^{l_Q} \int_Q |\gamma_\lambda(x)|^2 \, dx \, \frac{d\lambda}{\lambda} \leqslant C \sup_{Q \in \mathcal{Q}} \int_0^{l_Q} \int_Q |\gamma_\lambda(x) \cdot \Phi_\lambda \nabla F_Q(x)|^2 \, dx \, \frac{d\lambda}{\lambda}, \tag{8.5}$$

où C désigne une constante.

Le « théorème $T(b)$ d'Auscher et Tchamitchian » est l'énoncé suivant [3].

THÉORÈME 8.1. — *Sous les hypothèses* (8.1) *à* (8.5)*, la mesure*

$$d\mu(x,\lambda) = |\gamma_\lambda(x)|^2\, dx \frac{d\lambda}{\lambda}$$

est une mesure de Carleson, ce qui implique la conjecture de Kato.

Ce résultat est capital, car il a défini la stratégie qui conduisit à la preuve complète de la conjecture de Kato. Aujourd'hui, la preuve du théorème 8.1 peut sembler assez naturelle. Nous devons estimer le membre de gauche de (8.5). Nous utilisons (8.3) et (7.11). Ceci, joint aux estimations de l'action « à longue portée » de l'opérateur $(1+\lambda^2 T)^{-1}$, fournit l'estimation désirée.

Mais il reste à construire les « fonctions de test » F_Q, $Q \in \mathcal{Q}$. Cette construction repose sur de remarquables méthodes de temps d'arrêts (stopping time arguments) découvertes par S. Hofmann et J. Lewis [5] dans un contexte très différent. Cela suffisait pour démontrer la conjecture de Kato en dimension 2. La preuve finale [1] doit beaucoup aux améliorations apportées par M. Lacey.

RÉFÉRENCES

[1] P. AUSCHER, S. HOFMANN, M. LACEY, A. MCINTOSH & PH. TCHAMITCHIAN – « The solution of the Kato square root problem for second order elliptic operators on $\mathbb{R}^n$ », à paraître aux Annals of Maths.

[2] P. AUSCHER, S. HOFMANN, J. LEWIS & PH. TCHAMITCHIAN – « Extrapolation of Carleson measures and the analyticity of Kato's square root operator », Acta Math., à paraître.

[3] P. AUSCHER & PH. TCHAMITCHIAN – *Square root problem for divergence operators and related topics*, Astérisque, vol. 249, Société Mathématique de France, 1998.

[4] M. CHRIST & J.-L. JOURNÉ – « Polynomial growth estimates for multilinear singular operators », *Acta Math.* **159** (1987), p. 51–80.

[5] S. HOFMANN & J.L. LEWIS – « The Dirichlet problem for parabolic operators with singular drift terms », à paraître aux Memoirs of the Amer. Math. Soc.

[6] J.-L. JOURNÉ – « Remarks on the square root problem », *Pub. Math.* **35** (1991), p. 299–321.

[7] T. KATO – « Fractional powers of dissipative operators », *J. Math. Soc. Japan* **13** (1961), p. 246–274.

[8] J.-L. LIONS – « Espaces d'interpolation et domaines de puissances fractionnaires », *J. Math. Soc. Japan* **14** (1962), p. 233–241.

[9] V.G. MAZ'YA, S.A. NAZAROV & B.A. PLAMENEVSKII – « Absence of the De Giorgi-type theorems for strongly elliptic equations with complex coefficients », *J. Math. Sov.* **28** (1985), p. 726–739.

[10] Y. Meyer & R. Coifman – *Wavelets, Calderón-Zygmund operators and multilinear operators*, Cambridge Studies in advanced mathematics, vol. 48, Cambridge University Press, 1997.

Yves MEYER

École Normale Supérieure
C.M.L.A.
61 avenue du Président Wilson
F-94235 CACHAN Cedex
E-mail : Yves.Meyer@cmla.ens-cachan.fr

Séminaire BOURBAKI Mars 2002
54e année, 2001-2002, nO 903, p. 207 à 224

ON THE NEWTON STRATIFICATION

by **Michael RAPOPORT**

INTRODUCTION

This is a report on algebraic geometry in characteristic p. Let A/S be a family of abelian varieties over a base scheme of characteristic p. For any prime number $\ell \neq p$ the family of Tate modules $T_\ell(A_{\overline{s}})$ ($\overline{s}$ ranging over the geometric points of S) defines a local system of $\mathbb{Z}_\ell$-modules on S. The replacement for $\ell = p$ of the Tate module $T_\ell(A_{\overline{s}})$ is the *Dieudonné module* $M(A_{\overline{s}})$ which is an *F-crystal*. However, in contrast to the ℓ-adic case, the Dieudonné module is not locally constant as $\overline{s}$ varies over the base. This leads to the *Newton stratification* of S into locally closed subsets where the isomorphism classes of the *rational Dieudonné modules* $M(A_{\overline{s}}) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ are constant. Recently de Jong and Oort proved some general qualitative facts on this stratification. These were applied by Oort to the universal family of abelian varieties over the Siegel moduli space. We formulate this result in imprecise terms as follows.

THEOREM 0.1. — *The Newton stratification of the moduli space of principally polarized abelian varieties of a fixed dimension g in characteristic p has the strong stratification property (the closure of a stratum is a union of strata). Furthermore, the jumps in this stratification all occur in codimension one.*

The main tools in the proof of this theorem are the purity theorem on families of F-isocrystals and deformation theory of p-divisible groups. The latter is based on the theory of displays for formal p-divisible groups, which has recently been completed by Zink.

The layout of the report is as follows. In section 1 we introduce the notion of an F-isocrystal over a base scheme S and of the corresponding Newton stratification. Section 2 is devoted to the general theorems on Newton stratifications, and section 3 to the particular case of the Siegel moduli space. In section 4 we give the main

theorem of display theory. In the final section 5 we comment on other moduli spaces of abelian varieties.

I thank Th. Zink for his help with this report; the presentation in section 2 is largely based on his explanations. I also thank G. Laumon, F. Oort and T. Wedhorn for useful comments.

The subject matter of this report has deep historical roots, with contributions by many mathematicians. I apologize in advance for any oversights and misrepresentations, which are not intentional but rather due to my ignorance.

1. F-CRYSTALS

DEFINITION 1.1. — *Let k be a perfect field of characteristic p, with ring of Witt vectors $W(k)$. Let L be the fraction field of $W(k)$ and denote by σ the Frobenius automorphisms on k, $W(k)$ and L.*

a) *A (non-degenerate) F-*crystal over $\operatorname{Spec} k$ *is a free $W(k)$-module M of finite rank with a σ-linear endomorphism $F : M \to M$ such that $M/F(M)$ has finite length.*

b) *An F-*isocrystal over $\operatorname{Spec} k$ *is a finite-dimensional L-vector space N with a σ-linear bijective endomorphism $F : N \to N$.*

Recall that $W(k)$ is the unique complete discrete valuation ring with residue field k and with p as uniformizer. An F-crystal (M, F) defines an F-isocrystal via $(N, F) := (M, F) \otimes_{W(k)} L$. Conversely, given an F-isocrystal (N, F), the corresponding set of F-crystals is the set of $W(k)$-lattices M in N such that $F(M) \subset M$ (such lattices need not exist).

The F-isocrystals over $\operatorname{Spec} k$ form a category in the obvious way which is abelian $\mathbb{Q}_p$-linear and noetherian and artinian. If k' is a perfect field extension of k, then an F-crystal over $\operatorname{Spec} k$ defines an F-crystal over $\operatorname{Spec} k'$ via base extension $\otimes_{W(k)} W(k')$.

THEOREM 1.2 (Dieudonné). — *Let k be algebraically closed. Then the category of F-isocrystals is semi-simple. The simple objects are parametrized by the set of rational numbers. To $\lambda \in \mathbb{Q}$ corresponds the simple object E_λ defined as follows. If $\lambda = r/s$, with $s, r \in \mathbb{Z}$, $s > 0$, $(r, s) = 1$, then*

$$E_\lambda = \left(L^s, F = \begin{bmatrix} 0 & & & p^r \\ 1 & & & \\ & \ddots & & \\ & & 1 & 0 \end{bmatrix} \cdot \sigma \right) .$$

Furthermore

$$\operatorname{End}(E_\lambda) = D_\lambda,$$

where D_λ is the division algebra with center $\mathbb{Q}_p$ and invariant equal to the image of λ in $\mathbb{Q}/\mathbb{Z}$. □

We may parametrize the F-isocrystals of rank n over the algebraically closed field k by their Newton polygons, or preferably but equivalently, by their *Newton vectors*.

COROLLARY 1.3. — *Let k be algebraically closed. Then there is an injection (the Newton map)*

$$\{\textit{isomorphism classes of } F\textit{-isocrystals of rank } n\} \longrightarrow (\mathbb{Q}^n)_+,\ (N,F) \mapsto \nu(N,F).$$

Here $(\mathbb{Q}^n)_+ = \{(\nu_1,\dots,\nu_n) \in \mathbb{Q}^n,\ \nu_1 \geqslant \cdots \geqslant \nu_n\}$. The Newton map sends (N,F) to $\nu(N,F) \in (\mathbb{Q}^n)_+$, where $\lambda \in \mathbb{Q}$ occurs in $\nu(N,F)$ with multiplicity equal to the dimension of the isotypical component of type λ. The image of the Newton map may be described as follows. Write $\nu \in (\mathbb{Q}^n)_+$ as

$$\nu = (\nu(1)^{m_1},\dots,\nu(r)^{m_r}) \quad \textit{with } \nu(1) > \cdots > \nu(r).$$

Then ν lies in the image if and only if ν satisfies the integrality condition

$$\nu(i)m_i \in \mathbb{Z},\ \forall\, i = 1,\dots,r. \qquad \square$$

The components of the *Newton vector* $\nu(N,F)$ (i.e. the types of the isotypical components occurring in (N,F)) are called the *slopes* of the F-isocrystal.

Let (N,F) be an F-isocrystal over a perfect field k. Then the Newton vector of the F-isocrystal $(N,F) \otimes_{W(k)} W(\overline{k})$ over $\operatorname{Spec} \overline{k}$ is independent of the algebraically closed field $\overline{k}$ containing k. We may therefore speak of the Newton vector of (N,F).

Let X be a p-divisible group over a perfect field k of characteristic p. Then one may associate to X its (contravariant) Dieudonné module $(M(X),F)$, which is an F-crystal over $\operatorname{Spec} k$, compare [D]. *In this way one obtains*

a) an *anti-equivalence* of the category of p-divisible groups over $\operatorname{Spec} k$ and the full subcategory of the category of F-crystals over $\operatorname{Spec} k$ consisting of those F-crystals (M,F) such that $pM \subset FM$,

b) an *anti-equivalence* of the category of p-divisible groups over $\operatorname{Spec} k$ up to isogeny and the full subcategory of all F-isocrystals over $\operatorname{Spec} k$ such that all slopes lie between 0 and 1.

Let S be a scheme of characteristic p. An *F-crystal over S* is a crystal $\mathcal{E}$ of finite locally free $\mathcal{O}_{S_{\text{cris}}}$-modules, with a morphism $F : \mathcal{E}^{(\sigma)} \to \mathcal{E}$ such that the kernel and cokernel of F are annihiliated by a power of p. Here $\mathcal{O}_{S_{\text{cris}}}$ denotes the structure sheaf on the big crystalline site of S over $\mathbb{Z}_p$. We often write $\mathcal{E}$ for the F-crystal $(\mathcal{E},F)$. This notion makes precise the intuitive concept of a family of F-crystals parametrized by the (perfect closures of the residue fields of) points of S. The F-crystals over S form a $\mathbb{Z}_p$-linear category. A morphism of F-crystals $f : \mathcal{E} \to \mathcal{E}'$ is an *isogeny* if there exists locally on S a morphism $g : \mathcal{E}' \to \mathcal{E}$ such that $gf = p^n$ and $fg = p^n$ for some n. The category of F-isocrystals over S is obtained by formally inverting isogenies of F-crystals.

Example 1.4. — Let X be a p-divisible group over S. Then crystalline Dieudonné theory [Me] associates to X an F-crystal over S. More precisely, the Lie algebra of the universal extension of X is a crystal, and its dual is an F-crystal where F is induced by the Frobenius $\mathrm{Fr} : X \to X^{(\sigma)}$, compare [Me], IV.2.5.

For the sequel it is not essential to have mastered the notion of an F-crystal over a scheme, in order to understand the resulting statements for p-divisible groups (although some proofs in this special case are based on general F-crystals for which one can perform the usual linear algebra operations like tensor products etc.).

The most basic statement about families of F-crystals is the following semi-continuity theorem. Recall the usual dominance order on $(\mathbb{Q}^n)_+$, for which $(\nu_1, \dots, \nu_n) \leqslant (\nu'_1, \dots, \nu'_n)$ if and only if

$$\sum_{j=1}^{r} \nu_i \leqslant \sum_{j=1}^{r} \nu'_i, \quad \forall\, r = 1, \dots, n-1, \quad \text{and } \sum_{j=1}^{n} \nu_j = \sum_{j=1}^{n} \nu'_j.$$

For $\nu = (\nu_1, \dots, \nu_n) \in \mathbb{Q}^n$ we set $\|\nu\| = \sum\limits_{j=1}^{n} \nu_j$.

THEOREM 1.5 (Grothendieck [G]). — *Let $(\mathcal{E}, F)$ be an F-isocrystal over a scheme S of characteristic p. Then the Newton vector of $(\mathcal{E}_s, F_s)$, for s ranging over the points of S, goes down under specialization. More precisely, let $(\mathcal{E}, F)$ be of constant rank n. Then the function $s \mapsto \|\nu(\mathcal{E}_s, F_s)\|$ is locally constant on S and for any $\nu_0 \in (\mathbb{Q}^n)_+$ the set*

$$\{s \in S;\ \nu(\mathcal{E}_s, F_s) \leqslant \nu_0\}$$

is Zariski closed in S.

We note here that the Newton vector of the fiber of $\mathcal{E}$ at a geometric point $\overline{s}$ of S only depends on the underlying point $s \in S$.

The proof by Katz in [Ka] relies on the relation between the Newton vector of an F-isocrystal and the divisibility by p of F with respect to an underlying F-crystal. For (the F-isocrystal associated to) a p-divisible group over S a simple proof is contained in [D]. □

Remark 1.6. — This theorem is reminiscent of a theorem on vector bundles on a compact Riemann surface. In this theory one associates to a vector bundle of rank n its *Harder-Narasimhan vector* in $(\mathbb{Q}^n)_+$, and it is a basic fact that the HN-vector *goes up* (!) under specialization [AB].

COROLLARY 1.7. — *Let $(\mathcal{E}, F)$ be an F-isocrystal over a noetherian scheme S of characteristic p. Then the set of points of S where the Newton vector is constant is locally closed in S and this defines a finite decomposition of S.*

Proof. — We may assume S connected. Let us only consider the case when $(\mathcal{E}, F)$ comes from a p-divisible group X over S. Then the height and the dimension of X are constant. The assertion then follows from the preceding theorem by the following two observations.

For any $\nu_0 \in (\mathbb{Q}^n)_+$ the set

$$\{\nu \in (\mathbb{Q}^n)_+ \ ; \ \nu \leqslant \nu_0 \text{ and } \nu \text{ satisfies the integrality condition in Cor. 1.3}\}$$

is finite.

If $\nu = \nu(M(X), F)$ for a p-divisible group X of dimension d and height n, then $\nu \leqslant (1^d, 0^{n-d})$. This is a consequence of Mazur's inequality between the Hodge vector of an F-crystal over a perfect field and the Newton vector of its underlying F-isocrystal, [Ka]. □

Let $(\mathcal{E}, F)$ be an F-isocrystal of rank n over a noetherian scheme S of characteristic p. Associating to a geometric point $\overline{s}$ of S the Newton vector of $(\mathcal{E}_{\overline{s}}, F_{\overline{s}})$, we obtain a map

$$S \longrightarrow (\mathbb{Q}^n)_+.$$

Let S_ν be the fiber of this map over $\nu \in (\mathbb{Q}^n)_+$ (with its reduced scheme structure). The corresponding disjoint decomposition of S, finite according to Corollary 1.7, is called the *Newton stratification of S associated to the F-isocrystal $(\mathcal{E}, F)$*. The subschemes S_ν are called the *Newton strata.*

We speak of a stratification *in the strong sense* if the closure of a stratum is a union of strata. In general the Newton stratification associated to an F-isocrystal is *not* a stratification in the strong sense.

2. PURITY OF THE NEWTON STRATIFICATION

Let $(\mathcal{E}, F)$ be an F-isocrystal of rank n over a scheme S of characteristic p, with associated Newton stratification $(S_\nu)_{\nu \in (\mathbb{Q}^n)_+}$ of S. The purity theorem states that the jumps in this stratification all occur in codimension one. The corresponding statement for families of vector bundles on a Riemann surface (*cf.* Remark 1.6) is false.

THEOREM 2.1 (de Jong, Oort [JO]). — *Let $(\mathcal{E}, F)$ be an F-isocrystal of rank n over a locally noetherian scheme S of characteristic p, with associated Newton stratification $(S_\nu)_{\nu \in (\mathbb{Q}^n)_+}$. Let $\nu \in \mathbb{Q}^n_+$. Let η be a generic point of the scheme $\overline{S}_\nu \smallsetminus S_\nu$. Then*

$$\dim \mathcal{O}_{\overline{S}_\nu, \eta} = 1 \quad .$$

An equivalent statement is the following.

THEOREM 2.2. — *Let $(\mathcal{E}, F)$ be an F-isocrystal over a locally noetherian scheme S of characteristic p. Let U be an open subset of S such that* $\operatorname{codim}(S \smallsetminus U) \geqslant 2$. *If the Newton vector of $(\mathcal{E}, F)$ is constant at all points of U, then it is constant on all of S.*

These theorems are referred to as *purity theorems* since they are reminiscent of the purity theorem of Nagata-Zariski on étale coverings. We shall be mainly interested in this statement when $(\mathcal{E}, F)$ comes from a p-divisible group over S. The structure of a p-divisible group with constant Newton vector is addressed in the following two results.

THEOREM 2.3 (de Jong, Oort [JO]). — *Let $S = \operatorname{Spec} A$, where A is a complete noetherian local ring of characteristic p with algebraically closed residue field k. Let X be a p-divisible group over S with constant* scalar *Newton vector, i.e., there is only one slope at all points of S* (isoclinic case). *Then X is isogenous to a* constant *p-divisible group, i.e., one of the form $X_0 \times_{\operatorname{Spec} k} S$ for a p-divisible group X_0 over $\operatorname{Spec} k$.*

The motivation for this theorem is the heuristic idea that a p-divisible group with constant scalar Newton vector is analogous to a local system over S. The hypotheses on S in Theorem 2.3 ensure that it behaves like a simply connected space. One may expect a similar result for general F-isocrystals with constant scalar Newton vector, compare [JO], Remark 2.18. The case where A is a complete discrete valuation ring with algebraically closed residue field is due to Katz, [Ka], Thm. 2.7.1. □

The constancy up to isogeny becomes false when the constant Newton vector has more than one slope, for there can be highly nontrivial extensions of constant F-isocrystals over $\operatorname{Spec} k[[t]]$ (e.g. the p-divisible group of the universal deformation of an ordinary elliptic curve is a nontrivial extension of $\mathbb{Q}_p/\mathbb{Z}_p$ by $\widehat{\mathbb{G}}_m$). When there is more than one slope, there is the following analogue of the Harder-Narasimhan filtration of vector bundles, *cf.* Remark 1.6.

THEOREM 2.4 (Zink [Z3]). — *Let S be a regular scheme of characteristic p. Let X be a p-divisible group over S with constant Newton vector $\nu \in (\mathbb{Q}^n)_+$. Then X is isogenous to a p-divisible group Y which admits a filtration by closed embeddings of p-divisible groups*

$$(0) = Y_0 \subset Y_1 \subset \cdots \subset Y_r = Y,$$

such that the following condition is satisfied. Let $\nu = (\nu(1)^{m_1}, \ldots, \nu(r)^{m_r})$ with $\nu(1) > \cdots > \nu(r)$, cf. integrality condition in Cor. 1.3. Then there are natural numbers $r_i \geqslant 0$, $s_i > 0$ $(i = 1, \ldots, r)$ such that $\nu(i) = r_i/s_i$ and such that

$$p^{-r_i} \operatorname{Fr}^{s_i} : Y_i \longrightarrow Y_i^{(\sigma^{s_i})} \quad \textit{is an isogeny}$$

and

$$p^{-r_i} \operatorname{Fr}^{s_i} : Y_i/Y_{i-1} \longrightarrow (Y_i/Y_{i-1})^{(\sigma^{s_i})} \quad \textit{is an isomorphism,}$$

$\forall\, i = 1, \ldots, r$.

The degree of the isogeny between X and Y may be bounded in terms of the height of X. The heuristic idea behind this theorem is that the isotypic *direct sum decomposition* of an F-isocrystal over an algebraically closed field is replaced in the case of a more general base scheme by a filtration. In ongoing work of Oort and Zink, the regularity hypothesis on S is weakened. The case where $S = \operatorname{Spec} k$ for an arbitrary field k is due to Grothendieck [G]. □

The proof of Theorem 2.2 is based on the following result which is of independent interest.

THEOREM 2.5 (de Jong, Oort [JO]). — *Let $S = \operatorname{Spec} A$, where A is a normal complete noetherian local ring of dimension 2 with algebraically closed residue field k. Let $U = S \smallsetminus \{s\}$, where s denotes the closed point. Let $\pi : \widetilde{S} \to S$ be a resolution of singularities, i.e. a proper morphism from a regular scheme which induces an isomorphism over U and such that $E = \pi^{-1}(s) = \bigcup_{i=1}^{m} E_i$ is a union of smooth divisors crossing each other normally. Identifying $\pi^{-1}(U)$ with U we have the restriction map*

$$H^1_{\text{ét}}(\widetilde{S}, \mathbb{Z}_p) \longrightarrow H^1_{\text{ét}}(U, \mathbb{Z}_p).$$

This map is an isomorphism.

In terms of the fundamental groups (w.r.t. some geometric point of U) the assertion is that

$$\operatorname{Hom}(\pi_1(\widetilde{S}), \mathbb{Z}_p) \xrightarrow{\sim} \operatorname{Hom}(\pi_1(U), \mathbb{Z}_p). \tag{1}$$

Since both $\pi_1(\widetilde{S})$ and $\pi(U)$ are factor groups of the Galois group of the fraction field of S, the homomorphism $\pi_1(U) \to \pi_1(\widetilde{S})$ is surjective. Therefore we have the injectivity of the map (1), and to prove the surjectivity we may replace $\mathbb{Z}_p$ by $\mathbb{Q}_p$ in (1). Topologically or when $p \neq \operatorname{char} k$, this surjectivity is easy to see. Indeed, we need the injectivity of

$$H^2_E(\widetilde{S}, \mathbb{Q}_p) \longrightarrow H^2(\widetilde{S}, \mathbb{Q}_p). \tag{2}$$

But $H^2_E(\widetilde{S}, \mathbb{Q}_p)$ has the classes $\operatorname{cls}(E_i)$, $i = 1, \ldots, m$, as basis (purity). The image of $\operatorname{cls}(E_i)$ under the composition

$$H^2_E(\widetilde{S}, \mathbb{Q}_p) \longrightarrow H^2(\widetilde{S}, \mathbb{Q}_p) \xrightarrow{\text{Res}} H^2(E_j, \mathbb{Q}_p)$$

is the intersection product $(E_i.E_j)$. The injectivity of (2) follows therefore from the negative-definiteness of the intersection matrix $(E_i.E_j)_{i,j=1,\ldots,m}$.

When $p = \operatorname{char} k$, the proof of Theorem 2.5 is much more difficult. (That the situation in this case is radically different is already apparent from the fact that $H^2_{\text{ét}}(\widetilde{S}, \mathbb{Z}_p) = (0)$ when $p = \operatorname{char} k$. This is easily checked using Artin-Schreier theory.) Suppose that A has characteristic p, i.e. $k \subset A$. In this case, using de Jong's technique of alterations there is a reduction to the following situation. Let $\mathcal{C} \to \operatorname{Spec} k[[t]]$ be a flat projective family of curves with smooth generic fiber and strict semistable

reduction. Let $\mathcal{C}'$ be the scheme obtained from $\mathcal{C}$ by collapsing a proper union E of irreducible components of the special fiber to a point P. Then A is the complete local ring of P.

One now starts with an element $\alpha \in H^1_{\text{ét}}(U, \mathbb{Z}_p)$ and first globalizes it into an element $\alpha_1 \in H^1_{\text{ét}}(\mathcal{C}' \smallsetminus \{P\}, \mathbb{Z}_p) = H^1_{\text{ét}}(\mathcal{C} \smallsetminus E, \mathbb{Z}_p)$. This element α_1 is then extended to $\alpha_2 \in H^1_{\text{ét}}(\mathcal{C}, \mathbb{Z}_p)$ by using de Jong's extension theorem on homomorphisms of p-divisible groups [J1], [J2]. According to this theorem, any homomorphism between the generic fibers of p-divisible groups over Spec $k[[t]]$ extends. □

Let X be a p-divisible group of height h over a scheme S of characteristic p, for which there exists $r \geqslant 0$, $s > 0$ such that

$$p^{-r}\,\mathrm{Fr}^s : X \longrightarrow X^{(\sigma^s)} \text{ is an isomorphism.} \tag{3}$$

To X we associate the lisse p-adic sheaf of $W(\mathbb{F}_{p^s})$-modules $C_X = \varprojlim C_{X,n}$ for the étale topology on S, such that for any affine S-scheme Spec R

$$C_{X,n}(\operatorname{Spec} R) = \{x \in M/p^n M;\ p^{-r}F^s(x) = x\}. \tag{4}$$

Here M denotes the $W(R)$-module defined by the Dieudonné crystal of X. The fibers of C_X are free $W(\mathbb{F}_{p^s})$-modules of rank h. The formation of C_X is compatible with base change and defines a functor from the category of p-divisible groups over S with (3) to the category of lisse p-adic sheaves of $W(\mathbb{F}_{p^s})$-modules on S. The corresponding $W(\mathbb{F}_{p^s}) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$-adic sheaf only depends on the isogeny class of X and corresponds to a representation of the fundamental group,

$$\varrho_X : \pi_1(S) \longrightarrow GL_h(W(\mathbb{F}_{p^s}) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p). \tag{5}$$

Let R be a discrete valuation ring of characteristic p, with uniformizer π, residue field k and fraction field K. Let X be a p-divisible group over R. After replacing X_K by an isogenous p-divisible group Y over K we have integers $r_i \geqslant 0$, $s > 0$ and a filtration $(0) = Y_0 \subset Y_1 \subset \cdots \subset Y_r = Y$ as in Theorem 2.4. Applying the preceding considerations to $S = \operatorname{Spec} K$ and Y_i/Y_{i-1} we therefore obtain p-adic Galois representations

$$\varrho_i = \varrho_{Y_i/Y_{i-1}} : \mathrm{Gal}\,(\overline{K}/K) \longrightarrow GL_{h_i}(W(\mathbb{F}_{p^s}) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p), \tag{6}$$

where $h_i = \text{height}(Y_i/Y_{i-1})$. The proof of Theorem 2.2 is based on the following lemma.

LEMMA 2.6. — *With the previous notation, the following conditions are equivalent.*

(i) $\wedge^{h_i}\varrho_i$ *is unramified,* $\forall\, i = 1, \ldots, r$.

(ii) The Newton vector of X *is constant.*

Under these conditions the representations ϱ_i *are also unramified.*

In the proof of this lemma, again, as in the proof of Theorem 2.5, de Jong's theorem on extension of homomorphisms of p-divisible groups (or rather the techniques entering into the proof) plays a key role. □

Proof of Theorem 2.2. — We limit ourselves to the case where $(\mathcal{E}, F)$ comes from a p-divisible group X on S. An easy reduction allows us to assume that $S = \operatorname{Spec} A$, where A is a complete normal noetherian local ring of dimension 2 and where $U = S \smallsetminus \{s\}$, with s denoting the special point. In proving Theorem 2.2, we may replace A by an A-algebra A' of the same kind such that the special point s' of $\operatorname{Spec} A'$ is the unique point mapping to s. Since the Newton vector of X is constant on the regular scheme U, we obtain via Theorem 2.4 p-adic Galois representations,

$$\varrho_i : \pi_1(U) \longrightarrow GL_{h_i}(W(\mathbb{F}_{p^s}) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p), \tag{7}$$

$i = 1, \ldots, r$. The determinant representation of each ϱ_i is a character of $\pi_1(U)$ with values in $(W(\mathbb{F}_{p^s}) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^\times$. Since $(W(\mathbb{F}_{p^s}) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p)^\times$ contains an open subgroup of the form $\mathbb{Z}_p^s$, we may assume by the initial remark that $\wedge^{h_i}\varrho_i$ is an s-tuple of homomorphisms from $\pi_1(U)$ to $\mathbb{Z}_p$. Let $\pi : \widetilde{S} \to S$ be a resolution of singularities. Then by Theorem 2.5 this s-tuple of homomorphisms factors through $\pi_1(\widetilde{S})$. Let E be an irreducible component of the exceptional fiber $\pi^{-1}(s)$. Applying Lemma 2.6 to the discrete valuation ring $\mathcal{O}_{\widetilde{S},E}$, the pullback of X to $\operatorname{Spec} \mathcal{O}_{\widetilde{S},E}$ has constant Newton vector, as had to be shown. □

3. THE SIEGEL MODULI SPACE

We fix a positive integer g. For an auxiliary integer $m \geqslant 3$ prime to p, we denote by $\mathcal{M} = \mathcal{M}_g = \mathcal{M}_{g,m}$ the *Siegel moduli space of genus g* over $\operatorname{Spec} \mathbb{F}_p$. It represents the functor which to a locally noetherian scheme S in characteristic p associates the set of isomorphism classes of triples (A, λ, η), where A is an abelian scheme of relative dimension g over S, and $\lambda : A \to \widehat{A}$ is a principal polarization, and η is a (full) level-m-structure on (A, λ).

The universal abelian scheme over $\mathcal{M}$ defines a p-divisible group X on $\mathcal{M}$. The existence of a polarization implies that the Newton vectors of the fibers of X lie in the subset $(\mathbb{Q}^{2g})^1_+$ of $(\mathbb{Q}^{2g})_+$,

$$(\mathbb{Q}^{2g})^1_+ = \{(\nu_1, \ldots, \nu_{2g}) \in (\mathbb{Q}^{2g})_+ ; \nu_i + \nu_{2g-i+1} = 1, \ \forall\, i = 1, \ldots, g, \\ 0 \leqslant \nu_i \leqslant 1, \ \forall\, i = 1, \ldots, 2g\}.$$

Let B_g be the set of elements in $(\mathbb{Q}^{2g})^1_+$ which satisfy the integrality condition in Cor. 1.3. Then B_g is a finite partially ordered set (*poset*), which has a unique maximal

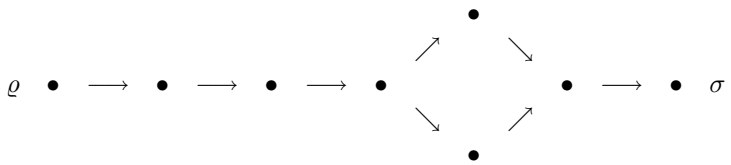
element and a unique minimal element,

$$\varrho = (1^g, 0^g) \text{ the ordinary Newton vector (maximal)}$$
$$\sigma = ((1/2)^{2g}) \text{ the supersingular Newton vector (minimal).}$$

For $\nu \in B_g$ we denote by $\mathcal{S}_\nu$ the corresponding Newton stratum in $\mathcal{M}$. By Grothendieck's semi-continuity theorem, Theorem 1.5, we have

$$\overline{\mathcal{S}}_\nu \subset \bigcup_{\nu' \leqslant \nu} \mathcal{S}_{\nu'}. \tag{8}$$

In particular $\mathcal{S}_\sigma$ is a closed subset and $\mathcal{S}_\varrho$ is an open subset.



Picture of B_4

For $\nu \in B_g$, with $\nu = (\nu_1, \dots, \nu_{2g})$, let

$$\Delta(\nu) = \{(i,j) \in \mathbb{Z}^2;\ 0 \leqslant i \leqslant g,\ \textstyle\sum_{\ell=1}^{i} \nu_{2g-\ell+1} \leqslant j < i\} \tag{9}$$

$$d(\nu) = \#\Delta(\nu) \tag{10}$$

Example 3.1. — $d(\varrho) = g(g+1)/2, \quad d(\sigma) = \lfloor g^2/4 \rfloor$.

LEMMA 3.2. — *Let $\nu, \nu' \in B_g$. Then*

(i) $\nu \leqslant \nu'$ if and only if $\Delta(\nu) \subseteq \Delta(\nu')$.

(ii) If $\nu \leqslant \nu'$, then any shortest chain in the poset B_g starting at ν and ending at ν' has length $d(\nu') - d(\nu)$. In particular, B_g is a catenary poset. □

THEOREM 3.3 (Oort [O2]). — *Each Newton stratum $\mathcal{S}_\nu$ is equidimensional of dimension $d(\nu)$. The Newton stratification of $\mathcal{M}$ has the strong stratification property,*

$$\overline{\mathcal{S}}_\nu = \bigcup_{\nu' \leqslant \nu} \mathcal{S}_{\nu'}.$$

Remark 3.4. — The analogous statement for moduli spaces of vector bundles on a Riemann surface is false. More precisely, consider the stack $\mathcal{M} = \mathcal{M}_X(r,d)$ of holomorphic vector bundles of rank r and degree d over a Riemann surface X. Then $\mathcal{M}$ is the disjoint union of its Harder-Narasimhan strata $\mathcal{M}_\nu$, and we have

$$\overline{\mathcal{M}}_\nu \subset \bigcup_{\nu' \geqslant \nu} \mathcal{M}_{\nu'},$$

cf. Remark 1.6. If $g(X) = 0$ we have equality here and the same holds for $g(X) = 1$, according to a recent paper of Friedman and Morgan [FM]. This fails for $g(X) \geqslant 2$ [FM]. It is conceivable that

$$\nu' \geqslant \nu \Longleftrightarrow \overline{\mathcal{M}}_\nu \cap \mathcal{M}_{\nu'} \neq \varnothing$$

and in loc. cit. this is proved, provided ν' and ν are adjacent. □

The proof of Theorem 3.3 is based on the following theorem.

THEOREM 3.5 (Oort [O2]). — *Let (X_0, λ_0) be a p-divisible group with principal polarization of dimension g and height $2g$ over an algebraically closed field k of characteristic p. Let $\nu_0 \in B_g$ be the Newton vector of X_0, and let $\nu \in B_g$ with $\nu_0 \leqslant \nu$. Then there exists a principally polarized p-divisible group (X, λ) over $k[[t]]$ with special fiber (X_0, λ_0) and with Newton vector of the generic fiber of X equal to ν.*

Theorem 3.5 implies via the Serre-Tate theorem the second statement in Theorem 3.3. Since, as is easily seen, $\mathcal{S}_\sigma \neq \varnothing$ (the supersingular locus), Theorem 3.5 implies $\mathcal{S}_\nu \neq \varnothing$ for all $\nu \in B_g$. The first statement in Theorem 3.3 now follows from Lemma 3.2 and the purity theorem. □

The two extreme cases in the Newton stratification deserve a separate discussion.

The supersingular stratum $\mathcal{S}_\sigma$. — The fact that $\mathcal{S}_\sigma$ is equidimensional of dimension $d(\sigma) = \left[g^2/4\right]$ (a special case of Theorem 3.3) was proved earlier by Li and Oort [LO], among other things. The supersingular Newton stratum is exceptional in several aspects. Every abelian variety occurring as a fiber of the universal abelian scheme at a point of $\mathcal{S}_\sigma$ is isogenous to E^g, where E is a supersingular elliptic curve. Moreover, let $x \in \mathcal{S}_\sigma$ and let (A, λ) be the fiber of x of the universal object on $\mathcal{M}$. Then one can represent (A, λ) in an almost canonical way as the quotient of E^g by a finite group scheme. This leads to the dimension formula for this stratum and in fact much more. There is the hope for an *explicit synthetic* description of $\mathcal{S}_\sigma$, like the one of Kaiser [K] for $g = 2$ (compare also [KR], and [R] for $g = 3$). In [LO] the number of irreducible components of $\mathcal{S}_\sigma$ is given.

The ordinary stratum $\mathcal{S}_\varrho$. — In contrast to the supersingular stratum, the ordinary stratum is quite amorphous and nonlinear, and there is no hope of an explicit description of it. According to Theorem 3.3, $\mathcal{S}_\varrho$ is open and dense in $\mathcal{M}$. This fact has been known for a long time, by more direct and easier proofs: 1) There is the proof by Mumford [M], and Norman and Oort [NoO] (compare also Chai and Faltings [FC]) using Cartier theory to construct deformations. 2) There is the proof by Koblitz [Kob], compare also [Ill], App. 2, who investigated by deformation-theoretic arguments the stratification of $\mathcal{M}$ by the p-rank of X. 3) There is the global proof using toroidal compactifications [FC]. 4) There is the proof of Ngo and Genestier [NG] who deduce the density result from a corresponding density result (of a combinatorial nature) in

bad reduction. Furthermore, Chai [C1] has proved the much stronger assertion that the orbit of an arbitrary ordinary point under the Hecke correspondences of degree prime to p is dense in $\mathcal{M}$. □

The proof of Theorem 3.5 is rather round-about. For a p-divisible group X over an algebraically closed field k, let

$$a(X) = \mathrm{Hom}_k(\alpha_p, X) = \dim_k M/(F(M) + V(M)). \tag{11}$$

Here (M, F) is the Dieudonné module of X and $V = pF^{-1}$. Hence $a(X) = 0$ if and only if $X \simeq \widehat{\mathbb{G}}_m^d \times (\mathbb{Q}_p/\mathbb{Z}_p)^{d'}$. In a first step one proves Theorem 3.5 under the additional assumption $a(X_0) = 1$, *cf.* Oort [O1] (the case where $a(X_0) = 0$, where $\nu_0 = \varrho$, is trivial). The technical tool for this is the theory of displays, compare Section 4, which allows one to write down explicitly by display equations deformations of a (polarized) formal p-divisible group. The hypothesis $a(X_0) = 1$ is then needed to read off from these display equations the Newton vectors of the p-divisible groups occurring in the deformation.

In a second step, one shows that (X_0, λ_0) can be deformed into a principally polarized p-divisible group (X, λ) with the same Newton vector and with $a(X) = 1$. This in turn is reduced to the following statement which is of independent interest.

THEOREM 3.6 (de Jong, Oort [JO]). — *Let X_0 be a p-divisible group over an algebraically closed field k of characteristic p such that its F-isocrystal is irreducible. Then there exists an irreducible scheme T over k and a p-divisible group X over T together with an isogeny $X_0 \times_{\mathrm{Spec}\, k} T \to X$ over T, such that any p-divisible group over k isogenous to X_0 occurs as a fiber of X at a k-rational point of T.*

An equivalent formulation of the previous theorem is the following. Let X_0 be as in the previous theorem. By [RZ] the following functor on $(Sch/\,\mathrm{Spec}\, k)$ is representable by a formal scheme $\mathcal{S}$ locally formally of finite type over $\mathrm{Spec}\, k$,

$$S \longmapsto \{\text{isomorphism classes of pairs } (X, \varrho), \text{ where } X \text{ is a } p\text{-divisible group over } S \text{ and } \varrho : X_0 \times_{\mathrm{Spec}\, k} S \longrightarrow X \text{ is a quasi-isogeny of height } 0\}.$$

Then $\mathcal{S}$ is irreducible. □

COROLLARY 3.7. — *Let X_0 be as in the previous theorem. Then there exists a deformation of X_0 into a p-divisible group X isogenous to X_0 with $a(X) = 1$.*

Indeed, the locus in T where $a(X) = 1$ is open. It therefore suffices to produce one point $t \in T(k)$ where $a(X_t) = 1$. This is easy. □

CONJECTURE 3.8. — *Let $\nu, \nu' \in B_g$ with $\nu' \leqslant \nu$. The closure of each irreducible component of $\mathcal{S}_\nu$ meets $\mathcal{S}_{\nu'}$.*

Conjecture 3.8 would certainly hold if Oort's conjecture [O2] was true, according to which for $\nu \neq \sigma$ the intersection of $\mathcal{S}_\nu$ with any connected component of $\mathcal{M}$ is irreducible.

4. DISPLAYS

Display equations for formal p-divisible groups were introduced by Mumford [M]. These techniques were applied to moduli problems of abelian varieties by Norman [No] and Norman and Oort [NoO]. We will follow here the recent formulation of the theory due to Zink [Z1].

Let R be a ring of characteristic p. We denote by $W(R)$ its ring of Witt vectors and by $x \mapsto {}^F x$ resp. $x \mapsto {}^V x$ its Frobenius resp. Verschiebung endomorphisms. Let $I_R \subset W(R)$ be the ideal of Witt vectors with trivial 0-component.

DEFINITION 4.1. — *A* not necessarily nilpotent display (= 3n-display) over R *is a quadruple* (P, Q, F, V^{-1}) *consisting of a finitely generated projective* $W(R)$*-module* P*, a submodule* $Q \subset P$ *and* ^{F}*-linear maps* $F : P \to P$ *and* $V^{-1} : Q \to P$*. The following conditions are required:*
(i) $I_R P \subset Q \subset P$ *and the quotient* P/Q *is a projective* R*-module.*
(ii) $V^{-1} : Q \to P$ *is a* ^{F}*-linear epimorphism.*
(iii) *For* $x \in P$ *and* $w \in W(R)$ *we have* ${}^V w \cdot x \in Q$ *and we require that*

$$V^{-1}({}^V w \cdot x) = w \cdot F(x).$$

We note that F is determined by the remaining data. There is no operator V. The reason for the notation comes from the following example.

Example 4.2. — Let $R = k$ be a perfect field. Then an F-crystal (M, F) over k such that $pM \subset FM$ defines a 3n-display (M, VM, F, V^{-1}). Here as usual $V = pF^{-1}$. This defines an equivalence of categories.

The notion of a display is obtained by imposing a nilpotency condition as follows. After localization in R there exists a $W(R)$-basis $e_1, \ldots, e_n$ of P such that

$$Q = I_R e_1 \oplus \cdots \oplus I_R e_d \oplus W(R) e_{d+1} \oplus \cdots \oplus W(R) e_n,$$

for some d with $0 \leqslant d \leqslant n$. Then there exists an invertible matrix $(\alpha_{ij}) \in GL_n(W(R))$ such that

$$Fe_j = \sum_{i=1}^{n} \alpha_{ij} e_i \quad \text{for } j = 1, \ldots, d$$

$$V^{-1} e_j = \sum_{i=1}^{n} \alpha_{ij} e_i \quad \text{for } j = d+1, \ldots, n.$$

Conversely, any $(\alpha_{ij}) \in GL_n(W(R))$ defines a 3n-display. Let $(\beta_{k\ell})$ be the inverse of (α_{ij}). Let $B \in M_{n-d}(R)$ be the image of $(\beta_{k\ell})_{k,\ell=d+1,\dots,n}$ under the 0-component map

$$M_{n-d}(W(R)) \longrightarrow M_{n-d}(R) \quad .$$

Let $B^{(p)}$ be the matrix obtained from B by raising its coefficients to the power p. The nilpotency condition can now be formulated: there exists N such that

$$B^{(p^N)} \cdot \dots \cdot B^{(p)} \cdot B = 0.$$

In the context of Example 4.2 the F-crystal (M, F) defines a display if and only if $pM \subset FM$ and if $V = pF^{-1}$ is topologically nilpotent on M.

THEOREM 4.3 (Zink [Z1]). — *We assume that the nilideal of R is nilpotent. Then there is a fully faithful functor BT from the category of displays over R to the category of formal p-divisible groups over R. This is an equivalence of categories if either R is an excellent local ring or an algebra of finite type over a field k.*

It is quite likely that this equivalence of categories holds for any noetherian ring R of characteristic p. The functor BT has the following properties: 1) It commutes with arbitrary base change. 2) Lie $BT(P, Q, F, V^{-1}) = P/Q$. 3) P can be identified with the value at $W(R)$ of the crystal defined by the universal extension of $BT(P, Q, F, V^{-1})$, *cf.* Example 1.4. 4) The passage from a formal p-divisible group to its dual p-divisible group can be expressed in terms of displays, provided that the dual p-divisible group is a formal group, i.e. has trivial étale part. □

The theory also works if p is only supposed to be nilpotent in R. For an extension of the theory to p-divisible groups with an étale part, compare [Z2].

5. OTHER MODULI SPACES OF ABELIAN VARIETIES

Let F be a finite-dimensional semisimple $\mathbb{Q}$-algebra equipped with a positive involution $*$ and let V be a finite F-module equipped with an alternating non-degenerate $\mathbb{Q}$-valued skew-hermitian pairing $\langle\ ,\ \rangle$. The F-linear similitudes of $(V, \langle\ ,\ \rangle)$ form an algebraic group G over $\mathbb{Q}$. We assume that G is a *connected* reductive algebraic group. We also fix a conjugacy class of algebraic homomorphisms $h : \mathbb{C}^\times \to G(\mathbb{R})$ satisfying the usual Riemann conditions. Let E be the corresponding Shimura field, i.e. the field of definition of the corresponding conjugacy class of cocharacters $\mu : \mathbb{G}_{m,\mathbb{C}} \to G_{\mathbb{C}}$. We assume that p is a prime of good reduction, in particular $G_{\mathbb{Q}_p}$ is unramified, and we choose a hyperspecial maximal compact subgroup K_p of $G(\mathbb{Q}_p)$. We choose a prime ideal of E over p with residue field κ. After a choice of some sufficiently small open compact subgroup $K^p \subset G(\mathbb{A}_f^p)$, Kottwitz [Kot1] has defined a moduli problem of abelian varieties which is representable by a smooth quasi-projective scheme $\mathcal{M} = \mathcal{M}(G, h)_K = \mathcal{M}(F, V, \langle\ ,\ \rangle, h, K^p.K_p)$ over Spec κ.

Let L be the fraction field of $W(\overline{\mathbb{F}}_p)$ and let $B(G)$ be the set of σ-conjugacy classes in $G(L)$. By associating to a point $s \in \mathcal{M}$ the *F-isocrystal with G-structure* defined by the fiber at (a geometric point over) s of the universal abelian scheme over $\mathcal{M}$ with its auxiliary structure (endomorphisms and polarization), we obtain a map

$$\mathcal{M} \longrightarrow B(G).$$

The conjugacy class μ defines a finite subset $B(G,\mu)$ of $B(G)$ ([Kot2], §6). It is defined by the group-theoretic version of Mazur's inequality, compare the proof of Corollary 1.7. The image of the map above is contained in $B(G,\mu)$ [RR]. (In the case of the Siegel moduli space $\mathcal{M}_g$ we have $G = GSp_{2g}$ and $B(G,\mu) = B_g$, *cf.* section 3.) Furthermore, $B(G,\mu)$ is partially ordered and the semicontinuity theorem 1.5 continues to hold in this context [RR]. We therefore obtain the *generalized Newton stratification of* $\mathcal{M}$ (by the locally closed subsets arising as inverse images of elements of $B(G,\mu)$),

$$\mathcal{M} = \bigcup_{b \in B(G,\mu)} \mathcal{M}_b.$$

Just as B_g, also $B(G,\mu)$ is a catenary poset [C2] with a unique minimal element b_0 (the *μ-basic element*) and a unique maximal element b_1 (the *μ-ordinary element*).

