Mise à jour en 2011 du chapitre IV de Physique du Sens, Editions du CNRS, Paris, 1992.

Éléments de théorie des singularités

(D'après Thom, Mather, Malgrange, Arnold, Zeeman et Boardman, Chenciner, Golubitsky, Guillemin, Milnor, Poénaru, Porteous, Ruelle, Sard, Siersma, Tougeron, Wall)

Jean Petitot*

Table des Matières

*École des Hautes Études en Sciences Sociales (EHESS) et CREA (École Polytechnique), Paris, France.
e-mail : petitot@poly.polytechnique.fr
URL : http://jean.petitot.pagesperso-orange.fr/

Introduction

Régulièrement, des collègues et des étudiants me demandent où ils pourraient trouver une introduction à la théorie des singularités. Il est difficile de leur répondre car il n'existe que des textes pour mathématiciens qui sont trop techniques ou des textes pour non mathématiciens qui sont trop élémentaires. C'est pourquoi il m'a paru utile de rendre à nouveau disponible une introduction à la théorie de Thom-Mather que j'ai rédigée à la fin des années 70 et dont une version réduite a été publiée en 1992 aux Presses du CNRS dans *Physique du Sens*, ouvrage depuis longtemps épuisé. Cette introduction concerne les fondements de la théorie, des travaux pionniers de Hassler Whitney jusqu'aux théorèmes fondamentaux de classification des singularités dites simples, et elle est centrée sur les résultats de René Thom, Bernard Malgrange, John Mather, Vladimir Arnold et Christopher Zeeman dont elle constitue une compilation. Le lecteur y rencontrera aussi les travaux d'autres grands mathématiciens comme Boardman, Golubitsky, Guillemin, Milnor, Poénaru, Porteous, Ruelle, Sard, Siersma, Tougeron ou Wall.

Le but de cette synthèse est de montrer comment, à travers une chaîne de théorèmes fondamentaux, difficiles et techniques, un modèle très général fondé sur quelques principes et concepts très abstraits peut se spécialiser en un résultat de classification exhaustif, précis et contraint des réalisations de ce modèle. Cette extraordinaire 'réduction de l'arbitraire" (pour parler comme René Thom) faisant apparaître de la nécessité dans une classe de structures possibles est, selon nous, constitutive du rôle des mathématiques dans les sciences. Le théorème de classification des singularités simples fait partie des grands résultats de classification qui dominent l'histoire des mathématiques depuis l'antique classification des solides platoniciens (*i.e.*, en langage moderne, des sous groupes finis du groupe $SO(3)$ des rotations directes de $\mathbb{R}^3$).

J'aimerais remercier deux camarades qui, depuis cette époque lointaine, sont devenus d'éminents spécialistes de théorie des singularités et sont restés d'excellents amis : Alain Chenciner (à qui l'on doit le magnifique article sur les singularités dans l'*Encyclopædia Universalis*) et Bernard Teissier (l'un des maîtres des liens entre géométrie algébrique et géométrie différentielle en matière de singularités).

1 Singularités, bifurcations et modèles morphodynamiques

Nous allons aborder la théorie des singularités en considérant, comme René Thom l'a souvent fait, les ressources mathématiques qu'elle fournit pour l'élaboration de modèles dynamiques de morphologies naturelles, modèles que nous appellons *morphodynamiques*. Nous commençons donc par exposer le modèle morphodynamique général puis ses différentes spécifications mathématiques. C'est la troisième spécification qui introduira la théorie des singularités de

façon naturelle.

1.1 Le contenu général des modèles morphodynamiques

Le modèle morphodynamique général peut être introduit de façon naturelle dans le cadre d'une théorie des systèmes contrôlés.

Soit S un système quelconque conçu comme une "boîte noire". Supposons que les hypothèses suivantes, toutes très générales, soient satisfaites.

(i) A l'intérieur de la boîte noire, il existe un processus interne (en général inobservable) X qui définit les *états internes* que le système S est susceptible d'occuper de façon stable. Pour des raisons de simplicité, on peut supposer que ceux-ci sont en nombre fini.

(ii) Le processus interne X définit globalement l'ensemble des états internes de S. Cette hypothèse est essentielle. Elle signifie que les états internes sont en compétition et donc que le choix de l'un d'eux comme état actuel virtualise les autres. Autrement dit, ces états n'existent pas en tant qu'entités isolées. Ils s'entredéterminent à travers des rapports de détermination réciproque, les états virtualisés par le choix de l'état actuel constituant autant d'alternatives possibles à ce dernier.

(iii) Il existe une instance de sélection I qui, sur la base de certains critères (spécifiques du système et pouvant varier notablement) sélectionne l'état actuel parmi les états internes possibles.

(iv) Enfin, autre hypothèse essentielle, le système S est contrôlé par un certain nombre de paramètres de contrôle, paramètres variant dans un espace W que, pour l'opposer au processus interne X, on appelle l'*espace externe* (ou espace de contrôle ou encore espace "substrat") de S. On suppose de plus que, en un sens intuitif non encore spécifié mathématiquement, le contrôle est "continu". Cela signifie que le processus interne X est un processus X_w qui dépend de la valeur w du contrôle, varie continûment lorsque w varie continûment dans W et, en se déformant, déforme la structure des états internes ainsi que leurs relations de détermination réciproque.

Soit alors $\mathcal{X}$ "l'espace" (qui sera en général mathématiquement un espace fonctionnel) des processus internes X possibles. Si les hypothèses ci-dessus sont vérifiées, le système S sera décrit d'abord par le champ (continu) $\sigma : W \to \mathcal{X}$ associant à $w \in W$ le processus X_w et ensuite par l'instance de sélection I. Nous considérons donc des situations $S = (W, \mathcal{X}, \sigma, I)$. Les phénomènes thermodynamiques de transitions de phases en offrent des exemples physiques typiques.

Dans le cas d'une transition de phase du premier ordre, le système S est le système thermodynamique considéré (par exemple un volume d'eau), les états

internes sont les phases thermodynamiques (solide, liquide, gaz), l'instance de sélection I est le principe de minimisation de l'énergie libre et les paramètres de contrôle W sont, par exemple, la pression et la température. Quant au processus interne X_w, très complexe, il est celui de la dynamique moléculaire. Or l'expérience la plus commune (faire bouillir ou geler de l'eau) indique que de tels systèmes subissent, lorsque les paramètres de contrôle traversent certaines valeurs particulières dites valeurs *critiques* (par exemple lorsque la température T traverse la valeur 100° sous la pression atmosphérique normale), des *discontinuités* de leurs qualités observables et donc des transformations brusques d'état interne. Les valeurs critiques constituent un sous-ensemble K de W, dit *diagramme de phase*. K partitionne W en domaines correspondant aux diverses phases que peut présenter S. Autrement dit, il le "catégorise" et y "externalise" sous la forme d'un système de discontinuités, de frontières et de seuils, la compétition (la détermination réciproque) des états internes. Il s'agit là d'une possibilité générale qui est une conséquence directe des hypothèses (i)-(iv) et n'a donc rien de spécifiquement thermodynamique.

Comment se manifeste en effet *phénoménologiquement* un système de type $S = (W, \mathcal{X}, \sigma, I)$? Phénoménologiquement parlant, l'apparaître du système est déterminé par les *qualités observables* $q^1 \dots q^n$ qu'il manifeste lorsqu'il occupe un certain état interne. Autrement dit, le processus interne X_w "s'extériorise" en qualités sensibles observables q^i_w. Lorsque le contrôle w varie continûment, la dynamique interne X_w varie continûment (hypothèse (iv)), et donc les qualités q^i_w aussi sauf si l'instance de sélection I fait changer le système S d'état interne. Mais, phénoménologiquement parlant, une variation continue n'est qu'un affaiblissement de l'invariance. Qualitativement, c'est une invariance. Elle n'est donc pas significative (de l'eau plus ou moins chaude reste de l'eau, cette variation continue de degré étant d'un tout autre type que ces variations discontinues que sont les transitions de phases). Appelons alors avec René Thom *point régulier* de W une valeur w du contrôle telle que les qualités observables q^i_w varient continûment – et restent donc qualitativement invariantes, c'est-à-dire *stables* – dans tout un voisinage U de w. Cela présuppose évidemment que l'on ait défini la notion de voisinage sur W, *i.e.* une *topologie*. Les points réguliers constituent par définition un ouvert R_W de W, l'ouvert de stabilité des qualités. Soit maintenant K_W le fermé complémentaire de R_W dans W. Par définition, les points de K_W sont les valeurs w du contrôle telles que au moins une qualité observable q^i_w traverse une discontinuité. Ce sont des valeurs critiques à la traversée desquelles le système S devient le support d'un phénomène critique. On les appelle aussi des valeurs "catastrophiques", le fermé K_W s'appelant "l'ensemble de catastrophes" de S. On appelle également les K_W des morphologies externes.

Comme René Thom y a souvent insisté, ce concept de morphologie est purement phénoménologique. Mais il est intimement solidaire du concept mathématique de transition ou de bifurcation d'état interne. Car quelle peut

être la cause efficiente des discontinuités qualitatives observées? Supposons que le contrôle w parcourt un chemin γ dans W. Soit A_w l'état interne actuel initial sélectionné par I. Au cours de la déformation de X_w le long de γ – et donc, d'après l'hypothèse (iv), de la structure de A_w et des relations de détermination réciproque qu'il entretient avec les états virtuels B_w, C_w, etc. suivant l'hypothèse (ii) – il peut fort bien se produire que, à la traversée d'une valeur particulière, critique précisément, A_w ne satisfasse plus aux critères de sélection imposés par I d'après l'hypothèse (iii). Le système bifurque donc spontanément de A_w vers un autre état actuel (jusque là virtuel) B_w. Cette transition catastrophique d'état interne se manifeste par une discontinuité de certaines des qualités observables q_w^i. Autrement dit, c'est la déstabilisation, relativement à l'instance I, des états internes actuels sous la variation du contrôle qui induit dans l'espace externe W un ensemble de discontinuités qualitatives K_W. Nous rencontrons là une dialectique interne/externe constitutive du modèle général, une dialectique de l'expression des conflits internes par les morphologies externes.

Dans les cas simples, l'ensemble K_W constituera un système d'interfaces, analogue à un diagramme de phases, partitionnant l'espace externe W en domaines dont chacun correspondra à la zone de W où domine l'un des états internes. Les systèmes $S = (W, \mathcal{X}, \sigma, I)$ partitionnent donc naturellement leurs espaces externes. Ce sont des modèles *d'émergence du discontinu par variation continue.*

1.2 Les spécifications mathématiques du modèle général

1.2.1 Première spécification : le concept de stabilité structurelle

Donnons maintenant un aperçu de la façon dont le contenu général du modèle peut être naturellement mathématisé et devenir ainsi la source d'un univers très diversifié de modèles.

La première spécification consiste d'abord à supposer que, eu égard à leur nature, les processus internes X_w constituent un espace $\mathcal{X}$ muni d'une *topologie* $\mathcal{T}$ "naturelle" (*i.e.* significative pour le type de processus étudié). Cela signifie que l'on sait dire quand deux processus internes X et Y sont voisins et donc que l'on sait définir rigoureusement la continuité du champ $\sigma : W \to \mathcal{X}$. En se déplaçant dans l'espace $\mathcal{X}$ on peut alors déformer ses éléments X.

La spécification consiste ensuite à supposer que l'on sait définir le *type qualitatif* des processus X. Le type qualitatif est une relation d'équivalence – qui est en général définie par l'action d'un groupe G sur $\mathcal{X}$ – et donc un affaiblissement de la relation d'identité. C'est une identité faible. Soit $\widetilde{X}$ la classe d'équivalence de X pour le type qualitatif (*i.e.* l'orbite de X sous l'action de G). On cherchera à caractériser ce qui demeure invariant lorsque X varie dans $\widetilde{X}$ (*i.e.* varie à type qualitatif constant) par une information discrète, par exemple par les valeurs d'un nombre fini d'invariants $\tau_1, \ldots, \tau_k$.

Au niveau des invariants – c'est-à-dire à un niveau "grossier", qualitatif – la variation dans une classe d'équivalence $\widetilde{X}$ se réduit à l'identité. Il n'y a donc de variation *qualitative* que lorsqu'une déformation dans $\mathcal{X}$ fait changer de classe d'équivalence. La variation se manifeste alors par une discontinuité de la valeur de certains invariants et l'on retrouve la "bonne" situation du modèle général.

La partition de $\mathcal{X}$ en classes d'équivalence suivant le type qualitatif est une classification des éléments $X \in \mathcal{X}$. Si l'on considère l'espace $\mathcal{X}$ comme un "genre" d'entités, alors les classes $\widetilde{X}$, $\widetilde{Y}$, etc. constituent autant "d'espèces". Le modèle général débouche ainsi sur la reprise originale d'une problématique très ancienne, celle de la logique genre/espèce ou générique/spécial. Les éléments X d'une classe C (si $X \in C$ alors $C = \widetilde{X}$) sont autant de spécialisations (de "tokens") d'un "type générique" associé à C.

L'idée de base est alors d'étudier les rapports entre cette logique classificatoire et la topologie $\mathcal{T}$. Par rapport à un point de vue standard qui consisterait à étudier, pour chaque système concret de type $S = (W, \mathcal{X}, \sigma, I)$, les processus internes X comme des entités isolées, on introduit un double déplacement de point de vue. D'abord, on considère comme objet d'étude non seulement les processus X mais également les *familles paramétrées* $(X_w)_{w \in W}$, et cela en se focalisant sur la géométrie des ensembles catastrophiques K_W induits dans les espaces externes W par la déstabilisation des états internes définis par les X_w. Ensuite, et surtout, on considère ces familles comme l'image de champs $\sigma : W \to \mathcal{X}$ "plongeant" l'espace externe W (qui, en général, sera un morceau d'espace standard $\mathbb{R}^n$) dans l'espace généralisé $\mathcal{X}$. Cela permet de réinterpréter la dialectique de l'expression des conflits internes par les morphologies externes à partir de la dialectique entre variation et invariance.

En effet, dès que l'on dispose sur un espace $\mathcal{X}$ d'une topologie $\mathcal{T}$ et d'une relation d'équivalence définissant le type qualitatif, on peut définir de façon naturelle une notion de *stabilité structurelle*.

Définition. – *Soit $X \in \mathcal{X}$. On dit que X est* structurellement stable *si tout Y assez voisin de X au sens de la topologie $\mathcal{T}$ est équivalent à X.*

X est donc structurellement stable si son type qualitatif "résiste" aux petites perturbations, ou encore si la classe d'équivalence $\widetilde{X}$ est *ouverte* (au sens de $\mathcal{T}$) en X.

Soit alors $K_{\mathcal{X}}$ le sous-ensemble fermé de $\mathcal{X}$ constitué des $X \in \mathcal{X}$ structurellement *instables*. $K_{\mathcal{X}}$ est une sorte d'ensemble catastrophique *intrinsèque*, canoniquement associé à l'espace fonctionnel $\mathcal{X}$ et au genre d'entités qu'il unifie. C'est une morphologie discriminante qui le catégorise et classifie les types qualitatifs de ses éléments structurellement stables. Autrement dit, $K_{\mathcal{X}}$ géométrise la classification interne à $\mathcal{X}$.

Soit alors $W \xrightarrow{\sigma} \mathcal{X}$ le champ caractéristique d'un système $S = (W, \mathcal{X}, \sigma, I)$. Soit $K'_W = \sigma^{-1}(K_{\mathcal{X}} \cap \sigma(W))$ la *trace* de $K_{\mathcal{X}}$ sur W par image inverse de σ. L'hypothèse de la modélisation est que l'ensemble catastrophique K_W de

S est déductible de K'_W à partir de l'instance de sélection I. Elle signifie qu'une valeur w du contrôle appartient à K_W (*i.e.* est une valeur critique) si et seulement si la situation en w est corrélée de la façon imposée par I à une situation appartenant à K'_W.

C'est par conséquent l'analyse (à la fois locale et globale) des ensembles catastrophiques "intrinsèques" $K_{\mathcal{X}}$ qui est au centre de la théorie. Si l'on introduit de plus l'hypothèse – évidente *a priori* – qu'un champ σ ne peut exister concrètement que s'il est structurellement stable, l'on est conduit à constater qu'une telle contrainte *borne* drastiquement *la complexité* susceptible d'être présentée par les K'_W. Dans les bons cas, on peut même accéder à une classification des structures locales des K'_W et donc des morphologies externes locales. La théorie fait ainsi apparaître des contraintes purement mathématiques contraignant le domaine des phénomènes morphologiques.

1.2.2 Deuxième spécification : la théorie des systèmes dynamiques

La spécification mathématique majeure du modèle général consiste à postuler que le processus interne X est un *système dynamique* différentiable sur un espace M de paramètres internes caractéristiques du système S considéré. On appelle M l'*espace interne*.

Dit très brièvement, le concept mathématique de système dynamique mathématise celui d'évolution temporelle déterministe d'un système à un nombre fini de degrés de liberté. On suppose donc que chaque état instantané[1] de S est descriptible par les valeurs d'un nombre fini de paramètres $x_1, \ldots, x_m$, autrement dit par un point parcourant une variété différentiable[2] M. Par exemple, en mécanique classique, un état instantané d'un système de N points matériels est décrit par $6N$ coordonnées : les coordonnées spatiales de chaque point ($3N$ coordonnées) et les composantes de la vitesse de chaque point ($3N$ coordonnées).[3]

Un système dynamique X sur M consiste d'abord à associer à chaque point x de M un vecteur tangent (un vecteur "vitesse") $X(x)$ à M en x, vecteur variant différentiablement avec x. X est donc un champ de vecteurs différentiable sur M, autrement dit, en termes des coordonnées locales $x_1, \ldots, x_m$, un système d'équations différentielles ordinaires :

$$\frac{dx_i}{dt} = f_i(x_1, \ldots, x_m)$$

[1] Il ne faut pas confondre "état instantané" et "état interne". *Cf.* plus bas.

[2] Une variété différentiable M de dimension m est un espace localement identique à $\mathbb{R}^m$ où l'on peut transporter toutes les notions (qui sont des notions locales) du calcul différentiel dans $\mathbb{R}^m$ et les globaliser par recollement. Pour plus de précisions, *cf.* plus bas. *Cf.* également Petitot [31] et [32].

[3] L'espace de représentation des états instantanés M (dit en mécanique espace de phases) n'est pas toujours un espace standard $\mathbb{R}^n$. Certaines coordonnées peuvent par exemple être angulaires et décrire des cercles.

où les composantes du champ f_i sont des fonctions différentiables des x_j et où t est le paramètre temporel.

Etant donné un tel champ, l'intégrer consiste à trouver dans M des courbes différentiables paramétrées par le temps t, *i.e.* des applications différentiables

$$\begin{aligned} \gamma : \mathbb{R} &\to M \\ t &\longmapsto \gamma(t) = (x_1(t), \ldots, x_m(t)) \end{aligned}$$

qui admettent en chaque point pour vecteur vitesse $\dfrac{dx}{dt} = \dfrac{d\gamma(t)}{dt}$ le vecteur du champ $X(x) = X(\gamma(t))$.

On dit que le champ est un système dynamique si :

(i) on peut intégrer les trajectoires sur un temps infini (*i.e.* si les trajectoires ne font pas "sortir" de M);

(ii) par tout point il ne passe qu'une et une seule trajectoire (principe de déterminisme : la condition initiale $x(0)$ au temps $t = 0$ détermine univoquement l'évolution future $x(t)$ pour $t > 0$ et l'évolution passée $x(t)$ pour $t < 0$);

(iii) les trajectoires varient différentiablement en fonction des conditions initiales.

Soit alors $\varphi_t : M \to M$ l'application qui associe à tout x de M le point au temps t de la trajectoire issue de x au temps $t = 0$. Il est facile de voir que φ_t est un *difféomorphisme* de M (autrement dit un automorphisme de sa structure de variété différentiable) et que l'application $\varphi : \mathbb{R} \to \operatorname{Diff} M$, $t \mapsto \varphi_t$ du groupe additif de $\mathbb{R}$ dans le groupe des difféomorphismes de M est un morphisme de groupe (*i.e.* $\varphi_0 =$ Identité de M, $\varphi_{t+t'} = \varphi_{t'} \circ \varphi_t$ et $\varphi_{-t} = (\varphi_t)^{-1}$). L'application φ s'appelle le *flot* du système dynamique X. C'est la version intégrale du champ de vecteurs X.

Pour définir les états internes des modèles $S = (W, \mathcal{X}, \sigma, I)$ où les processus internes X_w sont des systèmes dynamiques, l'idée de base est d'introduire une différence entre dynamique *rapide* et dynamique *lente*, c'est-à-dire entre deux échelles de temps, l'une interne rapide, l'autre externe lente. Comme l'a expliqué René Thom,

> "L'idée philosophique essentielle sous-jacente à la théorie des catastrophes est que tout phénomène, toute forme spatio-temporelle doit son origine à une distinction qualitative des modes d'action du *temps* dans les choses. Toute distinction d'apparence qualitative dans un espace W (le substrat), peut être attribuée à deux modes d'action du temps : un mode "rapide" qui crée dans un espace interne des "attracteurs" qui spécifient la *qualité* phénoménologique

locale du substrat; et un mode "lent" agissant dans l'espace substrat W lui-même".[4]

Autrement dit, on suppose – condition dite "d'adiabaticité" – que la dynamique interne d'évolution des états instantanés est "infiniment" rapide par rapport aux dynamiques externes d'évolution dans les espaces externes W. Seuls comptent donc les états *asymptotiques* (pour $t \to +\infty$) définis par les X_w, *i.e.* les régimes limites. Or l'analyse de ces états asymptotiques s'est révélée être d'une difficulté inattendue et redoutable. En effet, la complexité d'un système dynamique est en général prodigieuse.

D'abord le déterminisme idéal qu'est le déterminisme mathématique n'implique en rien un déterminisme au sens concret du terme (le déterminisme comme prédictibilité).[5] Concrètement, une condition initiale ne peut en effet être définie qu'approximativement. Elle n'est pas représentée par un point x_0 de M mais par un petit domaine U "épaississant" x_0. Pour que le déterminisme soit concret, il faut donc que les trajectoires issues de U forment un petit tube "épaississant" la trajectoire γ issue de x_0. Techniquement parlant, cela signifie que la trajectoire γ est *stable* relativement à de petites perturbations de sa condition initiale.[6]

Un système dynamique concrètement déterministe est donc un système dynamique (par définition idéalement déterministe) dont les trajectoires sont stables. Mais cela n'a aucune raison d'être le cas en général. Il existe même des systèmes dynamiques (par exemple les systèmes de géodésiques sur les variétés riemanniennes de courbure négative) qui présentent la propriété que *toutes* leurs trajectoires sont instables, et qui présentent qui plus est cette propriété de façon *structurellement stable*. Ainsi que l'a noté Vladimir Arnold dans les années 1970,

> "l'éventualité de systèmes structurellement stables à mouvements compliqués dont chacun est exponentiellement instable en soi est à mettre au rang des plus importantes découvertes faites ces dernières années en théorie des équations différentielles. (...) Jadis on supposait que dans les systèmes d'équations différentielles génériques ne pouvaient exister que des régimes limites stables simples : des positions d'équilibre et des cycles. Si le système était plus compliqué (conservatif par exemple), on admettait que sous l'effet d'une faible modification des équations (par exemple si l'on tenait compte des petites perturbations non conservatives) les mouvements compliqués se "décomposaient" en mouvements simples. Maintenant

[4] Thom [53].

[5] *Cf.* par exemple Ruelle [38].

[6] Cette stabilité des trajectoires est relative à un seul système dynamique X. Elle ne doit pas être confondue avec la stabilité structurelle de X qui est, elle, relative à l'espace fonctionnel $\mathcal{X}$.

> nous savons qu'il en va autrement et que dans l'espace fonctionnel des champs de vecteurs il existe des domaines composés de champs où les courbes de phase [les trajectoires] sont plus complexes. Les conclusions qui en découlent couvrent un grand nombre de phénomènes dans lesquels les objets déterministes ont un comportement 'stochastique'".[7]

L'indéterminisme concret (le chaos, le hasard, l'aléatoire, etc.) est donc parfaitement compatible avec le déterminisme mathématique. Comme le remarque Thom,

> "ce que l'on appelle 'lois du hasard' ne sont en fait que des propriétés du système déterministique le plus général".[8]

Revenons à la spécification du modèle général. En termes de systèmes dynamiques, les états internes de S sont les *attracteurs* de X_w. La notion très délicate d'attracteur généralise celle de point d'équilibre stable (*i.e.* attractif). Intuitivement, un attracteur A de X est un régime asymptotique stable. C'est un ensemble fermé, X-invariant et indécomposable pour ces deux propriétés (*i.e.* minimal) qui attire (*i.e.* qui capture asymptotiquement) toutes les trajectoires issues des points d'un de ses voisinages. Le plus grand voisinage de A, $B(A)$, ayant cette propriété s'appelle le *bassin d'attraction* de A. Dans les cas simples, les attracteurs auront une structure topologique simple (point attractif ou cycle attractif), seront en nombre fini et leurs bassins seront de "bons" domaines (de forme simple) séparés par des séparatrices. Mais cette image est trop optimiste car :

(i) les attracteurs peuvent être en nombre infini;

(ii) les bassins peuvent être intriqués les uns dans les autres de façon inextricable;

(iii) les attracteurs peuvent avoir une topologie très compliquée (attracteurs dits "étranges").

Sur un attracteur, les trajectoires d'un système dynamique présentent de la *récurrence.*[9] Intuitivement, la récurrence d'une trajectoire γ signifie que si $x \in \gamma$, γ repasse aussi près que l'on veut de x après un délai aussi grand que l'on veut et que γ revient donc indéfiniment sur elle-même. Les cas triviaux de récurrence sont les points fixes de X (les points où X s'annule, *i.e.* les

[7] Arnold [4], pp. 314-315.

[8] Thom [52], p. 124.

[9] Cette notion capitale est due à Poincaré et à Birkhoff, les fondateurs de la Dynamique qualitative.

trajectoires réduites à un point) et les cycles de X (les trajectoires fermées). Mais il existe en général de la récurrence non triviale. Si γ est une trajectoire récurrente "compliquée" et si A est sa fermeture topologique, A est un domaine entier de M (un fermé d'intérieur non vide) où γ se répand de façon dense.

Quoi qu'il en soit de ces difficultés, on peut supposer pour simplifier que, pour *presque toute* condition initiale $x_0 \in M$ (seulement presque toute car il faut tenir compte des séparatrices entre bassins), la trajectoire issue de x_0 est capturée asymptotiquement par un attracteur A_w de la dynamique interne X_w. Cela correspond à une hypothèse *d'équilibre local*[10] : la dynamique interne "rapide" entraîne le système vers un régime asymptotique stable correspondant à un état interne.

Une fois admises ces diverses hypothèses qui, dans le cadre différentiel, ne sont que très peu restrictives et s'imposent d'elles-mêmes, le modèle général se convertit de lui-même en un programme mathématique. Il s'agit en effet d'analyser en détail des problèmes du type suivant :

(i) structure générale des systèmes dynamiques (Dynamique qualitative ou Global Analysis);

(ii) caractérisation géométrique des systèmes dynamiques structurellement stables et de leurs attracteurs;

(iii) analyse des propriétés ergodiques des trajectoires sur les attracteurs "étranges";

(iv) analyse des causes possibles d'instabilité;

(v) analyse des déformations (des perturbations) des systèmes structurellement instables;

(vi) étude de la géométrie (qui peut être d'une extrême complexité) des ensembles catastrophiques "intrinsèques" $K_{\mathcal{X}}$; etc.

Ce programme, que l'on pourrait appeler le programme de Thom-Smale, prolonge celui de Poincaré et Birkhoff. C'est en fait celui de la Dynamique qualitative moderne. A travers lui, la Morphodynamique se révèle être solidaire d'un domaine et d'une tradition mathématiques d'une importance éminente tant sur le plan technique que sur le plan historique. Ceci dit, le programme de Thom-Smale, s'il est d'une immense portée, est également d'une immense difficulté. C'est pourquoi Thom a proposé de le simplifier drastiquement.

[10] *Cf.* Thom [53], p. 2.

1.2.3 Troisième spécification : la théorie des singularités

La complexité des systèmes dynamiques généraux étant trop grande pour être maîtrisée, on peut en faire une étude "grossière", de type thermodynamique. Celle-ci consiste à ne pas tenir compte de la structure fine (de la topologie compliquée) des attracteurs. Elle est d'autant plus nécessaire que les ensembles catastrophiques empiriques K_W sont en général beaucoup plus simples que ceux induits par les bifurcations de systèmes dynamiques généraux. Il faut donc comprendre comment des systèmes peuvent être "intérieurement" chaotiques (stochasticité des attracteurs définissant les états internes) et "extérieurement" ordonnés (simplicité des morphologies observables), autrement dit ce que Thom appelle

> "l'émergence du descriptible à partir de l'indescriptible".[11]

L'idée est de tenter de généraliser aux systèmes généraux ce qui se passe dans le cas des systèmes de *gradient*, à savoir l'existence de lignes de pente et de variétés de niveau. Pour cela, on utilise le fait que, si A est un attracteur d'un système dynamique X sur une variété M, on peut construire sur le bassin $B(A)$ de A une fonction positive f – dite *fonction de Liapounov* – qui décroît strictement sur les trajectoires dans $B(A)-A$ et s'annule sur A. Cette fonction est une sorte d'entropie locale exprimant que, au cours du temps, $B(A)$ se "contracte" sur A de façon analogue à un système de gradient. Mais elle ne dit rien sur la structure interne de l'attracteur.

L'idée est alors de ne retenir des bifurcations d'attracteurs que les bifurcations associées de leurs fonctions de Liapounov. Cette réduction ressemble à un moyennage thermodynamique. Elle correspond à un changement de niveau d'observation faisant passer du niveau "fin" décrit par les X_w au niveau "grossier" décrit par les f_w. Elle est analogue à celle que l'on trouve dans la théorie du champ moyen (théorie de Landau) en théorie des transitions de phases. Thom en donne la justification suivantes :

> "Personnellement, j'aime à penser que ce qui joue un rôle, ce n'est pas la notion – trop fine – d'attracteur, mais une classe d'équivalence d'attracteurs "équivalents" parce qu'encapsulés dans la variété de niveau d'une fonction de Liapounov (un quasi-potentiel) pourvu que l'attracteur échappe à des implosions de caractère exceptionnel. Telle serait, selon nous, la voie par où trouver une définition mathématiquement satisfaisante de la notion de régime asymptotique stationnaire d'une dynamique. Dans une telle optique, la structure fine interne de l'attracteur n'a que peu d'importance : seule importe la fonction de Liapounov qui l'enserre dans une de

[11] *Cf.* Thom [52], p. 124.

ses variétés de niveau. Mais on peut concevoir que seule la structure du tube enfermant l'attracteur a phénoménologiquement de l'importance, et on retrouve ainsi une problématique proche de la théorie des catastrophes élémentaires".[12]

D'où l'intérêt de passer des quasi-potentiels que sont les fonctions de Liapounov aux systèmes de gradient dérivés d'un "vrai" potentiel. On suppose par conséquent que la dynamique interne X_w est en fait la dynamique de gradient associée à une fonction potentiel différentiable $f_w : M \to \mathbb{R}$. Les états internes déterminés par f_w sont alors ses *minima* (si f est assimilée à une "énergie", ce principe est celui de la minimisation de l'énergie du système).

Mathématiquement, ces modèles – baptisés "statiques" par René Thom – font donc partie intégrante de la théorie des bifurcations des fonctions potentiel. Or, pour les potentiels, il existe une caractérisation simple de la stabilité structurelle (théorème de Morse, *cf.* plus bas section 5.3). Sous l'hypothèse que la variété M soit compacte, $f : M \to \mathbb{R}$ est structurellement stable si et seulement si :

(i) ses points critiques, *i.e.* ses minima, ses maxima et ses cols (dans le produit cartésien $M \times \mathbb{R}$, le graphe de f constitué de l'ensemble des couples $(x, f(x))$ est comme un "relief" au-dessus de M) sont non dégénérés, c'est-à-dire ne sont pas des fusions de plusieurs minima, maxima ou cols; et si

(ii) ses valeurs critiques (*i.e.* les valeurs $f(x)$ de f pour x critique) sont distinctes.

Il y a donc deux causes d'instabilité structurelle :

(i) la dégénérescence de points critiques, correspondant aux catastrophes dites *de bifurcation*;

(ii) l'égalité de deux valeurs critiques, correspondant aux catastrophes dites *de conflit*.

A ces deux types bien distincts de catastrophes correspondent respectivement deux types d'instances de sélection I, ce que Thom a appelé deux *conventions* :

(i) la convention du "retard parfait" selon laquelle le système S demeure dans un état interne (un minimum de f_w) tant que celui-ci existe : il n'y a donc catastrophe que lorsqu'un minimum disparaît par fusion avec un autre point critique (bifurcation);

[12] Thom [53], p. 5.

(ii) la convention de Maxwell selon laquelle le système S occupe toujours le minimum absolu de f_w : il n'y a donc catastrophe que lorsqu'un autre minimum devient à son tour le minimum absolu (conflit).

L'une des causes majeures d'instabilité structurelle des fonctions potentiel est ainsi l'existence de singularités. Ce sont donc celles-ci qu'il faut étudier en prorité. C'est ce que nous nous proposons de faire maintenant.

Notre plan sera le suivant. Après avoir rappelé certains résultats généraux de base nous indiquerons les résultats fondamentaux de John Mather permettant de caractériser la stabilité structurelle. Nous esquisserons ensuite la théorie de Thom-Boardman permettant d'analyser la géométrie des applications en utilisant des stratifications canoniques des espaces de jets. Cela nous conduira au théorème de classification des singularités dites "simples" (au sens d'Arnold) et à l'établissement des formes normales de leurs déploiements universels. Nous analyserons alors avec quelques détails la géométrie des singularités les plus élémentaires.

2 Généralités sur les singularités

On veut étudier les modèles morphodynamiques de type $\sigma : W \to \mathcal{F}$ où $\mathcal{F}$ est l'espace fonctionnel des potentiels (différentiables) $f : M \to \mathbb{R}$ sur un espace interne M. On considère donc des familles f_w paramétrées par des espaces externes W. On peut associer à f_w l'application différentiable, dite *déploiement*

$$\begin{aligned} F : M \times W &\to \mathbb{R} \times W \\ (x, w) &\mapsto (f_w(x), w) \end{aligned}$$

Dire que f_w est un déploiement *universel* de $f_0 = f$, c'est dire que F est une application autant que faire se peut structurellement stable de $M \times W$ dans $\mathbb{R} \times W$ où la dimension de W est minimale. Il est par suite naturel de généraliser le problème initial et de se proposer l'analyse de la structure des applications différentiables $f : M \to N$ entre deux variétés différentiables quelconques.

Nous allons présenter un résumé des bases de la théorie sans entrer dans les difficultés trop techniques. Ces difficultés sont nombreuses et en grande partie dues au fait que les espaces fonctionnels d'applications différentiables sont des espaces compliqués sur lesquels il est impossible de faire directement de la géométrie. La stratégie générale sera donc de réduire, autant que faire se peut, les situations rencontrées à des situations *algébriques* de dimension *finie* sur lesquelles il devient possible de calculer et de faire de la "bonne" géométrie.

2.1 Le concept de variété différentiable

Rappelons d'abord brièvement le concept de *variété différentiable.*[13] Une variété différentiable de dimension m est un espace obtenu par *recollement* de "morceaux" d'espace standard $\mathbb{R}^m$. C'est un espace topologique[14] qui peut être recouvert par un ensemble $(U_i)_{i\in I}$ d'ouverts tels que

(i) chaque U_i soit homéomorphe[15], *via* un homéomorphisme $\varphi_i : U_i \to V_i$, à un ouvert V_i de $\mathbb{R}^m$;

(ii) pour tout couple (i, j), les applications induites par $\varphi_j \circ \varphi_i^{-1}$ entre l'image dans V_i de l'intersection $U_i \cap U_j$ et son image dans V_j soient différentiables[16] (*cf.* figure 1).

Les ouverts U_i sont appelés *cartes locales* et les fonctions différentiables $\varphi_j \circ \varphi_i^{-1}$ des changements de cartes locales. A travers les φ_i, on peut transporter aux U_i les systèmes de coordonnées de $\mathbb{R}^m$ et donc raisonner sur M en termes de *coordonnées locales*, les changements de cartes s'identifiant à des changements différentiables de coordonnées locales. Cela permet de prolonger aux variétés différentiables toutes les notions, les entités et les constructions qui relèvent de la structure différentiable des espaces standard $\mathbb{R}^m$ et qui sont de nature locale. Si par exemple $f : M \to \mathbb{R}^n$ est une application de M dans $\mathbb{R}^n$, on dira qu'elle est différentiable si pour tout $x \in M$ il existe une carte locale $\varphi : U \to V$ contenant x telle que $f \circ \varphi^{-1} : V \to U \to \mathbb{R}^n$ soit différentiable. On définit de même les applications différentiables entre variétés différentiables.[17]

Remarque. – Ainsi, une variété différentiable M de dimension m est un espace qui est identifiable à $\mathbb{R}^m$ seulement localement et non pas globalement. D'où la nécessité de développer des méthodes permettant de "mesurer" la non

[13]Pour les lecteurs non mathématiciens désireux de se documenter sur ce concept, signalons que l'article "Variétés différentiables" de l'*Encyclopædia Universalis* est excellent.

[14]Un espace topologique M est un ensemble où l'on a défini la notion de voisinage entre éléments. Un ouvert U de M est alors un sous-ensemble qui est un voisinage de chacun de ses éléments. Dans l'espace euclidien $\mathbb{R}^m$ les ouverts de base sont les boules ouvertes $B(a, \ell)$ ensemble des points $x \in \mathbb{R}^m$ dont la distance à $a \in \mathbb{R}^m$ est strictement inférieure à $\ell > 0$ fixé et U est ouvert si et seulement si pour tout $a \in U$ il existe $\ell > 0$ tel que $B(a, \ell) \subset U$. Là encore l'article "Topologie" de l'*Encyclopædia Universalis* est excellent et fournit une bonne introduction pour les non mathématiciens.

[15]Si M et N sont deux espaces topologiques, les applications $f : M \to N$ qui sont compatibles à leur structure topologique (*i.e.* qui préservent la relation de voisinage) sont les applications continues. Une application continue est un homéomorphisme si c'est un isomorphisme pour la structure topologique, c'est-à-dire si c'est une bijection continue dont l'inverse est également continu.

[16]Une application $f : V \to \mathbb{R}^n$ d'un ouvert V de $\mathbb{R}^m$ dans $\mathbb{R}^n$ est dite différentiable si ses composantes $f_i(x_1, \ldots, x_m)$, $i = 1, \ldots, n$, admettent des dérivées partielles à tous les ordres. Il s'agit là d'une notion locale.

[17]Suivant l'ordre de différentiabilité 1, 2, ..., k, ..., ∞, on parle alors d'applications de classe C^1, C^2, ..., C^k, ..., C^∞.

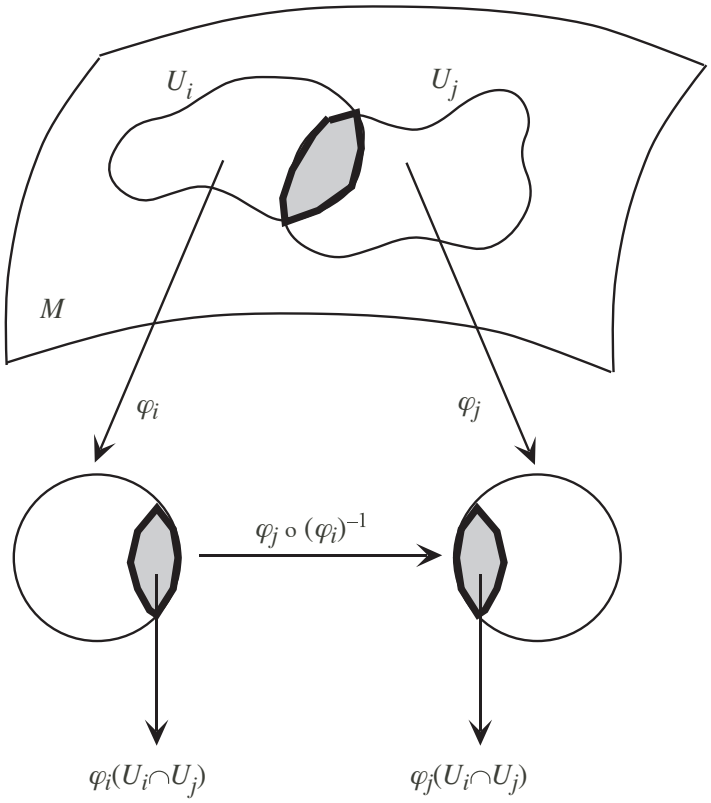


Figure 1: Recollement différentiable de deux ouverts d'une variété différentiable M.

trivialité globale de M. Introduites par Poincaré, ces méthodes dites homotopiques, homologiques et cohomologiques constituent une part essentielle de la géométrie différentielle moderne.

2.2 Vecteurs tangents et fibré tangent

Soit M une variété différentiable. On peut définir la notion de *vecteur tangent* à M en un point a de la façon suivante. Considérons des courbes différentiables de M passant par a, c'est-à-dire des applications différentiables $\gamma : (I, 0) \to (M, a)$ d'un segment ouvert $I =]-\varepsilon, +\varepsilon[$ de $\mathbb{R}$ dans M envoyant 0 sur a. Si $(x_1, \ldots, x_m)$ est un système de coordonnées locales en a, γ est définie par les nombres $x_i(t)$, $i = 1, \ldots, m$, et l'on peut poser que deux courbes γ et γ' sont équivalentes si les dérivées $\left.\dfrac{dx_i}{dt}\right|_{t=0}$ et $\left.\dfrac{dx'_i}{dt}\right|_{t=0}$ sont égales pour tout i, c'est-à-dire si leurs vecteurs vitesse en $t = 0$ sont identiques.[18] On appellera alors vecteur tangent en a à M une classe d'équivalence de courbes pour cette relation d'équivalence. Si ξ est un vecteur tangent en a à M, ses composantes sont les m nombres $\xi_i = \left.\dfrac{dx_i}{dt}\right|_{t=0}$. Il est facile de voir que les vecteurs tangents en a à M constituent un espace vectoriel sur $\mathbb{R}$ de dimension m. Celui-ci est dit *espace tangent* en a à M et noté T_aM.

La meilleure façon d'appréhender les vecteurs tangents à une variété différentiable M est de les traiter comme des *opérateurs différentiels* opérant sur les fonctions différentiables $f : M \to \mathbb{R}$. Soient en effet f une telle fonction, $\xi = (\xi_1, \ldots, \xi_m)$ un vecteur de T_aM et $\gamma : (I, 0) \to (M, a)$ une courbe différentiable représentant ξ. Restreinte à γ, f devient une fonction $f(\gamma(t))$ de t dont la dérivée en $t = 0$ est donnée par:

$$\left.\frac{df}{dt}\right|_{t=0} = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(a) \left.\frac{dx_i}{dt}\right|_{t=0} = \sum_{i=1}^{m} \xi_i \frac{\partial f}{\partial x_i}(a) \ . \tag{1}$$

Cette formule permet d'identifier ξ à l'opérateur $\sum_{i=1}^{m} \xi_i \dfrac{\partial}{\partial x_i} = D_\xi$ associant à chaque fonction différentiable $f : M \to \mathbb{R}$, sa dérivée $\xi(f) = D_\xi(f)$ relativement à ξ. Il est facile de vérifier que, ainsi conçu, ξ est un opérateur de dérivation, c'est-à-dire un opérateur linéaire s'annulant sur les constantes et satisfaisant la formule lebnizienne du produit :

$$\xi(fg) = f\,\xi(g) + g\,\xi(f). \tag{2}$$

Dans cette optique T_aM est donc l'espace vectoriel de base $\left(\dfrac{\partial}{\partial x_1}, \ldots, \dfrac{\partial}{\partial x_m}\right)$ associée aux coordonnées locales $(x_1, \ldots, x_m)$.

[18] Il est trivial de vérifier qu'il s'agit bien là d'une relation d'équivalence indépendante des coordonnées locales choisies.

La base $\left(\dfrac{\partial}{\partial x_i}\right)$ de T_aM étant déterminée par le choix des coordonnées locales (x_i), l'ensemble des espaces tangents T_aM est identifiable dans un voisinage U de a au produit direct $U \times \mathbb{R}^m$ de coordonnées (x_i, ξ_i). Mais *globalement* il n'en est rien. L'ensemble TM des espaces tangents T_aM est canoniquement muni d'une structure de variété différentiable de dimension $2m$ qui a la propriété qu'il existe une projection $\pi : TM \to M$ (associant à chaque vecteur tangent $\xi \in TM$ son origine $a \in M$) dont chaque fibre $\pi^{-1}(a) = T_aM$ est un espace vectoriel. Une telle structure s'appelle un *fibré vectoriel* de base M et, muni de sa structure de fibré vectoriel, TM s'appelle le *fibré tangent* de M. Ce fibré est localement trivial, un fibré trivial étant par définition difféomorphe[19] en tant que fibré au produit direct de sa base M et de sa fibre. Mais il n'est pas en général globalement trivial. La "mesure" de sa non trivialité fournit même des renseignements précieux sur la structure globale de M. Par exemple, si TM est trivial, on dit que M est *parallélisable*. L'on peut montrer (théorème d'Adams) que les seules sphères parallélisables sont S^1, S^3 et S^7. Ce résultat profond est intimement lié au fait que les espaces $\mathbb{R}^2, \mathbb{R}^4$ et $\mathbb{R}^8$ (dont S^1, S^3 et S^7 sont les sphères unité) sont les seuls espaces vectoriels sur $\mathbb{R}$ possédant une structure multiplicative en faisant des corps ($\mathbb{R}^2$ étant le corps des nombres complexes, $\mathbb{R}^4$ celui des quaternions et $\mathbb{R}^8$ celui des nombres de Cayley).

2.3 Applications linéaires tangentes et jacobiens

La définition des fibrés tangents permet de prolonger aux applications différentiables $f : M \to N$ entre variétés différentiables la notion de dérivée d'une fonction $f : \mathbb{R} \to \mathbb{R}$, c'est-à-dire de son approximation linéaire. Soient en effet $a \in M$, $f(a)$ son image et ξ un vecteur de T_aM. ξ est un opérateur de dérivation sur les $g : M \to \mathbb{R}$. Par composition avec f on lui associe un opérateur de dérivation $D_af(\xi)$ sur les $h : N \to \mathbb{R}$ défini par

$$D_af(\xi)(h) = \xi(h \circ f) \ . \tag{3}$$

Pour voir que $D_af(\xi)$ est bien un vecteur tangent η à N en $f(a)$, il suffit de considérer des coordonnées locales $(x_1, \ldots, x_m)$ en a et $(y_1, \ldots, y_n)$ en $y = f(a)$. Dans ces coordonnées, f est donnée par un système

$$y_j = f_j(x), \ j = 1, \ldots, n$$

[19] Une application différentiable $f : M \to N$ est un difféomorphisme si c'est un isomorphisme pour la structure différentiable, *i.e.* une bijection différentiable dont l'inverse est également différentiable.

et l'on a

$$\xi(h\circ f)=\sum_{i=1}^{m}\ \xi_i\ \frac{\partial h(f(x))}{\partial x_i}=\sum_{i=1}^{m}\ \xi_i\ \sum_{j=1}^{n}\ \frac{\partial h}{\partial y_j}\ \frac{\partial f_j\,(x)}{\partial x_i}\ ,\ i.e.$$

$$\xi(h\circ f)=\sum_{j=1}^{n}\ \left(\sum_{i=1}^{m}\ \xi_i\ \frac{\partial f_j(x)}{\partial x_i}\right)\ \frac{\partial h}{\partial y_j}=\sum_{j=1}^{n}\ \eta_j\ \frac{\partial h}{\partial y_j}\ .$$

On a donc, pour tout $h : N \to \mathbb{R}$, $\xi(h\circ f) = D_a f(\xi)(h) = \eta(h)$ et par suite $D_a f(\xi) = \eta$ avec

$$\eta_j = \sum_{i=1}^{m}\ \xi_i\ \frac{\partial f_j}{\partial x_i}\ . \tag{4}$$

La formule (4) montre que $D_a f$ est une application linéaire de l'espace vectoriel $T_a M$ dans l'espace vectoriel $T_{f(a)}N$ dont la matrice, dite *jacobien* de f en a, dans les bases $\left(\dfrac{\partial}{\partial x_i}\right)$ et $\left(\dfrac{\partial}{\partial y_j}\right)$ est la matrice $m\times n$

$$J=\begin{pmatrix}\dfrac{\partial f_1}{\partial x_1}\cdots\dfrac{\partial f_1}{\partial x_m}\\ \vdots\qquad\quad\vdots\\ \dfrac{\partial f_n}{\partial x_1}\cdots\dfrac{\partial f_n}{\partial x_m}\end{pmatrix}\ .$$

$D_a f$ est appelée *application linéaire tangente* à f en a. Lorsque a varie, on obtient un morphisme $Df : TM \to TN$ des fibrés tangents, c'est-à-dire une application différentiable

$$\begin{array}{ccc} TM & \overset{Df}{\to} & TN\\ \downarrow_{\pi_M} & & \downarrow_{\pi_N}\\ M & \underset{f}{\to} & N\end{array}$$

compatible aux projections π_M et π_N, linéaire sur les fibres et rendant commutatif le diagramme ci-dessus.

2.4 Jets et développements de Taylor

L'on peut généraliser aux applications différentiables la notion fondamentale de développement de Taylor bien connue pour les fonctions $f : \mathbb{R} \to \mathbb{R}$. Soit $f : \mathbb{R} \to \mathbb{R}$ de classe C^∞ (*i.e.* indéfiniment différentiable). On sait que la meilleure approximation locale de f en x par des polynômes de degré k est donnée par :

$$j^k f(x) = f(x) + hf'(x) + \ldots + h^k \frac{f^{(k)}(x)}{k!}$$

où $f^{(k)}$ est la $k^{\text{ème}}$ dérivée de f et où $k!$ est la factorielle de k égale au produit $1.2.3.\ldots.k$. La fonction f étant C^∞, on peut donc lui associer en tout point $x \in \mathbb{R}$ la série formelle $T_f(x)$, dite série de Taylor de f en x et donnée par

$$T_f(x) = f(x) + hf'(x) + \ldots + h^k \frac{f^{(k)}(x)}{k!} + \ldots$$

Pour un polynôme P de degré k, il est trivial de vérifier :

(i) que les séries de Taylor s'arrêtent toutes à l'ordre k, *i.e.* que $T_P(x) = j^k P(x)$;

(ii) que P est localement égal en tout point à sa série de Taylor, *i.e.* que $P(x+h) = T_P(x) = j^k P(x)$ pour h assez petit;

(iii) que P est égal en tout point à l'une quelconque de ses séries de Taylor, *i.e.* que l'approximation à l'ordre k est non seulement une égalité locale mais également une égalité globale.

Il s'agit là d'une propriété fondamentale des polynômes : la connaissance d'un polynôme P et de ses dérivées en *un* point, *i.e.* sa connaissance locale, détermine sa connaissance globale.

Il en va tout autrement pour les fonctions C^∞.

(i) Lorsqu'une fonction C^∞ f est égale partout à la somme de sa série de Taylor $T_f(x)$ en un point, on dit que f est *entière*. Mais une fonction C^∞ n'est pas entière en général.

(ii) Lorsque f est égale au voisinage de chaque point $x \in \mathbb{R}$ à la somme de sa série de Taylor $T_f(x)$ en ce point, on dit que f est *analytique*. Mais une fonction C^∞ n'est pas analytique en général.

(iii) En général f ne sera même pas localement représentable par sa série de Taylor et cela pour deux raisons : soit $T_f(x)$ sera une série divergente et ne représentera donc aucune fonction, soit $T_f(x)$ sera une série convergente dont la somme sera différente de f.

On peut résumer ces caractères fondamentaux des fonctions C^∞ en disant qu'au niveau différentiable il n'y a pas de solidarité entre le local et le global.

Soient alors $f : M \to N$ et $a \in M$. Soient $(x_1, \ldots, x_m)$ un système de coordonnées locales en a et $(y_1, \ldots, y_n)$ un système de coordonnées locales en $b = f(a)$. f équivaut localement à la donnée de n fonctions $y_i = f_i(x)$ des m variables $x = (x_j)$. L'on sait donc définir les dérivées partielles successives de f en a. Par exemple la dérivée d'ordre 1 est la matrice *jacobienne* $\left(\dfrac{\partial f_i}{\partial x_j}(a)\right)$. La

dérivée seconde est le système des $\left(\dfrac{\partial^2 f_i}{\partial x_j \partial x_k}(a)\right)$, etc. Cela semble permettre de définir le développement de Taylor de f en a. Mais, sauf pour le jacobien, les dérivées ne sont pas des entités intrinsèques. C'est pour pallier cette difficulté que Charles Ehresmann a introduit la notion de *jet*. Bien que les dérivées ne soient pas des entités intrinsèques, la propriété pour deux fonctions f et g d'avoir le même développement de Taylor en a jusqu'à un certain ordre k est en revanche une propriété invariante par changement de coordonnées locales. L'on peut donc définir l'équivalence de f et g en a à l'ordre k, la classe d'équivalence de f s'appelant jet d'ordre k de f en a et se notant $j^k f(a)$. On note $J^k(M,N)_{a,b}$ l'ensemble des jets d'ordre k des applications telles que $f(a) = b$. Les $J^k(M,N)_{a,b}$ sont des espaces vectoriels tous isomorphes qui varient différentiablement avec (a,b) et qui engendrent, lorsque (a,b) parcourt $M \times N$, un fibré localement trivial noté $J^k(M,N)$. Si $f \in C^\infty(M,N)$ on peut alors lui associer son jet d'ordre k, $j^k f$, qui est l'application différentiable

$$\begin{aligned} j^k f : M &\to J^k(M,N) \\ a &\mapsto j^k f(a) \end{aligned}$$

Notre objet d'étude sera donc l'espace fonctionnel $\mathcal{F} = C^\infty(M,N)$ des applications indéfiniment différentiables entre une variété différentiable M de dimension m et une variété différentiable N de dimension n. On cherchera d'abord à caractériser géométriquement la stabilité structurelle des éléments $f \in \mathcal{F}$, puis ensuite à décrire, lorsqu'elle existe, la géométrie stratifiée des déploiements universels des f instables.

2.5 Stabilité structurelle, équivalence différentiable et détermination finie

Contrairement aux niveaux de structure analytique et algébrique, le niveau différentiable ne manifeste aucune solidarité entre le local et le global. Il est donc vain de chercher à classifier directement les éléments de $\mathcal{F}$. Hassler Whitney a par exemple montré que, pour tout fermé F de M, il existe une fonction différentiable $f : M \to \mathbb{R}$ dont F est l'ensemble des zéros ($F = f^{-1}(0)$). Mais si F n'est pas trivial, une telle fonction est infiniment instable. La propriété de stabilité structurelle rétablit par conséquent une certaine forme de solidarité entre le local et le global. D'où l'idée de chercher à classer d'abord les applications stables puis ensuite les applications présentant des "degrés" finis croissants d'instabilité jusqu'à en arriver, à la limite, au sous-ensemble de $\mathcal{F}$ constitué des applications (non classifiables) infiniment instables (par exemple les fonctions constantes sur des domaines entiers de M).

La notion de stabilité structurelle utilisée sera essentiellement celle associée à l'*équivalence différentiable*. On dit que $f, g : M \to N$ sont différentiablement équivalentes si il existe des difféomorphismes $\varphi \in \mathrm{Diff}(M)$ et $\psi \in \mathrm{Diff}(N)$ tels que $g \circ \varphi = \psi \circ f$, autrement dit si f et g ne diffèrent que par des

changements de coordonnées à la source et au but.[20] Parfois on utilisera une équivalence plus stricte en imposant $\psi = Id_N$ ou une équivalence plus faible comme l'équivalence topologique.

Cette notion d'équivalence permet de renforcer considérablement l'utilité du concept de développement de Taylor. En effet, même si f n'est pas représentée localement en a par sa série de Taylor $T_f(a)$, rien n'interdit qu'elle soit équivalente, localement en a, à l'un de ses jets $j^k f(a)$. On dit alors qu'elle est *déterminée à l'ordre* k en a. Cela signifie que la différence entre f et $j^k f$, y compris quand f n'est pas représentée par T_f, est résorbable dans des changements de coordonnées C^∞.

2.6 Localisation et germes

D'autre part, on sera souvent conduit à localiser les constructions et à considérer des *germes* d'applications. Pour définir le germe d'une application[21] $f : M \to N$ en $a \in M$, on introduit une équivalence entre les $f \in \mathcal{F}$ en posant que f et g sont équivalentes si elles coïncident sur un voisinage de a. Le germe $\eta_f(a)$ de f en a est alors la classe d'équivalence de f. Intuitivement, $\eta_f(a)$ exprime la structure "infinitésimale" de f en a.

2.7 Topologie de Whitney

Notons enfin que les notions de stabilité utilisées dépendent de la *topologie* choisie sur $\mathcal{F}$. Dans la mesure où l'on se propose de travailler au niveau différentiable, il est naturel de choisir la topologie de la convergence uniforme de f et de toutes ses dérivées : f et g sont voisines si elles sont uniformément voisines ainsi que toutes leurs dérivées. Dans le cas $\mathcal{F} = C^\infty(M, \mathbb{R})$ avec M *compacte*, cette topologie est facile à définir. Comme M est compacte, l'image $f(M)$ de M par f est compacte dans $\mathbb{R}$, donc bornée, et l'on peut définir la norme de f par $\|f\| = \max_{x \in M} |f(x)|$ (où $|\quad|$ est la valeur absolue sur $\mathbb{R}$). Comme $\mathcal{F}$ hérite de $\mathbb{R}$ une structure d'espace vectoriel, on peut donc définir la *distance* entre $f, g \in \mathcal{F}$ par $d(f, g) = \|f - g\|$. Comme toutes les dérivées partielles des $f \in \mathcal{F}$ sont des éléments de $\mathcal{F}$ cela permet de définir la relation de voisinage entre f et g par leur distance ainsi que par celles de leurs dérivées partielles.

Dans le cas général où M n'est pas compacte, la topologie "naturelle" sur $\mathcal{F}$ est plus difficile à définir. Il s'agit de la topologie dite de *Whitney* qui est une topologie de la convergence uniforme sur les compacts adaptée aux

[20] Il est trivial de vérifier que, les ensembles $\mathrm{Diff}(M)$ et $\mathrm{Diff}(N)$ des difféomorphismes de M et de N étant des groupes, la relation ainsi définie est bien une équivalence. C'est l'équivalence naturelle pour toutes les entités qui sont des morphismes de structure $f : M \to N$, $\mathrm{Diff}(M)$ et $\mathrm{Diff}(N)$ étant alors remplacés par les groupes d'automorphismes $\mathrm{Aut}(M)$ et $\mathrm{Aut}(N)$.

[21] Par la suite nous omettrons le qualificatif "différentiable" lorsque le contexte ne prêtera pas à confusion.

approximations successives d'une application par ses jets (qui sont des applications polynômiales de degré croissant). Soit k un ordre fixé. La topologie C^k de Whitney est la topologie de la convergence uniforme des $j^k f$ sur les compacts de M avec stationarité "à l'infini" : une suite f_n converge vers f si il existe un compact K de M tel que les $j^k f_n$ convergent uniformément vers $j^k f$ et tel que $f_n = f$ à l'extérieur de K sauf éventuellement pour un nombre *fini* de f_n (contrôle "à l'infini"). La topologie de Whitney est la topologie C^∞ limite des C^k-topologies. Elle est très fine et tout résultat sur la densité d'un sous-ensemble $\mathcal{H}$ de $\mathcal{F}$ est donc un résultat très fort.

Pour les propriétés de généricité que nous aborderons par la suite, le résultat suivant est préjudiciel.

Théorème 1. – Muni de la topologie de Whitney, l'espace fonctionnel $\mathcal{F} = C^\infty(M, N)$, est un espace de Baire, autrement dit toute intersection dénombrable d'ouverts denses est dense.[22]

Dans un espace de Baire $\mathcal{B}$ les sous-ensembles denses qui sont intersection dénombrables d'ouverts denses sont dits *résiduels*. Une propriété d'éléments de $\mathcal{B}$ est alors dite *générique* si elle est satisfaite sur un ensemble résiduel.

L'intérêt principal de la topologie de Whitney est que, munis de cette topologie, les espaces vectoriels fonctionnels que nous rencontrerons (espaces $C^\infty(M, \mathbb{R})$, espaces des champs de vecteurs sur M, etc.) seront, si M est compacte, des *espaces de Fréchet* et les espaces fonctionnels non vectoriels des *variétés de Fréchet*. Un espace vectoriel topologique est dit espace de Fréchet s'il est *complet* pour une métrique compatible avec sa structure vectorielle, les limites de suites devant converger (suites dites de Cauchy) existant effectivement. La complétude permet alors de définir la différentiabilité d'une application continue $f : \mathcal{F}_1 \to \mathcal{F}_2$ entre deux espaces de Fréchet exactement comme l'on définit celle d'une application continue $f : \mathbb{R}^m \to \mathbb{R}^n$. La notion de variété de Fréchet généralise dans ce contexte celle de variété différentiable. Il suffit de prendre pour cartes locales des ouverts d'espaces de Fréchet.

3 Trivialité locale et théorème des fonctions implicites

Pour analyser la structure d'une application $f : M \to N$ il faut d'abord disposer d'une idée de ce que l'on pourrait appeler la "bonne situation" et ensuite y comparer la situation réelle.

Soient donc $f : M \to N$ et $a \in M$. Considérons f au voisinage de a et de $f(a)$, c'est-à-dire la situation locale en $(a, f(a))$. Suivant les dimensions respectives m et n de M et de N, il est clair que les "bonnes situations" sont les suivantes :

[22] Pour une preuve, *cf.* Golubitsky, Guillemin [14], p. 44. On trouvera dans cet ouvrage la démonstration des propriétés fondamentales de la topologie de Whitney. Rappelons que si E est un espace topologique, un sous-ensemble U de E est dit *dense* si sa fermeture topologique $\overline{U}$ est égale à E tout entier. Cela signifie que tout point de E est approximable avec une précision aussi grande que l'on veut par des points de U.

(i) $m = n$, f est localement l'identité $\mathbb{R}^m \to \mathbb{R}^n$;

(ii) $m < n$, f est localement l'injection canonique $\mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^{n-m}$;

(iii) $m > n$, f est localement la projection canonique $\mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^{m-n} \to \mathbb{R}^n$.

Nous dirons que f est *localement triviale* si elle est, localement en $(a, f(a))$, différentiablement équivalente à la "bonne situation" qu'imposent les dimensions m et n.

Un théorème fondamental, dit *théorème des fonctions implicites*, affirme que la trivialité locale ne dépend que de l'application linéaire tangente $D_a f$ de f en a et est réalisée dès qu'elle peut l'être c'est-à-dire dès que $D_a f$, qui est une application linéaire du vectoriel $T_a M$ de dimension m dans le vectoriel $T_{f(a)} N$ de dimension n, *est de rang maximal* : m et n si $m = n$, m si $m < n$ et n si $m > n$.

Définition 2. – Soient $f : M \to N$ et $a \in M$. On dit que a est un point régulier *de f si l'application linéaire tangente $D_a f$ est de rang maximal.*

1. Considérons d'abord le cas $m = n$.

Théorème 3 (théorème d'inversion locale). – Si $m = n$ et si $a \in M$ est un point régulier de f alors f est, localement en a, inversible (i.e. est, localement en a, un difféomorphisme).[23]

Avant d'aborder les cas $m < n$ et $m > n$, donnons le corollaire fondamental du théorème d'inversion locale.

Soit $f : U_1 \times U_2 \to \mathbb{R}^n$ une application différentiable où U_1 est un ouvert de $\mathbb{R}^k$ et U_2 un ouvert de $\mathbb{R}^n$. Soient $x_0 \in U_1$ et $y_0 \in U_2$ et f_{x_0} l'application $f_{x_0} : U_2 \to \mathbb{R}^n$ donnée par $f_{x_0}(y) = f(x_0, y)$.

Théorème 4 (théorème des fonctions implicites). – Si $f(x_0, y_0) = y_0$ et si l'application linéaire tangente $D_{y_0} f_{x_0}$ de f_{x_0} en y_0 est de rang maximal n, alors, quitte à restreindre U_1 et U_2, il existe une application différentiable $g : U_1 \times U_2 \to U_2$ telle que $g(x_0, y_0) = y_0$ et $f(x, g(x, y)) = y$ pour tout $x \in U_1$ et tout $y \in U_2$.

Preuve. – Intuitivement, le théorème des fonctions implicites dit que si f est définie sur un produit $U_1 \times U_2$ et si sa dérivée à x constant est inversible alors on peut exprimer y en fonction de f. On le démontre en considérant l'application $F : U_1 \times U_2 \to \mathbb{R}^k \times \mathbb{R}^n$ définie par $F(x, y) = (x, f(x, y))$. La matrice jacobienne de F en (x, y) est la matrice $(k + n) \times (k + n)$:

$$\begin{pmatrix} I_k & 0 \\ D_x f_y & D_y f_x \end{pmatrix}$$

où I_k est la matrice identité $k \times k$. Comme $D_{y_0} f_{x_0}$ est de rang maximal n, DF est inversible en (x_0, y_0). D'après le théorème 3, F est donc localement

[23] Ce théorème est particulièrement fondamental car il affirme que ce qui est vrai infinitésimalement l'est aussi localement.

inversible au voisinage de son point fixe (x_0, y_0). Soit G son inverse, et soient g_1, g_2 les composantes de G. On a

$$(x, y) = F(G(x, y)) = F(g_1(x, y), g_2(x, y)) = (g_1(x, y), f(g_1(x, y), g_2(x, y)))\ .$$

Donc $x = g_1(x, y)$ et $y = f(x, g_2(x, y))$. Il suffit alors de prendre $g = g_2$. $\square$

2. Cas $m < n$.

Définition 5. – Supposons $m < n$ et soit a un point régulier de f (i.e. $D_a f$ est de rang maximal m). On dit alors que f est une immersion *en a. f est une immersion si c'est une immersion en tout point de M.*

Proposition 6. – Soit $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ une immersion en $a \in U$. Alors par changement de carte dans le but $\mathbb{R}^n$, f peut se ramener localement à l'injection canonique $\mathbb{R}^m \to \mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$ restreinte à U. Autrement dit, f est localement triviale et peut être linéarisée.

Preuve. – Soit en effet $D_a f$ l'application linéaire tangente de f en a. Par permutation de l'ordre des coordonnées, on peut supposer, puisque $D_a f$ est de rang maximal m, que le jacobien partiel $J = \left(\dfrac{\partial f_i}{\partial x_j}(a)\right)\ i, j = 1, \ldots, m$ est inversible. Soit $F : U \times \mathbb{R}^{n-m} \to \mathbb{R}^m \times \mathbb{R}^{n-m}$ l'application $F(x, y) = f(x) + (0, y)$. Sa matrice jacobienne en (x, y) est la matrice:

$$\begin{pmatrix} J & 0 \\ \frac{\partial f_{m+i}}{\partial x_j} & I_{n-m} \end{pmatrix}$$

Comme J est inversible, DF est inversible en $(a, 0)$ et donc F est un difféomorphisme au voisinage de $(a, 0)$ (que F envoie sur $f(a)$). Mais $f = F(x, 0)$. Comme $F^{-1}F = Id$, l'action de F^{-1} sur f linéarise f. Autrement dit, une immersion est localement triviale. $\square$

Localement, une immersion f est un *plongement*, c'est-à-dire intuitivement un isomorphisme de la variété source M sur une sous-variété de la variété but N. Mais elle n'est pas forcée de l'être globalement. D'abord elle peut ne pas être injective. Ensuite elle peut, bien qu'injective, être telle que la topologie induite par N sur son image $f(M)$ ne soit pas la même que celle induite par M. C'est par exemple le cas d'une droite de pente irrationnelle spiralant indéfiniment de façon dense sur le tore obtenu à partir du carré unité en identifiant les côtés opposés (mouvement quasi-périodique). On est donc conduit à introduire un concept plus restrictif. On appelle *plongement* une immersion injective rendant M homéomorphe à son image $f(M)$ munie de la topologie induite par N. Il est facile de montrer qu'une immersion injective *propre* (c'est-à-dire telle que l'image réciproque de tout compact de N soit un compact de M) est nécessairement un plongement. En particulier si M est compacte, toute application continue $f : M \to N$ est propre et donc toute immersion injective $f : M \to N$ est un plongement.

3. Cas $m > n$.

Définition 7. – Supposons $m > n$ et soit a un point régulier de f (i.e. $D_a f$ est de rang maximal n). On dit alors que f est une submersion *en a. f est une submersion si c'est une submersion en tout point de M.*

Proposition 8. – Soit $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ une submersion en $a \in U$. Alors par un changement de carte locale en a, f peut se ramener localement à la projection canonique $\mathbb{R}^{m-n} \times \mathbb{R}^n \to \mathbb{R}^n$ (restreinte à U). Autrement dit, f est localement triviale et peut être linéarisée.

