ANNALES DE L'INSTITUT FOURIER

PATRICK ASSOUAD

Densité et dimension

Annales de l'institut Fourier, tome 33, nº 3 (1983), p. 233-282
<http://www.numdam.org/item?id=AIF_1983__33_3_233_0>

Ann. Inst. Fourier, Grenoble
33, 3 (1983), 233-282

DENSITÉ ET DIMENSION

par Patrick ASSOUAD

Introduction.

1. Je vais essayer d'indiquer d'abord comment et pourquoi s'introduisent les *classes de Vapnik-Cervonenkis* :

– soit $\mathfrak{S}$ une classe de parties d'un ensemble infini X ; une façon assez naturelle de mesurer l'ordre de grandeur de $\mathfrak{S}$ est d'étudier le comportement à l'infini de la fonction $\Delta^{\mathfrak{S}}$:

$$r \longrightarrow \operatorname{Sup}\{|A \cap \mathfrak{S}| \mid A \subset X, |A| = r\}$$

(où $A \cap \mathfrak{S}$ désigne l'ensemble des parties de A de la forme $A \cap S$ avec $S \in \mathfrak{S}$) ;

– si $\mathfrak{S}$ est l'ensemble de toutes les parties de X, la fonction $\Delta^{\mathfrak{S}}$ est simplement l'exponentielle $r \longrightarrow 2^r$;

– si $\mathfrak{S}$ est l'ensemble $\binom{X}{n}$ de toutes les parties de X de cardinal n, la fonction $\Delta^{\mathfrak{S}}$ est un polynôme de degré n (car on a $\Delta_{\mathfrak{S}}(r) = \binom{r}{n}$ pour tout r) ;

– on appelle classes de Vapnik-Cervonenkis les classes $\mathfrak{S}$ telles que la fonction $\Delta^{\mathfrak{S}}$ soit majorée par un polynôme ; ce sont donc, en un certain sens, les *classes "polynomialement bornées" de parties de* X.

– la classe de tous les demi-espaces affines de $\mathbf{R}^n$ est un autre exemple $\left(\text{moins trivial que } \binom{X}{n}\right)$ de classe de Vapnik-Cervonenkis.

Les classes de Vapnik-Cervonenkis ont été introduites (comme ne l'indique pas tout à fait le nom que nous leur donnons) par Sauer [30] et, indépendamment, par Vapnik et Cervonenkis [39].

De plus ces classes ont effectivement un *intérêt en Calcul des Probabilités* : en effet on peut établir, pour chaque classe de Vapnik-Cervonenkis $\mathfrak{S}$ convenablement mesurable et chaque loi de probabilité sur la tribu engendrée par $\mathfrak{S}$, une loi des grands nombres (Vapnik, Cervonenkis [39]) et un théorème central limite (Dudley [14]) pour les lois empiriques ; lorsque $\mathfrak{S}$ est la classe des demi-espaces affines $]-\infty, t]$ de $\mathbf{R}$, les deux résultats précédents se réduisent respectivement au théorème de Glivenko-Cantelli et au théorème de Kolmogorov-Smirnov pour les fonctions de répartition empiriques.

2. Le présent travail est, bien que son titre ne l'indique pas, un *essai d'exposé de synthèse sur les classes de Vapnik-Cervonenkis.* C'est que nous définirons, pour chaque classe $\mathfrak{S}$ de parties d'un ensemble X, *quatre indices de densité,* la finitude d'un quelconque de ces quatre indices équivalant à ce que $\mathfrak{S}$ soit une classe de Vapnik-Cervonenkis (ces quatre indices sont, il faut y insister, *de nature combinatoire*).

Le caractère d'exposé systématique de ce travail m'entraînera donc souvent à présenter (et parfois à démontrer) des résultats d'autres auteurs. Il n'est donc pas inutile de préciser quelle est, pour l'essentiel, ma propre contribution :

a) j'introduis, pour chaque classe $\mathfrak{S}$ de parties d'un ensemble X, *deux indices de densité duaux,* notés respectivement $\mathrm{Dens}^*_X(\mathfrak{S})$ et $\mathrm{dens}^*(\mathfrak{S})$, liés au comportement de la fonction

$$\Delta_{\mathfrak{S}} : r \longrightarrow \mathrm{Sup}\{|\mathrm{At}(\mathfrak{C}|X)| \mid \mathfrak{C} \subset \mathfrak{S}, |\mathfrak{C}| = r\}$$

(où $\mathrm{At}(\mathfrak{C}|X)$ est l'ensemble des atomes de l'algèbre de parties de X engendrée par $\mathfrak{C}$) ; ces indices duaux possèdent les propriétés suivantes :

a1) la finitude $\mathrm{Dens}^*_X(\mathfrak{S})$ ou de $\mathrm{dens}^*(\mathfrak{S})$ équivaut à ce que $\mathfrak{S}$ soit une classe de Vapnik-Cervonenkis ;

a2) les deux indices de densité duaux de la classe $\mathfrak{S}$ sont exactement les indices de densité directs (voir ci-après) de la classe duale $\mathfrak{S}^*$;

b) les deux indices de densité directs, notés Dens($\mathcal{S}$) et dens($\mathcal{S}$), sont liés au comportement de la fonction $\Delta^{\mathcal{S}}$ (voir 1 ci-dessus); la terminologie "densité" et le premier indice Dens($\mathcal{S}$) proviennent de Sauer [30]; le second indice dens($\mathcal{S}$) (que j'appelle *densité réelle de* $\mathcal{S}$) se trouve implicitement dans Dudley [14], où il joue un rôle important; je développe systématiquement les *propriétés de la densité réelle*, en démontrant notamment les résultats suivants :

b1) on a toujours

$$\mathrm{dens}(\mathcal{S}_1 \cup \mathcal{S}_2) = \mathrm{Sup}(\mathrm{dens}(\mathcal{S}_1), \mathrm{dens}(\mathcal{S}_2));$$

notant de plus $\mathcal{S}_1 \underline{\times} \mathcal{S}_2$ l'ensemble des produits d'un élément de $\mathcal{S}_1$ par un élément de $\mathcal{S}_2$, on obtient

$$\mathrm{dens}(\mathcal{S}_1 \underline{\times} \mathcal{S}_2) \leqslant \mathrm{dens}(\mathcal{S}_1) + \mathrm{dens}(\mathcal{S}_2);$$

enfin, si on note $\mathcal{S}^{(k)}$ la réunion des tribus engendrées par k éléments de $\mathcal{S}$, on trouve $\mathrm{dens}(\mathcal{S}^{(k)}) \leqslant k\,\mathrm{dens}(\mathcal{S})$;

b2) la classe $\mathcal{A}_n$ de tous les demi-espaces affines fermés de $\mathbf{R}^n$ est de densité réelle n (alors que sa densité entière Dens ($\mathcal{A}_n$) est égale à $n+1$);

b3) par contre la classe $\mathcal{B}_{\mathrm{E}}$ de toutes les boules fermées d'un espace normé $\mathrm{E} = (\mathbf{R}^n, \|\ \|)$ est en général de densité infinie;

b4) cependant, si n est $\leqslant 2$, ou si E est un sous-espace normé de L^{2p} (avec p entier), alors $\mathcal{B}_{\mathrm{E}}$ est de densité réelle $n+1$;

b5) observons que la densité réelle coïncide, en (b2) et en (b4), avec le nombre "réel" de paramètres dont dépend la classe étudiée; mais il ne faudrait pas croire que la densité réelle est toujours à valeurs entières : je donne un exemple où elle vaut 3/2 (je montre par contre qu'elle ne peut jamais appartenir à $]0,1[$);

c) si P est une loi de probabilité sur la tribu $\mathcal{T}$ engendrée par $\mathcal{S}$, on note d_P l'écart $\mathrm{S},\mathrm{S}' \longrightarrow \mathrm{P}(\mathrm{S}\,\Delta\,\mathrm{S}')$ et $\overline{\dim}(\mathcal{S}, d_P)$ la *dimension d'entropie* de l'espace $(\mathcal{S}, d_P)$; alors la borne supérieure de $\overline{\dim}(\mathcal{S}, d_P)$, pour toutes les lois P sur $\mathcal{T}$, est égale à la densité réelle de la classe $\mathcal{S}$ (Dudley [14] avait montré qu'elle était inférieure ou égale); on peut donc déduire de ce qui précède des *calculs de dimension d'entropie*, dont voici un exemple : soit X un polygone convexe compact de $\mathbf{R}^2$, muni d'une distance d

continue et arguésienne, c'est à dire telle que

$$d(x\,,y) = d(x\,,z) + d(z\,,y)\,,$$

quels que soient $x\,,y \in \mathrm{X}$ et $z \in [x\,,y]$; la dimension d'entropie de l'espace $(\mathrm{X}\,,d)$ est alors inférieure ou égale à 2.

3. Voici donc le *plan de ce travail* :

§ 1 Densité réelle et densité entière (un paragraphe de définitions, incluant le lemme fondamental de Sauer-Vapnik-Cervonenkis),

§ 2 Propriétés des indices de densité (où on trouvera notamment a), b1) et b5)),

§ 3 Exemples de classes de Vapnik-Cervonenkis (avec en particulier b2), b3) et b4)

§ 4 Dimension et densité (où nous développerons le point (c) ci-dessus),

§ 5 $\mathfrak{S}$-indépendance et nombres de Radon (un Appendice sur deux notions liées à ce qui précède).

Remerciements. – Je tiens à remercier Jean Bretagnolle qui, ayant lu la première version, très défectueuse, de ce travail, m'a encouragé à la reprendre entièrement.

1. Densité réelle et densité entière.

Nous indiquons, dans ce paragraphe, la définition de divers indices de densité, dont les relations seront précisées au paragraphe 2.

Bien qu'on en trouve plusieurs démonstrations indépendantes dans la littérature, nous démontrons de plus le Lemme fondamental (le lemme 1.8 ci-dessous) qui est la base de la théorie des classes de Vapnik-Cervonenkis.

Fixons d'abord *quelques notations* :

1.1. – a) si X est un ensemble, on note 2^{X} l'ensemble de toutes les parties de X et $\binom{\mathrm{X}}{n}$ l'ensemble des parties de X de cardinal n ; si X est fini, son cardinal est noté $|\mathrm{X}|$; si X est infini, on pose $|\mathrm{X}| = +\infty$.

b) si $\mathfrak{S}$ est une partie de 2^X (c'est-à-dire une classe de parties de X) et si A est une partie de X, on note $A \cap \mathfrak{S}$ l'ensemble de toutes les parties de A de la forme $A \cap S$ avec $S \in \mathfrak{S}$ ($A \cap \mathfrak{S}$ est appelée la *trace de* $\mathfrak{S}$ *sur* A) ;

c) si Y est une partie de X et si $\mathcal{A} = \{S_1, \ldots, S_r\}$ est une partie finie de 2^X, on note $\mathrm{At}(\mathcal{A} \mid Y)$ l'ensemble des atomes de l'algèbre de parties de Y engendrée par $\mathcal{A} \cap Y$, c'est-à-dire l'ensemble des parties non vides de Y de la forme

$$\Big(\bigcap_{i \in I} S_i \cap Y\Big) \cap \Big(\bigcap_{j \in J} Y \backslash S_j\Big),$$

pour toutes les partitions (I, J) de $\{1, \ldots, r\}$ (les éléments de $\mathrm{At}(\mathcal{A} \mid Y)$ sont appelés les *atomes de* $\mathcal{A}$ *dans* Y).

Nous allons indiquer maintenant deux façons (non équivalentes) de calculer l'ordre "à l'infini" d'une classe $\mathfrak{S}$ de parties d'un ensemble X :

1.2. – Soient $\mathfrak{S}$ une partie de 2^X et Y une partie de X :

a) pour chaque $r \in \mathbf{N}$, on note $\Delta^{\mathfrak{S}}_{Y}(r)$ la quantité

$$\mathrm{Sup}\{|A \cap \mathfrak{S}| \mid A \subset Y, |A| = r\};$$

b) pour chaque $r \in \mathbf{N}$, on note $\Delta^{Y}_{\mathfrak{S}}(r)$ la quantité

$$\mathrm{Sup}\{|\mathrm{At}(\mathcal{A} \mid Y)| \mid \mathcal{A} \subset \mathfrak{S}, |\mathcal{A}| = r\}$$

(on écrira souvent $\Delta^{\mathfrak{S}}(r)$ pour $\Delta^{\mathfrak{S}}_{X}(r)$, et $\Delta_{\mathfrak{S}}(r)$ pour $\Delta^{X}_{\mathfrak{S}}(r)$) ;

c) on dit qu'une partie finie A de l'ensemble X est *pulvérisée par* $\mathfrak{S}$ si on a $|A \cap \mathfrak{S}| = 2^{|A|}$;

d) on dit qu'une partie finie $\mathcal{A}$ de 2^X est *indépendante dans* Y si on a $|\mathrm{At}(\mathcal{A} \mid Y)| = 2^{|\mathcal{A}|}$
(on peut noter que les classes indépendantes infinies jouent un rôle important dans l'étude des espaces de Banach qui ne contiennent pas ℓ^1 (voir Rosenthal [29]), une classe infinie de parties de X étant dite indépendante dans X si chacune de ses sous-classes finies est indépendante dans X).

Voici maintenant la *définition des indices de densité* d'une classe de parties d'un ensemble X :

1.3. – Soit $\mathfrak{S}$ une partie non vide de 2^X :

a) on appelle *densité entière de* $\mathcal{S}$ *dans* X la quantité

$$\mathrm{Dens}(\mathcal{S}) = \mathrm{Sup}\{r \in \mathbf{N} \mid \Delta_X^{\mathcal{S}}(r) = 2^r\},$$

autrement dit la borne supérieure des cardinaux des parties de X pulvérisées par $\mathcal{S}$;

b) on appelle *densité réelle de* $\mathcal{S}$ *dans* X la quantité

$$\mathrm{dens}(\mathcal{S}) = \mathrm{Inf}\{s \in \mathbf{R}^+ \mid \exists\, C \in]0, +\infty[,\ \forall\, r \in \mathbf{N},\ \Delta_X^{\mathcal{S}}(r) \leqslant C\, r^s\};$$

c) on appelle *densité duale entière de* $\mathcal{S}$ *dans* X la quantité

$$\mathrm{Dens}_X^*(\mathcal{S}) = \mathrm{Sup}\{r \in \mathbf{N} \mid \Delta_{\mathcal{S}}^X(r) = 2^r\},$$

autrement dit la borne supérieure des cardinaux des parties de $\mathcal{S}$ indépendantes dans X ;

d) on appelle *densité duale réelle de* $\mathcal{S}$ *dans* X la quantité

$$\mathrm{dens}^*(\mathcal{S}) = \mathrm{Inf}\{s \in \mathbf{R}^+ \mid \exists\, C \in]0, +\infty[,\ \forall\, r \in \mathbf{N},\ \Delta_{\mathcal{S}}^X(r) \leqslant C\, r^s\}$$

(j'ai remplacé ici les adjectifs "forte" et "faible" de ma note [6] respectivement par "entière" et "réelle") ;

e) on pose de plus par définition

$$\mathrm{Dens}(\phi) = \mathrm{dens}(\phi) = \mathrm{Dens}_X^*(\phi) = \mathrm{dens}_{\mathcal{S}}^*(\phi) = -1 .$$

1.4. – Soit $\mathcal{S}$ une partie de 2^X ; nous dirons (en suivant Dudley [14]) que $\mathcal{S}$ est une *classe de Vapnik-Cervonenkis* si sa densité entière $\mathrm{Dens}(\mathcal{S})$ est finie.

1.5. – Indiquons brièvement d'où proviennent ces notions :

a) la densité entière a été introduite par Sauer [30] et Vapnik, Cervonenkis [39], la terminologie "densité" que nous employons vient de [30], car Vapnik et Cervonenkis préfèrent employer l'indice $S(\mathcal{S}) = \mathrm{Inf}\{r \in \mathbf{N} \mid \Delta_X^{\mathcal{S}}(r) < 2^r\} = \mathrm{Dens}(\mathcal{S}) + 1$ (indice qui est appelé l'*indice de Vapnik* de $\mathcal{S}$ par Dudley [14]) ; signalons à ce propos que $\mathrm{Dens}(\mathcal{S}) - 1$ est la dimension, au sens de [35], du complexe simplicial formé par les parties de X pulvérisées par $\mathcal{S}$;

b) la densité réelle joue un rôle important dans [30], [39] et surtout dans [14] mais elle n'y est pas explicitement introduite, ni étudiée ;

c) les deux densités duales (entières et réelles) sont dues à l'auteur de ces lignes (voir [6]) ; ce sont des notions nouvelles, bien que les évaluations de nombre d'atomes interviennent beaucoup en théorie de l'approximation (Lorentz [25], Shapiro [32], Warren [41]) et que les classes indépendantes remontent à Fichtenholz, Kantorovitch [19] et Szpilrajn [37] ;

d) signalons à ce propos qu'une classe $\mathcal{A} = \{S_1, \ldots, S_r\}$ de parties de X est indépendante dans X si et seulement s'il existe une loi de probabilité sur la tribu engendrée par $\mathcal{A}$ telle que $S_1, \ldots, S_r$ soient indépendantes (au sens probabiliste) et de probabilités différentes de 0 et de 1.

Nous n'étudierons les densités duales que plus tard, et la fin de ce paragraphe est consacrée aux densités "directes", dont voici *quelques propriétés élémentaires* :

1.6. – a) $\mathrm{Dens}(\mathcal{S})$ est finie dès que $\mathrm{dens}(\mathcal{S})$ est finie (car pour tout $C \in]0, +\infty[$ et tout $s \in [0, +\infty[$, il existe $r \in \mathbf{N}$ avec $C r^s < 2^r$) ;

b) $\mathrm{dens}(\mathcal{S})$ s'annule dès que $|\mathcal{S}|$ est fini (car on peut alors prendre $s = 0$ et $C = |\mathcal{S}|$) ;

c) on a $\mathrm{Dens}\left(\bigcup_{k=0}^{n} \binom{X}{k}\right) = n$ dès que $0 \leqslant n \leqslant |X|$;

d) $\mathrm{Dens}(\mathcal{S})$ s'annule si et seulement si $|\mathcal{S}| = 1$.

