Introduction aux couplages en théorie ergodique

T. de la Rue

Master 2 MFA Rouen 2024-2025

On considère ici des systèmes dynamiques mesurés de la forme (X, μ, T), où T est une transformation inversible, bi-mesurable, préservant la mesure de probabilité μ sur l'espace de Lebesgue $\big(X, \mathscr{B}(X), \mu\big)$. Quand il n'y aura pas d'ambiguïté, on désignera par le seul symbole T le système dynamique (X, μ, T).

Le but de ce chapitre est l'étude des liens que peuvent avoir différents systèmes dynamiques entre eux, à travers une notion fondamentale en théorie ergodique, introduite par Furstenberg en 1967 [1] : les couplages.

1 Couplages de systèmes dynamiques mesurés

1.1 Produits et facteurs : arithmétique des systèmes dynamiques

Dans ce célèbre article, Furstenberg observe qu'une sorte d'arithmétique peut être faite avec les systèmes dynamiques. En effet, il existe deux opérations naturelles en théorie ergodique qui présentent quelques analogies avec des opérations sur les nombres entiers.

À partir de deux systèmes dynamiques donnés (X, μ, T) et (Y, ν, S), on peut construire leur *produit direct* $(X \times Y, \mu \otimes \nu, T \times S)$ où $T \times S : \ (x, y) \longmapsto (Tx, Sy)$. Remarquons que le produit cartésien $X \times Y$ est ici muni de la mesure *produit* $\mu \otimes \nu$, qui est toujours $T \times S$-invariante. Opérant sur les classes d'équivalences de systèmes isomorphes, le produit direct est commutatif, associatif, et possède un élément neutre : le système trivial réduit à un point.

Inversement, on dit que (Y, ν, S) est un *facteur* du système (X, μ, T) si il existe une application mesurable $\pi : X \to Y$ satisfaisant :
— $\pi_*(\mu) = \nu$;
— $\pi \circ T = S \circ \pi$.
Une telle application est appelée un *homomorphisme* de systèmes dynamiques. L'existence de cette application peut être interprétée de la façon suivante : on peut « voir » le système (Y, ν, S) à l'intérieur de (X, μ, T) en observant $\pi(x)$.

Exercice 1.1. *Soit (Y, ν, S) un facteur de (X, μ, T). Prouver que*

1. *T ergodique $\Longrightarrow$ S ergodique.*
2. *T faiblement mélangeant $\Longrightarrow$ S faiblement mélangeant.*
3. *T mélangeant $\Longrightarrow$ S mélangeant.*

Il est clair par exemple que les deux systèmes (X, μ, T) et (Y, ν, S) sont facteurs du produit direct $(X \times Y, \mu \otimes \nu, T \times S)$. En effet chacun de ces systèmes peut être vu dans le produit en observant seulement une des coordonnées. Il est évident aussi que le système trivial réduit à un seul point est facteur de n'importe quel autre système, via un homomorphisme constant.

Facteur engendré par une fonction mesurable

Il existe une manière simple d'obtenir un facteur d'un système dynamique (X, μ, T) : à partir de n'importe quelle fonction mesurable $f : X \to \mathbb{R}$, on peut considérer la fonction $\varphi : X \to \mathbb{R}^{\mathbb{Z}}$ définie à partir de f par

$$\varphi(x) := \big(f(T^n x)\big)_{n \in \mathbb{Z}}.$$

La mesure image $\nu := \varphi_*(\mu)$ est une probabilité sur $\mathbb{R}^{\mathbb{Z}}$ qui est invariante par le décalage des coordonnées S (parce que μ est invariante par T). Et par construction, on a évidemment $\varphi \circ T = S \circ \varphi$. Autrement dit, φ est un homomorphisme entre le système dynamique (X, μ, T) et le système dynamique $\big(\mathbb{R}^{\mathbb{Z}}, \mathscr{B}(\mathbb{R}^{\mathbb{Z}}), \nu, S\big)$, et ce dernier système est un facteur de T. On l'appelle le *facteur de T engendré par f*.

Dans cette construction, on peut clairement remplacer $\mathbb{R}$ par n'importe quel espace polonais (métrisable avec une métrique complète et séparable).

Il est important de noter que, dès que f n'est pas μ-p.s. constante, le facteur obtenu n'est pas le système réduit à un seul point (qui est de manière triviale facteur de n'importe quel système dynamique). Mais certains systèmes dynamiques T ont la propriété d'être « premiers » (encore une analogie avec l'arithmétique !), dans le sens où, si $\pi : X \to Y$ est un homomorphisme entre les systèmes dynamiques T et S, alors ou bien S est le système réduit à un seul point, ou bien π est en fait un isomorphisme.

Systèmes premiers entre eux ?

En poursuivant l'analogie avec les entiers, on peut voir deux façons de définir deux système « premiers entre eux » :

Propriété 1. *(X, μ, T) et (Y, ν, S) n'ont pas de facteur commun en-dehors du système trivial réduit à un point.*

Propriété 2. *Si (X, μ, T) et (Y, ν, S) sont facteurs d'un même système dynamique, alors le produit direct $(X \times Y, \mu \otimes \nu, T \times S)$ est aussi un facteur de ce système dynamique.*

Pour être tout-à-fait précis, la seconde propriété devrait être formulée plutôt comme suit : Si (X, μ, T) et (Y, ν, S) apparaissent comme facteurs d'un même système dynamique (Z, ρ, R), à travers des homomorphismes π_X et π_Y, alors le produit direct $(X \times Y, \mu \otimes \nu, T \times S)$ apparaît lui-même comme facteur, via l'homomorphisme $\pi_X \times \pi_Y$.