Preuve. – Cette proposition est une conséquence immédiate du théorème des fonctions implicites. □

L'intérêt fondamental des submersions est de préserver par image réciproque la structure de variété des sous-variétés de leur but ainsi que leur codimension, *i.e.* la différence entre la dimension de la variété ambiante et celle de la sous-variété.

Proposition 9. – Soit $f : M \to N$ une submersion. Alors $f^{-1}(b)$ est une sous-variété de M de codimension n pour tout $b \in N$ (b est de dimension 0 et donc de codimension n dans N) et, plus généralement, si W est une sous-variété de N de codimension c, $f^{-1}(W)$ est une sous-variété de M de même codimension.

Preuve. – Le fait d'être une sous-variété étant une propriété locale, il suffit, d'après la proposition 8, de vérifier l'énoncé pour la projection canonique $\mathbb{R}^{m-n} \times \mathbb{R}^n \to \mathbb{R}^n$ ce qui est évident puisque $f^{-1}(W) = \mathbb{R}^{m-n} \times W$. □

On remarquera que, dans le cas où $M = \mathbb{R}^m$ et $N = \mathbb{R}^n$, les théorèmes de linéarisation des immersions et des submersions signifient que, au voisinage d'un point régulier, une application est différentiablement équivalente à sa partie linéaire $f(a) + D_a f(x - a)$: par changement de carte locale dans la source ou le but, on peut éliminer les termes du développement de Taylor de f en a de degré ≥ 2. Au voisinage d'un point régulier une application est donc *déterminée par son jet d'ordre 1.*

4 Le théorème de Sard

Au voisinage de ses points réguliers une application $f : M \to N$ est donc localement triviale. Mais elle ne l'est pas en général globalement et cela pour des raisons très différentes. Soit, bien que partout régulière, elle immerge M dans N sans la plonger, soit elle la plonge de façon complexe (cas par exemple des nœuds comme plongements non triviaux du cercle $\mathbb{S}^1$ dans $\mathbb{R}^3$) ou encore n'a pas la structure d'un produit (cas par exemple des fibrés vectoriels non triviaux sur N). Soit elle admet des points non réguliers, dits aussi *singuliers* ou *critiques*, dont les images s'appellent des valeurs singulières ou des valeurs critiques. La théorie des singularités s'intéresse avant tout à cette seconde possibilité.

Les singularités d'une application peuvent être extrêmement complexes. Mais, d'après un théorème remarquable dû à Sard, elles sont nécessairement rares.

Théorème 10 (Sard). – Soit $f : M \to N$ une application différentiable. Alors l'ensemble des valeurs critiques de f est de mesure nulle dans N et donc l'ensemble des valeurs régulières est dense *dans N.*

Il est facile de vérifier ce théorème lorsque $m < n$. En effet, une application ne peut augmenter la dimension de sa source et si $m < n$, $f(M)$ (bien qu'éventuellement fort complexe) sera de dimension au plus m dans N et donc de mesure nulle. Ce qui rend le théorème de Sard non trivial est le cas $m \geq n$.

Comme nous le verrons plus tard, le théorème de Sard, par sa généralité, est préjudiciel pour toute la théorie des singularités d'applications différentiables. Ainsi que l'a noté Alain Chenciner:

> "Ce lemme est le seul théorème de structure global applicable à toute fonction C^∞ (...); bien que d'énoncé peu géométrique (...), il est très porteur de géométrie : par l'intermédiaire des théorèmes de transversalité de René Thom dans les espaces de jets, il permet de faire de la géométrie sur presque toute fonction C^∞."[24]

Preuve. – L'idée de la démonstration est la suivante. Soit $C^f = C$ l'ensemble des points critiques de f. Il faut montrer que l'image $f(C)$ est de mesure nulle dans N. La degré de criticalité de f est mesuré par l'ordre jusqu'auquel les dérivées partielles des composantes $(f_1, \ldots, f_n)$ de f s'annulent. Soit C_i l'ensemble des points critiques où elles s'annulent jusqu'à l'ordre i. Les C_i constituent évidemment une suite décroissante. Une union finie d'ensembles de mesure nulle étant de mesure nulle, il suffit donc de montrer

(i) que $f(C - C_1)$ est de mesure nulle,

(ii) que $f(C_i - C_{i+1})$ est de mesure nulle,

(iii) qu'il existe un i tel que $f(C_i)$ est de mesure nulle.

Pour démontrer le point (i), considérons un point critique $a \in C - C_1$. Il existe par hypothèse une dérivée partielle de f en a qui est non nulle. Supposons que ce soit $\frac{\partial f_1}{\partial x_1}$et considérons l'application $h(x_1, \ldots, x_m) = (f_1(x), x_2, \ldots, x_m)$. La matrice jacobienne de h étant inversible, h est localement inversible. On peut donc remplacer f par $g = f \circ h^{-1}$. Or g possède la propriété que

$$g_1(y_1, \ldots, y_m) = f_1 \circ h^{-1}(y_1, \ldots, y_m) = y_1 \ ,$$

[24] Chenciner [10].

autrement dit g est l'identité sur la première composante. Mais cela implique par un raisonnement par récurrence que $g(C^g)$ soit de mesure nulle. Si $m = 1$, l'affirmation est triviale. Si $m \geq 2$, on considère l'application à $n-1$ composantes paramétrée par $t = y_1$, définie par $g_t(y') = (g_2(t, y'), \ldots, g_n(t, y'))$ avec $y' = (y_2, \ldots, y_m)$. Comme $g_1(y) = y_1$, y' est point critique de g_t si et seulement si (t, y') est point critique de g. Si donc i_t est l'injection $i_t : \mathbb{R}^{m-1} \to \mathbb{R}^m$ définie par $i_t(y') = (t, y')$, l'ensemble critique de g_t est $i_t^{-1}(C^g)$. Comme on est descendu en dimension $m-1$ on peut (hypothèse de récurrence) supposer que $g_t\left(i_t^{-1}(C^g)\right)$ est de mesure nulle pour tout t. Mais $g_t\left(i_t^{-1}(C^g)\right) = i_t^{-1}(g(C^g))$. Or il existe un théorème classique, le théorème de Fubini, disant que si dans un $\mathbb{R}^k = \mathbb{R} \times \mathbb{R}^{k-1}$ toutes les sections A_t d'un sous-ensemble $A \subset \mathbb{R}^k$ sont de mesure nulle dans $\{t\} \times \mathbb{R}^{k-1}$ alors A est lui-même de mesure nulle dans $\mathbb{R}^k$.

Pour démontrer le point (ii) on considère une dérivée partielle g d'ordre i telle que $\frac{\partial g}{\partial x_1}(a) \neq 0$ (il en existe une par hypothèse si $a \in C_i - C_{i+1}$). L'application $h(x) = (g(x), x_2, \ldots, x_m)$ est alors un difféomorphisme local en a. Par hypothèse, $g(C_i) = 0$ et donc $h(C_i) \subset \{0\} \times \mathbb{R}^{m-1}$ est $(m-1)$-dimensionnel. Soit alors $\ell : \mathbb{R}^{m-1} \to \mathbb{R}^m$ la restriction de $f \circ h^{-1}$ à $\{0\} \times \mathbb{R}^{m-1}$. L'image $f(C_i)$ est incluse dans $k\left(C^\ell\right)$. Par hypothèse de récurrence, $f(C_i)$ est donc de mesure nulle au voisinage de tout $a \in C_i - C_{i+1}$. Mais comme l'on peut extraire de tout recouvrement ouvert de $C_i - C_{i+1}$ un recouvrement dénombrable et qu'une union dénombrable d'ensembles de mesure nulle est encore de mesure nulle, $f(C_i - C_{i+1})$ est de mesure nulle.

Enfin, pour démontrer le point (iii) on utilise la propriété que si $x \in C_k$ alors $|f(y) - f(x)| \leq K|x-y|^{k+1}$. Cela permet d'inclure l'image $f(C_k)$ dans un volume V_δ dépendant de m, n, de k et du pas δ utilisé pour quadriller M. On montre alors que, pour k assez grand, $V_\delta \underset{\delta \to 0}{\to} 0$ et donc que $f(C_k)$ est alors de mesure nulle. $\square$

5 Les théorèmes de transversalité de Thom

Par définition, la propriété de stabilité structurelle est une propriété *ouverte*. Mais pour qu'elle constitue bien une détermination mathématique du concept correspondant, il faudrait encore que toute application instable soit "stabilisable" c'est-à-dire approximable par des applications stables. Mathématiquement, un tel réquisit s'exprime en disant que la propriété de stabilité structurelle doit être dense, et même *générique*.

Pour démontrer cette propriété de généricité, Thom a introduit l'idée profonde consistant à interpréter les phénomènes de stabilité en termes de situations de *transversalité* et à démontrer ensuite que, de façon très générale, les propriétés de transversalité sont génériques.

5.1 Le théorème général de transversalité

La notion de transversalité est très intuitive et, dans son intuition même, indissociable de celle de généricité. Elle précise la vieille notion géométrique de "position générale". Considérons par exemple deux courbes régulières γ_1 et γ_2 dans un plan $\mathbb{R}^2$ et soit x un de leurs points d'intersection. γ_1 et γ_2 s'intersectent transversalement en x si elles ne sont pas tangentes. Il est clair que cette propriété est générique dans la mesure où, si γ_1 et γ_2 sont tangentes en x, on peut, par petites déformations, faire "exploser" ce point de tangence en un certain nombre (éventuellement nul) d'intersections transversales. Mais supposons que γ_1 et γ_2 sont plongées dans $\mathbb{R}^3$. Le fait de s'intersecter transversalement n'est plus générique car, disposant d'une dimension supplémentaire, on peut, par petites déformations, disjoindre γ_1 et γ_2.

Soit alors N une variété différentiable de dimension n et soient N_1 et N_2 deux sous-variétés de dimensions respectives n_1 et n_2. N_1 et N_2 sont dites *transverses* si en tout point x de leur intersection $N_1 \cap N_2$ l'espace tangent T_xN en x à N est la somme des espaces tangents T_xN_1 et T_xN_2 : $T_xN = T_xN_1 + T_xN_2$. Cette définition implique :

(i) Si N_1 et N_2 sont disjointes, elles sont transverses.

(ii) Si $n_1 + n_2 < n$, N_1 et N_2 ne peuvent être transverses que si elles sont disjointes. Autrement dit, si l'on ne peut rejoindre la dimension de l'espace ambiant N à partir de celles des sous-espaces N_1 et N_2, la condition de transversalité implique que N_1 et N_2 "s'évitent".

(iii) Si $n_1 + n_2 = n$ (*i.e.* si les dimensions n_1 et n_2 sont complémentaires) alors N_1 et N_2 ne sont transverses que si tous leurs points d'intersection sont des points isolés d'intersection transversale.

(iv) Si $n_1 + n_2 > n$ la condition de transversalité signifie que la dimension de l'intersection $N_1 \cap N_2$ est égale à l'excès $n_1 + n_2 - n$ de $n_1 + n_2$ sur n et que, le long de cette intersection, l'intersection est transversale.

Si maintenant W est une sous-variété de N et si $f : M \to N$ est une application, on dit que f est transverse sur W (notation $f \pitchfork W$) si, pour tout $x \in M$ tel que $f(x) \in W$, on a $T_{f(x)}N = T_{f(x)}W + D_xf(T_xM)$. La transversalité implique en particulier que $f(M)$ "*évite*" toute sous-variété de N dont la codimension est strictement supérieure à la dimension de M.

Une des propriétés les plus importantes de la transversalité est de préserver par image réciproque la structure de variété des sous-variétés du but d'une application ainsi que leur codimension.

Proposition 11. – Soit $f : M \to N$ une application transverse sur une sous-variété W de N. Alors $f^{-1}(W)$ est une sous-variété de M de même codimension que W.

Preuve. – Ce résultat généralise la proposition 9 sur les submersions. En fait, celle-ci en est un corollaire immédiat puisqu'il est clair qu'une submersion est par définition transverse sur *toute* sous-variété de N. La démonstration se ramène d'ailleurs à un cas particulier de la proposition 9. Soit en effet $a \in M$ tel que $f(a) \in W$. Localement en $f(a)$, on peut identifier N à un produit direct $W \times L$ et donc considérer la submersion $\varphi : W \times L \to L$. Il est facile de voir que $f \pitchfork W$ en a si et seulement si $\varphi \circ f$ est une submersion en a. Mais $f^{-1}(W) = (\varphi \circ f)^{-1}(f(a))$ et est donc d'après la proposition 9 une sous-variété de M de codimension égale à la codimension de a dans L, *i.e.* égale à la dimension de L, elle-même égale à la codimension de W. □

Le théorème fondamental de transversalité de Thom dit que la transversalité dans les espaces de jets est générique. Il s'applique en fait aux sous-variétés *stratifiées* des espaces de jets. Nous ne définissons pas ici rigoureusement le concept, pourtant des plus centraux, de stratification. Informellement, un espace E est stratifié s'il est la réunion d'un ensemble de "bons" sous-espaces "réguliers", appelés strates, chaque strate admettant pour bord un ensemble de strates de dimension inférieure. On impose des conditions de finitude et des bonnes propriétés d'incidence des strates entre elles. L'intuition sous-jacente est, par exemple en dimension 3, celle d'un espace "cloisonné" en domaines, les domaines étant séparés par des "cloisons" se rejoignant le long "d'arêtes" se rejoignant elles-mêmes en des "sommets".[25]

Théorème 12 (théorème de transversalité). – Soit W une sous-variété stratifiée de l'espace des jets $J^k(M,N)$. L'ensemble T_W des $f \in \mathcal{F} = C^\infty(M,N)$ dont le k-jet $j^k f$ est transverse sur W est résiduel.

Preuve. – Pour démontrer ce théorème, on peut raisonner strate par strate et supposer que W est une sous-variété (non stratifiée) de $J^k(M,N)$. L'on recouvre ensuite W par une famille dénombrable d'ouverts $(W_i)_{i\in I}$ dont la fermeture $\overline{W}_i \subset W$ est *compacte* et dont les projections sur $M \times N$ sont incluses dans des produits $U_i \times V_i$ de cartes locales d'adhérences $\overline{U}_i$ compactes.[26] Soit T_i l'ensemble des $f \in \mathcal{F}$ dont le k-jet $j^k f$ est transverse sur W aux points de $\overline{W}_i$. Il est clair que $T_W = \bigcap_{i\in I} T_i$. Pour montrer que T_W est résiduel, il suffit donc de montrer que les T_i sont des ouverts denses de $\mathcal{F}$. Montrons d'abord que T_i est *ouvert*.

Lemme 1. – Soit W une sous-variété fermée de $J^k(M,N)$. Alors l'ensemble des $f \in \mathcal{F}$ dont le k-jet $j^k f$ évite W (i.e. $j^k f(M) \cap W = \varnothing$) est ouvert pour la topologie de Whitney.

A partir de ce lemme on peut montrer l'ouverture des T_i. Si en effet $\dim M < \operatorname{codim} W$, la transversalité équivaut à $j^k f(M) \cap \overline{W}_i = \varnothing$ et l'ouverture

[25] *Cf.* Thom [51], §3.2.

[26] Les $\overline{W}_i$ sont en général en nombre dénombrable non fini car en général W n'est ni ouverte ni fermée.

de T_i découle directement du lemme. Si $\dim M \geq \operatorname{codim} W$, on considère l'application $j^1(j^k f) : M \to J^1(M, J^k(M,N))$. On a

$$j^1(j^k f)(x) = (x, j^k f(x), D_x j^k f) \ .$$

Or la transversalité de $j^k f$ sur W signifie que si $z = j^k f(x) \in W$ alors $T_z J^k(M,N) = T_z W + D_x j^k f(T_x M)$. Elle équivaut donc à la disjonction de $j^1(j^k f)$ et de la sous-variété F_i des points de $J^1(M, J^k(M,N))$ se projetant sur $\overline{W}_i$ et tels que les dérivées partielles du jet en ce point adjointes à une base de l'espace tangent à W en ce point constituent une matrice qui n'est pas de rang maximum. Or F_i est fermée. D'après le lemme 1, l'ensemble $\mathcal{U}_i$ des $g : M \to J^k$ telles que $j^1 g$ évite F_i est donc ouvert dans $C^\infty(M, J^k)$. Mais l'application $j^k : C^\infty(M,N) \to C^\infty(M, J^k)$ est continue pour la topologie de Whitney. Donc $T_i = (j^k)^{-1}(\mathcal{U}_i)$ est ouvert dans $\mathcal{F} = C^\infty(M,N)$.

Montrons maintenant la *densité* des T_i. Soit $f \in \mathcal{F}$. Dans tout voisinage de f il doit exister des éléments de T_i. L'idée est de considérer des *déformations* bien choisies f_p de f paramétrées par des espaces externes P et d'utiliser la conséquence fondamentale de la transversalité et du théorème de Sard qu'exprime le lemme suivant.

Lemme 2. – Soit $F : P \times M \to N$ une déformation de f, $F(p,x) = f_p(x)$. Soit $\Phi : P \times M \to J^k(M,N)$ la déformation des jets, $\Phi(p,x) = j^k f_p(x)$. Si Φ est transverse sur une sous-variété W de $J^k(M,N)$, alors l'ensemble des $p \in P$ tels que $j^k f_p$ soit transverse sur W est dense dans P.

Ce lemme découle lui-même d'un autre.

Lemme 3. – Si une famille d'applications paramétrées par un espace P est transverse sur une sous-variété W alors, pour un ensemble dense des valeurs des paramètres, les applications individuelles intersectent aussi W transversalement.

Ce lemme découle directement du théorème de Sard. Il est la forme fonctionnelle du

Lemme 4. – Si une sous-variété produit $V \times P$ de N est transverse sur W, alors l'ensemble des $p \in P$ tels que $V_p = V \times \{p\}$ soit transverse sur W est dense dans P.

En effet, soit π la projection canonique $\pi : V \times P \to P$. Les V_p sont les "fibres" de π. Si les dimensions $d_v + d_p$ de $V \times P$ et d_w de W sont complémentaires ($d_v + d_p + d_w = n$) alors $V \times P$ et W s'intersectent transversalement en un nombre au plus dénombrable de points isolés w_i. Soit $p_i = \pi(w_i)$. En dehors des p_i, V_p et W sont disjointes et il y a donc transversalité (*cf.* figure 2(a)). Si $d_v + d_p + d_w > n$ alors les p pour lesquels il existe un point de non transversalité entre V_p et W sont les valeurs critiques de la restriction π' de π à $(V \times P) \cap W$ (comme $V \times P$ et W sont transverses, $(V \times P) \cap W$ est une sous-variété de $V \times P$). D'après le théorème de Sard, l'ensemble de ces p est de mesure nulle dans P et son complémentaire est dense (*cf.* figure 2(b)).

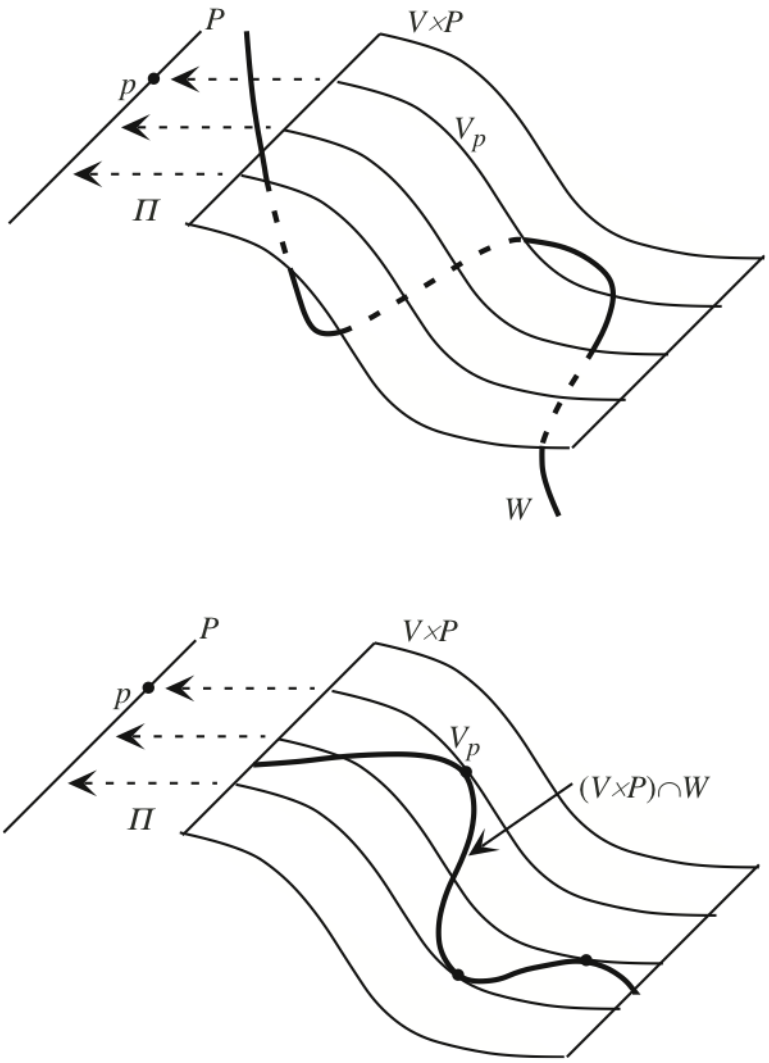


Figure 2: Illustration du lemme 4.

Pour démontrer le lemme 3 pour une famille d'applications $f_p : M \to N$ paramétrées par P telle que $F : M \times P \to N$, $F(x,p) = f_p(x)$, soit transverse sur W, on considère l'image réciproque $W_F = F^{-1}(W)$ de W par F (qui est une sous-variété de $M \times P$ par transversalité de F) et l'on montre que si $\dim W_F \geq \dim P$ (le cas $\dim W_F < \dim P$ étant trivial) alors les valeurs régulières de la restriction π' à W_F de la projection $\pi : M \times P \to P$ sont des valeurs de p pour lesquelles f_p est transverse sur W. Or d'après le théorème de Sard ces valeurs régulières sont denses dans P.

D'après ces lemmes, tout le problème devient donc de trouver de bonnes déformations f_p de f ayant la propriété que la déformation des jets $\Phi : P \times M \to J^k(M,N)$ est transverse sur $\overline{W}_i$. Soit $\mathcal{P}^k$ l'espace des polynômes $p : \mathbb{R}^m \to \mathbb{R}^n$ de degré k. Comme le problème est local, on peut supposer $M = \mathbb{R}^m$ et $N = \mathbb{R}^n$. Soit φ (resp. φ') une fonction égale à 0 à l'extérieur de U_i (resp. V_i) et égale à 1 sur la projection de $\overline{W}_i$ dans U_i (resp. dans V_i). $\overline{W}_i$ étant compact on peut trouver φ et φ' satisfaisant ces propriétés. On posera $f_p = f$ si $x \notin U_i$ ou si $f(x) \notin V_i$ et $f_p = \varphi(x)\varphi'(f(x))p(x) + f(x)$ si $x \in U_i$ et $f(x) \in V_i$. f_p est donc une perturbation polynômiale locale de f. Or, puisque les jets d'ordre k sont précisément les polynômes d'ordre k, on peut trouver un voisinage P de l'origine de $\mathcal{P}^k$ tel que l'application $\Phi : P \times M \to J^k(M,N)$ soit transverse sur W, tout simplement parce qu'elle est un difféomorphisme. $\square$

Par des méthodes analogues, on démontre le théorème suivant de "transversalité au but" destiné à tenir compte de la non injectivité de f. Soit $M^{(s)}$ le sous-ensemble des s-uples de points $(x^1, \ldots, x^s) \in M^s$ tels que $x^i \neq x^j$ pour $i \neq j$ (*i.e.* on considère s points distincts de M qui peuvent avoir même image par f). Soit $J^k_s(M,N)$ l'ensemble des s-uples de k-jets dont la projection sur M^s est $M^{(s)}$. $J^k_s(M,N)$ est un ouvert de $(J^k(M,N))^s$. Si $f \in \mathcal{F}$, on définit son *multi-jet* $j^k_s f : M^{(s)} \to J^k_s(M,N)$ par $j^k_s f(x^1, \ldots, x^s) = (j^k f(x^1), \ldots, j^k f(x^s))$.

Théorème 13. – Soit W une sous-variété stratifiée de l'espace des multi-jets $J^k_s(M,N)$. L'ensemble T_W des $f \in \mathcal{F}$ dont le multi-jet j^k_s est transverse sur W est résiduel.

Les théorèmes de transversalité de Thom ont un nombre considérable de conséquences. Citons-en quelques unes de tout à fait immédiates qui sont de simples conséquences de l'arithmétique des dimensions.

5.2 Les théorèmes d'immersion et de plongement de Whitney

Soit $f : M \to N$. Exprimons en termes de transversalité dans un espace de jets le fait qu'elle soit une immersion. Dire que f est une immersion en $x \in M$, c'est dire, d'après la définition 5, que $D_x f$ est de rang maximal $m \leq n$. Soit $J^1(M,N)$ le fibré des 1-jets. Sa fibre F est l'espace des matrices jacobiennes $D_x f$ et est donc de dimension mn. Quant à $J^1(M,N)$, il est de dimension $m + n + mn$. Soit H le sous-espace de F composé des matrices A qui ne sont

pas de rang maximal. H est défini par les conditions que tous les mineurs $m \times m$ de A s'annulent et il est donc de codimension $r = n - m + 1$ dans F. Lorsque (x, y) varie dans $M \times N$, H décrit un sous-fibré S de codimension r de $J^1(M \times N)$ et f est une immersion si et seulement si $j^1 f(M)$ n'intersecte pas S. Mais si $m < n - m + 1$, *i.e.* si $2m \leq n$, cette condition est une condition de transversalité et en appliquant le théorème 12 de transversalité, on obtient:

Corollaire 14 (théorème d'immersion de Whitney). – Si $n \geq 2m$, les applications $f : M \to N$ sont génériquement des immersions.

Exprimons maintenant en termes de transversalité le fait que f soit *injective*. Si f n'est pas injective, il existe $x, x' \in M$, $x \neq x'$ tels que $f(x) = f(x')$. Cela signifie que $j_2^0 f : M^{(2)} \to J_2^0(M, N)$ intersecte la sous-variété $W = M^{(2)} \times \Delta$ de $J_2^0(M, N) = M^{(2)} \times N^2$, où Δ est la diagonale de N^2. Or Δ étant de codimension n dans N^2, W est de codimension n dans $J_2^0(M, N)$. Comme $\dim M^{(2)} = 2m$, si $2m < n$, le fait que f soit injective équivaut à la transversalité de $j_2^0 f$ sur W. D'après le théorème 13 de transversalité au but, si $n \geq 2m+1$, alors f est génériquement injective. Comme d'après le corollaire 14 elle est aussi génériquement une immersion, on obtient le :

Corollaire 15 (théorème des immersions injectives). – Si $n \geq 2m + 1$, les applications $f : M \to N$ sont génériquement des immersions injectives.

A partir de ces corollaires, on montre facilement que si $n \geq 2m + 1$, les *plongements* $f : M \to N$ forment un ouvert dense de l'ensemble des applications *propres* de M dans N. S'il existe des applications propres (ce qui est trivialement le cas si M est compacte) il existe alors des plongements.

Corollaire 16 (théorème de plongement de Whitney). – Si M est compacte et si $n \geq 2m + 1$, il existe un plongement de M dans N et les applications $f : M \to N$ sont même génériquement des plongements.

Conséquences directes des théorèmes de transversalité, les théorèmes de Whitney montrent que si la dimension du but N est assez grande on peut toujours, par petite déformation, rendre immersive (et injective) n'importe quelle application. Les inégalités de dimensions qui y interviennent sont les meilleures possibles. Il est par exemple intuitif que si $m = 1$ et $n = 2m = 2$, une courbe du plan présentant une boucle avec deux branches s'intersectant transversalement (immersion non injective) ne peut pas devenir injective par déformation. En revanche si $n = 2m + 1 = 3$ alors on peut la désingulariser en écartant les branches dans la 3ème dimension et faire disparaître la non injectivité.

5.3 La théorie de Morse

Mais à l'autre extrême des relations entre les dimensions m et n, il y a le cas où n, loin d'être grand relativement à m, est au contraire minimal, *i.e.* égal à 1. Ce cas est celui de la théorie de Morse étudiant la structure des fonctions potentiel $f : M \to \mathbb{R}$.

Soit $f : M \to \mathbb{R}$. Le point $x \in M$ est un *point critique* de f si et seulement si la $(m \times 1)$-matrice jacobienne $\left(\dfrac{\partial f}{\partial x_1}, \ldots, \dfrac{\partial f}{\partial x_m}\right)$ est nulle en x, *i.e.* si $j^1 f(x)$ appartient à la sous-variété S_1 de $J^1(M, \mathbb{R})$ qui est sa section 0 (J^1 est de dimension $m + 1 + m = 2m + 1$, sa fibre est de dimension m, S_1 est de dimension $m+1$ et de codimension m). On dit qu'un point critique x est *non dégénéré* si le hessien de f en x – c'est-à-dire la matrice $m \times m$ des dérivées secondes $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(x)$ – est non dégénéré c'est-à-dire de rang maximal m. Or on peut facilement interpréter cette propriété en termes de transversalité.

Proposition 17. – Un point $x \in M$ est un point critique non dégénéré de f si et seulement si $j^1 f(x) \in S_1$ et $j^1 f$ est transverse sur S_1 en $j^1 f(x)$.

Preuve. – En termes de coordonnées locales $(x_1, \ldots, x_m)$ en x, $J^1(M, \mathbb{R})$ est de coordonnées $(x_1, \ldots, x_m, y, \xi_1, \ldots, \xi_m)$ et $j^1 f$ est l'application de M dans $J^1(M, \mathbb{R})$ qui à $x = (x_1, \ldots, x_n)$ associe

$$\left(x_1, \ldots, x_m, f(x), \frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_m}(x)\right).$$

Dire que x est un point critique de f, c'est dire que $\xi_1 = 0, \ldots, \ \xi_m = 0$. S_1 est de dimension $m + 1$ dans $J^1(M, \mathbb{R})$ qui est de dimension $2m + 1$. Quant à $j^1 f(M)$ c'est un sous-espace de dimension au plus m. Les dimensions $m + 1$ et m étant complémentaires pour $J^1(M, \mathbb{R})$, $j^1 f(M)$ ne peut intersecter S_1 transversalement qu'en des points isolés $z = j^1 f(x)$ où l'égalité

$$T_z S_1 + D(j^1 f)(x)(T_x M) = T_z J^1(M, \mathbb{R}) \tag{5}$$

est satisfaite. Soit $z = (x_1, \ldots, x_m, f(x), 0, \ldots, 0)$ un tel point et soient

$$(X_1, \ldots, X_m, Y, \Xi_1, \ldots, \Xi_m)$$

les coordonnées de $T_z J^1(M, \mathbb{R})$. L'application linéaire tangente en x de $j^1 f$ est l'application linéaire qui à un vecteur tangent $\alpha = (\alpha_1, \ldots, \alpha_m)$ de $T_x M$ associe le vecteur tangent $\left(\alpha, \sum_{i=1}^{i=m} \alpha_i \frac{\partial f}{\partial x_i}(x), H\alpha\right)$ de $T_z J^1(M, \mathbb{R})$ où H est le hessien de f en x. Or $T_z S_1$ correspond aux $m + 1$ premières coordonnées $(X_1, \ldots, X_m, Y)$ de $T_z J^1(M, \mathbb{R})$. Pour que l'égalité (6) puisse être satisfaite, il faut donc que $D(j^1 f)(x)(T_x M)$ "occupe" toutes les autres dimensions, *i.e.* que l'espace des $H\alpha$ lorsque α varie dans $T_x M$ soit de dimension maximale m. Mais cela signifie précisément que le hessien H de f en x est de rang maximal et que donc x est un point non dégénéré de f. □

D'après le théorème 12 de transversalité on a par conséquent :

Corollaire 18. – Soit $f : M \to \mathbb{R}$. Génériquement, tous les points critiques de f sont non dégénérés.

On appelle *fonction de Morse* les fonctions dont tous les points critiques sont non dégénérés. Toute fonction est donc transformable en fonction de Morse par petites déformations. Si f est de Morse, l'ensemble de ses points critiques $(j^1 f)^{-1}(S_1)$ est une sous-variété de codimension m de M d'après la proposition 11 et donc une sous-variété de *dimension* 0. Si M est compacte, les fonctions de Morse possèdent par conséquent un nombre *fini* de points critiques non dégénérés *isolés* .

Considérons maintenant les valeurs critiques de $f \in \mathcal{F} = C^\infty(M, \mathbb{R})$ et exprimons en termes de transversalité le fait qu'elles soient toutes distinctes. Pour cela, considérons $j^1_2 f : M^{(2)} \to J^1_2(M, \mathbb{R})$. Soit S la sous-variété de $J^1_2(M, \mathbb{R})$ constituée des couples de 1-jets (d'origines différentes) appartenant à S_1 et de buts différents. S est de codimension $2m+1$ dans $J^1_2(M, \mathbb{R})$. Comme $\dim M^{(2)} = 2m$, la transversalité de $j^1_2 f$ sur S équivaut à $j^1_2(f)(M^{(2)}) \cap S = \varnothing$. Mais cela signifie exactement que toutes les valeurs critiques de f sont distinctes. En appliquant le théorème 13 de transversalité au but on obtient donc :

Corollaire 19. – Génériquement, une fonction $f : M \to \mathbb{R}$ est une fonction de Morse dont toutes les valeurs critiques sont distinctes. On appelle une telle fonction une fonction de Morse excellente.

On voit bien sur cet exemple l'extrême puissance des théorèmes de transversalité. Toute fonction $f : M \to \mathbb{R}$ est approximable par une fonction possédant comme seules singularités des points critiques non dégénérés de valeurs critiques toutes distinctes, *i.e.* dont le graphe $G(f) \subset M \times \mathbb{R} \to M$ est un "relief" au-dessus de M avec ses sommets, ses bassins et ses cols. Qui plus est, les propriétés de transversalité dans les espaces de jets sont stables par déformations de f et donc si $a \in M$ est un point critique non dégénéré de f, *i.e.* un point d'intersection transversale de $j^1 f(M)$ et de S_1, pour toute fonction $g : M \to \mathbb{R}$ assez voisine de f pour la topologie de Whitney, $j^1 g(M)$ admettra aussi un point d'intersection transversale avec S_1 voisin de a, et donc g admettra aussi un point critique non dégénéré voisin de a. Il y a par conséquent stabilité des points critiques non dégénérés.

Avant de conclure cette section, rappelons aussi qu'il existe une *forme normale* des points critiques non dégénérés.

Théorème 20 (Théorème de Morse). – Soit a un point critique non dégénéré de f. Il existe un système de coordonnées locales $(x_1, \ldots, x_m)$ en a tel que $f(x) = f(0) - (x_1^2 + \ldots + x_k^2) + x_{k+1}^2 + \ldots + x_m^2$ (où $0 = a$ et donc $f(0) = f(a)$). Qui plus est, le nombre k est défini intrinsèquement. Il s'appelle l'indice *de a.*

En raison de leur forme normale, les points critiques non dégénérés sont aussi dits *quadratiques*. Le théorème affirme qu'en un point singulier quadratique, une fonction $f : M \to \mathbb{R}$ est déterminée par son jet d'ordre 2. Il généralise le résultat selon lequel, en un point régulier, une fonction est déterminée par son jet d'ordre 1 (*i.e.* est linéarisable par changement de coordonnées) :

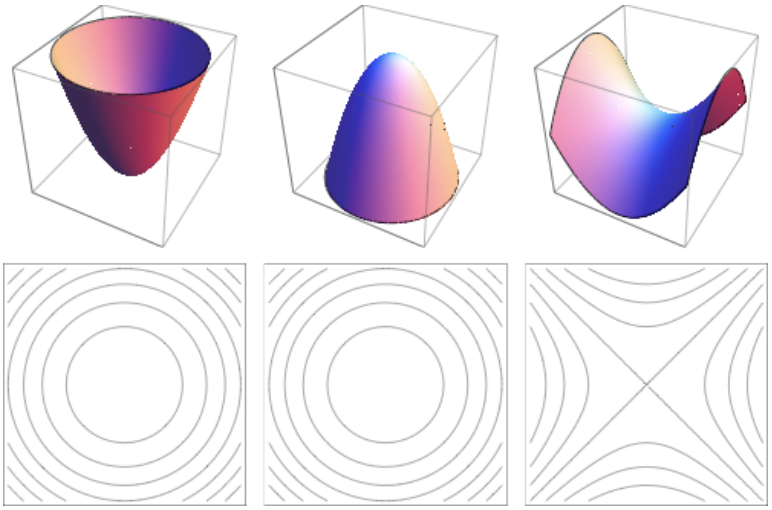


Figure 3: Divers types de points critiques non dégénérés : bassin, sommet et col, avec leurs lignes de niveau.

les points critiques non dégénérés sont les points singuliers les plus “ réguliers” possibles. Il se démontre en montrant qu'il existe un changement de coordonnées locales éliminant tous les termes du développement de Taylor de f en a de degré > 2. Si (x) est un tel système alors $f(x) = f(0) + x^t\, Hx$ [27] où H est le hessien de f en a. Il suffit dès lors, par un changement linéaire de coordonnées, de ramener la forme quadratique $x^t Hx$ à ses axes principaux pour ramener f à sa forme normale de Morse. Les 3 types de points quadratiques de corang 2 sont représentés à la figure 3.

Le théorème 20 des formes normales est un cas particulier d'un résultat plus général dû à Jean-Claude Tougeron et Pierre Samuel. Notons $\mathcal{E}_m$ l'espace des germes en 0 d'applications $f : \mathbb{R}^m \to \mathbb{R}$. Cet $\mathcal{E}_m$ hérite la structure additive et multiplicative du but $\mathbb{R}$ et est donc un anneau commutatif ayant pour élément unité le germe constant 1. On peut facilement montrer que le seul idéal maximal de $\mathcal{E}_m$ est l'idéal $\mathfrak{m}_m$ engendré par les germes des fonctions coordonnées $(x_1, \ldots, x_m)$, germes que l'on note encore $(x_1, \ldots, x_m)$.[28] Il est clair que $\mathfrak{m}_m$ est l'idéal des germes η tels que $\eta(0) = 0$ puisqu'une fonction $f : U \to \mathbb{R}$ s'annule à l'origine si et seulement si elle s'écrit sous la forme $f(x) = \sum_{i=1}^m x_i g_i(x)$. On voit que si $\eta \in \mathcal{E}_m$, la suite de ses images dans les quotients successifs $\mathcal{E}_m/\mathfrak{m}_m^k$ n'est rien d'autre que la suite de ses jets, *i.e.* la

[27] x est ici le vecteur colonne $\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ et son transposé x^t le vecteur ligne $(x_1, \ldots, x_m)$.

[28] D'une façon générale, on appelle anneau *local* un anneau ne possédant qu'un seul idéal maximal. Un morphisme d'anneaux $\varphi : \mathfrak{a} \to \mathfrak{b}$ d'un anneau local $\mathfrak{a}$ d'idéal maximal $\mathfrak{m}$ dans un anneau local $\mathfrak{b}$ d'idéal maximal $\mathfrak{n}$ est dit local si $\varphi(\mathfrak{m}) \subset \mathfrak{n}$.

suite des tronquages à l'ordre k de sa série de Taylor.

Si $\eta \in \mathcal{E}_m$, on peut lui associer l'idéal jacobien $J(\eta) = \left(\frac{\partial \eta}{\partial x_1}, \ldots, \frac{\partial \eta}{\partial x_m}\right)$ engendré par ses dérivées partielles du premier ordre (*i.e.* par les germes des dérivées partielles d'un représentant quelconque $f : U \to \mathbb{R}$ de η). On montre que 0 est un point critique non dégénéré de η si et seulement si $J(\eta) = \mathfrak{m}_m$. En effet dire que 0 est un point critique, c'est dire que le gradient en 0 $\left(\frac{\partial \eta}{\partial x_i}(0)\right) = 0$, et donc que $J(\eta) \subset \mathfrak{m}_m$. Dire que 0 est un point critique non dégénéré, c'est dire que le hessien en 0 $\left(\frac{\partial^2 \eta}{\partial x_i \partial x_j}(0)\right)$ est de rang maximal m. D'après le théorème 3 d'inversion locale, cela signifie que l'application $F : \mathbb{R}^m \to \mathbb{R}^m$ associant au point $(x_1, \ldots, x_m)$ le point $\left(\frac{\partial \eta}{\partial x_1}, \ldots, \frac{\partial \eta}{\partial x_m}\right)$ est un difféomorphisme local en 0. Or cela équivaut à $J(\eta) = \mathfrak{m}_m$ puisque cela équivaut à dire que les $\frac{\partial \eta}{\partial x_i}$ sont des coordonnées locales en 0.

Mais le théorème de Tougeron et Samuel[29] dit que si $\eta, \xi \in \mathcal{E}_m$ et si $\xi - \eta \in \mathfrak{m}_m J(\eta)^2$ alors η et ξ sont équivalents à gauche (*i.e.* uniquement par l'action d'un changement de coordonnées locales à la source). Donc si 0 est un point critique non dégénéré de η et si ξ est la partie quadratique de η, $\xi - \eta \in \mathfrak{m}_m^3$ et donc, puisque $J(\eta) = \mathfrak{m}_m$, $\xi - \eta \in \mathfrak{m}_m J(\eta)^2$ et par conséquent η et ξ sont équivalents. Ce que dit précisément le théorème 20.

6 Les divers types de stabilité

Comme nous l'avons noté plusieurs fois, un des points essentiels du programme de Thom est de caractériser la propriété de stabilité structurelle. La définition générale "$f : M \to N$ est structurellement stable si et seulement si pour toute application $g : M \to N$ assez voisine de f il existe $\varphi \in \mathrm{Diff}(M)$ et $\psi \in \mathrm{Diff}(N)$ tels que $\psi \circ f = g \circ \varphi$" n'est pas en effet de manipulation facile. La stratégie consistera donc à définir des notions plus maniables de stabilité et à montrer qu'elles sont équivalentes à celle de stabilité structurelle. Cette stratégie se heurte à une foule de difficultés de haute technicité. Elle est dominée par les travaux de René Thom, Vladimir Arnold et John Mather[30] et nous en exposerons les rudiments en suivant l'ouvrage d'introduction déjà cité de Martin Golubitsky et Victor Guillemin.

6.1 La stabilité infinitésimale

Soient $\mathcal{F} = C^\infty(M, N)$ et $G = \mathrm{Diff}(M) \times \mathrm{Diff}(N)$. La stabilité structurelle des éléments de $\mathcal{F}$ est associée à l'action du groupe G sur $\mathcal{F}$ donnée par

[29] Tougeron [54].

[30] *Cf.* Thom [46], [47], [48], [49], [50], Levine [16], Mather [20], [21], [22], [23], [24], [25], [26], Chenciner [9], Arnold [1], [2], [3].

$(\varphi, \psi)f = \psi \circ f \circ \varphi^{-1}$, action dont les orbites sont les types différentiables. La classe d'équivalence différentiable $\tilde{f}$ de f est l'image de G par l'application $\gamma_f : G \to \mathcal{F}$ définie par $\gamma_f(\varphi, \psi) = (\varphi, \psi)f = \psi \circ f \circ \varphi^{-1}$.

Toutes les difficultés techniques viennent du fait que G et $\mathcal{F}$ sont des "mauvais" espaces fonctionnels et que la situation est donc beaucoup plus complexe que celle rencontrée lorsqu'un groupe de Lie agit différentiablement sur une variété différentiable.[31] Mais, conceptuellement, la situation est bien du même type et il est naturel de s'inspirer d'abord du cas de dimension finie pour ensuite essayer de s'y ramener.

Soient $f \in \mathcal{F}$ et $\tilde{f}$ son orbite image de G par γ_f. Si nous appliquons purement conceptuellement les notions valides en dimension finie, nous sommes conduits à considérer "l'application linéaire tangente" $D_e\gamma_f$ de γ_f à l'origine $e = (1_M, 1_N)$ de G.[32] Comme $\gamma_f(e) = f$, on voit que $D_e\gamma_f$ envoie "l'espace tangent" T_eG à G en e sur "l'espace tangent" $T_f\,\tilde{f}$ à $\tilde{f}$ en f. Or, dire que f est structurellement stable, c'est dire que $\tilde{f}$ est un voisinage de f, et par suite que γ_f est une "submersion" en e, ou encore que $D_e\gamma_f$ envoie T_eG *sur* $T_f\mathcal{F}$.

Mais qu'est-ce qu'un "vecteur tangent" en f à $\mathcal{F}$? C'est "une tendance infinitésimale" de déformation de f, *i.e.* la donnée pour tout $f(x)$ d'un vecteur tangent $X(x)$ à N en $f(x)$.

Définition 21. – Un champ de vecteurs sur N le long de $f : M \to N$ *est une application différentiable* $X : M \to TN$ *telle que le diagramme*

$$\begin{array}{ccc} & & TN \\ & X \nearrow & \downarrow \pi_N \\ M & \underset{f}{\to} & N \end{array}$$

soit commutatif (i.e. $X(x)$ *est un vecteur tangent à* N *en* $f(x)$*).*

Si l'on note $C_f^\infty(M, TN)$ l'espace fonctionnel des champs de vecteurs sur N le long de f, on a $C_f^\infty(M, TN) = T_f\mathcal{F}$. $T_f\mathcal{F}$ est aussi l'espace $\Gamma\left(f^*(TN)\right)$ (noté aussi $\Gamma(f^*TN)$ pour simplifier) des sections différentiables du fibré $f^*(TN) \to M$ image réciproque par f du fibré $TN \to N$. D'autre part, les vecteurs tangents à G en e sont les couples (R, S) d'un champ de vecteurs R sur M et d'un champ de vecteurs S sur N. Autrement dit, T_eG est la somme directe $T_eG = \mathcal{X}(M) \oplus \mathcal{X}(N)$, ou encore $T_eG = \Gamma(TM) \oplus \Gamma(TN)$ puisque, si M est une variété, l'espace $\mathcal{X}(M)$ de ses champs de vecteurs est par définition celui des sections différentiables du fibré $TM \to M$, espace noté habituellement $\Gamma(TM)$.

$D_e\gamma_f$ opère de la façon suivante sur T_eG.

1. Soit $R \in \Gamma(TM)$. On a $D_e\gamma_f(R) = Df \circ R \in \Gamma(f^*TN)$:

[31] Cette situation de dimension finie est déjà hautement non triviale puisque, lorsque G est le tout simple groupe additif $\mathbb{R}$, elle correspond à la théorie des systèmes dynamiques.

[32] 1_M et 1_N sont les applications identité de M et de N.

$$\begin{array}{ccc} TM & \overset{Df}{\to} & TN \\ R\uparrow\downarrow & \nearrow & \downarrow \\ M & \underset{f}{\to} & N \end{array}$$

Notons $\theta_f : \Gamma(TM) \to \Gamma(f^*TN)$ cette application.

2. Soit $S \in \Gamma(TN)$. On a $D_e\gamma_f(S) = S \circ f \in \Gamma(f^*TN)$:

$$\begin{array}{ccc} TM & \overset{Df}{\to} & TN \\ \downarrow & \nearrow & \downarrow\uparrow S \\ M & \underset{f}{\to} & N \end{array}$$

Notons $\Omega_f : \Gamma(TN) \to \Gamma(f^*TN)$ cette application.

Donc, si $(R, S) \in T_eG$, on a $D_e\gamma_f(R, S) = \theta_f(R) + \Omega_f(S)$. $D_e\gamma_f$ est donc surjective si et seulement si $\theta_f + \Omega_f$ l'est.

Définition 22. – Une application $f : M \to N$ est dite infinitésimalement stable *si l'application $\theta_f + \Omega_f : \Gamma(TM) \oplus \Gamma(TN) \to \Gamma(f^*TN)$ est surjective.*

Le résultat fondamental sur la notion de stabilité infinitésimale est que, si M est compacte (ou si f est propre), elle est équivalente à celle, a priori beaucoup plus forte, de stabilité structurelle. C'est le fameux théorème de Mather. Si f est structurellement stable il est clair qu'elle est infinitésimalement stable. C'est donc la réciproque qui fait problème.

Pour montrer à quel point cette notion de stabilité infinitésimale est maniable, donnons quelques exemples où l'on suppose M compacte.

1. Si $f : M \to N$ est une submersion alors elle est infinitésimalement stable – et donc structurellement stable d'après le théorème de Mather. En effet $D_xf : T_xM \to T_{f(x)}N$ est surjective pour tout $x \in M$. On peut montrer que son noyau K_x est de dimension localement constante et que, lorsque x varie dans M, il engendre un sous-fibré K de TM. Soit H un sous-fibré supplémentaire de K dans TM. Pour tout $x \in M$, $D_xf : H_x \to T_{f(x)}N$ est un *isomorphisme*. Si donc X est un champ de vecteurs sur N le long de f, on peut lui associer le champ de vecteurs R sur M défini par $R(x) = (D_xf)^{-1}(X(x)) \in H_x$. Clairement $\theta_f(R) = X$ et donc θ_f est surjective. *A fortiori*, $\theta_f + \Omega_f$ est surjective.

2. Si $f : M \to N$ est une immersion injective alors elle est infinitésimalement stable – et donc structurellement stable d'après le théorème de Mather. En effet un champ de vecteurs X sur N le long de f s'identifie à un champ sur $f(M)$. Or un tel champ peut être prolongé à N et donc Ω_f est surjective. *A fortiori* $\theta_f + \Omega_f$ est surjective. Ce résultat implique que si $n \geq 2m + 1$, alors $f : M \to N$ est stable si et seulement si c'est une immersion injective. En effet si f est stable, comme les immersions injectives sont génériques pour

$n \geq 2m+1$ d'après le corollaire 15, elle est équivalente à une immersion injective et est donc une immersion injective. Réciproquement, si f est une immersion injective elle est stable. On voit bien sur cet exemple la nécessité de l'hypothèse de compacité. Un mouvement quasi-périodique sur un tore est une immersion injective de source $M = \mathbb{R}$ non compacte et est extrêmement *instable*.

3. Soit $f : M \to \mathbb{R}$ une fonction structurellement stable. Alors c'est une fonction de Morse excellente. En effet, d'après le corollaire 19, les fonctions de Morse excellentes sont génériques et il en existe donc dans tout voisinage de f. La fonction f étant structurellement stable, toute fonction assez voisine lui est équivalente et f est donc équivalente à une fonction de Morse excellente. Mais une fonction équivalente à une fonction de Morse excellente est elle-même une fonction de Morse excellente.

Réciproquement, toute fonction de Morse excellente est infinitésimalement stable – et donc structurellement stable d'après le théorème de Mather.

Théorème 23 (Morse). – Si M est compacte, $f : M \to \mathbb{R}$ est structurellement stable si et seulement si c'est une fonction de Morse excellente. D'après le corollaire 19, la stabilité structurelle est donc générique *pour les fonctions. Comme elle est qui plus est ouverte par définition, elle est par suite ouverte et dense.*

Ce théorème est le premier que nous rencontrons qui caractérise explicitement des entités structurellement stables. On le démontre de la façon suivante.

Preuve.[33] – Soient $f : M \to \mathbb{R}$ une fonction de Morse excellente et $X : M \to T\mathbb{R} \simeq \mathbb{R}\times\mathbb{R}$ un champ de vecteurs sur $\mathbb{R}$ le long de f. X équivaut à la donnée d'un couple (f, ξ) de fonctions $M \to \mathbb{R}$, $\xi(x)$ étant le vecteur $X(x)$ d'origine $f(x)$. Soit R un champ de vecteurs sur M. $\theta_f(R)$ est alors le champ de vecteurs sur $\mathbb{R}$ le long de f défini par $\xi = L_R f$ où L_R est l'opérateur de dérivation associé à R. Soit S un champ de vecteurs sur $\mathbb{R}$, *i.e.* une application $S : \mathbb{R} \to \mathbb{R}$. $\Omega_f(S)$ est alors le champ de vecteurs sur $\mathbb{R}$ le long de f défini par $\xi = S \circ f$. Dire que f est infinitésimalement stable, c'est donc dire que toute fonction $\xi : M \to \mathbb{R}$ peut s'écrire comme une somme $\xi = L_R f + S \circ f$ pour un champ R sur M et une application $S : \mathbb{R} \to \mathbb{R}$. Comme f est une fonction de Morse excellente sur une variété compacte, elle n'admet qu'un nombre fini $c_1, \ldots, c_n$ de points critiques non dégénérés dont toutes les valeurs $v_1, \ldots, v_n$ sont distinctes. On peut donc choisir S de façon à ce que $\xi(c_i) = S \circ f(c_i)$ pour $i = 1, \ldots, n$. La fonction $\xi' = \xi - S \circ f$ est alors une fonction s'annulant sur les points critiques de f et il suffit de montrer que ξ' peut toujours s'écrire sous la forme $\xi' = L_R f$ avec un champ R construit à partir de f et de ξ'.

Choisissons autour de chaque point régulier x de f, une carte U_x où Df ne s'annule pas et autour de chaque point critique c_i une carte U_i où f s'écrit sous forme normale. Comme M est compacte, on peut extraire de ce recouvrement

[33] *Cf.* Golubitsky-Guillemin [14], p. 79.

ouvert de M un sous-recouvrement fini $U_1, \dots U_k$ de M correspondant à un nombre fini de points (réguliers ou critiques) $a_1, \dots, a_k$. Soit $g_1, \dots, g_k$ une *partition de l'unité* subordonnée à ce recouvrement (il en existe toujours), c'est-à-dire un ensemble de fonctions $g_1, \dots, g_k : M \to \mathbb{R}$ telles que :

(i) le support de g_i (*i.e.* la fermeture de l'ensemble des x tels que $g_i(x) \neq 0$) soit inclus dans U_i,

(ii) $\sum_{i=1}^k g_i(x) = 1$ pour tout $x \in M$.

Si a_i est régulier, choisissons un champ local w_i sur U_i tel que $Df(w_i) \neq 0$ sur U_i et associons lui le champ global R_i bien défini sur M par :

$$\begin{cases} R_i = \dfrac{\xi' g_i w_i}{L_{w_i} f} \text{ sur } U_i, \\ R_i = 0 \text{ sur } M - U_i. \end{cases}$$

Si a_i est critique, alors $\xi'(a_i) = 0$ et donc $g_i \xi'$ peut s'écrire sur U_i sous la forme $g_i \xi' = \sum_{j=1}^m x_j h_j$. Considérons le champ global R_i bien défini sur M par :

$$\begin{cases} R_i = \displaystyle\sum_{j=1}^m \frac{\varepsilon_j h_j}{2} \frac{\partial}{\partial x_j} \text{ (où } \varepsilon_j \text{ est le signe de } x_j \text{ dans la forme normale de } f \\ \qquad \text{au voisinage de } a_i) \text{ sur } U_i \\ R_i = 0 \text{ sur } M - U_i. \end{cases}$$

Le champ somme $R = \sum_{i=1}^k R_i$ satisfait à $L_R f = \xi'$. En effet, si a_i est régulier, on a :

$$\begin{cases} L_{R_i} f = \xi' \, g_i \text{ sur } U_i \ , \\ L_{R_i} f = 0 \text{ sur } M - U_i \ , \end{cases}$$

soit $L_{R_i} f = \xi' g_i$. Si a_i est critique, on a:

$$\begin{cases} L_{R_i} f = \displaystyle\sum_{j=1}^m \frac{\varepsilon_j h_j}{2} \frac{\partial}{\partial x_j} (\textstyle\sum_{k=1}^m \varepsilon_k x_k^2) \text{ sur } U_i \ , \\ \qquad = \displaystyle\sum_{j=1}^m h_j x_j = \xi' g_i \text{ sur } U_i \ , \\ \qquad = 0 \text{ sur } M - U_i \ , \end{cases}$$

soit $L_{R_i} f = \xi' \, g_i$. Donc $L_R f = \sum_{i=1}^k L_{R_i} f = \sum_{i=1}^k \xi' g_i = \xi'$. □

4. Par des techniques du même ordre, on peut montrer qu'une immersion $f : M \to N$ est stable si et seulement si elle est *à croisements normaux* c'est-à-dire si et seulement si, pour chaque s-uple $(x^1, \dots, x^s)$ de points de M tel que $f(x^1) = \dots = f(x^s)$, l'application $f^s : M^{(s)} \to N^s$ est *transverse* sur la diagonale $\Delta N^s = \{(y, \dots, y) \in N^s, y \in N\}$.

5. Nous avons vu avec l'exemple 1 que les submersions sont stables. Mais si M est compacte il n'existe aucune submersion $f : M \to \mathbb{R}$ (puisque $f(M)$ étant

compact admet un minimum et un maximum). On est donc conduit à raffiner le concept de submersion pour tenir compte de ces cas importants. Soient M et N telles que $m \geq n$ et soit $k = \dim M - \dim N$ la dimension des fibres régulières de $f : M \to N$. Soit S_1 la sous-variété de $J^1(M, N)$ constituée des 1-jets dont la matrice jacobienne au lieu d'être de rang maximal n (*i.e.* de corang 0) est de rang $n-1$ (*i.e.* de corang 1). Soit $f : M \to N$ une application telle que $j^1 f$ soit transverse sur S_1. Alors $S_1(f) = (j^1 f)^{-1}(S_1)$ est une sous-variété de M de codimension égale à la codimension de S_1 dans $J^1(M, N)$, *i.e.* égale à $k+1$. On dit que $x \in S_1(f)$ est un *pli* de f si $T_x M = T_x S_1(f) + \operatorname{Ker} D_x f$, *i.e.* si $\operatorname{Ker} D_x f$ et $T_x S_1(f)$ sont transverses (en effet comme $\dim(\operatorname{Ker} D_x f) = k+1$, ils sont de dimensions complémentaires). On dit que f est une *submersion avec plis* si $j^1 f$ est transverse sur S_1 et si tous ses points critiques sont des plis. On montre alors que si f est une submersion avec plis elle est (infinitésimalement) stable si et seulement si la restriction $f\,|_{S_1(f)}$ de f à son ensemble singulier $S_1(f)$ est une immersion à croisements normaux.

6.2 La stabilité infinitésimale locale

La notion de stabilité infinitésimale peut évidemment se localiser. Soient $f : M \to N$, $a \in M$ et $b = f(a) \in N$. Soit η le germe de f en a. L'application f est localement infinitésimalement stable en a si η est infinitésimalement stable, c'est-à-dire si et seulement si, pour tout germe χ en a de champ de vecteurs sur N le long de f, il existe un germe ρ en a de champ de vecteurs sur M et un germe σ en b de champ de vecteurs sur N tels que

$$\chi = (D_a f)\,\rho + \sigma \circ \eta.$$

Si $(x_1, \ldots, x_m)$ et $(y_1, \ldots, y_n)$ sont des systèmes de coordonnées locales en a et b, si $\chi(x) = \sum_{i=1}^n \chi_i(x) \dfrac{\partial}{\partial y_i}$, $\rho = \sum_{j=1}^m \rho_j \dfrac{\partial}{\partial x_j}$ et $\sigma = \sum_{i=1}^n \sigma_i \dfrac{\partial}{\partial y_i}$, on doit donc pouvoir résoudre au voisinage de $x = 0$ et $y = 0$, les équations :

$$\chi_i(x) = \sum_{j=1}^m \rho_j \frac{\partial f}{\partial x_j} + \sigma_i(f_1, \ldots, f_n), \quad i = 1, \ldots, n. \tag{6}$$

6.3 La stabilité par déformations

Notre présentation intuitive du paradigme morphodynamique, nous a montré l'importance des déploiements ou des déformations d'une application sur un espace externe T, c'est-à-dire des familles d'applications f_t paramétrées par un espace pointé $(T, 0)$ et telles que $f_0 = f$.

Définition 24. – Soit T un voisinage de l'origine de $\mathbb{R}^k$. Une application différentiable $F : M \times T \to N \times T$ est dite une k-déformation *de f si (i) elle est compatible aux projections $M \times T \to T$ et $N \times T \to T$ (i.e. si F envoie $M \times \{t\}$ dans $N \times \{t\}$) et équivaut par conséquent à une famille f_t donnée*

par $F(x,t) = (f_t(x), t)$*, et (ii) si* $f_0 = f$*. Une* k*-déformation est* triviale *s'il existe des difféomorphismes* $\varphi : M \times U \to M \times U$ *et* $\Psi : N \times U \to N \times U$ *déformant respectivement* 1_M *et* 1_N *sur un voisinage de* 0*,* $U \subset T$*, et tels que le diagramme :*

$$\begin{array}{ccc} M \times U & \xrightarrow{F} & N \times U \\ \varphi \downarrow & & \downarrow \psi \\ M \times U & \underset{f \times 1_U}{\to} & N \times U \end{array}$$

soit commutatif.

On dit alors que f est *stable par* k*-déformations* si toute k -déformation de f est triviale. La stabilité par 1-déformations s'appelle aussi *stabilité homotopique*. La notion de stabilité par k-déformations a été introduite par René Thom et Harold Levine.

6.4 La stabilité transversale

Les éléments de théorie de la stabilité que nous avons déjà rencontrés montrent qu'il faut chercher à caractériser la stabilité par des propriétés de transversalité dans des espaces de jets. Il est donc naturel d'introduire une notion de stabilité directement liée à de telles propriétés et de chercher à montrer ensuite qu'elle équivaut à la stabilité structurelle.

Soit $J^k(M,N)$ un espace de jets. Le groupe $G = \operatorname{Diff} M \times \operatorname{Diff} N$ opère évidemment sur J^k à travers les k-jets de ses éléments. Soit j un élément de $J^k(M,N)_{a,b}$. Si $(\varphi,\psi) \in G$ alors (φ,ψ) opère sur j par $(\varphi,\psi)(j) = j^k\psi(b) \circ j \circ j^k(\varphi^{-1})(\varphi(a))$ selon le diagramme :

$$\begin{array}{ccccc} M \xrightarrow{f} & N & a & \xmapsto{f} & f(a) = b \\ \varphi \downarrow & \downarrow \psi & \downarrow & & \downarrow \\ M \to & N & \varphi(a) & \mapsto & \psi(b) \end{array}$$

Par réduction aux k-jets, l'action de G sur $J^k(M,N)$ devient *algébrique* ce qui entraîne que les orbites $\tilde{j}$ soient des *sous-variétés* de $J^k(M,N)$.

Définition 25. – Soient $f : M \to N$ *et* $a \in M$*. On dit que* f *est* localement transversalement stable en a *si son* n*-jet* $j^n f$ *(où* $n = \dim N$*) est transverse sur* $\widetilde{j^n f(a)}$*.*

On peut évidemment donner une version multi-jets de cette notion. G opère aussi sur les s-uples $(j_1 \ldots, j_s) \in J^k_s(M,N)$ et, si $j \in J^k_s(M,N)$, son orbite $\tilde{j}$ est également une sous-variété de $J^k_s(M,N)$.

Définition 26. – Soit $f : M \to N$*. On dit que* f *est* transversalement stable *si pour tout multi-jet* $j = (j_1, \ldots, j_s) \in J^n_s(M,N)$ *tel que* $1 \leq s \leq n+1$ *et* $j_1, \ldots, j_s$ *soient de même but* $b \in N$*, le* s*-multi-jet* $j^n_s f$ *est transverse sur* $\tilde{j}$*.*

7 L'équivalence des diverses stabilités et le théorème de Thom-Mather

Le résultat central de la théorie de la stabilité de Thom-Mather est que les diverses notions de stabilité introduites précédemment sont équivalentes.

Théorème 27 (théorème de stabilité de Thom-Mather). – Soient M compacte et $f : M \to N$. Alors les conditions suivantes sont équivalentes.

(i) *f est structurellement stable,*

(ii) *f est infinitésimalement stable,*

(iii) *f est stable par déformations,*

(iv) *f est homotopiquement stable,*

(v) *f est transversalement stable.*

7.1 Etapes de la preuve

La preuve de ce théorème est assez labyrinthique et beaucoup trop technique pour être présentée ici.[34] Elle repose sur un théorème fondamental dû à Bernard Malgrange et connu sous le nom de *théorème de préparation différentiable* (théorème 39 ci-dessous). Ses étapes successives sont les suivantes.

1. On montre d'abord, à l'aide du théorème de Malgrange, que la stabilité infinitésimale locale peut se ramener à une condition sur les jets, *i.e.* peut être descendue en dimension finie.

2. On montre ensuite que la stabilité infinitésimale est une version multi-jet de la stabilité infinitésimale locale.

3. On montre ensuite que si f est infinitésimalement stable alors toute g assez voisine de f est localement infinitésimalement stable.

4. On montre ensuite que si f est infinitésimalement stable et stable par k-déformations pour k assez grand alors toute g assez voisine de f est infinitésimalement stable.

5. On montre ensuite que si f est stable par déformations et si toute g assez voisine de f est aussi stable par déformations alors f est structurellement stable.

6. Cela implique que si la stabilité infinitésimale est équivalente à la stabilité par déformations, alors elle est équivalente à la stabilité structurelle. En effet, d'après (4), f admet un voisinage composé de g infinitésimalement stables donc stables par déformations. Comme f est stable par déformations, elle est structurellement stable d'après (5).

[34] *Cf.* les articles de Mather cités en bibliographie et Golubitsky-Guillemin [14].

7. On caractérise ensuite les déformations triviales.

8. Cela permet de montrer que la stabilité par déformations implique la stabilité infinitésimale.

9. On montre ensuite la réciproque.

10. On montre enfin que la stabilité transversale est équivalente à la stabilité structurelle.

Esquissons quelques éléments de cette preuve.

7.2 Finitude de la stabilité infinitésimale locale

Pour demontrer que f est localement infinitésimalement stable en $a \in M$, il faut, nous l'avons vu, pouvoir résoudre localement les équations (6)

$$\chi_i(x) = \sum_{j=1}^{m} \rho_j \frac{\partial f}{\partial x_j} + \sigma_i(f_1, \ldots, f_n), \quad i = 1, \ldots, n$$

où $\chi(x) = \sum_{i=1}^{n} \chi_i(x) \dfrac{\partial}{\partial y_i}$ est un germe en a de champ de vecteurs sur N le long de f, où $\rho = \sum_{j=1}^{m} \rho_j \dfrac{\partial}{\partial x_j}$ est un germe de champ de vecteurs sur M et où $\sigma = \sum_{i=1}^{n} \sigma_i \dfrac{\partial}{\partial y_i}$ est un germe en $b = f(a)$ de champ de vecteurs sur N.

Le germe χ est défini par ses composantes χ_i et donc l'espace des χ, $\Gamma(f^*TN)_a$ (germes en a), que nous noterons $\mathcal{T}$, est isomorphe à la somme directe de n exemplaires de $\mathcal{E}_m$. C'est donc un $\mathcal{E}_m$-module libre de type fini. Soit $\mathcal{A}$ le sous-module des $(D_x f)\rho$ pour $\rho \in \mathcal{X}_a(M) = \Gamma(TM)_a$, *i.e.* l'image de $\Gamma(TM)_a$ par l'application θ_f définie plus haut. Pour montrer (6), à savoir que

$$\theta_f + \Omega_f : \Gamma(TM)_a \oplus \Gamma(TN)_b \to \Gamma(f^*TN)_a$$

est surjective, l'idée est de quotienter par $\mathcal{A}$ et de montrer que Ω_f devient surjective. Soit donc $\mathcal{B} = \mathcal{T}/\mathcal{A}$ le quotient de $\mathcal{T}$ par $\mathcal{A}$. C'est un $\mathcal{E}_m$-module de type fini. Mais comme f est une application de M dans N, elle induit un morphisme d'anneaux locaux

$$f^* : C_b^\infty(N) = \mathcal{E}_n \to C_a^\infty(M) = \mathcal{E}_m$$

et, à travers f^*, $\mathcal{B}$ devient un $\mathcal{E}_n$-module. Soient $e_1, \ldots, e_n$ les projections dans $\mathcal{B}$ des $f^* \dfrac{\partial}{\partial y_i}$. D'après (6), f est infinitésimalement stable en a si et seulement si les e_i engendrent $\mathcal{B}$ comme $\mathcal{E}_n$-module. Or pour ce faire, il suffit que cela soit vrai modulo $\mathfrak{m}_m^{n+1}$. En effet, un corollaire du théorème de préparation différentiable est la proposition suivante :

Proposition 28. – Soit $\mathcal{B}$ un $\mathcal{E}_m$-module de type fini. Soient $e_1, \ldots, e_k$ des éléments de $\mathcal{B}$. Alors ils engendrent $\mathcal{B}$ comme $\mathcal{E}_n$-module si et seulement si ils engendrent $\mathcal{B}/\mathfrak{m}_m^{k+1}\mathcal{B}$ comme $\mathcal{E}_n$-module.

Autrement dit, pour résoudre les équations (6), il suffit de les résoudre jusqu'à l'ordre n, *i.e.* il suffit de résoudre les équations

$$\chi_i(x) = \sum_{j=1}^{m} \rho_j \frac{\partial f}{\partial x_j} + \sigma_i(f_1, \ldots, f_n) + h_i, \quad i = 1, \ldots, n \tag{7}$$

où les h_i sont des fonctions différentiables dont la série de Taylor commence à l'ordre $n+1$.