Malgré 1.6a, on ne peut majorer $\mathrm{Dens}(\mathcal{S})$ à l'aide de $\mathrm{dens}(\mathcal{S})$; c'est ce que montre 1.6c (ou, si on préfère, 1.6b et 1.6d).

La comparaison en sens inverse, qui est beaucoup plus intéressante, est l'objet du *Lemme fondamental* de Sauer [30] et de Vapnik, Cervonenkis [39] (voir aussi Shelah [33], p. 254 et Eckhoff [18], p. 181) :

1.7. – Soit X un ensemble fini de cardinal r ; pour chaque entier $n \geqslant 0$, on note $\binom{r}{n}$ le cardinal de $\binom{X}{n}$ et on pose

$$\varphi(r, n) = \sum_{k=0}^{n} \binom{r}{n} ;$$

on pose de plus $\varphi(r, n) = 0$ pour chaque entier $n < 0$.

LEMME 1.8 (Lemme fondamental). — *Soit* $\mathfrak{S}$ *une classe de parties d'un ensemble* X *vérifiant* Dens$(\mathfrak{S}) = n$. *On a alors* $\Delta_X^{\mathfrak{S}}(r) \leqslant \varphi(r, n)$, *quel que soit* $r \in \mathbf{N}$.

Démonstration. — Montrons, par récurrence sur $r \in \mathbf{N}$, la propriété (P_r) suivante :

(P_r) $\Delta_X^{\mathfrak{S}}(r) \leqslant \varphi(r, n)$, quels que soient l'ensemble X et l'entier n et quelle que soit la partie $\mathfrak{S}$ de 2^X de densité entière $\leqslant n$.

Soient donc X un ensemble et $\mathfrak{S}$ une partie de 2^X avec Dens$(\mathfrak{S}) \leqslant n$.

a) On a $\Delta_X^{\mathfrak{S}}(0) = 1$ si $\mathfrak{S}$ n'est pas vide (c'est-à-dire si $n \geqslant 0$) et $\Delta_X^{\mathfrak{S}}(0) = 0$ sinon, ce qui établit la propriété (P_0).

b) Soit r un entier $\geqslant 1$ et supposons la propriété (P_k) établie pour chaque $k = 1, \ldots, r-1$. Soient $A \subset X$ avec $|A| = r$ et soit s un point de A ; on note $B = A \backslash \{s\}$ et on pose

$$\mathfrak{T} = \{T \in \mathfrak{S} \mid \exists S \in \mathfrak{S}, B \cap S = B \cap T \quad \text{et} \quad A \cap S \neq A \cap T\}.$$

Si C est une partie de B pulvérisée par $\mathfrak{T}$, alors $C \cup \{s\}$ est pulvérisée par $\mathfrak{S}$; on a donc Dens$(B \cap \mathfrak{T}) \leqslant n - 1$. Or on a

$$|A \cap \mathfrak{S}| = |B \cap \mathfrak{S}| + |B \cap \mathfrak{T}| ;$$

on a donc (en appliquant l'hypothèse de récurrence) :

$$|A \cap \mathfrak{S}| \leqslant \varphi(r-1, n) + \varphi(r-1, n-1) = \varphi(r, n),$$

ce qui établit la Proposition.

□

Voici quelques précisions sur le lemme ci-dessus :

1.9. — a) la majoration de $\Delta_X^{\mathfrak{S}}$ obtenue en 1.8 ne peut être améliorée : en effet soit X un ensemble infini et posons $\mathfrak{S} = \bigcup_{k=1}^{n} \binom{X}{k}$; on a alors Dens$(\mathfrak{S}) = n$ (voir 1.6c) et $\Delta_X^{\mathfrak{S}}(r) = \varphi(r, n)$ pour chaque entier $r \geqslant 0$;

b) on peut noter que (lorsque n est fixé) la fonction $r \longrightarrow \varphi(r, n)$ est un polynôme en r de degré n ; on a même $\varphi(r, n) \leqslant r^n$ pour tout $r \geqslant 2$ et tout $n \geqslant 1$ (on peut le montrer par récurrence sur $n \geqslant 1$: c'est clair pour $n = 1$, et la récurrence se poursuit en écrivant

$$\varphi(r\,,n) = \varphi(r\,,n-1) + \binom{r}{n} \leqslant r^{n-1} + \frac{r^n}{n!} \leqslant \frac{r^n}{2} + \frac{r^n}{2} = r^n\,).$$

On obtient donc le résultat suivant (qui justifie nos notations) :

COROLLAIRE 1.10. – *Soit* $\mathcal{S}$ *une classe de parties d'un ensemble* X. *On a alors* $\mathrm{dens}(\mathcal{S}) \leqslant \mathrm{Dens}(\mathcal{S})$. *En particulier la classe* $\mathcal{S}$ *est une classe de Vapnik-Cervonenkis si et seulement si sa densité réelle* $\mathrm{dens}(\mathcal{S})$ *est finie.*

Démonstration. – L'inégalité $\mathrm{dens}(\mathcal{S}) \leqslant \mathrm{Dens}(\mathcal{S})$ provient du Lemme 1.8 et de la remarque 1.9b ci-dessus. On en déduit (en tenant compte de 1.6a) que $\mathrm{dens}(\mathcal{S})$ est finie si et seulement si $\mathrm{Dens}(\mathcal{S})$ est finie.

□

Anticipant sur le paragraphe qui suit, mentionnons qu'on montrera aussi qu'une classe $\mathcal{S}$ de parties d'un ensemble X est une classe de Vapnik-Cervonenkis si et seulement si sa densité entière duale $\mathrm{Dens}^*_{\mathrm{X}}(\mathcal{S})$ (resp. sa densité réelle duale $\mathrm{dens}^*(\mathcal{S})$) est finie.

2. Propriétés des indices de densité.

Indiquons (sans démonstration) les propriétés les plus évidentes :

2.1. – a) soit Y une partie de X et soit $\mathcal{S}$ une partie de 2^{Y} ; alors $\mathrm{Dens}(\mathcal{S})$ et $\mathrm{dens}(\mathcal{S})$ ne changent pas selon qu'on considère la classe $\mathcal{S}$ comme une partie de 2^{Y} ou une partie de 2^{X} (et cela justifie bien nos notations) ;

b) soit Y une partie de X et soit $\mathcal{S}$ une partie de 2^{Y} ; alors $\mathrm{dens}^*(\mathcal{S})$ ne change pas selon qu'on considère $\mathcal{S}$ comme une partie de 2^{Y} ou une partie de 2^{X} (le nombre d'atomes dans Y ne diffère du nombre d'atomes dans X que d'un au plus) ; mais il n'en est pas de même de $\mathrm{Dens}^*_{\mathrm{X}}(\mathcal{S})$, ce qui justifie nos notations (car une famille $\mathrm{S}_1, \ldots, \mathrm{S}_r$ indépendante dans X cesse de l'être dans $\mathrm{Y} = \bigcup_{i=1}^{r} \mathrm{S}_i$) ;

c) soit $\mathcal{S}$ une partie de 2^{X} et soit $\mathcal{C}$ une partie de $\mathcal{S}$; chacun des quatre indices de densité de la classe $\mathcal{C}$ est alors inférieur ou égal à l'indice correspondant de la classe $\mathcal{S}$;

d) soit $\mathfrak{S}$ une partie de 2^X ; on note $X\backslash\mathfrak{S}$ la classe des parties de X de la forme $X\backslash S$ avec $S \in \mathfrak{S}$; chacun des quatre indices de densité de la classe $X\backslash\mathfrak{S}$ est alors égal à l'indice correspondant de la classe $\mathfrak{S}$.

En fait la principale propriété de stabilité des indices de densité est leur *stabilité par image réciproque* :

PROPOSITION 2.2. – *Soit $\mathfrak{S}$ une partie de 2^Y et soit f une application de X dans Y ; on note $f^{-1}(\mathfrak{S})$ la classe des parties de X de la forme $f^{-1}(S)$ avec $S \in \mathfrak{S}$; chacun des quatre indices de densité de la classe $f^{-1}(\mathfrak{S})$ est alors inférieur ou égal à l'indice correspondant de la classe $\mathfrak{S}$. Si de plus f est surjective, ces inégalités deviennent des égalités.*

Démonstration. – a) Soit A une partie finie de X avec $|A| = r$; posons $B = f(A)$ et notons g la restriction de f à A. On a alors pour chaque $S, S' \in \mathfrak{S}$

$$\begin{aligned}(A \cap f^{-1}(S))\, \Delta (A \cap f^{-1}(S')) &= A \cap f^{-1}(S \,\Delta\, S') \\ &= g^{-1}(B \cap (S \,\Delta\, S')) = g^{-1}((B \cap S)\, \Delta (B \cap S')).\end{aligned}$$

Comme g est surjective, $g^{-1}(V)$ est vide si et seulement si V est vide ; on a donc $|A \cap f^{-1}(\mathfrak{S})| = |B \cap \mathfrak{S}|$. Comme on a $|B| \leqslant r$, cela entraîne $\Delta_X^{f^{-1}(\mathfrak{S})}(r) \leqslant \Delta_X^{\mathfrak{S}}(r)$, avec égalité si f est surjective (car alors on peut choisir, pour chaque $B \subset Y$, une partie A de X avec $f(A) = B$ et $|A| = |B|$) ; cela donne le résultat annoncé pour les densités "directes".

b) Soit $\mathfrak{A}$ une partie finie de $f^{-1}(\mathfrak{S})$ avec $|\mathfrak{A}| = r$; soit $\mathfrak{B}$ une partie de $\mathfrak{S}$ avec $\mathfrak{A} = f^{-1}(\mathfrak{B})$ et $|\mathfrak{B}| = r$. L'application $V \longrightarrow f^{-1}(V)$, qui est un morphisme d'algèbres, transforme chaque atome de $\mathfrak{B}$ dans Y, ou bien en un atome de $f^{-1}(\mathfrak{B})$ dans X, ou bien dans l'ensemble vide (cette seconde éventualité ne se produisant pas si f est surjective). On a donc

$$\mathrm{At}(\mathfrak{A}|X) \subset f^{-1}(\mathrm{At}(\mathfrak{B}|Y)$$

et donc $|\mathrm{At}(\mathfrak{A}|X)| \leqslant |\mathrm{At}(\mathfrak{B}|Y)|$ (car $V \longrightarrow f^{-1}(V)$ envoie parties disjointes sur parties disjointes), avec égalité si f est surjective. Cela entraîne $\Delta^X_{f^{-1}(\mathfrak{S})}(r) \leqslant \Delta^X_{\mathfrak{S}}(r)$ (avec égalité si f est surjective) et cela donne le résultat annoncé pour les densités duales.

□

On peut noter que f est une application simpliciale (au sens de [35], p. 109) du complexe des parties de X pulvérisées par $f^{-1}(\mathcal{S})$ dans le complexe des parties de Y pulvérisées par $\mathcal{S}$.

Dans le cas où f est l'injection naturelle dans X d'une partie Y de X, la Proposition précédente donne :

2.3. – Soient $\mathcal{S}$ une partie de 2^X et Y une partie de X ; chacun des quatre indices de densité de $Y \cap \mathcal{S}$ est alors inférieur ou égal à l'indice correspondant de $\mathcal{S}$.

La densité réelle se comporte mieux que la densité entière vis-à-vis de beaucoup d'opérations (voir 2.5 et 2.6) ; c'est le cas déjà pour la *réunion de deux classes* :

PROPOSITION 2.4. – *Soient* $\mathcal{S}_1$ *et* $\mathcal{S}_2$ *deux parties de* 2^X. *On pose* $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ *(il s'agit de la partie* $\mathcal{S}$ *de* 2^X *dont les éléments appartiennent, ou à* $\mathcal{S}_1$, *ou à* $\mathcal{S}_2$).

On a alors $\operatorname{dens}(\mathcal{S}) = \operatorname{Sup}(\operatorname{dens}(\mathcal{S}_1), \operatorname{dens}(\mathcal{S}_2))$

et $\operatorname{Dens}(\mathcal{S}) \leqslant \operatorname{Dens}(\mathcal{S}_1) + \operatorname{Dens}(\mathcal{S}_2) + 1$.

Démonstration. – a) Il est clair qu'on a $\operatorname{dens}(\mathcal{S}) \geqslant \operatorname{Sup}(\operatorname{dens}(\mathcal{S}_1), \operatorname{dens}(\mathcal{S}_2))$ (à cause de 2.1c). Pour montrer l'inégalité inverse, on note qu'on a, pour chaque $A \subset X$ (avec $|A| = r$) :

$$|A \cap \mathcal{S}| \leqslant |A \cap \mathcal{S}_1| + |A \cap \mathcal{S}_2| \leqslant C_1 r^s + C_2 r^t \leqslant (C_1 + C_2) r^{\operatorname{Sup}(s,t)}$$

(où s et t majorent respectivement les densités réelles de $\mathcal{S}_1$ et $\mathcal{S}_2$).

b) Notons s et t les densités entières de $\mathcal{S}_1$ et de $\mathcal{S}_2$. On a alors, pour chaque partie A de X de cardinal $r \geqslant s + t + 2$,

$$|A \cap \mathcal{S}| \leqslant |A \cap \mathcal{S}_1| + |A \cap \mathcal{S}_2| \leqslant \varphi(r, s) + \varphi(r, t) = \sum_{k=0}^{s} \binom{r}{k} + \sum_{k=r-t}^{r} \binom{r}{k} < 2^r ;$$

une telle partie A ne peut être pulvérisée par $\mathcal{S}$; donc

$$\operatorname{Dens}(\mathcal{S}) \leqslant s + t + 1 .$$

□

Un autre exemple de calcul de densités concerne les *classes de produits* :

PROPOSITION 2.5. — a) *Soient* $\mathcal{S}_1$ *une partie de* 2^X *et* $\mathcal{S}_2$ *une partie de* 2^Y ; *on note* $\mathcal{S}_1 \underline{\times} \mathcal{S}_2$ *la classe des parties de* $X \times Y$ *de la forme* $S_1 \times S_2$ *avec* $S_1 \in \mathcal{S}_1$ *et* $S_2 \in \mathcal{S}_2$; *on a alors*

$$\operatorname{dens}(\mathcal{S}_1 \underline{\times} \mathcal{S}_2) \leqslant \operatorname{dens}(\mathcal{S}_1) + \operatorname{dens}(\mathcal{S}_2).$$

b) *Posons de plus* $F(k) = \operatorname{Sup}\{r \in \mathbf{N} \mid \varphi(2r, k) \geqslant 2^r\}$ *pour chaque* $k \in \mathbf{N}$; *on a alors* $\operatorname{Dens}(\mathcal{S}_1 \underline{\times} \mathcal{S}_2) \leqslant F(\operatorname{Dens}(\mathcal{S}_1) + \operatorname{Dens}(\mathcal{S}_2))$.

Démonstration. — a) Soit A une partie de $X \times Y$. On note A' le produit des projections A_1 et A_2 de A sur X et sur Y. On a alors $|A| \geqslant \operatorname{Sup}(|A_1|, |A_2|)$ et

$$|A \cap (\mathcal{S}_1 \underline{\times} \mathcal{S}_2)| \leqslant |A' \cap (\mathcal{S}_1 \underline{\times} \mathcal{S}_2)| = |A_1 \cap \mathcal{S}_1| \times |A_2 \cap \mathcal{S}_2|.$$

Cela donne la majoration cherchée de $\operatorname{dens}(\mathcal{S}_1 \underline{\times} \mathcal{S}_2)$.

b) Pour majorer $\operatorname{Dens}(\mathcal{S}_1 \underline{\times} \mathcal{S}_2)$, on pose $s = \operatorname{Dens}(\mathcal{S}_1)$ et $t = \operatorname{Dens}(\mathcal{S}_2)$; si A est une partie de $X \times Y$ de cardinal $r > F(s+t)$, on a donc

$$|A \cap (\mathcal{S}_1 \underline{\times} \mathcal{S}_2)| \leqslant \varphi(r, s)\, \varphi(r, t) \leqslant \varphi(2r, s+t) \leqslant 2^r ;$$

une telle partie A ne peut donc pas être pulvérisée par $\mathcal{S}_1 \underline{\times} \mathcal{S}_2$. □

(La forme initiale de 2.5b était inexacte ; l'erreur m'a été signalée par R.M. Dudley, et subsiste malheureusement dans ma note [6] ; la majoration de $\operatorname{Dens}(\mathcal{S}_1 \underline{\times} \mathcal{S}_2)$ que je donne ici est maintenant exacte, bien que non optimale).

Si $\mathcal{S}$ est une classe de Vapnik-Cervonenkis, il en est de même de la classe $\mathcal{S}^{(k)}$ obtenue à partir de $\mathcal{S}$ par un *calcul à* k *arguments* (l'idée en revient à Dudley [14], p. 922, mais nous allons de plus évaluer la densité réelle de $\mathcal{S}^{(k)}$) :

PROPOSITION 2.6. — *Soit* $\mathcal{S}$ *une partie de* 2^X *et soit* k *un entier* $\geqslant 1$. *On note* $\mathcal{S}^{(k)}$ *la réunion de toutes les algèbres (de parties de* X*) engendrées par* k *éléments de* $\mathcal{S}$. *On a alors*

$$\operatorname{déns}(\mathcal{S}^{(k)}) \leqslant k \operatorname{dens}(\mathcal{S}).$$

Démonstration. – Soit A une partie de X et notons $(A \cap \mathfrak{S})^{(k)}$ la réunion de toutes les algèbres (de parties de A) engendrées par k éléments de $A \cap \mathfrak{S}$. On a alors $A \cap \mathfrak{S}^{(k)} = (A \cap \mathfrak{S})^{(k)}$. Or une algèbre à k générateurs a au plus 2^{2^k} éléments. On a donc

$$|A \cap \mathfrak{S}^{(k)}| \leqslant 2^{2^k} |A \cap \mathfrak{S}|^k ,$$

ce qui donne le résultat.

□

Structure d'incidence et dualité.

