Introduction à la théorie des modèles

• É. Bouscaren

Dans ce texte, notre but est tout d'abord de présenter, en nous plaçant dans un contexte particulier qui nous permet de simplifier beaucoup les définitions, les principales notions de base de la théorie des modèles, ses outils et ses résultats les plus classiques. Cela sera fait dans la première section. Dans cette section, les trois théorèmes que nous allons donner comme exemples d'applications, de même que les théorèmes de théorie des modèles que l'on utilise pour les démontrer, datent des années 60, et utilisent uniquement des outils très basiques de théorie des modèles.

Dans la seconde section, nous changerons un peu de style afin de présenter « en survol » les développements modernes et plus sophistiqués de la théorie des modèles.

1. Notions de base à travers quelques exemples classiques

1.1 – Exemples d'applications

Nous commençons par énoncer trois exemples d'applications des méthodes basiques de théorie des modèles.

Les démonstrations des Théorèmes 1 et 2 sont à peu près identiques et très simples. Le Théorème 3 est plus compliqué. Il a d'ailleurs fallu attendre plusieurs années pour qu'en soit donnée une démonstration purement géométrique (la preuve « théorie des modèles » date de 1965, il faudra attendre 2011 pour que Jan Denef ([9]) présente une preuve purement géométrique.

Nous verrons un peu plus loin les stratégies et les ingrédients des démonstrations, qui nous serviront de motivation et de fil conducteur pour les définitions de théorie des modèles que nous donnerons. Une fois les outils de théorie des modèles introduits dans la section 1.2, nous pourrons donner (section 1.3) les démonstrations complètes des Théorèmes 1 et 2. En revanche, pour ce qui est du Théorème 3 nous nous contenterons de quelques indications dans cette section et dans l'appendice (section 3).

Les trois applications

Théorème 1. (Ax 1969) *Soit f une application polynomiale de $\mathbb{C}^n$ dans $\mathbb{C}^n$ ($n \geqslant 1$). Si f est injective, f est surjective.*

(L'application f est polynomiale :

$$f(x_1,\dots,x_n) = (f_1(x_1,\dots,x_n),\dots f_n(x_1,\dots,x_n))$$

avec, pour chaque i, $f_i \in \mathbb{C}[X_1,\dots,X_n]$.)

Le théorème, énoncé ainsi, est seulement un cas particulier du Théorème d'Ax, qui considère les endomorphismes des variétés algébriques (et en fait des schémas de type fini ([1])). Le cas particulier de l'espace affine tout entier, c'est-à-dire $\mathbb{C}^n$ avait déjà été remarqué dans [5]. À peu près à la même époque il y a eu plusieurs démonstrations dont par exemple Borel ([6]) utilisant la cohomologie, Grothendieck, puis plus tard Rudin ([18]).

Après avoir vu la démonstration via la théorie des modèles, dans le cas de $\mathbb{C}^n$, on doit pouvoir refaire sans aucune difficulté le cas où f est une application polynomiale de V dans V, avec $V \subset \mathbb{C}^n$, une variété affine,

$$V = \{a \in \mathbb{C}^n : P_1(a) = P_2(a) = \dots = P_m(a) = 0\}$$

pour $P_1,\dots,P_m \in \mathbb{C}[X_1,\dots,X_n]$.

Remarquons que l'hypothèse que f est polynomiale est essentielle, il y a par exemple des contre-exemples de fonctions holomorphes [7].

On peut exactement de la même manière que pour le théorème d'Ax, démontrer à très peu de frais le résultat classique suivant [4] :

Théorème 2. *Soit G un groupe fini d'ordre p^m, pour p premier, et γ une action algébrique de G sur $\mathbb{C}^n$. Alors l'action γ a un point fixe.*

On fixe un groupe G d'ordre p^m. On a donc une action γ,

$$\gamma : G \times \mathbb{C}^n \mapsto \mathbb{C}^n$$

et puisque γ est algébrique, pour chaque $g \in G$, il

existe $f_{g_1},\dots,f_{g_n} \in \mathbb{C}[X_1,\dots,X_n]$, tels que pour tout $(x_1,\dots,x_n) \in \mathbb{C}^n$,

$$\gamma(g,(x_1,\dots,x_n)) = g.(x_1,\dots,x_n) = (f_{g_1}(x_1,\dots,x_n),\dots,f_{g_n}(x_1,\dots,x_n)).$$

On dira que l'action est de degré inférieur ou égal à d si chacun des polynômes f_{g_i} est de degré inférieur ou égal à d.

Pour ces deux théorèmes, on procède par transfert (Théorème 4), à partir de cas où la propriété est très facile à démontrer (en l'occurrence les corps finis).

Enfin, plus compliqué,

Conjecture d'Artin pour les corps p-adiques. *Pour tout p, pour tout $d \geqslant 1$, tout polynôme homogène de degré d en $n > d^2$ variables, à coefficients dans $\mathbb{Q}_p$, a un zéro non trivial dans $\mathbb{Q}_p$.*

Ce résultat avait été montré (Carlitz, Lang) pour tout p, pour les corps des séries formelles à coefficients dans $\mathbb{F}_p$, $\mathbb{F}_p((t))$. En 1965, Ax et Kochen dans [3] ont montré une version « asymptotique » de cette conjecture. On a d'ailleurs maintenant des contre-exemples qui montrent que seule la version asymptotique est vraie ([22]).

Théorème 3 (Conjecture d'Artin asymptotique). *Soit un entier $d \geqslant 1$. Il existe un nombre premier $p_0(d)$ tel que pour tout $p \geqslant p_0(d)$, tout polynôme sur $\mathbb{Q}_p$, homogène de degré d en $n > d^2$ variables, a un zéro non trivial dans $\mathbb{Q}_p$.*

Ce théorème se montre aussi par transfert (Théorème 5), à partir du résultat connu pour les corps $\mathbb{F}_p((t))$.

Faisons tout de suite une remarque qui permet de simplifier les énoncés que l'on veut transférer de $\mathbb{F}_p((t))$ à $\mathbb{Q}_p$. Pour tout corps K, s'il est vrai que, pour tout $d \geqslant 1$, tout polynôme homogène en d^2+1 variables a un zéro non trivial, alors tout polynôme en plus de d^2+1 variables a aussi un zéro non trivial : il suffit de prendre $n-(d^2+1)$ variables égales à zéro.

Comment montrer les Théorèmes 1 et 2

Il faut évidemment quand même utiliser un petit peu d'algèbre.

Quelques rappels sur les corps algébriquement clos.

1. Un corps K est *algébriquement clos* si tout polynôme de $K[X]$ de degré $\geqslant 1$ a une racine dans K.

2. Soit k un corps, il existe une extension de k qui est algébriquement close et algébrique sur k, qu'on appellera une *clôture algébrique* de k et qu'on notera $\bar{k}$. Deux clôtures algébriques de k sont isomorphes au-dessus de k.

3. Cas de la caractéristique p. Fixons une clôture algébrique du corps fini $\mathbb{F}_p$, $\overline{\mathbb{F}_p}$. Dans cette clôture algébrique, $\mathbb{F}_p$ a une seule extension de degré n, $\mathbb{F}_p \subset \mathbb{F}_{p^n}$. De plus, $\mathbb{F}_{p^n} \subset \mathbb{F}_{p^m}$ si et seulement si n divise m. On obtient donc $\overline{\mathbb{F}_p}$ comme limite inductive de la famille $(\mathbb{F}_{p^n})_{n\geqslant 1}$. En particulier, *tout sous-corps finiment engendré de $\overline{\mathbb{F}_p}$ est fini.*

4. Cas de la caractéristique 0. On utilisera ici seulement un seul fait bien connu, tout corps algébriquement clos de caractéristique zéro et de cardinalité identique à celle de $\mathbb{C}$ (c'est-à-dire de cardinalité le continu, la cardinalité de $\mathcal{P}(\mathbb{N})$) est isomorphe à $\mathbb{C}$. On s'en servira un peu plus tard.

On va traiter ensemble les preuves des deux théorèmes car elles sont très semblables. On commence par regarder les cas évidents.

- Les corps finis. Soit K un corps fini ($K = \mathbb{F}_{p^r}$, pour p premier et $r \geqslant 1$).
  1. Alors, évidemment, si g est *n'importe quelle* application injective de K^n dans K^n, g est surjective, simplement car il s'agit d'une application injective d'un ensemble fini dans lui-même.
  2. C'est un simple calcul sur le nombre d'orbites de vérifier que si un groupe G de cardinalité p^m agit sur un ensemble fini de cardinalité q^r, avec q et p premiers distincts, alors il doit y avoir un point fixe.
- Passage aux clôtures algébriques des corps finis.
  Cela repose sur le fait que tout sous-corps finiment engendré de $K_p := \overline{\mathbb{F}_p}$ est fini.
  1. Si f est polynomiale injective de K_p^n dans K_p^n, soit a un élément quelconque de K_p^n. On considère le sous-corps K_0 engendré par a et par le nombre fini de coefficients des polynômes de f. Alors K_0 est fini et la restriction de f, de K_0 dans K_0, est injective, donc surjective, et a est donc bien dans l'image de f.
  2. Si G, groupe d'ordre p^m, agit sur K_q^n, pour $p \neq q$ premiers, soit K_0 le sous-corps engendré par les coefficients des polynômes qui définissent l'action algébrique de G sur K_q^n. Alors l'action de G restreinte aux éléments de K_0 définit une action de G sur K_0, qui est un corps fini, et a donc un point fixe.

Après ces remarques d'algèbre élémentaires, on *invoque la théorie des modèles pour conclure directement* car

Théorème 4. *Si une propriété qui s'exprime par un* énoncé du premier ordre dans le langage des anneaux *est vraie dans la clôture algébrique de $\mathbb{F}_p$ pour tout p premier sauf un nombre fini, elle est vraie dans $\mathbb{C}$, le corps des complexes.*

Dans ce texte, nous allons essayer de donner un sens précis à tout cela,

- définir ce qu'est un *énoncé du premier ordre* et vérifier que les propriétés énoncées dans les deux théorèmes sont bien « du premier ordre » (section 1.2),
- démontrer ce théorème de transfert, en utilisant *les ultraproduits* (section 1.3).

Ce théorème de transfert est en fait un cas particulier du théorème fondamental de la logique du premier ordre, *le théorème de compacité*, que l'on montrera après, en utilisant également les ultraproduits, dans la section 1.3.