THEOREM 5.1 (Wedhorn[W]). — *The μ-ordinary locus $\mathcal{M}_{b_1}$ is open and dense in $\mathcal{M}$.* □

This is about the only known general statement in direction of the following conjecture.

CONJECTURE 5.2. — (i) *The generalized Newton stratification of $\mathcal{M} = \mathcal{M}(G,h)_K$ has the strong stratification property.*

(ii) *The generalized Newton stratum corresponding to $b \in B(G,\mu)$ is equidimensional of dimension $d(b) = \dim \mathcal{M} - c(b)$, where $c(b)$ is the length of a chain joining b to b_1.*

(iii) *Let $b, b' \in B(G,\mu)$ with $b' \leqslant b$. The closure of each irreducible component of $\mathcal{M}_b$ meets $\mathcal{M}_{b'}$.*

We note that Chai [C2] has given a group theoretical formula for $d(b)$. When G is a group of unitary similitudes, there are results supporting (i) and (ii) of this conjecture:

THEOREM 5.3 (Oort). — *Let F be an imaginary quadratic field such that p splits in F. Then* (i) *and* (ii) *of Conjecture 5.2 hold true for $\mathcal{M} = \mathcal{M}(F, V, \langle\ ,\ \rangle, h, K^p.K_p)$.*

The proof is analogous to the proof of Theorem 3.3 (which proves (i) and (ii) of Conjecture 5.2 for the Siegel moduli space). The analogue of Theorem 3.5 is the following statement which confirms a conjecture of Grothendieck [G]. Its proof is similar to that of Theorem 3.5, but simpler.

THEOREM 5.4 (Oort[O2]). — *Let X_0 be a p-divisible group of height n and dimension d over an algebraically closed field k of characteristic p, with Newton vector $\nu_0 \in (\mathbb{Q}^n)_+$. Let $\nu \in (\mathbb{Q}^n)_+$ such that ν satisfies the integrality condition of Corollary 1.3 and with $\nu_0 \leqslant \nu \leqslant (1^d, 0^{n-d})$. Then there exists a p-divisible group X over $k[[t]]$ with special fiber X_0 and with Newton vector of the generic fiber of X equal to ν.* □

Using the Serre-Tate theorem, Theorem 5.4 implies that property (i) holds in Theorem 5.3. Using Honda-Tate theory one shows that the μ-basic locus of $\mathcal{M}$ is non-empty, compare [Z4]. Therefore as in the proof of Theorem 3.3, Theorem 5.4 implies that $\mathcal{M}_b$ is non-empty for all $b \in B(G, \mu)$, and the purity theorem allows one now to deduce also property (ii) in Theorem 5.3 from Theorem 5.4. □

We mention that when F is an imaginary quadratic field such that $p \neq 2$ is inert, Bültel and Wedhorn [BW] have proved (i) and (ii) of Conjecture 5.2, provided that the signature of the skew-hermitian form $\langle\ ,\ \rangle$ on V is of the form $(n-1, 1)$. On the other hand, the conjecture seems to be open even for such classical moduli spaces as the Hilbert-Blumenthal varieties.

REFERENCES

[AB] M.F. ATIYAH & R. BOTT – The Yang-Mills equations over Riemann surfaces, *Philos. Trans. Roy. Soc. London Ser. A* **308** (1983), p. 523–615.

[BW] O. BÜLTEL & T. WEDHORN – Congruence relations for Shimura varieties associated to some unitary groups, `math.AG/0202026`.

[C1] C.L. CHAI – Every ordinary symplectic isogeny class in positive characteristic is dense in the moduli, *Invent. Math.* **121** (1995), p. 439–479.

[C2] ———, Newton polygons as lattice points, *Amer. J. Math.* **122** (2000), p. 967–990.

[D] M. DEMAZURE – *Lectures on p-divisible groups*, Lecture Notes in Mathematics, vol. 302, Springer-Verlag, Berlin, 1986, Reprint of the 1972 original.

[FC] G. FALTINGS & C.L. CHAI – *Degeneration of abelian varieties*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 22, Springer-Verlag, Berlin, 1990, With an appendix by David Mumford.

[FM] R. FRIEDMAN & J.W. MORGAN – On the converse to a theorem of Atiyah and Bott, `math.AG/0006086`.

[G] A. GROTHENDIECK – *Groupes de Barsotti-Tate et cristaux de Dieudonné*, Sém. Math. Sup., vol. 45, Presses de l'Univ. de Montreal, 1970.

[Ill] L. ILLUSIE – Déformations de Groupes de Barsotti-Tate, in *Séminaire sur les pinceaux arithmétiques: la conjecture de Mordell*, Astérisque, vol. 127, Soc. Math. France, Paris, 1985, p. 151–198.

[J1] A.J. DE JONG – Barsotti-Tate groups and crystals, in *Proceedings of the International Congress of Mathematicians,* vol. II *(Berlin 1998)*, Doc. Math., Deutsche Math. Verein., 1998, p. 259–265.

[J2] ———, Homomorphisms of Barsotti-Tate groups and crystals in positive characteristic, *Invent. Math.* **134** (1998), p. 301–333.

[JO] A.J. de Jong & F. Oort – Purity of the stratification by Newton polygons, *J. Amer. Math. Soc.* **13** (2000), p. 209–241.

[K] C. Kaiser – *Ein getwistetes fundamentales Lemma für die* GSp_4, Bonner Mathematische Schriften, vol. 303, Universität Bonn, Mathematisches Institut, Bonn, 1997, 71 pp.

[Ka] N. Katz – Slope filtration of F-crystals, in *Journées de géométrie algébrique de Rennes (Rennes 1978)*, Astérisque, vol. 63, Soc. Math. France, Paris, 1979, p. 113–163.

[Kob] N. Koblitz – p-adic variation of the zeta function over families of varieties defined over finite fields, *Compositio Math.* **31** (1975), p. 119–218.

[Kot1] R. Kottwitz – Points on some Shimura varieties over finite fields, *J. Amer. Math. Soc.* **5** (1992), p. 373–444.

[Kot2] ———, Isocrystals with additional structure. II, *Compositio Math.* **109** (1997), p. 255–339.

[KR] S. Kudla & M. Rapoport – Cycles on Siegel threefolds and derivatives of Eisenstein series, *Ann. Sci. École Norm. Sup. (4)* **33** (2000), p. 695–756.

[LO] K.Z. Li & F. Oort – *Moduli of supersingular abelian varieties*, Lecture Notes in Mathematics, vol. 1680, Springer-Verlag, Berlin, 1998, iv+116 pp.

[Me] W. Messing – *The crystals associated to Barsotti-Tate groups: with applications to abelian schemes*, Lecture Notes in Mathematics, vol. 264, Springer-Verlag, Berlin-New York, 1972.

[M] D. Mumford – Bi-extensions of formal groups, in *Algebraic Geometry, Bombay Colloquium 1968*, Tata Inst. Fund. Research and Oxford Univ. Press, 1969, p. 307–322.

[NG] B.C. Ngo & A. Genestier – Alcôves et p-rang des variétés abéliennes, `math.AG/0107223`, 2001.

[No] P. Norman – An algorithm for computing local moduli of abelian varieties, *Ann. of Math. (2)* **101** (1975), p. 499–509.

[NoO] P. Norman & F. Oort – Moduli of abelian varieties, *Ann. of Math. (2)* **112** (1980), p. 413–439.

[O1] F. Oort – Newton polygons and formal groups: conjectures by Manin and Grothendieck, *Ann. of Math. (2)* **152** (2000), p. 183–206.

[O2] ———, Newton polygon strata in the moduli space of abelian varieties, in *Moduli of abelian varieties, (Texel Island 1999)*, Progr. Math., vol. 195, Birkhäuser, Basel, 2001, p. 417–440.

[RR] M. Rapoport & M. Richartz – On the classification and specialization of F-isocrystals with additional structure, *Compositio Math.* **103** (1996), p. 153–181.

[RZ] M. Rapoport & T. Zink – *Period spaces for p-divisible groups*, Annals of Mathematics Studies, vol. 141, Princeton University Press, Princeton, NJ, 1996, xxii+324 pp.

[R] M. Richartz – Klassifikation von selbstdualen Dieudonnégittern in einem dreidimensionalen polarisierten supersingulären Isokristall, Diss., Universität Bonn, 1998.

[W] T. Wedhorn – Ordinariness in good reductions of Shimura varieties of PEL-type, *Ann. Sci. École Norm. Sup. (4)* **32** (1999), p. 575–618.

[Z1] T. Zink – The display of a formal p-divisible group, in *Périodes p-adiques* (I), Astérisque, vol. 278, Soc. Math. France, Paris, 2002, p. 127–248.

[Z2] ———, A Dieudonné theory for p-divisible groups, *Advanced Studies in Pure Mathematics* **30** (2001), p. 139–160.

[Z3] ———, On the slope filtration, *Duke Math. J.* **109** (2001), p. 79–95.

[Z4] ———, Isogenieklassen von Punkten von Shimuramannigfaltigkeiten mit Werten in einem endlichen Körper, *Math. Nachr.* **112** (1983), p. 103–124.

Michael RAPOPORT

Mathematisches Institut
der Universität zu Köln
Weyertal 86-90
D-50931 Köln
Germany
E-mail : rapoport@mi.uni-koeln.de

Séminaire BOURBAKI Juin 2002
54e année, 2001-2002, nO 904, p. 225 à 242

DYNAMIQUES GÉNÉRIQUES : HYPERBOLICITÉ ET TRANSITIVITÉ

par **Christian BONATTI**

INTRODUCTION

Pour structurer la dynamique globale d'un difféomorphisme f d'une variété compacte M, on cherche à caractériser les parties de M qui sont « indécomposables » pour f. La notion la plus naturelle d'indécomposabilité est sans doute la notion d'*ensemble minimal* : un compact K invariant par f est dit *minimal* s'il est minimal pour l'inclusion parmi les compacts invariants, ce qui se caractérise par le fait que toute orbite d'un point de K est dense dans K. Le Théorème de Zorn entraîne que toute orbite contient un minimal dans son adhérence. En fait, il y a en général trop de minimaux pour structurer la dynamique : le classique « Fer à Cheval » de Smale contient un ensemble non dénombrable de ces minimaux.

Ensembles transitifs maximaux

Une notion moins restrictive d'indécomposabilité est la transitivité. Un ensemble compact K invariant par f est dit *transitif* s'il vérifie l'une des conditions équivalentes suivantes :

– il existe un point x de K dont l'orbite positive $\{f^n(x), n \geqslant 0\}$ est dense dans K ;
– pour tout couple U, V d'ouverts de K, il existe $n > 0$ tel que $f^n(U) \cap V$ soit non vide ;
– l'ensemble des points de K dont l'orbite positive et l'orbite négative sont denses dans K contient un G_δ dense dans K. On dira que l'orbite d'un *point générique de* K est dense dans K.

On s'intéresse aux parties transitives qui ne sont pas une sous-partie d'une autre plus grande : outre l'indécomposabilité, on souhaite la maximalité. Si $\{K_i\}_{i\in\mathcal{I}}$ est une famille de compacts f-invariants transitifs, totalement ordonnée pour $\subset$, on vérifie facilement que $\overline{\bigcup_{i\in\mathcal{I}} K_i}$ est encore un compact f-invariant transitif. L'ensemble des

compacts transitifs invariants, ordonné par $\subset$, est donc inductif et le Théorème de Zorn implique que tout compact transitif est contenu dans un *transitif maximal*.

Cette notion a un inconvénient : les transitifs maximaux d'un difféomorphisme ne sont pas toujours disjoints(1). Pour cette raison, on introduit la notion plus forte de *transitif saturé* : un compact K f-invariant transitif est dit *saturé* s'il contient tout compact transitif qui l'intersecte. Deux transitifs saturés sont donc disjoints ou confondus.

Le Théorème de décomposition spectrale de Smale

Cette approche est déjà celle de la théorie de Smale des dynamiques hyperboliques(2) : le théorème de décomposition spectrale de Smale (voir [Sm]) articule la dynamique des difféomorphismes hyperboliques autour des pièces basiques qui, outre l'indécomposabilité et la maximalité (elles sont des transitifs maximaux), possèdent trois autres propriétés fondamentales : l'isolement, la finitude, et la robustesse. Plus précisément :

– La variété admet une *filtration adaptée à* f, c'est-à-dire une famille $\varnothing = M_{k+1} \subset M_k \subset \cdots \subset M_1 = M$ de sous-variétés compactes à bord M_i, de même dimension que M, strictement invariante par f, c'est-à-dire que $f(M_i)$ est contenu dans l'intérieur $\overset{o}{M}_i$ de M_i.

– Considérons l'ensemble Λ_i des points dont l'orbite (positive et négative) reste dans la « tranche » $M_i \setminus \overset{o}{M}_{i+1}$: en formule $\Lambda_i = \bigcap_{n\in\mathbb{Z}} f^n(M_i \setminus \overset{o}{M}_{i+1})$. On dit que Λ_i est *l'ensemble maximal invariant dans* $M_i \setminus \overset{o}{M}_{i+1}$. Alors Λ_i est un compact transitif.

– Il existe un C^1-voisinage $\mathcal{U}_f$ de f tel que pour tout $g \in \mathcal{U}_f$ la filtration M_i est adaptée à g et l'ensemble maximal invariant $\Lambda_i(g)$ de g dans $M_i \setminus \overset{o}{M}_{i+1}$ est un compact transitif.

On sait que les dynamiques hyperboliques ne sont pas denses dans l'ensemble des difféomorphismes de classe C^1 pour les variétés de dimension trois ou plus(3). On aimerait pouvoir donner une description analogue à celle donnée par le théorème de

(1) Exemple : soit A un difféomorphisme d'Anosov du tore T^2. Notons f_0 le difféomorphisme du tore $T^5 = T^2 \times T^2 \times S^1$ défini par $f_0(x, y, t) = (A(x), A(y), t)$. Soit φ une fonction de T^5 dans $[0, +\infty[$ qui est nulle exactement sur $T^4 \times \{1/2\}$ et sur $(T^2 \times \{0\} \cup \{0\} \times T^2) \times \{0\}$. Finalement soit $f = g \circ f_0$ où g est le temps 1 du flot de $\varphi \frac{\partial}{\partial t}$.
Les ensembles $T^2 \times \{0\} \times \{0\}$ et $\{0\} \times T^2 \times \{0\}$ sont des transitifs maximaux de f qui s'intersectent au point $0 \in T^5$.

(2) Dans cet exposé, j'appellerai *difféomorphismes hyperboliques* les difféomorphismes qui vérifient l'axiome A (hyperbolicité de l'ensemble $\Omega(f)$ des points non errants et densité dans $\Omega(f)$ de l'ensemble des points périodiques) et la transversalité forte (transversalité de toute intersection entre une variété stable et une variété instable). Rappelons que l'hyperbolicité est équivalente à la C^1-stabilité stucturelle, d'après les théorèmes de Robbin, Robinson et Mañé.

(3) En dimension 2, les difféomorphismes hyperboliques ne sont pas C^1-denses si l'on exige la transversalité forte (nécessaire à la stabilité structurelle). La C^1-densité des difféomorphismes axiome A

décomposition spectrale, pour un ensemble aussi large que possible de dynamiques, si possible une partie dense ou générique de $\mathrm{Diff}^1(M)$.

Pièces élémentaires de dynamiques

Plusieurs notions apparaissent comme les candidats naturels pour substituer les pièces basiques des dynamiques hyperboliques, suivant que l'on désire garder telle ou telle de leurs propriétés, la propriété indispensable restant la transitivité (indécomposabilité) :

– (Indécomposabilité, maximalité) Les ensembles maximaux transitifs et transitifs saturés, le problème étant de les caractériser.

– (Indécomposabilité, isolement, robustesse) L. Díaz, E. Pujals et R. Ures ont défini dans ce but la notion *d'ensemble robustement transitif* : si U est un ouvert de M et $\mathcal{U}$ un C^1-voisinage de f tel que, pour tout $g \in \mathcal{U}$ l'ensemble maximal invariant $\Lambda_g = \bigcap_{n\in\mathbb{Z}} g^n(\bar{U})$ est un compact transitif contenu dans U, on dit que Λ_f est robustement transitif.

Si U est un *voisinage filtrant* de f, (i.e. $U = M_i \setminus \overset{o}{M}_{i+1}$ pour une filtration M_i adaptée à f), l'ensemble Λ_f est naturellement aussi maximal.

– Finalement, de façon moins conceptuelle mais plus constructive, à toute orbite périodique selle p on associe sa *classe homocline* $H(p, f)$ comme étant l'adhérence des points d'intersection transverse de ses variétés invariantes (stable et instable). Les classes homoclines sont des ensembles transitifs canoniquement associés aux points périodiques hyperboliques de type selle.

Dans le cas des dynamiques hyperboliques ces trois notions coïncident : les pièces basiques du théorème de décomposition spectrale sont à la fois les classes homoclines, les ensembles maximaux transitifs (et saturés), et sont robustement transitifs. Dans [BD], nous posions le problème suivant :

Pour tout difféomorphisme générique[(4)] f*, les classes homoclines sont-elles les maximaux transitifs de* f *?*

Ensemble maximaux transitifs des dynamiques génériques

Récemment, des idées de R. Mañé ([Ma1]) et des lemmes de perturbations en topologie C^1 (Lemme de connexion de Hayashi et ses généralisations par Hayashi, Xia et Wen et Arnaud) ont permis de répondre complètement au problème ci-dessus, et

sans cycles reste un problème ouvert. Cependant l'ensemble des difféomorphismes axiome A n'est pas dense en topologie C^2 (Théorème de Newhouse voir [N, PT]).

(4) Dans tout ce texte nous utiliserons des abus de langage du type « *tout difféomorphisme générique de* M *vérifie une propriété* $\mathcal{P}$ » pour dire « *il existe une partie* $\mathcal{R} \subset \mathrm{Diff}^1(M)$ *résiduelle (i.e. contenant un* G_δ *dense) telle que* $\mathcal{P}$ *est vraie pour tout* f *de* $\mathcal{R}$. », sauf si nous voulons expliciter le résiduel $\mathcal{R}$ où $\mathcal{P}$ est vérifiée.

de préciser les liens entre une forme affaiblie d'hyperbolicité (*décomposition dominée*, encore appelée *hyperbolicité projective*) et la transitivité robuste.

Voici une présentation cohérente de ces résultats qui donne un panorama général de la dynamique des difféomorphismes C^1-génériques[(5)] :

THÉORÈME 1

– [Ar, CMP] *Il existe un ensemble résiduel* $\mathcal{R} \subset \mathrm{Diff}^1(M)$ *de difféomorphismes* f *pour lesquels toute classe homocline* $H(p,f)$, $p \in \mathrm{Per}(f)$ *est un ensemble maximal transitif, saturé (en particulier deux classes homoclines sont disjointes ou confondues).*

– [Ab] *Le cardinal* $n(f) \in \mathbb{N} \cup \{\infty\}$ *de l'ensemble* $\{H(p,f),\, p \in \mathrm{Per}(f)\}$ *est localement constant sur* $\mathcal{R}$ *(tout* $f \in \mathcal{R}$ *possède un voisinage* U_f *tel que* $n(f) = n(g)$ *pour tout* $g \in U_f \cap \mathcal{R}$*).*

Nous dirons qu'un difféomorphisme $f \in \mathcal{R}$ est apprivoisé si $n(f) \in \mathbb{N}$ et qu'il est sauvage si $n(f) = +\infty$. Le théorème de décomposition spectrale de Smale se généralise aux dynamiques apprivoisées, justifiant ainsi leur nom :

THÉORÈME 2

– [Ab] *Si* $f \in \mathcal{R}$ *est apprivoisé alors, quitte à restreindre* U_f*, il existe une filtration* $\varnothing = M_{k+1} \subset M_k \subset \cdots \subset M_1 = M$*,* $k = n(f)$*, adaptée à tous les difféomorphismes* $g \in U_f$*. De plus, pour tout* $g \in U_f \cap \mathcal{R}$*, l'ensemble maximal invariant* $\Lambda_i(g)$ *du voisinage filtrant* $M_i \setminus \overset{o}{M}_{i+1}$ *est un transitif saturé qui est une classe homocline.*

– [CM] *L'union des bassins d'attractions des* $\Lambda_i(g)$ *qui sont des attracteurs topologiques est dense dans* M*.*

Ce résultat amène naturellement à affaiblir la notion d'ensemble robustement transitif : on parlera d'ensemble *génériquement transitif*, si la transitivité n'est obtenue que pour les difféomorphismes génériques au voisinage de f. Les ensembles génériquement transitifs possèdent une forme affaiblie d'hyperbolicité (voir l'appendice A pour les définitions des formes affaiblies d'hyperbolicité).

THÉORÈME 3 ([Ma1, DPU, BDP, Ab]). — *Tout ensemble génériquement transitif* K *est* hyperbolique en volume*, c'est-à-dire qu'il possède une décomposition dominée* $TM_K = E_1 \oplus \cdots \oplus E_\ell$ *telle que la différentielle* f_* *contracte uniformément le volume dans* E_1 *et dilate uniformément le volume dans* E_ℓ*.*

Il n'existe pour l'instant aucun exemple connu d'ensemble génériquement transitif qui ne soit pas robustement transitif, et nous conjecturons que ces deux notions coïncident, sur un ouvert dense de difféomorphismes.

(5) Cette présentation est un assemblage de travaux de, par ordre alphabétique : Abdenur, Arnaud, Bonatti, Carballo, Díaz, Hayashi, Morales, Pacifico, Pujals, Rocha, Ures, Viana, Wen, Xia... et Mañé, omniprésent dans l'esprit de ces travaux.

On espère que l'hyperbolicité en volume, nécessaire à la transitivité robuste, permettra de donner une description de ces ensembles génériquement transitifs. À l'opposé, le Théorème 3 permet de construire des exemples de dynamiques sauvages. Il existe en effet des mécanismes maintenant assez bien compris, permettant de construire des classes homoclines n'ayant, de façon robuste, aucune décomposition dominée. Ceci nous a permis de montrer :

THÉORÈME 4 ([BD2]). — *Pour toute variété compacte M de dimension $\geqslant 3$, il existe un ouvert $\mathcal{V}$ de $\mathrm{Diff}^1(M)$ et une partie résiduelle $\mathcal{W}$ de $\mathcal{V}$, telle que tout $f \in \mathcal{W}$ admet un ensemble non dénombrable d'ensembles transitifs saturés qui ne contiennent aucune orbite périodique.*

Structure de cet exposé

Dans la suite de cet exposé je chercherai à donner une intuition de ces résultats et les idées qui ont permis de les démontrer, sans chercher à résumer les preuves, parfois longues et techniques.

– La première partie construit des exemples de difféomorphismes robustement transitifs non hyperboliques. Ces exemples ont eu un grand rôle dans la théorie, en permettant de comprendre jusqu'où on pouvait relâcher l'hyperbolicité.

– La seconde présente les idées de Mañé qui permettent de montrer qu'un peu d'hyperbolicité est nécessaire à la transitivité robuste.

– La troisième met en parallèle les récents théorèmes de connexions et leurs conséquences sur les classes homoclines des difféomorphismes génériques.

– La quatrième décrit un exemple à l'opposé de ceux de la première partie : l'absence robuste de toute forme d'hyperbolicité casse la dynamique en une infinité de « petites dynamiques » indépendantes.

1. EXEMPLES DE DIFFÉOMORPHISMES ROBUSTEMENT TRANSITIFS

On dit qu'un difféomorphisme f d'une variété compacte M est *robustement transitif* s'il appartient à l'intérieur C^1 de l'ensemble des difféomorphismes transitifs ; en d'autres termes et suivant la terminologie introduite ci-dessus, f est robustement transitif si la variété M est elle-même un ensemble robustement transitif.

Plus généralement, nous dirons qu'une propriété de f est *robuste* si elle est valide sur un C^1-voisinage de f.

1.1. Exemples classiques

Voyons d'abord deux exemples très différents de difféomorphismes transitifs, sur le tore T^2 :

– Si $a = (a_1, a_2)$ est un vecteur de $\mathbb{R}^2$ dont les coordonnées a_1 et a_2 sont irrationnelles et indépendantes sur $\mathbb{Q}$, la translation $(x, y) \mapsto (x + a_1, y + a_2)$ induit un difféomorphisme du tore T^2 dont toutes les orbites sont denses.

– Si $A \in SL(2, \mathbb{Z})$ est une matrice de trace différente de ± 2 (par exemple la célèbre matrice $\left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$) elle induit sur le tore T^2 un difféomorphisme transitif, appelé *difféomorphisme d'Anosov*, possédant une infinité d'orbites périodiques.

La dynamique d'une translation irrationnelle peut paraître plus parfaite : elle est minimale (toutes les orbites sont denses) et uniquement ergodique. Elle est cependant très fragile : la transitivité ne résiste pas aux perturbations de la dynamique.

Voici une démonstration rapide et élémentaire (n'utilisant pas la puissance de la théorie hyperbolique de Smale) du résultat très classique suivant :

THÉORÈME 5. — *Le difféomorphisme d'Anosov $A \in \mathrm{Diff}^1(T^2)$ est robustement transitif.*

Démonstration. — Soit E^s et E^u les directions propres de la matrice A, associées aux valeurs propres λ^{-1} et λ de A, de module respectivement < 1 et > 1. On écrit tout vecteur de la forme $v = v^s + v^u$ et on note $\mathcal{C}^s$ et $\mathcal{C}^u$ les cônes définis par

$$v \in \mathcal{C}^u \quad \text{si} \quad \|v^u\| \geqslant 2\|v^s\| \quad \text{et} \quad v \in \mathcal{C}^s \quad \text{si} \quad \|v^s\| \geqslant 2\|v^u\|.$$

Voici deux remarques simples :

(1) Il existe $L > 0$ tel que, si γ^s et γ^u sont des segments immergés dans T^2, tangents respectivement aux cônes stable et instable, et chacun de longueur supérieure à L, alors $\gamma^s \cap \gamma^u \neq \varnothing$.

(2) Il existe un C^1-voisinage $\mathcal{U}$ de A tel que tout difféomorphisme $f \in \mathcal{U}$ vérifie :

– pour tout vecteur $v \in \mathcal{C}^u \subset T_p(T^2)$ le vecteur $w = f_*(v)$ appartient encore au cône $\mathcal{C}^u$, et de plus $\|w^u\| \geqslant \sqrt{\lambda}\|v^u\|$;

– pour tout vecteur $v \in \mathcal{C}^s \subset T_p(T^2)$ le vecteur $w = f_*^{-1}(v)$ appartient encore au cône $\mathcal{C}^s$, et de plus $\|w^u\| \geqslant \sqrt{\lambda}\|v^u\|$.

La seconde de ces remarques montre que tout ouvert O_1 de T^2 possède des itérés positifs $f^n(O_1)$, $n > 0$, contenant des segments de longueur L tangents au cône instable. De la même façon les itérés négatifs grands d'un ouvert O_2 contiennent des segments de longueur L tangents au cône stable. La première remarque assure alors que $f^{n+m}(O_1) \cap O_2$ est non vide, prouvant la transitivité de f. □

1.2. Difféomorphismes robustement transitifs non hyperboliques

Voici quelques remarques simples :

– Si un difféomorphisme transitif possède deux points périodiques hyperboliques d'indices de Morse (dimension de la variété instable) différents, il ne peut pas être hyperbolique.

– Dans le même esprit : pour qu'un difféomorphisme transitif d'une variété compacte de dimension 4 ne soit pas partiellement hyperbolique (i.e. ne laisse invariant aucun sous-fibré uniformément contractant ou dilatant) il suffit qu'il possède trois points fixes de type selle : un point d'indice 1, un point d'indice 3, et finalement un point d'indice 2 ayant deux valeurs propres complexes (non réelles) (une de module > 1 et une de module < 1).

– Pour qu'un difféomorphisme f soit transitif, il suffit qu'il possède un point périodique hyperbolique x de type selle dont les variétés stable et instable sont denses dans la variété M : en effet, si U et V sont deux ouverts de M, U contient alors un disque transverse à la variété stable de x, et le classique « λ-Lemma » implique alors que les itérés positifs de ce disque s'accumulent sur toute la variété instable de x, et rencontrent donc V.

– Soit f un difféomorphisme d'une variété compacte M, partiellement hyperbolique, possédant un fibré uniformément dilatant. Soit p un point périodique hyperbolique de type selle de f. On suppose que toute orbite de feuille du feuilletage instable fort $\mathcal{F}^u$ coupe la variété stable $W^s(p)$ de l'orbite de p. Alors la variété stable de p est robustement dense.

Démonstration. — En effet cette propriété implique qu'il existe L tel que le disque de rayon L centré en p dans $W^s(p)$ coupe tout disque de rayon L d'une feuille de $\mathcal{F}^u$. La famille des disques instables de rayon $2L$ est une famille compacte, variant continûment avec le difféomorphisme. On en déduit que, pour tout difféomorphisme g suffisamment C^1-proche de f, le disque de rayon $2L$ centré en p de $W^s(p,g)$ coupe tout disque de $\mathcal{F}^u(g)$ de rayon $2L$. Par invariance par g^{-1} on en déduit que $W^s(p,g)$ coupe tout disque instable arbitrairement petit, et donc est dense. □

Ces quelques idées permettent de reconstruire simplement des exemples de difféomorphismes robustement transitifs non hyperboliques, mais partiellement hyperboliques, ayant à la fois un fibré uniformément contractant et un fibré uniformément dilatant. Voici en quelques mots l'exemple construit par M. Shub dans [Sh] :

Considérons une famille différentiable $f_x\colon T^2 \to T^2$ de difféomorphismes de T^2, dont le paramètre x varie dans le tore T^2. On suppose que, pour $x = 0 \in T^2$, le difféomorphisme f_0 est un difféomorphisme d'Anosov.

Soit $A\colon T^2 \to T^2$ un difféomorphisme d'Anosov admettant 0 comme point fixe. Alors pour tout n assez grand, le difféomorphisme $F\colon T^2 \times T^2 \to T^2 \times T^2$ défini par $F(x,y) = (A^n(x), f_x(y))$ est partiellement hyperbolique, les feuilletages stable et instable forts de F se projetant sur les feuilletages stable et instable du difféomorphisme A.

THÉORÈME 6. — *Le difféomorphisme F défini ci-dessus est robustement transitif.*

Démonstration. — Soit 0 un point fixe du difféomorphisme f_0, si bien que $p_f = (0,0)$ est un point fixe de F. Nous allons montrer que sa variété stable et sa variété instable sont robustement denses dans T^4, ce qui conclut, d'après la 3ème des remarques ci-dessus.

Pour cela, on remarque que la fibre $T_F = \{0\} \times T^2$ est une variété compacte, F-invariante et normalement hyperbolique. Un théorème de Hirsch, Pugh et Shub [HPS] montre que cette variété invariante a une continuation G-invariante T_G, pour tout difféomorphisme G suffisamment C^1-proche de F. On peut alors considérer les variétés stable et instable du tore T_G, notées $W^s(T_G)$ et $W^u(T_G)$, unions des feuilles stables et instables fortes, respectivement, passant par T_G. On remarque que $W^s(T_F)$ coupe toutes les feuilles instables fortes, ce qui prouve qu'elle est robustement dense dans le tore T^4, d'après notre 4ème remarque ci-dessus. De même $W^u(T_F)$ est robustement dense dans T^4.

Pour tout G suffisamment C^1-proche de F, la restriction de G à T_G est un difféomorphisme d'Anosov conjugué à f_0 et nous noterons p_G le point fixe de G dans T_G qui est la continuation de p_F. On remarque que la variété stable de p_G dans T_G est dense dans T_G et on en déduit que l'union des feuilles stables fortes s'appuyant sur les points de la variété stable de p_G dans T_G est dense dans $W^s(T_G)$ et donc dans T^4. La variété stable de p_G est donc robustement dense dans T^4 et il en est de même pour sa variété instable, ce qui prouve la transitivité robuste. □

Pour construire un exemple où F est un difféomorphisme robustement transitif et robustement non hyperbolique, il suffit que, pour un point fixe $x \neq 0$ de A, l'application f_x ait un point fixe y attracteur ou répulseur : le point (x,y) est alors un point fixe selle d'indice 1 ou 3 de F, et la première remarque montre que F n'est pas hyperbolique.

Ce type d'argument utilise fortement l'existence d'un feuilletage stable fort et d'un feuilletage instable fort, et ne peut pas servir pour construire des difféomorphismes robustement transitifs non partiellement hyperboliques. Un autre type d'argument a été utilisé par Mañé ([Ma2]) pour construire un difféomorphisme robustement transitif, robustement non hyperbolique de T^3 : il considère un difféomorphisme d'Anosov linéaire de T^3 ayant trois valeurs propres réelles différentes, $\lambda_1 < 1 < \lambda_2 < \lambda_3$. Il fait alors une perturbation de cette application, à support dans une toute petite boule B,

et qui conserve une décomposition invariante en trois fibrés unidimensionnels, un instable fort, un stable fort et un central. En utilisant l'expansion uniforme dans la direction instable forte, il montre que tout ouvert contient au moins un point dont les itérés positifs ne passent qu'un nombre fini de fois dans B. Comme en dehors de B la dynamique coïncide avec la dynamique du difféomorphisme d'Anosov initial, les itérés d'un petit disque centré en x dans la direction instable vont contenir un disque de rayon fixé, qui coupera tout segment stable fort de longueur assez grande. La transitivité robuste est obtenue comme dans la preuve, donnée ci-dessus, de la transitivité robuste des difféomorphismes d'Anosov du tore (Théorème 5). La non hyperbolicité est obtenue en créant, dans B, un point fixe d'indice de Morse égal à 1.

Cet argument a été adapté par M. Viana et moi-même (voir [BV]) pour construire sur le tore T^4 des difféomorphismes robustement transitifs qui n'admettent aucun sous-fibré invariant uniformémement contractant ou dilatant : les difféomorphismes que nous construisons admettent une unique décomposition dominée $E^1 \oplus E^2$ (E_i de dimension 2) telle que l'aire est uniformément contractée dans E^1 et uniformément dilaté dans E^2. Comme dans l'exemple de Mañé, nous les construisons C^1-proches d'un difféomorphisme d'Anosov en dehors de deux petites boules B_1 et B_2. Nous adaptons le raisonnement de Mañé en utilisant l'expansion uniforme du volume au lieu de la dilatation uniforme, pour montrer que tout ouvert contient un point dont seul un nombre fini d'itérés positifs passent dans les boules B_i et un point dont seul un nombre fini d'itérés négatifs passent dans les boules B_i. On obtient alors la transitivité robuste comme dans l'exemple de Mañé.

L'absence de fibré hyperbolique invariant est alors donnée par la présence d'un point fixe d'indice 1 dans B_1, d'un point d'indice 3 dans B_2 et d'un point ayant des valeurs propres complexes (comme dans les remarques).

2. DE L'HYPERBOLICITÉ EN VOLUME POUR LES DYNAMIQUES ROBUSTEMENT TRANSITIVES

2.1. En dimension 2 : l'argument de Mañé

Dans [Ma1], R. Mañé montre que tout difféomorphisme robustement transitif d'une surface compacte est un difféomorphisme d'Anosov du tore T^2. Sa preuve se décompose en deux étapes :

PROPOSITION 2.1. — *Soit f un difféomorphisme d'une surface compacte S et soit $\mathcal{E} = \{x_i\}$ un ensemble d'orbites périodiques de type selle de f. Alors l'une des affirmations suivantes est vérifiée :*

(1) *ou bien la décomposition naturelle de TM au-dessus de $\mathcal{E}$ donnée par les directions stable $E^s(x)$ et instable $E^u(x)$ des points x de $\mathcal{E}$ est dominée ;*

(2) *ou bien, pour tout $\varepsilon > 0$, il existe un point $x \in \mathcal{E}$ et une ε-C^1-perturbation g de f à support sur un voisinage arbitrairement petit de l'orbite de x de façon que x soit un puits ou une source de g.*

PROPOSITION 2.2. — *Supposons à présent que $\mathcal{E}$ possède une décomposition dominée $TM_{\mathcal{E}} = E \oplus F$ mais que $\bar{\mathcal{E}}$ n'est pas hyperbolique. Alors il existe g aussi C^1-proche de f que l'on veut tel que g possède un puits ou une source contenue dans un voisinage arbitrairement petit de $\bar{\mathcal{E}}$.*

Les preuves de ces deux étapes sont de natures très différentes :

Pour la première étape, un lemme de Franks permet réaliser toute perturbation de la différentielle de f au-dessus d'un ensemble fini, sans modifier f sur cet ensemble, par une C^1-perturbation de f à support dans un voisinage arbitraire de cet ensemble. Mañé utilise alors deux idées simples (je supposerai ici que f préserve l'orientation) :

– Si une matrice hyperbolique $A \in GL_+(2, \mathbb{R})$ a ses directions propres qui font entre elles un angle inférieur à ε, alors il existe $\alpha \in [-\varepsilon, \varepsilon]$ tel que $R_\alpha \circ A$ possède une valeur propre complexe, où R_α est la rotation d'angle α. Remarquons qu'une orbite périodique en dimension 2 qui a une valeur propre complexe (non réelle) est (ou peut être perturbée en) un puits ou une source.

– Si la décomposition $E^s \oplus E^u$ n'est pas dominée, c'est que certaines orbites tardent à voir leur hyperbolicité : pour tout $n > 0$, il existe $x \in \mathcal{E}$ tel que f^n_* ne dilate pas les vecteurs de $E^u(x)$ deux fois plus que ceux de $E^s(x)$. Mañé montre que l'on peut alors perturber la différentielle de f le long de l'orbite de x, de façon que l'angle entre les nouvelles directions stable et instable au point $f^n(x)$ soit très petit. Le premier item permet alors de créer un puits ou une source, par une nouvelle perturbation.

Pour la deuxième étape, Mañé a trouvé une démonstration très astucieuse et profonde. Bien qu'elle ait certains côtés un peu techniques, je ne résiste pas au plaisir de vous la présenter :

Si E n'est pas uniformément contractant, c'est qu'il existe des points de $\mathcal{E}$ qui tardent à voir la contraction. Mañé construit alors une mesure μ à support sur $\bar{\mathcal{E}}$ telle que l'intégrale du logarithme de la dérivée de f dans la direction E est positive ou nulle. Cette propriété reste vraie pour au moins une des composantes ergodiques de μ : on peut donc supposer μ ergodique. On en déduit que μ-presque tout point a un exposant de Lyapunov positif dans la direction E.

Mañé prouve alors un raffinement important du Lemme de Fermeture de Pugh ; il montre que μ-presque tout point x possède une infinité de temps $n > 0$ tels que l'on peut « fermer l'orbite de x au temps n » :

il existe g C^1-proche de f telle que x est périodique pour g de période n et $g^i(x)$ et $f^i(x)$ restent très proches l'un de l'autre pour tout $i \in \{1, \dots, n\}$.

C'est le Lemme de Fermeture Ergodique. En fermant ainsi l'orbite d'un point ayant un exposant positif ou nul dans la direction de E, il crée un nouveau point périodique

ayant une valeur propre proche de 1 (en valeur absolue) dans la direction de E. Une nouvelle perturbation rend cette valeur propre strictement plus grande que 1 : le point devient donc une source, ce qui conclut la preuve.

2.2. Généralisation en dimension plus grande

Si l'on applique directement les techniques de Mañé en dimension plus grande, on obtient que, en l'absence de décomposition dominée, il existe des points périodiques dont on va pouvoir faire changer l'indice de Morse par une petite perturbation. Les exemples de difféomorphismes robustement transitifs non hyperboliques montrent que ceci n'est pas suffisant pour briser la transitivité. Pour obtenir un puits ou une source (qui, eux, brisent la transitivité), [DPU] et [BDP] vont utiliser une propriété dynamique supplémentaire de l'ensemble des orbites périodiques homocliniquement liées à un point périodique p : si x et y sont de telles orbites, il existe des orbites périodiques z qui accompagnent x un temps arbitraire, puis s'approchent de y un temps uniformément borné et accompagnent y en un temps arbitraire, etc. Cette propriété permet, en quelque sorte, de multiplier les différentielles de f (à la période) des points x et y.

Cette propriété des classes homoclines a permis de montrer :

THÉORÈME 7 ([BDP]). — *Soit f un difféomorphisme d'une variété compacte et soit p un point périodique hyperbolique de type selle f. On suppose que la classe homocline $H(p,f)$ n'admet aucune décomposition dominée. Alors, pour tout $\varepsilon > 0$, il existe un point périodique x de f homocliniquement lié à p, ayant la propriété suivante :*

Pour tout voisinage U de l'orbite de x, il existe un difféomorphisme g, ε-C^1-proche de f, coïncidant avec f hors de U et le long de l'orbite de x, tel que la différentielle $g^n_*(x)$ soit une homothétie, où n est la période de x.

Ce résultat montre l'existence d'une décomposition dominée pour toute dynamique robustement transitive. La démonstration de l'hyperbolicité en volume (dilatation et contraction uniforme du volume dans les fibrés extrémaux de la décomposition dominée la plus fine) est alors une adaptation facile de l'argument de Mañé en dimension 2.

3. CLASSES HOMOCLINES DES DIFFÉOMORPHISMES GÉNÉRIQUES

3.1. Classes homoclines des points périodiques selles

Considérons un difféomorphisme f d'une variété compacte M et p un point périodique hyperbolique de f. On définit la classe homocline $H(p,f)$ de p comme l'adhérence des points d'intersection transverse des variétés stable et instable de p. En formule :

$$H(p,f) = \overline{W^s(p) \pitchfork W^u(p)}.$$

C'est bien sûr un compact invariant par f. Rappelons la démonstration du fait classique suivant :

LEMME 3.1. — *La classe homocline $H(p,f)$ est un ensemble transitif.*

Démonstration. — Considérons deux ouverts U, V de M rencontrant K. L'ouvert U contient un point homocline transverse et contient donc un disque instable D^u transverse à $W^s(p)$ en un point $x \in H(p,f) \cap U$, et de même V contient un disque D^s tranverse à $W^u(p)$ en un point $y \in H(p,f) \cap V$. Le λ-lemma implique que les itérés $f^n(D^u)$ convergent quand $n \to +\infty$ vers toute la variété instable de p, et ceci en topologie C^1. On en déduit qu'il existe n tel que $f^n(D^u)$ coupe transversalement D^s en un point proche de y. Ce point est un point homocline transverse et donc appartient à $H(p,f)$. On vient de montrer que $f^n(H(p,f) \cap U)$ rencontre $H(p,f) \cap V$, ce qui montre la transitivité de $H(p,f)$. □

Voici une autre façon de voir la classe homocline : on dit que deux points périodiques selles x et y sont *homocliniquement liés* si la variété stable et la variété instable de x possèdent chacune une intersection transverse avec la variété instable et la variété stable, respectivement, de y. Ceci implique que x et y ont même indice de Morse. Le λ-lemma permet de montrer simplement que cette relation est une relation d'équivalence.

La classe homocline de p coïncide avec l'adhérence de l'ensemble des points périodiques de même indice que p et homocliniquement liés à p.

On a ainsi associé de façon canonique un ensemble transitif à toute orbite périodique hyperbolique.

3.2. Difféomorphismes génériques

Les propriétés des difféomorphismes génériques sont intimement liées aux théorèmes de perturbations, et la raison essentielle pour laquelle les propriétés des difféomorphismes génériques sont pour la topologie C^1 est que le lemme de fermeture de Pugh, le lemme de connexion de Hayashi et ses généralisations, et enfin un lemme de Franks utilisent tous la topologie C^1 (voir [Pu, Ha, Ar, WX, F]).

L'un des lemmes de perturbations les plus célèbres est le lemme de fermeture de C. Pugh :

THÉORÈME ([Pu]). — *Si f est un difféomorphisme d'une variété compacte et si x est un point non errant de f, alors il existe g arbitrairement C^1-proche de f tel que x est périodique pour g.*

Le lemme de fermeture et le théorème de Kupka-Smale permettent de montrer :

COROLLAIRE. — *Pour tout difféomorphisme générique, l'ensemble $\Omega(f)$ des points non errants de f coïncide avec l'adhérence de l'ensemble des points périodiques de f, et de plus ceux-ci sont tous hyperboliques, et toute intersection entre leurs variétés invariantes est transverse.*

Longtemps conjecturé, le lemme de connexion d'Hayashi a permis de reprendre l'étude des dynamiques génériques :

THÉORÈME ([Ha]). — *Soient p et q deux points périodiques hyperboliques et supposons qu'il existe une suite de points x_n convergeant vers un point x de $W^u_{\mathrm{loc}}(p)$ et des itérés $y_n = f^{m(n)}(x_n)$, $m(n) \geqslant 0$, convergeant vers un point $y \in W^s_{\mathrm{loc}}(q)$.*

Il existe alors g, arbitrairement C^1-proche de f, tel que les points x et y soient sur une même orbite hétérocline des points p et q ; en d'autres termes :

$$x \in W^u_{\mathrm{loc}}(p, g),\ y \in W^s_{\mathrm{loc}}(q, g) \text{ et il existe } n > 0 \text{ tel que } g^n(x) = y.$$

Si p et q appartiennent à un même ensemble transitif K, les suites de points x_n et y_n (de l'énoncé ci-dessus) sont obtenues naturellement à l'aide d'une orbite positive dense dans K. Ceci a permis de montrer :

COROLLAIRE 3.2 ([BD]). — *Pour tout difféomorphisme générique, deux points périodiques p et q appartiennent à un même ensemble transitif si et seulement si leurs classes homoclines sont égales.*

Voici maintenant une généralisation du lemme de connexion d'Hayashi que l'on peut trouver dans [WX] ou dans [Ar] :

THÉORÈME 8. — *Soit x un point non périodique de f, et soit $\varepsilon > 0$ fixé. Il existe $N > 0$ tel que, pour tout voisinage U du segment d'orbite $\{f^{-N}(x), \dots, f^N(x)\}$, il existe un voisinage V de x avec la propriété suivante :*

Soient p et q deux points hors de U tels qu'il existe n, m deux entiers positifs tels que $f^n(p) \in V$ et $f^{-m}(q) \in V$. Il existe un difféomorphisme g, ε-C^1-proche de f, coïncidant avec f hors de U et il existe $k > 0$ tel que $g^k(p) = q$.

