

# Résultats récents sur les courbes optimales \*

Héctor J. Sussmann †‡

## Sommaire

On se propose de faire un bilan de quelques résultats récents sur les conditions nécessaires pour qu'une courbe soit une solution du problème de minimisation d'une fonctionnelle réelle définie sur une famille de courbes dans  $\mathbb{R}^n$ . On verra comment, par une série de généralisations successives, ces conditions ont évolué à partir de l'équation d'Euler-Lagrange jusqu'au « principe du maximum » de la théorie du contrôle, et comment celui-ci a été à son tour généralisé dans plusieurs directions. Vers 1990, cette évolution avait abouti à une prolifération de versions, sous des hypothèses techniques différentes et pour des problèmes divers, qui ne s'inscrivaient pas dans un cadre unique permettant de dériver tous les résultats à partir d'un petit nombre de principes généraux. Plus récemment, on s'est aperçu que la plupart de ces versions n'étaient en réalité que des cas particuliers d'un seul théorème sur les variations des flots, leur propagation par des « différentielles généralisées », et l'existence de points d'intersection de deux ensembles « transversaux ». Ce théorème sera énoncé pour des « calculs différentielles généralisés » (CDGs) quelconques. On donnera un exemple nouveau de CDG, en définissant la notion de « quotient différentiel généralisé », qui permet de démontrer une version très forte du principe du maximum.

## 1 Introduction

Le but de cet exposé est de présenter quelques idées récentes sur le très vieux problème de la minimisation d'une fonctionnelle réelle  $\mathcal{F}$  définie sur un ensemble (ou sur une famille paramétrisée)  $\mathcal{C}$  de courbes dans  $\mathbb{R}^n$ .

Plus précisément, il s'agit d'étudier ou de caractériser les solutions du problème suivant : on se donne une famille  $\mathcal{C} = \{\xi_\alpha\}_{\alpha \in A}$  de courbes et une fonctionnelle  $\mathcal{F} : A \rightarrow \mathbb{R}$ , et on cherche les valeurs  $\bar{\alpha}$  du paramètre tels que  $\mathcal{F}(\bar{\alpha}) \leq \mathcal{F}(\alpha)$  quel que soit  $\alpha \in A$ . À ce niveau de généralité, cette question est évidemment identique à celle de la minimisation d'une fonction réelle arbitraire  $f$  sur un ensemble arbitraire  $S$ , mais la spécificité du problème pour les *familles de courbes*

---

\*Travail présenté à la journée annuelle de la Société Mathématique de France du 17 juin 2000.

†Department of Mathematics, Rutgers University, Piscataway, NJ 08854, USA

‡Research supported in part by NSF Grant DMS-9803411 and AFOSR Grant 0923

provient de ce que l'on s'intéresse à des propriétés des solutions liées à une *structure dynamique* qui donne lieu à la famille  $\mathcal{C}$ , comme on verra clairement dans la suite. (Par exemple, on cherchera des énoncés du type : « si  $\bar{\alpha}$  est une solution, et la courbe  $\xi_{\bar{\alpha}}$  est définie sur l'intervalle  $[a, b]$ , il existe une fonction  $\pi : [a, b] \rightarrow \mathbb{R}^n$  qui vérifie l'équation différentielle  $\frac{d\pi}{dt}(t) = \frac{\partial L}{\partial x}(\xi_{\bar{\alpha}}(t), \frac{d\xi_{\bar{\alpha}}}{dt}(t))$  aussi bien que l'identité  $\pi(t) = \frac{\partial L}{\partial \dot{x}}(\xi_{\bar{\alpha}}(t), \frac{d\xi_{\bar{\alpha}}}{dt}(t))$  ».)

Le lecteur trouvera d'abord une esquisse très abrégée de l'histoire du problème, que nous diviserons en trois grandes périodes, dont la première — qui s'étend de la préhistoire jusqu'au début du XVIII<sup>e</sup> siècle — est celle des « problèmes particuliers étudiés par des méthodes *ad hoc* ». De cette époque, nous choisissons cinq exemples de problèmes de minimisation, dont on se servira plus tard pour montrer comment les développements ultérieurs — y compris ceux des dix dernières années — permettent une meilleure compréhension des solutions et donnent lieu à des résultats nouveaux même pour des questions très classiques.

Vers la fin du XVII<sup>e</sup> siècle et le début du XVIII<sup>e</sup>, les frères Bernoulli, Euler et plus tard Lagrange, réussirent à dégager une structure qui leur permit d'étudier ces problèmes par des techniques générales, qui plus tard seraient nommées (par Euler) « calcul des variations ». La deuxième époque est donc celle du calcul des variations, qui se prolongera jusqu'à la naissance du *contrôle optimal*, vers 1950-60.

Pour préparer notre description de la transition du calcul des variations au contrôle optimal, nous illustrerons les résultats du calcul de variations classique en suivant l'évolution des conditions nécessaires pour un minimum d'un point de vue très particulier. Partant de l'équation d'Euler-Lagrange, nous prendrons parti pour la logique plutôt que pour l'histoire, en passant directement à la *formulation hamiltonienne non classique*, qui à notre avis représente ce qui aurait dû être fait vers 1830, mais ne fut fait que vers 1950. Dans cette formulation, le « hamiltonien » est une « fonction de la position, de la vitesse et de l'impulsion » (c'est-à-dire une fonction définie sur le produit fibré des fibrés tangent et cotangent de l'espace des configurations  $X$ ). Afin de convaincre le lecteur de l'inévitabilité de cette reformulation — qui joue un rôle crucial dans le « principe du maximum » de la théorie du contrôle — nous soulignerons les limitations de l'équation d'Euler-Lagrange et de la forme classique des équations de Hamilton, où le hamiltonien est une « fonction de la position et de l'impulsion », c'est-à-dire une fonction définie sur le fibré cotangent de  $X$ , appelé *espace des phases*.

Pour la troisième période, qui s'étend des années 1950 à nos jours, nous choisissons trois sujets : (a) l'énoncé de la version classique — « lisse » — du « principe du maximum de Pontriaguine », (b) les résultats obtenus à partir de 1970 par plusieurs auteurs — notamment F. Clarke, A. Ioffe, B. Mordukhovich, R.T. Rockafellar, R. Vinter, J. Warga, Q. Zhu — sur les problèmes « non-lisses » (ce qui dans ce contexte veut dire « lipschitziens ») par des méthodes « duales » inspirées par des considérations d'analyse convexe, et, surtout, (c) notre version récente du principe du maximum.

Le cœur de notre exposé est le traitement, au §4, du progrès des dix dernières années vers la formulation d’une version unifiée et générale des conditions nécessaires. Cette version — qui est l’aboutissement d’un chemin tracé par un groupe de chercheurs<sup>1</sup> d’orientation plutôt « primale » — marque un retour à la tradition de Weierstrass et des créateurs du contrôle optimal, puisqu’elle se démontre en utilisant les *variations ponctuelles* où *variations aiguille*. Elle contient la plupart des théorèmes qui ont été démontrés par d’autres techniques et, en plus, elle s’applique aussi à des problèmes « très non-lisses » et « hybrides » où les hypothèses des autres théorèmes ne sont pas vérifiées. En outre, elle donne très souvent des conclusions plus fortes pour des problèmes où les autres résultats sont applicables.

Les trois idées centrales de notre méthode découlent des trois observations suivantes : (1) La preuve du principe du maximum se sert des différentielles de façon décisive (en calculant d’abord des vecteurs tangents qui représentent la « partie de premier ordre de l’effet d’une variation », et puis en utilisant la différentielle du « flot de référence » le long de la trajectoire de référence pour transporter tous ces vecteurs tangents au point terminal), mais elle ne dépend en réalité que d’un très petit nombre de propriétés de la différentielle. Elle peut donc être adaptée facilement à d’autres « calculs différentiels généralisés » (abrégés : CDGs). (2) Dans les versions usuelles du principe du maximum, les champs de vecteurs ne jouent que le rôle de *générateurs de flots*, ce qui permet de récrire le principe sous la forme d’un *énoncé sur des familles de flots* plutôt que sur des familles de champs de vecteurs. Puisque les propriétés de régularité des flots sont typiquement plus fortes que celles des champs qui les engendrent, l’énoncé pour les flots est un théorème plus fort. (3) Le principe du maximum doit être perçu comme un résultat géométrique sur la séparation de deux ensembles. Plus précisément, le vrai problème est celui de trouver des conditions suffisantes pour que l’intersection de deux sous-ensembles  $S_1, S_2$  de  $\mathbb{R}^n$  tels que  $0 \in S_1 \cap S_2$  contienne en outre d’autres points de  $\mathbb{R}^n$ . Il s’agit évidemment d’une question de *transversalité*, que nous aborderons en proposant une notion de « transversalité » pour deux cônes convexes et une généralisation de cette notion au cas de deux « multicônes » convexes. (Un *multicône* est un ensemble de cônes. Les multicônes sont les objets qui, dans notre théorie, jouent de façon naturelle le rôle d’« objets tangents » à un ensemble en un point, généralisant la notion de sous-espace tangent d’une sous-variété.) Pour pouvoir démontrer le principe du maximum pour un CDG donné, il suffit que ce CDG possède une *propriété de transversalité* : du fait que deux ensembles ont des « multicônes tangents » fortement transversaux<sup>2</sup> il faut que l’on puisse déduire que l’intersection des deux ensembles ne se réduit pas à un seul point. Cette propriété est à son tour une conséquence d’un « théorème de l’application directionnellement ouverte », qu’on démontre, pour tous les CDGs introduits ici, par des arguments de point fixe.

<sup>1</sup>Notamment A. Bressan, A. Cellina, A. Fryszkowski, H. Halkin, B. Kaskosz, S. Lojasiewicz, et T. Rzeżuchowski

<sup>2</sup>Deux cônes sont fortement transversaux s’ils sont transversaux et leur intersection contient une demi-droite; pour la généralisation aux multicônes, cf. la définition 4.4.2.

## 2 Quelques notations et définitions

Les symboles  $\mathbb{Z}, \mathbb{Z}_+, \bar{\mathbb{Z}}_+, \mathbb{N}, \bar{\mathbb{N}}, \mathbb{R}, \mathbb{R}_+, \bar{\mathbb{R}}_+$  désignent, respectivement, l'ensemble des nombres entiers, l'ensemble  $\{n : n \in \mathbb{Z} \wedge n \geq 0\}$ , la réunion  $\mathbb{Z}_+ \cup \{+\infty\}$ , l'ensemble  $\{n : n \in \mathbb{Z} \wedge n > 0\}$ , la réunion  $\mathbb{N} \cup \{+\infty\}$ , l'ensemble des nombres réels, l'intervalle  $[0, +\infty[$ , et l'intervalle  $[0, +\infty]$ .

On notera  $X^\dagger$  le dual algébrique d'un espace vectoriel réel  $X$ . Les vecteurs  $x$  de  $\mathbb{R}^n$  seront identifiés à des matrices  $n \times 1$ , c'est-à-dire à des vecteurs colonne. L'expression  $\mathbb{R}_n$  désignera le dual  $(\mathbb{R}^n)^\dagger$ , dont les éléments seront identifiés à des matrices  $1 \times n$ , c'est-à-dire à des vecteurs ligne. Il s'ensuit que les expressions  $p \cdot x$  et  $p(x)$  désignent le même nombre réel, qui sera aussi noté  $\langle p, x \rangle$ .

Si  $X, Y$  sont des espaces vectoriels réels,  $L(X, Y)$  désignera l'espace des applications linéaires de  $X$  dans  $Y$ . Si  $X = \mathbb{R}^n$  et  $Y = \mathbb{R}^m$ ,  $L(\mathbb{R}^n, \mathbb{R}^m)$  sera identifié à l'espace  $\mathbb{R}^{m \times n}$  des matrices réelles à  $m$  lignes et  $n$  colonnes. Ainsi donc, si  $p \in \mathbb{R}_m$  et  $M \in \mathbb{R}^{m \times n}$ , les expressions  $p \cdot L$  et  $p \circ L$  désignent le même élément de  $\mathbb{R}_n$ , qui sera aussi noté  $\langle p, L \rangle$ .

Toutes les variétés seront supposées de classe  $C^k$ ,  $k \in \bar{\mathbb{Z}}_+$ , de dimension finie, séparées, sans bord, et paracompactes.

Si  $k \in \bar{\mathbb{N}}$ ,  $X$  est une variété de classe  $C^k$ , et  $x \in X$ , on notera  $TX, T^*X, T_xX, T_x^*X$ , respectivement, les fibrés tangent et cotangent de  $X$  —qui sont donc des variétés de classe  $C^{k-1}$ — et les espaces tangent et cotangent de  $X$  en  $x$ . Lorsque la variété  $X$  est appelée l'*espace des positions*, ou *espace des configurations*, ou *espace des états*, d'un système, on appellera  $TX$  l'*espace des positions* (ou *des configurations*, ou *des états*) et des vitesses, et  $T^*X$  sera appelé l'*espace des positions* (ou *des configurations*, ou *des états*) et des impulsions, ou l'*espace des phases*.

Si  $k, \ell \in \bar{\mathbb{Z}}_+$ ,  $\ell \leq k$ , et  $E, F$  sont deux fibrés vectoriels de classe  $C^\ell$  sur une variété  $X$  de classe  $C^k$ , on notera  $E \times_X F$  le *produit fibré* de  $E$  et  $F$ , c'est-à-dire, le fibré vectoriel  $G$  de classe  $C^\ell$  sur  $X$  tel que, pour  $x \in X$ , la fibre  $G_x$  est le produit des fibres  $E_x$  et  $F_x$ . Dans ce travail, l'exemple le plus important de produit fibré est  $TX \times_X T^*X$ , que nous appellerons l'*espace des vitesses et des phases* associé à la variété  $X$ . Si  $X$  est de dimension  $n$ ,  $TX \times_X T^*X$  est évidemment de dimension  $3n$ .

La notation  $f : A \dashrightarrow B$  indiquera que  $f$  est une application partielle de  $A$  dans  $B$ , c'est-à-dire une application dont le domaine  $\text{Dom}(f)$  et l'image  $\text{Im}(f)$  vérifient  $\text{Dom}(f) \subseteq A$  et  $\text{Im}(f) \subseteq B$ .

Une *courbe* dans un espace topologique  $X$  est une application continue  $\xi$  d'un sous-intervalle non vide  $I$  (le *domaine* de  $\xi$ , noté  $\text{Dom}(\xi)$ ) de  $\mathbb{R}$  dans  $X$ . Si  $X$  est une variété,  $\xi$  est une courbe dans  $X$  et  $t \in \text{Dom}(\xi)$ , on notera  $\dot{\xi}(t)$  la dérivée de  $\xi$  en  $t$ , si elle existe. Le vecteur  $\dot{\xi}(t)$  appartient donc à  $T_{\xi(t)}X$ , et  $\dot{\xi} : \text{Dom}(\dot{\xi}) \dashrightarrow TX$  est un champ de vecteurs partiel le long de  $\xi$ . Si  $a, b \in \mathbb{R}$ ,  $a < b$ , et  $\text{Dom}(\xi) = [a, b]$ , le point  $\partial\xi \stackrel{\text{déf}}{=} (\xi(a), \xi(b)) \in X \times X$  est la *valeur au bord* de  $\xi$ .

Si  $A, B$  sont des ensembles, on notera  $FNS(A, B)$  l'ensemble de toutes

les fonctions partielles  $f : A \times \mathbb{R} \dashrightarrow B$ . (Les éléments de  $FNS(A, B)$  sont donc des fonctions  $A \times \mathbb{R} \ni (x, t) \rightarrow f(x, t) \in B$ , c'est-à-dire des « fonctions de  $x$  qui dépendent aussi du temps », ou des « fonctions non stationnaires de  $x$  », d'où la notation  $FNS(A, B)$ .) Si  $f \in FNS(A, B)$  et  $B$  est un espace mesurable, on dit que  $f$  est *mesurable par rapport au temps* si la fonction partielle  $\mathbb{R} \ni t \dashrightarrow f(x, t) \in B$  est mesurable (c'est-à-dire, l'ensemble

$$f_x^{-1}(] - \infty, \alpha]) \stackrel{\text{d\u00e9f}}{=} \{t \in \mathbb{R} : (x, t) \in \text{Dom}(f) \wedge f(x, t) \in M\}$$

appartient à la tribu de Lebesgue de  $\mathbb{R}$  pour tout sous-ensemble mesurable  $M$  de  $B$ ) quel que soit  $x \in A$ . Si  $A$  et  $B$  sont des espaces topologiques, on dit que  $f$  est *continue par rapport à l'état* si la fonction partielle

$$A \ni x \dashrightarrow f^t(x) \stackrel{\text{d\u00e9f}}{=} f(x, t) \in B$$

est une application continue de  $\text{Dom}(f^t)$  dans  $B$  quel que soit  $t \in \mathbb{R}$ , et que  $f$  est *localement bornée* si pour tout sous-ensemble compact  $K$  de  $\text{Dom}(f)$  l'image  $f(K)$  est relativement compacte dans  $B$ . Si  $A$  et  $B$  sont des variétés de classe  $C^k$ , on dit que  $f$  est *de classe  $C^k$  par rapport à l'état*<sup>3</sup> si pour tout  $t \in \mathbb{R}$  le domaine de la fonction partielle  $f^t$  est une partie ouverte de  $A$  et  $f^t$  est de classe  $C^k$  sur ce domaine, et que  $f$  et toutes ses dérivées partielles d'ordre  $\leq k$  par rapport à l'état<sup>3</sup> sont localement bornées si  $f$  est de classe  $C^k$  par rapport à l'état et localement bornée, et la fonction  $\text{Dom}(f) \ni (x, t) \rightarrow \left( V_1 V_2 \cdots V_\ell (\varphi \circ f^t) \right)(x) \in \mathbb{R}$  est localement bornée quels que soient  $\ell \in \{1, \dots, k\}$ , les champs de vecteurs  $V_1, \dots, V_\ell$  de classe  $C^{k-1}$  sur  $A$ , et la fonction  $\varphi : B \rightarrow \mathbb{R}$  de classe  $C^k$ .

### 3 L'histoire du problème, des étireurs de cordes égyptiens à nos jours

#### 3.1 L'ère des problèmes particuliers résolus par des méthodes *ad hoc*

**Exemple 3.1.1** *Le plus court chemin entre deux points donnés.* C'est le plus ancien problème de minimisation pour les courbes. On montre aisément que cette question est équivalente au « problème dual » suivant : parmi tous les chemins  $\xi : [0, 1] \rightarrow \mathbb{R}^n$  d'une longueur donnée  $L$ , trouver ceux qui maximisent la distance de  $\xi(0)$  à  $\xi(1)$ .

L'observation que les solutions sont les segments de droite se remonte aux plus anciennes civilisations. (En Égypte antique il y avait des travailleurs — appelés « harpenodaptai », ou « étireurs de cordes », cf. Gandz [7] — spécialisés dans l'art d'étirer des cordes pour produire des longs segments de

<sup>3</sup>Bien sûr, si  $A$  est le fibré tangent (resp. cotangent)  $TX$  (resp.  $T^*X$ ) d'une variété  $X$ , on dira « par rapport à l'état et la vitesse » (resp. « par rapport à l'état et l'impulsion ») plutôt que « par rapport à l'état ».

droite, ce qui montre qu'on comprenait parfaitement que le segment de droite résout le problème dual, et aussi qu'on savait *appliquer* cette observation à la construction de segments. On voit donc bien que l'application des mathématiques à l'ingénierie<sup>4</sup>, dont la théorie du contrôle est un exemple moderne, est une activité très ancienne, dont les premières manifestations sont peut-être antérieures à l'application des mathématiques à la physique.)

Notons que notre problème, si on l'écrit dans le langage mathématique d'aujourd'hui, devient celui de minimiser l'intégrale  $\int_0^1 \|\dot{\xi}(t)\| dt$  dans l'ensemble de tous les chemins absolument continus  $\xi : [0, 1] \rightarrow \mathbb{R}^n$  tels que  $\xi(0) = A$  et  $\xi(1) = B$ , où  $A$  et  $B$  sont des points fixés. Il s'agit donc d'un *problème aux variations* tout à fait classique, mais *avec un lagrangien non-lisse*. Il s'ensuit que l'équation d'Euler-Lagrange n'est pas directement applicable pour trouver toutes les solutions, y compris les chemins constants, qui correspondent précisément à la vitesse nulle, c'est-à-dire, aux points où le lagrangien n'est pas différentiable. Par conséquent, pour trouver la solution complète du problème on est obligé d'employer, outre le calcul des variations, des transformations *ad hoc* ou des arguments spéciaux pour tenir compte de tous les cas possibles. On verra plus tard (cf. Exemple 3.3.4) que ce problème s'inscrit de façon naturelle —sans qu'on ait besoin d'aucun argument supplémentaire— dans le cadre de la théorie du contrôle optimal, et que la supériorité du contrôle optimal devient indéniable dès que l'on passe de la norme euclidienne à une norme générale, qui pourrait ne pas être différentiable partout sur  $\mathbb{R}^n \setminus \{0\}$ .  $\diamond$

**Exemple 3.1.2 La loi de la réflexion.** Héron d'Alexandrie (I<sup>er</sup> siècle de notre ère) montra dans son œuvre *La Catoptrique* que la loi de la réflexion de la lumière par un miroir —c'est-à-dire l'égalité des angles d'incidence et réflexion— peut être déduite d'un principe variationnel dont l'énoncé dit, très simplement, que la lumière suit le plus court chemin.