Les liens entre ces deux propriétés ne sont pas a priori évidentes, ils seront abordés plus loin dans le cours (section 3). La propriété 2 a été nommée *disjonction* des systèmes (X, μ, T) et (Y, ν, S) par Furstenberg. Cette propriété concerne toutes les façons possibles de voir simultanément deux systèmes donnés dans un même troisième. C'est précisément l'objet de la théorie des couplages que nous allons exposer maintenant.

1.2 Couplages

Définition 1.1. *Soient (X, μ, T) et (Y, ν, S) deux systèmes dynamiques. Un* couplage *de ces systèmes est une mesure de probabilité λ sur le produit cartésien $X \times Y$, dont les marginales sur X et Y sont respectivement μ et ν, et qui est invariante par la transformation produit $T \times S : (x, y) \mapsto (Tx, Sy)$.*

L'ensemble des couplages de T et S, noté $J(T, S)$, n'est jamais vide : en effet il contient toujours la mesure produit $\mu \otimes \nu$. Chaque $\lambda \in J(T, S)$ permet de définir un nouveau système dynamique $(X \times Y, \lambda, T \times S)$, que nous noterons

$(T \times S)_\lambda$ pour éviter la confusion avec le produit direct $T \times S$. Les systèmes T et S sont tous les deux facteurs de $(T \times S)_\lambda$, via les homomorphismes donnés par les projections sur X et Y respectivement. Inversement, si T et S apparaissent comme facteurs d'un troisième système dynamique (Z, ρ, R) à travers des homomorphismes $\pi_X : Z \to X$ et $\pi_Y : Z \to Y$, on peut toujours considérer l'application $\pi : Z \to X \times Y$ définie par $\pi(z) := (\pi_X(z), \pi_Y(z))$. Alors la mesure image $\lambda := \pi_*(\rho)$ sur $X \times Y$ est un couplage de T et S, et le système $(T \times S)_\lambda$ est un facteur de R.

Cette définition peut êtres généralisée de manière évidente à la notion de *couplage d'une famille finie ou dénombrable de systèmes dynamiques* $(T_i)_{i \in I}$. Si tous les systèmes dynamiques T_i sont des copies d'un même système dynamique T, on parle d'*autocouplages* de T. En particulier, pour tout $k \geq 2$, un couplage de k copies du système T est appelé *autocouplage d'ordre* k de T. L'ensemble des autocouplages d'ordre k de T est noté $J_k(T)$.

Couplages ergodiques

On note $J_e(T, S)$ l'ensemble des couplages *ergodiques* de T et S. Puisque chaque facteur d'un système ergodique est lui-même ergodique, une condition nécessaire pour que $J_e(T, S)$ ne soit pas vide est que T et S soient tous deux ergodiques. La proposition suivante montre que la réciproque est vraie.

Proposition 1.2. *Si T et S sont deux systèmes ergodiques, alors il existe au moins un couplage ergodique T et S.*

Démonstration. On part du seul couplage dont on connaît l'existence à coup sûr : la mesure produit $\mu \otimes \nu$. Ce couplage peut ne pas être ergodique, mais alors on peut considérer sa *décomposition en composantes ergodiques* : Il existe une mesure de probabilité P sur l'ensemble de toutes les mesures de probabilités sur $X \times Y$, $T \times S$-invariantes et ergodiques, telle que

$$\mu \otimes \nu \; = \; \int \lambda \, dP(\lambda).$$

En notant λ_X (respectivement λ_Y) la marginale de la composante ergodique λ sur X (respectivement sur Y), on obtient

$$\mu \; = \; \int \lambda_X \, dP(\lambda),$$

et

$$\nu = \int \lambda_Y \, dP(\lambda).$$

Comme μ et ν sont ergodiques, ceci implique que pour P-presque tout λ, $\lambda_X = \mu$ et $\lambda_Y = \nu$. Autrement dit, P-presque tout λ est un couplage ergodique de T et S. $\square$

Notons que cette preuve donne en fait un résultat plus fort que celui annoncé : si T et S sont ergodiques, alors la décomposition ergodique d'un couplage de T et S ne fait intervenir que des couplages ergodiques de T et S.

2 De la disjonction à l'isomorphisme

Dans cette section, on s'intéresse aux deux cas extrêmes concernant les couplages de deux systèmes dynamiques T et S. Le premier a lieu quand l'ensemble des couplages $J(T, S)$ est réduit au singleton $\{\mu \otimes \nu\}$. Heuristiquement, cela signifie que T et S n'ont rien en commun qui pourrait conduire à un couplage non trivial. (La question de ce que S et T ont en commun quand ils ne sont pas disjoints est étudiée dans la section 3.2.) À l'opposé, si T et S sont isomorphes, on peut construire des couplages de T et S très spéciaux, supportés sur les graphes des isomorphismes de T et S.

2.1 Mesure produit et disjonction

Définition 2.1. *Les systèmes dynamiques (X, μ, T) et (Y, ν, S) sont* disjoints *si la mesure produit $\mu \otimes \nu$ est l'unique couplage de T et S. On note dans ce cas $T \perp S$.*

On laisse en exercice le soin de vérifier l'équivalence de cette définition avec la propriété 2 (voir la formulation précise dans le texte).