2.7. – Soient X et Y deux ensembles et soit T une partie de $X \times Y$; considérée comme une *structure d'incidence sur* (X, Y) (au sens de Dembowski [13]), cette partie T permet de définir une classe $T(X)$ de parties de Y et une classe $T^{-1}(Y)$ de parties de X, de la façon suivante :

– pour chaque $x \in X$ et chaque $y \in Y$, on pose

$$T_x = \{z \in Y \mid (x, z) \in T\} \quad \text{et} \quad T^y = \{z \in X \mid (z, y) \in T\}$$

– on note alors $T(X)$ la classe $\{T_x \mid x \in X\}$ et $T^{-1}(Y)$ la classe $\{T^y \mid y \in Y\}$.

2.8. – Il va de soi que toute partie $\mathfrak{S}$ de 2^X peut être obtenue de cette manière. En effet posons $[X, \mathfrak{S}] = \{(x, S) \mid x \in S \in \mathfrak{S}\}$; c'est une structure d'incidence sur $(X, \mathfrak{S})$ et on a $[X, \mathfrak{S}]^{-1}(\mathfrak{S}) = \mathfrak{S}$; la classe $[X, \mathfrak{S}](X)$ est notée $\mathfrak{S}_X^*$ et appelée la *classe duale de* $\mathfrak{S}$ *sur* X. En général si $T(X)$ et $T^{-1}(Y)$ sont les deux classes associées à une structure d'incidence T sur (X, Y). on dira que $T(X)$ et $T^{-1}(Y)$ sont *deux classes duales* (relativement à T).

2.9. – a) Soit $\mathfrak{S}$ une partie de 2^X ; la classe duale de $\mathfrak{S}_X^*$ sur $\mathfrak{S}$ est notée $\mathfrak{S}_X^{**}$ (c'est une classe de parties de $\mathfrak{S}_X^*$) et appelée *classe biduale de* $\mathfrak{S}$ *sur* X.

b) On dit que deux points x et x' de X sont *séparés par* $\mathfrak{S}$ s'il existe $S \in \mathfrak{S}$ avec $(x \in S, x' \notin S)$ ou avec $(x \notin S, x' \in S)$; x et x' sont donc séparés par $\mathfrak{S}$ si et seulement si $[X, \mathfrak{S}]_x$ est distinct de $[X, \mathfrak{S}]_{x'}$.

c) Notons $p_{\mathfrak{S}}$ la surjection $x \longrightarrow [X, \mathfrak{S}]_x$ de X sur $\mathfrak{S}_X^{*'}$; on vérifie immédiatement l'égalité $\mathfrak{S} = p_{\mathfrak{S}}^{-1}(\mathfrak{S}_X^{**})$. La proposition 2.2 implique donc que chacun des quatre indices de densité de la classe $\mathfrak{S}_X^{**}$ est égal à l'indice correspondant de la classe $\mathfrak{S}$.

LEMME 2.10. – *Soient* T *une partie de* $X \times Y$ *et* A *une partie finie de* X. *On pose* $T(A) = \{T_x \mid x \in A\}$. *On a alors*

$$|\mathrm{At}(T(A) \mid Y)| = |A \cap T^{-1}(Y)|.$$

Démonstration. – Les atomes de $T(A)$ dans Y sont exactement les ensembles non vides de la forme

$$\alpha(B) = \Big(\bigcap_{x \in B} T_x\Big) \cap \Big(\bigcap_{x \in A \setminus B} Y \setminus T_x\Big).$$

Soient y un point de Y et B une partie de A ; l'ensemble $\alpha(B)$ contient y si et seulement si $A \cap T^y = B$; cela démontre le Lemme.

□

PROPOSITION 2.11. – *Soit* T *une partie de* $X \times Y$; *posons* $\mathfrak{S} = T^{-1}(Y)$ *et* $\mathfrak{C} = T(X)$. *On a alors* $\Delta^Y = \Delta^{\mathfrak{S}}_X$ *et, en particulier,* $\mathrm{dens}^*(\mathfrak{C}) = \mathrm{dens}(\mathfrak{S})$, $\mathrm{Dens}^*_Y(\mathfrak{C}) = \mathrm{Dens}(\mathfrak{S})$.

Démonstration. – a) Soit r un entier $\geqslant 0$.

– Si A est une partie de X avec $|A| = r$, on a

$$|\mathrm{At}(T(A) \mid Y)| = |A \cap \mathfrak{S}|$$

(grâce à 2.10) ; or on a $|T(A)| \leqslant r$. On a donc $\Delta^Y_{\mathfrak{C}}(r) \geqslant \Delta^{\mathfrak{S}}_X(r)$.

– Si $\mathfrak{A}$ est une partie de $\mathfrak{C}$ avec $|\mathfrak{A}| = r$, il existe $A \subset X$ avec $\mathfrak{A} = T(A)$ et $|A| = r$; or on a $|\mathrm{At}(\mathfrak{A} \mid Y)| = |A \cap \mathfrak{S}|$ (grâce à 2.10). On a donc $\Delta^Y_{\mathfrak{C}}(r) \leqslant \Delta^{\mathfrak{S}}_X(r)$.

b) On a donc $\Delta^Y_{\mathfrak{C}}(r) = \Delta^{\mathfrak{S}}_X(r)$, quel que soit $r \in \mathbf{N}$, et cela entraîne les deux autres conclusions.

□

On peut même comparer les densités entières de deux classes duales $T(X)$ et $T^{-1}(Y)$:

PROPOSITION 2.12. – *Soit* T *une partie de* $X \times Y$; *posons* $\mathfrak{S} = T^{-1}(Y)$ *et* $\mathfrak{C} = T(X)$. *On a alors*

$$\mathrm{Dens}(\mathfrak{C}) < 2^{1+\mathrm{Dens}(\mathfrak{S})} \quad et \quad \mathrm{Dens}(\mathfrak{S}) < 2^{1+\mathrm{Dens}(\mathfrak{C})}.$$

Démonstration. – a) Si la densité entière de $\mathfrak{C}$ est 0 ou -1, on a clairement $\mathrm{Dens}(\mathfrak{C}) < 2^{1+\mathrm{Dens}(\mathfrak{S})}$. Sinon soit r un entier $\geqslant 0$ avec $2^{r+1} > \mathrm{Dens}(\mathfrak{C}) \geqslant 2^r$; il existe donc $A \subset Y$ avec

$|A| = 2^r$ et $A \cap \mathcal{C} = 2^A$. Or, parmi toutes les parties d'un ensemble A de cardinal 2^r, il est clair qu'on peut trouver une famille indépendante dans A formée de r parties (utiliser la remarque 1.5d) ; on peut donc trouver $B \subset X$ avec

$$|A \cap T(B)| = r \text{ et } |At(T(B)|A)| = 2^r$$

($A \cap T(B)$ est la famille indépendante mentionnée ci dessus) ; sans changer $T(B)$, on peut même supposer de plus $|B| = r$. On a donc $|At(T(B)|Y)| = 2^r$ et donc (grâce à 2.10) $|B \cap \mathcal{S}| = 2^r$. B est donc pulvérisée par $\mathcal{S}$. On a donc $\mathrm{Dens}(\mathcal{S}) \geqslant r$ et cela entraîne

$$\mathrm{Dens}(\mathcal{S}) < 2^{r+1} \leqslant 2^{1+\mathrm{Dens}(\mathcal{S})}.$$

b) La seconde inégalité se déduit de la première en remplaçant T par la structure d'incidence inverse

$$T^{-1} = \{(y,x) \in Y \times X \mid (x,y) \in T\}.$$

□

Voici *quelques conséquences immédiates des résultats précédents* :

2.13. – Soit $\mathcal{S}$ une partie de 2^X :

a) on a

$$\mathrm{Dens}_X^*(\mathcal{S}) = \mathrm{Dens}(\mathcal{S}_X^*), \; \mathrm{dens}^*(\mathcal{S}) = \mathrm{dens}(\mathcal{S}_X^*),$$

$$\mathrm{Dens}^*(\mathcal{S}_X^*) = \mathrm{Dens}(\mathcal{S}) \text{ et } \mathrm{dens}^*(\mathcal{S}_X^*) = \mathrm{dens}(\mathcal{S})$$

(appliquer 2.11 à la structure d'incidence $[X, \mathcal{S}]$) ;

b) on a $\mathrm{Dens}_X^*(\mathcal{S}) < 2^{1+\mathrm{Dens}(\mathcal{S})}$ et $\mathrm{Dens}(\mathcal{S}) < 2^{1+\mathrm{Dens}_X^*(\mathcal{S})}$ (c'est une conséquence de a) et 2.12) ;

c) supposons $\mathrm{Dens}_X^*(\mathcal{S}) = n$; on a alors $\Delta_{\mathcal{S}}^X(r) \leqslant \varphi(r,n)$, quel que soit $r \in \mathbf{N}$ (appliquer a), 1.8 et 2.11) ; on a donc $\mathrm{dens}^*(\mathcal{S}) \leqslant \mathrm{Dens}_X^*(\mathcal{S})$;

d) $\mathrm{Dens}_X^*(\mathcal{S})$ est finie dès que $\mathrm{dens}^*(\mathcal{S})$ est finie (c'est une conséquence de a) et 1.6a) ;

e) la classe $\mathcal{S}$ est une classe de Vapnik-Cervonenkis si et seulement si $\mathrm{Dens}_X^*(\mathcal{S})$ (resp. $\mathrm{dens}^*(\mathcal{S})$) est finie (c'est une conséquence de b), c) et d)) ;

f) soit Y une partie de X et supposons que $\mathcal{S}$ soit une partie de 2^Y ; on a alors $\mathrm{Dens}_Y^*(\mathcal{S}) \leqslant \mathrm{Dens}_X^*(\mathcal{S}) \leqslant \mathrm{Dens}_Y^*(\mathcal{S}) + 1$ (observer qu'on a $\mathcal{S}_Y^* \subset \mathcal{S}_X^* \subset \mathcal{S}_Y^* \cup \{\phi\}$ et appliquer a), 2.1c, 1.6d et 2.4).

Classe complétée pour la convergence simple.

2.14. – Soit $\mathcal{S}$ une partie de 2^X. On appelle *classe complétée de* $\mathcal{S}$ (et on note $\overline{\mathcal{S}}$) la classe de toutes les parties Z de X qui sont limites simples d'une famille filtrante $(S_i)_{i\in I}$ d'éléments de $\mathcal{S}$ (autrement dit dont l'indicatrice 1_Z est limite simple des indicatrices 1_{S_i}).

PROPOSITION 2.15. – *Soit* $\mathcal{S}$ *une partie de* 2^X *et soit* $\overline{\mathcal{S}}$ *la classe complétée de* $\mathcal{S}$. *Chacun des quatre indices de densité de la classe* $\overline{\mathcal{S}}$ *est alors égal à l'indice correspondant de la classe* $\mathcal{S}$.

Démonstration. – a) En effet, si A est une partie finie de X, la classe $A\cap\overline{\mathcal{S}}$ est la classe complétée de $A\cap\mathcal{S}$; or $A\cap\mathcal{S}$ est fermée pour la convergence simple (car A est fini) ; donc $A\cap\overline{\mathcal{S}} = A\cap\mathcal{S}$. Cela démontre la Proposition pour les indices de densité "directs".

b) Soit $\mathfrak{A} = \{\overline{S}_1, \ldots, \overline{S}_r\}$ une partie de $\overline{\mathcal{S}}$; on peut donc trouver, pour chaque $k = 1, \ldots, r$, une famille filtrante $(S_{ki})_{i\in I}$ dont la limite est $\overline{S}_k$ (au sens de la convergence simple). Pour chaque $i\in I$, on pose $\mathfrak{A}_i = \{S_{1i}, \ldots, S_{ri}\}$. Or, lorsque r est un entier fixé, le cardinal de l'ensemble $\mathrm{At}(\{T_1, \ldots, T_r\}\,|\,X)$ est une fonction semi-continue inférieurement de $(T_1, \ldots, T_r)$ pour la convergence simple (car il est égal à $\operatorname{Sup}\{|\mathrm{At}(\{T_1, \ldots, T_r\}\,|\,A)|\;|\;|A| = 2^r\}$). On a donc $|\mathrm{At}(\mathfrak{A}|X)| \leqslant \liminf_{i\in I} |\mathrm{At}(\mathfrak{A}_i|X)|$.

La fonction $\Delta^X_{\overline{\mathcal{S}}}$ est donc inférieure ou égale à $\Delta^X_{\mathcal{S}}$; elle est même égale (car $\mathcal{S}$ est incluse dans $\overline{\mathcal{S}}$), ce qui démontre la Proposition.

□

Séparation d'une classe par une autre.

2.16. – Soient $\mathcal{S}$ et $\mathcal{T}$ deux parties de 2^X ;

a) on dit que $\mathcal{T}$ *sépare* $\mathcal{S}$ si, pour chaque $S, S'\in\mathcal{S}$, il existe $T, T'\in\mathcal{T}$ avec $S\backslash S'\subset T\backslash T'$ et $S'\backslash S\subset T'\backslash T$;

b) la séparation est un préordre ; de plus $\mathcal{S}$ sépare toute sous-classe de $\mathcal{S}$;

c) soit A une partie de X ; si $\mathcal{T}$ sépare $\mathcal{S}$, alors $A\cap\mathcal{T}$ sépare $A\cap\mathcal{S}$;

d) si $\mathscr{T}$ sépare 2^X, alors $\mathscr{T}$ est égale à 2^X (en effet, pour chaque partition (S, S') de X, il existe dont $T, T' \in \mathscr{T}$ avec $S \subset T \backslash T'$ et $S' \subset T' \backslash T$; on a donc $S = T$ et $S' = T'$; donc S et S' appartiennent à $\mathscr{T}$).

Lemme 2.17. – *Soient $\mathscr{S}$ et $\mathscr{T}$ deux parties de 2^X. On suppose que $\mathscr{T}$ sépare $\mathscr{S}$. On a alors* $\mathrm{Dens}(\mathscr{S}) \leqslant \mathrm{Dens}(\mathscr{T})$.

Démonstration. – Soit A une partie de X pulvérisée par $\mathscr{S}$; on a donc $A \cap \mathscr{S} = 2^A$; or $A \cap \mathscr{T}$ sépare $A \cap \mathscr{S}$ (par 2.16c) ; donc $A \cap \mathscr{T}$ est égal à 2^A (par 2.16d) ; donc A est pulvérisée par $\mathscr{T}$.

□

Les classes totalement ordonnées et les classes disjointes sont des classes de densité entière $\leqslant 1$ (voir [39] et [43]) ; on peut le démontrer d'une façon très semblable à 2.17 (et c'est pour cela que nous le démontrons maintenant) :

2.18. – Soit $\mathscr{S}$ une partie de 2^X ;

a) si $\mathscr{S}$ est une *classe totalement ordonnée* (c'est-à-dire si, pour chaque $S, S' \in \mathscr{S}$ on a $S \subset S'$ ou $S' \subset S$), alors $\mathrm{Dens}(\mathscr{S}) \leqslant 1$;

b) si $\mathscr{S}$ est une *classe disjointe* (c'est-à-dire si, pour chaque $S, S' \in \mathscr{S}$, on a $S \cap S' = \phi$ ou $S = S'$), alors $\mathrm{Dens}(\mathscr{S}) \leqslant 1$ (en effet, soit A une partie de X pulvérisée par une classe $\mathscr{S}$ totalement ordonnée ou disjointe ; on a donc $A \cap \mathscr{S} = 2^A$; or $A \cap \mathscr{S}$ est, elle aussi, totalement ordonnée ou disjointe ; on a donc $|A| \leqslant 1$).

Les valeurs de la densité réelle.

On a vu (en 1.6b) que la densité réelle $\mathrm{dens}(\mathscr{S})$ s'annule dès que la classe $\mathscr{S}$ est finie ; voici la réciproque de ce résultat :

Proposition 2.19. – *Soit $\mathscr{S}$ une partie de 2^X avec* $\mathrm{dens}(\mathscr{S}) < 1$; *alors la classe $\mathscr{S}$ est finie.*

Démonstration. – a) Soit $\mathscr{S}$ une partie de 2^X avec $\mathrm{dens}(\mathscr{S}) < 1$. On pose $\mathscr{S}_1 = \mathscr{S} \cup (X \backslash \mathscr{S})$; on a donc encore $\mathrm{dens}(\mathscr{S}_1) < 1$ (voir 2.4). Il existe donc $C \in]0, +\infty[$ et $s \in]0,1[$ tels que

$$|A \cap \mathscr{S}_1| \leqslant C|A|^s,$$

quel que soit A inclus dans X. En particulier, il existe un entier $m \geqslant 0$ tel que $|A \cap \mathcal{S}_1| < m$ dès que $|A| = m$; nous allons montrer que cette propriété implique que $\mathcal{S}_1$ (et donc $\mathcal{S}$) est finie.

b) Nous établissons d'abord le résultat auxiliaire suivant : soient k un entier $\geqslant 1$, $\mathcal{S}_k$ la classe des réunions de k éléments de $\mathcal{S}_1$ et $\mathcal{A}$ une partie de $\mathcal{S}_k$. Il existe alors une partie $\mathcal{F}$ de $\mathcal{A}$ telle que $|\mathcal{F}| < m^k$ et que tout élément de $\mathcal{A}$ contienne l'intersection des éléments de $\mathcal{F}$.

