



# Journées mathématiques X-UPS

Année 1995

## Aspects géométriques et combinatoires de la convexité

Michèle AUDIN

**Matrices hermitiennes et convexité**

*Journées mathématiques X-UPS* (1995), p. 1-22.

<https://doi.org/10.5802/xups.1995-01>

© Les auteurs, 1995.



Cet article est mis à disposition selon les termes de la licence

LICENCE INTERNATIONALE D'ATTRIBUTION CREATIVE COMMONS BY 4.0.

<https://creativecommons.org/licenses/by/4.0/>

Les Éditions de l'École polytechnique  
Route de Saclay  
F-91128 PALAISEAU CEDEX  
<https://www.editions.polytechnique.fr>

Centre de mathématiques Laurent Schwartz  
CMLS, École polytechnique, CNRS,  
Institut polytechnique de Paris  
F-91128 PALAISEAU CEDEX  
<https://portail.polytechnique.edu/cmls/>



Publication membre du

Centre Mersenne pour l'édition scientifique ouverte

[www.centre-mersenne.org](http://www.centre-mersenne.org)

## MATRICES HERMITIENNES ET CONVEXITÉ

*par*

Michèle Audin

---

et la photo du grand-père dans son cadre ovale, et la cheminée avec, entre les deux pots à pieds coniques, la lampe dont le socle est un cube d'opaline et le poisson rouge dans son bocal sphérique, et la table ronde avec le dessous de plat métallique en forme de losange, le cendrier octogonal de faïence blanche et la petite boîte à cigarettes en bois sculpté affectant la forme d'un tonneau

G. P.

### Table des matières

|                                                                                       |    |
|---------------------------------------------------------------------------------------|----|
| 1. Autour du théorème de Toeplitz-Hausdorff.....                                      | 2  |
| 1.1. Le théorème de Toeplitz-Hausdorff.....                                           | 2  |
| 1.2. Des vecteurs unitaires aux matrices hermitiennes .                               | 3  |
| 1.3. Des polygones et des polytopes convexes.....                                     | 5  |
| 2. Démonstration des théorèmes de convexité.....                                      | 8  |
| 2.1. Stratégie de démonstration : comment démontrer<br>un théorème de convexité?..... | 8  |
| 2.2. Fonctions de Morse et opérations de groupes.....                                 | 12 |
| 2.3. Récapitulons.....                                                                | 17 |
| 3. Polygones et polytopes convexes.....                                               | 17 |
| 3.1. Le cas des matrices normales.....                                                | 18 |
| 3.2. D'autres théorèmes de convexité.....                                             | 20 |
| 3.3. Quelques compléments sur Toeplitz-Hausdorff.....                                 | 21 |
| Indications bibliographiques.....                                                     | 22 |
| Références.....                                                                       | 22 |

Dans ces notes, je vais expliquer comment on peut utiliser des techniques de topologie différentielle, la théorie de Morse, qui traite des points critiques de fonctions numériques, pour démontrer des théorèmes de convexité. Je focaliserai l'étude sur des théorèmes mettant en jeu des matrices hermitiennes, et notamment sur des généralisations du théorème de Toeplitz-Hausdorff en dimension finie.

## 1. Autour du théorème de Toeplitz-Hausdorff

**1.1. Le théorème de Toeplitz-Hausdorff.** Un problème classique, dit de Toeplitz-Hausdorff, est celui de la détermination de l'*image numérique* d'un opérateur  $A$  sur un espace de Hilbert  $\mathcal{H}$ . Il s'agit tout simplement de l'ensemble  $W(A)$  des valeurs prises par le produit  $\langle Av, v \rangle$  quand  $v$  parcourt la sphère unité de  $\mathcal{H}$ . Je me limiterai ici au cas des matrices, i.e., au cas d'un espace de dimension finie : je supposerai que  $\mathcal{H} = \mathbb{C}^n$  avec la forme hermitienne standard. Ainsi l'opérateur  $A$  est-il une matrice  $n \times n$  complexe.

*Exemples.* L'image numérique peut prendre des formes relativement variées. Je m'inspire ici des exemples donnés dans [7], considérant les matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix},$$

$$D = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

pour lesquelles on vérifie (plus ou moins facilement) que les images numériques sont celles représentées sur la figure 1.

Le point commun le plus évident entre ces différentes parties de  $\mathbb{C}$  est, bien entendu, leur convexité. C'est ce qu'affirme le théorème de Toeplitz-Hausdorff :

**Théorème 1.** *L'image numérique  $W(A)$  est convexe.*

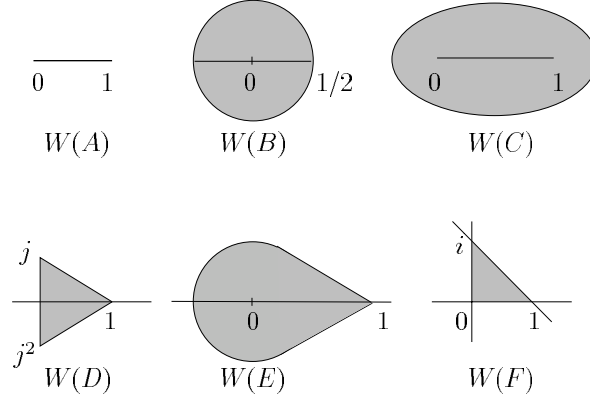


FIGURE 1. Images numériques d'opérateurs

**Remarque.** Ce théorème reste vrai dans un espace de Hilbert quelconque. Comme en dimension infinie la sphère unité n'est plus compacte, on peut trouver des convexes non fermés (voir [7] et les références citées dans cet ouvrage).

### 1.2. Des vecteurs unitaires aux matrices hermitiennes

*Des vecteurs unitaires à l'espace projectif.* L'application  $v \mapsto \langle Av, v \rangle$  dont nous contemplons l'image depuis le début de cet article est définie sur la sphère unité

$$S = \{v \in \mathbb{C}^n \mid \|v\| = 1\}$$

mais à vrai dire, la valeur de  $\langle Av, v \rangle$  ne dépend que de la droite complexe  $\ell$  que le vecteur  $v$  engendre : si  $w = \lambda v$ ,

$$\langle Aw, w \rangle = |\lambda|^2 \langle Av, v \rangle = \langle Av, v \rangle$$

quand  $v$  et  $w$  sont unitaires. En d'autres termes, notre application est en fait bien définie sur l'espace projectif

$$\mathbb{P}^{n-1}(\mathbb{C}) = \{\text{droites vectorielles } \ell \text{ de } \mathbb{C}^n\},$$

le quotient de la sphère  $S$  par la relation d'équivalence

$$v \sim w \Leftrightarrow \exists \lambda \in \mathbb{C}, |\lambda| = 1 \text{ tel que } w = \lambda v.$$

On a ainsi une application :

$$\begin{aligned} f_A : \mathbb{P}^{n-1}(\mathbb{C}) &\longrightarrow \mathbb{C} \\ \ell &\longmapsto \langle Av, v \rangle \end{aligned}$$

(où  $v$  est un vecteur unitaire de  $\ell$ ).

