

# Chapitre VI

## Développements

### Sommaire

---

<b>VI.1 Entiers <math>p</math>-adiques, espaces ultramétriques</b>	<b>113</b>
<b>VI.2 Dendrogrammes</b>	<b>121</b>
<b>VI.3 Introduction au transport optimal</b>	<b>123</b>
VI.3.1 Problème d'affectation et problème de Monge Kantorovich discret	123
VI.3.2 Transport optimal, cas général	126
<b>VI.4 Distance de Gromov-Wasserstein</b>	<b>127</b>
<b>VI.5 Propagation d'opinion et flot de gradient</b>	<b>129</b>
<b>VI.6 Modèles macroscopiques de trafic routier</b>	<b>134</b>
<b>VI.7 Autour de la notion de complétude, théorème de Banach-Steinhaus</b>	<b>137</b>
VI.7.1 Lemme de Baire	137
VI.7.2 Théorème de Banach Steinhaus	138
VI.7.3 Théorèmes de l'application ouverte, du graphe fermé	139

---

### VI.1 Entiers $p$ -adiques, espaces ultramétriques

Cette section est consacrée à la construction du complété de  $\mathbb{N}$  pour une certaine métrique, appelée  $p$ -adique. Nous détaillons cette construction pour le cas  $p = 2$ , qui aboutit à l'espace dit des entiers 2-adiques, noté  $\mathbb{Z}_2$ , dont on présente ci-dessous certaines propriétés.

#### Distance 2-adique sur $\mathbb{N}$

**Définition VI.1.1.** (Valuation et valeur absolue 2-adiques)

Tout  $a \in \mathbb{Z}$  non nul peut s'écrire de façon unique  $a'2^k$ , où  $a' \in \mathbb{Z}$  est impair. On appelle  $v_2(a) = k$  la *valuation* 2-adique de  $a$ , et l'on définit la valeur absolue 2-adique

$$|a|_2 = 2^{-v_2(a)}.$$

On pose  $v_2(0) = +\infty$ , et  $|0|_2 = 0$ .

On se place maintenant sur  $X = \mathbb{N}$  l'ensemble des entiers naturels, et l'on définit

$$(a, b) \in \mathbb{N}^2 \mapsto d(a, b) = |b - a|_2.$$

On peut donner une expression équivalente de cette définition, en considérant que tout entier naturel  $a$  s'écrit (écriture dyadique)

$$a = \sum_{n=0}^{+\infty} a_n 2^n, \quad a_n \in \{0, 1\}, \quad a_n \text{ nul au delà d'un certain rang.} \quad (\text{VI.1.1})$$

Deux nombres  $a$  et  $b$  peuvent ainsi être écrits en base 2 sous la forme de suites finies  $(a_n)$  et  $(b_n)$  de 0 et de 1, et l'on a  $d(a, b) = 2^{-n_{ab}}$ , où

$$n_{ab} = \min \{n, a_n \neq b_n\}, \quad (\text{VI.1.2})$$

pris égal à  $+\infty$  si les bits de  $a$  et  $b$  s'identifient (i.e. si  $a = b$ ). En effet, par définition de  $n_{ab}$ , on a

$$a - b = 2^{n_{ab}} c,$$

où  $c \in \mathbb{Z}$  est impair.

**Proposition VI.1.2.** L'application  $d(\cdot, \cdot)$  est une distance sur  $\mathbb{N}$ , qui est *ultramétrique*, c'est-à-dire qu'elle vérifie l'inégalité triangulaire renforcée :

$$d(a, b) \leq \max(d(a, c), d(b, c)) \quad \forall a, b, c \in X.$$

*Démonstration.* La séparation est immédiate par hypothèse, et l'on a bien  $d(a, b) = d(b, a)$ . Pour l'inégalité triangulaire renforcée, on utilise la formulation (VI.1.2). Pour tous entiers  $a, b, c$ , on considère les écritures dyadiques associées. Les  $n_{ab} - 1$  premiers bits de  $a$  et  $b$  s'identifient par définition, ainsi que les  $n_{bc} - 1$  premiers bits de  $b$  et  $c$ . Les premiers bits de  $a$  et  $c$  s'identifient donc au moins sur les  $\min(n_{ab}, n_{bc}) - 1$  indices, d'où

$$n_{ac} \geq \min(n_{ab}, n_{bc}),$$

et ainsi

$$d(a, c) = 2^{-n_{ac}} \leq \max(2^{-n_{ab}}, 2^{-n_{bc}}) = \max(d(a, b), d(b, c)),$$

qui est l'inégalité ultramétrique annoncée, qui entraîne l'inégalité triangulaire usuelle.  $\square$

Cette distance munit  $\mathbb{N}$  d'une distance, appelée *ultramétrique*, aux propriétés particulières (voir proposition VI.1.6 et les suivantes).

Le diamètre de  $(\mathbb{N}, d)$  est 1, et ce diamètre est atteint en de multiples couples : tout entier impair est en particulier diamétralement opposé à 0, ainsi qu'à tout entier pair. La sphère unité (sphère de centre 0 et de rayon 1) est l'ensemble des nombres impairs. De façon générale, la sphère de centre 0 et de rayon  $2^{-k}$  est l'ensemble des nombres du type  $2^k b$ , avec  $b$  impair. Ces sphères de centre 0 et de rayons 1,  $1/2$ ,  $1/4$ ,  $\dots$ , et 0, réalisent une partition de  $\mathbb{N}$  (comme un emboîtement infini de poupées russes).

La suite  $(2^n)$  tend vers 0, ainsi que toute suite telle que l'exposant de 2 dans la décomposition en facteurs premiers des termes tend vers  $+\infty$ .

Nous terminons cette section par une propriété négative sur  $(\mathbb{N}, d)$ , qui justifie la démarche de complétion qui suit.

**Proposition VI.1.3.** L'espace ultramétrique  $(\mathbb{N}, d)$  n'est pas complet.

*Démonstration.* Considérons la suite  $(2^n - 1)$ . Cette suite est de Cauchy pour  $d(\cdot, \cdot)$ . Si elle converge vers  $a \in \mathbb{N}$ , alors  $2^n$  converge vers  $a + 1$ , d'où  $a + 1 = 0$ , équation qui n'admet pas de solution dans  $\mathbb{N}$ .  $\square$

**Remarque VI.1.4.** Noter que l'on aurait pu choisir comme espace de départ  $X = \mathbb{Z}$  au lieu de  $\mathbb{N}$ , auquel cas le contre-exemple ci-dessus n'en est plus un, puisque  $a + 1$  admet une solution dans  $\mathbb{Z}$ . Pour se convaincre qu'il manque plus à  $(\mathbb{N}, d)$  que les nombres négatifs pour être complet, on peut considérer par exemple la suite

$$a_N = \sum_{n=0}^{2N} a_n 2^n,$$

avec  $a_n = n + 1 \bmod 2$  (alternance périodique de 1 et de 0). On a

$$a_N = \sum_{n=0}^N 2^{2n} = \frac{1 - 4^{N+1}}{1 - 4} = \frac{4^{N+1} - 1}{3}.$$

Cette suite est de Cauchy (comme toute série partielle de ce type). S'il existe  $a \in \mathbb{N}$  tel que  $a_N$  converge vers  $a$ , alors on a

$$\left| \frac{4^{n+1} - 1}{3} - a \right|_2 \rightarrow 0 \Rightarrow |4^{n+1} - 1 - 3a|_2 = 0,$$

d'où  $3a + 1 = 0$ , équation qui n'admet pas de solution dans  $\mathbb{N}$ , ni dans  $\mathbb{Z}$ .

### Complété de $\mathbb{N}$

**Définition VI.1.5.** Le complété de  $\mathbb{N}$  (voir théorème A.2.4) pour la distance définie ci-dessus est appelé ensemble des entiers 2-adiques. Il est noté  $\mathbb{Z}_2$ .

L'écriture (VI.1.1), qui permet de représenter tout entier comme une suite finie de 0 ou de 1 (écriture en base 2), permet de se faire une meilleure idée de  $\mathbb{Z}_2$ , et de préciser à quoi correspondent certains de ses éléments. Considérons une suite  $(a^k)$  dans  $\mathbb{N}$ , de Cauchy pour  $d(\cdot, \cdot)$ . On peut écrire les termes

$$a^k = (a_0^k, a_1^k, a_2^k, \dots)$$

Si une suite est de Cauchy alors, comme dans le cas des réels en écriture décimale (voir la démonstration de la proposition A.1.39), chacune des suites  $(a_n^k)_k$  finit par se stabiliser à une valeur  $a_n$ . On peut donc représenter la limite par une suite (infinie) de 0 et de 1. On écrira le nombre correspondant somme d'une série, ou en écriture flottante, avec la convention de placer les bits *avant* la virgule :

$$a = \sum_{n=0}^{+\infty} a_n 2^n, \quad \text{ou} \quad a = \dots a_3 a_2 a_1 a_0, 0.$$

Considérons par exemple le nombre  $\dots 111, 0$ . On a

$$\dots 111, 0 = \lim_{N \rightarrow \infty} \left( \sum_{n=0}^N 2^n \right) = \lim_{N \rightarrow \infty} \left( \frac{1 - 2^{N+1}}{1 - 2} \right) = -1 - \lim_{N \rightarrow \infty} 2^{N+1}$$

qui tend donc selon les règles de calcul usuel vers un nombre qui n'est pas dans l'espace de départ, et qu'il est tentant de noter  $-1$ .

La métrique induite sur  $\mathbb{Z}_2$  est définie essentiellement de la même manière que sur  $\mathbb{N}$ . Pour deux éléments  $a$  et  $b$  de  $\mathbb{Z}_2$ , si l'on note  $n_{ab}$  le plus petit indice pour lequel les bits diffèrent ( $n_{ab} = +\infty$  si  $a = b$ ), la distance entre  $a$  et  $b$  est simplement définie par  $d(a, b) = 2^{-n_{ab}}$ .

La distance sur  $\mathbb{Z}_2$  définie ci-dessus hérite des propriétés ultramétriques de la distance de départ sur  $\mathbb{N}$ . L'espace  $\mathbb{Z}_2$  possède donc des propriétés propres aux espaces ultramétrique, telles que celles énoncées ci-après.

### Espaces ultramétriques généraux

**Proposition VI.1.6.** Soit  $(X, d)$  un espace ultramétrique. Tout point d'une boule (ouverte ou fermée) est centre de cette boule.

*Démonstration.* Soient  $x \in X$  et  $r \in \mathbb{R}_+$ . On considère  $x' \in B_f(x, r)$ , i.e. tel que  $d(x, x') \leq r$ . Pour tout  $y \in B_f(x', r)(x, r)$ , on a

$$d(y, x') \leq \max(d(y, x), d(x, x')) \leq r, \quad \text{d'où} \quad y \in B_f(x', r).$$

On a donc  $B_f(x, r) \subset B_f(x', r)$ , et on a l'inclusion inverse en intervertissant les rôles de  $x$  et  $x'$ . La démonstration est identique pour une boule ouverte.  $\square$

**Corollaire VI.1.7.** Soit  $(X, d)$  un espace ultramétrique. L'intersection entre deux boules est soit vide soit l'une des deux boules. Si les rayons sont les mêmes, deux boules sont donc soit disjointes, soit identiques.

**Démonstration:** Soit  $B(x, r)$  et  $B(x', r')$ . Si  $y$  appartient aux deux boules, il est aussi centre de ces deux boules d'après ce qui précède, la plus petite est donc incluse dans la plus grande, avec égalité si les rayons sont les mêmes.

**Proposition VI.1.8.** Soit  $(X, d)$  un espace ultramétrique. Toute boule ouverte est un fermé.

*Démonstration.* Si  $r = 0$  (alors  $B(x, r) = \emptyset$ ) ou si  $B(x, r) = X$ , c'est immédiat. Sinon, pour tout  $y \in B(x, r)^c$ ,  $B(x, r) \cap B(y, r) = \emptyset$  d'après la proposition précédente, d'où  $B(y, x) \subset B(x, r)^c$ . Le complémentaire de  $B(x, r)$  est donc un ouvert, la boule ouverte elle-même est donc un fermé.  $\square$

**Proposition VI.1.9.** Un espace ultramétrique est totalement discontinu, au sens où chaque point s'identifie à sa propre composante connexe (définition A.2.17).

*Démonstration.* Soient  $x$  et  $y \neq x$  dans  $X$ , et  $0 < r < d(x, y)$ . La boule ouverte  $B(x, r)$  et son complémentaire (qui est aussi un ouvert) réalisent une partition de  $X$  en 2 ouverts,  $x$  et  $y$  ne peuvent donc appartenir à la même composante connexe.  $\square$

**Proposition VI.1.10.** Soit  $(X, d)$  un espace ultramétrique. Tout triangle dans  $X$  est isocèle (avec deux grands côtés et un petit côté).

**Proposition VI.1.11.** On considère trois points dans  $X$ , et l'on note  $\ell_1, \ell_2$ , et  $\ell_3$  les longueurs des côtés. On a  $\ell_1 \leq \max(\ell_2, \ell_3)$ , et les mêmes propriétés obtenues en permutant les indices. Si les longueurs sont distinctes deux à deux, on a par exemple  $\ell_1 < \ell_2 < \ell_3$ , qui invalide  $\ell_3 \leq \max(\ell_1, \ell_2)$ . Deux des longueurs au moins sont donc identiques, par exemple  $\ell_1 = \ell_2$ . Le troisième côté est de longueur  $\ell_3 \leq \max(\ell_1, \ell_2) = \ell_1$ .

**Proposition VI.1.12.** Soit  $(X, d)$  un espace ultramétrique. Une suite  $(x_n)$  est de Cauchy si et seulement si

$$\lim_{n \rightarrow +\infty} d(x_{n+1}, x_n) = 0.$$

**Exemple VI.1.1.** On se reportera à la section VI.2 pour la description d'un procédé qui permet de construire un espace ultramétrique à partir d'un espace métrique fini, par une démarche de *clustering* très utilisée en analyse de données.

## Retour sur l'espace des entiers 2-adiques

**Addition sur  $\mathbb{Z}_2$ .** On peut définir une addition sur  $\mathbb{Z}_2$ , qui est "l'addition de l'écolier" en partant de la droite dans l'écriture 2-adique ci-dessus. On considère deux entiers dyadiques  $a$  et  $b$ , et l'on cherche à construire la somme  $c = a + b$ , qui étende la somme sur  $\mathbb{N}$ . Si  $a_0 + b_0 = 0$  ou 1, on affecte à  $c_0$  cette valeur, et on passe au rang suivant. Si la somme vaut 2 on pose  $c_0 = 0$ , et l'on garde une retenue de 1 pour le rang suivant. Au rang 1 on se retrouve dans la même situation, avec éventuellement une retenue de 1 en plus. La somme peut donc maintenant valoir 3. Si c'est le cas on pose  $c_1 = 1$  et on garde 1 de retenue pour le rang suivant, etc ...

De façon assez frappante, le fait de prendre le complété de  $\mathbb{N}$  pour la distance choisie, qui est une démarche purement topologique, conduit à un espace qui possède une structure algébrique que l'espace de départ n'avait pas.