En effet, sinon, $\mathcal{A}$ contiendrait une famille $(T_i)_{i=1,\dots,m^k}$ telle que $\left(\bigcap_{j<i} T_j\right) \backslash T_i$ soit non vide quel que soit $i = 1, \dots, m^k$ (en particulier tel que $X \backslash T_1$ soit non vide) ; on choisirait alors $A = \{x_1, \dots, x_{m^k}\}$ avec $x_i \in \left(\bigcap_{j<i} T_j\right) \backslash T_i$ quel que soit $i = 1, \dots, m^k$; on aurait donc $|A \cap \mathcal{S}_k| \geqslant m^k > |A \cap \mathcal{S}_1|^k$, ce qui est impossible.

c) Soit $\mathcal{B}$ une partie finie de $\mathcal{S}_1$. Soit α un atome de $\mathcal{B}$ dans X ; notons $\mathcal{A}_\alpha$ la classe des éléments de $\mathcal{B} \cup (X \backslash \mathcal{B})$ qui contiennent α ; on applique b) (avec $k = 1$) à la classe $\mathcal{A}_\alpha$. et on trouve ainsi que l'atome α est intersection de $(m - 1)$ éléments de $\mathcal{S}_1$. Notons maintenant $\mathcal{A}$ l'ensemble des complémentaires de tous les atomes de $\mathcal{B}$ dans X ($\mathcal{A}$ est donc inclus dans $\mathcal{S}_{m-1}$). On applique b) (avec $k = m - 1$) à la classe $\mathcal{A}$, et on voit ainsi que $\mathcal{A}$ contient une classe de cardinal $\leqslant N = m^{m-1} - 1$ et d'intersection vide. La classe $\mathcal{B}$ a donc au plus N atomes. Ceci valant pour chaque partie finie $\mathcal{B}$ de $\mathcal{S}_1$, on a donc $|\mathcal{S}_1| \leqslant 2^N$.

□

Ce résultat (dont la démonstration est peut-être un peu longue) montre que la densité réelle ne prend jamais de valeur dans $]0,1[$; indiquons maintenant *un exemple de classe de densité réelle* $\frac{3}{2}$ (l'observation qui suit m'a été suggérée par P. Frankl) :

2.20. – a) Soient (P, L) une partition d'un ensemble X et E une partie de $P \times L$; on dit alors que le couple $G = (X, E)$ est un *graphe biparti porté par* X ; les éléments de la classe

$$\mathcal{S}(G) = \{\{p, \ell\} \mid (p, \ell) \in E\}$$

sont appelés les *arêtes de* G ;

b) si de plus E ne contient aucune partie de la forme

$$\{p, p'\} \times \{\ell, \ell'\}$$

(avec $p \neq p'$ et $\ell \neq \ell'$), on dit alors que le graphe biparti $G = (X, E)$ est *sans circuit de longueur 4* ;

c) or il existe (ainsi que l'ont montré Kövari, Sos, Turan [24] un nombre $C \in]0, +\infty[$ tel que chaque graphe biparti $G = (X, E)$ sans circuit de longueur 4 possède au plus $C|X|^{3/2}$ arêtes ;

d) soit $\mathbf{F}_q$ le corps fini à q éléments (q est donc une puissance d'un nombre premier) et soit L_q l'ensemble des droites affines de l'espace vectoriel $P_q = (\mathbf{F}_q)^2$; on note X_q la réunion disjointe de P_q et de L_q ; enfin on pose $E_q = \{(p, \ell) \mid p \in \ell \in L_q\}$; alors le graphe biparti $G_q = (X_q, E_q)$ est sans circuit de longueur 4 (par deux points il ne passe qu'une droite) ; de plus on a $|P_q| = q^2$, $|L_q| = q^2 + q$ et $|E_q| = q|L_q|$, donc $|E_q| \geqslant q^3$ et $|X_q| \leqslant 3q^2$;

e) soit $G = (X, E)$ le graphe biparti qui est la réunion disjointe des graphes bipartis $G_q = (X_q, E_q)$ (pour tous les entiers q puissances d'un nombre premier) ; la classe $\mathscr{S}(G)$ de toutes les arêtes de G a une densité réelle égale à $\frac{3}{2}$: en effet, soient A une partie de X et $G|_A$ le graphe biparti porté par A dont les arêtes sont les arêtes de G incluses dans A ; le graphe biparti $G|_A$ est sans circuit de longueur 4 et possède donc au plus $C|A|^{3/2}$ arêtes ; on a donc

$$|A \cap \mathscr{S}(G)| \leqslant 1 + |A| + C|A|^{3/2} \leqslant (C+2)\,|A|^{3/2} ;$$

mais on a aussi $|X_q \cap \mathscr{S}(G)| = |E_q| > \frac{1}{8}|X_q|^{3/2}$, quel que soit l'entier q puissance d'un nombre premier on a donc $\operatorname{dens}(\mathscr{S}(G)) = \frac{3}{2}$.

3. Exemples de classes de Vapnik-Cervonenkis.

Nous avons déjà donné, dans les paragraphes précédents, quelques exemples de classes de Vapnik-Cervonenkis (voir 1.6, 2.18 et 2.20) et de nombreux moyens d'en engendrer à partir de ces exemples.

Indiquons d'abord *quelques exemples de classes de densité infinie* (il est commode d'appeler classe de densité infinie toute classe qui n'est pas une classe de Vapnik-Cervonenkis) :

3.1. a) si X est un ensemble infini, alors 2^X est une classe de densité infinie ;

b) soit $\mathcal{S}$ la classe des progressions arithmétiques de la forme $p\mathbf{Z}$ avec p premier ; la classe $\mathcal{S}$ est alors une classe de densité infinie (en effet, si $p_1, \ldots, p_r$ sont des nombres premiers distincts, la classe $\mathcal{C} = \{p_1\mathbf{Z}, \ldots, p_r\mathbf{Z}\}$ a 2^r atomes, par le théorème des restes chinois [36] p. 72) ;

c) soit n un entier $\geqslant 2$, alors la classe Γ_n de tous les sous-ensembles convexes de $\mathbf{R}^n$ est une classe de densité infinie (en effet l'ensemble des sommets d'un polygone convexe est pulvérisé par Γ_n) ; par contre la classe Γ_1 de tous les intervalles de $\mathbf{R}$ est une classe de Vapnik-Cervonenkis, comme cela résulte de 2.18a et 2.6.

Nous nous servirons plus tard de l'exemple suivant :

LEMME 3.2. – *Soit* Σ *la classe de tous les polygones convexes de* $\mathbf{R}^2$ *qui sont symétriques (autour de* 0) *et à sommets rationnels ; alors* Σ *est une classe de densité infinie.*

Démonstration. – Soit A l'ensemble des sommets d'un polygone convexe symétrique à $2r$ sommets ; alors $A \cap \Sigma$ contient tous les sous-ensembles symétriques de A ; on a donc

$$|A \cap \Sigma| \geqslant 2^r = (2^{|A|})^{1/2} ;$$

la fonction $r \longrightarrow \Delta^\Sigma(2r)$ ne peut donc être majorée par un polynôme.

□

Demi-espaces affines et demi-espaces vectoriels dans $\mathbf{R}^n$.

3.3. – Soit n un entier $\geqslant 0$; on appelle *demi-espace affine* fermé (resp. ouvert) dans $\mathbf{R}^n$ l'ensemble des points de $\mathbf{R}^n$ où une fonction affine non constante est $\geqslant 0$ (resp. est > 0) ; si cette fonction affine est une forme linéaire, on parle alors de *demi-espace vectoriel* fermé (resp. ouvert) ; on notera :

$\mathcal{A}_n$ l'ensemble de tous les demi-espaces affines fermés de $\mathbf{R}^n$,

$\mathscr{A}_n^0$ l'ensemble des éléments de $\mathscr{A}_n$ qui ne contiennent pas 0, et $\mathscr{A}_n^1$ la trace de la classe $\mathscr{A}_n^0$ sur $\mathbf{R}^n \setminus \{0\}$.

On va voir que les demi-espaces affines fermés de $\mathbf{R}^n$ forment une classe de Vapnik-Cervonenkis (c'est ce résultat qui est la motivation initiale de [39]) ; il s'agit d'ailleurs d'un résultat très ancien, au moins sous sa forme "majoration du nombre d'atomes de la tribu engendrée par r demi-espaces de $\mathbf{R}^n$".

Cela va résulter du *Théorème de Radon,* que nous allons interpréter comme un calcul de densité entière (ce théorème est susceptible d'autres interprétations : voir par exemple [12], [18], [7] et cidessous §5) :

Proposition 3.4 (Théorème de Radon). – *Soit* A *une partie de* $\mathbf{R}^n$ *avec* $|A| = n + 2$. *Alors* A *n'est pas pulvérisée par la classe* $\mathscr{A}_n$; *plus précisément il existe* $B \subset A$ (*avec* $B \neq A, B \neq \phi$) *telle que ni* B *ni* $A \setminus B$ *n'appartienne à* $A \cap \mathscr{A}_n$.

Démonstration (nous suivons [12]. – Soient $a_1, \ldots, a_{n+2}$ les éléments de A. Considérons l'opérateur

$$u : x \longrightarrow \left(\sum_{i=1}^{n+2} x_i a_i, \sum_{i=1}^{n+2} x_i\right) \quad \text{de } \mathbf{R}^{n+2} \text{ dans } \mathbf{R}^n \times \mathbf{R} ;$$

comme u ne peut être injectif, il existe un élément x de $\mathbf{R}^{n+2}$ non nul et tel que $\sum_{i=1}^{n+2} x_i a_i = 0$ et $\sum_{i=1}^{n+2} x_i = 0$. Soient alors

$B = \{a_i \mid x_i > 0\}$ et $c = \sum_{i=1}^{n+2} x_i^+$; le point $\frac{1}{c} \sum_{i=1}^{n+2} x_i^+ a_i$ appartient donc à l'enveloppe convexe de B et à celle de $A \setminus B$. Donc ni B ni $A \setminus B$ n'appartient à la classe $A \cap \mathscr{A}_n$ (car sinon B et $A \setminus B$ seraient séparés strictement par un hyperplan).

□

(L'énoncé usuel du théorème de Radon est le suivant : dans une partie A de $\mathbf{R}^n$ avec $|A| = n + 2$, on peut trouver une partition (B, B') telle que les enveloppes convexes de B et de B' se coupent).

On peut en déduire la densité entière de $\mathscr{A}_n$:

COROLLAIRE 3.5. – *La densité entière de la classe* $\mathcal{A}_n$ *est égale à* $n+1$. *La densité entière de chacune des classes* $\mathcal{A}_n^0$, $\mathcal{A}_n^1$ et $\mathcal{A}_n^1 \cup \{\phi\}$ *est égale à* n.

Démonstration. – a) Soit $(x_1, \ldots, x_{n+1})$ une base affine de $\mathbf{R}^n$ (c'est-à-dire les sommets d'un simplexe de $\mathbf{R}^n$). Alors $\mathcal{A}_n$ pulvérise $\{x_1, \ldots, x_{n+1}\}$; on a donc $\mathrm{Dens}(\mathcal{A}_n) \geqslant n+1$ et le théorème de Radon implique $\mathrm{Dens}(\mathcal{A}_n) = n+1$.

b) Soit $(x_1, \ldots, x_n)$ une base vectorielle de $\mathbf{R}^n$; alors $\mathcal{A}_n^0$, $\mathcal{A}_n^1$ et $\mathcal{A}_n^1 \cup \{\phi\}$ pulvérisent $\{x_1, \ldots, x_n\}$; leurs densités entières sont donc $\geqslant n$. Inversement, soit A une partie de $\mathbf{R}^n$ de cardinal $n+1$:

– si 0 appartient à A, A n'est évidemment pas pulvérisée par $\mathcal{A}_n^0$;

– si 0 n'appartient pas à A, il existe (par le théorème de Radon) une partie B de A non vide et n'appartenant pas à $(\mathrm{A} \cup \{0\}) \cap \mathcal{A}_n$, donc n'appartenant pas à $\mathrm{A} \cap \mathcal{A}_n^0$, ni à $\mathrm{A} \cap \mathcal{A}_n^1$, ni à $\mathrm{A} \cap (\mathcal{A}_n^1 \cup \{\phi\})$.

□

Avant de déterminer les autres indices de densité de la classe $\mathcal{A}_n$, nous devons faire quelques observations :

3.6. – a) la classe $\mathcal{A}_n^0$ est la classe $\mathrm{T}^{-1}(\mathbf{R}^n \backslash \{0\})$ engendrée par la structure d'incidence $\mathrm{T} = \{(x, y) \in \mathbf{R}^n \times (\mathbf{R}^n \backslash \{0\}) \mid (x|y) \geqslant 1\}$; la classe duale $\mathrm{T}(\mathbf{R}^n)$ est $\mathcal{A}_n^1 \cup \{\phi\}$

b) soit $\mathcal{S}$ une partie non vide de 2^{X} ; les classes $\mathcal{S}$ et $\mathcal{S} \cup \{\phi\}$ ont alors même densité entière duale et même densité réelle duale (car $\mathfrak{A}$ et $\mathfrak{A} \cup \{\phi\}$ ont mêmes atomes, dès que $\mathfrak{A}$ est une partie non vide de $\mathcal{S}$) ; de plus on a $\mathrm{dens}(\mathcal{S}) = \mathrm{dens}(\mathcal{S} \cup \{\phi\})$, comme cela résulte de 2.4.

PROPOSITION 3.7. – *Soit* n *un entier* $\geqslant 1$. *Excepté* $\mathrm{Dens}(\mathcal{A}_n)$ (*qui est égal à* $n+1$), *tous les indices de densité des classes* $\mathcal{A}_n$, $\mathcal{A}_n^0$ et $\mathcal{A}_n^1$ *sont égaux à* n.

Démonstration. – a) Observons d'abord qu'on a :

$$\mathrm{Dens}(\mathscr{A}_n^1) \leqslant \mathrm{Dens}(\mathscr{A}_n^0) = \mathrm{Dens}^*_{\mathbf{R}^n\setminus\{0\}}(\mathscr{A}_n^1 \cup \{\phi\}) = \mathrm{Dens}^*_{\mathbf{R}^n\setminus\{0\}}(\mathscr{A}_n^1),$$

$$\mathrm{dens}(\mathscr{A}_n^1) \leqslant \mathrm{dens}(\mathscr{A}_n^0) = \mathrm{dens}^*(\mathscr{A}_n^1 \cup \{\phi\}) = \mathrm{dens}^*(\mathscr{A}_n^1),$$

$$\mathrm{Dens}^*_{\mathbf{R}^n\setminus\{0\}}(\mathscr{A}_n^1) \leqslant \mathrm{Dens}^*_{\mathbf{R}^n}(\mathscr{A}_n^0) = \mathrm{Dens}(\mathscr{A}_n^1 \cup \{\phi\}) = \mathrm{Dens}(\mathscr{A}_n^1),$$

$$\mathrm{dens}^*(\mathscr{A}_n^1) \leqslant \mathrm{dens}^*(\mathscr{A}_n^0) = \mathrm{dens}(\mathscr{A}_n^1 \cup \{\phi\}) = \mathrm{dens}(\mathscr{A}_n^1)$$

(chaque fois, l'inégalité provient de 2.3, la première égalité provient de 3.6a et de 2.11, la seconde égalité provient de 3.5 ou de 3.6b).

b) Soit $\mathfrak{A} = \{S_1, \dots, S_r\}$ une partie de $\mathscr{A}_n^0$; pour chaque $i = 1, \dots, r$ on note ℓ_i le bord de S_i ; le nombre d'atomes de $\mathfrak{A}$ dans $\mathbf{R}^n$ est donc le nombre de composantes connexes de $\mathbf{R}^n \setminus \bigcup_{i=1}^{r} \ell_i$; ce nombre est donc $\leqslant \varphi(r, n)$, avec égalité si les hyperplans ℓ_i sont en position générale (voir par exemple [11]); on a donc $\mathrm{dens}^*(\mathscr{A}_n^0) = n$.

c) On voit donc, en tenant compte de a), de b) et de 3.5, que les indices de densité des classes $\mathscr{A}_n^0$ et $\mathscr{A}_n^1$ sont tous égaux à n.

d) Enfin soient $\overline{\mathscr{A}}_n$ et $\overline{\mathscr{A}}_n^0$ les classes complétées de $\mathscr{A}_n$ et $\mathscr{A}_n^0$ pour la convergence simple. La classe $\overline{\mathscr{A}}_n$ est la réunion de $\overline{\mathscr{A}}_n^0$ et de $\mathbf{R}^n \setminus \overline{\mathscr{A}}_n^0$, et on a donc

$$\mathrm{dens}^*(\overline{\mathscr{A}}_n) = \mathrm{dens}^*(\overline{\mathscr{A}}_n^0), \quad \mathrm{Dens}^*_{\mathbf{R}^n}(\overline{\mathscr{A}}_n) = \mathrm{Dens}^*_{\mathbf{R}^n}(\mathscr{A}_n^0)$$

(car $\mathfrak{A}$ et $\mathfrak{A} \cup (X \setminus \mathfrak{A})$ ont mêmes atomes, dès que $\mathfrak{A}$ est une partie non vide de $\overline{\mathscr{A}}_n^0$). On a aussi $\mathrm{dens}(\overline{\mathscr{A}}_n) = \mathrm{dens}(\overline{\mathscr{A}}_n^0)$ (appliquer 2.1d et 2.4). En utilisant enfin 2.15, on obtient

$$\mathrm{dens}(\mathscr{A}_n) = \mathrm{Dens}^*_{\mathbf{R}^n}(\mathscr{A}_n) = \mathrm{dens}^*(\mathscr{A}_n).$$

□

En tenant compte de 2.2, on retrouve le résultat suivant (dû à [15] et [32] et fort utile pour obtenir des classes de Vapnik-Cervonenkis) :

COROLLAIRE 3.8. – *Soit* F *un espace vectoriel de fonctions réelles sur un ensemble* X ; *on note* $\mathscr{C}_F$ *la classe de toutes les parties de* X *de la forme* $\{x \in X \mid \varphi(x) \geqslant 0\}$ (*pour tout* $\varphi \in F$). *On suppose que l'espace vectoriel* F *est de dimension* n. *On a alors* $\mathrm{Dens}(\mathscr{C}_F) \leqslant n$ (Dudley [15]) *et* $\mathrm{Dens}^*_X(\mathscr{C}_F) \leqslant n$ (Shapiro [32]).

Démonstration. – Soit $(\varphi_1, \ldots, \varphi_n)$ une base de l'espace vectoriel F ; notons f l'application $x \longrightarrow (\varphi_1(x), \ldots, \varphi_n(x))$ de X dans $\mathbf{R}^n$. Soit $\mathcal{V}_n$ la classe de tous les demi-espaces vectoriels fermés de $\mathbf{R}^n$; la classe $\mathcal{V}_n \cup \{\mathbf{R}^n\}$ est donc incluse dans la classe complétée (pour la convergence simple) de la classe $\mathbf{R}^n \backslash \mathscr{A}_n^0$. Or $\mathcal{P}_F$ est l'image réciproque de $\mathcal{V}_n \cup \{\mathbf{R}^n\}$ par f. Le corollaire résulte donc de 3.7 et de divers résultats du paragraphe 2.