Il existe d'autres théorèmes de transferts classiques que nous n'aurons pas le temps de voir ici, en particulier « La complétude de la théorie des corps algébriquement clos de caractéristique fixée » : tout énoncé du premier ordre (définition dans la section 1.2) qui est vrai dans un corps algébriquement clos est vrai dans tout corps algébriquement clos de même caractéristique. Ici nous avons choisi une preuve du théorème d'Ax qui n'utilise pas directement ce second principe de transfert.

La propriété de complétude de la théorie des corps algébriquement clos de caractéristique fixée est souvent utilisée de manière un peu informelle en géométrie algébrique sous le nom de *Principe de Lefschetz*, par exemple si on veut montrer « quelque chose » sur un corps algébriquement clos de caractéristique zéro, on sait qu'il suffit de le faire pour $\mathbb{C}$.

De même, pour le transfert de propriétés vraies en caractéristique p pour (presque) tout p à la caractéristique 0. Mais souvent cela est fait au cas par cas, de manière *ad hoc*. Toutefois Serre dans un article assez récent traite de ces questions de manière plus systématique [19].

Comment montrer le Théorème 3

Le Théorème 3 repose aussi sur un théorème de transfert général dont nous dirons ici juste quelques mots et un tout petit peu plus dans l'appendice (section 3).

Théorème 5. *Une propriété qui s'exprime par un énoncé du premier ordre dans le langage des corps (valués) est vraie dans $\mathbb{Q}_p$ pour tout p sauf un nombre fini si et seulement si elle est vraie dans $\mathbb{F}_p((t))$ pour tout p sauf un nombre fini.*

Or, on sait que le résultat sur les polynômes homogènes est vrai pour tous les corps $\mathbb{F}_p((t))$, donc, une fois vérifié que l'existence de zéros pour les polynômes homogènes s'exprime par des énoncés du premier ordre, le théorème précédent nous donne le Théorème 3.

Ce second théorème de transfert repose sur un autre théorème, théorème emblématique de la théorie des modèles des corps valués, le théorème d'Ax-Kochen-Eršov, montré indépendamment en 1965 par Ax et Kochen d'une part, par Eršov d'autre part ([3], [10]).

Nous en donnons l'énoncé ici et renvoyons à l'appendice (section 3) pour quelques définitions de base concernant les corps valués.

Théorème 6 (Ax-Kochen-Eršov). *Deux corps valués henséliens d'équicaractéristique* 0 *satisfont aux mêmes énoncés du premier ordre dans le langage des corps valués si et seulement si leurs corps résiduels satisfont aux mêmes énoncés dans le langage des corps et leurs groupes de valeur satisfont aux mêmes énoncés dans le langage des groupes ordonnés.*

Le Théorème d'Ax-Kochen-Eršov, a depuis été généralisé, soit en relaxant les hypothèses sur les corps valués (équicaractéristique zéro) soit en élargissant le type d'énoncés que l'on peut transférer [1].

1.2 – Théorie des modèles

La théorie des modèles est l'étude des structures abstraites, dont les structures mathématiques classiques sont des cas particuliers : groupes, anneaux, corps, corps ordonnés, corps différentiels...

Une première approche intuitive de la notion de structure

Nous commençons par une première approche qui a l'avantage de mettre en lumière les objets

1. Par exemple, dans [8] Cluckers et Loeser montrent que l'on peut aussi transférer des égalités entre intégrales p-adiques.

qui nous intéressent vraiment, les ensembles définissables. Mais qui a l'inconvénient, comme on le verra juste après, qu'on ne peut pas vraiment travailler avec, et il faudra donc introduire d'autres définitions.

Définition 1 (Première définition abstraite d'une structure, ensembles basiques). Une *structure* $\mathcal{M} = \langle M, (B_i)_{i\in I}\rangle$ est la donnée d'un ensemble (non vide) M et d'une famille $(B_i)_{i\in I}$ de sous-ensembles de $\bigcup_{n\geqslant 1} M^n$ (la réunion des puissances cartésiennes de M), famille contenant la diagonale de M^2.

Les ensembles B_i sont appelés les sous-ensembles *basiques* ou *atomiques* de la structure $\mathcal{M}$.

Les sous-ensembles basiques sont ceux qui sont obtenus à partir des fonctions et relations de base de la structure à laquelle on s'intéresse. Par exemple, si on prend les réels et qu'on s'intéresse à $(\mathbb{R},+)$ en tant que groupe, on prendra pour ensembles basiques les sous-ensembles de la forme $\{(a_1,\dots,a_n)\in\mathbb{R}^n : m_1a_1+\dots+m_na_n = 0 \text{ pour } m_i\in\mathbb{Z}\}$. Si on s'intéresse à $\mathbb{R}$ en tant que corps, donc à $(\mathbb{R},+,.,0,1)$, ou à n'importe quel corps K, on prendra pour ensembles basiques les solutions d'une équation polynomiale, c'est-à-dire les ensembles de la forme $\{(a_1,\dots,a_n)\in K^n; P(a_1,\dots,a_n) = 0\}$, pour $P\in K[X_1,\dots X_n]$. Pour $\mathbb{R}$, corps ordonné, on prendra les ensembles de la forme suivante, pour $P\in\mathbb{R}[X_1,\dots,X_n]$,

$$\{a\in\mathbb{R}^n; P(a) = 0\} \text{ et } \{a\in\mathbb{R}^n; P(a) > 0\}.$$

Ensuite la famille des *sous-ensembles définissables* de la structure sera obtenue à partir de la famille des sous-ensembles basiques en composant un nombre fini de fois les opérations auxquelles on peut s'attendre. Plus précisément,

Définition 2 (La famille des ensembles définissables). Si $\mathcal{M} = \langle M, (B_i)_{i\in I}\rangle$ est une structure, on définit la famille des *sous-ensembles définissables* de la structure $\mathcal{M}$, notée $Def(\mathcal{M})$ comme la plus petite famille de sous-ensembles de $\bigcup_{n\geqslant 1} M^n$ avec les propriétés de clôture suivantes.

1. Pour tout $i\in I$, $B_i\in Def(\mathcal{M})$.
2. (Opérations booléennes finies) Si $A,B\in Def(\mathcal{M})$, alors $A\cup B, A\cap B, M^n\setminus A$ appartiennent à $Def(\mathcal{M})$.
3. (Permutation de coordonnées) Si $A\subset M^n$, $A\in Def(\mathcal{M})$ et σ est une permutation de $\{1,\dots,N\}$, alors $\sigma(A) = \{(a_{\sigma(1)},\dots,a_{\sigma(n)},;(a_1,\dots,a_n)\in A\} \in Def(\mathcal{M})$.
4. (Projection) Si $A\subset M^{n+m}$, $A\in Def(\mathcal{M})$, et π_n est la projection de A sur M^n, alors $\pi_n(A)\in Def(\mathcal{M})$.
5. (Spécialisation) Si $A\subset M^{n+m}$, $A\in Def(\mathcal{M})$, et si $\bar{c}\in M^n$, alors $A(\bar{c}) := \{\bar{b}; (\bar{c},\bar{b})\in A\}\in Def(\mathcal{M})$.

Reprenons les exemples donnés plus haut. Pour un corps K, à partir des ensembles basiques, quand on clôt par intersection finie, on obtient les fermés de Zariski (ou ensembles algébriques de K^n). Ensuite si on clôt par combinaisons booléennes finies (complémentaire, intersection, réunion) on obtient exactement les ensembles constructibles de la topologie de Zariski. Pour obtenir les ensembles définissables il faut également clore par projection. Si le corps K est algébriquement clos alors il est bien connu (c'est un théorème de Chevalley, ou, du point de vue de la théorie des modèles, c'est le théorème d'élimination des quantificateurs de Tarski) que la projection d'un ensemble constructible est encore constructible, et donc les ensembles définissables du corps K sont exactement les ensembles constructibles. En revanche dès que K est infini mais non algébriquement clos, cela n'est plus vrai (théorème de Macintyre en 1971 [16]) ; par exemple, en caractéristique p, si K n'est pas parfait, c'est-à-dire s'il existe $b\in K$ sans racine $p^{\text{ième}}$, alors $K^p := \{a\in K; \exists b\in K, b^p = a\}$ est la projection d'un constructible mais n'est pas constructible.

Maintenant pour la structure $\mathbb{R}$ comme corps ordonné, à partir des ensembles basiques, par combinaisons booléennes, on obtient les ensembles *semi-algébriques*. Là aussi il est bien connu (Théorème de Tarski-Seidenberg) que la projection d'un semi-algébrique est semi-algébrique. Cela est vrai car $\mathbb{R}$ est un corps réel clos (tout polynôme de degré impair a une solution et -1 n'est pas un carré). Pour une exposition de la théorie des modèles des corps réels clos et plus généralement de la géométrie dite « modérée », on pourra regarder l'article de Thomas Scanlon, *o-minimality*, de ce même numéro.

Comme déjà dit plus haut, cette définition des structures, directement à partir d'une famille de sous-ensembles « basiques », est assez naturelle mais n'est pas très utile en général. En effet ce cadre ne permet pas de traiter des questions essentielles, comme celles de l'extension des ensembles définissables. Prenons un exemple simple. On considère les deux extensions de $\mathbb{Q}$, $\mathbb{Q}\subset\tilde{\mathbb{Q}}\subset\mathbb{C}$, où $\tilde{\mathbb{Q}}$ est la clôture algébrique de $\mathbb{Q}$. On définit les ensembles basiques de la même manière dans ces trois corps, ce sont les zéros des polynômes. Il est naturel de vouloir étudier les relations entre $Def(\mathbb{Q})$, $Def(\tilde{\mathbb{Q}})$ et $Def(\mathbb{C})$. Y-a-t-il une manière canonique d'étendre un ensemble définissable de $\mathbb{Q}$ à un ensemble définissable de $\tilde{\mathbb{Q}}$ ou de $\mathbb{C}$? Réciproquement, si X est un sous-ensemble définissable dans $\mathbb{C}^n$, qu'en est-il de $X\cap\tilde{\mathbb{Q}}^n$ ou $X\cap\mathbb{Q}^n$? Tout corps dénombrable de

caractéristique 0 se plonge dans $\mathbb{C}$, existe-t-il toujours de telles structures « universelles » dans toute classe de structures ? Peut-on définir proprement ce qu'on entend par des structures de *type similaire* (la classe des corps par exemple) ? Comment énoncer et démontrer rigoureusement les théorèmes de transfert utilisés plus haut ?