D'après le théorème de préparation différentiable (théorème 39 ci-dessous), la stabilité infinitésimale locale de f en a ne dépend ainsi que du jet $j^{n+1}f(a)$. On peut donc "descendre" cette notion en dimension *finie* dans les espaces de jets :

Corollaire 29. – f est localement infinitésimalement stable en a si et seulement si $J^n(f^*TN)_a = (D_a f) J^n(TM)_a + f^* J^n(TN)_b$.

Nous sommes ainsi passés, par réduction à la dimension finie, d'espaces fonctionnels compliqués à de "bonnes" variétés algébriques.

7.3 Caractérisation de la stabilité infinitésimale

Pour caractériser la stabilité infinitésimale en termes de stabilité infinitésimale locale, on montre par des techniques analogues que si $S = \{a_1, \ldots, a_n\}$ est un ensemble fini de points de même image b par f (non injectivité), f doit être simultanément localement infinitésimalement stable aux points de S.

Proposition 30. – f est infinitésimalement stable si et seulement si, pour tout $b \in N$ et tout $S \subset f^{-1}(b)$ de cardinal $\leq n+1$, on a : $J^n(f^*TN)_S = (Df) J^n(TM)_S + f^* J^n(TN)_b$.

7.4 Ouverture de la stabilité infinitésimale locale

Après avoir caractérisé ces types de stabilité, on montre qu'il s'agit de propriétés ouvertes.

Proposition 31. – Si f est infinitésimalement stable, alors toute application g assez voisine de f est localement infinitésimalement stable.

Preuve. – En effet, soit $a \in M$ et $b = f(a)$. Par hypothèse

$$J^n(f^*TN)_a = (D_a f) J^n(TM)_a + f^* J^n(TN)_b .$$

Mais l'application

$$\theta_f + \Omega_f = Df + f^* : J^n(TM)_a \oplus J^n(TN)_b \to J^n(f^*TN)_a$$

est une application linéaire entre espaces vectoriels dépendant continûment de a et de $j^{n+1}f$. Or la surjectivité est une propriété stable des applications linéaires. Il existe donc un voisinage U_a de a dans M et un voisinage W_a de f dont tous les éléments sont localement infinitésimalement stables sur U_a.

Comme M est compacte, un nombre fini de U_a recouvrent M et l'intersection des W_a correspondants est un voisinage de f dont tous les éléments sont partout localement infinitésimalement stables. $\square$

7.5 Ouverture de la stabilité infinitésimale

Si dans la proposition 31 ci-dessus on voulait montrer que les g assez voisines de f sont *globalement* infinitésimalement stables, il faudrait utiliser le critère de la proposition 30. Mais celui-ci fait intervenir des variétés $M^{(s)}$ qui sont *non compactes* même si M est compacte. L'argument précédent n'est donc plus applicable tel quel. Cependant l'on peut montrer que si f est infinitésimalement stable alors toutes les applications g assez voisines le sont également à condition que f satisfasse la condition supplémentaire :

$(*)$ Pour chaque $a \in M$, il existe un voisinage U_a de a dans M et un voisinage W_a de f tels que pour tout $g \in W_a$ et pour tout $S = \{a_1, \ldots, a_s\} \subset g^{-1}(b) \cap U_a$ on ait :

$$J^n(g^*TN)_S = (Dg)\, J^n(TM)_S + g^* J^n(TN)_b \tag{8}$$

Proposition 32. – Si f est infinitésimalement stable et stable par k-déformations pour k assez grand, alors elle satisfait la condition $(*)$.

Preuve. – Sans entrer dans les (nombreux) détails techniques donnons néanmoins une idée de la preuve. Soit $a \in M$. Pour se localiser en a, on utilise une application $\lambda : M \to \mathbb{R}$ égale à 1 sur un compact K inclus dans un voisinage U_a de a dans M et égale à 0 sur $M - U_a$. On se restreint ensuite à un voisinage W_ε de f dont les éléments g sont à une distance de f qui est $< \varepsilon$, la distance $d(f,g) = \|f - g\|$ étant déduite de la norme $\|h\|$ qui est le maximum sur le compact K de la valeur absolue des composantes de h et de toutes ses dérivées partielles jusqu'à l'ordre $n+1$. On déforme alors f en une famille F sur la base de l'espace P des polynômes $p : \mathbb{R}^m \to \mathbb{R}^n$ de degré $\leq n$ en prenant $f_p = f + \lambda p$ (*cf.* la démonstration du théorème 12 de tranversalité). Comme f est stable par déformations par hypothèse, F est triviale sur un voisinage Z de l'origine de l'espace P. Il est alors aisé de montrer puisque $\lambda = 1$ sur V_a, que l'on peut choisir ε de façon à ce que, si $S = (a_1, \ldots, a_s) \in V_a^{(s)}$ avec $s \leq n+1$ et $V_a \subset K$ et si $g \in W_\varepsilon$, alors il existe $p \in Z$ tel que $j^{n+1}(g - f)(a_i) = j^{n+1}p(a_i)$, *i.e.* $j^{n+1}(g)(a_i) = j^{n+1}(f + p)(a_i)$, pour tous les a_i de S. Comme F est triviale sur Z et comme f est infinitésimalement stable, f_p l'est également et satisfait donc (8) d'après la proposition 30. Comme cette condition ne dépend que des $(n+1)$-jets sur S et que $j^{n+1}(g)(a_i) = j^{n+1}f_p(a_i)$ sur S, g satisfait (8) et f satisfait la condition $(*)$. $\square$

Corollaire 33. – Si f est infinitésimalement stable et stable par déformations, alors toute application assez voisine de f est également infinitésimalement stable.

7.6 Stabilité homotopique et stabilité structurelle

Pour arriver au nerf de la preuve que la stabilité infinitésimale implique la stabilité structurelle, il faut donc en dernier lieu montrer la :

Proposition 34. – Si $f : M \to N$ est homotopiquement stable (i.e. stable par 1-déformations) et admet un voisinage W constitué de g homotopiquement stables, alors f est structurellement stable.

Preuve. – On peut supposer que W est connexe par arcs, *i.e.* que pour tout $g \in W$ il existe une homotopie $F = (f_t) : M \times I \to N \times I$ (avec $I = [0, 1]$) telle que $f_0 = f$, $f_1 = g$ et $f_t \in W$ pour tout $t \in I$. Soient donc $g \in W$ et $F = (f_t)$ une homotopie de f à g. Comme par hypothèse f_t est homotopiquement stable pour tout $t \in I$, l'homotopie F est localement triviale. Mais comme I est connexe, elle globalement triviale et g est donc équivalente à f. □

7.7 Le critère de trivialité de Thom-Levine

Nous pouvons maintenant aborder la démonstration du fait que la stabilité infinitésimale équivaut à la stabilité par déformations et donc, d'après (5), qu'elle équivaut aussi à la stabilité structurelle. Pour ce faire, il faut disposer d'un critère de trivialité pour les déformations de f.

Soit $F = (f_t) : M \times T \to N \times T$ une k-déformation de f où T est un voisinage de l'origine de $\mathbb{R}^k$. Un vecteur tangent ς à $M \times T$ (resp. $N \times T$) en (x, t) (resp. (y, t)) est un couple (ξ, u) (resp. (η, v)) d'un vecteur tangent ξ à M en x et d'un vecteur tangent u à T en t (resp. d'un vecteur tangent η à N en y et d'un vecteur tangent v à T en t). Si l'on identifie ξ à $(\xi, 0)$ et u à $(0, u)$, on peut écrire $\zeta = \xi + u$.

Considérons alors les k champs de vecteurs sur $N \times T$ le long de F définis par :

$$\tau_i = (DF)\left(\frac{\partial}{\partial t_i}\right) - F^*\left(\frac{\partial}{\partial t_i}\right) \tag{9}$$

où le premier $\frac{\partial}{\partial t_i}$ est considéré sur le T de $M \times T$ et le second sur le T de $N \times T$. Ces champs "mesurent" en quelque sorte la dépendance de f_t par rapport à t. En effet, l'application linéaire tangente DF est donnée dans les bases $\left\{\frac{\partial}{\partial x_j}, \frac{\partial}{\partial t_i}\right\}$ de $T(M \times T)$ et $\left\{\frac{\partial}{\partial y_\ell}, \frac{\partial}{\partial t_i}\right\}$ de $T(M \times T)$ par la matrice $(m + k) \times (n + k)$

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_m} & \frac{\partial F_1}{\partial t_1} & \cdots & \frac{\partial F_1}{\partial t_k} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_m} & \frac{\partial F_n}{\partial t_1} & \cdots & \frac{\partial F_n}{\partial t_k} \\ \\ 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix}$$

et, par ailleurs, F agit comme l'identité sur T. Donc, pour $(x,t) \in M \times T$, $\tau_i(x,t)$ est le vecteur tangent à $N \times T$ en $F(x,t) = (f_t(x), t)$ donné par

$$\begin{aligned} \tau_i &= \left(\sum_{\ell=1}^{\ell=n} \frac{\partial f_{t,\ell}}{\partial t_i} \frac{\partial}{\partial y_i} \right) + \frac{\partial}{\partial t_i} - \frac{\partial}{\partial t_i} \\ &= \sum_{\ell=1}^{\ell=n} \frac{\partial f_{t,\ell}}{\partial t_i} \frac{\partial}{\partial y_i} \end{aligned} \tag{10}$$

Autrement dit, $F = (f_t)$ est égale à $f \times 1_T$ si et seulement si tous les champs τ_i sont identiquement nuls.

Dire que F est triviale, c'est dire que sur un voisinage U de 0 dans T, F est ramenable à $f \times 1_T$ par des difféomorphismes $\varphi_t \in \mathrm{Diff}(M)$ et $\psi_t \in \mathrm{Diff}(N)$ déformant 1_M et 1_N sur U. Cela signifie que $f_t = (\psi_t)^{-1} \circ f \circ \varphi_t$. Calculons donc les champs τ_i dans ce cas. Soit $(x,t) \in M \times T$ et considérons le vecteur tangent $\frac{\partial}{\partial t_i}$ en (x,t). Notons $\varphi \in \mathrm{Diff}(M \times U)$ et $\psi \in \mathrm{Diff}(N \times U)$ les familles (φ_t) et (ψ_t). Appliquons $D\varphi$ à $\frac{\partial}{\partial t_i}$. On obtient un vecteur tangent $(D\varphi)\frac{\partial}{\partial t_i}$ en $(\varphi_t(x), t) \in M \times T$. Prenons sa M-composante et appliquons-lui $D\varphi^{-1}$. On obtient ainsi un nouveau vecteur tangent ξ_i en (x,t) qui est "vertical", *i.e.* tangent à $M \times \{t\}$ (pas de T-composante). Soit maintenant $(y,t) \in N \times T$ et considérons le vecteur tangent $\frac{\partial}{\partial t_i}$ en $\psi(y,t) = (\psi_t(y), t) \in N \times T$. Appliquons-lui $D\psi^{-1}$. On obtient un vecteur tangent η_i en (y,t) qui est "vertical", *i.e.* tangent à $N \times \{t\}$.

Remarquons que si l'on applique DF au champ vertical ξ_i on obtient un champ vertical sur $N \times T$ puisque F étant une déformation elle envoie fibre sur fibre. Remarquons aussi que les champs τ_i sont verticaux puisque le terme $-F^*\left(\frac{\partial}{\partial t_i}\right)$ annule la T-composante de $(DF)\left(\frac{\partial}{\partial t_i}\right)$.Un calcul simple montre alors que l'on a :

$$\tau_i = (DF)(\xi_i) + F^*(\eta_i) \tag{11}$$

Or la décomposition (11) des τ_i est caractéristique des déformations triviales. On a en effet le critère de trivialité de Thom-Levine :

Théorème 35. – Une k-déformation F est triviale si et seulement si il existe un voisinage U de 0 dans T et des champs verticaux ξ_i sur $M \times U$ et η_i sur $N \times U$, $i = 1, \ldots, k$, satisfaisant $\tau_i = (DF)(\xi_i) + F^*(\eta_i)$.

Preuve. – Ce théorème se démontre en considérant un compact $K \subset U$, un voisinage compact de $F(M \times K)$ dans $N \times U$, en intégrant les champs ξ_i et η_i et en montrant qu'ils définissent des déformations φ_t et ψ_t de 1_M et 1_N trivialisant la k-déformation F. □

7.8 La stabilité homotopique implique la stabilité infinitésimale

Théorème 36. – Si $f : M \to N$ est stable par déformations alors elle est infinitésimalement stable.

Preuve. – Soit $f : M \to N$ une application stable par déformations. Pour montrer qu'elle est infinitésimalement stable on doit considérer un champ de vecteurs τ sur N le long de f et montrer qu'il existe un champ ξ sur M et un champ η sur N tels que

$$\tau = Df(\xi) + f^*(\eta) \tag{12}$$

Il faut donc arriver à déduire cette condition (12) d'une condition de type (11) $\tau_i = (DF)(\xi_i) + F^*(\eta_i)$ pour une déformation F bien choisie.

Soit $\mathrm{Gr}(f) \subset M \times N$ le graphe de f. Le champ τ est un champ sur $\mathrm{Gr}(f)$ "parallèle" à N, *i.e.* "vertical" relativement à la projection canonique $M \times N \to M$. Comme M est compacte, $\mathrm{Gr}(f)$ est plongé dans $M \times N$ et l'on peut donc étendre τ en un champ $\widetilde{\tau}$ sur $M \times N$ qui est vertical et à support compact. Soit θ_t le flot sur $M \times N$ de $\widetilde{\tau}$. Comme $\widetilde{\tau}$ est vertical, θ_t est en fait une famille de groupes à un paramètre $\theta_{x,t} : N_x \to N_x$ paramétrisée par $x \in M$ où $N_x = \{x\} \times N$ est la fibre de la projection canonique $M \times N \to M$ au-dessus de x. On définit alors une homotopie de f, $F = (f_t) : M \times \mathbb{R} \to N \times \mathbb{R}$ par $f_t(x) = \theta_{x,t}(f(x))$. Comme f est homotopiquement stable par hypothèse, il existe d'après le critère de Thom-Levine (théorème 35) un voisinage U de l'origine de $\mathbb{R}$, un champ vertical ξ_t sur $M \times U$ et un champ vertical η_t sur $N \times U$ tels que $\tau_t(f) = DF(\xi_t) + F^*(\eta_t)$. Mais τ_t est le champ vertical sur $M \times U$ le long de F défini par $\tau_t = \frac{\partial f_t}{\partial t}$. Or par construction de F, on a $\tau_0 = \tau$ et par conséquent $\tau = \tau_0 = Df(\xi_0) + f^*(\eta_0)$, ce qui est la condition (12). □

7.9 La stabilité infinitésimale implique la stabilité par déformations

La réciproque du théorème 36 est au cœur de la démonstration de Mather que la stabilité infinitésimale implique la stabilité structurelle. Elle exige une nouvelle application du théorème de préparation différentiable de Malgrange (théorème 39 ci-dessous).

Théorème 37. – Soit $f : M \to N$ une application infinitésimalement stable. Alors f est stable par k-déformations pour tout k.

Preuve. – Soit $f : M \to N$ une application infinitésimalement stable et considérons une k-déformation $F = (f_t) : M \times T \to N \times T$ de f. Pour démontrer la trivialité de F, il suffit d'après le critère de Thom-Levine de montrer qu'il existe un voisinage U de l'origine de T et des champs verticaux ξ_i sur $M \times U$ et η_i sur $N \times U$, $i = 1, \ldots, k$, tels que la condition (11), $\tau_i = (DF)(\xi_i) + F^*(\eta_i)$, $i = 1, \ldots, k$, soit satisfaite.

On localise d'abord la situation en se restreignant aux germes en $(a,0) \in M \times U$. Comme dans la démonstration de la finitude de la stabilité infinitésimale locale, on considère l'espace $\mathcal{T}$ des germes en $(a,0)$ des champs verticaux sur $M \times U$ le long de F ainsi que le sous-ensemble $\mathcal{A}$ de ces champs qui sont de la forme $DF(\xi)$ où ξ est un germe en $(a,0)$ de champ vertical sur $M \times U$. Si $\mathcal{E}_{m+1}$ (resp. $\mathcal{E}_{n+1}$) est l'anneau local des germes en $(a,0)$ de fonctions $g : M \times \mathbb{R} \to \mathbb{R}$ (resp. en $(b,0)$ de fonctions $h : N \times \mathbb{R} \to \mathbb{R}$), $\mathcal{T}$ est un $\mathcal{E}_{m+1}$-module de type fini engendré par les $F^*\left(\frac{\partial}{\partial y_i}\right)$, $i = 1, \ldots, n$, où $(y_1, \ldots, y_n)$ est un système e coordonnées locales en $b = f(a)$. Comme $\mathcal{A}$ est un sous-$\mathcal{E}_{m+1}$-module de $\mathcal{T}$, $\mathcal{B} = \mathcal{T}/\mathcal{A}$ est un $\mathcal{E}_{m+1}$-module de type fini qui, à travers F^*, devient un $\mathcal{E}_{n+1}$-module.

Lemme. – *$\mathcal{B}$ est un $\mathcal{E}_{n+1}$-module de type fini engendré par les images $e_1, \ldots, e_n$ dans $\mathcal{B}$ des générateurs $F^*\left(\frac{\partial}{\partial y_i}\right)$ de $\mathcal{T}$.*

Soit $\rho \in \mathcal{T}$. On a $\rho(x,t) = \rho_0(x) + \sum_{i=1}^{i=k} t_i \rho_i(x,t)$. Comme ρ est un champ sur $N \times U$ le long de F, ρ_0 est un champ sur N le long de f. Comme f est localement infinitésimalement stable par hypothèse, il existe des germes ξ_0 et η_0 de champs sur M et N tels que $\rho_0 = Df(\xi_0) + f^*(\eta_0)$. Soient ξ et η les extensions triviales de ξ_0 et η_0 à $M \times U$ et $N \times U$. On a :

$$\rho = DF(\xi) + F^*(\eta) + \sum_{i=1}^{i=k} t_i \rho_i' \ .$$

Considérons alors le quotient $\mathcal{B}/(t_1, \ldots, t_k)\mathcal{B} = \mathcal{B}_0$. L'image de ρ dans $\mathcal{B}_0$ est $F^*(\eta)$. Mais comme $\eta = (\eta_1, \ldots, \eta_n)$ est une extension triviale de η_0, $F^*(\eta) = \sum_{i=1}^{i=n} \eta_i f \frac{\partial}{\partial y_i}$. Autrement dit, les images des $F^*\left(\frac{\partial}{\partial y_i}\right)$ engendrent $\mathcal{B}_0$. D'autre part, soit $\mathfrak{m}_{n+1}$ l'idéal maximal de l'anneau local $\mathcal{E}_{n+1}$. Considérons le quotient $\mathcal{B}_1 = \mathcal{B}/\mathfrak{m}_{n+1}\mathcal{B}$. Comme l'idéal $(t_1, \ldots, t_k) \subset \mathfrak{m}_{n+1}$, $\mathcal{B}_1$ est un quotient de $\mathcal{B}_0$ et est donc engendré par les images des $e_1, \ldots, e_n$. D'après le théorème 39 de Malgrange exposé plus bas, $\mathcal{B}$ est engendré comme $\mathcal{E}_{n+1}$-module par les $e_1, \ldots, e_n$.

Revenons au théorème 37. Dire que $\mathcal{B}$ est engendré comme $\mathcal{E}_{n+1}$-module par les $e_1, \ldots, e_n$, c'est dire précisément que, modulo $\mathcal{A}$, τ peut s'écrire sous la forme $\tau \equiv F^*(\eta)$. Or par définition de $\mathcal{A}$ cela revient à dire que la condition (11) est satisfaite localement (*i.e.* pour les germes).

Suivant les mêmes techniques, on montre que la condition (11) peut être satisfaite non seulement localement mais multi-localement, *i.e.* simultanément au voisinage de s points $(a_1, \ldots, a_s)$ de même image b par f. Il reste alors à globaliser. Pour cela on montre d'abord que l'on peut globaliser la condition (11) à un voisinage dans $M \times U$ de l'ensemble critique $\Sigma \times \{0\}$ de f. L'application f étant infinitésimalement stable, si b est une valeur critique, l'ensemble $\Sigma_b = f^{-1}(b) \cap \Sigma$ est *fini*. On peut donc, d'après le résultat multi-local, résoudre (11) au voisinage de Σ_b pour toute valeur critique b de f. Σ

étant compact (car fermé dans un compact), son image $f(\Sigma)$ est compacte et la globalisation se fait en recollant les solutions multi-locales à travers des partitions de l'unité bien choisies. Le champ ξ obtenu est global mais n'est vertical et ne satisfait (11) qu'au voisinage de $\Sigma \times \{0\}$.

Pour terminer la démonstration, on suppose $m \geq n$ (car si $m < n$, $\Sigma = M$ et le théorème est démontré). Soient ξ et η les champs donnés par la solution au voisinage de $\Sigma \times \{0\}$. Soit $\sigma = \tau - DF(\xi) - F^*(\eta)$. Ce champ σ est un champ vertical sur $M \times U$ le long de F qui *s'annule* sur un voisinage W de $\Sigma \times \{0\}$ induisant un voisinage W_0 de Σ dans M. Or sur $M - W_0$, f est une *submersion* puisque $m \geq n$ et donc, puisque les submersions sont stables, F est une submersion sur $M \times U - W$ si U est assez petit. Or cela implique que l'on puisse trouver un champ ξ' vertical sur $M \times U - W$ tel que $DF(\xi') = \sigma$ sur $M \times U - W$. En recollant ξ' à la restriction de ξ à W, on obtient une solution globale de (11). $\square$

Le théorème 37 achève la démonstration de l'équivalence, sous l'hypothèse de compacité de M, de la stabilité infinitésimale, de la stabilité homotopique, de la stabilité par déformations et de la stabilité structurelle.

7.10 Panorama conceptuel de la preuve du théorème de Thom-Mather

Nous venons, en suivant Golubitsky et Guillemin [14], d'esquisser les étapes de la démonstration du théorème de Mather sur l'équivalence des différentes sortes de stabilité. Nous allons en reprendre, en suivant cette fois John Mather lui-même, un panorama plus conceptuel.

D'après la définition 22, une application différentiable $f : M \to N$ est infinitésimalement stable si l'application

$$\theta_f + \Omega_f : \Gamma(TM) \oplus \Gamma(TN) \to \Gamma(f^*(TN))$$

est surjective, *i.e.* si tout champ de vecteurs τ sur N le long de f peut s'écrire sous la forme $\tau = Df(\xi) + f^*(\eta)$, ξ étant un champ de vecteurs sur M et η un champ de vecteurs sur N. Soient $\Gamma(M)$ et $\Gamma(N)$ les anneaux topologiques des fonctions $g : M \to \mathbb{R}$ et $h : N \to \mathbb{R}$. Les espaces fonctionnels $\Gamma(TM)$, $\Gamma(TN)$, $\Gamma(f^*(TN))$ sont des espaces vectoriels topologiques naturellement munis d'une structure de module sur ces anneaux : $\Gamma(TM)$ et $\Gamma(f^*(TN))$ sont naturellement des $\Gamma(M)$-modules et $\Gamma(TN)$ est naturellement un $\Gamma(N)$-module. L'application $f : M \to N$ induit par composition un morphisme f^* d'anneaux topologiques $f^* : \Gamma(N) \to \Gamma(M)$. L'application $\theta_f : \Gamma(TM) \to \Gamma(f^*(TN))$ est un morphisme de $\Gamma(M)$-modules et l'application $\Omega_f : \Gamma(TN) \to \Gamma(f^*(TN))$ est un morphisme d'un $\Gamma(N)$-module dans un $\Gamma(M)$-module au-dessus de f^*. On se trouve donc en présence d'une situation abstraite bien définie constituée :

1. d'un morphisme d'anneaux $\varphi : \mathfrak{B} \to \mathfrak{A}$ ($\mathfrak{A} = \Gamma(M)$, $\mathfrak{B} = \Gamma(N)$, $\varphi = f^*$);

2. de deux $\mathfrak{A}$-modules $\mathfrak{M}$ et $\mathfrak{N}$ ($\mathfrak{M} = \Gamma(TM)$ et $\mathfrak{N} = \Gamma(f^*(TN))$) et d'un $\mathfrak{B}$-module $\mathfrak{P}$ ($\mathfrak{P} = \Gamma(TN)$);

3. d'un morphisme $\theta : \mathfrak{M} \to \mathfrak{N}$ de $\mathfrak{A}$-modules ($\theta = \theta_f$) et d'un morphisme $\Omega : \mathfrak{P} \to \mathfrak{N}$ au-dessus de φ ($\Omega = \Omega_f$).

John Mather a appelé la donnée d'un tel quintuple $(\theta, \Omega, \mathfrak{M}, \mathfrak{N}, \mathfrak{P})$ un *morphisme mixte* au-dessus de $\varphi : \mathfrak{B} \to \mathfrak{A}$. Un tel morphisme mixte est dit *surjectif* si $\theta + \Omega : \mathfrak{M} \oplus \mathfrak{P} \to \mathfrak{N}$ est surjectif.

Soit alors $f : M \to N$ infinitésimalement stable. On veut montrer que f est structurellement stable. Il faut donc trouver un voisinage W de f dans l'espace fonctionnel $\mathcal{F} = C^\infty(M, N)$ tel que, pour tout $g \in W$ on puisse construire des difféomorphismes $\varphi \in \mathrm{Diff}(M)$ et $\psi \in \mathrm{Diff}(N)$ satisfaisant $f = \psi \circ g \circ \varphi^{-1}$. Comme on peut supposer W connexe par arcs, on peut partir d'une déformation f_t de f sur $I = [0, 1]$ conduisant de $f_0 = f$ à $f_1 = g$ et chercher des déformations φ_t et ψ_t de 1_M et 1_N telles que, pour tout $t \in I$, $f = \psi_t \circ f_t \circ \varphi_t^{-1}$.L'idée naturelle qui s'impose est évidemment de déduire les déformations φ_t et ψ_t de *l'intégration de champs verticaux* ξ_t et η_t sur $M \times I$ et $N \times I$ et d'utiliser l'hypothèse de stabilité infinitésimale pour montrer que de tels champs existent. Pour cela, il faut d'abord disposer d'un critère assurant que l'on a bien $f = \psi_t \circ f_t \circ \varphi_t^{-1}$.

Lemme. – *Si les conditions* (13) *et* (14) *suivantes sont réalisées alors* $f = \psi_t \circ f_t \circ \varphi_t^{-1}$ *pour tout* $t \in I$.

$$\frac{\partial \varphi_t^{-1}}{\partial t} \circ \varphi_t = \xi_t \text{ et } \frac{\partial \psi_t^{-1}}{\partial t} \circ \psi_t = \eta_t \text{ pour tout } t \in I \ , \tag{13}$$

$$\frac{\partial f_t}{\partial t} = -Df_t \circ \xi_t + \eta_t \circ f_t \ . \tag{14}$$

Preuve. – Il s'agit d'un simple calcul formel. Come évidemment $f = \psi_0 \circ f_0 \circ \varphi_0^{-1} = 1_N \circ f \circ 1_M$, il suffit de montrer que $\frac{\partial}{\partial t}\left(\psi_t \circ f_t \circ \varphi_t^{-1}\right) = 0$. Mais si l'on se donne une composition $v_t \circ u_t$ d'applications différentiables $U \xrightarrow{u_t} V \xrightarrow{v_t} W$, on a :

$$\frac{\partial}{\partial t}(v_t \circ u_t) = \frac{\partial v_t}{\partial t} \circ u_t + Dv_t \circ \frac{\partial u_t}{\partial t} \tag{15}$$

$$\begin{array}{ccccc} & & TV & \xrightarrow{Dv_t} & TW \\ & \frac{\partial u_t}{\partial t} \nearrow & \downarrow & \frac{\partial v_t}{\partial t} \nearrow & \downarrow \\ U & \xrightarrow{u_t} & V & \xrightarrow{v_t} & W \end{array}$$

D'après (15) ,

$$\frac{\partial}{\partial t}\left(\psi_t \circ f_t \circ \varphi_t^{-1}\right) = \frac{\partial \psi_t}{\partial t} \circ f_t \circ \varphi_t^{-1} + D\psi_t \circ \frac{\partial f_t}{\partial t} \circ \varphi_t^{-1} + D\psi_t \circ Df_t \circ \frac{\partial \varphi_t^{-1}}{\partial t} \ .$$

Mais comme $\psi_t \circ \psi_t^{-1} = 1_N$ est une application constante, $\frac{\partial \psi_t}{\partial t} \circ \psi_t^{-1} + D\psi_t \circ \frac{\partial \psi_t^{-1}}{\partial t} = 0$ et donc $\frac{\partial \psi_t}{\partial t} = -D\psi_t \circ \frac{\partial \psi_t^{-1}}{\partial t} \circ \psi_t$. En substituant on obtient :

$$\frac{\partial}{\partial t}\left(\psi_t \circ f_t \circ \varphi_t^{-1}\right) = D\psi_t \circ \left(-\frac{\partial \psi_t^{-1}}{\partial t} \circ \psi_t \circ f_t + \frac{\partial f_t}{\partial t} + Df_t \circ \frac{\partial \varphi_t^{-1}}{\partial t} \circ \varphi_t\right) \circ \varphi_t^{-1} .$$

Mais comme, d'après l'hypothèse (13), $\frac{\partial \varphi_t^{-1}}{\partial t} \circ \varphi_t = \xi_t$ et $\frac{\partial \psi_t^{-1}}{\partial t} \circ \psi_t = \eta_t$, on obtient :

$$\frac{\partial}{\partial t}\left(\psi_t \circ f_t \circ \varphi_t^{-1}\right) = D\psi_t \circ \left(-\eta_t \circ f_t + \frac{\partial f_t}{\partial t} + Df_t \circ \xi_t\right) \circ \varphi_t^{-1} .$$

Or d'après l'hypothèse (14) on a précisément $\frac{\partial f_t}{\partial t} = -Df_t \circ \xi_t + \eta_t \circ f_t$ et donc $\frac{\partial}{\partial t}\left(\psi_t \circ f_t \circ \varphi_t^{-1}\right) = 0$ et $f = \psi_t \circ f_t \circ \varphi_t^{-1}$ pour tout $t \in I$. □

Ce calcul formel élémentaire explique l'importance des conditions du type $\tau = Df(\xi) + f^*(\eta)$ que nous avons constamment rencontrées.

Ainsi, pour montrer que f est structurellement stable, il faut d'abord montrer que, quelle que soit l'homotopie f_t de f, on peut toujours trouver des champs verticaux ξ_t sur $M \times I$ et η_t sur $N \times I$ qui satisfont la condition (14) $\frac{\partial f_t}{\partial t} = -Df_t \circ \xi_t + \eta_t \circ f_t$. Or cette condition est précisément le critère de trivialité de Thom-Levine (théorème 35) pour $k = 1$. D'où la relation fondamentale (en fait, nous l'avons vu, une équivalence) entre la stabilité structurelle et la stabilité homotopique.

Supposons que les champs verticaux ξ_t et η_t existent. Soient φ_τ et ψ_τ leurs flots respectifs sur $M \times I$ et $N \times I$.[35] On a $\frac{d\varphi_\tau}{d\tau}(x,t) = \xi(\varphi_\tau(x,t))$ et $\frac{d\psi_\tau}{d\tau}(x,t) = \eta(\psi_\tau(x,t))$. Si $t = \tau$, on obtient donc la condition (13) $\frac{\partial \varphi_t^{-1}}{\partial t} \circ \varphi_t = \xi_t$ et $\frac{\partial \psi_t^{-1}}{\partial t} \circ \psi_t = \eta_t$ où l'on a changé φ_t en φ_t^{-1} et ψ_t en ψ_t^{-1}. On voit que la construction des φ_t et ψ_t permettant de démontrer la stabilité structurelle de f comprend deux parties :

(i) la démonstration pour toute homotopie f_t de l'existence de ξ_t et η_t satisfaisant (14),

(ii) la démonstration de l'intégrabilité des ξ_t, η_t.

Le point (ii) ne pose pas de problèmes conceptuels particuliers et se résout par des techniques standard à partir de l'hypothèse de compacité de M. Le véritable problème de fond est donc de démontrer le point (i), c'est-à-dire la résolubilité de l'équation (14).

[35] Le paramètre t est un paramètre de déformation alors que le paramètre τ est la variable temporelle des systèmes dynamiques ξ_t et η_t.

Mais la condition qu'exprime cette équation est une condition de stabilité infinitésimale *pour les homotopies*. Il est donc naturel de l'exprimer par la surjectivité d'un morphisme mixte approprié et pour ce faire il suffit de généraliser au cas des germes de déformations toutes les entités intervenant dans le morphisme mixte $\mu = (\theta_f, \Omega_f, \Gamma(TM), \Gamma(f^*(TN)), \Gamma(TN))$ dont la surjectivité exprime la stabilité infinitésimale. Soit $(T, 0)$ une base de déformation. On considère des germes en 0 de déformations f_T de f. Ils constituent un espace $\mathcal{F}_T = C^\infty(T, \mathcal{F})$. De même, les déformations g_T de fonctions $g : M \to \mathbb{R}$ ou h_T de fonctions $h : N \to \mathbb{R}$ constituent des anneaux $\Gamma_T(M_T)$ et $\Gamma_T(N_T)$. Les champs sur $M \times T$ le long d'une déformation f_T de f constituent un $\Gamma_T(M_T)$-module $\Gamma_T(f_T^*(TN_T))$, etc. La condition (14) de stabilité infinitésimale par déformation s'exprime donc en disant que le morphisma mixte $\mu_T = (\theta_{f_T}, \Omega_{f_T}, \Gamma_T(M_T), \Gamma_T(f_T^*(TN_T)), \Gamma_T(N_T))$ sur $f_T^* : \Gamma_T(N_T) \to \Gamma_T(M_T)$ est surjectif.

Pour des situations de ce genre généralisant aux homotopies et aux déformations une situation initiale il existe évidemment une évaluation en $t = 0$ qui n'est rien d'autre que la situation initiale : le morphisme mixte μ est l'évaluation en $t = 0$ du morphisme mixte μ_T.

Le théorème clé permettant de montrer que la stabilité infinitésimale implique la stabilité structurelle est alors le théorème suivant de Mather :

Théorème 38 (Mather). – Si μ est surjectif alors μ_T est surjectif.

Ce théorème reformule le théorème 37 via le critère de stabilité par déformations de Thom-Levine. Comme nous allons le voir, le théorème de préparation différentiable y joue un rôle déterminant.

7.11 Le théorème de préparation différentiable

7.11.1 Le théorème et le lemme de Nakayama

Conjecturé par René Thom et démontré par Bernard Malgrange en 1963, le théorème de préparation différentiable concerne la situation suivante. Soient $f : M \to N$, $a \in M$, $b = f(a) \in N$, soient $\mathcal{E}_m = \Gamma_a(M)$ et $\mathcal{E}_n = \Gamma_b(N)$ les anneaux locaux des germes de fonctions sur M et N en a et b et soit enfin $\mathfrak{B}$ un $\mathcal{E}_m$-module de type fini. A travers $f^* : \mathcal{E}_n \to \mathcal{E}_m$, $\mathfrak{B}$ devient un $\mathcal{E}_n$-module et la question est de savoir *s'il est également de type fini*. Le théorème de préparation dit essentiellement que, pour cela, il suffit de vérifier les conditions trouvées jusqu'à un ordre *fini*.

Theorème 39. – Soit $\mathcal{B}$ un $\mathcal{E}_m$-module de type fini. Alors $\mathcal{B}$ est de type fini comme $\mathcal{E}_n$-module si et seulement si $\mathcal{B}/\mathfrak{m}_n\mathcal{B}$ est un espace vectoriel de dimension finie.

Montrons d'abord que ce théorème implique bien, comme nous l'avons affirmé plus haut, la Proposition 28 :

Proposition 28. – Soit $\mathcal{B}$ un $\mathcal{E}_m$-module de type fini. Soient $e_1, \ldots, e_k$ des éléments de $\mathcal{B}$. Alors ils engendrent $\mathcal{B}$ comme $\mathcal{E}_n$-module si et seulement si ils

engendrent $\mathcal{B}/\mathfrak{m}_m^{k+1}\mathcal{B}$ comme $\mathcal{E}_n$-module.

Preuve. – Supposons en effet que $e_1, \ldots, e_k$ engendrent $\mathcal{B}/\mathfrak{m}_m^{k+1}\mathcal{B}$ comme $\mathcal{E}_n$-module et considérons le quotient $\mathcal{B}' = \mathcal{B}/\left(\mathfrak{m}_m^{k+1}\mathcal{B} + \mathfrak{m}_n\mathcal{B}\right)$. $\mathcal{B}'$ est un module sur $\mathcal{E}_n/\mathfrak{m}_n \simeq \mathbb{R}$, c'est-à-dire un $\mathbb{R}$-espace vectoriel. Comme il est engendré par les images de $e_1, \ldots, e_k$, il est de dimension $\leq k$. Soit alors $\mathcal{B}_\ell = \mathfrak{m}_m^\ell \mathcal{B}'$. Lorsque ℓ varie de $\ell = 0$ à $\ell = k+1$, les $\mathcal{B}_\ell$ constituent une suite décroissante d'espaces vectoriels allant de $\mathcal{B}_0 = \mathcal{B}'$ à $\mathcal{B}_{k+1} = \mathfrak{m}_m^{k+1}\mathcal{B}' = 0$. Pour des raisons dimensionnelles ($k+2$ vectoriels pour une différence de dimension d'au plus k), il existe donc un des $\ell \in \{0, 1, \ldots, k+1\}$ tel que $\mathcal{B}_\ell = \mathcal{B}_{\ell+1}$, *i.e.* tel que

$$\mathfrak{m}_m^\ell \mathcal{B} + \mathfrak{m}_n\mathcal{B} = \mathfrak{m}_m^{\ell+1}\mathcal{B} + \mathfrak{m}_n\mathcal{B}$$

Or d'après un lemme classique d'algèbre, sur lequel nous allons revenir, dit *lemme de Nakayama*, cela implique $\mathfrak{m}_m^\ell \mathcal{B} \subset \mathfrak{m}_n\mathcal{B}$.

Lemme de Nakayama. – Soit $\mathcal{B}$ un $\mathcal{E}_m$-module de type fini. Si $\mathcal{B} = \mathfrak{m}_m\mathcal{B}$ alors $\mathcal{B} = \{0\}$. Autrement dit, si $\mathcal{B}/\mathfrak{m}_m\mathcal{B} = \{0\}$, alors $\mathcal{B} = \{0\}$.

Comme $(e_1, \ldots, e_k)$ engendrent $\mathcal{B}/\mathfrak{m}_m^{k+1}\mathcal{B}$ comme $\mathcal{E}_n$-module par hypothèse, ils engendrent par conséquent aussi $\mathcal{B}/\mathfrak{m}_n\mathcal{B}$ comme espace vectoriel. D'après le théorème de Malgrange, $\mathcal{B}$ est donc de type fini comme $\mathcal{E}_n$-module. Or, toujours d'après le lemme de Nakayama, si $\mathcal{B}$ est un $\mathcal{E}_n$-module de type fini et si $(u_1, \ldots, u_k)$ est une base de l'espace vectoriel $\mathcal{B}/\mathfrak{m}_n\mathcal{B}$, alors, si $(e_1, \ldots, e_k)$ relèvent $(u_1, \ldots, u_k)$ dans $\mathcal{B}$, ils engendrent $\mathcal{B}$ sur $\mathcal{E}_n$. □

7.11.2 Le cas des submersions

Revenons au théorème de Malgrange. Il formule en termes de propriétés de finitude de modules sur des anneaux locaux de germes de fonctions un résultat fondamental et profond concernant la possibilité de *diviser* sous certaines conditions des fonctions différentiables par d'autres fonctions différentiables.

Supposons d'abord que f soit une *submersion* en a avec $m = n+1$ et soit $\mathcal{B}$ un $\mathcal{E}_m$-module de type fini tel que l'espace vectoriel $V = \mathcal{B}/\mathfrak{m}_n\mathcal{B}$ soit de dimension finie. Soit $(e_1, \ldots, e_k)$ des éléments de $\mathcal{B}$ dont les images dans V constituent une base de V. Conformément à la proposition 8, on suppose que l'on a choisi des coordonnées locales trivialisantes $(x_1, \ldots, x_m)$ permettant d'identifier f à la projection canonique $\mathbb{R}^m = \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, $(x_1, \ldots, x_m) \mapsto (x_2, \ldots, x_m)$. L'idéal maximal $\mathfrak{m}_n = (x_2, \ldots, x_m)\,\mathcal{E}_n$ de $\mathcal{E}_n$ s'identifie alors à un sous-ideal de $(x_1, \ldots, x_m)\,\mathcal{E}_m$. Comme par hypothèse les $(e_1, \ldots, e_k)$ engendrent $V = \mathcal{B}/\mathfrak{m}_n\mathcal{B}$ et donc, a fortiori, $V' = \mathcal{B}/\mathfrak{m}_m\mathcal{B}$, nous allons voir que, d'après le lemme de Nakayama, ils engendrent $\mathcal{B}$ comme $\mathcal{E}_m$-module.

Le lemme de Nakayama dit que si $\mathcal{B}/\mathfrak{m}_m\mathcal{B} = \{0\}$, alors $\mathcal{B} = \{0\}$. Il est une conséquence directe du fait que dans tout anneau *local* $\mathcal{A}$ les éléments de l'idéal maximal $\mathfrak{m}$ sont exactement les éléments *non inversibles* de $\mathcal{A}$. Soit en effet $s \in \mathfrak{m}$. Si s était inversible on aurait $s.s^{-1} = 1 \in \mathfrak{m}$ et donc $\mathfrak{m} = \mathcal{A}$, ce qui

est impossible. Réciproquement, si $s \notin \mathfrak{m}$, alors l'idéal principal (s) n'est pas inclus dans $\mathfrak{m}$ et donc $(s) = \mathcal{A}$. Il existe par conséquent un t tel que $s.t = 1$ et s est inversible.

Supposons alors que les $(e_1, \ldots, e_k)$ engendrent $\mathcal{B}$. Comme $\mathcal{B} = \mathfrak{m}\mathcal{B}$ par hypothèse, les e_i s'écrivent $e_i = \sum_{j=1}^{j=k_i} m_{ij} a_{ij}$ avec $m_{ij} \in \mathfrak{m}$ et $a_{ij} \in \mathcal{B}$. Mais par ailleurs $a_{ij} = \sum_{\ell=1}^{\ell=k} r_{ij\ell} e_\ell$ et donc, en définitive,

$$e_i = \sum_{j=1}^{j=k_i} \sum_{\ell=1}^{\ell=k} m_{ij} r_{ij\ell} e_\ell = \sum_{\ell=1}^{\ell=k} \left(\sum_{j=1}^{j=k_i} m_{ij} r_{ij\ell} \right) e_\ell = \sum_{\ell=1}^{\ell=k} s_{i\ell} e_\ell \ .$$

Cela implique $\sum_{\ell=1}^{\ell=k} (\delta_{i\ell} - s_{i\ell}) e_\ell = 0$ ($\delta_{i\ell}$ étant le symbole de Kronecker) et donc, si la matrice $k \times k$ $E = (\delta_{i\ell} - s_{i\ell})$ est *inversible*, $e_1 = \ldots = e_k = 0$ et par conséquent $\mathcal{B} = \{0\}$. Mais le déterminant de E est de la forme $1 + s$ avec $s \in \mathfrak{m}$ et il est donc inversible dans $\mathcal{A}$ car sinon il serait dans $\mathfrak{m}$ et $1 + s - s = 1$ serait dans $\mathfrak{m}$. Donc la matrice E est inversible. $\square$

Une conséquence immédiate est le résultat que nous avons utilisé plus haut (pour un $\mathcal{E}_n$-module) que si $\mathcal{B}$ est un $\mathcal{E}_m$-module de type fini et si $(u_1, \ldots, u_k)$ est une base de l'espace vectoriel $\mathcal{B}/\mathfrak{m}_m\mathcal{B}$, alors, si les $(e_1, \ldots, e_k)$ relèvent $(u_1, \ldots, u_k)$ dans $\mathcal{B}$, ils engendrent $\mathcal{B}$ sur $\mathcal{E}_m$. En effet, soit $\mathcal{A}$ le sous-module engendré par $(e_1, \ldots, e_k)$ et $\mathcal{C} = \mathcal{B}/\mathcal{A}$. $\mathcal{C}$ est un $\mathcal{E}_m$-module de type fini et $\mathcal{C}/\mathfrak{m}_m\mathcal{C} = \{0\}$. D'après le lemme de Nakayama, $\mathcal{C} = \{0\}$ et donc $\mathcal{B} = \mathcal{A}$.

Par conséquent, dans la preuve du théorème de Malgrange, puisque les $(e_1, \ldots, e_k)$ engendrent $V' = \mathcal{B}/\mathfrak{m}_m\mathcal{B}$, ils engendrent bien $\mathcal{B}$ comme $\mathcal{E}_m$-module. Donc tout élément $a \in \mathcal{B}$ peut s'écrire sous la forme $a = \sum_{i=1}^{i=k} h_i e_i$ avec $h_i \in \mathcal{E}_m$. En fait on peut préciser cette forme : tout élément $a \in \mathcal{B}$ peut s'écrire sous la forme $a = \sum_{i=1}^{i=k} (c_i + f_i) e_i$ avec $c_i \in \mathbb{R}$ et $f_i \in \mathfrak{m}_n\mathcal{E}_m$. En effet, comme les $(e_1, \ldots, e_k)$ engendrent $V = \mathcal{B}/\mathfrak{m}_n\mathcal{B}$, a peut s'écrire $a = \sum_{i=1}^{i=k} c_i e_i + a'$ avec $c_i \in \mathbb{R}$ et $a' \in \mathfrak{m}_n\mathcal{B}$, $a' = \sum_{i=1}^{i=k} m_i b_i$, $m_i \in \mathfrak{m}_n$, $b_i \in \mathcal{B}$. Mais comme les $(e_1, \ldots, e_k)$ engendrent $\mathcal{B}$ comme $\mathcal{E}_m$-module, $b_i = \sum_{j=1}^{j=k} h_{ij} e_j$, $h_{ij} \in \mathcal{E}_m$. Donc $a' = \sum_{i,j=1}^{i,j=k} m_i h_{ij} e_j = \sum_{j=1}^{j=k} f_j e_j$ avec $f_j \in \mathfrak{m}_n\mathcal{E}_m$.

Il s'agit donc maintenant de montrer, à partir du résultat que tout $a \in \mathcal{B}$ peut s'écrire sous la forme $a = \sum_{i=1}^{i=k} (c_i + f_i) e_i$ avec $c_i \in \mathbb{R}$ et $f_i \in \mathfrak{m}_n\mathcal{E}_m$, que $\mathcal{B}$ est de type fini comme $\mathcal{E}_n$-module. Pour ce faire, il faut montrer qu'il existe un système fini $(\alpha_1, \ldots, \alpha_r)$ d'éléments de $\mathcal{B}$ tels que tout $a \in \mathcal{B}$ puisse s'écrire sous la forme $a = \sum_{i=1}^{i=r} d_i \alpha_i$ avec $d_i \in \mathcal{E}_n$. Comparons cette forme à la précédente. Comme $d_i \in \mathcal{E}_n$, d_i est le germe en 0 d'une fonction $d_i(x_2, \ldots, x_m)$ et l'on a donc $d_i = d_i(0) + d_i'$ avec $d_i' \in \mathfrak{m}_n$. Cette forme est analogue à la précédente à ceci près, qui est précisément le point crucial, que $d_i' \in \mathfrak{m}_n$ alors que $f_i \in \mathfrak{m}_n\mathcal{E}_m$. On voit que la différence porte sur la variable x_1. Si l'on développe $f_i(x_1, \ldots, x_m)$ en série de x_1, on a formellement $f_i(x_1, \ldots, x_m) = \sum_{j=0}^{j=\infty} x_1^j f_{ij}(x_2, \ldots, x_m)$ avec $f_{ij}(x_2, \ldots, x_m) \in \mathfrak{m}_n$. Pour montrer que $\mathcal{B}$ est de type fini sur $\mathcal{E}_n$, il suffirait donc de montrer qu'il existe un entier s tel que *toute*

fonction $f_i \in \mathfrak{m}_n\mathcal{E}_m$ puisse s'écrire comme une somme *finie* $f_i(x_1,\ldots,x_m) = \sum_{j=0}^{j=s} x_1^j f_{ij}(x_2,\ldots,x_m)$ avec $f_{ij}(x_2,\ldots,x_m) \in \mathfrak{m}_n$. En effet, dans ce cas, les $(e_1,\ldots,e_k)$ et les $x_1^j e_i$, $i=1,\ldots,k$, $j=1,\ldots,s$, engendreraient alors $\mathcal{B}$ comme $\mathcal{E}_n$-module.

Mais une telle exigence est évidemment impossible à satisfaire. En effet, $x_1^\ell \in \mathfrak{m}_n\mathcal{E}_m$ pour tout ℓ et, si $\ell > s$, x_1^ℓ ne peut évidemment pas s'écrire sous le forme $\sum_{j=0}^{j=s} x_1^j f_{ij}(x_2,\ldots,x_m)$. Il est donc nécessaire de raffiner la procédure. Comme $x_1 e_i \in \mathfrak{m}_n\mathcal{B}$ pour $i=1,\ldots,k$, on a $x_1 e_i = \sum_{j=1}^{j=k}(c_{ij}+f_{ij})e_j$ avec $c_{ij} \in \mathbb{R}$ et $f_{ij} \in \mathfrak{m}_n\mathcal{E}_m$. Ce système d'équations linéaires est équivalent à $Ee=0$ avec $e=(e_1,\ldots,e_k)$ et E la matrice $(x_1\delta_{ij} - c_{ij} - f_{ij})$. Soit $P(x_1,\ldots,x_m)$ le déterminant de E. Il est associé à la donnée de e et, comme $Ee=0$, il satisfait $Pe_i = 0$ pour $i=1,\ldots,k$. D'autre part, comme $f_{ij} \in \mathfrak{m}_n\mathcal{E}_m$, $f_{ij}(x_1,0,\ldots,0)=0$ et donc $P(x_1,0,\ldots,0)$ est un *polynôme* en x_1 de degré $\leq k$. Il existe par conséquent un entier $s \leq k$ tel que $P(x_1,0,\ldots,0) = x_1^s g(x_1)$ avec $g(0) \neq 0$.

Soit alors $a = \sum_{i=1}^{i=k}(c_i+f_i)e_i$ (avec $c_i \in \mathbb{R}$ et $f_i \in \mathfrak{m}_n\mathcal{E}_m$) un élément de $\mathcal{B}$. Supposons que l'on puisse *diviser* les $f_i \in \mathfrak{m}_n\mathcal{E}_m$ par P, *i.e.* que l'on puisse écrire f_i sous la forme $f_i = q_i P + \sum_{j=0}^{j=s} x_1^j R_{ij}(x_2,\ldots,x_m)$. Cela résoudrait le problème puisque, étant donné que $Pe_i=0$, on aurait $f_i e_i = \sum_{j=0}^{j=s} x_1^j R_{ij}(x_2,\ldots,x_m)e_i$ et les $(e_1,\ldots,e_k)$ et les $x_1^j e_\ell$, $\ell=1,\ldots,k$, $j=1,\ldots,s$, engendreraient alors $\mathcal{B}$ comme $\mathcal{E}_n$-module. C'est donc la division des $f_i \in \mathfrak{m}_n\mathcal{E}_m$ par P qui constitue le cœur de la preuve. D'où la deuxième étape reposant sur un théorème de division.

7.11.3 Le théorème de division

Théorème 40 (théorème de division différentiable). – Soit P une fonction différentiable au voisinage de l'origine de $\mathbb{R}^m = \mathbb{R}\times\mathbb{R}^n$ (en particulier un polynôme) telle que $P(x_1,0,\ldots,0) = x_1^s g(x_1)$ avec $g(0)\neq 0$. Alors, étant donnée une fonction différentiable f quelconque au voisinage de l'origine de $\mathbb{R}^m$, on peut diviser f par P, c'est-à-dire trouver des fonctions q et r telles que (i) $f = qP + r$, et (ii) $r = \sum_{j=0}^{j=s-1} x_1^j r_j(x_2,\ldots,x_m)$.

Preuve. – Dans le cas où l'on considère des fonctions *holomorphes* sur $\mathbb{C}^m = \mathbb{C}\times\mathbb{C}^n$, ce résultat est connu depuis longtemps sous le nom de *théorème de division de Weierstrass* et se démontre assez facilement. En revanche, dans le cas différentiable, il est hautement non trivial. Remarquons que si on l'applique à la fonction x_1^s on obtient $x_1^s = qP + \sum_{j=0}^{j=s-1} x_1^j r_j(x_2,\ldots,x_m)$, autrement dit $qP = x_1^s - \sum_{j=0}^{j=s-1} x_1^j r_j(x_2,\ldots,x_m)$ avec $q(0)\neq 0$. Ce résultat est connu sous le nom de *théorème de préparation de Weierstrass*. Il dit que si une fonction holomorphe P est telle que $P(x_1,0,\ldots,0) = x_1^s g(x_1)$ *avec* $g(0)\neq 0$ alors, en multipliant P par une fonction convenable q inversible en 0 ($q(0)\neq 0$), on peut ramener la dépendance de P par rapport à x_1 à une dépendance polynomiale de degré s.

Pour démontrer le théorème de division, on considère d'abord le cas des polynômes $P_{s,\lambda} = x_1^s + \sum_{j=1}^{j=s-1} \lambda_j x_1^j$ et l'on montre que les fonctions q et r_j dépendent différentiablement de λ. Cela permet de démontrer facilement le cas général. Il existe de nombreuses preuves. La plus élégante est sans doute celle de Nirenberg cherchant à prolonger directement le cas holomorphe sur $\mathbb{C}$ au cas différentiable sur $\mathbb{R}$.[36] Le principe en est d'étendre f définie sur $\mathbb{R}^m = \mathbb{R} \times \mathbb{R}^n$ à une fonction différentiable $\widetilde{f_\lambda}$ définie sur $\mathbb{C} \times \mathbb{C}^n \times \mathbb{C}^s$ et à valeurs dans $\mathbb{C}$.On applique alors les théorèmes de Weierstrass pour trouver les fonction q et r. Ces fonctions font intervenir des intégrales doubles intégrant $\frac{\partial \widetilde{f}}{\partial \overline{z_1}}/P_s$ sur des domaines pouvant contenir des zéros de P_s. En ces zéros il y a obstruction à l'intégration et à la construction de q et r sauf si $\frac{\partial \widetilde{f}}{\partial \overline{z_1}}$ s'annule lorsque $\overline{z_1} = x_1$ est réel. Qui plus est, on montre que pour que q et r soient différentiables à l'ordre k il faut que $\frac{\partial \widetilde{f}}{\partial \overline{z_1}}$ s'annule à l'ordre k sur les zéros de P_s lorsque $\overline{z_1} = x_1$ est réel. Le noeud de la preuve est alors de montrer que de telles extensions existent (théorème d'extension de Nirenberg). $\square$

7.11.4 Le cas général

Le théorème de division permet de démontrer le théorème de préparation dans le cas où f est une submersion. Dans le cas où f est une immersion, $f^* : \mathcal{E}_n \to \mathcal{E}_m$ est surjectif et le théorème est trivial puisque $\mathcal{B}$ est alors nécessairement de type fini.

Si $f : M \to N$ est une application quelconque, on raisonne par récurrence. On considère l'immersion $\widetilde{f} : M \to M \times N$, $x \mapsto \widetilde{f}(x) = (x, f(x))$. Si $\mathcal{E}_{m+n}$ est l'anneau des germes $\Gamma_{(a,b)}(M \times N)$ pour $(a,b) \in M \times N$, $\mathcal{B}$ est un $\mathcal{E}_{m+n}$-module de type fini. Supposons que $\mathcal{B}/\mathfrak{m}_n\mathcal{B}$ soit de dimension finie. Localement, M est identifiable à $\mathbb{R}^m$ et on peut considérer la projection canonique $\mathbb{R}^m = \mathbb{R} \times \mathbb{R}^{m-1} \to \mathbb{R}^{m-1}$. Comme $\mathfrak{m}_n\mathcal{B} \subset \mathfrak{m}_{(a,b)}(\mathbb{R}^{m-1} \times N)$, on a une surjection canonique $\mathcal{B}/\mathfrak{m}_n\mathcal{B} \to \mathcal{B}/\mathfrak{m}_{(a,b)}(\mathbb{R}^{m-1} \times N)$ et le quotient $\mathcal{B}/\mathfrak{m}_{(a,b)}(\mathbb{R}^{m-1} \times N)$ est donc de dimension finie. En appliquant le théorème de préparation à la submersion $\mathbb{R}^m \times N \to \mathbb{R}^{m-1} \times N$, on prouve que $\mathcal{B}$ est de type fini sur $\Gamma_{(a,b)}(\mathbb{R}^{m-1} \times N)$ et on conclut par récurrence.

7.11.5 L'interprétation de René Thom

Le théorème de préparation étant un résultat particulièrement profond et utile, nous nous permettrons de rappeler la façon dont René Thom [51] en a explicité la signification dans des termes plus adaptés à la théorie des déploiements universels que nous aborderonsplus loin.

Une des formulations du théorème de préparation de Weierstrass dit que si $f(z)$ est une fonction holomorphe sur un voisinage U de l'origine de $\mathbb{C}$ et $P(z)$ un polynôme de degré s dont toutes les racines sont dans U, alors f s'écrit sous la forme $f = qP + r$ où le reste r est un polynôme de degré $\leq s-1$. En termes algébro-géométriques, cela signifie que l'idéal (P) engendré par P dans la $\mathbb{C}$-

[36] Cf. Nirenberg [28] et Mather [26].

algèbre $\mathcal{H}(U)$ des fonctions holomorphes sur U est un sous-espace vectoriel de codimension s admettant pour supplémentaire le sous-espace vectoriel R, de dimension s, des polynômes r de degré $\leq s-1$. Autrement dit, $\mathcal{H}(U) = (P)+R$ et la décomposition $f = qP + r$ est tout simplement celle de f relativement à cette somme. On peut prendre en particulier $P = z^s$. On peut aussi prendre pour P une fonction holomorphe admettant un zéro d'ordre s en 0 (i.e. $P = z^s g$ avec $g(0) \neq 0$). On peut également adjoindre des variables supplémentaires, les coefficients du reste r en devenant des fonctions holomorphes. C'est le cas que nous avons rencontré.

Le théorème de Weierstrass est assez intuitif. En effet pour qu'une fonction holomorphe $f \in \mathcal{H}(U)$ soit un multiple de P, il faut et il suffit que f s'annule sur les racines de P avec au moins la même multiplicité que P. Cette condition est la réunion de s conditions linéaires linéairement indépendantes et l'idéal (P) est donc de codimension s. Si f est quelconque, r est le polynôme (unique) défini par le comportement de $f - P$ aux racines de P.

La transversalite de (P) et de R est *stable* relativement aux déformations de P. Si l'on part de $P = z^s$ et si l'on déforme P en $P_w = z^s + \sum_{i=1}^{i=s-1} w_i z^i$, R reste un supplémentaire de l'idéal (P_w). Si $f \in \mathcal{H}(U)$, son reste r sera donc une fonction $r(w)$ de w. La seconde partie du théorème de Weierstrass dit que $r(w)$ est une fonction holomorphe de la multi-variable w parcourant un voisinage W de l'origine de $\mathcal{C}^{s-1}$.

Lorsque l'on passe du domaine complexe au domaine réel, la situation se trouve considérablement compliquée par le fait que le nombre de racines de P_w n'est plus constant dans W. L'idéal (P_w) est toujours de codimension s lorsque P_w a ses s racines réelles. Mais il devient de codimension $< s$ lorsque P_w admet des racines complexes. Le reste $r(w)$ est donc une fonction différentiable de w sur la partie W_R de W où P_w a toutes ses racines réelles. Le théorème de Malgrange dit essentiellement qu'il est possible d'étendre $r(w)$ à W *tout entier*.

7.12 L'intérêt de la stabilité transversale

D'après la caractérisation de la stabilité infinitésimale à travers le théorème 39 de Malgrange, la stabilité d'une application $f : M \to N$ ne dépend que de son $(n+1)$-jet. Il est donc normal de chercher à l'interpréter en termes de *transversalité* dans les espaces de jets. L'idée la plus naturelle consiste à partir de la décomposition des espaces de jets $J^k(M,N)_{(a,b)}$ en orbites sous l'action du groupe $G = \mathrm{Diff}(M) \times \mathrm{Diff}(N)$ en remarquant qu'elle est *universelle* (*i.e.* indépendante de f et dépendante uniquement des dimensions m et n et de k) et à imposer que les jets de f soient transverses sur ces orbites. Tel est le contenu de la stabilité transversale.

Soient donc $f : M \to N$, $a \in M$, $b = f(a) \in N$.

Proposition 41. – Si f est stable, alors f est localement transversalement stable en a.

Preuve. – D'après la définition 25, il faut montrer que le n-jet $j^n f$ de f est transverse en a sur l'orbite $\widetilde{j^n f(a)}$ de $j^n f(a)$ dans $J^n(M,N)$. Or f étant stable et la transversalité sur $\widetilde{j^n f(a)}$ étant générique d'après le théorème de transversalité de Thom, il existe g équivalente à f et transverse sur $\widetilde{j^n f(a)}$. Si, dans l'équivalence entre f et g, a' correspond à a, $j^n g$ est transverse sur $\widetilde{j^n f(a)}$ en a' et donc $j^n f$ est transverse sur $\widetilde{j^n f(a)}$ en a. □

Réciproquement, on a le :

Théorème 42. – Si f est localement transversalement stable en a, alors f est localement infinitésimalement stable en a.

Preuve. – Par hypothèse, $j^n f$ est transverse sur l'orbite $\widetilde{j^n f(a)}$ et, d'après le corollaire 29 caractérisant la stabilité infinitésimale locale, il faut montrer que

$$J^n(f^*TN)_a = (D_a f)(J^n(TM)_a) + f^*(J^n(TN)_b) \ .$$

Soit donc $\tau_n \in J^n(f^*TN)_a$ le n-jet en a d'un champ de vecteurs τ sur M le long de f. Nous devons trouver localement en a, un champ ξ sur M de n-jet ξ_n et un champ η sur N de n-jet $\eta_n$tels que l'on ait $\tau_n = D_a f(\xi_n) + f^*(\eta_n)$. Considérons une homotopie f_t de f telle que $\frac{df_t}{dt}\big|_{t=0} = \tau$. Il lui correspond une courbe $t \longmapsto j^n f_t(a)$ de $J^n(M,N)$ passant par $j^n f(a)$. Notons $\lambda(\tau)$ son vecteur tangent en $j^n f(a)$. On définit ainsi une application $\lambda : J^n(f^*TN)_a \to T_{j^n f(a)} J^n(M,N)$ qui, c'est facile de le voir, est une injection linéaire. Comme f est localement transversalement stable en a, il existe, par définition de la transversalité, des vecteurs tangents $w \in T_{j^n f(a)}\widetilde{j^n f(a)}$ et $v \in T_a M$ tels que $\lambda(\tau_n) = w + D_a j^n f(v)$. Or l'espace tangent $T_\sigma \widetilde{\sigma}$ en $\sigma = j^n f(a)$ à l'orbite $\widetilde{\sigma} = \widetilde{j^n f(a)}$ est décomposable. Soit η un germe de champ sur N en b, $\eta \in \Gamma_b(TN)$. Intégrons localement η. On obtient un germe de groupe à un paramètre ψ_t et $j^n \psi_t(b).\sigma$ est une courbe de l'orbite $\widetilde{\sigma}$. Soit θ_1 l'application $\theta_1 : \Gamma_b(TN) \to T_\sigma \widetilde{\sigma}$ ainsi définie. Il est facile de vérifier que le diagramme suivant est commutatif :

$$\begin{array}{ccc} \Gamma_b(TN) & \xrightarrow{\theta_1} & T_\sigma \widetilde{\sigma} \\ \downarrow & & \uparrow \lambda \\ J^n(TN)_b & \xrightarrow{f^*} & J^n(f^*TN)_a \end{array}$$

De même, soit ξ un germe de champ sur M en a. En intégrant ξ localement, on obtient un germe de groupe à un paramètre φ_t et une courbe $\sigma.j^n \varphi_t^{-1}(\varphi_t(a))$ de $\widetilde{\sigma}$. Soit θ_2 l'application $\theta_2 : \Gamma_a(TM) \to T_\sigma \widetilde{\sigma}$ ainsi définie. Il est facile de vérifier que le diagramme suivant est commutatif :

$$\begin{array}{ccc} \Gamma_a(TM) & \xrightarrow{\theta_2} & T_\sigma \widetilde{\sigma} \\ \downarrow & & \uparrow \Lambda \\ J^n(TM)_a & \xrightarrow{Id} & J^n(TM)_a \end{array}$$

où $\Lambda : J^n(TM)_a \to T_\sigma\widetilde{\sigma}$ est l'application qui, à $\xi_n \in J^n(TM)_a$, associe $-\lambda.D_a f(\xi) + D_a j^n f(\xi_n(a))$.

Lemme. – $\theta_1 + \theta_2 : \Gamma_b(TN) \oplus \Gamma_a(TM) \to T_\sigma\widetilde{\sigma}$ *est surjective.*

On notera la ressemblance de cette condition avec les conditions de surjectivité des morphismes mixtes caractérisant la stabilité infinitésimale et la stabilité par déformations.

Ce lemme permet d'écrire le terme w de $\lambda(\tau_n) = w + D_a j^n f(v)$ sous la forme $w = -\theta_2(\xi) + \theta_1(\eta)$ avec $\xi \in \Gamma_a(TM)$ et $\eta \in \Gamma_b(TN)$. En prenant les n-jets ξ_n et η_n et en utilisant la commutativité des diagrammes, on obtient

$$\begin{aligned}\lambda(\tau_n) &= -\Lambda(\xi_n) + \lambda(f^*(\eta_n)) + D_a j^n f(v)\\ &= \lambda(D_a f(\xi_n)) - D_a j^n f(\xi_n(a)) + \lambda(f^*(\eta_n)) + D_a j^n f(v)\ .\end{aligned}$$

Mais on peut montrer que, nécessairement, $\xi_n(a) = v$. Par conséquent, $\lambda(\tau_n) = \lambda(D_a f(\xi_n)) + \lambda(f^*(\eta_n))$ et donc, comme λ est injective, $\tau_n = D_a f(\xi_n) + f^*(\eta_n)$. □

En passant à une version multi-locale de ce théorème et en utilisant la caractérisation 30 de la stabilité infinitésimlale, on obtient le théorème :

Théorème 43. – *f est (infinitésimalement) stable si et seulement si f est transversalement stable.*

On voit ainsi, qu'en définitive, la stabilité structurelle d'une application différentiable $f : M \to N$ signifie essentiellement que les jets de f sont transverses aux orbites de l'action des J^kG ($G = \mathrm{Diff}(M) \times \mathrm{Diff}(N)$) sur les espaces de jets $J^k(M,N)$. Cette action est universelle. Pour comprendre la géométrie de f, il faut donc d'abord analyser la géométrie de la décomposition en orbites de $J^k(M,N)$. Cela peut se faire en se situant au niveau des fibres de ces fibrés. Soit $J^k(m,n)$ la fibre de $J^k(M,N)$ en (a,b). C'est un espace vectoriel ne dépendant que de m et de n et égal à $J^k(\mathbb{R}^m,\mathbb{R}^n)_{(0,0)}$. Soit $L^k(p)$ l'ensemble des k-jets des germes de difféomorphismes de $\mathbb{R}^p$ en 0. Le groupe de Lie $L^k(m,n) = L^k(m) \times L^k(n)$ opère naturellement sur $J^k(m,n)$. Si l'on connaît la géométrie de cette action, on connaîtra par globalisation celle de l'action de J^kG sur $J^k(M,N)$ et donc celle des application stables puisque, dans ce cas, la géométrie de f sera l'image réciproque par les j^kf de la géométrie de $J^k(M,N)$ et que la transversalité "préserve" la géométrie par image réciproque.

Mais, comme nous allons le voir dans la prochaine section, $J^k(m,n)$ est en général décomposé en une *infinité non dénombrable* d'orbites par $L^k(m,n)$. Pour que f soit stable, elle devra donc satisfaire une infinité non dénombrable de conditions de transversalité et cela explique pourquoi la stabilité structurelle *n'est pas forcément* une propriété générique. Dans un théorème remarquable, John Mather [25] a donné la liste des dimensions (m,n) pour lesquelles la stabilité structurelle est générique.

Faisons une dernière remarque à propos de la stabilité. Une application stable possède *toutes* les propriétés qui sont génériques et ne dépendent que

du type différentiable. Appelons encore "génériques" de telles propriétés. Le fait que la stabilité transversale implique la stabilité signifie qu'il suffit qu'une applications possède *certaines* propriétés génériques (la transversalité sur les orbites des espaces de jets) pour les posséder *toutes*.