Donnons maintenant l'exemple le plus élémentaire de disjonction : l'identité est disjointe de tout système dynamique ergodique.

Proposition 2.2. *Si T est la transformation identité sur X, et si S est un système dynamique ergodique, alors $T \perp S$.*

Démonstration. Soit λ un couplage de $T = \mathrm{Id}$ et de S. Pour tous boréliens $A \subset X$ et $B \subset Y$, l'invariance de λ par $T \times S$ donne

$$\begin{aligned}\forall n \geq 1, \quad \lambda(A \times B) &= \lambda(A \times S^{-n}B) \\ &= \frac{1}{n}\sum_{k=0}^{n-1} \lambda(A \times S^{-k}B) \\ &= \int_{A \times Y} \frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_B(S^k y)\, d\lambda(x,y).\end{aligned}$$

Mais comme S est ergodique, le théorème de Birkhoff montre que la convergence

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_B(S^k y) \xrightarrow[n\to\infty]{} \nu(B)$$

a lieu ν-presque partout, donc aussi λ-preque partout. Comme ces moyennes ergodiques sont toutes majorées par 1, le théorème de convergence dominée permet de passer à la la limite, et on obtient

$$\lambda(A \times B) = \mu(A)\nu(B),$$

d'où $\lambda = \mu \otimes \nu$. □

Une formulation alternative souvent utile du résultat précédent est donnée dans la proposition qui suit.

Proposition 2.3. *Soient T et S deux systèmes dynamiques, avec S ergodique, et soit $\lambda \in J(T,S)$. Si λ est invariante par* $\mathrm{Id} \times S$*, alors $\lambda = \mu \otimes \nu$.*

En fait, la disjonction entre l'identité et un système ergodique est un exemple d'une notion un peu plus forte : la disjonction spectrale.

Disjonction spectrale

Rappelons que le type spectral maximal d'un système dynamique (X, μ, T) est défini comme le type spectral maximal de l'action de l'opérateur de Koopman U_T sur $L^2_0(X,\mu)$ (sous-espace de $L^2(X,\mu)$ des fonctions d'intégrale nulle).

Définition 2.4. *Les systèmes T et S sont dits* spectralement disjoints *si leurs types spectraux maximaux respectifs sont mutuellement singuliers.*

Un exemple élémentaire est celui donné au paragraphe précédent : si T est l'identité, le type spectral maximal de T est la masse de Dirac en 0. Par ailleurs, le système S est ergodique si et seulement si son type spectral maximal ne charge pas $\{0\}$. Ainsi, l'identité est spectralement disjointe de tout système ergodique.

Exercice 2.1. *Les réels α et β sont dits* rationnellement indépendants *si une égalité du type $m\alpha + n\beta = 0 \mod 1$ où m et n sont des entiers entraîne $m = n = 0$. Prouver que si α et β sont rationnellement indépendants, alors les rotations d'angles respectifs $2\pi\alpha$ et $2\pi\beta$ sur le cercle sont spectralement disjointes.*

Exercice 2.2. *Prouver que tout système à spectre discret est spectralement disjoint de tout système faiblement mélangeant.*

Proposition 2.5. *Si T et S sont spectralement disjoints, alors ils sont disjoints.*

Démonstration. Soit λ un couplage des deux systèmes spectralement disjoints T et S. Prenons $A \subset X$ et $B \subset Y$ mesurables. On a $f := \mathbb{1}_A - \mu(A) \in L_0^2(X, \mu)$ et $g := \mathbb{1}_B - \nu(B) \in L_0^2(Y, \nu)$. Comme λ se projette sur μ et ν respectivement, dans le système couplé $(T \times S)_\lambda$, la mesure spectrale de la fonction $F : (x, y) \mapsto f(x)$ est celle de f, et la mesure spectrale de la fonction $G : (x, y) \mapsto g(y)$ est celle de g. Par hypothèse, ces deux mesures spectrales sont donc mutuellement singulières, et il s'en suit que les fonctions F et G sont orthogonales dans $L^2(X \times Y, \lambda)$. On a donc

$$\begin{aligned}
0 &= \int_{X \times Y} F(x, y)\, \overline{G(x, y)}\, d\lambda(x, y) \\
&= \int_{X \times Y} \Big(\mathbb{1}_A(x) - \mu(A)\Big)\Big(\mathbb{1}_B(x) - \nu(B)\Big)\, d\lambda(x, y) \\
&= \lambda(A \times B) - \mu(A)\nu(B),
\end{aligned}$$

et ceci prouve que λ est la mesure produit. □

Corollaire 2.6. *Si α et β sont deux nombres réels rationnellement indépendants, alors les rotations du cercle d'angle $2\pi\alpha$ et $2\pi\beta$ sont disjointes. Tout système à spectre discret est disjoint de tout système faiblement mélangeant.*

Attention, la réciproque de la proposition 2.5 est fausse : il existe des exemples de systèmes disjoints mais spectralement isomorphes.