Pour pouvoir répondre à ce type de questions, il nous faut une manière plus constructive de décrire les ensembles définissables, qui ne dépende pas d'une structure précise donnée ($\mathbb{R}$ par exemple) mais uniquement du « type » de la structure, les corps dans ce cas, ou bien les corps ordonnés... C'est pourquoi on introduit les « langages », et les *formules*. Les ensembles définissables deviennent alors les ensembles des « solutions » d'une formule dans une structure donnée, de la même manière que l'on considère, en géométrie, les points K-rationnels d'une variété, quand on fait varier le corps K.

Comme expliqué dans l'introduction, nous n'allons pas ici donner les définitions dans le cadre général qui est celui de la logique du premier ordre, ni démontrer le théorème de compacité en général. Pour éviter le formalisme et les complications techniques, nous allons nous placer dès le départ dans le cas particulier qui nous intéresse, c'est-à-dire définir ce que nous appellerons une formule du premier ordre pour la classe des anneaux commutatifs. Il s'agit là de définitions et d'un cadre « ad hoc » qui permettent d'admettre sans problèmes toute une série de détails techniques qu'il nous faudrait considérer sérieusement autrement.

Définition de l'ensemble des formules pour les anneaux commutatifs

Nous allons nous placer tout de suite dans la classe de tous les *anneaux commutatifs unitaires*. Nous avons donc des structures R munies de l'addition $+ : R \times R \mapsto R$, de l'inverse de l'addition, $- : R \mapsto R$, de la multiplication $. : R \times R \mapsto R$ et de deux éléments distingués, 0 et 1.

Dans ce cas, le « langage » de la structure R, ce sont deux fonctions binaires, une fonction unaire et deux constantes.

On va définir, pas à pas, à la fois l'ensemble des formules pour le langage des anneaux commutatifs et pour une formule, ce que cela signifie qu'elle est « vraie » dans un anneau R donné.

Soyons francs, ces définitions peuvent sembler naïves ou même « tautologiques » à première vue et surtout dans le cadre particulier que nous avons choisi. Mais c'est un peu normal, puisqu'il s'agit de définir de manière plus formelle des notions que l'on utilise tous les jours en mathématiques. De plus nous passons sous silence ici le versant *syntactique* de la logique du premier ordre : le fait que, quand on formalise la notion de démonstration, on obtient bien qu'une formule est vraie si et seulement si il en existe une démonstration formelle [2], ce qui fait la force de ces définitions.

On définit tout d'abord l'ensemble des formules *basiques* ou *atomiques*, $\mathcal{F}_0$: ce sont toutes les expressions de la forme $P(x_1,\dots,x_n) = 0$ où P est un polynôme dans $\mathbb{Z}[X_1,\dots,X_n]$.
Les polynômes peuvent être de degré zéro : par exemple « $p = 0$ » est une formule, pour chaque nombre premier p.

On dira que la variable x_i est *libre* dans la formule $P(x_1,\dots,x_n) = 0$ si x_i est de degré supérieur ou égal à 1 dans le polynôme P.

Pour une formule ϕ de $\mathcal{F}_0$, on écrira $\phi(x_1,\dots,x_n)$ pour indiquer, quand cela sera important, que les variables libres dans ϕ sont parmi $\{x_1,\dots,x_n\}$.

Si R est un anneau commutatif, si $\phi(x_1,\dots,x_n)$ est une formule de la forme « $P(x_1,\dots,x_n) = 0$ », et si $a_1,\dots a_n$ sont des éléments de R, on dira que « R satisfait à la formule $\phi(a_1,\dots,a_n)$ », ou que la suite $a_1,\dots,a_n$ satisfait à la formule $\phi(x_1,\dots,x_n)$ dans R, ou que ϕ est vraie pour $(a_1,\dots,a_n)$, noté

$$R \models \phi(a_1,\dots,a_n),$$

si, dans R, on a $P(a_1,\dots,a_n) = 0$.

Pour les formules sans variables libres, elles sont vraies ou fausses dans un anneau donné. Par exemple la formule « $p = 0$ » est vraie dans les anneaux de caractéristique égale à p.

On dira que le sous-ensemble de R^k, $\{(a_1,\dots,a_k) \in R^k : R \models \phi(a_1,\dots,a_k)\}$ est « défini » ou *définissable* par la formule ϕ. Grâce aux formules basiques on récupère ainsi pour chaque anneau R, et pour chaque n, la famille d'ensembles basiques que nous avions choisie plus haut.
Il nous faut maintenant, pour obtenir l'ensemble de toutes les formules, clore l'ensemble des formules basiques par les opérations suivantes :

– la *négation* (correspondant au complémentaire). Si $\phi(x_1,\dots,x_n)$ est une formule, $\neg\phi(x_1,\dots,x_n)$, la négation de ϕ, est une formule (lue « non » ϕ), qui définit le complémentaire de l'ensemble défini par

2. Théorème de complétude de Gödel

ϕ : sans surprise, si $a_1,\dots a_n$ sont des éléments de R, on dira que la formule $\neg\phi(a_1,\dots,a_n)$ est vraie dans R ($R \models \phi(a_1,\dots,a_n)$), si la formule $\phi(a_1,\dots,a_n)$ n'est pas vraie dans R ($R \not\models \phi(a_1,\dots,a_n)$). Les variables libres de $\neg\phi$ sont les mêmes que celles de ϕ ;

– la *conjonction* (correspondant à l'intersection). Si $\phi_1(x_1,\dots,x_n)$ et $\phi_2(x_1,\dots x_n)$ sont des formules alors $\phi_1(x_1,\dots,x_n) \wedge \phi_2(x_1,\dots x_n)$ (lu ϕ_1 « et » ϕ_2 et noté aussi $\phi_1 \wedge \phi_2(x_1,\dots,x_n)$) est une formule, qui définit l'intersection des ensembles définis par ϕ_1 et ϕ_2 : si $a_1,\dots a_n$ sont des éléments de R, la formule $\phi_1 \wedge \phi_2(a_1,\dots,a_n)$ est vraie dans R ($R \models \phi_1 \wedge \phi_2(a_1,\dots,a_n)$) si $R \models \phi_1(a_1,\dots,a_n)$ *et* $R \models \phi_2(a_1,\dots,a_n)$. Une variable x est libre dans $\phi_1 \wedge \phi_2$ si elle est libre dans ϕ_1 ou libre dans ϕ_2 ;

– la *quantification existentielle* (correspondant à la projection). Si $\phi(x_1,\dots,x_n)$ est une formule, alors $\exists x_1 \phi(x_1,\dots,x_n)$ est une formule (lue « il existe » x_1 ϕ), qui définit la projection de l'ensemble défini par ϕ : si $(b_2,\dots b_n)$ sont des éléments de R, alors la formule $\exists x_1 \phi(x_1,b_2,\dots,b_n)$ est vraie dans R ($R \models \exists x_1 \phi(x_1,b_2,\dots,b_n)$) s'il existe $a \in R$, tel que $R \models \phi(a_1,b_2,\dots,b_n)$. La variable x est libre dans $\exists y \phi(y,\dots,x)$, si x est libre dans $\phi(y,\dots,x)$, mais x n'est plus une variable libre de $\exists x \phi(x,\dots)$.

On peut aussi comme d'habitude, définir la disjonction de deux formules $\phi_1 \vee \phi_2$, qui correspond à la réunion, et le quantificateur universel $\forall x \phi$, pour tout x. Mais on les obtient grâce aux équivalences habituelles qui expriment la disjonction et le quantificateur universel à partir des trois précédents.

L'ensemble des formules du premier ordre est donc la clôture de l'ensemble des formules atomiques par trois opérations : la conjonction, la négation, la quantification existentielle (et donc aussi la disjonction et la quantification universelle). Au niveau des sous-ensembles définissables cela correspond à clore par intersection finie, complémentaire et projection.

Il est essentiel de remarquer que les formules sont obtenues à partir d'un nombre fini d'opérations, et que pour vérifier la satisfaction d'une formule dans un anneau R, on remplace les variables par des éléments de R, jamais par des sous-ensembles. Ce sont ces deux propriétés qui caractérisent la logique du premier ordre.

Il est également important de réaliser que pour que toutes ces définitions marchent bien, il y aurait un certain nombre de lemmes techniques à vérifier, et de conventions d'écriture à définir. Ici, on se permet de ne pas le faire car on utilise l'intuition que nous avons tous de ce qu'est un polynôme dans un anneau commutatif, des opérations que l'on peut faire avec, et des conventions habituelles de notation sur les indéterminées ou les variables dans ce cas.

Maintenant, si on veut étudier les anneaux commutatifs ordonnés, on s'autorise aussi à utiliser l'ordre et donc les formules atomiques dans ce cas seront les expressions de la forme $P(X_1,\dots,X_n) = 0$ et aussi celles de la forme $P(X_1,\dots,X_n) < 0$.

Dans son article sur l'o-minimalité, Scanlon va parler de langages qui étendent encore plus celui des anneaux commutatifs ordonnés. En gros, il considérera quatre types de structures particulièrement importantes. Tout d'abord, classiquement, les réels comme corps ordonné, $\langle \mathbb{R}, +, -, ., 0, 1, < \rangle$, dans le langage des anneaux commutatifs ordonnés. Ensuite, les réels avec l'exponentielle : $\mathbb{R}_{exp}$ dans laquelle on autorise aussi les formules atomiques du type $f(x_1,\dots,x_n) = 0$ et $f(x_1,\dots,x_n) > 0$, où f n'est plus simplement un polynôme mais est un polynôme exponentiel, c'est-à-dire obtenu par composition des opérations basiques, l'addition, l'opposé, la multiplication et la fonction exponentielle réelle. Une troisième structure, $\mathbb{R}_{an}$, dans laquelle on autorise toutes les fonctions analytiques réelles restreintes, c'est-à-dire, toutes les fonctions $h : \mathbb{R}^m \mapsto \mathbb{R}$ dont la restriction à $[-1,1]^m$ est analytique réelle et qui est nulle en dehors. Enfin, la structure $\mathbb{R}_{an,exp}$, dans laquelle on peut composer toutes les fonctions précédentes pour définir les formules atomiques.

Si l'on s'intéresse aux groupes (non nécessairement commutatifs), on prendra juste la loi de groupe, l'inverse et l'élément neutre. Les formules atomiques à n variables libres seront les identités du type

$$\prod_{1 \leqslant i \leqslant n} (x_i)^{e_{1,i}} \dots \prod_{1 \leqslant i \leqslant n} (x_i)^{e_{k,i}} = 1$$

avec $k \geqslant 1$, et $e_{j,i} \in \mathbb{Z}$, pour $1 \leqslant j \leqslant k$ et $1 \leqslant i \leqslant n$.