Voici un exemple d'application directe qui est sans doute de lecture plus facile :

THÉORÈME 9. — *Soient p_0 et q_0 deux points périodiques de f tels que $\overline{W^s(p_0, f)} \cap \overline{W^u(q_0, f)}$ contient un point non périodique x. Alors il existe g aussi C^1-proche que l'on veut de f tel que $x \in W^s(p_0, f) \cap W^u(q_0, f)$.*

Cette généralisation a permis à M.-C. Arnaud de montrer :

COROLLAIRE 3.3 ([Ar]). — *Pour tout difféomorphisme f générique et tout $p \in \mathrm{Per}(f)$ on a :*

$$H(p, f) = \overline{W^u(p)} \cap \overline{W^s(q)}.$$

Une variante de cette généralisation a permis à C. Carballo, C. Morales et M.J. Pacifico de montrer :

COROLLAIRE 3.4 ([CMP]). — *Pour tout difféomorphisme f générique et tout point périodique p de f, $\overline{W^u(p)}$ est stable au sens de Lyapunov* [(6)] *et $\overline{W^s(p)}$ est stable au sens de Lyapunov pour f^{-1}.*

De ces deux corollaires on déduit le corollaire suivant qui est la première partie du Théorème 1 :

COROLLAIRE 3.5 ([CMP]). — *Pour tout difféomorphisme f générique et tout $p \in \mathrm{Per}(f)$, la classe homocline $H(p,f)$ est un ensemble transitif saturé.*

Démonstration. — Soit K un compact transitif rencontrant $H(p,f)$, et soit x une orbite positivement et négativement dense dans K. L'orbite de x s'approche donc arbitrairement près d'un point de $\overline{W^u(p)}$. Comme $\overline{W^u(p)}$ est stable au sens de Lyapunov, on en déduit que K est inclus dans des voisinages arbitrairement petits de $\overline{W^u(p)}$, et donc finalement est inclus dans $\overline{W^u(p)}$. De même K est inclus dans $\overline{W^s(p)}$, et donc dans $\overline{W^u(p)} \cap \overline{W^s(p)}$. D'après le Corollaire 3.3, K est donc inclus dans $H(p,f)$. □

Voici finalement quelques remarques simples qui sont à la base de la preuve de la seconde partie du Théorème 1 :

– La classe homocline $H(p,f)$ est semi-continue inférieurement pour la distance de Hausdorff : en effet les disques compacts de $W^s(p,f)$ ou de $W^u(p,f)$ varient continûment avec f et les intersections transverses entre ces disques compacts varient donc elles aussi de façon continue avec f.

– Le même argument montre que les adhérences $\overline{W^s(p,f)}$ et $\overline{W^u(p,f)}$ varient de façon semi-continue inférieure avec f.

– En conséquence des items précédents, pour f générique les classes homoclines $H(p,f)$ et les adhérences $\overline{W^s(p,f)}$ et $\overline{W^u(p,f)}$ varient continûment avec f.

4. EXEMPLES DE DYNAMIQUES SAUVAGES

On considère un ouvert $\mathcal{U} \subset \mathrm{Diff}^1(M)$ tel qu'il existe un point périodique p_f variant continûment avec $f \in \mathcal{U}$ ayant les propriété suivantes :

– Pour tout $f \in \mathcal{U}$ la classe homocline de p_f contient deux points d'indice différent et possédant chacun une valeur propre complexe (non réelle) (contractante pour l'un et dilatante pour l'autre).

– Pour tout $f \in \mathcal{U}$ il existe deux points périodiques de même indice que p_f et homocliniquement liés à p_f tels que l'un soit de jacobien > 1 et l'autre de jacobien < 1.

(6) Un compact invariant K est dit stable au sens de Lyapunov si, pour tout voisinage U de K, il existe un voisinage V de K dont les orbites positives sont incluses dans U : de façon équivalente, K possède une base de voisinages positivement invariants.

Le premier item implique que la classe homocline $H(p_f, f)$ n'a aucune décomposition dominée. Le Théorème 7 montre que $H(p_f, f)$ possède des points périodiques dont la dérivée peut être perturbée de façon à devenir une homothétie. Le deuxième item permet de choisir ce point de façon que son jacobien soit aussi proche de 1 que l'on veut, si bien que l'on obtient finalement un point périodique dont la dérivée (à la période) est l'identité.

Ceci permet de montrer :

THÉORÈME 10 ([BD2]). — *Il existe une partie résiduelle $\mathcal{R}$ de $\mathcal{U}$ telle que tout $f \in \mathcal{R}$ possède une infinité de disques D_n périodiques (de période t_n), d'orbites deux à deux disjointes, avec la propriété universelle suivante : pour tout ouvert $\mathcal{O}$ de difféomorphisme de D^3 dans l'intérieur $\overset{o}{D}{}^3$, il existe n tel que la restriction de f^{t_n} au disque D_n est différentiablement conjuguée à un élément de $\mathcal{O}$.*

En remarquant que l'ensemble des difféomorphismes de D^3 dans $\overset{o}{D}{}^3$ contient un ouvert $\mathcal{U}_0$ possédant les propriétés de $\mathcal{U}$ décrites ci-dessus, on obtient un procédé de renormalisation : il existe une partie résiduelle de $\mathcal{U}$ qui contient une infinité de disques périodiques D_n qui contiennent chacun une infinité de disques périodiques qui contiennent eux-mêmes une infinité de disques périodiques, etc.

On crée ainsi un arbre dont chaque branche est une suite, décroissante pour l'inclusion, d'orbites de disques strictement périodiques et dont la période tend vers l'infini. L'intersection de cette suite décroissante est un compact transitif, Lyapunov stable et donc saturé, conjugué à un *odomètre* (adding machine) et donc sans orbite périodique. L'ensemble des branches infinies de cet arbre étant non dénombrable, on obtient le Théorème 4.

A. HYPERBOLICITÉ : UNIFORME, PARTIELLE, PROJECTIVE, EN VOLUME

Soit f un difféomorphisme d'une variété compacte et soit $\mathcal{E}$ un compact de M invariant par f. Voici une définition « minimaliste » de la notion de décomposition dominée (i.e. ayant le moins d'hypothèses inutiles possibles) :

DÉFINITION A.1. — *Soit $\mathcal{E}$ une partie de M invariante par f (i.e. $f(\mathcal{E}) = \mathcal{E}$). Considérons une décomposition $TM_x = E_1(x) \oplus \cdots \oplus E_k(x)$, $x \in \mathcal{E}$, de l'espace tangent en tout point de $\mathcal{E}$. On dira que cette décomposition est dominée si elle vérifie les propriétés suivantes :*

(1) *Pour tout $i \in \{1, \ldots, k\}$, la dimension de $E_i(x)$ ne dépend pas de $x \in \mathcal{E}$.*

(2) *La décompostion est invariante par l'action naturelle de la différentielle f_* de f : en d'autres termes $E_i(f(x)) = f_*(E_i(x))$.*

(3) *Il existe $\ell \in \mathbb{N}$ tel que, pour tout $x \in \mathcal{E}$ tout couple $1 \leqslant i < j \leqslant k$ et tout couple $u \in E_i(x) \setminus \{0\}, v \in E_j(x) \setminus \{0\}$, on a l'inégalité suivante :*

$$\frac{\|f_*^\ell(u)\|}{\|u\|} \leqslant \frac{\|f_*^\ell(v)\|}{2\|v\|}.$$

(Certains auteurs parlent d'*hyperbolicité projective* en présence d'une décomposition dominée.)

Voyons quelques propriétés élémentaires des décompositions dominées :

– (Continuité) Toute décomposition dominée sur un ensemble $\mathcal{E}$ est continue et s'étend de façon unique à l'adhérence de $\mathcal{E}$.

– (Extension à un voisinage) Il existe un voisinage U de $\bar{\mathcal{E}}$ tel que l'ensemble maximal invariant $\Lambda(\bar{U}, f)$ dans $\bar{U}$ possède une décomposition dominée prolongeant celle définie sur $\mathcal{E}$.

– (Robustesse) Il existe un C^1-voisinage $\mathcal{U}_f$ de f tel que pour tout $g \in \mathcal{U}_f$ l'ensemble maximal invariant $\Lambda(\bar{U}, g)$ possède une décomposition dominée variant continûment avec g.

– (Unicité) Si $\mathcal{E}$ admet une décomposition dominée, alors il existe une (unique) décomposition dominée $TM|_{\mathcal{E}} = E_1 \oplus \cdots \oplus E_k$, appelée *la décomposition dominée la plus fine*, telle que toute autre décomposition dominée $F_1 \oplus \cdots \oplus F_l$ de $\mathcal{E}$ s'obtient en regroupant les fibrés E_i.

On dira qu'un des fibrés E_i est *uniformément contractant* si (quitte à augmenter l'entier ℓ dans la définition ci-dessus) on a :

$$\frac{\|f_*^\ell(u)\|}{\|u\|} \leqslant \frac{1}{2} \quad \text{pour tout } x \in \mathcal{E} \text{ et tout } u \in E_i(x) \setminus \{0\}.$$

Bien sûr, la domination implique que, si E_i est uniformément contractant, alors tous les E_j, $j \leqslant i$ le sont aussi.

De même on dit que E_i est *uniformément dilatant* si $\|f_*^\ell(u)\|/\|u\| \geqslant 1/2$ pour tout $x \in \mathcal{E}$ et tout $u \in E_i(x) \setminus \{0\}$. Dans ce cas les fibrés E_j, $j \geqslant i$, sont aussi uniformément dilatants.

Un compact f-invariant K est *hyperbolique* s'il existe une décomposition dominée $TM|_K = E^s \oplus E^u$ où E^s est uniformément contractant et E^u est uniformément dilatant.

Un compact f-invariant K sera dit *partiellement hyperbolique* s'il possède une décomposition dominée et si l'un des sous-fibrés E_i de sa décomposition dominée la plus fine est uniformément contractant ou dilatant. On notera alors E^s et E^u la somme des sous-fibrés uniformément contractants et dilatants, respectivement, et E^c la somme des autres sous-fibrés. On obtient ainsi une nouvelle décomposition dominée du type $E^s \oplus E^c$, $E^c \oplus E^u$ ou $E^s \oplus E^c \oplus E^u$, ces fibrés étant appelés respectivement fibrés stable, central et instable.

Si f est un difféomorphisme partiellement hyperbolique d'une variété compacte M possédant une décomposition dominée $E_1 \oplus \cdots \oplus E_k$ telle que E_k soit uniformément dilatant, Brin et Pesin ([BrPe]) montrent l'existence d'un unique feuilletage $\mathcal{F}^u$ f-invariant tangent à E_k et dont les feuilles sont aussi différentiables que f, appelé *feuilletage instable fort*.

RÉFÉRENCES

[Ab] F. Abdenur – « Generic robustness of spectral decompositions », preprint IMPA, 2001.

[Ar] M.-C. Arnaud – « Création de connexions en topologie C^1 », *Ergod. Th. & Dynam. Systems* **21** (2001), p. 339–381.

[BD] Ch. Bonatti & L.J. Díaz – « Connexions hétéroclines et généricité d'une infinité de puits ou de sources », *Ann. Scient. Éc. Norm. Sup.,* 4^e *série* **32** (1999), p. 135–150.

[BD2] ———, « On maximal transitive sets of generic diffeomorphisms », *Publ. Math. Inst. Hautes Études Sci.* **96** (2002), p. 171–197.

[BDP] Ch. Bonatti, L.J. Díaz & E. Pujals – « A C^1-generic dichotomy for diffeomorphisms : weak forms of hyperbolicity or infinitely many sinks or sources », à paraître aux *Annals of Math.*

[BV] Ch. Bonatti & M. Viana – « SRB for partially hyperbolic attractors : the contracting case », *Israel Journal of Math.* **115** (2000), p. 157–193.

[BrPe] M. Brin & Ya. Pesin – « Partially hyperbolic dynamical systems », *Izv. Acad. Nauk. SSSR* **1** (1974), p. 177–212.

[CM] C. Carballo & C. Morales – « Homoclinic classes and finitude of attractors for vector fields on n-manifolds », Preprint, 2001.

[CMP] C. Carballo, C. Morales & M.J. P. fico – « Homoclinic class for $\mathcal{C}^1$-generic vector fields », preprint PUC-Rio, à paraître à *Ergod. Th. & Dynam. Systems*, 2000.

[DPU] L.J. Díaz, E. Pujals & R. Ures – « Partial hyperbolicity and robust transitivity », *Acta Math.* **183** (1999), p. 1–43.

[F] J. Franks – « Necessary conditions for stability of diffeomorphisms », *Trans. Amer. Math. Soc.* **158** (1971), p. 301–308.

[Ha] S. Hayashi – « Connecting invariant manifolds and the solution of the C^1-stability and Ω-stability conjectures for flows », *Ann. of Math.* **145** (1997), p. 81–137.

[HPS] M. Hirsch, C. Pugh & M. Shub – *Invariant manifolds*, Lecture Notes in Math., vol. 583, Springer-Verlag, 1977.

[Ma1] R. Mañé – « An ergodic closing lemma », *Annals of Math.* **116** (1982), p. 503–540.

[Ma2] ———, « Contributions to the stability conjecture », *Topology* **17** (1978), p. 386–396.

[N] S. Newhouse – « Diffeomorphisms with infinitely many sinks », *Topology* **13** (1974), p. 9–18.

[PT] J. Palis & F. Takens – *Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations*, Cambridge Studies in Advanced Mathematics, vol. 35, Cambridge University Press, 1993.

[Pu] C. Pugh – « The closing lemma », *Amer. J. Math.* **89** (1967), p. 956–1009.

[Sh] M. Shub – « Topological transitive diffeomorphism on T^4 », Lect. Notes in Math., vol. 206, Springer-Verlag, 1971, p. 39.

[Sm] S. Smale – « Differentiable dynamical systems », *Bull. Amer. Math. Soc.* **73** (1967), p. 747–817.

[WX] L. Wen & Z. Xia – « C^1 connecting lemmas », *Trans. Amer. Math. Soc.* **352** (2000), p. 5213–5230.

Christian BONATTI

Université de Bourgogne
UMR 5584 du C.N.R.S.
Laboratoire de Topologie
UFR de Sciences et Technologies
9, avenue Alain Savary
B.P. 47870
F-21078 Dijon Cedex
E-mail : bonatti@satie.u-bourgogne.fr

Séminaire BOURBAKI Juin 2002
54e année, 2001-2002, nO 905, p. 243 à 266

VARIÉTÉS RATIONNELLEMENT CONNEXES
[d'après T. Graber, J. Harris, J. Starr et A.J. de Jong]

par **Olivier DEBARRE**

1. LE THÉORÈME POUR LES VARIÉTÉS COMPLEXES

Une variété complexe[(1)] X est *rationnellement connexe* si deux points généraux de X peuvent être reliés par une courbe rationnelle, c'est-à-dire sont dans l'image d'un morphisme $\mathbf{P}^1 \to X$. On peut donner les exemples suivants.

• Toute variété propre rationnelle, ou même seulement unirationnelle, c'est-à-dire dominée par un espace projectif, est rationnellement connexe. Toutes ces notions sont équivalentes pour les variétés propres et lisses de dimension au plus 2.

• Une variété de Fano (c'est-à-dire une variété projective et lisse dont le déterminant du fibré tangent est ample) est rationnellement connexe ([C1], [KMM1], [KMM2]). En particulier, une hypersurface lisse de $\mathbf{P}^n$ de degré au plus n est rationnellement connexe.

En dimension au moins 3, la connexité rationnelle devient une propriété plus générale que la rationalité (et, conjecturalement, que l'unirationalité) qui s'est révélée être aussi beaucoup plus maniable : pour les familles de variétés complexes propres et lisses, elle est ouverte et fermée (*cf.* § 2) et il en existe, au moins conjecturalement, une caractérisation numérique (*cf.* conjecture 3.3). On peut faire remonter au moins à Mumford au début des années 80 l'idée (orale) de la définition et de cette conjecture, mais c'est dans [KMM2] que Kollár, Miyaoka et Mori donnent véritablement son essor à cette théorie. Le résultat central qui restait manquant, la dernière pièce du puzzle, qui montre en un certain sens que l'intuition de Mumford, Kollár, Miyaoka et Mori était la bonne, et qui couronne leur théorie, est le théorème ci-dessous, conjecturé dans [KMM2], et qui fait l'objet de cet exposé. Il confirme la place centrale qu'occupe la connexité rationnelle dans l'étude des variétés de dimension supérieure.

THÉORÈME 1.1 (Graber–Harris–Starr). — *Un morphisme propre d'une variété complexe sur une courbe lisse dont les fibres générales sont rationnellement connexes a une section.*

(1) C'est-à-dire un schéma séparé, intègre et de type fini sur $\mathbf{C}$.

Ce résultat est démontré dans [GHS][(2)] , et généralisé par de Jong et Starr dans [dJS] au cas d'un corps de base algébriquement clos quelconque (th. 2.1). La démonstration dans le cas complexe est expliquée dans le § 4, tandis que les modifications à apporter dans le cas général font l'objet du § 5.

Soit K le corps des fonctions d'une courbe complexe, c'est-à-dire une extension de $\mathbf{C}$ de degré de transcendance 1. Le théorème 1.1 peut s'énoncer ainsi :

1.2.— *Toute variété X définie sur K, propre et rationnellement connexe*[(3)]*, a un point rationnel sur K.*

L'origine de ce genre de résultat remonte à M. Noether, qui montre dans [N] qu'une surface fibrée en coniques sur $\mathbf{P}^1$ admet une section, résultat généralisé ensuite par Enriques au cas d'une courbe base irrationnelle ([En]), puis enfin par Tsen, qui montre dans [T] qu'étant donné le corps des fonctions K d'une courbe, toute hypersurface de $\mathbf{P}^n_K$ de degré au plus n a un point rationnel sur K (dans la terminologie d'E. Artin, K est « quasi algébriquement clos », dans celle de Lang, K est C_1). On trouve aussi des cas particuliers du théorème de Tsen dans des articles postérieurs de Conforto et Severi[(4)]. Enfin, Campana, Peternell et Pukhlikov donnent dans la prépublication [CPP] une démonstration différente de 1.2 lorsque X est une variété de Fano de dimension 3.

L'énoncé 1.2 était aussi déjà connu lorsque X est une surface propre, lisse et rationnellement connexe et K un corps C_1 ([M], [CT]) ou un espace homogène sous un groupe algébrique linéaire connexe et K un corps C_1 parfait (théorème de Chevalley–Springer ; *cf.* [S], III-14-16 et [B], 18.2).

Il est naturel de poser la question suivante : l'énoncé 1.2 reste-t-il valable pour tout corps C_1 ? Esnault y répond par l'affirmative dans [Es] lorsque K est un corps fini[(5)].

---

(2) Graber, Harris, Mazur et Starr démontrent dans [GHMS] la « réciproque » suivante au théorème 1.1 : étant donné un morphisme propre $u : X \to B$ entre variétés complexes, avec B normale et quasi-projective, il existe un plongement projectif de B tel que les deux propriétés suivantes sont équivalentes :

(i) le morphisme u a une section au-dessus d'une courbe section linéaire très générale de B ;
(ii) il existe une sous-variété Y de X telle que les fibres générales de $u|_Y : Y \to B$ soient rationnellement connexes (*cf.* note 6).

Ce résultat leur permet de construire une famille sans section de surfaces d'Enriques paramétrée par une courbe lisse : le théorème 1.1 ne se généralise donc pas aux variétés complexes X vérifiant par exemple la condition plus faible (*cf.* 3.4) $H^m(X, \mathscr{O}_X) = 0$ pour tout $m > 0$. Ceci répond négativement à une question de Serre ([GS], p. 152).

(3) C'est-à-dire que si $\overline{K}$ est une clôture algébrique de K, la variété $X_{\overline{K}}$ satisfait à la définition ci-dessus ; *cf.* aussi § 2.

(4) L'article de Tsen, paru en allemand dans une revue chinoise en 1936, n'a pas dû avoir à cette époque la diffusion qu'il méritait.

(5) Plus précisément, Esnault montre, pour toute variété X propre et lisse définie sur un corps fini $\boldsymbol{k}$ de cardinal q qui vérifie $\mathrm{CH}_0(X \times \mathrm{Spec}(\overline{\boldsymbol{k}(X)})) \simeq \mathbf{Z}$, la congruence $\mathrm{Card}(X(\boldsymbol{k})) \equiv 1 \pmod{q}$. Cela s'applique aux variétés propres, lisses et rationnellement connexes définies sur $\boldsymbol{k}$.

Quelques conventions : une *variété* X est un schéma séparé géométriquement intègre de type fini défini sur un corps $\boldsymbol{k}$; pour toute extension $\boldsymbol{k}'$ de $\boldsymbol{k}$, on note $X_{\boldsymbol{k}'}$ la variété $X \times_{\mathrm{Spec}(\boldsymbol{k})} \mathrm{Spec}(\boldsymbol{k}')$. On note Ω_X le faisceau des différentielles sur X ; lorsque X est une variété lisse, c'est le dual du fibré tangent T_X. Une *courbe* est un schéma séparé géométriquement réduit et connexe, de dimension 1, défini sur un corps, qui n'est pas nécessairement irréductible.

On dit qu'un point *général* (resp. *très général)* d'une variété vérifie une propriété donnée si l'ensemble des points qui la vérifie contient un ouvert non vide (resp. une intersection dénombrable d'ouverts non vides).

Je voudrais remercier T. Graber et J. Kollár de leurs patientes explications, ainsi que P. Baumann, A. Beauville, K. Behrend, J.-F. Boutot, A. Chambert-Loir, J.-L. Colliot-Thélène, H. Esnault, O. Gabber, J. Harris, J. de Jong, V. Kharlamov, Y. Laszlo, L. Moret-Bailly, M. Perret, M. Raynaud, Ph. Satgé et J. Starr de leurs conseils et suggestions.

2. VARIÉTÉS RATIONNELLEMENT CONNEXES ET SÉPARABLEMENT RATIONNELLEMENT CONNEXES

La définition de la connexité rationnelle donnée au début du § 1, valable sur le corps des complexes (ou sur tout corps algébriquement clos non dénombrable), se traduit sur un corps quelconque de la façon suivante : une variété X définie sur un corps $\boldsymbol{k}$ est *rationnellement connexe* s'il existe un $\boldsymbol{k}$-schéma de type fini T et un morphisme

$$F : T \times \mathbf{P}^1 \longrightarrow X$$

(auquel il faut penser comme à une famille de courbes rationnelles sur X paramétrée par T) tels que le morphisme

$$\begin{aligned} T \times \mathbf{P}^1 \times \mathbf{P}^1 &\longrightarrow X \times X \\ (t, u, u') &\longmapsto \big(F(t,u), F(t,u')\big) \end{aligned} \tag{1}$$

soit dominant[(6)]. Si X est lisse et $\boldsymbol{k}$ algébriquement clos de caractéristique nulle, c'est équivalent à demander que son application tangente soit surjective en un point (t_0, u_0, u_0'). Notant f la courbe $F(t_0, \cdot) : \mathbf{P}^1 \to X$, on vérifie ([D2], Cor. 4.17) que c'est le cas si et seulement si $H^1(\mathbf{P}^1, f^*T_X(-2)) = 0$; on dit que la courbe rationnelle f est *très libre*.

La présence d'une courbe rationnelle très libre sur une variété rationnellement connexe, souvent essentielle dans les applications, n'est pas assurée en caractéristique

(6) Un point est rationnellement connexe, le vide ne l'est pas.

non nulle[7]. Cela conduit à dire qu'une variété X est *séparablement rationnellement connexe* s'il existe T et $F : T \times \mathbf{P}^1 \to X$ comme ci-dessus tels que le morphisme (1) soit *génériquement lisse*. Une variété lisse X est alors séparablement rationnellement connexe si et seulement s'il existe sur X une courbe rationnelle très libre. En caractéristique nulle, les deux notions sont bien sûr identiques.

On peut maintenant énoncer la version du théorème 1.1 valable en toute caractéristique.

THÉORÈME 2.1 (de Jong–Starr). — *Un morphisme propre d'une variété lisse sur une courbe lisse, dont les fibres générales sont des variétés lisses et séparablement rationnellement connexes, a une section.*

De nouveau, ce résultat peut s'énoncer : *toute variété propre, lisse et séparablement rationnellement connexe définie sur le corps des fonctions d'une courbe a un point rationnel sur ce corps.* Il sera démontré dans le § 5.

Un produit fini de variétés (séparablement) rationnellement connexes est (séparablement) rationnellement connexe, un revêtement étale fini d'une variété (séparablement) rationnellement connexe est (séparablement) rationnellement connexe. La connexité rationnelle (séparable) est une propriété birationnelle parmi les variétés propres. Enfin, pour les familles de variétés propres et lisses, la propriété d'être séparablement rationnellement connexe est ouverte[8] et, en caractéristique nulle, fermée[9].

Il existe beaucoup de courbes rationnelles sur une variété séparablement rationnellement connexe, comme le montre le résultat suivant.

THÉORÈME 2.2. — *Soit X une variété propre et lisse définie sur un corps algébriquement clos. Les conditions suivantes sont équivalentes :*

(i) *la variété X est séparablement rationnellement connexe ;*

(ii) *étant donnés des points distincts $p_1, \dots, p_r$ de X et des directions tangentes $\ell_1, \dots, \ell_r$ en ces points, il existe une courbe rationnelle très libre $\mathbf{P}^1 \to X$ non ramifiée qui passe par chaque p_i dans la direction ℓ_i ;*

En caractéristique nulle, elles sont de plus équivalentes à :

(iii) *deux points généraux de X peuvent être reliés par une chaîne de courbes rationnelles.*

La démonstration se trouve dans [KMM2], 2.4, ou [Ko1], Th. IV.3.9.4, sauf pour ce qui concerne les directions tangentes, point pour lequel on pourra consulter [D2], Ex. 4.8.6.

[7] Kollár construit dans [Ko2] des variétés de Fano, donc projectives, lisses et rationnellement connexes, sans courbe rationnelle très libre (c'est-à-dire non séparablement rationnellement connexes).

[8] Cela résulte de leur caractérisation par l'existence d'une courbe rationnelle très libre.

[9] Cela résulte de leur caractérisation par la condition (iii) ci-dessous.

3. QUELQUES COROLLAIRES

Le théorème 1.1 est énoncé comme « Problem » dans [Ko1], IV.6, et l'on trouve déjà dans ce livre les corollaires suivants ([Ko1], Propositions IV.5.6.3 et IV.5.7).

3.1. Morphismes à base et fibres rationnellement connexes

COROLLAIRE 3.1. — *Soit $u : X \to Y$ un morphisme propre entre variétés complexes. Si Y et les fibres générales de u sont rationnellement connexes, il en est de même de X.*

Preuve. — Soient x_1 et x_2 des points généraux de X et $\mathbf{P}^1 \to Y$ une courbe rationnelle joignant $u(x_1)$ et $u(x_2)$. On note X' l'unique composante irréductible de $X \times_Y \mathbf{P}^1$ qui domine $\mathbf{P}^1$ et $\widetilde{X}' \to X'$ une désingularisation. Les fibres générales de $\widetilde{u}' : \widetilde{X}' \to \mathbf{P}^1$ sont propres et birationnelles à des fibres générales de u, donc sont rationnellement connexes. Le théorème entraîne que $\widetilde{u}'$ a une section $\sigma : \mathbf{P}^1 \to \widetilde{X}'$. On peut donc connecter deux points généraux $\widetilde{x}'_1$ et $\widetilde{x}'_2$ de $\widetilde{X}'$ par une chaîne de trois courbes rationnelles : une pour joindre $\widetilde{x}'_1$ à $\sigma(\widetilde{u}'(\widetilde{x}'_1))$, la section σ, et une pour joindre $\sigma(\widetilde{u}'(\widetilde{x}'_2))$ à $\widetilde{x}'_2$. Le théorème 2.2 entraîne que $\widetilde{X}'$ est rationnellement connexe, donc que x_1 et x_2 peuvent être joints par une courbe rationnelle. □

3.2. Le quotient rationnel n'est pas uniréglé

Soit X une variété complexe propre et lisse ; Campana montre dans [C2] (*cf.* aussi [D2], Chap. 5) qu'il existe un ouvert dense X^0 de X, une variété lisse $R(X)$ et un morphisme *propre* $\rho : X^0 \to R(X)$ tels que :

(Q1) les fibres de ρ sont rationnellement connexes ;

(Q2) toute courbe rationnelle de X rencontrant une fibre très générale de ρ est contenue dans cette fibre.

La variété $R(X)$ est uniquement déterminée à isomorphisme birationnel près ; on l'appelle le *quotient rationnel* de X. La variété X est rationnellement connexe si et seulement si $R(X)$ est un point ; elle est uniréglée[(10)] si et seulement si $\dim(R(X)) < \dim(X)$.

COROLLAIRE 3.2. — *Le quotient rationnel d'une variété complexe propre et lisse n'est pas uniréglé.*

[(10)] Une variété X de dimension n est uniréglée s'il existe une variété Y de dimension $n-1$ et une application rationnelle dominante $\mathbf{P}^1 \times Y \dashrightarrow X$ (un point n'est donc pas uniréglé). Pour une variété complexe propre X, cela signifie simplement qu'il passe une courbe rationnelle par chaque point de X.

Preuve. — Soient X une variété complexe propre et lisse et $\rho : X \dashrightarrow R(X)$ son quotient rationnel. Quitte à prendre des modèles birationnels (le quotient rationnel est un invariant birationnel pour les variétés propres et lisses), on peut supposer que $R(X)$ est projectif et que ρ est un morphisme. Soient $\mathbf{P}^1 \to R(X)$ une courbe rationnelle passant par un point très général de $R(X)$ et X' une désingularisation de l'unique composante irréductible de $X \times_{R(X)} \mathbf{P}^1$ qui domine $\mathbf{P}^1$. Le théorème 1.1 entraîne que $X' \to \mathbf{P}^1$ a une section, ce qui contredit la propriété (Q2) du quotient rationnel. □

3.3. Caractérisations numériques

Comme mentionné dans l'introduction, on espère pouvoir caractériser numériquement les variétés complexes rationnellement connexes.

CONJECTURE 3.3. — *Une variété complexe projective et lisse X est rationnellement connexe si et seulement si $H^0(X, \Omega_X^{\otimes m}) = 0$ pour tout entier $m > 0$.*

On a une conjecture analogue pour les variétés uniréglées.

CONJECTURE 3.4. — *Une variété complexe projective et lisse X est uniréglée si et seulement si $H^0(X, (\det(\Omega_X))^{\otimes m}) = 0$ pour tout entier $m > 0$.*

Pour chacune de ces conjectures, le sens direct est élémentaire ([D2], Cor. 4.17 et Cor. 4.12). La réciproque n'est connue qu'en dimension au plus 3 ([KMM2]) ; elle est conséquence en toute dimension du Programme du Modèle Minimal de Mori.

COROLLAIRE 3.5. — *La conjecture 3.4 entraîne la conjecture 3.3.*

Preuve. — Soient X une variété complexe projective et lisse qui n'est pas rationnellement connexe et $\rho : X \dashrightarrow R(X)$ son quotient rationnel. Quitte à prendre des modèles birationnels (les espaces vectoriels $H^0(X, \Omega_X^{\otimes m})$ sont des invariants birationnels pour les variétés propres et lisses), on peut supposer que $R(X)$ est projectif et lisse, de dimension $r > 0$, et que ρ est un morphisme. Si la conjecture 3.4 est vérifiée, le corollaire 3.2 entraîne qu'il existe une section non nulle de $(\det(\Omega_{R(X)}))^{\otimes m}$ pour un $m > 0$. Comme ce fibré en droites est facteur direct de $\Omega_{R(X)}^{\otimes rm}$, ce dernier a aussi une section non nulle, qui se relève en une section non nulle de $\Omega_X^{\otimes rm}$. □

3.4. Groupe fondamental

Soient X une variété propre, lisse et séparablement rationnellement connexe et m un entier strictement positif. Comme la restriction de Ω_X à une courbe rationnelle très libre est somme directe de fibrés en droites de degré négatif, toute section de $\wedge^m \Omega_X$ s'annule sur cette courbe. Les courbes rationnelles très libres recouvrant un ouvert dense de X, les groupes $H^0(X, \wedge^m \Omega_X)$ sont tous nuls.

En caractéristique nulle, la dualité de Hodge entraîne $H^m(X, \mathscr{O}_X) = 0$, de sorte que $\chi(X, \mathscr{O}_X) = 1$. Étant donné un revêtement étale connexe $u : Y \to X$, la variété Y est

encore séparablement rationnellement connexe, d'où $\chi(Y,\mathscr{O}_Y)=1$, de sorte que u est de degré $\chi(Y,\mathscr{O}_Y)/\chi(X,\mathscr{O}_X)=1$: la variété X est donc algébriquement simplement connexe. Sur $\mathbf{C}$, on montre même que X est simplement connexe (*cf.* [D2], Cor. 4.18).

En caractéristique non nulle, l'annulation de $H^m(X,\mathscr{O}_X)$ n'est pas connue[(11)]. On peut cependant, comme me l'a signalé Kollár, déduire la simple connexité de X du théorème 2.1[(12)].

COROLLAIRE 3.6 (Kollár). — *Une variété propre, lisse et séparablement rationnellement connexe est algébriquement simplement connexe.*

Preuve. — Soit X une variété propre, lisse et séparablement rationnellement connexe sur un corps algébriquement clos $\boldsymbol{k}$ de caractéristique p ; il s'agit de montrer que tout revêtement étale connexe galoisien $Y\to X$ est trivial. Si ce n'est pas le cas, il existe une factorisation $Y\xrightarrow{u}X'\to X$, où le groupe de Galois de u est cyclique d'ordre ℓ premier. Dans cette situation, X' est aussi propre, lisse et séparablement rationnellement connexe. La classification des revêtements cycliques d'une variété (théories de Kummer et d'Artin–Schreier) nous apprend qu'il existe un recouvrement de X' par des ouverts affines U_i et,

- si $\ell\neq p$, des fonctions régulières $h_i : U_i\to\boldsymbol{k}^*$ et $e_{ij} : U_i\cap U_j\to\boldsymbol{k}^*$ telles que $e_{ij}e_{jk}e_{ki}=1$ sur $U_i\cap U_j\cap U_k$ et $e_{ij}^\ell=h_ih_j^{-1}$ sur $U_i\cap U_j$, l'ouvert affine $u^{-1}(U_i)$ étant défini par l'équation $h_i=t_i^\ell$ dans $U_i\times\mathbf{A}^1_{\boldsymbol{k}}$ et le recollement se faisant par $t_i=e_{ij}t_j$;
- si $\ell=p$, des fonctions régulières $h_i : U_i\to\boldsymbol{k}$ et $e_{ij} : U_i\cap U_j\to\boldsymbol{k}$ telles que $e_{ij}+e_{jk}+e_{ki}=0$ sur $U_i\cap U_j\cap U_k$ et $e_{ij}^p-e_{ij}=h_i-h_j$ sur $U_i\cap U_j$, l'ouvert affine $u^{-1}(U_i)$ étant défini par l'équation $h_i=t_i^p-t_i$ dans $U_i\times\mathbf{A}^1_{\boldsymbol{k}}$ et le recollement se faisant par $t_i=e_{ij}+t_j$.

Posons $B=\operatorname{Spec}\boldsymbol{k}[t]$; on construit une sous-variété $\mathscr{Y}$ de $X'\times B$ comme la réunion des variétés affines définies dans $U_i\times\mathbf{A}^1_{\boldsymbol{k}}\times B$ de la façon suivante :

- si $\ell\neq p$, par l'équation $th_i=t_i^\ell$, le recollement se faisant par $t_i=e_{ij}t_j$;
- si $\ell=p$, par l'équation $h_i=t_i^p-t^{p-1}t_i$, le recollement se faisant par $t_i=te_{ij}+t_j$.

Les fibres de $\mathscr{Y}\to B$ hors de 0 sont isomorphes à Y, donc sont séparablement rationnellement connexes. La fibre en 0 est isomorphe à X', compté avec multiplicité ℓ. Ce morphisme admet une section par le théorème 2.1, ce qui entraîne $\ell=1$. □

(11) En dimension 3, elle est connue pour les variétés séparablement unirationnelles ([Ny]) et pour les variétés de Fano ([SB]).

(12) La trivialité du p-sous-groupe de Sylow du groupe fondamental était déjà connue ([E] et [Su]). Noter qu'il existe des surfaces unirationnelles qui ne sont pas simplement connexes ([Sh], prop. 5).

3.5. Déformation de morphismes

Nous aurons à plusieurs reprises besoin d'étudier les déformations d'un morphisme $f : C \to X$ d'une courbe C projective (variable) à points doubles ordinaires vers une variété projective fixe X, lisse sur l'image de f.

Notons E^i l'espace vectoriel $\mathrm{Ext}^i(f^*\Omega_X \to \Omega_C, \mathscr{O}_C)$; les déformations au premier ordre de f sont paramétrées par E^1 et les obstructions sont dans E^2 : cela signifie que les déformations de f sont paramétrées par un schéma séparé quasi-projectif qui peut être défini au voisinage de $[f]$ par $\dim(E^2)$ équations dans une variété lisse de même dimension que E^1. Le noyau de la différentielle $f^*\Omega_X \to \Omega_C$ est localement libre ; son dual, le *fibré normal* à C dans X, est noté $N_{C/X}$. On a une suite exacte

$$0 \longrightarrow N^*_{C/X} \longrightarrow f^*\Omega_X \longrightarrow \Omega_C \longrightarrow \Omega_{C/X} \longrightarrow 0.$$

Exemples 3.7. — (1) Si X est une courbe et que f est lisse sur un ouvert dense de C, le complexe $f^*\Omega_X \to \Omega_C$ est quasi-isomorphe au complexe gratte-ciel $0 \to \Omega_{C/X}$ et les E^i sont nuls pour $i \neq 1$. Il n'y a pas d'obstruction et l'espace vectoriel E^1 est de dimension le degré du diviseur de ramification de f. Un décompte de paramètres montre qu'une déformation générale de f est un revêtement de X de même degré d que f qui est *simple* (c'est-à-dire avec au moins $d-1$ points dans chaque fibre).

(2) Lorsque f n'est pas ramifié, c'est-à-dire que $\Omega_{C/X}$ est nul, le complexe $f^*\Omega_X \to \Omega_C$ est quasi-isomorphe au complexe $N^*_{C/X} \to 0$ et E^i est isomorphe à $H^{i-1}(C, N_{C/X})$.

4. DÉMONSTRATION DU THÉORÈME DANS LE CAS COMPLEXE

Dans ce qui suit, on se donne, sur le corps des complexes, une variété propre X et une courbe B projective et lisse, ainsi qu'un morphisme surjectif $u : X \to B$ dont les fibres générales sont rationnellement connexes ; on désire montrer que u a une section.

On peut supposer[13] B *rationnelle* : choisissons un revêtement ramifié quelconque $g : B \to \mathbf{P}^1$ de degré d, auquel correspond un morphisme de $\mathbf{P}^1$ dans le produit symétrique $B^{(d)}$. Posant[14] $Y = \mathbf{P}^1 \times_{B^{(d)}} X^{(d)}$, on obtient un morphisme $Y \to \mathbf{P}^1$ dont la fibre au-dessus d'un point t de $\mathbf{P}^1$ est le produit

$$\prod_{b \in g^{-1}(t)} u^{-1}(b)$$

[13] Cela ne sert que pour le théorème 4.1.

[14] Pour ceux qui préfèrent la version « birationnelle », on peut aussi, en posant $L = K(B)$ et $K = K(\mathbf{P}^1)$, considérer la restriction « à la Weil » $Y_K = R_{L/K}X_L$. C'est une variété définie sur K qui vérifie $Y_{\overline{L}} \simeq X^d_{\overline{L}}$ et $Y_K(K) = X_L(L)$. Elle est donc rationnellement connexe, et l'existence d'un L-point de X_L est équivalente à celle d'un K-point de Y_K.

qui est donc rationnellement connexe. Si $Y \to \mathbf{P}^1$ a une section, il en est de même de $X \to B$.

Quitte à remplacer X par une désingularisation, on peut aussi supposer que X est *projective et lisse.*

Les grandes lignes de la démonstration sont les suivantes : partant d'une multisection de u, c'est-à-dire d'une courbe lisse C dans X qui domine B, on montre que la réunion de C et de suffisamment de queues rationnelles très libres contenues dans les fibres de u (ici intervient la connexité rationnelle de ces fibres) admet suffisamment de déformations lisses $C' \to X$ pour que ceux des points de ramification du revêtement $C' \to B$ induit par *u qui appartiennent au lieu où u est lisse* se déplacent dans des directions arbitraires sur X.

Supposons un instant qu'il existe une multisection C de u contenue dans le lieu de lissité de u. La construction précédente fournit une autre multisection C' dont les déformations dans X ont des lieux de branchement sur B qui sont généraux. Comme B est rationnelle, les revêtements $C' \to B$ forment un espace irréductible (appelé *schéma de Hurwitz*) et dégénèrent en des revêtements qui ont des sections. La considération des espaces de Kontsevich des *courbes stables* permet de compactifier ces espaces de déformations de courbes et de relever ces dégénérescences sur X. On obtient ainsi une section de u.

Le problème reste de construire une multisection de u qui évite le lieu où u n'est pas lisse. Si ce lieu est de codimension au moins 2, cela résulte du théorème de Bertini. En revanche, si une multisection C rencontre une composante multiple d'une fibre de u, elle y est ramifiée au-dessus de B et aucune déformation de C ne pourra déplacer le point de branchement correspondant sur B. C'est là le cœur de l'argument de [GHS] : l'idée est de créer de nouveaux points de ramification que l'on fait ensuite converger vers la ramification située dans les composantes multiples de fibres. Toute composante « horizontale » de la courbe stable limite évite alors les composantes multiples de fibres et l'on peut appliquer la construction précédente.

4.1. Où l'on se ramène à la construction d'une multisection vérifiant certaines propriétés

L'objet de ce paragraphe est de montrer que l'existence d'une multisection de u possédant « suffisamment » de déformations (conditions (C2) ci-dessous) suffit à assurer l'existence d'une section. On ne se sert pas ici de la connexité rationnelle des fibres de u.

THÉORÈME 4.1. — *Soient X une variété et C une courbe, complexes, projectives et lisses. On se donne des morphismes $u : X \to \mathbf{P}^1$ et $f : C \to X$ tels que :*

(C1) *la ramification R de $u \circ f$ est simple et contenue dans l'image inverse par f du lieu lisse de u ;*

(C2) *l'espace vectoriel $H^1(C, N_{C/X}(-R))$ est nul.*

Alors u a une section.

Preuve. — La condition (C2) entraîne $H^1(C, N_{C/X}) = 0$; par Ex. 3.7(2), les déformations de C dans X sont donc paramétrées par un schéma qui est lisse de dimension $h^0(C, N_{C/X})$ au voisinage de $[C]$. De plus, la restriction

$$H^0(C, N_{C/X}) \longrightarrow (N_{C/X})|_R$$

est surjective ; cela signifie que par déformation $f' : C' \to X$ de f, les images dans X des points de R peuvent être indépendamment déplacées dans des directions normales à C arbitraires. Comme, par (C1), la différentielle de $u \circ f$ induit en chaque point p de R une surjection $N_{C/X,p} \to T_{\mathbf{P}^1, u\circ f(p)}$, les points de *branchement* de $u \circ f'$ peuvent aussi être indépendamment déplacés dans $\mathbf{P}^1$ dans des directions arbitraires. Le revêtement $u \circ f'$ est en particulier simple. Or les revêtements simples de $\mathbf{P}^1$ de degré d et genre g fixés forment un schéma lisse *irréductible*[15] de dimension $2d+2g-2$ (un pour chaque point de branchement) que l'on notera $M_g(\mathbf{P}^1, d)^h$: par déformation de $f : C \to X$, on remplit donc, en composant par u, un ouvert dense de $M_g(\mathbf{P}^1, d)^h$.

Comme on l'a expliqué plus haut, l'idée est de faire dégénérer ces revêtements et de montrer que ces dégénérescences se relèvent à X. On introduit pour cela des compactifications des espaces de courbes, construites de la façon suivante.

Soient β un élément de $H_2(X, \mathbf{Z})$ et g un entier positif. Une courbe quasi-stable de genre g de degré β sur X est la donnée d'une courbe connexe projective C de genre g dont les singularités sont des points doubles ordinaires et d'un morphisme $f : C \to X$ tels que $f_*C = \beta$. Elle est *stable* s'il n'y a qu'un nombre fini d'automorphismes $\sigma : C \to C$ tels que $f \circ \sigma = f$, c'est-à-dire si toute composante rationnelle lisse de C contractée par μ contient au moins 3 points singuliers de C. On peut toujours *stabiliser* une courbe quasi-stable $f : C \to X$ en contractant dans C les composantes rationnelles lisses contractées par f et contenant au plus 2 points singuliers de C.

On définit de façon analogue les familles plates de courbes stables de genre g de degré β sur X ; il existe un espace de modules grossier, noté $\overline{M}_g(X, \beta)$, qui est un schéma projectif[16], et un morphisme de composition ([BM], Th. 3.6)

$$u_* : \overline{M}_g(X, \beta) \longrightarrow \overline{M}_g(\mathbf{P}^1, u_*\beta)$$

décrit ensemblistement de la façon suivante : si $f : C \to X$ est une courbe stable de genre g et de degré β, l'image $u_*([f])$ est le point de $\overline{M}_g(\mathbf{P}^1, u_*\beta)$ associé à la stabilisée de la courbe quasi-stable $u \circ f : C \to \mathbf{P}^1$.

[15] C'est ici que l'hypothèse $B = \mathbf{P}^1$ intervient ; *cf.* [Cl] ; [H] ; [F], § 1.

[16] La meilleure référence semble être [FP] : on y trouve (Th. 1, p. 55) une construction détaillée de $\overline{M}_g(X, \beta)$ (comme sous-schéma fermé de $\overline{M}_g(\mathbf{P}^n, d)$, une fois choisi un plongement de X dans un espace projectif $\mathbf{P}^n$). Les auteurs mentionnent aussi (sans la mener à terme) une construction de $\overline{M}_g(\mathbf{P}^n, d)$ via la théorie géométrique des invariants de Mumford, qui devrait permettre de construire $\overline{M}_g(X, \beta)$ comme *champ algébrique*. Lorsque X est un point (ou que $\beta = 0$), on retrouve l'espace des modules $\overline{M}_g$ de Deligne–Mumford des courbes stables de genre g.

Le point essentiel est que par Ex. 3.7(1), l'adhérence $\overline{M}_g(\mathbf{P}^1,d)^h$ de $M_g(\mathbf{P}^1,d)^h$ dans $\overline{M}_g(\mathbf{P}^1,d)$ contient tous les morphismes *finis* $C\to\mathbf{P}^1$ de degré d, où C est une courbe projective connexe de genre g à points doubles ordinaires, donc en particulier des revêtements $C\to\mathbf{P}^1$ qui ont une section : il suffit de prendre d exemplaires de $\mathbf{P}^1$ et d'identifier $d+g-1$ paires de points situés sur des exemplaires distincts.

On peut alors terminer la démonstration du théorème 1.1 : on a construit une multisection de u dont les déformations dans X remplissent un ouvert dense de $M_g(\mathbf{P}^1,d)^h$. Cela signifie que l'image du morphisme propre u_* défini ci-dessus contient $\overline{M}_g(\mathbf{P}^1,d)^h$, donc des revêtements $C\to\mathbf{P}^1$ avec section. Ceux-ci se relèvent par construction à X, ce qui termine la démonstration. □

4.2. Où l'on attache des queues rationnelles très libres à une multisection pour la déformer

Le but de ce paragraphe est de construire des multisections de u vérifiant la condition (C2) du théorème 4.1.