2.2 Couplages portés par des graphes et isomorphismes

Supposons que S soit un facteur de T, et soit $\pi : X \to Y$ un homomorphisme. Alors π permet de construire un couplage très particulier de T et S. Considérons l'application

$$\Psi_\pi : x \in X \mapsto \big(x, \pi(x)\big) \in X \times Y,$$

et définissons la mesure de probabilité Δ_π sur $X \times Y$ par

$$\Delta_\pi := (\Psi_\pi)_* (\mu)$$

(mesure image de μ par Ψ_π). Autrement dit, pour tout ensemble mesurable C dans le produit cartésien $X \times Y$,

$$\Delta_\pi(C) = \mu \left\{x \in X : \big(x, \pi(x)\big) \in C\right\}).$$

En particulier, Δ_π est caractérisée par les valeurs qu'elle prend sur les pavés mesurables de $X \times Y$: pour tous boréliens $A \subset X$ et $B \subset Y$,

$$\Delta_\pi(A \times B) \ = \ \mu(A \cap \pi^{-1}B).$$

Exercice 2.3. *En utilisant les propriétés de l'homomorphisme π, vérifier que Δ_π est bien un couplage de T et S.*

Dans le produit cartésien $X \times Y$, on considère deux sous-σ-algèbres particulières :

- la σ-algèbre engendrée par la projection sur la première coordonnée, $\mathscr{F}_X := \{A \times Y : \ A \in \mathscr{B}(X)\}$,
- la σ-algèbre engendrée par la projection sur la seconde coordonnée, $\mathscr{F}_Y := \{X \times B : \ B \in \mathscr{B}(Y)\}$, engendrée par la projection sur la seconde coordonnée.

Lemme 2.7. *Le couplage Δ_π possède les deux propriétés suivantes :*

- $\Delta_\pi(G) = 1$, *où* $G := \left\{\big(x, \pi(x)\big), \ x \in X\right\}$ *est le graphe de* π *;*
- $\mathscr{F}_Y \subset \mathscr{F}_X \mod \Delta_\pi$, *où la notation « $\mathscr{C} \subset \mathscr{D} \mod \lambda$ » signifie que pour chaque C dans la σ-algèbre $\mathscr{C}$, il existe D dans la σ-algèbre $\mathscr{D}$ tel que $\lambda(C \vartriangle D) = 0$.*

Démonstration. Par définition de Δ_π, il est clair que cette mesure est portée par le graphe de π. Pour la seconde propriété, prenons $B \in \mathscr{B}(Y)$, et considérons $A := \pi^{-1}(B)$. On vérifie alors facilement que $\Delta_\pi(A \times Y \vartriangle B \times X) = 0$. $\square$

Il est assez remarquable que la réciproque du résultat précédent soit vraie, *i.e.* on peut caractériser le fait que S soit un facteur de T par l'existence d'un couplage particulier.

Proposition 2.8. *Soient T et S deux systèmes dynamiques. Supposons qu'il existe un couplage $\lambda \in J(T,S)$ tel que*

$$F_Y \subset \mathscr{F}_X \mod \lambda.$$

Alors S est un facteur de T, et il existe un homomorphisme $\pi : X \to Y$ tel que λ soit supportée par le graphe de π.

Démonstration. Commençons par désintégrer λ suivant la première coordonnée : pour μ-presque tout x, il existe une mesure de probabilité λ_x sur $(Y, \mathscr{B}(Y))$ telle que, pour tout $A \in \mathscr{B}(X)$ et tout $B \in \mathscr{B}(Y)$,

$$\lambda(A \times B) = \int_A \lambda_x(B)\, d\mu(x).$$

Considérons sur (Y, ν) une suite croissante $(\mathscr{P}_n)_{n\geq 0}$ de partitions finies qui sépare les points de Y : si $y_1 \neq y_2$, il existe n tel que y_1 et y_2 ne sont pas dans le même atome de $\mathscr{P}_n$. En particulier, comme (Y,ν) est un espace de Lebesgue, cela implique que la tribu $\bigvee_{n\geq 0} \mathscr{P}_n$ engendrée par cette suite de partitions coïncide avec $\mathscr{B}(Y)$ modulo ν. Par hypothèse sur λ, chaque partition $\mathscr{P}_n$ de Y peut être identifiée via λ à une partition $\overline{\mathscr{P}_n}$ de X, au sens où pour tout atome P de $\mathscr{P}_n$, il existe un atome $\overline{P}$ de $\overline{\mathscr{P}_n}$ avec $\lambda\Big((X \times P) \vartriangle (\overline{P} \times Y)\Big) = 0$. Cela implique que pour μ-presque tout $x \in \overline{P}$, la mesure conditionnelle λ_x est portée par P. On en déduit que pour μ-presque tout $x \in X$, pour chaque $n \geq 0$, il existe un atome $P_n(x)$ de la partition $\mathscr{P}_n$ tel que la mesure conditionnelle λ_x soit portée par $P_n(x)$. On a évidemment $P_{n+1}(x) \subset P_n(x)$. Comme la suite de partitions $(\mathscr{P}_n)_{n\geq 0}$ sépare les points de Y, $\bigcap_{n\geq 0} P_n(x)$ est soit vide, soit un singleton. Mais comme λ_x est μ-presque sûrement une mesure de probabilité, on en déduit que pour μ-presque tout x, $\bigcap_{n\geq 0} P_n(x)$ est un singleton, que l'on note $\pi(x)$. L'application $x \in X \mapsto \pi(x) \in Y$ est alors mesurable, car (à un négligeable près), si P est un atome de $\mathscr{P}_n$, $\pi^{-1}(P)$ est l'atome $\overline{P}$ de $\overline{\mathscr{P}_n}$ associé à P via λ. Ceci prouve également que $\pi_*(\mu) = \nu$, puisque $\mu(\overline{P}) = \nu(P)$ (cela suffit car la suite de partition $(\mathscr{P}_n)_{n\geq 0}$ engendre la tribu $\mathscr{B}(Y)$). Il est clair aussi que λ est supportée par le graphe de π car λ_x est portée par le singleton $\pi(x)$. Enfin, l'invariance de λ par $T \times S$ permet de montrer que l'on a μ-p.s. $\pi \circ T = S \circ \pi$. $\square$