Si l'on s'intéresse, comme c'est le cas pour le théorème d'Ax-Kochen-Eršov (Théorème 6) aux *corps valués*, on peut rajouter la relation « $val(x) \leqslant val(y)$ », comme nous l'expliquons dans l'appendice (section 3).

Les théorèmes de théorie des modèles que nous allons montrer dans la suite (définition des ultraproduits, théorème de Łos, théorème de compacité) sont vrais dans le contexte général avec les mêmes démonstrations.

Quelques premiers exemples : $\mathbb{R} \not\models \exists x(x^2 = -1)$, mais $\mathbb{C} \models \exists x(x^2 = -1)$. Un anneau R satisfait à la formule $\forall x\,(x = 0) \vee (\exists y\, xy = 1)$ si et seulement si R

est un corps. Si p est un nombre premier,

$$R \models \text{“}p = 0 \wedge \forall x \forall y (xy \neq 0 \vee x = 0 \vee y = 0)\text{”}$$

ssi R est un anneau intègre de caractéristique égale à p.

Il est facile de vérifier que les formules sont préservées par isomorphisme (d'anneau) :

Lemme 1. *Si f est un isomorphisme (d'anneau) de R sur R', si $\phi(x_1,\dots,x_n)$ est une formule, alors pour tous $a_1,\dots,a_n \in R$,*

$$R \models \phi(a_1,\dots,a_n) \text{ ssi } R' \models \phi(f(a_1),\dots,f(a_n)).$$

On introduit quelques notations habituelles.

- On écrira $P(x_1,\dots,x_n) \neq 0$ à la place de $\neg(P(x_1,\dots,x_n) = 0)$;
- On note $(\phi \to \psi)$ la formule $(\neg\phi \vee \psi)$; on note $(\phi \leftrightarrow \psi)$ la formule $((\phi \to \psi) \wedge (\psi \to \phi))$.

Définition. Deux formules, $\phi_1(x_1,\dots,x_n)$ et $\phi_2(x_1,\dots,x_n)$ sont équivalentes (pour la classe des anneaux commutatifs) si, pour tout anneau commutatif R, et toute suite $a_1,\dots,a_n$ d'éléments de R,

$$R \models \phi_1(a_1,\dots,a_n) \text{ ssi } R \models \phi_2(a_1,\dots,a_n).$$

Les équivalences dont on a l'habitude sont bien sûr vraies, par exemple : $(\phi_1 \vee \phi_2)$ est équivalente à $\neg(\neg\phi_1 \wedge \neg\phi_2)$, $\forall x \phi$ est équivalente à $\neg\exists x \neg\phi$ ...

Exemples, théories, modèles

Définitions et remarques. Une formule sans variable libre est appelée un *énoncé*. Un énoncé est soit vrai, soit faux dans un anneau donné (et pas les deux en même temps !). Si l'anneau R satisfait l'énoncé θ, on dira que R est un *modèle* de θ. Si Σ est un ensemble d'énoncés et si R est modèle de chacun des énoncés de Σ, on dira que R est modèle de Σ, noté $R \models \Sigma$.
Un ensemble (éventuellement infini) d'énoncés qui a un modèle sera appelé *une théorie*.
On dit que deux anneaux R et R' sont *élémentairement équivalents* s'ils satisfont exactement aux mêmes énoncés du premier ordre, noté $R \equiv R'$. Évidemment, si R et R' sont isomorphes, ils sont en particulier élémentairement équivalents.
On dira qu'une théorie Σ est *complète* si tous ses modèles sont élémentairement équivalents.
On dit que θ est *conséquence* de Σ si tout modèle de Σ est également modèle de θ, noté $\Sigma \vdash \theta$.

Regardons tout de suite quelques exemples qui vont permettre de comprendre ce que l'on peut exprimer dans la logique du premier ordre et ce que l'on ne peut pas exprimer.
Les modèles de l'énoncé $\Theta_c := \forall x(x = 0) \vee (\exists y x.y = 1)$ sont exactement les anneaux commutatifs qui sont des corps.

Si p est un nombre premier, les corps modèles de l'énoncé $\sigma_p :=$ « $p = 0$ » sont exactement les corps de caractéristique p.

Les anneaux modèles de $\Sigma := \{\theta_c, \sigma_p\}$ (ou de l'énoncé « $(\theta_c \wedge \sigma_p)$ » sont exactement les corps de caractéristique p.

Peut-on trouver un énoncé (du premier ordre) dont les modèles soient exactement tous les corps de caractéristique zéro ? Si l'on essaye, on voit qu'on ne trouve pas. En revanche on peut trouver un ensemble infini d'énoncés qui fait l'affaire. Soit

$$\Sigma_0 = \{\neg\sigma_p : p \text{ premier }\}.$$

Alors un corps K est de caractéristique zéro si et seulement si, pour chaque p, $K \models$ « $p \neq 0$ », c'est-à-dire, $K \models \Sigma_0$.

Pour dire qu'un corps est de caractéristique zéro, on dit donc qu'il n'est de caractéristique p pour aucun nombre premier p. Et on va voir bientôt que le théorème de compacité prouve qu'effectivement il ne peut pas exister un énoncé (ou un nombre fini d'énoncés, ce qui est la même chose) pour dire qu'un corps est de caractéristique zéro. D'ailleurs le théorème de transfert que nous allons utiliser (Théorème 4) ne pourrait être vrai s'il existait un énoncé F disant « être de caractéristique zéro ». En effet alors $\neg F$ serait vrai dans tous les $\overline{\mathbb{F}_p}$ et, par le théorème, devrait être vrai dans $\mathbb{C}$.

Le même problème se pose si l'on cherche un énoncé dont les modèles seraient exactement tous les anneaux finis. En revanche on peut facilement dire, pour $n \geqslant 2$, qu'un anneau R a exactement n éléments. Soit, pour chaque entier $n \geqslant 2$,

$$G_n := \exists x_1,\dots \exists x_n [(\wedge_{1\leqslant i<j\leqslant n} x_i \neq x_j) \wedge (\forall y \vee_{1\leqslant i\leqslant n} y = x_i)]$$

$$F_n := \exists x_1,\dots \exists x_n (\wedge_{1\leqslant i<j\leqslant n} x_i \neq x_j).$$

Un anneau R est de cardinalité égale à n si et seulement s'il satisfait à l'énoncé G_n. Un anneau R est de cardinalité supérieure ou égale à n si et seulement s'il satisfait à l'énoncé F_n.

Un anneau R sera infini si et seulement si, pour tout n, $R \models F_n$. À nouveau, le théorème de compacité nous dira que l'on ne peut pas remplacer cet ensemble infini d'énoncés par un nombre fini.

Intuitivement, pour dire que R est fini, on aurait envie de dire « il existe un entier n tel que F_n », mais ceci n'est pas un énoncé du premier ordre. Dire « il existe un entier n tel que F_n » revient à faire une disjonction infinie $\vee_{n\geqslant 1} F_n$. Ceci n'est pas un énoncé du premier ordre.

Théorie des corps algébriquement clos, énoncé du théorème d'Ax

Nous allons vérifier qu'il existe une théorie, c'est-à-dire un ensemble (infini) d'énoncés, dont les modèles sont exactement les corps algébriquement clos et également que le théorème d'Ax s'énonce par une infinité d'énoncés du premier ordre.

Soit $n \geqslant 1$ un entier fixé, on peut écrire un énoncé, θ_n, dont les modèles sont exactement les anneaux R tels que tout polynôme unitaire de degré n à coefficients dans R a une solution dans R :

$$\forall y_0 \ldots \forall y_{n-1} \exists x (x^n + \sum_{i=0}^{n-1} y_i . x^i) = 0.$$

Les corps algébriquement clos sont alors exactement les modèles de la théorie

$$T_{CAC} := \{\theta_c\} \cup \{\theta_n ; n \geqslant 1\}.$$

Pour p premier, on note T_{CAC_p} la théorie des corps algébriquement clos de caractéristique p, c'est-à-dire $T_{CAC} \cup \{\sigma_p\}$. Et T_{CAC_0} sera la théorie des corps algébriquement clos de caractéristique zéro, c'est-à-dire $T_{CAC} \cup \{\neg\sigma_p : p \text{ premier}\}$.

Parmi les résultats fondamentaux sur la théorie des corps algébriquement clos, nous avons déjà mentionné que dans un corps algébriquement clos, les sous-ensembles définissables sont les ensembles constructibles, c'est-à-dire ceux définis sans recourir à la projection. Au niveau des formules, cela veut dire que dans la théorie des corps algébriquement clos, toute formule est équivalente à une formule sans quantificateurs. On dit que la théorie T_{CAC} *élimine les quantificateurs*. Autre résultat fondamental, deux corps algébriquement clos de même caractéristique sont élémentairement équivalents (voir le Principe de Lefschetz mentionné après l'énoncé du théorème 4), c'est-à-dire les théories T_{CAC_0} et T_{CAC_p}, pour tout p premier, sont des théories complètes.

Enfin, il nous faut vérifier que le *théorème d'Ax s'énonce par une infinité d'énoncés du premier ordre.*

Ce que l'on peut dire, par une formule (sans variable libre) $J_{n,d}$, c'est que, pour n et d fixés, toute application polynomiale de K^n dans K^n, de degré inférieur ou égal à d, qui est injective, est surjective. Alors il sera vrai, dans K, que toute application polynomiale injective est surjective si et seulement si, pour chaque n et chaque d, l'énoncé $J_{n,d}$ est vrai dans K.

Comme au-dessus, il va falloir quantifier sur les coefficients des polynômes, pour dire « pour tout polynôme de degré $\leqslant d$ », c'est donc un peu fastidieux à écrire. Je l'écris ici dans un cas simple $f = (f_1, f_2)$, application de degré $\leqslant 2$ de $\mathbb{C}^2$ dans $\mathbb{C}^2$ et je laisse le soin au lecteur de se convaincre que cela marche pareil pour n et d plus grands.

On a :

$f_1(x_1, x_2) = \sum_{i+j\leqslant 2} y_{ij} {x_1}^i {x_2}^j$ et $f_2(x_1, x_2) = \sum_{i+j\leqslant 2} z_{ij} {x_1}^i {x_2}^j$.