Soit C la normalisation d'une multisection de u ; on suppose que le morphisme correspondant $f:C\to X$ n'est pas ramifié. Choisissons des points généraux $p_1,\dots,p_n$ sur C et, pour chaque i, une courbe rationnelle très libre non ramifiée $L_i\to X$ dans la fibre $u^{-1}(u(f(p_i)))$, passant par $f(p_i)$, à direction tangente générale dans cette fibre (th. 2.2). On forme ainsi un « peigne » $\overline{C}=C\cup L_1\cup\cdots\cup L_n$ muni d'un morphisme non ramifié $\overline{f}:\overline{C}\to X$.

Proposition 4.2. — *Soit n_0 un entier. Sous les hypothèses et les notations précédentes, pour n assez grand, le morphisme $\overline{f}:\overline{C}\to X$ se déforme en $C'\to X$, où C' est une courbe projective et lisse qui vérifie*

$$H^1(C',N_{C'/X}(-D'))=0$$

pour tout diviseur D' sur C' de degré au plus n_0.

Si on applique la proposition en prenant pour n_0 le degré de la ramification de $u\circ f$, on obtient une courbe $f':C'\to X$ de même genre que C, le morphisme $u\circ f'$ est de même degré que $u\circ f$ et son diviseur de ramification est de degré n_0. La condition (C2) est donc réalisée. De plus, si C satisfait à la condition (ouverte) (C1), il en est de même de C'. Il suffira donc de construire une courbe $C\to X$ satisfaisant à la condition (C1) pour montrer le théorème 1.1.

Cela permet déjà de démontrer le théorème (dans le cas complexe) lorsque les fibres de u sont réduites. En effet, le lieu singulier de u est alors de codimension au moins 2 dans X et un simple décompte de paramètres montre que l'intersection d'un nombre adéquat de sections hyperplanes générales de X est une courbe lisse dans X entièrement contenue dans le lieu de lissité de u, simplement ramifiée au-dessus de B.

Preuve de la proposition. — Notons L la réunion (disjointe) des L_i et P le diviseur $p_1 + \cdots + p_n$; on a une suite exacte

$$0 \longrightarrow N_{L/X} \longrightarrow N_{\overline{C}/X} \otimes \mathscr{O}_L \longrightarrow \bigoplus_{i=1}^{n} T_{C,p_i} \longrightarrow 0 \tag{2}$$

Comme $N_{L/X}(-P)$ est positif sur chaque L_i, on en déduit l'annulation de $H^1(L, N_{\overline{C}/X} \otimes \mathscr{O}_L(-P))$ donc, grâce à la suite exacte

$$0 \longrightarrow \mathscr{O}_L(-P) \longrightarrow \mathscr{O}_{\overline{C}} \longrightarrow \mathscr{O}_C \longrightarrow 0$$

tensorisée avec $N_{\overline{C}/X}$, la surjectivité de la restriction

$$H^0(\overline{C}, N_{\overline{C}/X}) \longrightarrow H^0(C, N_{\overline{C}/X} \otimes \mathscr{O}_C)$$

et la bijectivité de

$$H^1(\overline{C}, N_{\overline{C}/X}) \longrightarrow H^1(C, N_{\overline{C}/X} \otimes \mathscr{O}_C) \tag{3}$$

Considérons d'autre part l'analogue

$$0 \longrightarrow N_{C/X} \longrightarrow N_{\overline{C}/X} \otimes \mathscr{O}_C \longrightarrow \bigoplus_{i=1}^{n} T_{L_i,p_i} \longrightarrow 0$$

de la suite exacte (2) ci-dessus. La déformation au premier ordre de $\overline{f}$ correspondant à un élément de $H^0(\overline{C}, N_{\overline{C}/X})$ « lisse » le point double p_i de $\overline{C}$ si et seulement si son image dans T_{L_i,p_i} n'est pas nulle. On a un diagramme

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N_{\overline{C}/X} \otimes \mathscr{O}_C(-p_i) & \longrightarrow & N_{\overline{C}/X} \otimes \mathscr{O}_C & \longrightarrow & N_{\overline{C}/X,p_i} & \longrightarrow & 0 \\ & & & & & \searrow & \downarrow & & \\ & & & & & & T_{L_i,p_i} & & \end{array}$$

Pour n assez grand, $H^1(C, N_{\overline{C}/X} \otimes \mathscr{O}_C(-p_i))$ est nul par le lemme ci-dessous. Cela signifie qu'une section générale de $N_{\overline{C}/X} \otimes \mathscr{O}_C$ a une image non nulle dans T_{L_i,p_i}, donc aussi dans chacun des T_{L_j,p_j} : une déformation au premier ordre générale de $\overline{C}$ dans X est lisse.

Lemme 4.3. — *Soit n_0 un entier. Pour n assez grand, on a*

$$H^1(C, N_{\overline{C}/X} \otimes \mathscr{O}_C(-D)) = 0$$

pour tout diviseur D sur C de degré $\leqslant n_0$.

Laissant provisoirement sa démonstration de côté, on déduit d'abord du lemme et de la bijectivité de (3) l'annulation de $H^1(\overline{C}, N_{\overline{C}/X})$, de sorte que pour une déformation générale $f' : C' \to X$ de $\overline{f}$, la courbe C' est lisse (de même genre g

que C). Appliquons de nouveau le lemme avec un diviseur D_0 de degré $g+n_0$ sur $C-\{p_1,\dots,p_n\}$; la courbe C' vérifie, par semi-continuité,

$$H^1(C', N_{C'/X}(-D'_0)) = 0$$

pour un certain diviseur D'_0 sur C' de degré $g+n_0$. Soit D' un diviseur sur C' de degré au plus n_0. Par le théorème de Riemann-Roch, on peut écrire $D'_0 \equiv D' + D''$, où D'' est un diviseur effectif, et la suite exacte

$$0 \longrightarrow N_{C'/X}(-D'_0) \longrightarrow N_{C'/X}(-D') \longrightarrow N_{C'/X}(-D') \otimes \mathscr{O}_{D''} \longrightarrow 0$$

entraîne $H^1(C', N_{C'/X}(-D')) = 0$, ce qui prouve la proposition. □

Preuve du lemme. — On a un diagramme commutatif de suites exactes

$$\begin{array}{ccccccccc}
0 & \longrightarrow & N_{C/X}(-D) & \longrightarrow & N_{\overline{C}/X}(-D)\otimes\mathscr{O}_C & \longrightarrow & \bigoplus_{i=1}^n T_{L_i,p_i} & \longrightarrow & 0 \\
 & & \| & & \cap & & \cap & & \\
0 & \longrightarrow & N_{C/X}(-D) & \longrightarrow & N_{C/X}(-D)\otimes\mathscr{O}_C(P) & \longrightarrow & \bigoplus_{i=1}^n N_{C/X,p_i} & \longrightarrow & 0
\end{array}$$

Il existe un entier n_1 tel que, pour tout $n \geqslant n_1$, on ait $H^1(C, N_{C/X}(P-D)\otimes\mathscr{O}_C) = 0$ pour tout diviseur D sur C de degré $\leqslant n_0$, de sorte que le cobord

$$\delta_n : \bigoplus_{i=1}^n N_{C/X,p_i} \longrightarrow H^1(C, N_{C/X}(-D))$$

est surjectif. Considérons des antécédents $(t_{1,j},\dots,t_{n_1,j})_{1\leqslant j\leqslant h}$ par δ_{n_1} d'une base $(\omega_j)_{1\leqslant j\leqslant h}$ de $H^1(C, N_{C/X}(-D))$; la restriction de $\delta_{n_1 h}$ à $\bigoplus_{i=1}^n \mathbf{C}t_{i,j}$ est surjective, donc aussi, pour $n \geqslant n_1 h$, la restriction de δ_n à $\bigoplus_{i=1}^n T_{L_i,p_i}$ puisque les points p_i et les directions T_{L_i,p_i} sont généraux. On en déduit $H^1(C, N_{\overline{C}/X}(-D)\otimes\mathscr{O}_C) = 0$, d'où le lemme. □

4.3. Où l'on attache des courbes rationnelles à une multisection pour créer de nouveaux points de ramification

On considère de nouveau une courbe projective et lisse C, de genre g, et un morphisme $f : C \to X$ non ramifié dont l'image domine B. Nous aurons besoin de créer de nouveaux points de ramification du morphisme $u\circ f : C \to B$. Cela se fait de la façon suivante.

Tout d'abord, la proposition 4.2 montre qu'étant donné un entier n_0, on peut supposer, quitte à remplacer C par une déformation de la réunion de C et de suffisamment de queues rationnelles très libres verticales,

$$H^1(C, N_{C/X}(-D)) = 0 \tag{4}$$

pour tout diviseur D sur C de degré $\leqslant n_0 + g + 1$. Étant donnés des points p_1 et p_2 de C dans une fibre générale de $u\circ f$, on peut relier leurs images par f par une courbe

rationnelle très libre non ramifiée $L \to X$ contenue dans la fibre correspondante de u (th. 2.2). On veut montrer qu'on peut déformer la courbe $\overline{C} = C \cup L \to X$ ainsi obtenue en une courbe lisse C' de genre $g+1$. Le même raisonnement qu'au § 4.2 montre que

$$H^0(\overline{C}, N_{\overline{C}/X}) \longrightarrow H^0(C, N_{\overline{C}/X} \otimes \mathscr{O}_C) \quad \text{est surjective,}$$
$$H^1(\overline{C}, N_{\overline{C}/X}) \longrightarrow H^1(C, N_{\overline{C}/X} \otimes \mathscr{O}_C) \quad \text{est bijective.}$$

On a d'autre part une suite exacte

$$0 \longrightarrow N_{C/X} \longrightarrow N_{\overline{C}/X} \otimes \mathscr{O}_C \longrightarrow T \longrightarrow 0.$$

où T est un sous-faisceau « diagonal » de $T_{C,p_1} \oplus T_{C,p_2}$. Comme dans le § 4.2, on déduit de (4) d'une part que $H^1(\overline{C}, N_{\overline{C}/X})$ est nul, d'autre part qu'une déformation générale $C' \to X$ de $\overline{C} \to X$ est lisse de genre $g+1$ et vérifie $H^1(C', N_{C'/X}(-D')) = 0$ pour tout diviseur D' sur C' de degré $\leqslant n_0$.

Étant donnés des entiers n_0 et a, on obtient, en itérant cette construction, une courbe non ramifiée $f' : C' \to X$ telle que :

(R1) la courbe C' est lisse de genre $g+a$;

(R2) le degré de $u \circ f'$ est le même que celui de $u \circ f$;

(R3) $H^1(C', N_{C'/X}(-D')) = 0$ pour tout diviseur D' sur C' de degré $\leqslant n_0$;

(R4) le diviseur de ramification de $u \circ f'$ a $2a$ paires de points de plus que celui de $u \circ f$ et pour chacun de ces points, la ramification est simple et la monodromie échange les deux feuillets.

4.4. Où l'on tue la ramification située sur les fibres multiples

Nous allons montrer qu'on peut construire par déformation, à partir d'une multisection $f : C \to X$ non ramifiée de u, une multisection qui vérifie la condition (C1) : sa ramification est simple et contenue dans le lieu lisse de u.

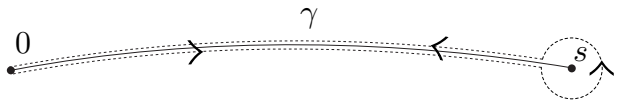
La construction du § 4.3 nous permet de créer autant de points de ramification simples qu'on le désire. Ceux-ci sont de plus mobiles ; en les faisant entrer en collision avec les points de ramification de $u \circ f$ situés sur les fibres non réduites, nous allons pouvoir « tuer » ces derniers.

Soit $S \subset B$ le lieu des points critiques de u. Si on choisit la multisection $f : C \to X$ de façon générale, f n'est pas ramifié et la ramification de $u \circ f$ hors de $(u \circ f)^{-1}(S)$ est simple. Parmi les courbes vérifiant ces propriétés, on choisit C de façon que le degré du diviseur de branchement D de $u \circ f$ dans S soit minimal. Soient g le genre de C et d le degré de $u \circ f$.

Choisissons un point 0 dans l'ouvert B^0 des points de $B - S$ au-dessus desquels $u \circ f$ n'est pas ramifiée. Pour tout b dans B^0, on pose $F_b = (u \circ f)^{-1}(b)$. On considère la représentation de monodromie

$$\rho : \pi_1(B^0, 0) \longrightarrow \mathrm{Aut}(F_0).$$

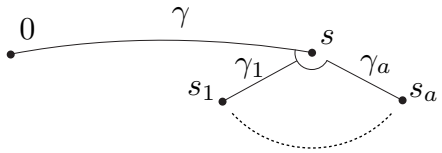
Choisissons un point s du support de D et un arc réel γ dans $B^0 \cup \{0, s\}$ joignant 0 à s. On définit la monodromie σ en s comme l'image par ρ du lacet en pointillé



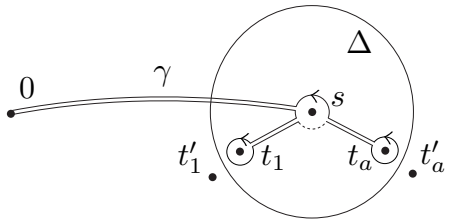
On la décompose en produit

$$\sigma = \tau_1 \cdots \tau_a$$

de transpositions (où a est l'ordre de s dans D). Choisissons des points $s_1, \ldots, s_a$ généraux proches de s ; les chemins $\gamma_1, \ldots, \gamma_a$ définis sur la figure



permettent d'identifier chaque F_{s_i} à F_0. La transposition τ_i échange deux points de F_0 ; pour chaque i, on relie les deux points correspondants de F_{s_i} par une courbe rationnelle très libre verticale. La construction du § 4.3 (avec $n_0 = 2(g+a)+2d-2$) permet d'obtenir par déformation une nouvelle multisection non ramifiée $f' : C' \to X$ de u vérifiant les propriétés (R1) à (R4), telle que $u \circ f'$ est simplement ramifié hors de S ; en particulier, pour chaque i, il y a deux points de branchement simples t_i et t'_i proches de s_i et, pour le choix de l'arc joignant 0 à t_i composé de l'arc γ_i suivi du « segment » $s_i t_i$, la monodromie en t_i est la transposition τ_i. L'image par la représentation de monodromie de C' du lacet



est $\sigma\tau_a \cdots \tau_1$, c'est-à-dire l'identité. La monodromie autour du bord du disque Δ est donc triviale. La condition (R3) entraîne que les déformations de $f' : C' \to X$ dans X sont paramétrées par une variété M lisse en $[f']$. Par minimalité du degré de D, le diviseur de branchement de $u \circ f'$ dans S est encore D et cela reste vrai pour une déformation générale de f'. L'application $\beta : M \to B^{(2g+2a+2d-2)}$ qui à une courbe associe son diviseur de branchement se factorise donc, au voisinage de $[f']$, en

$$\beta : M \xrightarrow{\beta'} D + B^{(2g+2a+2d-2-\deg(D))} \hookrightarrow B^{(2g+2a+2d-2)}.$$

La condition (R3) entraîne que la différentielle de β' est alors surjective. L'image de β' est donc dense, et celle de β l'est aussi dans $D' + \Delta^{(a)}$, où D' est le diviseur de branchement de $u \circ f'$ privé des points $t_1, \dots, t_a$. Il existe donc une suite $([f'_n])_n$ de points de M telle que les diviseurs de branchement des courbes correspondantes convergent vers $D' + as$. Une suite extraite converge dans l'espace compact $\overline{M}_{g(C')}(X, [C'])$ vers une courbe stable limite $f'_\infty : C'_\infty \to X$. Soit C_∞ la normalisation d'une composante de C'_∞ sur laquelle $u \circ f'_\infty$ n'est pas constante et soit $f_\infty : C_\infty \to X$ le morphisme induit par f'_∞.

Pour tout petit disque Δ_ε de centre s dans Δ, les morphismes $u \circ f'_n$, pour tout n assez grand, ne sont pas ramifiés au-dessus de $\Delta - \Delta_\varepsilon$ et la monodromie sur le bord de Δ étant triviale, le revêtement $(u \circ f'_n)^{-1}(\Delta - \Delta_\varepsilon) \to \Delta - \Delta_\varepsilon$ est trivial. Il s'ensuit que le seul point de branchement possible de $u \circ f_\infty$ dans Δ est s mais que, la monodromie étant triviale, il n'y a en fait pas de ramification au-dessus de s. La multisection f_∞ n'est donc pas ramifiée et son diviseur de branchement sur B est de degré strictement inférieur à celui de D, ce qui contredit les choix faits.

Cela signifie que la courbe $C \to X \to B$ n'est pas ramifiée au-dessus de S. Elle vérifie la condition (C1) et, comme on l'a vu au § 4.2, cela achève la démonstration du théorème.

Remarque 4.4. — Comme le remarquent les auteurs de [GHS], l'argument ci-dessus paraît un peu inquiétant : comment la courbe C', qui rencontre des composantes multiples de fibres, peut-elle dégénérer en une courbe qui les évite ? La réponse est simple : à la limite, la courbe C_∞ ne rencontre que des composantes réduites de fibres et la courbe C'_∞ contient des composantes verticales qui connectent C_∞ à ces composantes multiples. Considérons par exemple la courbe C_t d'équation $y^2 = x(x-t)$ dans $\mathbf{C}^2$; la première projection $u : \mathbf{C}^2 \to \mathbf{C}$ fait de C_t un revêtement double de $\mathbf{C}$. Éclatons l'origine dans $\mathbf{C}^2$, puis le point critique du morphisme induit par u. La fibre en 0 du morphisme $\widetilde{\mathbf{C}}^2 \to \mathbf{C}$ induit par u s'écrit $F + 2E + G$ où F est le transformé strict de $u^{-1}(0)$. Le transformé strict $\widetilde{C}_t$ de C_t rencontre E transversalement en un point de ramification qui converge, lorsque t tend vers 0, vers le point $E \cap G$. À la limite, la courbe $\widetilde{C}_0$ est la réunion de G et de deux composantes étales sur $\mathbf{C}$.

5. DÉMONSTRATION DU THÉORÈME SUR UN CORPS QUELCONQUE

La démonstration de de Jong et Starr du théorème 2.1, plus générale que la démonstration exposée ci-dessus, est aussi en un certain sens plus simple (mais moins géométrique et plus technique) ; elle ne fait pas appel à l'espace des modules des applications stables et construit la dégénérescence cherchée « à la main ».

5.1. Où l'on se ramène à la construction d'une multisection vérifiant certaines propriétés

On se donne un morphisme propre u d'une variété lisse X sur une courbe lisse B, dont les fibres générales sont des variétés lisses et séparablement rationnellement connexes. On peut supposer que la courbe B est projective et lisse et que la variété X est propre et normale. Il s'agit de montrer que u a une section.

Comme dans le cas complexe, on montre qu'il suffit pour cela de trouver une multisection de u « suffisamment mobile » ; plus précisément, on montre que le théorème 4.1 est encore valable (mais on a besoin de la connexité rationnelle des fibres de u).

THÉORÈME 5.1. — *Soient X une variété propre et B et C des courbes projectives et lisses. On se donne des morphismes $u : X \to B$, à fibres générales lisses et rationnellement connexes, et $f : C \to X$ tels que :*

(C1) *l'image de f est contenue dans le lieu lisse de u et $u \circ f$ est génériquement lisse ;*

(C2) *l'espace vectoriel $H^1(C, f^*T_{X/B})$ est nul.*

Alors u a une section.

Le faisceau tangent relatif $T_{X/B}$ est défini comme le noyau de l'application tangente $T_X \to u^*T_B$. Il est localement libre sur le lieu lisse de u ; si R est le diviseur de ramification de $u \circ f$, le faisceau $f^*T_{X/B}$ est isomorphe à $N_{C/X}(-R)$.

Preuve du théorème. — Pour éviter d'avoir recours aux schémas de Hurwitz (qui ne sont pas toujours irréductibles en caractéristique non nulle) et aux espaces de courbes stables, on construit « à la main » la dégénérescence d'un revêtement en un revêtement avec une section.

On se donne donc un morphisme fini $f : C \to B$ de degré d entre courbes projectives et lisses et un ensemble fini S de points de B qui contient le lieu de branchement de f. La proposition suivante montre que, quitte à identifier des paires de points de C situés dans les mêmes fibres de points hors de S, on peut déformer le revêtement obtenu $\mathscr{C}_0 \to B$ en un revêtement $\mathscr{C}_\infty \to B$ avec une section.

LEMME 5.2. — *Il existe une surface projective irréductible $\mathscr{C}$ et des morphismes surjectifs*

$$\begin{array}{ccc} \mathscr{C} & \xrightarrow{F} & B \\ {\scriptstyle p}\downarrow & & \\ \mathbf{P}^1 & & \end{array}$$

vérifiant, en notant $\mathscr{C}_t$, pour tout t dans $\mathbf{P}^1$, la fibre $p^{-1}(t)$,

(1) *la courbe $\mathscr{C}_t$ est lisse pour t général ;*

(2) *la surface $\mathscr{C}$ est lisse le long de $\mathscr{C}_0$, qui est une courbe intègre à points doubles ordinaires de normalisée C ; le morphisme F y induit f et $F(\mathrm{Sing}(\mathscr{C}_0)) \cap S = \varnothing$;*

(3) *la surface $\mathscr{C}$ est lisse en un point général d'une composante irréductible B' de $\mathscr{C}_\infty$, qui est envoyée isomorphiquement par F sur B ;*

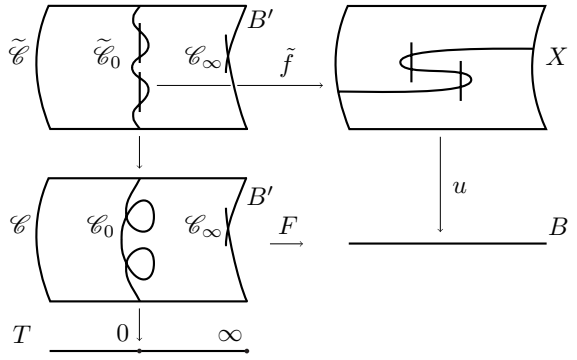
(4) *pour tout s dans S et tout $t \neq \infty$, les structures locales formelles des revêtements $F_t : \mathscr{C}_t \to B$ et $f : C \to B$ au voisinage de la fibre de s sont les mêmes, c'est-à-dire que l'on a un isomorphisme*

$$\mathscr{C}_t \times_B \mathrm{Spec}\, \widehat{\mathscr{O}}_{B,s} \simeq C \times_B \mathrm{Spec}\, \widehat{\mathscr{O}}_{B,s}$$

de $\mathrm{Spec}\, \widehat{\mathscr{O}}_{B,s}$-schémas.

Seul le point (4) n'est pas élémentaire. Il ne nous servira que plus tard. Remettant la démonstration du lemme à la fin de ce §, nous poursuivons celle du théorème.

Considérons un revêtement double $T \to \mathbf{P}^1$ ramifié en 0 et l'éclatement $\widetilde{\mathscr{C}} \to \mathscr{C} \times_{\mathbf{P}^1} T$ des points singuliers qui apparaissent au-dessus des points doubles ordinaires de $\mathscr{C}_0$. La fibre du morphisme $\widetilde{\mathscr{C}} \to T$ en le point, noté encore 0, de T situé au-dessus de 0 est réduite, réunion de C et de courbes rationnelles coupant chacune C transversalement en deux points. Notons $\widetilde{F} : \widetilde{\mathscr{C}} \to \mathscr{C} \to B$ le morphisme composé ; si on joint les paires de points de C identifiés dans $\mathscr{C}_0$ par des courbes rationnelles très libres non ramifiées contenues dans les fibres de u (th. 2.2), on définit un morphisme $\widetilde{f} : \widetilde{\mathscr{C}_0} \to X$. On a ainsi la figure



La fibre de $\widetilde{\mathscr{C}} \times_B X$ en 0 contient le graphe Γ de $\widetilde{f}$, dont les déformations comme sous-T-schéma de $\widetilde{\mathscr{C}} \times_B X$ sont paramétrées par une composante du T-schéma de Hilbert relatif $\mathrm{Hilb}(\widetilde{\mathscr{C}} \times_B X/T)$.

Les déformations au premier ordre sont paramétrées par l'espace vectoriel des sections du fibré normal $N_{\Gamma/(\widetilde{\mathscr{C}} \times_B X)_0}$ et les obstructions sont dans $H^1(\Gamma, N_{\Gamma/(\widetilde{\mathscr{C}} \times_B X)_0})$. Ce fibré s'identifie à $\widetilde{f}^* T_{X/B}$ sur $\widetilde{\mathscr{C}_0}$. Par construction, la restriction de ce fibré aux queues rationnelles est ample ; on montre comme dans le § 4.3 que l'annulation de $H^1(C, f^* T_{X/B})$ entraîne alors celle de $H^1(\widetilde{\mathscr{C}_0}, \widetilde{f}^* T_{X/B})$.

Le morphisme propre $\mathrm{Hilb}(\widetilde{\mathscr{C}} \times_B X/T) \to T$ est alors lisse en $[\Gamma]$. Quitte à effectuer un changement de base fini sur T, on peut donc supposer qu'il a une section, c'est-à-dire qu'il existe un sous-schéma Γ' de $\widetilde{\mathscr{C}} \times_B X$, plat sur T, tel que $\Gamma'_0 = \Gamma$. Dans le diagramme

$$\begin{array}{ccccc} \Gamma' & \subset & \widetilde{\mathscr{C}} \times_B X & \longrightarrow & \widetilde{\mathscr{C}} \\ & \searrow & & \swarrow & \\ & & T & & \end{array}$$

la ligne horizontale est un isomorphisme au-dessus du point 0 de T, donc est birationnelle. Son inverse $\widetilde{\mathscr{C}} \dashrightarrow \Gamma'$ est défini hors du complémentaire d'un ensemble fini dans le lieu normal de $\widetilde{\mathscr{C}}$, donc en le point générique de B'. Dans le diagramme commutatif

$$\begin{array}{ccccc} \Gamma' & \hookrightarrow & \widetilde{\mathscr{C}} \times_B X & \longrightarrow & X \\ & \nwarrow & \downarrow & & \downarrow u \\ B' & \hookrightarrow & \widetilde{\mathscr{C}} & \xrightarrow{\widetilde{F}} & B \\ \downarrow & & \downarrow & & \\ \{\infty\} & \hookrightarrow & T & & \end{array}$$

l'image de B' dans X est une section de u et ceci démontre le théorème 5.1. □

Revenons maintenant à la démonstration du lemme 5.2.

Preuve du lemme 5.2. — Soient m un grand entier et $g : C \to \mathbf{P}^1$ le morphisme associé à deux sections générales de $\mathscr{O}_C(mf^*S)$. On vérifie que :

- le morphisme $(f, g) : C \to B \times \mathbf{P}^1$ est birationnel sur son image C_0 ;
- les singularités de la courbe C_0 sont des points doubles ordinaires situés hors de $S \times \mathbf{P}^1$;
- la courbe C_0 est dans le système linéaire $|p_1^*\mathscr{O}_C(dmS) \otimes p_2^*\mathscr{O}_{\mathbf{P}^1}(d)|$.

Ce système linéaire contient aussi le diviseur

$$C_\infty = (dmS \times \mathbf{P}^1) + (B \times \{u_1, \ldots, u_d\}),$$

où $u_1, \ldots, u_d$ sont des points généraux de $\mathbf{P}^1$. Considérons l'application rationnelle

$$B \times \mathbf{P}^1 \dashrightarrow \mathbf{P}^1$$

définie par le pinceau (C_0, C_∞). Si $\mathscr{C} \to B \times \mathbf{P}^1$ est l'éclatement du schéma $C_0 \cap C_\infty$ et $p : \mathscr{C} \to B \times \mathbf{P}^1 \dashrightarrow \mathbf{P}^1$ et $F : \mathscr{C} \to B \times \mathbf{P}^1 \to B$ les morphismes composés, on vérifie facilement les points (1), (2) et (3).

Soient s un point de S et π une uniformisante de l'anneau de valuation discrète $\mathscr{O}_{B,s}$. Dans $\operatorname{Spec} \mathscr{O}_{B,s} \times \mathbf{P}^1$, la courbe C_0 est définie par une équation σ_0 (avec

$\sigma_0(s,u_i)\neq 0$ pour chaque i), la courbe C_∞ par l'équation $\pi^{md}\prod(u-u_i)$, donc la courbe $\mathscr{C}_t$ par $\sigma_0+t\pi^{md}\prod(u-u_i)$. Modulo π^{md}, les fibres en s des morphismes $\mathscr{C}_t\to B$ induits par F sont donc isomorphes pour $t\neq\infty$. Comme ceux-ci sont finis (cela résulte de la forme explicite de l'« équation » de $\mathscr{C}_t$ dans $B\times\mathbf{P}^1$), le lemme suivant entraîne (4). □

LEMME 5.3. — *Soient A un anneau de valuation discrète complet, d'uniformisante π et corps des fractions K. Soient K' une extension séparable de K et A' la fermeture intégrale de A dans K'. Il existe un entier positif m_0 tel que, pour tout anneau de valuation discrète A'' fini sur A, les conditions suivantes sont équivalentes :*

(i) *les A-algèbres A' et A'' sont isomorphes ;*

(ii) *il existe un entier $m>m_0$ tel que les A-algèbres A'/π^mA' et A''/π^mA'' sont isomorphes.*

Preuve. — Comme l'extension K' de K est séparable, il existe un élément a' de A' tel que $K'=K(a')$; son polynôme minimal P est à coefficients dans A et vérifie $P'(a')\neq 0$. Il existe donc un entier positif m_1 tel que $P'(a')\in\pi^{m_1}A'-\pi^{m_1+1}A'$. On pose $m_0=2m_1$.

Montrons (ii) $\Rightarrow$ (i). Comme l'anneau A'' est la fermeture intégrale de A dans le corps des fractions K'' de A'', il suffit de montrer que les K-extensions K' et K'' sont isomorphes. Les A-modules A' et A'' sont libres, de même rang puisque $A'/\pi A'\simeq A''/\pi A''$. Les K-extensions K' et K'' ont donc même degré et il suffit de montrer qu'il existe un K-homomorphisme de K' dans K''.

Soit a'' un élément de A'' relevant l'image de a' par la composée $A'\to A'/\pi^mA'\simeq A''/\pi^mA''$. On a alors

$$P(a'')\in\pi^mA''\quad\text{et}\quad P'(a'')\in\pi^{m_1}A''-\pi^{m_1+1}A''$$

L'anneau A'' est un anneau de valuation discrète complet pour la topologie π-adique. D'après le lemme de Hensel ([Bo], chap. III, § 4, n° 5, cor. 1), il existe $b''\in A''$ tel que $P(b'')=0$ et $b''\equiv a''\pmod{\pi^{m-m_1}}$. Le morphisme $A[X]/(P)\to A''$ qui s'en déduit induit le K-homomorphisme $K'\to K''$ cherché. □

5.2. Où l'on se ramène au cas où les fibres de u sont réduites

Pour éviter le délicat argument de monodromie employé dans le cas complexe lorsque u a des fibres non réduites, de Jong et Starr démontrent (c'est l'objet de la proposition ci-dessous) un résultat *a priori* surprenant, à savoir qu'il suffit de montrer l'existence d'une section après tout changement de base sur B tel que les fibres de la famille induite soient réduites. L'hypothèse que le corps de base $\boldsymbol{k}$ n'est pas dénombrable n'est pas un problème[(17)] : on peut par exemple remarquer que les sections de $u:X\to B$ sont paramétrées par un $\boldsymbol{k}$-schéma $S(u)$ et que si $\boldsymbol{k}'$ est une extension

[(17)] *Cf.* aussi note 18.

algébriquement close de $\boldsymbol{k}$, les sections de $u_{\boldsymbol{k}'} : X_{\boldsymbol{k}'} \to B_{\boldsymbol{k}'}$ sont paramétrées par le $\boldsymbol{k}'$-schéma $S(u)_{\boldsymbol{k}'}$: pour montrer que $S(u)$ n'est pas vide, on peut le faire après extension (algébriquement close non dénombrable) du corps de base.

Lorsque les fibres de u sont réduites et que X est projectif, le théorème de Bertini montre que dans ce cas, il existe une section vérifiant (C1). La condition (C2) est obtenue par la proposition 4.2, et le théorème 2.1 résulte alors du théorème 5.1. Lorsque X n'est que propre, il faut un argument différent, pour lequel je renvoie à [dJS].

PROPOSITION 5.4. — *Soient X une variété propre et normale, B une courbe projective et lisse et $u : X \to B$ un morphisme dont les fibres générales sont réduites, définis sur un corps algébriquement clos non dénombrable*[(18)].

(1) *Il existe une courbe projective et lisse C et un morphisme fini $C \to B$ génériquement lisse tels que les fibres du morphisme induit*

$$u_C : (C \times_B X)^{\mathrm{norm}} \longrightarrow C$$

soient réduites.

(2) *Si, pour tout morphisme fini génériquement lisse $C \to B$ de courbes lisses tel que les fibres de u_C soient réduites, u_C a une section, alors u a une section.*

Preuve. — Le point (1) se trouve en fait déjà dans la littérature ([Ep] ; [BLR], th. 2.1′, p. 368) et nous ne démontrerons donc que (2).

Considérons un morphisme $f : C \to B$ satisfaisant les propriétés de (1) et appliquons-lui le lemme 5.2, en prenant pour S l'ensemble des points de branchement de f et des points de B dont la fibre sous u n'est pas réduite. Il fournit un diagramme

$$\begin{array}{ccc} \mathscr{C} \times_B X & \longrightarrow & X \\ \downarrow & & \downarrow u \\ \mathscr{C} & \xrightarrow{F} & B \\ p\downarrow & & \\ \mathbf{P}^1 & & \end{array}$$

Considérons les fibres de $u_{\mathscr{C}_t} : (\mathscr{C}_t \times_B X)^{\mathrm{norm}} \to \mathscr{C}_t$ pour $t \neq \infty$. Hors de S, elles sont réduites puisque celles de u l'étaient déjà. Au-dessus d'un point de S, les fibres formelles sont réduites puisqu'elles sont isomorphes à celles de u_C. Toutes les fibres de $u_{\mathscr{C}_t}$ sont donc réduites. Pour t général, $\mathscr{C}_t$ est lisse et il existe par hypothèse une

(18) Cette dernière hypothèse, que l'on n'utilise que pour (2), peut être supprimée si l'on remplace l'hypothèse de (2) par : pour toute extension algébriquement close $\boldsymbol{k}'$ de $\boldsymbol{k}$ et tout morphisme fini génériquement lisse $C \to B_{\boldsymbol{k}'}$ de $\boldsymbol{k}'$-courbes lisses, tels que les fibres de u_C soient réduites, u_C a une section.

section σ_t de $\mathscr{C}_t \times_B X \to \mathscr{C}_t$. Comme le corps de base n'est pas dénombrable et que le schéma de Hilbert relatif $\mathrm{Hilb}(\mathscr{C} \times_B X/\mathbf{P}^1) \to \mathbf{P}^1$, qui est propre, n'a qu'un nombre dénombrable de composantes irréductibles, l'une d'elle contient presque tous les points correspondants aux images des sections σ_t, donc contient une courbe propre irréductible T qui domine $\mathbf{P}^1$.

Pour alléger les notations, nous désignerons par un indice T le produit $T \times_{\mathbf{P}^1} \cdot$. Il existe donc un sous-schéma fermé Γ de $(\mathscr{C} \times_B X)_T$, plat sur T, tel que la projection $\Gamma_t \to \mathscr{C}_t$ soit un isomorphisme pour $t \in T$ très général, de sorte que le morphisme $\Gamma \to \mathscr{C}_T$ est birationnel. Comme $\mathscr{C}$, donc aussi $\mathscr{C}_T$, est lisse en un point général de B', l'inverse $\mathscr{C}_T \dashrightarrow \Gamma$ est défini en un point général de B'. On retrouve un diagramme commutatif analogue à celui de la page 261 :

$$\begin{array}{ccccccc}
\Gamma & \hookrightarrow & (\mathscr{C} \times_B X)_T & \longrightarrow & \mathscr{C} \times_B X & \longrightarrow & X \\
 & \nwarrow & \downarrow & & \downarrow & & \downarrow u \\
B' & \hookrightarrow & \mathscr{C}_T & \longrightarrow & \mathscr{C} & \xrightarrow{F} & B \\
\downarrow & & \downarrow & & \downarrow & & \\
\{\infty\} & \hookrightarrow & T & \longrightarrow & \mathbf{P}^1 & &
\end{array}$$

où de nouveau, l'image de B' dans X est une section de u. □

RÉFÉRENCES

[BM] K. Behrend & Y. Manin – « Stacks of stable maps and Gromov-Witten invariants », *Duke Math. J.* **85** (1996), p. 1–60.

[B] A. Borel – *Linear algebraic groups*, Benjamin, 1969.

[BLR] S. Bosch, W. Lütkebohmert & M. Raynaud – « Formal and rigid geometry. IV. The reduced fibre theorem », *Invent. Math.* **119** (1995), p. 361–398.

[Bo] N. Bourbaki – *Éléments de mathématique. Algèbre commutative. Chapitres 1 à 4*, Masson, Paris, 1985.

[C1] F. Campana – « Connexité rationnelle des variétés de Fano », *Ann. Sc. Éc. Norm. Sup.* **25** (1992), p. 539–545.

[C2] ———, « Coréduction algébrique d'un espace analytique compact faiblement kählérien », *Invent. Math.* **63** (1981), p. 187–223.

[CPP] F. Campana, T. Peternell & A. Pukhlikov – « Generalized Tsen theorem and rationally connected Fano fibrations », prépublication `math.AG/0110017`.

[Ch] C. Chevalley – « Démonstration d'une hypothèse de E. Artin », *Abh. Math. Sem. Hansischen Univ.* **119** (1935), p. 73.

[Cl] A. Clebsch – « Zur Theorie der Riemann'schen Fläche », *Math. Ann.* **6** (1872), p. 216–230.

[CT] J.-L. Colliot-Thélène – « Arithmétique des variétés rationnelles et problèmes birationnels », in *Proc. Int. Congr. Math. (Berkeley, Calif. 1986)*, A.M.S., Providence, RI, 1987, p. 641–653.

[GS] P. Colmez & J.-P. Serre – *Correspondance Grothendieck-Serre*, Documents Mathématiques, vol. 2, Soc. Math. France, Paris, 2001.

[D1] O. Debarre – « Variétés de Fano », in *Séminaire Bourbaki 1996/97*, Astérisque, vol. 245, Soc. Math. France, Paris, 1997, Exposé 827, p. 197–221.

[D2] _____, *Higher-Dimensional Algebraic Geometry*, Universitext, Springer-Verlag, 2001.

[E] T. Ekedahl – « Sur le groupe fondamental d'une variété unirationnelle », *C. R. Acad. Sci. Paris* **297** (1983), p. 627–629.

[En] F. Enriques – « Sopra le superficie algebriche che contengono un fascio di curve razionali », *Math. Ann.* **52** (1899), p. 449–456.

[Ep] H. Epp – « Eliminating wild ramification », *Invent. Math.* **19** (1973), p. 235–249.

[Es] H. Esnault – « Varieties over a finite field with trivial Chow group of 0-cycles have a rational point », *Invent. Math.* **151** (2003), p. 187–191.

[F] W. Fulton – « Hurwitz schemes and irreducibility of moduli of algebraic curves », *Ann. of Math.* **90** (1969), p. 542–575.

[FP] W. Fulton & R. Pandharipande – « Notes on stable maps and quantum cohomology », in *Algebraic geometry (Santa Cruz 1995), Part 2*, Proc. Sympos. Pure Math., vol. 62, Amer. Math. Soc., Providence, RI, 1997, p. 45–96.

[GHMS] T. Graber, J. Harris, B. Mazur & J. Starr – « Rational connectivity and sections of families over curves », prépublication `math.AG/0210225`.

[GHS] T. Graber, J. Harris & J. Starr – « Families of Rationally Connected Varieties », *J. Amer. Math. Soc* **16** (2003), p. 57–67.

[H] A. Hurwitz – « Ueber Riemann'sche Flächen mit gegebenen Verzweigungspunkten », *Math. Ann.* **39** (1891), p. 1–61.

[dJS] A.J. de Jong & J. Starr – « Every rationally connected variety over the function field of a curve has a rational point », *Amer. J. Math.* **125** (2003), p. 567–580.

[Ko1] J. Kollár – *Rational Curves on Algebraic Varieties*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 32, Springer-Verlag, Berlin, 1996.

[Ko2] _____, « Nonrational hypersurfaces », *J. Amer. Math. Soc.* **8** (1995), p. 241–249.

[KMM1] J. Kollár, Y. Miyaoka & S. Mori – « Rational Curves on Fano Varieties », in *Proc. Alg. Geom. Conf. Trento*, Springer Lecture Notes, vol. 1515, Springer-Verlag, 1992, p. 100–105.

[KMM2] _____, « Rationally Connected Varieties », *J. Alg. Geom.* **1** (1992), p. 429–448.

[M] Y. Manin – *Cubic forms. Algebra, geometry, arithmetic*, 2ème éd., North-Holland Mathematical Library, vol. 4, North-Holland Publishing Co., Amsterdam-New York, 1986.

[N] M. Noether – « Ueber Flächen, welche Schaaren rationaler Curven besitzen », *Math. Ann.* **3** (1870), p. 161–226.

[Ny] N. Nygaard – « On the fundamental group of a unirational 3-fold », *Invent. Math.* **44** (1978), p. 75–86.

[S] J.-P. Serre – *Cohomologie Galoisienne*, Springer Lecture Notes, vol. 5, Springer-Verlag, Berlin, 1964.

[SB] N. Shepherd-Barron – « Fano threefolds in positive characteristic », *Compositio Math.* **105** (1997), p. 237–265.

[Sh] N. Shioda – « On unirationality of supersingular surfaces », *Math. Ann.* **225** (1977), p. 155–159.

[Su] N. Suwa – « A note on the fundamental group of a unirational variety », *Proc. Japan Acad. Ser. A Math. Sci.* **59** (1983), p. 98–99.

[T] C. Tsen – « Quasi-algebraisch-abgeschlossene Funktionenkörper », *J. Chinese Math.* **1** (1936), p. 81–92.

Olivier DEBARRE

Institut de Recherche Mathématique Avancée
Université Louis Pasteur
7, rue René Descartes
F-67084 Strasbourg Cedex
E-mail : debarre@math.u-strasbg.fr

Séminaire BOURBAKI Juin 2002
54e année, 2001-2002, nO 906, p. 267 à 284

TRAVAUX DE FRENKEL, GAITSGORY ET VILONEN SUR LA CORRESPONDANCE DE DRINFELD–LANGLANDS

par **Gérard LAUMON**

En 1967, R. Langlands a proposé une vaste extension de la théorie du corps de classes abélien de E. Artin et J. Tate. Plus précisément, il a conjecturé une correspondance naturelle entre représentations automorphes d'un groupe réductif G sur un corps global F et représentations galoisiennes de F à valeurs dans le groupe algébrique ${}^{\mathrm{L}}G$ dual de G. La composante neutre $\widehat{G}$ de ${}^{\mathrm{L}}G$ est le groupe réductif complexe dont les racines sont les co-racines de G et vice-versa.

Si G est le groupe linéaire $\mathrm{GL}(n)$ sur F, $\widehat{G}$ n'est autre que $\mathrm{GL}(n,\mathbb{C})$ et la correspondance de Langlands globale a été démontrée par V. Drinfeld [DR1] ($n=2$) et L. Lafforgue [La] (n arbitraire) lorsque F un corps de fonctions, c'est-à-dire une extension finie de $\mathbb{F}_p(t)$. La correspondance de Langlands globale sur les corps de nombres reste un des grands problèmes ouverts en mathématiques.

La *correspondance de Drinfeld-Langlands*, dite aussi *de Langlands géométrique*, est un analogue conjectural de la correspondance de Langlands pour un groupe réductif déployé G sur une extension finie F de $k(t)$, où k est un corps arbitraire. Si X est une courbe algébrique quasi-projective et lisse sur k, de corps des fonctions F, cette correspondance met en dualité un espace de modules de G-fibrés sur X et un espace de modules de $\widehat{G}$-systèmes locaux sur X.

Pour $G=\mathrm{GL}(1)$ la correspondance de Drinfeld-Langlands n'est autre que la théorie du corps de classes géométrique de M. Rosenlicht et S. Lang, exposée par J.-P. Serre dans [Se]. Le cas $G=\mathrm{GL}(2)$ a été traité par V. Drinfeld dans les articles [DR2] et [DR3] qui sont à l'origine de la théorie.

De nombreux travaux ont été consacrés à divers aspects de la correspondance de Drinfeld-Langlands, en particulier ceux de A. Beilinson et V. Drinfeld ([B-D]), de A. Braverman et D. Gaitsgory ([B-G]), et de S. Lysenko ([Ly1] et [Ly2]). Dans cet exposé je n'évoquerai que les travaux récents de E. Frenkel, D. Gaitsgory et K. Vilonen dans le cas où X est projective (correspondance partout non ramifiée) et $G=\mathrm{GL}(n)$, travaux qui généralisent ceux de V. Drinfeld dans [DR2].

Je remercie S. Lysenko pour son aide dans la préparation de cet exposé.

0. PRÉLIMINAIRES

Dans le cas partout non ramifié qui fait l'objet de cet exposé, on espère établir une correspondance de Drinfeld-Langlands pour chaque triplet formé d'un corps de base k, d'un corps des coefficients C et d'une théorie cohomologique à coefficients dans C pour la catégorie des schémas de type fini sur k.

Les trois triplets principaux (corps de base, corps des coefficients, théorie cohomologique) sont :

(Betti) $k = \mathbb{C}$, C algébriquement clos de caractéristique 0 et la théorie des faisceaux constructibles de C-espaces vectoriels pour la topologie classique,

(De Rham) k de caractéristique nulle, $C = k$ et la théorie des $\mathcal{D}$-Modules holonomes pour la topologie de Zariski,

(ℓ-adique) k arbitraire, $C = \overline{\mathbb{Q}}_\ell$ pour un nombre premier ℓ inversible dans k et la théorie des faisceaux ℓ-adiques pour la topologie étale.

La donnée première de la correspondance de Drinfeld-Langlands est celle d'une courbe algébrique sur le corps de base k. À cette courbe on attache des espaces de modules qui sont en général des champs algébriques sur k. J'utiliserai donc librement le langage des champs algébriques (*cf.* [L-M]).

Les théories cohomologiques ci-dessus n'ont pas été développées de manière systématique pour la catégorie des champs algébriques et les références sont parcellaires. La situation est assez satisfaisante pour le triplet (ℓ-adique) utilisé par Frenkel, Gaitsgory et Vilonen dans leurs articles (en fait, ils supposent de plus que k est de caractéristique $p > 0$ pour disposer d'une transformation de Fourier géométrique, mais cette restriction n'est pas nécessaire). Le triplet (De Rham) est utilisé dans les travaux [B-D] de Beilinson et Drinfeld où on trouvera une définition de l'anneau des opérateurs différentiels pour un champ algébrique. Le triplet (Betti), qui est en principe le plus élémentaire, a été étudié par Bernstein et Luntz [B-L], mais la théorie des champs analytiques reste à écrire.