**Remarque.** Considérons par exemple le cas où  $n = 3$ , de sorte que  $\mathbb{P}^{n-1}(\mathbb{C})$  est le plan projectif, qu'on peut considérer comme l'espace des coordonnées homogènes  $[x, y, z]$  ou comme le plan affine  $\{[x, y, 1] \mid (x, y) \in \mathbb{C}^2\}$  complété par sa droite de l'infini  $\{[x, y, 0]\}$  comme sur la figure 2 où la droite oblique est celle de l'infini et le

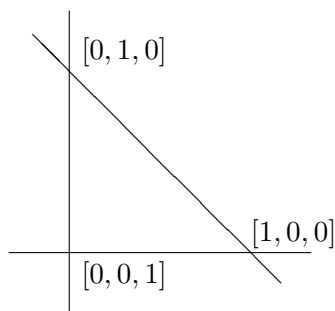


FIGURE 2. Le plan projectif

point  $[x, y, z]$  est celui dont les coordonnées barycentriques dans le triangle sont  $x$ ,  $y$  et  $z$ . Il y a beaucoup plus qu'une analogie entre cette figure et l'image numérique  $W(F)$  représentée sur la figure 1 : on pourra vérifier que les sommets du triangle sont envoyés sur les sommets et les droites projectives complexes schématisées par les trois droites de la figure 2 sur les côtés du triangle  $W(F)$ .

*De l'espace projectif aux matrices hermitiennes.* Étant donné une droite vectorielle  $\ell$  de  $\mathbb{C}^n$  et deux nombres réels  $\lambda$  et  $\mu$ , on construit un opérateur *hermitien*  $V : \mathbb{C}^n \rightarrow \mathbb{C}^n$  de la façon suivante : on décompose  $\mathbb{C}^n$  en  $\ell \oplus \ell^\perp$  et on définit  $V|_\ell = \lambda \text{Id}_\ell$ ,  $V|_{\ell^\perp} = \mu \text{Id}_{\ell^\perp}$ . Le spectre de  $V$  est  $(\lambda, \underbrace{\mu, \dots, \mu}_{n-1})$  par construction. Prenons par exemple  $\lambda = 1$  et

$\mu = 0$ . On vérifie alors simplement que, pour tout vecteur unitaire  $v$  de  $\ell$  et pour toute matrice complexe  $A$ ,

$$\langle Av, v \rangle = \text{tr}(AV)$$

de sorte qu'on peut être tenté de remplacer la sphère unité par l'ensemble des matrices hermitiennes de spectre  $(\lambda, \mu, \dots, \mu)$  *fixé* et, pourquoi pas, par l'ensemble *isospectral* plus général

$$\begin{aligned} \mathcal{H}_{(\lambda_1, \dots, \lambda_n)} &= \mathcal{H}_\Lambda \\ &= \{\text{matrices hermitiennes } V \text{ de valeurs propres } \lambda_1, \dots, \lambda_n\} \end{aligned}$$

pour tout  $(\lambda_1, \dots, \lambda_n) = \Lambda \in \mathbb{R}^n$ .

Le problème auquel on arrive ainsi est celui de la description de l'image de l'application

$$\begin{aligned} f_A : \mathcal{H}_\Lambda &\longrightarrow \mathbb{C} \\ V &\longmapsto \text{tr}(AV) \end{aligned}$$

qui est celui de Toeplitz-Hausdorff quand  $\Lambda = (1, 0, \dots, 0)$ . Il est remarquable que le théorème énoncé ci-dessus reste vrai dans cette généralité, comme l'a démontré Viktor Ginzburg dans [4] :

***Théorème 1 bis.*** *L'image de  $\mathcal{H}_\Lambda$  par  $f_A$  est une partie convexe de  $\mathbb{C}$ .*

**1.3. Des polygones et des polytopes convexes.** On aura remarqué, parmi les exemples ci-dessus, les cas de  $W(A)$ ,  $W(D)$  et  $W(F)$  qui sont un peu mieux que des convexes : ce sont des polygones convexes<sup>(1)</sup>. Les matrices qui les définissent ont une propriété spécifique : elles sont *normales* au sens où elles commutent avec leur adjointe. La démonstration du théorème suivant est en fait un exercice facile, que je laisse au lecteur :

***Théorème 2.*** *L'image numérique d'une matrice normale est l'enveloppe convexe de ses valeurs propres.*

***Remarque.*** En dimension infinie, c'est l'adhérence de l'image numérique qui a cette propriété.

---

<sup>(1)</sup>Le mot poly-gone (resp. -èdre, -tope) convexe désigne l'enveloppe convexe d'un nombre fini de points dans un espace de dimension 2 (resp. 3 ou quelconque).

La démonstration de Ginzburg du théorème 1 bis donne aussi, sans beaucoup de fatigue supplémentaire, la généralisation suivante du théorème 2, un peu moins triviale :

**Théorème 2 bis.** *Si  $A$  est une matrice normale dont les valeurs propres sont  $a_1, \dots, a_n$ , l'image de  $\mathcal{H}_A$  par  $f_A$  est l'enveloppe convexe des  $n!$  nombres complexes*

$$\sum_{j=1}^n a_j \lambda_{\sigma(j)},$$

où  $\sigma$  parcourt le groupe  $\mathfrak{S}_n$  des permutations de  $\{1, \dots, n\}$ .

*Exemples.* On a déjà remarqué que les  $W(A)$ ,  $W(D)$  et  $W(F)$  de la figure 1 entrent dans ce cadre. Même en restant dans les matrices  $3 \times 3$ , on a maintenant une panoplie d'exemples plus variés. Gardons la matrice

$$F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

mais changeons de spectre en considérant un  $\Lambda$  arbitraire. D'après le théorème 2 bis, l'image de  $f_F$  est formée de l'enveloppe convexe des points  $\lambda_k + i\lambda_l$  (pour  $k \neq l$ ). C'est un hexagone quand le spectre  $\Lambda$  est simple (figure 3), un triangle, voire un point, sinon.