Le calcul des plus courts chemins pour la réflexion est un *problème de contrôle optimal hybride* : si  $M$  désigne la surface du miroir, on cherche, pour deux points  $A, B$  qui se trouvent du même côté de  $M$ , le chemin le plus court de  $A$  à  $B$  qui passe par un point de  $M$ . Notons cependant que *le problème de la réflexion ne s'inscrit ni dans le cadre du calcul des variations classique ni dans celui du principe du maximum*.  $\diamond$

**Exemple 3.1.3 La loi de la réfraction.** Au XVII<sup>e</sup> siècle, Fermat généralisa l'observation de Héron, et énonça le *principe du temps minimum*, selon lequel les rayons de lumière dans un milieu à vitesse de lumière variable suivent les chemins qui demandent le temps le plus bref.

Leibniz et Huygens montrèrent qu'à partir du principe de Fermat on déduit la *loi de Snell*<sup>5</sup> sur la réfraction, qui dit que les sinus des angles d'incidence et de réfraction —lorsque la lumière traverse un plan  $P$  qui sert de frontière

<sup>4</sup>Religieuse aussi bien que civile, puisqu'il fallait bâtir des temples, des autels et des pyramides aussi bien que des palais.

<sup>5</sup>ou de « Snell-Descartes »

entre deux milieux à vitesses de lumière constantes mais différentes — sont proportionnels aux vitesses de la lumière dans les deux milieux.

Notons que, comme dans le cas de la réflexion, *le calcul des variations classique ne s'applique pas au problème de minimisation pour un milieu où la vitesse de la lumière est discontinue*. Par contre, ce problème se prête naturellement à l'application de la version récente du principe du maximum pour les flots.  $\diamond$

**Exemple 3.1.4 La caténaire.** On cherche à déterminer la configuration d'une corde  $C$  de longueur donnée et de densité de masse constante, dont les extrémités  $A$  et  $B$  sont fixées. Mathématiquement, si on admet que  $C$  est décrite par une fonction  $[a, b] \ni x \rightarrow y(x) \in \mathbb{R}$  telle que  $A = (a, y(a))$  et  $B = (b, y(b))$ , on veut trouver la fonction  $y(\cdot)$  qui minimise l'ordonnée du centre de gravité de  $C$ . Il s'agit donc de minimiser l'intégrale  $I = \int_a^b y(x) \sqrt{1 + y'(x)^2} dx$ , dans l'ensemble  $\mathcal{Y} = \mathcal{Y}(a, b, \alpha, \beta, \Lambda)$  de toutes les fonctions  $y(\cdot) \in W^{1,1}([a, b], \mathbb{R})$  telles que

$$y(a) = \alpha, \quad y(b) = \beta, \quad \text{et} \quad \int_a^b \sqrt{1 + y'(x)^2} dx = \Lambda, \quad (1)$$

où  $a, b, \alpha, \beta, \Lambda$  sont des nombres réels fixés, tels que  $a < b$  et  $\Lambda \geq 0$ . L'ensemble  $\mathcal{Y}$  étant évidemment vide si  $\Lambda < \sqrt{(b-a)^2 + (\beta-\alpha)^2} \stackrel{\text{def}}{=} \Delta$ , on supposera que  $\Lambda \geq \Delta$ .

Jean Bernoulli trouva la solution en 1691. Elle est donnée (si  $\Lambda > \Delta$ ) par la formule

$$y(x) = \gamma + \kappa \cosh\left(\frac{x - x_0}{\kappa}\right), \quad a \leq x \leq b,$$

où  $\gamma, \kappa, x_0$  sont des nombres réels tels que  $\kappa > 0$ , dont les valeurs sont déterminés par (1). Plus précisément, on définit  $\mathcal{C}$  et  $\mathcal{D}$  par

$$\tanh \mathcal{C} = \frac{\beta - \alpha}{\Lambda}, \quad \frac{\sinh \mathcal{D}}{\mathcal{D}} = \frac{\sqrt{\Lambda^2 - (\beta - \alpha)^2}}{b - a},$$

et on détermine  $\gamma, \kappa, x_0$  à partir de  $\mathcal{C}$  et  $\mathcal{D}$  par les formules

$$\kappa = \frac{b - a}{2\mathcal{D}}, \quad x_0 = \frac{a + b}{2} - \frac{(b - a)\mathcal{C}}{2\mathcal{D}}, \quad \gamma = \alpha - \frac{b - a}{2\mathcal{D}} \cosh(\mathcal{C} - \mathcal{D}).$$

Il en résulte que

$$y(x) = \alpha + \frac{b - a}{2\mathcal{D}} \left( \cosh\left(\mathcal{C} + \frac{\mathcal{D}(2x - a - b)}{b - a}\right) - \cosh(\mathcal{C} - \mathcal{D}) \right), \quad (2)$$

toujours sous l'hypothèse  $\Lambda > \Delta$ . Un calcul très simple montre que la limite  $\bar{y}(\cdot)$  de la solution lorsque  $\Lambda \downarrow \Delta$  est donnée par

$$\bar{y}(x) = \alpha + \frac{\beta - \alpha}{b - a} \cdot (x - a), \quad (3)$$

qui est d'ailleurs la seule fonction appartenant à  $\mathcal{Y}(a, b, \alpha, \beta, \Delta)$ , puisque le cas « singulier »  $\Lambda = \Delta$  correspond à la situation où la longueur de la corde est

exactement égale à la distance entre ses deux extrémités. La fonction  $\bar{y}$  est donc évidemment la solution du problème de minimisation lorsque  $\Lambda = \Delta$ .

Nous verrons plus tard que *dans le cas singulier, la fonction  $\bar{y}$ , qui est la seule solution du problème, ne vérifie pas les conditions nécessaires du calcul des variations classique, mais elle vérifie une version modifiée de ces conditions, contenant le « multiplicateur anormal ».*  $\diamond$

**Exemple 3.1.5 La brachistochrone et la brachistochrone réfléchie.** Le problème de la brachistochrone —posé par Jean Bernoulli en 1696— est mathématiquement équivalent à la question de trouver les rayons de lumière dans le demi-plan fermé  $\{(x, y) : y \geq 0\}$  lorsque la vitesse  $c$  de la lumière est donnée par  $c(x, y) = \sqrt{y}$ . Jean Bernoulli montra en 1697 que les solutions sont des cycloïdes.

Le *problème de la brachistochrone réfléchie* est une version légèrement modifiée de la question de Jean Bernoulli, que nous proposons parce qu'elle nous semble tout à fait naturelle du point de vue mathématique, et parce qu'elle donne lieu à des résultats intéressants : on pose exactement la même question mais sur le plan tout entier,  $c$  étant donné par  $c(x, y) = \sqrt{|y|}$ .

Pour la brachistochrone réfléchie, lorsque les extrémités  $A$  et  $B$  du rayon cherché ne sont pas du même côté de la droite  $L = \{(x, y) : y = 0\}$ , on peut appliquer les conditions nécessaires données par l'équation d'Euler-Lagrange, ou le principe du maximum, pour montrer que les solutions sont composées de deux arcs de cycloïde  $\gamma_1, \gamma_2$ , tels que  $\gamma_1$  rejoint  $A$  à un point  $P$  de  $L$ , et  $\gamma_2$  rejoint  $P$  à  $B$ . Mais ces conditions ne disent rien sur  $P$ , parce que la vitesse n'est ni lisse ni lipschitzienne au voisinage de  $P$ . Par contre, notre version pour les flots s'applique directement, et donne la condition qui manquait sur  $P$  : le point  $P$  doit être tel que, si  $R_1, R_2$  sont les rayons des cercles roulants  $C_1, C_2$  qui engendrent les arcs de cycloïde  $\gamma_1, \gamma_2$ , alors il faut que  $R_1 = R_2$ .  $\diamond$

## 3.2 Le calcul des variations classique

Si  $k \in \bar{\mathbb{N}}$  et  $X$  est une variété différentiable de classe  $C^k$ , un *lagrangien* sur  $X$  est une fonction partielle  $TX \times \mathbb{R} \ni (x, v, t) \mapsto L(x, v, t)$ . Si  $\ell \in \bar{\mathbb{Z}}_+$  et  $\ell + 1 \leq k$ , on notera  $\text{Lag}^{\ell, lb}(X)$  l'ensemble de tous les lagrangiens  $L$  sur  $X$  qui sont mesurables par rapport au temps, de classe  $C^\ell$  par rapport à l'état et la vitesse, et tels que  $L$  et toutes ses dérivées partielles d'ordre  $\leq \ell$  par rapport à l'état et la vitesse son localement bornées.

### 3.2.1 L'équation d'Euler-Lagrange

Soit  $X$  une variété différentiable de classe  $C^1$ . Si  $a, b \in \mathbb{R}$ ,  $a < b$ ,  $S$  est un sous-ensemble de  $X \times X$  et  $L$  est un lagrangien sur  $X$ , on appellera *courbe admissible pour les données  $X, a, b, S, L$*  une application  $\xi \in W^{1, \infty}([a, b], X)$  telle que (a)  $\partial \xi \in S$ , et (b) il existe un sous-ensemble compact  $K$  de  $\text{Dom}(L)$  et un voisinage relatif  $\Omega$  de  $\{(\xi(t), t) : a \leq t \leq b\}$  dans  $X \times [a, b]$  tels que, pour presque

tout  $t \in [a, b]$ , la condition  $(x, t) \in \Omega$  implique  $(x, \dot{\xi}(t), t) \in K$ . On notera  $\mathcal{Y}_{X,a,b,S,L}$  l'ensemble de toutes les courbes admissibles pour  $X, a, b, S, L$ .

Sous des hypothèses de régularité très faibles — par exemple, si (i) l'ensemble  $\text{Dom}(L) \cap (TX \times [a, b])$  est une partie ouverte de  $TX \times [a, b]$  et (ii)  $L$  est continu ou, plus généralement,  $L \in \text{Lag}^{0,lb}(X)$  — l'intégrale

$$J_{X,a,b,S,L}(\xi) \stackrel{\text{d\u00e9f}}{=} \int_a^b L(\xi(t), \dot{\xi}(t), t) dt \quad (4)$$

existe pour tout  $\xi \in \mathcal{Y}_{X,a,b,S,L}$ . La fonctionnelle  $J_{X,a,b,S,L} : \mathcal{Y}_{X,a,b,S,L} \rightarrow \mathbb{R}$  est le *co\u00fbt associ\u00e9 au lagrangien  $L$ , l'intervalle  $[a, b]$ , et la condition au bord  $\partial\xi \in S$* .

Supposons maintenant que  $X$  soit de classe  $C^2$  et que  $L$  soit un lagrangien sur  $X$  tel que  $L \in \text{Lag}^{1,lb}(X)$ . Si  $\xi \in \mathcal{Y}_{X,a,b,S,L}$ , l'*impulsion*  $\pi_{L,\xi}$  le long de  $\xi$  par rapport \u00e0  $L$  est le champ de covecteurs

$$[a, b] \ni t \rightarrow \pi_{L,\xi}(t) \stackrel{\text{d\u00e9f}}{=} \frac{\partial L}{\partial \dot{x}}(\xi(t), \dot{\xi}(t), t) \in T_{\xi(t)}^*X.$$

(Remarquons que, sous nos hypoth\u00e8ses, la d\u00e9riv\u00e9e partielle de  $L$  par rapport \u00e0 la vitesse  $v$  en un point  $(\bar{x}, \bar{v}, \bar{t})$  est un covecteur appartenant \u00e0  $T_{\bar{x}}^*X$ . Comme on fait d'habitude dans la litt\u00e9rature de m\u00e9canique analytique et de calcul des variations, la variable « vecteur vitesse » sera souvent not\u00e9e  $\dot{x}$  plut\u00f4t que  $v$ .) L'application partielle  $[a, b] \ni t \dashrightarrow \Pi_{\xi,L}(t) \stackrel{\text{d\u00e9f}}{=} (\xi(t), \pi_{\xi,L}(t)) \in T^*X$  est le *rel\u00e8vement-impulsion de  $\xi$  par rapport \u00e0  $L$* .

L'*\u00e9quation d'Euler-Lagrange associ\u00e9e au lagrangien  $L$*  est l'\u00e9galit\u00e9 formelle

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) = \frac{\partial L}{\partial x}. \quad (5)$$

On dira qu'une courbe  $\xi : [a, b] \rightarrow X$  est une *solution de l'\u00e9quation d'Euler-Lagrange associ\u00e9e au lagrangien  $L$*  s'il existe une courbe absolument continue  $[a, b] \ni t \rightarrow \Pi(t) = (\xi(t), \pi(t)) \in T^*M$  telle que

$$\left. \begin{array}{l} \pi(t) = \pi_{\xi,L}(t) \\ \text{et } \dot{\pi}(t) = \frac{\partial L}{\partial x}(\xi(t), \dot{\xi}(t), t) \end{array} \right\} \text{ pour presque tout } t \in [a, b]. \quad (6)$$

L'interpr\u00e9tation de la deuxi\u00e8me identit\u00e9 de (6) est \u00e9vidente localement, par rapport \u00e0 un syst\u00e8me de coordonn\u00e9es. Lagrange montra que cette identit\u00e9 est en effet invariante par des changements non-lin\u00e9aires de coordonn\u00e9es, d'o\u00f9 il r\u00e9sulte que la condition (6) est en effet compl\u00e8tement intrins\u00e8que sur les vari\u00e9t\u00e9s. (Plus g\u00e9n\u00e9ralement, bien que les expressions  $\dot{\pi}(t)$  et  $\frac{\partial L}{\partial x}(\xi(t), \dot{\xi}(t), t)$ , s\u00e9par\u00e9ment, ne soient pas intrins\u00e8ques, leur diff\u00e9rence  $\dot{\pi}(t) - \frac{\partial L}{\partial x}(\xi(t), \dot{\xi}(t), t)$  est toujours un covecteur, et l'application  $t \rightarrow \dot{\pi}(t) - \frac{\partial L}{\partial x}(\xi(t), \dot{\xi}(t), t)$  est un champ de covecteurs le long de  $\xi$ , qu'on appelle le *champ d'Euler-Lagrange* de  $\xi$ .)

**Théorème 3.2.1** Soit  $L$  un lagrangien sur une variété différentiable  $X$  de classe  $C^2$ . Soient  $a, b \in \mathbb{R}$  tels que  $a < b$ . Supposons que  $\text{Dom}(L) \cap (TX \times [a, b])$  soit ouvert dans  $TX \times [a, b]$ ,  $L \in \text{Lag}^{1,lb}(X)$ , et que  $S$  soit une sous-variété de  $X \times X$  de classe  $C^1$ . Notons  $\mathcal{Y} = \mathcal{Y}_{X,a,b,S,L}$ ,  $J = J_{X,a,b,S,L}$ . Si  $\xi_* \in \mathcal{Y}$  est tel que  $J(\xi_*) \leq J(\xi)$  quel que soit  $\xi \in \mathcal{Y}$ , alors  $\xi_*$  est une solution de l'équation d'Euler-Lagrange associée à  $L$  vérifiant la **condition de transversalité**

$$(-\pi_{\xi_*,L}(a), \pi_{\xi_*,L}(b)) \in (T_{\partial\xi_*}S)^\perp. \quad (7)$$

Remarquons que le théorème 3.2.1 contient une *propriété de régularité très forte* pour les solutions du problème de minimisation. En effet, le problème étant posé, en principe, dans l'espace  $W^{1,\infty}$  des courbes Lipschitziennes, le fait qu'une courbe  $\xi_*$  est une solution implique que  $\dot{\xi}_*(t)$  existe presque partout. Par conséquent, l'impulsion  $\pi_{\xi_*,L}(t)$  est définie presque partout, mais cela ne suffit pas pour conclure que  $\pi_{\xi_*,L}$  admet un prolongement qui est un champ de covecteurs absolument continu —ou même continu— le long de  $\xi$ . L'équation d'Euler-Lagrange dit, d'abord, que ce prolongement existe, et puis que sa dérivée obéit la deuxième identité de (6). Dans le cas où  $\xi_* \in C^1$  et  $\frac{\partial L}{\partial \dot{x}}$  est continue par rapport à  $x$ ,  $\dot{x}$  et  $t$ , la définition de  $\pi_{\xi_*,L}$  implique directement que  $\pi_{\xi_*,L}$  est continu, mais pour en déduire que  $\pi_{\xi_*,L}$  est absolument continu on aurait besoin d'hypothèses de régularité plus fortes. (Par exemple, il faudrait supposer que  $\xi_*$  appartienne à  $W^{2,\infty}$  et que  $\frac{\partial L}{\partial \dot{x}}$  soit de classe  $C^1$  par rapport à  $x$ ,  $\dot{x}$  et  $t$ .) Mais, même sous des hypothèses très fortes sur  $L$  —par exemple,  $L \in C^\infty(TX \times \mathbb{R}, \mathbb{R})$ — on ne pourra jamais démontrer la continuité de l'impulsion à moins qu'on suppose que  $\xi_* \in C^1$ . Or, l'hypothèse que  $\xi_* \in C^1$  n'est pas naturelle, parce que les théorèmes d'existence de solutions demandent typiquement qu'on se situe dans un espace fonctionnel tel que  $W^{1,\infty}$ , où la famille des sous-ensembles compacts est suffisamment riche. Dans des cas très importants —par exemple, pour les géodésiques riemanniennes— on démontre l'existence de façon naturelle dans un espace du type  $W^{1,\infty}$ . L'équation d'Euler-Lagrange implique alors que l'impulsion est absolument continue, ce qui permet de démontrer (en vertu de la formule bien connue pour « descendre les indices ») que la vitesse  $\dot{\xi}_*$  est elle-même absolument continue. Il en résulte —en utilisant de nouveau l'équation d'Euler-Lagrange— que  $\xi_*$  est la solution de l'équation différentielle ordinaire du second ordre dite « équation des géodésiques », ce qui permet de déduire des propriétés de régularité de  $\xi_*$  encore plus fortes.

### 3.2.2 Les deux formalismes hamiltoniens

Définissons d'abord une fonction  $(x, v, p, t) \mapsto H(q, v, p, t)$  de quatre groupes de variables : la position  $x$ , la vitesse  $v$ , l'impulsion  $p$ , et le temps  $t$ . (Plus précisément,  $H$  est une fonction partielle sur  $\hat{T}X \times \mathbb{R}$ , où  $\hat{T}X$  est le produit fibré de  $TX$  et  $T^*X$ , c'est-à-dire le fibré sur  $X$  tel que la fibre  $\hat{T}_x X$  de chaque point  $x \in X$  est le produit  $T_x X \times T_x^* X$ .) On définit  $H$  par l'identité

$$H(x, v, p, t) = \langle p, v \rangle - L(x, v, t), \quad (8)$$

où  $(v, p) \rightarrow \langle p, v \rangle$  est l'application de dualité de  $T_x X \times T_x^* X$  dans  $\mathbb{R}$ .

Il s'ensuit que  $\frac{\partial H}{\partial p} = v$ . Ainsi donc, l'identité

$$\dot{\xi}_*(t) = \frac{\partial H}{\partial p}(\xi_*(t), \dot{\xi}_*(t), \pi(t), t) \quad \text{pour presque tout } t \in \text{Dom}(\xi_*) \quad (9)$$

est vérifiée le long de n'importe quelle courbe absolument continue  $\xi_*$ , quel que soit le champ de covecteurs  $\pi$  le long de  $\xi_*$ .

D'ailleurs, en supposant, pour l'instant, qu'on ait fixé des coordonnées, il résulte de l'identité  $\frac{\partial H}{\partial x} = -\frac{\partial L}{\partial x}$  que la condition d'Euler-Lagrange (6) équivaut à

$$\dot{\pi}(t) = -\frac{\partial H}{\partial x}(\xi_*(t), \dot{\xi}_*(t), \pi(t), t) \quad \text{pour presque tout } t \in \text{Dom}(\xi_*). \quad (10)$$

Finalement, la formule  $\frac{\partial H}{\partial v} = p - \frac{\partial L}{\partial v} = p - \frac{\partial L}{\partial \dot{x}}$  montre que la définition de l'impulsion  $\pi_{\xi_*, L}$  équivaut — si  $\pi = \pi_{\xi_*, L}$  — à l'égalité :

$$\frac{\partial H}{\partial v}(\xi_*(t), \dot{\xi}_*(t), \pi(t), t) = 0, \quad (11)$$

Le système de trois équations (9), (10), (11), qui peut être écrit sous la forme

$$\frac{dx}{dt} = \frac{\partial H}{\partial p}, \quad \frac{dp}{dt} = -\frac{\partial H}{\partial x}, \quad \frac{\partial H}{\partial v} = 0, \quad (12)$$

est donc *exactement équivalent* — sans qu'on ait besoin d'aucune restriction additionnelle — à l'équation d'Euler-Lagrange.