Dans cet exposé, j'utiliserai néanmoins ce dernier triplet (Betti). Le corps de base sera donc $k = \mathbb{C}$ et, pour tout champ algébrique $\mathcal{S}$ de type fini, ou plus généralement localement de type fini (sur $\mathbb{C}$), je noterai simplement $D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{S})$ la catégorie dérivée des complexes de faisceaux de C-espaces vectoriels à cohomologie bornée et constructible sur le champ analytique associé à $\mathcal{S}$. Lorsque $\mathcal{S}$ est présenté comme un quotient d'un schéma S par l'action d'un groupe algébrique G, ce champ analytique n'est autre que le champ quotient de $S(\mathbb{C})^{\mathrm{an}}$ par l'action de $G(\mathbb{C})^{\mathrm{an}}$ et $D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{S})$ est la catégorie dérivée $D^{\mathrm{b}}_{G(\mathbb{C})^{\mathrm{an}}}(S(\mathbb{C})^{\mathrm{an}})$ introduite par Bernstein et Luntz dans [B-L].

Pour tous les champs algébriques $\mathcal{S}$ considérés dans ce texte, la catégorie dérivée $D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{S})$ est munie d'un produit tensoriel et d'un foncteur de dualité $D : D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{S})^{\mathrm{opp}} \to D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{S})$; pour tous les morphismes représentables $f : \mathcal{S} \to \mathcal{T}$ considérés entre tels champs, on a des foncteurs images directes $f_*, f_! : D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{S}) \to D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{T})$ et images inverses

$f^*, f^! : D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{T}) \to D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{S})$ satisfaisant au formalisme des six opérations de Grothendieck. La catégorie $D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{S})$ est munie de la t-structure pour la perversité intermédiaire dont le cœur est la catégorie $\mathrm{Perv}(\mathcal{S})$ des faisceaux pervers sur $\mathcal{S}$ (*cf.* [B-B-D]). Rappelons que cette sous-catégorie pleine de $D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{S})$ est abélienne, noethérienne, artinienne et auto-duale pour la dualité D.

Si $\pi : \mathcal{V} \to \mathcal{S}$ est un fibré vectoriel de rang constant r, on peut considérer le $\mathcal{S}$-champ quotient $\overline{\mathcal{V}}$ de $\mathcal{V}$ par l'action par homothétie du groupe multiplicatif $\mathbb{G}_{\mathrm{m}}$. Si $\mathcal{V}^\circ$ est l'ouvert complémentaire dans $\mathcal{V}$ de la section nulle, l'inclusion $\mathcal{V}^\circ \subset \mathcal{V}$ induit une immersion ouverte

$$j : \mathbb{P}(\mathcal{V}) = [\mathcal{V}^\circ/\mathbb{G}_{\mathrm{m}}] = \overline{\mathcal{V}}^{\,\circ} \hookrightarrow \overline{\mathcal{V}}$$

où $\mathbb{P}(\mathcal{V}) \to \mathcal{S}$ est le fibré projectif des droites de $\mathcal{V} \to \mathcal{S}$. Le fermé complémentaire de cette immersion ouverte est le quotient de la section nulle de $\mathcal{V}$ par l'action triviale de $\mathbb{G}_{\mathrm{m}}$, c'est-à-dire le champ classifiant $B(\mathbb{G}_{\mathrm{m}}/\mathcal{S})$.

Les complexes $\mathbb{G}_{\mathrm{m}}$-équivariants constructibles de C-espaces vectoriels sur $\mathcal{V}$ sont par définition les objets de $D^{\mathrm{b}}_{\mathrm{c}}(\overline{\mathcal{V}})$.

On a une transformation de Fourier géométrique pour ces complexes $\mathbb{G}_{\mathrm{m}}$-équivariants, appelée la *transformation de Fourier homogène* associée au fibré vectoriel $\mathcal{V}/\mathcal{S}$ (*cf.* [Lau3]). Cette transformation de Fourier est une équivalence de catégories dérivées

$$\mathrm{Four}_{\overline{\mathcal{V}}/\mathcal{S}} : D^{\mathrm{b}}_{\mathrm{c}}(\overline{\mathcal{V}}) \longrightarrow D^{\mathrm{b}}_{\mathrm{c}}(\overline{\mathcal{V}}^\vee),$$

d'inverse $\mathrm{Four}_{\overline{\mathcal{V}}^\vee/\mathcal{S}}$, où $\overline{\mathcal{V}}^\vee$ est le $\mathcal{S}$-champ quotient du fibré vectoriel dual $\pi^\vee : \mathcal{V}^\vee \to \mathcal{S}$ par l'action par homothétie du groupe multiplicatif $\mathbb{G}_{\mathrm{m}}$. Elle commute à la dualité et elle est t-exacte. On peut la définir de la façon suivante.

Sur le champ quotient $\mathcal{A} = [\mathbb{A}^1/\mathbb{G}_{\mathrm{m}}]$ quotient de la droite affine par l'action par homothétie du groupe multiplicatif, on a le complexe $\Psi = \beta_* C \in \mathrm{ob}\, D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{A})$, où $\beta : \mathrm{Spec}(\mathbb{C}) = [\mathbb{G}_{\mathrm{m}}/\mathbb{G}_{\mathrm{m}}] \hookrightarrow \mathcal{A}$ est l'immersion ouverte induite par l'inclusion $\mathbb{G}_{\mathrm{m}} \subset \mathbb{A}^1$. Ce complexe admet pour faisceaux de cohomologie non triviaux

$$\mathcal{H}^0(\Psi) = C$$

et

$$\mathcal{H}^1(\Psi) = \alpha_* C$$

où $\alpha : B(\mathbb{G}_{\mathrm{m}}) \hookrightarrow \mathcal{A}$ est l'immersion fermée complémentaire de β induite par l'inclusion de l'origine dans $\mathbb{A}^1$, et c'est en fait un faisceau pervers non irréductible, extension du faisceau pervers ponctuel $\alpha_* C[-1]$ par le faisceau pervers constant C.

Le morphisme d'accouplement naturel entre le fibré vectoriel $\mathcal{V}$ et son fibré dual passe au quotient en un morphisme de champs algébriques

$$\mu : \overline{\mathcal{V}}^\vee \times_{\mathcal{S}} \overline{\mathcal{V}} \longrightarrow \mathcal{A}$$

et, si on note $p : \overline{\mathcal{V}}^\vee \times_{\mathcal{S}} \overline{\mathcal{V}} \to \overline{\mathcal{V}}$ et $p^\vee : \overline{\mathcal{V}}^\vee \times_{\mathcal{S}} \overline{\mathcal{V}} \to \overline{\mathcal{V}}^\vee$ les deux projections canoniques, on a par définition

$$\mathrm{Four}_{\overline{\mathcal{V}}/\mathcal{S}}(K) = (p^\vee)_!(p^*K \otimes \mu^*\Psi)[r-1],\ \forall K \in D^{\mathrm{b}}_{\mathrm{c}}(\overline{\mathcal{V}}). \tag{0.1}$$

Les morphismes p et $p^\vee$ ne sont pas représentables et $(p^\vee)_!$ ne respecte pas $D^{\mathrm{b}}_{\mathrm{c}}$: il envoie $D^{\mathrm{b}}_{\mathrm{c}}(\overline{\mathcal{V}}^\vee \times_{\mathcal{S}} \overline{\mathcal{V}})$ dans $D^-_{\mathrm{c}}(\overline{\mathcal{V}}^\vee)$. Cependant, on peut vérifier que $\mathrm{Four}_{\overline{\mathcal{V}}/\mathcal{S}}$ respecte lui $D^{\mathrm{b}}_{\mathrm{c}}$.

Définition 0.2. — *Un faisceau pervers L sur $\mathbb{P}(\mathcal{V})$ est dit* propre *si la flèche d'oubli des supports $j_!L \to j_*L$ est un isomorphisme. Si tel est le cas $j_!L \cong j_*L$ est en fait un faisceau pervers sur $\overline{\mathcal{V}}$ qui n'est autre que le prolongement intermédiaire $j_{!*}L$ de L.*

Lemme 0.3. — *Soit L un faisceau pervers sur $\mathbb{P}(\mathcal{V})$. Alors, L est propre si et seulement si $(\overline{\pi}^\circ)_!L = 0$ où $\overline{\pi}^\circ : \mathbb{P}(\mathcal{V}) \to \mathcal{S}$ est la projection canonique.* □

La transformation de Radon géométrique associée au fibré projectif $\mathbb{P}(\mathcal{V}) \to \mathcal{S}$ est le foncteur

$$\mathrm{Rad}_{\mathbb{P}(\mathcal{V})/\mathcal{S}} : D^{\mathrm{b}}_{\mathrm{c}}(\mathbb{P}(\mathcal{V})) \longrightarrow D^{\mathrm{b}}_{\mathrm{c}}(\mathbb{P}(\mathcal{V}^\vee))$$

défini par

$$\mathrm{Rad}_{\mathbb{P}(\mathcal{V})/\mathcal{S}}(L) = (q^\vee)_! q^* L[r-1]$$

où $q : \mathcal{H} \to \mathbb{P}(\mathcal{V})$ et $q^\vee : \mathcal{H} \to \mathbb{P}(\mathcal{V}^\vee)$ sont les deux projections canoniques du champ d'incidence

$$\mathcal{H} \subset \mathbb{P}(\mathcal{V}^\vee) \times_{\mathcal{S}} \mathbb{P}(\mathcal{V})$$

des couples formés d'une droite et d'un hyperplan de $\mathcal{V} \to \mathcal{S}$ tels que la droite soit contenue dans l'hyperplan. On renvoie à la monographie de Brylinski [Br] pour une étude détaillée de cette transformation.

Lemme 0.4. — *Soit L un faisceau pervers sur $\mathbb{P}(\mathcal{V})$ que l'on suppose propre. Alors $\mathrm{Four}_{\overline{\mathcal{V}}/\mathcal{S}}(j_!L)$ est aussi propre et sa restriction à l'ouvert $\mathbb{P}(\mathcal{V}^\vee) = \overline{\mathcal{V}}^{\vee\circ}$ complémentaire de la section nulle dans $\overline{\mathcal{V}}^\vee$ est égale à $\mathrm{Rad}_{\mathbb{P}(\mathcal{V})/\mathcal{S}}(L)$.* □

1. CORPS DE CLASSES GÉOMÉTRIQUE

Soit X une courbe algébrique connexe, projective et lisse (une surface de Riemann compacte connexe) de genre g. Fixons un point base x_0, une base $(\delta_1, \ldots, \delta_{2g})$ du $\mathbb{Z}$-module libre

$$H_1(X, \mathbb{Z}) = \pi_1^{\mathrm{ab}}(X) := \pi_1(X, x_0)/[\pi_1(X, x_0), \pi_1(X, x_0)]$$

et une base $(\omega_1, \ldots, \omega_g)$ de l'espace vectoriel complexe $H^0(X, \Omega^1_X)$. Les périodes

$$\Pi_i = \left(\int_{\delta_i} \omega_1, \ldots, \int_{\delta_i} \omega_g \right) \in \mathbb{C}^g,\ i = 1, \ldots, 2g,$$

sont linéairement indépendantes sur $\mathbb{Z}$ et engendrent un réseau $\Lambda \subset \mathbb{C}^g$. La jacobienne de X est le tore complexe de dimension g,

$$J(X) = \mathbb{C}^g/\Lambda.$$

On a une application analytique

$$\varphi : X \longrightarrow J(X),\ x \longmapsto \left(\int_{x_0}^{x} \omega_1, \ldots, \int_{x_0}^{x} \omega_g\right) + \Lambda,$$

qui envoie x_0 sur l'élément neutre 0 de $J(X)$.

Soit $\mathrm{Pic}(X)$ le schéma de Picard qui paramètre les classes d'isomorphie de fibrés en droites sur X. C'est un schéma en groupes pour la structure de groupe induite par le produit tensoriel des fibrés en droites, qui s'insère dans la suite exacte

$$1 \longrightarrow \mathrm{Pic}^0(X) \longrightarrow \mathrm{Pic}(X) \xrightarrow{\deg} \mathbb{Z} \longrightarrow 0$$

où $\deg(\mathcal{L})$ est le degré du fibré en droites $\mathcal{L}$, et sa composante neutre $\mathrm{Pic}^0(X)$ est une variété abélienne.

THÉORÈME 1.1 (Abel-Jacobi). — *La jacobienne $J(X)$ est le tore complexe sous-jacent à la variété abélienne $\mathrm{Pic}^0(X)$ et le morphisme analytique*

$$\varphi : X \longrightarrow J(X) \cong \mathrm{Pic}^0(X)$$

n'est autre que le morphisme algébrique qui envoie $x \in X$ sur la classe d'isomorphie de $\mathcal{O}_X([x] - [x_0])$ dans $\mathrm{Pic}^0(X)$. □

L'homomorphisme

$$\pi_1(X, x_0) \longrightarrow \pi_1(J(X), 0)$$

induit par φ est l'application quotient

$$\pi_1(X, x_0) \twoheadrightarrow \pi_1(X, x_0)/[\pi_1(X, x_0), \pi_1(X, x_0)] = \Lambda.$$

Pour tout caractère $\chi : \pi_1(X, x_0) \to C^\times$ il existe donc un unique caractère $\mathrm{Aut}_\chi : \pi_1(\mathrm{Pic}^0(X), 0) \to C^\times$ tel que $\chi = \mathrm{Aut}_\chi \circ \varphi_*$.

Comme un système local E (de C-espaces vectoriels) sur une variété n'est autre qu'une représentation de dimension finie sur C du groupe fondamental de cette variété, on a montré par voie analytique :

THÉORÈME 1.2. — *Pour tout système local E de rang 1 sur X, il existe un unique système local (rigidifié à l'origine) Aut_E de rang 1 sur $\mathrm{Pic}^0(X)$ tel que $\varphi^* \mathrm{Aut}_E = E$.*

De plus, si $m : \mathrm{Pic}^0(X) \times \mathrm{Pic}^0(X) \to \mathrm{Pic}^0(X)$ est la loi de groupe, on a un isomorphisme canonique

$$m^* \mathrm{Aut}_E \cong \mathrm{Aut}_E \boxtimes \mathrm{Aut}_E\,.$$

Démonstration algébrique du théorème 1.2. — Pour chaque entier $d \geqslant 0$, le groupe symétrique $\mathfrak{S}_d$ agit de manière évidente sur le produit

$$X^d = \overbrace{X \times \cdots \times X}^{d}.$$

Le schéma quotient

$$X^{(d)} = X^d/\mathfrak{S}_d,$$

qui est connexe, projectif et lisse de dimension d, est l'espace de modules des diviseurs $D = \sum_x d_x[x]$ effectifs ($d_x \geqslant 0$) et de degré $\sum_x d_x = d$ sur X, et le morphisme φ induit un morphisme

$$\varphi^{(d)} : X^{(d)} \longrightarrow \mathrm{Pic}^0(X),\ D \longmapsto \mathcal{O}_X(D - d[x_0]).$$

Dès que $d \geqslant 2g-1$, $\varphi^{(d)}$ est un fibré projectif de rang $d-g$ et a donc ses fibres simplement connexes.

Pour chaque système local E de rang n sur X et chaque entier $d \geqslant 0$, le produit tensoriel externe $E^{\boxtimes d}$ est un système local de rang n^d sur X^d muni d'une action de $\mathfrak{S}_d$ qui relève celle sur X^d. Si on note $r : X^d \longrightarrow\!\!\!\!\!\rightarrow X^d/\mathfrak{S}_d = X^{(d)}$ le morphisme quotient, le faisceau constructible de C-espaces vectoriels $r_*E^{\boxtimes d}$ est donc muni d'une action de $\mathfrak{S}_d$ et on note

$$E^{(d)} = (r_*E^{\boxtimes d})^{\mathfrak{S}_d} \subset E^{\boxtimes d}$$

le sous-faisceau des vecteurs fixes par cette action. On vérifie que la fibre en $D \in X^{(d)}$ de $E^{(d)}$ est égale à

$$(E^{(d)})_D = \bigotimes_x \mathrm{Sym}^{d_x} E_x. \tag{1.3}$$

Si maintenant E est de rang 1 sur X, $E^{(d)}$ est en fait un système local de rang 1 sur $X^{(d)}$. Pour tout $d \geqslant 2g-1$, $E^{(d)}$ est donc constant sur les fibres simplement connexes du morphisme $\varphi^{(d)} : X^{(d)} \to \mathrm{Pic}^0(X)$ et se descend en un système local sur $\mathrm{Pic}^0(X)$ qui n'est autre que Aut_E. □

C'est le cas GL(1) de la correspondance de Drinfeld-Langlands partout non ramifiée.

2. LE THÉORÈME PRINCIPAL

Soit n un entier $\geqslant 2$. La généralisation naturelle de $\mathrm{Pic}(X)$, ou plutôt du champ quotient du schéma $\mathrm{Pic}(X)$ par l'action triviale de $\mathbb{G}_\mathrm{m}$, est le champ $\mathrm{Fib}_n(X)$ des fibrés vectoriels de rang n sur la surface de Riemann X. Ce champ est algébrique, lisse purement de dimension $n^2(g-1)$ et réunion croissante d'ouverts de type fini. Si $g \geqslant 2$, il contient un ouvert dense qui est une $\mathbb{G}_\mathrm{m}$-gerbe sur le schéma quasi-projectif de modules des fibrés stables sur X. Les composantes connexes $\mathrm{Fib}_n^d(X)$ de $\mathrm{Fib}_n(X)$ sont découpées par le degré d du fibré vectoriel universel.

Un fibré vectoriel $\mathcal{L}$ de rang n sur X peut aussi être vu comme un $\mathcal{O}_X$-Module localement libre de rang n. Une *modification inférieure élémentaire* de $\mathcal{L}$ en un point $x \in X$ est un fibré vectoriel $\mathcal{L}'$ de rang n sur X qui est contenu dans $\mathcal{L}$ en tant que sous-$\mathcal{O}_X$-Module de telle sorte que le $\mathcal{O}_X$-Module cohérent $\mathcal{L}/\mathcal{L}'$ soit un faisceau gratte-ciel de longueur 1 concentré en x. Le degré de $\mathcal{L}'$ est donc égal à

$$\deg(\mathcal{L}') = \deg(\mathcal{L}) - 1.$$

Le *champ de Hecke* est le champ modulaire $\mathrm{Hecke}_n(X)$ des triplets $(x, \mathcal{L}, \mathcal{L}' \subset \mathcal{L})$ où $x \in X$, $\mathcal{L} \in \mathrm{Fib}_n(X)$ et $\mathcal{L}' \subset \mathcal{L}$ est une modification élémentaire inférieure de $\mathcal{L}$ en x.

La *correspondance de Hecke* est le diagramme

$$\begin{array}{ccc} & \mathrm{Hecke}_n(X) & \\ {\scriptstyle (q,p')}\swarrow & & \searrow{\scriptstyle p} \\ X \times \mathrm{Fib}_n(X) & & \mathrm{Fib}_n(X) \end{array}$$

où les projections (q, p') et p sont données par

$$(q, p')(x, \mathcal{L}, \mathcal{L}' \subset \mathcal{L}) = (x, \mathcal{L}')$$

et

$$p(x, \mathcal{L}, \mathcal{L}' \subset \mathcal{L}) = \mathcal{L},$$

et sont des morphismes représentables projectifs et lisses de dimension relative $n-1$ et n respectivement.

Cette correspondance induit un *opérateur de Hecke* sur la catégorie dérivée $D^{\mathrm{b}}_{\mathrm{c}}(\mathrm{Fib}_n(X))$ qui envoie $K \in D^{\mathrm{b}}_{\mathrm{c}}(\mathrm{Fib}_n(X))$ sur

$$H(K) = (q, p')_* p^* K[n-1] \in D^{\mathrm{b}}_{\mathrm{c}}(X \times \mathrm{Fib}_n(X)).$$

On peut « itérer » H et on obtient en particulier l'opérateur

$$H \circ H : D^{\mathrm{b}}_{\mathrm{c}}(\mathrm{Fib}_n(X)) \longrightarrow D^{\mathrm{b}}_{\mathrm{c}}(X \times X \times \mathrm{Fib}_n(X)),$$

dont la restriction au complémentaire $(X \times X - \Delta_X) \times \mathrm{Fib}_n(X)$ de la diagonale est $\mathfrak{S}_2$-équivariante pour l'action qui permute les deux copies de X.

Définition 2.1. — *Soit E un système local irréductible de rang n sur X. Un faisceau pervers irréductible K sur* $\mathrm{Fib}_n(X)$ *est dit avoir la* propriété de Hecke relativement à *E s'il existe un isomorphisme*

$$H(K) \cong E \boxtimes K$$

tel que la restriction de l'isomorphisme induit

$$(H \circ H)(K) \cong E \boxtimes E \boxtimes K$$

à l'ouvert $(X \times X - \Delta_X) \times \mathrm{Fib}_n(X)$ soit $\mathfrak{S}_2$-équivariant.

THÉORÈME 2.2 (Drinfeld pour $n = 2$, [DR2] ; Frenkel, Gaitsgory et Vilonen en général, [F-G-V2] et [Ga])

Pour tout système local irréductible E de rang n sur X, il existe un faisceau pervers Aut_E *sur* $\mathrm{Fib}_n(X)$ *dont la restriction à chaque composante connexe* $\mathrm{Fib}_n^d(X)$ *est irréductible et qui a la propriété de Hecke relativement à E.*

La démonstration du théorème se fait en deux temps.

Premièrement on construit un complexe de faisceaux constructibles Aut'_E sur l'espace de modules $\mathrm{Fib}'_n(X)$ des couples (L, s) formés d'un fibré vectoriel L de rang n sur X et d'une section non nulle à homothétie près

$$s : (\Omega^1_X)^{\otimes(n-1)} \hookrightarrow L.$$

Cette construction est inspirée par le développement en série de Fourier d'une forme automorphe cuspidale pour $\mathrm{GL}(n)$ et la formule de Shintani pour les coefficients de cette série de Fourier, compte tenu du dictionnaire fonctions-faisceaux de Grothendieck ; mais ce n'en est pas une simple transposition.

Puis on montre que, sur un gros ouvert de $\mathrm{Fib}'_n(X)$, Aut'_E est un faisceau pervers qui se descend par le morphisme $\mathrm{Fib}'_n(X) \to \mathrm{Fib}_n(X)$ d'oubli de la section s, en un faisceau pervers Aut_E sur $\mathrm{Fib}_n(X)$ qui a les propriétés requises. C'est bien entendu dans cette deuxième étape que réside toute la difficulté.

3. LES FAISCEAUX PERVERS $\mathcal{L}_E^d$

Pour chaque entier $d \geqslant 0$, soit $\mathrm{Coh}_0^d(X)$ le champ des $\mathcal{O}_X$-Modules cohérents $\mathcal{M}$ de rang générique 0, c'est-à-dire à support fini, et de longueur $\dim H^0(X, \mathcal{M}) = d$. On note $\mathrm{Coh}_0(X)$ la réunion disjointe des $\mathrm{Coh}_0^d(X)$.

La définition du champ $\mathrm{Coh}_0^d(X)$ a un sens pour toute courbe algébrique X quasi-projective ; en particulier, on peut considérer le champ $\mathrm{Coh}_0^d(\mathbb{A}^1)$. Comme la donnée d'un $\mathcal{O}_{\mathbb{A}^1}$-Module cohérent de rang générique 0 et de longueur d équivaut à la donnée d'un espace vectoriel de dimension d muni d'un endomorphisme, $\mathrm{Coh}_0^d(\mathbb{A}^1)$ n'est autre que le champ quotient $[\mathrm{gl}(d)/\mathrm{GL}(d)]$ de l'espace affine des matrices carrées de taille $d \times d$ par l'action par conjugaison de $\mathrm{GL}(d)$. Comme toute courbe algébrique lisse X est localement pour la topologie étale isomorphe à la droite affine $\mathbb{A}^1$, $\mathrm{Coh}_0^d(X)$ est localement pour la topologie étale isomorphe à $[\mathrm{gl}(d)/\mathrm{GL}(d)]$.

Chaque diviseur effectif D sur X définit un $\mathcal{O}_X$-Module de torsion

$$\mathcal{O}_{X,D} = \mathcal{O}_X/\mathcal{O}_X(-D)$$

de longueur le degré de D et, pour chaque entier $d \geqslant 0$, le morphisme

$$X^{(d)} \longrightarrow \mathrm{Coh}_0^d(X),\ D \longmapsto \mathcal{O}_{X,D},$$

est lisse de dimension relative d. Tout $\mathcal{M} \in \mathrm{Coh}_0(X)$ est isomorphe à $\bigoplus_{i\in I} \mathcal{O}_{X,D_i}$ pour une famille finie $(D_i)_{i\in I}$ de diviseurs effectifs sur X et, pour chaque entier $d \geqslant 0$, on a un morphisme *déterminant*

$$\det : \mathrm{Coh}_0^d(X) \longrightarrow X^{(d)}$$

qui envoie $\mathcal{M} \cong \bigoplus_{i\in I} \mathcal{O}_{X,D_i}$ sur $\sum_{i\in I} D_i$. On note $X^{(d)}_{\mathrm{rss}}$ l'ouvert de $X^{(d)}$ formé des D sans multiplicité ($d_x = 0$ ou 1 quel que soit x) et $\mathrm{Coh}^d_{0,\mathrm{rss}}(X) = \det^{-1}(X^{(d)}_{\mathrm{rss}})$ l'ouvert correspondant de $\mathrm{Coh}^d_0(X)$.

Soit E un système local de rang n sur X. La formule (1.3) pour la fibre en $D \in X^{(d)}$ de $E^{(d)}$ montre que le rang de cette fibre varie avec D, mais que, au-dessus de l'ouvert $X^{(d)}_{\mathrm{rss}}$, ce rang est constant et la restriction de $E^{(d)}$ est un système local de rang n^d.

LEMME 3.1. — *Pour chaque entier $d \geqslant 0$, le complexe $E^{(d)}[d] \in D^{\mathrm{b}}_{\mathrm{c}}(X^{(d)})$, constitué du faisceau constructible $E^{(d)}$ placé en degré $-d$, est un faisceau pervers, qui est le prolongement intermédiaire de sa restriction à $X^{(d)}_{\mathrm{rss}}$.*

Le faisceau pervers $E^{(d)}[d]$ est irréductible (resp. semi-simple) si E l'est. □

La restriction

$$\mathcal{L}^d_{E,\mathrm{rss}} = \det{}^* E^{(d)}|\,\mathrm{Coh}^d_{0,\mathrm{rss}}(X)$$

du faisceau constructible $\det^* E^{(d)}$ à l'ouvert dense $\mathrm{Coh}^d_{0,\mathrm{rss}}(X)$ de $\mathrm{Coh}^d_0(X)$ est aussi un système local de rang n^d, qui est irréductible (resp. semi-simple) si E l'est.

DÉFINITION 3.2. — *Pour chaque entier $d \geqslant 0$, le faisceau pervers $\mathcal{L}^d_E$ associé à un système local E sur X est le prolongement intermédiaire de $\mathcal{L}^d_{E,\mathrm{rss}}$ à $\mathrm{Coh}^d_0(X)$ tout entier.*

Bien sûr, $\mathcal{L}^d_E$ est irréductible (resp. semi-simple) si E l'est. On vérifie en outre que l'image réciproque décalée de d de $\mathcal{L}^d_E$ par le morphisme $X^{(d)} \to \mathrm{Coh}^d_0,\ D \mapsto \mathcal{O}_{X,D}$, n'est autre que le faisceau pervers $E^{(d)}[d]$.

4. LE COMPLEXE Aut'_E

On suppose dorénavant que le genre g de X est $\geqslant 2$. En genre 0 ou 1, la correspondance de Drinfeld-Langlands partout non ramifiée pour $\mathrm{GL}(n)$ avec $n \geqslant 2$ est vide puisqu'il n'y a pas de système local irréductible de rang $\geqslant 2$ sur X.

On se propose dans cette section de construire un candidat Aut'_E pour la restriction de Aut_E au champ algébrique $\mathrm{Fib}'_n(X)$ des fibrés vectoriels $\mathcal{L}$ de rang n sur X munis d'une section $s : (\Omega^1_X)^{\otimes n-1} \to \mathcal{L}$ à homothétie près.

Considérons le champ algébrique $\mathcal{U}_n$ des triplets $(\mathcal{L}, \mathcal{L}_\bullet, s_\bullet)$ où $\mathcal{L}_n$ est un fibré vectoriel de rang n,

$$\mathcal{L}_\bullet = ((0) = \mathcal{L}_0 \subset \mathcal{L}_1 \subset \cdots \subset \mathcal{L}_n = \mathcal{L})$$

est un drapeau complet de sous-fibrés vectoriels et $s_\bullet$ est une famille d'isomorphismes de fibrés en droites

$$s_j : (\Omega^1_X)^{\otimes j-1} \xrightarrow{\sim} \mathcal{L}_{n-j+1}/\mathcal{L}_{n-j},\ j = 1, \dots, n,$$

cette famille étant prise à homothétie près : $(s_j)_{j=1,\dots,n} \sim (ts_j)_{j=1,\dots,n}$ pour tout $t \in \mathbb{G}_{\mathrm{m}}$. On a des morphismes

$$\begin{array}{ccc} \mathcal{U}_n & \xrightarrow{g_n} & \mathbb{A}^1 \\ \downarrow{\scriptstyle f_n} & & \\ \mathrm{Fib}'^{n(n-1)(g-1)}_n(X) & & \end{array}$$

où f_n envoie le triplet $(\mathcal{L}, \mathcal{L}_\bullet, s_\bullet)$ sur le fibré $\mathcal{L}$ muni de la section à homothétie près induite par

$$s_n : (\Omega^1_X)^{\otimes n-1} \xrightarrow{\sim} \mathcal{L}_1 \subset \mathcal{L}_n = \mathcal{L}$$

et g_n envoie ce même triplet sur le scalaire somme des classes d'extensions

$$0 \longrightarrow \mathcal{L}_j/\mathcal{L}_{j-1} \longrightarrow \mathcal{L}_{j+1}/\mathcal{L}_{j-1} \longrightarrow \mathcal{L}_{j+1}/\mathcal{L}_j \longrightarrow 0$$

dans

$$\mathrm{Ext}^1_{\mathcal{O}_X}(\mathcal{L}_{j+1}/\mathcal{L}_j, \mathcal{L}_j/\mathcal{L}_{j-1}) \cong \mathrm{Ext}^1_{\mathcal{O}_X}((\Omega^1_X)^{\otimes n-j-1}, (\Omega^1_X)^{\otimes n-j}) \cong \mathrm{Ext}^1_{\mathcal{O}_X}(\mathcal{O}_X, \Omega^1_X) \cong \mathbb{C}.$$

Si l'on fait agir le groupe multiplicatif $\mathbb{G}_{\mathrm{m}}$ sur $\mathcal{U}_n$ par

$$t \cdot (\mathcal{L}, \mathcal{L}_\bullet, s_\bullet) = (\mathcal{L}, \mathcal{L}_\bullet, t \cdot s_\bullet)$$

où $t \cdot s_\bullet = (t^{n-j} s_j)_{j=1,\dots,n}$ et par homothétie sur $\mathbb{A}^1$, le morphisme g_n est $\mathbb{G}_{\mathrm{m}}$-équivariant. En passant au quotient par cette action on obtient des morphismes

$$\begin{array}{ccc} \overline{\mathcal{U}}_n & \xrightarrow{\overline{g}_n} & [\mathbb{A}^1/\mathbb{G}_{\mathrm{m}}] = \mathcal{A} \\ \overline{f}_n \downarrow & & \\ \mathrm{Fib}'^{n(n-1)(g-1)}_n(X) & & \end{array}$$

et on peut former le complexe de faisceaux constructibles

$$L_n = (\overline{f}_n)_! (\overline{g}_n)^* \Psi[\dim \overline{\mathcal{U}}_n]$$

où $\Psi \in D^{\mathrm{b}}_{\mathrm{c}}(\mathcal{A})$ est le complexe défini dans la section 0.

Pour chaque entier d, on a une correspondance

$$\begin{array}{ccccc} & & \mathrm{Mod}'^d_n(X) & \xrightarrow{q} & \mathrm{Coh}^d_0(X) \\ & \swarrow^{p'} & & \searrow_{p} & \\ \mathrm{Fib}'^{n(n-1)(g-1)}_n(X) & & & & \mathrm{Fib}'^{n(n-1)(g-1)+d}_n(X) \end{array}$$

où $\mathrm{Mod}_n'^d(X)$ est le champ des diagrammes

$$(\Omega_X^1)^{\otimes n-1} \longrightarrow \mathcal{L}' \subset \mathcal{L}$$

formés de deux fibrés vectoriels de rang n et de degrés $n(n-1)(g-1)$ et $n(n-1)(g-1)+d$ respectivement et d'injections $\mathcal{O}_X$-linéaires $(\Omega_X^1)^{\otimes n-1} \hookrightarrow \mathcal{L}'$ (à homothétie près) et $\mathcal{L}' \hookrightarrow \mathcal{L}$, et où p, p' et q envoient un tel diagramme sur $(\Omega_X^1)^{\otimes n-1} \hookrightarrow \mathcal{L}'$ (à homothétie près), $(\Omega_X^1)^{\otimes n-1} \hookrightarrow \mathcal{L}$ (à homothétie près) et $\mathcal{L}/\mathcal{L}'$ respectivement.

Si E est un système local sur X, on définit le foncteur

$$M_{n,E}'^d : D_{\mathrm{c}}^{\mathrm{b}}(\mathrm{Fib}_n'^{n(n-1)(g-1)}(X)) \longrightarrow D_{\mathrm{c}}^{\mathrm{b}}(\mathrm{Fib}_i'^{n(n-1)(g-1)+d}(X))$$

par

$$M_{n,E}'^d(K) = p_* p'^*(K \otimes q^* \mathcal{L}_E^d)[nd]$$

où $\mathcal{L}_E^d$ est le faisceau pervers sur $\mathrm{Coh}_0^d(X)$ construit dans la section 3.

Définition 4.1. — *Si E est un système local sur X, on définit le complexe Aut_E' sur $\mathrm{Fib}_n'(X)$ par*

$$\mathrm{Aut}_E' \,|\, \mathrm{Fib}_n'^{n(n-1)(g-1)+d}(X) = M_{n,E}'^d(L_n)$$

si $d \geqslant 0$ et

$$\mathrm{Aut}_E' \,|\, \mathrm{Fib}_n'^{n(n-1)(g-1)+d}(X) = (0)$$

si $d < 0$.

5. CONSTRUCTION PAR TRANSFORMATIONS DE FOURIER

Pour chaque $i = 1, \ldots, n$, soient $\mathrm{Coh}_i(X)$ le champ algébrique des $\mathcal{O}_X$-Modules cohérents $\mathcal{M}_i$ de rang générique i et $\mathrm{Coh}_i'(X)$ le champ algébrique des $\mathcal{O}_X$-Modules cohérents $\mathcal{M}_i$ de rang générique i munis d'une injection $\mathcal{O}_X$-linéaire $s_i : (\Omega_X^1)^{\otimes i-1} \hookrightarrow \mathcal{M}_i$ à homothétie près. Ces champs admettent pour ouverts denses les champs $\mathrm{Fib}_i(X)$ et $\mathrm{Fib}_i'(X)$ et, tout comme ces derniers champs, leurs composantes connexes $\mathrm{Coh}_i^d(X)$ et $\mathrm{Coh}_i'^d(X)$ sont découpées par le degré d de $\mathcal{M}_i$.

Lemme 5.1. — *Il existe une constante $c(g,n)$ qui a la propriété suivante : pour tout entier $d \geqslant c(g,n)$ et tout $\mathcal{L} \in \mathrm{Fib}_n^d(X)$ tel que $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{L}, (\Omega_X^1)^{\otimes n+1}) \neq (0)$, $\mathcal{L}$ est très instable au sens où il existe une décomposition non triviale en somme directe $\mathcal{L} = \mathcal{L}_1 \oplus \mathcal{L}_2$ pour laquelle $\mathrm{Ext}_{\mathcal{O}_X}^1(\mathcal{L}_1, \mathcal{L}_2) = (0)$.* □

Pour chaque $i = 0, \ldots, n$, soient $\mathcal{C}_i \subset \mathrm{Coh}_i(X)$ l'ouvert formé des $\mathcal{M}_i$ de degré $\geqslant c(n,g) + i(i-1)(g-1)$ tels que $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{M}_i, (\Omega_X^1)^{\otimes n+1}) = (0)$, et donc a fortiori tels que $\mathrm{Hom}_{\mathcal{O}_X}(\mathcal{M}_i, (\Omega_X^1)^{\otimes i}) = (0)$ et que $\mathrm{Ext}_{\mathcal{O}_X}^1((\Omega_X^1)^{\otimes i-1}, \mathcal{M}_i) = (0)$ par dualité de Serre. Soient $\mathcal{V}_i$ le champ algébrique des couples $(\mathcal{M}_i, s_i)$ où $\mathcal{M}_i \in \mathcal{C}_i$ et $s_i \in \mathrm{Hom}_{\mathcal{O}_X}((\Omega_X^1)^{\otimes i-1}, \mathcal{M}_i)$ et $\mathcal{V}_i^\vee$ le champ algébrique des extensions

$$0 \longrightarrow (\Omega_X^1)^{\otimes i} \longrightarrow \mathcal{M}_{i+1} \longrightarrow \mathcal{M}_i \longrightarrow 0$$

de $\mathcal{O}_X$-Modules cohérents avec $\mathcal{M}_i \in \mathcal{C}_i$. On a des projections naturelles $\pi_i : \mathcal{V}_i \to \mathcal{C}_i$ et $\pi_i^\vee : \mathcal{V}_i^\vee \to \mathcal{C}_i$ qui sont des fibrés vectoriels en dualité.

Introduisons les ouverts

$$\mathcal{V}_i^\circ = \{(\mathcal{M}_i, s_i) \mid s_i \text{ est injective}\} \subset \mathcal{V}_i$$

et

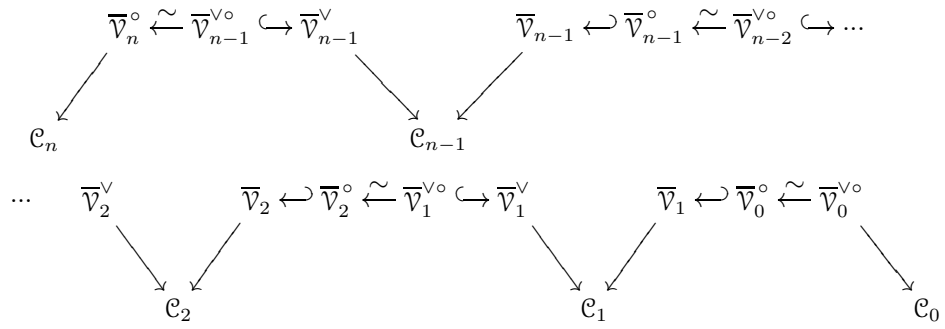
$$\mathcal{V}_i^{\vee\circ} = \{(0 \longrightarrow (\Omega_X^1)^{\otimes i} \longrightarrow \mathcal{M}_{i+1} \longrightarrow \mathcal{M}_i \longrightarrow 0) \mid \mathcal{M}_{i+1} \in \mathcal{C}_{i+1}\} \subset \mathcal{V}_i^\vee.$$

Bien entendu, on a un isomorphisme canonique

$$\mathcal{V}_i^{\vee\circ} \xrightarrow{\sim} \mathcal{V}_{i+1}^\circ$$

qui envoie $0 \to (\Omega_X^1)^{\otimes i} \to \mathcal{M}_{i+1} \to \mathcal{M}_i \to 0$ sur l'injection $(\Omega_X^1)^{\otimes i} \to \mathcal{M}_{i+1}$.

Le groupe multiplicatif agit par homothétie sur les fibrés vectoriels $\mathcal{V}_i$ et $\mathcal{V}_i^\vee$ et par passage au quotient on obtient le diagramme fondamental



où les flèches obliques sont les projections $\overline{\pi}_n^\circ : \overline{\mathcal{V}}_n^\circ = [\mathcal{V}_n^\circ/\mathbb{G}_{\mathrm{m}}] \to \mathcal{C}_n$, $\overline{\pi}_i : \overline{\mathcal{V}}_i = [\mathcal{V}_i/\mathbb{G}_{\mathrm{m}}] \to \mathcal{C}_i$, $\overline{\pi}_i^\vee : \overline{\mathcal{V}}_i^\vee = [\mathcal{V}_i^\vee/\mathbb{G}_{\mathrm{m}}] \to \mathcal{C}_i$ et $\overline{\pi}_0^{\vee\circ} : \overline{\mathcal{V}}_0^{\vee\circ} = [\mathcal{V}_0^{\vee\circ}/\mathbb{G}_{\mathrm{m}}] \to \mathcal{C}_0$ et où les flèches horizontales sont d'une part les immersions ouvertes $j_i : \overline{\mathcal{V}}_i^\circ \hookrightarrow \overline{\mathcal{V}}_i$ et $j_i^\vee : \overline{\mathcal{V}}_i^{\vee\circ} \hookrightarrow \overline{\mathcal{V}}_i^\vee$ induites par les inclusions $\mathcal{V}_i^\circ \subset \mathcal{V}_i$ et $\mathcal{V}_i^{\vee\circ} \subset \mathcal{V}_i^\vee$ et d'autre part les isomorphismes $\iota_i : \overline{\mathcal{V}}_i^{\vee\circ} \xrightarrow{\sim} \overline{\mathcal{V}}_{i+1}^\circ$ induits par les isomorphismes $\mathcal{V}_i^{\vee\circ} \xrightarrow{\sim} \mathcal{V}_{i+1}^\circ$ ci-dessus.

Pour tout système local E de rang n sur X on définit alors les complexes $K_{E,i} \in D_{\mathrm{c}}^{\mathrm{b}}(\overline{\mathcal{V}}_i^\circ)$, $i = 1, \ldots, n$, par

$$K_{E,1} = (\iota_0)_*(\overline{\pi}_0^{\vee\circ})^* \mathcal{L}_E,$$

où $\mathcal{L}_E$ est le complexe sur $\mathcal{C}_0$ dont la restriction à chaque composante connexe $\mathcal{C}_0^d$ de $\mathcal{C}^0$ est le faisceau pervers $\mathcal{L}_E^d$ défini en 3.2, décalé de $d-1$, et par la relation de récurrence

$$K_{E,i+1} = (\iota_i)_*(j_i^\vee)^* \operatorname{Four}_{\overline{\mathcal{V}}_i/\mathcal{C}_i}((j_i)_! K_{E,i}),\ i = 1, \ldots, n-1$$

où $\operatorname{Four}_{\overline{\mathcal{V}}_i/\mathcal{C}_i} : \mathrm{D}_{\mathrm{c}}^{\mathrm{b}}(\overline{\mathcal{V}}_i) \to \mathrm{D}_{\mathrm{c}}^{\mathrm{b}}(\overline{\mathcal{V}}_i^\vee)$ est la transformation de Fourier homogène pour le fibré vectoriel $\mathcal{V}_i \to \mathcal{C}_i$ définie dans la section 0.

Les champs $\overline{\mathcal{V}}_i^\circ$ et $\mathrm{Fib}'_i(X)$ sont des ouverts de $\mathrm{Coh}'_i(X)$ et on peut donc considérer leur intersection.

LEMME 5.2. — *Soit E un système local de rang n sur X. Pour chaque $i = 1, \ldots, n$ et chaque entier $d \geqslant 0$, la restriction de $K_{E,i}$ à l'ouvert $\overline{\mathcal{V}}_i^\circ \cap \mathrm{Fib}_i'^{i(i-1)(g-1)+d}(X)$ de $\mathrm{Coh}_i'^{i(i-1)(g-1)+d}(X)$ est égale à*

$$M_{i,E}'^d(L_i)$$

avec les notations de la section 4.

En particulier, pour chaque entier $d \geqslant 0$, les restrictions de $K_{E,n}$ et de Aut'_E à l'ouvert $\overline{\mathcal{V}}_n^\circ \cap \mathrm{Fib}_i'^{n(n-1)(g-1)+d}(X)$ de $\mathrm{Coh}_i'^{n(n-1)(g-1)+d}(X)$ coïncident. □

6. LE THÉORÈME D'ANNULATION

Soient $i \geqslant 1$ et $d \geqslant 0$ des entiers. Considérons le diagramme

$$\begin{array}{ccccc} & & \mathrm{Mod}_i^d(X) & \xrightarrow{q} & \mathrm{Coh}_0^d(X) \\ & \swarrow^{p'} & & \searrow_{p} & \\ \mathrm{Fib}_i(X) & & & & \mathrm{Fib}_i(X) \end{array}$$

où $\mathrm{Mod}_i^d(X)$ est le champ des couples $(\mathcal{L}, \mathcal{L}' \subset \mathcal{L})$ formé d'un fibré vectoriel de rang i et d'une modification inférieure $\mathcal{L}' \subset \mathcal{L}$ de co-longueur d, où les projections p' et p envoient $(\mathcal{L}, \mathcal{L}' \subset \mathcal{L})$ sur $\mathcal{L}'$ et $\mathcal{L}$ respectivement et où q envoie $(\mathcal{L}, \mathcal{L}' \subset \mathcal{L})$ sur le $\mathcal{O}_X$-Module cohérent $\mathcal{L}/\mathcal{L}'$ de rang générique 0 et de longueur d.

Pour tout système local E sur X, on a la variante suivante du foncteur $M_{i,E}'^d$ de la section 4

$$M_{i,E}^d : D_{\mathrm{c}}^{\mathrm{b}}(\mathrm{Fib}_i(X)) \longrightarrow D_{\mathrm{c}}^{\mathrm{b}}(\mathrm{Fib}_i(X)),\ M_{i,E}^d(K) = p_* p'^*(K \otimes q^* \mathcal{L}_E^d)[id],$$

où $\mathcal{L}_E^d$ est le faisceau pervers sur $\mathrm{Coh}_0^d(X)$ défini en 3.2.

THÉORÈME 6.1 (Gaitsgory, [Ga]). — *Soit E un système local* irréductible *de rang n. Alors, pour tout $i = 1, \ldots, n-1$ et tout entier $d > in(2g-2)$, le foncteur $M_{i,E}^d$ est identiquement nul.* □

Remarque 6.2. — La preuve de ce théorème est longue et technique. Elle dépasse largement le cadre de cet exposé. Signalons cependant que, pour $i = 1$, cet énoncé n'est autre qu'une reformulation d'un résultat de Deligne (*cf.* [DR2]) qui dit que

$$(\varphi^{(d)})_* E^{(d)} = 0$$

pour tout système local irréductible de rang $n > 1$ sur X et tout entier $d > n(2g-2)$, où $\varphi^{(d)} : X^{(d)} \to \mathrm{Pic}^0(X)$ est le fibré projectif considéré dans la section 1. □

Frenkel, Gaitsgory et Vilonen utilisent le théorème 6.1 pour démontrer la proposition suivante.

PROPOSITION 6.3. — *Soit E un système local* irréductible *de rang n. Alors, pour chaque $i = 1, \dots, n-1$, la flèche canonique*

$$(j_i)_! K_{E,i} \longrightarrow (j_i)_* K_{E,i}$$

est un isomorphisme.