8 La stratification des espaces de jets et la théorie de Thom-Boardman

Nous venons de voir que pour comprendre la structure géométrique des applications stables il faut d'abord analyser la géométrie des orbites des espaces de jets. Mais celle-ci n'est pas en général une "bonne" géométrie, ne serait-ce que parce qu'il peut y avoir des continua d'orbites. D'où l'idée de considérer des *sous-variétés invariantes* (*i.e.* des réunions d'orbites) des espaces de jets qui aient un sens géométrique associé à la géométrie des ensembles singuliers des applications. Cette idée est à la base de la théorie de Thom-Boardman fournissant une première classification des singularités des applications stables. Nous allons en exposer brièvement quelques grandes lignes en suivant Levine [16] et Boardman [6].

8.1 La structure des espaces de jets $J^k(m,n)$

8.1.1 La stratification de $J^k(m,n)$

Fixons les dimensions m et n et considérons les espaces de jets "fibres" $J^k(m,n) = J^k(\mathbb{R}^m,\mathbb{R}^n)_{(0,0)}$. Le groupe de Lie $L^k(m,n) = L^k(m)\times L^k(n)$ opère sur $J^k(m,n)$. Il est clair que si $l > k$, il existe une projection canonique $\pi_{l,k} : J^l(m,n) \to J^k(m,n)$ consistant à tronquer les l-jets à l'ordre k. $L^1(m)$ est le groupe linéaire $GL(m,\mathbb{R})$ des automorphismes linéaires de $\mathbb{R}^m$ puisque le 1-jet d'un germe de difféomorphisme en 0 de $\mathbb{R}^m$ est sa matrice jacobienne (inversible). On a $L^k(m) = \pi_{k,1}^{-1}(L^1(m))$. Les espaces de jets considérés sont des espaces de polynômes et les actions considérées sont *algébriques*. Autrement dit, la structure des espaces $J^k(m,n)$ est un problème de géométrie algébrique sur des espaces vectoriels.

Un résultat classique affirme que si G est un groupe de Lie opérant sur une variété différentiable M, les orbites sont des sous-variétés immergées de M. Les orbites de $J^k(m,n)$ sont donc des sous-variétés immergées.

Soit $\sigma \in J^k$ et f un représentant de σ.[37] Supposons que 0 soit un point régulier de f, *i.e.* que la matrice jacobienne $J(f)$ de f en 0 soit de rang maximal. On dira alors que σ est régulier. D'après le théorème des fonctions implicites, f est triviale (immersion, submersion ou difféomorphisme local) et est déterminée par son 1-jet (*cf.* section 2.). Soit ρJ^k l'ensemble des éléments réguliers de J^k. L^k opère *transitivement* sur ρJ^k, ρJ^k étant l'orbite

[37] Nous noterons $J^k(m,n)$ simplement par J^k lorsqu'il n'y aura pas de risque de confusion sur les dimensions m et n.

de l'application $y_1 = x_1, \ldots, y_n = x_n$ si $m \geq n$ (submersion) ou de l'application $y_1 = x_1, \ldots, y_m = x_m, y_{m+1} = 0, \ldots, y_n = 0$ si $m \leq n$ (immersion). ρJ^k est ouvert et dense dans J^k et donc, si l'on dit qu'une orbite S est *incidente* à une orbite S' si $S \subset \overline{S}'$, toutes les orbites de J^k sont incidentes à ρJ^k. ρJ^k s'appelle la *strate régulière* de J^k.

Le problème est de stratifier les espaces J^k par des stratifications naturelles possédant un sens *géométrique* pour les singularités des applications.

8.1.2 Le cas de J^1

Considérons par exemple l'espace J^1. C'est l'espace des matrices $m \times n$. Si $\sigma \in J^1$, $\sigma = (\sigma_{i,j})$ $(i = 1, \ldots, m; j = 1, \ldots, n)$, un représentant de σ est une application $f : (\mathbb{R}^m, 0) \to (\mathbb{R}^n, 0)$ dont la matrice jacobienne $J(f)$ en 0 est égale à σ. Or nous avons vu qu'une des données essentielles de la structure locale de f en 0 est le *rang* de sa matrice jacobienne. Il est donc naturel de stratifier J^1 par le rang de ses éléments. Soit $p = \min(m, n)$. Le rang de $\sigma \in J^1$ peut aller de p (rang maximal) à 0. Soit S_i l'ensemble des $\sigma \in J^1$ de rang $p - i$. $S_0 = \rho J^1$ est la strate régulière de J^1 et $S_p = \{0\}$. Les S_i constituent clairement une *partition* de J^1 par des sous-ensembles L^1-invariants (puisque le rang d'une matrice est un invariant pour l'action de L^1). En fait :

Proposition 44. – $S_0, S_1, \ldots, S_p$ sont les orbites de J^1. $\overline{S}_i = \bigcup_j S_{i+j}$, $j = 0, \ldots, p - j$ et S_i est une sous-variété plongée de J^1 de codimension $c_i = (m-p+i)(n-p+i)$, i.e. $(m-n+i)i$ si $m \geq n$ et $i(n-m+i)$ si $m \leq n$. Qui plus est, si $j^1 f \in S_i$ il existe des coordonnées locales $(x_1, \ldots, x_{p-i}, x_{p-i+1}, \ldots, x_m)$ et $(y_1, \ldots, y_{p-i}, y_{p-i+1}, \ldots, y_n)$ telles que f s'écrive sous la forme :

$$\begin{cases} y_i = x_i \quad \text{pour } i = 1, \ldots, p - i \ , \\ y_j = \varphi_j(x) \quad \textit{pour } j = p - i + 1, \ldots, n \ , \end{cases}$$

avec $\dfrac{\partial \varphi_j}{\partial x_\ell}(0) = 0$ pour $j = p - i + 1, \ldots, n$ et $\ell = 1, \ldots, m$. Dans ces coordonnées, la matrice jacobienne $J(f)$ est de la forme

$$\begin{pmatrix} I_{p-i} & 0 \\ 0 & 0 \end{pmatrix}$$

On remarquera que si $j^1 f \notin S_0$, f n'est pas déterminée par son 1-jet puisque φ est arbitraire. On remarquera aussi que les fermetures $\overline{S}_i$ des S_i sont des ensembles *algébriques* de J^1 puisqu'ils sont définis par l'annulation des mineurs des matrices $\sigma \in J^1$. Qui plus est, le lieu singulier de $\overline{S}_i$ est $\overline{S}_{i+1}$ et $S_i = \overline{S}_i - \overline{S}_{i+1}$.

8.1.3 Le cas de J^2 pour les fonctions

Tentons maintenant de stratifier J^2. Soit $\pi = \pi_{2,1}$ la projection canonique $J^2 \to J^1$. Soit $S_i^2 = \pi^{-1}(S_i)$. Considérons d'abord le cas $n = 1$ des *fonctions* $f : (\mathbb{R}^m, 0) \to (\mathbb{R}, 0)$. J^1 ne comprend que deux strates ρJ^1 et $\{0\}$. En effet, les

matrices jacobiennes $J^1(f)$ étant des matrices lignes $\sigma = (\sigma_1, \ldots, \sigma_m)$ elles sont soit de rang 1 (et f est une submersion) soit de rang 0 (et 0 est point critique de f). Un élément $\tau \in J^2$ est un couple $\tau = (\sigma, H)$ où $\sigma \in J^1$ et H est une forme quadratique, à savoir le hessien de f, $H = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$. Comme $S_0^2 = \pi^{-1}(S_0) = \pi^{-1}(\rho J^1)$ est la strate régulière de J^2 sur laquelle les f sont déterminées par leur 1-jet, il reste à considérer l'action de L^2 sur $S_1^2 = \{(0, H) \in J^2\}$. Cela revient à classer les *formes quadratiques* sous l'action de L^2. Pour cela, il faut d'abord connaître l'action de $L^2(m, 1)$. H est la forme quadratique intervenant dans le développement de Taylor $f(0) + \sigma \cdot x + x^t H x + O(3)$ où x est le vecteur colonne $(x_1, \ldots, x_m)$, x^t le vecteur ligne transposé et $O(3)$ indique les termes de degré ≥ 3. Soit $\psi \in L^2(1)$. C'est le 2-jet d'une fonction $g : (\mathbb{R}, 0) \to (\mathbb{R}, 0)$ définie par $g(y) = uy + O(2)$ avec $u \neq 0$ car g doit être inversible en 0. Soit $\varphi \in L^2(m)$. C'est le 2-jet d'une application $h : (\mathbb{R}^m, 0) \to (\mathbb{R}^m, 0)$ définie par $h(x) = Ax + O(2)$ avec A inversible. On vérifie que l'action de (φ, ψ) sur $(0, H)$ donne $(0, u \cdot A^t H A)$. Il faut donc classer les formes quadratiques H (*i.e.* les matrices symétriques (h_{ij}) $i, j = 1, \ldots, m$) en considérant comme équivalentes H et $u \cdot A^t H A$ où $u \neq 0$ et $A \in GL(m, \mathbb{R})$. Cette classification possède deux invariants :

(i) *Le rang* de H. Si r est le rang de H, il existe un sous-espace de dimension r de $\mathbb{R}^m$ sur lequel la restriction H_r de H est une forme quadratrique non dégénérée.

(ii) Si le rang de H est r, le second invariant est *l'indice* i de H_r, *i.e.* le nombre de signes négatifs intervenant dans la forme normale de H_r lorsque l'on choisit des axes principaux.

On a donc une décomposition de $J^2(m, 1)$ en orbites :

$$J^2 = \rho J^2 \cup H_{r,i}, \quad r = 0, 1, \ldots, m; \; i = 0, 1, \ldots r.$$

Remarquons que le théorème de Morse (théorème 20) dit que si $H(f) \in H_{m,i}$ (point critique non dégénéré), f est déterminée par son 2-jet.

8.1.4 Le cas général de J^2

Dans le cas où n est quelconque, nous allons voir qu'il peut y avoir une *infinité non dénombrable* d'orbites dans $J^2(m, n)$. Considérons la strate la plus singulière S_p de J^1, *i.e.* le point (0) (matrice jacobienne $J(f) = 0$). Soit $\Sigma = S_p^2 = \pi^{-1}(S_p)$. La décomposition de Σ en orbites fournit le renseignement essentiel sur la structure de J^2. Elle indique en effet comment se lève la dégénérescence des f admettant en 0 un point "maximalement" critique. Si $F \in \Sigma$, $F = (0, H_1, \ldots, H_n)$ où $0 \in J^1$ et H_i, $i = 1, \ldots, n$, sont les hessiens en $0 \in \mathbb{R}^m$ des composantes $f_1, \ldots, f_n$ de f. Soient $\varphi \in L^2(m)$ et $\psi \in L^2(n)$. Comme φ et ψ sont des germes de difféomorphismes, on peut les

linéariser par changement de coordonnées locales et donc écrire $\varphi(x) = U^{-1}(x)$ et $\psi(y) = V(y)$ avec $U \in GL(m, \mathbb{R})$ et $V \in GL(n, \mathbb{R})$. Il est facile de vérifier que, sous l'action de φ et ψ, F devient $F' = (0, H'_1, \ldots, H'_n)$ avec $H'_k = \sum_{j=1}^{n} V_{k,j} U^t H_j U$. Un simple calcul de dimensions montre alors qu'il peut exister une infinité non dénombrable d'orbites dans Σ. En effet la dimension de Σ est $\dfrac{m(m+1)n}{2}$ ($\dfrac{m(m+1)}{2}$ dimensions pour chaque hessien symétrique H_i). D'autre part, la dimension d'une orbite est nécessairement inférieure à celle de $L^1(m) \times L^1(n)$, *i.e.* $m^2 + n^2$. Si donc

$$m^2 + n^2 < \frac{mn(m+1)}{2}$$

il y a une infinité non dénombrable d'orbites dans Σ. Remarquons que si $n = 1$, cette inégalité est fausse. Elle équivaut en effet à $m^2 - m + 2 < 0$, or $m^2 - m + 2$ est toujours positif.

Nous voyons donc que la situation devient vite assez peu maniable. Bien que L^k opère algébriquement sur J^k, il peut se faire que certaines orbites aient une fermeture topologique qui ne soit pas un sous-ensemble algébrique de J^k. Tel est par exemple le cas des orbites $H_{r,i}$ de $J^2(m, 1)$. $\overline{H_{r,i}}$ n'est pas un ensemble algébrique de $J^2(m, 1)$ car le fait d'être d'indice i n'est pas définissable par des équations polynômiales des coefficients de la matrice H.[38]

D'où l'idée suivante de René Thom. Soit $f : (\mathbb{R}^m, 0) \to (\mathbb{R}^n, 0)$. Considérons d'abord son 1-jet. Nous avons vu que l'espace J^1 est "bien stratifié" par ses orbites $S_0, S_1, \ldots, S_p$ sous L^1. Supposons que $j^1 f(0) \in S_i$ et que $j^1 f$ soit *transverse* sur S_i. On dira alors que f présente transversalement la singularité S_i en 0. Cela ne sera possible que sous certaines conditions de dimension et de codimension et sera toujours le cas si f est stable. Si f présente la singularité S_i transversalement, alors $S_i(f) = j^1 f^{-1}(S_i)$ est une sous-variété (en fait un germe de sous-variété) de $(\mathbb{R}^m, 0)$ de même codimension dans $\mathbb{R}^m$ que S_i dans J^1 d'après la proposition 11. On peut donc définir la *restriction* $\widehat{f}_i = f|_{S_i}$ de f à S_i. $\widehat{f}_i : S_i \to \mathbb{R}^n$ est une nouvelle application différentiable et l'on peut considérer son 1-jet en 0 . La strate à laquelle appartient $j^1 \widehat{f}_i(0)$ dépend du 2-jet de f et l'on va chercher des sous-variétés $S_{i,j}$ de $J^2(m, n)$ possédant la propriété suivante : si $j^1 f$ présente transversalement S_i en 0 alors $j^1 \widehat{f}_i(0)$ appartient à la strate S_j de $J^1(S_i(f), \mathbb{R}^n)_{(0,0)}$ si et seulement si $j^2 f(0)$ appartient à $S_{i,j}$.

La théorie de Thom-Boardman consiste d'abord à construire les $S_{i,j}$ qui sont donc des *types universaux* de singularités, puis à *itérer* le processus afin de construire des $S_{i,j,k\ldots}$ eux aussi universaux. Pour l'expliciter, esquissons la façon dont elle intervient en dimensions $m = n = 2$. Il s'agit là en effet d'un cas déjà traité dans les années 1940 par Hassler Whitney et qui est à l'origine du "programme de Thom".

[38] Pour des précisions sur cette difficulté, *cf.* Levine [16].

8.2 Le théorème de Whitney : plis et cusps

Soit donc $f : (\mathbb{R}^2, 0) \to (\mathbb{R}^2, 0)$ un germe d'application du plan sur le plan. Considérons d'abord son 1-jet. $J^1(2,2)$ est l'espace vectoriel de dimension 4 des matrices 2×2 : $\sigma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Il est décomposé en 3 strates par l'action de $L^1(2,2)$.

(i) La strate régulière $S_0 = \rho J^1$ des matrices σ de rang 2, *i.e.* inversibles. f est alors un difféomorphisme local d'après le théorème 3 d'inversion locale.

(ii) La strate S_1 des matrices de rang 1. La fermeture $\overline{S_1}$ de S_1 est l'ensemble des matrices de déterminant nul. C'est la sous-variété algébrique de $J^1 \simeq \mathbb{R}^4$ d'équation $ad - bc = 0$. Elle admet une singularité à l'origine et est de codimension 1.

(iii) La strate $S_2 = \{0\}$ de la matrice nulle.

Comme S_2 est de dimension 0 et donc de codimension 4 et que $\mathbb{R}^2$ est de dimension 2, $j^1 f$ ne peut intersecter transversalement S_2 que si elle *l'évite*. Autrement dit, f ne peut présenter transversalement que la singularité S_1. Si tel est le cas, $S_1(f)$ est une sous-variété de codimension 1 ($= \operatorname{codim} S_1$) de $\mathbb{R}^2$, *i.e.* une courbe régulière Σ de $\mathbb{R}^2$. Considérons alors la restriction $\widehat{f_1}$ de f à Σ, $\widehat{f_1} : \Sigma \to \mathbb{R}^2$. Elle nous ramène à la stratification de $J^1(1,2)$. $J^1(1,2)$ est le vectoriel de dimension 2 des couples $\left(\frac{\partial h_1}{\partial x}, \frac{\partial h_2}{\partial x}\right)$ où x est une coordonnée locale sur $(\Sigma, 0)$ et (h_1, h_2) sont les composantes d'une application $h : (\Sigma, 0) \to (\mathbb{R}^2, 0)$. $J^1(1,2)$ comprend deux strates:

(i) la strate S_0 des matrices de rang 1 (plan épointé de l'origine);

(ii) la strate $S_1 = \{0\}$.

Comme application *quelconque* de Σ dans $\mathbb{R}^2$, $\widehat{f_1}$ ne pourrait pas présenter transversalement S_1 car $\operatorname{codim} S_1 = 2 > \dim \Sigma = 1$. Mais $\widehat{f_1}$ n'est pas quelconque. Supposons d'abord que 0 soit un point régulier de $\widehat{f_1} : \Sigma \to \mathbb{R}^2$, *i.e.*, si l'on suppose les $S_{i,j}$ construits, que $0 \in S_{1,0}(f)$. $J^2(2,2)$ est un espace de dimension $4 + 3 + 3 = 10$ et le 2-jet de f est donné par (J, H_1, H_2) où J est la matrice jacobienne et H_1 et H_2 sont les hessiens des composantes (f_1, f_2) de f. Dire que $0 \in S_1(2,2)$ c'est dire que J est de rang 1 et donc que l'application linéaire tangente $D_0 f$ admet un noyau unidimensionnel K. Il y a par suite en $0 \in M = \mathbb{R}^2$, deux sous-espaces unidimensionnels de $T_0 M$: d'une part $T_0\Sigma$ et d'autre part $K = \operatorname{Ker} D_0 f$. D'où une double possibilité (*cf.* figure 4).

1. T_0 et K sont transverses en 0;
2. T_0 et K sont confondus en 0.

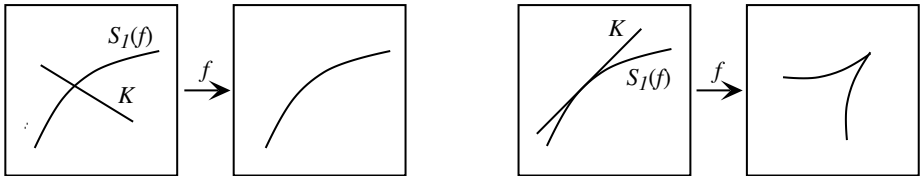


Figure 4: Les deux possibilités pour une singularité S_1 présentée transversalement.

Montrons que le cas 1 équivaut au fait que $0 \in S_{1,0}(f)$. f sera alors une submersion avec plis en 0 (*cf.* l'exemple 5 de la section 6.1). Dire que $0 \in S_{1,0}(f)$ c'est dire que $\widehat{f_1} : \Sigma \to \mathbb{R}^2$ est de rang 1 en 0, *i.e.* une *immersion.* Mais cela signifie précisément que $T_0\Sigma$ n'est pas le noyau K de $D_0 f$, *i.e.* que $T_0\Sigma$ et K sont transverses.

Théorème 45 (Whitney). – Si $0 \in S_{1,0}(f)$, on peut trouver des coordonnées à la source (x, y) et des coordonnées au but (X, Y) telles que f s'écrive sous la forme normale $X = x^2$ et $Y = y$.

Preuve. – $\widehat{f_1}$ étant une immersion, $f(\Sigma)$ est une courbe régulière de $N = \mathbb{R}^2$ et on peut donc la prendre pour coordonnée Y (*i.e.* la définir par $X = 0$). Comme $\widehat{f_1} : \Sigma \to f(\Sigma)$ est un difféomorphisme, on peut choisir y et Y de façon à ce que f s'exprime relativement à y par $Y = y$ et que Σ soit d'équation $x = 0$. Dans ces coordonnées f prend la forme: $(x, y) \to (g(x, y), y)$ et la matrice jacobienne J de f s'écrit $\begin{pmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ 0 & 1 \end{pmatrix}$. Elle est de rang 1 si $\frac{\partial g}{\partial x} = 0$. Il faut donc que $\frac{\partial g}{\partial x}$ s'annule sur Σ, *i.e.* pour $x = 0$, alors que, par construction, g s'annule pour $x = 0$. Cela implique que l'on puisse prendre $g = x^2$. $\square$

Comme pour les points critiques non dégénérés en théorie de Morse (*cf.* Théorème 20), il existe ainsi une *forme normale* pour les points singuliers de type $S_{1,0}$ des applications du plan dans le plan qui présentent S_1 transversalement. De tels points singuliers s'appellent des singularités *pli.* Cette appellation est naturelle si l'on se réfère à la figure 5.

Considérons maintenant le cas 2 où $T_0\Sigma$ et $K = \operatorname{Ker} D_0 f$ sont confondus, c'est-à-dire où $0 \in S_{1,1}(f)$. Dire que f présente transversalement la singularité $S_{1,1}$ en 0 c'est dire que le contact (la tangence) en 0 de Σ et du champ $\operatorname{Ker} D_x f$ pour $x \in \Sigma$ est un contact simple. Dans ce cas on a le :

Théorème 46 (Whitney). – Si f présente transversalement en 0 une singularité $S_{1,1}$ (ce qui est une condition sur son 3 -jet), on peut trouver des coordonnées à la source (x, y) et des coordonnées au but (X, Y) telles que f s'écrive sous la forme normale $X = xy + x^3$, $Y = y$.

Preuve. – Comme f est de rang 1 en 0, on peut choisir (x, y) et (X, Y) de façon à ce que f s'écrive sous la forme $Y = y$ et $X = h(x, y)$ avec $\frac{\partial h}{\partial x}(0) = \frac{\partial h}{\partial y}(0) = 0$ et $D\left(\frac{\partial h}{\partial x}\right) \neq 0$ (car sinon f ne présenterait pas S_1 transversalement).

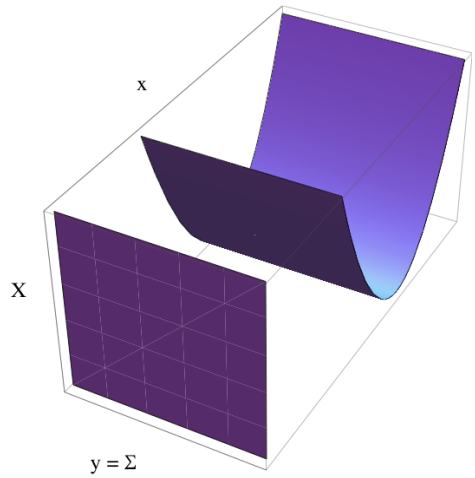


Figure 5: La singularité pli (la projection est la projection sur le plan (y, X) parallèlement à la direction horizontale x).

Comme la matrice jacobienne est la matrice $\begin{pmatrix} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \\ 0 & 1 \end{pmatrix}$ de déterminant $\frac{\partial h}{\partial x}$, $S_1(f)$ est la courbe $\frac{\partial h}{\partial x} = 0$. $\operatorname{Ker} D_0 f$ est engendré par $\frac{\partial}{\partial x}$. Comme $T_0\Sigma = \operatorname{Ker} D_0 f$, on a $\frac{\partial^2 h}{\partial x^2}(0) = 0$ et, pour que $S_{1,1}$ soit présentée transversalement (contact simple entre Σ et le champ $\operatorname{Ker} D_x f$), il faut que l'on ait $\frac{\partial^3 h}{\partial x^3}(0) \neq 0$. h est donc une fonction telle que :

1. $h(0,0) = 0$,
2. $\dfrac{\partial h}{\partial x}(0) = \dfrac{\partial h}{\partial y}(0) = 0$,
3. $\dfrac{\partial^2 h}{\partial x^2}(0) = 0$ et
4. $\dfrac{\partial^3 h}{\partial x^3}(0) \neq 0$.

Soit Γ l'anneau des germes en (0) d'applications $a : \mathbb{R}^2 \to \mathbb{R}^2$. Via f, Γ devient un Γ-module défini de la façon suivante : si $a \in \Gamma$ comme anneau et $b \in \Gamma$ comme Γ-module, on pose $a \cdot b(x,y) = a(f(x,y)) \cdot b(x,y)$. D'après les propriétés de h, le quotient $\Gamma / ((y,h) + \mathfrak{m}^4)$ est engendré par $1, x$ et x^2. D'après le théorème de Malgrange, Γ est donc engendré par $1, x$ et x^2 comme Γ-module. On a en particulier:

$$x^3 = 3a_2(y,h)x^2 + a_1(y,h)x + a_0(y,h)$$

où a_0, a_1 et a_2 s'annulent en 0. Autrement dit, on peut écrire $(x-a)^3+b(x-a) = c$ où a, b, c sont 3 fonctions de (y, h) s'annulant en 0. Or il est facile de vérifier que si l'on prend $x' = x - a(y, h)$, $y' = b(y, h)$, $X' = c(X, Y)$ et $Y' = b(X, Y)$ on obtient bien de nouvelles coordonnées. Dans ces coordonnées, f s'écrit sous la forme normale de Whitney. □

Une singularité $S_{1,1}$ présentée transversalement s'appelle un *point cusp*. Les singularités cusp sont nécessairement *isolées* pour des raisons de codimension.

Ces résultats se globalisent facilement. Et comme nous avons indiqué dans l'exemple 5 de la section 6.1 qu'une submersion avec plis n'est stable que si l'image $f(S_1(f))$ de son ensemble de points plis est une immersion à croisements normaux, nous arrivons à une caractérisation géométrique complète des applications stables en dimensions $m = 2$, $n = 2$.

Théorème 47 (Whitney-Thom). – Soit $f : M \to N$ une application structurellement stable entre surfaces. Alors f ne peut posséder comme singularités que des singularités pli et des singularités cusp. Les premières constituent des courbes régulières Σ de M dont les secondes sont des points isolés. L'image de Σ par f est un ensemble de courbes de N n'admettant comme seules singularités que des cusps et des croisements normaux.

8.3 La construction des S_I et les travaux de Boardman

Le théorème 23 de Morse résout à un premier niveau le problème de la structure des applications $f : M \to N$ pour m quelconque et $n = 1$. Le théorème 47 de Whitney-Thom le résout pour $m = 2$ et $n = 2$. Il s'agit de généraliser ces résultats aux dimensions quelconques.

Le théorème fondamental de Boardman consiste à montrer que, pour tout k-indice $I = (i_1, \ldots, i_k)$ avec $i_1 \geq i_2 \geq \ldots \geq i_k \geq 0$, on peut définir un sous-fibré universel S_I du fibré $J^k(M, N) \to M \times N$ ayant la propriété que si $j^\ell f$ est transverse sur tous les S_J où J est un ℓ-indice, $\ell < k$, alors $j^k f(x) \in S_I$ équivaut à $x \in S_{i_k}(S_{i_{k-1}}(\ldots(S_{i_1}(f))))$. Sa démonstration étant très technique, nous ne l'aborderons pas ici. Nous nous bornerons à en indiquer le cadre conceptuel (théorème des extensions jacobiennes).[39]

8.3.1 L'espace $J^\infty(M, N)$ des jets d'ordre infini et le fibré tangent $\mathcal{D}$

Comme le montre le théorème de Whitney, si $I = (i_1, \ldots, i_k)$, si $I' = (i_1, \ldots, i_{k-1})$ et si $j^{k-1}f(a) \in S_{I'}$, le fait que que $j^{k-1}f$ soit transverse sur $S_{I'}$ en $j^{k-1}f(a)$ est une condition portant sur le k-jet $j^k f$ de f. Si l'on veut pouvoir définir par récurrence les S_I quelle que soit la longueur $|I| = k$ de I, il faut donc pouvoir passer à volonté de J^k à J^{k+1}. Cela conduit naturellement, si l'on veut une formulation élégante et conceptuelle du problème, à travailler dans l'espace des jets d'ordre infini $J^\infty(M, N)$ que nous noterons simplement $J(M, N)$ ou même J. Cet espace J n'est plus un fibré de dimension finie dont les fibres sont des

[39] *Cf.* Boardman [6].

espaces de polynômes. C'est un fibré de dimension dénombrable dont les fibres sont des espaces de *séries formelles* (les séries de Taylor). En négligeant les détails techniques, nous ferons l'hypothèse que l'on peut y travailler comme dans un espace de jets de dimension finie.

D'autre part, si l'on considère l'application jet $j^k f : M \to J^k(M,N)$, elle admet une application linéaire tangente $D_a\left(j^k f\right) : T_a M \to T_\sigma J$ en tout point $a \in M$, avec $\sigma = j^k f\,(a)$. Il est clair qu'il existe une relation entre $D\left(j^k f\right)$ et le $(k+1)$-jet $j^{k+1} f$. Ensuite, les vecteurs tangents aux espaces de jets qui interviennent sont toujours des vecteurs tangents particuliers, images par des $D\left(j^k f\right)$ de vecteurs tangents à M, *i.e.* obtenus obtenus à partir de dérivations en les seules variables $(x_1, \ldots, x_m)$. Enfin, lorsque l'on travaille dans $J(M,N)$ on travaille à chaque fois dans un $J^k(M,N)$ et si la considération de $J(M,N)$ est nécessaire c'est que l'on ne peut pas borner k a priori.

Ces remarques ont conduit Boardman à considérer sur $J(M,N)$ un sous-espace $\mathcal{D}$ du fibré tangent TJ qui soit adapté à la situation. Pour le présenter de façon commode, il est utile, $(x_1, \ldots, x_m)$ et $(y_1, \ldots, y_n)$ étant des systèmes de coordonnées, de traiter les x_i et les y_j comme des fonctions coordonnées $x_i : M \to \mathbb{R}$ et $y_j : N \to \mathbb{R}$. Les composantes $(f_1, \ldots, f_n)$ de f ne sont alors rien d'autre que les composées $f_j = y_j \circ f$, $M \xrightarrow{f} N \xrightarrow{y_j} \mathbb{R}$.[40]

8.3.2 Les fonctions différentiables sur $J^\infty(M,N)$

Appelons alors *fonction différentiable* sur un ouvert U de l'espace de dimension infinie J, une application $\varphi : U \to \mathbb{R}$ qui se factorise localement par un J^k, *i.e.* telle que $\varphi = \psi \circ \pi_k$ où $\pi_k : J \to J^k$ est la projection canonique (troncage des jets à l'ordre k) et $\psi : V \to \mathbb{R}$ une application différentiable d'un ouvert $V \supset \pi_k\,(U)$ de J^k dans $\mathbb{R}$. Un telle factorisation doit exister pout tout $\sigma \in U$, mais le nombre k qui y intervient peut ne pas être borné lorsque σ parcourt U.

Notons $\mathcal{F}\,(U)$ l'espace des fonctions différentiables sur $U \subset J$. Les champs de vecteurs tangents sur U associés à l'espace tangent $\mathcal{D}$ qu'il s'agit de construire seront interprétés comme des dérivations sur $\mathcal{F}\,(U)$. Or remarquons que si π_M et π_N sont les projections canoniques $\pi_M : J \to M$ (application "source" des jets) et $\pi_N : J \to N$ (application "but" des jets), les fonctions de J dans $\mathbb{R}$ données par

$$\begin{cases} X_i = x_i \circ \pi_M,\ i = 1, \ldots, m \\ Y_j = y_j \circ \pi_N,\ j = 1, \ldots, n \\ Z_{j,\alpha}\sigma = d^\alpha\,(y_j \circ \sigma)\,(a)\,,\ j = 1, \ldots, n \end{cases}$$

où $\alpha = (\alpha_1, \ldots, \alpha_m)$ est un multi-indice et où d^α est la dérivée partielle $\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \circ \ldots \circ \frac{\partial^{\alpha_m}}{\partial x_m^{\alpha_m}}$, sont des coordonnées locales de J au voisinage du jet $\sigma \in J_a$.

[40] Pour simplifier les notations, nous faisons comme si les coordonnées étaient globales mais évidemment on localise quand c'est nécessaire.

L'espace tangent $\mathcal{D}$ est l'espace engendré par les seules dérivations $D_i = \frac{\partial}{\partial x_i}$, $i = 1, \ldots, m$, opérant sur les coordonnées par

$$\begin{cases} D_i X_j = \delta_{ij} \\ D_i Z_{j,\alpha} = Z_{j,\beta} \end{cases} \tag{16}$$

où β est égal à α à ceci près que α_i y est remplacé par α_{i+1}. Autrement dit, les sections de $\mathcal{D}$ sur un ouvert U de J (*i.e.* les champs de vecteurs sur U) forment un $\mathcal{F}(U)$-module libre engendré sur $\mathcal{F}(U)$ par les D_i : $D \in \mathcal{D}$ si seulement si $D = \sum_{i=1}^{i=m} \varphi_i D_i$ avec $\varphi_i \in \mathcal{F}(U)$. On remarquera que $Z_{j,\alpha} = D^\alpha Y_j$ et que, si l'on considère des multi-indices α tels que $\alpha_1 + \ldots + \alpha_m \leq k$, les X_i, Y_j et $Z_{j,\alpha}$ constituent des coordonnées locales en $\pi_k(\sigma)$ de J^k.

Le fibré vectoriel $\mathcal{D}$ sur $J(M, N)$ n'est ainsi rien d'autre que l'image réciproque par l'application "source" $\pi_M : J \to M$ du fibré tangent TM : $\mathcal{D} = \pi_M^*(TM)$

$$\begin{array}{ccc} \mathcal{D} & \longrightarrow & TM \\ \downarrow & & \downarrow \\ J & \xrightarrow{\pi_M} & M \end{array}$$

Si V est un ouvert de M et d un champ de vecteurs sur V, il lui correspond un champ $D \in \mathcal{D}(U)$ (où $U = \pi_M^{-1}(V)$) défini par $D\varphi \circ jf = d(\varphi \circ jf)$ avec $\varphi \in \mathcal{F}(U)$. On a $D = \pi_M^*(d)$ et cette correspondance définit l'application linéaire tangente du jet jf que nous noterons δjf pour éviter des confusions de notations

$$\begin{array}{ccc} TM & \underset{\longleftarrow}{\overset{\delta jf}{\longrightarrow}} & \mathcal{D} \\ \downarrow & & \downarrow \\ M & \underset{\xleftarrow{\pi_M}}{\overset{jf}{\longrightarrow}} & J \end{array}$$

Si $\xi \in T_a M$ est un vecteur tangent en a à M, si $\xi = \sum_{i=1}^{i=m} a_i d_i$ (où $d_i = \frac{\partial}{\partial x_i}$), on a $\delta jf(\xi) = \sum_{i=1}^{i=m} a_i D_i$.

8.3.3 Rang et corang

Après avoir ainsi défini le fibré tangent $\mathcal{D}$, Boardman généralise les notions de *rang* et de *corang* d'une application. Soit $f : V \to N$ une application différentiable d'un ouvert Vde M dans N ($N = \mathbb{R}^n$). La donnée de f équivaut à celle de ses composantes $(f_1, \ldots, f_n)$, *i.e.* à un ensemble de cardinal n de *fonctions* $f_i : V \to \mathbb{R}$. Réciproquement, si A est un ensemble de cardinal $|A|$ de fonctions $h_\alpha : V \to \mathbb{R}$, on peut lui associer une application $h_A : V \to \mathbb{R}^{|A|}$. Le rang $r_a A$ et le corang $c_a A$ de A en $a \in M$ seront alors ceux de l'application linéaire tangente $T_a h_A$, cette définition gardant un sens même si A est infini. Evidemment, $r_a A + c_a A = m$.

Il en va de même sur $J(M,N)$. Si $\sigma \in U \subset J$, où U est un ouvert de J, et si A est un sous-ensemble de $\mathcal{F}(U)$, on lui associe une application $h_A : U \to \mathbb{R}^{|A|}$ ainsi que son application linéaire tangente (au sens de $\mathcal{D}$) $Th_A : \mathcal{D} \to \mathbb{R}^{|A|}$. Le rang et le corang de A en σ sont alors ceux de l'application linéaire $T_\sigma h_A$. On a encore, étant donné que $\mathcal{D} = \pi_M^*(TM)$ est engendré par les D_i comme $\mathcal{F}(U)$-module, $r_\sigma A + c_\sigma A = m$. On voit que cette construction "relève" dans $(J, \mathcal{D})$ ce qui se passe dans (M, TM). Soient $\sigma \in U \subset J$, $a = \pi_M(\sigma)$, $b = \pi_N(\sigma)$ et $f : (M,a) \to (N,b)$ telle que $jf = \sigma$. Soit A un sous-ensemble de $\mathcal{F}(U)$ et $jf^*(A)$ son image réciproque par $jf : M \to \mathcal{F}$. Alors $r_\sigma A = r_a(jf^*(A))$ et $c_\sigma A = c_a(jf^*(A))$.

La notion de rang permet de définir l'*indépendance* des fonctions $\varphi \in \mathcal{F}(U)$. Si $\varphi_1, \ldots, \varphi_k \in \mathcal{F}(U)$, on dira qu'elles sont indépendantes en σ si le rang de leur système en σ est maximal, *i.e.* égal à k.

Lemme 1. – Si $\varphi_1, \ldots, \varphi_m \in \mathcal{F}(U)$ sont indépendantes sur un ouvert U de J (i.e. *en tout $\sigma \in U$), alors il existe m champs de vecteurs $D_1, \ldots, D_m \in \mathcal{D}(U)$ qui satisfont $D_i\varphi_j = \delta_{ij}$ et forment une base de $\mathcal{D}(U)$ comme $\mathcal{F}(U)$-module. Autrement dit, on peut substituer les fonctions φ_i aux fonctions coordonnées X_i.*

8.3.4 Extensions jacobiennes

La première définition des ensembles universels S_I se fait alors à travers ce que l'on appelle les *extensions jacobiennes*. Ces extensions expriment en termes d'algèbre linéaire différentielle le fait que les $S_{i_1 \cdots i_k}(f)$ sont des

$$S_{i_k}\left(S_{i_{k-1}}\left(\cdots\left(S_{i_1}(f)\right)\right)\right),$$

autrement dit, le fait que l'on considère des itérations de restrictions de f à des $S_{I'}$, et que, à chaque fois, on utilise uniquement la stratification de l'espace des 1-jets , *i.e.* le rang des matrices jacobiennes.

Pour comprendre comment s'introduisent naturellement ces extensions jacobiennes, il faut d'abord démontrer le résultat suivant. Considérons une sous-variété W de $J(M,N)$. Soient $a \in M$ et $f : M \to N$ telle que l'application jet jf soit *transverse* sur W en a. Au voisinage de a, l'image réciproque $Z = jf^{-1}(W)$ est une sous-variété de M de même codimension que celle de W dans $J(M,N)$. Or, étant donnés une sous-variété W d'une variété P et un point $x \in W$, il est naturel de considérer l'idéal de l'anneau $\Gamma_x(P)$ des germes en x de fonctions $h : P \to \mathbb{R}$ constitué des germes qui s'annulent sur W. Cet idéal contient en effet l'information fonctionnelle traduisant la structure de W en x.

Soit donc $\sigma = jf(a)$ et $\mathfrak{A}$ l'idéal de $\mathcal{F}(\sigma)$ (l'espaces des germes en σ de fonctions différentiables $h : J \to \mathbb{R}$) associé à W. L'idéal $\mathfrak{A}$ est l'espace des germes en σ de fonctions différentiables $h : J \to \mathbb{R}$ s'annulant sur W.

Lemme 2. – L'application linéaire tangente en a de la restriction de f

à $Z = jf^{-1}(W)$ *a un corang égal au corang en* σ *de l'idéal* $\mathfrak{A} + \pi_N^*(\mathfrak{m}_b)$ $(b = f(a))$.

Preuve. – En effet, en σ, W est l'image réciproque par $\pi_k : J \to J^k$ d'une sous-variété W^k de J^k (redescente en dimension finie). La transversalité implique que $jf^*(\mathfrak{A})$ engendre l'idéal des germes en a de fonctions $h : M \to \mathbb{R}$ s'annulant sur Z. Mais $c_\sigma(\mathfrak{A} + \pi_N^*(\mathfrak{m}_b)) = c_a(jf^*(\mathfrak{A}) + f^*(\mathfrak{m}_b))$ où $\mathfrak{m}_b$ est l'idéal maximal de l'anneau local $\Gamma_b(N)$. Or un vecteur $\xi \in T_aM$ n'annule $jf^*(\mathfrak{A})$ que s'il est tangent à Z et n'annule $f^*(\mathfrak{m}_b)$ que s'il appartient au noyau de T_af. Les $\xi \in T_aM$ qui s'annulent sur $jf^*(\mathfrak{A}) + f^*(\mathfrak{m}_b)$ sont donc ceux du noyau de l'application linéaire tangente en a de la restriction de f à la sous-variété Z. □

Ce lemme justifie en partie la définition suivante des extensions jacobiennes, la justification devenant complète à travers leur théorème de structure (théorème 50).

Définition 48. – *Soit* A *un sous-ensemble de* $\mathcal{F}(U)$. *Sa* k*-ième extension jacobienne* $\Delta^k A$ *est par définition l'idéal de* $\mathcal{F}(U)$ *engendré par* A *et par tous les mineurs* $\ell \times \ell$ *de type* $\det(D_i\alpha_j)$ *avec* $D_i \in \mathcal{D}(U)$, $\alpha_j \in A$, $i, j = 1, \ldots, \ell$ *et* $\ell = m - k + 1$.

Soient alors $\sigma \in J$ et $b = \pi_N(\sigma)$ son but. A travers $\pi_N : J \to N$, on peut considérer $\mathfrak{m}_b$ comme un sous-espace de $\mathcal{F}(\sigma)$. Les extensions jacobiennes $\Delta^{i_1}\mathfrak{m}_b$, $\Delta^{i_2}\Delta^{i_1}\mathfrak{m}_b$, etc. sont donc bien définies.

Définition 49. – *Soit* $I = \{i_1, \ldots, i_k\}$ *un multi-indice. L'ensemble universel* S_I *est l'ensemble des jets* $\sigma \in J$ *tels que l'on ait* $c_\sigma(\mathfrak{m}_b) = i_1$, $c_\sigma(\Delta^{i_1}\mathfrak{m}_b) = i_2$, $\ldots$, $c_\sigma(\Delta^{i_{k-1}} \ldots \Delta^{i_1}\mathfrak{m}_b) = i_k$.

8.3.5 Exemples simples

Cette définition est évidemment très abstraite. On peut l'expliciter dans les cas simples. Soit d'abord $I = \{i\}$. Un jet $\sigma \in J$ sera un élément de S_i si et seulement si $c_\sigma(\mathfrak{m}_b) = i$. Mais $c_\sigma(\mathfrak{m}_b) = c_a(jf^*(\mathfrak{m}_b))$. Or, étant donné que l'on a considéré $\mathfrak{m}_b$ comme plongé dans $\mathfrak{m}_\sigma$ à travers π_N et que $\pi_N \circ jf = f$, on a $jf^*(\mathfrak{m}_b) = f^*(\mathfrak{m}_b)$. Il s'agit donc de calculer le corang en $a \in M$ de l'idéal de l'image réciproque de $\mathfrak{m}_b$ par f. Mais ce corang est tout simplement celui de l'application linéaire tangente D_af de f en a et donc $\sigma \in S_i$ si et seulement si D_af est de corang i, *i.e.* si et seulement si $a \in S_i(f)$. On retrouve ainsi la définition précédente des S_i mais en ayant fait le détour par les idéaux $\mathfrak{m}_a$, $\mathfrak{m}_b$ et $f^*(\mathfrak{m}_b)$ et par leur "relèvement" dans $\mathcal{F}(\sigma)$.

Considérons maintenant les ensembles universels de type S_{ij} correspondant à $I = \{i, j\}$. Pour qu'un jet σ appartienne à S_{ij}, il faut $a \in S_i(f)$ et $c_\sigma(\Delta^i\mathfrak{m}_b) = j$. Par définition, $\Delta^i\mathfrak{m}_b$ est l'idéal de $\mathcal{F}(\sigma)$ engendré par $\mathfrak{m}_b$ et par les déterminants $\ell \times \ell$ $\det(D_k\alpha_h)$ avec $\alpha_h \in \mathfrak{m}_b$, $D_k \in \mathcal{D}(\sigma)$ et $\ell = m - i + 1$. Puisque sur $\mathfrak{m}_b$ les D_k s'identifient aux d_k qui sont les dérivations en les seules variables $x_1, \ldots, x_m$, on voit qu'il s'agit bien de passer de fonctions à des ma-

trices jacobiennes.

Cette définition des S_I par des extensions jacobiennes n'est évidemment pas très maniable. Le second point important développé par Boardman concerne la possibilité de les exprimer de façon plus simple en faisant apparaître leur *stratification* naturelle. Il s'agit là du point le plus technique.

Remarquons d'abord que l'on peut considérablement *restreindre* le nombre de générateurs d'une extension jacobienne. Cela est important car dans la définition des S_I on use d'extension jacobiennes d'*idéaux*.

Lemme 3. – Soit $A \subset \mathcal{F}(U)$ et $\mathfrak{A}$ l'idéal de $\mathcal{F}(U)$ engendré par A. Soit Ω un système générateur de $\mathcal{D}(U)$ comme $\mathcal{F}(U)$-module. L'extension jacobienne $\Delta^k\mathfrak{A}$ est engendrée par A et les déterminants $\det(D_i\alpha_j)$ *avec* $D_i \in \Omega$ et $\alpha_j \in A$.

Si nous revenons alors au cas des S_{ij}, on voit que, puisque $\mathfrak{m}_b$ est engendré par les fonctions coordonnées y_j et que les $\frac{\partial}{\partial x_i}$ engendrent $\mathcal{D}(U)$ comme $\mathcal{F}(U)$-module, il s'agit de calculer le corang en $a \in M$ de l'idéal engendré par les composantes de f et certains déterminants $(m-i+1)\times(m-i+1)$ de leurs dérivées partielles du premier ordre.

8.3.6 Les théorèmes de structure

Le lemme 3 permet de démontrer le théorème fondamental de structure des extensions jacobiennes. Soit C un sous-ensemble de $\mathcal{F}(U)$ composé de $m-k$ éléments *indépendants* sur U : $C = \{\varphi_1, \ldots, \varphi_{m-k}\}$. Considérons les champs tangents sur U (au sens de $\mathcal{D}$) qui annihilent C, *i.e.* les D_i tels que $D_i\varphi_j = 0$. Ces champs constituent un sous-fibré K de $\mathcal{D}\,|_U$ dont la fibre est de dimension k. En effet, les fonctions φ_j sont des coordonnées locales puisqu'elles sont indépendantes. Supposons qu'elles soient égales à $X_1, \ldots, X_{m-k}$. Alors K est engendré par les $\frac{\partial}{\partial x_1}, \ldots \frac{\partial}{\partial x_{m-k}}$. Soit $\Gamma_K = \Gamma_K(U)$ le $\mathcal{F}(U)$-module des sections de K.

Théorème 50 (structure des extensions jacobiennes). – Pour tout idéal $\mathfrak{A}$ de $\mathcal{F}(U)$ engendré par un ensemble A contenant C, on a $\Delta^k\mathfrak{A} = \mathfrak{A} + \Gamma_K\mathfrak{A}$. Qui plus est, l'idéal $\Gamma_K\mathfrak{A}$ est indépendant du choix de C.

Preuve. – En effet, écrivons $\mathcal{D} = K \oplus L$ avec L un supplémentaire du sous-fibré K. Les fonctions φ_j étant indépendantes, il existe des champs sur U, $D_1, \ldots, D_{m-k}$, tels que $D_i\varphi_j = \delta_{ij}$ pour $i, j = 1, \ldots, m-k$. Ces champs forment une base de L sur U. Appliquons le lemme 3. On n'utilise donc que les fonctions de A, les champs D_i et les champs de Γ_K dans les déterminants $(m-k+1)\times(m-k+1)$. Mais comme il n'existe que $(m-k)$ D_i indépendants, dans tous ces déterminants il faut au moins une colonne composée de $D\alpha$ avec $D \in \Gamma_K$ et donc tous ces déterminants sont dans $\Gamma_K\mathfrak{A}$. En fait on obtient ainsi tout $\Gamma_K\mathfrak{A}$. □

Ce théorème permet d'expliciter la structure stratifiée des ensembles universels S_I définis par les corangs des extensions jacobiennes successives de

l'idéal $\mathfrak{m}_b$. L'idée est de considérer des suites indexées par I de sous-fibrés Γ_K annihilant des éléments indépendants de $\mathcal{F}(U)$.

*Définition 51. – Soient $I = \{i_1, \ldots, i_k\}$ et U un ouvert de J. Un I-*drapeau *sur U est une suite $\mathcal{D}\mid_U = K_0 \supset K_1 \supset \cdots \supset K_k$ de sous-fibrés de $\mathcal{D}\mid_U$ tels qu'il existe des $C_1, \ldots, C_k$ de cardinal $|C_r| = i_{r-1} - i_r$ satisfaisant*

1. *K_r a des fibres de dimension i_r;*
2. *K_r est le fibré des champs annulant $C_1 \cup \cdots \cup C_r$;*
3. *$C_1 \cup \cdots \cup C_r$ est un système indépendant sur U;*
4. *$C_1 \subset \Gamma(\pi_N(U))$ et $C_r \subset \Gamma_{r-1} \cdots \Gamma_1 \Gamma(\pi_N U)$ (où Γ_r est le module des sections de K_r);*
5. *les fonctions de C_r sont en fait des fonctions sur $\pi_{r-1}(U) \subset J^{r-1}$.*

Etant donné un I-drapeau K sur U, on lui associe son *idéal* $\mathfrak{A}_K$ défini par

$$\mathfrak{A}_K = \Gamma_1 \Gamma(\pi_N U) + \Gamma_2 \Gamma_1 \Gamma(\pi_N U) + \cdots + \Gamma_{k-1} \cdots \Gamma_1 \Gamma(\pi_N U)$$

On dit que K est *nul* en $\sigma \in U$ si $\mathfrak{A}_K \subset \mathfrak{m}_\sigma$.

Théorème 52. – Soient $\sigma \in U$ et *$b = \pi_N(\sigma)$. Soit K un I-drapeau sur U. Alors $\Delta^I \mathfrak{m}_b = \mathfrak{A}_K + \mathcal{F}(U).\mathfrak{m}_b$.*

Ce théorème donne la structure des extensions jacobiennes successives intervenant dans la définition des ensembles S_I.

Théorème 53. – Un jet σ appartient à S_I si et seulement si il existe un I-drapeau K sur U nul en σ. Tout I-drapeau sur U est alors nul en σ.

Ces deux théorèmes fournissent la structure des ensembles universels S_I et impliquent qu'ils ne sont non vides que si $i_1 \geq \cdots \geq i_k \geq 0$, $m \geq i_1 \geq m - n$ et, dans le cas où $i_1 = m - n$, si $i_1 = \cdots = i_k$. A partir de là, il reste à montrer que si $f : M \to N$ est une application telle que jf soit transverse sur S_I, alors $S_{I,j}(f) = S_j\left(f\mid_{S_I(f)}\right)$. Pour cela, il faut disposer d'outils assez raffinés et en particulier de ce que l'on appelle les *dérivées intrinsèques* (rappelons que les dérivées partielles ne sont pas intrinsèques).

8.3.7 La stratification de l'espace source

Soit $f : M \to N$ une application stable. Elle satisfait les conditions de transversalité du théorème de Boardman et les sous-variétés $S_I(f) = j^k f^{-1}(S_I)$ sont donc bien définies. Qui plus est, elle satisfait une condition de type "croisements normaux" analogue à celle que nous avons rencontrée en 6.1 à propos des submersions avec plis. Plus précisément, si $I_1, \ldots, I_s$ sont des multi-indices, si $x_i \in S_{I_i}(f)$, $i = 1, \ldots, s$, sont des points de même image y respectivement de type $I_1, \ldots, I_s$ et si T_i est l'espace tangent à $S_{I_i}(f)$ en x_i, les sous-espaces

$D_{x_1}f(T_1), \ldots, D_{x_s}f(T_s)$ doivent être en position générale dans T_yN. Si l'on appelle *application de Boardman* une application $f : M \to N$ satisfaisant les conditions de transversalité permettant de définir les $S_I(f)$ ainsi que cette condition de croisements normaux, on a donc la

Proposition 54. – Si $f : M \to N$ est une application stable alors c'est une application de Boardman.

Evidemment, il serait particulièrement intéressant de pouvoir montrer la réciproque puisque cela caractériserait géométriquement les applications stables. Mais cela est impossible pour des raisons que Golubitsky et Guillemin explicitent de la façon suivante. Si f est une application de Boardman, f stratifie son espace source M, les strates de la stratification étant :

(i) les composantes connexes de l'ensemble des points réguliers $M - \bigcup_i S_i(f)$;

(ii) les ensembles de Boardman $S_{i_1,\ldots,i_{k-1},0}(f)$.

Restreinte à chaque strate, f est localement triviale, *i.e.* est une submersion ou une immersion à croisements normaux. f peut donc être considérée comme un "recollement" de ses restrictions localement triviales aux strates. Pour démontrer que f est structurellement stable, l'idée serait de partir de g voisine de f et, en utilisant la stabilité des propriétés de transversalité, de montrer que la stratification de g est équivalente à celle de f. Mais cela n'impliquerait pas pour autant que g soit équivalente à f car la connaissance de la stratification de f et de son type (submersion ou immersion à croisements normaux) sur chaque strate ne suffit pas à caractériser f à équivalence près. Il suffit par exemple de considérer les fonctions de Morse $g, f : \mathbb{R}^2 \to \mathbb{R}$ données par $f(x, y) = x^2 + y^2$ et $g(x, y) = x^2 - y^2$. Dans les deux cas, la stratification se réduit à la strate régulière $\mathbb{R}^2 - \{0\}$ sur laquelle f et g sont des submersions et à la strate singulière $\{0\}$ sur laquelle f et g sont des immersions. Pourtant f et g n'ont même pas le même type topologique puisque pour f le point critique non dégénéré est un minimum alors que pour g il est un col.

Si la stratification de Thom-Boardman n'est pas suffisante pour caractériser le type différentiable d'une application c'est qu'elle est définie en termes d'extensions jacobiennes et que celles-ci ne permettent pas d'exprimer des invariants comme par exemple la signature du hessien en un point critique non dégénéré d'une fonction de Morse. Rappelons que c'était d'ailleurs cette même difficulté qui conduisait dans la stratification de J^2 à des orbites dont les fermetures topologiques n'étaient pas des ensembles algébriques de J^2 (*cf.* section 8.1.4).

9 Codimension et détermination

Dans les précédentes sections, nous avons indiqué d'une part comment la stabilité structurelle pouvait se caractériser comme stabilité par déformations et comme stabilité infinitésimale, c'est-à-dire par la surjectivité de certains morphismes mixtes et d'autre part comment les applications de Boardman (et

donc en particulier les applications stables) se décomposaient, via la stratification de leur espace source, en "composantes" localement triviales. Dans les sections qui suivent, nous tenterons de préciser les autres idées directrices du paradigme morphodynamique et en particulier les notions de détermination, de codimension, de modèle transverse et de déploiement universel. Cela nous conduira à la classification et à la géométrie des catastrophes élémentaires.

9.1 Le problème général

Rappelons la situation générale des modèles statiques (*cf.* section 1.2.3). On considère un champ $\sigma : W \to \mathcal{F}$ envoyant un espace externe W dans l'espace fonctionnel $\mathcal{F} = C^\infty(M, \mathbb{R})$ des potentiels sur un espace interne M ($\mathcal{F}$ est un espace de Fréchet si M est compacte) et l'on cherche la trace sur W de l'ensemble catastrophique intrinsèque $K_\mathcal{F}$ constitué des fonctions structurellement instables.

On supposera M *compacte* (ce qui permettra d'utiliser le théorème de stabilité 27 de Thom-Mather) et, dans un premier temps, on se localisera dans W au voisinage d'un point catastrophique $w_0 \in \sigma^{-1}(K_\mathcal{F}) \cap W$. Soit $f = f_0 = f_{w_0}$. D'après le théorème 23 de Morse, une fonction $f : M \to \mathbb{R}$ est stable si et seulement si c'est une fonction de Morse excellente. Comme $f \in K_\mathcal{F}$, f admet des points critiques dégénérés et/ou des valeurs critiques égales. Il s'agit d'étudier la structure de $\mathcal{F}$ au voisinage de f. Une telle étude ne pourra s'effectuer correctement que dans le cas où :

(i) l'orbite $\tilde{f}$ de f sous l'action de $G = \mathrm{Diff}(M) \times \mathrm{Diff}(\mathbb{R})$ admet en f une section transverse de dimension finie (cas dit de *codimension finie*);

(ii) l'on peut "descendre" dans un espace de jets $J^k(M, \mathbb{R})$ où la situation devient algébrique (cas dit de *détermination finie*).

Si tel est le cas, on se trouvera ramené à l'analyse des modèles transverses à $j^k f$ dans J^k (théorie des *déploiements universels*), *i.e.* à l'analyse de germes de déformations f_T de base $(T, 0)$ où T est de dimension égale à la codimension de f.

Si, munis des résultats d'une telle étude, on revient à la situation générale $\sigma : W \to \mathcal{F}$ équivalente à la donnée d'une déformation f_w, $w \in W$, de $f_0 = f$, on peut considérer dans $M \times W$ le sous-espace Σ constitué des points critiques des f_w. Si la déformation f_w est générique, *i.e.* si σ est *transverse* sur $K_\mathcal{F}$, les f_w seront génériquement des fonctions de Morse excellentes et Σ sera donc un espace de même dimension que W constitué de nappes au-dessus de W engendrées par la variation des points critiques de f_w lorsque w varie dans W. Cette situation généralise celle en dimension 2 traitée par le théorème de Whitney (théorèmes 45, 46, 47). La restriction à Σ de la projection $\pi : M \times W \to W$ s'appelle *l'application catastrophique* de la déformation f_w et l'on peut espérer arriver, dans les cas simples, à une classification des

applications catastrophiques ainsi qu'à des formes normales. C'est ce parcours que nous allons esquisser. Il repose essentiellement sur la traduction des situations rencontrées en termes d'anneaux locaux, d'idéaux et de modules et sur l'usage répétitif des deux outils algébriques fondamentaux que sont le lemme de Nakayama (théorème 39, lemme) et le théorème de préparation différentiable de Bernard Malgrange (théorème 39). Notre présentation suivra étroitement le texte de John Mather [21].

9.2 L'anneau local de f en a

Soit $f : M \to N$ une application différentiable *instable*. Pour étudier la structure de l'espace fonctionnel $\mathcal{F} = C^\infty(M, N)$ au voisinage de f, on étudie la contribution de ses points critiques à son instabilité et pour cela on se multilocalise en un ensemble fini $S = \{a^1, \ldots, a^s\}$ de points de M de même image $b = f(S) \in N$ et on étudie la structure, au voisinage du germe, noté δ_f, de f, de l'espace $\mathcal{G}_{S,b} = \mathcal{G}$ des germes en S de fonctions $g : M \to N$ telles que $g(S) = b$.

Pour simplifier on supposera dans la suite que $S = \{a\}$. Le passage à un S fini quelconque s'effectue sans difficulté.

Comme dans les sections précédentes, on travaillera avec les entités algébriques suivantes. D'abord l'anneau $\Gamma_a(M) = \Gamma_a$ des germes en a de fonctions $h : M \to \mathbb{R}$. C'est un anneau local d'idéal maximal $\mathfrak{m}_a$ constitué des $h : M \to \mathbb{R}$ s'annulant en a. Si $(x_1, \ldots, x_m)$ est un système de coordonnées locales en a, traitées comme des fonctions $x_i : U \subset M \to \mathbb{R}$, $\mathfrak{m}_a$ est engendré par les x_i (en fait les germes δ_{x_i}) et le germe $\delta_h \in \mathfrak{m}_a^k$ si et seulement si h et toutes ses dérivées partielles jusqu'à l'ordre $k-1$ s'annulent en a.[41] On associe également à f l'anneau local $\Gamma_b(N) = \Gamma_b$ d'idéal maximal $\mathfrak{m}_b$ et le morphisme d'anneaux $f^* : \Gamma_b \to \Gamma_a$ défini par la composition avec f qui envoie $\mathfrak{m}_b$ dans $\mathfrak{m}_a$ et transforme tout Γ_a-module en un Γ_b-module. L'anneau local quotient $Q_f(a) = \Gamma_a / f^*(\mathfrak{m}_b)\Gamma_a$ s'appelle l'*anneau local* de f en a. C'est un anneau local d'idéal maximal $\mathfrak{m}_f = \mathfrak{m}_a / f^*(\mathfrak{m}_b)\Gamma_a$ qui, géométriquement, est l'idéal des germes des fonctions $h : (M, a) \to \mathbb{R}$ qui s'annulent sur la fibre $f^{-1}(b)$. Autrement dit, $Q_f(a)$ est l'anneau local des restrictions des fonctions $h : (M, a) \to \mathbb{R}$ à la fibre $f^{-1}(b)$.

Par un choix de coordonnées locales $(x_1, \ldots, x_m)$ en a et $(y_1, \ldots, y_n)$ en b, $\Gamma_b(N)$ devient isomorphe à $\mathcal{E}_n$ et $\Gamma_a(M)$ à $\mathcal{E}_m$ et l'étude de la structure locale de f au voisinage de a et b est donc celle du morphisme d'anneaux locaux $f^* : \mathcal{E}_n \to \mathcal{E}_m$ faisant de $\mathcal{E}_m$ un $\mathcal{E}_n$-module et, en particulier, celle de l'anneau local $Q_f : \mathcal{E}_m / f^*(\mathfrak{m}_n)\mathcal{E}_m$.

Donnons quelques exemples d'anneaux Q_f.

1. Si $f : (M, a) \to (N, b)$ est une immersion ($m \leq n$) alors, d'après la

[41] Si $S = \{a^1, \ldots, a^s\}$, Γ_S est l'anneau produit direct des $\Gamma_i = \Gamma_{a_i}$ et l'idéal $\mathfrak{m}_S$ des germes en S de $h : M \to \mathbb{R}$ s'annulant en S est le produit direct des idéaux maximaux $\mathfrak{m}_i = \mathfrak{m}_{a_i}$.

Proposition 6, dans des coordonnées locales linéarisant f, f s'écrit sous la forme normale $y_i = x_i$ pour $i = 1, \ldots, m$ et $y_j = 0$ pour $j = m+1, \ldots, n$. On a donc $f^*(y_i) = y_i \circ f = x_i$ si $i = 1, \ldots, m$ et $f^*(y_j) = y_j \circ f = 0$ si $j = m+1, \ldots, n$. Par conséquent, $f^*(\mathfrak{m}_n) = \mathfrak{m}_m$, $\mathfrak{m}_f = 0$ et $Q_f : \mathcal{E}_m / f^*(\mathfrak{m}_n)\mathcal{E}_m = \mathcal{E}_m / \mathfrak{m}_m \simeq \mathbb{R}$.

2. Si $f : (M, a) \to (N, b)$ est une submersion ($m \geq n$) alors, d'après la Proposition 8, dans des coordonnées locales linéarisant f, f s'écrit sous la forme normale $y_i = x_i$ pour $i = 1, \ldots, n$. On a donc $f^*(y_i) = y_i \circ f = x_i$ si $i = 1, \ldots, n$ et par conséquent, $f^*(\mathfrak{m}_n)\mathcal{E}_m$ est l'idéal de $\mathcal{E}_m$ engendré par $(x_1, \ldots, x_n)$. $Q_f : \mathcal{E}_m / f^*(\mathfrak{m}_n)\mathcal{E}_m$ est donc l'anneau local des germes en 0 des fonctions des $(m-n)$ variables restantes $(x_{n+1}, \ldots, x_m) \simeq \mathcal{E}_{m-n}$.

3. Supposons maintenant que f soit une fonction de Morse avec un point critique non dégénéré en a avec une forme normale $H(x) = \sum_{i=1}^{i=m} \varepsilon_i x_i^2$ ($\varepsilon_i = \pm 1$, $i = 1, \ldots, m$). Alors $f^*(\mathfrak{m}_n)\mathcal{E}_m$ est l'idéal principal des fonctions $g \in \mathcal{E}_m$ qui s'écrivent comme produits $g = Hg'$ et Q_f s'obtient en annihilant ces fonctions. Comme $f^*(\mathfrak{m}_n)\mathcal{E}_m \subset \mathfrak{m}_m^2$, $\mathfrak{m}_f / \mathfrak{m}_f^2$ est isomorphe à $\mathfrak{m}_m / \mathfrak{m}_m^2$ et est donc l'espace vectoriel de dimension m des fonctions linéaires en les x_i. Quant à $\mathfrak{m}_f^2 / \mathfrak{m}_f^3$, c'est l'espace des fonctions quadratiques en les x_i quotienté par l'idéal principal (H). Comme le produit tensoriel de deux fonctions linéaires en les x_i est une forme quadratique en les x_i, le hessien H de f en 0 est le noyau de l'application linéaire $(\mathfrak{m}_m / \mathfrak{m}_m^2) \otimes (\mathfrak{m}_m / \mathfrak{m}_m^2) \to \mathfrak{m}_f^2 / \mathfrak{m}_f^3$ et est donc dérivable de l'anneau local Q_f. Cela montre que l'anneau local Q_f est un invariant *plus fin* que la stratification de Thom-Boardman.

4. Considérons maintenant un point pli a de $f : (\mathbb{R}^2, a) \to (\mathbb{R}^2, b)$. D'après le théorème 45 de Whitney, f peut s'écrire sous la forme normale $X = x^2, Y = y$ et donc $f^*(\mathfrak{m}_n)\mathcal{E}_m = (x^2, y)$. Cela implique $Q_f \simeq \mathbb{R}[x] / (x^2)$ et donc l'anneau local Q_f est l'espace vectoriel bidimensionnel des fonctions affines $h(x) = a_0 + a_1 x$ muni du produit

$$(a_0 + a_1 x)(b_0 + b_1 x) = a_0 b_0 + (a_0 b_1 + a_1 b_0) x$$

où x devient nilpotent ($x^2 = 0$). Quant à $\mathfrak{m}_f$, c'est l'idéal des $h(x) = a_1 x$ ($a_0 = 0$).

5. Si enfin a est un point cusp de $f : (\mathbb{R}^2, a) \to (\mathbb{R}^2, b)$, d'après le théorème 46 de Whitney, f peut s'écrire sous la forme normale $X = xy + x^3, Y = y$ et donc $f^*(\mathfrak{m}_n)\mathcal{E}_m = (xy + x^3, y)$. Mais l'idéal $(xy + x^3, y)$ est trivialement égal à l'idéal (x^3, y) et par conséquent $Q_f \simeq \mathbb{R}[x] / (x^3)$. Autrement dit, l'anneau local Q_f est l'espace vectoriel tridimensionnel des polynômes du second degré $h(x) = a_0 + a_1 x + a_2 x^2$ muni du produit

$$\begin{aligned}\left(a_0 + a_1 x + a_2 x^2\right)\left(b_0 + b_1 x + b_2 x^2\right) = \\ a_0 b_0 + (a_0 b_1 + a_1 b_0) x + (a_0 b_2 + a_1 b_1 + a_2 b_0) x^2\end{aligned}$$

où x devient nilpotent ($x^3 = 0$). Quant à $\mathfrak{m}_f$, c'est l'idéal des $h(x) = a_1 x + a_2 x^2$ ($a_0 = 0$).

Notons par ailleurs que les jets $j^k f(a)$ sont déductibles de f^*. En effet, comme $f^*(\mathfrak{m}_n) \subset \mathfrak{m}_m$, f^* passe au quotient et induit pour chaque k un morphisme $f_k^* : \mathcal{E}_n/\mathfrak{m}_n^{k+1} \to \mathcal{E}_m/\mathfrak{m}_m^{k+1}$ qui est une façon d'exprimer $j^k f(a)$. La correspondance $f_k^* \leftrightarrow j^k f(a)$ est un isomorphisme entre $J^k(M,N)_{a,b}$ et l'espace des morphismes d'anneaux locaux entre $\mathcal{E}_n/\mathfrak{m}_n^{k+1}$ et $\mathcal{E}_m/\mathfrak{m}_m^{k+1}$.

Notons également qu'il existe un isomorphisme linéaire canonique entre $\mathfrak{m}_m/\mathfrak{m}_m^2$ et T_a^*M le dual de l'espace tangent T_aM. Considérons en effet l'application linéaire $\theta : \mathfrak{m}_m \to T_a^*M$ qui associe au germe $\eta = \delta g$ d'une fonction $g : (M,a) \to \mathbb{R}$ l'application linéaire tangente $D_a g$ (qui est une forme linéaire sur T_aM puisque $T_{g(a)}\mathbb{R} \simeq \mathbb{R}$). Il est trivial de vérifier que θ est surjective car si $g = \sum_{i=1}^{i=m} x_i \frac{\partial g}{\partial x_i}(a) + \cdots$ est un élément de $\mathfrak{m}_m$, $\theta(g) = \sum_{i=1}^{i=m} \frac{\partial g}{\partial x_i}(a)\, dx_i = dg$ et réciproquement. Il est évident que le noyau de θ est $\mathfrak{m}_m^2$. A travers ces isomorphismes, le morphisme $f_1^* : T_b^*N \to T_a^*M$ s'identifie à l'application transposée de l'application linéaire tangente $D_a f : T_aM \to T_bN$.