Si on suppose en plus que T et S sont *isomorphes* (autrement dit, que l'homomorphisme π de X dans Y est inversible, et que son inverse π^{-1} est aussi un homomorphisme de systèmes dynamiques ; dans ce cas on utilisera plutôt le symbole ϕ au lieu de π), le couplage Δ_ϕ vérifie alors

$$\mathscr{F}_X = \mathscr{F}_Y \mod \Delta_\phi,$$

où la notation « $\mathscr{C} = \mathscr{D} \mod \lambda$ » signifie que l'on a à la fois $\mathscr{C} \subset \mathscr{D} \mod \lambda$ et $\mathscr{D} \subset \mathscr{C} \mod \lambda$. L'analogue de la proposition 2.8 est vrai également, et l'on peut résumer ces résultats dans le théorème suivant.

Théorème 2.9. *Soient (X, μ, T) et (Y, ν, S) deux systèmes dynamiques.*

S est un facteur de T si et seulement si il existe un couplage $\lambda \in J(T, S)$ tel que $\mathscr{F}_Y \subset \mathscr{F}_X \mod \lambda$; dans ce cas λ est supporté par le graphe d'un homomorphisme.

T et S sont isomorphes si et seulement si il existe un couplage $\lambda \in J(T, S)$ tel que $\mathscr{F}_Y = \mathscr{F}_X \mod \lambda$; dans ce cas λ est supporté par le graphe d'un isomorphisme.

Application : isomorphisme de systèmes à spectre discret

On va voir une jolie application du théorème 2.9 en donnant l'élégante preuve, due à Mariusz Lemańczyk (publiée dans [3]), d'un théorème originellement obtenu par Halmos et von Neumann.

Théorème 2.10. *Soient T et S deux systèmes ergodiques à spectre discret. S'ils sont spectralement isomorphes, alors ils sont isomorphes.*

Démonstration. Soit λ un couplage ergodique de T et S. Pour une valeur propre α de T, soit f un vecteur propre associé dans le système T, et soit g associé à α dans le système S. Alors les deux fonctions $(x, y) \mapsto f(x)$ et $(x, y) \mapsto g(y)$ sont des vecteurs propres dans le système $(T \times S)_\lambda$, associés à la valeur propre α. Mais $(T \times S)_\lambda$ est un système ergodique, donc α est valeur propre simple dans ce système. On en déduit que les fonctions $(x, y) \mapsto g(y)$ et $(x, y) \mapsto f(x)$ sont proportionnelles. Ceci implique que $(x, y) \mapsto f(x)$ appartient à $L^2(X \times Y, \mathscr{F}_Y, \lambda)$. Mais comme T est à spectre discret, les vecteurs propres engendrent un sous-espace linéaire dense dans $L^2(\mathscr{A}, \mu)$, et on obtient $L^2(X \times Y, \mathscr{F}_X, \lambda) \subset L^2(X \times Y, \mathscr{F}_Y, \lambda)$. Pour les mêmes raisons, l'inclusion réciproque est valide également, ce qui prouve que $\mathscr{F}_X = \mathscr{F}_Y \mod \lambda$, et les deux systèmes sont isomorphes. □

Remarque : la preuve ci-dessus montre aussi que, sous les hypothèses du théorème, les seuls couplages ergodiques entre T et S sont portés par des graphes d'isomorphismes.

Autocouplages, commutant et propriété des autocouplages minimaux

Si on étudie les autocouplages d'un système dynamique T, on est évidemment dans une situation où les systèmes à coupler entre eux sont isomorphes. Mais qu'est-ce qu'un isomorphisme entre un système T et lui-même ? C'est une application mesurable inversible de (X, μ) dans lui-même, qui préserve la mesure μ (autrement dit, c'est un *automorphisme* de (X, μ)), qui commute avec T. On appelle *commutant de* T, et on note $C(T)$, l'ensemble

$$C(T) \;:=\; \{S \in \mathrm{Aut}(X, \mu) : S \circ T = T \circ S\}.$$

Pour chaque $S \in C(T)$, notons Δ_S l'autocouplage de T porté par le graphe de S : il est caractérisé par

$$\forall A, B \in \mathscr{B}(X), \quad \Delta_S(A \times B) = \mu(A \cap S^{-1}B).$$

Cela conduit à identifier $C(T)$ avec un certain sous-ensemble de $J_2(T)$: l'ensemble des $\lambda \in J_2(T)$ qui vérifient

$$\mathscr{F}_1 = \mathscr{F}_2 \mod \lambda,$$

où la notation $\mathscr{F}_1$ (respectivement $\mathscr{F}_2$) désigne dans $X \times X$ la sous-σ-algèbre engendrée par la première (respectivement la seconde) coordonnée.