Idée de la démonstration. — Au-dessus de l'ouvert $\mathrm{Fib}_i(X) \cap \mathcal{C}_i$ de $\mathcal{C}_i$, $\mathcal{V}_i^\circ \subset \mathcal{V}_i$ est le complémentaire de la section nulle dans le fibré vectoriel $\mathcal{V}_i \to \mathcal{C}_i$. D'après le lemme (0.3), pour démontrer l'assertion au-dessus de cet ouvert, il suffit donc de montrer que $(\overline{\pi}_i^\circ)_! K_{E,i} = (0)$ où $\overline{\pi}_i^\circ : \overline{\mathcal{V}}_i^\circ \to \mathcal{C}_i$ est la projection, ou ce qui revient au même la projection $\mathbb{P}(\mathcal{V}_i) \to \mathcal{C}_i$ puisque les deux projections coïncident au-dessus de $\mathrm{Fib}_i(X) \cap \mathcal{C}_i$. Or, d'après le lemme 5.2, pour tout entier $d \geqslant 0$, on a

$$(\overline{\pi}_i^\circ)_! K_{E,i} | (\mathcal{C}_i \cap \mathrm{Fib}_i^{i(i-1)(g-1)+d}(X)) = M_{i,E}^d(\pi_i' L_i) | (\mathcal{C}_i \cap \mathrm{Fib}_i^{i(i-1)(g-1)+d}(X))$$

où $\pi' : \mathrm{Fib}_i'(X) \to \mathrm{Fib}_i(X)$ est le morphisme d'oubli de la section $s : (\Omega_X^1)^{\otimes n-1} \to \mathcal{L}$, d'où la conclusion d'après le théorème 6.1. □

Il résulte de la proposition 6.3 que, si $K_{E,i}$ est un faisceau pervers irréductible, il en est de même de $(j_i)_! K_{E,i}$ puisque c'est alors le prolongement intermédiaire de $K_{E,i}$. Comme la transformation de Fourier homogène préserve la perversité et l'irréductibilité, on voit par récurrence sur i que la restriction de $K_{E,i}$ à chaque composante connexe de $\overline{\mathcal{V}}_i^\circ$ est un faisceau pervers irréductible. En particulier, on a démontré :

COROLLAIRE 6.4. — *Soit E un système local irréductible de rang n. La restriction de $K_{E,n}$ à chaque composante connexe de $\overline{\mathcal{V}}_n^\circ$ est un faisceau pervers irréductible.* □

7. LA DESCENTE

Commençons par rappeler deux résultats généraux. Pour tout schéma (ou champ algébrique) S et tout $K \in D_{\mathrm{c}}^{\mathrm{b}}(S)$, notons

$$\chi_K : S \longrightarrow \mathbb{Z}$$

la fonction *caractéristique d'Euler-Poincaré de K* définie par

$$\chi_K(s) = \sum_i (-1)^i \dim_C H^i(K_s), \ \forall s \in S,$$

où K_s est la fibre de K en s.

LEMME 7.1 (Deligne, [Il] Corollaire 2.10). — *Soit $f : Y \to Z$ un morphisme propre de schémas, voire un morphisme représentable et propre de champs algébriques. Alors, si $K, K' \in D^{\mathrm{b}}_{\mathrm{c}}(Y)$ sont localement isomorphes sur Y, les complexes f_*K et f_*K' ont même fonction caractéristique d'Euler-Poincaré $\chi_{f_*K} = \chi_{f_*K'}$ sur Z.* □

LEMME 7.2. — *Soit S un schéma, voire un champ algébrique, soit $V \to S$ un fibré vectoriel de rang constant fini r, définissant un fibré projectif $\pi : P = \mathbb{P}(V) \to S$, et soit K un faisceau pervers irréductible sur P. Alors, il existe un faisceau pervers (nécessairement irréductible) L sur S tel que K soit isomorphe à $\pi^*(L)[r-1]$ si et seulement si la fonction caractéristique d'Euler-Poincaré χ_K est constante le long des fibres de π.*

Démonstration. — La partie « seulement si » est triviale. Supposons donc χ_K constante le long des fibres de π. Par le théorème de structure des faisceaux pervers irréductibles, on a $K \cong i_* j_{!*} E[\dim Q]$ où $i : Q \hookrightarrow P$ est le fermé irréductible support de K, $j : Q^\circ \hookrightarrow Q$ est un ouvert dense et E est un système local irréductible sur Q°.

On vérifie dans un premier temps que $Q = \pi^{-1}(T)$ pour un fermé irréductible T de S, puis que l'on peut choisir Q° de la forme $\pi^{-1}(T^\circ)$ pour un ouvert dense T° de T et enfin que E provient d'un système local irréductible F sur T° puisque les fibres de π sont simplement connexes.

Le faisceau pervers L sur S cherché est alors le prolongement par zéro à S tout entier du prolongement intermédiaire de $F[\dim T]$ à T. □

Étudions maintenant la descente de Aut'_E en un faisceau pervers Aut_E sur $\mathrm{Fib}_n(X)$. Pour cela, considérons le champ $\mathcal{Y}^d_n$ introduit par Drinfeld, qui classifie les couples $(\mathcal{L}, s^\bullet)$ où $\mathcal{L}$ est un fibré vectoriel de degré $n(n-1)(g-1)+d$ et $s^\bullet = (s^1, \dots, s^n)$ est une suite d'injections de $\mathcal{O}_X$-Modules

$$\begin{array}{c} (\Omega^1_X)^{\otimes n-1} \xhookrightarrow{s^1} \mathcal{L} \\ \cdots \\ (\Omega^1_X)^{\otimes (n-1)+\cdots+(n-i)} \xhookrightarrow{s^i} \bigwedge^i \mathcal{L} \\ \cdots \\ (\Omega^1_X)^{\otimes n(n-1)/2} \xhookrightarrow{s^n} \bigwedge^n \mathcal{L} \end{array}$$

qui satisfont les relations de Plücker faisant que la suite $s^\bullet$ définisse un drapeau complet de sous-espaces vectoriels dans la fibre de $\mathcal{L}$ au point générique de X.

Le champ $\mathcal{Y}^d_n$ est une « compactification partielle » du champ des triplets $(\mathcal{L}, \mathcal{L}_\bullet, s_\bullet)$ où

$$\mathcal{L}_\bullet = ((0) = \mathcal{L}_0 \subset \mathcal{L}_1 \subset \cdots \subset \mathcal{L}_n)$$

est un drapeau de sous-$\mathcal{O}_X$-Modules de $\mathcal{L}$ et $s_\bullet$ est une suite d'isomorphismes de $\mathcal{O}_X$-Modules $s_j : (\Omega^1_X)^{\otimes j-1} \xrightarrow{\sim} \mathcal{L}_{n-j+1}/\mathcal{L}_{n-j}$ pour $j = 1, \dots, n$. En effet, le morphisme

$$(\mathcal{L}, \mathcal{L}_\bullet, s_\bullet) \longmapsto (\mathcal{L}, s^\bullet)$$

défini par

$$s^i = s_n \otimes s_{n-1} \otimes \cdots \otimes s_{n-i+1} : (\Omega^1_X)^{\otimes(n-1)+(n-2)+\cdots+(n-i)} \xrightarrow{\sim} \mathcal{L}_1 \otimes (\mathcal{L}_2/\mathcal{L}_1) \otimes \cdots \otimes (\mathcal{L}_i/\mathcal{L}_{i-1}) \cong \overset{i}{\bigwedge} \mathcal{L}_i$$

identifie ce champ des triplets à l'ouvert $\mathcal{Y}^{d\,\circ}_n \subset \mathcal{Y}^d_n$ formé des $(\mathcal{L}, s^\bullet)$ tels que les conoyaux des injections $s^i : (\Omega^1_X)^{\otimes(n-1)+\cdots+(n-i)} \hookrightarrow \bigwedge^i \mathcal{L}$ pour $i = 1, \ldots, n-1$ n'aient pas de torsion.

Le tore $\mathbb{G}^n_{\mathrm{m}}$ agit sur $\mathcal{Y}^d_n$ par

$$(t_1, \ldots, t_n) \cdot (\mathcal{L}, (s^1, \ldots, s^n)) = (\mathcal{L}, (t_1 s^1, \ldots, t_n s^n)).$$

Cette action respecte l'ouvert $\mathcal{Y}^{d\,\circ}_n$ et est donnée sur le champ des triplets par

$$(t_1, \ldots, t_n) \cdot (\mathcal{L}, \mathcal{L}_\bullet, (s_1, \ldots, s_{n-1}, s_n)) = (\mathcal{L}, \mathcal{L}_\bullet, (t_n t_{n-1}^{-1} s_1, \ldots, t_2 t_1^{-1} s_{n-1}, t_1 s_n)).$$

En particulier, on peut identifier le champ algébrique $\mathcal{U}_n$ (resp. $\overline{\mathcal{U}}_n = [\mathcal{U}_n/\mathbb{G}_{\mathrm{m}}]$) introduit dans la section 4, au quotient de $\mathcal{Y}^{0\,\circ}_n$ par l'action de $\mathbb{G}_{\mathrm{m}}$ (resp. $\mathbb{G}^2_{\mathrm{m}}$) à travers le plongement $\mathbb{G}_{\mathrm{m}} \hookrightarrow \mathbb{G}^n_{\mathrm{m}},\ t \to (t, t^2, \ldots, t^n)$ (resp. $\mathbb{G}^2_{\mathrm{m}} \hookrightarrow \mathbb{G}^n_{\mathrm{m}},\ (t, t') \mapsto (t, t^2 t', t^3 t'^3, \ldots, t^n t'^{n(n-1)/2})$).

Proposition 7.3 ([B-G] Proposition 1.2.2). — *Le morphisme d'oubli* $(\mathcal{L}, s^\bullet) \to (\mathcal{L}, s^1)$ *passe au quotient en un morphisme de champs algébriques*

$$\widetilde{f} : \widetilde{\mathcal{Y}}^d_n := [\mathcal{Y}^d_n/\mathbb{G}^n_{\mathrm{m}}] \longrightarrow \mathrm{Fib}'^{\,n(n-1)(g-1)+d}_n(X)$$

qui est représentable et propre. □

Considérons le complexe $(\overline{g}_n)^*\Psi$ sur $\overline{\mathcal{U}}_n = [\mathcal{Y}^{0\,\circ}_n/\mathbb{G}^2_{\mathrm{m}}]$ introduit dans la section 4, et notons $\Phi \in D^{\mathrm{b}}_{\mathrm{c}}(\widetilde{\mathcal{Y}}^0_n)$ son image directe à supports propres par le morphisme composé de l'application quotient $[\mathcal{Y}^{0\,\circ}_n/\mathbb{G}^2_{\mathrm{m}}] \to [\mathcal{Y}^{0\,\circ}_n/\mathbb{G}^n_{\mathrm{m}}]$ et de l'inclusion $[\mathcal{Y}^{0\,\circ}_n/\mathbb{G}^n_{\mathrm{m}}] \subset [\mathcal{Y}^0_n/\mathbb{G}^n_{\mathrm{m}}] = \widetilde{\mathcal{Y}}^d_n$.

Pour tout système local E de rang n sur X, on a une variante du foncteur $M^d_{n,E}$

$$N^d_{n,E} : D^{\mathrm{b}}_{\mathrm{c}}(\widetilde{\mathcal{Y}}^0_n) \longrightarrow D^{\mathrm{b}}_{\mathrm{c}}(\widetilde{\mathcal{Y}}^d_n),\ K \longrightarrow p_* p'^*(K \otimes q^* \mathcal{L}^d_E)[nd],$$

définie à l'aide du diagramme

$$\begin{array}{ccccc} & & \widetilde{\mathcal{Z}}^d_n & \xrightarrow{q} & \mathrm{Coh}^d_0(X) \\ & \swarrow^{p'} & & \searrow_{p} & \\ \widetilde{\mathcal{Y}}^0_n & & & & \widetilde{\mathcal{Y}}^d_n \end{array}$$

où

$$\widetilde{\mathcal{Z}}^d_n = \widetilde{\mathcal{Y}}^0_n \times_{\mathrm{Fib}_n(X), p'} \mathrm{Mod}^d_n(X)$$

est le quotient par $\mathbb{G}^n_{\mathrm{m}}$ du champ algébrique des triplets $(\mathcal{L}', s'^\bullet, \mathcal{L}' \subset \mathcal{L})$ avec $(\mathcal{L}', s'^\bullet) \in \mathcal{Y}^0_n$ et $(\mathcal{L}' \subset \mathcal{L}) \in \mathrm{Mod}^d_i(X)$, et où p, p' et q sont induits par les flèches qui

envoient un tel triplet sur $(\mathcal{L}, s^\bullet)$, $(\mathcal{L}', s'^\bullet)$ et $\mathcal{L}/\mathcal{L}'$ respectivement avec s^i le composé de s'^i et de l'inclusion $\bigwedge^i \mathcal{L}' \subset \bigwedge^i \mathcal{L}$.

LEMME 7.4. — *Soit E un système local de rang n sur X. Pour tout entier $d \geqslant 0$, on a*

$$\mathrm{Aut}'_E \,|\, \mathrm{Fib}'^{\,n(n-1)(g-1)+d}_n = \widetilde{f}_* N^d_{n,E}(\Phi).$$

Pour montrer que Aut'_E se descend par le morphisme $\mathrm{Fib}'_n(X) \to \mathrm{Fib}_n(X)$ d'oubli de la section, Frenkel, Gaitsgory et Vilonen utilisent le lemme 7.4, la proposition 7.3 et le lemme 7.1 pour montrer que la fonction caractéristique d'Euler-Poincaré de Aut'_E ne dépend pas de E. Puis ils calculent cette fonction d'Euler-Poincaré quand E est le système local trivial de rang n, et enfin ils utilisent le lemme 7.2 pour conclure.

RÉFÉRENCES

[B-B-D] A. BEILINSON, J. BERNSTEIN & P. DELIGNE – *Faisceaux pervers*, Astérisque, vol. 100, Soc. Math. France, Paris, 1982.

[B-D] A. BEILINSON & V. DRINFELD – « Quantization of Hitchin's Integrable System and Hecke Eigensheaves », prépublication, `http://www.math.uchicago.edu/~benzvi/BD/hitchin.ps.gz`.

[B-L] J. BERNSTEIN & V. LUNTS – *Equivariant Sheaves and Functors*, Lecture Notes in Mathematics, vol. 1578, Springer-Verlag, 1974.

[B-G] A. BRAVERMAN & D. GAITSGORY – « Geometric Eisenstein series », *Invent. Math.* **150** (2002), p. 287–384.

[Br] J.-L. BRYLINSKI – « Transformations canoniques, dualité projective, théorie de Lefschetz, transformations de Fourier et sommes trigonométriques », in *Géométrie et analyse microlocales*, Astérisque, vol. 140–141, Soc. Math. France, Paris, 1986, p. 3–134.

[DR1] V.G. DRINFELD – « Langlands' conjecture for GL(2) over functional fields », in *Proceedings of the International Congress of Mathematicians (Helsinki, 1978)*, Acad. Sci. Fennica, Helsinki, 1980, p. 565–574.

[DR2] ———, « Two-dimensional ℓ-adic representations of the fundamental group of a curve over a finite field and automorphic forms on GL(2) », *Amer. J. Math.* **105** (1983), p. 85–114.

[DR3] ———, « Two-dimensional ℓ-adic representations of the Galois group of a global field of characteristic p and automorphic forms on GL(2) », *J. of Soviet Math.* **36** (1987), p. 93–105.

[F-G-K-V] E. FRENKEL, D. GAITSGORY, D. KAZHDAN & K. VILONEN – « Geometric realization of Whittaker functions and the Langlands conjecture », *J. Amer. Math. Soc.* **11** (1998), p. 451–484.

[F-G-V1] E. FRENKEL, D. GAITSGORY & K. VILONEN – « Whittaker patterns in the geometry of moduli spaces of bundles on curves », *Annals of Math.* **153** (2001), p. 699–748.

[F-G-V2] ———, « On the geometric Langlands conjecture », *J. Amer. Math. Soc.* **15** (2002), p. 367–417.

[Ga] D. Gaitsgory – « On a vanishing conjecture appearing in the geometric Langlands correspondence », prépublication, `http://arXiv.org/abs/math/0204081`, 2002.

[Il] L. Illusie – « Théorie de Brauer et caractéristiques d'Euler-Poincaré d'après Deligne », in *Caractéristique d'Euler-Poincaré (séminaire ENS, 1978–1979)*, Astérisque, vol. 82–83, Soc. Math. France, Paris, 1981, p. 161–172.

[La] L. Lafforgue – « Chtoucas de Drinfeld et correspondance de Langlands », *Invent. Math.* **147** (2002), p. 1–241.

[Lau1] G. Laumon – « Correspondance de Langlands géométrique pour les corps de fonctions », *Duke Math. J.* **54** (1987), p. 309–359.

[Lau2] ———, « Faisceaux automorphes pour GL_n : la première construction de Drinfeld », prépublication, `http://arXiv.org/abs/math/9511004`, 1995.

[Lau3] ———, « Transformation de Fourier homogène », *Bull. Soc. math. France* **131** (2003), p. 527–551.

[L-M] G. Laumon & L. Moret-Bailly – *Champs algébriques*, Springer-Verlag, 1999.

[Ly1] S. Lysenko – « Local geometrized Rankin-Selberg method for $\mathrm{GL}(n)$ », *Duke Math. J.* **111** (2002), p. 451–493.

[Ly2] ———, « Global geometrized Rankin-Selberg method for $\mathrm{GL}(n)$ », `http://arxiv.org/abs/math.AG/0108208`, 2001.

[Se] J.-P. Serre – *Groupes algébriques et corps de classes*, Hermann, 1975.

Gérard LAUMON

Université Paris-Sud
UMR 8628 du CNRS
Mathématique, Bât. 425
F-91405 Orsay Cedex
E-mail : `Gerard.Laumon@math.u-psud.fr`

Séminaire BOURBAKI Juin 2002
54e année, 2001-2002, nO 907, p. 285 à 305

DESSINS D'ENFANTS

par **Joseph OESTERLÉ**

> Cette découverte, qui techniquement se réduit à si peu de choses, a fait sur moi une impression très forte, et elle représente un tournant décisif dans le cours de mes réflexions, un déplacement notamment de mon centre d'intérêt en mathématique, qui soudain s'est trouvé fortement localisé. Je ne crois pas qu'un fait mathématique m'ait jamais autant frappé que celui-là, et ait eu un impact psychologique comparable. Cela tient certainement à la nature tellement familière, non technique, des objets considérés, dont tout dessin d'enfant griffonné sur un bout de papier (pour peu que le graphisme soit d'un seul tenant) donne un exemple parfaitement explicite. À un tel dessin se trouvent associés des invariants arithmétiques subtils, qui seront chamboulés complètement dès qu'on y rajoute un trait de plus.
>
> Alexander Grothendieck, *Esquisse d'un programme,* 1984.

En 1984, Alexander Grothendieck présente un programme de recherche, intitulé *Esquisse d'un programme* ([9]), pour demander son détachement au CNRS (qu'il obtiendra, et conservera jusqu'à son départ en retraite en 1988). Grothendieck y utilise le terme de dessin d'enfant (dans son sens courant) comme un analogue imagé de certaines cartes cellulaires ; il explique que « toute carte orientée finie se réalise canoniquement sur une courbe algébrique complexe », et que « le groupe de Galois de $\overline{\mathbf{Q}}$ sur $\mathbf{Q}$ opère sur la catégorie de ces cartes de façon naturelle » : cela se déduit de la comparaison de différents points de vue sur les revêtements de $\mathbf{P}_1 - \{0, 1, \infty\}$. Depuis, le terme de dessin d'enfant a été souvent repris, avec un sens mathématique variable suivant les auteurs, pour désigner les objets (ou classes d'isomorphisme d'objets) intervenant dans l'un ou l'autre de ces points de vue. Nous ne chercherons pas à le définir ici, et nous nous contenterons de l'utiliser pour désigner l'ensemble de la théorie.

Voici quelques raisons qui plaident pour porter une attention particulière aux revêtements finis de la courbe $\mathbf{P}_1 - \{0, 1, \infty\}$:

a) C'est la courbe algébrique la plus simple dont le groupe fondamental n'est pas commutatif.

b) Elle a beaucoup de revêtements sur $\overline{\mathbf{Q}}$: d'après un théorème de Belyĭ, toute courbe algébrique intègre sur $\overline{\mathbf{Q}}$ possède un ouvert de Zariski non vide qui se réalise comme un tel revêtement.

c) Elle s'identifie à l'espace des modules $M_{0,4}$ des courbes de genre 0 munies de 4 points marqués. L'étude de l'action de $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ sur son π_1 est le point de départ de l'étude de la tour de Grothendieck-Teichmüller (formée des groupoïdes fondamentaux de tous les espaces de modules $M_{g,n}$, sur lesquels opère $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$).

Notations

a) Si K est un corps, $\mathbf{P}_\mathrm{K}$ désigne la droite projective, considérée comme une courbe algébrique sur K, et $\mathbf{P}(\mathrm{K}) = \mathrm{K} \cup \{\infty\}$ l'ensemble de ses points rationnels ; on note $\mathbf{P}'_\mathrm{K}$ et $\mathbf{P}'(\mathrm{K})$ les complémentaires de $\{0, 1, \infty\}$ dans $\mathbf{P}_\mathrm{K}$ et $\mathbf{P}(\mathrm{K})$.

b) On note $\overline{\mathbf{Q}}$ l'ensemble des nombres complexes qui sont algébriques sur $\mathbf{Q}$. C'est une clôture algébrique de $\mathbf{Q}$. On note $\mathrm{G}_\mathbf{Q}$ son groupe de Galois $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$; c'est un groupe profini.

1. REVÊTEMENTS FINIS DE $\mathbf{P}_1 - \{0, 1, \infty\}$

1.1. Le point de vue topologique et le point de vue algébrique complexe

Les revêtements finis de la droite projective complexe privée de 0, 1, ∞ peuvent être considérés de deux points de vue :

— *le point de vue topologique* : ce sont les revêtements finis de l'espace topologique $\mathbf{P}'(\mathbf{C}) = \mathbf{C} - \{0, 1\}$;

— *le point de vue algébrique* : ce sont les revêtements étales de la courbe algébrique complexe $\mathbf{P}'_\mathbf{C}$.

Ces deux points de vue sont équivalents. En effet, d'après un théorème de Grothendieck (*cf.* [13], XII, th. 5.1), pour tout schéma S de type fini sur $\mathbf{C}$, le foncteur « passage aux points complexes » est une équivalence de la catégorie des revêtements étales de S sur celle des revêtements (topologiques) finis de $\mathrm{S}(\mathbf{C})$. Dans le cas particulier qui nous intéresse, celui où $\mathrm{S} = \mathbf{P}'$, on peut bien sûr déduire directement ce résultat du théorème d'existence de Riemann.

> *Remarque.* — L'anneau des fonctions régulières de $\mathbf{P}'_\mathbf{C}$ est $\mathbf{C}[z, z^{-1}, (1-z)^{-1}]$ (où z est une indéterminée). Tout revêtement étale de $\mathbf{P}'_\mathbf{C}$ est affine. Le foncteur « algèbre des fonctions régulières » est une équivalence de la catégorie des revêtements étales de $\mathbf{P}'_\mathbf{C}$ sur la catégorie opposée à celle des algèbres étales et finies sur $\mathbf{C}[z, z^{-1}, (1-z)^{-1}]$.

1.2. Le point de vue des revêtements ramifiés

Un *revêtement ramifié fini* d'une surface topologique compacte S est par définition un couple (X, p), où X est une surface topologique compacte et $p : X \to S$ une application continue, dont le germe en chaque point $x \in X$ est isomorphe à celui de $z \mapsto z^{e(x)}$ au voisinage de 0 dans $\mathbf{C}$, pour un entier $e(x) \geqslant 1$. L'entier $e(x)$ s'appelle *l'indice de ramification* de p en x ; s'il est égal à 1, on dit que p est *non ramifié* en x. Tout revêtement fini du complémentaire d'une partie finie de S se prolonge de manière unique (à isomorphisme unique près) en un revêtement ramifié fini de S.

Le foncteur de restriction est donc une équivalence de la catégorie des revêtements ramifiés finis de $\mathbf{P}(\mathbf{C})$, non ramifiés au-dessus de $\mathbf{P}'(\mathbf{C})$, dans celle des revêtements finis de $\mathbf{P}'(\mathbf{C})$.

On a des résultats analogues dans la situation algébrique : les revêtements ramifiés considérés dans ce cas sont les couples (X, p), où X est une courbe algébrique projective et lisse sur $\mathbf{C}$ et $p : X \to \mathbf{P}_{\mathbf{C}}$ un morphisme fini, étale au-dessus de $\mathbf{P}'_{\mathbf{C}}$.

> *Remarque.* — Le corps des fonctions rationnelles de $\mathbf{P}_{\mathbf{C}}$ est $\mathbf{C}(z)$. Le foncteur « algèbre des fonctions rationnelles » est une équivalence de la catégorie des revêtements ramifiés de $\mathbf{P}_{\mathbf{C}}$, étales au-dessus de $\mathbf{P}'_{\mathbf{C}}$, sur la catégorie opposée à celle des algèbres réduites de dimension finie sur $\mathbf{C}(z)$, non ramifiées en dehors de $0, 1, \infty$.

Le foncteur « passage aux points complexes » permet de passer, pour les revêtements ramifiés, du cadre algébrique au cadre topologique. Les fibres en 0, 1 et ∞ des revêtements ramifiés considérés dans les deux cadres s'identifient canoniquement ; les notions d'indices de ramification définies dans les deux cadres coïncident.

1.3. Le point de vue des groupes fondamentaux (cas topologique)

Rappelons la définition du groupoïde fondamental $\varpi_1(S)$ d'un espace topologique S : ses points sont ceux de S ; les flèches reliant un point a à un point b sont les classes d'homotopie strictes de chemins reliant a à b ; l'ensemble de ces flèches est noté $\pi_1(S; a, b)$, ou $\pi_1(S, a)$ si $a = b$. La composée de deux flèches $\gamma \in \pi_1(S; a, b)$ et $\gamma' \in \pi_1(S; b, c)$ sera notée $\gamma'\gamma$ (contrairement à l'usage en topologie où on l'écrit plutôt $\gamma\gamma'$).

Soit (Y, q) un revêtement de S. La famille $(q^{-1}(a))_{a \in S}$ des fibres de q est un $\varpi_1(S)$-ensemble : chaque élément de $\pi_1(S; a, b)$ définit, par relèvement des chemins, une bijection de $q^{-1}(a)$ sur $q^{-1}(b)$. Supposons que S soit localement contractile (ou plus généralement que chaque point $a \in S$ possède un voisinage U connexe par arcs tel que l'homomorphisme $\pi_1(U, a) \to \pi_1(S, a)$ soit nul). Le foncteur $(Y, q) \mapsto (q^{-1}(a))_{a \in S}$ est alors une équivalence de la catégorie des revêtements de S dans celle des $\varpi_1(S)$-ensembles ; les revêtements finis correspondent aux $\varpi_1(S)$-ensembles $(E_a)_{a \in S}$ formés d'ensembles finis.

On peut remplacer dans ce qui précède le groupoïde fondamental par un sous-groupoïde plein, contenant au moins un point dans chaque composante connexe par arcs de S. En particulier, si S est connexe par arcs et $a \in S$, le foncteur « fibre en a »

est une équivalence de la catégorie des revêtements finis de S sur celle des $\pi_1(\mathrm{S}; a)$-ensembles finis.

Ceci s'applique en particulier au cas où $\mathrm{S} = \mathbf{P}'(\mathbf{C})$: pour tout $a \in \mathbf{P}'(\mathbf{C})$, le foncteur « fibre en a » est une équivalence de la catégorie des revêtements finis de $\mathbf{P}'(\mathbf{C})$ sur celle des $\pi_1(\mathbf{P}'(\mathbf{C}), a)$-ensembles finis.

Remarques. — 1) On peut faire jouer le rôle de point-base à un sous-ensemble contractile de $\mathbf{P}'(\mathbf{C})$. Prenons par exemple l'intervalle $]0, 1[$ et notons π le groupe fondamental correspondant ; il possède deux générateurs canoniques c_0 et c_1 (correspondant à un tour effectué dans le sens trigonométrique autour de 0 et 1 respectivement) ; le π-ensemble associé à (Y, q) s'identifie dans ce cas à l'ensemble des composantes connexes de $q^{-1}(]0, 1[)$.

2) On peut également faire jouer le rôle de point-base à un germe d'ensemble contractile (par exemple le germe en 0 de l'intervalle $]0, 1[$) ; le rôle de la fibre est alors tenu par l'ensemble des relèvements continus de ce germe.

3) Le groupe des permutations de $\{0, 1, \infty\}$ opère sur $\mathbf{P}'(\mathbf{C})$. Plutôt que de choisir un seul point-base, il est parfois plus commode d'en prendre un nombre fini, stable par ce groupe de symétrie. Dans *Esquisse d'un programme*, Grothendieck propose le choix suivant : on interprète $\mathbf{P}'(\mathbf{C})$ comme l'espace de modules $M_{0,4}(\mathbf{C})$ paramétrant les classes d'isomorphisme de courbes de genre 0 (projectives, lisses, irréductibles sur $\mathbf{C}$) munies (d'une suite ordonnée) de 4 points marqués (deux à deux distincts) : un point $a \in \mathbf{P}'(\mathbf{C})$ correspond à la classe d'isomorphisme de la droite projective, munie de $(0, 1, \infty, a)$.

Considérons une courbe de genre 0 munie de 4 points marqués ; en général, le groupe des automorphismes de la courbe qui stabilisent l'ensemble des points marqués est d'ordre 4 (et isomorphe au groupe de Klein). Il n'est plus grand que lorsque la classe d'isomorphisme de la courbe, munie de ses points marqués, correspond à l'un des points -1, $\frac{1}{2}$, 2, ρ, $\overline{\rho}$ de $\mathbf{P}'(\mathbf{C})$ (avec $\rho = \frac{1+i\sqrt{3}}{2}$). Grothendieck propose de prendre ces 5 points de $\mathbf{P}'(\mathbf{C})$ comme points-base.

Les classes d'isomorphisme des courbes de genre 0 munies de 4 points marqués, et qui possèdent une structure réelle pour laquelle l'ensemble des points marqués est stable par conjugaison complexe, correspondent dans $\mathbf{P}'(\mathbf{C})$ à la réunion des deux droites et des deux cercles représentés sur la figure 1 ci-dessous. En ne conservant que les portions de ces courbes reliant les points-base choisis, on obtient 6 chemins reliant respectivement chacun des trois points -1, $\frac{1}{2}$, 2 à chacun des deux points ρ, $\overline{\rho}$. Le sous-groupoïde plein du groupoïde fondamental $\mathbf{P}'(\mathbf{C})$ bâti sur ces 5 points-base est le groupoïde libre engendré par les classes de ces six chemins.

Cette construction qui peut sembler pédante dans le cas de $\mathbf{P}'(\mathbf{C})$ prend par contre tout son sel lorsqu'on essaie de la généraliser aux espaces de modules $M_{g,n}$.

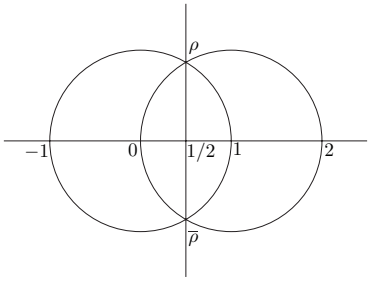


FIGURE 1

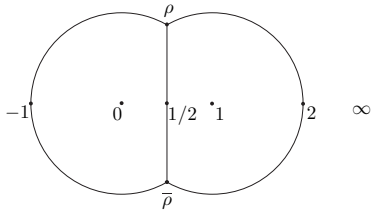


FIGURE 2

1.4. Le point de vue des groupes fondamentaux (cas algébrique)

Le groupoïde fondamental a un analogue en géométrie algébrique. Rappelons-en brièvement la définition, due à Grothendieck. Soit S un schéma. Un *point géométrique* de S est un point a de S à valeurs dans un corps séparablement clos ; sa donnée équivaut à celle d'un point s de S et d'une extension séparablement close k du corps résiduel de s. On définit la fibre en a d'un revêtement étale (Y, q) de S comme l'image réciproque de a par l'application $\mathrm{Y}(k) \to \mathrm{S}(k)$ déduite de q. On obtient ainsi un foncteur « fibre en a » de la catégorie des revêtements étales de S dans celle des ensembles finis.

Le groupoïde fondamental de S, noté $\varpi_1^{ét}(\mathrm{S})$, ou simplement $\varpi_1(\mathrm{S})$ lorsque cela ne prête pas à confusion, a pour points les points géométriques de S (en limitant les corps où ils prennent leurs valeurs à un univers, si on veut que ces points forment un ensemble) ; les flèches reliant un point a à un point b sont par définition les isomorphismes du foncteur « fibre en a » sur le foncteur « fibre en b » de la catégorie des revêtements étales de S dans celle des ensembles finis ; l'ensemble de ces flèches est noté $\pi_1^{ét}(\mathrm{S}; a, b)$, ou $\pi_1^{ét}(\mathrm{S}, a)$ si $a = b$ (la mention « *ét* » pouvant être omise lorsque cela ne prête pas à confusion). On munit $\pi_1^{ét}(\mathrm{S}; a, b)$ de la topologie de la convergence simple dans les fibres en a des revêtements étales de S ; c'est un ensemble profini. La composition et l'inversion des flèches sont continues.

Supposons que le schéma S n'ait qu'un nombre fini de composantes connexes (ce qui est le cas par exemple s'il est nœthérien). Le foncteur qui à un revêtement étale de S associe la famille de ses fibres géométriques, munie de l'opération du groupoïde $\varpi_1^{ét}(\mathrm{S})$, est une équivalence de la catégorie des revêtements étales de S dans celle des familles (indexées par l'ensemble des points géométriques de S) d'ensembles finis, munies d'une opération continue de $\varpi_1^{ét}(\mathrm{S})$. On peut remplacer dans ce qui précède le groupoïde fondamental par tout sous-groupoïde plein contenant au moins un point dans chaque composante connexe de S. En particulier, si S est connexe et que a est un point géométrique de S, le foncteur « fibre en a » est une équivalence de la catégorie des revêtements étales de S sur celle des ensembles finis, munis d'une opération continue du groupe profini $\pi_1^{ét}(\mathrm{S}, a)$.

Lorsque S est un schéma de type fini sur $\mathbf{C}$, tout point a de $\mathrm{S}(\mathbf{C})$ est un point géométrique de S et l'équivalence de catégorie rappelée au n° 1.1 permet d'associer à toute classe de chemins reliant a à un point b de $\mathrm{S}(\mathbf{C})$ un élément de $\pi_1^{ét}(\mathrm{S}; a, b)$. Cela définit un morphisme canonique du groupoïde topologique $\varpi_1(\mathrm{S}(\mathbf{C}))$ dans le groupoïde algébrique $\varpi_1^{ét}(\mathrm{S})$. Pour $a, b \in \mathrm{S}(\mathbf{C})$, $\pi_1^{ét}(\mathrm{S}; a, b)$ s'identifie au séparé complété de $\pi_1(\mathrm{S}(\mathbf{C}); a, b)$ pour la structure uniforme suivante : un système fondamental d'entourages est formé des ensembles $\{(\gamma, \delta) \mid \delta \in \gamma \mathrm{U}\}$, où U parcourt l'ensemble des sous-groupes d'indice fini de $\pi_1(\mathrm{S}(\mathbf{C}), a)$.

Ceci s'applique en particulier au cas où $\mathrm{S} = \mathbf{P}'_{\mathbf{C}}$: pour tout point géométrique a de $\mathbf{P}'_{\mathbf{C}}$, le foncteur « fibre en a » est une équivalence de la catégorie des revêtements

étales de $\mathbf{P}'_{\mathbf{C}}$ sur celle des ensembles finis, munis d'une opération à gauche continue de $\pi_1^{\text{ét}}(\mathbf{P}'_{\mathbf{C}}, a)$. Lorsque $a \in \mathbf{P}'(\mathbf{C})$, le groupe topologique $\pi_1^{\text{ét}}(\mathbf{P}'_{\mathbf{C}}, a)$ s'identifie au complété profini de $\pi_1(\mathbf{P}'(\mathbf{C}), a)$. Comme $\pi_1(\mathbf{P}'(\mathbf{C}), a)$ est isomorphe au groupe libre à deux générateurs, l'homomorphisme $\pi_1(\mathbf{P}'(\mathbf{C}), a) \to \pi_1^{\text{ét}}(\mathbf{P}'_{\mathbf{C}}, a)$ est injectif.

Remarques. — 1) Soient S une courbe algébrique lisse sur un corps de caractéristique 0, $\overline{\mathrm{S}}$ sa compactifiée lisse, et ξ un vecteur tangent non nul à $\overline{\mathrm{S}}$ en un point géométrique a de $\overline{\mathrm{S}} - \mathrm{S}$. On peut faire jouer à ξ le rôle d'un point-base géométrique de S (*cf.* [6]). La fibre en ξ d'un revêtement étale (Y, q) de S peut être décrite dans ce cas comme suit : on prolonge le revêtement étale en un morphisme fini $(\mathrm{X}, p) \to \overline{\mathrm{S}}$, où X est la compactifiée lisse de Y ; si b est un point géométrique de X au-dessus de a, et e l'indice de ramification de p en b, la partie principale de p en b est une application homogène t_b de degré e de l'espace tangent à X en b dans l'espace tangent à $\overline{\mathrm{S}}$ en a ; on prend alors pour fibre de ξ la réunion disjointe des ensembles $t_b^{-1}(\xi)$, pour b au-dessus de a. Le groupe fondamental de S en ξ peut encore être défini comme le groupe des automorphismes du foncteur « fibre en ξ ». Nous en donnerons une description galoisienne ultérieurement (*cf.* 2.5, remarque).

2) Pour $i \neq j$ dans $\{0, 1, \infty\}$, définissons un vecteur tangent à $\mathbf{P}_{\mathbf{C}}$ en i comme suit : $\overrightarrow{01}$ est le vecteur tangent $\frac{d}{dz}$ en 0 ; $\overrightarrow{ij}$ est son image par l'automorphisme de $\mathbf{P}_{\mathbf{C}}$ qui permute $\{0, 1, \infty\}$ en appliquant 0 sur i et 1 sur j. La remarque 1 s'applique par exemple en prenant pour S la courbe $\mathbf{P}'_{\mathbf{C}}$ et pour ξ l'un des six vecteurs tangents $\overrightarrow{ij}$. Le groupe fondamental de $\mathbf{P}'_{\mathbf{C}}$ en $\overrightarrow{01}$ est canoniquement isomorphe au complété profini du groupe fondamental (topologique) de $\mathbf{P}'(\mathbf{C})$ obtenu en faisant jouer le rôle de point-base soit à $]0, 1[$, soit au germe de $]0, 1[$ en 0 (*cf.* n° 1.3, remarques 1 et 2).

1.5. Le point de vue des cartes triangulaires orientées tricoloriées

La surface topologique $\mathbf{P}(\mathbf{C})$ possède une décomposition cellulaire canonique, dont le 1-squelette est $\mathbf{P}(\mathbf{R})$ et l'ensemble des sommets $\{0, 1, \infty\}$.

Cela permet d'associer à tout revêtement ramifié (X, p) de $\mathbf{P}(\mathbf{C})$, non ramifié au-dessus de $\mathbf{P}'(\mathbf{C})$, une carte triangulaire orientée tricoloriée $\mathrm{G} = (\mathrm{X}, \mathrm{K}, \mathrm{S}, t)$ (*cf.* Appendice, n° 2, pour la définition de ces cartes) : on munit la surface topologique X de l'orientation déduite par p de celle de $\mathbf{P}(\mathbf{C})$, on prend pour K l'ensemble fermé $p^{-1}(\mathbf{P}(\mathbf{R}))$, pour S l'ensemble fini $p^{-1}(\{0, 1, \infty\})$, et pour $t : \mathrm{S} \to \{0, 1, \infty\}$ l'application déduite de p.

Le foncteur ainsi défini est une équivalence de la catégorie des revêtements ramifiés de $\mathbf{P}(\mathbf{C})$, non ramifiés au-dessus de $\mathbf{P}'(\mathbf{C})$, sur la catégorie isotopique des cartes triangulaires orientées tricoloriées. (La définition des morphismes de cette catégorie, un peu technique, est précisée dans *loc. cit.*).

Exercice. — Considérons les ensembles d'arêtes de G de type $\{\infty, 0\}$, $\{0, 1\}$ et $\{1, \infty\}$ respectivement, et les ensembles de faces de G, positives et négatives respectivement. La relation d'incidence définit une bijection de chacun des trois premiers ensembles sur chacun des deux derniers. Vérifier que ces cinq ensembles et six bijections définissent un ϖ-ensemble, où ϖ est le groupoïde considéré dans la remarque 3 de 1.3, et que ce ϖ-ensemble est canoniquement isomorphe à celui associé au revêtement de $\mathbf{P}'(\mathbf{C})$ déduit de (X, p).

1.6. Le point de vue des cartes cellulaires orientées bicoloriées

À tout revêtement ramifié fini (X, p) de $\mathbf{P}(\mathbf{C})$, non ramifié au-dessus de $\mathbf{P}'(\mathbf{C})$, on associe une carte cellulaire orientée bicoloriée (X, K, S, b) (*cf.* Appendice, n° 3) en munissant la surface topologique X de la même orientation qu'au n° 1.5, en prenant pour K l'ensemble fermé $p^{-1}([0,1])$, pour S l'ensemble fini $p^{-1}(\{0,1\})$, et pour b : $S \to \{0,1\}$ l'application déduite de p.

Le foncteur ainsi défini est une équivalence de la catégorie des revêtements ramifiés fini de $\mathbf{P}(\mathbf{C})$, non ramifiés au-dessus de $\mathbf{P}'(\mathbf{C})$, sur la catégorie isotopique des cartes cellulaires orientées bicoloriées (*cf. loc. cit.*).

1.7. Le point de vue combinatoire

À tout revêtement fini (Y, q) de $\mathbf{P}'(\mathbf{C})$, associons l'ensemble E des composantes connexes de $q^{-1}(]0,1[)$, muni des permutations σ_0 et σ_1 suivant lesquelles les générateurs canoniques du groupe fondamental $\pi = \pi_1(\mathbf{P}'(\mathbf{C}),]0,1[)$ opèrent sur E (*cf.* remarque 1 du n° 1.3).

On définit ainsi une équivalence de la catégorie des revêtements finis de $\mathbf{P}'(\mathbf{C})$ dans celle des ensembles finis munis de deux permutations.

> *Remarque.* — Soit (X, p) l'unique (à isomorphisme unique près) revêtement ramifié de $\mathbf{P}(\mathbf{C})$ prolongeant (Y, q) et soit G la carte triangulaire orientée tricoloriée qui lui a été associée au n° 1.5. Le triplet (E, σ_0, σ_1) peut se décrire à partir de G comme suit : E est l'ensemble des arêtes de G de type $\{0,1\}$; si s est un sommet de type 0 (resp. de type 1) de G, les arêtes de G de type $\{0,1\}$ dont s est une extrémité sont munies d'un ordre cyclique[(1)] déduit de l'orientation de X en s, et σ_0 (resp. σ_1) applique chacune de ces arêtes sur la suivante pour cet ordre cyclique.

Notons $\langle\sigma_0\rangle$ et $\langle\sigma_1\rangle$ les sous-groupes de $\mathfrak{S}_E$ engendrés par σ_0 et σ_1 respectivement. Leurs orbites dans E paramètrent les points des fibres $p^{-1}(0)$ et $p^{-1}(1)$, *i.e.* les sommets de type 0 et 1 de G. On peut paramétrer les points de $p^{-1}(\infty)$, *i.e.* les sommets de type ∞ de G, par les orbites de $\langle\sigma_\infty\rangle$, où $\sigma_\infty = \sigma_0^{-1}\sigma_1^{-1}$, en associant à chaque sommet s de type ∞ l'ensemble des arêtes de type $\{0,1\}$ bordant une face positive dont s est un sommet. Les indices de ramification de p en les points au-dessus de 0, 1 et ∞ sont les cardinaux des orbites correspondantes de $\langle\sigma_0\rangle$, $\langle\sigma_1\rangle$ et $\langle\sigma_\infty\rangle$.

Pour que Y (ou ce qui revient au même X) soit connexe, il faut et il suffit que le groupe engendré par σ_0 et σ_1 opère transitivement sur E. La formule de Riemann-Hurwitz permet alors de « lire » le genre g de X sur (E, σ_0, σ_1) : on a $2g - 2 = n - n_0 - n_1 - n_\infty$, où n est le cardinal de E (égal au degré du revêtement), et n_0, n_1, n_∞ les nombres d'orbites respectifs de σ_0, σ_1, σ_∞ dans E.

(1) Un *ordre cyclique sur* un ensemble fini A est une permutation σ de A qui engendre un sous-groupe transitif de $\mathfrak{S}_A$; si a est un élément de A, on dit que $\sigma(a)$ est l'élément suivant pour cet ordre cyclique.

1.8. Le point de vue des sous-groupes d'indice fini de $\mathbf{SL}_2(\mathbf{Z})$

Notons $\mathfrak{H}$ le demi-plan de Poincaré, *i.e.* l'ensemble des nombres complexes de partie imaginaire >0. Le groupe $\mathbf{SL}_2(\mathbf{Z})$ opère sur $\mathfrak{H}$ par $\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right), \tau\right) \mapsto \frac{a\tau+b}{c\tau+d}$.

Notons $\Gamma(2)$ le sous-groupe d'indice 6 de $\mathbf{SL}_2(\mathbf{Z})$ formé des matrices congrues à la matrice unité modulo 2. La surface de Riemann $\mathrm{Y}(2) = \Gamma(2)\backslash\mathfrak{H}$ paramètre les classes d'isomorphisme de courbes elliptiques sur $\mathbf{C}$, munies de deux points d'ordre 2 distincts : à l'élément $\Gamma(2)\tau$ de $\mathrm{Y}(2)$ correspond la classe d'isomorphisme de la courbe elliptique $\mathbf{C}/(\mathbf{Z}+\mathbf{Z}\tau)$, munie des points d'ordre 2 images de $\frac{\tau}{2}$ et $\frac{1}{2}$. Il existe un unique nombre complexe $\lambda(\tau)$ tel que la courbe elliptique d'équation $y^2 = x(x-1)(x-\lambda(\tau))$, munie de ses points $(0,0)$ et $(1,0)$, soit isomorphe à la précédente.

La fonction λ ainsi définie est holomorphe sur $\mathfrak{H}$, et invariante par $\Gamma(2)$: c'est une fonction modulaire de poids 0 pour $\Gamma(2)$. Elle définit par passage au quotient un isomorphisme de $\mathrm{Y}(2)$ sur $\mathbf{P}'(\mathbf{C})$, qui se prolonge en un isomorphisme de la surface de Riemann $\mathrm{X}(2)$ (compactifiée de $\mathrm{Y}(2)$ par adjonction des pointes) sur $\mathbf{P}(\mathbf{C})$; celui-ci applique les trois pointes $\Gamma(2)\infty$, $\Gamma(2)0$ et $\Gamma(2)1$ de $\mathrm{X}(2)$ sur 0, 1 et ∞ respectivement. L'image par λ de la demi-droite $i]0,+\infty[$ est l'intervalle $]0,1[$.