**Proposition VI.1.13.** L'espace  $\mathbb{Z}_2$  muni de l'addition définie ci-dessus est un groupe additif.

**Proposition VI.1.14.** L'espace  $\mathbb{Z}_2$  n'est pas dénombrable.

*Démonstration.* Nous avons vu que tout élément de  $\mathbb{Z}_2$  peut se représenter de façon unique par une suite infinie de 0 ou de 1. On peut donc identifier (d'un point de vue ensembliste)  $\mathbb{Z}_2$  à l'ensemble des parties de

$\mathbb{N}$ , en considérant la suite  $(a_n)$  comme la fonction indicatrice d'une partie de  $\mathbb{N}$ . L'ensemble  $\mathbb{Z}_2$  n'est donc pas dénombrable d'après le théorème A.1.10.  $\square$

**Définition VI.1.15.** (Ordre lexicographique)

On peut définir sur  $\mathbb{Z}_2$  un ordre *lexicographique*, en considérant, pour deux éléments différents  $\dots a_2 a_1 a_0, 0$  et  $\dots b_2 b_1 b_0, 0$  le plus petit indice  $n$  pour lequel les deux diffèrent, et poser  $a < b$  si  $a_n < b_n$ .

Noter que l'ordre défini ci-dessus est très différent de l'ordre usuel. Comme  $\mathbb{Z} \subset \mathbb{Z}_2$ , on peut comparer de ce point de vue deux éléments de  $\mathbb{Z}$ , on retrouve certaines propriétés usuelles du type  $1 < 3$ ,  $1 < 5$ , mais aussi des choses plus déroutantes, comme  $5 < 3$  et, plus globalement,

$$0 \leq a \leq -1 \quad \forall a \in \mathbb{Z},$$

ce qui permet d'écrire

$$\mathbb{Z} = [0, -1].$$

**Proposition VI.1.16.** L'espace  $\mathbb{Z}_2$  est compact.

*Démonstration.* Considérons une suite  $(a^k)$  dans  $\mathbb{Z}_2$ . On considère la suite  $(a_0^k)_k$  de 0 et de 1. Cette suite visite une infinité de fois 0 ou 1 (ou les deux). On prend  $a_0$  égal à une valeur visitée une infinité de fois, et l'on extrait la sous-suite  $(a^{\varphi_0(k)})$  correspondante. On procède de même avec  $(a_1^{\varphi_0(k)})$ , pour extraire une suite  $(a^{\varphi_0 \circ \varphi_1(k)})$ . On extrait ainsi des sous-suites emboîtées les unes dans les autres. On définit maintenant (selon le processus d'extraction diagonale dit de Cantor)

$$\varphi(k) = \varphi_0 \circ \varphi_1 \circ \dots \circ \varphi_k(k).$$

La suite ainsi construite est telle que  $a_n^{\varphi(k)}$  a une valeur constante  $a_n \in \{0, 1\}$  au-delà d'un certain rang (pour tout  $k \geq n$ ), on a donc convergence de cette suite extraite vers  $a = \dots a_3 a_2 a_1 a_0, 0$ .  $\square$

### Système projectif, limite projective (●●●)

Un *système projectif* désigne une famille  $(X_n)$  d'ensemble muni d'une famille d'applications  $(f_n^m)_{n \leq m}$ , avec  $f_n^m : X_m \rightarrow X_n$ , qui vérifient les propriétés suivantes

- (1) L'application  $f_n^n$  est l'identité sur  $X_n$
- (2) Pour tous  $n \leq m \leq q$ , on a  $f_n^m \circ f_m^q = f_n^q$ .

On appelle limite projective l'ensemble des éléments du produit infini  $X_0 \times X_1 \times \dots$  dont les projections sont compatibles avec les  $f_i^j$  au sens suivant :

$$\varprojlim \left( (X_n), (f_i^j) \right) = \{ x = (x_0, x_1, x_2, \dots) \in X_0 \times X_1 \times \dots, f_n^m(x_m) = x_n \quad \forall i \leq j \}.$$

Pour tous  $n \leq m$ , on peut définir canoniquement<sup>1</sup> une surjection de  $\mathbb{Z}/2^m\mathbb{Z}$  dans  $\mathbb{Z}/2^n\mathbb{Z}$ . On note  $f_n^m$  cette surjection. On peut identifier ainsi  $\mathbb{Z}_2$  à la limite projective du système  $(\mathbb{Z}/2^n\mathbb{Z}, (f_n^m))$ .

### Représentation des arbres dyadiques, application au poumon humain

Le poumon humain se présente comme un arbre constitués de bronches (appelée bronchioles pour les plus petites), structuré de façon dyadique : la trachée se divise en deux, chacune des branches filles se divise elle-même en deux, etc... Pour l'arbre respiratoire d'un adulte, le nombre de bifurcations est de l'ordre de  $2^3$ , soit autour de  $2^{23} \approx 8 \times 10^8$  bronchioles terminales, dont on appelle *feuilles* les extrémités libres. Pour diverses raison (construction d'un modèle homogénéisé du parenchyme, construction d'un opérateur de la ventilation, qui à un champ de pression aux feuilles associe un champ de flux, etc ...), il peut être intéressant

1. On peut considérer la relation d'équivalence sur  $\mathbb{Z}/2^m\mathbb{Z}$  définie par  $z \mathcal{R} z' \iff z \equiv z' [2^n]$ . L'espace quotient s'identifie à  $\mathbb{Z}/2^n\mathbb{Z}$ .

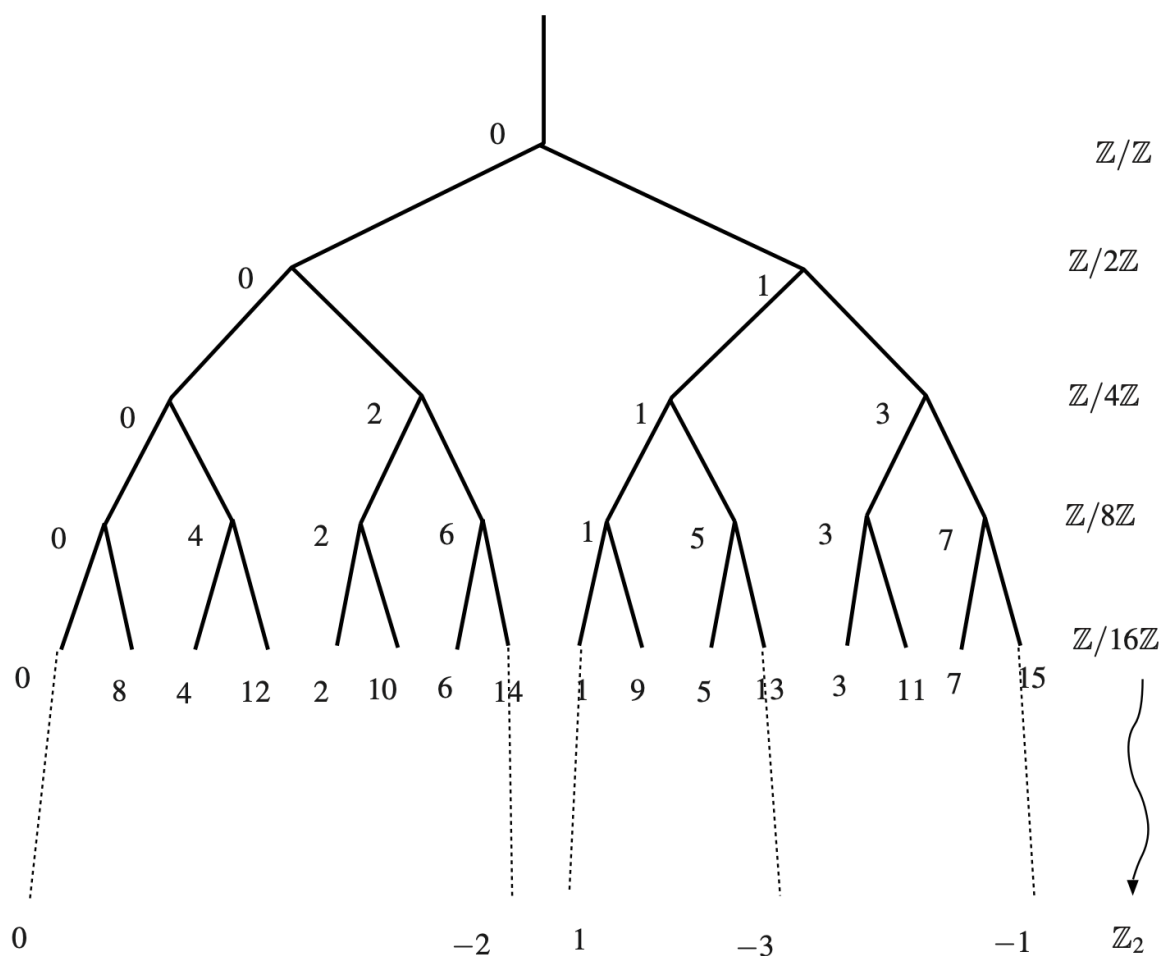


FIGURE VI.1.1 – Numérotation 2-adique

*d'extrapoler* cet arbre vers un nombre de générations infini. La notion de feuille est alors remplacée par celle de *bout*, on parle de *l'espace des bouts*, chacun de ces bouts correspondant à un chemin centrifuge s'éloignant de la racine (entrée de la trachée). Il est naturel de coder chacun de ces chemins par une suite infinie de 0 et de 1, par exemple selon la convention par exemple (on se figurera une représentation avec la racine en haut, comme sur la figure VI.1.1) : partant de la racine, si l'on part à gauche, on prend  $a_0 = 0$ , et  $a_0 = 1$  si c'est à droite. On encode le "choix" à chaque étape par un nombre entre 0 et 1.

La figure VI.1.1 représente la numérotation obtenue à chaque génération, où l'on a représenté chaque mot  $(a_N \dots, a_0)$  par sa représentation entière  $a = \sum a_n$ .

On notera que cette numérotation ne correspond pas du tout à l'énumération linéaire 0, 1, ... Cette numérotation présente un énorme avantage par rapport à l'énumération linéaire, en cela qu'elle respect la structure de l'arbre. Plus précisément, considérons deux feuilles de la 4-ième génération, d'indices  $a$  et  $b$ . Leur proximité *vis à vis de l'arbre*, qui correspondrait à un degré de parenté s'il s'agissait d'un arbre généalogique, ne dépend que de  $|a - b|$ , contrairement à ce qui se passerait pour la numérotation linéaire, comme on peut s'en convaincre facilement. Par exemple deux feuilles dont la différence d'indices est impaire appartiennent nécessairement à des lobes différents. Si la différence est divisible par 2, et pas par 4, les deux point appartiennent nécessairement tous deux à l'un des 4 sous-arbres issus de la génération 2, etc...

Plus précisément, si l'on considère l'arbre à 4 générations représenté sur la figure, qui présente  $2^4 = 16$  feuilles, on peut identifier la métrique dyadique sur l'ensemble des feuilles (identifié à  $\llbracket 0, 15 \rrbracket$  ou  $\mathbb{Z}/2^4\mathbb{Z}$ ) avec la métrique du plus court chemin au travers de l'arbre considéré comme un graphe pondéré. Pour retrouver

exactement la distance 2-adique, on peut vérifier qu'il suffit de considérer que les arêtes des deux dernières générations (3 et 4) ont pour longueur  $1/16$ , longueur  $1/8$  pour la génération 2 (4 branches), et  $1/4$  pour la génération 1.

Si l'on considère maintenant le poumon infini, on voit apparaître des indices déjà identifiés pour certains bouts. Par exemple le bout le plus à gauche est clairement 0. Pour le bout correspondant au chemin le plus à droite, on retrouve notre  $\dots 1111, 0 = -1$ . Noter que  $-1$  est le plus "grand" élément de  $\mathbb{Z}_2$  pour la relation d'ordre lexicographique (voir définition VI.1.15), et 0 le plus petit. De façon générale, la représentation de la figure VI.1.1 correspond à cet ordre (croissant de la gauche vers la droite).

On notera que, pour tout sous-arbre infini, le bout le plus à gauche est un entier positif, et le bout le plus à droite un entier négatif.

### Nombres entiers $p$ -adiques

On peut généraliser cette approche en remplaçant 2 par un nombre  $p$  premier quelconque : tout  $a \in \mathbb{Z}$  non nul peut s'écrire de façon unique  $a'p^k$ , où  $a' \in \mathbb{Z}$  n'est pas divisible par  $p$ . On appelle  $v_p(a) = k$  la *valuation*  $p$ -adique de  $a$ , et l'on définit la valeur absolue  $p$ -adique

$$|a|_p = p^{-v_p(a)}.$$

On pose  $v_p(0) = +\infty$ , et  $|0|_p = 0$ . Cette valeur absolue vérifie  $|ab|_p = |a|_p |b|_p$ .

On peut définir de la même manière que précédemment une distance  $d_p$  sur  $\mathbb{N}$ , et considérer le complété de  $\mathbb{N}$  pour cette distance, que l'on note  $\mathbb{Z}_p$ .

L'identification des éléments de  $\mathbb{Z}_p$  peut se faire à partir de la décomposition d'un entier en base  $p$  :

$$a = \sum_{n=0}^{+\infty} a_n p^n, \quad a_n \in [0, p-1] \quad \forall n, \quad a_n = 0 \text{ au-delà d'un certain rang.}$$

Un élément de  $\mathbb{Z}_p$  peut ainsi s'écrire comme une série infinie du type de celle ci-dessus, ou simplement codée par ses coefficients (on garde la convention d'une infinité de chiffres *avant* la virgule)

$$a = \dots a_2 a_1 a_0, 0.$$

Noter que cette complétion et l'identification à des nombres écrits en base  $p$  avec une infinité de chiffres avant la virgule ne nécessite pas que  $p$  soit premier. Si  $p$  n'est pas premier, on perd la propriété  $|ab|_p = |a|_p |b|_p$ , par exemple  $|2 \times 5|_{10} = 10^{-1}$  alors que  $|2|_{10} = |5|_{10} = 1$ . Mais tant que l'on s'en tient à des aspects métriques, et que l'on se contente d'additionner les nombres entre eux, la construction est valide. On peut en particulier construire l'objet  $\mathbb{Z}_{10}$  des entiers 10-adiques, qui présente une forte analogie apparente (le codage semble le même, à symétrie près) avec l'ensemble des réels de l'intervalle  $[0, 1]$ , mais qui présente des propriétés très différentes. Pour l'anecdote, si l'on considère la relation d'ordre lexicographique sur  $\mathbb{Z}_{10}$ , son plus grand nombre est (on notera que l'infinité de 9 avant la virgule n'est pas ici pathologique, le codage de  $\mathbb{Z}_{10}$  est tout à fait injectif, contrairement au codage décimal des réels)

$$\dots 9999, 0 = \sum_{n=0}^{+\infty} 9 \times 10^n = 9 \frac{1}{1-10} = -1,$$

de telle sorte que l'on peut écrire  $\mathbb{Z}_{10} = [0, -1]$  (*sic*).

On se restreint néanmoins en général aux nombres premiers, qui permettent des développements très féconds dans un cadre algébrique. Si l'on se restreint ainsi aux nombres premiers, on a une formule à la fois spectaculaire et très simple à établir, qui relie toutes les valeurs absolue  $p$ -adiques entre elles. Plus précisément, si l'on note  $|\cdot|_\infty$  la valeur absolue usuelle sur  $\mathbb{Z}$ , on a

$$\forall a \in \mathbb{Z}, \quad |a|_\infty \prod_{p \text{ premier}} |a|_p = 1.$$

On notera que pour  $p$  assez grand (notamment plus grand que  $|a|$ ), on a  $|a|_p = 1$ , le produit ci-dessus peut donc en fait se ramener à un produit fini, dont la définition ne nécessite donc pas d'arguments topologiques.