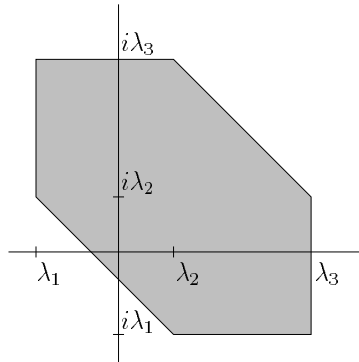


FIGURE 3.

Toujours dans cet exemple, regardons maintenant ce qu'est effectivement l'application  $V \mapsto \text{tr}(FV)$  : à une matrice hermitienne  $V = (v_{k,l})$  de spectre  $\Lambda$  elle associe  $v_{11} + iv_{22} \in \mathbb{C}$  ou  $(v_{11}, v_{22}) \in \mathbb{R}^2$ , ce qui est équivalent à donner les termes diagonaux de la matrice puisque la trace  $\sum v_{ii} = \sum \lambda_i$  est fixée sur  $\mathcal{H}_\Lambda$ . Le théorème 2 bis appliqué à cet exemple est donc une caractérisation de ceux des nombres  $(v_1, v_2, v_3) \in \mathbb{R}^3$  qui sont les termes diagonaux d'une matrice hermitienne  $3 \times 3$  de spectre  $(\lambda_1, \lambda_2, \lambda_3)$  fixé. Précisément ici, si on range les  $\lambda_i$  comme les  $v_i$  par ordre croissant,

$$\begin{cases} \lambda_1 \leq v_1 \leq \lambda_3 \\ \lambda_1 + \lambda_2 \leq v_1 + v_2 \leq \lambda_2 + \lambda_3 \\ v_1 + v_2 + v_3 = \lambda_1 + \lambda_2 + \lambda_3. \end{cases}$$

*Le théorème de Schur et Horn.* Il existe des inégalités analogues à celles que nous venons d'obtenir pour les diagonales des matrices hermitiennes de toute taille. Il s'agit d'un théorème de Schur [9] et Horn [8]. Il est bien plus simple d'écrire un théorème de convexité tel que le théorème 2 bis qu'une liste d'inégalités plus ou moins compréhensibles :

**Théorème 3.** *L'application qui, à une matrice hermitienne, associe sa diagonale envoie  $\mathcal{H}_{(\lambda_1, \dots, \lambda_n)}$  sur l'enveloppe convexe dans  $\mathbb{R}^n$  de l'ensemble des points  $(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(n)})$  (pour  $\sigma \in \mathfrak{S}_n$ ).*

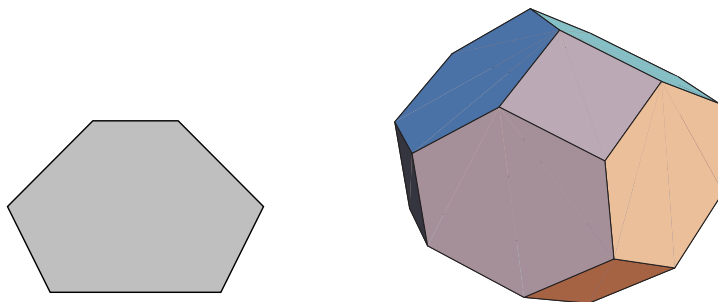


FIGURE 4. Diagonales des matrices  $n \times n$  hermitiennes, cas  $n = 3$  et  $n = 4$



La figure 4 représente cette image dans l'hyperplan  $\sum v_i = \sum \lambda_i$  de  $\mathbb{R}^n$  pour  $n = 3$  (hexagone) et  $n = 4$  (permutaèdre)<sup>(2)</sup> dans le cas où les  $\lambda_i$  sont distincts.

## 2. Démonstration des théorèmes de convexité

Les théorèmes que j'ai décrits ci-dessus appartiennent à une famille assez générale de « théorèmes de convexité en géométrie symplectique » — quel que soit le sens que ça puisse avoir. Le prototype en est un théorème publié simultanément en 1982 par Atiyah [2] et Guillemin et Sternberg [6]. Si le théorème de Schur-Horn (notre théorème 3) en est une application directe, celui de Ginzburg (nos théorèmes 1 bis et 2 bis) est plutôt une conséquence des idées de la démonstration d'Atiyah. Je vais expliquer maintenant ces idées et démontrer le théorème 1 bis (et donc en particulier le théorème 1).

### 2.1. Stratégie de démonstration : comment démontrer un théorème de convexité ?

*De la connexité à la convexité.* Imaginons qu'on veuille montrer que l'image d'une certaine application continue  $f$  d'un espace topologique  $W$  dans  $\mathbb{R}^m$  est convexe. Par définition un convexe est une partie de  $\mathbb{R}^m$  dont l'intersection avec toutes les droites est vide ou connexe. En conséquence, pour montrer que  $f(W)$  est convexe, il suffit de montrer que pour toute application linéaire  $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}^{m-1}$  et pour tout  $u \in \mathbb{R}^{m-1}$ ,  $(\varphi \circ f)^{-1}(u)$  est vide ou connexe : en effet

$$f(W) \cap \varphi^{-1}(u) = f((\varphi \circ f)^{-1}(u))$$

sera alors vide ou connexe.

Dans le cas des théorèmes de type 1 ou 1 bis ci-dessus, l'espace  $\mathbb{R}^m$  est  $\mathbb{C}$ , autrement dit  $m = 2$  et  $m - 1 = 1$  et les  $\varphi \circ f$  sont des fonctions numériques. On veut donc montrer que les niveaux<sup>(3)</sup> de certaines applications sont vides ou connexes.

---

<sup>(2)</sup>Je remercie J. Béthery, professeur au lycée Albert Schweizer au Raincy, de m'avoir fourni cette figure.

<sup>(3)</sup>Les *niveaux* de  $g$  sont les  $g^{-1}(u)$ .

*Comment démontrer que tous les niveaux d'une fonction sont connexes ?* Commençons par considérer les deux exemples représentés sur la figure 5.