Observons que pour tout système T, le commutant de T contient au moins les puissances T^n, $n \in \mathbb{Z}$. De plus, chaque système $(T \times T)_{\Delta_S}$ pour $S \in C(T)$ est isomorphe à T (un isomorphisme est donné par $x \longmapsto (x, Sx)$). En particulier, $(T \times T)_{\Delta_S}$ est ergodique dès que T est ergodique. Cela donne pour chaque T ergodique une famille d'autocouplages ergodiques « évidents » : les Δ_{T^n}, $n \in \mathbb{Z}$. Dans le cas où T est faiblement mélangeant, la mesure produit $\mu \otimes \mu$ est aussi un autocouplage ergodique évident. En 1979, Rudolph a défini dans [2] la notion importante d'*autocouplages minimaux d'ordre 2*, qui est définie par l'absence d'autocouplages ergodiques en dehors de ces autocouplages évidents.

De nombreux exemples de systèmes ayant la propriété des autocouplages minimaux à l'ordre 2 sont connus. On peut citer notamment la transformation de Chacon.

Exercice 2.4. *Montrer que si T a la propriété des autocouplages minimaux à l'ordre 2, alors T n'a pas de* racine carrée, *c'est-à-dire il n'existe pas d'automorphisme S de (X, μ) tel que $S \circ S = T$.*

3 Couplages et facteurs

Dans cette section seront discutés les liens et les différences entre les deux propriétés proposées dans l'introduction pour définir, par analogie avec les entiers, la propriétés « les systèmes T et S sont premiers entre eux ». Le point de départ est l'existence d'un couplage particulier dès que T et S ont un facteur commun.

3.1 Couplage relativement indépendant au-dessus d'un facteur commun

Soient (X, μ, T) et (Y, ν, S) deux systèmes ayant comme facteur commun un troisième système (Z, ρ, R). Plus précisément, soit $\pi_X : X \to Z$ un homomorphisme entre T et R, et $\pi_Y : Y \to Z$ un homomorphisme entre S et R. On considère les couplages Δ_{π_X} sur $X \times Z$ et Δ_{π_Y} sur $Y \times Z$, portés respectivement par les graphes de π_X et π_Y. Puis on désintègre ces deux mesures de probabilités par rapport à Z : pour ρ-presque tout $z \in Z$, il existe une mesure de probabilité λ_z^X sur X et une mesure de probabilité λ_z^Y sur Y satisfaisant, pour tout $A \in \mathscr{B}(X)$, tout $B \in \mathscr{B}(Y)$ et tout $C \in \mathscr{B}(Z)$:

$$\Delta_{\pi_X}(A \times C) = \int_C \lambda_z^X(A)\, d\rho(z), \quad \text{et } \Delta_{\pi_Y}(B \times C) = \int_C \lambda_z^Y(B)\, d\rho(z).$$

On peut alors définir une mesure de probabilité, notée $\mu \otimes_Z \nu$ sur $X \times Y$ en posant, pour tout $D \subset X \times Y$ mesurable,

$$\mu \otimes_Z \nu(D) := \int_Z (\lambda_z^X \otimes \lambda_z^Y)(D)\, d\rho(z). \tag{1}$$

Exercice 3.1. *Vérifier que la mesure de probabilité $\mu \otimes_Z \nu$ est bien un couplage de S et T.*

Le système couplé défini par cette mesure de probabilité sur $X \times Y$ est appelé *produit relativement indépendant de T et S au-dessus de leur facteur commun R*, et noté $T \times_R S$. Il est important de noter que ce couplage dépend non seulement du facteur R mais aussi des homomorphismes π_X et π_Y choisis. Ce couplage possède une propriété remarquable énoncée dans la proposition suivante.

Proposition 3.1. *Avec les hypothèses et notations du paragraphe précédent, on a pour $\mu \otimes_Z \nu$-presque tout $(x, y) \in X \times Y$*

$$\pi_X(x) = \pi_Y(y).$$

Ainsi le couplage relativement indépendant identifie le facteur R vu dans chacun des deux systèmes.

Démonstration. On sait par construction de Δ_{π_X} que cette mesure est portée par le graphe de π_X. Donc pour ρ-presque tout z, λ_z^X est portée par l'ensemble des $x \in X$ vérifiant $\pi_X(x) = z$. De même, pour ρ-presque tout z, λ_z^Y est portée par l'ensemble des $y \in Y$ vérifiant $\pi_Y(y) = z$. Il s'en suit que pour ρ-presque tout z, la mesure produit $\lambda_z^X \otimes \lambda_z^Y$ est portée par l'ensemble des $(x, y) \in X \times Y$ tels que $\pi_X(x) = \pi_Y(y) = z$. Si $D = \{(x, y) \in X \times Y : \pi_X(x) = \pi_Y(y)\}$, il s'en suit par (1) que $\mu \otimes_Z \nu(D) = 1$. □

Corollaire 3.2. *Si R n'est pas le système trivial réduit à un seul point, la mesure $\mu \otimes_Z \nu$ est différente de la mesure produit.*

Ainsi, si T et S ont un facteur commun non réduit à un point, ces deux systèmes ne sont pas disjoints. En d'autres termes, la propriété 2 entraîne la propriété 1. La question de la réciproque n'est pas complètement évidente, mais on peut montrer qu'il existe des exemples de systèmes T et S sans facteur commun non trivial qui ne sont pas disjoints.