## Mesure sur $\mathbb{Z}_2$ , sur $\mathbb{Z}_p$

Nous décrivons dans cette section la démarche permettant de construire une mesure sur  $\mathbb{Z}_2$ . Le rôle joué par les intervalles dans le cas réel est ici joué par les boules. Nous privilégierons les boules fermées  $B_f(a, 2^{-k})$ , en gardant en tête que ce sont aussi des boules ouvertes. En effet les valeurs prises par la distance étant quantifiées (ce sont les  $2^{-k}$ ), on a  $B(a, 2^{-k}) = B_f(a, 2^{-(k-1)})$ . On notera que la boule fermée de centre  $a$  et de rayon  $2^{-k}$  s'écrit aussi

$$a + 2^k \mathbb{Z}_2 = \{a + 2^k z, z \in \mathbb{Z}_2\} = \{\dots a_{k-1} a_{k-2} \dots a_0\}$$

c'est-à-dire l'ensemble des éléments de  $\mathbb{Z}_2$  dont l'écriture 2-adique commence comme celle de  $a$  (au moins jusqu'au rang  $k-1$ ).

Nous noterons  $\mathcal{B} = \mathcal{B}(\mathbb{Z}_2)$  la tribu des boréliens sur  $\mathbb{Z}_2$ , i.e. la tribu engendrée par les ouverts de  $\mathbb{Z}_2$ .

**Proposition VI.1.17.** La tribu  $\mathcal{B}$  des boréliens est engendrée par les  $a + 2^k \mathbb{Z}_2$ , avec  $a$  et  $k$  parcourant  $\mathbb{N}$ .

*Démonstration.* Montrons en premier lieu que tout ouvert est réunion dénombrable de boules (fermées!). Soit un ouvert  $U$  de  $\mathbb{Z}_2$ , et  $a \in U$ . Il existe  $k$  tel que  $a + 2^k \mathbb{Z}_2 \subset U$ . Notons  $\bar{a} \in \mathbb{N}$  la troncature de  $a$  au rang  $k$ , i.e.

$$\bar{a} = \sum_{n=0}^{k-1} a_n 2^n.$$

On a  $a + 2^k \mathbb{Z}_2 = \bar{a} + 2^k \mathbb{Z}_2$ . L'ouvert  $U$  est donc union de boules du type  $a + 2^k \mathbb{Z}_2$ , il s'agit donc d'une union dénombrable<sup>2</sup>. Cette propriété implique que la tribu engendrée par les  $a + 2^k \mathbb{Z}_2$ , contient la tribu des boréliens, et donc s'identifie à elle du fait qu'il s'agit d'ouverts.  $\square$

On cherche maintenant à définir une mesure  $\mu$  sur  $\mathbb{Z}_2$ , qui affecte une masse 1 à  $\mathbb{Z}_2$ . Nous verrons qu'il est possible de définir une telle mesure sur la tribu borélienne. Du fait que  $\mathbb{Z}_2$  est union disjointe de  $2^k$  boules fermés de rayon  $2^{-k}$ , on affecte le volume  $2^{-k}$  à toute boule fermée de rayon  $2^{-k}$ . Il peut sembler étonnant d'affecter un volume égal au *rayon*, mais on se souviendra que dans ce contexte ultramétrique, le rayon est égal au diamètre.

On se propose maintenant de construire, à partir de cette définition du volume sur le  $\pi$ -système des  $a + 2^k \mathbb{Z}_2$ , une mesure extérieure sur  $\mathbb{Z}_2$ , en suivant la démarche décrite sur  $\mathbb{R}$  dans la proposition B.4.6, page 180.

**Proposition VI.1.18.** Pour tout  $A \subset \mathbb{Z}_2$ , on note  $C_A$  l'ensemble des suites de boules fermées dont l'union recouvre  $A$  :

$$C_A = \left\{ (a_i + 2^{k_i} \mathbb{Z}_2)_{i \in \mathbb{N}}, A \subset \bigcup_{\mathbb{N}} (a_i + 2^{k_i} \mathbb{Z}_2) \right\}.$$

On autorise les rayons à être nul (i.e.  $k = +\infty$ ), ce qui autorise à considérer des collections avec un nombre fini de boules de volume non nul. On définit alors  $\mu^* : \mathcal{P}(\mathbb{Z}_2) \rightarrow [0, +\infty]$  par

$$\lambda^*(A) = \inf_{C_A} \left( \sum_i 2^{-k_i} \right). \quad (\text{VI.1.3})$$

Cette application est une mesure extérieure, et elle attribue à toute boule fermée de rayon  $2^{-k}$  la valeur  $2^{-k}$ .

*Démonstration.* Cette démonstration est parfaitement analogue<sup>3</sup> à celle de la proposition B.4.6, page 180. En premier lieu  $\mu^*(\emptyset) = 0$ . Pour la monotonie,

2. Noter, même si ça n'est pas nécessaire dans la démonstration, qu'on peut ne garder qu'une seule boule par  $a \in \mathbb{N}$  impliqué, du fait que deux boules sont soit disjointes soit concentriques (l'une dans l'autre). On peut donc "coder" un ouvert de  $\mathbb{Z}_2$  par une suite  $(a_i, k_i)_i$  dans  $\mathbb{N}^2$ .

3. Elle s'applique d'ailleurs à toute construction d'une mesure extérieure par l'extérieur précisément, basée sur ce principe de recouvrement par des objets donc on a fixé la taille.



Considérons maintenant une collection  $(A_n)$  de parties de  $\mathbb{Z}_2$ . On peut trouver pour chaque partie une collection de boules qui la recouvre, et telle que la somme des volumes approche  $\mu^*(A_n)$  à  $\varepsilon/2^n$  près. Comme dans la démonstration de la proposition B.4.6, page 180, cela permet d'établir que

$$\mu^*\left(\bigcup A_n\right) \leq \sum \mu^*(A_n) + 2\varepsilon,$$

et ce pour tout  $\varepsilon$ . On en déduit la sous-additivité.

Montrons maintenant que cette mesure extérieure affecte  $2^{-k}$  aux boules fermées de rayon  $2^{-k}$ . Soit  $B = a + 2^k\mathbb{Z}_2$  une telle boule. En premier lieu, la boule est recouverte par elle-même, d'où  $\mu^*(B) \leq 2^{-k}$ . Maintenant considérons un recouvrement de  $B$  par des boules. Ces boules fermées étant des ouverts, et  $B$  étant compacte, on peut en extraire un recouvrement fini, on peut enlever les boules qui sont incluses dans une autre du recouvrement, pour obtenir (d'après le corollaire VI.1.7) une partition finie, telle que la somme des volumes est exactement  $2^{-k}$ .  $\square$

On considère maintenant  $\mathcal{A}$  la tribu des parties mesurables pour  $\mu^*$  (selon la définition B.4.2, page 177), et l'on note  $\mu$  la mesure définie sur  $\mathcal{A}$  issue de  $\mu^*$ , selon le théorème B.4.4, page 178. Il s'agit maintenant de montrer que la tribu  $\mathcal{A}$  contient la tribu des boréliens. Comme dans le cas réel (voir proposition B.4.8, page 181), il suffit de vérifier que les boules fermées, qui engendrent  $\mathcal{A}$ , sont mesurables.

**Proposition VI.1.19.** La tribu  $\mathcal{A}$  des parties mesurables pour  $\mu^*$  contient les boules fermées, donc les boréliens. La mesure  $\mu$  introduite ci-dessus est ainsi définie sur la tribu des boréliens.

*Démonstration.* Soit  $B = a + 2^k\mathbb{Z}_2$  une boule fermée. Il s'agit de montrer que pour toute partie  $A \subset \mathbb{Z}_2$ ,

$$\mu^*(A) \geq \mu^*(A \cap B) + \mu^*(A \cap B^c).$$

On considère un recouvrement de  $A$  par des boules fermées  $A_n = a_n + 2^{k_n}\mathbb{Z}_2$ , avec

$$\sum 2^{-k_n} \leq \mu^*(A) + \varepsilon.$$

Il s'agit de distribuer au mieux chacune de ces boules entre  $B$  et  $B^c$ . Là encore, la preuve est plus simple que dans le cas réel. Soit une telle boule  $A_n$ . Si elle ne rencontre pas  $B$ , alors elle est incluse dans  $B^c$ , auquel cas on l'affecte à  $B^c$ . Si elle rencontre  $B$ , alors soit elle est incluse dans  $B$ , auquel cas on "l'affecte à  $B$ " soit, du fait de l'ultramétrie, elle contient  $B$ . Son rayon est  $2^{-k_n}$ , avec  $k_n < k$ . La boule  $A_n$  peut alors s'écrire, comme toute boule fermée, comme réunion disjointe de  $2^{k-k_n}$  boules fermées de rayon  $2^{-k_n}$  (l'une d'elle étant  $B$ ). On atomise donc la boule  $A_n$  en  $2^{k-k_n}$  boules plus petites, on affecte la boule qui s'identifie à  $B$  à  $B$ , et l'on affecte les  $2^{k-k_n} - 1$  autres à  $B^c$ , ce qui se fait sans augmenter la masse totale. On construit ainsi à partir du recouvrement de  $A$  deux recouvrements de  $A \cap B$  et  $A \cap B^c$ , respectivement, sans changer la masse totale, d'où l'on déduit que

$$\mu^*(A \cap B) + \mu^*(A \cap B^c) \leq \sum 2^{-k_n} \leq \mu^*(A) + \varepsilon$$

pour tout  $\varepsilon$ , d'où la propriété.  $\square$

## VI.2 Dendrogrammes

On décrit ici un procédé de construction d'un arbre à partir d'un espace métrique fini, qui induit une nouvelle métrique sur cet espace, de nature ultramétrique, et permet de visualiser d'une certaine manière la structure de l'espace. Il s'agit d'un procédé constructif et graphique, dont nous donnons ici une définition abstraite

**Définition VI.2.1.** (Dendrogramme)

On considère un espace métrique fini  $(X, d)$ , de cardinal  $N$ . On appelle dendrogramme de  $X$  une suite finie  $(X^k)_{0 \leq k < N}$  de partitions de  $X$  :

$$X^k = \{X_1^k, \dots, X_{N-k}^k\}, \quad X_i^k \in \mathcal{P}(X), \quad X_i^k \cap X_j^k = \emptyset \quad \forall i \neq j, \quad X = \bigcup_{j=1}^{N-k} X_j^k,$$

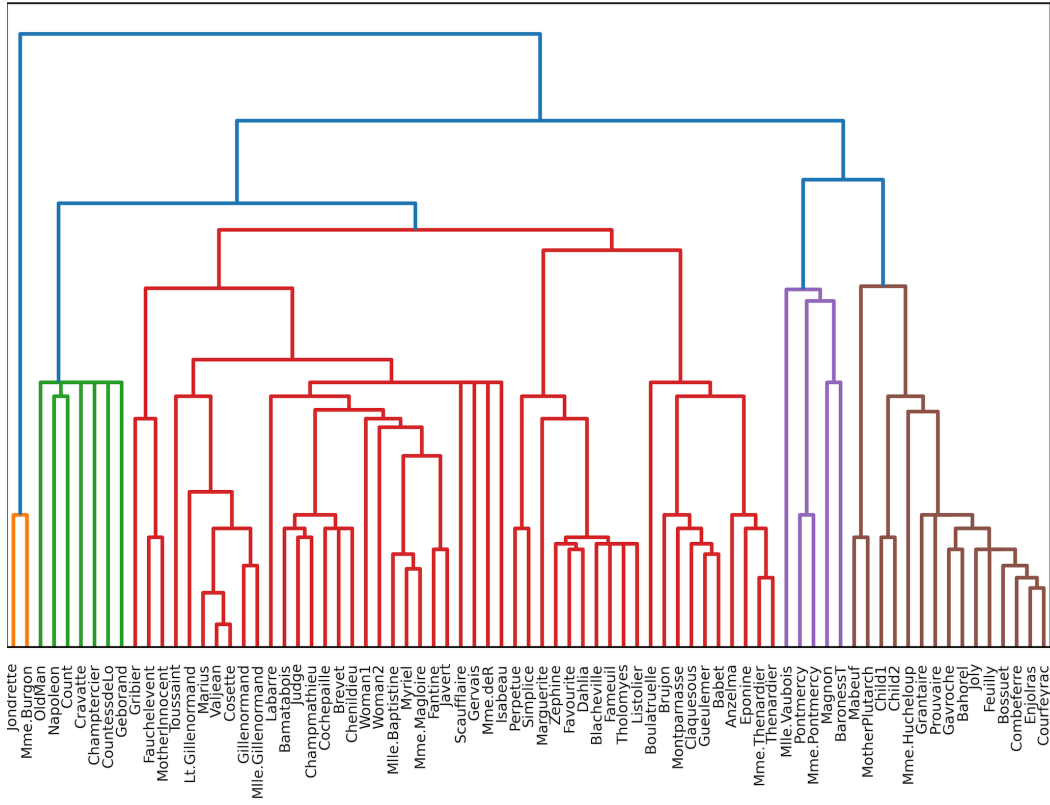


FIGURE VI.2.1 – Dendrogramme des Misérables

ainsi qu'une suite réelle  $(D^k)_{1 \leq k \leq N}$ , construits selon le processus d'agrégation suivant. La première partition  $X_0$  est la plus fine, de cardinal  $N$ , et  $D^0 = 0$ . Supposons  $X^k$  connue, de cardinal  $N - k$ , avec  $k \geq 0$ . On construit la partition suivante en agrégeant 2 éléments de la partition, selon le principe suivant. On définit

$$D_{ij}^{k+1} = d(X_i^k, X_j^k) = \inf_{x \in X_i^k, y \in X_j^k} d(x, y),$$

et l'on choisit un couple  $(i, j)$  qui minimise cette quantité. On note  $D^{k+1}$  le minimiseur, et l'on définit la partition  $X^k$  en conservant les éléments autres que  $i$  et  $j$ , auxquels on adjoint la réunion de  $X_i^k$  et  $X_j^k$ . Le cardinal de la partition est donc  $N - k - 1$ . La suite  $D^k$  est croissante par construction, avec  $D^1 > 0$ , et  $X^{N-1}$  est la partition triviale ( $X^{N-1} = \{X\}$ ).