Mather a fait de l'anneau local Q_f un usage systématique d'une grande efficacité. En effet ses quotients successifs $Q_f^k = Q_f/\mathfrak{m}_m^{k+1}$ (développements de Taylor des restrictions ci-dessus) fournissent des *invariants algébriques* qui sont plus fins, nous l'avons vu, que la stratification de Thom-Boardman (ils permettent par exemple de retrouver le hessien de f en a). Dans le cas de stabilité, ils permettent même de classifier les types différentiables. Nous verrons en effet plus bas (théorème 63) le résultat que si $Q_f^{n+1} \simeq Q_g^{n+1}$ alors f et g sont différentiablement équivalentes.

9.3 Equivalences, détermination et codimension

Une fois posés ces préliminaires, revenons à la notion de *détermination*. Rappelons d'abord :

Définition 55. – On dit que f est déterminée à l'ordre k en a si :

(i) *toute fonction g telle que $j^k g(a) = j^k f(a)$ est équivalente à f en a, i.e. si l'égalité des k-jets implique l'équivalence; et*

(ii) *si k est le plus petit entier satisfaisant cette propriété.*

On voit que la notion de détermination dépend de l'équivalence choisie. Or il existe plusieurs notions naturelles d'équivalence.

1. On dit que f et g sont d-équivalentes (équivalence à droite notée $f \underset{d}{\sim} g$) s'il existe un difféomorphisme $\varphi : (M,a) \to (M,a)$ tel que $g = \varphi^{-1} \circ f$. La d-équivalence est liée au groupe $G_M = \operatorname{Diff} M$ des changements de coordonnées à la source.

2. On dit que f et g sont g-équivalentes (équivalence à gauche notée $f \underset{g}{\sim} g$) s'il existe un difféomorphisme $\psi : (N,b) \to (N,b)$ tel que $g = \psi \circ f$. La g-équivalence est liée au groupe $G_N = \operatorname{Diff} N$ des changements de coordonnées au but.

3. L'équivalence différentiable, notée $f \sim g$, est la d-g-équivalence associée au groupe $G = G_M \times G_N$.

4. Il existe une autre équivalence, plus technique, dénommée par Mather *équivalence de contact* et correspondant à l'isomorphisme des anneaux locaux $Q_f(a)$ et $Q_g(a)$. On peut en donner une interprétation géométrique en disant que les graphes[42] respectifs Gr_f et Gr_g de f et g ont "même contact" avec $M_b = M \times \{b\}$ dans $M \times N$, *i.e.* qu'il existe un difféomorphisme local de $M \times N$ en (a, b) laissant fixe M_b et échangeant Gr_f et Gr_g. Cette équivalence est elle aussi définissable par l'action d'un groupe G_C sur $\mathcal{F} = C^\infty(M, N)$. En effet, on commence par considérer des familles de difféomorphismes $\psi_x \in G_N$ dépendant différentiablement de $x \in M$ et laissant b fixe. Une telle famille s'identifie à un difféomorphisme ψ_M de $M \times N$ compatible à la projection canonique $\pi : M \times N \to M$ et laissant $M \times \{b\}$ invariant point par point. Les ψ_M constituent un groupe G' agissant naturellement sur $\mathcal{F}$ par $\psi_M(f) = g : M \to N$ avec $g(x) = \psi_x(f(x))$.

Proposition 56. – f et g sont G'-équivalentes si et seulement si $f^*(\mathfrak{m}_b)\Gamma_a = g^*(\mathfrak{m}_b)\Gamma_a$.

Soit en effet $I_{a,b,f} = I_f$ l'idéal des germes en (a, b) de fonctions $h : M \times N \to \mathbb{R}$ s'annulant sur le graphe Gr_f. Soit i l'injection canonique $i : M \simeq M \times \{b\} \hookrightarrow M \times N$. A i se trouve associé le morphisme $i^* : \Gamma_{a,b}(M \times N) \to \Gamma_a(M)$. On montre d'abord que $i^*(I_f) = f^*(\mathfrak{m}_b)\Gamma_a$. En effet, étant donnée l'équation $y = f(x)$ de Gr_f, I_f est engendré dans $\Gamma_{a,b}(M \times N)$ par les applications $\pi_2^*(y_i) - \pi_1^*(f^*(y_i))$, $i = 1, \ldots, n$, où $\pi_1 : M \times N \to M$ et $\pi_2 : M \times N \to N$ sont les projections canoniques. Et comme $i^*(\pi_2^*(y_i) - \pi_1^*(f^*(y_i))) = -f^*(y_i)$, on a bien $i^*(I_f) = f^*(\mathfrak{m}_b)\Gamma_a$. Supposons ensuite que f et g soient G'-équivalentes. On a bien :

$$g^*(\mathfrak{m}_b)\Gamma_a = i^*(I_g) = i^*\psi_M(I_g) = i^*(I_f) = f^*(\mathfrak{m}_b)\Gamma_a \ .$$

Réciproquement, supposons que l'on ait $f^*(\mathfrak{m}_b)\Gamma_a = g^*(\mathfrak{m}_b)\Gamma_a$. Alors il existe des matrices $n \times n$, $U = (u_{ij})$ et $V = (v_{ij})$, sur Γ_a telles que $f^*(y_i) = \sum_{j=1}^{j=n} u_{ij} g^*(y_j)$ et $g^*(y_i) = \sum_{j=1}^{j=n} v_{ij} f^*(y_j)$. On peut supposer U inversible car il existe toujours une matrice C telle que $W = C(I - VU) + U$ soit inversible et l'on a $Wg^*(y) = Ug^*(y) = f^*(y)$. On peut donc remplacer s'il le faut U par W. Il est alors facile de voir que le difféomorphisme ψ_M donné par $\psi_x(y)_i = \sum_{j=1}^{j=n} u_{ij} y_j$ établit une équivalence entre f et g. □

5. La G'-équivalence est définie par des familles ψ_M de difféomorphismes de (N, b) paramétrés par M et ne tient donc pas compte des changements de coordonnées dans la source M. Le groupe G_C de l'équivalence de contact est le produit semi-direct de G' et de G_M dont les éléments sont des familles de

[42] Le graphe d'une application $f : M \to N$ est l'ensemble $Gr_f = \{(x, f(x))\}_{x \in M} \subset M \times N$. Il peut être vu comme l'application $\widehat{f} : M \to M \times N$ définie par $\widehat{f}(x) = (x, f(x))$.

difféomorphismes non pas de N_x mais entre N_x et $N_{\varphi(x)}$ avec $\varphi \in G_M$ (laissant a fixe). Si $\theta \in G_C$, il opère sur f par $\theta(f)(x) = \theta_{\varphi^{-1}(x)} f(\varphi^{-1}(x))$.

Définition 57. – f et g sont dites C-équivalentes (équivalence de contact) si g appartient à la G_C-orbite de f.

Cela équivaut à dire qu'il existe $\varphi \in G_M$ tel que $f \circ \varphi$ et g soient G'-équivalentes.

L'importance de l'équivalence de contact est d'être directement liée aux anneaux locaux $Q_f(a)$. En effet, dire que f et g sont C-équivalentes revient à dire que les graphes Gr_f et Gr_g ont *même contact* avec M_b dans $M \times N$ au sens de la définition suivante.

Définition 58. – Soient M, H_1 et H_2 trois sous-variétés équidimensionnelles en z d'une variété ambiante Z. On dit que H_1 et H_2 ont le même contact avec M en z si il existe un germe de difféomorphisme θ de (Z, z) induisant l'identité sur M et échangeant H_1 et H_2.

Soit $I_z(H_1)$ l'idéal des germes en z de fonctions sur Z s'annulant sur H_1. Si $i : M \hookrightarrow Z$ est l'inclusion canonique, i induit un morphisme d'anneaux entre $\Gamma_z(Z)$ et $\Gamma_z(M)$ et l'on pose $Q_{M,H_1} = \Gamma_z(M)/i^*(I_z(H_1))$. Q_{M,H_1} est un invariant du contact entre M et H_1 et f et g sont C-équivalentes si les anneaux $Q_{M_b,Gr_f}(a,b)$ et $Q_{M_b,Gr_g}(a,b)$ sont isomorphes. Or si $\widehat{f} : M \to M \times N$ est l'application graphe, il est facile de vérifier que $\widehat{f}^*\left(Q_{M_b,Gr_f}(a,b)\right) = Q_f(a)$. Autrement dit, f et g sont C-équivalentes si et seulement si leurs anneaux locaux sont isomorphes.

Nous sommes donc en présence de 5 groupes G_M, G_N, $G = G_M \times G_N$, G' et G_C définissant 5 types d'équivalence sur $\mathcal{F}$. Pour ces 5 groupes G_i nous appliquons la stratégie qui consiste à définir "l'espace tangent" en f à l'orbite $G_i(f)$ dans "l'espace tangent" en f à $\mathcal{F}$. "L'espace tangent" en f à $\mathcal{F}$ est l'espace vectoriel $\Gamma(f^*(TN))$ des champs de vecteurs sur N le long de f. Nous avons vu en 6.1 à propos de la définition de la stabilité infinitésimale que "l'espace tangent" à la G-orbite de f est donné par l'application linéaire $\theta_f + \Omega_f : \Gamma(TM) \oplus \Gamma(TN) \to \Gamma(f^*(TN))$ qui, à chaque couple (ξ, η) d'un champ sur M et d'un champ sur N, associe le champ sur N le long de f défini par $Df \circ \xi + \eta \circ f$. En ce qui concerne les G'-orbites, soit $\psi_{M,t}$ un germe de courbe à l'origine e de G'. C'est une famille $\psi_{x,t}$ de difféomorphismes de $N_x \subset M \times N$. Comme $\frac{\partial \psi_{x,t}}{\partial t}$ est un champ η_x sur N_x, un élément de $T_e G'$ est d'abord un champ sur $M \times N$ "vertical" relativement à la projection $\pi : M \times N \to M$. Soit $\Gamma(\pi)$ l'espace vectoriel de ces champs sur N le long de π. On veut de plus que les $\psi_{x,t}$ laissent b fixe. Cela signifie que η_x doit appartenir à $\mathfrak{m}_b \Gamma(\pi)$. On peut alors vérifier que, si $\gamma_f : (G', e) \to (\mathcal{F}, f)$ est l'application définissant la G'-orbite de f, un tel champ η_x est envoyé par l'application linéaire tangente $D_e \gamma_f$ sur un champ de $f^*(\mathfrak{m}_b)\Gamma(f^*(TN))$ et que $D_e \gamma_f$ est surjective sur cet espace qui s'identifie par conséquent à "l'espace tangent" en f à la G'-orbite de f. On montre de même que "l'espace tangent" en f à la G_C-orbite de f est l'espace

$\theta_f(\Gamma(TM)) + f^*(\mathfrak{m}_b)\Gamma(f^*(TN))$.

Connaissant les "espaces tangents" aux diverses orbites de f, il est alors naturel d'appeler *codimension* de f relativement à G_i la dimension du quotient de "l'espace tangent" total $T_f\mathcal{F}$ par "l'espace-tangent" à la G_i-orbite de f.

Définition 59. – Les différentes codimensions sont :

$$\begin{aligned}
\operatorname*{codim}_d f &= \dim_{\mathbb{R}} \Gamma(f^*(TN)) / \theta_f(\Gamma(TM)), \\
\operatorname*{codim}_g f &= \dim_{\mathbb{R}} \Gamma(f^*(TN)) / \Omega_f(\Gamma(TN)), \\
\operatorname{codim} f &= \dim_{\mathbb{R}} \Gamma(f^*(TN)) / (\theta_f(\Gamma(TM)) + \Omega_f(\Gamma(TN))), \\
\operatorname*{codim}_{G'} f &= \dim_{\mathbb{R}} \Gamma(f^*(TN)) / f^*(\mathfrak{m}_b)\Gamma(f^*(TN)), \\
\operatorname*{codim}_C f &= \dim_{\mathbb{R}} \Gamma(f^*(TN)) / (\theta_f(\Gamma(TM)) + f^*(\mathfrak{m}_b)\Gamma(f^*(TN))).
\end{aligned}$$

Le résultat fondamental en matière de codimension et de détermination finies dit que, quel que soit G_i, une application f est de détermination finie si et seulement si elle est de codimension finie. Dans le cas de codimension finie, on peut donc, puisqu'il y a détermination finie, "descendre" en dimension finie en se ramenant à des espaces de jets où la situation devient algébrique et calculable.

9.4 Le théorème fondamental de finitude

Théorème 60. – Soit $f \in \mathcal{F}$. f est de détermination finie relativement à G_i si et seulement si f est de codimension finie relativement à G_i.

Preuve. – Pour démontrer ce résultat, on montre d'abord que le lemme de Nakayama et le théorème de Malgrange permettent de caractériser facilement la codimension finie. Pour unifier et alléger les notations notons $\operatorname{codim}(f, G_i)$ la codimension de f relativement à G_i, Γ ou $\Gamma(f)$ l'espace $\Gamma(f^*(TN))$ et $K(G_i)$ le sous-espace par lequel on quotiente dans la définition 59 de la codimension. On a donc dans tous les cas

$$\operatorname{codim}(f, G_i) = \dim_{\mathbb{R}} \Gamma / K(G_i). \tag{17}$$

Proposition 61. – $\operatorname{codim}(f, G_i)$ *est finie si et seulement si il existe k tel que* $K(G_i) \supset \mathfrak{m}_a^k \Gamma$.

Preuve. – La condition est suffisante car $\Gamma/\mathfrak{m}_a^k\Gamma$ est l'espace des $(k-1)$-jets de germes en a de champs sur N le long de f et est donc un espace de dimension finie. Or si $K(G_i) \supset \mathfrak{m}_a^k\Gamma$ alors $\operatorname{codim}(f, G_i) \leq \dim_{\mathbb{R}} \Gamma/K(G_i)$.

La condition est nécessaire. Considérons par exemple le cas de l'équivalence de contact associée à G_C. D'après le lemme de Nakayama, il suffit de montrer que la codimension finie implique qu'il existe k tel que

$$\theta_f(\Gamma(TM)) + \left(f^*(\mathfrak{m}_b) + \mathfrak{m}_a^{k+1}\right)\Gamma \supset \mathfrak{m}_a^k\Gamma\ .$$

De même, dans le cas de l'équivalence différentiable associée à G, d'après le théorème e Malgrange, il suffit de montrer que la codimension finie implique qu'il existe ℓ (dépendant de k) tel que

$$\theta_f\left(\Gamma(TM)\right) + \Omega_f\left(\Gamma(TN)\right) + \mathfrak{m}_a^\ell \Gamma \supset \mathfrak{m}_a^k \Gamma \ .$$

Mais soit $c_p = \dim_{\mathbb{R}} \Gamma / \left(K(G_i) + \mathfrak{m}_a^p \Gamma\right)$. Les c_p constituent une suite infinie croissante d'entiers $\leq \operatorname{codim}(f, G_i)$. Si $\operatorname{codim}(f, G_i)$ est finie, cette suite est nécessairement stationnaire à partir d'un certain rang k ce qui implique les conditions ci-dessus. $\square$

A partir de là, montrons d'abord la suffisance dans le théorème 60, à savoir que si f est de codimension finie elle est de détermination finie. Prenons l'exemple de G_C (équivalence de contact). Il faut montrer que, si f et g ont même k-jet pour un certain k, alors elles sont G_C-équivalentes, et pour ce faire il faut construire un difféomorphisme assurant cette équivalence. Suivant une technique que nous avons déjà rencontrée, Mather considère une homotopie f_t, $t \in I = [0,1]$, allant de $f_0 = f$ à $f_1 = g$ et construit des homotopies des identités rendant f_t équivalente à f pour tout t. Pour cela il intègre des champs de vecteurs appropriés et utilise l'hypothèse de codimension finie pour montrer que de tels champs existent. Rappelons le calcul formel développé en 7.10 montrant que, si f_t est une homotopie telle qu'il existe des déformations φ_t de 1_M et ψ_t de 1_N satisfaisant $f = \psi_t \circ f_t \circ \varphi_t^{-1}$, alors on a

$$\frac{\partial \varphi_t^{-1}}{\partial t} \circ \varphi_t = \xi_t, \ \frac{\partial \psi_t^{-1}}{\partial t} \circ \psi_t = \eta_t, \ \frac{\partial f_t}{\partial t} = -Df_t \circ \xi_t + \eta_t \circ f_t$$

où ξ_t et η_t sont des champs "verticaux" définis respectivement sur $p_M : M \times I \to I$ et $p_N : N \times I \to I$.

Parmi les homotopies entre f et g la plus simple est $f_t = (1-t)\, f + tg$. Notons-la $F : (M \times I, a \times I) \to (N \times I, b \times I)$. Notons $F_N : (M \times I, a \times I) \to (N, b)$ la composée $p_N \circ F$. Si $h \in \Gamma_a(M)$, on peut l'identifier à la fonction égale à h sur toutes les fibres de $p_M : M \times I \to I$. Si ξ est un champ de vecteurs sur M, on peut de même l'identifer à un champ vertical égal à ξ sur toutes les fibres de $TM \times I \to I$, etc. D'autre part si ξ_t est une famille de champs sur M, on peut définir $\theta_F(\xi_t)$ qui est un champ sur N le long de F_N comme on a défini $\theta_f(\xi)$. De même pour $\Omega_F(\eta_t)$. On a alors le lemme d'approximation suivant :

Lemme. – Supposons que f et g aient le même k-jet.

1. *Si $\xi \in \Gamma(TM)$, alors $\theta_f(\xi) - \theta_F(\xi) \in \mathfrak{m}_a^{k-1}\Gamma(F_N)$.*
2. *Si $\eta \in \Gamma(TN)$, alors $\Omega_f(\eta) - \Omega_F(\eta) \in \mathfrak{m}_a^k \Gamma(F_N)$.*
3. *Si $h \in \Gamma_b(N)$, alors $f^*(h) - F^*(h) \in \mathfrak{m}_a^k \Gamma_{a,t}(M \times I)$.*

Ce lemme permet de montrer la suffisance dans le théorème 60. Supposons en effet que $\operatorname{codim}_C f = c < \infty$. Par hypothèse,

$$c = \dim_{\mathbb{R}} \Gamma(f) / \left(\theta_f(\Gamma(TM)) + f^*(\mathfrak{m}_b)\Gamma(f)\right) \ . \tag{18}$$

D'après le lemme de Nakayama, cela implique

$$\mathfrak{m}_a^c \Gamma(f) \subset \theta_f(\Gamma(TM)) + f^*(\mathfrak{m}_b)\Gamma(f) \ . \tag{19}$$

Si g a le même k-jet que f pour un $k > c$ assez grand et bien choisi, alors, d'après le lemme, on a :

$$\mathfrak{m}_a^c \Gamma(F_N) \subset \theta_F(\Gamma_{\mathbb{R}}(TM)) + F^*(\mathfrak{m}_b)\Gamma(F_N) + \mathfrak{m}_a^{c+1}\mathfrak{d}(F_N) \tag{20}$$

où $\Gamma_{\mathbb{R}}(TM)$ est l'espace des familles ξ_t de champs sur M. Toujours d'après le lemme de Nakayama, (20) implique

$$\mathfrak{m}_a^c \Gamma(F_N) \subset \theta_F(\Gamma_{\mathbb{R}}(TM)) + F^*(\mathfrak{m}_b)\Gamma(F_N) \tag{21}$$

et donc

$$\mathfrak{m}_a^k \Gamma(F_N) \subset \theta_F\left(\mathfrak{m}_a^{k-c}\Gamma_{\mathbb{R}}(TM)\right) + F^*(\mathfrak{m}_b)\,\mathfrak{m}_a^{k-c}\Gamma(F_N) \ . \tag{22}$$

Comme f et g ont le même k-jet, $\frac{\partial F}{\partial t} \in \mathfrak{m}_a^k \Gamma(F_N)$ et d'après (22) il existe $\xi_t \in \mathfrak{m}_a^{k-c}\Gamma_{\mathbb{R}}(TM)$ et $\widetilde{\eta}_t \in F^*(\mathfrak{m}_b)\,\mathfrak{m}_a^{k-c}\Gamma(F_N)$ tels que $\frac{\partial F}{\partial t} = \theta_F(\xi_t) + \widetilde{\eta}_t$. Mais comme $\Gamma(F_N)$ est engendré sur $\Gamma_{a,t}(M \times I)$ par $\frac{\partial}{\partial y_1} \circ F_N, \ldots, \frac{\partial}{\partial y_n} \circ F_N$, $\widetilde{\eta}_t$ peut s'écrire sous la forme

$$\widetilde{\eta}_t = \sum_{i,j=1}^{i,j=n} F^*(y_j)\, u_{ij} \left(\frac{\partial}{\partial y_i} \circ F_N\right), \ u_{ij} \in \mathfrak{m}_a^{k-c-1}\Gamma_{a,t}(M \times I) \ . \tag{23}$$

Si l'on considère la famille de champs sur $M \times N$, $\eta_t = \sum_{i,j=1}^{i,j=n} y_j u_{ij} \frac{\partial}{\partial y_i} - \xi_t$, et si l'on intègre les ξ_t et les η_t, on définit une famille d'éléments de G_C, $\left(\psi_{x,t}, \varphi_t\right)$, qui trivialise l'homotopie F. Les applications f et g sont donc C-équivalentes et f est de détermination finie relativement à l'équivalence de contact.

Pour démontrer la nécessité dans le théorème 60, nous prendrons l'exemple de l'équivalence différentiable associée au groupe G. Il faut montrer que si f est de détermination finie alors sa codimension est également finie. Soient k l'ordre de détermination de f, $\ell > k$ et $\pi : J^\ell \to J^k$ la projection canonique tronquant les ℓ-jets à l'ordre k. Soient σ le ℓ-jet de f et σ^k son k-jet. Comme f est k-déterminée, $\pi^{-1}\left(\sigma^k\right)$ est inclus dans l'orbite $\widetilde{\sigma}$ de σ dans J^ℓ sous l'action de G (qui est une action algébrique de groupe de Lie). John Mather définit alors une projection $\widetilde{\pi}$ de $\mathfrak{m}_a\Gamma(f)$ sur $T_\sigma J^\ell$ de la façon suivante. Soit $\zeta \in \mathfrak{m}_a\Gamma(f)$ un champ le long de f considéré comme dérivée $\frac{\partial F}{\partial t}\big|_{t=0}$ d'une déformation $F : M \times \mathbb{R} \to N \times \mathbb{R}$ de f. F définit une courbe σ_t dans J^ℓ et on pose $\widetilde{\pi}(\zeta) = \frac{d\sigma_t}{dt}\big|_{t=0}$. Il est facile de voir que le noyau de $\widetilde{\pi}$ est $\mathfrak{m}_a^{\ell+1}\Gamma(f)$.

Or l'espace tangent en σ à $\pi^{-1}\left(\sigma^k\right)$ est donné par $\widetilde{\pi}\left(\mathfrak{m}_a^{k+1}\Gamma\left(f\right)\right)$ et l'espace tangent en σ à $\widetilde{\sigma}$ est donné par $\widetilde{\pi}\left(\theta_f\left(\mathfrak{m}_a\Gamma(TM)\right)+\Omega_f\left(\mathfrak{m}_b\Gamma(TN)\right)\right)$. Cela implique

$$\mathfrak{m}_a^{k+1}\Gamma\left(f\right)\subset\theta_f\left(\mathfrak{m}_a\Gamma(TM)\right)+\Omega_f\left(\mathfrak{m}_b\Gamma(TN)\right)+\mathfrak{m}_a^{\ell+1}\Gamma\left(f\right) \tag{24}$$

ce qui, d'après le théorème de Malgrange, implique à son tour

$$\mathfrak{m}_a^{k+1}\Gamma\left(f\right)\subset\theta_f\left(\mathfrak{m}_a\Gamma(TM)\right)+\Omega_f\left(\mathfrak{m}_b\Gamma(TN)\right)$$

et par conséquent la détermination finie d'après la proposition 61. □

9.5 L'anneau local des applications stables

9.5.1 Détermination des applications stables

Soit $f:(M,a)\to(N,b)$. Si f est (infinitésimalement) stable on a $\theta_f\left(\Gamma(TM)\right)+\Omega_f\left(\Gamma(TN)\right)=\Gamma(f)$. Donc, d'après la définition 59 de la codimension, f est de G-codimension 0. D'après le théorème 60, f est de G-détermination finie. En fait on a le résultat beaucoup plus précis suivant :

Théorème 62 (Mather). – Si f est stable, f est déterminée à l'ordre $n+1$ relativement à l'équivalence différentiable.

Ce théorème généralise les théorèmes 23 de Morse et 47 de Whitney-Thom. Lorsque $n=1$ (théorème de Morse) une fonction n'est localement stable que si ses points critiques sont non dégénérés et, en ces points, f est effectivement déterminée à l'ordre $2=n+1$. Lorsque $n=2$ et $m=2$ (théorème de Whitney-Thom), une fonction f n'est (localement) stable que si ses points critiques sont des plis ou des cusps et, en ces points, f est effectivement déterminée à l'ordre $3=n+1$.

Preuve. – Esquissons la démonstration du théorème de Mather [21]. Soit $k\geq n+1$ l'ordre de G-détermination de f. Soit π la projection canonique $\pi:J^k\to J^{n+1}$ et V l'image réciproque $\pi^{-1}\left(j^{n+1}\left(f\right)\right)$ du jet $j^{n+1}\left(f\right)$ dans J^k. Comme f est k-déterminée, on peut se situer dans J^k, ce qui réduit le problème à la dimension finie. Pour montrer que f est $(n+1)$-déterminée, il faut montrer dans J^k que si g a le même $(n+1)$-jet que f, alors g est G-équivalente à f, autrement dit que le k-jet $j^k\left(g\right)$ de g appartient à l'orbite de $j^k\left(f\right)$ dans J^k sous l'action du groupe G^k des k-jets de difféomorphismes de G. Mais cela revient à montrer que V est inclus dans une orbite de cette action. Or on peut montrer le lemme technique suivant :

Lemme. – Soit $\pi:J\to J'$ une submersion compatible à l'action d'un groupe de Lie G sur les variétés J et J'. Soit $\sigma'\in J'$ et $V=\pi^{-1}\left(\sigma'\right)$. Pour que V soit inclus dans une orbite de G, il faut et il suffit que pour tout $\sigma\in V$ on ait $T_\sigma V\subset T_\sigma\widetilde{\sigma}$ où $\widetilde{\sigma}$ est la G-orbite de σ dans J.

Appliquons ce lemme. Si g est un représentant de $\sigma \in V$, l'espace tangent $T_\sigma \widetilde{\sigma}$ est donné par

$$T_\sigma \widetilde{\sigma} = \pi^k \left(\theta_g \left(\mathfrak{m}_a \Gamma(TM)\right) + \Omega_g \left(\mathfrak{m}_b \Gamma(TN)\right)\right) \tag{25}$$

où π^k est la projection naturelle $\pi^k : \mathfrak{m}_a \Gamma(g) \to T_\sigma J^k$. D'après (25), l'inclusion $T_\sigma V \subset T_\sigma \widetilde{\sigma}$ est équivalente à

$$\mathfrak{m}_a^{n+1} \Gamma(g) \subset \theta_g \left(\mathfrak{m}_a \Gamma(TM)\right) + \Omega_g \left(\mathfrak{m}_b \Gamma(TN)\right) + \mathfrak{m}_a^{k+1} \Gamma(g) \ . \tag{26}$$

En supposant que f est stable, il faut montrer que (26) est valable pour tout g de même $(n+1)$-jet que f. Or on sait déjà, d'après le théorème de Malgrange, que la stabilité de f ne dépend que de son $(n+1)$-jet. Si g a le même $(n+1)$-jet que f, alors g est également stable. Mais $c \subset g^* \left(\mathfrak{m}_b \Gamma(g)\right)$ et par conséquent

$$\operatorname{codim}(\theta_g \left(\Gamma(TM)\right) + g^* \left(\mathfrak{m}_b \Gamma(g)\right)) \leq \operatorname{codim}(\theta_g \left(\Gamma(TM)\right) + \Omega_g \left(\mathfrak{m}_b \Gamma(TN)\right))$$

la codimension étant calculée dans $\Gamma(g)$. Comme g est stable, on a $\Gamma(g) = \theta_g \left(\Gamma(TM)\right) + \Omega_g \left(\mathfrak{m}_b \Gamma(TN)\right)$ et donc $\operatorname{codim}(\theta_g \left(\Gamma(TM)\right) + \Omega_g \left(\mathfrak{m}_b \Gamma(TN)\right)) \leq n$. Mais cela implique $\operatorname{codim}(\theta_g \left(\Gamma(TM)\right) + g^* \left(\mathfrak{m}_b \Gamma(g)\right)) \leq n$ et par suite, d'après le lemme de Nakayama,

$$\mathfrak{m}_a^n \Gamma(g) \subset \theta_g \left(\Gamma(TM)\right) + g^* \left(\mathfrak{m}_b \Gamma(g)\right) \ . \tag{27}$$

On a donc, en multipliant par $\mathfrak{m}_a$,

$$\mathfrak{m}_a^{n+1} \Gamma(g) \subset \theta_g \left(\mathfrak{m}_a \Gamma(TM)\right) + g^* \left(\mathfrak{m}_b \Gamma(g)\right) \ .$$

Mais comme g est stable,

$$\begin{aligned} g^* \left(\mathfrak{m}_b \Gamma(g)\right) &= g^* \left(\mathfrak{m}_b \left(\theta_g \left(\Gamma(TM)\right) + \Omega_g \left(\Gamma(TN)\right)\right)\right) \\ g^* \left(\mathfrak{m}_b \Gamma(g)\right) &= \theta_g \left(g^* \left(\mathfrak{m}_b \Gamma(TM)\right)\right) + \Omega_g \left(\mathfrak{m}_b \Gamma(TN)\right) \end{aligned}$$

ce qui implique

$$\mathfrak{m}_a^{n+1} \Gamma(g) \subset \theta_g \left(\mathfrak{m}_a \Gamma(TM)\right) + \Omega_g \left(\mathfrak{m}_b \Gamma(TN)\right)$$

et donc a fortiori (26). □

9.5.2 Le rôle de l'anneau local

Nous avons déjà noté plus haut que l'anneau local de f en a, $Q_f(a) = \Gamma_a / f^* \left(\mathfrak{m}_b\right) \Gamma_a$, est associé à l'équivalence de contact de groupe G_C et que, si $Q_f^k = Q_f / \mathfrak{m}_a^{k+1}$, les germes $j^k f$ et $j^k g$ sont C-équivalents (pour G_C^k) si et seulement si $Q_f^k = Q_g^k$. Nous avions alors évoqué un résultat fondamental de John Mather [23] montrant que ces anneaux locaux permettent de *classifier* les germes d'applications *stables* pour l'équivalence différentiable.

Théorème 63. – Soient f et g deux applications stables. Si $Q_f^{n+1} \simeq Q_g^{n+1}$ alors f et g sont différentiablement équivalentes.

Preuve. – f et g étant stables, elles sont déterminées d'après le théorème 62 par leurs $(n+1)$-jets $\sigma = j^{n+1}f$ et $\delta = j^{n+1}g$. Si $Q_f^{n+1} \simeq Q_g^{n+1}$, σ et δ sont C-équivalents (sous l'action de G_C^{n+1} dans l'espace de jets J^{n+1}). Il faut montrer qu'ils sont aussi G-équivalents, *i.e.* qu'ils appartiennent à une même orbite de l'action de G^{n+1} dans J^{n+1}. Soit r le rang de f en a et choisissons des coordonnées locales $(x_1, \ldots, x_m)$ en a et $(y_1, \ldots, y_n)$ en b telles que f s'écrive sous la forme :

$$\begin{cases} y_i \circ f = f_i = x_i, \ i = 1, \ldots, r \\ d(f_j)(a) = 0, \ j = r+1, \ldots, n \end{cases} \tag{28}$$

Notons $\overline{x}$ les coordonnées $(x_1, \ldots, x_r)$ par rapport auxquelles f se réduit à l'identité et x' les coordonnées résiduelles $(x_{r+1}, \ldots, x_m)$. On peut ramener la situation à une situation ne portant que sur ces coordonnées résiduelles, *i.e.* à une situation où le rang r de f est nul. Notons en effet $\Gamma_{a'}$ l'anneau local des germes en 0 de fonctions en x' et considérons un module libre $\mathcal{E}$ à $n-r$ générateurs $\varepsilon_1, \ldots, \varepsilon_{n-r}$. Si $h \in \Gamma_a$, on peut lui associer canoniquement $h' \in \Gamma_{a'}$ défini par $h'(x_{r+1}, \ldots, x_m) = h(0, \ldots, 0, x_{r+1}, \ldots, x_m)$. Le Γ_a-module $\theta_f(\Gamma(TM)) + \Omega_f(\Gamma(TN))$ est engendré sur Γ_a par les $D_a(f)\frac{\partial}{\partial x_i}$ et les $\frac{\partial}{\partial y_j} \circ f$. Si f est de la forme (28), il est en fait engendré par les $D_a(f)\frac{\partial}{\partial x_i}$ pour $i = 1, \ldots, r$ et les $\frac{\partial}{\partial y_j} \circ f$ pour $j = r+1, \ldots, n$. Au champ

$$\zeta = \sum_{i=1}^{i=r} u_i D_a(f) \frac{\partial}{\partial x_i} + \sum_{j=r+1}^{j=n} u_j \frac{\partial}{\partial y_j} \circ f \in \Gamma(f)$$

on peut donc associer $\pi(\zeta) = \sum_{j=r+1}^{j=n} u'_j \varepsilon_{j-r} \in \mathcal{E}$, ce qui définit une application $\pi : \Gamma(f) \to \mathcal{E}$. Notons $f_\bullet = (f_{r+1}, \ldots, f_n)$ et $f'_\bullet = \left(f'_{r+1}, \ldots, f'_n\right)$ la réduction de f à $x' = (x_{r+1}, \ldots, x_m)$ et $y' = (y_{r+1}, \ldots, y_n)$. On peut exprimer la stabilité de f uniquement en termes de $f'_\bullet$. En effet, dire que f est stable c'est dire que pour $k \geq n$ on a :

$$\Gamma(f) = \theta_f(\Gamma(TM)) + \Omega_f(\Gamma(TN)) + \left(f^*(\mathfrak{m}_b) + \mathfrak{m}_a^{k+1}\right)\Gamma(f) \ . \tag{29}$$

Soit $\Omega(f'_\bullet) = \pi(\theta_f(\Gamma(TM)) + f^*(\mathfrak{m}_b)\Gamma(f))$. $\Omega(f'_\bullet)$ ne dépend que de $f'_\bullet$ et est le sous-$\Gamma_{a'}$-module de $\mathcal{E}$ engendré par les $\partial_i f = \sum_{j=r+1}^{j=n} \frac{\partial f'_j}{\partial x_i} \varepsilon_{j-r}$ pour $i = r+1, \ldots, n$ et les $f'_i \varepsilon_j$ pour $i = r+1, \ldots, n$, $j = 1, \ldots, n-r$. Soit ρ^k la projection canonique $\rho_k : \mathcal{E} \to \mathcal{E}/\mathfrak{m}_{a'}^{k+1}\mathcal{E} = \mathcal{E}^k$ (où $\mathfrak{m}_{a'}$ est l'idéal maximal de $\Gamma_{a'}$). Supposons que l'on ait

$$\Omega(f'_\bullet)^k + [\partial f]^k + [\varepsilon_1, \ldots, \varepsilon_{n-r}]^k = \mathcal{E}^k \tag{30}$$

où $[v_1, \ldots, v_s]$ désigne l'espace vectoriel engendré par les $v_i \in \mathcal{E}$ dans $\mathcal{E}$. Par image réciproque par $\rho_k \circ \pi$, (30) équivaut à (29), *i.e.* à la stabilité structurelle

de f. Mais comme on a toujours $\Omega\left(f'_\bullet\right)^k + \left[\partial f\right]^k \subset \mathfrak{m}_{a'}\mathcal{E}^k$, (30) équivaut en fait à (31) :

$$\Omega\left(f'_\bullet\right)^k + \left[\partial f\right]^k = \mathfrak{m}_{a'}\mathcal{E}^k \ . \tag{31}$$

D'où le lemme :

Lemme 1. – Si $k \geq n$ et si f est sous la forme (28), *f est structurellement stable si et seulement si l'égalité* (31) *est satisfaite.*

Ce lemme permet de ramener les calculs à des calculs sur la situation résiduelle exprimée en termes des coordonnées x' et y'. Dans ce nouveau cadre, il faut montrer que les $(n+1)$-jets σ et δ de f et g sont G-équivalents dans J^{n+1} si f et g (et donc σ et δ) sont stables. Soit donc S^{n+1} l'ensemble G-invariant des jets stables de J^{n+1}. Il suffit de montrer que si $\sigma \in S^{n+1}$ on a :

$$G^{n+1}\sigma = G_C^{n+1}\sigma \cap S^{n+1} \ . \tag{32}$$

Or G est un sous-groupe de G_C dont les orbites stables sont de même dimension. Cela implique que $G^{n+1}\sigma$ est *ouvert* dans G_C^{n+1}. Dans la partition de $G_C^{n+1}\sigma \cap S^{n+1}$ en orbites de G^{n+1}, ces orbites sont donc *à la fois* ouvertes et fermées. Autrement dit, $G^{n+1}\sigma$ est une *composante connexe* de $G_C^{n+1}\sigma \cap S^{n+1}$. Cette remarque est la base du théorème de Mather dont les idées directrices sont les suivantes.

1. On continue à travailler localement. Soient M' la sous-variété de M d'équations $x_1 = \cdots = x_r = 0$ et N' la sous-variété de N d'équations $y_1 = \cdots = y_r = 0$. On a $\Gamma_{a'} = \Gamma_a\left(M'\right)$ et $\mathfrak{m}_{a'}\mathcal{E}$ s'identifie à l'espace des germes en a d'applications $f' : (M', a) \to (N', b)$. Ce faisant, $\mathfrak{m}_{a'}\mathcal{E}^k$ s'identifie à l'espace J'^k des k-jets de telles applications et la formule (31) fournit une décomposition de cet espace. On peut alors définir les groupes G' et G'_C et analyser leur action sur J'^{n+1}.

Soit Λ_r l'ensemble des jets $\sigma \in J^{n+1}$ provenant d'applications de la forme (28) et soit $\lambda : \Lambda_r \to \mathfrak{m}_{a'}^2\mathcal{E}^{n+1}$ l'application définie par $\lambda\left(\sigma\right) = j^{n+1}\left(f'_\bullet\right)$ (jet d'ordre $(n+1)$ de l'application résiduelle $f'_\bullet$, qui est de rang 0). λ est l'application jet dans J'^{n+1} d'applications $f' : (M', a) \to (N', b)$ de rang 0.

2. Soit alors $V = G_C^{n+1}\sigma \cap \Lambda_r$. Pour prouver (32) on peut se restreindre à Λ_r et il suffit donc de prouver (33)

$$V \cap S^{n+1} \subset G^{n+1}\rho \text{ pour tout } \rho \in V \cap S^{n+1} \ . \tag{33}$$

La formule (33) signifie que si ρ est un $(n+1)$-jet stable G_C-équivalent à σ et de forme (28) alors il est G-équivalent à σ. Soit $\rho' = \lambda\left(\rho\right)$, *i.e.* le $(n+1)$-jet résiduel de ρ. Les ρ' parcourent $V' = \lambda\left(V\right)$ qui est l'orbite résiduelle de σ pour l'équivalence de contact relativement à la forme (28). Il est clair que V' est la G'_C -orbite de σ' dans J'^{n+1}.

3. Soit p le nombre de composantes connexes de G'_C. Le nombre de composantes connexes de V' est $\leq p$ puisque V' est une G'_C -orbite. La première

partie technique de la preuve de Mather consiste à montrer le lemme fondamental suivant :

Lemme 2. – Si V'_0 est une composante connexe de V', alors $\lambda^{-1}(V'_0) \cap S^{n+1}$ est inclus dans une seule orbite de G^{n+1}.

4. Pour démontrer le théorème, Mather remarque d'abord que G'^{n+1} rencontre chaque composante connexe de G'^{n+1}_C. Soient alors $\rho, \rho_1 \in V \cap S^{n+1}$. Comme V' est une orbite de G'^{n+1} dans J'^{n+1}, il existe $\theta' \in G'^{n+1}_C$ tel que $\lambda(\rho_1) = \theta'(\lambda(\rho))$. Soit $\theta'_1 \in G'^{n+1}$ appartenant à la même composante de G'^{n+1}_C que θ'. θ'_1 est le $(n+1)$-jet d'un élément $(\varphi', \psi') \in G'$, *i.e.* du couple d'un difféomorphisme φ' de M' et d'un difféomorphisme ψ' de N'. Etendons φ' et ψ' en $\theta_1 = (\varphi, \psi) \in G$ par l'identité sur les variables $\overline{x} = (x_1, \ldots, x_r)$ et $\overline{y} = (y_1, \ldots, y_r)$. Clairement, $\lambda(\theta_1(\rho)) = \theta'_1(\lambda(\rho))$. Donc $\lambda(\theta_1(\rho))$ et $\lambda(\rho_1)$ sont dans la même composante connexe V'_0 de V'. Mais comme, d'après le lemme 2, $\lambda^{-1}(V'_0) \cap S^{n+1}$ est inclus dans une seule orbite de G^{n+1}, $\theta_1(\rho)$ et ρ sont G-équivalents. Donc ρ et ρ_1 sont G-équivalents.

5. Reste à démontrer le lemme 2. On considère la restriction $\lambda : V \cap S^{n+1} \to V'$ et on montre que c'est une fibration localement triviale dont chaque fibre est contenue dans une seule orbite de G^{n+1}. Comme ces orbites sont des composantes connexes de G^{n+1}_C, cela implique le lemme 2. Pour montrer que chaque fibre est contenue dans une seule orbite de G^{n+1}, il faut connaître les composantes connexes des fibres et montrer que, pour chaque fibre, il existe une orbite en intersectant chaque composante connexe. Pour ce faire, on utilise la caractérisation (31) des jets stables : ρ est stable si et seulement si

$$\Omega(\lambda(\rho)) + [\partial\rho] = \mathfrak{m}_{a'}\mathcal{E}^n \, (= J'^n) \tag{34}$$

Or la codimension $c(\rho)$ de $\Omega(\lambda(\rho))$ dans $\mathfrak{m}_{a'}\mathcal{E}^n = J'^n$ est égale à celle de $\Omega(f'_\bullet)$ dans $\mathfrak{m}_{a'}\mathcal{E}$ et celle-ci est égale par construction à $c - (n-r)$ où c est la G_C-codimension de f. Mais l'on peut montrer que la codimension d de ρ dans J^n est égale à $c + m - n$. La codimension $c(\rho)$ est donc égale à $d - m + r$. D'après (34), comme $[\partial\rho]$ est un $\mathbb{R}$-vectoriel engendré par r éléments, $c(\rho) \leq r$ ($d \leq m$ et $c \leq n$).

La fibre de la fibration $\lambda : V \cap S^{n+1} \to V'$ provient du terme $[\partial\rho]$ de (34). Or $[\partial\rho]$ est le $\mathbb{R}$-vectoriel de $\mathfrak{m}_{a'}\mathcal{E}^n$ engendré par les lignes de la matrice $n \times (n-r)$ (u_{ij}) où u_{ij} est l'élément de $\mathfrak{m}_{a'}/\mathfrak{m}^{n+1}_{a'}$ égal au n-jet de $\left(\frac{\partial f_{j+r}}{\partial x_i}\right)'$, $i = 1, \ldots, r$, $j = 1, \ldots, n-r$. Une telle ligne peut être effectivement considérée comme un élément de $\mathfrak{m}_{a'}\mathcal{E}^n$ puisque c'est un $(n-r)$-uple d'éléments de $\mathfrak{m}_{a'}/\mathfrak{m}^{n+1}_{a'}$ et que $\mathcal{E}$ est de dimension $n-r$ sur $\Gamma_{a'}$. La description des fibres permet alors de montrer qu'elles sont connexes si $c(\rho) < r$ ou si $r = 0$ et qu'elles ont deux composantes connexes si $c(\rho) = r > 0$.

Si $c(\rho) < r$ ou si $r = 0$, la fibre est incluse dans une seule orbite de G^{n+1} car elle est connexe et G^{n+1} est une composante connexe de G^{n+1}_C. Si en revanche $c(\rho) = r > 0$, on vérifie que l'élément (φ, ψ) de G^{n+1}, où φ change x_1 en $-x_1$

et ψ change y_1 en $-y_1$ fait passer de la première à la seconde composante de la fibre. Cela termine la démonstration du théorème 63 de Mather : si f est une application structurellement stable, son type différentiable est déterminé par l'algèbre finie Q_f^{n+1}. $\square$

9.6 Classification des germes stables

Pour un germe stable, l'anneau local Q_f, et même un de ses quotients Q_f^{n+1}, caractérise donc le type différentiable. Cela permet de remplacer Q_f par l'anneau $\widehat{Q}_f$ limite projective des Q_f^{n+1}. $\widehat{Q}_f$ est le quotient de l'anneau de séries formelles $\mathbb{R}[[x_1, \ldots, x_m]]$ par l'idéal engendré par les séries de Taylor T_{f_i} en a des composantes $(f_1, \ldots, f_n)$ de f (séries T_{f_i} que nous continuerons à noter f_i par simplicité). La G_C-codimension de f peut se calculer directement sur $\widehat{Q}_f$. Soit en effet I l'idéal de $\left(\widehat{Q}_f\right)^n$ engendré par les m n-uples $\left(\frac{\partial f_1}{\partial x_i}, \ldots, \frac{\partial f_n}{\partial x_i}\right)$, $i = 1, \ldots, m$, *i.e.* par les lignes de la matrice jacobienne. Posons $\mu(f) = \dim_{\mathbb{R}}\left(\left(\widehat{Q}_f\right)^n / I\right)$. On a le résultat :

Proposition 64. – $\mu(f) = c$ (G_C-codimension de f).

Qui plus est, ce nombre ne dépend pas de la présentation de $\widehat{Q}_f$ comme quotient. Si en particulier f est de rang r et de la forme (28), il peut se déduire de $\widehat{Q}_{f'_\bullet} \simeq \widehat{Q}_f$.

John Mather a *caractérisé* les quotients des anneaux de séries formelles qui sont du type $\widehat{Q}_f$. Soit $A = \mathbb{R}[[x_1, \ldots, x_p]]/I$ un tel anneau quotient. Soit $i(A) = p - q$ où $q \leq p$ est le nombre minimal de générateurs de I. Pour chaque $c \leq i(A)$ on peut définir un nombre $\mu_c(A)$ tel que pour $A = \widehat{Q}_f$ on ait $\mu_{m-n}\left(\widehat{Q}_f\right) = \mu(f)$.

Théorème 65. – *Soient m, n des entiers et A un quotient d'anneau de séries formelles. Il existe un germe stable d'application $f : (M, a) \to (N, b)$ tel que $A \simeq \widehat{Q}_f$ si et seulement si :*

1. $i(A) \geq m - n$,
2. $\mu_{m-n}(A) \leq n$.

10 La classification des singularités de codimension ≤ 5

Les résultats techniques des paragraphes précédents permettent de classer ce que René Thom a appelé les "catastrophes élémentaires", à savoir les déploiements universels des singularités de *fonctions* potentiel de petite codimension. On considère le germe instable d'une application $f : M \to \mathbb{R}$ en un point critique a qui est dégénéré (puisque les points critiques non dégénérés sont stables) et on étudie ses modèles transverses. Comme on travaille localement, on peut supposer $M = \mathbb{R}^m$. On continue à noter $\mathcal{E}$ ($= \mathcal{E}_m$) l'anneau local

Γ_a et $\mathfrak{m}$ $(= \mathfrak{m}_m)$ son idéal maximal. On peut supposer $f(0) = 0$, *i.e.* $f \in \mathfrak{m}$. Dire que 0 est un point critique de f c'est dire que les dérivées partielles $\dfrac{\partial f}{\partial x_i}$, $i = 1, \ldots, m$ s'annulent, *i.e.* que $f \in \mathfrak{m}^2$. Rappelons que le corang de f est celui de son hessien $H = \left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(0)\right)$.

10.1 Précisions sur la codimension et la détermination

Proposition 66 (théorème des singularités résiduelles). – Soit $f \in \mathfrak{m}^2$ un germe de codimension finie et de corang s. Alors il existe des coordonnées locales $(x_1, \ldots, x_s, y_1, \ldots, y_{m-s})$ telles que f s'écrive sous la forme $f = H(y) + g(x)$ où H est une forme quadratique en les $y_1, \ldots, y_{m-s}$ et g une fonction en les $x_1, \ldots, x_s$ de hessien nul (i.e. dont 0 est un point critique totalement *dégénéré) définie à équivalence près.*

Comme les points quadratiques non dégénérés sont stables, on voit que seule la singularité résiduelle $g(x)$ intervient dans les déploiements de f. La proposition 66 dit que l'on peut *séparer les variables*, se restreindre à un nombre de variables égal au corang et supposer que $f \in \mathfrak{m}^3$ (point critique totalement dégénéré).

L'équivalence que l'on utilise dans le cas de fonctions potentiel est en général l'équivalence à droite associée aux changements de coordonnées locales à la source. La codimension sera donc dans ce cas :

$$\dim_{\mathbb{R}} \Gamma\left(f^*(T\mathbb{R})\right) / \theta_f\left(\Gamma(TM)\right) \ .$$

Les difféomorphismes intervenant dans cette formule devant laisser $0 \in M$ et $0 \in \mathbb{R}$ fixes, les champs associés doivent s'annuler en ces points.

Soit y la coordonnée de $\mathbb{R}$. Un champ sur $\mathbb{R}$ le long de f équivaut à la donnée d'une fonction (d'un germe) $u \in \mathfrak{m}$. En effet $T\mathbb{R} = \mathbb{R} \times \mathbb{R}$ et à $u \in \mathfrak{m}$ correspond le champ $\zeta : M \to T\mathbb{R}$ donné par $\zeta(x) = (f(x),\, u(x)\dfrac{\partial}{\partial y})$. Autrement dit, $\Gamma\left(f^*(T\mathbb{R})\right)$ est l'espace $\mathfrak{m}\dfrac{\partial}{\partial y}$ identifiable à $\mathfrak{m}$. D'autre part, si $\xi \in TM$, $\xi = \sum_{i=1}^m u_i(x)\dfrac{\partial}{\partial x_i}$, avec $u_i \in \mathfrak{m}$, alors $\theta_f(\xi)$ est le champ le long de f donné par $\theta_f(\xi) = \left(\sum_{i=1}^m u_i(x)\dfrac{\partial f(x)}{\partial x_i}\right)\dfrac{\partial}{\partial y}$. Donc $\theta_f\Gamma(TM)$ (*i.e.* "l'espace tangent" à l'orbite de f pour l'équivalence à droite) s'identifie dans ce cas au produit $\mathfrak{m}\Delta$ où Δ est l'idéal jacobien engendré par les dérivées partielles $\left(\dfrac{\partial f}{\partial x_i}\right)$. La codimension de f est par suite le nombre $\dim_{\mathbb{R}} \mathfrak{m}/\mathfrak{m}\Delta$. Mais comme l'on suppose $f \in \mathfrak{m}^2$, on a $\Delta \subset \mathfrak{m}$, $\mathfrak{m}\Delta \subset \mathfrak{m}^2$ et l'on préfère appeler codimension de f sa codimension dans $\mathfrak{m}^2$. Ce sera dans ce cas le nombre $\dim_{\mathbb{R}} \mathfrak{m}^2/\mathfrak{m}\Delta$ égal au nombre $\dim_{\mathbb{R}} \mathfrak{m}/\Delta$.

Si l'on veut tenir compte également de l'équivalence à gauche (*i.e.* de l'équivalence différentiable), il faut introduire le terme $\Omega_f(\Gamma(T\mathbb{R}))$. Or si $\eta = u\dfrac{\partial}{\partial y}$ est un (germe de) champ sur $\mathbb{R}$ avec $u \in \mathfrak{m}_1$, on a $\Omega_f(\eta) = u \circ f\dfrac{\partial}{\partial y}$. Cela montre que $\Omega_f(\Gamma(T\mathbb{R})) = f^*(\mathfrak{m}_1)$. La codimension de f sera dans ce cas $\dim_{\mathbb{R}} \mathfrak{m}/(\mathfrak{m}\Delta + f^*(\mathfrak{m}_1))$ dans $\mathfrak{m}$ ou $\dim_{\mathbb{R}} \mathfrak{m}^2/(\mathfrak{m}\Delta + f^*\mathfrak{m}_1) = \dim_{\mathbb{R}} \mathfrak{m}/(\Delta + f^*(\mathfrak{m}_1))$ dans $\mathfrak{m}^2$.

Par les techniques standard déjà souvent utilisées (lemme de Nakayama et théorème de Malgrange), on montre alors le théorème :

Théorème 67.

(i) *Si* $\mathfrak{m}^{k+1} \subset \mathfrak{m}^2\Delta + \mathfrak{m}^{k+2}$ *alors* f *est* k*-déterminée pour la d-équivalence.*

(ii) *Si* f *est* k*-déterminée pour la d-équivalence alors* $\mathfrak{m}^{k+1} \subset \mathfrak{m}\Delta$ *(et donc* a fortiori $\mathfrak{m}^{k+1} \subset \mathfrak{m}\Delta + \mathfrak{m}^{k+2}$).

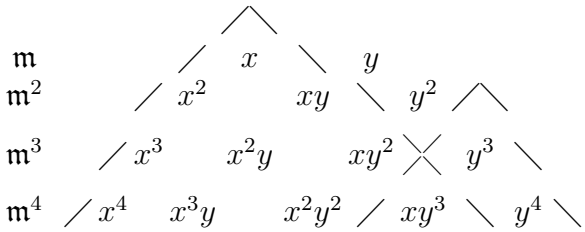
(iii) *Si* $\mathfrak{m}^{k+1} \subset \mathfrak{m}^2\Delta + f^*(\mathfrak{m}_1) + \mathfrak{m}^{k+2}$ *alors* f *est* k*-déterminée pour l'équivalence différentiable.*

(iv) *Si* f *est* k*-déterminée pour l'équivalence différentiable, alors* $\mathfrak{m}^{k+1} \subset \mathfrak{m}\Delta + f^*(\mathfrak{m}_1)$.

Ce théorème rend particulièrement clair le fait que f est de détermination finie si et seulement si f est de codimension finie (théorème 60). Il fournit aussi un critère très simple pour la détermination finie.

Corollaire 68. – f est de détermination finie si et seulement si il existe ℓ tel que $\mathfrak{m}^\ell \subset \Delta$ (cf. Proposition 61).

Ainsi que Dirk Siersma l'a montré,[43] le théorème 67 est d'un usage très simple pour le corang 2 (et évidemment pour le corang 1). Considérons par exemple $f = x^2 + y^4$ (qui d'après le théorème 66 des singularités résiduelles est en fait de corang 1). On a $\dfrac{\partial f}{\partial x} = 2x$ et $\dfrac{\partial f}{\partial y} = 4y^3$. Si l'on écrit la suite des $\mathfrak{m}$, $\mathfrak{m}^2$, $\mathfrak{m}^3$, etc. sous la forme de leurs monômes générateurs et si l'on considère les multiples dans $\mathfrak{m}$ de x et de y^3, on construit facilement Δ en utilisant le diagramme suivant :



[43] Siersma [41].

On voit que $\mathfrak{m}/\Delta = \mathbb{R}y + \mathbb{R}y^2$. La codimension dans $\mathfrak{m}^2$ est donc égale à 2. D'autre part $\mathfrak{m}^3 \subset \Delta$, $\mathfrak{m}^5 \subset \mathfrak{m}^2\Delta$ et f est 4-déterminée d'après 67(i).

Remarquons que f se réduisant à y^4, elle ne peut pas être déterminée à un ordre inférieur à 4 et que le théorème 67 ne peut donc pas être amélioré. Remarquons aussi que si l'on veut traiter la détermination pour l'équivalence différentiable, il faut également considérer l'idéal $f^*(\mathfrak{m}_1) = (f)$. Or dans ce cas $(f) \subset \Delta$ et f est donc aussi de codimension 2 et de détermination 4 pour l'équivalence différentiable. Remarquons enfin que les implications (ii) et (iv) du théorème 67 ne sont pas réversibles. En effet, comme $\mathfrak{m}^3 \subset \Delta$, $\mathfrak{m}^4 \subset \mathfrak{m}\Delta$ alors que f n'est pourtant pas 3-déterminée.

Considérons comme autre exemple la singularité dite *double cusp* $f = x^4 + y^4$. On a $\dfrac{\partial f}{\partial x} = 4x^3$ et $\dfrac{\partial f}{\partial y} = 4y^3$. D'après le diagramme de Siersma suivant, on voit que la codimension de f est 8 et que $\mathfrak{m}^5 \subset \Delta$ ce qui implique que f soit au moins 6-déterminée. Mais en fait $\mathfrak{m}^5 \subset \mathfrak{m}^2\Delta$ et f est déterminée à l'ordre 4.

$$
\begin{array}{lccccccccccccccc}
\mathfrak{m} & & & & & & & x & & y & & & & & & \\
\mathfrak{m}^2 & & & & & \wedge & & x^2 & & xy & & & y^2 & & \wedge & \\
\mathfrak{m}^3 & & & & / & & x^3 & & \backslash & & x^2y & & & xy^2 & / & & y^3 & & \backslash \\
\mathfrak{m}^4 & & & / & & x^4 & & & x^3y & & \backslash & & x^2y^2 & & / & & xy^3 & & & & y^4 & & \backslash \\
\mathfrak{m}^5 & / & & & x^5 & & & x^4y & & & & x^3y^2 & & \backslash\!/ & & x^2y^3 & & & & xy^4 & & & & y^5 & \backslash
\end{array}
$$

Comme $(f) \subset \Delta$, f est aussi de codimension 8 et de détermination 4 pour l'équivalence différentiable.

Considérons maintenant $f = x^2y$. Comme $\dfrac{\partial f}{\partial x} = 2xy$ et $\dfrac{\partial f}{\partial y} = x^2$, toutes les puissances de y sont à l'extérieur de Δ. f est donc de codimension et de détermination infinies.

Considérons enfin $f = \dfrac{x^3}{3} + xy^3$. Comme $\dfrac{\partial f}{\partial x} = x^2 + y^3$ et $\dfrac{\partial f}{\partial y} = 3xy^2$, $\mathfrak{m}^2\Delta$ est l'idéal engendré par: $x^4 + x^2y^3$, x^3y^2, $x^3y + xy^4$, x^2y^3, $x^2y^2 + y^5$, xy^4, *i.e.* par: x^4, x^3y^2, x^3y, x^2y^3, $x^2y^2 + y^5$, xy^4. Les idéaux $\mathfrak{m}$, $\mathfrak{m}^2$ et $\mathfrak{m}^3$ ne peuvent pas être inclus dans $\mathfrak{m}^2\Delta$ pour des raisons de degré. $\mathfrak{m}^4$ ne peut pas y être inclus à cause de x^2y^2, xy^3 et y^4. $\mathfrak{m}^5$ à cause de y^5. Mais $\mathfrak{m}^6$ y est inclus car $y(x^2y^2 + y^5) - x^2y^3 = y^6 \in \mathfrak{m}^2\Delta$. f est donc 5-déterminée. Comme $\mathfrak{m}^5 \not\subseteq \mathfrak{m}^2\Delta$, cela prouve que les implications (i) et (iii) du théorème 67 ne sont pas réversibles. Pour calculer la codimension de f, il est préférable d'utiliser $\dim_{\mathbb{R}} \mathfrak{m}^2/\mathfrak{m}\Delta$ plutôt que $\dim_{\mathbb{R}} \mathfrak{m}/\Delta$. En effet $\mathfrak{m}\Delta$ est l'idéal engendré par $x^3 + xy^3$, x^2y^2, $x^2y + y^4$, xy^3, *i.e.* par x^3, xy^3, x^2y^2, $x^2y + y^4$. Considérons alors le diagramme de Siersma :

$$
\begin{array}{lccccccccccccc}
\mathfrak{m} & & & & & x & & & & y & & & & \\
\mathfrak{m}^1 & & & & \diagup & & \diagdown & & & & & & & \\
 & & & & x^2 & & & & xy & & & & y^2 & \\
\mathfrak{m}^2 & & & \diagup & & \diagdown & & \diagup & & \diagdown & & \diagup & & \diagdown \\
 & & x^3 & & & & x^2y & & & & xy^2 & & & & y^3 \\
\mathfrak{m}^3 & \diagup & & & & \diagdown & & \diagup & & \diagdown & & \diagup & & \diagdown \\
 & x^4 & & & x^3y & & & & x^2y^2 & & & & xy^3 & & & & y^4
\end{array}
$$

On voit qu'il laisse ouverte la question de x^2y, xy^2, et des puissances de y. Or comme $y(x^2y + y^4) - x^2y^2 = y^5$, $y^5 \in \mathfrak{m}\Delta$. $\mathfrak{m}^2/\mathfrak{m}\Delta$ est donc de dimension $7 - 1 = 6$ (car il faut tenir compte de la relation $x^2y + y^4 \equiv 0$).

Ces quelques exemples montrent comment l'on peut calculer explicitement dans le cas de finitude la codimension de f et son degré de détermination. On peut même, comme l'a fait Christopher Zeeman, préciser le rapport entre ces deux nombres.[44]

Proposition 69. – Si $\operatorname{codim}(f)$ *et* $\det(f)$ *sont finies, on a*

$$\det(f) \leq \operatorname{codim}(f) + 2.$$

Preuve. – Considérons la suite décroissante de sous-espaces

$$\mathfrak{m} = \mathfrak{m} + \Delta \supset \mathfrak{m}^2 + \Delta \supset \ldots \supset \mathfrak{m}^k + \Delta \supset \ldots$$

Comme f est de détermination finie, cette suite est stationnaire à partir d'un certain k où l'on a $\mathfrak{m}^{k-1} + \Delta = \mathfrak{m}^k + \Delta$. Cela implique que f soit k-déterminée et donc $\det(f) \leq k$. D'autre part, $\operatorname{codim}(f) = \dim_{\mathbb{R}} \mathfrak{m}/\Delta \leq \dim_{\mathbb{R}} \mathfrak{m}/\mathfrak{m}^{k-1}$ car $\Delta \subset \mathfrak{m}^{k-1}$. La suite d'inclusions

$$\mathfrak{m}/\Delta = (\mathfrak{m} + \Delta)/\Delta \supset \left(\mathfrak{m}^2 + \Delta\right)/\Delta \supset \ldots \supset \left(\mathfrak{m}^{k-1} + \Delta\right)/\Delta = 0$$

est stricte. Comme elle comprend $k - 2$ termes, $\operatorname{codim}(f) \geq k - 2$ et donc $\operatorname{codim}(f) \geq \det(f) - 2$. Cette inégalité ne peut pas être améliorée puisque, dans le cas des fonctions stables (fonctions de Morse) $\operatorname{codim}(f) = 0$ et $\det(f) = 2$. $\square$

D'autre part les espaces de jets s'identifient (puisqu'on se situe dans $\mathfrak{m}^2$) aux quotients $I^k = \mathfrak{m}^2/\mathfrak{m}^{k+1}$. Comme $\dim_{\mathbb{R}} \mathcal{E}/\mathfrak{m}^{k+1} = \dfrac{(m+k)!}{m!k!}$, on a $\dim_{\mathbb{R}} I^k = \dfrac{(m+k)!}{m!k!} - (m+1)$. $\mathfrak{m}^2$ et les I^k sont *stratifiés par la codimension.* Si c est une codimension, on peut noter Σ_c la "strate" des f de codimension c, Ω_c l'union $\bigcup\limits_{c' \leq c} \Sigma_{c'}$ et Γ_c l'union $\bigcup\limits_{c' \geq c} \Sigma_{c'}$. On a alors la proposition :

Proposition 70. – Si $0 \leq c \leq k - 2$, $I^k = \Omega_c^k \cup \Gamma_{c+1}^k$ *(union disjointe) et* Γ_{c+1}^k *est une variété algébrique. Plus précisément,* I^k *est l'union disjointe* $I^k = \Sigma_0^k \cup \Sigma_1^k \cup \ldots \cup \Sigma_{k-2}^k \cup \Gamma_{k-1}^k$, *chaque strate* Σ_c^k *étant la différence* $\Gamma_c^k - \Gamma_{c+1}^k$ *de deux variétés algébriques.*

[44] Pour la suite de ce paragraphe, *cf.* Zeeman [56].

Preuve. – Soit en effet $\sigma \in I_k$ le k-jet de f en 0. Posons

$$\tau(\sigma) = \dim_{\mathbb{R}} \mathfrak{m}/\left(\Delta + \mathfrak{m}^k\right) \ .$$

Par définition $\tau(\sigma) \leq \operatorname{codim}(f)$. Supposons $\tau(\sigma) > c$. Cela implique $c < \operatorname{codim}(f)$ et donc $\sigma \in \Gamma^k_{c+1}$. Supposons au contraire $\tau(\sigma) \leq c$. Comme $c \leq k-2$, cela implique $\tau(\sigma) \leq k-2$. Or comme la suite croissante

$$0 = \mathfrak{m}/(\mathfrak{m}) = \mathfrak{m}/(\Delta + \mathfrak{m}) \subset \mathfrak{m}/\left(\Delta + \mathfrak{m}^2\right) \subset \cdots \subset \mathfrak{m}/\left(\Delta + \mathfrak{m}^k\right)$$

comprend k termes et doit aller de la dimension 0 à la dimension $\tau(\sigma) \leq k-2$, elle doit comprendre deux termes identiques $\Delta + \mathfrak{m}^{i-1} = \Delta + \mathfrak{m}^i$. D'après le lemme de Nakayama, on aura donc $\mathfrak{m}^k \subset \mathfrak{m}^{i-1} \subset \Delta$ et par suite $\Delta + \mathfrak{m}^k = \Delta$, *i.e.* $\tau(\sigma) = \operatorname{codim}(f)$, soit $\sigma \in \Omega^k_c$.

Pour montrer que Γ^k_{c+1} est une variété algébrique, on remarque qu'elle est l'ensemble des jets $\sigma \in I_k$ tels que $\dim_{\mathbb{R}}\left(\Delta + \mathfrak{m}^k\right)/\mathfrak{m}^k$ soit inférieure à un certain nombre K. Or cette propriété revient à dire qu'une matrice en les coefficients des polynômes de I^k est de rang $< K$ et une telle condition est algébrique. □

Sous les mêmes conditions, on peut montrer que les orbites d'un jet $\sigma \in I^k$ pour l'équivalence à droite sont des sous-variétés de I^k de la "bonne" codimension.

Proposition 71. – Soit $c = \operatorname{codim}(f)$, $c \leq k-2$. Alors l'orbite $\widetilde{\sigma} = G^k_M(\sigma)$ du k-jet σ de f est une sous-variété de codimension c de I^k.

Preuve. – On a $\det(f) - 2 \leq \operatorname{codim}(f)$ d'après la proposition 69 et $\operatorname{codim}(f) \leq k-2$ par hypothèse. Donc $\det(f) \leq k$ et f est k-déterminée. D'après le théorème 67(ii), cela implique $\mathfrak{m}^{k+1} \subset \mathfrak{m}\Delta$. Or le plan tangent à l'orbite de f est $\mathfrak{m}\Delta$ et donc est $\mathfrak{m}\Delta/\mathfrak{m}^{k+1}$ dans I^k (ce ne serait pas le cas si f n'était pas k-déterminée). Un calcul élémentaire donne alors $\operatorname{codim}(\widetilde{\sigma}) = \dim_{\mathbb{R}} \mathfrak{m}^2/\mathfrak{m}^{k+1} - \dim_{\mathbb{R}} \mathfrak{m}\Delta/\mathfrak{m}^{k+1} = c$. □

10.2 Le théorème de classification

Les résultats précédents permettent de classer les singularités de petite codimension, en fait de codimension ≤ 5 (nous verrons au paragraphe suivant la raison de cette limitation à 5) et de trouver leurs formes normales.[45] D'après le théorème 66 des singularités résiduelles, on peut supposer $f \in \mathfrak{m}^3$. Or la codimension ≤ 5 implique trivialement que le corang soit ≤ 2. f est donc au plus une fonction de 2 variables (x, y).

Proposition 72. – Si le corang de f est ≥ 2, $\operatorname{codim}(f) \geq 3$ et si le corang de f est ≥ 3, $\operatorname{codim}(f) \geq 6$.

Preuve. – Supposons que $f(x, y)$ soit de corang 2. Comme $f \in \mathfrak{m}^3$, $\Delta \subset \mathfrak{m}^2$. Dans le quotient $\mathfrak{m}/\Delta$, les termes linéaires x et y ne sont donc pas

[45] *Cf.* Siersma [41] et Zeeman [56].

atteignables à partir de Δ. Mais d'autre part $\mathfrak{m}^2$ est engendré par (x^2, xy, y^2) et Δ est engendré par seulement 2 éléments $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$. On a par conséquent $\dim_{\mathbb{R}}(\mathfrak{m}^2/\Delta) \geq 1$ et donc $\operatorname{codim}(f) = \dim_{\mathbb{R}}(\mathfrak{m}/\Delta) \geq 2 + 1 = 3$. De même, si f est de corang 3, $\operatorname{codim}(f) \geq 3 + (6 - 3) = 6$. □

10.2.1 Corang 1

Proposition 73. – Si f est de corang 1, alors f est équivalente à $\pm x^k$ (ou est de détermination et de codimension infinies).