□

Boules d'un espace normé.

3.9. – Soit $E = (\mathbf{R}^n, \|\ \|)$ un espace normé de dimension n ; on note $\mathcal{B}_E$ la classe de toutes les boules fermées de l'espace normé E.

Bien que $\mathcal{B}_E$ ne soit pas toujours une classe de Vapnik-Cervonenkis, nous allons indiquer des cas où la densité réelle de $\mathcal{B}_E$ est égale à $n + 1$ (c'est-à-dire au nombre "réel" de paramètres : n pour le centre et un pour le rayon) ; c'est ce qui va se passer notamment en dimension 2.

LEMME 3.10. – *Soit* B *une partie convexe compacte non vide de* $\mathbf{R}^2$ *et soit* $B' = \lambda B + c$ (*avec* $\lambda \in]0, +\infty[$ et $c \in \mathbf{R}^2$). *Il existe alors deux demi-espaces affines fermés* T *et* T' *tels que* $B \backslash B' \subset T \backslash T'$ *et* $B' \backslash B \subset T' \backslash T$.

Démonstration. – Si $B \subset B'$ ou $B' \subset B$, le résultat est évident ; on suppose donc que $B \backslash B'$ et $B' \backslash B$ sont non vides.

a) Si λ est différent de 1, on peut (par translation) se ramener à $c = 0$; alors 0 n'appartient ni à B, ni à B' (car sinon on aurait $B \subset B'$) ou $B' \subset B$) ; on peut donc trouver une application projective π envoyant 0 à l'infini et n'envoyant aucun point de B à l'infini (envoyer à l'infini une droite passant par 0 et ne coupant pas B). L'application π transforme donc B et B' en deux parties convexes compactes de $\mathbf{R}^2$, translatées l'une de l'autre.

b) On s'est donc ramené au cas $\lambda = 1$; on peut donc (quitte à changer de coordonnée) supposer qu'on a

$$B = \{(x, t) \in \mathbf{R}^2 \mid f(x) \leqslant t \leqslant g(x)\}$$

et
$$B' \doteq \{(x, t) \in \mathbf{R}^2 \mid f(x) \leqslant t - 1 \leqslant g(x)\},$$

où f est une application convexe de $\mathbf{R}$ dans $\mathbf{R} \cup \{+\infty\}$ et g

une application concave de $\mathbf{R}$ dans $\mathbf{R} \cup \{-\infty\}$. Soit h_0 la plus petite fonction concave $\geqslant \operatorname{Inf}(g, f+1)$ et h_1 la plus grande fonction convexe $\leqslant \operatorname{Sup}(f, g+1)$. L'ensemble $\{x \in \mathbf{R} \mid g(x) > f(x)+1\}$ est, s'il n'est pas vide, un intervalle ouvert $]a, b[$ et on a, pour chaque $t \in]0,1[$, $h_0((1-t)a+tb) = h_1((1-t)a+tb) = (1-t)g(a)+tg(b)$. On a donc $h_0(x) \leqslant h_1(x)$, quel que soit $x \in \mathbf{R}$. Soit alors h une application affine de $\mathbf{R}$ dans $\mathbf{R}$, avec $h_0 \leqslant h \leqslant h_1$; on pose

$$\mathrm{T} = \{(x, t) \in \mathbf{R}^2 \mid h(x) \geqslant t\} \quad \text{et} \quad \mathrm{T}' = \{(x, t) \in \mathbf{R}^2 \mid h(x) \leqslant t\}$$

et on a donc $\mathrm{B} \setminus \mathrm{B}' \subset \mathrm{T} \setminus \mathrm{T}'$ et $\mathrm{B}' \setminus \mathrm{B} \subset \mathrm{T}' \setminus \mathrm{T}$.

□

COROLLAIRE 3.11. – *Soit* $\mathrm{E} = (\mathbf{R}^2, \| \ \|)$ *un espace normé de dimension 2 ; on a alors* $\operatorname{Dens}(\mathcal{B}_{\mathrm{E}}) \leqslant 3$.

Démonstration. – La classe $\mathcal{B}_{\mathrm{E}}$ est séparée par la classe $\mathcal{A}_2$ des demi-espaces affines fermés de $\mathbf{R}^2$ (ainsi qu'on l'a vu en 3.10). Or la densité entière de $\mathcal{A}_2$ est égale à 3 (théorème de Radon). La densité entière de $\mathcal{B}_{\mathrm{E}}$ est donc $\leqslant 3$ (voir 2.17).

□

3.12. – En fait on a même $\operatorname{Dens}(\mathcal{B}_{\mathrm{E}}) = 3$, à cause de la remarque suivante : on a $\operatorname{Dens}(\mathcal{B}_{\mathrm{E}}) \geqslant n+1$, quel que soit l'espace normé $\mathrm{E} = (\mathbf{R}^n, \| \ \|)$ de dimension n (en effet $\mathcal{A}_n$ est inclus dans la classe complétée $\overline{\mathcal{B}}_{\mathrm{E}}$ de $\mathcal{B}_{\mathrm{E}}$, et 2.15 implique donc

$$n+1 = \operatorname{Dens}(\mathcal{A}_n) \leqslant \operatorname{Dens}(\overline{\mathcal{B}}_{\mathrm{E}}) = \operatorname{Dens}(\mathcal{B}_{\mathrm{E}})).$$

La propriété de séparation des boules par des demi-espaces affines reste vraie pour les boules euclidiennes dans $\mathbf{R}^n$, quel que soit n :

PROPOSITION 3.13 (Dudley [15]. – *Soit n un entier $\geqslant 1$ et soit $\mathcal{B}_n$ la classe de toutes les boules euclidiennes fermées dans* $\mathbf{R}^n$; *on a alors* $\operatorname{Dens}(\mathcal{B}_n) = n+1$.

Démonstration. – La classe $\mathcal{B}_n$ est séparée par la classe $\mathcal{A}_n$ des demi-espaces affines fermés de $\mathbf{R}^n$ (en effet, si B et B' sont deux boules euclidiennes, l'intersection de $\partial\mathrm{B}$ et de $\partial\mathrm{B}'$ est portée par un hyperplan affine). On applique alors le théorème de Radon, puis 2.17 et 3.12, et on obtient $\operatorname{Dens}(\mathcal{B}_n) = n+1$.

□

Mais cette propriété de séparation n'est plus vraie pour un espace normé quelconque de dimension $n \geqslant 3$, comme le montre l'exemple suivant.

3.14. – Soit $E = (\mathbf{R}^n, \| \ \|)$ un espace normé de dimension $n \geqslant 3$; on suppose que la boule unité B de E est un polyèdre.

a) Soit B′ une boule de rayon assez petit centrée au milieu d'une arête de B ; alors on ne peut séparer B et B′ (au sens de 2.17) par des semi-espaces affines.

b) On a cependant $\mathrm{dens}(\mathcal{B}_E) \leqslant kn$, où k est le nombre de facettes de B ($\mathcal{B}_E$ est inclus dans $(\mathcal{A}_n)^{(k)}$; cela résulte donc de 2.6).

La majoration obtenue en 3.14b est probablement assez mauvaise ; elle est en tout cas très éloignée de la minoration suivante :

Proposition 3.15. – *Soit* $E = (\mathbf{R}^n, \| \ \|)$ *un espace normé de dimension* n. *On a alors* $\mathrm{dens}(\mathcal{B}_E) \geqslant n + 1$.

Démonstration. – On la trouvera à la fin du paragraphe 4.

□

La densité réelle de $\mathcal{B}_E$ est donc égale au nombre "réel" de paramètres quand E est un espace euclidien ou lorsque sa dimension est 2 (comme cela résulte de 3.11 et de 3.13). Voici un autre exemple de cette situation :

3.16. – On dit qu'un espace normé $E = (\mathbf{R}^n, \| \ \|)$ est *à norme polynomiale,* s'il existe un entier pair p tel que la fonction $x \longrightarrow \|x\|^p$ soit un polynôme ; c'est le cas si l'espace normé E est un sous-espace normé d'un espace L^p (avec p entier pair), mais on sait (Reznick [26]) qui existe d'autres espaces normés à norme polynomiale.

Proposition 31.17. – *Soit* $E = (\mathbf{R}^n, \| \ \|)$ *un espace normé de dimension* n *à norme polynomiale. On a alors* $\mathrm{dens}(\mathcal{B}_E) \leqslant n + 1$.

Démonstration. – On la trouvera ci-dessous, après 3.29.

□

Mais nous allons voir que $\mathcal{B}_E$ n'est pas toujours une classe de Vapnik-Cervonenkis.

LEMME 3.18. – *Il existe une partie convexe compacte symétrique* B_0 *de* $\mathbf{R}^3$ *vérifiant la propriété suivante :*

(3.18.1) *la famille des intersections du plan*

$$H = \{x = (x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_3 = 0\}$$

avec les translatés de B_0 *contient un homothétique de chaque élément de la classe* Σ *(rappelons que* Σ *est la classe de tous les polygones convexes compacts de* $\mathbf{R}^2$ *qui sont symétriques autour de* 0 *et à sommets rationnels).*

Démonstration. – Soit $(P_0, P_1, P_2, \ldots, P_n, \ldots)$ une énumération des éléments de Σ. Pour chaque $i \in \mathbf{N}$, on se donne $a_i, b_i \in]0, +\infty[$ avec $a_i B \subset P_i \subset b_i B$ (où B est la boule euclidienne de centre 0 et de rayon 1). On considère maintenant $P_0, P_1, P_2, \ldots, P_n, \ldots$ comme des parties de H et on note v le vecteur $(0, 0, 1)$. Pour chaque entier $i \geqslant 0$, on pose

$$\lambda_i = \frac{1}{b_i} \prod_{k=0}^{i-1} \frac{a_k}{2b_k}, \quad \mu_i = \frac{1}{2} \sum_{k=i}^{+\infty} a_k\, b_k\, \lambda_k^2$$

et $c_i = (1 - \mu_i)\, v$. Soit B_0 l'enveloppe convexe fermée de $\bigcup_{i \in \mathbf{N}} ((\lambda_i P_i + c_i) \cup (\lambda_i P_i - c_i))$; on voit aisément que B_0 est une partie convexe compacte symétrique de $\mathbf{R}^3$ et que l'intersection de $B_0 - c_i$ avec H est égale à $\lambda_i P_i$.

□

PROPOSITION 3.19. – *Soit* $E = (\mathbf{R}^3, \|\ \|)$ *un espace normé dont la boule unité* B_0 *possède la propriété* (3.18.1). *Alors* $\mathcal{B}_E$ *n'est pas une classe de Vapnik-Cervonenkis.*

Démonstration. – Soit H le plan

$$\{x = (x_1, x_2, x_3) \in \mathbf{R}^3 \mid x_3 = 0\}.$$

La propriété (3.18.1) implique que la classe $H \cap \mathcal{B}_E$ contient la classe Σ. Or la classe Σ n'est pas une classe de Vapnik-Cervonenkis (voir 3.2) ; cela entraîne le résultat.

□

Voici quelques autres observations au sujet de la famille des boules d'un espace normé.

3.20. – a) Il est facile de construire (à partir du convexe B_0 de $\mathbf{R}^3$ construit ci-dessus, en 3.18) une partie convexe compacte symétrique de $\mathbf{R}^4$ dont la classe des translatés soit de densité infinie. Mais existe-t-il une partie convexe de $\mathbf{R}^3$ dont la classe des translatés soit de densité infinie ?

b) La Proposition 3.11 ci-dessus donne, pour la dimension 2, un début de solution au problème suivant : pour un entier n donné, caractériser les espaces métriques (X, d) qui sont sous-espaces métriques d'un espace normé de dimension n (c'est la version isométrique du problème de plongement lipschitzien dans $\mathbf{R}^n$ partiellement résolu par [4]).

Jusqu'à présent, les exemples de classes de Vapnik-Cervonenkis que nous avons donnés et les calculs de densité que nous avons effectués pouvaient tous être déduits, par une des opérations du § 2, des exemples 2.18a (classes totalement ordonnées) et 3.7 (demi-espaces affines dans $\mathbf{R}^n$). Nous allons envisager maintenant d'autres exemples.

Arrangements de pseudodroites.

3.21. – a) On appelle *arrangement de pseudodroites* (voir Grünbaum [21]) toute famille $(L_i)_{i \in I}$ de courbes du plan affine vérifiant :

– pour chaque $i \in I$, la courbe L_i partage le plan en 2 régions connexes et non bornées (l'adhérence de chacune de ces 2 régions s'appelle un *pseudodemiplan*),

– le cardinal de $L_i \cap L_j$ est égal à 1, quels que soient $i, j \in I$ (avec $i \neq j$).

b) Tout arrangement de 8 pseudodroites est l'image, par un homémorphisme du plan, d'un arrangement de 8 droites (Goodman, Pollack [20]). Par contre, il existe un arrangement de 9 pseudodroites en position générale non isomorphe (au sens ci-dessus) à un arrangement de 9 droites (Ringel [27]).

LEMME 3.22. – *Soit* $\mathfrak{S}$ *la classe des pseudodemiplans correspondant à un arrangement* $(L_i)_{i \in I}$ *de pseudodroites ; on a alors* :

a) $\mathrm{Dens}^*(\mathfrak{S}) \leqslant 2$; b) $\mathrm{Dens}(\mathfrak{S}) \leqslant 3$

((a) *est dû à* [21] *et* b) *est l'analogue du théorème de Radon*).

Démonstration. – a) On peut montrer en effet (voir [21], p. 45) que 3 pseudodemiplans déterminent au plus 7 atomes.

b) Soit A une partie de $\mathbf{R}^2$ de cardinal 4 ; si les 7 partitions non triviales de A étaient déterminées par 7 pseudodroites, l'arrangement de ces 7 pseudodroites serait (à cause de 3.21b) isomorphe à un arrangement de 7 droites, ce qui est impossible à cause du théorème de Radon.

□

Classe de toutes les variétés d'un matroïde.

Nous suivons la présentation de Cartier [11] (voir aussi Welsh [42]) :

3.23. – a) Un *matroïde* $M = (X, \mathcal{H})$ est un ensemble fini X muni d'une classe $\mathcal{H}$ de parties de X (appelées les *hyperplans* de M) satisfaisant aux conditions suivantes :

– l'ensemble X lui même n'est pas un hyperplan de M ;

– si H_1 et H_2 sont des hyperplans de M, alors $H_1 \subset H_2$ implique $H_1 = H_2$;

– si H_1 et H_2 sont des hyperplans de M, x un point de $X \setminus (H_1 \cup H_2)$ et y un point de $H_1 \setminus H_2$, alors il existe un hyperplan H_3 contenant $H_1 \cap H_2$ et x, mais ne contenant pas y.

b) Soit $M = (X, \mathcal{H})$ un matroïde ; une partie A de X est dite *indépendante dans* M si $A \setminus \{x\}$ appartient à $A \cap \mathcal{H}$, quel que soit $x \in A$;

c) pour chaque partie A de X, on note $\rho(A) + 1$ la borne supérieure des cardinaux des parties de A indépendantes dans M et l'entier $\rho(A)$ est appelé le *rang de* A *dans* M (l'entier $\rho(X)$ est appelé la *dimension du matroïde* M) ;

d) on appelle *variété de* M ("flat" en anglais) toute intersection d'un nombre fini d'hyperplans de M (en particulier X est une variété de M).

On a alors le résultat facile suivant (où X pourrait très bien être infini) :

LEMME 3.24. – *Soit* $M = (X, \mathcal{H})$ *un matroïde de dimension* n. *Alors la classe de toutes les variétés de* M *a une densité entière égale à* $n + 1$.

Démonstration. – Elle sera donnée après 5.2, car la structure de matroïde n'y joue réellement pas de rôle.

□

Il serait beaucoup plus intéressant d'évaluer la densité de la classe des demi-espaces dans un matroïde orienté :

3.25. – Orienter un matroïde $M = (X, \mathcal{H})$ consiste à se donner, pour chaque hyperplan H de M, une partition de X\H en deux demi-espaces ouverts, de façon à satisfaire à une version "orientée" des axiomes 3.23a (voir Bland, Las Vergnas [8]). Conserve-t-on alors les résultats démontrés en 3.7 pour l'espace affine de dimension n (qui est un matroïde orienté de dimension n) ?

Densités des classes de demi-espaces algébriques :

3.26. – Soient n, d deux entiers $\geqslant 0$; on note $\mathcal{P}_{n,d}$ la classe des *demi-espaces algébriques de* $\mathbf{R}^n$ *de degré* $\leqslant d$, c'est-à-dire des parties de $\mathbf{R}^n$ de la forme $\{x \in \mathbf{R}^n \mid p(x) \geqslant 0\}$, où p est un polynôme réel à n variables et de degré $\leqslant d$.

Notons que les polynômes réels à n variables et de degré $\leqslant d$ forment un espace vectoriel de dimension $\binom{n+d}{n}$; on pourrait donc déduire de 3.8 l'évaluation $\operatorname{dens}^*(\mathcal{P}_{n,d}) \leqslant \binom{n+d}{n}$.

En fait on peut obtenir une bien meilleure évaluation de la densité réelle duale de la classe $\mathcal{P}_{n,d}$. C'est une conséquence directe du beau résultat suivant (dû à Warren [40], [41]) :

3.27. – a) Si t est un nombre réel, on pose $\operatorname{sgn}(t) = 1$ si $t > 0$, et $\operatorname{sgn}(t) = -1$ sinon ;

b) soit $F = (f_1, \ldots, f_r)$ une famille de r applications de X dans $\mathbf{R}$; on appelle *suite de signes engendrée par* F *sur* X toute suite de signes de la forme $(\operatorname{sgn}(f_1(x)), \ldots, \operatorname{sgn}(f_r(x)))$, où x est un point de X.