**Proposition VI.2.2.** (Métrique ultramétrique associée à un dendrogramme)

On définit

$$\delta : (x, y) \in X \times X \mapsto \delta(x, y) \in \mathbb{R}_+$$

de la façon suivante : pour  $x$  et  $y$  dans  $X$  donné, on définit  $k_{xy}$  comme le plus petit entier tel que  $x$  et  $y$  sont dans une même sous-partie de  $X^k$ , et l'on fixe  $\delta(x, y) = D^k$ . L'application  $\delta(\cdot, \cdot)$  ainsi définie est une distance, qui est ultra-métrique, c'est à dire qu'elle vérifie l'inégalité triangulaire renforcée

$$\delta(x, y) \leq \max(\delta(x, z), \delta(z, y)) \quad \forall x, y, z \in X$$

*Démonstration.* L'application  $\delta$  prend bien des valeurs positives, on a  $\delta(x, y) = 0$  si et seulement si  $k_{xy} = 0$ , si et seulement si  $x = y$ . Elle est symétrique par construction. Soient maintenant  $x, y, z$ , et les  $k_{xy}$  et  $k_{yz}$  associés. Pour  $k \geq \max(k_{xy}, k_{yz})$ ,  $x$  et  $z$  sont dans la même composante, d'où

$$k_{xz} \leq \max(k_{xy}, k_{yz}),$$

d'où l'inégalité ultramétrique par croissance de la suite  $(D_k)$ .  $\square$

À titre d'illustration, la figure VI.2.1 représente graphiquement le dendrogramme obtenu à partir de l'ensemble  $X$  des personnages des Misérables, muni d'une métrique qui prend en compte la proximité dans le roman des personnages (la distance est d'autant plus petite que les 2 personnages partagent un nombre important de scènes).

## VI.3 Introduction au transport optimal

### VI.3.1 Problème d'affectation et problème de Monge Kantorovich discret

Le problème d'affectation se formule comme suit :

**Problème VI.3.1.** On considère 2 ensembles de même cardinal  $N \in \mathbb{N}$ , tous deux identifiés à  $\{1, \dots, N\}$ , et l'on se donne une collection de *coûts*  $c_{ij} \in \mathbb{R}$ . Le problème consiste à trouver une bijection  $\varphi$  qui minimise la quantité

$$\sum_{i=1}^N c_{i\varphi(i)}.$$

Le problème ci-dessus ne présente pas d'intérêt théorique particulier : l'ensemble des bijections (groupe symétrique  $S_N$ ) est fini, le problème admet bien (au moins) une solution. Mais la recherche effective de ce minimum peut extrêmement laborieuse, car le cardinal de l'ensemble des candidats croît comme  $N!$ .

Nous allons considérer une version relaxée du problème ci-dessus<sup>4</sup>, qui peut se formuler intuitivement de la façon suivante, dans un contexte de transport : on considère le premier ensemble comme contenant des positions dans un certain espace (il n'est pas nécessaire de préciser lequel ici), et le second ensemble aussi comme une collection de positions dans un espace (éventuellement le même, mais pas forcément). On note  $c_{ij}$  ce que cela coûte de transporter une quantité unitaire de matière de  $x_i$  vers  $y_j$ . Le problème précédent consistant à considérer que l'on avait une même quantité de matière en chaque point (par exemple  $1/N$ ), et que l'on cherchait à transporter cette matière vers le second ensemble en envoyant toute la matière de chaque point vers une destination unique. Nous allons considérer maintenant qu'il est possible de distribuer la matière venant d'un point vers plusieurs destinations. Cette relaxation du problème permet de lever la contrainte d'avoir le même nombre de points au départ et à l'arrivée. Dans ce qui suit on notera  $\gamma_{ij}$  la quantité de matière allant de  $i$  vers  $j$ . On appellera  $\gamma = (\gamma_{ij})$  un *plan de transport*.

#### Problème de Monge-Kantorovich discret

On considère 2 ensembles<sup>5</sup> finis  $X$  et  $Y$ , de cardinaux respectifs  $N$  et  $M \in \mathbb{N}$  et l'on se donne une collection de coûts  $c_{ij} \in \mathbb{R}$ . On se donne deux mesures de probabilités discrètes  $\mu$  et  $\nu$  sur  $X$  et  $Y$ , respectivement ( $\mu_i$  est la masse portée par  $i$ , avec  $\sum \mu_i = 1$ , de même pour  $\nu$ ). On supposera tous les poids strictement positifs<sup>6</sup>. Le problème s'écrit

$$\min_{\Pi_{\mu\nu}} C(\gamma) \tag{VI.3.1}$$

avec

$$C(\gamma) = \sum_{i,j} c_{ij} \gamma_{ij}, \quad \Pi_{\mu,\nu} = \left\{ \gamma \in \mathbb{R}_+^{N \times M}, \sum_j \gamma_{ij} = \mu_i \quad \forall i, \sum_i \gamma_{ij} = \nu_j \quad \forall j \right\}$$

4. Cette approche a été proposée par L.V. Kantorovich en 1942. On trouvera une traduction du papier original sur <http://www.math.toronto.edu/mccann/assignments/477/Kantorovich42.pdf>

5. Il n'y a pas lieu de préciser ici les points d'arrivée et points de départ. Nous nous intéresserons plus loin au transport entre points d'un espace euclidien, mais ici on peut tout aussi bien concevoir le transport d'une essence vers le *concept de néant* chez Sartre.

6. On peut toujours se ramener à cette situation en supprimant de  $X$  et  $Y$  les points non chargés.

**Remarque VI.3.1.** On peut formuler ce problème en termes probabilistes, en considérant  $\gamma$  comme une loi de probabilité sur l'espace produit  $X \times Y$ , dont les mesures images par les projections sur  $X$  et  $Y$  sont respectivement  $\mu$  et  $\nu$ . Parmi de telles lois, on cherche celle(s) qui minimise(nt) l'espérance de la "fonction"  $c = (c_{ij})$  sur  $X \times Y$ .

**Remarque VI.3.2.** L'ensemble admissible est non vide, il contient en particulier le plan correspondant à une loi de probabilité sur  $X \times Y$  pour deux variables indépendantes, qui s'écrit

$$\gamma_{ij} = \mu_i \nu_j.$$

**Proposition VI.3.3.** Le problème VI.3.1 admet un minimiseur.

*Démonstration.* Les  $\gamma_{ij}$  sont positifs, et chacun d'eux est majoré par le max des  $\mu_i$ , l'ensemble  $\Pi$  est donc borné, il est évidemment fermé donc compact : la fonction continue (car linéaire)  $C(\cdot)$  admet donc un minimiseur sur  $\Pi$ .  $\square$

**Remarque VI.3.4.** Dans le cas d'un coût du type  $c_{ij} = a_i + b_j$ , le problème est fortement dégénéré, puisque tout transport de  $\mu$  vers  $\nu$  réalise le même coût. Par ailleurs, pour deux ensembles de même cardinal  $N$ , avec  $\mu$  et  $\nu$  lois uniformes sur  $X$  et  $Y$ , si l'on se donne une bijection  $\varphi$  de  $S_n$ , on peut construire une famille de coûts telle que le plan associé à la bijection<sup>7</sup> soit l'unique minimiseur, en prenant par exemple  $c_{i\varphi(i)} = -1$ , et  $c_{ij} = 0$  si  $j \neq \varphi(i)$ .

**Question VI.3.1.** Étant donnée une collection de coût  $(c_{ij})$ , existe-t-il des ensembles  $X$  et  $Y$  de points de  $\mathbb{R}^d$  tels que  $c_{ij} = |y_j - x_i|$ ? (on pourra aussi considérer  $c_{ij} = |y_j - x_i|^p$ ,  $c_{ij} = \psi(|y_j - x_i|)$  avec  $\psi$  croissante et nulle en 0.)

**Question VI.3.2.** Le problème VI.3.1 admet-il une solution unique "en général"? (on s'attachera à exprimer précisément ce que l'on entend par unicité générique.)

### Lien avec le problème d'affectation

Dans le cas où les cardinaux sont les mêmes, et les mesures équidistribuées, on peut préciser le lien entre le modèle relaxé basé sur les plans de transports et le problème d'affectation. Pour simplifier les notations, on considère ici la situation où chaque point porte une masse unitaire, de telle sorte que la masse totale des mesures considérées est égale au nombre de points. Il ne s'agit donc plus de mesure de probabilité, mais on peut s'y ramener en divisant la mesure par le nombre de points.

**Proposition VI.3.5.** On se place dans le cas  $N = M$  (même nombre de points de part et d'autre, et  $\mu_i = \nu_j \equiv 1$ ), et l'on note  $\Pi_S$  l'ensemble des plans de transport associés à une affectation, i.e.  $\gamma_{ij} = \delta_{i\varphi(i)}$ , où  $\varphi$  est une permutation du groupe symétrique. L'ensemble des points extrémaux<sup>8</sup> de  $\Pi$  s'identifie à  $\Pi_S$ .

*Démonstration.* Tout point de  $\Pi_S$  est de façon évidente extrémal pour  $\Pi$ . Réciproquement, considérons un plan générique (i.e. qui n'est pas associé à une bijection)  $\gamma$ . On considère dans un premier temps les indices  $i$  pour lesquels  $\gamma_{ij}$  est nul pour tous les indices  $j$  sauf un (qui vaut donc 1). Cette sous-famille des points de départ est en bijection avec les points d'arrivées  $j$  correspondants, pour lesquels, symétriquement,  $\gamma_{ij}$  est nul pour tous les  $i$  sauf 1. On note  $I$  (resp.  $J$ ) l'ensemble des indices non concernés dans l'espace de départ (resp. d'arrivée). Les ensembles  $I$  et  $J$  sont de même cardinal, et non vides par hypothèse. La restriction du plan  $\gamma$  à  $X_I \times Y_J$  est diffuse, au sens que pour tout  $i$ ,  $\gamma_{ij} \in ]0, 1[$  pour au moins 2 indices  $j \in J$ , et pour tout  $j \in J$ , on a  $\gamma_{ij} \in ]0, 1[$  pour au moins 2 indices  $i \in I$ . On part d'un indice  $i_0 \in I$ , et l'on choisit  $j_0$  tel que  $\gamma_{i_0 j_0} > 0$ . On choisit ensuite  $i_1 \neq i_0$  tel que  $\gamma_{i_1 j_0} > 0$ , puis  $j_1 \neq j_0$  tel que  $\gamma_{i_1 j_1} > 0$ . On construit ainsi une

7. C'est à dire :  $\gamma_{i\varphi(i)} = 1/N$ , et  $\gamma_{ij} = 0$  si  $j \neq \varphi(i)$ .

8. On dit que  $\gamma \in \Pi \subset \mathbb{R}^d$  est point extrémal de  $\Pi$  si  $\gamma = (\gamma^1 + \gamma^2)/2$ , avec  $\gamma^1, \gamma^2 \in \Pi$ , implique  $\gamma^1 = \gamma^2 = \gamma$ .

suite d'indices

$$i_0, j_0, i_1, \dots, i_{n-1}, i_n,$$

que l'on peut voir comme un chemin dans le graphe sur  $I \cup J$  associé au plan  $\gamma$ , chemin qui ne contient pas d'aller-retour. L'ensemble des indices étant fini, il existe forcément un  $n$  tel que  $i_n$  correspond à un indice  $i_\ell \neq i_{n-1}$  déjà visité. On considère alors la variation

$$h = \sum_{k=\ell}^{n-1} (\pi_{i_k, j_k} - \pi_{i_{k+1}, j_k}),$$

avec  $i_n = i_\ell$ , et où  $\pi_{i,j}$  est l'élément de  $\mathbb{R}^{NM}$  qui vaut 1 sur la composante  $(i,j)$ , et qui est nul pour les autres couples. Pour  $\eta$  suffisamment petit,  $\gamma \pm \eta h$  est positif, et par construction  $\gamma \pm \eta h$  vérifie les contraintes de marginales, les deux perturbations sont donc dans  $\Pi_{\mu, \nu}$ , et  $\gamma$  est moyenne non triviale de ces deux plans de transport, il ne s'agit donc pas d'un point extrême.

Les seuls points extrêmes correspondent donc aux permutations.  $\square$

**Corollaire VI.3.6.** L'ensemble  $\Pi$  des plans de transport admissibles est l'enveloppe convexe de  $\Pi_S$ .

*Démonstration.* Il s'agit d'une conséquence du théorème de Krein-Milman en dimension finie, qui assure que tout convexe compact d'un espace affine de dimension finie est l'enveloppe convexe de ses points extrêmes.  $\square$

**Proposition VI.3.7.** On se place comme précédemment dans la situation de mesures équidistribuées sur des ensembles de même cardinal. Le problème de Monge Kantorovich discret VI.3.1 admet au moins une solution dans  $\Pi_S$ , i.e. une solution optimale du type permutation.

*Démonstration.* D'après la proposition VI.3.3, le problème VI.3.1 admet un minimiseur  $\gamma$ . D'après la proposition VI.3.5, ce minimiseur s'écrit comme combinaison convexe de plans associés à des permutations  $\varphi_1, \dots, \varphi_K$  :

$$\gamma = \sum \theta_k \gamma^k$$

(on ne garde dans la somme ci-dessus que les termes non triviaux, de telle sorte que  $\theta_k > 0$  pour tout  $k$ ). Le coût étant linéaire, on a

$$C(\gamma) = \sum \theta_k C(\gamma^k).$$

Comme chaque  $C(\gamma^k)$  est supérieur ou égal à  $C(\gamma)$ , et que  $\sum \theta_k = 1$  avec  $\theta_k > 0$  pour tout  $k$ , la combinaison convexe ci-dessus implique que  $C(\gamma^k)$  est égal à  $C(\gamma)$  pour tout  $k$ . Chaque permutation impliquée dans la combinaison réalise donc le minimum.  $\square$

### Distance sur l'ensemble des mesures atomiques

On note  $\mathcal{A} = \mathcal{A}(\mathbb{R}^d)$  l'ensemble des mesures de probabilités atomiques sur  $\mathbb{R}^d$  à support fini, c'est à dire l'ensemble des  $\mu$  de la forme

$$\mu = \sum_{i=1}^N \mu_i \delta_{x_i}, \quad \mu_i > 0, \quad \sum_{i=1}^N \mu_i = 1.$$

L'entier  $N$  n'est pas fixé, mais on ne considère ici que des sommes finies. Pour  $p \in [1, +\infty[$  fixé,  $\mu$  et  $\nu$  dans  $\mathcal{A}$ , on note

$$W_p(\mu, \nu) = \left( \inf_{\gamma \in \Pi(\mu, \nu)} \sum \gamma_{ij} |y_j - x_i|^p \right)^{1/p},$$

où l'infimum correspond au problème de MK discret (VI.3.1), pour lequel l'existence d'un plan minimisant est établie dans VI.3.3. On se propose de montrer que  $W_p$  est une distance sur  $\mathcal{A}_d$ .

**Théorème VI.3.8.** La fonction  $W_p(\cdot, \cdot)$  définie ci-dessus sur  $\mathcal{A} \times \mathcal{A}$  est une distance.