Soit $I(\bar{y}, \bar{z})$ la formule :

$$\begin{aligned}\forall x_1 \forall x_2 \forall x_1' \forall x_2' [(&(\sum_{i+j\leqslant 2} y_{ij} {x_1}^i {x_2}^j \\ &= \sum_{i+j\leqslant 2} y_{ij} {x_1'}^i {x_2'}^j) \wedge (\sum_{i+j\leqslant 2} z_{ij} {x_1}^i {x_2}^j \\ &= \sum_{i+j\leqslant 2} z_{ij} {x_1'}^i {x_2'}^j)) \rightarrow (x_1 = x_1' \wedge x_2 = x_2')].\end{aligned}$$

Alors $R \models I((\bar{y}, \bar{z})$ ssi f est injective.

Soit $S(\bar{y}, \bar{z})$ la formule

$$\begin{aligned}\forall v \forall w \exists x_1 \exists x_2 ((&\sum_{i+j\leqslant 2} y_{ij} {x_1}^i {x_2}^j = v) \\ &\wedge (\sum_{i+j\leqslant 2} z_{ij} {x_1}^i {x_2}^j = w)).\end{aligned}$$

Alors $R \models S(\bar{y}, \bar{z})$ ssi f est surjective.

Maintenant soit

$$J_{2,2} := (\forall \bar{y} \forall \bar{z} (I(\bar{y}, \bar{z}) \rightarrow S(\bar{y}, \bar{z}))).$$

Alors $R \models J_{2,2}$ si et seulement si toute application polynomiale de degré inférieur ou égal à 2, de R^2 dans R^2, qui est injective, est surjective.

Nous laisserons aussi le lecteur se convaincre que l'énoncé du Théorème 2 s'exprime également par une infinité de formules du premier ordre, pour chacune on fixe m, n et le degré des polynômes qui définissent l'action.

Pour montrer les Théorèmes 1 et 2, il nous suffit donc maintenant de démontrer le Théorème 4, ce que nous allons faire dans la section 1.3. En effet, nous avons vu (section 1.1) que, pour tout p, toute application polynomiale injective de $\overline{\mathbb{F}_p}^n$ dans $\overline{\mathbb{F}_p}^n$

est surjective. Par ce qui précède, pour tout n et d, l'énoncé $J_{n,d}$ est donc vrai dans $\overline{\mathbb{F}_p}$. Le théorème de transfert nous dit alors que chacun de ces énoncés $J_{n,d}$ est également vrai dans le corps $\mathbb{C}$, et donc, que toute application polynomiale injective de $\mathbb{C}^n$ dans $\mathbb{C}^n$ est surjective. Le raisonnement est identique pour le Théorème 2.

1.3 – Ultraproduits et démonstration du théorème de compacité

Théorème de compacité et conséquences

Nous allons donner la démonstration du théorème de compacité grâce aux ultraproduits un peu plus loin, mais voyons ce qu'il dit et ses premières conséquences.

Théorème 7 (Théorème de compacité 1). *Soit Σ un ensemble d'énoncés tel que tout sous-ensemble fini de Σ a un modèle. Alors Σ a un modèle.*

Une formulation équivalente :

Théorème 8 (Théorème de compacité 2). *Soit Σ un ensemble d'énoncés, et θ un énoncé. Si $\Sigma \vdash \theta$, il existe un sous-ensemble fini F de Σ tel que $F \vdash \theta$.*

Corollaire 1. *Si T est une théorie qui a des modèles finis de cardinalité arbitrairement grande (c'est-à-dire telle que, pour chaque $n \geqslant 1$, T a un modèle fini de cardinalité supérieure à n), alors T a un modèle infini.*

Preuve. Considérons l'énoncé $F_n := \exists x_1, \ldots \exists x_n \wedge_{1 \leqslant i < j \leqslant n} x_i \neq x_j$. Un modèle de F_n est de cardinalité au moins n. Soit $T' := T \cup \{F_n : n \geqslant 2\}$. Un modèle de T', s'il en existe, est un modèle de T de cardinalité supérieure à n pour chaque n, donc infini. Si F est un sous-ensemble fini de T', alors il existe k tel que $F \subset T \cup \{F_n : 2 \leqslant n \leqslant k\}$. Par hypothèse, T a un modèle de cardinalité au moins égale à k, R_k. Celui-ci est un modèle de $T \cup \{F_n : 2 \leqslant n \leqslant k\}$ et donc de F. Le théorème de compacité entraîne bien que T' a un modèle. □

Corollaire 2. *Si σ est vrai dans tous les corps de caractéristique zéro, alors il existe un nombre premier q tel que σ est vrai dans tous les corps de caractéristique supérieure ou égale à q.*

Preuve. Soit $T = \{\theta_c\} \cup \{p \neq 0 : p \text{ premier}\}$. Les modèles de T sont tous les corps de caractéristique 0. Si σ est vrai dans tous les corps de caractéristique 0, alors $T \vdash \sigma$. Par le théorème de compacité 2, il existe F fini, $F \subset T$, tel que $F \vdash \sigma$. Puisque F est fini, il existe un nombre premier q tel que $F \subset \{\theta_c\} \cup \{p \neq 0 : p < q\}$. Si K est un corps de caractéristique $\geqslant q$, K est modèle de F, et donc aussi modèle de σ. □

Le corollaire précédent s'énonce aussi de manière équivalente : *si σ est un énoncé qui est vrai dans des corps de caractéristique p, pour p arbitrairement grand, alors il existe un corps de caractéristique zéro dans lequel σ est vrai.*

Corollaire 3. *Si un énoncé σ est vrai dans tous les corps algébriquement clos de caractéristique 0, alors il existe p_0 tel que σ est vrai dans tous les corps algébriquement clos de caractéristique $p > p_0$.*

Preuve. Si σ est vrai dans tous les corps algébriquement clos de caractéristique zéro, alors σ est conséquence de la théorie que nous avons appelée T_{CAC_0} (section 1.2) donc, par le théorème de compacité, d'une partie finie de cette théorie. Il y a donc p_0 tel que σ est vrai dans tout modèle de $T_{CAC} \cup \{p \neq 0 : p \leqslant p_0\}$, donc en particulier dans tout corps algébriquement clos de caractéristique supérieure à p_0. □

On pourrait déduire maintenant le Théorème 4, du Corollaire 3 et du fait que tous les corps algébriquement clos de même caractéristique sont élémentairement équivalents. Toutefois, la preuve directe du Théorème 4 que nous allons donner, en utilisant les ultraproduits, nous semble particulièrement « parlante ».

En effet, la construction par ultraproduits permet de donner un sens précis à la notion de « limites » de structures, de propriétés vraies « asymptotiquement » et permet aussi de transformer la notion d'équivalence élémentaire en une notion plus algébrique d'isomorphisme.

Ultraproduits d'ensembles

Définitions. Soit I un ensemble non vide, un filtre sur I est une famille $\mathcal{F}$ de sous-ensembles de I, avec les propriétés suivantes : $I \in \mathcal{F}$, $\emptyset \notin \mathcal{F}$, si $X, Y \in \mathcal{F}$, alors $X \cap Y \in \mathcal{F}$, si $X \in \mathcal{F}$ et $X \subset Z \subset I$, alors $Z \in \mathcal{F}$. Si $\emptyset \neq X_0 \subset I$, $\mathcal{F}_{X_0} = \{Y \subset I : X_0 \subset Y\}$ est un filtre, qui est appelé un *filtre principal*. Si I est infini, l'ensemble des parties cofinies de X, $\mathcal{FR} = \{X \subset I ; I \setminus X \text{ est fini }\}$, est un filtre, appelé le *filtre de Fréchet*, qui n'est jamais principal. Un ultrafiltre est un filtre maximal

pour l'inclusion [3]. Tout filtre est contenu dans un ultrafiltre, par le lemme de Zorn. On vérifie facilement qu'un ultrafiltre est non principal si et seulement si il contient le filtre de Fréchet. Il suit de la définition que pour tout $X \subset I$, X est dans $\mathcal{U}$ ou bien son complémentaire est dans $\mathcal{U}$ (mais pas les deux) !

Maintenant, soit $(A_i)_{i\in I}$ une famille d'ensembles non vides, $\mathcal{A} = \prod_{i\in I} A_i$. On considère sur les éléments de $\mathcal{A}$, la relation d'équivalence $\equiv_{\mathcal{U}}$: $a \equiv_{\mathcal{U}} b$ ssi $\{i \in I; a(i) = b(i)\} \in \mathcal{U}$. On notera $a_{\mathcal{U}}$ la classe de a.
L'ensemble quotient, $\prod_{i\in I} A_i/\mathcal{U} = \{a_{\mathcal{U}} : a \in \prod_{i\in I} A_i\}$ est appelé *l'ultraproduit* des A_i (par $\mathcal{U}$).

Le lemme suivant sur les cardinalités nous sera utile.

Lemme 2. *Soit $(A_i)_{i\in\mathbb{N}}$ une famille d'ensembles non vides, dénombrables infinis. Soit $\mathcal{U}$ un ultrafiltre non principal sur $\mathbb{N}$. Alors $\prod_{i\in I} A_i/\mathcal{U}$ est de cardinalité égale au continu, c'est-à-dire à $|\mathbb{N}^{\mathbb{N}}| = |\mathcal{P}(\mathbb{N})|$.*

Preuve. Il est clair que $\prod_{i\in I} A_i/\mathcal{U}$ est de cardinalité inférieure ou égale au continu. Ce qui nécessite une preuve, c'est qu'on atteint le continu. On commence par construire un ensemble E d'applications de $\mathbb{N}$ dans $\mathbb{N}$, de cardinalité $|\mathcal{P}(\mathbb{N})|$ et tel que si $f, g \in E$, et $f \neq g$, alors $\{i \in \mathbb{N} : f(i) = g(i)\}$ est fini. Pour chaque application δ de $2^{\mathbb{N}}$ ($\delta : \mathbb{N} \mapsto \{0,1\}$), on définit une application f_δ de $\mathbb{N}$ dans $\mathbb{N}$ en posant

$$f_\delta(n) = \sum_{m<n} \delta(m)2^m.$$

La famille $E = \{f_\delta : \delta \in 2^{\mathbb{N}}\}$ a la propriété requise. Maintenant, pour chaque $i \in \mathbb{N}$, on fixe une énumération de $A_i = (e_{i,n})_{n\in\mathbb{N}}$.