PROPOSITION 3.28 [41]. – *Soient* $p_1, \ldots, p_r$ *des polynômes réels à* n *variables et de degré* $\leqslant d$. *Alors le nombre de suites de signes différentes engendrées par* $(p_1, \ldots, p_r)$ *sur* $\mathbf{R}^n$ *est inférieur* *ou égal à* $\left(\frac{4e}{n}\, dr\right)^n$. *En particulier, pour* $d \geqslant 2$, *il est strictement inférieur à* 2^r *dès que* $r \geqslant 8n\,\mathrm{Log}_2(d)$.

Démonstration. – On renvoie à [41].

□

Clairement les suites de signes $(\mathrm{sgn}(f_1(x)), \ldots, \mathrm{sgn}(f_r(x)))$ correspondent aux atomes de la classes $\mathcal{P}_{\mathrm{F}}$ de toutes les parties de X de la forme $\{x \in \mathrm{X} \mid f_i(x) \geqslant 0\}$ (pour $i = 1, \ldots, r$). En d'autres termes on a donc :

COROLLAIRE 3.29. – *Soient* n, d *deux entiers* $\geqslant 1$; *on a alors* $\mathrm{dens}^*(\mathcal{P}_{n,d}) = n$. *De plus, pour* $d \geqslant 2$, *on a*

$$\mathrm{Dens}^*_{\mathbf{R}^n}(\mathcal{P}_{n,d}) < 8n\,\mathrm{Log}_2(d).$$

Enfin, pour $d = 1$, *on a* $\mathcal{P}_{n,d} = \mathcal{A}_n$ *et donc* $\mathrm{Dens}^*_{\mathbf{R}^n}(\mathcal{P}_{n,d}) = n$.

Démonstration. – a) Il résulte de 3.28 qu'on a $\mathrm{dens}^*(\mathcal{P}_{n,d}) \leqslant n$. De plus $\mathcal{P}_{n,d}$ contient $\mathcal{P}_{n,1}$, c'est-à-dire $\mathcal{A}_n$; on a donc

$$\mathrm{dens}^*(\mathcal{P}_{n,d}) = n.$$

b) Les deux autres assertions résultent directement de 3.28 et de 3.7.

□

Le caractère remarquable de 3.28 tient à ce que la densité réelle duale de la classe $\mathcal{P}_{n,d}$ ne dépend que du nombre de variables, et non du degré. Nous en déduisons immédiatement la démonstration de 3.17 :

Démonstration de la Proposition 3.17. – Soit $\mathrm{E} = (\mathbf{R}^n, \|\ \|)$ un espace normé à norme polynomiale ; de façon précise, soit p un entier pair tel que la fonction $x \longrightarrow \|x\|^p$ soit un polynôme (qui est donc de degré p). On pose alors

$$\mathrm{T} = \{(x,(y,t)) \in \mathbf{R}^n \times (\mathbf{R}^n \times \mathbf{R}^+) \mid \|x-y\|^p \leqslant t\}\ ;$$

c'est une structure d'incidence sur $(\mathbf{R}^n, \mathbf{R}^n \times \mathbf{R}^+)$; la classe $\mathrm{T}^{-1}(\mathbf{R}^n \times \mathbf{R}^+)$ n'est autre que $\mathcal{B}_{\mathrm{E}}$; sa classe duale $\mathrm{T}(\mathbf{R}^n)$ est incluse dans $\mathcal{P}_{n+1,p}$. On a donc

$$\mathrm{dens}(\mathcal{B}_{\mathrm{E}}) = \mathrm{dens}^*(\mathrm{T}(\mathbf{R}^n)) \leqslant \mathrm{dens}^*(\mathcal{P}_{n+1,p}) \leqslant n+1$$

(observons qu'on a aussi $\mathrm{dens}^*(\mathcal{B}_{\mathrm{E}}) \leqslant n$, car $\mathcal{B}_{\mathrm{E}}$ est incluse dans $\mathcal{P}_{n,p}$).

□

Signalons enfin qu'une motivation pour établir des résultats tels que 3.28 est leur utilité pour minorer des vitesses d'approximation (voir [25], [32] et [40])

4. Dimension et densité.

Je vais rappeler un peu de terminologie concernant les espaces métriques en suivant [5] :

4.1. – Soit X un ensemble muni d'un écart d :

a) une partie Y de X est dite *ϵ-discernable dans* (X, d) (pour $\epsilon \in]0, +\infty[$) si on a $d(y, y') \geqslant \epsilon$, pour chaque couple (y, y') de points distincts de Y ;

b) soit Z une partie de X ; la quantité $\mathrm{Sup}\,\{d(z, z') \mid z, z' \in \mathrm{Z}\}$ est alors appelée le *diamètre de* Z *dans* (X, d) ;

c) la borne supérieure des cardinaux des parties ϵ-discernables dans (X, d) est notée $\mathrm{N}(\mathrm{X}, d\,; \epsilon)$;

d) la *dimension d'entropie de* (X, d), notée $\overline{\dim}(\mathrm{X}, d)$, est la borne inférieure des nombres $s \in]0, +\infty[$ vérifiant la propriété suivante : pour chaque $b \in]0, +\infty[$, il existe $\mathrm{C}(b) \in]0, +\infty[$ tel qu'on ait $\mathrm{N}(\mathrm{Z}, d\,; \epsilon) \leqslant \mathrm{C}(b)\,\epsilon^{-s}$, quel que soit $\epsilon \in]0, b]$ et quelle que soit la partie Z de (X, d) de diamètre $\leqslant b$.

La définition de la dimension d'entropie se simplifie évidemment pour un espace de diamètre fini ; c'est le cas notamment lorsque cet espace est une partie $\mathcal{S}$ de 2^{X} munie de *l'écart de la différence symétrique pour une loi* P :

4.2. – Soit $\mathcal{S}$ une partie de 2^{X} :

a) on appellera (un peu abusivement) *loi sur* $(\mathrm{X}, \mathcal{S})$ toute loi de probabilité P sur X muni de la σ-algèbre engendrée par $\mathcal{S}$;

b) le noyau $\mathrm{S}, \mathrm{S}' \longrightarrow \mathrm{P}(\mathrm{S} \,\Delta\, \mathrm{S}')$ est alors un écart sur $\mathcal{S}$, qui est noté d_p (rappelons que la *différence symétrique* $\mathrm{S} \,\Delta\, \mathrm{S}'$ est l'ensemble $(\mathrm{S}\backslash\mathrm{S}') \cup (\mathrm{S}'\backslash\mathrm{S})$).

Nous allons montrer que la borne supérieure de $\overline{\dim}(\mathcal{S}, d_p)$ (pour toutes les lois P sur $(X, \mathcal{S})$) est égale à la densité réelle de la classe $\mathcal{S}$. Ce résultat, le principal de ce paragraphe, sera établi en deux temps :

– la dimension d'entropie de $(\mathcal{S}, d_p)$ est toujours inférieure ou égale à $\mathrm{dens}(\mathcal{S})$ (ce résultat est dû à Dudley [14] ; nous le démontrons pour être complet) ;

– on peut trouver des lois P sur $\mathcal{S}$ telles que $\overline{\dim}(\mathcal{S}, d_p)$ soit arbitrairement proche de $\mathrm{dens}(\mathcal{S})$ (ce résultat est nouveau).

PROPOSITION 4.3 ([14] p. 922). – *Soit* $\mathcal{S}$ *une partie de* 2^X *avec* $\mathrm{dens}(\mathcal{S}) = s$. *On a alors* $\overline{\dim}(\mathcal{S}, d_p) \leqslant s$, *quelle que soit la loi* P *sur* $(X, \mathcal{S})$.

Démonstration. – Soit $\{S_1, S_2, \dots, S_N\}$ une partie finie ϵ-discernable de $(\mathcal{S}, d_p)$ (avec $\epsilon \in]0,1]$) ; autrement dit on a $P(S_i \Delta S_j) \geqslant \epsilon$ pour tous $i, j \in \{1, \dots, N\}$ $(i \neq j)$.

a) Soit n un entier $\geqslant 1$ à déterminer plus tard ; on note P^n le produit tensoriel de n lois égales à P ; on a alors :

$$P^n \{(x_1, \dots, x_n) \in X^n \mid \exists\, i, j \in \{1, \dots, N\} (i \neq j),$$

$$\forall\, k \in \{1, \dots, N\}, x_k \notin S_i \Delta S_j\}$$

$$\leqslant \sum_{i=1}^{N} \sum_{j=1}^{N} (P(X \backslash (S_i \Delta S_j))^n \, 1_{i \neq j} \leqslant N^2 (1 - \epsilon)^n .$$

On fixe r, t avec $s < r < t < +\infty$; puisque $\mathcal{S}$ est de densité réelle s, il existe $C \in]0, +\infty[$ tel qu'on ait $|A \cap \mathcal{S}| \leqslant C |A|^r$, quelle que soit A partie finie de X. Si on a $N^2 (1 - \epsilon)^n < 1$, l'inégalité ci-dessus montre qu'il existe $(x_1, \dots, x_n) \in X^n$ tel que les intersections de $S_1, \dots, S_N$ avec $A = \{x_1, \dots, x_n\}$ soient distinctes. On a donc $|A \cap \mathcal{S}| \geqslant N$; or on a $|A| \leqslant n$; on a donc $N \leqslant C n^r$.

b) On pose maintenant $\eta = \dfrac{1}{r} - \dfrac{1}{t}$ et on choisit n tel que $(n - 1)\epsilon \leqslant N^\eta - 1 < n\epsilon$. On fixe d'autre part un entier N_0 tel qu'on ait $M^2 \leqslant \exp(M^\eta - 1)$ pour tout $M \geqslant N_0$.

Il y a alors deux possibilités :

– ou bien N est inférieur à N_0 ;

– ou bien on a $N \geqslant N_0$, et donc $N^2 \leqslant \exp(N^\eta - 1)$; dans ce cas, on a alors $N^\eta - 1 < n\epsilon$ (par définition de n), donc $N^2 < \exp(n\epsilon)$ (à cause de l'inégalité ci-dessus), donc $N^2 (1-\epsilon)^n < 1$ (car on a $1 - \epsilon \leqslant \exp(-\epsilon)$), donc $N \leqslant C n^r$ (voir (a) ci-dessus) ; or on a $(n-1)\epsilon \leqslant N^\eta - 1$ (par définition de n), donc $n\epsilon \leqslant N^\eta$, donc $n^{rt} \leqslant \epsilon^{-rt} N^{t-r}$; on a donc $N^t \leqslant C^t \epsilon^{-rt} N^{t-r}$, c'est à dire $N \leqslant C^{t/r} \epsilon^{-t}$.

c) On a donc, quel que soit le cas considéré, $N \leqslant (N_0 + C^{t/r})\epsilon^{-t}$. Cela implique qu'on a $N(X, d ; \epsilon) \leqslant (N_0 + C^{t/r})\epsilon^{-t}$, pour chaque $\epsilon \in]0,1]$. Comme t peut être pris arbitrairement près de s, cela donne le résultat.

□

On peut, dans le résultat précédent, remplacer les lois de probabilité par des *mesures* $\geqslant 0$ *de masse totale finie et non nécessairement σ-additives* :

4.4. – a) On peut parfaitement supposer en 4.3 que P est seulement une mesure additive (positive et de masse totale 1) sur l'algèbre de parties de X engendrée par la classe $\mathfrak{S}$: en effet, on se ramène alors aux hypothèses de 4.3 en passant au compactifié de Stone (ce qui signifie, comme on le vérifie aisément : prendre la classe duale de $\mathfrak{S}$ (2.8), la complèter (2.14) et reprendre la classe duale).

b) On peut même supposer que P est seulement une mesure de masse totale finie (positive et additive) sur l'algèbre de parties de X engendrée par la classe $\mathfrak{S}$ (en effet la dimension d'entropie de $(\mathfrak{S}, d_{\lambda P})$ est la même que celle de $(\mathfrak{S}, d_p)$).

c) Mais la conclusion de 4.3 ne subsiste pas si on passe aux mesures de masse totale infinies, méme σ-finies (prendre $\mathfrak{S} = \binom{\mathbf{N}}{1}$, et pour mesure la cardinalité).

Voici maintenant la réciproque (annoncée dans [5]) de la Proposition 4.3 :

PROPOSITION 4.5. – *Soit* $\mathfrak{S}$ *une partie de* 2^X *avec* $\operatorname{dens}(\mathfrak{S}) > s$. *On peut alors trouver une loi* P *sur* $(X, \mathfrak{S})$ *telle que* $\overline{\dim}(\mathfrak{S}, d_p) > s$.

Démonstration. – On fixe un nombre t avec $s < t < \operatorname{dens}(\mathcal{S})$; on peut donc trouver, pour chaque entier $j \geqslant 1$, une partie finie et non vide A_j de l'ensemble X vérifiant $|A_j \cap \mathcal{S}| \geqslant j\,|A_j|^t$; cela entraîne en particulier $\lim\limits_{j \to +\infty} |A_j| = +\infty$.

a) On pose $B_0 = A_1$, puis on choisit un entier $k \geqslant 2\,|B_0 \cap \mathcal{S}|$ tel que $|A_k \backslash B_0| \geqslant 2$, et on pose $B_1 = A_k \backslash B_0$. On a donc

$$|B_1 \cap \mathcal{S}| \geqslant \frac{2}{k}\,|B_0 \cap \mathcal{S}|\,|B_1 \cap \mathcal{S}| \geqslant \frac{2}{k}\,|A_k \cap \mathcal{S}| \geqslant 2\,|B_1|^t .$$

b) Supposons qu'on ait construit $B_0, \ldots, B_{n-1}$ inclus dans X, disjoints et vérifiant $|B_j \cap \mathcal{S}| \geqslant 2^j\,|B_j|^t$ et $|B_j| \geqslant 2^j$, quel que soit $j = 0, \ldots, n-1$. On choisit alors un entier

$$\ell \geqslant 2^n \left|\left(\bigcup_{j=0}^{n-1} B_j\right) \cap \mathcal{S}\right| \quad \text{tel que} \quad \left|A_\ell \backslash \bigcup_{j=0}^{n-1} B_j\right| \geqslant 2^n ,$$

et on pose $B_n = A_\ell \backslash \left(\bigcup\limits_{j=0}^{n-1}\right) B_j$. Les ensembles $B_0, \ldots, B_n$ sont donc disjoints et on a $|B_n| \geqslant 2^n$; on a aussi

$$|B_n \cap \mathcal{S}| \geqslant \frac{2^n}{\ell}\left|\left(\bigcup_{j=0}^{n-1} B_j\right) \cap \mathcal{S}\right|\,|B_n \cap \mathcal{S}| \geqslant \frac{2^n}{\ell}\,|A_\ell \cap \mathcal{S}| \geqslant 2^n\,|B_n|^t .$$

c) On fixe maintenant $r \in]s, t[$ et on pose $\alpha_j = |B_j|^{-t/r}$, pour tout entier $j \geqslant 0$. on pose $\sigma = \sum\limits_{j=0}^{+\infty} \alpha_j\,|B_j|$ (σ est inférieur ou égal à $\sum\limits_{j=0}^{+\infty} 2^{n-n(t/r)}$ et est donc fini). On considère alors la loi de probabilité P sur X portée par $\bigcup\limits_{j=0}^{+\infty} B_j$ et donnant la masse $\frac{1}{\sigma}\,\alpha_j$ à chaque point de B_j (quel que soit $j \in \mathbf{N}$).

d) On a alors $\overline{\dim}(\mathcal{S}, d_p) \geqslant r$: en effet sinon il existerait $C \in]0, +\infty[$ tel que toute partie ϵ-discernable de $(\mathcal{S}, d_p)$ soit de cardinal $\leqslant C\,\epsilon^{-r}$, quel que soit $\epsilon \in]0,1[$. Or $B_j \cap \mathcal{S}$ est la trace sur B_j d'une partie $\frac{1}{\sigma}\,\alpha_j$-discernable de $(\mathcal{S}, d_p)$, quel que soit $j \in \mathbf{N}$. On devrait donc avoir

$$2^j\,|B_j|^t \leqslant |B_j \cap \mathcal{S}| \leqslant C\left(\frac{1}{\sigma}\,\alpha_j\right)^{-r} = C\,\sigma^r\,|B_j|^t ,$$

quel que soit $j \in \mathbf{N}$, ce qui est impossible. □

Tenant compte de 4.3 et de 4.5, on obtient :

PROPOSITION 4.6. – *Soit* $\mathcal{S}$ *une partie de* 2^X ; *on a alors*

$$\operatorname{dens}(\mathcal{S}) = \operatorname*{Sup}_{P} \overline{\dim}(\mathcal{S}, d_p),$$

où la borne supérieure est prise sur toutes les lois P *sur* $(X, \mathcal{S})$.

Extension éventuelle à d'autres indices de dimension :

Voici maintenant quelques observations concernant d'éventuels analogues de 4.6 pour d'autres indices de dimension (nous suivons [5]) :

4.7. – Soit X un ensemble muni d'un écart d :

a) on note $\dim(X, d)$ la *dimension de Hausdorff* de l'espace métrique quotient de l'espace (X, d) (voir par exemple Rogers [28]) ;

b) on note $\operatorname{Dim}(X, d)$ la *dimension métrique* de (X, d) (voir [4]), c'est-à-dire la borne inférieure des nombres $s \in]0, +\infty[$ vérifiant la propriété suivante : il existe $C \in]0, +\infty[$ tel qu'on ait $N(Z, d\ ; \epsilon) \leqslant C\left(\frac{\epsilon}{b}\right)^{-s}$, quel que soit $b \in]0, +\infty[$, quel que soit $\epsilon \in]0, b]$ et quelle que soit la partie Z de (X, d) de diamètre $\leqslant b$ (comparer à la définition de la dimension d'entropie) ;

c) on a alors toujours $\dim(X, d) \leqslant \overline{\dim}(X, d) \leqslant \operatorname{Dim}(X, d)$ (voir [5]), ce qui justifie bien les notations que nous employons.