Il existe néanmoins un résultat important qui permet de déduire de la non-disjonction de deux systèmes une information sur leurs facteurs. C'est l'objet du paragraphe suivant

3.2 Lemme fondamental de non-disjonction et applications

Théorème 3.3. *Si T et S ne sont pas disjoints, alors S a un facteur commun non trivial avec un autocouplage d'ordre infini de T.*

Démonstration. Commençons avec un couplage $\lambda \in J(T, S)$. Évidemment, S est un facteur du système couplé $(T \times S)_\lambda$ via l'homomorphisme donné par la projection sur Y. Désintégrons la mesure de probabilité λ par rapport à Y : pour ν-presque tout $y \in Y$, il existe une mesure de probabilité λ_y sur X vérifiant, pour tout $A \in \mathscr{A}$ et tout $B \in \mathscr{B}(Y)$,

$$\lambda(A \times B) = \int_B \lambda_y(A)\, d\nu(y).$$

La construction qui suit consiste à prendre une famille dénombrable de copies du système couplé $(T \times S)_\lambda$ et à considérer leur couplage relativement indépendant au-dessus de leur facteur commun S. Comme ce couplage relativement indépendant identifie la projection sur Y dans chacun des système, on peut en fait le définir comme une probabilité λ_∞ sur le produit cartésien $Y \times X^{\times\mathbb{N}}$, en posant pour tout $B \in \mathscr{B}(Y)$ et tout C mesurable dans $X^{\mathbb{N}}$

$$\lambda_\infty(B \times C) := \int_B \lambda_y^{\otimes\mathbb{N}}(C) d\nu(y).$$

Autrement dit, la loi de $(y, x_0, x_1, x_2, \ldots) \in Y \times X^{\mathbb{N}}$ sous λ_∞ satisfait les propriétés suivantes :

- — la coordonnée y est distribuée suivant ν,
- — conditionnellement à y, les coordonnées x_k sont indépendantes et toutes distribuées suivant λ_y.

En particulier, pour chaque k le couple (y, x_k) est distribué suivant le couplage de départ λ. Il est clair que cette probabilité λ_∞ est invariante par la transformation

$$\sigma : (y, x_0, x_1, x_2, \ldots) \longmapsto (y, x_1, x_2, x_3 \ldots).$$

Soit $\mathscr{I}_\sigma$ la σ-algèbre des événements invariants par σ :

$$\mathscr{I}_\sigma := \left\{ I \in \mathscr{B}(Y) \otimes \mathscr{B}(X)^{\otimes\mathbb{N}} : I = \sigma^{-1} I \right\}.$$

Notons comme précédemment $\mathscr{F}_Y$ la σ-algèbre des événements mesurables par rapport à la coordonnée y :

$$\mathscr{F}_Y := \{B \times X^{\mathbb{N}} : B \in \mathscr{B}(Y)\}.$$

Puisque la coordonnée y reste inchangée par σ, on a clairement $\mathscr{F}_Y \subset \mathscr{I}_\sigma$. Inversement, soit $I \in \mathscr{I}_\sigma$ un événement σ-invariant, et pour tout $y \in Y$, considérons

$$I^y := \left\{ (x_0, x_1 \ldots) \in X^{\mathbb{N}} : (y, x_0, x_1 \ldots) \in I \right\}.$$

En notant σ_X le décalage des coordonnées sur $X^{\mathbb{N}}$, on vérifie aisément que pour tout $y \in Y$, I^y est un ensemble σ_X-invariant. Par la loi du 0–1 de Kolmogorov, on en déduit que pour ν-presque tout y,

$$\lambda_y^{\otimes\mathbb{N}}(I^y) \in \{0,1\}.$$

Posons

$$B := \left\{ y \in Y : \lambda_y^{\otimes\mathbb{N}}(I^y) = 1 \right\}.$$

On voit facilement que I coïncide modulo λ_∞ avec $B \times X^{\mathbb{N}}$, qui est un événement y-mesurable. Ainsi on a démontré que $\mathscr{I}_\sigma = \mathscr{F}_Y \mod \lambda_\infty$.

Soit maintenant $f : X \to \mathbb{R}$ une fonction mesurable bornée, et soit $F : Y \times X^{\mathbb{N}} \to \mathbb{R}$ définie par $F(y, x_0, x_1, \ldots) := f(x_0)$. Le théorème ergodique de Birkhoff appliqué à F pour la transformation σ donne, pour λ_∞-presque tout $(y, x_0, x_1, \ldots)$:

$$\begin{aligned}
\frac{1}{n}\sum_{k=0}^{n-1} f(x_k) &= \frac{1}{n}\sum_{k=0}^{n-1} F\big(\sigma^k(y, x_0, x_1, \ldots)\big) \\
&\xrightarrow[n\to\infty]{} \mathbb{E}_{\lambda_\infty}\left[F | \mathscr{I}_\sigma\right] \\
&= \mathbb{E}_{\lambda_\infty}\left[F | \mathscr{F}_Y\right] \\
&= \int_X f(x)\, d\lambda_y(x).
\end{aligned}$$

Ainsi, λ_∞ permet d'indentifier deux facteurs de $(S \times T^{\times\mathbb{N}})_{\lambda_\infty}$:

- celui engendré par $\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(x_k)$, qui est aussi, en notant λ_∞^X la marginale de λ_∞ sur $X^{\mathbb{N}}$, un facteur de $(T^{\times\mathbb{N}})_{\lambda_\infty^X}$,
- celui engendré par $\int_X f(x)\, d\lambda_y(x)$, qui est aussi un facteur de S.