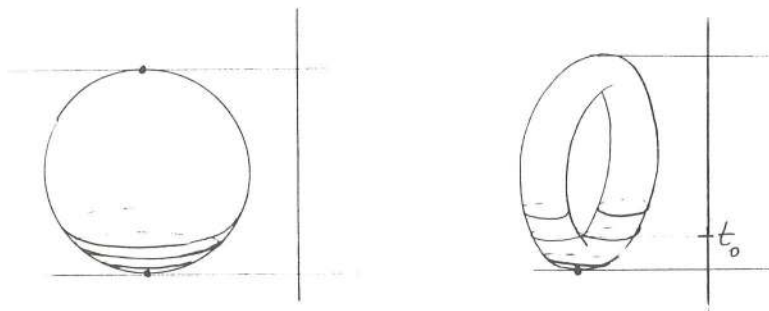


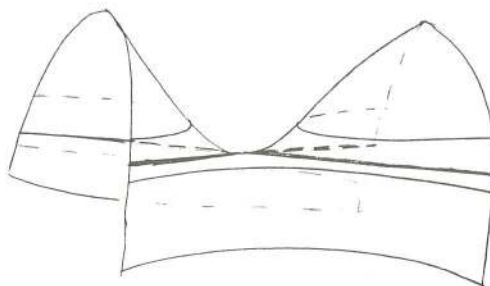
FIGURE 5.

Dans les deux cas, la fonction considérée est l'altitude, la hauteur. Il y a un minimum à partir duquel les niveaux sont non vides et connexes. Dans le cas de la sphère (à gauche) tous les niveaux non vides sont effectivement connexes. Dans le cas du tore (à droite) ce n'est vrai que jusqu'au niveau  $t_0$  au dessus duquel les niveaux ont deux composantes connexes.

En traversant la valeur minimale, le type topologique du niveau change (il était vide, il devient un cercle). Il se passe quelque chose d'analogue en chaque valeur critique. La valeur  $t_0$  correspond à un *col* (figure 6).

L'étude de ce type de phénomènes est l'objet de la *théorie de Morse*, outil précieux utilisé dans des contextes variés : on considère une *fonction de Morse*, une application dont tous les points critiques sont non dégénérés (voir la définition juste en dessous) et on a des modèles locaux très similaires au cas simple du col représenté sur la figure 6 qui permettent de décrire comment le type topologique des niveaux change quand on traverse une valeur critique.

En un point critique, par définition, la différentielle de la fonction est nulle.

FIGURE 6. le col  $z = t_0 + x^2 - y^2$ 

**Définition.** On dit que le point critique  $c$  est *non-dégénéré* si la dérivée seconde est une forme quadratique non-dégénérée. Une *fonction de Morse* est une fonction dont tous les points critiques sont non-dégénérés.

Il est équivalent<sup>(4)</sup> de dire que le point critique  $c$  est non-dégénéré ou qu'il existe des coordonnées locales  $(x_1, \dots, x_m)$  sur  $W$  dans lesquelles la fonction  $f$  s'écrit

$$(1) \quad f(x_1, \dots, x_m) = f(c) + Q(x_1, \dots, x_m)$$

où  $Q$  est une forme quadratique non-dégénérée.

Du coup on a des modèles locaux pour les changements de topologie. Par exemple, le cas d'un minimum local est celui d'une forme  $Q$  de signature  $(m, 0)$  (le bas des dessins de la figure 5 représente la cas  $(2, 0)$  et le haut le cas  $(0, 2)$ ). La figure 6 représente le cas d'une forme de signature  $(1, 1)$  et la figure 7 celui d'une forme de signature  $(2, 1)$ .

Plus généralement, le changement local de type topologique est décrit très explicitement par la transformation de la quadrique  $Q(x_1, \dots, x_m) = t$  pour  $t \in [-\varepsilon, \varepsilon]$ . De cette description explicite, on déduit que, pour être certain que tous les niveaux d'une fonction de Morse sont connexes<sup>(5)</sup>, il suffit de s'assurer que celle-ci n'a aucun

---

<sup>(4)</sup>C'est ce qu'affirme le théorème qui fonde la théorie de Morse, le célèbre *lemme de Morse* (voir [5] par exemple).

<sup>(5)</sup> $W$  est bien entendu supposée connexe ici.

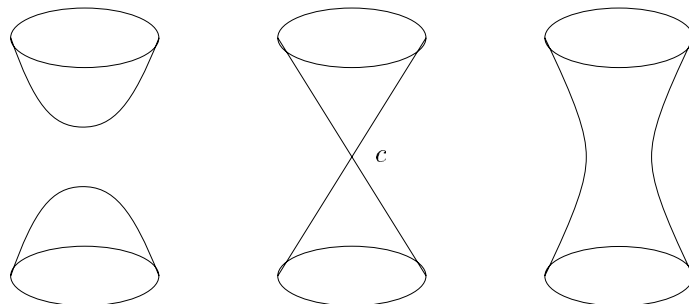


FIGURE 7.

point critique en lequel la dérivée seconde a pour signature  $(m-1, 1)$  ou  $(1, m-1)$ .

**Proposition.** *Soit  $f : W \rightarrow \mathbb{R}$  une fonction de Morse n'ayant aucun point critique en lequel la signature est  $(m-1, 1)$  ou  $(1, m-1)$ . Si  $W$  est connexe, alors tous les niveaux de  $f$  sont vides ou connexes.*

**Remarque.** Au début de ce §, je parlais d'un espace topologique  $W$ . Il va sans dire que les techniques de théorie de Morse s'appliquent à une fonction  $W \rightarrow \mathbb{R}$  qui doit être au moins de classe  $\mathcal{C}^2$ ... et donc que l'espace  $W$  doit être assez « lisse » pour que ça ait un sens. La bonne « catégorie » d'espaces est celle des *variétés* différentielles. Ceux que nous étudions ici sont les  $\mathcal{H}_\Lambda$ , qui sont des *sous-variétés* de l'espace vectoriel des matrices  $n \times n$  : ce n'est pas plus compliqué que la sphère ou le tore de l'espace de dimension 3 qu'on a pu contempler sur la figure 5.

*Points critiques d'une application et points fixes de l'action d'un groupe.* Regardons d'un peu plus près l'exemple de la sphère sur la figure 5 : les points critiques de la fonction considérée ne sont autres que les points fixes du groupe des rotations autour de l'axe vertical, dont nos niveaux sont d'ailleurs les orbites.

Il se trouve que toutes les fonctions dont nous souhaitons montrer que les niveaux sont connexes sont aussi reliées à des actions de

groupes. Je vais expliquer maintenant pourquoi ceci est vrai, et aussi pourquoi cette propriété donne la connexité des niveaux.

**2.2. Fonctions de Morse et opérations de groupes.** Le cadre général des théorèmes de convexité en géométrie symplectique dont je parle ici<sup>(6)</sup> est celui d'une variété symplectique  $W$  — ce sera notre  $\mathcal{H}_\Lambda$  — munie de l'opération d'un tore — un sous-groupe compact, connexe et abélien d'un groupe de matrices<sup>(7)</sup> — et d'une fonction liée à cette opération. Je préfère expliciter tous ces objets dans la famille d'exemples à laquelle je m'intéresse ici plutôt que de donner une liste abstraite de définitions.