4.8. – Soit $\mathcal{S}$ une partie de 2^X ; notons σ-dens$(\mathcal{S})$ la borne inférieure des nombres $s \in]0, +\infty[$ tels qu'il existe une partition $(\mathcal{S}_i)_{i \in \mathbf{N}}$ de la classe $\mathcal{S}$ telle qu'on ait $\operatorname{dens}(\mathcal{S}_i) \leqslant s$, quel que soit $i \in \mathbf{N}$. On a évidemment $\operatorname*{Sup}_{P} \dim(\mathcal{S}, d_p) \leqslant \sigma\text{-dens}(\mathcal{S}) \leqslant \operatorname{dens}(\mathcal{S})$. Est-il possible d'obtenir la quantité $\operatorname*{Sup}_{P} \dim(\mathcal{S}, d_p)$ d'une manière "purement combinatoire" ?

Nous allons voir que, pour une classe $\mathcal{S}$ de Vapnik-Cervonenkis, la dimension métrique de $(\mathcal{S}, d_p)$ n'est pas toujours finie.

LEMME 4.9. – *Soit* Q_2 *la classe des quadrants orientés dans* $\mathbf{R}^2$, *c'est-à-dire la classe des parties de* $\mathbf{R}^2$ *de la forme* $]-\infty, u] \times]-\infty, v]$ (*pour tous* $u, v \in \mathbf{R}$).

a) *La classe* Q_2 *est une classe de Vapnik-Cervonenkis.*

b) *Soit* P *la loi uniforme sur le segment d'extrémités* $e_1 = (1,0)$ et $e_2 = (0,1)$; *la dimension métrique de l'espace* (Q_2, d_p) *est alors infinie.*

Démonstration. – a) Il résulte de 2.5, par exemple, que Q_2 est une classe de Vapnik-Cervonenkis (pour plus de précisions, voir [43]).

b) En effet, soit $b \in]0, \frac{1}{2}]$; notons R_t le quadrant orienté de sommet $t e_1 + (b + 1 - t) e_2$, quel que soit $t \in [b, 1]$; on pose $Z_b = \{R_t \mid t \in [b, 1]\}$. On a alors

$$d_p(R_s, R_t) = 2 \operatorname{Inf}(|s - t|, b),$$

quels que soient $s, t \in [b, 1]$. Donc Z_b est de diamètre $\leqslant 2b$ dans l'espace (Q_2, d_p) et $N(Z_b, d_p; b)$ est supérieur ou égal à la partie entière de $\dfrac{1-b}{b}$. Donc $\operatorname{Dim}(Q_2, d_p)$ est infini.

□

On peut cependant faire l'observation suivante :

LEMME 4.10. – *Soit* $\mathfrak{S}$ *une partie de* 2^X ; *on suppose que* $\mathfrak{S}$ *est réunion d'un nombre fini de classes totalement ordonnées. On a alors* $\operatorname{Dim}(\mathfrak{S}, d_p) \leqslant 1$ *pour toute loi* P *sur* $(X, \mathfrak{S})$.

Démonstration. – Rappelons que la dimension métrique d'une réunion finie est la borne supérieure des dimensions métriques (voir [4]). On peut donc supposer que $\mathfrak{S}$ est totalement ordonnée ; on a alors $d_p(S, S') = |P(S) - P(S')|$, quels que soient $S, S' \in \mathfrak{S}$; l'application $S \longrightarrow P(S)$ est donc un plongement isométrique (au sens de [5]) de l'espace $(\mathfrak{S}, d_p)$ dans $[0,1]$ muni de la distance euclidienne. Cela donne le résultat.

□

Dimension des espaces arguésiens.

4.11. – a) Un *espace arguésien* (Busemann [10]) est une partie convexe X de $\mathbf{R}^n$ munie d'un écart d vérifiant $d(x, y) = d(x, z) + d(z, y)$, quels que soient $x, y \in X$ et quel que soit le point z du segment $[x, y]$. En d'autres termes, dans un espace arguésien (X, d), chaque corde (c'est-à-dire l'intersection de chaque droite avec X) est une géodésique dans l'espace (X, d).

Donnons quelques *exemples d'espaces arguésiens* (en suivant [10]) :

4.12. – a) Soit X une partie convexe ouverte de $\mathbf{R}^n$; pour chaque paire (x, y) de points distincts de X, on note u et v les intersections (éventuellement à l'infini) de la droite xy avec ∂X, et on pose $d(x, y) = \left|\operatorname{Log}\left(\frac{\|x-u\|}{\|x-v\|} \times \frac{\|y-v\|}{\|y-u\|}\right)\right|$ (module du logarithme du module du birapport) ; alors d est un écart sur X (dit *écart de Hilbert sur* X) et l'espace (X, d) est un espace arguésien.

b) La boule euclidienne ouverte (de rayon 1) dans $\mathbf{R}^n$, munie de l'écart de Hilbert, est *l'espace hyperbolique de dimension* n (c'est une représentation due à Klein, où les cordes sont les géodésiques, voir [9] p. 259).

Un exemple particulièrement important est le suivant :

4.13. – a) Soient X une partie convexe de $\mathbf{R}^n$ et $\mathscr{A}_n^0$ la classe des demi-espaces affines fermés de $\mathbf{R}^n$ qui ne contiennent pas 0 ; notons $\mathscr{A}_X^0$ la classe $X \cap \mathscr{A}_n^0$ et $[X]$ la structure d'incidence $[X, \mathscr{A}_X^0]$ (suivant la définition 2.8).

b) Soient $x, x' \in X$; alors $[X]_x \,\Delta\, [X]_{x'}$ est l'ensemble des éléments S de $\mathscr{A}_X^0$ qui *séparent* x *et* x', c'est-à-dire tels qu'on ait $(x \in S, x' \notin S)$ ou $(x \notin S, x' \in S)$.

c) Si μ est une mesure $\geqslant 0$ sur $\mathscr{A}_X^0$, le noyau

$$d_\mu : x, x' \longrightarrow \mu([X]_x \,\Delta\, [X]_{x'})$$

est un écart sur X et (X, d_μ) est un espace arguésien (en effet, soit S un élément de $\mathscr{A}_X^0$ séparant deux points x et y de X ; alors chaque point z de X est séparé par S, soit de x, soit de y ; si de plus z appartient à $[x, y]$, alors S ne peut à la fois séparer z de x et z de y).

d) Soit (X, d) un espace arguésien ; s'il existe une mesure positive μ sur $\mathscr{A}_X^0$ telle que $d = d_\mu$, on dit que μ est une *mesure de Crofton pour* (X, d).

On a alors le résultat suivant (qu'on donne, pour simplifier, seulement dans le cas d'un écart continu) :

PROPOSITION 4.14. – (Alexander [1], Ambartzumian [3]. – *Soit* (X, d) *un espace arguésien, où* X *est une partie convexe de* $\mathbf{R}^2$, *et où* d *est un écart continu sur* X *(pour la topologie usuelle). Alors* (X, d) *possède une mesure de Crofton* μ. *De plus, pour chaque polygone convexe compact* K *inclus dans* X, *la quantité* $2\mu\{S \in \mathscr{A}_X^0 \mid$ K *coupe* S *et* $X \backslash S\}$ *est égale au périmètre de* K *dans* (X, d) *(le périmètre de* K *dans* (X, d) *est la somme des diamètres dans* (X, d) *des arêtes de* K).

4.15. – Par contre, soit (X, d) un espace arguésien, où X est une partie convexe de $\mathbf{R}^n$ et où d est un écart continu sur X. Alors (X, d) ne possède pas toujours de mesure de Crofton (par exemple si (X, d) est un espace normé non plongeable dans L^1). Pour un tel espace, les assertions suivantes sont équivalentes (Alexander [2]) :

a) l'espace (X, d) possède une mesure de Crofton,

b) l'espace (X, d) est un sous-espace métrique d'un espace L^1,

c) l'écart d est de type négatif (au sens de Schoenberg [31].

PROPOSITION 4.16. – *Soit* (X, d) *un espace arguésien, où* X *est un polyèdre convexe compact dans* $\mathbf{R}^n$, *et où* d *est un écart continu sur* X *(pour la topologie usuelle). On a alors* $\overline{\dim}(X, d) \leqslant n$, *dès qu'on est dans l'un des deux cas suivants* :

a) $n = 2$,

b) (X, d) *est un sous-espace métrique d'un espace* L^1.

Démonstration. – L'espace (X, d) possède une mesure de Crofton μ (voir 4.14 et 4.16) ; les ensembles ϕ et X ne séparant aucun couple de points de X, on peut même supposer que μ est portée par $\mathscr{A}_X^0 \backslash \{\phi, X\}$ (c'est ce que nous supposerons).

i) La masse totale de μ est finie :

– en effet, pour $n = 2$, la quantité $\mu(\mathscr{A}_X^0 \backslash \{\phi, X\})$ n'est autre que le demi périmètre de X dans (X, d) (voir 4.14) ; elle est donc finie ;

– dans le cas b), la quantité $\mu(\mathscr{A}_X^0 \backslash \{\phi, X\})$ est de toutes façons majorée par la somme des diamètres dans (X, d) des arêtes du polyèdre

X (car tout élément de $\mathscr{A}^0_X \setminus \{\phi , X\}$ sépare certainement deux sommets adjacents de X) ; elle est donc encore finie.

ii) Rappelons que [X] est une structure d'incidence sur $(X, \mathscr{A}^0_X)$;

– la classe $[X]^{-1}(\mathscr{A}^0_X)$ n'est autre que $\mathscr{A}^0_X$, c'est-à-dire $X \cap \mathscr{A}^0_n$;

– la classe duale $[X](X)$, que nous allons noter $\mathscr{C}$, est la classe de toutes les parties de $\mathscr{A}^0_X$ de la forme $[X]_x$ (pour tout $x \in X$).

iii) De 2.11, il résulte qu'on a

$$\mathrm{dens}(\mathscr{C}) = \mathrm{dens}^*(\mathscr{A}^0_X) \leqslant \mathrm{dens}^*(\mathscr{A}^0_n) ;$$

on a donc $\mathrm{dens}(\mathscr{C}) \leqslant n$ (voir 3.5). Or l'espace $(\mathscr{C}, d_\mu)$ est isométrique à (X, d) (car μ est une mesure de Crofton pour (X, d)). Comme μ est de masse totale finie, on a donc

$$\overline{\dim}(X, d) \leqslant \mathrm{dens}(\mathscr{C}) \leqslant n .$$

□

Cela suggère deux questions :

4.17. – a) Soit (X, d) un espace arguésien, où X est un polyèdre convexe compact dans $\mathbf{R}^n$, et où d est un écart continu sur X. On a montré qu'on a $\overline{\dim}(X, d) \leqslant n$, dès que (X, d) est plongeable dans un espace L^1 (cas a) et b) ci-dessus). Ce résultat subsiste-t-il même si (X, d) n'est pas plongeable dans L^1 (il subsiste en tout cas lorsque (X, d) est un espace normé) ?

b) Soit (X, d) un sous espace métrique d'un espace L^1, tel que $\overline{\dim}(X, d)$ soit finie. Existe-t-il une classe de Vapnik-Cervonenkis $\mathscr{S}$, incluse dans 2^Y, et une mesure μ positive et de masse totale finie sur $(Y, \mathscr{S})$ telles que l'espace $(\mathscr{S}, d_\mu)$ soit isométrique à l'espace (X, d) ?

Terminons ce paragraphe avec la démonstration de la Proposition 3.15 :

Démonstration de la Proposition 3.15. – Soit B_1 la boule fermée de centre 0 et de rayon 1 dans l'espace normé $E = (\mathbf{R}^n, \| \ \|)$; on note $\mathscr{B}$ la classe des éléments de $\mathscr{B}_E$ qui sont inclus dans B_1, et P la restriction de la mesure de Lebesgue à B_1.

Nous allons montrer que $\overline{\dim}(\mathcal{B}, d_p)$ est supérieure ou égale à $n + 1$ (par 4.3, cela suffit à établir la Proposition, puisque P est de masse totale finie)

a) Pour chaque $\epsilon \in]0,1[$ et chaque entier $i \in \left[1, \frac{1}{\epsilon}\right]$, on pose $r_i(\epsilon) = (i\epsilon)^{1/n}$ et $\alpha_i(\epsilon) = \frac{1}{i} r_i(\epsilon)$; on pose de plus $r_0(\epsilon) = 0$;

– on a alors

$$\alpha_i(\epsilon)\,(r_i(\epsilon))^{n-1} = \epsilon \quad \text{et} \quad [r_i(\epsilon) - r_{i-1}(\epsilon)]\,(r_i(\epsilon))^{n-1} \geqslant \frac{\epsilon}{n} ;$$

– on choisit une partie $\alpha_i(\epsilon)$-discernable $Y_i(\epsilon)$ dans la boule B_1 munie de la distance $x, x' \longrightarrow \|x - x'\|$; plus précisément, on la choisit maximale et contenant 0 ;

– on note $\mathcal{C}_i(\epsilon)$ l'ensemble des éléments de $\mathcal{B}$ de rayon $r_i(\epsilon)$ centrés en un point de $Y_i(\epsilon)$.

b) Il existe trois nombres $C_1, C_2, C_3 \in]0, +\infty[$ tels qu'on ait, pour chaque $\epsilon \in]0,1[$ et chaque entier $i \in \left[1, \frac{1}{\epsilon}\right]$, les propriétés suivantes :

b1) on a $d_p(B, B') \geqslant C_1\, \alpha_i(\epsilon)\,(r_i(\epsilon))^{n-1}$, pour tous $B, B' \in \mathcal{C}_i(\epsilon)$ $(B \neq B')$;

b2) on a $d_p(B, B') \geqslant C_2\, [r_i(\epsilon) - r_{i-1}(\epsilon)]\,(r_i(\epsilon))^{n-1}$, pour tout $B \in \mathcal{C}_i(\epsilon)$ et tout $B' \in \bigcup_{j=1}^{i-1} \mathcal{C}_j(\epsilon)$;

b3) on a $|\mathcal{C}_i(\epsilon)| \geqslant C_3(\alpha_i(\epsilon))^{-n}$;

(b1) et b2) nécessitent une démonstration, que nous ne développerons pas).

c) Posons $C = \operatorname{Inf}(C_1, \frac{1}{n} C_2)$. Pour chaque $\epsilon \in]0,1[$, on note k la partie entière de $\frac{1}{\epsilon}$; la classe $\mathcal{C}(\epsilon) = \bigcup_{i=1}^{k} \mathcal{C}_i(\epsilon)$ est $C\epsilon$-discernable dans $(\mathcal{B}, d_p)$ (c'est ce qu'on a montré en b)) ; de plus on obtient :

$$|\mathcal{C}(\epsilon)| \geqslant C_3 \sum_{i=1}^{k} (\alpha_i(\epsilon))^{-n} = C_3 \frac{1}{\epsilon} \sum_{i=1}^{k} i^{n-1} \geqslant C_3 \frac{1}{\epsilon} \frac{k^n}{n} .$$

Comme k est la partie entière de $\frac{1}{\epsilon}$, la Proposition est démontrée.

□

5. $\mathcal{S}$ - Indépendance et nombres de Radon.

Ces deux notions ont des liens avec les classes de Vapnik-Cervonenkis ; nous essayons d'abord d'indiquer leur utilité :

– la $\mathcal{S}$-indépendance intervient naturellement dans l'étude des martingales à plusieurs indices (voir Hajek [22]) et dans les matroïdes ;

– la théorie combinatoire des convexes et de leurs nombres de Radon recèle des problèmes difficiles, auquels le résultat de Jamison [23] que nous démontrons peut servir d'introduction.

$\mathcal{S}$-indépendance.

L'étude des martingales indiciées par une famille de parties d'un ensemble (notamment les martingales à deux indices) a conduit Hajek [22] à utiliser la notion d'indépendance suivante, déjà rencontrée en 3.23b pour les matroïdes :

5.1. – Soit $\mathcal{S}$ une partie de 2^X et soit A une partie finie de X ; on dit que A est *$\mathcal{S}$-indépendante* si $A \setminus \{x\}$ appartient à $A \cap \mathcal{S}$, quel que soit $x \in A$

Une partie $\mathcal{S}$-indépendante n'est pas autre chose qu'une partie pulvérisée par une classe convenable :

LEMME 5.2. – *Soit $\mathcal{S}$ une partie de 2^X. Soit $\mathcal{S}_d$ la classe stabilisée de $\mathcal{S}$ par intersection finie, c'est-à-dire la classe des parties de X de la forme $\bigcap_{i \in I} S_i$, où $(S_i)_{i \in I}$ est une famille finie d'éléments de $\mathcal{S}$. Soit A une partie finie de X. A est $\mathcal{S}$-indépendante si et seulement si A est pulvérisée par $\mathcal{S}_d$.*

Démonstration. – a) Supposons que A soit $\mathcal{S}$-indépendante. Pour chaque $x \in A$, il existe alors $S_x \in \mathcal{S}$ avec $A \cap S_x = A \setminus \{x\}$. Pour chaque $B \subset A$, on pose $S_B = \bigcap_{x \in B} S_x$. S_B appartient alors à $\mathcal{S}_d$ et on a $A \cap S_B = B$. Donc A est pulvérisée par $\mathcal{S}_d$.

b) Supposons que A soit pulvérisée par $\mathcal{S}_d$. Pour chaque $x \in A$, il existe alors une famille finie $(S_i)_{i \in I}$ d'éléments de $\mathcal{S}$ telle que $A \cap (\bigcap_{i \in I} S_i) = A \backslash \{x\}$. Il existe donc $i \in I$ tel que $A \cap S_i = A \backslash \{x\}$. A est donc $\mathcal{S}$-indépendante.

$\square$

L'indépendance dans un matroïde $M = (X, \mathcal{H})$ n'est pas autre chose que la $\mathcal{H}$-indépendance. Nous allons donc démontrer maintenant le Lemme 3.24 :

Démonstration du Lemme 3.24. – Soit $M = (X, \mathcal{H})$ un matroïde de dimension n. L'entier $n + 1$ est donc la borne supérieure des cardinaux des parties $\mathcal{H}$-indépendantes de X (voir 3.23c), donc la borne supérieure des cardinaux des parties de X pulvérisées par $\mathcal{H}_d$ (voir 5.2). Donc $\mathcal{H}_d$, qui n'est autre que la classe de toutes les variétés de M, est de densité entière $n + 1$.