*Démonstration.* On a de façon évidente  $W_p(\mu, \nu) = 0$  si et seulement si  $\mu = \nu$ , et la distance est symétrique par construction (le problème de recherche d'un plan de coût minimal est symétrique par rapport aux mesures). Pour l'inégalité triangulaire, on considère trois mesures  $\mu^1, \mu^2$ , et  $\mu^3$  de  $\mathcal{A}$ . On note  $\gamma^{12}$  et  $\gamma^{23}$  des plans qui réalisent la distance de 1 vers 2 et de 2 vers 3, respectivement. On note  $\gamma^{123}$  le "plan à trois" défini de la façon suivante<sup>9</sup>

$$\gamma_{i_1 i_2 i_3}^{123} = \frac{1}{\mu_{i_2}^2} \gamma_{i_1 i_2}^{12} \gamma_{i_2 i_3}^{23}.$$

On note  $\gamma^{13}$  le plan défini de façon naturelle par

$$\gamma_{i_1 i_3}^{13} = \sum_{i_2} \gamma_{i_1 i_2 i_3}^{123}.$$

On a

$$\begin{aligned} W_p(\mu^1, \mu^3) &\leq \left( \sum_{i_1 i_3} \gamma_{i_1 i_3}^{13} |x_{i_3}^3 - x_{i_1}^1|^p \right)^{1/p} = \left( \sum_{i_1 i_2 i_3} \gamma_{i_1 i_2 i_3}^{123} |x_{i_3}^3 - x_{i_1}^1|^p \right)^{1/p} \\ &\leq \left( \sum_{i_1 i_2 i_3} \gamma_{i_1 i_2 i_3}^{123} |x_{i_2}^2 - x_{i_1}^1|^p \right)^{1/p} + \left( \sum_{i_1 i_2 i_3} \gamma_{i_1 i_2 i_3}^{123} |x_{i_3}^3 - x_{i_2}^2|^p \right)^{1/p} \end{aligned}$$

d'après l'inégalité de Minkowski, d'où finalement

$$W_p(\mu^1, \mu^3) \leq \left( \sum_{i_1 i_2} \gamma_{i_1 i_2}^{12} |x_{i_2}^2 - x_{i_1}^1|^p \right)^{1/p} + \left( \sum_{i_2 i_3} \gamma_{i_2 i_3}^{23} |x_{i_3}^3 - x_{i_2}^2|^p \right)^{1/p} = W_p(\mu^2, \mu^3) + W_p(\mu^1, \mu^2),$$

ce qui termine la preuve.  $\square$

### VI.3.2 Transport optimal, cas général

On considère ici deux espaces mesurables  $(X, \mathcal{A})$  et  $(X', \mathcal{A}')$ . On se donne deux mesures de probabilité  $\mu$  et  $\mu'$  sur ces espaces respectifs. On introduit l'ensemble  $\Lambda_{\mu, \mu'}$  des applications  $T$  mesurables qui poussent  $\mu$  vers  $\mu'$ , c'est à dire telles que  $T_{\#}\mu = \mu'$ . On définit une fonction de *coût* mesurable pour la tribu produit  $\mathcal{A} \otimes \mathcal{A}'$

$$c : X \times X' \longrightarrow [0, +\infty]$$

La fonction  $x \mapsto c(x, T(x))$  est  $\mathcal{A}$  mesurable. En effet, pour tout  $b \in \mathbb{R}$ , l'ensemble

$$B = \{(x, x'), c(x, x') \leq b\} \subset X \times X',$$

est dans  $\mathcal{A} \otimes \mathcal{A}'$  par mesurabilité de  $c$ . L'image réciproque de  $] - \infty, b]$  par  $x \mapsto c(x, T(x))$  est donc l'image réciproque de  $B$  par l'application  $x \mapsto (x, T(x))$ , il suffit donc de vérifier que cette dernière application est mesurable. Il suffit pour cela de vérifier que l'image réciproque de tout rectangle  $A \times A'$ , avec  $A \in \mathcal{A}$  et  $A' \in \mathcal{A}'$ , est dans  $\mathcal{A}$ . Or cette image réciproque est

$$A \cap T^{-1}(A'),$$

qui est mesurable par mesurabilité de  $T$ .

On définit maintenant le coût de transport associé à  $T$  comme

$$C(T) = \int_X c(x, T(x)) d\mu(x),$$

qui est bien défini d'après ce qui précède. Le problème abstrait de Monge consiste à minimiser  $C(T)$  parmi les transports admissibles, il s'écrit

$$\min_{T \in \Lambda_{\mu, \mu'}} C(T).$$

<sup>9</sup>. On peut voir  $\gamma^{123}$  comme la loi d'une variable aléatoire sur  $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d$  dont les projections ont pour lois respectives  $\mu^1, \mu^2$  et  $\mu^3$ .

Ce problème est très fortement non linéaire vis-à-vis de l'inconnue  $T$ . Il est par ailleurs évidemment mal posé dans certains cas, par exemple si la mesure de départ est atomique (combinaison de masses ponctuelles), alors que la mesure d'arrivée est diffuse, ou elle-même atomique mais avec un nombre différent de masse. Une forme relaxée de ce problème a été proposé par Leonid Kantorovich dans les années 1940. L'idée consiste à autoriser que la masse partant d'un point puisse être envoyée vers plusieurs points. Cette relaxation du problème initial peut se formaliser en considérant non plus des applications de transport, mais des plans de transport, que l'on peut voir comme des mesures de probabilités sur l'espace produit  $X \times X'$ . On considère dans cet esprit l'ensemble des mesures  $\gamma \in \mathcal{P}(X \times X')$  qui admettent pour marginales les mesures de départ et d'arrivée, respectivement :

$$\Pi_{\mu, \mu'} = \left\{ \gamma \in \mathcal{P}(X \times X'), P_{\#}^X \gamma = \mu, P_{\#}^{X'} \gamma = \mu' \right\},$$

où  $P_{\#}^X \gamma$  est la mesure image de  $\gamma$  par la projection sur  $X$ , i.e.

$$\forall A \in \mathcal{A}, P_{\#}^X \gamma(A) = \gamma(A \times X').$$

Le coût total associé à un plan de transport  $\gamma$  est alors

$$C(\gamma) = \int_{X \times X'} c(x, x') d\gamma(x, x'),$$

et le problème s'écrit

$$\min_{\gamma \in \Pi_{\mu, \mu'}} C(\gamma).$$

La fonctionnelle est maintenant linéaire en l'inconnue, et l'existence d'une solution est assurée sous des hypothèses très générales.

## VI.4 Distance de Gromov-Wasserstein

**Définition VI.4.1.** Soient  $(X, d, \mu)$  et  $(X', d', \mu')$  deux espaces métriques finis probabilisés (i.e. munis d'une mesure définie sur la tribu discrète, de masse totale 1). On suppose que les mesures chargent tous les points<sup>10</sup> On dit que ces espaces sont isomorphes s'il existe une bijection de  $X$  vers  $X'$  qui préserve les structures de distance et de mesure, i.e. s'il existe une isométrie  $T$  telle que

$$T_{\#} \mu = \mu' \quad \text{i.e.} \quad \mu'(A') = \mu(T^{-1}(A')) \quad \forall A' \in \mathcal{P}(X').$$

Soient  $(X, d, \mu)$  et  $(X', d', \mu')$  deux espaces métriques probabilisés et finis, de cardinaux respectifs  $N$  et  $N'$ , munis de leurs tribus discrètes respectives. On suppose que  $\mu$  et  $\mu'$  sont toutes deux de masse 1 (mesures de probabilité). On note  $\Pi$  l'ensemble des plans de transport entre  $\mu$  et  $\mu'$  :

$$\Pi = \left\{ \gamma = (\gamma_{xx'}) \in \mathbb{R}_+^{N \times N'}, \sum_x \gamma_{xx'} = \mu'_{x'}, \sum_{x'} \gamma_{xx'} = \mu_x \right\}.$$

**Définition VI.4.2.** Pour  $p \in [1, +\infty[$ , on définit

$$d_{GW_p}(X, X') = \inf_{\gamma \in \Pi} \left( \sum_{xx'} \sum_{yy'} |d(x, y) - d(x', y')|^p \gamma_{xx'} \gamma_{yy'} \right)^{1/p}.$$

**Lemme VI.4.3.** L'infimum de la définition précédente est atteint.

*Démonstration.* L'ensemble admissible  $\Pi$  est compact, et l'application

$$\gamma \mapsto \left( \sum_{xx'} \sum_{yy'} |d(x, y) - d(x', y')|^p \gamma_{xx'} \gamma_{yy'} \right)^{1/p}$$

est continue. □

**Proposition VI.4.4.** La quantité définie ci-dessus est une distance sur l'ensemble des espaces métriques probabilisés finis (quotienté par les isomorphismes au sens de la définition VI.4.1)

*Démonstration.* Dans ce qui suit on note  $\gamma$  un plan qui réalise l'infimum ci-dessus. Si l'on a  $d_{GWp}(X, X') = 0$  alors, pour tous  $x, x', y, y'$  tels que  $\gamma_{xx'} \neq 0$  et  $\gamma_{yy'} \neq 0$ , on a  $d(x, y) = d(x', y')$ . Soit maintenant  $x, x', y'$ , tels que  $x$  envoie de la masse à  $x'$  et  $y'$ . On a  $d(x', y') = d(x, x) = 0$ , d'où  $x' = y'$ . Pour chaque  $x$  il existe donc unique  $x'$  tel que  $\gamma_{xx'} \neq 0$ , et l'on a donc  $\gamma_{xx'} = \mu_x$ . On peut mener le même raisonnement dans l'autre sens : pour chaque  $x'$  il existe un unique  $x$  tel que  $\gamma_{xx'} \neq 0$ , et l'on a donc  $\gamma_{xx'} = \mu_{x'}$ . Le plan de transport  $\gamma$  correspond donc à une bijection  $T$  entre  $X$  et  $X'$ , et l'on a

$$\begin{aligned} 0 = d_{GWp}(X, X') &= \left( \sum_{xx'} \sum_{yy'} |d(x, y) - d(x', y')|^p \gamma_{xx'} \gamma_{yy'} \right)^{1/p} \\ &= \left( \sum_x \sum_y |d(x, y) - d(T(x), T(y))|^p \right)^{1/p} \end{aligned}$$

La symétrie est immédiate d'après la définition.

Pour l'inégalité triangulaire, on considère 3 espaces métriques probabilisés  $X, X',$  et  $X''$ , et des plans de transport  $(\gamma_{xx'})$  et  $(\gamma_{x'x''})$  qui réalisent les distance entre  $X$  et  $X'$  et entre  $X'$  et  $X''$ , respectivement. On construit à partir de ces plans un plan entre  $X$  et  $X''$  (non nécessairement optimal, mais qui suffira pour l'inégalité triangulaire) en "collant" dans un premier temps les plans, puis en condensant l'espace  $X'$  intermédiaire. Plus précisément, on introduit

$$\gamma_{xx'x''} = \frac{\gamma_{xx'} \gamma_{x'x''}}{\mu_{x'}},$$

et l'on définit

$$\gamma_{xx''} = \sum_{x' \in X'} \gamma_{xx'x''}.$$

On a

$$\sum_{xx''} \sum_{yy''} |d(x, y) - d(x'', y'')| \gamma_{xx''} \gamma_{yy''} = \sum_{xx''} \sum_{yy''} |d(x, y) - d(x'', y'')| \sum_{x'} \frac{\gamma_{xx'} \gamma_{x'x''}}{\mu_{x'}} \sum_{y'} \frac{\gamma_{yy'} \gamma_{y'y''}}{\mu_{y'}}.$$

On écrit  $|d(x, y) - d(x'', y'')| \leq |d(x, y) - d(x', y')| + |d(x', y') - d(x'', y'')|$ , ce qui permet de majorer la quantité de départ par deux termes. le premier s'écrit

$$\begin{aligned} &\sum_{xx''} \sum_{yy''} |d(x, y) - d(x', y')| \sum_{x'} \frac{\gamma_{xx'} \gamma_{x'x''}}{\mu_{x'}} \sum_{y'} \frac{\gamma_{yy'} \gamma_{y'y''}}{\mu_{y'}} \\ &= \sum_x \sum_y |d(x, y) - d(x', y')| \sum_{x'} \frac{\gamma_{xx'}}{\mu_{x'}} \underbrace{\sum_{x''} \gamma_{x'x''}}_{\mu_{x'}} \sum_{y'} \frac{\gamma_{yy'}}{\mu_{y'}} \underbrace{\sum_{y''} \gamma_{y'y''}}_{\mu_{y'}} \\ &= \sum_x \sum_y |d(x, y) - d(x', y')| \sum_{x'} \gamma_{xx'} \gamma_{yy'} = d_{GW}(X, X'). \end{aligned}$$

Le second terme s'identifie de la même manière à  $d_{GW}(X'X'')$  □



## VI.5 Propagation d'opinion et flot de gradient

On s'intéresse ici à un modèle simple de propagation d'opinion sur un réseau. Les nœuds de ce réseaux sont des personnes, ou "agents", auxquels on affecte un nombre représentant une opinion sur un certain sujet à un instant donné. Ce nombre peut par exemple représenter la tendance qu'a un individu à voter pour tel ou tel candidat au second tour d'une élection présidentielle, ou l'idée que l'on peut se faire de la probabilité de gain d'une équipe nationale à une finale de coupe du monde. Dans un autre contexte, on pourra penser à la valeur d'une quantité qui fait l'objet d'un débat public, comme l'augmentation de la température moyenne sur la planète dans 20 ans.

On considère un ensemble  $V$  de  $N$  individus, on note  $u_x^k$  l'opinion de l'individu  $x$  à l'instant  $k$ , et par  $u^k = (u_x^k)_{x \in V}$  la collection des opinions. L'influence de  $y \in V$  sur  $x$  est quantifiée par un coefficient  $K_{xy} \in [0, 1]$ , et l'on suppose :

$$\sum_{y \in V} K_{xy} = 1 \quad \forall x \in V.$$

La collection de l'ensemble des coefficient est donc encodée par une matrice (sans choix de numérotation des sommets)  $(K_{xy})_{(x,y) \in V^2}$  *stochastique*.

On notera que toute l'information sur le graphe est dans la collection des influences :  $E \subset V \times V$  est défini par

$$E = \text{supp}(K_{xy}) = \{(x, y) \in V \times V, K_{xy} > 0\}.$$

On conservera néanmoins la notation (redondante)  $(V, E, K)$  pour désigner le réseau.

On écrira  $x \rightarrow y$  si  $K_{xy} > 0$ , qui signifie que  $x$  écoute  $y$ , ou  $x$  suit  $y$ , ou plus généralement  $x$  est influencé par  $y$ . Avec cette convention, l'opinion / influence remonte le sens des flèches.

Le coefficient diagonal  $K_{xx}$  peut être non nul (présence de boucles dans le réseau), ce qui correspond à une certaine inertie de  $x$ , ou résistance de  $x$  à modifier son opinion sous l'effet d'influences extérieures, jusqu'à éventuellement ne plus se préoccuper de l'opinion des autres (cas extrême  $K_{xx} = 1$ ) On note  $\Gamma \subset V$  l'ensemble des sommets qui ne pointent que vers eux-mêmes

$$\Gamma = \{x \in V, K_{xx} = 1\}, \quad (\text{VI.5.1})$$

et par  $\hat{V} = V \setminus \Gamma$  l'ensemble des sommets intérieurs. La frontière  $\Gamma$  correspond aux individus qui n'écoutent qu'eux mêmes, et étant éventuellement suivis par d'autres. Si l'on considère que ces agents affichent une opinion dans le dessein de modifier l'opinion d'autres agents du réseau, on peut voir ces individus comme des *influenceurs*<sup>11</sup>, ou plus simplement, s'ils ne nourrissent aucun dessein particulier, de personnes non influençables, ou *têtes*.

### Modèle discret d'évolution

On se donne une collection d'opinions initiales  $(u_x^0)_{x \in V}$ , et l'on considère que l'opinion évolue d'un jour à l'autre selon la relation

$$u_x^{k+1} = \sum_{x \rightarrow y} K_{xy} u_y^k \quad \forall x \in V. \quad (\text{VI.5.2})$$

qui peut s'écrire aussi matriciellement On notera que l'indication ' $x \rightarrow y$ ' n'est pas à strictement parler obligatoire, du fait que  $K_{xy} = 0$  dès que  $(x, y) \notin E$ . Elle sera parfois omise, et l'on écrira alors simplement  $\sum K_{xy} u_y^k$ .