Preuve. – Soit $f \in \mathfrak{m}^k$, $f = ax^k + bx^{k+1} + \ldots$, $a \neq 0$. $\Delta(f) = \mathfrak{m}^{k-1}$, $\mathfrak{m}^{k+1} \subset \mathfrak{m}^2\Delta(f)$ et f est k-déterminée d'après le théorème 67(i) : $f \underset{d}{\sim} ax^k$. Si k est pair, x^k est invariant par le changement $x \to -x$ et x^k et $-x^k$ ne sont pas équivalents. En revanche si k est impair x^k et $-x^k$ sont équivalents.

Les singularités de type x^k sont dites *cuspoïdes ou queues d'aronde.* Leur d-codimension est $\dim_{\mathbb{R}}(x)/(x^{k-1}) = k-2$. Comme $x^k \in \Delta(x^k)$, leur codimension pour l'équivalence différentiable est la même. On les note A_{k-1}. □

10.2.2 Corang 2

Soit $f(x, y) \in \mathfrak{m}^3$. Le 3-jet de f est une cubique homogène en (x, y) : $j^3 f = ax^3 + bx^2y + cxy^2 + dy^3$. Ces cubiques constituent un espace C, isomorphe à $\mathbb{R}^4$, de coordonnées (a, b, c, d) et il faut décomposer C en orbites pour l'action des changements de coordonnées (x, y).

Lemme 1. – C comprend 5 orbites :

(i) *l'orbite de $x^3 + y^3$ ou $x^3 + xy^2$ (de dimension 4 et donc de codimension 0);*

(ii) *l'orbite de $x^3 - xy^2$ (de codimension 0);*

(iii) *l'orbite de x^2y (de codimension 1);*

(iv) *l'orbite de x^3 (de codimension 2);*

(v) *l'orbite $\{0\}$ de 0 (de codimension 4).*

Preuve. – Soit $P(x, y) = j^3 f$. Comme polynôme du 3ème degré, P peut posséder :

(i) soit deux racines complexes (conjuguées) et une racine réelle;

(ii) soit trois racines réelles distinctes;

(iii) soit une racine double et une racine simple;

(iv) soit une racine triple;

(v) soit être nul.

• Le cas (v) est trivial.
• Pour le cas (iv), $P = (\alpha x + \beta y)^3$ et, par changement linéaire de coordonnées, on peut ramener P à la forme x^3.
• Pour le cas (iii), $P = (\alpha x + \beta y)^2(\gamma x + \delta y)$ et, par changement linéaire de coordonnées, on peut ramener P à la forme x^2y.
• Cas (ii). Soit $P = p_1p_2p_3$ avec $p_i = a_ix + b_iy$, $i = 1, 2, 3$. Considérons les déterminants $k_1 = \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$, $k_2 = \begin{vmatrix} a_3 & b_3 \\ a_1 & b_1 \end{vmatrix}$ et $k_3 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$. Ils sont tous différents de 0 car les formes linéaires p_i sont linéairement indépendantes. Posons alors $u' = k_1p_1$, $v' = k_2p_2$. Le déterminant de ce changement linéaire de coordonnées est égal à $\begin{vmatrix} k_1a_1 & k_1b_1 \\ k_2a_2 & k_2b_2 \end{vmatrix} = k_1k_2k_3 \neq 0$. Ce changement est donc admissible. Si l'on pose de plus, $u + v = u'$, $u - v = v'$, on obtient $u^3 - uv^2 \sim -k_1k_2k_3P \sim P$.
• Cas (i). Il suffit de poser $a_2 = \overline{a}_1, b_2 = \overline{b}_1, a_3$ et b_3 restant réels. Cela implique $k_1 = -\overline{k}_2$ et k_3 imaginaire pur. Si l'on pose $iu + v = k_1p_1$, $iu - v = k_2p_2$, il est facile de vérifier que (u, v) est réel et que le changement $(x, y) \to (u, v)$ est admissible. Par ce changement P devient équivalent à $2(u^3 + 3uv^2)$ soit à $u'^3 + v'^3$ si l'on pose $u' = u + v$ et $v' = u - v$. □

Profitons de ce lemme de classification des formes cubiques pour rappeler la différence (et évidemment aussi le lien) qu'il y a entre une classification réduite comme ici à une énumération de formes normales et une classification réalisée géométriquement comme *stratification*. Dans une certaine mesure, le cas des cubiques est exemplaire de la stratégie de la théorie des singularités où il s'agit à la fois de trouver des formes normales (algébriques) pour les singularités et de maîtriser la géométrie des stratifications qui expriment leurs relations d'incidence. La stratification S de l'espace $\mathbb{R}^4 = (a, b, c, d)$ des formes cubiques consiste à étudier la géométrie du *discriminant* D de l'équation générale du 3e degré $aX^3 + bX^2 + cX + d = 0$, c'est-à-dire de l'hypersurface[46]

$$D = 4(ac^3 + b^3d) + 27a^2d^2 - b^2c^2 - 18abcd = 0 \ .$$

Christopher Zeeman l'a analysée de la façon suivante.[47]

D'abord D étant homogène, S est une stratification conique de sommet (0) (strate la plus singulière) et il suffit donc d'en connaître la trace sur la sphère unité S^3 de $\mathbb{R}^4$. On peut identifier S^3 à $\mathbb{R}^3$ en enlevant un "point à l'infini". On montre que sur $\mathbb{R}^3$, S est donnée par le "bracelet" Σ engendré par une hypocycloïde à trois rebroussements tournant d'un tiers de tour en accomplissant une révolution autour de son âme (*cf.* figure 6).

[46] Pour l'équation $X^3 + pX + q = 0$, D se réduit à la forme bien connue $4p^3 + 27q^2$.
[47] *Cf.* Zeeman [57].

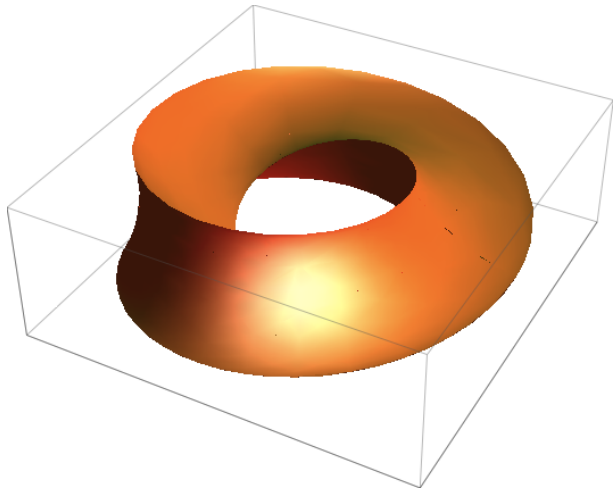


Figure 6: La stratification des formes cubiques dans S^3 d'après Zeeman.

L'orbite de $x^3 + y^3 \sim x^3 + xy^2$ est l'extérieur de Σ, celle de $x^3 - xy^2$ l'intérieur, celle de x^2y la surface Σ moins l'arête de rebroussement et celle de x^3 est l'arête de rebroussement.

Une fois classés à équivalence près les 3-jets de f, il reste à vérifier dans quels cas f est 3-déterminée et sinon quels termes de degré supérieur il faut ajouter pour obtenir la détermination. Les singularités ainsi obtenues sont dites *ombilics* car elles sont présentées par le fibré normal des surfaces de $\mathbb{R}^3$ aux points traditionnellement appelés points ombilicaux.

L'ombilic hyperbolique. Soit $f = x^3 + y^3$. On a $\dfrac{\partial f}{\partial x} = 3x^2$, $\dfrac{\partial f}{\partial y} = 3y^2$ et $f \in \Delta = (x^2, y^2)$. Sa codimension est 3 (car $\mathfrak{m}/\Delta = \mathbb{R}x + \mathbb{R}y + \mathbb{R}xy$) et comme $\mathfrak{m}^4 \subset \mathfrak{m}^2\Delta$, f est déterminée à l'ordre 3.

L'ombilic elliptique. Soit $f = x^3 - xy^2$. On a $\dfrac{\partial f}{\partial x} = 3x^2 - y^2$, $\dfrac{\partial f}{\partial y} = -2xy$ et sa codimension est 3. D'autre part, comme $\mathfrak{m}^2\Delta$ est engendré par $3x^4 - x^2y^2$, x^3y, $3x^3y - xy^3$, x^2y^2, $3x^2y^2 - y^4$, xy^3 et que $(3x^4 - x^2y^2) + x^2y^2 = 3x^4$ et $(3x^2y^2 - y^4) - 3x^2y^2 = -y^4$, $\mathfrak{m}^4 \subset \mathfrak{m}^2\Delta$ et f est également 3-déterminée.

L'ombilic parabolique. Soit $f = x^2y$. Comme $\dfrac{\partial f}{\partial x} = 2xy$ et $\dfrac{\partial f}{\partial y} = x^2$, les puissances de y n'appartiennent pas à Δ et f est de codimension et de détermination infinies. Il faut donc considérer les jets d'ordre supérieur. Supposons que le premier jet de f différent de x^2y soit le jet d'ordre k. $j^kf = x^2y + g$ où g est un polynôme homogène de degré $k \geq 4$, $g = a_0x^k + a_1x^{k-1}y + \ldots + a_ky^k$.

Lemme 2. – Si $j^kf = x^2y + g$ où g est un polynôme comme ci-dessus, alors j^kf est équivalent à $x^2y + a_ky^k$.

Preuve. – Considérons le changement de variables $x = x' + p$, $y = y' + q$ où $p = -\dfrac{1}{2}(a_1 x'^{k-2} + \ldots + a_{k-1} y'^{k-2}) \in \mathfrak{m}^{k-2}$ et $q = -a_0 x'^{k-2} \in \mathfrak{m}^{k-2}$. Dans un monôme $x^{k-i}y^i = (x'+p)^{k-i}(y'+q)^i$ interviennent des termes de la forme $x'^{k-i-r}p^r y'^{i-s}q^s$ de degré $k-i-r+r(k-2)+i-s+s(k-2) = k+(r+s)(k-3)$. Comme $k \geq 4$, $k-3 \geq 1$ et ce degré ne peut donc être égal à k que si $r = s = 0$. Autrement dit, on a à l'ordre k :

$$j^k f = (x'+p)^2(y'+q) + a_0 x'^k + a_1 x'^{k-1} y' + \cdots + a_k y'^k$$

où l'on ne compte dans $(x'+p)^2(y'+q)$ que les termes de degré $\leq k$. Or

$$(x'+p)^2(y'+q) = x'^2 y' + 2x'y'p + x'^2 q + y'p^2 + 2x'pq + p^2 q$$

le premier terme étant de degré 3, le 2ème et le 3ème de degré $2+k-2 = k$, le 4ème et le 5ème de degré $1 + 2(k-2) > k$ (car $k > 3$) et le 6ème de degré $3(k-2) > k$ (car $k > 3$). A l'ordre k on a donc $(x'+p)^2(y'+q) = x'^2 y' + 2x'y'p + x'^2 q$ et

$$j^k f = x'^2 y' + 2x'y'p + x'^2 q + a_0 x'^k + a_1 x'^{k-1} y' + \cdots + a_k y'^k.$$

Comme $x'^2 q = -a_0 x'^k$ et $2x'y'p = -[a_1 x'^{k-1} y' + \cdots + a_{k-1} x' y'^{k-1}]$, on a $j^k f = x'^2 y' + a_k y'^k$. □

Lemme 3. – *$f = x^2 y \pm y^k$ est k-déterminée.*

Preuve. – On a $\dfrac{\partial f}{\partial x} = 2xy$ et $\dfrac{\partial f}{\partial y} = x^2 \pm k y^{k-1}$. Appliquons le critère $\mathfrak{m}^{k+1} \subset \mathfrak{m}^2\Delta$. $\mathfrak{m}^2\Delta$ est engendré par

$$x^3 y, x^2 y^2, x y^3, x^4 \pm k x^2 y^{k-1}, x^3 y \pm k x y^k, x^2 y^2 \pm k y^{k+1}.$$

D'autre part $\mathfrak{m}^{k+1}$ est engendré par x^{k+1}, $x^k y$, $x^{k-1}y^2$, $\ldots$, xy^k et y^{k+1}. Comme $k \geq 4$, tous ces termes, à l'exception de x^{k+1} et y^{k+1}, sont déjà dans l'idéal (x^3y, x^2y^2, xy^3). Pour x^{k+1} on a $x^{k+1} = x^{k+1-4}(x^4 \pm k x^2 y^{k-1}) \mp k x^{k-1} y^{k-1}$ et, comme $k \geq 4$, $x^{k+1} \in \mathfrak{m}^2\Delta$. Il en va de même pour y^{k+1}. □

Ces deux lemmes montrent que si $j^3 f = x^2 y$ et si f est de détermination finie, alors f est équivalente à $x^2 y \pm y^k$. De telles singularités s'appellent des *ombilics paraboliques* surtout pour $k = 4$. On voit (quitte à échanger x et y dans la forme normale des ombilics hyperboliques et elliptiques) que les ombilics sont tous de la forme $x^2 y \pm y^k$ pour $k \geq 3$. On les note D_{k+1}.

Lemme 4. – *Les singularités D_{k+1} sont de codimension k.*

Preuve. – Tous les monômes de type $x^r y^s$ peuvent s'obtenir facilement comme combinaisons linéaires de $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ à l'exception de x, x^2, y, $\ldots$, y^{k-1}($k+1$ termes). Comme il faut tenir compte de la relation $\dfrac{\partial f}{\partial y} \equiv 0$, $\mathfrak{m}/\Delta$ est de dimension $2 + (k-1) - 1 = k$. □

Les singularités exceptionnelles $\mathbf{E}_6, \mathbf{E}_7, \mathbf{E}_8$. Supposons maintenant que $j^3 f = x^3$. Comme fonction de 2 variables, f est trivialement de détermination et de codimension infinies. Il faut donc considérer les jets d'ordre supérieur. Par des calculs analogues, mais plus complexes, à ceux développés ci-dessus on peut montrer le résultat suivant.[48]

Proposition 74. – (Arnold, Siersma). – Si $j^3 f = x^3$ et si $\operatorname{codim}(f) \leq 7$, *alors f est équivalente à l'une des singularités suivantes :*

(i) $x^3 \pm y^4$, *singularité E_6 de codimension 5 et déterminée à l'ordre 4.*

(ii) $x^3 + xy^3$, *singularité E_7 de codimension 6 et déterminée à l'ordre 4.*

(iii) $x^3 + y^5$, *singularité E_8 de codimension 7 et déterminée à l'ordre 5.*

Ce résultat achève le théorème de classification des singularités de fonctions de codimension ≤ 5.

10.3 Le problème des modules

René Thom s'était proposé de classer les singularités de codimension ≤ 4 car ce sont elles qui, sous le nom de catastrophes élémentaires, peuvent intervenir stablement dans les modèles dont l'espace de contôle est l'espace-temps. Depuis, le théorème de classification des 7 catastrophes élémentaires (les 4 queues d'arondes x^3, x^4, x^5, x^6 et les trois ombilics) a été notablement amélioré. D. Siersma a par exemple classé les singularités jusqu'à la codimension 9.

La limitation à la codimension 4 est quelque peu arbitraire. Mais il existe une sous-classe de singularités, composée des queues d'arondes A_{k-1}, des ombilics D_{k+1}, et des singularités exceptionnelles E_6, E_7, E_8,[49] qui peut être définie de façon *intrinsèque*. En effet, un résultat fondamental d'Arnold[50] dit que ces singularités sont les seules singularités *simples* au sens où, à leur voisinage, il n'existe qu'un nombre *fini* d'orbites. Pour toutes les autres singularités, il existe ce que l'on appelle des *modules*, c'est-à-dire des homotopies f_t de f (des déformations unidimensionnelles) où le type différentiable de f_t varie *continûment* avec t, toutes les f_t étant dans des orbites différentes.

Ce phénomène subtil est lié à la chaîne d'implications

$$\mathfrak{m}^{k+1} \subset \mathfrak{m}^2\Delta + \mathfrak{m}^{k+2} \Longrightarrow f \text{ est } k\text{-déterminée} \Longrightarrow \mathfrak{m}^{k+1} \subset \mathfrak{m}\Delta + \mathfrak{m}^{k+2}$$

dans le théorème 67. Si $\mathfrak{m}^{k+1} \subset \mathfrak{m}\Delta + \mathfrak{m}^{k+2}$, alors f est $(k+1)$-déterminée. Pour savoir si f est en fait k-déterminée, on considère des déformations $f_p = j^k f + p_{k+1}$ où p est un polynôme homogène de degré $k+1$ et l'on cherche

[48] *Cf.* Siersma [41].

[49] Les symboles de ces singularités proviennent du rapport que celles-ci entretiennent avec les groupes de Lie. *Cf.* Slodowy [42] et [43].

[50] *Cf.* Arnold [2].

à montrer que toutes ces déformations sont triviales. D'après ce que nous avons vu en 7.6, l'étape essentielle concerne les homotopies f_t. Parmi celles-ci, on considère essentiellement celles de la forme $f_t = f + th$. En utilisant les techniques de Mather et le critère de trivialité de Thom-Levine (théorème 35), on montre d'abord la proposition :

Proposition 75. – Soit $f_t = f + th$, $t \in I = [0, 1]$. Si $h \in \mathfrak{m}\Delta(f_t)$ pour tout t, alors l'homotopie f_t est triviale. Pour cela, il suffit qu'il existe k tel que, pour tout t,

1. $\mathfrak{m}^{k+1} \subset \mathfrak{m}\Delta(f_t) + \mathfrak{m}^{k+2}$,
2. $h \in \mathfrak{m}\Delta(f_t) + \mathfrak{m}^{k+1}$.

Soit alors f_t une homotopie. Avec Siersma appelons t un "invariant local" – un module – si, pour tout $t_0 \in I$, il existe un voisinage U de t_0 tel que toutes les f_t, $t \in U$, appartiennent à des orbites différentes, *i.e.* tel que le type différentiable de f_t varie continûment avec t.

Théorème 76.[51] *– Si f_t est k-déterminée pour tout $t \in I$ et si pour tout t on a $h \notin \mathfrak{m}\Delta(f_t) + \mathfrak{m}^{k+1}$, alors t est un invariant local (un module) pour l'homotopie f_t.*

Preuve. – Comme les f_t sont k-déterminées on peut se situer dans l'espace des jets J^k. Soit C la droite de J^k donnée par $j^k(f) + \mathbb{R}h$. Soient $\widetilde{f}$ l'orbite de $j^k(f)$ dans J^k et Z_f la fermeture de $\widetilde{f}$ au sens de la topologie de Zariski, topologie intrinsèquement adaptée à la nature algébrique de J^k. Par définition, Z_f est l'intersection de toutes les sous-variétés algébriques de J^k contenant l'orbite $\widetilde{f}$ (si P est un polynôme sur $\widetilde{f}$ il s'annule sur Z_f). Z_f est un sous-ensemble algébrique et l'on peut montrer que $\widetilde{f}$ est ouverte dans Z_f. Or C étant une droite, soit C intersecte Z_f en un nombre fini de points, soit elle est incluse dans Z_f. Dans le second cas, $C \cap \widetilde{f}$ est un ensemble d'intervalles ouverts de C et donc f_t est localement triviale. En revanche dans le premier cas, C intersecte $\widetilde{f}$ en un nombre fini de points et donc, au voisinage de $t = 0$, $f_0 = f$ est la seule fonction f_t à avoir le type de f. Cela est évidemment également valable si l'on remplace f par f_{t_0}. Or pour que C intersecte $\widetilde{f}$ en un ensemble d'intervalles ouverts, il faut que dans un voisinage de t_0, la direction de C, *i.e.* $j^k(h)$, soit contenue dans l'espace tangent en $j^k(f_{t_0})$ à $\widetilde{f_{t_0}}$. Cette condition équivaut à $h \in \mathfrak{m}\Delta(f_{t_0}) + \mathfrak{m}^{k+1}$. Si donc, pour tout $t \in I$, $h \notin \mathfrak{m}\Delta(f_t) + \mathfrak{m}^{k+1}$, alors, pour tout $t \in I$, f_t est la seule fonction à avoir son type différentiable dans un voisinage de t et t est donc bien un invariant local. □

On peut préciser ce résultat de la façon suivante :

Théorème 77. – Supposons que l'homotopie $f_t = f + th$ soit telle que

[51] *Cf.* Siersma [41].

1. $\mathfrak{m}^{k+1} \subset \mathfrak{m}\Delta(f) + \mathfrak{m}^{k+2}$,

2. $h \notin \mathfrak{m}\Delta(f) + \mathfrak{m}^{k+1}$,

3. $\operatorname{codim}(f_t)$ *est constante,*

alors t est un invariant local (un module) au voisinage de $t = 0$.

Comme nous le verrons à la section suivante, lorsque l'on considère les déploiements universels W des singularités simples, les ensembles catastrophiques K_W de ces déploiements sont naturellement stratifiés par une stratification *conique* sur f, la strate la plus singulière se réduisant à $\{f\}$. Lorsqu'il existe des modules il n'en va plus de même. La strate la plus singulière de K_W ne se réduit plus à $\{f\}$. Elle est une déformation à μ paramètres (μ étant le nombre de modules) de f, les f_μ étant de types différentiables différents tout en étant exactement de même "complexité". Pour éclairer ce dernier point donnons un exemple de module. Considérons l'homotopie $f_t = x^4 + y^4 + tx^2y^2$ du double cusp $x^4 + y^4$ où $h = x^2y^2$. Les f_t sont déterminées à l'ordre 4 et de codimension constante 8. Il est en plus facile de vérifier que $h \notin \mathfrak{m}\Delta(f) + \mathfrak{m}^5$. En effet h est de degré 4 et dans $\mathfrak{m}\Delta(f) + \mathfrak{m}^5$ les seuls termes de degré 4 sont des combinaisons linéaires de x^4, y^4, x^3y et xy^3. D'après le théorème 77, cela implique que t est un module (au voisinage de $t = 0$).

Dans cet exemple, la raison de ce phénomène étrange est la suivante. Comme les f_t sont 4-déterminées, on peut se situer dans $J^4(2,1)$ et y considérer les orbites des f_t. Mais comme les f_t sont *homogènes* seule intervient à l'ordre 4 la partie linéaire des difféomorphismes de G. f_t étant homogène de degré 4, elle est le produit de 4 droites complexes sur lesquelles les changements linéaires de coordonnées dans $\mathbb{R}^2$ opèrent comme des transformations *projectives*. Ces changements laissent donc invariant l'invariant projectif qu'est le *birapport* de ces 4 droites. Or le paramètre t est fonction de ce birapport.

Notons toutefois que les phénomènes de modules disparaîssent pour les fonctions (*i.e.* $n = 1$) lorsque l'on considère, au lieu de l'équivalence différentiable (ou de l'équivalence à droite), l'équivalence plus grossière qu'est l'équivalence topologique.

10.4 La classification topologique et le lemme d'isotopie de Thom

Si l'on remplace les difféomorphismes de M et $\mathbb{R}$ par des homéomorphismes, on obtient l'équivalence topologique. René Thom avait conjecturé qu'il n'existait dans $J^k(m,1)$ qu'un nombre *fini* de types topologiques et donc que, dans les homotopies f_t manifestant les modules, le type différentiable variait à type topologique constant. Cette conjecture a été démontrée par Takuo Fukuda. Toutefois, elle est fausse pour $J^k(m,n)$ si $n \geq 3$ car il existe alors des types

topologiques variant eux-mêmes de façon continue.[52]

Nous allons résumer la preuve de Fukuda car, outre son intérêt propre, elle nous permettra de préciser le concept si fondamental de stratification et d'évoquer le théorème fondamental de Thom appelé "lemme d'isotopie".

Précisons d'abord le concept de stratification au sens de Whitney. Soit M une variété différentiable (compacte). On peut toujours la considérer comme plongée dans $\mathbb{R}^N$ pour N assez grand (théorème de plongement de Whitney, Corollaire 16) et donc supposer $M = \mathbb{R}^m$. Une stratification de $\mathbb{R}^m$ est alors une partition de $\mathbb{R}^m$ en sous-variétés connexes X_i, appelées strates, satisfaisant trois types de propriétés.

1. Une propriété de finitude : si $x \in \mathbb{R}^m$, il existe un voisinage de x ne rencontrant qu'un nombre fini de strates.

2. Une propriété de frontière : si $\overline{X_i} \cap X_j \neq \varnothing$ alors $X_j \subset \overline{X_i} - X_i = \partial X_i$. Autrement dit, si la strate X_j contient des points adhérents à X_i alors elle est totalement incluse dans la frontière $\partial X_i = \overline{X_i} - X_i$ de X_i. On dit que X_i et X_j sont *incidentes* et on note $X_j \prec X_i$. C'est la propriété d'incidence qui est la plus caractéristique de la notion de stratification. Intuitivement, une strate de codimension k correspond à la satisfaction de k conditions indépendantes de codimension 1 et une strate incidente revient à satisfaire une condition supplémentaire. Les stratifications que nous avons rencontrées (et qui géométrisent la notion de classification d'objets hierarchisés par un degré d) sont de ce type : stratification de Thom-Boardman et stratification des espaces de jets $J^k(m, 1)$ par la codimension. Dans ce cas, si E est l'espace total des entités f considérées et si E_d est l'ensemble des éléments de degré $\geq d$, les strates de degré 0 sont les composantes connexes de $E - E_1$, les strates de degré 1 les composantes connexes de $E_1 - E_2$, etc.

3. Deux propriétés de régularité d'incidence appelées conditions A et B de Whitney. Soit (X_i, X_j) une paire de strates incidentes, $X_j \subset \overline{X_i} - X_i$, et soit y un point de X_j.

Condition A. Si (x_n) est une suite de points de X_i convergeant vers y et telle que la suite des espaces tangents $T_{x_n}(X_i)$ admettent une limite T, alors $T_y(X_j) \subset T$.

Condition B. Si (x_n, y_n) est une suite de points de $X_i \times X_j$ convergeant vers (y, y) et telle que la droite $x_n y_n$ tende vers une droite limite ℓ et que l'espace tangent $T_{x_n}(X_i)$ tende vers une limite T, alors $\ell \subset T$.

Une telle stratification s'appelle aussi complexe de Whitney ou W-complexe. Si M et N sont deux ensembles stratifiés, une application $f : M \to N$ sera dite stratifiée si

1. l'image d'une strate de M est contenue dans une strate de N;

2. la restriction de f à chaque strate est une submersion.

[52] *Cf.* Fukuda [11]. *Cf.* aussi Thom [51] §3.2.

L'un des intérêts principaux de la notion de stratification est qu'il existe des stratifications naturelles des applications polynomiales dont les strates sont des ensembles *semi-algébriques*. Elles correspondent à l'itération de l'opération "prendre le lieu singulier" (idée que nous avons déjà rencontrée à propos de la stratification de Thom-Boardman). Rappelons qu'un ensemble algébrique X de $\mathbb{R}^m$ est l'ensemble $X = \{x \in \mathbb{R}^m \mid p_1(x) = \cdots = p_r(x) = 0\}$ des zéros communs à un ensemble fini de polynômes $p_1, \ldots, p_r$. La classe $\mathfrak{A}$ des ensembles algébriques n'est pas stable par les opérations de complémentarité $X \to \mathbb{R}^m - X$ et d'image $X \to f(X)$ où f est un polynôme $f : \mathbb{R}^m \to \mathbb{R}^n$. Si l'on considère en revanche les ensembles semi-algébriques définis par des égalités $p_i(x) = 0$ et des inégalités $p_j(x) > 0$ alors leur classe $\mathfrak{S}$ devient stable par réunion, intersection, complémentation et image (théorème de Tarski-Seidenberg). On peut montrer (théorème de Chevalley-Lojasiewicz) que si $X \in \mathfrak{S}$, alors $\overline{X} \in \mathfrak{S}$ et que si X_{reg} est l'ensemble des points réguliers de X, *i.e.* des points au voisinage desquels X est une variété, alors $X_{\text{reg}} \in \mathfrak{S}$.

Il est par conséquent naturel d'appeler $\mathfrak{S}$-stratification une stratification dont toutes strates sont des ensembles semi-algébriques. Si X et Y sont deux strates incidentes avec $Y \subset \overline{X} - X = \partial X$, notons $K_A(X, Y)$ (resp. $K_B(X, Y)$) l'ensemble des points de Y où la condition A de Whitney (resp. la condition B) n'est pas satisfaite (*i.e.* les points où l'incidence est singulière). Whitney a démontré que les lieux singuliers K_A et K_B sont semi-algébriques et que $\dim(K_A) \leq \dim(K_B) < \dim(Y)$.

Considérons maintenant une application $f : M \to N$ dont la source M est stratifiée. Soit (X, Y) une paire de strates incidentes de M, $Y \subset \overline{X} - X$. Supposons que les restrictions $f|_X$ et $f|_Y$ de f aux strates soient de rang constant. On dira que (X, Y) possède la propriété α_f d'être *sans éclatement* si, étant donnée une suite (x_n) de points de X convergeant vers $y \in Y$ et telle que les noyaux $\operatorname{Ker}(D_{x_i}(f|_X))$ convergent vers $T \subset T_y M$, on a $\operatorname{Ker}(D_y(f|_Y)) \subset T$. D'autre part, si $f : \mathbb{R}^m \to \mathbb{R}^n$ est une application polynomiale, on peut, dans la ligne de ce que nous avons vu à propos de la stratification de Thom-Boardman, considérer la décomposition de la source par le rang de l'application linéaire tangente. Plus précisément, si $M \subset \mathbb{R}^m$ est une sous-variété $\mathfrak{S}$-stratifiée et si X est une strate de M, notons $S(f|_X)$ l'ensemble des $x \in X$ où le rang de $D_x(f|_X)$ n'est pas égal à sa valeur maximum sur X. Comme la non maximalité du rang s'exprime par une condition alébrique d'annulation de déterminants, $S(f|_X)$ est semi-algébrique. On peut donc *raffiner* la $\mathfrak{S}$-stratification de M de façon à ce que f soit de rang constant sur les strates. On peut alors montrer que l'ensemble des points où les (X, Y) ne satisfont pas la propriété α_f sont également semi-algébriques.

Cet ensemble de résultats permet de comprendre que, par raffinements successifs, on puisse stratifier une application polynomiale sur un ensemble semi-algébrique M de $\mathbb{R}^m$. On stratifie d'abord M en prenant les composantes connexes de M_{reg}. Puis on stratifie par récurrence son lieu singulier M_{sing} en

prenant d'abord les composantes de $(M_{\text{sing}})_{\text{reg}}$, puis celles de $\left((M_{\text{sing}})_{\text{sing}}\right)_{\text{reg}}$, etc. On raffine cette stratification de façon à ce que les conditions A et B de Whitney soit satisfaites. Ensuite, on raffine encore la stratification obtenue de façon à ce que f soit de rang constant sur les strates. Enfin, on raffine une dernière fois pour que la propriété α_f soit partout satisfaite. Si de plus on considère que le but N de f est également $\mathfrak{S}$-stratifié, on peut raffiner encore la stratification de façon à ce que f devienne une application stratifiée.

Dans ce cadre général, pour démontrer le théorème de Fukuda sur la finitude du nombre de types topologiques, on raisonne de la façon suivante. Soit $J^k = J^k(m,1)$ l'espace des polynômes sur $\mathbb{R}^m$ de degré $\leq k$. Un polynôme $p(x) \in J^k$ s'écrit sous la forme $\sum_{r=1}^{r=N} a_r x^{(\alpha)_r}$ où $\alpha = (\alpha_1, \ldots, \alpha_m)$ est un multi-indice de degré $\alpha_1 + \cdots + \alpha_m \leq k$, $x^{(\alpha)} = x_1^{\alpha_1} \cdots x_m^{\alpha_m}$ et N est le nombre de coefficients d'un $p(x)$ général. Autrement dit, on identifie J^k à $\mathbb{R}^N$. Soit F l'application $F : \mathbb{R}^N \times \mathbb{R}^m \to \mathbb{R}^N \times \mathbb{R}$ qui à $(p,x) \in \mathbb{R}^N \times \mathbb{R}^m$ associe l'évaluation $(p, p(x)) \in \mathbb{R}^N \times \mathbb{R}$. Considérons la suite

$$\mathbb{R}^N \times \mathbb{R}^m \xrightarrow{F} \mathbb{R}^N \times \mathbb{R} \xrightarrow{\pi} \mathbb{R}^N$$

où π est la projection canonique. Il s'agit de la stratifier de façon à pouvoir appliquer l'outil fondamental qu'est le lemme d'isotopie de Thom.

Pour énoncer le lemme d'isotopie, donnons d'abord la définition suivante. Soit $f : M \to N$ une application stratifiée. On dit que f est *sans éclatement* si tout couple (X,Y) de strates incidentes satisfait la propriété α_f. Soient $f : E \to F$ et $g : F \to V$ deux applications stratifiées où V est une variété *connexe* trivialement stratifiée. On dit que f est un *morphisme de Thom* au-dessus de g si, pour tout point $a \in V$, $f_a : E_a \to F_a$ est une application stratifiée sans éclatement où f_a, E_a, F_a sont les fibres au-dessus de a. Autrement dit, un morphisme de Thom est une déformation sur V de morphismes sans éclatement.

Théorème 78 (second lemme d'isotopie de Thom). – Soit $f : E \to F$ un morphisme de Thom propre au-dessus de $g : F \to V$ propre. Alors les applications fibres $f_a : E_a \to F_a$ ont toutes le même type topologique.

On peut lever la condition de propreté en se localisant dans les fibres.

Théorème 79 (lemme d'isotopie locale). – Soit $f : E \to F$ un morphisme de Thom au-dessus de $g : F \to V$. Si p et q sont deux points d'une même strate de E, si $a = g \circ f(p)$ et $b = g \circ f(q)$, alors le germe en p de l'application fibre $f_a : E_a \to F_a$ et le germe en q de $f_b : E_b \to F_b$ ont le même type topologique.

Si on applique ce lemme d'isotopie locale à la suite $\mathbb{R}^N \times \mathbb{R}^m \xrightarrow{F} \mathbb{R}^N \times \mathbb{R} \xrightarrow{\pi} \mathbb{R}^N$, on voit que la fibre au-dessus de $p \in \mathbb{R}^N = J^k$ est tout simplement le polynôme $p : \mathbb{R}^m \to \mathbb{R}$. Le point essentiel consiste donc à montrer qu'il existe une $\mathfrak{S}$-stratification $S(\mathbb{R}^N)$ de $\mathbb{R}^N$ telle que, pour chaque strate X de $S(\mathbb{R}^N)$ il existe des $\mathfrak{S}$-stratifications de $X \times \mathbb{R}^m$ et $X \times \mathbb{R}$ pour lesquelles la restriction

de F à $X \times \mathbb{R}^m$ soit un morphisme de Thom au-dessus de la restriction de π à $X \times \mathbb{R}$. En effet, la finitude découlera alors, d'après le lemme d'isotopie, de celle des $\mathfrak{S}$-stratifications. La partie technique du travail de Fukuda [11] consiste en la construction de ces $\mathfrak{S}$-stratifications.

11 Les déploiements universels

Après avoir classé les singularités simples, il nous reste à envisager leur déploiements universels. Intuitivement, si f est k-déterminée, il s'agit de se situer dans $J^k(m,1)$, de considérer d'abord l'orbite $\tilde{f}$ de f (orbite dont l'espace tangent en f est l'image dans J^k de $\mathfrak{m}\Delta(f)$) puis, si $c = \mathrm{codim}(f)$, une section W de dimension c transverse à $\tilde{f}$ en f. W étant isomorphe à un voisinage de l'origine de $\mathbb{R}^c$, cette section est identifiable à une déformation – à un déploiement – f_w de $f_0 = f$, $w \in W$. Trois questions se posent alors :

(i) Montrer que tous ces déploiements sont équivalents.

(ii) Montrer autant que faire se peut qu'ils sont stables en tant qu'applications $F : \mathbb{R}^m \times \mathbb{R}^c \to \mathbb{R} \times \mathbb{R}^c$.

(iii) Montrer qu'ils sont universels au sens où tout déploiement f_t de f de base T peut se déduire de f_w par image réciproque d'une application, si possible unique, $\theta : T \to W$.

Soit (f_t) un déploiement de f, c'est-à-dire une application $F' : \mathbb{R}^m \times \mathbb{R}^\ell \to \mathbb{R}$, $F'(x,t) = f_t(x)$. Si $(g_t) = G'$ est un déploiement de même base, on dit que F' et G' sont équivalents si il existe un difféomorphisme H de $\mathbb{R}^m \times \mathbb{R}^\ell$ au-dessus d'un difféomorphisme φ de $(\mathbb{R}^\ell, 0)$ (*i.e.* de la forme $H(x,t) = (H_1(x,t), \varphi(t))$ et une application $\psi : \mathbb{R}^\ell \to \mathbb{R}$ tels que :

(i) $H(x,0) = Id_{\mathbb{R}^m}$ et

(ii) $G'(x,t) = F'(H(x,t)) + \psi(t)$.

On considère donc deux déploiements comme équivalents si l'on peut les ramener l'un à l'autre :

(i) par un difféomorphisme φ de la base;

(ii) par une famille de difféomorphismes H_t de la source $\mathbb{R}^m$ variant différentiablement avec t et envoyant la fibre $\mathbb{R}^m \times (t)$ sur la fibre $\mathbb{R}^m \times (\varphi(t))$;

(iii) par une famille de translations $\psi(t)$ du but $\mathbb{R}$ dépendant de t.

Si l'on associe à F' et G' les applications $F : \mathbb{R}^m \times \mathbb{R}^\ell \to \mathbb{R} \times \mathbb{R}^\ell$ et $G : \mathbb{R}^m \times \mathbb{R}^\ell \to \mathbb{R} \times \mathbb{R}^\ell$ données par $F(x,t) = (f_t(x), t)$ et $G(x,t) = (g_t(x), t)$, il est facile de vérifier que F' et G' sont équivalents comme déploiements si et seulement si F et G sont équivalents pour une restriction de l'équivalence de contact. Dire en effet que F et G sont C-équivalents, c'est dire par définition qu'il existe un difféomorphisme θ de $\mathbb{R}^m \times \mathbb{R}^\ell$ et une famille $Q_{(x,t)}$ de difféomorphismes entre $\mathbb{R} \times \mathbb{R}^\ell \times \{(x,t)\}$ et $\mathbb{R} \times \mathbb{R}^\ell \times \{\theta(x,t)\}$ tels que l'on ait $G(x,t) = Q_{(x,t)}(F(\theta(x,t)))$. Or si F' et G' sont équivalents comme déploiements, on a

$$\begin{aligned} G(x,t) &= (G'(x,t), t) = (F'(H_1(x,t), \varphi(t)) + \psi(t), t) \\ &= Q_{(x,t)}(F(\theta(x,t))) = Q_{(x,t)}(F'(\theta_1(x,t), t)) \end{aligned}$$

où $\theta = H$ est un difféomorphisme tenant compte de la structure de produit de $\mathbb{R}^m \times \mathbb{R}^\ell$ et où $Q_{(x,t)}$ est une translation indépendante de x.

Faisons d'ailleurs une remarque à ce propos. L'idée de déploiement universel consiste à plonger optimalement une application $f : M \to N$ structurellement instable dans une application $F : M \times \mathbb{R}^\ell \to N \times \mathbb{R}^\ell$ qui soit, elle, structurellement stable. Mais dans cette optique, F doit être stable au sens des déploiements, *i.e.* au sens d'applications compatibles aux projections canoniques $M \times \mathbb{R}^\ell \to \mathbb{R}^\ell$, $N \times \mathbb{R}^\ell \to \mathbb{R}^\ell$. D'où la question de savoir quel type de stabilité il faudrait considérer pour que les déploiements universels soient structurellement stables non pas simplement comme déploiements mais aussi comme applications. C'est là que l'équivalence de contact introduite par Mather prend tout son sens. Si l'on considère en effet l'équivalence de contact, un déploiement universel de f est une application structurellement stable $F : M \times \mathbb{R}^\ell \to N \times \mathbb{R}^\ell$. Ce résultat est à l'origine de la classification des germes stables par l'anneau local de leur fibre $f^{-1}(0)$.[53]

Munis de la définition de l'équivalence des déploiements, on peut alors développer une théorie de la stabilité en tout point analogue à celle que nous avons esquissée plus haut. C'est bien la stabilité qui est fondamentale puisqu'en déployant une singularité f on cherche précisément à la plonger dans une famille f_t qui soit stable comme famille (l'opération inverse consistant à partir d'une application stable "fibrée" au-dessus d'une base et à chercher les fibres instables). Donnons quelques idées des résultats obtenus.

D'abord on ramène la condition de stabilité à une condition de transversalité dans les espaces de jets, *i.e.* à une condition de stabilité infinitésimale, et l'on montre réciproquement que la stabilité infinitésimale implique la stabilité. Ensuite on prouve un critère de stabilité. Soit $\rho(f)$ le plus petit k tel que $\mathfrak{m}^{k+1} \subset \Delta(f)$. D'après le corollaire 68, $\rho(f)$ existe si f est de codimension finie et $\rho(f) \leq \operatorname{codim}(f)$. Puisque l'on cherche des déploiements universels, on ne considère que le cas $\ell \geq \operatorname{codim}(f)$. Disons que ℓ germes $h_1, \ldots, h_\ell$ constituent

[53] *Cf.* Chenciner [10] §9 et Thom [51] §3.2.5.

un *système transverse* pour f si, avec les fonctions constantes, ils engendrent $\mathcal{E}$ comme espace vectoriel sur $\mathbb{R}$ modulo $\Delta(f) + \mathfrak{m}^{\rho(f)+1}$.

Proposition 80. – Le déploiement (f_t) est stable si et seulement si les $\dfrac{\partial f_t}{\partial t_i}$, $i = 1, \ldots, \ell$ constituent un système transverse pour f et, plus précisément, si et seulement si f_t est (à équivalence près) de la forme $f_t(x) = f(x) + \sum_{i=1}^{\ell} t_i h_i(x)$ où les h_i constituent un système transverse pour f et sont des polynômes de degré $\leq \rho(f)$.

Si f_t est un déploiement, notons V_t le sous-espace vectoriel de $\mathcal{E}$ engendré par 1 et les $\dfrac{\partial f_t}{\partial t_i}$, $i = 1, \ldots, \ell$. Disons que f_t est k-transversal si $\mathcal{E} = \Delta + V_t + \mathfrak{m}^{k+1}$. D'après la proposition 80, f_t est stable si et seulement si il est $\rho(f)$-transversal. Cette condition peut s'exprimer de façon très géométrique dans les espaces de jets J^k (où k est supérieur à l'ordre de détermination de f). Soit f un germe de $\mathfrak{m}^2$ et σ son k-jet. σ ne tient compte que du germe de f en 0 et ne dit rien sur les germes de f en $x_0 \neq 0$. Pour pallier cette difficulté, on peut, pour $x_0 \neq 0$, considérer la fonction $f(x_0 + x) - f(x_0)$ dont le germe en x_0 est un élément f_{x_0} de $\mathfrak{m}$ et définir le *prolongement* de f comme le germe de l'application qui à $x_0 \in \mathbb{R}^m$ associe f_{x_0}. On définit de même le prolongement du k-jet de f, le prolongement d'un déploiement et le prolongement du k-jet d'un déploiement.

Proposition 81. – Le déploiement f_t est k-transversal si et seulement si le prolongement de son k-jet est transverse sur l'orbite de $j^k f$ dans J^k.

Ce résultat permet facilement de construire des déploiements transversaux (et donc stables) des singularités f de codimension finie. Il suffit de considérer une base $h_1, \ldots, h_c$ ($c = \mathrm{codim}(f)$) de $\mathfrak{m}/\Delta$ et de définir f_w par

$$f_w = f + \sum_{i=1}^{c} w_i h_i. \tag{35}$$

Ces déploiements sont transversaux pour tout $k \geq 0$.

Le résultat fondamental sur les déploiements est que tous les déploiements transverses de bonne dimension $c = \mathrm{codim}(f)$ sont non seulement stables mais universels et tous équivalents, autrement dit, qu'à équivalence près, il n'existe qu'*un* déploiement universel d'une singularité de codimension finie.

Théorème 82.[54] *– Si f est de détermination finie k et de codimension finie c, alors :*

(i) *f admet des déploiements universels de dimension minimale c,*

(ii) *tout déploiement k-transversal est universel,*

[54] *Cf.* par exemple Zeeman [56].

(iii) *tous ces déploiements sont équivalents s'ils ont même dimension.*

Preuve. – Esquissons la démonstration. Il est d'abord assez facile de vérifier que si f_t est un déploiement universel de base $T = \mathbb{R}^\ell$, alors il est k-transversal pour tout $k > 0$ et $\ell \geq \operatorname{codim}(f)$. Soit en effet f_w le déploiement (35). Comme f_t est universel, il existe un morphisme $W \to T$ tel que f_w s'obtienne par image réciproque de f_t. On montre que cela implique $V_w \subset \Delta + V_t$. Or comme f_w est transversal, on a $\mathcal{E} = \Delta + V_w$ et donc $\mathcal{E} = \Delta + V_t$, *i.e.* f_t est k-transversal pour tout $k > 0$. D'autre part, $\ell \geq \dim(V_t) - 1 \geq \dim(\mathfrak{m}/\Delta) = \operatorname{codim}(f)$.

Mais le point le plus délicat de la preuve est de montrer que si f_t et g_u sont deux déploiements k-transversaux de f ($k \geq \operatorname{codim}(f)$) de même dimension ℓ, alors ils sont équivalents. Notons V_t' et V_u' les espaces vectoriels $V_t - \{\text{constantes}\}$ et $V_u - \{\text{constantes}\}$, *i.e.* les sous-espaces vectoriels de $\mathfrak{m}$ engendrés par les $\frac{\partial f_t}{\partial t_i}$ et les $\frac{\partial g_u}{\partial u_i}$ pour $i = 1, \ldots, \ell$. Dire que f_t est k-transversal c'est dire que $\mathfrak{m} = \Delta + V_t' + \mathfrak{m}^{k+1}$. Comme f est k-déterminée, $\mathfrak{m}^{k+1} \subset \mathfrak{m}\Delta \subset \Delta$ et donc $\mathfrak{m} = \Delta + V_t'$. Pour les mêmes raisons $\mathfrak{m} = \Delta + V_u'$. Autrement dit les systèmes des $\frac{\partial f_t}{\partial t_i}$ et des $\frac{\partial g_u}{\partial u_i}$, $i = 1, \ldots, \ell$, engendrent tous deux $\mathfrak{m}/\Delta$. On montre alors que l'on peut, en prenant un déploiement universel équivalent à f_t, se ramener au cas où $\frac{\partial f_t}{\partial t_i} = \frac{\partial g_u}{\partial u_i}$, $i = 1, \ldots, \ell$. On considère ensuite l'homotopie $f_{t,s} = (1-s)\, f_t + s g_u$, $s \in I = [0,1]$, entre f_t et g_u et on montre que $\frac{\partial f_t}{\partial t_i} = \frac{\partial g_u}{\partial u_i}$, $i = 1, \ldots, \ell$, implique que $f_{t,s}$ soit k-transversal pour tout $s \in I$ et que cette condition de transversalité permet de construire les difféomorphismes d'une équivalence déformant l'identité sur I et faisant passer de f_t à $f_{t,s}$ et donc à g_u pour $s = 1$.

Une fois démontré ce résultat clé, il est trivial de vérifier que deux déploiements universels f_t et g_u de même dimension ℓ sont isomorphes. Ils sont en effet k-transversaux pour tout $k > 0$. Réciproquement, on montre que si f_t est k-transversal ($k \geq \det(f)$), il est universel. Enfin, le déploiement f_w de (35) est universel, transversal et de dimension minimale. □

Pour conclure cette section, donnons les déploiements universels des singularités simples.

1. Les cuspoïdes ou queues d'arondes $A_{k-1} = x^k$, $k \geq 3$:

$$f_w = x^k + a_2 x^{k-2} + \cdots + a_{k-1} x.$$

2. Les ombilics $D_{k+1} = x^2 y \pm y^k$, $k \geq 3$:

$$f_w = x^2 y \pm y^k + a_1 y^{k-1} + \cdots + a_{k-1} y + bx.$$

3. $E_6 = x^3 \pm y^4$:

$$f_w = x^3 \pm y^4 + axy^2 + bxy + cy^2 + dy + ex.$$

4. $E_7 = x^3 + xy^3$:

$$f_w = x^3 + xy^3 + ay^4 + by^3 + cy^2 + dxy + ex + fy.$$

5. $E_8 = x^3 + y^5$:

$$f_w = x^3 + y^5 + axy^3 + bxy^2 + cy^3 + dy^2 + exy + fx + gy.$$

12 La géométrie des catastrophes élémentaires

12.1 Du concept à la géométrie

Nous avons vu comment les concepts constitutifs du modèle général pouvaient être progressivement transformés en une théorie géométrique. Les étapes de cette progression du conceptuel vers le géométrique sont les suivantes :

(i) spécification du modèle général en termes d'étude de la structure des espaces fonctionnels $\mathcal{F}$ (analyse des ensembles catastrophiques $K_{\mathcal{F}}$);

(ii) réduction à la dimension finie et démonstration de l'existence de formes normales algébriques;

(iii) étude de la géométrie des applications catastrophiques associées à ces formes normales.

D'après les théorèmes difficiles évoqués dans les sections précédentes, cette géométrie est essentiellement de type "géométrie des discriminants". Elle doit être interprétée à équivalence différentiable près.

Soit $f \in \mathfrak{m}^2$ une singularité de codimension finie $c = \mathrm{codim}(f)$. Soit (f_w) un déploiement de f de dimension ℓ ayant pour base W un voisinage de l'origine de $\mathbb{R}^\ell$. Soit Σ_W le lieu critique de (f_w), *i.e.* l'ensemble des $(x, w) \in \mathbb{R}^m \times \mathbb{R}^\ell$ tels que x soit un point critique de f_w.[55] Σ_W est défini (localement) dans $\mathbb{R}^m \times \mathbb{R}^\ell$ par les équations $\dfrac{\partial f_w}{\partial x_i} = 0$, $i = 1, \ldots, m$. Comme $f \in \mathfrak{m}^2$, $0 \in \Sigma_W$ et l'on peut considérer (le germe en 0 de) l'application $\chi_W : \Sigma_W \to \mathbb{R}^\ell$ restriction à Σ_W de la projection canonique $\mathbb{R}^m \times \mathbb{R}^\ell \to \mathbb{R}^\ell$. χ_W s'appelle l'*application catastrophique* (ou la catastrophe) induite par (f_w).

Considérons le cas le plus élémentaire – dit "catastrophe pli" – celui d'un point critique dégénéré de type x^3 (point d'inflexion). La singularité x^3 est déterminée à l'ordre 3. Elle est de codimension 1 et de déploiement universel $f_u(x) = x^3 + ux$. L'équation $f'_u(x) = 0$ donnant les points critiques est ici l'équation du second degré $3x^2 + u = 0$ et donc, pour $u > 0$, f_u est sans point critique alors que, pour $u < 0$, f_u admet deux points critiques, un maximum et un minimum (*cf.* figure 7). On peut considérer le graphe total V de la fonction à deux variables $f(x, u) = f_u(x)$. Ce graphe est un "relief" au-dessus du plan (x, u) qui a la forme caractéristique d'une surface "fronce" obtenue par évanouissement de deux plis. Son équation est $y = x^3 + ux$, soit $x^3 + ux - y = 0$ (*cf.* figure 8).

[55] Rappelons que Σ_W s'appelle aussi l'(hyper)surface caractéristique ou l'(hyper)surface de catastrophe.

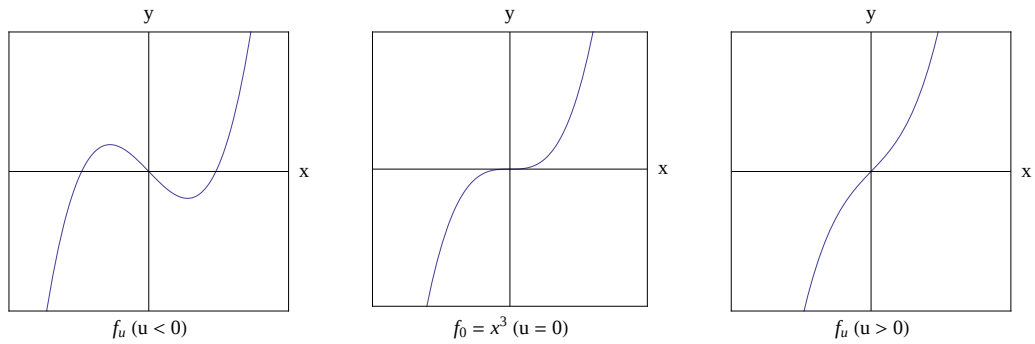


Figure 7: Déploiement universel de la singularité x^3.

Le lieu critique Σ du déploiement est d'équation $f'_u(x) = 3x^2 + u = 0$. C'est une parabole. L'application χ est la projection de Σ sur l'axe u. On remarquera que le sommet de la parabole Σ est en $(u = 0, x = 0)$ et que, pour cette valeur, la fibre $\chi^{-1}(u)$ change de type : pour $u > 0$, $\chi^{-1}(u)$ est vide alors que pour $u < 0$, $\chi^{-1}(u)$ est constituée de 2 points. Le collapse des deux points critiques de f_u $(u < 0)$ en $u = 0$ peut s'exprimer de façon imagée en disant que Σ se "plisse" au-dessus de $u = 0$. D'où le nom de *catastrophe pli* donnée à la catastrophe de bifurcation associée à la singularité x^3. On remarquera aussi que l'ensemble catastrophique K de W (W est ici réduit à l'axe des u) est le point $u = 0$ et que ce point est le *contour apparent* du lieu critique Σ relativement à la projection $\chi : \Sigma \to W$ c'est-à-dire l'image par χ de l'ensemble Γ des points de Σ où la direction de projection est tangente à Σ. En effet, dire que le point (x, u) de Σ appartient à Γ, c'est dire non seulement que x est point critique de f_u mais aussi qu'en ce point $f''_u(x) = 0$, que x est donc un point critique *dégénéré* de f_u et donc que $u \in K$. D'ailleurs, le lieu critique Σ peut lui-même s'obtenir à partir d'un autre lieu critique C, celui de la surface V relativement à la projection π du $\mathbb{R}^3$ de coordonnées (x, u, y) sur le plan (u, y). V est en effet d'équation $f_u(x) - y = 0$. Son lieu critique C, relativement à π, est donc d'équation $f'_u(x) = 0$ et la projection de C sur le plan (x, u) redonne bien Σ (*cf.* figure 8).

Enfin, on peut considérer le graphe des *valeurs critiques* de f_u, c'est-à-dire l'ensemble des points du plan (u, y) tels que y soit la valeur $f_u(x)$ de f_u en un de ses points critiques. Il est facile de vérifier que ce graphe est le contour apparent $\pi(C)$ de V relativement à la projection π (*cf.* figure 8).

Supposons $f \in \mathfrak{m}^3$ et considérons un déploiement universel (f_w) de f. Les x_i, $i = 1, \ldots, m$ appartiennent à $\mathfrak{m}/\Delta$ puisque $\Delta \subset \mathfrak{m}^2$ et l'on peut donc supposer dans la formule (35) que $h_i = x_i$ pour $i = 1, \ldots, m$, les h_j étant de degré ≥ 2 pour $m < j \leq c$. Alors $\dfrac{\partial f_w}{\partial x_i} = \dfrac{\partial f}{\partial x_i} + w_i + \sum_{j=m+1}^{c} w_j \dfrac{\partial h_j}{\partial x_i}$ et Σ_W est donnée par m-équations polynomiales $w_i = \psi(x_1, \ldots, x_m, w_{m+1}, \ldots, w_c)$,

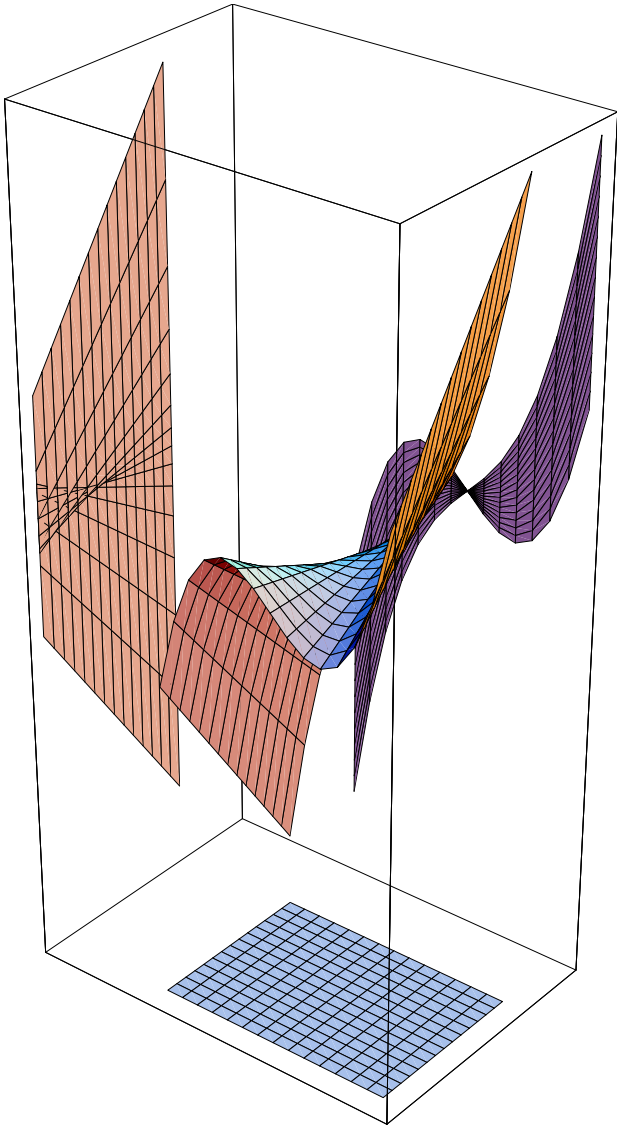


Figure 8: La catastrophe pli.

i.e. par le graphe de ψ. Or un tel graphe est difféomorphe à sa source. Autrement dit, Σ_W est difféomorphe à $\mathbb{R}^c$ et la catastrophe χ_W est une application différentiable entre les variétés équidimensionnelles Σ_W et $\mathbb{R}^c$.

Les applications catastrophiques se transforment naturellement par les morphismes de déploiements. En particulier, si deux déploiements sont équivalents, leurs catastrophes sont différentiablement équivalentes et donc, si l'on se restreint aux déploiements universels (qui sont tous équivalents), il n'existe essentiellement qu'*une* catastrophe déployant une singularité f de codimension finie. Qui plus est, il est facile de montrer que le type différentiable de cette catastrophe ne dépend que du type différentiable de f. Les catastrophes des singularités simples de petite codimension sont appelées *élémentaires*. Elles sont au nombre de 7 en codimension ≤ 4 et au nombre de 11 en codimension ≤ 5. Comme l'on connaît les déploiements universels sous forme normale algébrique, elles sont explicitement (sinon facilement) calculables.

Les catastrophes élémentaires déployant les singularités simples sont stables au sens où si χ_W est induite par un déploiement universel (f_w), pour tout déploiement f_u assez voisin de f_w, $u \in U$, χ_U sera équivalente à χ_W. Mais cela ne signifie pas pour autant qu'elles soient stables dans l'espace des applications $\chi : \mathbb{R}^c \to \mathbb{R}^c$. Elles ne sont stables qu'en tant qu'applications de $\mathbb{R}^c \to \mathbb{R}^c$ *induites* par des déploiements (f_w). Certaines singularités peuvent apparaître stablement dans une application χ quelconque sans pouvoir pour autant apparaître dans les χ_W. Réciproquement, certaines catastrophes stables comme catastrophes peuvent être instables si on les considère comme applications quelconques. En codimension 2, il y a équivalence car, d'après le théorème 47 de Whitney-Thom, les deux seules singularités stables $\chi : \mathbb{R}^2 \to \mathbb{R}^2$ sont le pli et le cusp c'est-à-dire précisément les deux catastrophes élémentaires. Mais en codimension 4, il y a divergence. Seules les 4 cuspoïdes ou queues d'aronde sont *à la fois* des singularités stables et des catastrophes élémentaires. Les ombilics ne sont pas des singularités stables et les singularités stables de type S_2 ne sont pas des catastrophes élémentaires.[56]

12.2 Linéarisation des catastrophes

Avant que d'en venir à la géométrie des catastrophes élémentaires, indiquons en suivant Christopher Zeeman [57] comment on peut simplifier une application catastrophique en la *linéarisant* en partie.

12.2.1 Le cadre général

Soient f_w un déploiement universel de f de dimension $c = \mathrm{codim}(f)$ finie et $\chi_W : \Sigma_W \to W$ la catastrophe induite. Soit k l'ordre de f: $f \in \mathfrak{m}^k$, $k \geq 3$, $f \notin \mathfrak{m}^{k+1}$. Le déploiement f_w est associé à une base $h_1, \ldots, h_c$ de $\mathfrak{m}/\Delta$ par la formule (35) : $f_w = f + \sum_{r=1}^{r=c} w_r h_r$. Σ_W est la sous-variété de $\mathbb{R}^m \times \mathbb{R}^c =$

[56] *Cf.* Zeeman [56].

$\mathbb{R}^m \times \mathfrak{m}/\Delta = \mathbb{R}^m \times W$ d'équations $\dfrac{\partial f_w}{\partial x_i} = 0$, $i = 1, \ldots, m$ et $\chi_W : \Sigma_W \to \mathfrak{m}/\Delta$ est la restriction à Σ_W de la projection canonique $\mathbb{R}^m \times \mathfrak{m}/\Delta \to \mathfrak{m}/\Delta$. L'idée est alors de linéariser Σ_W et de factoriser χ_W à travers cette linéarisation.

Considérons l'application $\widetilde{f} : \mathbb{R}^m \times \mathfrak{m}/\Delta \to \mathfrak{m}$ qui associe à (x, w) le germe de la fonction $\widetilde{f}_{(x,w)}(\xi) = f_w(x + \xi) - f_w(x)$, $\xi \in \mathbb{R}^m$. $\widetilde{f}$ est le prolongement de f_w obtenu en translatant l'origine $(0, 0)$ en (x, w). Or dans les espaces $\mathfrak{m}^k$ et $\mathfrak{m}^k/\mathfrak{m}^\ell$ il existe des stratifications canoniques S_k et $S_{k,\ell}$ données par les orbites pour l'équivalence (à droite). Elles ne dépendent pas de f mais uniquement de m, k, ℓ.

Lemme 1. – $\Sigma_W = \widetilde{f}^{-1}(\mathfrak{m}^2)$.

Preuve. – Considérons en effet le développement de Taylor de $\widetilde{f}_{(x,w)}$. On a $\widetilde{f}_{(x,w)}(\xi) = \xi \frac{\partial \widetilde{f}_{(x,w)}}{\partial \xi}(0) + \cdots$ et donc :

$$\widetilde{f}_{(x,w)} \in \mathfrak{m}^2 \Leftrightarrow \frac{\partial \widetilde{f}_{(x,w)}}{\partial \xi}(0) = 0 \tag{36}$$

Mais comme $\frac{\partial \widetilde{f}_{(x,w)}}{\partial \xi} = \left(\frac{\partial f_w}{\partial x_1}(x + \xi), \ldots, \frac{\partial f_w}{\partial x_m}(x + \xi)\right)$, (36) est équivalente à $\dfrac{\partial f_w}{\partial x_i}(x) = 0$, $i = 1, \ldots, m$, *i.e.* à $(x, w) \in \Sigma_W$. □

Ceci dit, il y a deux façons de stratifier χ_W et sa source Σ_W. Soit on considère χ_W comme une application (stable) entre variétés équidimensionnelles et on lui applique les techniques de Thom-Boardman, soit on considère en chaque point $(x, w) \in \Sigma_W$ le type de $\widetilde{f}_{(x,w)}$ et on stratifie Σ_W par l'image réciproque $\widetilde{f}^*(S_1)$ de la stratification canonique de $\mathfrak{m}$ par $\widetilde{f}$. Dans ce cas, puisque $\Sigma_W = \widetilde{f}^{-1}(\mathfrak{m}^2)$ d'après le lemme 1, $\mathrm{Strat}(\Sigma_W) = \widetilde{f}^*(S_2)$. Le rapport entre ces deux points de vue est le suivant. Considérons d'abord $S_1(\chi_W)$, *i.e.* le lieu singulier de χ_W : $(x, w) \in \mathrm{Sing}(\chi_W) = S_1(\chi_W)$ si et seulement si l'application linéaire tangente $D_{(x,w)}\chi_W$ n'est pas de rang maximal c en (x, w). Or Σ_W est l'intersection (régulière) des hypersurfaces Σ_i, $i = 1, \ldots, m$, de $\mathbb{R}^m \times W = \mathbb{R}^m \times \mathfrak{m}/\Delta$ d'équations respectives $\dfrac{\partial f_w}{\partial x_i}(x) = 0$. Les vecteurs tangents à Σ_i en (x, w) sont les vecteurs (ξ, μ) de $T_{(x,w)}(\mathbb{R}^m \times \mathfrak{m}/\Delta)$ satisfaisant :

$$d\left(\frac{\partial f_w}{\partial x_i}\right)(\xi, \mu) = 0 \ , \ \textit{i.e.} \tag{37}$$
$$\left(\frac{\partial^2 f_w}{\partial x_i \partial x}\middle|\, \xi\right) + \left(\frac{\partial^2 f_w}{\partial x_i \partial w}\middle|\, \mu\right) = 0 \ .$$

où $\left(\dfrac{\partial^2 f_w}{\partial x_i \partial x}\middle|\, \xi\right)$ est le produit scalaire $\sum_{j=1}^{j=m} \dfrac{\partial^2 f_w}{\partial x_i \partial x_j}\xi_j$ et $\left(\dfrac{\partial^2 f_w}{\partial x_i \partial w}\middle|\, \mu\right)$ le produit scalaire $\sum_{r=1}^{r=c} \dfrac{\partial^2 f_w}{\partial x_i \partial w_r}\mu_r$. Dire que $(x, w) \in \mathrm{Sing}(\chi_W) = S_1(\chi_W)$ c'est

dire que $D_{(x,w)}\chi_W$ annule un vecteur de $T_{(x,w)}(\Sigma_W)$. Mais comme χ_W est la restriction d'une *projection* $\mathbb{R}^m \times W \to W$, cela signifie qu'il existe un vecteur $(\xi, \mu) \in T_{(x,w)}(\Sigma_W)$ qui est *vertical*, *i.e.* de la forme $(\xi, 0)$. D'après (37), cela équivaut à dire que l'on a $\left(\frac{\partial^2 f_w}{\partial x_i \partial x}\middle|\xi\right) = 0$ pour $i = 1, \ldots, m$, soit :

$$\sum_{j=1}^{j=m} \frac{\partial^2 f_w}{\partial x_i \partial x_j} \xi_j = 0, \text{pour } i = 1, \ldots, m. \tag{38}$$

Or cela signifie que le hessien $H(f_w)$ de f_w en x n'est pas de rang maximal, autrement dit que x est un point critique dégénéré de f_w ou encore que 0 est un point critique dégénéré de $\widetilde{f}_{(x,w)}$. La stratification $S_1(\Sigma_W)$ est donc l'image réciproque par $\widetilde{f}$ du complémentaire dans $\mathfrak{m}^2$ des strates régulières de S_2. Qui plus est, comme le type de χ_W en (x, w) ne dépend que du type de $\widetilde{f}_{(x,w)}$, $\mathrm{Strat}(\chi_W)$ au sens de $\widetilde{f}^*(S_2)$ est *plus fine* que $\mathrm{Strat}(\chi_W)$ au sens de Thom-Boardman. Nous avons vu par exemple que la stratification de Thom-Boardman ne distingue pas les points critiques non dégénérés suivant leur indice.

Supposons $f \in \mathfrak{m}^k$, $k \geq 3$. On a $\Delta \subset \mathfrak{m}^{k-1}$ et $\mathfrak{m}\Delta \subset \mathfrak{m}^k \subset \mathfrak{m}^2$. On peut donc considérer la projection canonique $\mathfrak{m}^2/\mathfrak{m}\Delta \to \mathfrak{m}^2/\mathfrak{m}^k$, la stratification canonique $S = S_{2,k}$ de $\mathfrak{m}^2/\mathfrak{m}^k$ et son image réciproque S^* dans $\mathfrak{m}^2/\mathfrak{m}\Delta$. Soit alors l'application composée

$$\theta' : \mathbb{R}^m \times \mathfrak{m}/\Delta \xrightarrow{\widetilde{f}} \mathfrak{m} \xrightarrow{\pi} \mathfrak{m}/\mathfrak{m}\Delta$$

où π est la projection canonique.

Lemme 2 (Zeeman). – θ' est un difféomorphisme local.