*Le sous-groupe  $\mathcal{T}_X$  de  $U(n)$ .* Soit  $X$  une matrice hermitienne fixée. La matrice  $iX$  est anti-hermitienne, de sorte que  $\exp(itX)$  est, pour tout  $t \in \mathbb{R}$ , une matrice unitaire. Quand  $t$  décrit  $\mathbb{R}$ , l'ensemble de ces matrices décrit un sous-groupe commutatif du groupe unitaire  $U(n)$ . L'adhérence de ce sous-groupe (figure 8) est un sous-groupe compact, connexe et abélien de  $U(n)$ , appelons le  $\mathcal{T}_X$ .

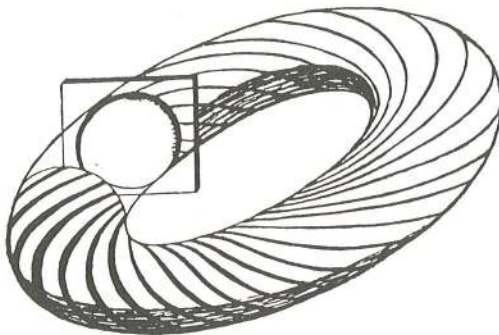


FIGURE 8. Le sous-groupe  $\{\exp(itX)\}$  et son adhérence

<sup>(6)</sup>On peut faire plus général (voir le § 3).

<sup>(7)</sup>Un tel sous-groupe est effectivement un *tore*, i.e., isomorphe à  $\mathbb{R}^r/\mathbb{Z}^r$  ou à  $\{(t_1, \dots, t_r) \in \mathbb{C}^r \mid |t_i| = 1\}$ .

*Une opération de  $\mathcal{T}_X$  sur  $\mathcal{H}_\Lambda$ .* Revenons à  $\mathcal{H}_\Lambda$ . Il est temps de remarquer que c'est une classe de conjugaison sous le groupe unitaire (c'est dire que toute matrice hermitienne est diagonalisable dans une base unitaire). Donc  $U(n)$  et tous ses sous-groupes opèrent sur  $\mathcal{H}_\Lambda$  par conjugaison. En particulier  $\mathcal{T}_X$ .

**Remarque.** De ce que  $\mathcal{H}_\Lambda$  est une orbite de  $U(n)$ , on déduit en particulier que c'est une sous-variété compacte et connexe de l'espace de toutes les matrices  $n \times n$ . Elle s'identifie au quotient de  $U(n)$  par le stabilisateur de la matrice diagonale  $(\lambda_1, \dots, \lambda_n)$ . Plus précisément, si les  $\lambda_i$  sont ordonnés en

$$\begin{aligned} \lambda_1 = \dots = \lambda_{n_1} &< \lambda_{n_1+1} \\ &= \dots = \lambda_{n_1+n_2} < \dots < \lambda_{n_1+\dots+n_{k-1}+1} = \dots = \lambda_{n_1+\dots+n_k} \end{aligned}$$

(avec  $n_1 + \dots + n_k = n$ ),  $\mathcal{H}_\Lambda$  s'identifie à  $U(n)/U(n_1) \times \dots \times U(n_k)$ . Son type topologique ne dépend que des multiplicités  $(n_1, \dots, n_k)$  et pas des valeurs des  $\lambda_i$ , c'est ce qu'on appelle une *variété de drapeaux*, généralisation de l'espace projectif.

*Une fonction sur  $\mathcal{H}_\Lambda$ .* Comme  $X$  est hermitienne,  $\text{tr}(XV)$  est réelle pour toute matrice hermitienne  $V$ , de sorte qu'on a une fonction à valeurs réelles

$$\begin{aligned} f_X : \mathcal{H}_\Lambda &\longrightarrow \mathbb{R} \\ V &\longmapsto \text{tr}(XV). \end{aligned}$$

Il se trouve que l'opération du groupe  $\mathcal{T}_X$  sur  $\mathcal{H}_\Lambda$  et la fonction  $f_X$  ont des liens très étroits. C'est là que la géométrie symplectique entre en jeu.

*Géométrie symplectique sur  $\mathcal{H}_\Lambda$ .* Sur  $\mathcal{H}_\Lambda$ , il y a une *forme symplectique*. C'est tout simplement, sur chaque espace tangent  $T_V \mathcal{H}_\Lambda$ , une forme bilinéaire alternée *non-dégénérée*  $\omega_V$ . Ce type de machine permet d'identifier vecteurs tangents et formes linéaires sur l'espace tangent. Ici on a justement, pour tout  $V \in \mathcal{H}_\Lambda$

- une forme linéaire  $T_V \mathcal{H}_\Lambda \rightarrow \mathbb{R}$ , la différentielle de  $f_X$  en  $V$ ,

- un vecteur de  $T_V\mathcal{H}_\Lambda$ , le vecteur tangent  $\underline{X}_V$  en  $t = 0$  de la courbe

$$t \longmapsto e^{itX} V e^{-itX}.$$

Le lien étroit entre l'opération de  $\mathcal{T}_X$  et la fonction  $f_X$ , c'est que, pour toute matrice  $V$  dans  $\mathcal{H}_\Lambda$ , ces deux objets se correspondent via la forme symplectique. Pour expliquer ça, il faut décrire la forme symplectique  $\omega_V$  — et tout d'abord l'espace tangent  $T_V\mathcal{H}_\Lambda$ .

*L'espace tangent à  $\mathcal{H}_\Lambda$  en  $V$ .* Comme je l'ai déjà dit,  $\mathcal{H}_\Lambda$  est une orbite de  $U(n)$ . Autrement dit,  $U(n)$  opère sur  $\mathcal{H}_\Lambda$  par conjugaison et l'opération est transitive. Ainsi  $\mathcal{H}_\Lambda$  s'identifie à  $U(n)/G_\Lambda$ , où  $G_\Lambda$  est le stabilisateur d'une matrice diagonale de l'orbite, et on peut décrire  $T_V\mathcal{H}_\Lambda$  comme quotient de  $T_{\text{Id}}U(n)$ . Comme  $U(n)$  est décrit par l'équation  ${}^t\overline{A}A = \text{Id}$ , son espace tangent en  $\text{Id}$  est l'espace  $\mathfrak{u}(n)$  des matrices antihermitiennes<sup>(8)</sup>.