Le fait que cela prenne un certain temps pour  $x$  d'absorber l'influence de ses voisins peut être modélisé en introduisant un paramètre d'inertie  $\theta \in [0, 1]$ , et en écrivant le modèle relaxé

$$u_x^{k+1} = (1 - \theta)u_x^k + \theta \sum_{x \rightarrow y} K_{xy} u_y^k. \quad (\text{VI.5.3})$$

11. Ces influenceurs peuvent aussi être vus comme des agents influencés par une entité extérieure (entreprise, groupe de pression, ...) qui les contrôle.

Pour  $\theta = 1$ , on retrouve le problème discret. Noter que ce nouveau problème rentre dans le cadre discret précédent, en introduisant les paramètres modifiés

$$K'_{xx} = (1 - \theta) + \theta K_{xx}, \quad K'_{xy} = \theta K_{xy} \quad \text{for } y \neq x.$$

### Équation différentielle

On considère que  $\theta$  s'écrit  $\varepsilon/\eta$ , où  $\eta$  est un temps de relaxation fixe (temps typique de propagation de l'influence), et  $\varepsilon$  un petit paramètre (également homogène à un temps). On a

$$\frac{u_x^{k+1} - u_x^k}{\varepsilon} = \frac{1}{\eta} \left( \sum_{x \rightarrow y} K_{xy} (u_y^k - u_x^k) \right).$$

L'évolution prend la forme de la discrétisation en temps d'un système d'équations différentielles ordinaires pour des quantités  $t \mapsto u_x^t \in \mathbb{R}$  qui varient continûment en temps, et vérifient le système d'équations

$$\frac{d}{dt} u_x^t = \frac{1}{\eta} \left( \sum_{x \rightarrow y} K_{xy} (u_y^t - u_x^t) \right).$$

Le problème continu en temps s'écrit donc

$$\frac{d}{dt} u^t = -\frac{1}{\eta} A u^t \quad \text{avec} \quad A = I - K. \quad (\text{VI.5.4})$$

On se propose de caractériser ici les cas où le problème d'évolution (VI.5.4) a une structure de flot de gradient pour un certain produit scalaire.

Le cas où l'on se restreint au produit scalaire euclidien canonique est immédiat :

**Proposition VI.5.1.** Le problème VI.5.4 a une structure de flot de gradient pour le produit scalaire canonique, i.e. il existe une fonctionnelle  $\Psi$  deux fois continûment différentiable telle que  $Au = \nabla \Psi(u)$ , si et seulement si  $A$  est symétrique.

*Démonstration.* Si  $A$  est symétrique on a  $Au = \nabla \Psi(u)$  avec

$$\Psi(u) = \frac{1}{2} \langle u | u \rangle.$$

Si  $Au = \nabla \Psi(u)$ , alors  $a_{ij} = \partial^2 \Psi / \partial x_i \partial x_j = \partial^2 \Psi / \partial x_j \partial x_i = a_{ji}$ . □

Plus généralement, si l'on considère une matrice  $M$  symétrique définie positive, on note  $\langle \cdot | \cdot \rangle_M$  le produit scalaire associé, i.e.

$$\langle u | v \rangle_M = \langle Mu | v \rangle.$$

Pour toute fonctionnelle  $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}$  continûment différentiable, on note  $\nabla_M \Phi(u) \in \mathbb{R}^N$  son gradient en  $u$  selon le produit scalaire associé à  $M$ , c'est à dire le vecteur tel que

$$\Phi(u + h) = \Phi(u) + \langle \nabla_M \Phi | h \rangle_M + o(h) = \Phi(u) + \langle M \nabla_M \Phi | h \rangle + o(h).$$

On a donc  $\nabla_M \Phi = M^{-1} \nabla \Phi$ , ce qui permet d'énoncer une première caractérisation des flots de gradient.

**Proposition VI.5.2.** Le problème VI.5.4 a une structure de flot de gradient pour le produit scalaire associé à la matrice s.d.p.  $M$ , i.e. il existe une fonctionnelle  $\Psi$  deux fois continûment différentiable telle que  $Au = \nabla_M \Psi(u)$ , si et seulement si  $A$  s'écrit  $M^{-1}B$ , où  $B$  est une matrice symétrique.

*Démonstration.* Si  $A = M^{-1}B$ , alors  $Au = \nabla_M \Psi(u)$  avec  $\Psi(u) = \frac{1}{2} \langle Bu | u \rangle$ . Si  $Au = \nabla_M \Psi(u)$  alors, comme précédemment,  $MA$  est nécessairement symétrique.  $\square$

Dans la suite, nous aborderons un cas particulier de systèmes présentant une structure de gradient, il s'agit que réseaux que nous appellerons *charismatiques* (voir la définition VI.5.4). Pour de tels réseaux, on aura  $K = M^{-1}B$ , où  $M$  est une matrice *diagonale*. Les coefficients diagonaux de  $M$  correspondent aux charismes  $(m_x)_{x \in V}$  des agents. D'un point de vue probabiliste, cette situation correspond au cas d'une chaîne de Markov *réversible*. Nous proposons maintenant une caractérisation plus exploitable des matrices  $A$  pour lesquelles l'équation (VI.5.4) a une structure de flot gradient (pour matrice s.d.p.  $M$  quelconque).

**Proposition VI.5.3.** Le problème (VI.5.4) a une structure de flot de gradient pour un certain produit scalaire si et seulement si  $A$  (ou, de façon équivalente,  $K$ ) est diagonalisable, et ses valeurs propres sont réelles.

*Démonstration.* Si  $Au$  est le gradient d'une fonctionnelle quadratique en  $u$  pour un produit scalaire  $\langle \cdot | \cdot \rangle_M$ , il existe une matrice symétrique  $B$  telle que  $A = M^{-1}B$ . Comme  $M$  est s.d.p., elle s'écrit  $M = UDU^{-1}$ , où  $U$  est une matrice orthogonale et  $D$  est diagonale. On définit alors  $M^{1/2}$  comme  $UD^{1/2}U^{-1}$ . La matrice  $A$  est semblable à

$$M^{1/2}AM^{-1/2} = M^{1/2}M^{-1}BM^{-1/2} = M^{-1/2}BM^{-1/2},$$

qui est symétrique car  $M^{-1/2}$  et  $B$  le sont. La matrice  $A$  est donc semblable à une matrice symétrique, elle est donc diagonalisable de valeurs propres réelles.

On suppose maintenant que  $A$  est diagonalisable, de valeurs propres réelles :  $A = PDP^{-1}$  où  $D$  est diagonale réelle, et  $P$  est une matrice inversible. On écrit

$$A = PDP^{-1} = PP^T P^{-T} DP^{-1} = M^{-1}B,$$

où  $M = P^{-T}P^{-1}$  est une matrice symétrique définie positive<sup>12</sup>, et  $B = P^{-T}DP^{-1}$  est symétrique réelle.  $\square$

## Réseaux charismatiques

Nous nous intéressons ici à des réseaux qui encodent des interactions d'une nature symétrique, qui correspondent comme nous verrons à la situation d'un modèle d'évolution de type flot de gradient, pour une matrice  $M$  *diagonale*. Plus précisément, les réseaux que nous considérons ici sont basés sur l'existence d'un paramètre afférent à chaque individu, un poids que nous appellerons *charisme* dans ce contexte, qui conditionne l'influence qu'il exerce sur les autres. Comme nous allons le voir, cette hypothèse rapprochera les réseaux qui la vérifie des réseaux résistifs ou, dans un contexte stochastique, des chaînes de Markov *réversibles*.

**Définition VI.5.4.** (Réseaux charismatiques)

On dit que le réseau  $(VE, K)$  est *charismatique* s'il existe un champ  $m = (m_x) \in ]0, +\infty[^V$  tel que, pour tous  $x, y \in V$ ,

$$m_x K_{xy} = m_y K_{yx}. \quad (\text{VI.5.5})$$

**Remarque VI.5.5.** On vérifie immédiatement que les réseaux charismatiques sont un cas particulier de flots de gradient. En effet, si l'on introduit la matrice  $C = (C_{xy})$ , avec  $C_{xy} = m_x K_{xy}$ , la matrice  $C$  est symétrique, et l'on a, avec  $M = \text{diag}(m_x)$ ,

$$C = MK \implies A = I - K = I - M^{-1}C = M^{-1}(M - C) = M^{-1}B,$$

où  $M$  est s.d.p. et  $B = M - C$  est symétrique, on est donc bien dans le cas d'un flot de gradient pour la métrique induite par  $M$ , dans le cas d'une matrice  $M$  diagonale. On a de plus

$$m_x K_{xy} = C_{xy} \implies \sum_{y \sim x} m_x K_{xy} = m_x = \sum_{y \sim x} C_{xy}.$$

<sup>12</sup>. Ce produit scalaire fait de la famille des vecteurs colonnes de  $P$  une base orthonormée.

**Remarque VI.5.6.** Si nous considérons  $K$  comme la matrice de transition d'une chaîne de Markov sur l'ensemble  $V$ , où  $K_{xy}$  est la probabilité de transition de  $x$  à  $y$ , alors la définition VI.5.4 correspond à celle d'une chaîne *reversible*, et  $m = (m_x)$  joue le rôle d'une mesure invariante.

**Remarque VI.5.7.** Si le réseau est charismatique,  $K_{xy} > 0$  si et seulement si  $K_{yx} > 0$ . En conséquence  $x \rightarrow y$  si et seulement si  $y \rightarrow x$ . On considèrera néanmoins le graphe sous-jacent comme orienté, étant entendu que  $(x, y) \in E \iff (y, x) \in E$ , mais l'on ne fait pas l'identification entre les deux arêtes. Par ailleurs, tout influenceur  $x \in \Gamma$  est isolé, et ne joue donc aucun rôle dans la dynamique d'opinion. Si l'on se restreint à des réseaux connexes, il ne peut donc pas y avoir d'influenceurs.

**Remarque VI.5.8.** L'identité (VI.5.5) peut s'écrire

$$K_{xy} = \frac{m_y}{m_x} K_{yx}.$$

L'influence que  $y$  exerce sur  $x$  dépend donc du rapport des charismes de  $x$  et de  $y$ , et de l'influence que  $x$  exerce sur  $y$ . Plus le charisme de  $y$  est grand comparé à celui de  $x$ , plus  $y$  influence  $x$  comparé à l'influence de  $x$  sur  $y$ . C'est ce qui justifie l'appellation *charisma* : plus le charisme est grand, plus l'influence exercée sur les autres est grande.

**Proposition VI.5.9.** Soit  $(V, E, K)$  un réseau charismatique. Si le réseau est connexe, alors le charisme est défini de façon unique à constante multiplicative près. En particulier il admet un unique charisme  $m = (m_x)$  qui est une loi de probabilité sur  $V$ , i.e. tel que

$$\sum_V m_x = 1.$$

*Démonstration.* Notons en premier lieu que, d'après la remarque VI.5.7, la connexité entraîne la forte connexité. Considérons  $m$  et  $m'$  deux champs de charisme sur  $V$ . Soit  $x \in V$  arbitraire. On pose  $\lambda = m'_x/m_x > 0$ . Pour tout  $y \sim x$ , on a

$$m'_y = m'_x \frac{K_{xy}}{K_{yx}} = m'_x \frac{m_y}{m_x} = \lambda m_y.$$

Cette relation de proportionnalité se propage de proche en proche, donc en tous les sommets par connexité du graphe, on a donc  $m' = \lambda m$ . Il existe donc en particulier un unique champ de charisme de masse totale unitaire.  $\square$

**Remarque VI.5.10.** A cardinal de  $V$  fixé, on peut établir une relation de bijection entre les réseaux charismatiques et l'ensemble des matrices symétriques à coefficient positifs ou nul, à constante positive multiplicative près. En effet, si  $K$  et  $m$  satisfont les relations (VI.5.5) alors la matrice  $C$  définie par  $C_{xy} = m_x K_{xy}$ , est symétrique. Si l'on note  $M$  la matrice  $\text{diag}(m_x)$ , on peut écrire  $K = M^{-1}C$ . Réciproquement, si  $C$  est une matrice symétrique dont les éléments sont positifs, et que l'on souhaite lui associer une matrice  $K = M^{-1}C$  encodant les influences d'un réseau charismatique, le seul choix possible est, étant donnée la contrainte de normalisation des lignes de  $K$ ,

$$m_x = \sum_y C_{xy}.$$

Si l'on prend pour  $M$  la matrice diagonale de coefficients  $(m_x)_{x \in V}$ , alors  $K = M^{-1}C$  correspond à un réseau charismatique.

Les réseaux charismatiques présentent une propriété de conservation particulière. Nous avons noté (voir remarque ??) que l'opinion totale n'est en général pas conservée. Dans le cas des réseaux charismatiques, une certaine quantité est pourtant conservée, il s'agit d'une certaine moyenne de l'opinion, plus précisément de l'espérance de l'opinion relativement à la mesure  $(m_x)$ .

**Proposition VI.5.11.** (Propriété de conservation)

Soit  $(V, E, K, m)$  un réseau charismatique, et  $(u^k)$  la suite des opinions associés au modèle (VI.5.2). L'opinion moyenne relativement à la mesure  $m$ , définie par

$$\bar{u}^k = \sum_{x \in V} m_x u_x^k,$$

se conserve au cours des itérations.

*Démonstration.* On a

$$\begin{aligned} \bar{u}^{k+1} &= \sum_{x \in V} m_x u_x^{k+1} = \sum_{x \in V} \sum_{y \leftarrow x} m_x K_{xy} u_y^k = \sum_{x \in V} \sum_{y \leftarrow x} m_y K_{yx} u_y^k \\ &= \sum_{y \in V} m_y u_y^k \sum_{x \leftarrow y} K_{yx} = \sum_{y \in V} m_y u_y^k = \bar{u}^k, \end{aligned} \quad (\text{VI.5.6})$$

qui établit la propriété de conservation annoncée.  $\square$

Cette propriété permet de caractériser les limites possibles du problème d'évolution.