Ensuite, pour chaque $f \in E$, on définit un élément a_f dans $\prod_{i\in\mathbb{N}} A_i$: $a_f(i) := e_{i,f(i)}$. Alors si $f \neq g$, $\{i \in \mathbb{N} : a_f(i) = b_g(i)\} = \{i \in \mathbb{N} : f(i) = g(i)\}$ est fini et n'est donc pas dans $\mathcal{U}$, ultrafiltre non principal. Donc, si $f \neq g$, $(a_f)_{\mathcal{U}} \neq (b_f)_{\mathcal{U}}$. □

Remarque. La famille d'applications E construite dans la preuve ci-dessus a également la propriété que, pour toute $f \in E$, et pour tout n, $f(n) < 2^n$. On peut en déduire de manière très semblable que, si $(A_i)_{i\in\mathbb{N}}$ est une famille d'ensembles finis, mais de cardinalités non bornées, c'est-à-dire tels que, pour chaque n, $\{i \in \mathbb{N} : |A_i| \leqslant n\}$ est fini, et si $\mathcal{U}$ est un ultrafiltre non principal sur $\mathbb{N}$, alors $\prod_{i\in I} A_i/\mathcal{U}$ est de cardinalité égale à $|\mathcal{P}(\mathbb{N})|$.

Maintenant, le but, en théorie des modèles, est de faire des ultraproduits de *structures*, et pas seulement d'ensembles. Sans surprise dans le cas de structures algébriques classiques on retrouve des constructions algébriques également classiques, mais il faut se souvenir que l'intérêt des ultraproduits c'est que c'est une *construction qui marche pour toutes les structures du premier ordre*, même si ici nous nous contenterons de regarder des ultraproduits d'anneaux commutatifs.

Ultraproduits d'anneaux

Soit $(R_i)_{i\in I}$ une famille d'anneaux commutatifs. Le produit cartésien de la famille, $\mathcal{R} := \Pi_{i\in I} R_i$, muni de l'addition et de la multiplication coordonnée par coordonnée, est un anneau commutatif.
En passant au quotient par un ultrafiltre sur I, on obtient encore un anneau commutatif.
En fait les ultraproduits dans ce cas sont des quotients par un idéal : l'ensemble $J_{\mathcal{U}}$ des éléments du produit qui sont nuls presque partout, c'est-à-dire $\{a \in \prod_{i\in I} R_i$: tels que $\{i \in I; a_i = 0\} \in \mathcal{U}\}$ est un idéal. Si les R_i sont tous des corps, alors $J_{\mathcal{U}}$ est un idéal maximal et donc l'ultraproduit est aussi un corps. Cela découle aussi du théorème général fondamental sur les ultraproduits, le Théorème de Łos, dont la preuve est en fait quasi immédiate à partir de la définition de l'ensemble des formules et des propriétés des ultrafiltres :

Théorème 9 (Théorème de Łos pour le cas des anneaux commutatifs). *Soit $(R_i)_{i\in I}$ une famille d'anneaux commutatifs et $\mathcal{U}$ un ultrafiltre sur I. Soit $\phi(x_1,\ldots,x_n)$ une formule, et $a_1,\ldots,a_n \in \Pi_{i\in I} R_i = \mathcal{R}$. Alors $\mathcal{R}/\mathcal{U} \models \phi(a_{1_{\mathcal{U}}},\ldots,a_{n_{\mathcal{U}}})$ ssi $\{i \in I : R_i \models \phi(a_1(i),\ldots,a_n(i))\} \in \mathcal{U}$.*

Remarque. Si I est infini et si $\mathcal{U}$ est un ultrafiltre principal sur I, $\mathcal{U}$ est engendré par $\{i_0\}$. Alors on peut vérifier que $\mathcal{R}$ est isomorphe à l'anneau R_{i_0}, par l'application $a \in R_{i_0} \mapsto b_{\mathcal{U}}$, où b est tel que $b(i_0) = a$ et $b(j) = 0$, pour $j \neq i_0$.

On peut maintenant montrer directement le théorème de transfert (section 1.1) qui implique le Théorème d'Ax (Théorème 1) et le Théorème 2.

Preuve (du Théorème 4). On considère P l'ensemble des nombres premiers et $\mathcal{U}$ un ultrafiltre non principal sur P. On a remarqué que $\mathcal{U}$ doit alors contenir le filtre de Fréchet sur P, donc tout ensemble de complémentaire fini. Soit $K_p := \overline{\mathbb{F}_p}$ et $\mathcal{K} := \prod_{p\in P} K_p/\mathcal{U}$. Quelles sont les propriétés du corps $\mathcal{K}$?

1. $\mathcal{K}$ est de caractéristique zéro : en effet, soit

3. Remarque : à tout ultrafiltre $\mathcal{U}$ sur I est associée une mesure finiment additive sur $\mathcal{P}(I)$, à valeur dans $\{0,1\}$, telle que $\mathcal{U}$ est la famille des ensembles de mesure 1.

p un premier, alors $\{q \in P : K_q \models p \neq 0\}$ est de complémentaire fini, donc est un élément du filtre de Fréchet et donc de $\mathcal{U}$. Par le Théorème de Łos, il suit que $\mathcal{K} \models p \neq 0$ aussi, et cela pour chaque p.

2. $\mathcal{K}$ est algébriquement clos. En effet, on a vu (section 1.2) qu'un corps est algébriquement clos si et seulement si, pour chaque $n \geqslant 1$, il satisfait à un certain énoncé θ_n. Par hypothèse, pour chaque p, K_p est algébriquement clos, donc satisfait à θ_n ; l'ensemble P est un élément de $\mathcal{U}$, donc par le Théorème de Łos, $\mathcal{K}$ est modèle de θ_n.

3. Quelle est la cardinalité de $\mathcal{K}$? Par le lemme 2, il est de la puissance du continu, c'est-à-dire a la même cardinalité que $\mathcal{P}(\mathbb{N})$, qui est aussi la cardinalité de $\mathbb{C}$. À isomorphisme près, il y a un unique corps algébriquement clos de caractéristique zéro et ayant la puissance du continu, c'est le corps des complexes $\mathbb{C}$.

Maintenant, soit σ un énoncé qui est vrai dans tous les K_p, alors, par le Théorème de Łos, il est vrai dans $\mathcal{K}$ qui est isomorphe à $\mathbb{C}$, donc il est vrai dans $\mathbb{C}$. □

Cela nous donne bien le Théorème 4.

Démonstration du théorème de compacité

Les ultraproduits permettent de donner une démonstration non syntactique du théorème de compacité, à partir de la famille d'anneaux R_F, où, si F est un sous-ensemble fini de l'ensemble infini d'énoncés Σ, R_F est un modèle de F. On va faire un ultraproduit des R_F et obtenir un anneau R dans lequel tous les énoncés de Σ seront vrais en même temps :

Théorème 10 (Théorème de compacité pour la classe des anneaux commutatifs). *Soit Σ un ensemble d'énoncés tel que tout sous-ensemble fini de Σ a un modèle. Alors Σ a un modèle.*

Preuve. Soit Σ un ensemble infini d'énoncés, tel que tout sous-ensemble fini F de Σ a un modèle R_F. Soit I l'ensemble de toutes les parties finies de Σ. Pour chaque énoncé θ de Σ, soit $V(\theta) := \{F \in I; \theta \in F\}$. On remarque que si $F \in V(\theta)$, $R_F \models \theta$. Dans la famille $\mathcal{V} := \{V(\theta); \theta \in \Sigma\}$, toute intersection finie est non vide. Il s'ensuit que l'on peut trouver un ultrafiltre (de l'ensemble des parties finies de I) qui la contient. On vérifie facilement que $(\prod_{F \in I} R_F)/\mathcal{U}$, l'ultraproduit des R_F, est, par le Théorème de Łos, un modèle de Σ. □

2. Un survol de théorie des modèles plus récente

2.1 – D'autres exemples d'applications des ultraproduits

Nous avons mentionné plus haut le fait que les ultraproduits permettent d'algébriser la notion d'équivalence élémentaire. En effet, on a le théorème suivant (Shelah) :

Théorème 11. *Soient M et N deux structures dans un langage L. Alors M et N sont élémentairement équivalentes si et seulement s'il existe un ensemble I et un ultrafiltre $\mathcal{U}$ sur I tel que $M^I/\mathcal{U}$ et $N^I/\mathcal{U}$ sont isomorphes.*

L'ultrapuissance de M, $M^I/\mathcal{U}$, est l'ultraproduit de $|I|$ copies de M. Un sens du théorème est immédiat, M et son ultrapuissance sont, par le Théorème de Łos, élémentairement équivalentes, donc, si M et N ont des ultrapuissances isomorphes, M et N sont élémentairement équivalentes. L'autre direction est nettement plus compliquée.

Enfin, l'un des principaux intérêts des ultraproduits, c'est qu'ils permettent d'obtenir des *résultats d'uniformité*.

Les applications sont très nombreuses. Nous citerons ici un seul exemple, simple à énoncer, concernant les groupes algébriques sur les corps finis.

Dans [14], Hrushovski et Pillay montrent plusieurs résultats sur des sous-groupes de $GL_n(\mathbb{F}_p)$, en utilisant la théorie des modèles des ultraproduits de corps finis (les corps dits *pseudo-finis*, étudiés originellement par Ax dans [2]).

En particulier, ils montrent :

Théorème 12. *Soit p un nombre premier et G un sous-groupe de $GL_n(\mathbb{F}_p)$ engendré par des éléments d'ordre p. Alors $G = \langle x_1 \rangle \dots \langle x_k \rangle$ pour des éléments $x_1, \dots, x_k$ de G d'ordre p, où l'entier k ne dépend que de n et pas de p.*

Ils donnent également de nouvelles preuves « modèle-théoriques » de plusieurs théorèmes déjà connus, dont, par exemple, le théorème de Nori-Weisfeiler qui montre (en particulier) qu'un sous-groupe de $SL_n(\mathbb{Z})$ est Zariski-dense si et seulement s'il se surjecte sur $SL_n(\mathbb{Z}/p\mathbb{Z})$ pour tous les nombres premiers p sauf peut-être un nombre fini d'entre eux.

Les preuves de ce type de résultats sont bien sûr beaucoup plus sophistiquées que les deux exemples très simples que nous avons donnés au début de ce texte et utilisent des outils et des notions de théo-

rie des modèles développés depuis les années 60. Dans notre dernière section, nous allons tenter de donner une toute petite idée de ce à quoi s'intéresse la théorie des modèles moderne.