Si on suppose que T et S ne sont pas disjoints, on peut prendre au départ un couplage λ qui n'est pas la mesure produit, puis on peut choisir f de telle façon que $\int_X f(x)\, d\lambda_y(x)$ ne soit pas une fonction ν-p.s. constante de y. Alors, le facteur de S engendré par $\int_X f(x)\, d\lambda_y(x)$ n'est pas le facteur trivial réduit un un seul point, et ceci achève la démonstration. □

3.2.1 Applications : exemples de disjonction

Comme application du théorème 3.3, on peut par exemple prouver la disjonction de deux classes de systèmes définies à l'aide de la notion d'entropie. L'argument repose sur la stabilité des systèmes d'entropie nulle par couplage et passage au facteur.

Lemme 3.4. *1. Soit T un système d'entropie nulle, et S un facteur de T. Alors S est d'entropie nulle.*

2. Soit $(X_i, \mu_i, T_i)_{i\in\mathbb{N}}$ une famille dénombrable de systèmes dynamiques vérifiant tous $h(T_i) = 0$, et soit $T = (T_0 \times T_1 \times T_2 \times \cdots)_\lambda$ un système obtenu comme un couplage des T_i. Alors $h(T) = 0$.

Démonstration. Pour la première partie du lemme, donnons-nous un homomorphisme $\pi : X \to Y$ entre T et S. Soit $\mathscr{P}$ une partition finie de Y, et $\overline{\mathscr{P}} := \pi^{-1}(\mathscr{P})$ la partition de X obtenue à partir de π. Alors pour tout $n \geq 0$, on a

$$\operatorname{dist}\left(\bigvee_{k=0}^{n} S^{-k}\mathscr{P}\right) = \operatorname{dist}\left(\bigvee_{k=0}^{n} T^{-k}\overline{\mathscr{P}}\right),$$

donc $h(\mathscr{P}, S) = h(\overline{\mathscr{P}}, T) = 0$.

Passons à la seconde partie du lemme. Pour tout $i \in \mathbb{N}$ et toute partition finie $\mathscr{P}_i$ de X_i, notons $\mathscr{P}_i^\times$ la partition de $X_0 \times X_1 \times \cdots$ dont les atomes sont de la forme

$$X_1 \times \cdots \times X_{i-1} \times P_i \times X_{i+1} \times X_{i+2} \times \cdots,$$

où P_i est un atome de $\mathscr{P}_i$. Pour toute partition finie $\mathscr{P}$ dans le produit $X_0 \times X_1 \times \cdots$ et tout $\varepsilon > 0$, il existe $n \geq 0$, des partitions finies $\mathscr{P}_0$ de X_0, $\mathscr{P}_1$ de X_1,…, $\mathscr{P}_n$ de X_n et une partition $\overline{\mathscr{P}}$ moins fine que $\mathscr{P}_1^\times \vee \mathscr{P}_2^\times \vee \cdots \vee \mathscr{P}_n^\times$ tels que $|\mathscr{P} - \overline{\mathscr{P}}| < \varepsilon$. On a alors

$$\begin{aligned} h(\overline{\mathscr{P}}, T) &\leq h(\mathscr{P}_1^\times \vee \mathscr{P}_2^\times \vee \cdots \vee \mathscr{P}_n^\times, T) \\ &\leq h(\mathscr{P}_1^\times, T) + \cdots + h(\mathscr{P}_n^\times, T) \\ &= h(\mathscr{P}_1, T_1) + \cdots + h(\mathscr{P}_n, T_n) = 0. \end{aligned}$$

On en déduit par continuité de l'entropie que $h(\mathscr{P}, T) = 0$. □

Rappelons que (Y, ν, S) est un K-système si, pour tout facteur non trivial T de S, on a $h(T) > 0$.

Théorème 3.5. *Si S est un K-système, et si $h(T) = 0$, alors $T \perp S$.*

Démonstration. Supposons que $h(T) = 0$, et que S ne soit pas disjoint de T. Par le théorème 3.3, S a un facteur commun non trivial avec un couplage d'une famille dénombrable de copies de T. Par le lemme 3.4, ce couplage d'une famille dénombrable de copies de T est d'entropie nulle, et donc son facteur commun avec S est lui aussi d'entropie nulle. Mais alors S a un facteur non trivial d'entropie nulle, et donc S n'est pas un K-système. □

Il est facile de voir que cet argument se généralise à toute situation où l'on remplace la propriété « être d'entropie nulle » par une propriété P stable par couplage dénombrable et passage au facteur. (Il faut alors aussi remplacer « K-système » par « sans facteur non trivial ayant la propriété P ».) Par exemple, prenons pour P la propriété « avoir spectre discret », et remarquons que les systèmes sans facteur non trivial à spectre discret sont précisément les systèmes faiblement mélangeants : on retrouve une nouvelle preuve de la disjonction entre les systèmes à spectre discret et ceux faiblement mélangeants.

Références

[1] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation*, Math. Systems Theory **1** (1967), 1–49.

[2] D.J. Rudolph, *An example of a measure preserving map with minimal self-joinings ans applications*, J. d'Analyse Math. **35** (1979), 97–122.

[3] J.P. Thouvenot, *Some properties and applications of joinings in ergodic theory*, Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), London Math. Soc. Lecture Note Ser., vol. 205, Cambridge Univ. Press, Cambridge, 1995, pp. 207–235.