**Proposition VI.5.12.** Soit  $(V, E, K, m)$  un réseau charismatique. On suppose que la suite des itérés du modèle discret converge vers un consensus associé à la valeur  $u^\infty$ . Alors cette valeur  $u^\infty$  correspond à la moyenne des opinions initiales relativement au charisme  $m$  normalisé à 1 selon la proposition VI.5.9 :

$$u^\infty = \sum_{x \in V} m_x u_x^0.$$

*Démonstration.* Si toutes les opinions  $u_x^k$  convergent vers  $u^\infty$ , on a, d'après la proposition VI.5.11,

$$\sum_{x \in V} m_x u_x^0 = \sum_{x \in V} m_x u_x^k \longrightarrow \sum_{x \in V} m_x u^\infty = u^\infty.$$

qui montre que l'opinion limite commune est bien la combinaison barycentrique des opinions initiales, pondérée par les charismes des agents.  $\square$

**Point de vue variationnel, flot de gradient & réseaux résistifs**

Une autre particularité du cadre charismatique est que le problème présente une structure variationnelle. Considérons un réseau charismatique  $(V, E, K)$ , de charisme normalisé  $m$ . Comme décrit dans la section ?? (voir aussi la remarque VI.5.5), nous sommes dans la situation où  $A$  s'écrit  $M^{-1}B$ , avec  $M = \text{diag}(m_x)$ , et  $B = M - C$ . Le problème d'évolution continu en temps est donc, d'après la proposition VI.5.2, un flot de gradient pour la fonctionnelle

$$\Psi(u) = \frac{1}{2} \langle (M - C)u | u \rangle = \frac{1}{2} \sum_x u_x \sum_{y \sim x} C_{xy} (u_x - u_y) = \frac{1}{2} \sum_{e \in E} C_{xy} (u_x - u_y)^2, \quad (\text{VI.5.7})$$

pour le produit scalaire défini par

$$\langle u | v \rangle_M = \sum_{x \in V} m_x u_x v_x.$$

En conséquence, l'énergie  $\Psi$  décroît au cours du temps, et le modèle exprime un principe d'évolution selon la ligne de plus grande pente vis à vis de  $\Psi$ , pour la métrique définie par  $M$ .

Cette énergie permet de faire un lien avec les réseaux résistifs. On peut penser  $u_x$  comme un potentiel en  $x$ , la quantité  $C_{xy} = m_x K_{xy}$  (qui est symétrique en  $x, y$ ) jouant le rôle d'une *conductance* (inverse d'une résistance) de l'arête (symétrique selon ce point de vue) joignant  $x$  et  $y$ . La quantité  $\Phi(u)$  correspond dans

cette analogie électrique à (la moitié de) l'énergie dissipée au sein du réseau aux conductances  $m_x K_{xy}$  et aux potentiels  $u_x$ . L'évolution tend donc à minimiser cette énergie dissipée, que l'on peut voir comme une estimation de l'écart à l'équilibre en termes d'opinion. Dans cette optique, les paramètres  $C_{xy}$  peuvent être interprétés comme des *coefficients de friction*, et les  $u_x$  comme des vitesses<sup>13</sup>.

En poursuivant cette analogie avec les systèmes mécaniques, concevoir l'opinion d'un sommet  $x$  comme une quantité scalaire de type vitesse, la quantité obtenue par multiplication par la "masse"  $m_x$ , donne une quantité de mouvement-opinion. On a bien un principe de Newton pour ce système mécanique : d'après la proposition VI.5.11, la quantité de mouvement-opinion globale pour ce système libre (non forcé de l'extérieur) se conserve. Le carré de la norme naturellement associée au modèle correspond à une énergie cinétique. On prendra garde en revanche que l'énergie globale  $\Phi$  dont dérive l'équation d'évolution, quadratique en les vitesses, n'a rien d'une énergie cinétique, il s'agit plutôt comme indiqué ci-dessus d'une somme de termes de nature frictionnelle, qui quantifieraient des puissances dissipées au niveau de chaque arête (relation entre deux individus), d'autant plus que les opinions divergent. Cette interprétation est étayée par un pseudo-bilan énergétique que l'on peut obtenir à partir de l'équation de conservation de la quantité de mouvement, on effectue le produit scalaire de

$$M \frac{du}{dt} = -\nabla \Psi(u)$$

avec la "vitesse"  $u$ , pour obtenir

$$\frac{d}{dt} \frac{1}{2} \langle Mu | u \rangle = -\langle \nabla \Psi(u) | u \rangle = \sum_{e \in E} C_{xy} |u_x - u_y|^2,$$

qui peut se lire : la dérivée en temps de l'énergie cinétique est égale à la puissance dissipée par friction entre opinions différentes. S'il s'agissait d'un système mécanique standard, cette énergie serait dissipée sous forme de chaleur au sein du système ou vers le monde extérieur (l'ajout d'un modèle thermique permettrait de préciser le devenir de cette énergie thermique).

## VI.6 Modèles macroscopiques de trafic routier

On considère l'évolution d'une population de piétons ou de véhicules sur une voie rectiligne, population représentée par une densité linéique  $\rho(x, t)$ . On considère que la vitesse des entités est fonction de la densité :  $v = v(\rho)$ . La manière la plus simple de prendre en compte le fait que la vitesse est d'autant plus faible que la densité est importante est  $v(\rho) = U(1 - \rho/\rho_{\max})$ . La conservation de la masse s'écrit alors

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} (\rho v(\rho)) = 0,$$

qui a la forme d'une équation de conservation que l'on peut écrire sous forme générale

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} f(\rho) = 0, \tag{VI.6.1}$$

où  $f$  est le *flux*.

### Propagation des perturbations

Si l'on considère une solution stationnaire  $\rho_{eq}$  de l'équation, et une solution perturbée  $\rho_{eq} + h$ , on obtient formellement une équation de transport sur la perturbation :

$$\partial_t h + f'(\rho_{eq}) \partial_x h = 0 \tag{VI.6.2}$$

qui exprime que les perturbations sont transportées à la vitesse  $f'(\rho_{eq})$ .

13. Comme deux objets en contact, allant à des vitesses différentes, sont soumis à une force d'interaction de type friction proportionnelle à leur vitesse relative, dissipant ainsi une énergie proportionnelle au carré de cette vitesse relative.

Supposons que  $\rho(x, t)$  est une solution régulière de cette équation. On appelle courbe caractéristique une courbe  $t \mapsto x(t)$  telle que

$$\dot{x}(t) = f'(\rho(x(t), t)).$$

On vérifie immédiatement que  $\rho$  est constant le long de telles courbes :

$$\frac{d}{dt}\rho(x(t), t) = \partial_t \rho(x(t), t) + \dot{x}(t) \partial_x \rho(x(t), t) = \partial_t \rho(x(t), t) + f'(\rho(x(t), t)) \partial_x \rho(x(t), t) = 0.$$

Comme  $\rho$  est constant le long de la caractéristique, la célérité (fonction de cette seule densité) elle-même est constante, et l'on a

$$t \mapsto x + t f'(\rho_0(x)).$$

Si l'on se donne une densité initiale  $\rho_0$ , on peut ainsi construire la solution associée en reportant la valeur de densité initiale le long des caractéristiques. Cette démarche n'est évidemment possible que tant que les caractéristiques ne se croisent pas.

Pour une densité initiale donnée, supposée lisse (continûment différentiable), on peut considérer le flot associé aux caractéristiques

$$\Phi_t : x \mapsto x + f'(\rho_0(x_0))t.$$

Si l'on suppose que la fonction  $f$  est  $C^2$ , on peut calculer le jacobien de la transformation

$$J(t, x) = 1 + t f''(\rho_0(x)) \rho'_0(x).$$

Ce Jacobien reste  $> 0$  (la transformation est un difféomorphisme, i.e. les trajectoires ne se croisent pas) pour tout  $t$  si  $f''(\rho_0(x)) \rho'_0(x) \geq 0$ . Si en revanche cette dernière quantité est négative, alors l'application ne sera régulière que pour

$$t < -\frac{1}{f''(\rho_0(x)) \rho'_0(x)}.$$

Le temps de vie de la solution lisse sera donc

$$T = \frac{1}{\max |(f''(\rho_0(x)) \rho'_0(x))_-|}$$

(inverse du max de la partie négative de  $f''(\rho_0(x)) \rho'_0(x)$ ).

Si l'on considère le flux indiqué précédemment  $f(\rho) = U\rho(1 - \rho/\rho_{\max})$ , on a  $f''(\rho) = -2U/\rho_{\max} < 0$ . On aura donc existence de solution lisse si  $\rho_0$  est décroissante, et croisement de caractéristique en temps fini si en revanche  $\rho_0$  est croissante.

**Remarque VI.6.1.** On prendra garde au fait que, bien que l'on ait considéré le Jacobien de l'application  $\Phi_t$ , ce qui suggère un transport de mesure, n'est aucunement associée à un quelconque transport conservatif de masse.

## Solutions faibles

Les considérations précédentes indiquent qu'il ne peut, en général, exister de solution lisse globale. Pour donner un sens aux solutions non lisses qui sont susceptibles d'apparaître spontanément, on définit la notion de solution faible :

**Définition VI.6.2.** On dit que  $\rho \in L^1_{loc}(\mathbb{R} \times ]0, T[)$  est une solution faible de (VI.6.1) sur  $\mathbb{R} \times ]0, T[$  si  $f(\rho) \in L^1_{loc}(\mathbb{R} \times ]0, T[)$  et si, pour tout  $\varphi$ , fonction  $C^1$  à support compact dans  $\mathbb{R} \times ]0, T[$ , on a

$$\int_{\mathbb{R}} \int_0^T \partial_t \varphi \rho(x, t) dx dt + \int_{\mathbb{R}} \int_0^T \partial_x \varphi f(\rho(x, t)) dx dt = 0.$$

On peut intégrer une condition initiale à cette définition. On dira que  $\rho$  est solution faible associée à la condition initiale  $\rho|_{t=0} = \rho^0 \in L^1_{loc}(\mathbb{R})$  si

$$\int_{\mathbb{R}} \int_0^T \partial_t \varphi \rho(x, t) dx dt + \int_{\mathbb{R}} \int_0^T \partial_x \varphi f(\rho(x, t)) dx dt + \int_{\mathbb{R}} \varphi(x, 0) \rho^0(x) dx = 0$$

pour toute fonction  $\varphi$  régulière à support compact dans  $\mathbb{R} \times [0, T[$

On vérifie immédiatement que toute solution régulière est solution faible. Mais cette définition peut s'appliquer à des solutions qui ne sont pas régulières. Considérons par exemple deux densités qui réalisent le même flux :  $F = f(\rho_-) = f(\rho_+)$ . La densité

$$\rho = \rho_- \mathbb{1}_{]-\infty, 0[} + \rho_+ \mathbb{1}_{]0, +\infty[}$$

est solution faible stationnaire de (VI.6.1), de même que la densité obtenue en intervertissant  $\rho_-$  et  $\rho_+$ . On peut construire des solutions non stationnaires de la façon suivante : on se donne deux densités  $\rho_L$  et  $\rho_R$ , et l'on cherche une solution  $\rho$  constante de part et d'autre d'un point de discontinuité  $s(t)$  variable en temps. On vérifie qu'une telle densité est solution faible dès que  $s$  vérifie une condition dite de *Rankine-Hugoniot*, comme l'exprime la

**Proposition VI.6.3.** (Relation de Rankine-Hugoniot)

On suppose la fonction flux  $f$  continue sur son intervalle de définition, et  $\rho_L$  et  $\rho_R$  deux valeurs sur cet intervalle. La densité

$$\rho = \rho_L \mathbb{1}_{]-\infty, s(t)[} + \rho_R \mathbb{1}_{]s(t), +\infty[}$$

est solution faible de (VI.6.1) si et seulement si la discontinuité  $s$  progresse à la vitesse constante

$$\dot{s} = \frac{f(\rho_L) - f(\rho_R)}{\rho_L - \rho_R}. \quad (\text{VI.6.3})$$

*Démonstration.* On utilise la définition d'une solution faible, en écrivant la première intégrale double

$$\int_{\mathbb{R}} \int_0^{+\infty} \partial_t \varphi \rho = \int_0^{+\infty} \left( \rho_L \int_{-\infty}^{s(t)} \partial_t \varphi + \rho_R \int_{s(t)}^{+\infty} \partial_t \varphi \right),$$

avec

$$\int_{-\infty}^{s(t)} \partial_t \varphi = \frac{d}{dt} \left( \int_{-\infty}^{s(t)} \varphi \right) - \dot{s}(t) \varphi(s(t), t), \quad \int_{s(t)}^{+\infty} \partial_t \varphi = \frac{d}{dt} \left( \int_{s(t)}^{+\infty} \varphi \right) + \dot{s}(t) \varphi(s(t), t).$$

La seconde intégrale double (avec la dérivée en espace sur la fonction test s'écrit

$$\begin{aligned} \int_{\mathbb{R}} \int_0^{+\infty} \partial_x \varphi f(\rho(x, t)) &= \int_0^{+\infty} \left( f(\rho_L) \int_{-\infty}^{s(t)} \partial_x \varphi + f(\rho_R) \int_{s(t)}^{+\infty} \partial_x \varphi \right) \\ &= \int_0^{+\infty} \varphi(s(t), t) (f(\rho_L) - f(\rho_R)). \end{aligned}$$

On obtient donc finalement

$$\int_0^{+\infty} \varphi(s(t), t) (-\dot{s}(t)(\rho_L - \rho_R) + f(\rho_L) - f(\rho_R)),$$

qui est identiquement nul pour toute fonction test  $\varphi$  si et seulement si la condition (VI.6.3) est identiquement vérifiée.  $\square$

**Remarque VI.6.4.** On peut retrouver la relation (VI.6.3) en écrivant simplement un bilan de masse au voisinage de la discontinuité.

**Remarque VI.6.5.** On peut voir cette formule comme la généralisation de la formule donnant la vitesse de propagation de perturbations au voisinage d'une densité uniforme, en prenant  $\rho_R = \rho_L + \varepsilon$ , ce qui donne  $\dot{s} \approx f'(\rho_L)$ .

On peut vérifier que, sous sa forme faible, l'équation n'est pas bien posée, au sens où elle admet en général plusieurs solutions. La théorie complète de telles équation dépasse le cadre de ce cours sous sa forme actuelle, disons simplement ici qu'il est possible d'imposer à la solution considérer de vérifier un critère



supplémentaire, dit *d'entropie*, qui permet de sélectionner la solution physique<sup>14</sup> parmi les nombreuses possibles. Ce critère n'est pertinent que pour discriminer des solutions qui présentent des discontinuités, on peut montrer que ces solutions acceptables sont telles que, lorsque la solution présente une discontinuité, les courbes caractéristiques doivent arriver vers la discontinuité, et non pas en partir. Le développement précédent donnant la vitesse de propagation de la discontinuité en fonction des états à gauche et à droite, on peut exprimer le fait que les caractéristiques vont vers la discontinuité de la façon suivante :

**Définition VI.6.6.** Soit  $\rho(x, t)$  une solution faible de l'équation de conservation (VI.6.1), avec  $f(\cdot)$  une fonction  $C^1$ , au sens de la définition VI.6.2. On suppose que  $\rho$  présente localement (au voisinage d'un point de l'espace temps) une discontinuité entre les valeurs  $\rho_L$  et  $\rho_R$ . On dit que cette discontinuité vérifie la condition d'entropie de Lax si

$$f'(\rho_L) > \frac{f(\rho_R) - f(\rho_L)}{\rho_R - \rho_L} > f'(\rho_R).$$

On notera que, dans le cas où  $f$  est convexe (ou  $f$  concave), la condition ci-dessus peut se limiter à l'inégalité entre les bornes.

## VI.7 Autour de la notion de complétude, théorème de Banach-Steinhaus

### VI.7.1 Lemme de Baire

**Définition VI.7.1.** (Parties maigres)

Soit  $X$  un espace topologique, et  $A$  une partie de  $X$ . On dit que  $A$  est maigre si  $A$  est contenue dans une réunion dénombrable de fermés d'intérieur vide.

**Définition VI.7.2.** (Espaces de Baire)

Un espace topologique est appelé espace de Baire si toute partie maigre est d'intérieur vide.

**Exercice VI.7.1.** Vérifier que l'espace des suites qui s'annulent au-delà d'un certain rang, muni de la norme  $\ell^\infty$ , n'est pas un espace de Baire.