Preuve. – Pour prouver ce lemme, il suffit de montrer, d'après le théorème 3 d'inversion locale, que l'application linéaire tangente $D_0\theta'$ de θ' en $(0, 0)$ est un isomorphisme. Pour cela, remarquons d'abord que $\mathbb{R}^m \times \mathfrak{m}/\Delta$ et $\mathfrak{m}/\mathfrak{m}\Delta$ sont de même dimension. En effet, $\dim_{\mathbb{R}}(\mathbb{R}^m \times \mathfrak{m}/\Delta) = m + c$ et d'autre part

$$\dim_{\mathbb{R}}(\mathfrak{m}/\mathfrak{m}\Delta) = \dim_{\mathbb{R}}(\mathfrak{m}/\Delta) + \dim_{\mathbb{R}}(\Delta/\mathfrak{m}\Delta) = c + \dim_{\mathbb{R}}(\Delta/\mathfrak{m}\Delta) \ .$$

Or $\dim_{\mathbb{R}}(\Delta/\mathfrak{m}\Delta) = m$ car les $\frac{\partial f}{\partial x_i}$, $i = 1, \ldots, m$, constituent une base de Δ modulo $\mathfrak{m}$. En effet, tout élément $g \in \Delta$ s'écrit comme $g = \sum_{i=1}^{i=m} a_i \frac{\partial f}{\partial x_i}$ avec $a_i \in \mathcal{E}$. Mais $a_i = a'_i + b_i$ avec $a'_i \in \mathbb{R}$ et $b_i \in \mathfrak{m}$. Donc, modulo $\mathfrak{m}\Delta$, $g = \sum_{i=1}^{i=m} a'_i \frac{\partial f}{\partial x_i}$ et les $\frac{\partial f}{\partial x_i}$ engendrent donc $\Delta/\mathfrak{m}\Delta$ comme $\mathbb{R}$-vectoriel. Or ils sont linéairement indépendants car s'ils étaient linéairement dépendants, il existerait des a'_i non tous nuls et des b_i tels que $\sum_{i=1}^{i=m} a'_i \frac{\partial f}{\partial x_i} = \sum_{i=1}^{i=m} b_i \frac{\partial f}{\partial x_i}$.

Si X était le champ sur $\mathbb{R}^m$ défini par $X = \sum_{i=1}^{i=m} (a_i - b_i) \dfrac{\partial f}{\partial x_i}$, on aurait donc $X(f) = 0$. Mais X serait régulier en 0 car les a_i ne sont pas tous nuls. On pourrait donc, par changement de coordonnées locales, écrire $X = \dfrac{\partial}{\partial x_1}$. Cela impliquerait que f ne dépende que de $(x_2, \ldots, x_m)$. Or cela est impossible puisque f est de codimension finie.

Revenons au lemme 2. Les vecteurs tangents en $(0,0)$ à $\mathbb{R}^m \times \mathfrak{m}/\Delta$ sont de la forme (ξ, μ). Pour les vecteurs de type $(\xi, 0)$, $D_0\theta'$ est un isomorphisme entre $T_0\mathbb{R}^m$ et le sous-espace de $\mathfrak{m}/\mathfrak{m}\Delta$ engendré par les $\dfrac{\partial f_w}{\partial x_i}$ pour $w = 0$, autrement dit, entre $T_0\mathbb{R}^m$ et $\Delta/\mathfrak{m}\Delta$. Pour les vecteurs de type $(0, \mu)$, $D_0\theta'$ est, par construction de f_w, un isomorphisme entre $T_0(\mathfrak{m}/\Delta)$ et un sous-espace supplémentaire de Δ dans $\mathfrak{m}/\mathfrak{m}\Delta$. Cela démontre le lemme. $\square$

Comme $\mathfrak{m}\Delta \subset \mathfrak{m}^2$ et comme, d'après le lemme 1, $\Sigma_W = \widetilde{f}^{-1}(\mathfrak{m}^2)$, θ' induit un difféomorphisme θ entre Σ_W et $\mathfrak{m}^2/\mathfrak{m}\Delta$, qui sont tous les deux de dimension c. Mais, contrairement à Σ_W, $\mathfrak{m}^2/\mathfrak{m}\Delta$ est un *espace vectoriel* et l'on peut donc dire que θ *linéarise* Σ_W. Par le biais de θ, on peut remplacer l'application catastrophique χ_W par une application $\chi = \chi_W \circ \theta^{-1} : \mathfrak{m}^2/\mathfrak{m}\Delta \to \mathfrak{m}/\Delta = W$ rendant commutatif le diagramme suivant :

$$\begin{array}{ccc} \Sigma_W & \xrightarrow{\theta} & \mathfrak{m}^2/\mathfrak{m}\Delta \\ \chi_W \searrow & & \swarrow \chi \\ & \mathfrak{m}/\Delta = W & \end{array}$$

L'application χ_W est une projection et donc une application "simple" (linéaire) entre une variété "compliquée" Σ_W et W. Au contraire, χ est une application "compliquée" entre une variété "simple" (linéaire) $\mathfrak{m}^2/\mathfrak{m}\Delta$ et W.

Dans $\mathfrak{m}^2/\mathfrak{m}\Delta$, on peut considérer la stratification canonique S^* image réciproque de $S_{2,k} = S$ par la projection $\mathfrak{m}^2/\mathfrak{m}\Delta \to \mathfrak{m}^2/\mathfrak{m}^k$. Il est facile de voir que la stratification de χ image par θ de $\mathrm{Strat}(\chi_W) = f^*(S_2)$ *raffine* S^*, le raffinement ne portant que sur les strates de S dont les jets sont non déterminés à l'ordre k.

12.2.2 Les principaux exemples

Explicitons les applications linéarisantes χ et la stratification S^* pour les catastrophes élémentaires.

Les cuspoïdes $\mathbf{A_{k-1}}$. Soit A_{k-1} la queue d'aronde x^k ou, mieux, $\frac{1}{k}x^k$. On a $\Delta = \left(x^{k-1}\right) = \mathfrak{m}^{k-1}$, $\mathfrak{m}\Delta = \mathfrak{m}^k$ et $\mathrm{codim}(f) = k-2$. Le déploiement universel est

$$f_w = \frac{1}{k}x^k + \frac{a_{k-2}}{k-2}x^{k-2} + \cdots + a_1 x$$

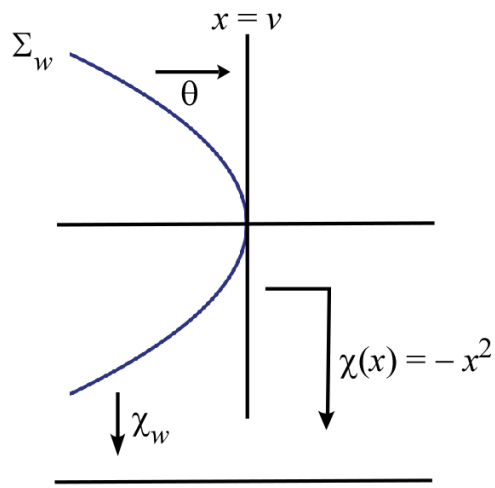


Figure 9: La linéarisation de la catastrophe pli (voir texte).

et $\mathfrak{m}/\Delta$ a pour base $\left(x, \frac{x^2}{2}, \ldots, \frac{x^{k-2}}{k-2}\right)$. Donc

$$\widetilde{f}_{(x,w)}(\xi) = \xi f'_w(x) + \frac{\xi^2}{2} f''_w(x) + \cdots + \frac{\xi^k}{k!} f^{(k)}_w(x)$$

et $\mathfrak{m}/\mathfrak{m}\Delta = \mathfrak{m}/\mathfrak{m}^k$ a pour base $\left(\xi, \xi^2, \ldots, \xi^{k-1}\right)$ et pour coordonnées $(u_1, \ldots, u_{k-1})$. Dans ces bases, θ' est l'application qui associe à $(x, a_1, \ldots, a_{k-2}) = (x, w) \in \mathbb{R} \times \mathfrak{m}/\Delta$ le $(k-1)$-uple $\left(f'_w(x), \ldots, \frac{f^{(k-1)}_w(x)}{(k-1)!}\right)$ et $(x, w) \in \Sigma_W$ si et seulement si $f'_w(x) = 0$, autrement dit, si et seulement si $\theta'(x, w) \in \mathfrak{m}^2/\mathfrak{m}\Delta$ qui est de base $\left(\xi^2, \ldots, \xi^{k-1}\right)$ et de coordonnées $(u_2, \ldots, u_{k-1})$, et dans ce cas, on a $\theta(x, w) = \left(\frac{f''_w(x)}{2}, \ldots, \frac{f^{(k-1)}_w(x)}{(k-1)!}\right)$. D'où le système d'équations :

$$\begin{cases} u_2 = \frac{1}{2}\left[(k-1)x^{k-2} + (k-3)a_{k-2}x^{k-4} + \cdots + 3a_4x^2 + 2a_3x + a_2\right] \\ u_3 = \frac{1}{6}\left[(k-1)(k-2)x^{k-3} + (k-3)(k-4)a_{k-2}x^{k-5} + \cdots + 6a_4x + 2a_3\right] \\ \vdots \\ u_{k-1} = x \end{cases} \tag{39}$$

Dans le cas du pli x^3, si l'on pose $a_1 = a$, $u_1 = u$ et $u_2 = v$, on obtient $\theta' : (x, a) \mapsto (u = x^2 + a, v = x)$. Σ_W est, dans le plan (x, a), la parabole $x^2 + a = 0$ et θ est la projection de Σ_W sur l'axe des x. Quant à χ_W, c'est la projection sur l'axe des a et χ est l'application d'équation $a = \chi(x) = -x^2$ (cf. figure 9).

La stratification canonique de $\mathfrak{m}^2/\mathfrak{m}\Delta = \mathfrak{m}^2/\mathfrak{m}^k = (u_2, \ldots, u_{k-1})$ est particulièrement évidente dans le cas des cuspoïdes. Elle comprend 2 strates régulières $u_2 > 0$ (minima) et $u_2 < 0$ (maxima), une strate de codimension 1 $u_2 = 0$ et $u_3 \neq 0$ (plis), 2 strates de codimension 2 $u_2 = 0$, $u_3 = 0$ et $u_4 > 0$ (cusps) et $u_2 = 0$, $u_3 = 0$ et $u_4 < 0$ (cusps duaux), etc. Autrement dit, c'est un drapeau alternant des espaces et des demi-espaces de codimension décroissante correspondant aux cuspoïdes A_ℓ pour $\ell \leq k-1$. Comme A_{k-1} est déterminée à l'ordre k, cette stratification ne comprend aucune strate indéterminée et la stratification de χ_W est simplement son image inverse par θ. Dans le cas du pli, la seule strate singulière est $u_2 = v = 0$ correspondant au sommet de la parabole.

Dans le cas du cusp $A_3 = \frac{x^4}{4}$, avec $a_2 = u$ et $a_3 = v$, le déploiement universel est donné par $\frac{x^4}{4} + u\frac{x^2}{2} + vx$ et θ' est donc l'application :

$$\begin{cases} u_1 = x^3 + ux + v \\ u_2 = \frac{3}{2}x^2 + \frac{u}{2} \\ u_3 = x \end{cases} \tag{40}$$

et θ est l'application :

$$\begin{cases} u_2 = \frac{3}{2}x^2 + \frac{u}{2} \\ u_3 = x \end{cases} \tag{41}$$

D'après (41), la surface Σ_W de $\mathbb{R}^3\,(x, u, v)$[57] est la surface d'équations paramétriques (en fonction de u_2 et u_3) :

$$\begin{cases} x = u_3 \\ u = 2u_2 - 3u_3^2 \\ v = 2u_3^3 - 2u_2u_3 \end{cases} \tag{42}$$

et χ est donc l'application de (u_2, u_3) dans (u, v) d'équations $u = 2u_2 - 3u_3^2$, $v = 2u_3^3 - 2u_2u_3$ (cf. figure 10).

L'ombilic hyperbolique $\mathbf{D_4^+}$. Considérons maintenant l'ombilic hyperbolique sous sa forme $f = x^2y + \frac{1}{3}y^3$. On a $\Delta = (2xy, x^2 + y^2)$ et $\mathfrak{m}\Delta = \mathfrak{m}^3$. Prenons pour base de $\mathfrak{m}/\Delta$ les fonctions $(-x, -y, x^2)$. Le déploiement universel de l'ombilic s'écrit (forme un peu différente de celle que nous avons donnée plus haut) :

$$f_w = x^2y + \frac{1}{3}y^3 + tx^2 - ux - vy\ . \tag{43}$$

Choisissons pour $\mathfrak{m}^2/\mathfrak{m}\Delta = \mathfrak{m}^2/\mathfrak{m}^3$ la base $\left(\xi^2, 2\xi\eta, \eta^2\right)$ et les coordonnées (a, b, c). On a

[57] $\mathbb{R}^3\,(x, u, v)$ dénote le $\mathbb{R}^3$ de coordonnées (x, u, v).

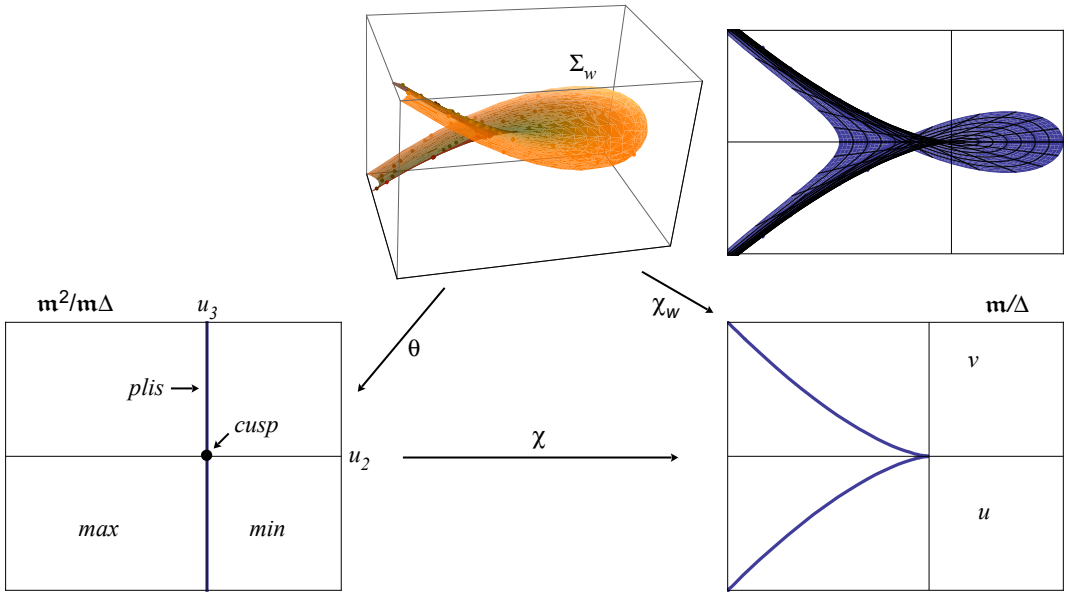


Figure 10: La linéarisation du cusp. On a utilisé la paramétrisation du cone C donnée par $a = r(1-\cos(\varphi))$, $b = r\sin(\varphi)$, $b = r(1+\cos(\varphi))$, $r \in [-1,1]$, $\varphi \in [0,2\pi]$. La surface fronce représentée est la surface d'équations paramétriques $x = u_3$, $u = 2u_2 - 3u_3^2$, $v = 2u_3^3 - 2u_2u_3$ avec $u_2 \in [-0.5, 0.5]$, $u_3 \in [-1,1]$. Sa projection sur le plan (u,v) est représentée en haut à droite. On y voit clairement son lieu singulier en cusp. D'après Zeeman [57].

$$\widetilde{f}_{(x,w)}(\xi,\eta) = \xi(2xy+2tx-u)+\eta\left(x^2+y^2-v\right)+\xi^2(y+t)+ \\ 2\xi\eta x+\eta^2 y+\xi^2\eta+\frac{\eta^3}{3}$$

ce qui donne le difféomorphisme $\theta' : \mathbb{R}^2 \times \mathfrak{m}/\Delta \to \mathfrak{m}/\mathfrak{m}\Delta = \mathfrak{m}/\mathfrak{m}^3$. La sous-variété Σ_W de dimension 3 est d'équations $2xy+2tx-u=0$ et $x^2+y^2-v=0$ dans le $\mathbb{R}^5(x,y,t,u,v)$. Elle est envoyée difféomorphiquement par θ sur $\mathfrak{m}^2/\mathfrak{m}\Delta = \mathbb{R}^3(a,b,c)$ par les équations

$$\begin{cases} a = y+t \\ b = x \\ c = y \end{cases} \tag{44}$$

et l'application $\chi : \mathfrak{m}^2/\mathfrak{m}\Delta(a,b,c) \to \mathfrak{m}/\Delta(t,u,v)$ est donc donnée par les équations :

$$\begin{cases} t = a-c \\ u = 2ab \\ v = b^2+c^2 \end{cases} \tag{45}$$

Les stratifications de χ et χ_W raffinent la stratification canonique $S = S_{2,3}$ de $\mathfrak{m}^2/\mathfrak{m}^3$, *i.e.* la stratification des polynômes homogènes de degré 2 : $p = a\xi^2+2b\xi\eta+c\eta^2$. Le groupe d'équivalence agit sur cet espace en tant que groupe *linéaire* car les polynômes sont homogènes. Or, à transformation linéaire près, les polynômes p sont classés dans $\mathbb{R}^3(a,b,c)$ par le discriminant $b^2-ac=0$ de l'équation du second degré $aX^2+2bX+c=0$. L'équation $b^2-ac=0$ est l'équation d'un cône C dans $\mathbb{R}^3(a,b,c)$ contenant les axes a $(b=c=0)$ et c $(b=a=0)$. S comprend (*cf.* figure 11) :

1. 3 strates régulières de codimension 0 (dimension 3) :
   (a) les minima de type x^2+y^2 correspondant à $ac>b^2$, $a>0$ (intérieur d'un des demi-cônes);
   (b) les maxima de type $-x^2-y^2$ correspondant à $ac>b^2$, $a<0$ (intérieur de l'autre demi-cône);
   (c) les cols de type x^2-y^2 correspondant à $ac<b^2$ (extérieur du cône);
2. 2 strates de codimension 1 (dimension 2)
   (a) les plis de type x^2 correspondant à $ac=b^2$, $a+c>0$ (surface du demi-cône des minima);
   (b) les plis de type $-x^2$ correspondant à $ac=b^2$, $a+c<0$ (surface du demi-cône des maxima);

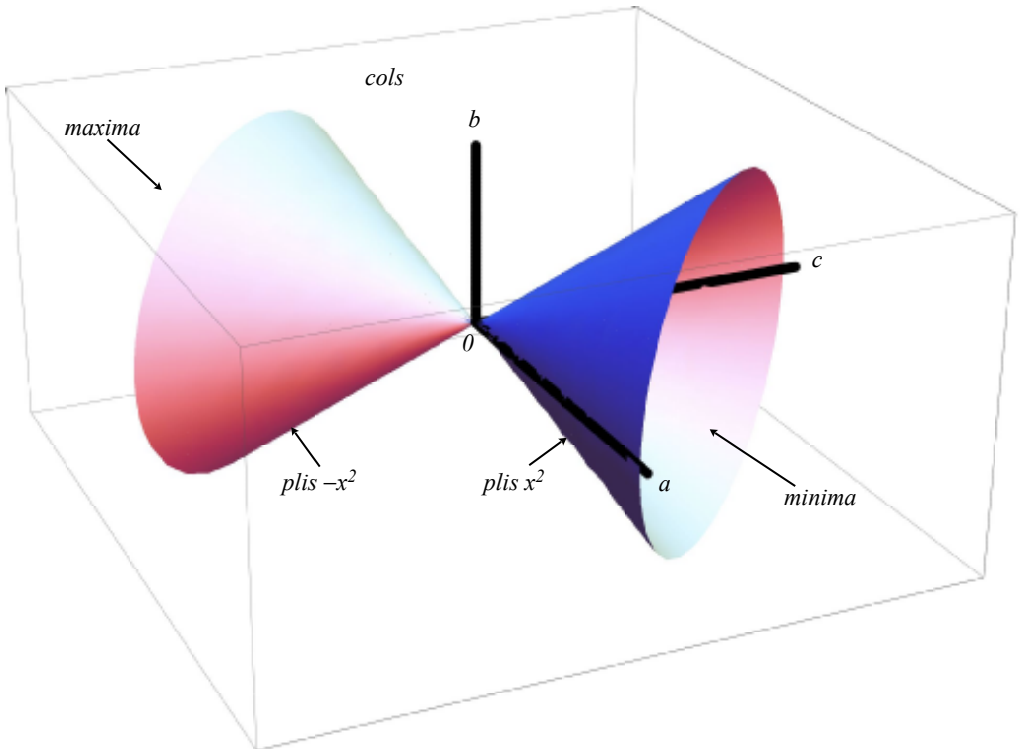


Figure 11: Le cone discriminant $b^2 - ac = 0$ dans $\mathfrak{m}^2/\mathfrak{m}^3$ et sa stratification canonique.

3. la strate singulière de codimension 3 qu'est l'origine $a = b = c = 0$.

On remarquera qu'il n'y a pas de strates de codimension 2 (dimension 1) à cause de la symétrie de rotation du cône autour de son axe et que les strates indéterminées de S sont celles du cône discriminant. Ce sont elles que χ raffine. Le jacobien de χ est la matrice 3×3

$$J = \begin{pmatrix} 1 & 0 & -1 \\ 2b & 2a & 0 \\ 0 & 2b & 2c \end{pmatrix}$$

dont le déterminant est $4\,(ac - b^2)$. Le cône C est donc bien le lieu singulier de χ, $C = S_1\,(\chi)$. Or l'on peut montrer que la restriction $\chi\mid_C$ de χ à C est singulière sur l'axe c (brisure de symétrie). Cette génératrice de C correspond à deux strates supplémentaires de codimension 2 (dimension 1) de χ (une pour chaque demi-axe) composées respectivement de points cusp et de points cusp duaux (cf. figure 12).

L'ombilic elliptique $\mathbf{D_4^-}$. Il en va de même pour l'ombilic elliptique $x^2y - \frac{1}{3}y^3$. L'application χ est donnée par le système :

$$\begin{cases} t = a + c \\ u = 2ab \\ v = b^2 - c^2 \end{cases} \tag{46}$$

Son lieu singulier est C et $\chi\mid_C$ est cette fois singulière le long de 3 génératrice de C, l'axe c et les droites de coefficients directeurs $\left(3, \pm\sqrt{3}, 1\right)$. A travers χ, le cône C se transforme en une déformation d'hypocycloïde à 3 rebroussements (cf. figure 13).

L'ombilic parabolique $\mathbf{D_5}$. Soit maintenant $f = x^2y + \frac{1}{4}y^4$ l'ombilic parabolique. Sa stratification est déjà d'une complexité notable et il est donc particulièrement intéressant de disposer de la stratification canonique qu'elle raffine. Coisissons pour base de $\mathfrak{m}/\Delta$ les fonctions $(-x, -y, x^2, y^2)$ et des coordonnées (t, w, u, v). Considérons le déploiement universel :

$$f_w = x^2y + \frac{1}{4}y^4 + tx^2 + wy^2 - ux - vy\ .$$

Comme $\Delta = (2xy, x^2 + y^3)$, $\mathfrak{m}\Delta$ est engendré par $x^2y, x^3 + y^4, xy^2, x^2y + y^4$. Comme $y^4 = (x^2y + y^4) - x^2y \in \mathfrak{m}\Delta$, $x^3 = (x^3 + y^4) - y^4 \in \mathfrak{m}\Delta$. En revanche $y^3 \notin \mathfrak{m}\Delta$. On peut donc choisir pour base de $\mathfrak{m}^2/\mathfrak{m}\Delta$ la base $\left(\xi^2, 2\xi\eta, \eta^2, \eta^3\right)$ et les coordonnées (a, b, c, d). On a

$$\widetilde{f}_{(x,w)}\,(\xi, \eta) = \xi\,(2xy + 2tx - u) + \eta\left(x^2 + y^3 + 2wy - v\right) + \xi^2\,(y + t) + \\ 2\xi\eta x + \frac{1}{2}\eta^2\left(3y^2 + 2w\right) + \xi^2\eta + \eta^3 + \frac{1}{4}\eta^4$$

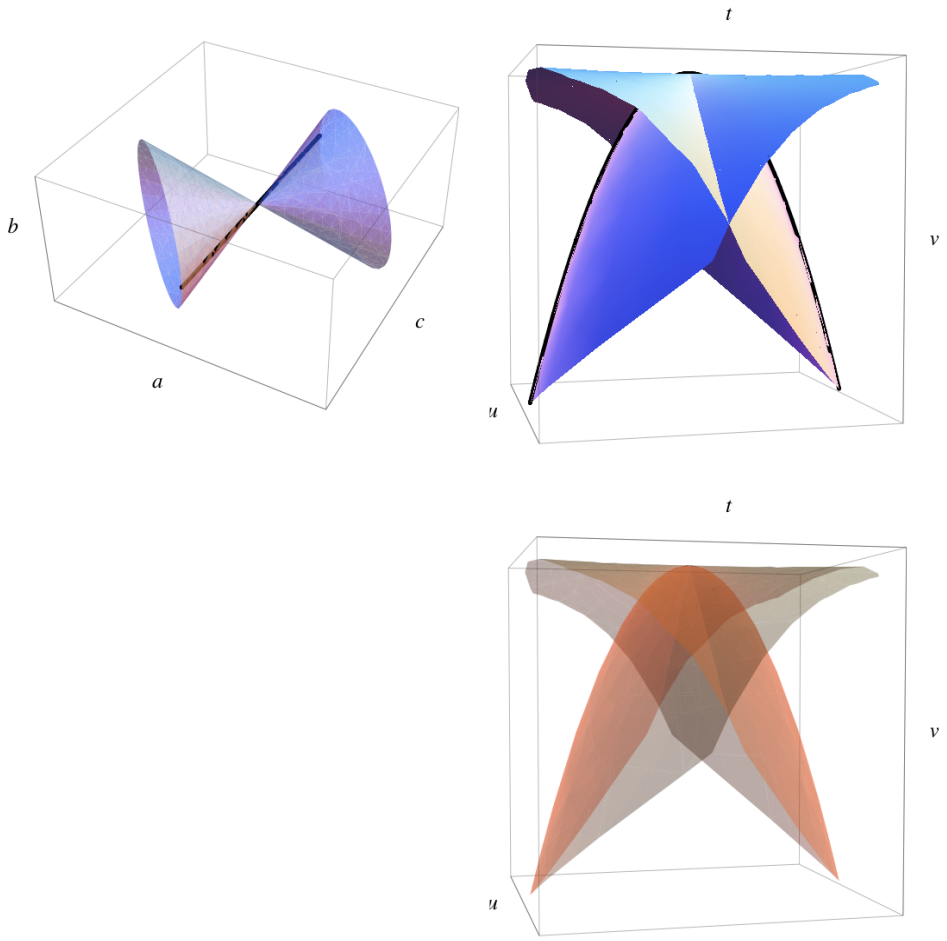


Figure 12: Stratification de l'ombilic hyperbolique. A gauche le cone C avec sa strate (génératrice) supplémentaire. En haut à droite sa transformée par l'application χ. On a représenté en bas à droite une version transparente de la surface. D'après Zeeman [57].

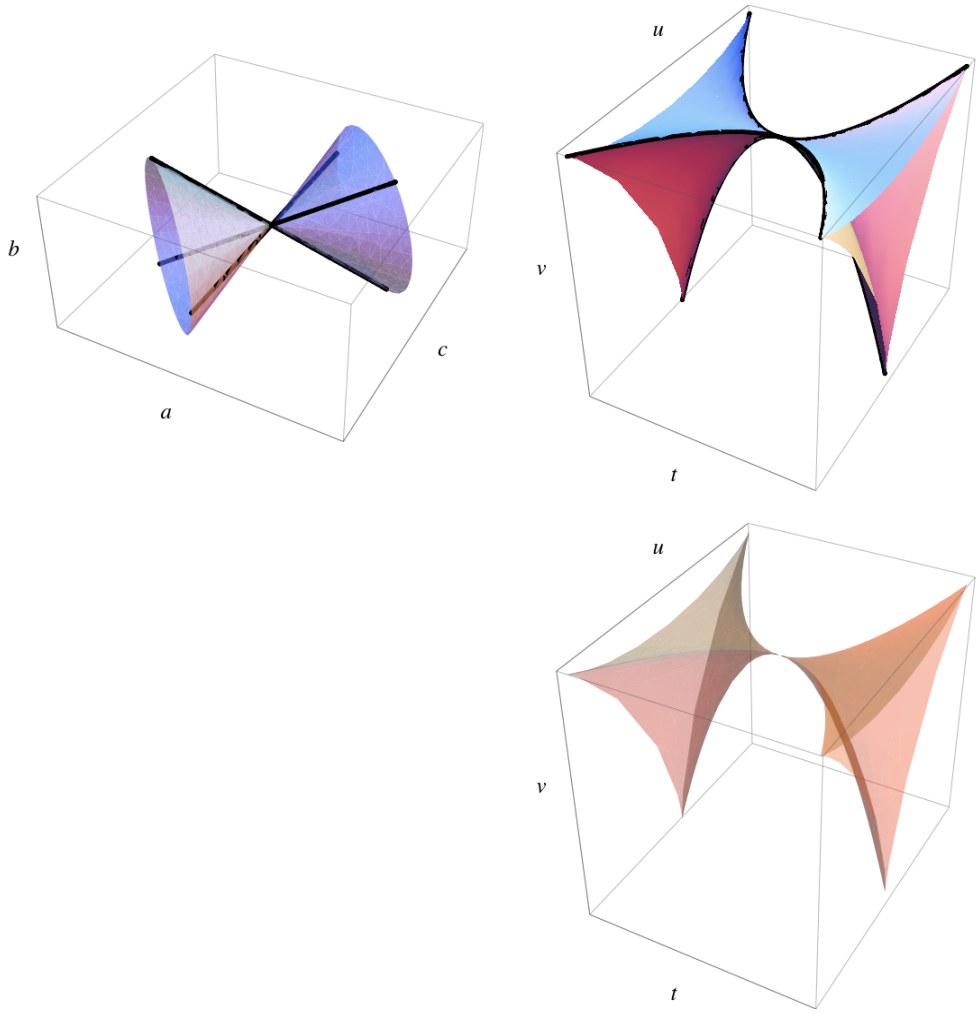


Figure 13: Stratification de l'ombilic elliptique. A gauche le cone C avec ses trois strates (génératrices) supplémentaires. En haut à droite sa transformée par l'application χ. On a représenté en bas à droite une version transparente de la surface. D'après Zeeman [57].

ce qui donne le difféomorphisme $\theta' : \mathbb{R}^2 \times \mathfrak{m}/\Delta \to \mathfrak{m}/\mathfrak{m}\Delta$ entre espaces de dimension 6. La sous-variété Σ_W de dimension 4 est d'équations

$$2xy + 2tx - u = 0 \text{ et } x^2 + y^3 + 2wy - v = 0 \tag{47}$$

dans le $\mathbb{R}^6\,(x, y, t, w, u, v)$ et elle est envoyée difféomorphiquement par θ sur $\mathfrak{m}^2/\mathfrak{m}\Delta = \mathbb{R}^3\,(a, b, c, d)$ par les équations

$$\begin{cases} a = y + t \\ b = x \\ c = \frac{1}{2}\,(3y^2 + 2w) \\ d = y \end{cases} \tag{48}$$

et l'application $\chi : \mathfrak{m}^2/\mathfrak{m}\Delta\,(a, b, c, d) \to \mathfrak{m}/\Delta\,(t, w, u, v)$ est donc donnée par les équations :

$$\begin{cases} t = a - d \\ u = 2ab \\ v = b^2 + 2cd - 2d^3 \\ w = c - \frac{3}{2}d^2 \end{cases} \tag{49}$$

La stratification canonique S^* de $\mathfrak{m}^2/\mathfrak{m}\Delta$ est l'image réciproque de la stratification canonique $S_{2,3}$ de $\mathfrak{m}^2/\mathfrak{m}^3$ par la projection $\pi : \mathfrak{m}^2/\mathfrak{m}\Delta \to m^2/\mathfrak{m}^3$. Comme $\mathfrak{m}^2/\mathfrak{m}\Delta = m^2/\mathfrak{m}^3 \oplus \mathbb{R}\eta^3$, elle est donc donnée par $C \times \mathbb{R}$. Il existe 3 strates raffinées par χ : $C_+ \times \mathbb{R}$, $C_- \times \mathbb{R}$ et $\{0\} \times \mathbb{R}$ (où C_+ et C_- sont respectivement les demi-cônes des plis x^2 et $-x^2$).

Le jacobien de χ est la matrice 4×4

$$J = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 2b & 2a & 0 & 0 \\ 0 & 2b & 2d & 2c - 6d^2 \\ 0 & 0 & 1 & -3d \end{pmatrix}$$

dont le déterminant est $4\,(ac - b^2)$. Le produit $C \times \mathbb{R}$ est donc bien le lieu singulier de χ.

Pour trouver le raffinement de $C \times \mathbb{R}$, il faut calculer le lieu singulier de la restriction $\chi\,|_{C\times\mathbb{R}}$. Le calcul aboutit aux raffinements suivants.

1. La strate $\{0\} \times \mathbb{R}$ est raffinée en 3 strates :
   (a) $\{0\} \times \mathbb{R}_+$ correspondant à $d > 0$ et composée d'ombilics hyperboliques;
   (b) $\{0\} \times \mathbb{R}_-$ correspondant à $d < 0$ et composée d'ombilics elliptiques;
   (c) $\{0\} \times \{0\}$ correspondant à $d = 0$ et est la strate de f.
2. La strate $C_- \times \mathbb{R}$ est raffinée en 2 strates :

(a) une strate de plis (ouverte, *i.e.* de dimension 3) correspondant à $ac = b^2 \neq -ad^2$ et formée de deux composantes;

(b) une strate de cusps de dimension 2 qui en est la frontière, correspond à $ac = b^2 = -ad^2$, $a^2 \neq 4c$, est formée de deux composantes et est adjacente aux strates ombilicales $\{0\} \times \mathbb{R}$;

3. La strate $C_+ \times \mathbb{R}$ est raffinée en 3 strates :

(a) une strate de plis comme $C_- \times \mathbb{R}$;

(b) une strate de cusps comme $C_- \times \mathbb{R}$ à ceci près que l'une des composante est formée de cusps duaux et que la seconde se dédouble en une composante de cusps duaux et une composante de cusps duaux à travers :

(c) une strate de queues d'aronde A_4 de dimension 1 correspondant à $ac = b^2 = -ad^2$, $a^2 = 4c$, $a \neq 0$ (cette strate ne peut pas apparaître dans $C_- \times \mathbb{R}$ car on y a $a^2 = 4ac = b^2 > 0$ et donc $a > 0$).

On obtient ainsi une stratification à 12 strates de χ. Pour plus de précisions, on se référera à l'article déjà cité [56] de Christopher Zeeman. Et pour une étude détaillée de l'ensemble catastrophique induit par χ (ou χ_W) dans W (ensemble encore plus compliqué que Strat (χ)), on pourra consulter le travail exhaustif de Godwin [13].

12.3 La géométrie des cuspoïdes

12.3.1 Le cusp

Il faut bien se convaincre du progrès considérable qu'apporte le concept de catastrophe par rapport aux conceptions taxinomiques plus anciennes où l'on cherchait seulement à énumérer les diverses façons qu'il y avait de stabiliser par petites déformations un potentiel instable.

Considérons un élément instable $f \in K_{\mathcal{F}}$ de codimension finie. D'après ce qui précède, il comprend un nombre fini de points critiques non dégénérés ou dégénérés de détermination finie, certaines des valeurs critiques associées pouvant être égales. Il est donc naturel de penser qu'on peut lui associer un "degré" d'instabilité qui est la somme des "degrés" de l'instabilité causée par chaque point critique dégénéré et chaque égalité de valeur critique. En fait, ce " degré" est précisément mesuré par la codimension c de f, c'est-à-dire par une dimension d'espace. Considérons alors un déploiement universel W de f. Il exprime toutes les façons qui existent de stabiliser progressivement f par petites déformations. Il est naturel de penser que ces stabilisations peuvent se faire par étapes successives en faisant décroître à chaque pas le degré d'instabilité d'une unité. Autrement dit, il est naturel de penser que l'on peut associer à f un "graphe" décrivant les types instables intermédiaires entre

Figure 14: Un point critique de type x^4, ses trois stabilisés, ses trois stabilisés partiels intermédiaires (avec égalité de valeurs critiques ou point d'inflexion).

f et tous les stabilisés possibles de f par petites déformations. Considérons par exemple un point critique dégénéré de corang 1 qui est un minimum où la tangente (horizontale) coupe le graphe de f non pas en 2 points confondus comme pour un minimum quadratique, ni même en 3 points confondus comme pour un point d'inflexion, mais en 4 points confondus. Le développement de Taylor de f en un tel point pris pour origine, commence par un terme en x^4. La singularité x^4 est déterminée à l'ordre 4. f est donc différentiablement équivalente à x^4 en un tel point. Par stabilisation, un tel point critique dégénéré "explose" en points critiques quadratiques. Pour connaître leur nombre il suffit de remarquer que les points critiques d'une fonction $f : \mathbb{R} \to \mathbb{R}$ sont les points où la dérivée $f'(x)$ s'annule, que, si $f(x) = x^n$, la dérivée est nx^{n-1} et que, par stabilisation, l'équation $nx^{n-1} = 0$ devient une équation générale de degré $n-1$ admettant donc $n-1$ racines distinctes soit réelles, soit imaginaires (les racines imaginaires allant toujours par paires de racines conjuguées). Pour $n = 4$, on obtient une équation du 3e degré admettant soit une racine réelle, soit 3 racines réelles. Cela permet d'énumérer immédiatement les trois types de stabilisés de f ainsi que les stabilisés partiels intermédiaires (*cf.* figures 14 et 15).

Il est ainsi trivial d'énumérer les stabilisés partiels ou complets de f par petites déformations. Mais il s'agit de passer du concept d'énumération à celui de classification et c'est ici qu'intervient le concept clef de stratification qui tire les conséquences du fait que la codimension d'une singularité "mesure" en quelque sorte son "degré" d'instabilité. Si l'on considère une singularité de codimension n, il est naturel de penser que dans un déploiement universel (donc de dimension n) de f

(i) les stabilisés de f constituent des cellules de dimension n (puisque "stable" = "codimension 0" = "dimension n");

(ii) les stabilisés partiels de degré 1 permettant de transformer ces stabilisés les uns dans les autres constituent des interfaces de dimension $n-1$ (puisque "degré 1" = "codimension 1" = "dimension $n-1$") séparant ces cellules;

(iii) les stabilisés partiels de degré 2 permettant de transformer les uns dans les autres les stabilisés partiels de degré 1 constituent des "arêtes" de dimension $n-2$ recollant deux interfaces de dimension $n-1$, etc.

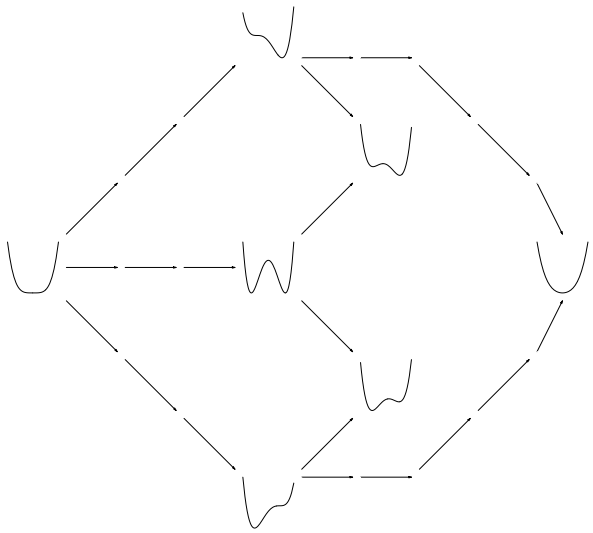


Figure 15: Le “graphe” (orienté) décrivant les chemins de stabilisation possibles.

Cette idée fondamentale est bien illustrée par la catastrophe cusp $f_{u,v}(x) = x^4/4 + ux^2/2 + vx$ déployant la singularité x^4 dans un voisinage W de l'origine du plan (u, v).

Pour représenter le graphe total V de ce déploiement donné par l'équation $f_{u,v}(x) - y = 0$, il nous faudrait disposer d'un espace quadridimensionnel (x, u, v, y). Nous nous bornerons donc à considérer (comme dans le cas du pli) le lieu critique Σ et son contour apparent sur W suivant l'application catastrophique $\chi : \Sigma \to W$ ainsi que le graphe G des valeurs critiques. G est le contour apparent $\pi(C)$ de $V \subset \mathbb{R}^4$ sur le $\mathbb{R}^3(u, v, y)$ suivant la projection $\pi : (x, u, v, y) \to (u, v, y)$ et Σ est la projection de C sur le $\mathbb{R}^3(x, u, v)$.

Dans le $\mathbb{R}^3(x, u, v)$, Σ est par définition l'ensemble des points critiques de $f_{u,v}$ et est donc la surface d'équation $f'_{u,v}(x) = x^3 + ux + v = 0$. On remarquera que cette équation devient celle du graphe total de la singularité *pli* x^3 si l'on change v en $-y$. De façon générale on peut ainsi identifier le lieu critique de x^n au graphe total de x^{n-1}. Σ est donc une surface “fronce” d'où le nom de singularité *fronce* donnée à x^4.

Pour construire l'ensemble catastrophique K de W ainsi que sa stratification, il suffit de remarquer que les stabilisés de $f = x^4$ sont classés en fonction du nombre de racines de l'équation dérivée $f'_{u,v}(x) = 0$ définissant Σ. Cette équation étant du 3e degré, si u et v sont des valeurs générales, elle aura soit une racine réelle, soit trois racines réelles distinctes. Les stabilisés partiels de f de type “point d'inflexion + minimum” correspondent donc aux valeurs exceptionnelles (catastrophiques) de (u, v) pour lesquelles $f'_{u,v}(x) = 0$ admet une racine double (*i.e.* pour lesquelles on a aussi $f''_{u,v}(x) = 0$). On sait que

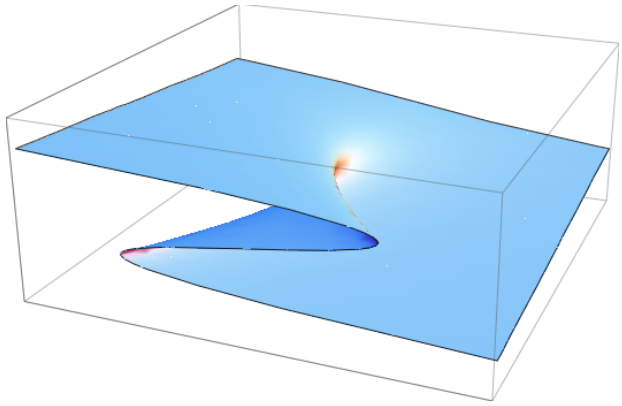


Figure 16: La surface fronce de la catastrophe cusp.

ces valeurs sont données par le discriminant de l'équation $f'_{u,v}(x) = 0$ (*i.e.* par $4u^3 + 27v^2 = 0$). Il est facile de vérifier cette affirmation. Si l'on a à la fois $x^3 + ux + v = 0$ et $3x^2 + u = 0$, alors $u = -3x^2, v = 2x^3$ et donc $4u^3 + 27v^2 = 0$.

Nous obtenons ainsi une première partie K_b de l'ensemble catastrophique K correspondant lorsqu'on le traverse aux catastrophes de bifurcation. K_b est une parabole semi-cubique présentant un rebroussement (*i.e.* un cusp) en $(u = 0, v = 0)$. D'où le nom de catastrophe *cusp* donné au déploiement de x^4. K_b est le contour apparent de Σ sur W relativement à la projection $\chi : \Sigma \to W$. Il est composé de trois strates, d'une part des deux branches du cusp épointées de l'origine (strates de codimension 1) et d'autre part de l'origine (strate de codimension 2 correspondant au centre organisateur $f = x^4$). Quant à K il comprend une strate supplémentaire de conflit K_c correspondant à l'égalité des valeurs des deux minima pouvant "émerger" de x^4 (*cf.* figures 16 et 17).

Considérons alors un chemin γ dans W parallèle à l'axe des v et de $u < 0$. Il est facile de vérifier que le graphe des valeurs critiques suivant ce chemin est de type "queue d'aronde". A l'extérieur du cusp, il ne comprend qu'une seule branche γ_1 (valeur du minimum). A la première traversée de K_b il présente un "point de naissance" d'où émergent deux nouvelles branches γ_2 et γ_3. A la traversée de K_c les branches γ_1 et γ_2 s'intersectent transversalement (croisement des valeurs critiques lors d'une catastrophe de conflit). A la seconde traversée de K_b, γ_1 et γ_3 disparaissent en un "point de mort", la branche γ_2 demeurant seule à l'extérieur du cusp. Lorsque u varie, ce graphe se déforme et, lorsque u s'annule, il devient un graphe à une seule branche.[58]

De même que le graphe des valeurs critiques du pli x^3 dans le plan (u, y) s'identifie à l'ensemble catastrophique K_b du cusp x^4 dans le plan (u, v) par la transformation $v \to -y$, de même le graphe des valeurs critiques de x^4 s'identifie au K_b de la singularité x^5 que nous allons maintenant étudier (*cf.*

[58] γ s'appelle le facteur normal du cusp et u le facteur de splitting (Zeeman).

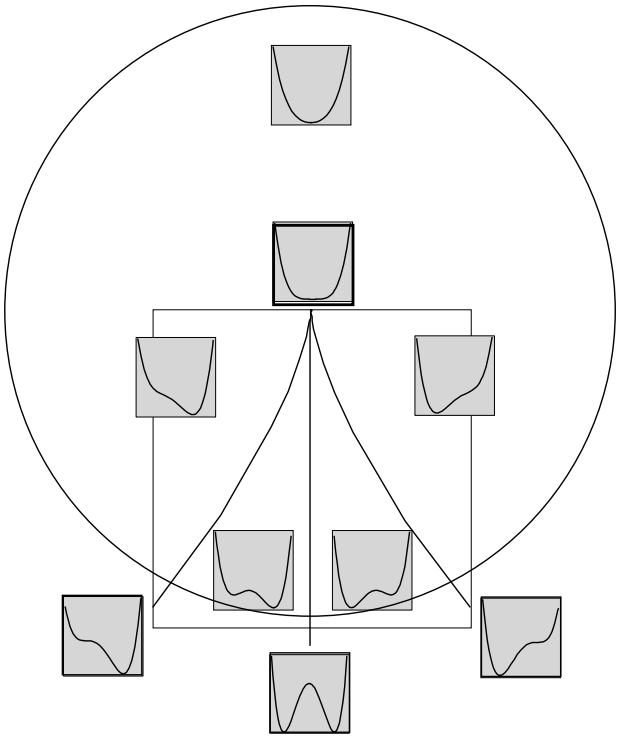


Figure 17: Déploiement universel et stratification de la singularité cusp x^4.

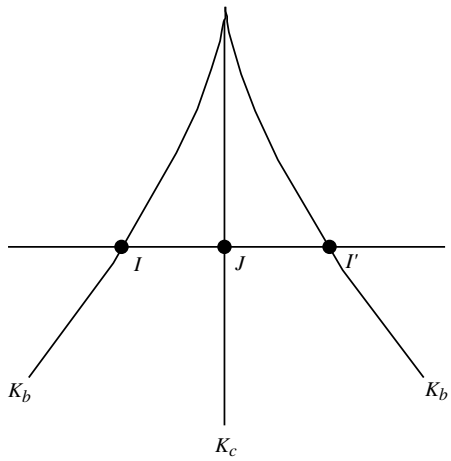


Figure 18: Chemin transverse à l'ensemble catastrophique K dans l'espace externe du cusp.

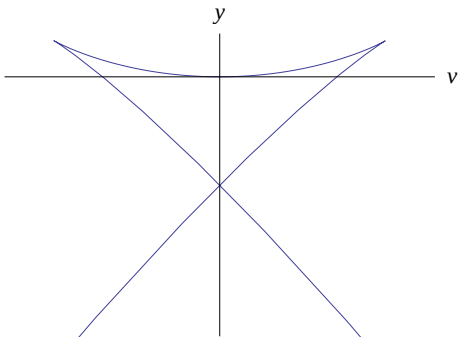


Figure 19: Le graphe des valeurs critiques le long du chemin de la figure 18.

figures 18, 19, 20).

12.3.2 La queue d'aronde

Si l'on considère le cas un peu plus complexe $f(x) = x^5$ (point d'inflexion dégénéré) il est facile d'énumérer les stabilisés partiels et les stabilisés de f et de construire le "graphe" associé. Par petites déformations x^5 peut "exploser" soit en 4 points critiques quadratiques, deux minima et deux maxima, soit en 2 points critiques quadratiques (un minimum et un maximum), soit en 0 point critique. Ces points critiques quadratiques peuvent "émerger" de la singularité x^5 (où ils " collapsent") et la déployer avec tous les rapports possibles entre leurs valeurs critiques. S'il y a 0 point critique, il n'y a qu'un seul type différentiable possible. Il en va de même s'il y a 2 points critiques. S'il y a en revanche 4 points critiques alors il y a 5 types différentiables possibles que l'on peut déformer les uns dans les autres à condition de traverser des types

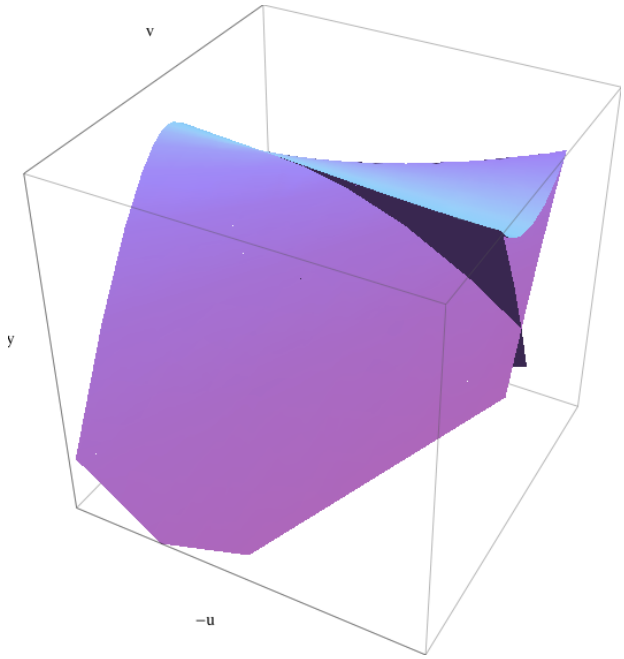


Figure 20: La surface “queue d’aronde” des valeurs critiques du cusp.

instables de degré 1 de l’espèce “égalité de valeurs critiques” (*cf.* figure 21).

Pour passer d’une des classes (a), (b), (c) de la figure 21 à l’autre, il faut faire naître ou mourir des paires de points critiques et donc traverser des types instables de degré 1 de l’espèce “point d’inflexion”. Six cas sont possibles (*cf.* figure 22).

Maintenant, pour passer d’un type instable de degré 1 à un autre, il faut traverser des types instables de degré 2 et il est trivial de vérifier que ceux-ci sont au nombre de six (*cf.* figure 23).

Il est ainsi très facile d’énumérer tous les stabilisés de $f(x) = x^5$ par petites déformations. Si l’on essaie de construire le “graphe génératif” des “chemins” de stabilisation, on s’aperçoit aussitôt que c’est un graphe qui est très loin d’être une arborescence, des stabilisés partiels différents de degré 2 (resp. 1) pouvant engendrer des stabilisés partiels de degré 1 (resp. 0) identiques (*cf.* figure 24).

Pour passer à la classification opérée par le déploiement universel (W, K) de x^5, il nous faut faire au préalable quelques remarques sur les “mélanges” d’instabilités de type “bifurcation” et de type “conflit”.

Le cas le plus simple est celui où l’instabilité de f est due à des singularités *indépendantes* de codimension 1. Cela signifie

(i) que f ne présente que des points d’inflexion simples et/ou des égalités de valeurs pour deux points critiques quadratiques et

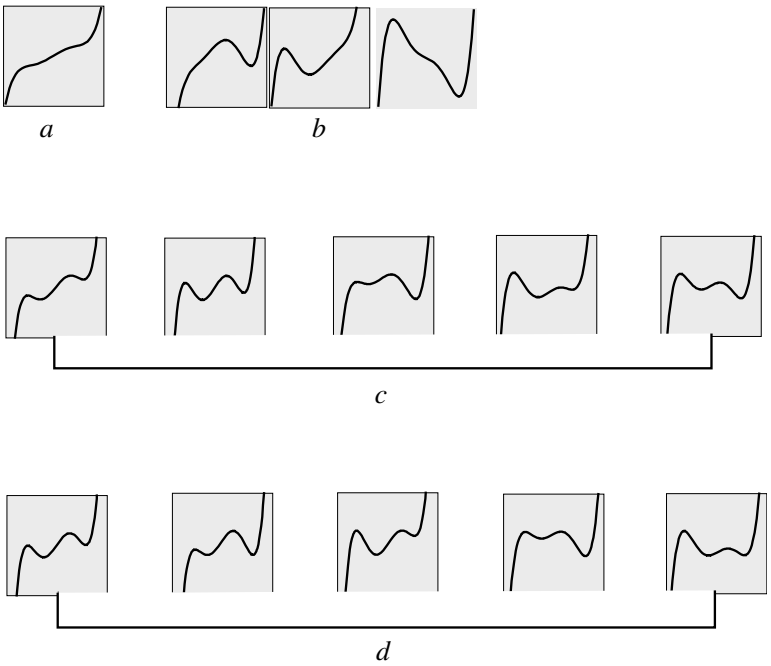


Figure 21: (a), (b), (c) Stabilisés de la singularité x^5. (d) Types instables de degré 1 permettant de transformer l'un dans l'autre les 5 types stables de la classe (c).

Figure 22: Les six types instables de l'espèce "point d'inflexion" permettant de passer d'une des classes (a), (b) ou (c) de la figure 21 à une autre.

Figure 23: Les types instables de degré 2 permettant de transformer l'un dans l'autre les types instables de degré 1 des figures 21 et 22.

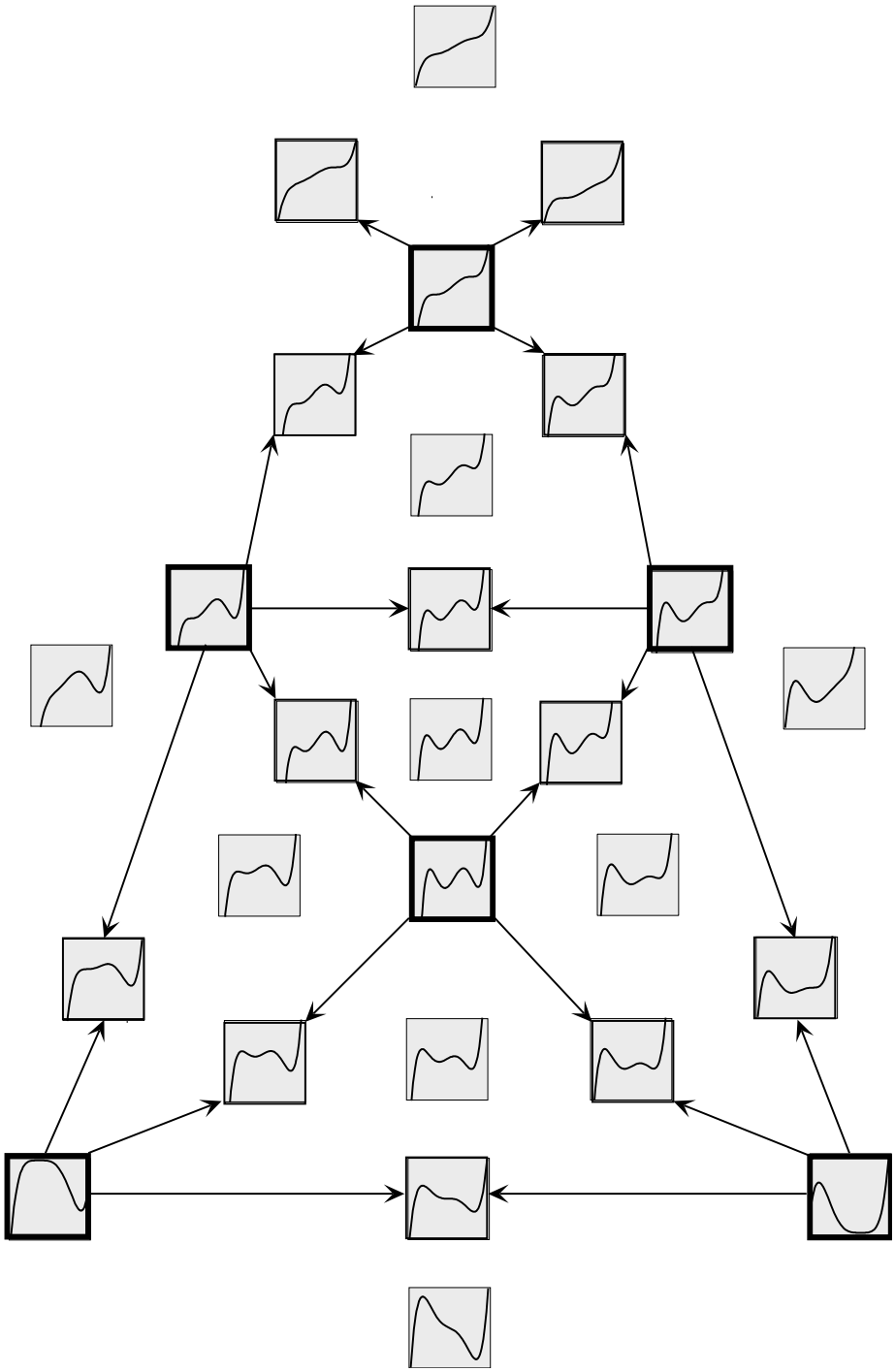


Figure 24: Graphe de stabilisation de $f(x) = x^5$.

(ii) que ces instabilités élémentaires sont toutes de valeurs critiques différentes.

Si f est de codimension n, alors l'ensemble catastrophique K de W sera constitué par n hypersurfaces H_i (de dimension $n-1$) passant par f et s'intersectant transversalement. Autrement dit, au voisinage de l'origine 0 de $W \simeq \mathbb{R}^n$, K sera isomorphe au "cloisonnement" de $\mathbb{R}^n$ par les hyperplans de coordonnées. Dans ce cas, la stratification de (W, K) est particulièrement évidente. L'instabilité de f est due à n instabilités indépendantes de codimension $1, S_1, \ldots, S_n$. Chacune de ces instabilités peut se stabiliser de deux façons, symbolisables par $+$ et $-$. Il y a donc 2^n stabilisés de f que l'on peut noter symboliquement $f_{\pm 1,\ldots,\pm n}$, $f_{+1,-2,+3,\ldots,+n}$ signifiant par exemple que S_1 s'est stabilisée "côté $+$", S_2 "côté $-$", S_3 "côté $+$ ", etc. L'hypersurface H_i est alors une hypersurface constituée de stabilisés partiels "finaux" de f (*i.e.* de stabilisés partiels de codimension 1 "précédant" immédiatement les stabilisés). Sur H_i toutes les singularités sont stabilisées sauf une, à savoir S_i. Si on assimile H_i à un hyperplan d'équation $\varphi_i = 0$ où φ_i est une forme linéaire sur $\mathbb{R}^n$, les deux demi-espaces séparés par H_i sont définis par les inégalités $\varphi_i > 0$ et $\varphi_i < 0$. Autrement dit, chaque stabilisé de f, par exemple $f_{+1,-2,+3,\ldots,+n}$, correspond à un " quadrant" *ouvert* de $\mathbb{R}^n$ défini par un système complet d'inégalités strictes, par exemple $\varphi_1 > 0$, $\varphi_2 < 0$, $\varphi_3 > 0$, $\ldots$, $\varphi_n > 0$. Ces quadrants ouverts sont les strates (les "cellules", les " chambres") de codimension 0 (de dimension n) de la stratification.

Considérons alors les "arêtes" de ces chambres, c'est-à-dire les intersections $H_i \cap H_j$ entre 2 "murs" H_i et H_j. Les $H_i \cap H_j$ pour i fixé et $j \neq i$ variant de 1 à n, décomposent H_i en 2^{n-1} quadrants, ouverts dans H_i, qui sont les strates de codimension 1 (de dimension $n-1$) associées aux stabilisés partiels "finaux" de types $f_{\pm 1,\ldots,\pm(i-1),0,\pm(i+1),\ldots,\pm n}$ (le 0 en position i signifiant que la singularité S_i n'a pas été déployée). De même, pour $i \neq j$ fixés et $k \neq i$, j variant de 1 à n, les $H_i \cap H_j \cap H_k$ décomposent $H_i \cap H_j$ en 2^{n-2} quadrants, ouverts dans $H_i \cap H_j$, qui sont les strates de codimension 2 (de dimension $n-2$) associées aux stabilisés partiels de type $f_{\pm 1,\ldots,\pm(i-1),0,\pm(i+1),\ldots,\pm(j-1),0,\pm(j+1),\ldots,\pm n}$.

Remarquons que dans ce cas élémentaire la stratification "archétype" est indépendante de la nature des singularités et ne dépend que de leur nombre. Dans la figure 25 on trouvera un exemple de stratification en codimension 2 déployant deux singularités indépendantes de type conflit. Elle intervient dans la queue d'aronde.

Il existe trois possibilités pour une fonction f de codimension n de ne pas correspondre à la combinatoire simple que nous venons d'exposer : soit elle inclut des singularités non indépendantes, soit elle inclut des singularités de codimension supérieure à 1, soit elle inclut les deux.

Un exemple typique de non indépendance est fourni par la singularité dite singularité *bec* comprenant un point d'inflexion simple (une catastrophe pli) dont la valeur critique est égale à la valeur d'un point critique quadratique.

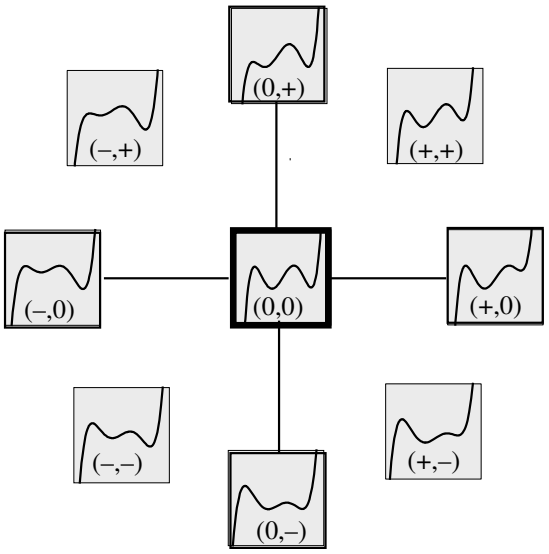


Figure 25: Déploiement de deux singularités indépendantes.

Soit f une telle singularité. En se déployant, le point d'inflexion S_1 peut soit disparaître, soit "exploser" en un minimum m_1 et un maximum M_1 et l'on peut obtenir plusieurs types stables suivant les rapports de hauteur entre m_1, M_1 et le point quadratique M, ainsi que les stabilisés partiels $f(m_1) = f(M)$ et $f(M_1) = f(M)$. Quant aux autres stabilisés partiels ils sont clairement de type pli (et donc de codimension 1) et correspondent à la levée de la dégénérescence $f(S_1) = f(M)$. La singularité bec est de codimension 2 et la stratification de son déploiement universel est représentée à la figure 26.

Revenons au déploiement et à la stratification de la singularité x^5 organisant la catastrophe élémentaire dite "queue d'aronde". Cette singularité indécomposable est de codimension 3 et son déploiement universel est donné par $x^5 + ux^3 + vx^2 + wx = f_{u,v,w}(x)$ (ou par $x^5/5 + ux^3/3 + vx^2/2 + wx$ si on veut normaliser). Nous connaissons maintenant les modèles transverses de tous ses stabilisés partiels de codimension 2 et 1. Si nous revenons aux figures 21, 22 et 23 nous voyons:

(i) qu'en codimension 2, nous trouvons (figure 23) (a) deux plis indépendants, (b) une singularité bec, (c) une autre singularité bec, (d) deux conflits indépendants, (e) un cusp, (f) un cusp "dual" (où c'est un maximum qui est de type cusp et non un minimum);

(ii) qu'en codimension 1 nous trouvons (figure 21(d) et figure 22) (a) 5 types de conflits, (b) 6 types de plis.

L'ensemble catastrophique $K \subset W$ de $f = x^5$ est la surface de W représentée à la figure 27.

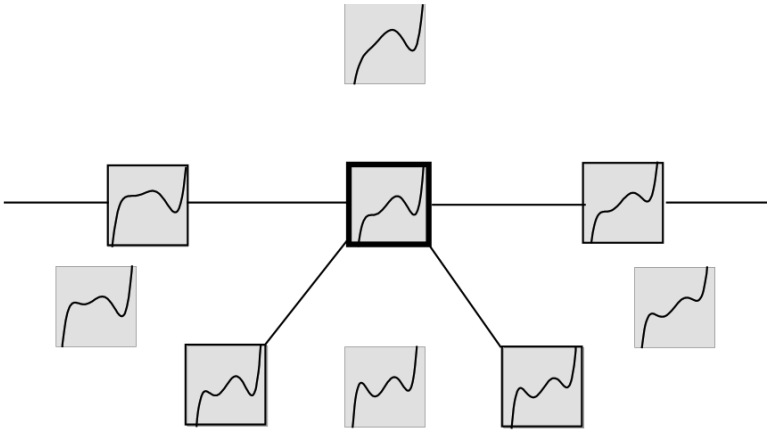


Figure 26: Stratification du déploiement universel d'une singularité bec.

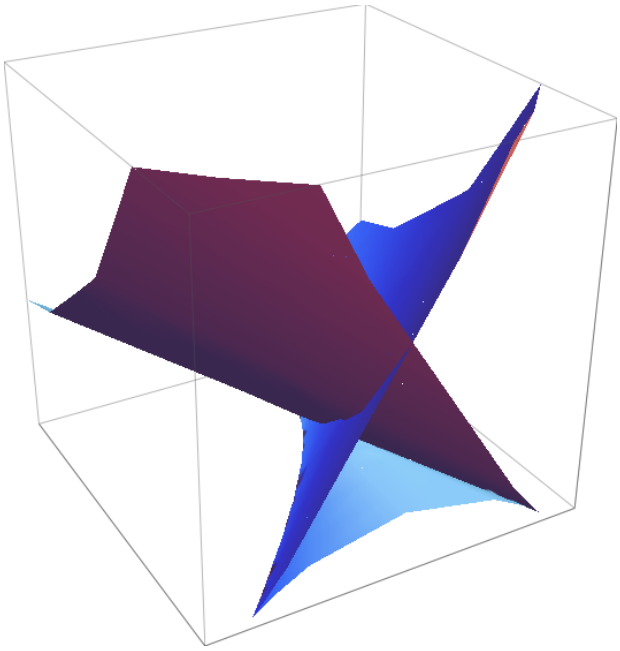


Figure 27: L'ensemble catastrophique $K \subset W$ de $f = x^5$ représenté comme la surface paramétrique de W d'équations $(-u, v = -4x^3 - 2ux, w = 3x^4 + ux^2)$, pour $u \in [-1, 1]$ et $x \in [-1.3, 1.3]$.

Pour montrer la façon dont cet ensemble catastrophique réalise géométriquement la classification des stabilisés partiels et complets de f, nous avons représenté figure 28 une section de dimension 2 de W transverse à K dans sa partie la plus complexe.

On voit se manifester clairement sur cet exemple le fait que la stratification de (W, K) est une stratification naturelle. En effet:

(i) les strates de codimension 0 (les cellules) sont les composantes connexes du complémentaire de K dans W;

(ii) les strates de codimension 1 sont les composantes connexes du complémentaire dans K du lieu singulier $\mathrm{Sing}(K)$ de K. Ce lieu singulier est constitué de lignes de self-intersection, d'arêtes de rebroussement et de lignes "becs";

(iii) les strates de codimension 2 sont les composantes connexes dans $\mathrm{Sing}(K)$ du lieu singulier $\mathrm{Sing}\,(\mathrm{Sing}(K)) = \{f\}$ de $\mathrm{Sing}(K)$: les lignes de self-intersection, les arêtes de rebroussement et les lignes de points "bec" de K admettent en f (origine de W) un point de rebroussement.

Autrement dit, K a une géométrie relativement typique. C'est un "empilement" de lieux singuliers de dimension croissante et sa stratification consiste à itérer l'opération Sing et à prendre les composantes connexes de $(\mathrm{Sing})^n - (\mathrm{Sing})^{n+1}$. Il est remarquable que cette stratification de K dérivée de sa structure géométrique coïncide (au voisinage de f) avec la trace sur W de la décomposition de $\mathcal{F}$ en G-orbites (où $G = \mathrm{Diff}\ \mathbb{R} \times \mathrm{Diff}\ \mathbb{R}$ est le groupe définissant les types différentiables). En général, la situation est plus complexe, la K-classification des stabilisés partiels ou complets de f assurée par la stratification de K étant strictement moins fine que la G-classification (problème des "modules"). Mais dans les cas élémentaires comme ceux présentés ici, la K-classification s'identifie à la G-classification et c'est en ce sens que l'on peut dire que la stratification de K, qui, insistons-y, est de nature purement géométrique, réalise géométriquement la classification des formes engendrées par le déploiement d'un centre organisateur.