La surjection canonique $\mathfrak{H} \to \mathrm{Y}(2)$ est un revêtement universel de $\mathrm{Y}(2)$. On en déduit que tout triplet $(\mathrm{Y}, q, \mathrm{A})$, où (Y, q) est un revêtement *connexe non vide* de $\mathbf{P}'(\mathbf{C})$ et A une composante connexe de $q^{-1}(]0,1[)$, est isomorphe à un unique triplet de la forme $(\mathrm{Y}_\Gamma, q_\Gamma, \mathrm{A}_\Gamma)$, où Γ est un sous-groupe d'indice fini de $\Gamma(2)$ contenant $\{\pm 1\}$, Y_Γ la surface de Riemann $\Gamma\backslash\mathfrak{H}$, q_Γ le composé de la surjection canonique $\mathrm{Y}_\Gamma \to \mathrm{Y}(2)$ et de l'isomorphisme $\mathrm{Y}(2) \to \mathbf{P}'(\mathbf{C})$ ci-dessus, et A_Γ l'image de la demi-droite $i]0,+\infty[$ dans Y_Γ. De plus l'isomorphisme entre ces deux triplets est unique.

> *Variante.* — Soient (X, p) un revêtement ramifié *connexe non vide* de $\mathbf{P}(\mathbf{C})$, non ramifié au-dessus de $\mathbf{P}'(\mathbf{C})$ et A une composante connexe de $p^{-1}(]0,1[)$. Supposons que les indices de ramification des points de $p^{-1}(1)$ divisent 2 et que ceux des points de $p^{-1}(\infty)$ divisent 3. Le triplet $(\mathrm{X}, p, \mathrm{A})$ est alors isomorphe à un unique triplet de la forme $(\mathrm{X}_\Gamma, p_\Gamma, \mathrm{A}_\Gamma)$, où Γ est un sous-groupe d'indice fini de $\mathbf{SL}_2(\mathbf{Z})$ contenant $\{\pm 1\}$, X_Γ est la surface de Riemann compactifiée de $\Gamma\backslash\mathfrak{H}$, p_Γ se déduit par passage au quotient de $\frac{1728}{j}$, où j est l'invariant modulaire, et A_Γ est l'image de la demi-droite $i]1,+\infty[$ dans X_Γ. De plus l'isomorphisme entre ces deux triplets est unique.
>
> *Exercice.* — Tout sous-groupe d'indice fini de $\Gamma(2)$ est aussi un sous-groupe d'indice fini de $\mathbf{SL}_2(\mathbf{Z})$. Interpréter cela, via les dictionnaires ci-dessus, en termes de subdivisions barycentriques de cartes cellulaires triangulaires tricoloriées, et en déduire sans calculs la relation $j = 2^8 \frac{(\lambda^2-\lambda+1)^3}{\lambda^2(\lambda-1)^2}$. Faire le lien avec la fig. 1.

1.9. Le point de vue algébrique sur $\overline{\mathbf{Q}}$

Soient k un corps algébriquement clos de caractéristique 0, k' une extension algébriquement close de k et S un schéma de type fini sur k. Soit S' le schéma sur k' déduit de S par extension des scalaires. Le foncteur « extension des scalaires de k à k' » est une équivalence de la catégorie des revêtements étales de S sur celle des revêtements étales de S'. De façon équivalente, si s' est un point géométrique de S' et s le point

géométrique correspondant de S, l'homomorphisme canonique $\pi_1^{ét}(\mathrm{S}', s') \to \pi_1^{ét}(\mathrm{S}, s)$ est un isomorphisme (*cf.* [13], XIII, cor. 3.5 et remarque 3.1.3).

En particulier, le foncteur « extension des scalaires de $\overline{\mathbf{Q}}$ à $\mathbf{C}$ » est une équivalence de la catégorie des revêtements étales de $\mathbf{P}'_{\overline{\mathbf{Q}}}$ sur celle des revêtements étales de $\mathbf{P}'_{\mathbf{C}}$. Comme précédemment, ces catégories sont aussi équivalentes aux suivantes :

— la catégorie des revêtements ramifiés de $\mathbf{P}_{\overline{\mathbf{Q}}}$, étales au-dessus de $\mathbf{P}'_{\overline{\mathbf{Q}}}$;

— la catégorie opposée à celle des algèbres étales et finies sur $\overline{\mathbf{Q}}[z, z^{-1}, (1-z)^{-1}]$;

— la catégorie opposée à celle des algèbres réduites de dimension finie sur $\overline{\mathbf{Q}}(z)$, non ramifiées en dehors de $0, 1, \infty$;

— la catégorie des ensembles finis munis d'une opération à gauche continue de $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, a)$, pour a un point géométrique de $\mathbf{P}'_{\overline{\mathbf{Q}}}$ (ou un vecteur tangent non nul à $\mathbf{P}_{\overline{\mathbf{Q}}}$ en 0, 1 ou ∞).

1.10. Le théorème de Belyĭ

Soient X une courbe algébrique projective et lisse sur $\overline{\mathbf{Q}}$ et p une fonction rationnelle sur X qui n'est constante sur aucune composante connexe de X. On peut considérer p comme un morphisme fini de X dans la droite projective $\mathbf{P}_{\overline{\mathbf{Q}}}$. On dit que p est une *fonction de Belyĭ* si ce morphisme est étale au-dessus de $\mathbf{P}'_{\overline{\mathbf{Q}}}$. Cette terminologie est justifiée par le très surprenant théorème suivant de Belyĭ ([1]) :

THÉORÈME 1. — *Soit* X *une courbe algébrique projective et lisse sur* $\overline{\mathbf{Q}}$. *Pour tout ensemble fini* $\mathrm{S} \subset \mathrm{X}(\overline{\mathbf{Q}})$, *il existe une fonction de Belyĭ* p *sur* X *telle que* $p(\mathrm{S}) \subset \{0, 1, \infty\}$.

Soit f une fonction rationnelle sur X, qui n'est constante sur aucune composante connexe de X, *i.e.* un morphisme fini de X dans $\mathbf{P}_{\overline{\mathbf{Q}}}$. L'ensemble R_f des points de $\mathrm{X}(\overline{\mathbf{Q}})$ en lesquels ce morphisme est ramifié est fini. Si q est une fonction de Belyĭ sur $\mathbf{P}_{\overline{\mathbf{Q}}}$ qui applique $f(\mathrm{S}) \cup f(\mathrm{R}_f)$ dans $\{0, 1, \infty\}$, on peut prendre $p = q \circ f$. Il nous suffit donc de traiter le cas où $\mathrm{X} = \mathbf{P}_{\overline{\mathbf{Q}}}$.

Notons dans ce cas d le degré maximum des points de S sur $\mathbf{Q}$ et e le nombre de points de S de degré d ; nous démontrerons le théorème par récurrence sur le couple (d, e) (pour l'ordre lexicographique). Traitons d'abord le cas où $d \geqslant 2$. Choisissons un point $a \in \mathrm{S}$ de degré d. Notons f son polynôme minimal ; il est de degré d, à coefficients rationnels. L'ensemble R_f se compose de ∞ et des racines du polynôme dérivé de f. Ses points sont donc de degré $\leqslant d-1$ sur $\mathbf{Q}$ et il en est de même des points de $f(\mathrm{R}_f)$. Quant aux points de $f(\mathrm{S})$, ils sont de degré $\leqslant d$ et il y en a au plus $e-1$ de degré d, puisque $f(a) = 0$. Le premier alinéa et l'hypothèse de récurrence permettent de conclure.

Traitons maintenant le cas où $d = 1$, c'est-à-dire où $\mathrm{S} \subset \mathbf{P}(\mathbf{Q})$. Si $e \leqslant 3$, on peut prendre pour p une homographie. Supposons $e \geqslant 4$. Par une homographie, on se ramène au cas où S contient $\{0, 1, \infty\}$ et au moins un point a dans l'intervalle $]0, 1[$.

Écrivons $a = \frac{m}{m+n}$ avec m, n entiers $\geqslant 1$ premiers entre eux et notons f le polynôme $t^m(t-1)^n$. On a $\mathrm{R}_f \subset \{0, 1, \infty, a\} \subset \mathrm{S}$ et $f(\mathrm{S})$ a au plus $e-1$ éléments, puisque $f(0) = f(1) = 0$. Le premier alinéa et l'hypothèse de récurrence permettent encore de conclure.

2. OPÉRATIONS DE $\mathrm{G}_{\mathbf{Q}}$ SUR LES DESSINS D'ENFANTS

2.1. Opération de $\mathrm{G}_{\mathbf{Q}}$ sur la catégorie des revêtements étales de $\mathbf{P}'_{\overline{\mathbf{Q}}}$

Posons $\mathrm{R} = \overline{\mathbf{Q}}[z, z^{-1}, (z-1)^{-1}]$. Pour $\sigma \in \mathrm{G}_{\mathbf{Q}}$, notons σ_{R} l'automorphisme de l'anneau R qui prolonge σ et fixe z. Soit A une algèbre sur R, *i.e.* un anneau muni d'un homomorphisme $\rho : \mathrm{R} \to \mathrm{A}$. Le même anneau, muni de l'homomorphisme $\rho \circ \sigma_{\mathrm{R}}^{-1}$ est une nouvelle R-algèbre que nous noterons ${}^{\sigma}\mathrm{A}$ (et qui est canoniquement isomorphe à celle déduite de A par l'extension des scalaires σ_{R}). Tout homomorphisme de R-algèbres $u : \mathrm{A} \to \mathrm{B}$ est aussi un homomorphisme de R-algèbres de ${}^{\sigma}\mathrm{A}$ dans ${}^{\sigma}\mathrm{B}$, que nous noterons ${}^{\sigma}u$. Si τ est un second élément de $\mathrm{G}_{\mathbf{Q}}$, on a ${}^{\sigma}({}^{\tau}\mathrm{A}) = {}^{\sigma\tau}\mathrm{A}$ et ${}^{\sigma}({}^{\tau}u) = {}^{\sigma\tau}u$. Nous avons ainsi défini une opération de $\mathrm{G}_{\mathbf{Q}}$ sur la catégorie des R-algèbres. Elle stabilise la sous-catégorie pleine formée des algèbres étales et finies sur R.

Explicitons l'opération correspondante de $\mathrm{G}_{\mathbf{Q}}$ sur la catégorie des revêtements étales de $\mathbf{P}'_{\overline{\mathbf{Q}}}$: si (Y, q) est un tel revêtement, le revêtement ${}^{\sigma}(\mathrm{Y}, q) = ({}^{\sigma}\mathrm{Y}, {}^{\sigma}q)$ (dit *conjugué de* (Y, q) *par* σ) est défini comme suit : ${}^{\sigma}\mathrm{Y}$ a même schéma sous-jacent que Y, mais son morphisme structural vers $\mathrm{Spec}(\overline{\mathbf{Q}})$ est celui de Y composé avec $\mathrm{Spec}(\sigma^{-1})$; le morphisme ${}^{\sigma}q : {}^{\sigma}\mathrm{Y} \to \mathbf{P}'_{\overline{\mathbf{Q}}} = \mathrm{Spec}(\mathrm{R})$ est le composé de q avec $\mathrm{Spec}(\sigma_{\mathrm{R}}^{-1})$.

> *Remarque.* — Le groupe $\mathrm{G}_{\mathbf{Q}}$ opère de manière analogue sur la catégorie des $\overline{\mathbf{Q}}(z)$ algèbres réduites de dimension finie, non ramifiées en dehors de $\{0, 1, \infty\}$ et sur celle des revêtements ramifiés de $\mathbf{P}_{\overline{\mathbf{Q}}}$, étales au-dessus de $\mathbf{P}'_{\overline{\mathbf{Q}}}$.
>
> Il opère aussi sur l'une quelconque des autres catégories $\mathscr{C}$ équivalentes à cette dernière (par exemple celle des ensembles finis munis de deux permutations), avec cependant dans ce cas les complications mineures venant du fait que, pour un objet D de $\mathscr{C}$ et $\sigma \in \overline{\mathbf{Q}}$, l'objet ${}^{\sigma}\mathrm{D}$ n'est que défini de manière unique à isomorphisme unique près, et que ${}^{\sigma}({}^{\tau}\mathrm{D})$ n'est pas égal, mais seulement canoniquement isomorphe à ${}^{\sigma\tau}\mathrm{D}$, etc.

2.2. Corps de définition, corps des modules d'un revêtement

Soient (Y, q) un revêtement étale de $\mathbf{P}'_{\overline{\mathbf{Q}}}$ et L un sous-corps de $\overline{\mathbf{Q}}$. Un *modèle* de (Y, q) sur L est un triplet $(\mathrm{Y}_{\mathrm{L}}, q_{\mathrm{L}}, \varphi)$, où $(\mathrm{Y}_{\mathrm{L}}, q_{\mathrm{L}})$ est un revêtement étale de $\mathbf{P}'_{\mathrm{L}}$ et φ un isomorphisme du revêtement $(\mathrm{Y}_{\overline{\mathbf{Q}}}, q_{\overline{\mathbf{Q}}})$ déduit de $(\mathrm{Y}_{\mathrm{L}}, q_{\mathrm{L}})$ par extension des scalaires de L à $\overline{\mathbf{Q}}$ sur le revêtement (Y, q).

On dit que L est un *corps de définition* de (Y, q) s'il existe un *modèle* de (Y, q) sur L. Il existe des corps de définition de (Y, q) qui sont de degré fini sur $\mathbf{Q}$.

Soit $(\mathrm{Y_L}, q_\mathrm{L}, \varphi)$ un modèle de (Y, q) sur L. Le revêtement $(\mathrm{Y}_{\overline{\mathbf{Q}}}, q_{\overline{\mathbf{Q}}})$ est canoniquement isomorphe à chacun de ses conjugués par $\mathrm{G_L} = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathrm{L})$, d'où par transport de structure un isomorphisme $u_\sigma : (\mathrm{Y}, q) \to {}^\sigma(\mathrm{Y}, q)$ pour tout $\sigma \in \mathrm{G_L}$. Ces isomorphismes satisfont la relation de cocycle $u_{\sigma\tau} = {}^\sigma u_\tau \circ u_\sigma$ pour $\sigma, \tau \in \mathrm{G_L}$.

Inversement, si des isomorphismes $u_\sigma : (\mathrm{Y}, q) \to {}^\sigma(\mathrm{Y}, q)$, pour $\sigma \in \mathrm{G_L}$, satisfont cette relation de cocycle, ils proviennent par la construction précédente d'un modèle de (Y, q) sur L, unique à isomorphisme unique près.

Remarque. — Soit $(\mathrm{Y_L}, q_\mathrm{L}, \varphi)$ un modèle de (Y, q) sur L. Alors $(\sigma, g) \mapsto u_\sigma^{-1}\, {}^\sigma g u_\sigma$ définit une opération continue de $\mathrm{G_L}$ sur le groupe des automorphismes de (Y, q). L'ensemble des classes d'isomorphisme de modèles de (Y, q) sur L est paramétré par $\mathrm{H}^1(\mathrm{G_L}, \mathrm{Aut}(\mathrm{Y}, q))$. Il est en particulier réduit à un seul élément si le groupe des automorphismes de (Y, q) est réduit à l'élément neutre.

L'ensemble des $\sigma \in \mathrm{G}_\mathbf{Q}$ tels que le revêtement ${}^\sigma(\mathrm{Y}, q)$ soit isomorphe à (Y, q) est un sous-groupe ouvert d'indice fini de $\mathrm{G}_\mathbf{Q}$. Il est donc égal à $\mathrm{G_K}$ pour un corps de nombres K. Le corps K est appelé le *corps des modules* du revêtement (Y, q). Il est contenu dans tout corps de définition de (Y, q), mais n'en est pas forcément un lui-même, sauf si le groupe des automorphismes de (Y, q) est réduit à l'élément neutre : en effet dans ce cas, il existe pour tout $\sigma \in \mathrm{G_K}$ un unique isomorphisme u_σ de (Y, q) sur ${}^\sigma(\mathrm{Y}, q)$, et la condition de cocycle est automatiquement vérifiée.

Remarques. — 1) Le corps des modules est l'intersection des corps de définition (*cf.* [3], ou [5], 3.4).

2) On trouvera dans [5] une description cohomologique des obstructions à ce que le corps des modules soit un corps de définition et diverses conditions suffisantes pour que ces obstructions soient triviales.

3) La notion de corps de définition, de corps des modules d'un revêtement étale de $\mathbf{P}'_{\overline{\mathbf{Q}}}$ ne dépend que de la classe d'isomorphisme de ce revêtement ; elle conserve donc un sens pour les objets (ou classes d'isomorphismes d'objets) de chacune des catégories équivalentes à celle des revêtements étales de $\mathbf{P}'_{\overline{\mathbf{Q}}}$ (par exemple celle des ensembles finis munis de deux permutations).

4) Soient (Y, q) un revêtement étale de $\mathbf{P}'_{\overline{\mathbf{Q}}}$ et $(\mathrm{E}, \sigma_0, \sigma_1)$ le triplet formé d'un ensemble fini muni de deux permutations qui lui est associé (*cf.* 1.7). Pour que le corps des modules de (Y, q) soit *réel*, il faut et il suffit qu'il existe une permutation u de E telle que $u \circ \sigma_0 = {\sigma_0}^{-1} \circ u$ et $u \circ \sigma_1 = {\sigma_1}^{-1} \circ u$. Pour que $\mathbf{R}$ soit un corps de définition de (Y, q), il faut et il suffit qu'il existe une *involution* de E satisfaisant la condition précédente ; les classes d'isomorphisme de modèles de (Y, q) sur $\mathbf{R}$ sont alors paramétrées par les classes de conjugaison de telles involutions dans le groupe des permutations de E.

5) Soit K un corps de nombres. Pour tout revêtement ramifié (X, p) de $\mathbf{P}_{\overline{\mathbf{Q}}}$, étale au-dessus de $\mathbf{P}'_{\overline{\mathbf{Q}}}$ et sans automorphismes non triviaux, de corps des modules K, considérons l'unique (à isomorphisme unique près) modèle $(\mathrm{X_K}, p_\mathrm{K})$ de (X, p) sur K. Une variante du théorème de Belyĭ, due à Couveignes ([4]), affirme que toute courbe projective et lisse sur K est isomorphe à une courbe $\mathrm{X_K}$ obtenue de cette manière. Il en résulte en particulier que lorsque X est connexe de genre 0 (donc isomorphe à $\mathbf{P}_{\overline{\mathbf{Q}}}$), $\mathrm{X_K}$ n'est pas forcément isomorphe à $\mathbf{P_K}$ (mais peut être une conique sur K sans points rationnels).

6) Soient (X, p) un revêtement ramifié de $\mathbf{P}_{\overline{\mathbf{Q}}}$, étale au-dessus de $\mathbf{P}'_{\overline{\mathbf{Q}}}$, a un point de $p^{-1}(0)$, ξ un point de la fibre du vecteur tangent $\overrightarrow{01}$ (*cf.* 1.4, remarques 1 et 2). Une variante des définitions de ce numéro permet d'introduire la notion de corps de définition et de corps des modules du triplet (X, p, a), et du triplet (X, p, ξ). Dans chacun de ces deux cas, le corps des modules est un corps de définition : dans le second cas, cela résulte du fait que (X, p, ξ) n'a pas d'automorphismes non triviaux ; pour le premier cas, *cf.* [2], th. 2.

2.3. Groupoïde fondamental d'un schéma de type fini sur un corps

Nous verrons dans la suite de cet exposé que l'ensemble des informations concernant l'opération de $G_{\mathbf{Q}}$ sur la catégorie des revêtements étales de $\mathbf{P}'_{\overline{\mathbf{Q}}}$ est codé dans le groupoïde fondamental de $\mathbf{P}'_{\mathbf{Q}}$ (et même dans son groupe fondamental en un point rationnel).

Nous allons rappeler dans ce numéro les liens entre le groupoïde fondamental d'un schéma S *géométriquement connexe* de type fini sur un corps k et celui du schéma $\overline{S}$ qui se déduit de S par extension des scalaires à une clôture algébrique $\overline{k}$ de k. Des morphismes de schémas $\overline{S} \to S$ et $S \to \mathrm{Spec}(k)$, on déduit des morphismes de groupoïdes $\varpi_1^{ét}(\overline{S}) \to \varpi_1(S)$ et $\varpi_1^{ét}(S) \to \varpi_1^{ét}(\mathrm{Spec}(k))$. En particulier, si $\overline{a}$ est un point géométrique de $\overline{S}$ (à valeurs dans un corps séparablement clos Ω, qui *ipso facto* est une extension de $\overline{k}$), et si a et c sont les points géométriques de S et de $\mathrm{Spec}(k)$ images de $\overline{a}$, on a une suite de groupes profinis et d'homomorphismes continus

$$1 \longrightarrow \pi_1^{ét}(\overline{S}, \overline{a}) \longrightarrow \pi_1^{ét}(S, a) \longrightarrow \pi_1^{ét}(\mathrm{Spec}(k), c) \longrightarrow 1\,. \tag{1}$$

Cette suite est exacte ([13], IX, th. 6.1) et $\pi_1^{ét}(\mathrm{Spec}(k), c)$ est canoniquement isomorphe au groupe de Galois $\mathrm{Gal}(k_s/k)$, où k_s est la fermeture séparable de k dans $\overline{k}$. Chaque élément de $\mathrm{Gal}(k_s/k)$ définit donc (en faisant opérer un de ses relèvements par conjugaison sur $\pi_1^{ét}(\overline{S}, \overline{a})$) un élément du groupe $\mathrm{Out}(\pi_1^{ét}(\overline{S}, \overline{a}))$ (où, par définition, le groupe $\mathrm{Out}(G)$ des automorphismes extérieurs d'un groupe profini G est le quotient du groupe des automorphismes continus de G par le sous-groupe formé des automorphismes intérieurs).

Si $\overline{b}$ est un second point géométrique de $\overline{S}$ (pas forcément à valeurs dans le même corps séparablement clos), l'ensemble $\pi_1^{ét}(\overline{S}; \overline{a}, \overline{b})$ est non vide et tout élément u de cet ensemble définit un isomorphisme $x \mapsto uxu^{-1}$ de $\pi_1^{ét}(\overline{S}, \overline{a})$ dans $\pi_1^{ét}(\overline{S}, \overline{b})$; celui-ci ne dépend pas de u à automorphisme intérieur près. On en déduit donc un isomorphisme canonique de $\mathrm{Out}(\pi_1^{ét}(\overline{S}, \overline{a}))$ sur $\mathrm{Out}(\pi_1^{ét}(\overline{S}, \overline{b}))$, ce qui permet légitimement de définir $\mathrm{Out}(\pi_1^{ét}(\overline{S}))$ sans référence au point-base.

L'homomorphisme de $\mathrm{Gal}(k_s/k)$ dans $\mathrm{Out}(\pi_1^{ét}(\overline{S}))$ déduit de (1) ne dépend pas non plus du point-base. En effet, soient σ' et σ'' des relèvements d'un élément $\sigma \in \mathrm{Gal}(k_s/k)$ dans $\pi_1^{ét}(S, a)$ et $\pi_1^{ét}(S, b)$. Pour $x \in \pi_1(\overline{S}, \overline{a})$, on a, en identifiant $\pi_1^{ét}(\overline{S}; \overline{a}, \overline{b})$ à une partie de $\pi_1^{ét}(S; a, b)$, $\sigma''(uxu^{-1})\sigma''^{-1} = v(\sigma' x \sigma'^{-1})v^{-1}$, où $v = \sigma'' u \sigma'^{-1}$; il suffit donc de démontrer que $u^{-1}v$ appartient à $\pi_1^{ét}(\overline{S}, \overline{a})$, *i.e.* opère trivialement sur la fibre en c de tout revêtement étale de $\mathrm{Spec}(k)$ de la forme $\mathrm{Spec}(k')$, où k' est une

extension séparable de degré fini de k, ce qui est immédiat. En conclusion, on dispose d'un homomorphisme canonique

$$\mathrm{Gal}(k_s/k) \longrightarrow \mathrm{Out}(\pi_1^{ét}(\overline{\mathrm{S}}))\,.$$

Lorsque le point géométrique $a : \mathrm{Spec}(\Omega) \to \mathrm{S}$ est rationnel sur k, *i.e.* se factorise par un morphisme $\mathrm{Spec}(k) \to \mathrm{S}$, ce morphisme définit une section continue $\pi_1^{ét}(\mathrm{Spec}(k), c) \to \pi_1^{ét}(\mathrm{S}, a)$ de l'extension (1), qui est donc canoniquement scindée. L'image par cette section d'un élément σ de $\mathrm{Gal}(k_s/k) \approx \pi_1(\mathrm{Spec}(k), c)$ sera notée σ_a.

Chaque point $a \in \mathrm{S}(k)$ définit un point géométrique de S et un point géométrique de $\overline{\mathrm{S}}$ à valeurs dans $\overline{k}$, que nous noterons encore a par abus. Pour $a, b \in \mathrm{S}(k)$, on définit une opération de $\mathrm{Gal}(k_s/k)$ sur l'ensemble $\pi_1^{ét}(\overline{\mathrm{S}}; a, b)$ en posant, pour $u \in \pi_1^{ét}(\overline{\mathrm{S}}; a, b)$,

$$^{\sigma}u = \sigma_b u \sigma_a^{-1}\,,$$

le calcul étant effectué dans $\pi_1^{ét}(\mathrm{S}; a, b)$. (On notera que la flèche figurant au second membre appartient bien à $\pi_1^{ét}(\overline{\mathrm{S}}; a, b)$, car $u^{-1}\sigma_b u \sigma_a^{-1}$ appartient à $\pi_1^{ét}(\overline{\mathrm{S}}, a)$ d'après ce que nous avons vu plus haut.) L'opération de $\mathrm{Gal}(k_s/k)$ ainsi définie est compatible avec la composition des flèches.

2.4. L'homomorphisme canonique $\mathrm{G}_{\mathbf{Q}} \to \mathrm{Out}(\widehat{\pi})$

Appliquons les résultats de 2.3 au cas particulier où $k = \mathbf{Q}$, $\overline{k} = \overline{\mathbf{Q}}$ et $\mathrm{S} = \mathbf{P}'_{\mathbf{Q}}$. On obtient un homomorphisme canonique $\mathrm{G}_{\mathbf{Q}} \to \mathrm{Out}(\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}))$, où pour définir $\mathrm{Out}(\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}))$ on peut prendre pour point-base un point géométrique arbitraire. En prenant un point de l'intervalle $]0, 1[$, par exemple $\frac{1}{2}$, on obtient donc un homomorphisme canonique

$$\rho : \mathrm{G}_{\mathbf{Q}} \longrightarrow \mathrm{Out}(\widehat{\pi})$$

où $\widehat{\pi}$ est le complété profini du groupe π considéré dans la remarque 1 du n° 1.3. Notons que $\widehat{\pi}$ est un groupe profini libre à deux générateurs (ceux-ci étant les images des éléments c_0 et c_1 de π).

THÉORÈME 2. — *L'homomorphisme canonique $\rho : \mathrm{G}_{\mathbf{Q}} \to \mathrm{Out}(\widehat{\pi})$ est injectif.*

Cela résulte des deux lemmes suivants :

Lemme 1. — Si σ appartient au noyau de ρ, tout revêtement étale (Y, q) de $\mathbf{P}'_{\overline{\mathbf{Q}}}$ est isomorphe à son conjugué $^{\sigma}(\mathrm{Y}, q)$.

Nous pouvons supposer Y connexe. Soient a un point de $\mathbf{P}'(\mathbf{Q})$ et σ_a le relèvement canonique de σ dans $\pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}}, a)$ (*cf.* 2.3). Si σ appartient au noyau de ρ, σ_a opère par conjugaison dans $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, a)$ suivant un automorphisme intérieur. Le groupe $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, a)$ opère transitivement dans les fibres géométriques $q^{-1}(a)$ et $^{\sigma}q^{-1}(a)$ et $x \mapsto \sigma(x)$ est une bijection de $q^{-1}(a)$ dans $^{\sigma}q^{-1}(a)$, équivariante pour l'automorphisme $g \mapsto \sigma_a g \sigma_a^{-1}$ de $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, a)$. Si donc H est le stabilisateur d'un point $x \in q^{-1}(a)$ dans $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, a)$, celui de $\sigma(x)$ est $\sigma_a \mathrm{H} \sigma_a^{-1}$, qui est conjugué à H dans $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, a)$; cela

implique que les $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, a)$-ensembles $q^{-1}(a)$ et ${}^{\sigma}q^{-1}(a)$ sont isomorphes, et donc que les revêtements (Y, q) et ${}^{\sigma}(\mathrm{Y}, q)$ le sont.

Lemme 2. — Le groupe $\mathrm{G}_{\mathbf{Q}}$ *opère fidèlement dans l'ensemble des classes d'isomorphismes de revêtements étales de* $\mathbf{P}'_{\overline{\mathbf{Q}}}$.

Pour tout $j \in \overline{\mathbf{Q}}$, il existe en effet d'après le théorème de Belyĭ (th. 1) un revêtement étale (Y, q) de $\mathbf{P}'_{\overline{\mathbf{Q}}}$, où Y est un ouvert de Zariski non vide d'une courbe elliptique sur $\overline{\mathbf{Q}}$ d'invariant modulaire j. Le corps des modules de ce revêtement contient $\mathbf{Q}(j)$.

Remarque. — H. W. Lenstra a démontré que $\mathrm{G}_{\mathbf{Q}}$ opère déjà fidèlement dans l'ensemble des classes d'isomorphismes de revêtements (X, p) de $\mathbf{P}_{\overline{\mathbf{Q}}}$, étales au-dessus de $\mathbf{P}'_{\overline{\mathbf{Q}}}$, possédant les propriétés suivantes : X est connexe, de genre 0 et a un seul point au-dessus de ∞ (ce qui équivaut à dire que (X, p) est isomorphe à un couple $(\mathbf{P}_{\overline{\mathbf{Q}}}, f)$, où la fonction de Belyĭ f est un polynôme, ou encore que le 1-squelette de la carte bicoloriée associée est un arbre). Pour la démonstration, *cf.* [12], th. II.4.

2.5. Le groupe fondamental de $\mathbf{P}'_{\mathbf{Q}}$ en $\overrightarrow{01}$

Nous allons maintenant préciser la structure du groupe fondamental de $\mathbf{P}'_{\mathbf{Q}}$, lorsqu'on fait jouer le rôle de point-base au vecteur tangent $\overrightarrow{01} = (0, \frac{d}{dz})$, ou plus généralement aux vecteurs tangents $\overrightarrow{ij}$ qui se déduisent de $\overrightarrow{01}$ par les automorphismes de $\mathbf{P}_{\mathbf{Q}}$ qui stabilisent $\{0, 1, \infty\}$ (*cf.* n° 1.4, remarque 2). La théorie est semblable à celle décrite en 2.3 pour les points-base rationnels sur le corps de base : on a en particulier une suite exacte de groupes profinis, canoniquement scindée,

$$1 \longrightarrow \pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01}) \longrightarrow \pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}}, \overrightarrow{01}) \longrightarrow \mathrm{G}_{\mathbf{Q}} \longrightarrow 1. \tag{2}$$

Le groupe $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$ est canoniquement isomorphe à $\pi_1^{ét}(\mathbf{P}'_{\mathbf{C}}, \overrightarrow{01})$, c'est-à-dire au complété profini $\widehat{\pi}$ du groupe fondamental de $\mathbf{P}'(\mathbf{C})$ où le rôle de point-base est tenu par le germe en 0 de $]0, 1[$, ou encore par $]0, 1[$ (1.3, remarques 1 et 2). C'est donc un groupe profini libre à deux générateurs (ceux-ci, que nous noterons x et y, correspondant par les isomorphismes précédents aux générateurs c_0 et c_1 de π).

Comme la suite exacte (2) est canoniquement scindée, le groupe $\mathrm{G}_{\mathbf{Q}}$ opère sur $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$, et $\pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}}, \overrightarrow{01})$ est canoniquement isomorphe au produit semi-direct de $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$ par $\mathrm{G}_{\mathbf{Q}}$. Toute la richesse de la structure du groupe $\pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}}, \overrightarrow{01})$ est donc codée dans la manière dont le groupe $\mathrm{G}_{\mathbf{Q}}$ opère sur $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$. Cette opération est définie par un homomorphisme

$$\mathrm{G}_{\mathbf{Q}} \longrightarrow \mathrm{Aut}(\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})) \tag{3}$$

qui relève l'homomorphisme $\rho : \mathrm{G}_{\mathbf{Q}} \to \mathrm{Out}(\widehat{\pi})$ considéré en 2.4 (lorsqu'on identifie $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$ à $\widehat{\pi}$), et qui par suite est injectif.

Remarque. — Avant d'étudier cet homomorphisme, nous allons donner une description alternative de cette suite exacte (2) en termes galoisiens. Choisissons une uniformisante locale t de $\mathbf{P}_{\mathbf{Q}}$ en 0 telle que $dt(\overrightarrow{01}) = 1$, par exemple $t = z$. Soit $\overline{\mathbf{Q}}\{\{t\}\}$

le corps des séries de Puiseux en t à coefficients dans $\overline{\mathbf{Q}}$. C'est une extension algébriquement close de $\overline{\mathbf{Q}}(z)$ (on plonge $\overline{\mathbf{Q}}(z)$ dans $\overline{\mathbf{Q}}\{\{t\}\}$ en associant à chaque fraction rationnelle son développement de Laurent en t au point 0). Notons M la plus grande extension algébrique de $\overline{\mathbf{Q}}(z)$ contenue dans $\overline{\mathbf{Q}}\{\{t\}\}$ et non ramifiée en dehors de $\{0,1,\infty\}$. C'est une extension galoisienne de $\mathbf{Q}(z)$. Les groupes profinis $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}},\overrightarrow{01})$ et $\pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}},\overrightarrow{01})$ sont canoniquement isomorphes à $\mathrm{Gal}(\mathrm{M}/\overline{\mathbf{Q}}(z))$ et $\mathrm{Gal}(\mathrm{M}/\mathbf{Q}(z))$, et la suite exacte (2) s'identifie à la suite exacte

$$1 \longrightarrow \mathrm{Gal}(\mathrm{M}/\overline{\mathbf{Q}}(z)) \longrightarrow \mathrm{Gal}(\mathrm{M}/\mathbf{Q}(z)) \longrightarrow \mathrm{Gal}(\overline{\mathbf{Q}}(z)/\mathbf{Q}(z)) \approx \mathrm{G}_{\mathbf{Q}} \longrightarrow 1\,.$$

Le relèvement canonique de $\sigma \in \mathrm{G}_{\mathbf{Q}}$ dans $\mathrm{Gal}(\mathrm{M}/\mathbf{Q}(z))$ opère dans M par $\Sigma a_\lambda t^\lambda \mapsto \Sigma\sigma(a_\lambda)t^\lambda$.

Si (E,p) est un revêtement étale connexe de $\mathbf{P}'_{\mathbf{Q}}$ et K son corps de fonctions rationnelles, il existe une bijection canonique de la fibre en $\overrightarrow{01}$ de ce revêtement (*cf.* 1.4, remarque 1) sur l'ensemble des plongements $\mathbf{Q}(z)$-linéaires de K dans M, par laquelle l'opération de $\pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}},\overrightarrow{01})$ sur la fibre s'identifie à l'opération $(g,u)\mapsto g\circ u$ de $\mathrm{Gal}(\mathrm{M}/\mathbf{Q}(z))$ sur l'ensemble de ces plongements.

Soit σ un élément de $\mathrm{G}_{\mathbf{Q}}$. L'élément σ opère sur les racines n-ièmes de l'unité par élévation à la puissance $\chi_n(\sigma)$-ième, où $\chi_n(\sigma)$ est un entier premier à n bien défini modulo n, *i.e.* un élément de $(\mathbf{Z}/n\mathbf{Z})^\times$. Les $\chi_n(\sigma)$, pour $n\geqslant 1$, définissent un élément $\chi(\sigma)$ de la limite projective des $(\mathbf{Z}/n\mathbf{Z})^\times$, c'est-à-dire du groupe multiplicatif $\widehat{\mathbf{Z}}^\times$ du complété profini $\widehat{\mathbf{Z}}$ de $\mathbf{Z}$. L'application $\chi : \mathrm{G}_{\mathbf{Q}} \to \widehat{\mathbf{Z}}^\times$ est un homomorphisme continu, appelé le *caractère de Teichmüller*.

Le groupe $\mathrm{G}_{\mathbf{Q}}$ opère sur l'ensemble $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}},\overrightarrow{01},\overrightarrow{10})$. La définition de cette opération est analogue à celle donnée en 2.3 dans le cas de points-base rationnels : pour $\sigma\in\mathrm{G}_{\mathbf{Q}}$ et $u\in\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}},\overrightarrow{01},\overrightarrow{10})$, on a

$$^{\sigma}u = \sigma_{\overrightarrow{10}}\, u\, \sigma_{\overrightarrow{01}}^{-1}\,,$$

où $\sigma_{\overrightarrow{01}}$ et $\sigma_{\overrightarrow{10}}$ sont les relèvements canoniques de σ dans $\pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}},\overrightarrow{01})$ et $\pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}},\overrightarrow{10})$ respectivement.

Un analogue de 1.4 montre que le chemin continu $c : t\mapsto t$ de $[0,1]$ dans $\mathbf{P}_1(\mathbf{C})$ définit un élément de $\pi_1^{ét}(\mathbf{P}'_{\mathbf{C}},\overrightarrow{01},\overrightarrow{10})$, donc de $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}},\overrightarrow{01},\overrightarrow{10})$. On notera encore c cet élément. Nous noterons f_σ l'élément $c^{-1}\,{}^{\sigma}c$ de $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}},\overrightarrow{01})$.

PROPOSITION 1. — *L'élément f_σ appartient à l'adhérence $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}},\overrightarrow{01})'$ du sous-groupe dérivé de $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}},\overrightarrow{01})$.*

Il s'agit de démontrer que f_σ opère trivialement sur la fibre en $\overrightarrow{01}$ des revêtements étales abéliens de $\mathbf{P}'_{\overline{\mathbf{Q}}}$, et il suffit pour cela de traiter le cas des revêtements induits par les revêtements ramifiés $\mathbf{P}_{\overline{\mathbf{Q}}}\to\mathbf{P}_{\overline{\mathbf{Q}}}$ de la forme $z\mapsto z^n$ et de la forme $z\mapsto 1-(1-z)^n$, où n est un entier $\geqslant 1$. Traitons par exemple le premier cas, le second étant similaire. Les fibres $\mathrm{F}_{\overrightarrow{01}}$ et $\mathrm{F}_{\overrightarrow{10}}$ en $\overrightarrow{01}$ et $\overrightarrow{10}$ du revêtement considéré se composent respectivement des vecteurs tangents $(0,\zeta\frac{d}{dz})$ et $(\zeta,-\zeta\frac{d}{dz})$, où ζ parcourt l'ensemble des racines n-ièmes de l'unité. L'élément $\sigma_{\overrightarrow{01}}$ de $\pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}},\overrightarrow{01})$ opère dans $\mathrm{F}_{\overrightarrow{01}}$ par $(0,\zeta\frac{d}{dz})\mapsto(0,\zeta^{\chi_n(\sigma)}\frac{d}{dz})$ et l'élément $\sigma_{\overrightarrow{10}}$ de $\pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}},\overrightarrow{10})$ opère dans

$\mathrm{F}_{\overrightarrow{10}}$ par $(\zeta, -\zeta\frac{d}{dz}) \mapsto (\zeta^{\chi_n(\sigma)}, -\zeta^{\chi_n(\sigma)}\frac{d}{dz})$; la bijection de $\mathrm{F}_{\overrightarrow{01}}$ sur $\mathrm{F}_{\overrightarrow{10}}$ définie par c est $(0, \zeta\frac{d}{dz}) \mapsto (\zeta, -\zeta\frac{d}{dz})$. Il en résulte aussitôt que $f_\sigma = c^{-1}\,{}^\sigma c = c^{-1}\sigma_{\overrightarrow{10}} c \sigma_{\overrightarrow{01}}^{-1}$ opère trivialement dans $\mathrm{F}_{\overrightarrow{01}}$.

Les deux éléments $\chi(\sigma) \in \widehat{\mathbf{Z}}^\times$ et $f_\sigma \in \pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})'$ que nous venons d'associer à σ contiennent toute l'information nécessaire à décrire la manière dont σ opère dans $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$. En effet :

Proposition 2. — *L'élément σ opère dans $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$ en appliquant les générateurs canoniques x et y sur $x^{\chi(\sigma)}$ et $f_\sigma^{-1} y^{\chi(\sigma)} f_\sigma$ respectivement.*

Identifions $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$ et $\pi_1^{ét}(\mathbf{P}'_{\mathbf{Q}}, \overrightarrow{01})$ aux groupes de Galois $\mathrm{Gal}(\mathrm{M}/\overline{\mathbf{Q}}(z))$ et $\mathrm{Gal}(\mathrm{M}/\mathbf{Q}(z))$, comme dans la remarque ci-dessus. Alors x et $\sigma_{\overrightarrow{01}}$ s'identifient respectivement aux automorphismes $\sum a_\lambda z^\lambda \mapsto \sum a_\lambda e^{2\pi i\lambda} z^\lambda$ et $\sum a_\lambda z^\lambda \mapsto \sum \sigma(a_\lambda) z^\lambda$ de M. On en déduit aussitôt que l'on a $\sigma_{\overrightarrow{01}} x \sigma_{\overrightarrow{01}}^{-1} = x^{\chi(\sigma)}$.

Lorsqu'on identifie $\widehat{\pi}$ à $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{10})$, les générateurs c_0 et c_1 de π s'identifient à des générateurs topologiques x' et y' de $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{10})$, et l'on a $x = c^{-1}x'c$, $y = c^{-1}y'c$. Du premier alinéa, on déduit par transport de structure par l'automorphisme $z \mapsto 1 - z$ de $\mathbf{P}_{\mathbf{Q}}$ que l'on a ${}^\sigma y' = y'^{\chi(\sigma)}$. Il en résulte que l'on a

$$ {}^\sigma y = ({}^\sigma c)^{-1} y'^{\chi(\sigma)} ({}^\sigma c) = f_\sigma^{-1} c^{-1} y'^{\chi(\sigma)} c f_\sigma = f_\sigma^{-1} y^{\chi(\sigma)} f_\sigma \,, $$

d'où la proposition.

2.6. Le groupe de Grothendieck-Teichmüller

Soit $\widehat{\mathrm{F}}$ un groupe profini libre à deux générateurs x et y. Il sera commode, si u est un homomorphisme continu de $\widehat{\mathrm{F}}$ dans un autre groupe profini et si a, b sont les images de x et y par u, de noter $f(a, b)$ l'image par u d'un élément f de $\widehat{\mathrm{F}}$. Avec ces notations, on a en particulier $f = f(x, y)$.

Posons $\mathrm{M} = \widehat{\mathbf{Z}}^\times \times \widehat{\mathrm{F}}$. Pour $(\lambda, f) \in \mathrm{M}$ et $g \in \widehat{\mathrm{F}}$, posons

$$ {}^{(\lambda,f)}g = g(x^\lambda, f^{-1} y^\lambda f). \tag{4} $$

L'application $g \mapsto {}^{(\lambda,f)}g$ est un endomorphisme continu du groupe profini $\widehat{\mathrm{F}}$. Définissons une loi de composition $\star$ sur M en posant

$$ (\lambda, f) \star (\mu, g) = (\lambda\mu, f.{}^{(\lambda,f)}g) \,. \tag{5} $$

On vérifie tout d'abord que l'on a

$$ {}^{(\lambda,f)\star(\mu,g)}h = {}^{(\lambda,f)}({}^{(\mu,g)}h) \tag{6} $$

pour $(\lambda, f) \in \mathrm{M}$, $(\mu, g) \in \mathrm{M}$ et $h \in \widehat{\mathrm{F}}$, puis que la loi de composition de M est associative. Elle admet pour élément neutre le couple $(1, 1)$. En d'autres termes, M, muni de la loi de composition $\star$, est un monoïde, et (4) définit une opération de ce monoïde dans le groupe $\widehat{\mathrm{F}}$.

Nous avons au numéro précédent associé à chaque élément $\sigma \in \mathbf{G}_{\mathbf{Q}}$ un élément $\chi(\sigma)$ de $\widehat{\mathbf{Z}}^\times$ et un élément f_σ de $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$. Identifions désormais $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$ au

groupe $\widehat{\mathrm{F}}$ (en identifiant les éléments de $\pi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$ notés x et y au numéro précédent aux éléments x, y de F).

PROPOSITION 3. — *L'application $\sigma \mapsto (\chi(\sigma), f_\sigma)$ de $\mathrm{G}_{\mathbf{Q}}$ dans* M *est un homomorphisme. Son image est contenue dans le groupe des éléments inversibles de* M. *De plus, pour tout $\sigma \in \mathrm{G}_{\mathbf{Q}}$ et tout $g \in \widehat{\mathrm{F}}$, on a ${}^{\sigma}g = {}^{(\chi(\sigma), f_\sigma)}g$.*

Il suffit de démontrer la dernière assertion lorsque g est un des deux générateurs topologiques x, y de $\widehat{\mathrm{F}}$, et elle résulte dans ce cas de la prop. 2. Si σ et τ sont deux éléments de $\mathrm{G}_{\mathbf{Q}}$, on a $\chi(\sigma\tau) = \chi(\sigma)\chi(\tau)$ et

$$f_{\sigma\tau} = c^{-1}.\,{}^{\sigma\tau}c = c^{-1}.\,{}^{\sigma}c\,.\,{}^{\sigma}({}^{\tau}c^{-1}) = f_\sigma\,.\,{}^{\sigma}f_\tau = f_\sigma\,.\,{}^{(\chi(\sigma), f_\sigma)}f_\tau$$

d'où la première assertion. La seconde en résulte aussitôt.

Il résulte du th. 2 et de la prop. 2 que l'homomorphisme $\sigma \mapsto (\chi(\sigma), f_\sigma)$ de $\mathrm{G}_{\mathbf{Q}}$ dans M est injectif. Un problème fondamental consiste à essayer d'en déterminer l'image, car cela fournirait une sorte d'écriture des éléments de $\mathbf{G}_{\mathbf{Q}}$ (un peu analogue à l'écriture décimale des nombres réels ou l'écriture des nombres p-adiques comme vecteurs de Witt).

> *Remarque.* — On peut donner de M une interprétation un peu plus intrinsèque : notons B l'ensemble à deux éléments $\{\overrightarrow{01}, \overrightarrow{10}\}$ et $\varpi_1^{\mathrm{B}}(\mathbf{P}'_{\overline{\mathbf{Q}}})$ le sous-groupoïde plein de $\varpi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}})$ ayant B pour ensemble de points. Alors M s'identifie à l'ensemble des endomorphismes u du groupoïde $\varpi_1^{\mathrm{B}}(\mathbf{P}'_{\overline{\mathbf{Q}}})$ qui fixent les deux points, et possèdent la propriété suivante : il existe un élément $\lambda \in \widehat{\mathbf{Z}}^{\times}$ tel que u applique l'élément x de $\pi_1(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$ sur x^{λ} et l'élément y' de $\pi_1(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{10})$ sur y'^{λ}. (Un tel endomorphisme est déterminé dès que l'on se donne λ et l'image de c, que l'on écrit cf ; il applique alors y sur $f^{-1}y^{\lambda}f$, c^{-1} sur $f^{-1}c^{-1}$ et x' sur $cfc^{-1}x'^{\lambda}cf^{-1}c$.)