La dérivée en  $\text{Id}$  de l'application d'orbite

$$\begin{aligned} U(n) &\longrightarrow \mathcal{H}_\Lambda \\ g &\longmapsto gVg^{-1} \end{aligned}$$

est une application  $\mathfrak{u}(n) \rightarrow T_V\mathcal{H}_\Lambda$  qui peut s'écrire

$$\begin{aligned} \mathfrak{u}(n) &\longrightarrow T_V\mathcal{H}_\Lambda \subset \{\text{matrices } n \times n\} \\ Y &\longmapsto [V, Y]. \end{aligned}$$

Ainsi  $T_V\mathcal{H}_\Lambda$  est-il le quotient de  $\mathfrak{u}(n)$  par le sous-espace des matrices  $Y$  telles que  $[V, Y] = 0$ .

*La forme symplectique.* Pour  $Y, Z \in \mathfrak{u}(n)$ , posons

$$\omega_V(Y, Z) = \text{tr}(iV[Y, Z]).$$

On vérifie que ça définit bien une forme bilinéaire alternée *non-dégénérée* sur  $T_V\mathcal{H}_\Lambda$ .

Maintenant, le vecteur tangent  $\underline{X}_V$  à la courbe  $e^{itX} V e^{-itX}$  est précisément l'image de  $iX \in \mathfrak{u}(n)$  dans  $T_V\mathcal{H}_\Lambda$ . La forme linéaire

---

<sup>(8)</sup>C'est l'algèbre de Lie du groupe  $U(n)$ .

$\omega_V(iX, \cdot)$  sur  $T_V\mathcal{H}_\Lambda$  qui lui est associée par la forme symplectique est

$$Y \longmapsto \operatorname{tr}(iV[iX, Y]) = -\operatorname{tr}(V[X, Y]).$$

Quant à la différentielle de  $f_X$  en  $V$ , comme  $f_X$  est linéaire, ce n'est pas très difficile, on écrit un vecteur tangent  $[V, Y]$  et on obtient

$$d_V f_X(Y) = \operatorname{tr}(X[V, Y]) = -\operatorname{tr}(V[X, Y]),$$

ce qu'on voulait :

**Proposition.** *On a*

$$d_V f_X(\underline{Y}) = \omega_V(iX, Y),$$

en d'autres termes la différentielle de  $f_X$  et le vecteur tangent  $\underline{X}$  défini par  $X$  se correspondent via la dualité définie par la forme symplectique.

*Exemple.* Choisissons  $n = 2$  et  $\Lambda = (1, -1)$ , ainsi  $\mathcal{H}_\Lambda$  est l'ensemble des matrices hermitiennes dont la trace est nulle et le déterminant vaut  $-1$ , i.e.,

$$\mathcal{H}_\Lambda = \left\{ V = \begin{pmatrix} a & \beta \\ \bar{\beta} & -a \end{pmatrix} \text{ telles que } a \in \mathbb{R}, \beta \in \mathbb{C} \text{ et } a^2 + |\beta|^2 = 1 \right\}.$$

Il s'agit de la sphère unité<sup>(9)</sup> de  $\mathbb{R}^3 = \mathbb{C} \times \mathbb{R}$ . Prenons pour  $X$  la matrice  $A$  du § 1.1 de sorte que  $\operatorname{tr}(XV) = a$  : la fonction  $f_X$  est la hauteur sur la sphère comme sur la figure 5. De plus,

$$\exp(itX) = \begin{pmatrix} e^{it} & 0 \\ 0 & 1 \end{pmatrix},$$

le groupe  $\mathcal{T}_X$  est un cercle et

$$e^{itX} \begin{pmatrix} a & \beta \\ \bar{\beta} & -a \end{pmatrix} e^{-itX} = \begin{pmatrix} a & e^{it}\beta \\ e^{-it}\bar{\beta} & -a \end{pmatrix},$$

ainsi, il opère sur la sphère unité par rotations autour de l'axe des  $a$ . C'est exactement la situation à laquelle j'ai fait allusion à la fin du § 2.1.

---

<sup>(9)</sup>Et bien sûr aussi de la droite projective  $\mathbb{P}^1(\mathbb{C})$ .



*Points critiques et points fixes.* Comme la forme symplectique est non-dégénérée, la différentielle  $d_V f_X$  est nulle si et seulement si le vecteur tangent correspondant est nul, i.e., exactement quand  $V$  est un *point fixe* de l'opération du groupe  $\mathcal{T}_X$ .

Or, il est facile d'étudier l'opération d'un tore au voisinage d'un point fixe  $c$  : on peut *linéariser* l'opération. Précisément, au voisinage de  $c$ , il existe des coordonnées locales  $(x_1, \dots, x_p, y_1, \dots, y_p)$  dans lesquelles l'action du tore est

$$g \cdot (z_1, \dots, z_p) = (g^{a_1} z_1, \dots, g^{a_p} z_p),$$

écriture dans laquelle

- $g$  désigne un élément du tore, je le considère comme un  $r$ -uplet  $(g_1, \dots, g_r)$  de nombres complexes de module 1
- j'ai regroupé les coordonnées  $x_j + iy_j = z_j$ , il ne faut pas s'étonner qu'elles soient en nombre pair ici : les formes bilinéaires alternées sont de rang pair et donc la dimension d'une variété symplectique est paire
- les  $a_i$  sont des multi-exposants,  $a_i \in \mathbb{Z}^r$ .

Mais ce n'est pas tout : dans les coordonnées en question, la forme symplectique s'écrit de façon assez standard. Je ne veux pas m'étendre sur cette écriture mais seulement sur la conséquence qui m'importe ici : la fonction  $f_X$  s'écrit, dans ces coordonnées

$$(2) \quad f_X(z_1, \dots, z_p) = f_X(c) + \sum_{i=1}^p \alpha_i |z_i|^2$$

pour certains réels  $\alpha_1, \dots, \alpha_p$  fonctions des  $a_i$  et de la position de  $\exp(itX)$  dans  $\mathcal{T}_X$ .

C'est une expression du même type que celle que nous a donnée le lemme de Morse (1)... à cela près que  $|z_i|^2 = x_i^2 + y_i^2$  et donc que la signature de notre forme quadratique est ici de la forme  $(2\rho, 2\sigma)$ ... *en particulier jamais*  $(1, m-1)$  *ou*  $(m-1, 1)$ , d'où l'on déduit :

**Proposition.** *Tous les niveaux de  $f_X$  sont vides ou connexes.*

**Remarque.** J'avoue que j'ai un peu triché : les fonctions  $f_X$  ne sont pas tout à fait des fonctions de Morse, autrement dit certains des  $\alpha_i$  qui figurent dans le deuxième membre de (2) peuvent très bien être nuls. Les points critiques de  $f_X$  ne sont peut-être pas isolés, mais on

a une propriété de non-dégénérescence transversalement à ces points critiques qui est suffisante pour démontrer ce que nous voulons.