Le théorème suivant traduit le fait que tout espace métrique complet est de Baire.

**Théorème VI.7.3.** (Lemme de Baire)

Soit  $X$  un espace métrique complet, et  $(X_n)_{n \in \mathbb{N}}$  une suite de fermés de  $X$ . On suppose que

$$\text{Int}(X_n) = \emptyset \quad \forall n \in \mathbb{N}.$$

On a alors

$$\text{Int}\left(\bigcup_{n=0}^{+\infty} X_n\right) = \emptyset.$$

On utilise pour la démonstration une formulation équivalente du théorème : soit  $X$  un espace métrique complet, et  $(U_n)_{n \in \mathbb{N}}$  une suite d'ouverts de  $X$  denses dans  $X$ . Alors l'intersection des  $U_n$  est dense dans  $X$ .

*Démonstration.* On introduit

$$U = \bigcap_{n \in \mathbb{N}} U_n,$$

---

14. Ce type de critère a été élaboré dans le cadre de la dynamique des gaz. Précisons que, dans le cadre du transport d'entités vivantes, sa légitimité est moins nette

et on se donne  $x \in X$ . Pour toute boule  $B(x, r)$ , on va construire une suite  $a_n$  qui converge vers une limite  $a$  dans  $B(x, r) \cap U$ , ce qui établira la densité de  $U$ . Comme  $U_o$  est un ouvert dense, il existe  $a_o \in U_o$  et  $r_o > 0$ , avec  $r_o \leq r/2$ , tel que

$$\overline{B}(a_o, r_o) \subset B(x, r).$$

On construit par récurrence la suite  $(a_n)$  de la façon suivante : supposons  $a_k$  et  $r_k$  construits pour  $k \leq n$ , la densité de l'ouvert  $U_{n+1}$  assure l'existence d'une boule fermée  $\overline{B}(a_{n+1}, r_{n+1})$  incluse dans  $U_{n+1} \cap B(a_n, r_n)$ , et telle que  $r_{n+1} \leq r_n/2$ . Les suites  $(a_n)_{n \in \mathbb{N}}$  et  $(r_n)_{n \in \mathbb{N}}$  ainsi construites, on vérifie immédiatement que

$$d(a_n, a_{n+1}) < r_n \leq \frac{r}{2^n},$$

d'où l'on déduit que  $(a_n)_{n \in \mathbb{N}}$  est de Cauchy, donc qu'elle converge vers une limite  $a \in X$ . Par construction,  $a$  est dans  $B(x, r)$  et dans chacune des boules fermées  $\overline{B}(a_n, r_n)$ , il est donc dans  $U$ , ce qui termine la démonstration.  $\square$

**Exercice VI.7.2.** Soit  $E$  un e.v.n. de dimension infinie. Montrer que si  $E$  est complet, alors sa dimension est non dénombrable, *i.e.* il n'existe pas de famille génératrice dénombrable.

### VI.7.2 Théorème de Banach Steinhaus

**Définition VI.7.4.** Soient  $E$  et  $F$  deux espaces vectoriels normés. On note  $\mathcal{L}(E, F)$  l'espace des applications linéaires **continues** de  $E$  dans  $F$ . C'est un e.v.n. muni de la norme

$$\|f\|_{\mathcal{L}(E, F)} = \sup_{x \neq 0} \frac{\|f(x)\|_F}{\|x\|_E}.$$

De plus, si  $F$  est complet, alors  $\mathcal{L}(E, F)$  est un espace de Banach.

**Théorème VI.7.5.** (Banach-Steinhaus)

Soient  $E$  et  $F$  deux espaces de Banach et  $(T_\alpha)_{\alpha \in A}$  une famille (non nécessairement dénombrable) d'opérateurs de  $\mathcal{L}(E, F)$ . On suppose

$$\sup_{\alpha \in A} \|T_\alpha x\|_F < +\infty \quad \forall x \in E.$$

On a alors

$$\sup_{\alpha \in A} \|T_\alpha\|_{\mathcal{L}(E, F)} < +\infty.$$

Ce théorème est parfois appelé “principe de la borne uniforme”. En effet, on déduit de majorations ponctuelles une majoration uniforme : ce théorème assure l'existence d'une constante  $c > 0$  telle que

$$\|T_\alpha x\|_F \leq c \|x\|_E \quad \forall x \in E \quad \forall \alpha \in A.$$

*Démonstration.* On introduit les ensembles

$$E_n = \left\{ x \in E, \sup_{\alpha \in A} \|T_\alpha x\|_F \leq n \right\}.$$

Les  $E_n$  sont des fermés de  $E$  comme intersection de fermés. D'autre part leur réunion est  $E$  tout entier, d'après l'hypothèse. L'un des  $E_n$  est donc d'intérieur non vide. Soit  $n_o$  tel que  $\text{Int}(E_{n_o}) \neq \emptyset$ . Il existe  $a \in E$  et  $\rho > 0$  tel que  $\overline{B}_\rho(a) \subset E_{n_o}$ , d'où

$$\|T_\alpha(a + \rho u)\| \leq n_o \quad \forall u \in B_E.$$

On a donc, pour tout  $\alpha \in A$ ,

$$\|T_\alpha\|_{\mathcal{L}(E, F)} \leq \frac{1}{\rho} \left( n_o + \sup_{\alpha \in A} \|T_\alpha(a)\|_F \right),$$

ce qui conclut la démonstration.  $\square$

Le théorème de Banach-Steinhaus est souvent utilisé sous la forme suivante :

**Corollaire VI.7.6.** Soient  $E$  et  $F$  deux Banach, et  $(T_n)_{n \in \mathbb{N}}$  une suite d'opérateurs de  $\mathcal{L}(E, F)$  telle que, pour tout  $x \in E$ ,  $T_n x$  converge vers un élément de  $F$ , que l'on note  $Tx$ . On a alors

1.  $\sup_{n \in \mathbb{N}} \|T_n\| < +\infty$ ,
2.  $T \in \mathcal{L}(E, F)$ ,
3.  $\|T\|_{\mathcal{L}(E, F)} \leq \liminf \|T_n\|_{\mathcal{L}(E, F)}$ .

La dernière inégalité du corollaire précédent peut être stricte. Considérer par exemple  $E = \ell^p$  avec  $p \in [1, +\infty[$ , et la suite des formes linéaires

$$T_k : x = (x_n)_{n \in \mathbb{N}} \mapsto x_k \in \mathbb{R}.$$

**Corollaire VI.7.7.** Soit  $G$  un espace vectoriel normé. Un sous-ensemble  $B$  de  $G$  est borné si et seulement si

$$f(B) = \{\langle f, x \rangle, x \in B\} \text{ est borné } \quad \forall f \in G'.$$

*Démonstration.* On applique le théorème VI.7.5 avec  $E = G'$ ,  $F = \mathbb{R}$ , et la famille d'applications, indexée par  $B$  lui-même,  $(T_x)_{x \in B}$  :

$$f \in G' \mapsto \langle f, x \rangle.$$

Le théorème assure l'existence d'une constante  $c$  telle que

$$|\langle f, x \rangle| \leq c \|f\| \quad \forall x \in B \quad \forall f \in G'.$$

Comme  $\|x\| = \sup_{f \in B_{G'}} \langle f, x \rangle$  (où  $B_{G'}$  est la boule unité fermée de  $G'$ ),  $\|x\| \leq c$  sur  $B$ . □

**Remarque VI.7.8.** Le corollaire précédent ne nécessite pas l'hypothèse  $G$  complet. L'espace qui joue le rôle de l'espace de départ dans le théorème de Banach-Steinhaus est le dual topologique de  $G$ , qui est toujours complet.

### VI.7.3 Théorèmes de l'application ouverte, du graphe fermé

**Théorème VI.7.9.** (Application ouverte)

Soient  $E$  et  $F$  deux espaces de Banach et soit  $T \in \mathcal{L}(E, F)$  surjectif. Alors il existe une constante  $c$  telle

$$B(0, c) \subset T(B_E).$$

La conclusion signifie que  $T$  transforme tout ouvert en un ouvert.

**Démonstration:** La démonstration s'effectue en deux étapes. On montre dans un premier temps que l'adhérence de  $T(B_E)$  contient une boule ouverte centrée en 0. On note  $C = \overline{T(B_E)}$  cette adhérence. Les  $C_n = nC$  sont des fermés de  $F$  par construction, et leur union est  $F$  tout entier (car  $T$  est surjective). L'espace d'arrivée  $F$  étant complet, il en existe donc un d'intérieur non vide (d'après le lemme de Baire), donc  $C$  lui-même est d'intérieur non vide (les  $C_n$  sont homothétiques à  $C$ ) :  $C$  contient une boule ouverte  $B = B(a, 2c)$ . Par symétrie de  $C$ ,  $-B$  est également dans  $C$ , et par convexité ( $C$  est l'adhérence de l'image d'un convexe par une application linéaire),

$$\begin{aligned} \frac{1}{2}B + \frac{1}{2}(-B) &= \left\{ \frac{1}{2}a - \frac{1}{2}a + \frac{1}{2}h_1 - \frac{1}{2}h_2, h_1, h_2 \in B(0, 2c) \right\} \\ &= \{0 + h, h \in B(0, 2c)\} \\ &= B(0, 2c) \subset C = \overline{T(B_E)}. \end{aligned}$$

On va maintenant montrer, en utilisant cette fois la complétude de l'espace de départ  $E$ , que  $C$  contient  $B(0, c)$ . On se donne  $y \in B(0, c)$ , et on cherche à construire un antécédent  $x$  dans  $B_E$ . D'après (VI.7.1), pour tout  $\varepsilon > 0$ , il existe  $z$  dans  $1/2B_E$  tel que  $\|y - Tz\| < \varepsilon$ . On construit ainsi  $z_1$  tel que

$$\|y - Tz_1\| < \frac{c}{2}, \quad \|z_1\| \leq \frac{1}{2}.$$

On construit de la même manière  $z_2$

$$\|(y - Tz_1) - Tz_2\| < \frac{c}{4}, \quad \|z_2\| \leq \frac{1}{4},$$

puis par récurrence les termes de la suite  $(z_n)$  tels que

$$\|(y - Tz_1 - \dots - Tz_{n-1}) - Tz_n\| < \frac{c}{2^n}, \quad \|z_n\| \leq \frac{1}{2^n}.$$

Par construction la suite  $x_n = z_1 + \dots + z_n$  est de Cauchy, donc converge vers un certain  $x$  de norme  $\leq 1$ , et on a bien  $y = Tx$  par continuité de  $T$ .  $\square$

On en déduit le

**Corollaire VI.7.10.** Soient  $E$  et  $F$  deux Banach et soit  $T \in \mathcal{L}(E, F)$  bijectif. Alors  $T^{-1}$  est continu de  $F$  dans  $E$ .

**Démonstration:** Comme  $T$  est surjectif, on a, d'après le théorème de l'application ouverte,  $\|Tx\| \geq c$  pour tout  $x$  de norme 1. On en déduit, par homogénéité,

$$\|Tx\| \geq c\|x\| \quad \forall x \in E,$$

d'où

$$\|T^{-1}y\| \leq \frac{1}{c}\|y\| \quad \forall y \in F.$$

$\square$

Dans le cas où  $T$  n'est pas surjectif, on peut néanmoins utiliser le théorème de l'application ouverte pour la même application vue comme un opérateur de  $E$  dans  $T(E)$ , sous réserve que  $T(E) \subset F$  soit un espace de Banach :

**Corollaire VI.7.11.** Soient  $E$  et  $F$  deux espaces de Banach, et  $T \in \mathcal{L}(E, F)$ . On suppose que l'image de  $T$  est fermée. Alors il existe  $\alpha > 0$  tel que

$$\forall y \in T(E), \exists x \in E, \|x\| \leq \alpha\|y\|, \quad y = Tx. \quad (\text{VI.7.1})$$

**Démonstration:** L'espace  $T(E)$  est fermé dans le complet  $F$ , il est donc complet. On note  $T_1$  l'application  $T$  vue comme surjection de  $E$  dans  $T(E)$ . On peut appliquer le théorème de l'application ouverte à  $T_1$  :

$$\exists c > 0, \quad B_{T(E)}(0, c) \subset T_1(B_E).$$

Pour tout  $y \in T(E)$ , on a  $cy/2\|y\| \in B_{T(E)}(0, c)$ . Il existe donc  $x \in B_E$  tel que

$$T_1x = Tx = \frac{cy}{2\|y\|},$$

d'où

$$T\left(\frac{2\|y\|}{c}x\right) = y, \quad \left\|\frac{2\|y\|}{c}x\right\| \leq \frac{2}{c}\|y\|,$$

ce qui termine la démonstration ( $\alpha = 2/c$ ). On remarque que, si  $T$  est de plus injectif, ce corollaire exprime simplement la continuité de l'application réciproque  $T^{-1}$  définie sur  $T(E)$ , conformément au corollaire précédent.  $\square$

**Définition VI.7.12.** (Graphe)

Soient  $E$  et  $F$  deux e.v.n et  $T$  un opérateur linéaire de  $E$  vers  $F$ . On appelle graphe de  $T$  l'ensemble

$$G(T) = \{(x, Tx), x \in E\} \subset E \times F.$$

**Théorème VI.7.13.** (Graphe fermé)

Soient  $E$  et  $F$  deux Banach, et  $T$  un opérateur linéaire de  $E$  vers  $F$ . On a

$$G(T) \text{ fermé} \iff T \in \mathcal{L}(E, F).$$

**Démonstration:** La condition suffisante est immédiate. Supposons maintenant  $G(T)$  fermé. On considère la nouvelle norme sur  $E$  :

$$\|x\|_1 = \|x\|_E + \|Tx\|_F.$$

Montrons que  $E$  est complet pour la norme  $\|\cdot\|_1$ . Toute suite de Cauchy  $(x_n)_{n \in \mathbb{N}}$  pour  $\|\cdot\|_1$  est de Cauchy pour  $\|\cdot\|_E$ , donc converge vers  $x$  dans  $E$ . De la même manière,  $(Tx_n)$  converge vers  $y \in F$ . Comme  $G(T)$  est fermé, on a  $y = Tx$ , et  $(x_n)_{n \in \mathbb{N}}$  converge bien vers  $x$  pour  $\|\cdot\|_1$ . Comme  $(E, \|\cdot\|_1)$  est complet, d'après le corollaire VI.7.10, l'identité  $(E, \|\cdot\|_1) \rightarrow (E, \|\cdot\|_E)$  est bicontinue. La norme  $\|\cdot\|_1$  est donc équivalente à la norme  $\|\cdot\|_E$ , d'où la continuité de  $T$ .  $\square$

**Exercice VI.7.3.** Soit  $E$  un espace de Banach, et  $T$  une application linéaire de  $E$  dans  $E'$  telle que

$$(Tx, x) \geq 0 \quad \forall x \in E.$$

Montrer que  $T$  est continue. (On pourra utiliser le théorème du graphe fermé : on considère  $(x_n, T(x_n)) \rightarrow (x, f)$ , et on cherche à montrer que  $(x, f) \in G(T)$ . On pourra montrer dans un premier temps que  $(f - Tz, x - z) \geq 0$  pour tout  $z \in E$ , puis prendre  $z$  de la forme  $z = x + \lambda h$ .)