Nous appellerons la fonction  $H$  le « vrai hamiltonien »<sup>6</sup>, et nous tenons à souligner tout de suite que *ce hamiltonien n'est pas le hamiltonien classique*, qui sera noté  $\mathcal{H}$  ci-dessous, et que *la forme (12)* — c'est-à-dire le système d'équations (9), (10), (11) — *des « équations de Hamilton » n'est pas le « système de Hamilton » classique.*

Le *hamiltonien classique* est la fonction partielle

$$T^* X \times \mathbb{R} \ni (x, p, t) \mapsto \mathcal{H}(x, p, t) \in \mathbb{R}$$

— qui est « une fonction de  $x$ ,  $p$  et  $t$  » — définie par

$$\mathcal{H}(x, p, t) = \langle p, \dot{x} \rangle - L(x, \dot{x}, t). \quad (13)$$

<sup>6</sup>Notre terminologie est choisie exprès pour signaler notre très profond désaccord avec le point de vue de plusieurs auteurs qui appellent  $H$  le « pseudo-hamiltonien », ou le « quasi-hamiltonien », laissant entendre que le hamiltonien classique est un objet mathématiquement plus naturel et le seul qui mériterait le nom de « vrai » hamiltonien. Pour nous,  $H$  est clairement supérieur, parce que (a) on peut toujours définir  $\mathcal{H}$  à partir de  $H$  — lorsque  $\mathcal{H}$  existe — mais il n'est pas possible de définir  $\mathcal{H}$  à partir de  $H$ , (b) le passage de  $H$  à  $\mathcal{H}$  est très souvent accompagné d'une perte d'information, et (c) parfois — par exemple, dans le cas très important de la dynamique d'une particule dans la relativité restreinte — ce passage n'est pas possible du tout. Nous donnerons plus tard d'autres arguments en faveur de  $H$ , cf. la remarque 3.2.3.

Dans cette identité, il est entendu que  $\dot{x}$  est en réalité une fonction de  $(x, p, t)$ , déterminée implicitement par la formule

$$p = \frac{\partial L}{\partial \dot{x}}(x, \dot{x}, t). \quad (14)$$

Le *système hamiltonien classique* est le système de deux équations

$$\frac{dx}{dt} = \frac{\partial \mathcal{H}}{\partial p}, \quad \frac{dp}{dt} = -\frac{\partial \mathcal{H}}{\partial x}, \quad (15)$$

qu'il ne faut pas confondre avec (12), malgré la ressemblance superficielle des deux expressions. On voit alors que *si l'application*  $(x, \dot{x}, t) \rightarrow (x, p, t)$  *définie par (14) est inversible, c'est-à-dire s'il est possible de « résoudre (14) pour trouver la fonction*  $\dot{x} = \dot{x}(x, p, t)$  *», alors (15) équivaut à (12). Mais lorsque l'application*  $\dot{x} \rightarrow \frac{\partial L}{\partial \dot{x}}(x, \dot{x}, t)$  *n'est pas inversible pour*  $(x, t)$  *fixé, la reformulation hamiltonienne classique de l'équation d'Euler-Lagrange n'est pas possible, tandis que la « vraie » reformulation hamiltonienne existe toujours, et équivaut toujours à (12).*

### 3.2.3 La condition de Weierstrass

Soit  $\tilde{TX}$  le produit fibré de  $TX$  avec  $TX$ , c'est-à-dire le fibré sur  $X$  tel que la fibre d'un point  $x \in X$  est le produit  $T_x X \times T_x X$ . Définissons une fonction partielle  $\mathcal{E} : \tilde{TX} \times \mathbb{R} \dashrightarrow \mathbb{R}$  par

$$\mathcal{E}(x, v, w, t) = L(x, w, t) - L(x, v, t) - \left\langle \frac{\partial L}{\partial \dot{x}}(x, v, t), w - v \right\rangle. \quad (16)$$

Weierstrass démontra essentiellement<sup>7</sup> le résultat suivant :

**Theorème 3.2.2** *Soient*  $X, a, b, L, S, \xi_*$  *des données vérifiant les hypothèses du théorème 3.2.1. Définissons*

$$V_{\xi_*, L}(t) \stackrel{\text{def}}{=} \left\{ w \in T_{\xi_*(t)} X : (\xi(t), w, t) \in \text{Dom}(L) \right\} \quad \text{pour } t \in \text{Dom}(\xi_*). \quad (17)$$

*On a alors la condition supplémentaire de Weierstrass*

$$\min \left\{ \mathcal{E}(\xi_*(t), \dot{\xi}_*(t), w, t) : w \in V_{\xi_*, L}(t) \right\} = 0 \quad (18)$$

*pour presque tout*  $t \in [a, b]$ . ◇

---

<sup>7</sup>Sans utiliser, bien sûr, la notion d'égalité « presque partout », qui n'avait pas encore été inventée. Il travailla avec des fonctions continument différentiables, ce qui rendait ses conclusions vraies partout.

### 3.2.4 La réécriture hamiltonienne de la condition de Weierstrass

En posant  $x = \xi_*(t)$  et  $v = \dot{\xi}_*(t)$  dans (16), et en faisant intervenir l'identité

$$\pi_{\xi_*,L}(t) = \frac{\partial L}{\partial \dot{x}}(\xi_*(t), \dot{\xi}_*(t), t), \quad (19)$$

or arrive à l'expression suivante pour la fonction  $\mathcal{E}$  de Weierstrass :

$$\begin{aligned} \mathcal{E}(\xi_*(t), \dot{\xi}_*(t), w, t) &= L(\xi_*(t), w, t) - L(\xi_*(t), \dot{\xi}_*(t), t) - \langle \pi_{\xi_*,L}(t), w - \dot{\xi}_*(t) \rangle \\ &= H(\xi_*(t), \dot{\xi}_*(t), \pi_{\xi_*,L}(t), t) - H(\xi_*(t), w, \pi_{\xi_*,L}(t), t). \end{aligned}$$

On peut donc récrire (18) sous la *forme hamiltonienne*

$$H(\xi_*(t), \dot{\xi}_*(t), \pi(t), t) = \max_{w \in V_{\xi_*,L}(t)} H(\xi_*(t), w, \pi(t), t) \text{ pour p.t. } t \in [a, b], \quad (20)$$

où  $\pi \equiv \pi_{\xi_*,L}$ .

**Remarque 3.2.3** *C'est grâce à l'introduction du « vrai » hamiltonien  $H$  que nous avons réussi à récrire la condition de Weierstrass sous la forme hamiltonienne (20). Une telle réécriture n'aurait guère été possible avec le hamiltonien classique  $\mathcal{H}$ . En effet, pour définir la fonction  $\mathcal{E}$  il est absolument nécessaire de comparer le comportement du lagrangien au voisinage du vecteur vitesse  $\dot{\xi}_*(t)$  le long de la courbe  $\xi_*$  avec le comportement au voisinage d'une valeur arbitraire  $w$  de la vitesse. Or, dans la définition de  $\mathcal{H}$  la vitesse est éliminée, ce qui rend *a priori* impossible l'utilisation de  $\mathcal{H}$  pour la comparaison.  $\diamond$*

### 3.2.5 La libération de l'impulsion et la disparition de la dérivée du lagrangien par rapport à la vitesse

Jusqu'ici, l'impulsion  $\pi$  associée à la courbe  $\xi_*$  et le lagrangien  $L$  était, par définition, égale à  $\pi_{\xi_*,L}$ . (Plus précisément,  $\pi$  le seul champ de covecteurs le long de  $\xi_*$  qui est continu et se confond avec  $\pi_{\xi_*,L}$  presque partout.) Or, la formule (19) définissant l'impulsion équivaut à la condition (11), c'est-à-dire à la troisième équation de (12). Mais, ***sous les hypothèses de l'énoncé du théorème 3.2.1, (11) est une conséquence immédiate de la condition (20) sur la maximisation du hamiltonien.*** Il est donc possible d'énoncer les résultats des théorèmes 3.2.1 et 3.2.2 de façon unifiée sous forme hamiltonienne, sans faire appel à la définition (19) de l'impulsion :

**Theorème 3.2.4** *Soit  $L$  un lagrangien sur une variété différentiable  $X$  de classe  $C^2$ , tel que  $L \in \text{Lag}^{1,lb}(X)$ . Supposons que  $a, b \in \mathbb{R}$  et  $a < b$ , que  $\text{Dom}(L) \cap (TX \times [a, b])$  soit ouvert dans  $TX \times [a, b]$ , et que  $S$  soit une sous-variété de  $X \times X$  de classe  $C^1$ . Notons  $\mathcal{Y} = \mathcal{Y}_{X,a,b,S,L}$ ,  $J = J_{X,a,b,S,L}$ . Soit  $\xi_* \in \mathcal{Y}$  tel que  $J(\xi_*) \leq J(\xi)$  quel que soit  $\xi \in \mathcal{Y}$ . Il existe alors un champ de covecteurs absolument continu  $\pi$  le long de  $\xi_*$ , vérifiant : (a) les **équations de Hamilton** (9) et (10), (b) la condition (20) de **maximisation du hamiltonien** (où  $V_{\xi_*,L}(t)$  est défini par (17)), et (c) la **condition de transversalité** (7).  $\diamond$*

Nous venons d'expliquer que, si nous gardons les hypothèses du théorème 3.2.4, on peut se débarrasser de l'identité

$$\pi \equiv \pi_{\xi_*, L} \quad (21)$$

parce qu'elle est de toute façon une conséquence de (20). Examinons maintenant la situation de plus près, faisant attention au rôle précis de nos hypothèses dans le passage de (20) à (21) ou à sa version équivalente (11). Ce passage fut possible en vertu de deux aspects de nos hypothèses. Premièrement, la condition que  $\text{Dom}(L) \cap (TX \times [a, b])$  est ouvert dans  $TX \times [a, b]$  fut utilisée pour déduire que, pour presque tout  $t \in [a, b]$ , l'ensemble  $V_{\xi_*, L}(t)$  défini par (17) est un voisinage de  $\dot{\xi}_*(t)$ . Deuxièmement, on montra que le hamiltonien  $H$  est différentiable par rapport à la vitesse, en vertu de (8) et de l'hypothèse de différentiabilité du lagrangien  $L$  par rapport à la vitesse. Le fait que la fonction  $T_{\xi_*(t)}X \ni v \rightarrow H(\xi_*(t), v, \pi(t), t)$  est différentiable en  $\dot{\xi}_*(t)$  et définie sur un voisinage de ce point nous a permis de déduire l'égalité (21) à partir de (20).

Notons, d'autre part, que dans l'énoncé du théorème 3.2.4 la seule référence à la différentiabilité du lagrangien par rapport à la vitesse se trouve dans l'hypothèse que  $L$  appartient à  $\text{Lag}^{1, lb}(X)$ , et que la dérivée partielle  $\frac{\partial L}{\partial \dot{x}}$ , qui semblait jouer un rôle fondamental dans l'équation d'Euler-Lagrange, a disparu complètement de la scène.

Il est donc naturel de chercher un énoncé plus général, où les deux hypothèses qui ont rendu possible le passage de (20) à (21) seraient remplacées par des conditions plus faibles. Bien entendu, dans cette nouvelle version l'impulsion  $\pi$  serait complètement libérée de l'obligation d'être égale à  $\pi_{\xi_*, L}$  (qui, d'ailleurs, pourrait ne pas exister), et deviendrait un champ de covecteurs le long de  $\xi_*$  vérifiant les équations de Hamilton et la maximisation du hamiltonien.

Plaçons-nous, pour l'instant, dans le cadre plus restrictif où

(H1) *L'espace des configurations  $X$  est une partie ouverte d'un espace euclidien  $\mathbb{R}^n$ ,  $a, b \in \mathbb{R}$ ,  $a < b$ , et  $L$  est un lagrangien sur  $X$ .*

Le produit  $X \times \mathbb{R}^n$  peut donc être identifié au fibré tangent  $TX$ .

Pour que les équations de Hamilton possèdent un sens précis, et qu'elles soient associées à un « bon » champ de vecteurs hamiltonien non autonome sur  $T^*X$ , on supposera que les conditions suivantes soient vérifiées :

(H2) *L'ensemble  $\{(x, t) \in X \times [a, b] : (x, v, t) \in \text{Dom}(L)\}$  est une partie ouverte de  $X \times [a, b]$  quel que soit  $v \in \mathbb{R}^n$ .*

(H3) *La fonction  $X \ni x \dashrightarrow L(x, v, t)$  — dont le domaine est ouvert dans  $X$  en vertu de (H2) — est de classe  $C^1$  quel que soit  $(v, t) \in \mathbb{R}^n \times [a, b]$ .*

(H4) *La fonction  $TX \ni (x, v) \dashrightarrow L(x, v, t) \in \mathbb{R}$  est continue quel que soit  $t \in [a, b]$ .*

(H5) *La fonction  $[a, b] \ni t \dashrightarrow L(x, v, t) \in \mathbb{R}$  est mesurable quel que soit le point  $(x, v) \in TX$ .*

(H6) Les fonctions  $L$  et  $\frac{\partial L}{\partial x}$  sont localement bornées sur  $X \times V \times [a, b]$ .

Rappelons que, par définition, si  $\xi \in \mathcal{Y}_{X,a,b,S,L}$ , il existe un sous-ensemble négligeable  $N$  de  $[a, b]$  et un sous-ensemble compact  $K$  de  $\text{Dom}(L)$  tels que  $\dot{\xi}(t)$  existe et  $(\xi(t), \dot{\xi}(t), t) \in K$  quel que soit  $t \in [a, b] \setminus N$ . En vertu des hypothèses (H3)-(H6), la fonction  $t \rightarrow L(\xi(t), \dot{\xi}(t), t)$  est mesurable et bornée. Ainsi donc, l'intégrale de (4) existe quel que soit  $\xi \in \mathcal{Y}_{X,a,b,S,L}$ .

La conjecture naturelle dit donc que l'énoncé du théorème 3.2.4 reste vrai sous ces hypothèses plus faibles. Cette conjecture s'avère presque correcte mais, pour arriver à un résultat rigoureusement vrai, on est obligé d'abord de tenir compte d'un nouveau phénomène, étudié déjà par Bolza en 1913, cf. [1].

### 3.3 De Weierstrass au principe du maximum de la théorie du contrôle

#### 3.3.1 Le multiplicateur anormal

Revenons à l'exemple 3.1.4. Dans notre traitement du problème de la caténaire, nous avons souligné l'existence du « cas singulier », correspondant à la situation où la longueur de la corde est exactement égale à la distance entre ses deux extrémités. Dans ce cas il n'y a évidemment qu'une configuration  $\bar{y}$  possible — donnée par la formule (3) — qui est donc nécessairement optimale. Nous allons montrer que *ce problème s'inscrit dans le cadre de la « conjecture naturelle »* de la section précédente mais que, malheureusement, *la conjecture n'est pas vraie dans le cas singulier*. (Un calcul assez long dont nous ne donnerons pas les détails permet de montrer que la conjecture est vraie dans le cas non singulier.)

Reprenons les notations de l'exemple 3.1.4. Définissons, pour chaque courbe  $y(\cdot) \in \mathcal{Y}$ , une nouvelle fonction  $s : [a, b] \rightarrow \mathbb{R}$  — la longueur d'arc — par la formule  $s(x) = \int_a^x \sqrt{1 + y'(r)^2} dr$ . Il existe alors une bijection naturelle entre l'espace  $\mathcal{Y}$  des fonctions  $y(\cdot)$  et l'ensemble  $\Xi$  des courbes absolument continues  $[a, b] \ni x \rightarrow \xi(x) = (y(x), s(x)) \in \mathbb{R}^2$  dans  $\mathbb{R}^2$  vérifiant  $\xi(a) = (\alpha, 0)$ ,  $\xi(b) = (\beta, \Lambda)$ , et  $\xi'(x) \in V$  pour presque tout  $x \in [a, b]$ , où  $V = \{(u, v) \in \mathbb{R}^2 : v = \sqrt{1 + u^2}\}$ .

Ainsi donc, nous avons un problème aux variations vérifiant les hypothèses (H1)-(H6), dont le lagrangien est donné par

$$\begin{aligned} \text{Dom}(L) &= \mathbb{R}^2 \times V \times \mathbb{R}, \\ L(y, s, u, v, t) &= yv \quad \text{pour } (y, s) \in \mathbb{R}^2, (u, v) \in V, t \in \mathbb{R}. \end{aligned}$$

Si la conjecture était vraie, il en résulterait que la solution  $y$  est donnée par la formule — énoncée sans démonstration au §3.1 —

$$y(x) = \gamma + \kappa \cosh\left(\frac{x - x_0}{\kappa}\right), \quad (22)$$

où  $\gamma$ ,  $x_0$ ,  $\kappa$  sont des constantes réelles, et  $\kappa > 0$ . Nous avons montré au §3.1 que les constantes  $\gamma$ ,  $x_0$ , et  $\kappa$ , sont déterminées de façon unique — par des formules explicites — par les paramètres  $a$ ,  $b$ ,  $\alpha$ ,  $\beta$ ,  $\Lambda$ .

Il est clair que la fonction  $y(\cdot)$  définie par (22) n'est jamais linéaire. Comme on a vu au §3.1, la *limite* de cette fonction lorsque  $\Lambda \downarrow \Delta$  est la fonction linéaire  $\bar{y}$  qui résout le problème singulier, mais cette limite est obtenue en faisant  $\kappa$  tendre vers  $+\infty$ . La fonction  $\bar{y}$  n'appartient pas à la famille de fonctions à trois paramètres définie par (22). Or, cette famille contient *toutes* les courbes vérifiant la conclusion de la conjecture du §3.2.5. Il s'ensuit que la conjecture doit être abandonnée ou modifiée.

En fait, il suffit de modifier la conjecture en multipliant  $L$  par une nouvelle variable scalaire  $p_0$  dans la définition du hamiltonien. Il en résulte la formule

$$H(x, v, p, p_0, t) = \langle p, v \rangle - p_0 L(x, v, t). \quad (23)$$

Les *équations de Hamilton* s'écrivent maintenant sous la forme

$$\dot{\xi}_*(t) = \frac{\partial H}{\partial p}(\xi_*(t), \dot{\xi}_*(t), \pi(t), \pi_0, t) \quad (24)$$

$$\dot{\pi}(t) = -\frac{\partial H}{\partial x}(\xi_*(t), \dot{\xi}_*(t), \pi(t), \pi_0, t) \quad (25)$$

et la condition de *maximisation du hamiltonien* devient

$$H(\xi_*(t), \dot{\xi}_*(t), \pi(t), \pi_0, t) = \max_{v \in V_{\xi_*, L}(t)} H(\xi_*(t), v, \pi(t), \pi_0, t), \quad (26)$$

$\pi_0$  étant un nombre réel non négatif, appelé le « multiplicateur anormal ».

Avec cette modification, la généralisation que nous avons proposée du théorème 3.2.4 devient vraie. Appelons *champ de covecteur positivement augmenté le long d'une courbe*  $\xi$  un couple  $(\pi, \pi_0)$  tel que  $\pi$  est un champ de covecteurs le long de  $\xi$  et  $\pi_0$  est un nombre réel tel que  $\pi_0 \geq 0$ . On a alors le résultat suivant :

**Théorème 3.3.1** *Supposons que les hypothèses (H1)-(H6) soient vérifiées. Soit  $S$  une sous-variété de classe  $C^1$  du produit  $X \times X$ . Notons  $\mathcal{Y} = \mathcal{Y}_{X,a,b,S,L}$ ,  $J = J_{X,a,b,S,L}$ . Soit  $\xi_* \in \mathcal{Y}$  tel que  $J(\xi_*) \leq J(\xi)$  quel que soit  $\xi \in \mathcal{Y}$ . Il existe alors un champ de covecteurs  $(\pi, \pi_0)$  le long de  $\xi_*$ , absolument continu et positivement augmenté, tel que*

$$(a) \quad \|\pi(b)\| + \pi_0 > 0,$$

$$(b) \quad (-\pi(a), \pi(b)) \in (T_{\partial\xi_*} S)^\perp.$$

(c) *les équations de Hamilton (24), (25) et la condition (26) de maximisation du hamiltonien sont vérifiées presque partout.*  $\diamond$

**Remarque 3.3.2** Le cœur de l'énoncé du théorème 3.3.1 se trouve dans la condition (a), de *non-trivialité*. En effet, si  $\xi_*$  est une courbe arbitraire — pas nécessairement optimale — on pourra toujours satisfaire aux autres conclusions en choisissant  $\pi(t) \equiv 0$ ,  $\pi_0 = 0$ . Le contenu du théorème est justement que, si  $\xi_*$  est optimale, il doit y exister au moins un couple  $(\pi, \pi_0)$  autre que  $(0, 0)$ .  $\diamond$

On appellera *multiplicateur* (pour  $\xi_*$ , par rapport à  $L$  et  $S$ ) un champ de covecteurs absolument continu positivement augmenté  $(\pi, \pi_0)$  le long de  $\xi_*$  vérifiant les conditions (a), (b), (c) de l'énoncé du théorème 3.3.1. Le multiplicateur  $(\pi, \pi_0)$  est *normal* si  $\pi_0 > 0$ , et *anormal* si  $\pi_0 = 0$ .

Une courbe  $\xi_*$  est une *extrémale* (par rapport à  $L$  et  $S$ ) si elle admet un multiplicateur. Le théorème 3.3.1 dit donc que *si une courbe est optimale alors elle est une extrémale*.

L'extrémale  $\xi_*$  est *normale* (resp. *anormale*) si elle admet un multiplicateur normal (resp. anormal). On remarquera qu'une courbe  $\xi_*$  peut bien être à la fois une extrémale normale et une extrémale anormale.