$\square$

Convexes et nombres de Radon.

Soit $\mathcal{S}$ une partie de 2^X stable par intersection et contenant ϕ et X ; on dit alors que $(X, \mathcal{S})$ est un *espace de convexité.* On y étudie notamment les notions suivantes (voir Eckhoff [18]) :

5.3. – a) Si Y est une partie de X, on appelle *enveloppe convexe de* Y *dans* $(X, \mathcal{S})$ (et on note $\overline{Y}^{\mathcal{S}}$) l'intersection des éléments de $\mathcal{S}$ contenant Y.

b) Soient A une partie finie de X et k un entier $\geqslant 2$; une partition $(A_1, \ldots, A_k)$ de A est appelée une *k-partition de Radon de* A *dans* $(X, \mathcal{S})$ si l'ensemble $\bigcap_{i=1}^{k} \overline{A}_i^{\mathcal{S}}$ est non vide.

c) On appelle $k^{\text{ième}}$ *nombre de Radon de* $(X, \mathcal{S})$ (et on note $r_k(\mathcal{S})$) la borne inférieure des entiers n tels que toute partie de X de cardinal $\geqslant n$ admette au moins une k-partition de Radon dans $(X, \mathcal{S})$.

Nous indiquerons dans un instant les liens entre le $2^{\text{ième}}$ nombre de Radon et la densité entière. Voici pour le moment quelques propriétés des nombres de Radon (en suivant [18]) :

5.4. – a) On a $r_\ell(\mathcal{S}) \geqslant r_k(\mathcal{S})$, quel que soit $\ell \geqslant k$.

b) Si X est un ensemble infini, alors $r_2(2^S)$ est infini.

c) Soit Γ_n la classe de toutes les parties convexes fermées de $\mathbf{R}^n$ (déjà considérée en 3.1c, où on a vu qu'elle est de densité infinie) ; on a alors $r_2(\Gamma_n) = n + 2$ (c'est le théorème de Radon, voir ci-dessus 3.4) et, plus généralement,

$$r_k(\Gamma_n) - 1 = (k - 1)\,[r_2(\Gamma_n) - 1],$$

quel que soit l'entier $k \geqslant 2$ (théorème de Tverberg [38].

d) Une importante question [18], toujours ouverte, est de savoir si le théorème de Tverberg subsiste pour tout espace de convexité $(X, \mathcal{S})$, autrement dit si on a $r_k(\mathcal{S}) - 1 \leqslant (k - 1)\,[r_2(\mathcal{S}) - 1]$, quel que soit l'entier $k \geqslant 2$, pour chaque classe $\mathcal{S}$ stable par intersection.

Nous allons essayer d'aborder la *comparaison entre* $r_2(\mathcal{S}) - 1$ *et* $\mathrm{Dens}(\mathcal{S})$:

5.5. – a) On sait évaluer la densité entière d'une classe image réciproque (voir 2.2). Pour les nombres de Radon, on a seulement, en sens inverse :

– soient $(Y, \mathcal{S})$ un espace de convexité et f une application de X dans Y ; alors $(X, f^{-1}(\mathcal{S}))$ est un espace de convexité ;

– si de plus f est surjective, alors on a $r_k(f^{-1}(\mathcal{S})) \geqslant r_k(\mathcal{S})$, quel que soit l'entier $k \geqslant 2$

(en effet, pour $A \subset Y, |A| = r_k(f^{-1}(\mathcal{S}))$, on peut trouver $B \subset X$ avec $f(B) = A, |B| = |A|$; l'ensemble B admet alors une k-partition de Radon dans $(X, f^{-1}(\mathcal{S}))$ et donc A admet une k-partition de Radon dans $(Y, \mathcal{S})$).

b) On a toujours $r_2(\mathcal{S}) - 1 \leqslant \mathrm{Dens}(\mathcal{S})$ (car une partie A qui n'admet pas de 2-partition de Radon dans $(X, \mathcal{S})$ est nécessairement pulvérisée par $\mathcal{S}$) ; mais l'exemple 5.4c montre qu'il n'y a pas en général égalité.

5.6. Cependant on va voir que $r_2(\mathcal{S}) - 1$ est bien une densité entière s'il existe une classe $\mathcal{C}$ vérifiant les conditions suivantes :

a) $\mathcal{C}$ *sépare faiblement* $\mathcal{S}$ (à rapprocher de 2.16), ce qui signifie : pour chaque $S, S' \in \mathcal{S}\ (S \cap S' = \phi)$, chaque partie finie B de S et chaque partie finie B′ de S′, il existe $T, T' \in \mathcal{C}$ avec $B \subset T \backslash T'$ et $B' \subset T' \backslash T$;

b) $A \cap T$ et $A \setminus T$ appartiennent à $A \cap \mathcal{S}$, quel que soit $T \in \mathcal{C}$ et quelle que soit A partie finie de X.

PROPOSITION 5.7. – *Soit* $(X, \mathcal{S})$ *un espace de convexité et soit* $\mathcal{C}$ *une partie de* 2^X

a) *Si* $\mathcal{C}$ *vérifie la propriété* 5.6a, *on a alors* $r_2(\mathcal{S}) - 1 \leqslant \mathrm{Dens}(\mathcal{C})$

b) *Si* $\mathcal{C}$ *vérifie la propriété* 5.6b, *on a alors* $r_2(\mathcal{S}) - 1 \geqslant \mathrm{Dens}(\mathcal{C})$.

Démonstration. – a) Soit A une partie finie de X n'admettant pas de 2-partition de Radon dans $(X, \mathcal{S})$. Pour chaque partition (B, B') de A, il existe donc $S, S' \in \mathcal{S}$ avec

$$B \subset S, B' \subset S', S \cap S' = \phi ;$$

donc (par 5.6a), il existe $T, T' \in \mathcal{C}$ avec $B \subset T \setminus T'$ et $B' \subset T' \setminus T$; on a donc $B = A \cap T$ et $B' = A \cap T'$; donc A est pulvérisée par $\mathcal{C}$.

b) Inversement soit A une partie finie de X pulvérisée par $\mathcal{C}$. Pour chaque partition (B, B') de A, il existe donc $T, T' \in \mathcal{C}$ avec $B = A \cap T, B' = A \cap T'$; B et B' sont donc respectivement inclus dans $T \setminus T'$ et dans $T' \setminus T$; or $A \cap (T \setminus T')$ et $A \cap (T' \setminus T)$ appartiennent à $A \cap \mathcal{S}$ (par 5.6b) ; donc A n'admet pas de 2-partition de Radon dans $(X, \mathcal{S})$.

(On peut noter que, si $\mathcal{S} = \Gamma_n$, la classe $\mathcal{C} = \mathcal{A}_n$ vérifie 5.6a et 5.6b).

Bien que 5.4d reste une conjecture, on sait cependant majorer $r_k(\mathcal{S})$ à l'aide de $r_2(\mathcal{S})$. Le résultat suivant, dû à Jamison [23], fournit une telle majoration ([23] étant peu accessible, nous en donnons une démonstration) :

PROPOSITION 5.8 [23]. – *Soient* $(X, \mathcal{S})$ *un espace de convexité et* k *un entier* $\geqslant 2$. *Notons* p *le plus petit entier* $\geqslant \mathrm{Log}_2(k)$. *On a alors* $r_k(\mathcal{S}) \leqslant (r_2(\mathcal{S}))^p$.

Démonstration. – On pose $Y = X \times \mathbf{N}$ et $\mathcal{C} = f^{-1}(\mathcal{S})$, où f est la projection de Y sur X.

a) On a $r_{ij}(\mathcal{C}) \leqslant r_i(\mathcal{C})\, r_j(\mathcal{C})$, quels que soient i et j, entiers $\geqslant 2$:

en effet, soient i et j des entiers $\geqslant 2$; posons $p = r_i(\mathfrak{C})$ et $q = r_j(\mathfrak{C})$. Soit A une partie de Y de cardinal $\geqslant pq$. On peut donc trouver une partition $(A_1, \ldots, A_p)$ de A telle que $|A_m| \geqslant q$, quel que soit $m = 1, \ldots, p$. Pour chaque $m = 1, \ldots, p$, on peut donc trouver une j-partition de Radon $(A_{m,1}, \ldots, A_{m,j})$ de A_m dans $(Y, \mathfrak{C})$ et choisir un point y_m de $\bigcap_{n=1}^{j} \overline{A}_{m,n}^{\mathfrak{C}}$. On peut de plus supposer (et c'est pour cela qu'on a substitué $(Y, \mathfrak{C})$ à $(X, \mathfrak{S})$) que l'ensemble $B = \{y_1, \ldots, y_p\}$ est de cardinal p. On peut donc trouver une i-partition de Radon $(B_1, \ldots, B_i)$ de B dans $(Y, \mathfrak{C})$ et choisir un point y de $\bigcap_{\ell=1}^{i} \overline{B}_{\ell}^{\mathfrak{C}}$. Pour chaque $\ell = 1, \ldots, i$ et chaque $n = 1, \ldots, j$, on note alors $C_{\ell,n}$ la réunion des ensembles $A_{m,n}$ tels que y_m appartienne à B_ℓ.

La partition $(C_{\ell,n})_{\ell=1,\ldots,i,\, n=1,\ldots,j}$ est une ij-partition de Radon de A dans $(Y, \mathfrak{C})$: en effet, pour chaque $\ell = 1, \ldots, i$ et chaque $n = 1, \ldots, j$, l'ensemble $\overline{C}_{\ell,n}^{\mathfrak{C}}$ contient $\overline{A}_{m,n}^{\mathfrak{C}}$ dès que y_m appartient à B_ℓ, donc $\overline{C}_{\ell,n}^{\mathfrak{C}}$ contient B_ℓ, donc il contient $\overline{B}_{\ell}^{\mathfrak{C}}$ qui contient y.

b) On a $r_2(\mathfrak{C}) = r_2(\mathfrak{S})$:

en effet, on sait déjà qu'on a $r_2(\mathfrak{C}) \geqslant r_2(\mathfrak{S})$ (voir 5.5a) ; inversement, si A n'admet pas de 2-partition de Radon dans $(Y, \mathfrak{C})$, la restriction de f à A doit être une injection, et donc $B = f(A)$ n'admet pas de 2-partition de Radon dans $(X, \mathfrak{S})$.

c) Il résulte de ce qui précède qu'on a :

$$r_k(\mathfrak{S}) \leqslant r_k(\mathfrak{C}) \leqslant r_{2^p}(\mathfrak{C}) \leqslant (r_2(\mathfrak{C}))^p = (r_2(\mathfrak{S}))^p .$$

□

Références.

– Pour les nombres de Radon et le théorème de Radon, on doit lire les articles de synthèse de Danzer, Grünbaum, Klee [12] (1963) et de Eckhoff [18] (1978).

– On trouvera, après la liste des références utilisées dans le présent travail, une bibliographie complémentaire concernant les classes de Vapnik-Cervonenkis : je signale notamment l'article de

Glazek (sur les questions d'indépendance, et avec une importante bibliographie) qui se trouve dans le volume de Colloquium Mathematicum dédié à la mémoire d'Edward Marczewski.

BIBLIOGRAPHIE

[1] R. ALEXANDER, Planes for which the lines are the shortest path, *Illinois J. of Math.*, 22 (1978), 177-190.

[2] R. ALEXANDER, Metrics on $\mathbf{R}^n$ which possess a Crofton formula, *Notices AMS*, 26 (1979), A-278.

[3] R.V. AMBARTZUMIAN, A note on pseudometrics on the plane, *Z. Warscheinlich keitstheorie*, 37 (1976), 145-155.

[4] P. ASSOUAD, *Espaces métriques, plongements, facteurs*, Thèse, Université de Paris-Sud, (1977).

[5] P. ASSOUAD, Pseudodistances, facteurs et dimension métrique, Séminaire d'Analyse Harmonique 1979/1980 (Orsay) *Publications Mathématiques d'Orsay* n° 81-07, pp. 1-33.

[6] P. ASSOUAD, Sur les classes de Vapnik-Cervonenkis, *C. R. A. S.*, Paris, 292 (1981), 921-924.

[7] E.G. BAJMOCZY, I. BARANY : On a common generalization of Borsuk's and Radon's theorems, *Acta Math. Acad. Scient. Hungar.*, 34 (1979), 347-350.

[8] R.G. BLAND, M. LAS VERGNAS, Orientability of matroïds, *J. of Comb. Th.*, B. 24 (1978), 94-123.

[9] K. BORSUK, W. SZMIELEW, *Foundations of geometry* (english transl.), North Holland (1960).

[10] H. BUSEMANN, Desarguesian Spaces, *Proc. of Symp. in Pure Math.*, 28 (1976), 131-141.

[11] P. CARTIER. Arrangements d'hyperplans : un chapitre de géométrie combinatoire, *Seminaire Bourbaki* (1980), exposé n° 651.

[12] L. DANZER, B. GRUNBAUM, V. KLEE, Helly's theorem and its relatives, *Proc. of Symp. in Pure Math.*, 7 (1963), 101-180.

[13] P. DEMBOWSKI, *Finite geometries,* Springer, 1968.

[14] R.M. DUDLEY, Central limit theorems for empirical measures, *Ann. of Prob.,* 6 (1978), 899-929.

[15] R.M. DUDLEY, Balls in $\mathbf{R}^k$ do not cut all subsets of $(k + 2)$ points, *Advances Math.,* 31 (1979), 306-308.

[16] R.M. DUDLEY, Vapnik-Cervonenkis Donsker classes of functions, preprint (1980).

[17] M. DURST, R.M. DUDLEY, Empirical processes, Vapnik Cervonenkis classes and Poisson processes, *Probability and Math. Stat.,* 1 (1981) to appear.

[18] J. ECKHOFF, Radon's theorem revisited, In Contributions to geometry, *Proc. of the Geometry Symp. in Siegen* 1978, Birkhaüser (1979) 164-185.

[19] G. FICHTENHOLZ, K. KANTOROVITCH, Sur les opérations linéaires dans l'espace des fonctions bornées, *Studia Math.,* 5 (1934), 69-98.

[20] J.E. GOODMAN, R. POLLACK, Proof of Grünbaum's conjecture on the stretchability of certain arrangements of pseudolines, *J. Comb. Th,* A 29 (1980), 385-390.

[21] B. GRUNBAUM, Arrangements and spreads, *CBMS Regional Conference Series in Math.,* n° 10 (1972).

[22] B. HAJEK, *Stochastic integration, Markov property and measure transformation of random fields,* Ph. D. Dissertation, Berkeley (1979).

[23] R.E. JAMISON, *A general theory of convexity,* Doct. Dissert., Univ. of Washington (Seattle, 1974).

[24] T. KOVARI, V.T. SOS, P. TURAN, On a problem of K. Zarankiewicz, *Colloquium Math.,* 3 (1954), 50-57.

[25] G.G. LORENTZ, Lower bound for the degree of approximation, *Trans. Amer, Math. Soc.,* 97 (1960), 25-34.

[26] B. REZNICK, Banach spaces with polynomial norms, *Pac. J. Math.,* 82 (1979), 223-235.

[27] G. RINGEL, Teilungen der Ebene durch Geraden oder topologische Geraden, *Math. Z.,* 64 (1955), 79-102.

[28] C.A. Rogers, *Hausdorff measures,* Cambridge University Press (1970).

[29] H.P. Rosenthal, A characterization of Banach spaces containing ℓ^1, *Proc. Nat. Acad. Sci. USA,* 71 (1974), 2411-2413.

[30] N. Sauer, On the density of families of sets, *J. Comb. Th.,* A 13 (1972) 145-147.

[31] I.J. Schoenberg, Metric spaces and positive definite functions, *Transact. Amer. Math. Soc.,* 44 (1938), 522-536.

[32] H.S. Shapiro, Some negative theorems of approximation theory, *Michigan Math. J.,* 11 (1964), 211-217.

[33] S. Shelah, A combinatorial problem ; stability and order for models and theories in infinitary languages, *Pacific J. Math.,* 41 (1972) 247-261.

[34] G. Sierksma, *Generalized Radon partitions in convexity spaces,* preprint (1980).

[35] E.H. Spanier, *Algebraic topology,* Mc Graw Hill (1966).

[36] H.M. Stark, An introduction to number theory, *Markham Publ. Co.,* 1971.

[37] E. Szpilrajn (Marczewski), Ensembles indépendants et mesures non séparables, *C.R.A.S.,* Paris, 207 (1938), 768-770.

[38] H. Tverberg, A generalization of Radon's theorem, *J. of the London Math. Soc.,* 41 (1966), 123-128.

[39] V.N. Vapnik, A. YA. Cervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theor. Prob. Appl.,* 16 (1971), 264-280.

[40] H.E. Warren, Partitions by real algebraic varieties and applications to questions of nonlinear approximation, *Bull. AMS,* 73 (1967) 192-194.

[41] H.E. Warren, Lower bounds for approximation by nonlinear manifolds, *Trans. AMS,* 133 (1968), 167-178.

[42] D.J.A. Welsh, *Matroïd theory,* Academic Press (1976).

[43] R.S. Wenocur, R.M. Dudley, Some special Vapnik-Cervonenkis classes, *Discrete Math.,* 33 (1981), 313-318.

Bibliographie complémentaire.

B.J. BIRCH : On 3N points in a plane, *Proc. Cambridge Phil. Soc.*, 55 (1959), 289-293.

J.A. BONDY : Induced subsets, *J. Comb. Th.*, B 12 (1972), 201-202.

K. GLAZEK : Some old and new problems in the independence theory, *Colloquium Math.*, 42 (1979), 127-189.

M.G. KARPOVSKY, V.D. MILMAN : Coordinate density of sets of vectors, *Discrete Math.*, 24 (1978), 177-184.

P. TOMASTA : "Dart calculus" of induced subsets, *Discrete Math.*, 34 (1981), 195-198.

Manuscrit reçu le 8 décembre 1981
révisé le 26 novembre 1982.

Patrice ASSOUAD,
Université de Paris Sud
Centre d'Orsay
Mathématiques
bâtiment 425
91405 Orsay Cedex.