On a représenté à la figure 29 la catastrophe queue d'aronde $f_{u,v,w}(x) = x^5/5 + ux^3/3 + vx^2/2 + wx$ pour $u = -1$. C'est la surface d'équation $f'_{u,v,w}(x) = x^4 - x^3 + vx + w = 0$ froncée au-dessus du plan de contrôle (v, w) avec $x \in [-1.5, 1.5]$, $v \in [-1, 1]$ et $w \in [-2, 1]$. Pour rendre sa structure plus visible nous en montrons les 4 "tranches" $v \in [-1, -0.5]$, $v \in [-1, 0]$, $v \in [-1, 0.5]$ et enfin $v \in [-1, 1]$. On voit clairement comment se déploie la queue d'aronde avec ses plis et ses cusps : initialement il n'y a qu'un pli, puis un cusp apparaît dont l'une des branches va progressivement devenir le pli terminal alors que l'autre branche capture le pli initial et disparaît avec lui en un second cusp.

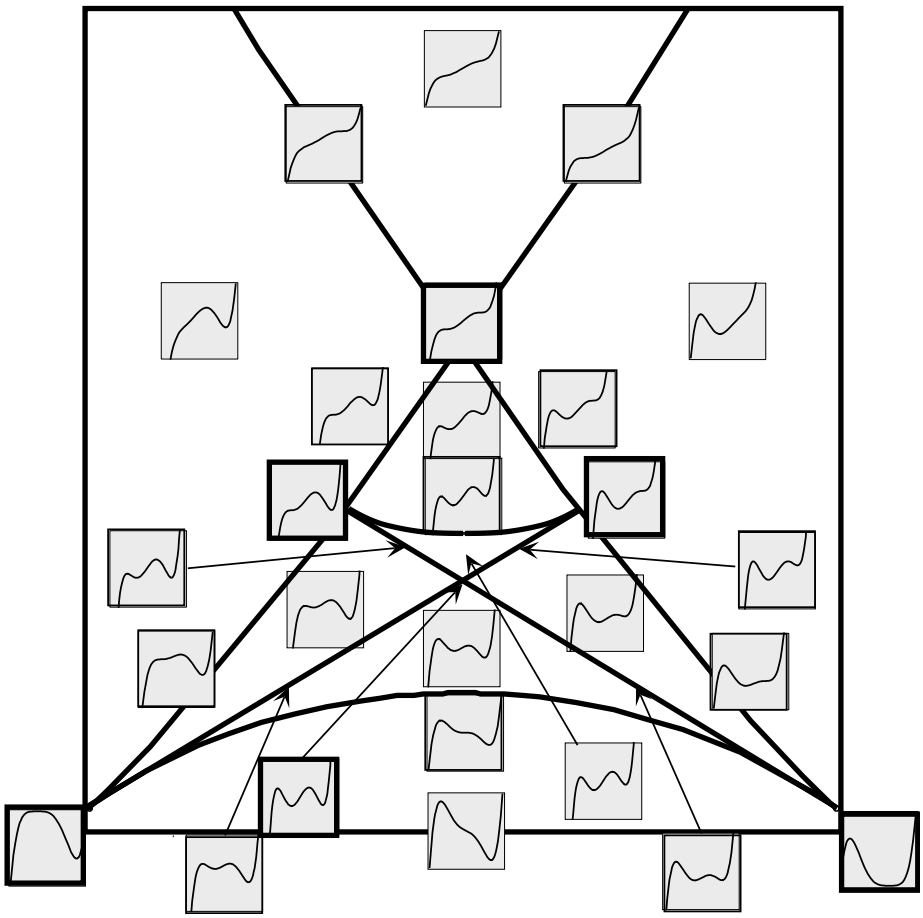


Figure 28: Section de dimension 2 de l'espace externe W de la queue d'aronde qui est transverse à l'ensemble catastrophique K dans sa partie la plus complexe

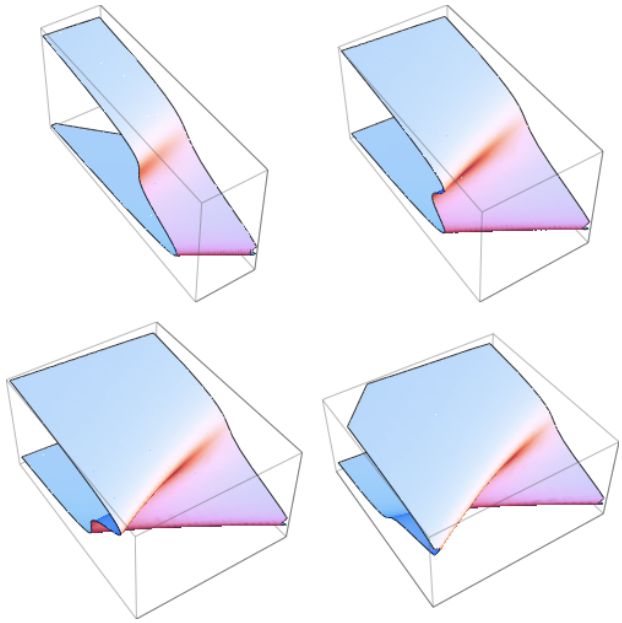
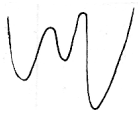


Figure 29: Les fronces de la queue d'aronde. Les différentes perspectives sont scalées automatiquement et ne sont donc pas à la même échelle.

12.3.3 Le papillon

La fonction $f = x^6$ présente à l'origine un minimum dégénéré obtenu par collapse de 5 points critiques quadratiques, 3 minima et 2 maxima. Il est donc facile d'énumérer les types qualitatifs intervenant dans son déploiement universel. Ce sont tous les types que l'on peut obtenir à partir d'une fonction de type  en faisant varier la hauteur des maxima et des minima et en les collapsant de toutes les façons possibles. Pourtant, comme nous allons le voir, la géométrie discriminant ces divers types est déjà très complexe.

Nous prendrons comme déploiement universel de $f_0 = f$ le déploiement normalisé :

$$f_\tau = \frac{x^6}{6} + t\frac{x^4}{4} + u\frac{x^3}{3} + v\frac{x^2}{2} + wx$$

où $\tau = (t, u, v, w)$ parcourt un voisinage W de l'origine de $\mathbb{R}^4$. Soit K_b l'ensemble catastrophique associé aux bifurcations de f_τ. K_b est le contour apparent sur W de l'application catastrophique $\chi : \Sigma \to W$.

Les strates de la stratification de K_b correspondent aux multiplicités possibles des racines de l'équation $f'_\tau = 0$ définissant Σ, c'est-à-dire aux diverses décompositions possibles de 5 en sommes d'entiers. Ces décompositions α étant $(1, 1, 1, 1, 1)$, $(1, 1, 1, 2)$, $(1, 1, 3)$ et $(1, 2, 2)$, $(2, 3)$ et $(1, 4)$, (5), l'hypersurface

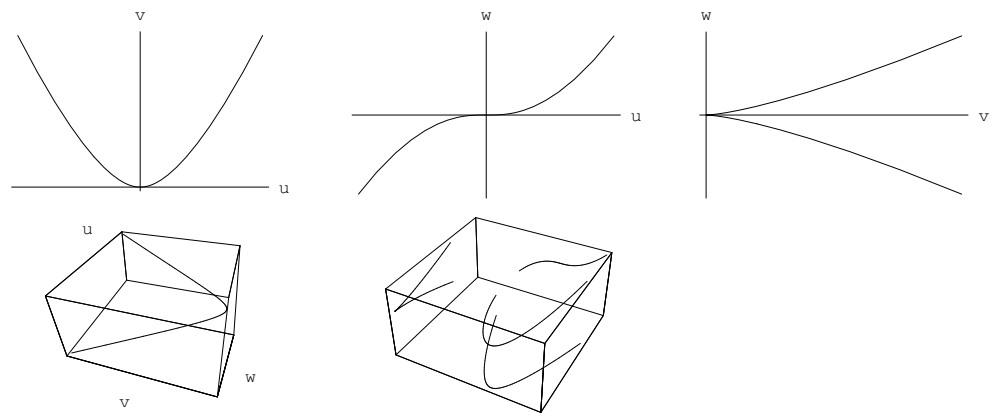


Figure 30: La section A_t de la surface de rebroussemment A pour $t > 0$. Nous représentons la courbe gauche A_t pour $t = 1$. Nous montrons les 3 projections sur les plans (u, v), (u, w) et (v, w), la courbe gauche et enfin la courbe munie de ses 3 projections.

K_b admet comme lieu singulier une surface de $\mathbb{R}^4$ qui recolle deux strates de codimension 2, une de points cusp ($\alpha = (1, 1, 3)$) et une de self-intersection ($\alpha = (1, 2, 2)$), suivant une courbe gauche admettant l'origine comme point de rebroussement ($\alpha = (5)$) et composée d'une strate de points queue d'aronde ($\alpha = (1, 4)$) et d'une strate de points bec ($\alpha = (2, 3)$).

Pour visualiser K_b, nous décrivons d'abord la surface A, fermeture de la strate des points cusp, en suivant ses sections A_t par l'hyperplan $t =$ cste. A_t est une courbe gauche dans l'espace (u, v, w). La surface A est le lieu des points τ pour lesquels f'_τ admet une racine triple. Ses équations sont donc :

$$\begin{cases} f'_\tau(x) = x^5 + tx^3 + ux^2 + vx + w = 0 \\ f''_\tau(x) = 5x^4 + 3tx^2 + 2ux + v = 0 \\ f'''_\tau(x) = 20x^3 + 6tx + 2u = 0 \end{cases}$$

d'où la représentation paramétrique:

$$u = -10x^3 - 3tx,\ v = 15x^4 + 3tx^2,\ w = -6x^5 - tx^3. \tag{50}$$

Pour $t > 0$, la section A_t est une courbe gauche régulière dont les projections sur les plans de coordonnées sont données à la figure 30. Elle possède l'allure d'une sorte de parabole "ventilée" dans la dimension w.

Pour $t = 0$, la section A_0 d'équations $u = -10x^3$, $v = 15x^4$, $w = -6x^5$ admet à l'origine un point de courbure infinie. Ses projections sur les plans de coordonnées sont données à la figure 31.

Pour $t < 0$, la section A_t admet des points singuliers pour les valeurs $x = \pm\sqrt{-\frac{t}{10}}$ du paramètre. En effet, d'après (50), $\frac{dv}{du} = -2x\frac{R}{R}$, $\frac{dw}{dv} = -\frac{x}{2}\frac{R}{R}$,

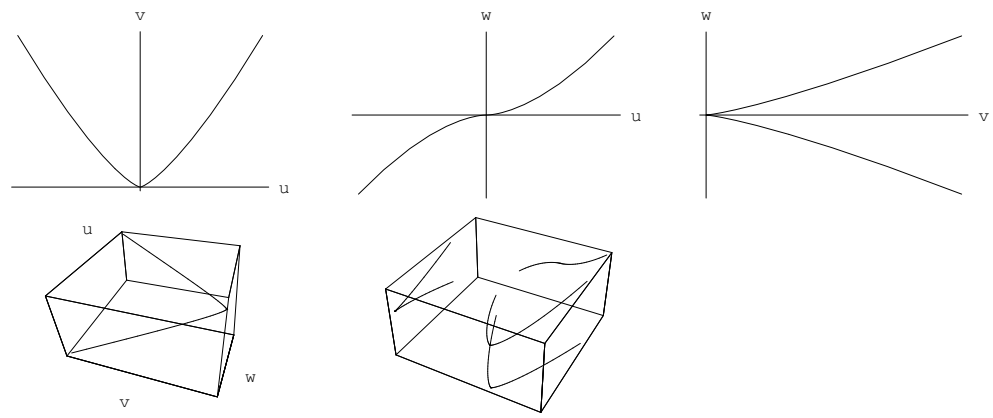


Figure 31: La section A_t de la surface de rebroussemment A pour $t = 0$.

$\dfrac{dw}{du} = x^2\dfrac{R}{R}$ où $R = 10x^2 + t$. Il s'agit de deux rebroussements Γ_1 et Γ_2 (points queue d'aronde) et A_t a l'allure d'une section K_u (pour $u < 0$) de la queue d'aronde, "ventilée" dans la dimension w (*cf.* figure 32).

Pour décrire l'hypersurface $K_b = \Delta$ admettant A comme surface de rebroussement, nous suivons ses sections Δ_t pour t variable et pour décrire Δ_t nous suivons ses sections $\Delta_{t,u}$ pour u variable. $\Delta_{t,u}$ est une courbe du plan (v, w).

Δ est définie par les équations :

$$\begin{cases} f'_\tau(x) = x^5 + tx^3 + ux^2 + vx + w = 0 \\ f''_\tau(x) = 5x^4 + 3tx^2 + 2ux + v = 0 \end{cases}$$

et $\Delta_{t,u}$ par les équations paramétriques :

$$\begin{cases} v = -5x^4 - 3tx^2 - 2ux \\ w = 4x^5 + 2tx^3 + ux^2 \end{cases}$$

Comme $\dfrac{dw}{dv} = -x\dfrac{R}{R}$ où $R = 10x^3 + 3tx + u$, les points singuliers de $\Delta_{t,u}$ sont les racines de l'équation $R = 0$ de discriminant $\dfrac{2}{5}t^3 + u^2$. Pour $t > 0$, $\Delta_{t,u}$ n'admet donc qu'un point singulier, un rebroussement qui décrit, lorsque u varie, la parabole "ventilée" A_t analogue à celle de la figure 30. Pour $t = 0$, Δ_0 est une surface analogue mais singulière à l'origine (*cf.* figures 33 et 34).

Pour $t < 0$, l'équation $R = 0$ possède trois racines réelles pour $-\sqrt{-\dfrac{2t^3}{5}} < u < \sqrt{\dfrac{-2t^3}{5}}$. Sur cet intervalle, $\Delta_{t,u}$ admet trois rebroussements et en dehors un seul, le lieu de ces rebroussements étant la queue d'aronde "ventilée" A_t de la figure 32 (*cf.* figure 35).

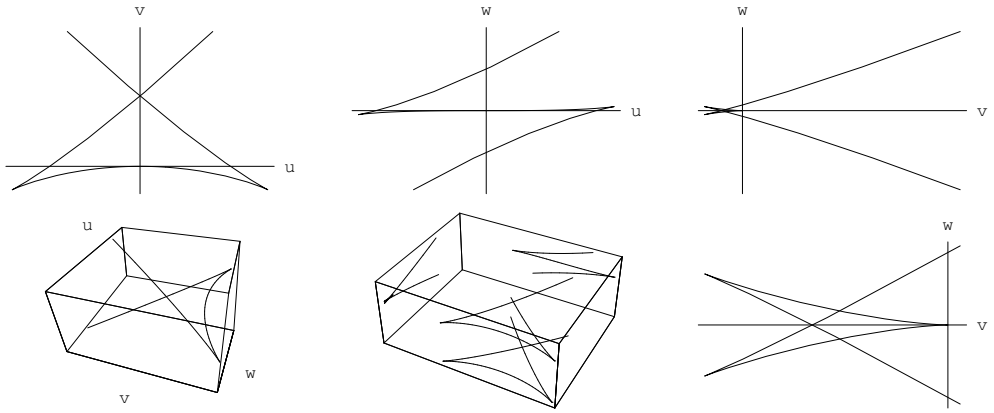


Figure 32: La section A_t de la surface de rebroussemment A pour $t < 0$. Nous avons zoomé à la dernière image sur la projection (v, w) de façon à rendre visible la structure de la singularité.

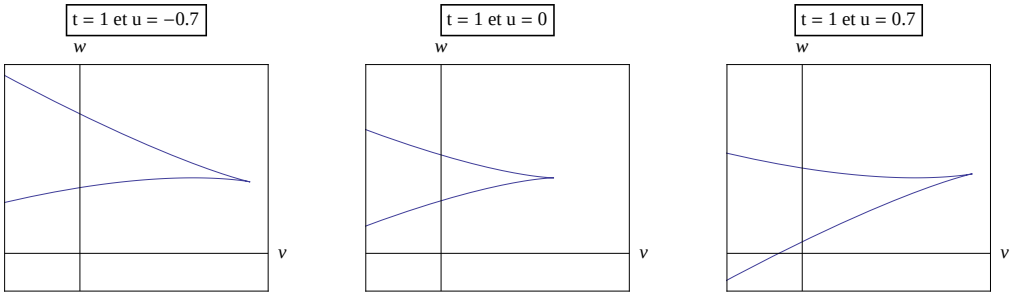


Figure 33: Sections $\Delta_{t,u}$ de Δ_t pour $t = 1$ et u variable.

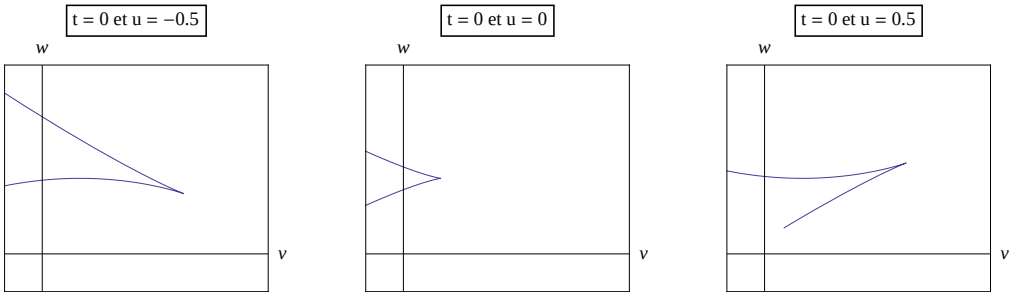


Figure 34: Sections $\Delta_{0,u}$ de Δ_0 pour $t = 0$ et u variable.

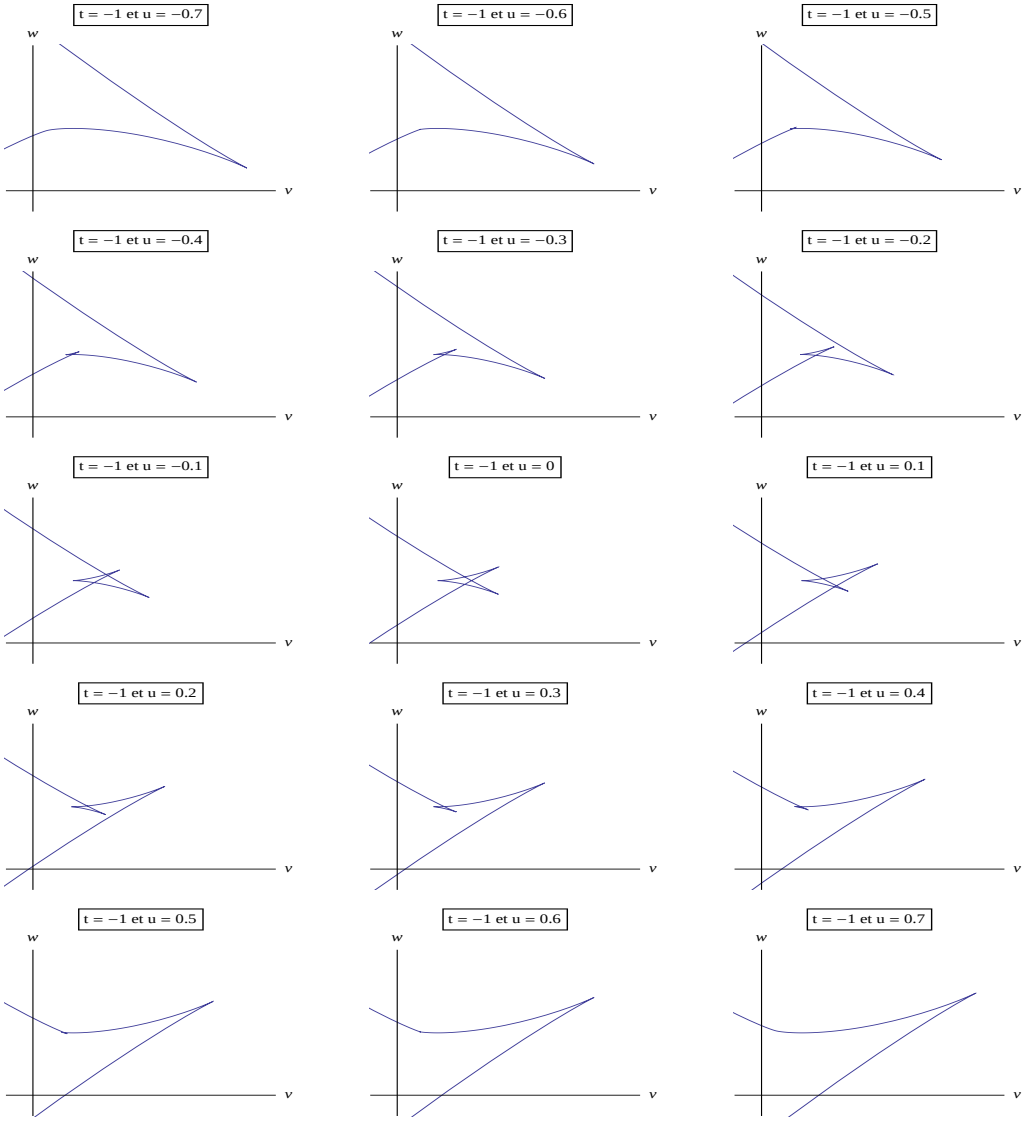


Figure 35: Sections $\Delta_{t,u}$ de Δ_t pour $t = 1$ et u variable. Les singularités queue d'aronde apparaîssent pour $u = \pm 0.6325$.

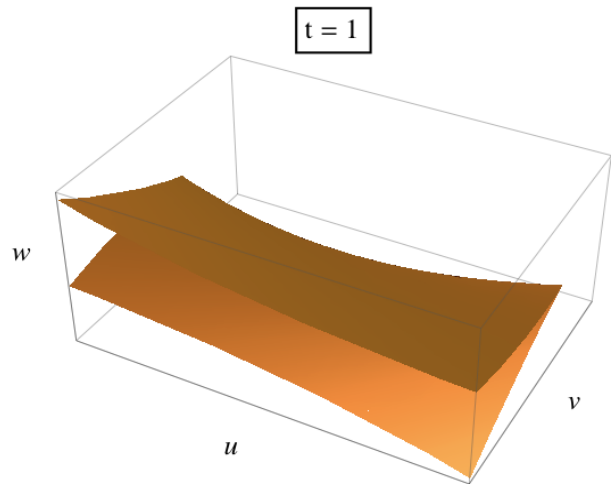
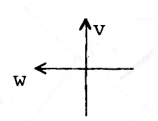


Figure 36: La surface Δ_t pour $t = 1$ avec $u \in [-0.7, 0]$, $v \in [-0.5, 0.5]$, $w \in [-0.3, 0.3]$.

Les points Γ_1 et Γ_2 sont les points queue d'aronde ($\alpha = (1, 4)$), ceux Q_1 et Q_2 les points bec où f_τ admet un cusp et un pli indépendants ($\alpha = (2, 3)$). Pour $|u| < \sqrt{\dfrac{-2t^3}{5}}$, $\Delta_{t,u}$ admet des points de self-intersection décrivant, lorsque u varie, la strate $\alpha = (1, 2, 2)$.

Pour $t > 0$, la surface Δ_t est comme un cylindre ayant pour section un cusp et pour génératrice l'axe des u. A la figure 36 nous représentons la surface Δ_t pour $t = 1$ avec $u \in [-0.7, 0]$, $v \in [-0.5, 0.5]$, $w \in [-0.3, 0.3]$.

Pour $t < 0$, la situation est plus compliquée. A la figure 37 nous représentons la surface Δ_t pour $t = -1$ d'abord avec $u \in [-0.7, 0]$, $v \in [-0.5, 1.3]$, $w \in [-0.8, 0.8]$, puis avec $u \in [-0.7, 0.7]$. Aux figures 38 et 39 nous représentons les surfaces fronces du papillon pour $t = -1$ et u variant de $u = -0.7$ à $u = 0.7$ avec un zoom sur un voisinage de $u = 0$ afin de mieux voir la structure des 3 cusps.

La connaissance de l'hypersurface K_b et des singularités correspondant à ses diverses strates permet alors de reconstruire par adjonction des strates de conflit l'ensemble catastrophique complet K. Nous traiterons une section K_t de K pour $t < 0$ en faisant varier u. Pour avoir plus de commodité dans les dessins, nous ferons varier u de $u = -\infty$ à $u = +\infty$ et nous orienterons les axes (v, w) de la façon suivante : 

- Pour $u < -\sqrt{-2t^3/5}$, nous sommes en présence d'un cusp (*cf.* figure 40).

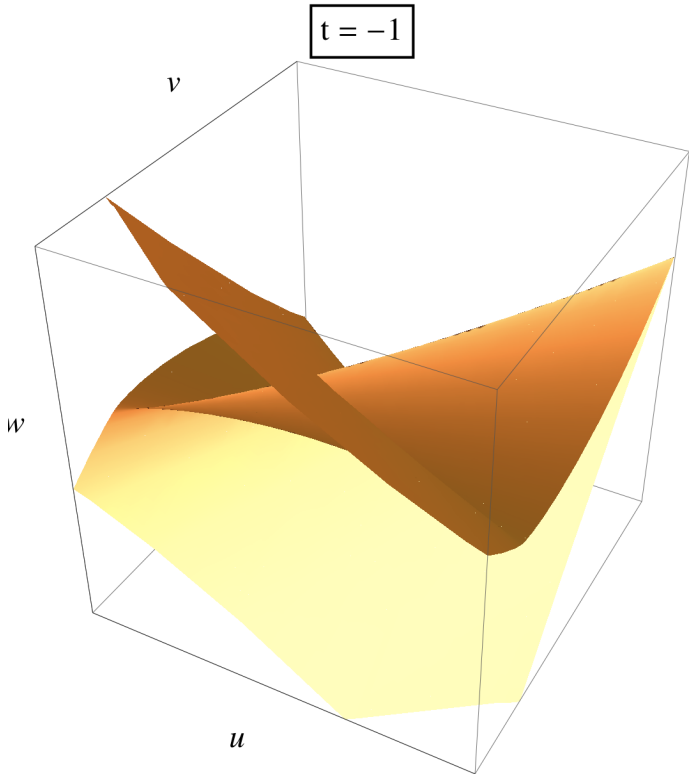


Figure 37: La surface Δ_t pour $t = -1$ avec $u \in [-0.7, 0.7]$, $v \in [-0.5, 1.3]$, $w \in [-0.8, 0.8]$.

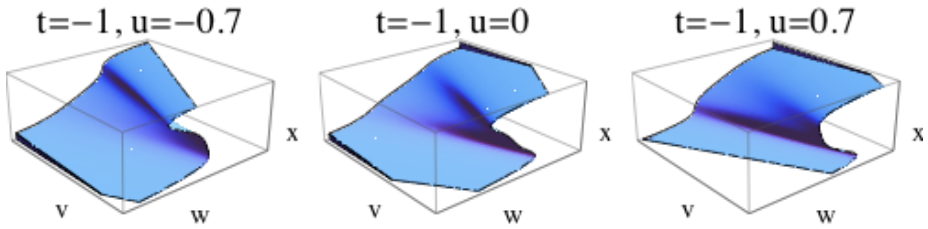


Figure 38: Les surfaces fronces du papillon pour $t = -1$ et u variant de $u = -0.7$ à $u = 0.7$.

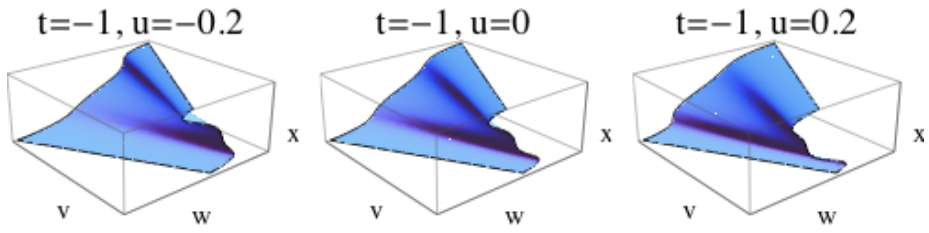


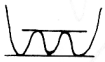
Figure 39: Zoom sur un voisinage de $u = 0$ dans la figure 38.

- Après l'apparition de la queue d'aronde Γ_2 sur la strate B_2', on obtient la section de la figure 41.
- Ensuite le cusp γ_1 traverse la strate de conflit initiale C_2 (cusp à un niveau critique :  ce qui fait naître un point triple T et deux nouveaux points bec β_2 et β_4 (*cf.* figure 42).
- Ensuite le cusp γ_1 traverse en Q_2 la seconde branche B_2 du cusp initial γ_2 ce qui déploie une singularité cusp-pli (singularités indépendantes)  (*cf.* figure 43).

A partir de cette configuration, on aboutit à la configuration symétrique pour $u > 0$ en franchissant le cas symétrique $u = 0$. Cette transition est représentée à la figure 44. Elle exige:

(i) que le cusp dual γ_3 traverse la strate de conflit C_2 : singularités indépendantes cusp-conflit  ;

(ii) que les points bec β_1 et β_2 traversent les strates de conflit et de bifurcation : singularités  et  ;

(iii) que le point triple T traverse la strate de conflit des maxima : singularité  (conflit maximal).

Les sections (a) et (b) de la figure 44 sont représentées aux figures 45 et 46. La transition entre les cas (a) et (c) de la figure 44 est quant à elle représentée à la figure 47.

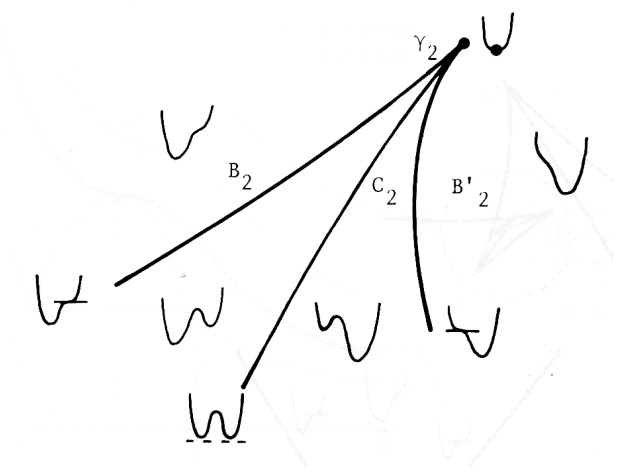


Figure 40: Cusp initial.

Pour préciser la structure du lieu critique Σ dont K est le contour apparent par l'application catastrophique $\chi : \Sigma \to W$, nous reprenons les sections $\Sigma_{t,u}$ correspondant à $K_{t,u}$ de la figure 38). Nous représentons les nappes de minima en traits pleins et celles des maxima en pointillés. Nous désignons respectivement par 1, 2, 3 et par a, b les minima et les maxima dans l'ordre suivant :

13 La dialectique local/global dans les déploiements

Partons du déploiement universel de la singularité "queue d'aronde" $f = x^5$ (figures 27 et 28). f étant de codimension 3 dans $\mathcal{F}$, elle reste de codimension 3 dans son déploiement W qui est de dimension 3. La trace sur W de sa G-orbite $\tilde{f}$ est donc de dimension 0 dans W. Autrement dit (puisqu'il n'y a pas dans ce cas de modules), la strate de f se réduit à $\{f\}$. Soit alors g un stabilisé partiel de f de codimension 2. La fonction g reste de codimension 2 dans W et sa strate $\overline{g}$ est donc de dimension $3 - 2 = 1$ dans W. Cette strate $\overline{g}$ est la trace sur W de la G-orbite $\tilde{g}$ de g dans $\mathcal{F}$: $\overline{g} = \tilde{g} \cap W$. On peut, en remplaçant $\mathcal{F}$ par W et $\tilde{g}$ par $\overline{g}$, appliquer *dans* W (et non plus dans $\mathcal{F}$) la procédure de construction des modèles transverses. Dans un voisinage de g dans W on considère :

(i) une section W_g de dimension 2 transverse en g à $\overline{g}$;

(ii) l'intersection K_g de K avec W_g.

Il est clair que (W_g, K_g) est un déploiement universel de g, *i.e.* un modèle transverse de g dans $\mathcal{F}$. En effet, nous sommes dans la "bonne situation"

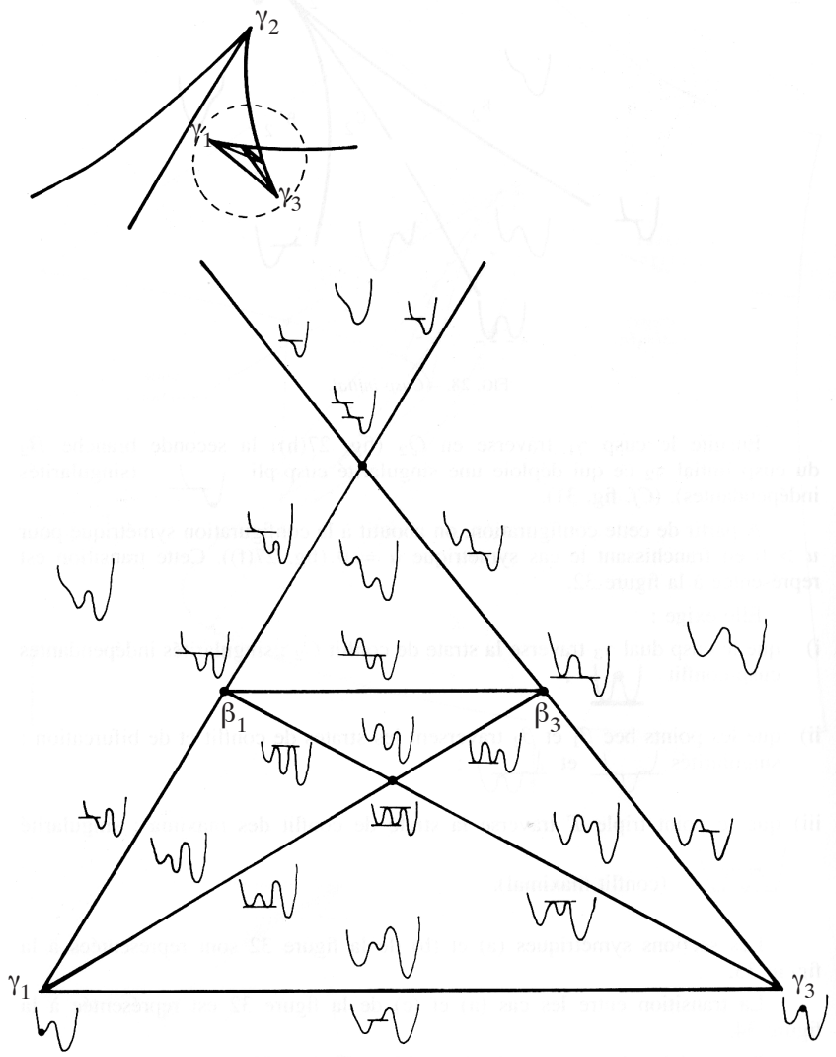


Figure 41: Section après la traversée de la queue d'aronde Γ_2. γ_1 est un cusp et γ_3 un cusp dual; β_1 et β_2 sont des points bec.

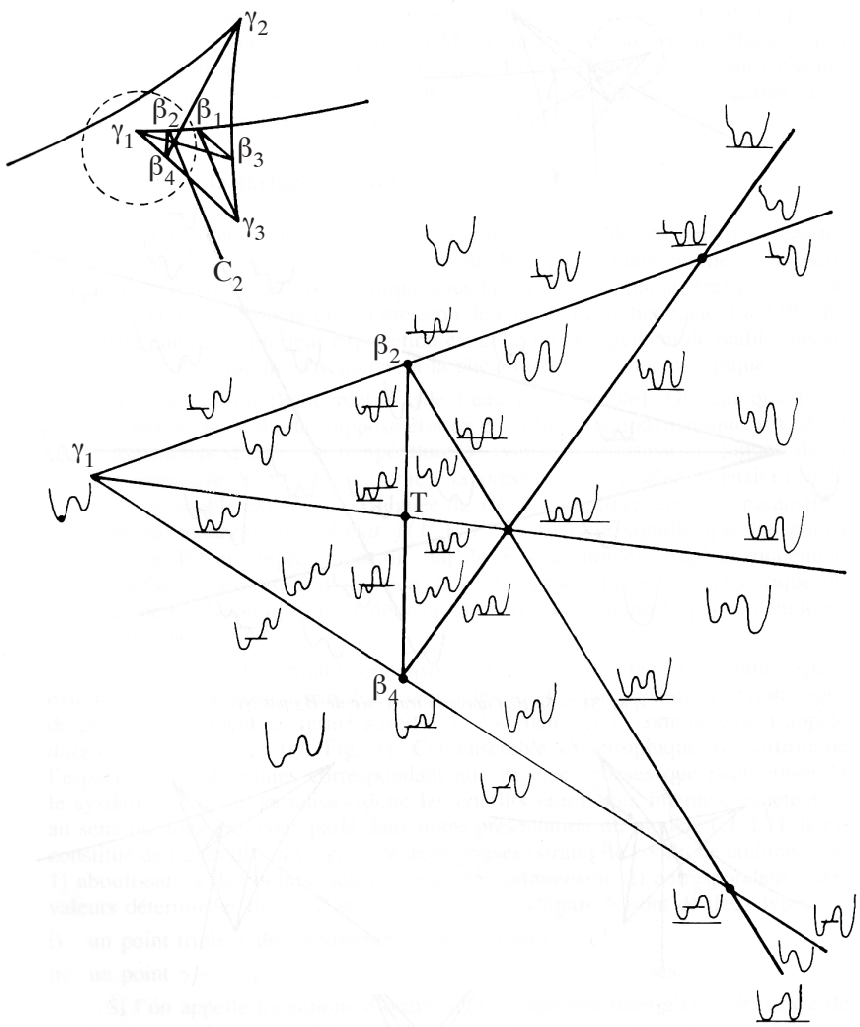


Figure 42: Section après la traversée de C_2 par γ_1.

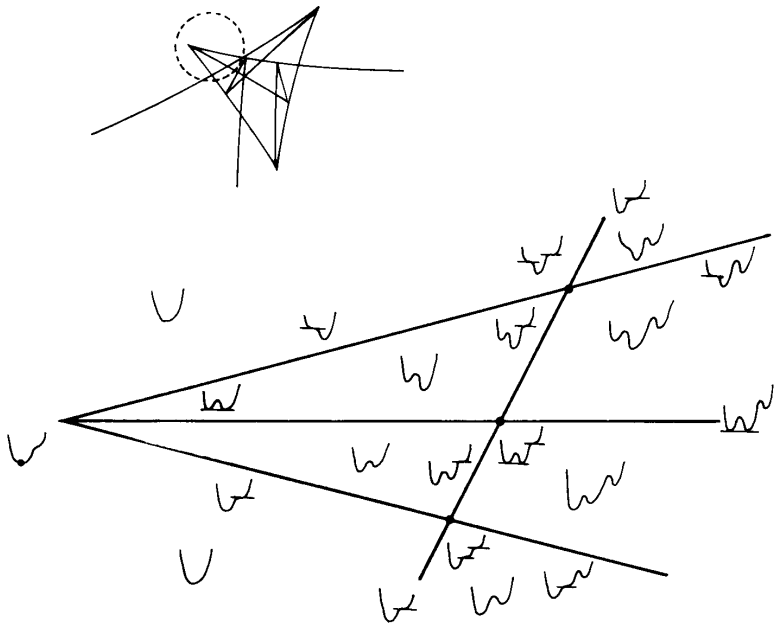


Figure 43: Section après la traversée de B_2 par γ_1.

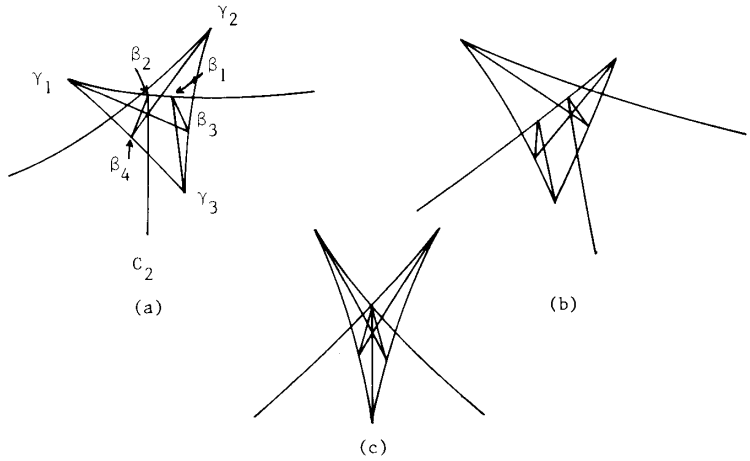


Figure 44: Transition par la section symétrique $u = 0$.

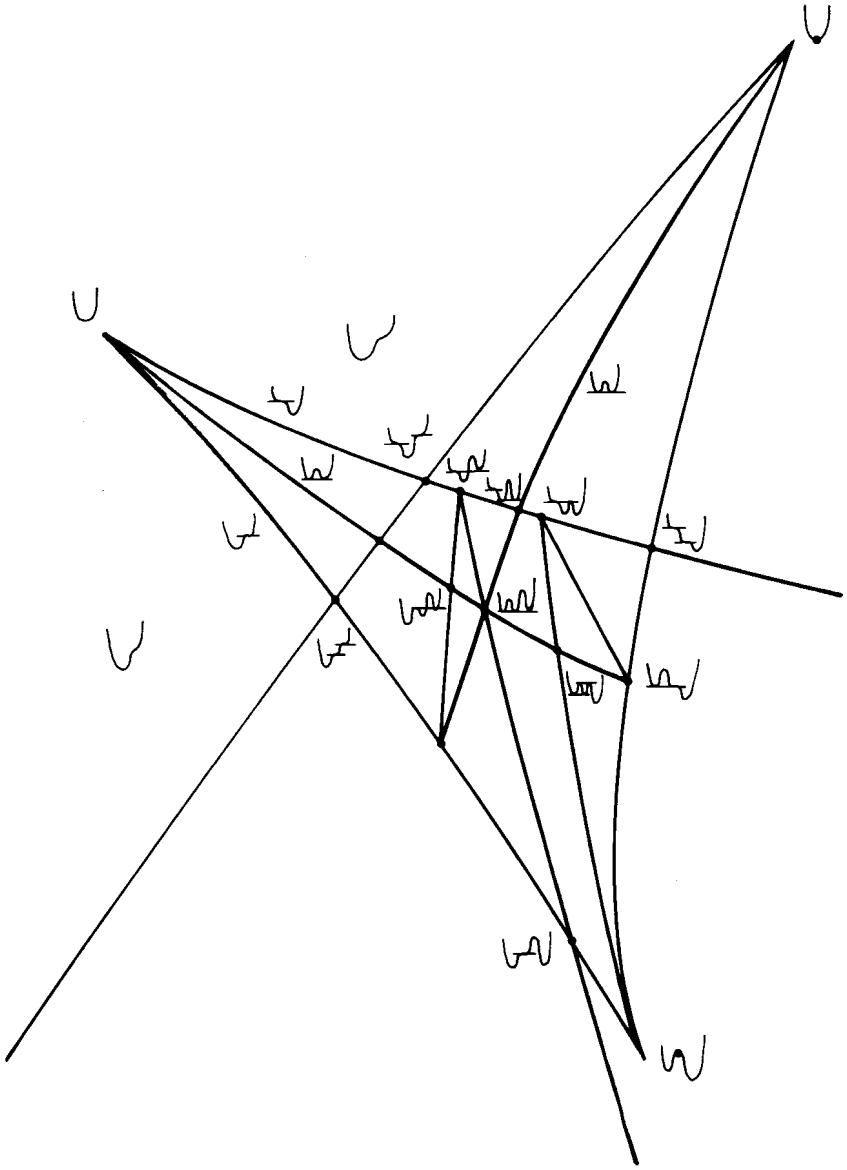


Figure 45: Section (a) de la figure 44.

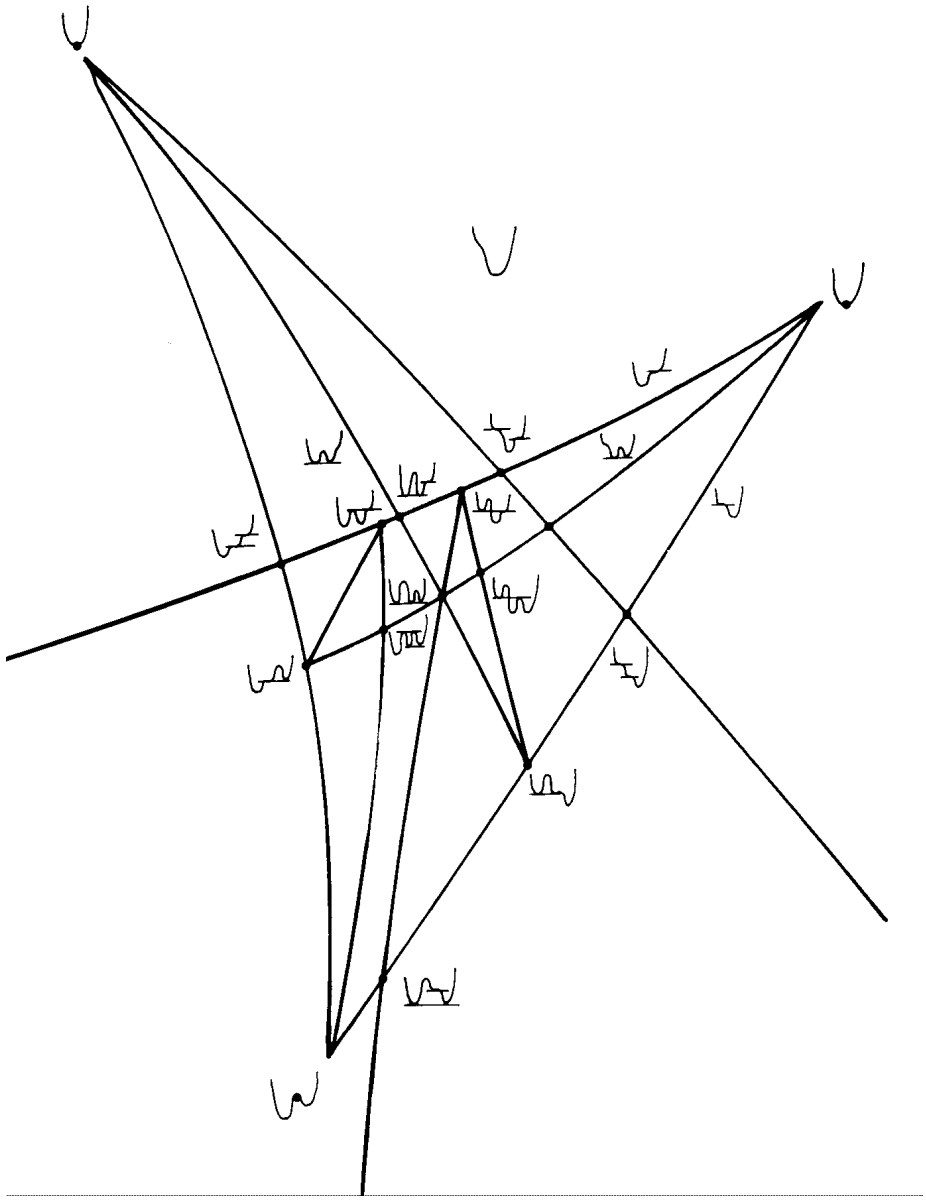


Figure 46: Section (b) de la figure 44.

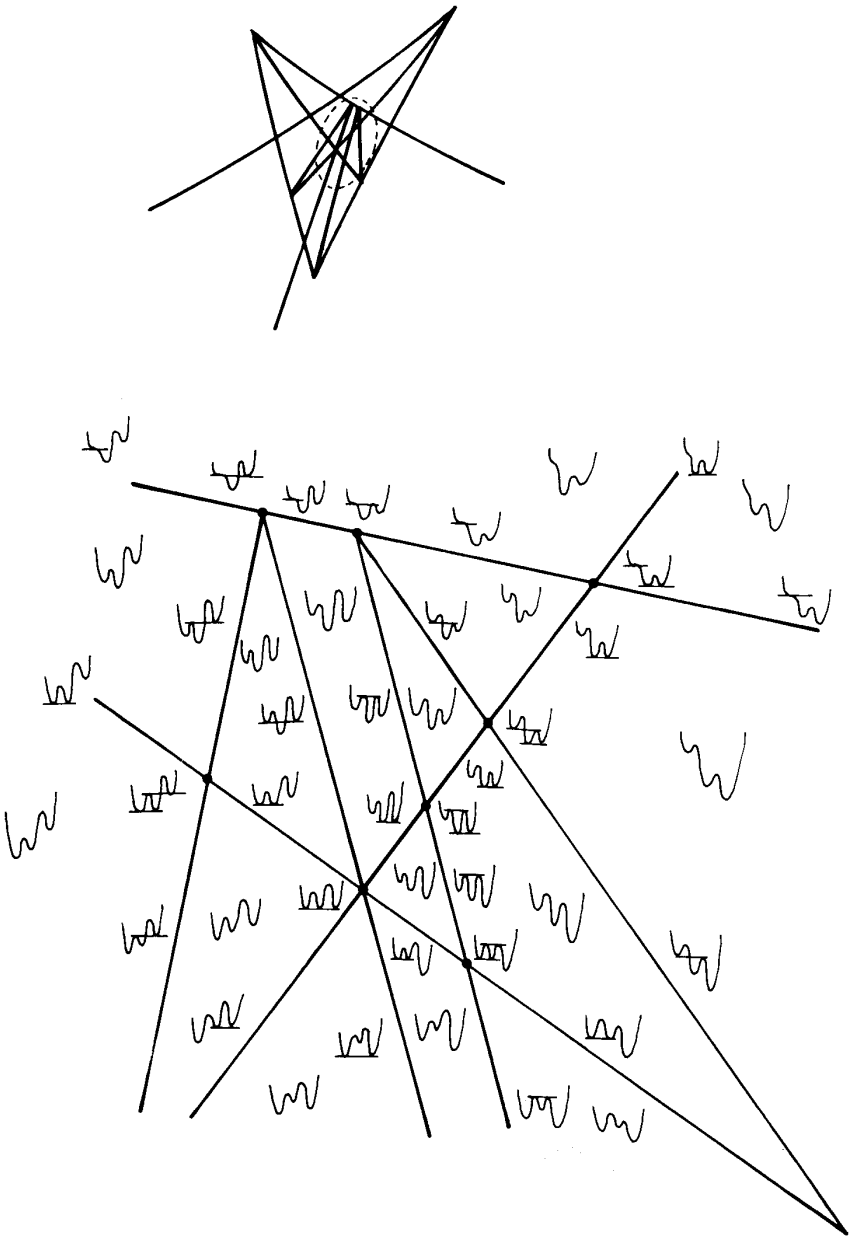


Figure 47: Transition entre les cas (a) et (c) de la figure 44.

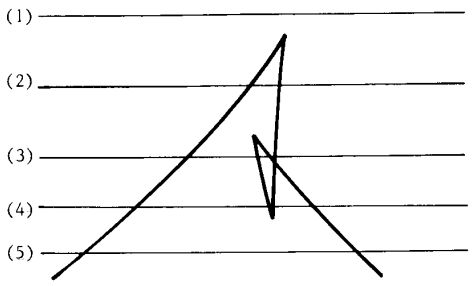


Figure 48: Cinq sections de $\Delta_{t,u}$ ($\Delta = K_b$) après l'apparition de la queue d'aronde Γ_2.

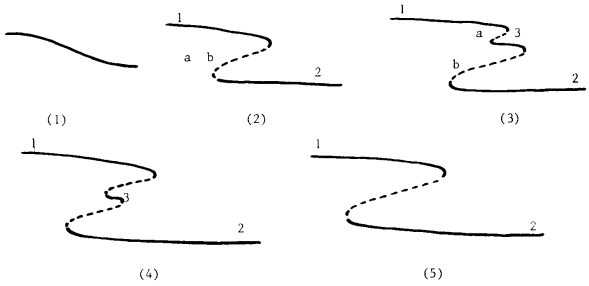


Figure 49: Sections correspondant à celles de la figure 48.

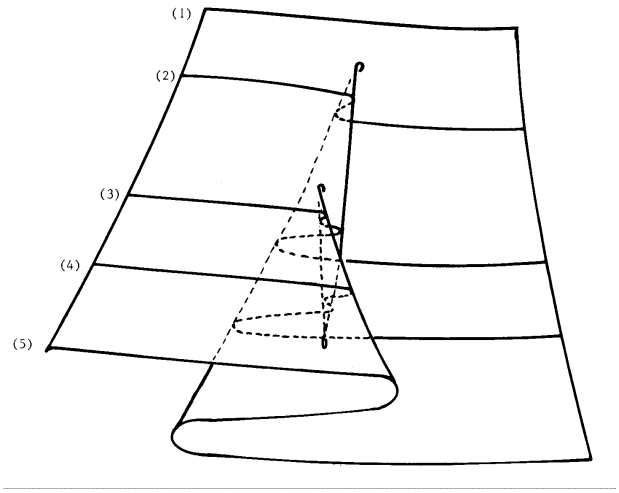


Figure 50: Surface $\Sigma_{t,u}$ complète associée à la figure 49.

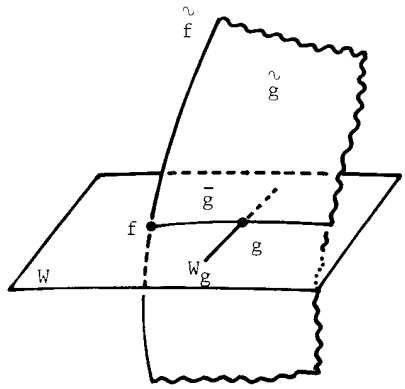


Figure 51: Le principe de la transitivité des déploiements universels.

où, au voisinage de f, $K_{\mathcal{F}}$ est le produit direct de K par $\tilde{f}$. La G-orbite $\tilde{g}$ est donc localement le produit direct de $\overline{g}$ par $\tilde{f}$ et, pour construire un modèle transverse de g (*i.e.* un supplémentaire en g de $\tilde{g}$ dans $\mathcal{F}$) il suffit de construire un supplémentaire en g de $\overline{g}$ dans W (*cf.* figure 51).

On peut exprimer ce phénomène en disant qu'il y a en quelque sorte *transitivité* des déploiements universels : si (W, K) est un modèle transverse de f et si g appartient à W, tout modèle transverse de g dans (W, K) est un modèle transverse de g dans $\mathcal{F}$. Autrement dit, dans la "bonne situation", le déploiement universel (W, K) d'une singularité f de codimension c peut être conçu comme le *recollement* des modèles transverses de ses stabilisés partiels (on dit aussi de ses singularités incidentes) : si $g \in K$, (W, K) est, localement en g, identifiable au produit direct de la strate $\overline{g}$ de g par un modèle transverse W_g de g dans (W, K). On trouvera aux figures 52 et 53 deux exemples de transitivité et l'on pourra facilement vérifier qu'il en va de même pour toutes les singularités incidentes à la queue d'aronde.

La transitivité des déploiements universels implique une conséquence conceptuelle tout à fait non classique en ce qui concerne la dialectique entre le local et le global. En effet, si l'on considère les déploiements comme des "objets", ces "objets" constituent un univers pour lequel les oppositions conceptuelles classiques simple/complexe, irréductible/composé, atome/assemblage, constituant/système, etc. ne sont pas pertinentes. Par exemple la section de dimension 2 du déploiement universel de la queue d'aronde représentée à la figure 28 peut être considérée comme un "mot", un "assemblage" ou un "système" complexe composé à partir des "lettres", des " atomes" ou des "constituants" simples "double pli", "bec", "bec", "double conflit", "cusp" et "cusp dual". Mais dans le même temps ce complexe peut être considéré comme engendré par la "lettre", "l'atome" ou le "constituant" simple "queue d'aronde".

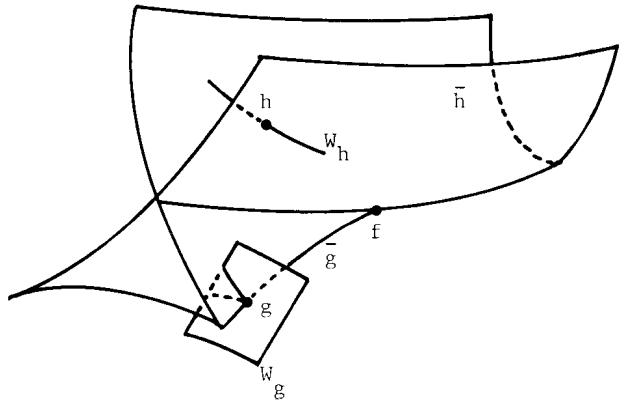


Figure 52: La transitivité des déploiements universels dans le cas de la queue d'aronde.

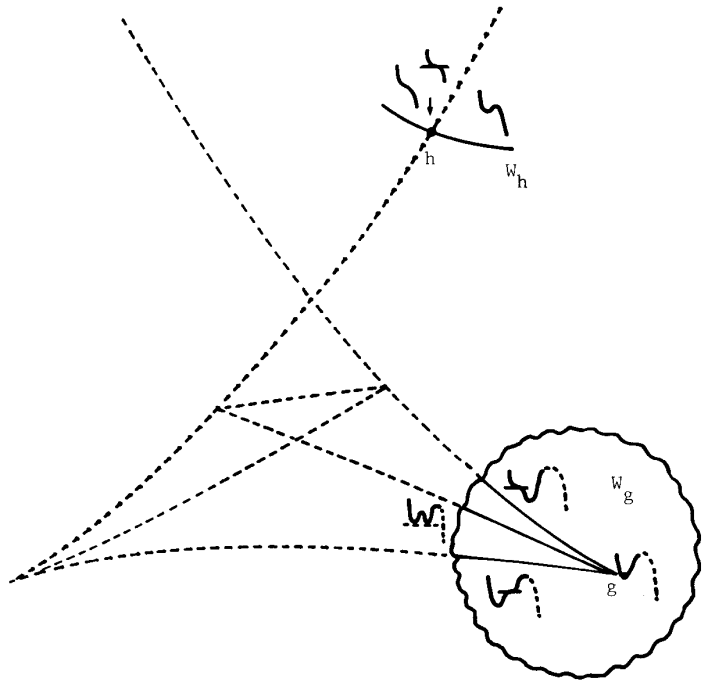


Figure 53: En g appartenant à une strate cusp, un modèle transverse en g à $\overline{g}$ redonne le déploiement universel du cusp de la figure 17. De même, en h appartenant à une strate pli, un modèle transverse en h à $\overline{h}$ redonne le déploiement universel du pli.

Dans la “ logique” des déploiements, on dispose donc d’une double relation allant du simple au complexe. D’une part de la relation classique de composition interprétée en termes de recollement. D’autre part de celle, non classique, de déploiement. En jouant sur cette double relation, on peut “ remonter” dans certains cas du complexe au simple, en traitant les complexes comme des sections de déploiements ayant des centres organisateurs de codimension plus grande. On pourrait dire que, en ce sens, la “grammaire” des déploiements est une grammaire “générative”.

Bibliographie

[1] Arnold, V.I., 1968. Singularities of Smooth Mappings, *Russian Mathematical Surveys*, 23, 1, 3-44.

[2] Arnold, V.I., 1972. Normal Forms for Functions near Degenerate Critical Points, the Weyl Groups A_k, D_k, E_k and Lagrangian Singularities, (trad. anglaise), *Functional Anal. Appl.*, 6, (1973), 254-272.

[3] Arnold, V.I., 1974. Normal Forms of Functions Near Degenerate Critical Points, *Russian Mathematical Surveys*, 29, 2, 11-49.

[4] Arnold, V.I., 1976(a). *Méthodes mathématiques de la Mécanique classique*, Editions Mir, Moscou.

[5] Arnold, V.I., 1976(b). Wave Front Evolution and Equivariant Morse Lemma, *Comm. Pure Appl. Math.*, 29, 557-582.

[6] Boardman, J.M., 1967. Singularities of Differentiable Mappings, *Publications Mathématiques de l'IHES*, 33, 21-57, Presses Universitaires de France, Paris.

[7] Cargèse, 1973. *Singularités en géométrie analytique*, (F. Pham ed.), *Astérisque*, 7-8, Société Mathématique de France.

[8] Cargèse, 1975. Rencontre de Cargèse sur les singularités et leurs applications, (F. Pham ed.), Institut d'études scientifiques de Cargèse.

[9] Chenciner, A., 1973. Travaux de Thom et Mather sur la stabilité topologique, *Séminaire Bourbaki*, 424.

[10] Chenciner, A., 1980. Singularités des fonctions différentiables, *Encyclopædia Universalis*, Paris.

[11] Fukuda, T., 1976. Types topologiques des polynômes, *Publications Mathématiques de l'IHES*, 46, 87-106, Presses Universitaires de France, Paris.

[12] Glaeser, G., 1971. L'interpolation des fonctions différentiables de plusieurs variables, *Liverpool Singularities Symposium*, (C.T.C. Wall ed.), Lecture Notes in Mathematics, 209, 1-29, Springer, Berlin, Heidelberg, New-York.

[13] Godwin, A.N., 1971. Three Dimensional Pictures for Thom's Parabolic Umbilic, *Publications Mathématiques de l'IHES*, 40, 117-138, Presses Universitaires de France, Paris.

[14] Golubitsky, M., Guillemin, V., 1973. *Stable Mappings and their Singularities*, Graduate Texts in Mathematics, 14, Springer, New-York, Heidelberg, Berlin.

[15] Gromoll, D., Meyer, W., 1969. On Differentiable Functions with Isolated Critical Points, *Topology*, 8, 361-370.

[16] Levine, H.I., 1971. Singularities of Differentiable Mappings, *Liverpool Singularities Symposium*, (C.T.C. Wall ed.), Lecture Notes in Mathematics, 192, 1-89, Springer, Berlin, Heidelberg, New-York.

[17] Malgrange, B., 1962-1963. Le théorème de préparation en géométrie différentielle, *Séminaire H. Cartan*, 15, exposés 11, 12, 13, 22.

[18] Malgrange, B., 1966. *Ideals of Differentiable Functions*, Oxford University Press, Oxford.

[19] Martinet, J., 1974. *Singularités des fonctions et applications différentiables*, Université Pontificale de Rio de Janeiro.

[20] Mather, J., 1968(a). Stability of C^∞ Mappings I: the Division Theorem, *Ann. of Math.*, 87, 89-104.

[21] Mather, J., 1968(b). Stability of C^∞ Mappings III: Finitely Determined Map Germs, *Publications Mathématiques de l'IHES*, 35, 127-156, Presses Universitaires de France, Paris.

[22] Mather, J., 1969(a). Stability of C^∞ Mappings II: Infinitesimal Stability Implies Stability, *Ann. of Math.*, 89, 254-291.

[23] Mather, J., 1969(b). Stability of C^∞ Mappings IV: Classification of Stable Germs by $\mathbb{R}$ -algebras, *Publications Mathématiques de l'IHES*, 37, 223-248, Presses Universitaires de France, Paris.

[24] Mather, J., 1970. Stability of C^∞ Mappings V: Transversality, *Advances in Math.*, 4, 301-336.

[25] Mather, J., 1971(a). Stability of C^∞ Mappings VI: the Nice Dimensions, *Liverpool Singularities Symposium*, (C.T.C. Wall ed.), Lecture Notes in Mathematics, 192, 207-253, Springer, Berlin, Heidelberg, New-York.

[26] Mather, J., 1971(b). On Nirenberg's Proof of Malgrange's Preparation Theorem, *Liverpool Singularities Symposium*, (C.T.C. Wall ed.), Lecture Notes in Mathematics, 192, 116-120, Springer, Berlin, Heidelberg, New-York.

[27] Milnor, J., 1968. Singular Points of Complex Hypersurfaces, *Princeton University Press*.

[28] Nirenberg, L., 1971. A proof of the Malgrange preparation theorem, *Liverpool Singularities Symposium*, (C.T.C. Wall ed.), Lecture Notes in Mathematics, 192, 97-105, Springer, Berlin, Heidelberg, New-York.

[29] Petitot, J., 1977. Introduction à la théorie des catastrophes, *Mathématiques et Sciences Humaines*, 59, Ecole des Hautes Etudes en Sciences Sociales, Paris.

[30] Petitot, J., 1978. Catastrophes (Théorie des), *Encyclopædia Universalis*.

[31] Petitot, J, 1979. Locale/Globale, *Enciclopedia Einaudi*, VIII, 429-490, Einaudi, Turin.

[32] Petitot, J., 1982. *Pour un schématisme de la structure : de quelques implications sémiotiques de la théorie des catastrophes*, 4 vol., Thèse, Ecole des Hautes Etudes en Sciences Sociales, Paris.

[33] Petitot, J., (ed.) 1989. *Logos et théorie des catastrophes*, Editions Patiño, Genève.

[34] Poénaru, V., 1976. Singularités C^∞ en présence de symétries, *Lecture Notes in Mathematics*, 510, Springer, Berlin, Heidelberg, New-York.

[35] Porteous, I.R., 1981. The Normal Singularities of Surfaces in $\mathbb{R}^3$, *Proceedings of the AMS Summer Institute on Singularities*, Arcata, Californie.

[36] Poston, T., Stewart, I, 1976. *Taylor Expansions and Catastrophes*, Research Notes in Mathematics, 7, Pitman Publishing, Londres.

[37] Poston, T., Stewart, I, 1978. *Catastrophe Theory and its Applications*, Pitman Publishing, Londres.

[38] Ruelle, D., 1984. Déterminisme et prédictibilité, *Pour la Science*, 82, 58 sq.

[39] Sard, A., 1942. The Measure of the Critical Values of Differentiable Maps, *Bull. Amer. Math. Soc.*, 48, 883-896.

[40] Siersma, D., 1973. Singularities of C^∞-functions of Right-codimension Smaller or Equal than Eight, *Indag. Math.*, 25, 31-37.

[41] Siersma, D., 1974. Classification and Deformation of Singularities, *Academic Service*, Vinkeveen, Hollande.

[42] Slodowy, P., 1980. Simple Singularities and Simple Algebraic Groups, *Lecture Notes in Mathematics*, 815, Springer, New-York, Heidelberg, Berlin.

[43] Slodowy, P., 1982. Platonic Solids, Kleinian Singularities, Elementary Catastrophes, and Lie Groups, *Logos et théorie des catastrophes*, (J. Petitot ed.), Editions Patiño, Genève, 1989, 71-98.

[44] Teissier, B., 1974. Introduction to Equisingularity Problems, *Proceedings of Symposia in Pure Mathematics*, Vol. XXIX, *Algebraic Geometry*, Arcata, American Mathematical Society, Providence.

[45] Teissier, B., 1976. The Hunting of Invariants in the Geometry of Discriminants, *Nordic Summer School*, Noordhoff et Sijthoff ed., Oslo.

[46] Thom, R., 1955. Les singularités des applications différentiables, *Ann. Institut Fourier*, 6, 43-87.

[47] Thom, R., 1964. Local Properties of Differentiable Mappings, *Differential Analysis* (Bombay Colloquium), 191-202, Oxford University Press, Oxford.

[48] Thom, R., 1969. Ensembles et morphismes stratifiés, *Bull. Amer. Math. Soc.*, 75, 2, 240-284.

[49] Thom, R., 1970. The Bifurcation Subset of a Space of Maps, *Manifold*, Lecture Notes in Mathematics, 197, 202-208, Springer, Berlin, Heidelberg, New-York.

[50] Thom, R., 1971. Stratified Sets and Morphisms, *Liverpool Singularities Symposium*, (C.T.C. Wall ed.), Lecture Notes in Mathematics, 192, 153-165, Springer, Berlin, Heidelberg, New-York.

[51] Thom, R., 1980(a). *Modèles mathématiques de la Morphogenèse*, (2e ed.), Christian Bourgois, Paris.

[52] Thom, R., 1980(b). Halte au hasard, silence au bruit, *Le Débat*, 3, 119-132, Gallimard, Paris.

[53] Thom, R., 1984. Classification des sciences et des techniques, *Séminaire de Philosophie et Mathématiques*, Ecole Normale Supérieure, Paris (16 Janvier 1984).

[54] Tougeron, J.C., 1972. *Idéaux de fonctions différentiables*, Springer, New-

York, Heidelberg, Berlin.

[55] Wall, C.T.C., 1971. Stratified Sets: a Survey, *Liverpool Singularities Symposium*, (C.T.C. Wall ed.), Lecture Notes in Mathematics, 192, 133-140, Springer, Berlin, Heidelberg, New-York.

[56] Zeeman, E.C., 1976(a) (avec D.J.A. Trotman). The classification of Elementary Catastrophes of Codimension ≤ 5, *Lecture Notes in Mathematics*, 525, 263-327, (repris dans Zeeman 1977).

[57] Zeeman, E.C., 1976(b). The Umbilic Bracelet and the Double-cusp Catastrophe, *Lecture Notes in Mathematics*, 525, 328-366, (repris dans Zeeman 1977).

[58] Zeeman, E.C., 1977. *Catastrophe Theory: Selected Papers 1972-1977*, Addison-Wesley, Mass.