Il est évident que, si  $(\pi, \pi_0)$  est un multiplicateur et  $r > 0$ , le couple  $(r\pi, r\pi_0)$  est aussi un multiplicateur. Ainsi donc, on pourra toujours supposer que le multiplicateur anormal  $\pi_0$  appartienne à  $\{0, 1\}$ . Le cas normal —c'est-à-dire  $\pi_0 = 1$ — est exactement celui des théorèmes 3.2.1 et 3.2.4. Démontrons la

**Proposition 3.3.3** *Sous les hypothèses du théorème 3.3.1, si  $V$  est une partie ouverte de  $\mathbb{R}^n$  —ou, plus généralement, si l'ensemble  $\{t : \dot{\xi}_*(t) \in \text{Int}(V)\}$  est de mesure positive— on a  $\pi_0 > 0$ , et on peut donc choisir  $\pi_0 = 1$ .*

PREUVE. Si  $\pi_0 = 0$ , on déduit de (a) que  $\pi(b) \neq 0$ . L'équation de Hamilton (25) devient  $\dot{\pi}(t) = 0$ . On a donc  $\pi(t) \neq 0$  quel que soit  $t$ . La condition (26) dit que la fonction linéaire  $V \ni v \rightarrow \langle \pi(t), v \rangle$  est maximisée par  $v = \dot{\xi}_*(t)$ , ce qui est clairement impossible si  $\dot{\xi}_*(t) \in \text{Int}(V)$ .  $\diamond$

Pour terminer, revenons à l'exemple de la caténaire, et montrons que *les courbes qui vérifient la conclusion du théorème 3.3.1 sont toutes les solutions, y compris celles du problème singulier, qui correspondent exactement au cas anormal*.

Le hamiltonien  $H$  est donné par

$$H(y, s, u, v, p_y, p_s, p_0) = p_y u + p_s v - p_0 y v,$$

où  $p_y, p_s$  sont les composantes du covecteur impulsion conjuguées aux variables  $y, s$ , et  $p_0$  est le multiplicateur anormal. (Nous écrivons  $H(y, s, u, v, p_y, p_s, p_0)$  au lieu de  $H(y, s, u, v, p_y, p_s, p_0, t)$  parce que notre problème est autonome, et  $H$  est indépendant du temps.) Dans le cas normal on peut choisir  $\pi_0 = 1$ , et cela donne les solutions pour le cas non singulier.

Considérons maintenant le cas anormal, où  $\pi_0 = 0$ . Le hamiltonien anormal —c'est-à-dire, la fonction  $(y, s, u, v, p_y, p_s) \rightarrow H(y, s, u, v, p_y, p_s, 0)$ — est égal à  $p_y u + p_s v$ . Les équations de Hamilton s'écrivent  $\dot{y} = u, \dot{s} = v, \dot{\pi}_y = 0, \dot{\pi}_s = 0$ , d'où nous déduisons que le covecteur impulsion  $(\pi_y, \pi_s)$  est constant. En vertu de la condition de nontrivialité (a), ce covecteur n'est pas nul. La condition de maximisation du hamiltonien se réduit à

$$\pi_y \dot{y}(x) + \pi_s \dot{s}(x) = \max\{\pi_y u + \pi_s v : (u, v) \in V\}. \quad (27)$$

Il est immédiat à partir de la définition de  $V$  que, si  $\lambda$  est une fonctionnelle linéaire sur  $\mathbb{R}^2$  telle que  $\lambda \neq 0$ , alors une des deux possibilités suivantes doit avoir

lieu : ou bien  $\sup\{\lambda(w) : w \in V\} = +\infty$ , ou bien il existe un point  $\bar{w} \in V$ , qui est d'ailleurs unique, tel que  $\sup\{\lambda(w) : w \in V\} = \lambda(\bar{w})$ . Notons  $w_\lambda$  le point  $\bar{w}$ , qui dépend de  $\lambda$ . En prenant  $\lambda = (\pi_y, \pi_s)$ , et en utilisant le fait que nous savons déjà que  $(\pi_y, \pi_s) \neq (0, 0)$ , on déduit que le cas où  $\sup\{\lambda(w) : w \in V\} = +\infty$  n'est pas possible, car (27) doit être vérifiée presque partout. Il en découle que  $(\dot{y}(x), \dot{s}(x)) = w_{(\pi_y, \pi_s)}$  presque partout. Nous avons donc établi que la dérivée  $\dot{y}(\cdot)$  est une fonction constante, ce qui montre que  $y(\cdot)$  est linéaire, et achève notre démonstration.

### 3.3.2 Deux exemples

Montrons la supériorité du théorème 3.3.1 sur le théorème 3.2.4, en étudiant deux exemples de problèmes de minimisation assez simples, où le calcul des variations classique ne suffit pas pour trouver toutes les solutions. Ces deux exemples correspondent exactement aux deux hypothèses qui ont joué un rôle crucial dans l'identification de l'impulsion  $\pi$  avec le champs de covecteurs  $\pi_{\xi, L}$ .

Le problème de notre premier exemple s'inscrit dans le cadre du théorème 3.3.1, mais pas dans celui des autres théorèmes, car le lagrangien  $L$  n'est pas différentiable par rapport à la vitesse.

**Exemple 3.3.4** Soit  $n \in \mathbb{N}$ , et soit  $\|\cdot\|$  une norme sur  $\mathbb{R}^n$ . Deux points  $A, B$  de  $\mathbb{R}^n$  étant fixés, on cherche une courbe lipschitzienne  $[0, 1] \ni t \rightarrow \xi(t) \in \mathbb{R}^n$  telle que

$$\xi(0) = A \quad \text{et} \quad \xi(1) = B, \quad (28)$$

qui rend l'intégrale  $\int_0^1 \|\dot{\xi}(t)\| dt$  minimale dans l'ensemble de toutes les courbes  $\xi(\cdot) \in W^{1, \infty}([0, 1], \mathbb{R}^n)$  vérifiant (28).

Pour appliquer le théorème 3.3.1, remarquons d'abord que  $V = \mathbb{R}^n$ . En vertu de la proposition 3.3.3, on est dans le cas normal. Le hamiltonien  $H$  est donné par  $H(x, v, p, p_0, t) = \langle p, v \rangle - p_0 \|v\|$ , et les équations de Hamilton pour un multiplicateur normal  $(\pi, 1)$  deviennent  $\dot{\xi} = v$ ,  $\dot{\pi} = 0$ . Il en résulte que l'impulsion  $\pi(t)$  est égale à un covecteur constant  $\bar{\pi}$ . Comme la fonction  $\mathbb{R}^n \ni v \rightarrow \langle p, v \rangle - \|v\|$  n'est pas bornée supérieurement si  $\|p\| > 1$ , on déduit de la condition de maximisation du hamiltonien que  $\|\bar{\pi}\| \leq 1$ . Si  $\|\bar{\pi}\| < 1$ , on a l'égalité  $\dot{\xi}_*(t) \equiv 0$ , correspondant aux *courbes constantes*, qui sont évidemment des solutions du problème de minimisation, pour  $A = B$ .

Il reste le cas d'une impulsion  $\bar{\pi}$  telle que  $\|\bar{\pi}\| = 1$ . Dans ce cas, l'ensemble  $M(\bar{\pi}) \stackrel{\text{déf}}{=} \{v \in \mathbb{R}^n : \langle \bar{\pi}, v \rangle = \|v\|\}$  est un cône convexe fermé non vide. Soit  $\Xi(\bar{\pi})$  l'ensemble de toutes les courbes lipschitziennes  $\xi : [0, 1] \rightarrow \mathbb{R}^n$  telles que  $\dot{\xi}(t) \in M(\bar{\pi})$  pour presque tout  $t \in [0, 1]$ . On déduit du théorème 3.3.1 que toutes les courbes optimales appartiennent à l'ensemble

$$\Xi \stackrel{\text{déf}}{=} \bigcup \{\Xi(\bar{\pi}) : \|\bar{\pi}\| = 1\}.$$

En outre, toutes les courbes  $\xi \in \Xi$  sont évidemment optimales. (Preuve :

supposons que  $\xi \in \Xi(\bar{\pi})$  et  $\|\bar{\pi}\| = 1$ . Soient  $A = \xi(0)$ ,  $B = \xi(1)$ . Il vient

$$\int_0^1 \|\dot{\xi}(t)\| dt = \int_0^1 \langle \bar{\pi}, \dot{\xi}(t) \rangle dt = \left\langle \bar{\pi}, \int_0^1 \dot{\xi}(t) dt \right\rangle = \langle \bar{\pi}, B - A \rangle \leq \|B - A\|,$$

ce qui établit l'optimalité de  $\xi$ .) Il en résulte que  $\Xi$  est exactement l'ensemble de toutes les solutions.

Il est clair que  $\Xi$  contient tous les segments de droite. Dans le cas d'une norme strictement convexe, les  $M(\bar{\pi})$  sont des demi-droites, et l'ensemble des chemins optimaux est exactement celui de tous les segments de droite (paramétrisés de façon arbitraire). Par contre, lorsqu'il s'agit d'une norme complètement générale, il y aura d'autres solutions. (Par exemple, soit  $n = 2$ , et choisissons la norme définie par  $\|(x_1, x_2)\| = |x_1| + |x_2|$ . Alors le chemin  $\xi : [0, 2] \rightarrow \mathbb{R}^2$  donné par  $\xi(t) = (t, 0)$  si  $0 \leq t \leq 1$  et  $\xi(t) = (1, t - 1)$  si  $1 \leq t \leq 2$  est optimal.)

**Le théorème 3.3.1 nous a permis de trouver toutes les solutions, ce qui n'aurait pas été possible avec le calcul des variations classique.**

Remarquons que, même dans le cas de la norme euclidienne, le lagrangien de ce problème n'est pas différentiable partout. Ainsi donc, le calcul des variations classique ne suffit pas à lui seul pour trouver toutes les solutions. Il s'ensuit que même la question du plus court chemin entre deux points donnés, qui est sans doute le plus ancien, le plus célèbre et le plus simple de tous les problèmes aux variations, fournit déjà un exemple d'un problème variationnel dont la solution complète n'est pas dérivable à partir du calcul de variations classique, mais qui s'encadre de façon naturelle dans la théorie plus moderne et plus générale du contrôle optimal.  $\diamond$

Dans notre second exemple, le lagrangien est lisse, mais le problème ne s'inscrit pas dans le cadre variationnel classique parce que l'ensemble des valeurs de la vitesse n'est pas ouvert. L'équation d'Euler-Lagrange n'est pas vérifiée, ce qui n'est pas contradictoire, puisque ce problème ne vérifie pas les hypothèses du théorème 3.2.2. Par contre, l'exemple vérifie les hypothèses du théorème 3.3.1.

**Exemple 3.3.5** Un nombre positif  $L$  étant donné, cherchons la fonction lipschitzienne  $[0, L] \ni t \rightarrow \xi_*(t) \in \mathbb{R}$  qui rend l'intégrale  $\int_0^L \xi(t)^2 dt$  minimale dans l'ensemble de toutes les fonctions absolument continues  $\xi$  vérifiant

$$\xi(0) = 1, \quad \xi(L) = 1, \quad \text{et } |\dot{\xi}(t)| \leq 1 \text{ pour presque tout } t \in [0, L]. \quad (29)$$

Un calcul très simple permet d'obtenir directement la solution  $\xi_*$ . Si  $L > 2$ , on a  $\xi_*(t) = 1 - t$  pour  $0 \leq t \leq 1$ ,  $\xi_*(t) = 0$  pour  $1 \leq t \leq L - 1$ ,  $\xi_*(t) = t + 1 - L$  pour  $L - 1 \leq t \leq L$ . (Si  $L \leq 2$ , on a  $\xi_*(t) = 1 - t$  pour  $0 \leq t \leq \frac{L}{2}$  et  $\xi_*(t) = t + 1 - L$  pour  $\frac{L}{2} \leq t \leq L$ .) Il est facile de vérifier que la courbe  $\xi_*$  est effectivement une extrémale. En outre, sans supposer qu'on ait déjà trouvé le candidat  $\xi_*$ , il est possible d'appliquer directement le théorème 3.3.1 pour trouver la solution  $\xi_*$ , par un calcul un peu long mais assez simple, dont nous ne donnerons pas les détails.  $\diamond$

### 3.3.3 Le principe du maximum de la théorie du contrôle

Pour faciliter la dérivation du théorème 3.3.1, nous avons abandonné le contexte des problèmes invariants sur des variétés et nous nous sommes placés, en introduisant l'hypothèse (H1), dans le cadre beaucoup plus étroit où l'espace d'états est un ouvert dans  $\mathbb{R}^n$ . Nous allons maintenant regagner l'invariance perdue en donnant une version invariante du théorème 3.3.1.

Remarquons d'abord que, dans la situation du théorème 3.3.1, nous pouvons associer à chaque point  $v$  de  $\mathbb{R}^n$  un champ de vecteurs constant  $G_v$ , défini par la formule  $G_v(x) = v$ . Le lagrangien  $L$  nous permet aussi d'associer à chaque  $v \in \mathbb{R}^n$  et à chaque  $t \in [a, b]$  une fonction scalaire partielle  $L_{v,t} : X \dashrightarrow \mathbb{R}$ , définie par  $L_{v,t}(x) = L(x, v, t)$ .

On appellera *champ de vecteurs augmenté* sur une variété différentiable  $X$  un couple  $(F, L)$  tel que  $F$  est un champ de vecteurs partiel sur  $X$  et  $L$  est une fonction réelle partielle sur  $X$ . À chaque champ de vecteurs augmenté  $(F, L)$  sur  $X$  on associe une fonction partielle  $H_{F,L} : T^*X \times \mathbb{R} \dashrightarrow \mathbb{R}$  par la formule  $H_{F,L}(x, p, p_0) = \langle p, F(x) \rangle - p_0 L(x)$ , le domaine de  $H_{F,L}$  étant, évidemment, l'ensemble des  $(x, p, p_0) \in T^*X \times \mathbb{R}$  tels que  $x \in \text{Dom}(F) \cap \text{Dom}(L)$ .

Revenant maintenant à la situation du théorème 3.3.1, notons que le lagrangien  $L$  et l'intervalle  $[a, b]$  donnent lieu à une *famille paramétrisée*

$$\mathbf{F} = \mathbf{F}_{L,a,b} = \left\{ (F_{v,t}, L_{v,t}) \right\}_{(v,t) \in \mathbb{R}^n \times [a,b]}$$

de champs de vecteurs augmentés, où  $F_{v,t} \equiv G_v$ .

Écrivons maintenant  $F(x, v, t)$  au lieu de  $F_{v,t}(x)$ . Il en résulte que

$$H(x, v, p, p_0, t) = H_{F_{v,t}, L_{v,t}}(x, p, p_0) = \langle p, F(x, v, t) \rangle - p_0 L(x, v, t).$$

L'ensemble  $\mathcal{Y}$  est essentiellement l'ensemble des courbes  $\xi$  telles qu'il existe une fonction  $[a, b] \ni t \rightarrow v(t) \in \mathbb{R}^n$  qui rend l'équation

$$\dot{\xi}(t) = F(\xi(t), v(t), t) \tag{30}$$

vraie presque partout. Le coût  $J(\xi)$  de la courbe  $\xi$  est donné par

$$J(\xi) = \int_a^b L(\xi(t), v(t), t) dt,$$

$t \rightarrow v(t)$  étant la fonction vérifiant (30), qui est évidemment unique en vertu de l'équation  $\dot{\xi}(t) = F(\xi(t), v(t), t) = v(t)$ .

***Pour obtenir la version intrinsèque du théorème 3.3.1, il suffit d'oublier que la famille paramétrisée  $\mathbf{F}$  est de la forme particulière de ce théorème, où chaque champ de vecteurs  $x \rightarrow F(x, v, t)$  est constant et le paramètre  $v$  est précisément la valeur constante de ce champ, et de considérer des famille paramétrisées complètement générales de champs de vecteurs augmentés.***

Le paramètre  $v$  — qu'on appellera le *contrôle*, ou la *commande* — ne sera plus assimilé à la vitesse  $\xi$ , et deviendra plutôt un paramètre à valeurs dans un ensemble abstrait sans aucune structure. Désormais, pour distinguer clairement la commande de la vitesse, la commande sera notée  $u$  plutôt que  $v$ .

Dans ce nouveau cadre, il n'est plus vrai que la fonction  $t \rightarrow v(t)$  est déterminée par la courbe  $\xi$ . La fonctionnelle  $J$  sera donc définie sur un espace — noté  $\mathcal{Z}$  — de « couples trajectoire-commande » plutôt que sur un espace de trajectoires.

L'hypothèse (H1) sera remplacée par la condition suivante :

(H<sub>c</sub>1) *L'espace des configurations  $X$  est une variété de classe  $C^2$ ,  $a \in \mathbb{R}$ ,  $b \in \mathbb{R}$ ,  $a < b$ ,  $U$  est un espace métrique séparable, et  $\mathbf{F} = \{ (F_{u,t}, L_{u,t}) \}_{(u,t) \in U \times [a,b]}$  est une famille de champs de vecteurs augmentés sur  $X$ .*

Écrivons  $F(x, u, t) \stackrel{\text{déf}}{=} F_{u,t}(x)$ ,  $L(x, u, t) \stackrel{\text{déf}}{=} L_{u,t}(x)$ ,  $\text{Dom}(\mathbf{F}) \stackrel{\text{déf}}{=} \text{Dom}(F) \cap \text{Dom}(L)$ . Si  $u \in U$ , on notera  $\mathbf{F}^u$  l'application partielle  $(x, t) \rightarrow (F(x, u, t), L(x, u, t))$ , de  $X \times [a, b]$  dans  $TX \times \mathbb{R}$ . Ainsi donc,

$$\text{Dom}(\mathbf{F}^u) = \{ (x, t) : (x, u, t) \in \text{Dom}(F) \cap \text{Dom}(L) \}.$$

Pour énoncer les hypothèses de régularité par rapport à  $x$ ,  $v$  et  $t$ , définissons d'abord le *hamiltonien*  $H$  par la formule

$$H(x, u, p, p_0, t) = \langle p, F(x, u, t) \rangle - p_0 L(x, u, t). \quad (31)$$

Il s'ensuit que  $H$  est une fonction partielle sur l'ensemble

$$D_{X,U,a,b} \stackrel{\text{déf}}{=} \left\{ (x, u, p, p_0, t) : (x, p, p_0, t, u) \in T^*X \times \mathbb{R} \times [a, b] \times U \right\},$$

dont le domaine est donné par

$$\text{Dom}(H) = \left\{ (x, u, p, p_0, t) \in D_{X,U,a,b} : (x, u, t) \in \text{Dom}(F) \cap \text{Dom}(L) \right\}.$$

On supposera vérifiées les hypothèses techniques suivantes :

(H<sub>c</sub>2) *Pour tout  $u \in U$ , le domaine de la fonction partielle*

$$X \times [a, b] \ni (x, t) \dashrightarrow (F(x, u, t), L(x, u, t)) \in TX \times \mathbb{R}$$

*est une partie ouverte de  $X \times [a, b]$ ,*

(H<sub>c</sub>3) *Pour tout  $(u, t)$  appartenant à  $U \times [a, b]$ , la fonction partielle*

$$T^*X \times \mathbb{R} \ni (x, p, p_0) \dashrightarrow H(x, u, p, p_0, t) \in \mathbb{R}$$

*— dont le domaine est ouvert dans  $T^*X \times \mathbb{R}$  en vertu de (H<sub>c</sub>2) — est de classe  $C^1$ ,*

- (H<sub>c</sub>4) la fonction  $T^*X \times \mathbb{R} \times U \ni (x, p, p_0, u) \mapsto H(x, u, p, p_0, t) \in \mathbb{R}$  est continue quel que soit  $t \in [a, b]$ ,
- (H<sub>c</sub>5) la fonction  $[a, b] \ni t \mapsto H(x, u, p, p_0, t) \in \mathbb{R}$  est mesurable quel que soit  $(x, p, p_0, u) \in T^*X \times \mathbb{R} \times U$ ,
- (H<sub>c</sub>6) les fonctions  $H$  et  $(-\frac{\partial H}{\partial p}, \frac{\partial H}{\partial x})$  sont localement bornées.

On appellera *couple trajectoire-commande* pour les données  $X, a, b, S, \mathbf{F}$  un couple  $(\xi, \eta)$  tel que (a)  $\partial\xi \in S$ , (b)  $\eta$  est une fonction mesurable sur  $[a, b]$  à valeurs dans  $U$ , (c)  $\xi$  est une application absolument continue de  $[a, b]$  dans  $X$ , (d) il existe un sous-ensemble compact  $K$  de  $\text{Dom}(\mathbf{F})$  et un voisinage relatif  $\Omega$  de  $\{(\xi(t), t) : a \leq t \leq b\}$  dans  $X \times [a, b]$  tels que, pour presque tout  $t \in [a, b]$ , la condition  $(x, t) \in \Omega$  implique  $(x, \eta(t), t) \in K$ , et (e)  $\dot{\xi}(t) = F(\xi(t), \eta(t), t)$  pour presque tout  $t \in [a, b]$ . Nous noterons  $\mathcal{Z}_{X,a,b,S,\mathbf{F}}$  l'ensemble de tous les couples trajectoire-commande pour  $X, a, b, S, \mathbf{F}$ .

En vertu des nos hypothèses, si  $(\xi, \eta)$  appartient à  $\mathcal{Z}_{X,a,b,S,\mathbf{F}}$ , la fonction  $[a, b] \ni t \mapsto L(\xi(t), \eta(t), t) \in \mathbb{R}$  est mesurable et bornée. Ainsi donc, l'intégrale

$$J_{X,a,b,S,\mathbf{F}}(\xi, \eta) = \int_a^b L(\xi(t), \eta(t), t) dt. \quad (32)$$

existe quel que soit  $(\xi, \eta) \in \mathcal{Z}_{X,a,b,S,\mathbf{F}}$ . On définit

$$U_{\xi_*, \mathbf{F}}(t) \stackrel{\text{déf}}{=} \left\{ u \in \mathbb{R}^n : (\xi_*(t), t) \in \text{Dom}(\mathbf{F}^u) \right\}, \quad (33)$$

L'énoncé suivant se réduit exactement à celui du théorème 3.3.1 dans le cas particulier où  $X$  est un ouvert de  $\mathbb{R}^n$ ,  $U = \mathbb{R}^n$ , et  $F(x, u, t) \equiv u$ . D'autre part, cet énoncé a la vertu d'avoir une interprétation précise —et complètement invariante— sur les variétés, ce qui est une raison assez forte pour soupçonner qu'il doit être vrai.