**2.3. Récapitulons.** Reprenons et démontrons le théorème 1 bis. On décompose la matrice  $A$  en  $A = X + iY$  où  $X$  et  $Y$  sont hermitiennes. Alors  $\text{tr}(AV) = \text{tr}(XV) + i \text{tr}(YV)$  de sorte que  $f_A = f_X + if_Y$ . On peut appliquer le traitement ci-dessus à chacune des coordonnées  $f_X$  et  $f_Y$ , mais aussi à  $f_{\alpha X + \beta Y}$  pour  $\alpha, \beta \in \mathbb{R}$ .

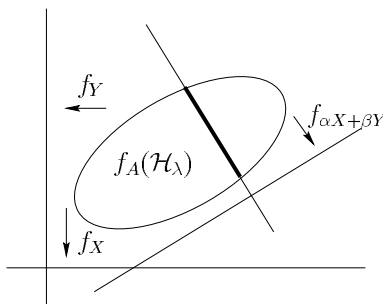


FIGURE 9. Convexité de  $f_A(\mathcal{H}_A)$

On en déduit que tous les niveaux de toutes les  $f_{\alpha X + \beta Y}$  sont vides ou connexes. Remarquons qu'on est exactement dans le cadre décrit au début du § 2.1 :

$$f_{\alpha X + \beta Y} = \varphi \circ f_A$$

où  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$  est l'application linéaire  $(x, y) \mapsto \alpha X + \beta Y$  (voir la figure 9). Ainsi, pour toute application linéaire  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ , et pour tout  $u \in \mathbb{R}$ ,  $(\varphi \circ f_A)^{-1}(u)$  est vide ou connexe. Donc  $f_A(\mathcal{H}_A)$  est convexe.

### 3. Polygones et polytopes convexes

En guise de conclusion, je vais donner une idée de démonstration des théorèmes 2 bis (et donc 2) et 3. La grosse différence est que l'image est maintenant un *polytope* convexe.

**3.1. Le cas des matrices normales.** Rappelons que  $A$  est *normale* si elle commute avec son adjointe. Il est équivalent de dire, en décomposant  $A$  en  $X+iY$  avec  $X$  et  $Y$  hermitiennes, que  $X$  et  $Y$  commutent. C'est justement ce qu'il faut pour que  $\exp(itX)$  et  $\exp(itY)$  ensemble engendrent un sous-groupe abélien de  $U(n)$ . Appelons  $\mathcal{T}$  l'adhérence de ce sous-groupe. Son opération sur  $\mathcal{H}_\Lambda$  se linéarise près des points fixes comme on l'a vu et chacune des fonctions  $f_X$ ,  $f_Y$  se met sous la forme (2) dans ces mêmes coordonnées.

De cette façon, les deux composantes de la fonction  $f_A$  sont données au voisinage d'un point fixe  $C$  par

$$(z_1, \dots, z_p) \mapsto \left( f_X(C) + \sum_{i=1}^p \alpha_i |z_i|^2, f_Y(C) + \sum_{i=1}^p \beta_i |z_i|^2 \right)$$

pour certains réels  $\alpha_i, \beta_i$ .

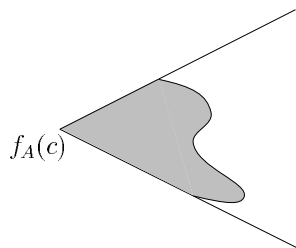


FIGURE 10.

Un voisinage de  $C$  a donc pour image un cône convexe. C'est une propriété de convexité locale. Ajoutée à la convexité (globale) que nous avons déjà démontrée, elle donne que l'image de  $f_A$  est l'enveloppe convexe des images  $f_A(C)$  des points fixes de l'opération du tore  $\mathcal{T}$ .

Pour finir la démonstration du théorème 2 bis, il reste à identifier ces points fixes. Il s'agit des matrices  $C$  telles que  $\exp(itA)$  commute avec  $C$  pour tout  $t$ , i.e., telles que  $A$  commute avec  $C$ . Remarquons que, comme  $A$  est normale, ses parties hermitienne ( $X$ ) et antihermitienne ( $iY$ ) commutent et donc qu'on peut les diagonaliser dans la même base unitaire. En d'autres termes, il existe une base unitaire

dans laquelle  $A$  est diagonale. Quitte à remplacer  $A$  par  $gAg^{-1}$  (et  $C$  par  $gVg^{-1}$ ), on est réduit à chercher les  $C$  de  $\mathcal{H}_\Lambda$  qui commutent avec la matrice *diagonale*

$$A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}.$$

Il existe un changement de base  $P$  tel que  $PCP^{-1}$  et  $PAP^{-1}$  sont diagonales<sup>(10)</sup>, en particulier

$$PCP^{-1} = \begin{pmatrix} \lambda_{\sigma(1)} & & \\ & \ddots & \\ & & \lambda_{\sigma(n)} \end{pmatrix} \quad \text{avec } \sigma \in \mathfrak{S}_n.$$

Comme  $f_A(C) = \text{tr}(AC) = \text{tr}((PAP^{-1})(PCP^{-1})) = \sum a_i \lambda_{\sigma(i)}$ , on en déduit le théorème 2 bis.

**Exemple.** Considérons à nouveau la matrice  $F$  du § 1.1 et l'application  $f_F$  sur  $\mathcal{H}_\Lambda$  pour  $\Lambda = (1, 0, 0)$ .

Ainsi  $\mathcal{H}_\Lambda$  est le plan projectif  $\mathbb{P}^2(\mathbb{C})$  : donner une matrice  $V$  de  $\mathcal{H}_\Lambda$  est équivalent à donner une droite  $\ell \subset \mathbb{C}^3$ , le sous-espace propre de  $V$  pour la valeur propre 1, l'autre espace propre, où  $V$  opère trivialement, est  $\ell^\perp$ .