Nous allons maintenant passer en revue trois conditions, découvertes par Drinfeld, qui permettent de restreindre le monoïde dans lequel l'homomorphisme $\mathrm{G}_{\mathbf{Q}} \to \mathrm{M}$ prend ses valeurs.

CONDITION I. — *Si $(\lambda, f) \in \mathrm{M}$ appartient à l'image de $\mathbf{G}_{\mathbf{Q}}$, on a $f(x,y)f(y,x) = 1$.*

Dans l'interprétation de M donnée dans la remarque, pour qu'un élément (λ, f) de M soit tel que $f(x,y)f(y,x) = 1$, il faut et il suffit que l'automorphisme du groupoïde $\varpi_1^{\mathrm{B}}(\mathbf{P}'_{\overline{\mathbf{Q}}})$ correspondant soit invariant par l'automorphisme de $\varpi_1^{\mathrm{B}}(\mathbf{P}'_{\overline{\mathbf{Q}}})$ déduit de $z \mapsto 1 - z$. Il est dès lors clair que l'ensemble de ces éléments est un sous-monoïde de M qui contient l'image de $\mathrm{G}_{\mathbf{Q}}$.

CONDITION II. — *Si $(\lambda, f) \in \mathrm{M}$ appartient à l'image de $\mathbf{G}_{\mathbf{Q}}$, on a, en posant $z = (xy)^{-1}$ et $m = \frac{1}{2}(\lambda - 1)$, $f(z,x)z^m f(y,z)y^m f(x,y)x^m = 1$.*

Notons C l'ensemble des six éléments $\overrightarrow{ij}$, pour $i \neq j$ dans $\{0, 1, \infty\}$ et $\varpi_1^{\mathrm{C}}(\mathbf{P}'_{\overline{\mathbf{Q}}})$ le sous-groupoïde plein de $\varpi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}})$ ayant C pour ensemble de points. Dans l'interprétation de M donnée dans la remarque, pour qu'un élément (λ, f) de M satisfasse la condition I et la condition II, il faut et il suffit que l'automorphisme du groupoïde

$\varpi_1^{\mathrm{B}}(\mathbf{P}'_{\overline{\mathbf{Q}}})$ correspondant se prolonge en un automorphisme u de $\varpi_1^{\mathrm{C}}(\mathbf{P}'_{\overline{\mathbf{Q}}})$, invariant par l'action du groupe des permutations de $\{0,1,\infty\}$ et tel que, si $d \in \varpi_1^{ét}(\mathbf{P}'_{\overline{\mathbf{Q}}}; \overrightarrow{01}, \overrightarrow{0\infty})$ est la classe d'une « boucle reliant $\overrightarrow{01}$ à $\overrightarrow{0\infty}$ dans le demi-plan supérieur », $d^{-1}u(d)$ appartienne au sous-groupe fermé de $\pi_1(\mathbf{P}'_{\overline{\mathbf{Q}}}, \overrightarrow{01})$ engendré par x (auquel cas on a automatiquement $d^{-1}u(d) = x^m$ avec $m = \frac{1}{2}(\lambda - 1)$). Il est dès lors clair que l'ensemble des éléments de M satisfaisant les conditions I et II est un sous-monoïde de M qui contient l'image de $\mathrm{G}_{\mathbf{Q}}$.

Condition III. — Cette condition, plus technique, fait intervenir les relations entre les espaces de modules $M_{0,4}$ et $M_{0,5}$. Je n'ai pas réussi à en donner une interprétation aussi simple que pour les conditions I et II. Je renvoie donc pour la discussion de cette condition à l'article original de Drinfeld ([7]) et à ceux d'Ihara ([10], [11]).

L'ensemble des éléments inversibles de M qui satisfont les conditions I, II, III, est un groupe appelé le groupe de Grothendieck-Teichmüller et noté $\widehat{\mathrm{GT}}$. Une question naturelle est de savoir si l'image de $\mathrm{G}_{\mathbf{Q}}$ dans M est égale à $\widehat{\mathrm{GT}}$. Récemment, Yves André a développé un analogue local de ce formalisme pour $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ et démontré que ce groupe se réalise comme le groupe des automorphismes d'un foncteur π_1 approprié en géométrie p-adique.

APPENDICE : CARTES

1. Cartes cellulaires

En suivant [14], déf 1.1.1, nous appellerons *carte cellulaire* (de dimension 2) un triplet $\mathrm{G} = (\mathrm{X}, \mathrm{K}, \mathrm{S})$, où X est une surface topologique compacte (sans bord), K une partie fermée de X et S une partie finie de K, telles que : $\mathrm{X} - \mathrm{K}$ a un nombre fini de composantes connexes, chacune homéomorphe à un 2-disque ouvert ; $\mathrm{K} - \mathrm{S}$ a un nombre fini de composantes connexes, chacune homéomorphe à un 1-disque ouvert ; tout point de S est adhérent à $\mathrm{K} - \mathrm{S}$. Lorsque la surface topologique X est munie d'une orientation, on dit que la carte G est *orientée*.

Les points de S, les composantes connexes de $\mathrm{K} - \mathrm{S}$ et les composantes connexes de $\mathrm{X} - \mathrm{K}$ s'appellent respectivement les *sommets*, les *arêtes* et les *faces* de G. Ils définissent une décomposition cellulaire de X. Le bord de chaque face est réunion d'une famille d'arêtes et de sommets. Le bord de chaque arête se compose d'un ou deux sommets. Chaque arête est contenue dans le bord d'une ou de deux faces.

Si A est une arête et x un point de A, la limite projective des ensembles $\pi_0(\mathrm{U} - \mathrm{A})$, où U parcourt l'ensemble des voisinages de x dans X, a deux éléments ; chacun d'eux s'appelle une *rive* (ou une *orientation transverse*) de A en x. Lorsque l'arête A et la surface X sont orientées en x, on peut parler des rives gauche et droite de A en x.

2. Cartes triangulaires tricoloriées

Une *carte triangulaire* est une carte cellulaire (X, K, S) dans laquelle l'adhérence de chaque face, munie de la décomposition cellulaire induite par celle de X, est isomorphe à un simplexe euclidien de dimension 2 muni de la décomposition cellulaire standard. Le bord de chaque arête se compose alors de deux sommets (appelés ses extrémités), et le bord de chaque face comprend trois arêtes et trois sommets.

On appelle *carte triangulaire tricoloriée* un quadruplet (X, K, S, t), où (X, K, S) est une carte triangulaire et t une application qui assigne à chaque sommet un type dans $\{0, 1, \infty\}$, les trois sommets adhérents à une même face étant de types distincts. Si A est une arête d'une telle carte, les extrémités de A sont de types i et j distincts, et l'on dit que A est de type $\{i, j\}$.

Si une carte triangulaire tricoloriée est *orientée*, ses faces se divisent en faces positives et négatives : une face est positive si l'ordre cyclique dans lequel se succèdent les types de ses sommets sur son bord orienté est $0, 1, \infty$.

Soient $G = (X, K, S, t)$ et $G' = (X', K', S', t')$ deux cartes triangulaires orientées tricoloriées. Notons $\mathscr{T}(G, G')$ l'ensemble des applications continues $f : X \to X'$ qui induisent un homéomorphisme de l'adhérence de chaque face de X sur l'adhérence d'une face de X', cet homéomorphisme respectant la décomposition cellulaire, le type des sommets et l'orientation.

La *catégorie isotopique des cartes triangulaires orientées tricoloriées* est la catégorie dont les objets sont les cartes triangulaires orientées tricoloriées, un morphisme de G dans G' étant une classe d'isotopie, c'est-à-dire une composante connexe par arcs, de l'ensemble $\mathscr{T}(G, G')$ muni de la topologie de la convergence compacte.

3. Cartes cellulaires bicoloriées

On appelle *carte cellulaire bicoloriée* un quadruplet (X, K, S, t), où (X, K, S) est une carte cellulaire et b une application qui assigne à chaque sommet un type dans $\{0, 1\}$, chaque arête possédant une extrémité de type 0 et une extrémité de type 1.

Lorsqu'une carte cellulaire bicoloriée est orientée, chaque arête est munie en chacun de ses points d'une rive gauche et d'une rive droite (déterminées par l'orientation de l'arête du sommet de type 0 vers le sommet de type 1 et par l'orientation donnée de X).

Soient $G = (X, K, S, b)$ et $G' = (X', K', S', b')$ deux cartes cellulaires bicoloriées orientées. Notons $\mathscr{B}(G, G')$ l'ensemble des applications continues $f : X \to X'$ telles que $f^{-1}(S') = S$ et $f^{-1}(K') = K$, qui respectent les types des sommets et le caractère gauche ou droit des rives des arêtes.

La *catégorie isotopique des cartes cellulaires bicoloriées orientées* est la catégorie dont les objets sont les cartes cellulaires bicoloriées orientées, un morphisme de G dans G' étant une classe d'isotopie, c'est-à-dire une composante connexe par arcs, de l'ensemble $\mathscr{B}(G, G')$ muni de la topologie de la convergence compacte.

Remarque. — On définit une équivalence de la catégorie isotopique des cartes triangulaires orientées tricoloriées sur la catégorie isotopique des cartes cellulaires bicoloriées orientées par $G = (X, K, S, t) \mapsto (X, K', S', b)$, où S' est l'ensemble des sommets de G de type 0 ou 1, b coïncide avec t dans S' et K' est la réunion de S' et des arêtes de type $\{0, 1\}$ de G.

RÉFÉRENCES

[1] G. Belyĭ – « Galois extensions of a maximal cyclotomic field », *Izv. Akad. Nauk SSSR, Ser. Mat.* **43** (1979), no. 2, p. 267–276, en russe ; traduction anglaise dans : *Math. USSR Izv.* **14** (1979), p. 247-256.

[2] B. Birch – « Noncongruence subgroups, covers and drawings », in *The Grothendieck theory of dessins d'enfants (Luminy, 1993)*, London Math. Soc. Lecture Note Ser., vol. 200, Cambridge Univ. Press, Cambridge, 1994, p. 25–46.

[3] K. Combes & D. Harbater – « Hurwitz families and arithmetic Galois groups », *Duke Math. J.* **52** (1985), p. 821–839.

[4] J.-M. Couveignes – « À propos du théorème de Belyĭ », *J. Théor. Nombres Bordeaux* **8** (1996), no. 1, p. 93–99.

[5] P. Dèbes & J.-C. Douai – « Algebraic covers : field of moduli versus field of definition », *Ann. scient. Éc. Norm. Sup.* 4^{e} *série* **30** (1997), p. 303–338.

[6] P. Deligne – « Le groupe fondamental de la droite projective moins trois points », in *Galois groups over* $\mathbf{Q}$, Publ. MSRI, vol. 16, Springer, 1989, p. 79–298.

[7] V. G. Drinfeld – « On quasi-triangular quasi-Hopf algebras and some group closely associated with $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ », *Leningrad Math. J.* **2** (1991), p. 829–860.

[8] A. Grothendieck – « Technique de descente et théorèmes d'existence en géométrie algébrique, I. Généralités. Descente par morphismes fidèlement plats », in *Séminaire Bourbaki 1959/1960*, Benjamin, 1966, exp. n^{o} 190 (réédité par la S.M.F en 1995).

[9] ———, « Esquisse d'un programme », in *Geometric Galois Actions*, London Math. Soc. Lecture Note Ser., vol. 242, Cambridge Univ. Press, Cambridge, 1997, p. 5–48.

[10] Y. Ihara – « Braids, Galois groups, and some arithmetic functions », in *Proceedings of the ICM (Kyoto, 1990)*, p. 99–120.

[11] ———, « On the embedding of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ into $\widehat{\mathrm{GT}}$ », in *The Grothendieck theory of dessins d'enfants (Luminy, 1993)*, London Math. Soc. Lecture Note Ser., vol. 200, Cambridge Univ. Press, Cambridge, 1994, p. 289–305.

[12] L. Schneps – « Dessins d'enfants », in *The Grothendieck theory of dessins d'enfants (Luminy, 1993)*, London Math. Soc. Lecture Note Ser., vol. 200, Cambridge Univ. Press, Cambridge, 1994, p. 47–77.

[13] *Revêtements étales et groupe fondamental, Séminaire de Géométrie Algébrique du Bois-Marie 1960/1961 (SGA 1)* – Lect. Notes in Math., vol. 224, Springer, 1970, dirigé par A. Grothendieck.

[14] C. Voisin & J. Malgoire – *Cartes cellulaires*, Cahiers mathématiques, vol. 12, Université de Montpellier, 1977.

Joseph OESTERLÉ

Institut de mathématiques de Jussieu
175, rue du Chevaleret
F-75013 Paris
E-mail : oesterle@math.jussieu.fr

Séminaire BOURBAKI Juin 2002
54^{e} année, 2001-2002, n^{o} 908, p. 307 à 317

KAM TECHNIQUES IN PDE

by **Ricardo PÉREZ-MARCO**

We present a partial account of recent application of KAM techniques in the context of PDEs. We don't address several other topics in Hamiltonian PDE as for instance those related to invariant Gibbs measures or Nekhoroshev bounds of diffusion for solutions of non-linear PDEs. We adopt a Dynamical Systems point of view (this just reflects the background and motivation of the author).

I am grateful to J. Bourgain, W. Craig, J.-C. Guillot, H. Eliasson and S. Kuksin for helpful discussions.

1. INTRODUCTION

1.1. Ancient historical motivation

The quasi-periodic motions of the planets in our Solar System was observed long, very long time ago, back when Astronomy, Physics and Mathematics were a single science. There is ample evidence from cuneiform clay tablets of Babylonian observations that go as far as 2000 before the Christian era (see the compilations by O. Neugebauer [NEU1] and also [NEU2]). One can also find traces of some accurate Babylonian observations in Ptolomy's Almagest [PTO]. The Babylonians were using the quasi-periodic evolution to forecast future positions of the planets. From a rich data of observations, it is fairly simple to notice the periodic character of the positions, plus an error which is itself periodic, which in turn has a much smaller periodic error, and so on...

It is revealing that this simple observation remains still unproved! This is certainly the oldest open problem in Mathematics.

1.2. Old and new K.A.M. theory

First came Newton's theory of gravitation. The next progress in the above historical problem was only achieved in the XXth century with the emergence of K.A.M. theory. A.N. Kolmogorov [KOL] discovered the persistence of invariant tori in Hamiltonian systems near completely integrable ones. This was one of the major achievements of Dynamical Systems in the XXth Century. K.A.M. theory, named after its founders Kolmogorov, Arnold and Moser, was developed during the late fifties and sixties. We refer to the comprehensive Bourbaki seminar by J.-B. Bost [BOS] for a survey and bibliography of this classical topic. Later, V.K. Melnikov [MEL] announced the persistence not only of mid-dimensional tori, but also of low dimensional tori. The first proofs appeared only in the late eighties by H. Eliasson [ELI1], J. Pöschel [POS1] and S.B. Kuksin [KUK1]. These results opened the door to the application of KAM techniques to infinite dimensional Hamiltonian systems (the approach of Kuksin is indeed infinite dimensional). These occur naturally in Hamiltonian PDE, some of which appear as perturbations of completely integrable ones. The introduction of KAM techniques in this field showed the existence of quasi-periodic solutions for non-linear and non-integrable PDE.

Unfortunately (or fortunately) there hasn't been any Bourbaki seminar in this topic, so there are indeed several good surveys available. The first book of S. Kuksin [KUK1] and, more recently, the one of W. Craig [CRA] and the second book of Kuksin [KUK3] cover largely the developments of the theory. Other surveys where one can find useful material are [BOU2], [KUK2], [POS2]. For this reason, after a preliminary introduction in the first sections, we concentrate on the more recent results and techniques developed by J. Bourgain, some of which are not yet published [BOU5]. Starting from a technique devised by W. Craig and E. Wayne to find periodic solutions, Bourgain pushed it to get quasi-periodic solutions and indeed a whole new approach to K.A.M. theory. It presents a certain number of advantages, for example, it handles the Schrödinger equation in higher dimension where the difficulties related to arithmetic approximation has stopped any progress for a long time. Recent simplifications using local uniformizations for semi-algebraic sets, and central multiscale arguments (inspired from the theory of the discrete Schrödinger equation) to control the inverse of high dimensional matrices with critical sites, solve these problems and yield a unified approach to classical results, for example those on low dimensional tori. We will illustrate this new approach in this simpler, finite dimensional setting. Applications to PDEs do not differ substantially.

We start with a brief survey on persistence of low dimensional tori in finite dimension, followed by some examples of classical PDE where the techniques have been applied. Next we describe Bourgain's new techniques, following [BOU5], and how they are used to prove the existence of low dimensional tori.

2. LOW DIMENSIONAL TORI

2.1. Introduction to Hamiltonian systems and K.A.M. theory

We consider a Hamiltonian system

$$\dot{q}_k = +\frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}$$

defined in $\mathbf{T}^n \times \mathbf{U} \subset \mathbf{T}^n \times \mathbf{R}^n \subset \mathbf{R}^{2n}$ where $U \subset \mathbf{R}^n$ is a bounded domain. In what follows all Hamiltonian systems are supposed to be real analytic (but most of the results persist with lower regularity).

When the Hamiltonian $H = H_0$ is independent of the angular variables $\mathbf{q}$, the system is completely integrable. In that case, the solutions of the system are periodic or quasi-periodic solutions

$$\mathbf{q}(t) = t\ \lambda + \mathbf{q}(0), \quad \lambda = \Big(\frac{\partial H_0}{\partial p_1}, \ldots, \frac{\partial H}{\partial p_n}\Big)$$

fill densely tori contained in the fibers $\{\mathbf{p} = \mathbf{p}(0)\}$. When the frequency vector λ is purely irrational, that is

$$\dim_{\mathbf{Q}}(1, \lambda) = n + 1,$$

then the invariant tori have maximal dimension n.

Many natural systems appear in nature as a small perturbation of completely integrable ones. A fundamental example of completely integrable system in Celestial Mechanics is the two-body problem. The three-body planar with a third small mass, and its different versions (planar, restricted) appear as a perturbation of the completely integrable system. Kolmogorov [KOL] announced in 1954 the persistence of many of these maximal tori for a perturbed system $H = H_0 + \varepsilon H_1$ where ε is small (H_1 may also depend on ε). More precisely, under the "twist" condition

$$\det\left[\frac{\partial^2 H_0}{\partial p_k \partial p_l}\right] \neq 0,$$

any torus with a frequency λ with good arithmetic will persist under a sufficiently small (depending on λ) perturbation. The condition on the arithmetic of the frequency ensures that $\langle \mathbf{k}, \lambda \rangle$ is not two small (modulo 1) depending on the size of $\mathbf{k} \in \mathbf{Z}^n$ and is necessary in this type of Small Divisors problem. The small divisors appear when one attempts to write down and solve the equation of the invariant tori (Moser's approach, see for example [S-M], section 32) for the simpler situation of invariant curves of the annulus) or write down approximate first integrals (Arnold's approach [ARN]).

We point out for later reference that the above Hamiltonian system can be written in a convenient form using complexified variables

$$u_k = p_k + iq_k, \quad v_k = \overline{u}_k = p_k - iq_k\,;$$

considering the real analytic Hamiltonian $H = H(\mathbf{u}, \overline{\mathbf{u}})$ the above equations read

$$i\dot{u}_k = 2\ \frac{\partial H}{\partial \overline{u}_k}.$$

2.2. Low dimensional tori

When the frequency λ is not purely irrational, the solution of the completely integrable system fills a low dimensional torus. These in general are unstable due to its normal degenerate character (nevertheless see [CHE] and [ELI2]). But one may consider the general problem of persistence of low dimensional tori with non-degenerate normal part. In general, mixed hyperbolic and elliptic character may be present in the normal direction (according to when the spectrum of the Hamiltonian vector field in the normal direction has non purely imaginary ($\notin i\mathbf{R}$) or purely imaginary ($\in i\mathbf{R}$) eigenvalues respectively). For a complete account in this situation, and state of the art theorems with minimal twist condition, we refer to the extensive article of H. Rüsmann [RUS]. The main new feature in the persistence of lower dimensional tori is that, in absence of external parameters, the frequencies of the dynamics on the persistent tori cannot be prescribed. At the opposite, these frequencies are employed as parameters to locate the persistent tori.

When the normal behavior is normally hyperbolic, the problem is simpler. The first result of this type is A.M. Liapounov's center theorem [LIA] of preservation of periodic solutions (or tori of dimension 1). For the higher dimensional results see [GRA], [MOS] and [ZEH].

We consider in what follows the purely elliptic situation with tangential frequencies $(\lambda_1, \ldots, \lambda_n)$ (that are used as parameters) and normal frequencies $(\mu_1, \ldots, \mu_m)$, for the linearization of the unperturbed system. Melnikov announced [MEL] during the sixties, and at the end of the eighties Eliasson [ELI1], Kuksin [KUK1], and Pöschel proved the persistence of these tori under the non-resonance conditions,

$$\langle \mathbf{k}, \lambda \rangle - \mu_j \neq 0$$

for all $k \in \mathbf{Z}^n$ and $1 \leqslant j \leqslant m$, and

$$\langle \mathbf{k}, \lambda \rangle + \mu_{j_1} - \mu_{j_2} \neq 0$$

for all $k \in \mathbf{Z}^n$ and $1 \leqslant j_1, j_2 \leqslant m$, $j_1 \neq j_2$. The second condition plays a role in the reduction to the preliminary normal form. J. Bourgain got rid of it using the Lyapunov-Schmidt approach we describe below that does not require this preliminary normal form (at the expense of having to control the inverses of non-diagonal linear operators). Since then J. You [YOU] has obtained an improvement of the original KAM approach by using a different normal form that also yields non-diagonal linearized operators.

Once understood the conditions that are required for the persistence of lower dimensional tori, the techniques were ready to be used for infinite dimensional Hamiltonian systems. In that way Kuksin established the first results for hamiltonian PDE's. His method is exposed in his book [KUK1].

3. SOME HAMILTONIAN PDEs

The methods presented below apply to a wide array of Hamiltonian PDEs. Sometimes part of the difficulty consists in finding a suitable Birkhoff normal form, in which the equation appears as a perturbation of a completely integrable PDE.

– A remarkable and extensively studied completely integrable PDE is Korteweg-de Vries (KDV) equation,

$$\partial_t u = -\partial_{xxx} u + 6u\partial_x u.$$

We refer to the upcoming book of T. Kappeler and J. Pöschel [K-P] for K.A.M. on KDV.

– The non-linear Schrödinger equation (NLS)

$$i\partial_t u - \partial_{xx} u + g(x)u + \varepsilon \partial_{\overline{u}} H(u, \overline{u}) = 0$$

which can be considered in the context of periodic or Dirichlet boundary conditions.

– The non-linear wave equation

$$\partial_{tt} u = \partial_{xx} u + g(x)u + \varepsilon h(x, u)$$

which can also be considered with periodic or Dirichlet boundary conditions.

We refer to chapter 2 of [CRA] for more examples and precisions, and more information on the Hamiltonian character and Birkhoff normal forms of these equations near an equilibrium position.

4. PERSISTENCE OF LOW DIMENSIONAL TORI

We use this problem to illustrate the techniques in [BOU5]. We follow section XVII of [BOU5]. This approach to Melnikov's theorem, without the simplifications of [BOU5], appeared first in [BOU1]. The application to the construction of quasi-periodic solutions of the non-linear Schrödinger equation and non-linear wave equation *in arbitrary dimension* can be found in sections XVIII and XIX of [BOU5] and are based on the same ideas.

4.1. Setup of the problem

The problem of persistence of low dimensional tori can be reduced to the perturbation of a linear Hamiltonian system (see [BOU2] chapter 6) by a standard procedure of writing the system in the appropriate Birkhoff normal form.

Complexifying coordinates we are led to find invariant tori for the perturbed system, $1 \leqslant j \leqslant N$ (real dimension $2N$),

$$\frac{1}{i}\dot{z}_j = \frac{\partial H}{\partial \overline{z}_j}$$

where

$$H(z,\overline{z}) = \sum_{j=1}^{N} \lambda_j |z_j|^2 + \varepsilon H_1(z,\overline{z})$$

We assume for simplicity that H_1 is independent of ε, but this is irrelevant. It is also assumed that the non-linear perturbation is polynomial. This plays a role on the argument with semi-algebraic sets, but the proof should go through in the general analytic case.

We consider $\lambda = (\lambda_1, \ldots, \lambda_n)$ with $1 \leqslant n < N$ and the solution of the unperturbed system,

$$z_j(t) = a_j e^{i\lambda_j t}$$

with $a_j = 0$ for $n < j \leqslant N$. When λ is purely irrational, this solution fills densely an n-dimensional torus. Note also that in this case we can assume the amplitudes (a_j) to be real and positive (shift t). For a substantial (when $\varepsilon \to 0$) set of values of λ, this torus survives the perturbation, but we may have to change slightly the frequency of the dynamics on it. We express this using the Fourier expansion of the solution. The theorem to prove is the following:

THEOREM. — *Let $\Omega \subset \mathbf{R}^n$. Given $\varepsilon_0 > 0$ sufficiently small, for each ε, $0 < |\varepsilon| < \varepsilon_0$, there is a compact set $\Omega_\varepsilon \subset \Omega$ with $|\Omega - \Omega_\varepsilon| \to 0$ when $\varepsilon \to 0$, and a smooth map $\Lambda_\varepsilon : \Omega \to \mathbf{R}^n$, $\lambda \to \Lambda_\varepsilon(\lambda) = \lambda'$, such that for each $\lambda \in \Omega_\varepsilon$ there is a quasi-periodic solution*

$$z(t) = \sum_{k \in \mathbf{Z}^n} \widehat{z}(t) e^{i\langle k, \lambda' \rangle t}$$

such that (we denote $(e_1, \ldots, e_n)$ the canonical base of $\mathbf{Z}^n$, and $c > 0$)

$$\begin{aligned} \widehat{z}_j(e_j) &= a_j, \qquad \forall 1 \leqslant j \leqslant n, \\ |\widehat{z}_j(k)| &\leqslant e^{-c|k|}, \quad \forall 1 \leqslant j \leqslant n, \ \forall k \in \mathbf{Z}^n \\ \sum_{(j,k) \notin \mathcal{R}} |\widehat{z}_j(k)| &\leqslant \sqrt{\varepsilon}, \end{aligned}$$

where $\mathcal{R} = \{(j, e_j); j = 1, \ldots, n\}$ is the resonant set.

4.2. Lyapunov-Schmidt decomposition

The idea of the proof is not to try to localize the invariant torus, following Moser, nor to try to seek the domains where approximate first integrals exist, following Arnold. The new idea is to catch the solution directly using its Fourier expansion. The invariant torus is then obtained as the closure of this quasi-periodic solution. In some sense, this torus is constructed from the inside.

This is a very natural idea and it is not surprising that it yields a powerful approach.

How to find the perturbed solution? Again, the natural idea is to expand in Fourier series the equation

$$\frac{1}{i}\dot{z}_j - \lambda_j z_j + \varepsilon \frac{\partial H_1}{\partial \overline{z}_j} = 0,$$

then identify the Fourier coefficients of both sides. This gives for $k \in \mathbf{Z}^n$ and $1 \leqslant j \leqslant N$,

$$(\langle k, \lambda' \rangle - \lambda_j)\widehat{z}_j(k) + \varepsilon \widehat{\frac{\partial H_1}{\partial \overline{q}_j}}(k) = 0.$$

We use the first n frequencies as a parameter $\lambda' \in \mathbf{R}^n$.

If we request $\widehat{z}_j(e_j) = a_j$ as specified by the theorem for $1 \leqslant j \leqslant n$, we get the Q-equations

$$\lambda'_j - \lambda_j + \frac{\varepsilon}{a_j}\widehat{\frac{\partial H_1}{\partial \overline{q}_j}}(e_j) = 0.$$

The other equations, for the non-resonant indexes, are called the P-equations. The division of the equations into P and Q equations constitutes the Lyapounov-Schmidt decomposition. W. Craig and E. Wayne [C-W] used this method to find periodic solutions (1-dimensional tori) for PDE's. This already involves dealing with Small Divisors. The method is as follows. One first solves the P-equations, which consists in inverting a non-singular infinite dimensional linear operator to get an approximate solution $z(\lambda, \lambda')$. Then, one plugs this solution into the Q-equations to determine λ' in function of λ using the implicit function theorem.

4.3. Main ideas in the proof

As expected, things are not as simple as described. First, solving the P-equations with proper estimates is not an easy task. The P-equations correspond to non-resonant indexes and do have a formal solution, but small divisors appear in the inversion of the linear operator, so one must control the arithmetic of λ' to get the desired bounds. These restrictions on the arithmetic of λ' make that after hard work the good bounds are only obtained on a closed set Ω_ε with empty interior. To use the implicit function theorem in such a set demands careful justification. The procedure, well known in K.A.M., consists in finding successive approximations $z^{(l)}$, regular everywhere, of the P-equations. We can then solve the Q-equation at each step to get and approximate values of λ'.

In order to find the successive approximations $z^{(l)}$ one uses a rapidly convergent Newton scheme. At each step the P-equations are truncated by considering only those k's such that $||k||_\infty \leqslant N_l$, where the sequence of scales (N_l) increases geometrically. Thus, at each step we are faced with the inversion of a finite dimensional (of growing dimension) linear operator. More precisely, in order to get good estimates for $z^{(l+1)}$ one needs to control the inverse of the operator

$$T_l = D_l + \varepsilon S_l\,,$$

where D_l is a diagonal operator with eigenvalues $\pm\langle k, \lambda'\rangle - \lambda_j$ (some, the critical sites, are small and give small divisors), and S_l is a self-adjoint operator with exponentially decaying off-diagonal entries. More precisely, $z^{(l+1)}$ is defined by

$$z^{(l+1)} = z^{(l)} - T_l^{-1}(F(z^{(l)}))\,,$$

where $F(z^{(l)})$ is the data on the P-equation when $z^{(l)}$ is plugged in.

Two main technical simplifications are present in [BOU5] with respect to the original approach [BOU1]. First, a multiscale analysis is performed on the complexification of the matrix T_l in order to obtain good bounds for its inverse. More precisely, the diagonal part is modified by shifting the eigenvalues of D_l into $\pm(\langle k, \lambda'\rangle + \sigma) - \lambda_j$ with a complex parameter σ. Using the analyticity on σ, plus some harmonic analysis estimates, and a multiscale argument incorporated in the induction, it is proved that the inverse has the proper bounds except for a small set of parameters σ. This constitutes the core of the estimate, and the common theme of [BOU5] which focus on the estimation of Lyapounov exponents. This type of multiscale analysis is well known in the theory of the discrete Schrödinger equation. The formal analogy between the two theories is brought a step closer with these common techniques.

The second step consists in showing that in the (λ, λ') plane not much is discarded in order to avoid these bad eigenvalues. One can write (using the Q-equation)

$$\lambda'_l = \lambda + \varepsilon\varphi_l(\lambda).$$

We want to avoid to have $\langle k, \varepsilon\varphi_l(\lambda)\rangle$ in the bad set of σ's, $\mathcal{S}_l$. One can say that in this step we adjust the arithmetic of λ' to avoid the small divisors. In previous works the arithmetic related to the resonant set was a major obstacle. For example, in [BOU4] it prevented the extension of the results to higher dimension, and it required arithmetic lemmas on the grouping and well separation of the critical sites. In [BOU5] these problems are treated with a more conceptual argument. With the previous approach, the main point is to notice that the set to be avoided by (λ, λ') is a semi-algebraic set for which we have an explicit control on the degree. This implies (in numerous senses) that the geometry is controlled. Since the condition becomes

$$\langle k, \varepsilon\varphi_l(\lambda)\rangle \notin \mathcal{S}_l,$$

with k large, the removed measure is small.

These ideas are very general and applicable to all sorts of PDEs, in particular to the non-linear Schrödinger equation and the wave equation in arbitrary dimension as treated in [BOU5].

4.4. Inversion of analytic matrices

We give the following illustrative example ([BOU5], Proposition 13.1; see also [B-G-S]), on how lower scale bad sites cannot affect the inverse for too many complex parameters, or, alternatively, how to use the space of holomorphy around to improve

the bound on the bad set of parameters. We denote for $S \subset \mathbf{N}$, R_S the "restriction to indexes in S" linear operator.

THEOREM. — *Let $A(\sigma) \in M_d(\mathbf{C})$ be a real analytic matrix function for $\sigma \in [-\delta, \delta]$, holomorphic in*

$$\{|\Re\sigma| < \delta\} \cap \{|\Im\sigma| < \gamma\},$$

and such that

$$||A(\sigma)|| \leqslant B_1.$$

For each $\sigma \in [-\delta, \delta]$, there is a subset $\Lambda \subset [1, d]$ such that

$$|\Lambda| \leqslant M$$

and

$$|| \left(R_{[1,d]-\Lambda}\ A(\sigma)\ R_{[1,d]-\Lambda}\right)^{-1} || \leqslant B_2$$

We assume also that

$$|\{\sigma \in [-\delta, \delta]; ||A(\sigma)^{-1}|| \geqslant B_3\}| \leqslant 10^{-3}\gamma(1 + B_1)^{-1}(1 + B_2)^{-1}$$

Then for $\kappa < (1 + B_1 + B_2)^{-10M}$, we have

$$|\{\sigma \in [-\delta/2, \delta/2]; ||A(\sigma)^{-1}|| \geqslant \kappa^{-1}\}| \leqslant \exp\left(-\frac{c \log \kappa^{-1}}{M \log(M + B_1 + B_2 + B_3)}\right).$$

Combining this result with the exponential decrease of the off-diagonal terms, one can prove the following type of result that fits into the induction described in the previous section.

THEOREM 4.1 ([BOU5], 13.31). — *Let A be an $N \times N$ with exponential off-diagonal decay,*

$$|A(n, n')| \leqslant e^{-c_0|n-n'|}.$$

We consider the sub-scale $\overline{N} = N^\tau$, $0 < \tau < 1$.

We assume that for all intervals $J \subset [1, N]$, $|J| \geqslant \overline{N}$,

$$||(R_J A R_J)^{-1}|| \leqslant e^{L^b},$$

with $0 < b < 1$, $\tau + b < 1$.

An $\overline{N}$-interval J is said to be good if moreover

$$|A_J^{-1}(n, n')| \leqslant e^{-c|n-n'|}$$

for $n, n' \in J$ with $|n - n'| \geqslant \overline{N}/10$ (with $0 < c < c_0/10$).

Assume that there are at most N^b disjoint bad $\overline{N}$-intervals.

Then for $|n - n'| \geqslant N/10$,

$$|A^{-1}(n, n')| \leqslant e^{-c'|n-n'|}$$

where $c' = c - N^{-\kappa}$, $\kappa = \kappa(\tau, b) > 0$.

REFERENCES

[ARN] V.I. Arnold – Proof of a theorem of A.N. Kolmogorov on the conservation of quasi-periodic motions under a small change of the Hamiltonian function, *Uspekhi Mat. Nauk.* **18** (1963), no. 5, p. 13–40, Russ. Math. Surv. **18** (1963), no. 5, p. 9-36.

[BOS] J.-B. Bost – Tores invariants des systèmes dynamiques hamiltoniens (d'après Kolmogorov, Arnold, Moser, Rüssmann, Zehnder, Herman, Pöschel, ...), in *Sém. Bourbaki 1984-85*, Astérisque, vol. 133-134, Soc. Math. France, Paris, 1985, exp. n° 639, p. 113–157.

[BOU1] J. Bourgain – Construction of quasi-periodic solutions of Hamiltonian perturbations of linear equations and applications to nonlinear PDE, *International Math. Res. Notices* **11** (1994), p. 475–497.

[BOU2] ———, Non-linear Schrödinger equations, Park City Lectures, July 1995.

[BOU3] ———, On Melnikov's persistence problem, *Mathematical Research Letters* **4** (1997), p. 445–458.

[BOU4] ———, Quasi-periodic solutions of Hamiltonian perturbations of 2D-linear Schrödinger equations, *Ann. Math.* **148** (1998), p. 363–439.

[BOU5] ———, Green's function estimates for lattice Schrödinger operators and applications, Manuscript, 2001.

[B-G-S] J. Bourgain, M. Goldstein & W. Schlag – Anderson localization for Schrödinger operators on $\mathbf{Z}^2$ with quasi-periodic potential, Preprint, 2000.

[CHE] C.-Q. Chen – Lower dimensional invariant tori in the region of instability for nearly integrable Hamiltonian systems, *Comm. Math. Phys.* **203** (1999), no. 2, p. 385–419.

[CRA] W. Craig – *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, Panoramas et Synthèses, vol. 9, Soc. Math. France, Paris, 2000.

[C-W] W. Craig & C.E. Wayne – Newton's method and periodic solutions of non-linear wave equations, *Comm. Pure Appl. Math.* **46** (1993), p. 1409–1498.

[ELI1] L.H. Eliasson – Perturbations of stable invariant tori for Hamiltonian systems, *Ann. Sco. Norm. Sup. Pisa, Sci. Fis. Mat., Ser. IV* **XV** (1988), no. 1, p. 115–147.

[ELI2] ———, Biasymptotic solutions of perturbed integrable Hamiltonian systems, *Bol. Soc. Mat. Bras.* **25** (1994), no. 1, p. 57–76.

[GRA] S.M. Graff – On the continuation of hyperbolic invariant tori for Hamiltonian systems, *J. Diff. Eq.* **15** (1974), p. 1–69.

[K-P] T. Kappeler & J. Pöschel – KDV and KAM, Book to appear in Springer-Verlag.

[KOL] A.N. Kolmogorov – On the conservation of conditionally periodic motions for a small change in Hamilton's function, *Dokl. Acad. Nauk. SSSR* **98** (1954), p. 525–530.

[KUK1] S.V. Kuksin – *Nearly integrable infinite-dimensional systems*, LNM, vol. 1556, Springer-Verlag, 1991.

[KUK2] ———, Elements of a qualitative theory of Hamiltonian PDEs, in *Proceedings ICM, Vol. II, Berlin (1998)*, Doc. Math., Deutsche Math. Verein., 1998, p. 819–829.

[KUK3] ———, *Analysis of Hamiltonian PDEs*, Oxford Lecture Series, vol. 19, Oxford University Press, 2000.

[LIA] A.M. Liapounov – *Problème général de la stabilité du mouvement*, Ann. Math. Studies, vol. 17, Princeton Univ. Press, Princeton, 1947.

[MEL] V.K. Melnikov – *Dokl. Akad. Nauk. SSSR* **165** (1965), no. 6, p. 1245–1248, *Dokl. Akad. Nauk. SSSR* **181** (1968) no. 3, p. 546-549.

[MOS] J. Moser – *Stable and random motions in dynamical systems*, Ann. Math. Studies, vol. 77, Princeton Univ. Press, Princeton, 1973.

[NEU1] O. Neugebauer – *Astronomical cuneiform texts* I, II, III, Springer-Verlag, 1983.

[NEU2] ———, *The exact sciences in antiquity*, Dover Publications, 1969.

[POS1] J. Pöschel – On elliptic lower dimensional tori in Hamiltonian systems, *Math. Z.* **202** (1989), p. 559–608.

[POS2] ———, Non-linear partial differential equations, Birkhoff normal forms, and KAM theory, in *European Congress of Mathematics, Vol. II (Budapest 1996)*, Progr. Math., vol. 169, Birkhäuser, Basel, 1998.

[PTO] Ptolomy – Almagest, ...

[RUS] H. Russmann – Invariant tori in non-degenerate nearly integrable hamiltonian systems, *Regular and Chaotic Dynamics* **6** (2001), no. 2, p. 119–204.

[S-M] C.L. Siegel & J. Moser – *Lectures on Celestial Mechanics*, Springer-Verlag, 1971.

[YOU] J. You – Perturbations of lower-dimensional tori for Hamiltonian systems, *J. Diff. Equ.* **152** (1999), no. 1, p. 1–29.

[ZEH] E. Zehnder – Generalized implicit function theorem with applications to some small divisors problems, I and II, *Comm. Pure Appl. Math.* **28** (1975), p. 91–140; (1976), p. 49–111.

Ricardo PÉREZ-MARCO

UCLA, Department of Mathematics
405, Hilgard Avenue
Los Angeles, CA 90024
USA
E-mail : ricardo@math.ucla.edu

ASTÉRISQUE

2003

290. SÉMINAIRE BOURBAKI, volume 2001/2002, exposés 894-908
289. P. SCOTT, G.A. SWARUP – *Regular neighbourhoods and canonical decompositions for groups*
288. M. REES – *Views of Parameter Space : Topographer and Resident*
287. Geometric Methods in Dynamics (II), volume in honor of Jacob Palis, W. DE MELO, M. VIANA, J.-C. YOCCOZ, editors
286. Geometric Methods in Dynamics (I), volume in honor of Jacob Palis, W. DE MELO, M. VIANA, J.-C. YOCCOZ, editors
285. PO HU – *Duality for Smooth Families in Equivariant Stable Homotopy Theory*
284. Autour de l'analyse microlocale, volume en l'honneur de Jean-Michel Bony, G. LEBEAU, éditeur

2002

283. L. ROBBIANO, C. ZUILY – *Analytic theory for the quadratic scattering wave front set and application to the Schrödinger equation*
282. SÉMINAIRE BOURBAKI, volume 2000/2001, exposés 880-893
281. T. DUQUESNE, J.-F. LE GALL – *Random Trees, Lévy Processes and Spatial Branching Processes*
280. A. MOKRANE, P. POLO, J. TILOUINE – *Cohomology of Siegel varieties*
279. Cohomologies p-adiques et applications arithmétiques (II), P. BERTHELOT, J.-M. FONTAINE, L. ILLUSIE, K. KATO, M. RAPOPORT, éditeurs
278. Cohomologies p-adiques et applications arithmétiques (I), P. BERTHELOT, J.-M. FONTAINE, L. ILLUSIE, K. KATO, M. RAPOPORT, éditeurs
277. B. RÉMY – *Groupes de Kac-Moody déployés et presque déployés*
276. SÉMINAIRE BOURBAKI, volume 1999/2000, exposés 865-879

2001

275. J.-M. BISMUT, S. GOETTE – *Families torsion and Morse functions*
274. A. BONNET, G. DAVID – *Cracktip is a global Mumford-Shah minimizer*
273. K. NISHIYAMA, H. OCHIAI, K. TANIGUCHI, H. YAMASHITA, S. KATO – *Nilpotent orbits, associated cycles and Whittaker models for highest weight representations*
272. M. BOILEAU, J. PORTI – *Geometrization of 3-orbifolds of cyclic type (avec la collaboration de M. HEUSENER)*
271. M. KASHIWARA, P. SCHAPIRA – *Ind-Sheaves*
270. M. BONK, J. HEINONEN, P. KOSKELA – *Uniformizing Gromov hyperbolic spaces*
269. J.-L. WALDSPURGER – *Intégrales orbitales nilpotentes et endoscopie pour les groupes classiques non ramifiés*

2000

268. J. FRANCHETEAU, G. MÉTIVIER – *Existence de chocs faibles pour des systèmes quasi-linéaires hyperboliques multidimensionnels*
267. R. CERF – *Large Deviations for three Dimensional Supercritical Percolation*
266. SÉMINAIRE BOURBAKI, volume 1998/1999, exposés 850-864
265. O. BIQUARD – *Métriques d'Einstein asymptotiquement symétriques*
264. L. LIPSHITZ, Z. ROBINSON – *Rings of Separated Power Series and Quasi-Affinoid Geometry*
263. C. SABBAH – *Équations différentielles à points singuliers irréguliers et phénomène de Stokes en dimension 2*
262. A. VASY – *Propagation of singularities in three-body scattering*
261. Géométrie complexe et systèmes dynamiques, colloque en l'honneur d'Adrien Douady, Orsay 1995, M. FLEXOR, P. SENTENAC et J.-C. YOCCOZ, éditeurs

1999

260. S. D. CUTKOSKY – *Local monomialization and factorization of morphisms*
259. R. KRIKORIAN – *Réductibilité des systèmes produits-croisés à valeurs dans des groupes compacts*
258. Structure Theory of Set Addition, J.-M. DESHOUILLERS, B. LANDREAU and A.A. YUDIN, editors
257. J.-P. LABESSE – *Cohomologie, stabilisation et changement de base (avec la collaboration de L. BREEN et L. CLOZEL)*
256. F. MOREL – *Théorie homotopique des schémas*
255. R.E. KOTTWITZ, D. SHELSTAD – *Foundations of twisted endoscopy*
254. C.J. BUSHNELL, G. HENNIART – *Local tame lifting for* $\mathrm{GL}(n)$ *II : wildly ramified supercuspidals*
253. B. MAGNERON – *Involutions complexes et vecteurs sphériques associés pour les groupes de Lie nilpotents réels*

1998

252. SÉMINAIRE BOURBAKI, volume 1997/1998, exposés 835-849
251. Nombre et répartition de points de hauteur bornée, E. PEYRE, éditeur
250. C. BONATTI, R. LANGEVIN – *Difféomorphismes de Smale des surfaces (avec la collaboration de E. JEANDENANS)*
249. P. AUSCHER, P. TCHAMITCHIAN – *Square root problem for divergence operators and related topics*
248. P. COLMEZ – *Intégration sur les variétés* p*-adiques*
247. G. PISIER – *Non-commutative vector valued* L_p*-spaces and completely* p*-summing maps*

1997

246. V. TARASOV, A. VARCHENKO – *Geometry of* q*-hypergeometric functions, quantum affine algebras and elliptic quantum groups*
245. SÉMINAIRE BOURBAKI, volume 1996/1997, exposés 820-834
244. J.-M. BISMUT – *Holomorphic families of immersions and higher analytic torsion forms*
243. L. LAFFORGUE – *Chtoucas de Drinfeld et conjecture de Ramanujan-Petersson*
242. N. BURQ – *Pôles de diffusion engendrés par un coin*
241. SÉMINAIRE BOURBAKI, volume 1995/1996, exposés 805-819

1996

240. A. SÀ BARETTO, R. B. MELROSE, M. ZWORSKI – *Semilinear diffraction of conormal waves*
239. J.-L. VERDIER – *Des catégories dérivées des catégories abéliennes*
238. A. BROISE, F. DAL'BO, M. PEIGNÉ – *Méthodes des opérateurs de transfert : transformations dilatantes de l'intervalle et dénombrement de géodésiques fermées*
237. SÉMINAIRE BOURBAKI, volume 1994/1995, exposés 790-804
236. Hommage à P. A. MEYER et J. NEVEU
235. J.-P. OTAL – *Le théorème d'hyperbolisation pour les variétés fibrées de dimension 3*
234. A. GENESTIER – *Espaces symétriques de Drinfeld*

1995

233. I. KRIZ, J.P. MAY – *Operads, algebras modules and motives*
232. Recent advances in operator algebras (Orléans, 1992)
231. J.-C. YOCCOZ – *Petits diviseurs en dimension 1*
230. J.-Y. CHEMIN – *Fluides parfaits incompressibles*
229. B. PERRIN-RIOU – *Fonctions* L p*-adiques des représentations* p*-adiques*
228. Columbia University number theory seminar (New-York, 1992)
227. SÉMINAIRE BOURBAKI, volume 1993/1994, exposés 775-789