**Théorème 3.3.6** *Supposons que les hypothèses (H<sub>c</sub>1)-(H<sub>c</sub>6) soient vérifiées. Soit  $S$  une sous-variété de classe  $C^1$  de  $X \times X$ . Notons  $\mathcal{Z} = \mathcal{Z}_{X,a,b,S,\mathbf{F}}$  et  $J = J_{X,a,b,S,\mathbf{F}}$ . Supposons que  $(\xi_*, \eta_*) \in \mathcal{Z}$  soit tel que  $J(\xi_*, \eta_*) \leq J(\xi, \eta)$  quel que soit  $(\xi, \eta) \in \mathcal{Z}$ . Il existe alors un champ de covecteurs  $(\pi, \pi_0)$  le long de  $\xi_*$ , absolument continu et positivement augmenté, vérifiant les conditions (a) et (b) de l'énoncé du théorème 3.3.1, tel que les **équations de Hamilton***

$$\dot{\xi}_*(t) = \frac{\partial H}{\partial p}(\xi_*(t), \eta_*(t), \pi(t), \pi_0, t), \quad (34)$$

$$\dot{\pi}(t) = -\frac{\partial H}{\partial x}(\xi_*(t), \eta_*(t), \pi(t), \pi_0, t) \quad (35)$$

et la condition de **maximisation du hamiltonien**

$$H(\xi_*(t), \eta_*(t), \pi(t), \pi_0, t) = \max \left\{ H(\xi_*(t), u, \pi(t), \pi_0, t) : u \in U_{\xi_*, \mathbf{F}}(t) \right\} \quad (36)$$

sont vérifiées presque partout.  $\diamond$

Le théorème 3.3.6 est en effet vrai. Il est en fait essentiellement identique<sup>8</sup> à la version classique du *principe du maximum de Pontriaguine*, ou *principe du maximum de la théorie du contrôle optimal*, cf. [14].

### 3.4 Généralisations de la première version du principe du maximum

Après la découverte du principe du maximum sous la forme de [14], c'est-à-dire du théorème 3.3.6, plusieurs généralisations ont été proposées. Ces généralisations sont des versions modifiées du théorème 3.3.6 où, dans l'énoncé du théorème, on a substitué :

- (1) *Des hypothèses techniques plus faibles.* Par exemple, on permet que l'ensemble  $S$  soit plus général qu'une sous-variété, ou on remplace l'hypothèse que  $F$  et  $L$  soient de classe  $C^1$  par rapport à l'état par des conditions de Lipschitz, ou par des conditions de différentiabilité en  $\xi_*(t)$  sans supposer que  $F$  ou  $L$  soient lipschitziens au voisinage de  $\xi_*(t)$ .
- (2) *Des conclusions plus fortes.* Par exemple, on obtient un multiplicateur qui, outre la maximisation du hamiltonien, vérifie d'autres inégalités, liées à des « variations d'ordre supérieur ».
- (3) *Un cadre plus général.* Par exemple, on considère des inclusions différentielles au lieu de champs de vecteurs. Ou bien on remplace l'optimisation par d'autres problèmes plus généraux, tels que celui de la « séparation de l'ensemble d'accessibilité d'un autre ensemble donné », qui contient le problème d'optimisation aussi bien que celui des optimums de Pareto et celui de la controllabilité locale le long d'une trajectoire.

Bien sûr, chacune de ces trois possibilités peut être réalisée de plusieurs façons et, en outre, les trois possibilités peuvent être combinées. Il s'ensuit que des très nombreuses généralisations sont logiquement possibles en principe, et qu'il ne faut pas s'étonner de ce que la littérature spécialisée contienne une vraie prolifération de versions.

L'espace limité dont nous disposons ne nous permettra pas de donner une liste exhaustive de ces versions. D'ailleurs, une telle liste ne serait utile que s'il y avait des bonnes raisons pour retenir chacun des résultats, comme ce serait le cas si, par exemple, il était impossible de les réduire tous à des cas particuliers d'un petit nombre de théorèmes généraux.

Or, le but de notre travail est précisément de montrer qu'une telle réduction est possible, et qu'il existe un théorème unique qui contient et généralise —très

---

<sup>8</sup>Nos hypothèses techniques sont légèrement différentes de celles des théorèmes de cf. [14], et un peu plus restrictives dans certains aspects. Nous choisissons la forme du théorème 3.3.6 parce qu'elle est un point de départ très naturel pour les généralisations qui sont le but de ce travail. Ces généralisations seront, de toute façon, beaucoup plus fortes que les résultats de [14].

fortement— tous les autres, à l'exception d'un petit nombre de résultats isolés qui jusqu'ici n'ont pas cédé à nos méthodes (cf. les remarques 3.4.7 et 4.8.4).

Nous nous limiterons donc à énumérer les éléments qui ont été incorporés dans ces versions, suivant un ordre logique plutôt que chronologique, ce qui nous mènera assez directement à l'énoncé général cherché.

**1. Familles générales de champs de vecteurs non-autonomes.** Dans l'énoncé du théorème 3.3.6, chaque commande  $\eta : [a, b] \rightarrow U$  donne lieu à un champ de vecteurs augmenté non autonome  $(f_\eta, \ell_\eta)$ , défini par les formules

$$\begin{aligned} f_\eta(x, t) &\stackrel{\text{déf}}{=} F(x, \eta(t), t) = F_{\eta(t), t}(x), \\ \ell_\eta(x, t) &\stackrel{\text{déf}}{=} L(x, \eta(t), t) = L_{\eta(t), t}(x). \end{aligned}$$

Appelons *système de champs de vecteurs augmentés* une triple  $(X, I, \mathbf{f})$  telle que  $X$  est une variété différentiable de classe  $C^2$ ,  $I$  est un sous-intervalle de  $\mathbb{R}$  de mesure positive, et  $\mathbf{f} = \{(f_\eta, \ell_\eta)\}_{\eta \in \mathcal{U}}$  est une famille complètement générale de champs de vecteurs augmentés non autonomes sur  $X$ , paramétrisée par un ensemble arbitraire  $\mathcal{U}$  sans aucune structure. Le symbole  $\eta$  désigne maintenant un élément de  $\mathcal{U}$  qui n'est plus nécessairement une fonction du temps. La notion de mesurabilité d'une commande étant donc dépourvue de sens, on la remplace par l'hypothèse

(A1) *La fonction partielle  $I \ni t \dashrightarrow (f_\eta(x, t), \ell_\eta(x, t)) \in TX \times \mathbb{R}$  est mesurable quel que soit  $(x, \eta) \in X \times \mathcal{U}$ .*

Au lieu de (H<sub>c</sub>2), on introduit les conditions suivantes :

(A2) *L'ensemble  $(X \times I) \cap \text{Dom}(f_\eta) \cap \text{Dom}(\ell_\eta)$  est une partie ouverte de  $X \times I$  quel que soit  $(t, \eta) \in I \times \mathcal{U}$ .*

(A3) *la fonction partielle  $X \ni x \dashrightarrow (f_\eta(x, t), \ell_\eta(x, t)) \in TX \times \mathbb{R}$  est de classe  $C^1$  quel que soit  $(t, \eta) \in I \times \mathcal{U}$ .*

Si  $M$  est l'ensemble de toutes les fonctions mesurables d'un intervalle  $[a, b]$  dans un espace métrique  $U$ , la *condition de remplacement* suivante est vérifiée : si  $c \in [a, b]$ ,  $\alpha \in M$  et  $\beta \in M$ , et  $\gamma(t) = \alpha(t)$  pour  $t < c$ ,  $\gamma(t) = \beta(t)$  pour  $t \geq c$ , alors  $\gamma \in M$ . Dans notre situation plus générale, on fera l'hypothèse suivante :

(A4) *Quel que soit  $c \in I$ , si  $\eta_1 \in \mathcal{U}$ ,  $\eta_2 \in \mathcal{U}$ , et on écrit*

$$\begin{aligned} f_3(x, t) &= f_{\eta_1}(x, t), & \ell_3(x, t) &= \ell_{\eta_1}(x, t) & \text{si } t < c, \\ f_3(x, t) &= f_{\eta_2}(x, t), & \ell_3(x, t) &= \ell_{\eta_2}(x, t) & \text{si } t \geq c, \end{aligned}$$

*alors il existe  $\eta_3 \in \mathcal{U}$  tel que  $(f_3, \ell_3) = (f_{\eta_3}, \ell_{\eta_3})$ .*

L'hypothèse (H<sub>c</sub>6) peut être remplacée par une condition d'intégrabilité :

(A5) Pour tout  $\eta$  appartenant à  $\mathcal{U}$  et tout sous-ensemble compact  $K$  de l'intersection  $(X \times I) \cap \text{Dom}(f_\eta) \cap \text{Dom}(\ell_\eta)$  il existe une fonction localement intégrable  $\varphi : I \rightarrow \mathbb{R}_+$  telle que

$$\|f_\eta(x, t)\| + |\ell_\eta(x, t)| + \left\| \frac{\partial f_\eta}{\partial x}(x, t) \right\| + \left\| \frac{\partial \ell_\eta}{\partial x}(x, t) \right\| \leq \varphi(t)$$

pour tout  $(x, t) \in K$ , où les normes sont définies par rapport à une métrique riemannienne sur  $X$ .

Soit  $(X, I, \mathbf{f}) = (X, I, \{(f_\eta, \ell_\eta)\}_{\eta \in \mathcal{U}})$  un système de champs de vecteurs augmentés. On appelle *couple trajectoire-commande* pour  $(X, I, \mathbf{f})$  un couple  $(\xi, \eta)$  tel que  $\eta \in \mathcal{U}$ ,  $\xi$  est une courbe absolument continue dans  $X$  telle que  $\text{Dom}(\xi)$  est un sous-intervalle compact de  $I$ ,  $(\xi(t), t) \in \text{Dom}(f_\eta) \cap \text{Dom}(\ell_\eta)$  pour tout  $t \in \text{Dom}(\xi)$ , et l'équation  $\dot{\xi}(t) = f_\eta(\xi(t), t)$  est vérifiée pour presque tout  $t \in \text{Dom}(\xi)$ . On notera  $\mathcal{Z}_{X, I, \mathbf{f}}$  l'ensemble de tous les couples trajectoire-commande pour  $(X, I, \mathbf{f})$ .

Si  $(\xi, \eta) \in \mathcal{Z}_{X, I, \mathbf{f}}$ , la fonction  $\text{Dom}(\xi) \ni t \rightarrow \ell_\eta(\xi(t), t)$  est intégrable, et nous pouvons définir le coût  $J(\xi, \eta) = \int_{\text{Dom}(\xi)} \ell_\eta(\xi(t), t) dt$ . Si l'on spécifie des temps  $a \in \mathbb{R}$ ,  $b \in \mathbb{R}$ , tels que  $a < b$ , et une sous-variété  $S$  de classe  $C^1$  de  $X \times X$ , on notera  $\mathcal{Z}_{X, I, \mathbf{f}; a, b, S}$  l'ensemble des  $(\xi, \eta) \in \mathcal{Z}_{X, I, \mathbf{f}}$  tels que  $\text{Dom}(\xi) = [a, b]$  et  $\partial \xi \in S$ , et  $J_{X, I, \mathbf{f}; a, b, S}$  désignera la restriction de  $J_{X, I, \mathbf{f}}$  à  $\mathcal{Z}_{X, I, \mathbf{f}; a, b, S}$ .

On définit le hamiltonien  $H_\eta$  associé à un  $\eta \in \mathcal{U}$  par la formule

$$H_\eta(x, p, p_0, t) = \langle p, f_\eta(x, t) \rangle - p_0 \ell_\eta(x, t).$$

**Théorème 3.4.1** Soit  $(X, I, \mathbf{f}) = (X, I, \{(f_\eta, \ell_\eta)\}_{\eta \in \mathcal{U}})$  un système de champs de vecteurs augmentés. Supposons que les hypothèses (A1)-(A5) soient vérifiées. Soit  $S$  une sous-variété de classe  $C^1$  de  $X \times X$ . Notons  $\mathcal{Z} = \mathcal{Z}_{X, I, \mathbf{f}; a, b, S}$  et  $J = J_{X, I, \mathbf{f}; a, b, S}$ . Soit  $(\xi_*, \eta_*) \in \mathcal{Z}$  tel que  $J(\xi_*, \eta_*) \leq J(\xi, \eta)$  pour tout  $(\xi, \eta) \in \mathcal{Z}$ . Il existe alors un champ de covecteurs  $(\pi, \pi_0)$  le long de  $\xi_*$ , absolument continu et positivement augmenté, vérifiant les conditions (a) et (b) de l'énoncé du théorème 3.3.1, et tel que

### 1. les équations de Hamilton

$$\dot{\xi}_*(t) = \frac{\partial H_{\eta_*}}{\partial p}(\xi_*(t), \pi(t), \pi_0, t) \quad (37)$$

$$\dot{\pi}(t) = -\frac{\partial H_{\eta_*}}{\partial x}(\xi_*(t), \pi(t), \pi_0, t) \quad (38)$$

ont lieu presque partout,

### 2. pour tout $\eta \in \mathcal{U}$ , on a l'inégalité

$$H_\eta(\xi_*(t), \pi(t), \pi_0, t) \leq H_{\eta_*}(\xi_*(t), \pi(t), \pi_0, t) \quad (39)$$

pour presque tout  $t \in \{s \in [a, b] : (\xi_*(s), s) \in \text{Dom}(f_\eta) \cap \text{Dom}(\ell_\eta)\}$ .  $\diamond$

**Remarque 3.4.2** L'idée de considérer des familles paramétrisées arbitraires de champs de vecteurs augmentés non autonomes est due à Kaskosz et Lojasiewicz, cf. [13].  $\diamond$

**2. Champs lipschitziens.** Au lieu de supposer les fonctions partielles  $x \dashrightarrow (f_\eta(x, t), \ell_\eta(x, t))$  de classe  $C^1$ , on peut supposer qu'elles soient localement lipschitziennes. Plus précisément, on remplace (A3) et (A5) par

(A3') *La fonction partielle  $X \ni x \dashrightarrow (f_\eta(x, t), \ell_\eta(x, t)) \in TX \times \mathbb{R}$  est localement lipschitzienne quel que soit  $(t, \eta) \in I \times \mathcal{U}$ .*

(A5') *Pour tout  $\eta$  appartenant à  $\mathcal{U}$  et tout sous-ensemble compact  $K$  de l'intersection  $(X \times I) \cap \text{Dom}(f_\eta) \cap \text{Dom}(\ell_\eta)$  il existe une fonction localement intégrable  $\varphi : I \rightarrow \mathbb{R}_+$  telle que  $\|f_\eta(x, t)\| + |\ell_\eta(x, t)| \leq \varphi(t)$  pour tout  $(x, t) \in K$ , et  $\|f_\eta(x, t) - f_\eta(y, t)\| + |\ell_\eta(x, t) - \ell_\eta(y, t)| \leq \varphi(t)\|x - y\|$  pour tout  $(x, t) \in K, (y, t) \in K$ , où les normes sont définies par rapport à une métrique riemannienne sur  $X$ .*

La deuxième équation de Hamilton doit être remplacée par la condition

$$-\dot{\pi}(t) \in \partial_x H_{\eta_*}(\xi_*(t), \pi(t), \pi_0, t) \text{ pour presque tout } t. \quad (40)$$

où  $\partial_x H_{\eta_*}(\xi_*(t), \pi(t), \pi_0, t)$  est le gradient généralisé de Clarke de la fonction  $x \rightarrow H_{\eta_*}(x, \pi(t), \pi_0, t)$  (cf. [5]), c'est-à-dire l'enveloppe convexe de toutes les limites  $\omega = \lim_{j \rightarrow \infty} \nabla_x H_{\eta_*}(x_j, \pi(t), \pi_0, t)$ , pour toutes les suites  $\{x_j\}_{j \in \mathbb{N}}$  telles que  $\omega_j = \nabla_x H_{\eta_*}(x_j, \pi(t), \pi_0, t)$  existe pour chaque  $j$  et la limite  $\lim_{j \rightarrow \infty} \omega_j$  existe.

On a alors le *principe du maximum non lisse de F. Clarke* (cf. [4, 5]) :

**Theorème 3.4.3** *Soit  $(X, I, \mathbf{f}) = (X, I, \{(f_\eta, \ell_\eta)\}_{\eta \in \mathcal{U}})$  un système de champs de vecteurs augmentés. Supposons que les hypothèses (A1), (A2), (A3'), (A4), (A5') soient vérifiées. Soit  $S$  une sous-variété de classe  $C^1$  de  $X \times X$ . Notons  $\mathcal{Z} = \mathcal{Z}_{X, I, \mathbf{f}; a, b, S}$  et  $J = J_{X, I, \mathbf{f}; a, b, S}$ . Soit  $(\xi_*, \eta_*) \in \mathcal{Z}$  tel que  $J(\xi_*, \eta_*) \leq J(\xi, \eta)$  quel que soit  $(\xi, \eta) \in \mathcal{Z}$ . Il existe alors un couple  $(\pi, \pi_0)$  vérifiant toutes les conclusions du théorème 3.4.1, à l'exception de l'équation de Hamilton (38), qui doit être est remplacée par l'inclusion différentielle (40).  $\diamond$*

**3. Champs de vecteurs continus.** S. Lojasiewicz remarqua que dans la conclusion du théorème 3.4.3 la condition de Lipschitz ne joue un rôle que pour le « champ de référence »  $(f_{\eta_*}, \ell_{\eta_*})$ . Pour les autres champs, il suffit de supposer qu'il soient continus. Plus précisément, on remplace (A3') et (A5') par

(A3' .a) *La fonction partielle  $X \ni x \dashrightarrow (f_\eta(x, t), \ell_\eta(x, t)) \in TX \times \mathbb{R}$  est continue quel que soit  $(t, \eta) \in I \times \mathcal{U}$ .*

(A3' .b) *La fonction partielle  $X \ni x \dashrightarrow (f_{\eta_*}(x, t), \ell_{\eta_*}(x, t)) \in TX \times \mathbb{R}$  est lipschitzienne au voisinage de  $\xi_*(t)$  quel que soit  $t \in I$ .*

(A5".a) Pour tout  $\eta$  appartenant à  $\mathcal{U}$  et tout sous-ensemble compact  $K$  de l'intersection  $(X \times I) \cap \text{Dom}(f_\eta) \cap \text{Dom}(\ell_\eta)$  il existe une fonction localement intégrable  $\varphi : I \rightarrow \mathbb{R}_+$  telle que  $\|f_\eta(x, t)\| + |\ell_\eta(x, t)| \leq \varphi(t)$  pour tout  $(x, t) \in K$ .

(A5".b) Il existe un voisinage relatif  $\Omega$  de  $\{(\xi_*(t), t) : a \leq t \leq b\}$  dans le produit  $X \times [a, b]$  tel que  $\Omega \subseteq \text{Dom}(f_{\eta_*}) \cap \text{Dom}(\ell_{\eta_*})$  et une fonction intégrable  $\varphi : [a, b] \rightarrow \mathbb{R}_+$  vérifiant

$$\|f_{\eta_*}(x, t) - f_{\eta_*}(y, t)\| + |\ell_{\eta_*}(x, t) - \ell_{\eta_*}(y, t)| \leq \varphi(t)\|x - y\|$$

pour tout  $(x, t) \in \Omega, (y, t) \in \Omega$ .

Sous ces hypothèses, on a le principe du maximum non lisse de Lojasiewicz<sup>9</sup> :

**Théorème 3.4.4** Soit  $(X, I, \mathbf{f}) = (X, I, \{(f_\eta, \ell_\eta)\}_{\eta \in \mathcal{U}})$  un système de champs de vecteurs augmentés. Supposons que les hypothèses (A1), (A2), (A3".a), (A4), (A5".a) soient vérifiées. Soit  $S$  une sous-variété de classe  $C^1$  du produit  $X \times X$ . Notons  $\mathcal{Z} = \mathcal{Z}_{X, I, \mathbf{f}; a, b, S}$  et  $J = J_{X, I, \mathbf{f}; a, b, S}$ . Soit  $(\xi_*, \eta_*) \in \mathcal{Z}$  tel que (A3".b) et (A5".b) sont vérifiées, et  $J(\xi_*, \eta_*) \leq J(\xi, \eta)$  quel que soit  $(\xi, \eta) \in \mathcal{Z}$ . Il existe alors un couple  $(\pi, \pi_0)$  vérifiant toutes les conclusions du théorème 3.4.3.  $\diamond$

**5. Champs de vecteurs de référence vérifiant une condition de différentiabilité généralisée.** Les conditions de Lipschitz (A3".b) et (A5".b) sont évidemment vérifiées lorsque les conditions plus fortes de (A3) et (A5) le sont, et dans ce cas le gradient généralisé de Clarke  $\partial_x H_{\eta_*}$  de (40) se confond avec la dérivé partielle usuelle  $\frac{\partial H_{\eta_*}}{\partial x}$ .