Bien sûr,  $F$  est normale, le sous-groupe  $\mathcal{T}$  est engendré par

$$e^{itX} = \begin{pmatrix} e^{it} & & \\ & 1 & \\ & & 1 \end{pmatrix} \quad \text{et} \quad e^{itY} = \begin{pmatrix} 1 & & \\ & e^{it} & \\ & & 1 \end{pmatrix}$$

de sorte que

$$\mathcal{T} = \left\{ \begin{pmatrix} e^{it} & & \\ & e^{iu} & \\ & & 1 \end{pmatrix} \text{ telles que } t, u \in \mathbb{R} \right\}.$$

Il opère sur  $\mathcal{H}_\Lambda$  par conjugaison. Traduite en termes de  $\mathbb{P}^2(\mathbb{C})$  et de ses coordonnées homogènes, cette opération est

$$(e^{it}, e^{iu}) \cdot [x, y, z] = [e^{it}x, e^{iu}y, z]$$

---

<sup>(10)</sup>Si les  $a_i$  sont distincts, on n'a pas besoin de  $P$ .

dont les points fixes sont  $[1, 0, 0]$ ,  $[0, 1, 0]$  et  $[0, 0, 1]$ . Ce sont bien ceux qui correspondent aux matrices diagonales

$$\begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & & \\ & 1 & \\ & & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix}$$

de  $\mathcal{H}_\Lambda$ . L'image de  $f_F$  est donc l'enveloppe convexe des points 1,  $i$  et 0 comme on l'a vu sur la figure 1 (et d'ailleurs aussi sur la figure 3).

**3.2. D'autres théorèmes de convexité.** La démonstration des théorèmes 1 et 1 bis que j'ai expliquée ici est un produit des idées utilisées par Atiyah dans [2] pour démontrer un théorème de convexité assez général, que je vais énoncer maintenant, même si je n'ai pas défini toutes les notions.

**Théorème 4.** *Soit  $W$  une variété compacte et connexe munie d'une forme symplectique. Soit  $\mathcal{T}$  un tore de dimension  $r$  opérant sur  $W$  avec application moment*

$$\mu : W \longrightarrow \mathbb{R}^r.$$

*Alors l'image de  $\mu$  est un polytope convexe dans  $\mathbb{R}^r$ , l'enveloppe convexe des images des points fixes de l'opération du tore.*

Le théorème 3 en est une application directe : la variété symplectique compacte et connexe est notre  $\mathcal{H}_\Lambda$ . Le tore  $\mathcal{T}$  est

$$\mathcal{T} = \left\{ \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix} \text{ telles que } |z_i| = 1 \right\} \subset U(n)$$

et il opère sur  $\mathcal{H}_\Lambda$  par conjugaison.

Le sous-groupe de  $\mathcal{T}$  défini par la  $j$ -ème coordonnée, ou par les équations  $z_1 = \cdots = z_{j-1} = z_{j+1} = \cdots = 1$  est exactement, avec les notations précédentes,

$$\mathcal{T}_{X_j} = \{\exp(itX_j)\}$$

pour

$$X_j = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \ddots \end{pmatrix}$$

de sorte qu'on a  $n$  fonctions  $f_{X_1}, \dots, f_{X_n}$  liées à l'opération de ces sous-groupes comme ci-dessus. L'application *moment*  $\mu$  dont il est question dans l'énoncé du théorème 4 n'est rien d'autre que

$$\mu = (f_{X_1}, \dots, f_{X_n})$$

et, plus concrètement

$$\mu(V) = (\text{tr}(X_1 V), \dots, \text{tr}(X_n V)) = (v_{11}, \dots, v_{nn}),$$

la diagonale de la matrice  $V$ .

La démonstration des théorèmes 3 ou 4 est fondée sur une récurrence (« connexité  $\Leftrightarrow$  convexité » — comme au début du § 2.1) et sur les arguments de théorie de Morse expliqués ci-dessus.

**3.3. Quelques compléments sur Toeplitz-Hausdorff.** Revenons aux exemples du § 1.1. Les matrices  $A$ ,  $D$  et  $F$  sont normales, on en déduit aisément les images numériques correspondantes sur la figure 1. Les images  $W(A)$ ,  $W(C)$  et  $W(E)$  sont sans doute plus intéressantes. Halmos signale dans [7] que, pour une matrice  $2 \times 2$ , l'image numérique est une ellipse dont les foyers sont les valeurs propres de cette matrice (d'où les images  $W(A)$ ,  $W(B)$  et  $W(C)$ ).

Plus directement accessibles par nos méthodes est le cas de la matrice  $E$ . Remarquons que c'est la somme directe

$$E = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \oplus (1).$$

La matrice  $2 \times 2$  est  $B$ , dont l'image numérique est le disque (figure 1). L'image numérique de la matrice  $(1)$  est évidemment le point  $1 \in \mathbb{C}$ , et  $W(E)$  est... l'enveloppe convexe de la réunion des deux.

### Indications bibliographiques

Le petit livre [5] est une introduction accessible (dans tous les sens du terme) à la théorie de Morse. On trouvera des rudiments de géométrie symplectique (entre autres choses) dans [1]. Le livre [3] est complètement consacré aux techniques et théorèmes évoqués dans ces notes.

Bien que je n'aie utilisé ici qu'une toute petite partie de [7], je profite de l'occasion pour conseiller sa lecture : l'idée directrice de ce livre est « c'est en faisant des mathématiques qu'on les apprend », idée certes élémentaire mais peu développée dans la littérature.

Les autres références citées en bibliographie sont les sources originales des théorèmes énoncés et/ou démontrés ici.

### Références

- [1] V. I. ARNOLD – *Méthodes mathématiques de la mécanique classique*, MIR, Moscou, 1974.
- [2] M. F. ATIYAH – « Convexity and commuting Hamiltonians », *Bull. London Math. Soc.* **23** (1982), p. 1–15.
- [3] M. AUDIN – *The topology of torus actions on symplectic manifolds*, Progress in Math., vol. 93, Birkhäuser, 1991.
- [4] V. A. GINZBURG – « Equivariant cohomology and Kähler geometry », *Funktsional. Anal. i Prilozhen.* **21** (1987), p. 271–283.
- [5] A. GRAMAIN – *Topologie des surfaces*, PUF, Paris, 1971.
- [6] V. GUILLEMIN & S. STERNBERG – « Convexity properties of the moment mapping I & II », *Invent. Math.* **67** (1982), p. 491–513, & **77** (1984), p. 533–546.
- [7] P. HALMOS – *A Hilbert space problem book*, Graduate Texts in Math., Springer, 1982.
- [8] A. HORN – « Doubly stochastic matrices and the diagonal of a rotation matrix », *Amer. J. Math.* **76** (1954), p. 620–630.
- [9] I. SCHUR – « Über eine Klasse von Mittelbildungen mit Anwendungen auf der Determinantentheorie », *Sitzungsberichte der Berliner Mathematischen Gesellschaft* **22** (1923), p. 9–20.

Michèle Audin, Institut de Recherche Mathématique Avancée, Université Louis Pasteur et URA 001 du CNRS, 7 rue René-Descartes, F-67084 Strasbourg Cedex