2.2 – La théorie des modèles géométrique

Pour ce survol nous allons, comme dans le reste de cet article, partir d'un exemple et reprendre la théorie des corps algébriquement clos que nous avions notée T_{CAC}. Pour simplifier les notations, prenons les corps de caractéristique 0, donc la théorie T_{CAC_0} (les définitions et les notations ont été introduites dans la section 1.2).
Nous y avons vu quelques propriétés :

– la théorie T_{CAC_0} est complète, c'est-à-dire, deux corps algébriquement clos de caractéristique zéro sont élémentairement équivalents ;

– la théorie T_{CAC_0} élimine les quantificateurs : les sous-ensembles définissables sont les ensembles constructibles, définis à partir des ensembles atomiques sans utiliser le quantificateur existentiel (la projection).

Dans le cas de ces deux propriétés, on part de la classe des corps algébriquement clos, et on démontre, pour cette classe, des résultats sur les formules qui sont vraies dans ces corps. Rappelons-nous maintenant le résultat « dans l'autre sens » que nous avions énoncé dans la section 1.2, celui de Macintyre : si un corps infini admet l'élimination des quantificateurs, alors c'est un corps algébriquement clos. Là, on part d'une propriété « abstraite » sur la famille des ensembles définissables et on en déduit des propriétés algébriques. On commence à faire de la théorie des modèles moderne...

En fait, on considère en général que le vrai tournant dans l'histoire de la théorie des modèles, le moment où l'on est passé de la théorie des modèles classique au développement intrinsèque de nouvelles notions et de nouveaux outils, plus sophistiqués, c'est en 1965, avec un célèbre théorème de Morley.

Avant d'énoncer ce théorème, reprenons nos corps algébriquement clos. Ils ont une autre propriété fondamentale : le type d'isomorphisme d'un corps K algébriquement clos de caractéristique zéro est donné par la cardinalité d'une base de transcendance sur $\mathbb{Q}$. Donc, comme le passage à la clôture algébrique ne change pas la cardinalité, il y aura une infinité dénombrable de corps algébriquement clos dénombrables à isomorphisme près :

$$\mathbb{Q}^{alg}, \mathbb{Q}(t_1)^{alg}, \ldots, \mathbb{Q}(t_1, \ldots, t_n)^{alg}, \ldots, \mathbb{Q}(t_1, \ldots, t_n, \ldots)^{alg}$$

où $t_1, \ldots, t_n, \ldots$ sont algébriquement indépendants. Mais, pour toute cardinalité λ non dénombrable, il y aura à isomorphisme près un seul corps de cardinalité λ, celui dont une base de transcendance est de cardinalité λ. Quand une théorie a un seul modèle de cardinalité λ à isomorphisme près, pour un cardinal λ infini, on dira que la théorie est *λ-catégorique*. La théorie T_{CAC_0} est donc λ-catégorique pour tout cardinal λ non dénombrable, et on voit bien pourquoi, parce qu'il y a une notion d'indépendance (l'indépendance algébrique) qui donne lieu, comme l'indépendance linéaire dans les espaces vectoriels, à l'existence de bases. Le théorème de Morley nous explique que cela n'est pas un hasard.

Théorème 13 (Théorème de Morley [17]). *Soit T une théorie complète dans un langage dénombrable, alors T est λ-catégorique pour un λ non dénombrable si et seulement si T est λ-catégorique pour tout cardinal λ non dénombrable.*

Pour arriver à ce résultat, Morley montre que l'hypothèse abstraite, « T est λ-catégorique pour un λ non dénombrable », entraîne que, dans tout modèle de T, on peut définir une bonne notion de dimension (le rang de Morley) dans la famille de tous les sous-ensembles définissables et une notion d'indépendance telle que, dans tout modèle, deux ensembles indépendants maximaux ont la même cardinalité (existence de bases). Chaque modèle est en quelque sorte une « clôture » de n'importe laquelle de ses bases ; et donc le type d'isomorphisme d'un modèle est déterminé uniquement par la cardinalité d'une base. Appliquée au cas particulier des fermés de Zariski (qui sont bien des sous-ensembles définissables) dans les corps algébriquement clos, le rang de Morley coïncide avec la dimension algébrique.

La théorie des modèles « pure » (par opposition à l'époque à la théorie des modèles « appliquée ») était née. Shelah, dans une œuvre monumentale ([20]), développe, à partir des années 70, une théorie encore plus générale de l'indépendance et de la dimension, appelée *théorie de la stabilité* ou *théorie de la classification*. L'une des idées maîtresses est de classifier les structures et leurs théories en fonction des objets combinatoires infinis qu'on peut y interpréter grâce aux formules : ordres, arbres ... Par exemple, si on ne peut pas définir d'ordre infini, on dit alors que la structure est *stable*, la notion d'indépendance abstraite se comporte particulièrement bien. C'est évidemment le cas des corps algébriquement clos et des espaces vectoriels, mais aussi des corps séparablement clos, des corps différentiellement clos, ou encore de toutes les théories

de modules. On sait que si un corps infini est stable et que tous ses sous-ensembles définissables ont un rang de Morley bien défini (on dit alors que le corps est ω-stable), il est algébriquement clos. En revanche la caractérisation des corps infinis qui sont seulement stables est toujours ouverte : on conjecture que les corps stables sont exactement les corps séparablement clos.

Dans la seconde moitié des années 80, sous l'influence de Zilber tout d'abord, puis de Hrushovski, la théorie de la stabilité a commencé à évoluer, en se concentrant sur l'étude locale fine des ensembles définissables de dimension finie. Dans la théorie de la stabilité à la Shelah, les méthodes utilisées étaient souvent celles de la combinatoire infinie. Avec Zilber et Hrushovski, les méthodes et les outils, bien que s'inscrivant dans la suite des travaux de Shelah, sont plutôt inspirés de la géométrie, à la fois dans le sens des géométries combinatoires (matroïdes) et de la géométrie algébrique. C'est la naissance de la *stabilité géométrique* ou *théorie des modèles géométrique* et de ses applications remarquées à la géométrie Diophantienne (qui ont commencé avec la preuve de la conjecture de Mordell-Lang pour les corps de fonctions en toute caractéristique par Hrushovski en 1994 [12]).

Maintenant, on classifie les théories en fonction des structures algébriques classiques qu'on peut y définir (groupes, corps...) et on cherche aussi à caractériser ces structures classiques par des propriétés de géométrie abstraites.

On obtient par exemple, mais dans un cadre beaucoup plus général, des résultats (la *configuration de groupe*) qui sont en droite ligne de deux théorèmes classiques bien connus.

1. Le vieux théorème de géométrie : une géométrie projective Desarguésienne de dimension au moins 3 est la géométrie projective sur un corps (on « construit » le corps à partir de la « géométrie combinatoire » donnée).

2. Le théorème de Weil qui construit un groupe algébrique à partir d'une opération générique rationnelle associative sur une variété algébrique.

Pour donner une idée intuitive des questions qui se posent dans ce cadre, reprenons notre exemple-guide. Après la complétude et la catégoricité, la théorie des corps algébriquement clos a en fait une propriété encore plus forte. Pour K algébriquement clos, regardons les sous-ensembles définissables dans K lui-même, c'est-à-dire ceux qui correspondent aux formules à une seule variable. Les ensembles définissables sont les ensembles constructibles, donc les sous-ensembles de K sont des combinaisons booléennes finies d'ensembles de solutions d'un polynôme à une variable. Il n'est pas difficile de vérifier qu'un tel sous-ensemble est soit fini, soit de complémentaire fini dans K.

Définition 3. Soit T une théorie complète (dans un langage dénombrable), on dira que T est *fortement minimale* si, pour toute formule $\phi(x, y_1, \ldots, y_n)$ et pour tout modèle M de T, pour tous $b_1, \ldots, b_n \in M$, l'ensemble $\{m \in M : M \models \phi(m, b_1, \ldots, b_n\}$ est fini ou de complémentaire fini.

C'est une condition très forte, qui entraîne l'existence d'une dimension pour laquelle tous les ensembles définissables sont de dimension finie, l'existence de bases, et donc en particulier que la théorie est catégorique en tout cardinal non dénombrable. De plus cette dimension est associée à une géométrie (au sens combinatoire) sur chaque modèle.

Il y a trois familles d'exemples classiques de théories fortement minimales et des géométries correspondantes.

1. Les théories dites *triviales*. La géométrie associée est triviale et, en particulier il n'y a pas de groupe infini définissable dans les modèles de T. Exemple typique : la théorie avec pour seule relation l'égalité dont les modèles sont exactement tous les ensembles infinis (et les isomorphismes sont les bijections entre modèles).

2. Les théories dites *localement modulaires*. La géométrie associée est affine ou projective, on montre qu'on peut définir un groupe infini abélien dans tout modèle de T et en fait la théorie T est presque la théorie des espaces vectoriels de dimension infinie sur un corps (non nécessairement commutatif) fixé.

3. La théorie des corps algébriquement clos, ou plus généralement une théorie telle que l'on puisse définir dans ses modèles un corps algébriquement clos.

Ce sont là les seuls exemples *classiques* de théories fortement minimales et la *Conjecture de trichotomie de Zilber* énonçait que, en gros, ce sont les seuls. Cette conjecture est restée ouverte plusieurs années et a été l'un des moteurs du développement du sujet. En 1993, Hrushovski a construit une famille de contre-exemples inattendus : des structures fortement minimales avec une géométrie non triviale mais dans lesquelles on ne peut pas définir de groupe (et donc encore moins de corps).

Le principe de trichotomie de Zilber est donc faux en général, mais en fait, avec quelques hypo-

thèses supplémentaires sur la théorie, plus précisément dans le contexte des *Structures de Zariski*, introduites par Hrushovki et Zilber également à la même époque ([15]), il est vrai. Ce principe de trichotomie, appliqué dans les corps différentiellement clos de caractéristique 0 et dans les corps séparablement clos non parfaits, a joué un rôle essentiel dans les premières applications de la théorie des modèles à la géométrie Diophantienne.

Nous nous arrêterons là pour ce survol du début de la stabilité géométrique. Les lecteurs remarqueront que les résultats que nous avons cités en exemple datent des années 90, bien des choses se sont passées depuis. Le développement de nouveaux outils a entraîné l'élargissement encore des champs d'application des méthodes de théorie des modèles, citons par exemple les travaux de Hrushovski et Loeser sur les espaces de Berkovich ([13]) ou les applications à la combinatoire additive ([11]).