Par contre, si une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est localement lipschitzienne, et la différentielle classique  $\nabla f(\bar{x})$  existe en un point  $\bar{x}$ , il se peut que  $\{\nabla f(\bar{x})\}$  —qui doit satisfaire à la relation  $\{\nabla f(\bar{x})\} \subseteq \partial_x f(\bar{x})$ — soit un sous-ensemble propre de  $\partial_x f(\bar{x})$ . (Prenons, par exemple,  $n = 1$  et  $f(x) = x^2 \sin \frac{1}{x}$ . Il vient  $f'(0) = 0$  et  $\partial f(0) = [-1, 1]$ .) Supposons que le champ augmenté  $x \rightarrow (f_{\eta_*}(x, t), \ell_{\eta_*}(x, t))$  admette une différentielle classique<sup>10</sup>  $D(t) = (D^f(t), D^\ell(t))$  en  $x = \xi_*(t)$  pour presque tout  $t \in [a, b]$ , et que cette différentielle soit une fonction intégrable (à valeurs matrices ou jets). On peut donc écrire formellement l'équation de Hamilton (38) associée à cette différentielle, qui prend la forme<sup>11</sup>

<sup>9</sup>dont l'énoncé et la démonstration nous ont été communiqués par Lojasiewicz au cours de deux conversations en 1992 et 1993. À notre connaissance, aucune version de ce résultat remarquable n'a été publiée, malgré son importance dans le développement de notre sujet, et le rôle décisif qu'il a joué en révélant la possibilité de démontrer des théorèmes non lisses par des méthodes « primales ».

<sup>10</sup>Dans le cas où  $X$  est un ouvert de  $\mathbb{R}^n$ , on a  $D(t) \in \mathbb{R}^{(n+1) \times n}$ ,  $D^f(t) \in \mathbb{R}^{n \times n}$ , et  $D^\ell(t) \in \mathbb{R}^{1 \times n} = \mathbb{R}^n$ . Lorsque  $X$  est une variété,  $D(t)$  est un 1-jet en  $\xi_*(t)$  de sections  $\sigma$  du fibré  $TX \times \mathbb{R}$  telles que  $\sigma(\xi_*(t)) = (f_{\eta_*}(\xi_*(t), t), \ell_{\eta_*}(\xi_*(t), t))$ .

<sup>11</sup>L'expression  $\pi_0 D^\ell(t)$  est évidemment un covecteur. Le caractère intrinsèque de la somme  $\dot{\pi}(t) + \pi(t) \cdot D^f(t)$  se démontre de la façon suivante. Écrivons  $x = \xi_*(t)$ ,  $v = \dot{\xi}_*(t)$ ,  $\gamma(s) = \xi_*(t + s)$ . Soient  $\sigma, \tau$  des sections des  $TX$  —c'est-à-dire des champs

$$\begin{aligned}
\dot{\pi}(t) &= -\frac{\partial H_{\eta_*}}{\partial x}(\xi_*(t), \pi(t), \pi_0, t) \\
&= -\pi(t) \cdot \frac{\partial f_{\eta_*}}{\partial x}(\xi_*(t), \pi(t), \pi_0, t) + \pi_0 \frac{\partial \ell_{\eta_*}}{\partial x}(\xi_*(t), \pi(t), \pi_0, t) \\
&= -\pi(t) \cdot D^f(t) + \pi_0 D^\ell(t).
\end{aligned} \tag{41}$$

Si  $\{\frac{\partial H_{\eta_*}}{\partial x}(\xi_*(t), \pi(t), \pi_0, t)\} \neq \partial_x H_{\eta_*}(\xi_*(t), \pi(t), \pi_0, t)$ , l'équation (41) est strictement plus restrictive que l'inclusion différentielle (40). In s'ensuit que la conclusion du principe du maximum avec (41) au lieu de (40) est strictement plus forte.

En outre, il y a des fonctions classiquement différentiables en un point  $x$  qui ne sont pas localement lipschitziennes au voisinage de  $x$ . Pour ces fonctions, on peut toujours écrire l'équation (41), tandis que (40) est dépourvue de sens.

H. Halkin trouva une nouvelle notion de « dérivée généralisée »—qu'il appela « shield »—dont la différentielle classique et le jacobien de Clarke sont des cas particuliers (cf. Halkin [8, 9, 10]).

En utilisant cette notion de différentielle, Halkin démontra une généralisation du théorème 3.4.3. Nous allons présenter une version encore plus générale en combinant l'idée de Halkin avec celle de Łojasiewicz.

Si  $X$  est une variété de class  $C^2$ ,  $x \in X$  et  $v \in T_x X$ , notons  $J_x^1(TX, v)$  l'ensemble des 1-jets en  $x$  de champs de vecteurs sur  $X$  dont le 0-jet en  $x$  est le vecteur  $v$ .

Nous dirons que  $D$  est un *générateur variationnel* pour le couple trajectoire-commande  $(\xi_*, \eta_*)$  et le système de champs de vecteurs augmentés  $(X, I, \mathbf{f})$  si

(GV.1)  $D$  est une fonction multivoque mesurable

$$[a, b] \ni t \rightarrow D(t) \subseteq J_{\xi_*(t)}^1(TX, \dot{\xi}_*(t)) \times T_{\xi_*(t)}^* X$$

à valeurs compacts convexes non vides,

(GV.2) Il existe un couple  $(\Omega, \mathbf{k})$  telle que

(GV.2.a)  $\Omega$  est un voisinage relatif de  $\{(\xi_*(t), t) : a \leq t \leq b\}$  dans le produit  $X \times [a, b]$ , tel que  $\Omega \subseteq \text{Dom}(f_{\eta_*}) \cap \text{Dom}(\ell_{\eta_*})$ ,

(GV.2.b)  $\mathbf{k}$  est une famille  $\mathbf{k} = \{k_\delta\}_{0 < \delta \leq \bar{\delta}}$  de fonctions intégrables sur  $[a, b]$  telle que  $\lim_{\delta \downarrow 0} \int_a^b k_\delta(t) dt = 0$ ,

---

de vecteurs—différentiables en  $x$  et telles que le 1-jet de  $\sigma$  en  $x$  est égal à  $(v, D^f(t))$  et  $[\sigma, \tau](x) = 0$ . La dérivée en  $s = 0$  de la fonction scalaire  $s \rightarrow \pi(t+s) \cdot \tau(\gamma(s))$  est égale, relativement à une carte  $\kappa$ , à  $\dot{\pi}(t) \cdot w + \pi(t) \cdot E \cdot v$ , où  $(w, E)$  est le 1-jet de  $\tau$  en  $x$ . Or,  $E \cdot v - D^f(t) \cdot w = [\sigma, \tau](x) = 0$ . Le dérivée est donc égale à  $(\dot{\pi}(t) + \pi(t) \cdot D^f(t)) \cdot w$ . Ceci montre que  $\lambda(w) = (\dot{\pi}(t) + \pi(t) \cdot D^f(t)) \cdot w$  ne dépend pas du choix de  $\kappa$ . Il s'ensuit que  $\lambda$  est bien un covecteur.

(GV.2.c) si  $0 < \delta \leq \bar{\delta}$ ,  $a \leq t \leq b$ , et  $\|x - \xi_*(t)\| \leq \delta$ , on a  $(x, t) \in \Omega$ , et il existe  $(M, \mu) \in D(t)$  tel que

$$\begin{aligned} & \|f_{\eta_*}(x, t) - f_{\eta_*}(\xi_*(t), t) - M \cdot (x - \xi_*(t))\| \\ & + \|\ell_{\eta_*}(x, t) - \ell_{\eta_*}(\xi_*(t), t) - \mu \cdot (x - \xi_*(t))\| \leq \delta k_\delta(t), \end{aligned}$$

(GV.2.d)  $\sup \left\{ \|M\| + \|\mu\| : (M, \mu) \in D(t) \right\} \leq k_{\bar{\delta}}(t)$  pour tout  $t \in [a, b]$ .

La généralisation suivante du théorème 3.4.4 combine les résultats de Halkin et Lojasiewicz :

**Théorème 3.4.5** Soit  $(X, I, \mathbf{f}) = (X, I, \{(f_\eta, \ell_\eta)\}_{\eta \in \mathcal{U}})$  un système de champs de vecteurs augmentés. Supposons que les hypothèses (A1), (A2), (A3".a), (A4), (A5".a) soient vérifiées. Soit  $S$  une sous-variété de classe  $C^1$  du produit  $X \times X$ . Notons  $\mathcal{Z} = \mathcal{Z}_{X, I, \mathbf{f}; a, b, S}$  et  $J = J_{X, I, \mathbf{f}; a, b, S}$ . Soit  $(\xi_*, \eta_*) \in \mathcal{Z}$  tel que  $J(\xi_*, \eta_*) \leq J(\xi, \eta)$  quel que soit  $(\xi, \eta) \in \mathcal{Z}$ . Soit  $D$  un générateur variationnel pour  $(\xi_*, \eta_*)$  et  $(X, I, \mathbf{f})$ . Il existe alors un couple  $(\pi, \pi_0)$  vérifiant toutes les conclusions du théorème 3.4.4, à l'exception de l'inclusion différentielle (40), qui doit être remplacée par la condition suivante : il existe une sélection mesurable  $t \rightarrow (M(t), \mu(t))$  de  $D$  vérifiant l'équation  $-\dot{\pi}(t) = \pi(t) \cdot D(t) - \pi_0 \cdot \mu(t)$  presque partout.  $\diamond$

**6. Une condition de transversalité plus générale.** Au lieu d'imposer une condition au bord de la forme  $\partial\xi \in S$ , où  $S$  est une sous-variété, on peut considérer des ensembles  $S$  plus généraux. Dans ce cas, la condition de transversalité est toujours formellement celle de (b) du théorème 3.3.1, mais on doit donner un sens précis à l'expression  $\ll (T_{\partial\xi_*} S)^\perp \gg$ . Nous interpréterons  $\ll (T_{\partial\xi_*} S)^\perp \gg$  comme le cône polaire de  $T_{\partial\xi_*} S$ , où  $T_{\partial\xi_*} S$  est un cône tangent à  $S$  au point  $\partial\xi_*$ . (Rappelons que le cône polaire d'un cône  $C$  dans un espace linéaire réel  $X$  de dimension finie est le sous-ensemble  $C^\dagger$  du dual  $X^\dagger$  défini par  $C^\dagger = \{\lambda \in X^\dagger : \langle \lambda, x \rangle \leq 0 \text{ pour tout } x \in C\}$ .) La notion de « cône tangent » se prête à plusieurs interprétations précises. On verra au §4 qu'à chaque calcul différentiel généralisé correspond une notion de cône (ou « multicône ») tangent à un ensemble en un point. Pour l'instant, nous choisissons la notion de cône tangent associée au CDG des différentielles classiques :

**Définition 3.4.6** Soit  $S$  un sous-ensemble d'un espace vectoriel réel  $X$  de dimension finie. On appelle cône tangent à  $S$  en un point  $s \in S$  un cône convexe  $C$  dans  $X$  tel qu'il existe  $n \in \mathbb{Z}_+$ , un cône convexe fermé  $D$  dans  $\mathbb{R}^n$ , un voisinage  $U$  de 0 dans  $\mathbb{R}^n$ , et une application continue  $F : U \cap D \rightarrow S$  telle que  $F(0) = s$ ,  $F$  est différentiable en 0, et  $DF(0) \cdot D = C$ .  $\diamond$

Avec cette définition de « cône tangent », le théorème 3.4.5 reste vrai.

**Remarque 3.4.7** Il y a des notions de « cône tangent » (par exemple, le cône tangent de Clarke) qui ne sont pas associées à des CDGs, et il y aussi des

définitions de cône normal où ce cône n'est pas le polaire d'un cône tangent (par exemple, des cônes normaux non convexes, comme le cône de Mordukhovich). Jusqu'à présent, nous n'avons pas réussi à dériver ces versions à partir de notre résultat général.  $\diamond$

**7. Champs de vecteurs discontinus.** L'hypothèse (A3".a) peut être remplacée par une condition encore plus faible. Il suffit que les champs de vecteurs  $f_\eta$  et les lagrangiens  $\ell_\eta$  vérifient une condition de « continuité intégrale » : la fonctionnelle  $(s, t, \xi) \rightarrow (\int_s^t f_\eta(\xi(r), r) dr, \int_s^t \ell_\eta(\xi(r), r) dr)$  doit être continue si  $\xi$  prend des valeurs dans l'espace des courbes absolument continues telles que  $\|\dot{\xi}(t)\| \leq \varphi(t)$  presque partout,  $\varphi$  étant une fonction intégrable telle que  $\|f_\eta(x, t)\| + |\ell_\eta(x, t)| \leq \varphi(t)$  pour tout  $x$  et presque tout  $t$ . Cette condition suffit pour démontrer toutes les propriétés des champs de vecteurs dont on a besoin pour la preuve du principe du maximum. (Par exemple, l'existence locale des trajectoires se démontre en utilisant le théorème de point fixe de Schauder.)

**8. Inclusions différentielles.** Les champs de vecteurs peuvent être remplacés par des inclusions différentielles, comme nous avons montré dans [18]. Ceci est possible en vertu d'une série de théorèmes de sélection (cf. for exemple [3] et [6]) qui impliquent, pour la classe des fonctions multivoques « presque semicontinues inférieurement » (cf. [18]), l'existence de un nombre suffisamment grand de sélections dans la classe des champs possédant la propriété de continuité intégrale. On arrive ainsi à démontrer tout ce qui est nécessaire pour prouver le principe du maximum, en dépit du fait que les fonctions multivoques presque semicontinues inférieurement (et même les fonctions multivoques continues) n'admettent pas en général de sélections continues.

**9. La géométrisation du problème.** Le problème de la *séparation de l'ensemble d'accessibilité d'un autre ensemble donné* est posé de la façon suivante : soit  $\Sigma = (I, X, \{f_\eta\}_{\eta \in \mathcal{U}})$  un système de champs de vecteurs (sans lagrangiens). Notons  $\mathcal{R}_{a,b}^\Sigma(q)$  l'ensemble d'accessibilité pour  $\Sigma$  sur un intervalle  $[a, b]$  à partir d'un point  $q$ . Soit  $S$  un sous-ensemble de  $X$ . Soit  $\xi_* : [a, b] \rightarrow X$  une trajectoire de  $\Sigma$  telle que  $\xi_*(b) \in S$ . On cherche alors de conditions nécessaires pour que  $\mathcal{R}_{[a,b]}^\Sigma(\xi_*(a)) \cap S$  soit égal à  $\{\xi_*(b)\}$ . La forme des conditions nécessaires est exactement celle des théorèmes pour le problème de contrôle optimal, sauf que cette fois il n'y a ni lagrangien ni multiplicateur anormal. Cette formulation « géométrique » du problème est évidemment plus élégante que la question de contrôle optimal, parce qu'elle ne fait intervenir qu'une classe d'objets fondamentaux — les champs de vecteurs — au lieu de deux classes — les champs de vecteurs et les lagrangiens —. On peut montrer par un calcul un peu long mais assez simple que le théorème « géométrique » de séparation implique le résultat pour le contrôle optimal. (Nous renvoyons le lecteur à [14] ou à [20] pour les détails). Ici nous nous bornerons à remarquer que la réduction du problème de minimisation à un problème de séparation se fait en « augmentant » la dynamique. On ajoute le « coût courant »  $x_0 = \int_a^{\cdot} \ell_\eta(\xi(t), t) dt$  aux autres variables d'état, et on montre que l'optimalité de  $\xi_*$  implique une condition de séparation pour le nouveau système. En appliquant la condition

nécessaire pour la séparation on obtient un multiplicateur  $(\pi, \pi_0)$ . Ceci explique clairement pourquoi le multiplicateur anormal devient nécessaire : il n'est que la composante de l'impulsion pour le nouveau système conjuguée à la variable d'état  $x_0$ .

## 4 Calcul des variations pour les flots

Nous allons maintenant présenter notre version récente du principe du maximum, qui contient et généralise tous les théorèmes du §3. Dans cette version les champs de vecteurs sont remplacés par des flots, l'« équation adjointe » s'écrit sous sa forme intégrée et devient une équation de transport associée aux différentielles du flot de référence, et les différentielles ordinaires sont remplacées par des différentielles généralisées.

### 4.1 Notations

Si  $\mathcal{C}$  est une catégorie, et  $A, B$  sont des objets de  $\mathcal{C}$ , on notera  $\text{Hom}_{\mathcal{C}}(A, B)$  l'ensemble de tous les  $\mathcal{C}$ -morphisms de  $A$  dans  $B$ . Nous appellerons  $\mathcal{C}$ -*multimorphisme* de  $A$  dans  $B$  un sous-ensemble de  $\text{Hom}_{\mathcal{C}}(A, B)$ , et nous noterons  $M\text{Hom}_{\mathcal{C}}(A, B)$  l'ensemble de tous les  $\mathcal{C}$ -multimorphisms de  $A$  dans  $B$ . Si  $A, B, C$  sont des objets de  $\mathcal{C}$ ,  $\mu \in M\text{Hom}_{\mathcal{C}}(A, B)$  et  $\nu \in M\text{Hom}_{\mathcal{C}}(B, C)$ , le  $\mathcal{C}$ -*multimorphisme composé*  $\nu \circ \mu$  est l'ensemble

$$\nu \circ \mu \stackrel{\text{d\u00e9f}}{=} \left\{ g \circ f : g \in \text{Hom}_{\mathcal{C}}(B, C), f \in \text{Hom}_{\mathcal{C}}(A, B) \right\}.$$

On notera  $M\mathcal{C}$  la catégorie définie par les conditions suivantes : (a) les objets de  $M\mathcal{C}$  sont les objets de  $\mathcal{C}$ , et (b) l'ensemble  $\text{Hom}_{M\mathcal{C}}(A, B)$  des morphismes entre deux objets  $A, B$ , de  $\mathcal{C}$  est donné par la formule

$$\text{Hom}_{M\mathcal{C}}(A, B) \stackrel{\text{d\u00e9f}}{=} M\text{Hom}_{\mathcal{C}}(A, B).$$

L'opération  $M$  peut être itérée. Ainsi, les éléments de  $\text{Hom}_{MM\mathcal{C}}(A, B)$  seront appelés *multimultimorphisms* de  $A$  dans  $B$ . Un multimultimorphisme de  $A$  dans  $B$  est donc un ensemble de sous-ensembles de  $\text{Hom}_{\mathcal{C}}(A, B)$ .

Nous appellerons *catégorie m-topologique* une catégorie  $\mathcal{C}$  telle que pour tout couple  $(A, B)$  d'objets de  $\mathcal{C}$  l'ensemble  $\text{Hom}_{\mathcal{C}}(A, B)$  est muni d'une structure d'espace topologique séparé, et la composition

$$\text{Hom}_{\mathcal{C}}(B, C) \times \text{Hom}_{\mathcal{C}}(A, B) \ni (\nu, \mu) \rightarrow \nu \circ \mu \in \text{Hom}_{\mathcal{C}}(A, C)$$

est une application continue quels que soient les objets  $A, B, C$  de  $\mathcal{C}$ . Si  $\mathcal{C}$  est une catégorie m-topologique, il est évident que le multimorphisme composé de deux multimorphisms compacts est compact. Dans ce cas, nous noterons  $M_c\mathcal{C}$  la catégorie  $\mathcal{D}$  telle que (a) les objets de  $\mathcal{D}$  sont les objets de  $\mathcal{C}$ , et (b) l'ensemble  $\text{Hom}_{\mathcal{D}}(A, B)$  des morphismes entre deux objets  $A, B$ , de  $\mathcal{D}$  est l'ensemble des  $\mathcal{C}$ -multimorphisms compacts non vides de  $A$  dans  $B$ .

Une *application multivoque* est une triple  $F = (A, B, G)$  telle que  $A, B$  sont des ensembles et  $G \subseteq A \times B$ . Si  $F = (A, B, G)$  est une application multivoque, les ensembles  $A, B, G$  seront appelés, respectivement, la *source*, la *cible*, et le *graphe* de  $F$ , et seront notés  $\text{So}(F)$ ,  $\text{Ci}(F)$ ,  $\text{Gr}(F)$ . Les ensembles

$$\text{Dom}(F) \stackrel{\text{déf}}{=} \{x : (\exists y)((x, y) \in \text{Gr}(F))\}, \quad \text{Im}(F) \stackrel{\text{déf}}{=} \{y : (\exists x)((x, y) \in \text{Gr}(F))\}$$

sont, respectivement, le *domaine* et l'*image* de  $F$ .

Si  $F$  est une application multivoque, et  $x$  est un objet quelconque, on notera  $F(x)$  l'ensemble  $\{y : (x, y) \in \text{Gr}(F)\}$ . Il s'ensuit que  $\text{Dom}(F) = \{x : F(x) \neq \emptyset\}$  et  $\text{Im}(F) = \bigcup_{x \in \text{Dom}(F)} F(x)$ .

Si  $A_i, i = 1, 2, 3$ , sont des ensembles, et  $F_i$  sont des applications multivoques de  $A_i$  dans  $A_{i+1}$  pour  $i = 1, 2$ , on définit l'*application multivoque composée*  $F_2 \circ F_1$  par la formule  $F_2 \circ F_1 = (A_1, A_3, G)$ , où

$$G = \text{Gr}(F_2 \circ F_1) = \{(x_1, x_3) : (\exists x_2)((x_1, x_2) \in \text{Gr}(F_1) \wedge (x_2, x_3) \in \text{Gr}(F_2))\}.$$

Il s'ensuit que  $F_2 \circ F_1$  est une application multivoque de  $A_1$  dans  $A_3$ . La composition d'applications multivoques est évidemment une opération associative.

Nous noterons  $\mathcal{E}\mathcal{N}\mathcal{S}-\mathcal{M}\mathcal{V}$  la catégorie  $\mathcal{C}$  telle que (a) les objets de  $\mathcal{C}$  sont les ensembles et (b) les  $\mathcal{C}$ -morphismes d'un ensemble  $A$  dans un ensemble  $B$  sont les applications multivoques de  $A$  dans  $B$ . Il s'ensuit que, si  $A, B$  sont des ensembles, alors  $\text{Hom}_{\mathcal{E}\mathcal{N}\mathcal{S}-\mathcal{M}\mathcal{V}}(A, B)$  est l'ensemble de toutes les applications multivoques de  $A$  dans  $B$ .

On notera  $\mathbf{C}^1\mathcal{M}\mathcal{V}$  la catégorie caractérisé par les deux conditions suivantes : (a) les objets de  $\mathbf{C}^1\mathcal{M}\mathcal{V}$  sont les variétés différentiables de classe  $C^1$ , (b) les  $\mathbf{C}^1\mathcal{M}\mathcal{V}$ -morphisms d'un objet  $A$  dans un objet  $B$  sont les applications multivoques de  $A$  dans  $B$ .

L'expression  $\mathcal{E}\mathcal{V}_{\mathbb{R}}\mathcal{D}\mathcal{F}$  désignera la catégorie  $\mathcal{C}$  telle que (a) les objets de  $\mathcal{C}$  sont les espaces vectoriels réels de dimension finie et (b) les  $\mathcal{C}$ -morphisms sont les applications  $\mathbb{R}$ -linéaires,

Si  $I$  est un ensemble totalement ordonné, on notera  $\leq_I$  (ou, tout simplement,  $\leq$ ) la relation d'ordre de  $I$ . Le symbole  $\hat{I}$  désignera l'ensemble

$$\hat{I} \stackrel{\text{déf}}{=} \{(s, t) \in I \times I : s \leq_I t\}.$$

Si  $J \subseteq I$ ,  $J$  est évidemment muni d'une structure naturelle d'ensemble totalement ordonné induite par  $\leq_I$ , qui sera notée  $\leq_J$ . Un sous-ensemble  $J$  de  $I$  est un *intervalle* si les relations  $r \leq s \leq t$ ,  $r \in J$  et  $t \in J$  impliquent  $s \in J$ . Si  $a, b \in I$  et  $a \leq b$ , on notera  $] - \infty, a]$ ,  $[a, b]$ ,  $[b, +\infty[$ , respectivement, les « intervalles fermés »  $\{t \in I : t \leq a\}$ ,  $\{t \in I : a \leq t \leq b\}$ ,  $\{t \in I : b \leq t\}$ .

Un ensemble totalement ordonné  $I$  est assimilé canoniquement à la catégorie  $\mathcal{C}_I$  telle que les objets de  $\mathcal{C}_I$  sont les éléments de  $I$ , et  $\text{Hom}_{\mathcal{C}_I}(s, t) = \emptyset$  si  $s \not\leq_I t$ ,  $\text{Hom}_{\mathcal{C}_I}(s, t) = \{o_{s,t}\}$  si  $s \leq_I t$  ( $o_{s,t}$  étant un objet arbitraire, par exemple le couple  $(s, t)$ ).

## 4.2 Calculs différentiels généralisés

Si  $n, m \in \mathbb{Z}_+$ , on notera  $\mathcal{A}_{n,m}$  l'ensemble de tous les couples  $(F, S)$  tels que  $F$  est une application multivoque de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  et  $S$  est un sous-ensemble de  $\mathbb{R}^n$ . Nous noterons  $\mathcal{A}$  l'ensemble  $\bigcup_{n,m \in \mathbb{Z}_+} \mathcal{A}_{n,m}$ .

**Definition 4.2.1** On appellera *calcul différentiel généralisé* (abrév.: « CDG ») une application  $\mathcal{D}$  possédant les propriétés suivantes :

- (1)  $\mathcal{D}$  associe à chaque couple  $(F, S) \in \mathcal{A}$  un ensemble  $\mathcal{D}(F; S)$  d'ensembles compacts non vides d'applications linéaires de  $\text{So}(F)$  dans  $\text{Ci}(F)$ <sup>12</sup>.
- (2) Si  $n, m \in \mathbb{Z}_+$ ,  $(F, S)$  et  $(F', S')$  appartiennent à  $\mathcal{A}_{n,m}$ ,  $K \subseteq K' \subseteq \mathbb{R}^{m \times n}$ ,  $\text{Gr}(F) \subseteq \text{Gr}(F')$ ,  $K'$  est compact, et  $K \in \mathcal{D}(F; S)$ , alors  $K' \in \mathcal{D}(F'; S)$ .
- (3) Si  $n, m \in \mathbb{Z}_+$ ,  $(F, S) \in \mathcal{A}_{n,m}$ ,  $(F', S') \in \mathcal{A}_{n,m}$ ,  $U$  et  $U'$  sont des voisinages ouverts de 0 dans  $\mathbb{R}^n$ ,  $V$  et  $V'$  sont des voisinages ouverts de 0 dans  $\mathbb{R}^m$ ,  $\Phi, \Psi$  sont des difféomorphismes de classe  $C^1$  de  $U$  sur  $U'$  et de  $V$  sur  $V'$ , respectivement<sup>13</sup>, tels que

$$\Phi(0) = 0, \quad \Psi(0) = 0, \quad \text{Gr}(\Psi^{-1} \circ F' \circ \Phi) = (U \times V) \cap \text{Gr}(F), \quad \text{et} \quad \Phi(U \cap S) = U' \cap S',$$

alors

$$D\Psi^{-1}(0) \circ \mathcal{D}(F', S') \circ D\Phi(0) = \mathcal{D}(F, S).$$

- (4) Si  $F \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ ,  $F(0) = 0$ ,  $S \subseteq \mathbb{R}^n$ , et il existe  $\nu \in \mathbb{Z}_+$  et une application multivoque  $G$  de  $\mathbb{R}^n$  dans  $\mathbb{R}^\nu$  tels que  $\mathcal{D}(G; S) \neq \emptyset$ , alors  $\{DF(0)\} \in \mathcal{D}(F; S)$ .
- (5) Si  $n, m \in \mathbb{Z}_+$ ,  $S$  est un cône convexe fermé dans  $\mathbb{R}^n$  ou  $S = \emptyset$ , et  $F$  est une application linéaire de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$ , alors  $\{F\} \in \mathcal{D}(F; S)$ .
- (6) Si (a)  $(F_1, S_1) \in \mathcal{A}_{n_1, n_2}$  et  $(F_2, S_2) \in \mathcal{A}_{n_2, n_3}$ , (b)  $\Lambda_1 \in \mathcal{D}(F_1; S_1)$  et  $\Lambda_2 \in \mathcal{D}(F_2; S_2)$ , (c)  $F_1(S_1) \subseteq S_2$ , et (d)  $S_2 = \mathbb{R}^{n_2}$  ou  $F_1$  est univoque, alors  $\Lambda_2 \circ \Lambda_1 \in \mathcal{D}(F_2 \circ F_1; S_1)$ .
- (7) Si  $(F_i, S_i) \in \mathcal{A}_{n_i, m_i}$  et  $\Lambda_i \in \mathcal{D}(F_i; S_i)$  pour  $i = 1, 2$ , alors  $\Lambda_1 \times \Lambda_2$  appartient à  $\mathcal{D}(F_1 \times F_2; S_1 \times S_2)$ .  $\diamond$

<sup>12</sup>Autrement dit, les éléments de  $\mathcal{D}(F; S)$  sont des sous-ensembles compacts non vides de  $\mathbb{R}^{m \times n}$ , où  $n, m$  sont tels que  $(F, S) \in \mathcal{A}_{n,m}$ . Rappelons que notre définition de la notion d'« application multivoque » garantit que, si  $(F, S) \in \mathcal{A}$ , alors la source  $\mathbb{R}^n = \text{So}(F)$  et la cible  $\mathbb{R}^m = \text{Ci}(F)$  sont complètement déterminées par  $F$ . Ainsi donc, les ensembles  $\mathcal{A}_{n,m}, \mathcal{A}_{n',m'}$  sont disjoints si  $(n', m') \neq (n, m)$ .

<sup>13</sup>Plus précisément,  $\Phi$  et  $\Psi$  sont des applications multivoques de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$  et de  $\mathbb{R}^m$  dans  $\mathbb{R}^m$ , telles que  $\text{Dom}(\Phi) = U$ ,  $\text{Dom}(\Psi) = V$ , et les restrictions  $\Phi|_U, \Psi|_V$ , de  $\Phi$  à  $U$  et de  $\Psi$  à  $V$  sont des difféomorphismes de classe  $C^1$  de  $U$  sur  $U'$  et de  $V$  sur  $V'$ .

On déduit de cette définition —en vertu, surtout, de la propriété d’invariance (3)— qu’un CDG  $\mathcal{D}$  admet toujours un prolongement canonique aux variétés, qui sera aussi noté  $\mathcal{D}$  : si  $M, N$  sont des variétés de classe  $C^1$ ,  $S \subseteq M$ ,  $x \in M$ ,  $y \in N$ , et  $F$  est une application multivoque de  $M$  dans  $N$ , alors  $\mathcal{D}(F; x, y; S)$  est un ensemble de sous-ensembles compacts non vides<sup>14</sup> de  $L(T_x M, T_y N)$ .

**Definition 4.2.2** Soit  $\mathcal{D}$  un CDG. On dira que  $\mathcal{D}$  possède la propriété de l’application directionnellement ouverte si la condition suivante est vérifiée :

(ADO) si  $n, m \in \mathbb{Z}_+$ ,  $(F, C) \in \mathcal{A}_{n,m}$ ,  $C$  est un cône convexe fermé dans  $\mathbb{R}^n$ ,  $v \in \mathbb{R}^m$ , et  $\Lambda \in \mathcal{D}(F; C)$ , alors la condition

$$v \in \text{Int}(L \cdot C) \quad \text{pour tout } L \in \Lambda \quad (42)$$

implique

- pour tout voisinage  $U$  de 0 dans  $\mathbb{R}^n$  il existe un cône convexe fermé  $D$  dans  $\mathbb{R}^m$  tel que  $v \in \text{Int}(D)$  et un  $r > 0$  tel que

$$D \cap \{y \in \mathbb{R}^m : \|y\| \leq r\} \subseteq F(C \cap U). \quad \diamond$$

**Remarque 4.2.3** Dans (ADO), on peut choisir  $C = \mathbb{R}^n$  et  $v = 0$ . Il s’ensuit que, si un CDG possède la propriété de l’application directionnellement ouverte, alors le théorème classique de l’application ouverte est vrai : si  $F$  est une application multivoque de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  et  $\Lambda \in \mathcal{D}(F; 0, 0; \mathbb{R}^n)$  est tel que tous les éléments de  $\Lambda$  sont des applications linéaires surjectives, alors pour tout voisinage  $U$  de 0 dans  $\mathbb{R}^n$  l’image  $F(U)$  est un voisinage de 0 dans  $\mathbb{R}^m$ .  $\diamond$

### 4.3 Quatre exemples de calculs différentiels généralisés

Nous donnons maintenant quatre exemples de calculs différentiels généralisés possédant la propriété de l’application directionnellement ouverte.

Le premier exemple est celui du CDG des *différentielles des applications de classe  $C^1$* , qui sera noté  $\mathcal{D}_{C^1}$ . Si  $n, m \in \mathbb{Z}_+$ ,  $S$  est un sous-ensemble de  $\mathbb{R}^n$  dont le germe en 0 coïncide avec le germe d’un ensemble fermé, et  $F$  est une application multivoque de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$ , alors  $\mathcal{D}_{C^1}(F; S)$  est l’ensemble de tous les sous-ensembles compacts non vides  $\Lambda$  de  $\mathbb{R}^{m \times n}$  tels qu’il existe  $L \in \Lambda$  vérifiant la condition suivante : il existe une application  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$  de classe  $C^1$ , telle que  $G(0) = 0$ , le germe en  $(0, 0)$  du graphe de  $G \upharpoonright S$  est inclus dans le germe en  $(0, 0)$  du graphe de  $F \upharpoonright S$ , et  $L = DG(0)$ .

Notre second exemple de CDG est celui des *différentielles classiques des applications continues*, qui sera noté  $\mathcal{D}_{DC-C^0}$ . La définition est presque identique à celle de  $\mathcal{D}_{C^1}$ , la seule différence étant que la condition que  $G \in C^1$  est remplacée par une propriété plus faible : on demande seulement que  $G$  soit continue au voisinage de 0 et différentiable en 0.

<sup>14</sup>Mais  $\mathcal{D}(F; x, y; S)$  peut bien être vide. Dans ce cas, on dira que  $F$  n’est pas  $\mathcal{D}$ -différentiable en  $(x, y)$  dans la direction de  $S$ .

Pour définir le CDG des *conteneurs de dérivées* — proposé par J. Warga, cf. [23] — il faudra d’abord définir la notion de « conteneur de dérivées » en 0 d’une application partielle  $f$  de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  telle que  $f(0) = 0$ . Soit  $S$  un sous-ensemble fermé de  $\mathbb{R}^n$ . On appellera *conteneur de dérivées de  $f$  en 0 dans la direction de  $S$*  un sous-ensemble compact  $\Lambda$  de  $\mathbb{R}^{m \times n}$  vérifiant : pour tout voisinage  $W$  de  $\Lambda$  dans  $\mathbb{R}^{m \times n}$  il existe un voisinage compact  $U$  de 0 dans  $\mathbb{R}^n$  tel que  $U \cap S \subseteq \text{Dom}(f)$  et une suite  $\{h_j\}_{j \in \mathbb{N}}$  d’applications de classe  $C^1$  de  $U \cap S$  dans  $\mathbb{R}^m$  telle que  $h_j \rightarrow f$  uniformément sur  $U \cap S$ , et  $Dh_j(x) \in W$  pour tout  $x \in U \cap S$ ,  $j$  quelconque. (La notion d’« application de classe  $C^1$  sur un fermé » est entendue dans le sens de Whitney, cf. par exemple Tougeron [22].) Lorsqu’il existe, le conteneur de dérivées n’est jamais unique, parce que tout sous-ensemble compact de  $\mathbb{R}^{m \times n}$  contenant un conteneur de dérivées donné est aussi un conteneur de dérivées. Pour que  $f$  admette un conteneur de dérivées dans la direction de  $S$  il faut et suffit que  $f$  soit définie et lipschitzienne sur un voisinage relatif de 0 dans  $S$ . Lorsque cette condition est vérifiée, il se peut qu’il n’existe pas de choix canonique d’un « conteneur de dérivées minimal ». Par contre, il existe toujours un *conteneur de dérivées minimal parmi tous les conteneurs de dérivées convexes*, qui n’est autre que le jacobien généralisé de Clarke de  $f$  en 0 dans la direction de  $S$ . En général il y aura aussi des conteneurs de dérivées non convexes strictement plus petits que le jacobien généralisé.

Pour définir le CDG des conteneurs de dérivées de Warga — qui sera noté  $\mathcal{D}_{CDW}$  — il suffit de copier la définition de  $\mathcal{D}_{C^1}$ , en remplaçant « de classe  $C^1$  » par « lipschitzien », et la différentielle ordinaire par les conteneurs de dérivées. On arrive ainsi à la formulation suivante : si  $n, m \in \mathbb{Z}_+$ ,  $S$  est un sous-ensemble de  $\mathbb{R}^n$  dont le germe en 0 coïncide avec le germe d’un ensemble fermé, et  $F$  est une application multivoque de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$ , alors  $\mathcal{D}_{CDW}(F; S)$  est l’ensemble de tous les sous-ensembles compacts non vides  $\Lambda$  de  $\mathbb{R}^{m \times n}$  tels qu’il existe une application localement lipschitzienne  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$  satisfaisant aux conditions : (a)  $G(0) = 0$ , (b) le germe en  $(0, 0)$  du graphe de  $G \upharpoonright S$  est inclus dans le germe en  $(0, 0)$  du graphe de  $F \upharpoonright S$ , et (c)  $\Lambda$  est un conteneur de dérivées de  $G \upharpoonright S$  en 0.

Finalement, notre quatrième exemple est le CDG des « semi-différentielles » des applications continues, que nous noterons  $\mathcal{D}_{SD-C^0}$ . (Nous avons défini la notion de « semidifférential », et celle encore plus générale de « multidifférential », dans [15, 16, 18, 19], pour les applications multivoques. Ici nous nous contenterons de rappeler la définition des « semidifférentiels » — qui est beaucoup plus simple que celle des « multidifférentiels » — et nous ne le ferons que pour le cas des applications univoques continues<sup>15</sup>, qui est plus simple que celui des applications multivoques.) H. Halkin s’aperçut en 1974 d’un très grave défaut dans les théories des dérivées généralisées qui avaient été proposées à l’époque suivant les idées de l’« analyse non lisse ». Le point de départ de ces théories était la notion de dérivée d’une application continument différentiable, qu’elles généralisait aux applications lipschitziennes. Mais elles ne s’accordaient

<sup>15</sup>Pour les applications univoques, ce que nous appelons « multidifférential » fut en fait introduit par Halkin en 1974 sous le nom de « screen », cf. Halkin [8] et aussi [9, 10]

pas bien avec une autre généralisation, tout à fait naturelle et très classique, dans une direction différente : celle des différentielles des applications qui ne sont pas de classe  $C^1$ . (Par exemple, la fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  définie par  $f(x) = |x|^{3/2} \sin \frac{1}{x}$  est différentiable en 0 mais pas lipschitzienne au voisinage de 0. La fonction  $g : \mathbb{R} \rightarrow \mathbb{R}$  donnée par  $g(x) = |x|^2 \sin \frac{1}{x}$  est différentiable en 0 et aussi lipschitzienne au voisinage de 0, mais la différentielle classique en 0 s'annule, tandis que le gradient généralisé de Clarke en 0 — qui dans ce cas se confond avec le plus petit conteneur de dérivées de Warga en 0 — est l'intervalle  $[-1, 1]$ .) Pour remédier à cette faiblesse, il faudrait trouver un « calcul différentiel » contenant les différentielles classiques aussi bien que les conteneurs de dérivées de Warga, et possédant les « bonnes propriétés » d'un tel calcul, notamment la règle de composition des différentielles (c'est-à-dire la condition (6) de notre définition 4.2.1). Il s'ensuit de ces conditions que ce calcul doit abandonner tout espoir de pouvoir définir une « différentielle » unique, puisque, par exemple, la fonction  $g$  définie par  $g(x) = x^2 \sin \frac{1}{x}$  doit posséder au moins deux « dérivées » : celle de la théorie classique, c'est-à-dire 0, et le plus petit des conteneurs de dérivées, c'est-à-dire l'intervalle  $[-1, 1]$ . En revanche, le nouveau calcul devrait nécessairement être applicable à une classe strictement plus grande que la réunion des domaines d'applicabilité des autres théories : par exemple, la fonction  $h$  donnée par  $f(x) = |x| + |x|^{3/2} \sin \frac{1}{x}$  — qui n'est ni classiquement différentiable en 0 ni lipschitzienne au voisinage de 0 — devrait devenir « différentiable » en 0, de telle sorte que l'intervalle  $[-1, 1]$  soit une de ses dérivées en 0. (Rappelons que la règle de la somme des différentielles se déduit de la règle de composition des différentielles et du fait que si  $F$  est une fonction linéaire alors  $F$  — ou l'ensemble  $\{F\}$  — est une différentielle de  $F$ .) La théorie des « screens » de Halkin donne une solution à ce problème. Notre notion — légèrement différente et bien plus simple — de « semi-différentielle » en donne une autre. On obtient la définition des semi-différentielles par la même opération qui permet de passer des différentielles des fonctions linéaires aux différentielles des fonctions arbitraires, mais en remplaçant les fonctions linéaires par les fonctions lipschitziennes. Il en résulte la définition suivante : si  $n, m \in \mathbb{Z}_+$ ,  $S$  est un sous-ensemble de  $\mathbb{R}^n$  dont le germe en 0 coïncide avec le germe d'un ensemble fermé, et  $F$  est une application multivoque de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$ , alors  $\mathcal{D}_{SD-C^0}(F; S)$  est l'ensemble de tous les sous-ensembles compacts non vides  $\Lambda$  de  $\mathbb{R}^{m \times n}$  tels qu'il existe une application localement lipschitzienne  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$  et une application continue  $H : \mathbb{R}^n \rightarrow \mathbb{R}^m$  satisfaisant aux conditions : (a)  $G(0)=0$ , (b) le germe en  $(0,0)$  du graphe de  $H \upharpoonright S$  est inclus dans le germe en  $(0,0)$  du graphe de  $F \upharpoonright S$ , (c)  $\Lambda$  est un conteneur de dérivées de  $G \upharpoonright S$  en 0, et (d)  $\lim_{x \rightarrow 0, x \in S \setminus \{0\}} \frac{\|G(x) - H(x)\|}{\|x\|} = 0$ . Autrement dit,  $\Lambda$  est une semi-différentielle de  $F$  en 0 dans la direction de  $S$  si le graphe de  $F \upharpoonright S$  contient le graphe d'une fonction continue  $H$  qui est approximable à un  $o(\|x\|)$  près par une application lipschitzienne  $G$  telle que  $\Lambda$  est un conteneur de dérivées de  $G \upharpoonright S$  en 0.

**Proposition 4.3.1** *Les théories  $\mathcal{D}_{C^1}$ ,  $\mathcal{D}_{DC-C^0}$ ,  $\mathcal{D}_{CDCW}$ , et  $\mathcal{D}_{SD-C^0}$ , sont des calculs différentiels généralisés possédant la propriété de l'application direction-*

nellement ouverte.