D'autres méthodes de théorie des modèles se sont également mises à jouer un rôle essentiel en Géométrie Diophantienne, avec les travaux de Pila et Wilkie dans le cadre des théories o-minimales. Ces développements plus récents sont exposés dans l'article de Scanlon de ce même numéro.

Enfin, ces dernières années, en cherchant à étendre les méthodes utilisées dans les théories stables à d'autres types de structures, instables elles (ensembles ordonnés, corps valués, graphe aléatoire), on est souvent revenu à des méthodes plus combinatoires que géométriques. Il en est un peu question dans l'article de Scanlon et également dans un exposé récent au séminaire Bourbaki [21] sur les théories dites NIP et leurs applications à la combinatoire.

3. Petit appendice sur les corps valués

Quelques définitions sur les corps valués.

- Un *corps valué* est un corps K, muni d'une application surjective $v : K^* \mapsto G$, où $(G, +, 0_G)$ est un groupe abélien ordonné, appelé le *groupe de valeurs* de (K, v), telle que, pour tous $x, y \in K^*$,
  - $v(xy) = v(x) + v(y)$
  - $v(x + y) \geqslant \min\{v(x), v(y)\}$ (l'inégalité ultramétrique).
- L'anneau des entiers de la valuation v, $O_{K,v} := \{x \in K^*; v(x) \geqslant 0\} \cup \{0\}$, a un unique idéal maximal, $M_{K,v} := \{x \in O_{K,v}; v(x) > 0\} \cup \{0\}$.
- Le corps résiduel de (K, v), $\overline{K_v} := O_{K,v}/M_{K,v}$. Un corps valué K est dit d'équicaractéristique 0 si K et son corps résiduel sont tous deux de caractéristique 0.
- Le corps valué (K, v) est dit *hensélien* si la valuation v a une seule extension sur toute extension algébrique de K.
- On considère $\mathbb{F}_p((t))$, avec la valuation habituelle : $v_1 : \mathbb{F}_p((t))^* \mapsto \mathbb{Z}$ où $v_1(\sum_{i \geqslant m} a_i t^i) :=$ le plus petit entier m tel que $a_m \neq 0$. Le groupe de valeur de v_1 est $\mathbb{Z}$, le corps résiduel est $\mathbb{F}_p$, et le corps valué $(\mathbb{F}_p((t)), v_1)$ est hensélien.
- Pour $\mathbb{Q}_p$, on considère v_p la valuation p-adique sur $\mathbb{Q}$. Pour $n \in \mathbb{Z}^*$, on définit $v_p(n)$ comme le plus grand m tel que p^m divise n ; v_p s'étend à $\mathbb{Q}$: $v_p(a/b) = v_p(a) - v_p(b)$. Le corps $\mathbb{Q}_p$ est le complété de $\mathbb{Q}$ pour la valuation v_p. Attention, on a plutôt l'habitude de considérer la *valeur absolue* p-adique sur $\mathbb{Q}$, $|x|_p = p^{-v_p(x)}$. Le groupe de valeur de v_p est $\mathbb{Z}$, le corps résiduel est $\mathbb{F}_p$ et le corps valué $(\mathbb{Q}_p, v_p)$ est hensélien.

Du point de vue de la théorie des modèles, dans un corps valué (K, v), on pourra prendre comme ensemble de formules atomiques, les formules de la forme : $P(x_1, \ldots, x_n) = 0$ et $v(P(x_1, \ldots, x_n) \leqslant v(Q(x_1, \ldots, x_n)$, pour $P, Q \in K[X_1, \ldots, X_n]$.

Dans ce langage, il existe une théorie, un ensemble (infini) d'énoncés, dont les modèles sont exactement les corps valués henséliens. Pour chaque nombre premier p, $\mathbb{Q}_p$ et $\mathbb{F}_p((t))$ sont tous deux des corps valués henséliens, ils ont même groupe de valeur et même corps résiduel, mais $\mathbb{Q}_p$ est de caractéristique 0 et $\mathbb{F}_p((t))$ est de caractéristique p.

Revenons brièvement sur les quelques explications que nous avions données dans la section 1.1 sur la démonstration, par Ax et Kochen, des Théorèmes 3 et 5.

Prenons un ultrafiltre non principal sur l'ensemble des nombres premiers, $\mathcal{U}$, et considérons d'une part K_1, l'ultraproduit des corps $\mathbb{Q}_p$, et d'autre part K_2, l'ultraproduit des corps $\mathbb{F}_p((t))$. Par le Théorème de Łos (Théorème 9), K_1 et K_2 sont tous deux des corps valués henséliens. De plus, leurs corps résiduels k_1 et k_2 vont être les mêmes (l'ultraproduit de copies de $\mathbb{F}_p$) et leurs groupes de valeur, G_1 et G_2 aussi (l'ultraproduit de copies de $\mathbb{Z}$). Par le Théorème de Łos également, puisque p devient de plus en plus grand, les corps K_1 et K_2 seront de caractéristique 0, et les corps résiduels, k_1 et k_2 aussi. Donc

par le théorème d'Ax-Kochen-Ersov (Théorème 6), K_1 et K_2 satisfont aux mêmes énoncés.

En appliquant le Théorème de Łos à nouveau, on en déduit immédiatement le Théorème 5 : soit σ un énoncé du premier ordre dans le langage des corps valués ; σ est vrai dans $\mathbb{Q}_p$ pour tout p sauf un nombre fini si et seulement si σ est vrai dans K_1 si et seulement si σ est vrai dans K_2 si et seulement si σ est vrai dans $\mathbb{F}_p((t))$ pour tout p sauf un nombre fini.

Ensuite, pour en déduire le Théorème 3, comme nous l'avons indiqué dans la section 1.1, il suffit de vérifier que « tout polynôme homogène de degré d en d^2+1 variables a un zéro non trivial » s'exprime, pour chaque $d \geqslant 1$, par un énoncé du premier ordre, ϕ_d. En effet, dans $\mathbb{F}_p((t))$, pour tout p, pour tout $d \geqslant 1$, l'énoncé ϕ_d est vrai. Par le Théorème 5, pour chaque $d \geqslant 1$, l'énoncé ϕ_d est donc vrai dans $\mathbb{Q}_p$ pour tout p sauf un nombre fini.

Références

[1] J. Ax. « Injective endomorphisms of varieties and schemes ». *Pacific J. Math.* **31** (1969), p. 1–7.

[2] J. Ax. « The elementary theory of finite fields ». *Ann. of Math. (2)* **88** (1968), p. 239–271.

[3] J. Ax et S. Kochen. « Diophantine problems over local fields. I ». *Amer. J. Math.* **87** (1965), p. 605–630.

[4] I. Bertuccioni. « Algebraic actions of p-groups ». *Arch. Math. (Basel)* **58**, n° 4 (1992), p. 329.

[5] A. Białynicki-Birula et M. Rosenlicht. « Injective morphisms of real algebraic varieties ». *Proc. Amer. Math. Soc.* **13** (1962), p. 200–203.

[6] A. Borel. « Injective endomorphisms of algebraic varieties ». *Arch. Math. (Basel)* **20** (1969), p. 531–537.

[7] G. T. Buzzard. « Tame sets, dominating maps, and complex tori ». *Trans. Amer. Math. Soc.* **355**, n° 6 (2003), p. 2557–2568.

[8] R. Cluckers et F. Loeser. « Ax-Kochen-Eršov theorems for p-adic integrals and motivic integration ». In : *Geometric methods in algebra and number theory*. Vol. 235. Progr. Math. Birkhäuser Boston, Boston, MA, 2005, p. 109–137.

[9] J. Denef. « Proof of a conjecture of Colliot-Thélène ». Preprint, arXiv:1108.6250. 2011.

[10] J. L. Eršov. « On elementary theories of local fields ». *Algebra i Logika Sem.* **4**, n° 2 (1965), p. 5–30.

[11] E. Hrushovski. « Stable group theory and approximate subgroups ». *J. Amer. Math. Soc.* **25**, n° 1 (2012), p. 189–243.

[12] E. Hrushovski. « The Mordell-Lang conjecture for function fields ». *J. Amer. Math. Soc.* **9**, n° 3 (1996), p. 667–690.

[13] E. Hrushovski et F. Loeser. *Non-archimedean tame topology and stably dominated types*. **192**. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 2016, p. vii+216.

[14] E. Hrushovski et A. Pillay. « Definable subgroups of algebraic groups over finite fields ». *J. Reine Angew. Math.* **462** (1995), p. 69–91.

[15] E. Hrushovski et B. Zilber. « Zariski geometries ». *J. Amer. Math. Soc.* **9**, n° 1 (1996), p. 1–56.

[16] A. Macintyre. « On ω_1-categorical theories of fields ». *Fund. Math.* **71**, n° 1 (1971), p. 1–25.

[17] M. Morley. « Categoricity in power ». *Trans. Amer. Math. Soc.* **114** (1965), p. 514–538.

[18] W. Rudin. « Injective polynomial maps are automorphisms ». *Amer. Math. Monthly* **102**, n° 6 (1995), p. 540–543.

[19] J.-P. Serre. « How to use finite fields for problems concerning infinite fields ». In : *Arithmetic, geometry, cryptography and coding theory*. Vol. 487. Contemp. Math. Amer. Math. Soc., Providence, RI, 2009, p. 183–193.

[20] S. Shelah. *Classification theory and the number of nonisomorphic models*. Second. **92**. Studies in Logic and the Foundations of Mathematics. North-Holland Publishing Co., Amsterdam, 1990.

[21] S. Starchenko. « NIP, Keisler measures and combinatorics [after S. Shelah, H. J. Keisler, E. Hrushovski, Y. Peterzil, A. Pillay, P. Simon, ...] » Séminaire Bourbaki: Vol. 2015/2016, Exp. No. 1119. 2016.

[22] G. Terjanian. « Un contre-exemple à une conjecture d'Artin ». *C. R. Acad. Sci. Paris Sér. A-B* **262** (1966), A612.



Élisabeth Bouscaren

Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, Centre national de la recherche scientifique (cnrs), Université Paris-Saclay, 91405 Orsay.

Élisabeth Bouscaren est Directrice de Recherche au cnrs, au Laboratoire de Mathématiques d'Orsay, dans l'équipe « Arithmétique et Géométrie Algébrique ». Elle est spécialiste de Théorie des Modèles et de ses applications à l'Algèbre et la Géométrie.

Merci à Tom Scanlon pour avoir accepté de m'accompagner dans cette aventure, et à Caroline Gruson, Martin Hils et Silvain Rideau, pour leurs relectures de ce texte et leurs remarques.