IDÉE DE LA DÉMONSTRATION. La vérification des conditions de la définition 4.2.1 est immédiate. Pour démontrer la propriété de l'application directionnellement ouverte, il suffit de le faire pour  $\mathcal{D}_{SD-C^0}$ , puisque cette théorie contient toutes les autres. On suppose que  $(F, C) \in \mathcal{A}_{n,m}$ ,  $C$  est un cône convexe fermé dans  $\mathbb{R}^n$ ,  $\Lambda \in \mathcal{D}_{SD-C^0}(F; S)$ ,  $v \in \mathbb{R}^m$ , et la condition (42) est vérifiée. On déduit de (42) que, si  $W$  est un voisinage suffisamment petit de  $\Lambda$  dans  $\mathbb{R}^{m \times n}$ , il existe un cône convexe fermé  $E$  dans  $\mathbb{R}^m$  tel que  $v \in \text{Int}(E)$ , un voisinage  $U$  de 0 dans  $\mathbb{R}^m$ , et une application continue  $W \times (U \cap E) \ni (L, y) \rightarrow \zeta(L, y) \in \mathbb{R}^n$ , lisse sur  $W \times ((U \cap E) \setminus \{0\})$ , positivement homogène de degré 1 par rapport à  $y$ , telle que  $\zeta(L, y) \in C$  et  $L \cdot \zeta(L, y) = w$  pour tout  $(L, y)$ . Choisissons  $G$  et  $H$  vérifiant les conditions de la définition des semi-différentielles. Soit  $\{G_j\}_{j \in \mathbb{N}}$  une suite de fonctions de classe  $C^2$  sur un voisinage  $V$  de 0 dans  $\mathbb{R}^n$  telle que les  $G_j$  convergent uniformément vers  $F$  sur  $V \cap C$  et que  $DG_j(x) \in W$  pour  $x \in V \cap C$  et  $j$  quelconque. On applique alors la technique de Ważewski (cf. [24]) : pour  $y \in U \cap E$ , soit  $t \rightarrow \xi_y^j(t)$  la solution du problème de Cauchy

$$\dot{\xi}(t) = \zeta(DG_j(\xi(t)), y), \quad \xi(0) = 0,$$

qui est unique, puisque  $\zeta$  est lisse et  $x \rightarrow DG_j(x)$  est de classe  $C^1$ . Il vient alors  $G_j(\xi_y^j(t)) = tw$ , puisque

$$\frac{d}{dt} \left( G_j(\xi_y^j(t)) \right) = DG_j(\xi_y^j(t)) \cdot \dot{\xi}_y^j(t) = DG_j(\xi_y^j(t)) \cdot \zeta(DG_j(\xi_y^j(t)), y) = y.$$

Il s'ensuit que  $G_j(\xi_y^j(1)) = y$ , pourvu que  $\xi_y(1)$  existe, ce qui sera toujours le cas si  $U$  est suffisamment petit. L'application continue  $y \rightarrow K_j(y) \stackrel{\text{déf}}{=} \xi_y^j(1)$  est donc une inverse à gauche de  $G_j$  (c'est-à-dire, elle vérifie  $G_j(K_j(y)) = y$ ), définie sur  $U \cap E$ , à valeurs dans  $C$ , et vérifiant l'inégalité  $\|K_j(y)\| \leq \kappa \|y\|$ ,  $\kappa$  étant une constante positive indépendante de  $j$ . On en déduit que

$$\|H(K_j(y)) - y\| \leq o(\|y\|) + \varepsilon_j \quad \text{pour } y \rightarrow 0, \quad y \in E$$

indépendamment de  $j$ , où  $\varepsilon_j \rightarrow 0$ . Soient  $D, D'$  des cônes convexes fermés dans  $\mathbb{R}^m$  tels que  $v \in \text{Int}(D)$ ,  $D \subseteq \text{Int}(D') \cup \{0\}$ , et  $D' \subseteq \text{Int}(E) \cup \{0\}$ . Il existe alors un  $\rho$  positif tel que la boule  $B(y) = \{z \in \mathbb{R}^m : \|z - y\| \leq \rho \|y\|\}$  vérifie  $B(y) \subseteq E$  quel que soit  $y \in D'$ . On peut évidemment supposer aussi que  $\rho \leq 1$ . En remplaçant  $U$  par un voisinage plus petit, on peut supposer que  $\|H(K_j(y)) - y\| \leq \frac{\rho \|y\|}{8} + \varepsilon_j$  pour  $y \in U \cap D'$ ,  $j$  quelconque. Il existe alors un  $\delta$  positif tel que si  $z \in D$  et  $\|z\| \leq \delta$  alors  $B(z) \subseteq D'$  et

$$\|H(K_j(y)) - y\| \leq \frac{\rho \|y\|}{8} + \varepsilon_j \leq \frac{(\rho + \rho^2) \|z\|}{8} + \varepsilon_j \leq \frac{\rho \|z\|}{4} + \varepsilon_j$$

pour tout  $y \in B(z)$ . Si  $z \neq 0$ , on peut choisir un  $j(z)$  tel que  $\varepsilon_j \leq \frac{\rho \|z\|}{4}$  pour  $j \geq j(z)$ . On aura donc  $\|H(K_j(y)) - y\| \leq \frac{\rho \|z\|}{2}$  quels que soient  $z \in D$  tel que

$0 < \|z\| \leq \delta$ ,  $y \in B(z)$ , et  $j \geq j(z)$ . Il s'ensuit que les applications continues  $B(z) \ni y \rightarrow \mu_{j,z}(y) \stackrel{\text{d\'ef}}{=} H(K_j(y)) \in \mathbb{R}^m$  sont telles que  $\|\mu_{j,z}(y) - y\| \leq \frac{r(z)}{2}$  pour tout  $y \in B(z)$ , pourvu que  $0 < \|z\| \leq \delta$  et  $j \geq j(z)$ , où  $r(z)$  est le rayon de la boule  $B(z)$ . En vertu d'un corollaire du théorème de point fixe de Brouwer, on a  $z \in \mu_j(B(z))$ . Il suffit donc de choisir  $j = j(z)$  et de prendre  $x = K_j(y)$  pour un  $y \in B(z)$  tel que  $\mu_j(y) = z$ , pour obtenir un  $x \in V \cap C$  tel que  $H(x) = z$ . Ainsi donc,  $z \in F(x)$ , ce qui achève la démonstration.  $\diamond$

## 4.4 Transversalité

**Definition 4.4.1** Soit  $X$  un espace vectoriel réel de dimension finie, et soient  $C_1, C_2$  deux cônes convexes dans  $X$ . On dira que  $C_1$  et  $C_2$  sont *transversaux* si  $C_1 - C_2 = X$ , où  $C_1 - C_2 = \{c_1 - c_2 : c_1 \in C_1, c_2 \in C_2\}$ . On dira que  $C_1$  et  $C_2$  sont *fortement transversaux* s'ils sont transversaux et  $C_1 \cap C_2$  contient au moins un vecteur non nul.  $\diamond$

On appellera *multicône* (resp. *multicône fermé*, *multicône convexe*) dans un espace vectoriel réel  $X$  un ensemble non vide de cônes (resp. cônes fermés, cônes convexes). Si  $\Lambda$  est un ensemble d'applications linéaires d'un autre espace  $Y$  dans  $X$ , et  $C$  est un cône dans  $Y$ , on notera  $\Lambda \cdot C$ , ou  $\Lambda(C)$ , le multicône  $\{L(C) : L \in \Lambda\}$ , qui est évidemment convexe quand  $C$  est convexe.

**Definition 4.4.2** Soit  $n \in \mathbb{Z}_+$ , et soit  $S$  un sous-ensemble de  $\mathbb{R}^n$ . Soit  $\mathcal{D}$  un CDG. On appellera *multicône  $\mathcal{D}$ -tangente* à  $S$  en 0 un multicône  $\mathcal{C}$  dans  $X$  tel qu'il existe  $\nu \in \mathbb{Z}_+$ ,  $(F, C) \in \mathcal{A}_{\nu, n}$ , vérifiant les conditions suivantes : (a)  $F(C) \subseteq S$ , (b)  $C$  est un cône convexe fermé dans  $\mathbb{R}^\nu$  et (c) il existe  $\Lambda \in \mathcal{D}(F; C)$  tel que  $\mathcal{C} = \Lambda \cdot C$ .  $\diamond$

En vertu de l'invariance des CDGs par les difféomorphismes locaux de classe  $C^1$ , on parlera librement de « multicônes  $\mathcal{D}$ -tangents » à un sous-ensemble  $S$  d'une variété différentiable  $Q$  en un point  $q \in Q$ . Un tel multicône sera évidemment un ensemble de cônes convexes dans l'espace tangent  $T_q Q$ . Il est clair que deux ensembles  $S_1, S_2$  dont les germes en  $q$  coïncident auront les mêmes multicônes tangents.

**Definition 4.4.3** Soit  $X$  un espace vectoriel réel de dimension finie, et soient  $\mathcal{C}_1, \mathcal{C}_2$  deux multicônes convexes dans  $X$ . On dira que  $\mathcal{C}_1$  et  $\mathcal{C}_2$  sont *transversaux* si  $C_1$  et  $C_2$  sont transversaux quels que soient  $C_1 \in \mathcal{C}_1, C_2 \in \mathcal{C}_2$ . On dira que  $\mathcal{C}_1$  et  $\mathcal{C}_2$  sont *fortement transversaux* s'ils sont transversaux et il existe une fonctionnelle linéaire  $\mu$  sur  $X$  telle que

$$C_1 \cap C_2 \cap \{x \in X : \mu(x) > 0\} \neq \emptyset \quad \text{quels que soient } C_1 \in \mathcal{C}_1, C_2 \in \mathcal{C}_2. \quad \diamond$$

**Théorème 4.4.4** Soit  $\mathcal{D}$  un CDG possédant la propriété de l'application directionnellement ouverte. Soient  $S_1, S_2$  deux sous-ensembles d'une variété  $Q$  de classe  $C^1$ , et soient  $\mathcal{C}_1, \mathcal{C}_2$  des multicônes  $\mathcal{D}$ -tangents à  $S_1, S_2$  en un point  $q \in Q$ .

Pour qu'il existe une suite  $\{q_j\}_{j \in \mathbb{N}}$  convergeant vers  $q$  telle que  $q_j \in S_1 \cap S_2$  et  $q_j \neq q$  pour tout  $j$ , il suffit que  $\mathcal{C}_1$  et  $\mathcal{C}_2$  soient fortement transversaux.

IDÉE DE LA DÉMONSTRATION. On peut supposer que  $Q = \mathbb{R}^n$ ,  $q = 0$ . Pour  $i = 1, 2$ , choisissons  $\nu_i \in \mathbb{Z}_+$ ,  $(F_i, E_i) \in \mathcal{A}_{\nu_i, n}$ ,  $\Lambda_i \in \mathcal{D}(F_i; C_i)$  tels que les  $E_i$  sont des cônes convexes fermés,  $\Lambda_i \cdot E_i = C_i$ , et  $F_i(C_i) \subseteq S_i$ .

Soit  $v \in \mathbb{R}^n$  tel que  $C_1 \cap C_2 \cap \{x \in \mathbb{R}^n : \langle v, x \rangle > 0\} \neq \emptyset$  quels que soient  $C_1 \in \mathcal{C}_1$ ,  $C_2 \in \mathcal{C}_2$ . Définissons une application multivoque  $H$  de  $\mathbb{R}^{\nu_1} \times \mathbb{R}^{\nu_2}$  dans  $\mathbb{R}^n \times \mathbb{R}$  par la formule

$$H(y_1, y_2) \stackrel{\text{déf}}{=} \{(x_1 - x_2, \mu(x_1)) : x_1 \in F_1(y_1), x_2 \in F_2(y_2)\}.$$

Soit  $\Xi$  l'ensemble de toutes les applications linéaires

$$\mathbb{R}^{\nu_1} \times \mathbb{R}^{\nu_2} \ni (z_1, z_2) \rightarrow h_{L_1, L_2}(z_1, z_2) \stackrel{\text{déf}}{=} (L_1 z_1 - L_2 z_2, \mu(L_1 z_1)) \in \mathbb{R}^n \times \mathbb{R},$$

pour  $(L_1, L_2) \in \Lambda_1 \times \Lambda_2$ . Notons  $E = E_1 \times E_2$ ,  $\omega = (0, 1) \in \mathbb{R}^n \times \mathbb{R}$ .

On vérifie aisément que  $\Xi \in \mathcal{D}(H; E)$ . Un calcul algébrique très simple montre ensuite que l'hypothèse de transversalité forte des  $\mathcal{C}_i$  équivaut exactement à la condition de nonsingularité suivante :  $\omega \in \text{Int}(\xi \cdot E)$  quel que soit  $\xi \in \Xi$ . En vertu de la propriété de l'application directionnellement ouverte, on déduit qu'il existe un nombre positif  $\rho$  tel que  $(0, r) \in H(E)$  si  $0 \leq r \leq \rho$ . In en résulte que pour chaque  $r \in ]0, \rho]$  il existe un couple  $(y_1(r), y_2(r))$  appartenant à  $E$  tel que  $(0, r) \in H(y_1(r), y_2(r))$ . Autrement dit, il existe  $x_1(r) \in F_1(y_1(r))$ ,  $x_2(r) \in F_2(y_2(r))$ , tels que  $(0, r) = (x_1(r) - x_2(r), \mu(x_1(r)))$ . Il s'ensuit d'abord que  $x_1(r) = x_2(r)$ , ce qui montre que le point  $x(r) = x_1(r)$  appartient à  $F_1(E_1) \cap F_2(E_2)$  et donc à  $S_1 \cap S_2$ , et puis que  $\mu(x_1(r)) = r$ , ce qui montre que  $x_1(r) \neq 0$ . Nous avons donc bien réussi à trouver un point  $x$  tel que  $x \in S_1 \cap S_2$  et  $x \neq 0$ .

Notre raisonnement étant applicable à l'intersection de  $S_1 \cap S_2$  avec un voisinage quelconque de  $0$ , l'existence de la suite  $\{q_j\}$  est démontrée.  $\diamond$

## 4.5 Flots

**Definition 4.5.1** Soit  $\mathcal{C}$  une catégorie. On appelle *flot dans  $\mathcal{C}$*  ou  *$\mathcal{C}$ -flot*, un couple  $\Phi = (I, \Phi)$  telle que  $I$  est un ensemble totalement ordonné et  $\Phi$  est un foncteur covariant de  $\mathcal{C}_I$  dans  $\mathcal{C}$ . Si  $\Phi = (I, \Phi)$  est un  $\mathcal{C}$ -flot, l'ensemble  $I$  sera appelé *l'intervalle de temps* de  $\Phi$ . On notera  $Q_t^\Phi$  — au lieu de  $\Phi(t)$  — l'objet de  $\mathcal{C}$  associé à  $t \in I$  par le foncteur  $\Phi$ . La famille  $Q^\Phi = \{Q_t^\Phi\}_{t \in I}$  sera appelée la *famille d'espaces d'états* du flot  $\Phi$ , et  $Q_t^\Phi$  est l'*espace d'états de  $\Phi$  au temps  $t$* . On écrira  $\Phi_{t,s}$  au lieu de  $\Phi(o_{s,t})$ .  $\diamond$

Il s'ensuit qu'un  $\mathcal{C}$ -flot est déterminé par les données suivantes :

1. un ensemble totalement ordonné  $I$ ,

2. une famille paramétrisée  $Q = \{Q_t\}_{t \in I}$  d'objets de  $\mathcal{C}$ ,
3. une famille paramétrisée  $\Phi = \{\Phi_{t,s}\}_{(s,t) \in \hat{I}}$  telle que  $\Phi_{t,s}$  est un  $\mathcal{C}$ -morphisme de  $Q_s$  dans  $Q_t$  pour tout  $(s,t) \in \hat{I}$ ,  $\Phi_{t,t} = \mathbb{1}_{Q_t}$  pour tout  $t \in I$ , et l'identité  $\Phi_{t_3,t_2} \circ \Phi_{t_2,t_1} = \Phi_{t_3,t_1}$  est vérifiée chaque fois que  $t_i \in I$  pour  $i = 1, 2, 3$  et  $t_1 \leq t_2 \leq t_3$ .

Nous assimilerons donc un  $\mathcal{C}$ -flot  $\Phi$  à la triple  $(I, Q, \Phi)$ , où  $Q = Q^\Phi$ , et nous écrirons parfois  $\Phi = (I, Q, \Phi)$  au lieu de  $\Phi = (I, \Phi)$ . Nous dirons aussi que le couple  $(I, Q)$  est la *famille d'espaces d'états* de  $\Phi$ , et que  $\Phi$  est un  $\mathcal{C}$ -flot sur  $(I, Q)$ .

On utilisera des noms spéciaux pour les flots dans des catégories particulières. Par exemple:

- Lorsque  $\mathcal{C}$  est la catégorie  $\mathcal{ENS}\text{-}\mathcal{MV}$ , on dira tout simplement *flot* au lieu de «  $\mathcal{C}$ -flot » ou «  $\mathcal{ENS}\text{-}\mathcal{MV}$ -flot ».
- Si  $\mathcal{C}$  est une catégorie, un flot dans  $M\mathcal{C}$  sera appelé *multiflot dans  $\mathcal{C}$* , ou  *$\mathcal{C}$ -multiflot*, et un flot dans  $MM\mathcal{C}$  sera appelé *multimultiflot dans  $\mathcal{C}$* , ou  *$\mathcal{C}$ -multimultiflot*.
- Si  $\mathcal{C}$  est une catégorie m-topologique, un flot dans  $M_c\mathcal{C}$  sera appelé *multiflot compact dans  $\mathcal{C}$* , ou  *$\mathcal{C}$ -multiflot compact*, et un flot dans  $MM_c\mathcal{C}$  sera appelé *multimultiflot compact dans  $\mathcal{C}$* , ou  *$\mathcal{C}$ -multimultiflot compact*.
- Un flot dans la catégorie  $\mathcal{EV}_{\mathbb{R}}\mathcal{DF}$  sera appelé *flot linéaire*.
- Un flot dans  $M\mathcal{EV}_{\mathbb{R}}\mathcal{DF}$  sera appelé *multiflot linéaire*.
- Un flot dans  $MM\mathcal{EV}_{\mathbb{R}}\mathcal{DF}$  sera appelé *multimultiflot linéaire*.
- Un flot dans  $M_c\mathcal{EV}_{\mathbb{R}}\mathcal{DF}$  sera appelé *multiflot linéaire compact*.

Si  $\Phi = (I, Q, \Phi)$  est un flot dans une catégorie  $\mathcal{C}$ , et  $J \subseteq I$ , on notera  $\Phi \lceil J$  la restriction de  $\Phi$  à  $J$ , c'est-à-dire le  $\mathcal{C}$ -flot  $\Phi \lceil J \stackrel{\text{déf}}{=} (J, Q \lceil J, \Phi \lceil J)$ , où

$$Q \lceil J = \{Q_t\}_{t \in J}, \quad \Phi \lceil J = \{\Phi_{t,s}\}_{(s,t) \in \hat{J}}.$$

**Definition 4.5.2** Soit  $\Phi = (I, Q, \Phi)$  un flot dans une catégorie  $\mathcal{C}$  dont les objets sont des ensembles et les morphismes sont des applications multivoques. On appellera *trajectoire* de  $\Phi$  une application  $\xi$  telle que  $\text{Dom}(\xi) = I$ ,  $\xi(t) \in Q_t$  pour tout  $t \in I$ , et  $\xi(t) \in \Phi_{t,s}(\xi(s))$  quel que soit  $(s,t) \in \hat{I}$ . On notera  $\text{Traj}(\Phi)$  l'ensemble de toutes les trajectoires de  $\Phi$ .  $\diamond$

Si  $\Phi = (I, Q, \Phi)$  est un flot dans  $\mathbf{C}^1\mathcal{MV}$ , et  $\xi \in \text{Traj}(\Phi)$ , nous noterons  $X^{\Phi, \xi}$  la famille  $\{X_t\}_{t \in I}$ , où  $X_t$  est l'espace tangent  $T_{\xi(t)}Q_t$ .

**Exemple 4.5.3** Soit  $\mathbf{C}^1$  la catégorie dont les objets sont les variétés différentiables de classe  $C^1$  et les morphismes sont les applications de classe  $C^1$ . Soit  $\Phi = (I, Q, \Phi)$  un flot dans  $\mathbf{C}^1$ . Si  $\xi \in \text{Traj}(\Phi)$ , on définit un flot linéaire  $D(\Phi, \xi)$  — que nous appellerons la *différentielle de  $\Phi$  le long de  $\xi$*  — de la façon suivante :  $D(\Phi, \xi) = (I, X, D\Phi)$ , où (a)  $X = \{X_t\}_{t \in I}$ , (b)  $X = X^{\Phi, \xi}$ , et (c) si  $(s, t) \in \hat{I}$ ,  $D\Phi_{t,s}$  est la différentielle de  $\Phi_{t,s}$  en  $\xi(s)$ , qui est évidemment une application linéaire de  $X_s$  dans  $X_t$ .

Plus généralement, on peut définir un flot linéaire  $D(\Phi, \xi)$  de la même manière si les  $Q_t$  sont des variétés de classe  $C^1$  et chaque  $\Phi_{t,s}$  est définie au voisinage de  $\xi(s)$  et différentiable en  $\xi(s)$ .  $\diamond$

**Definition 4.5.4** Soit  $\Phi = (I, Q, \Phi)$  un multiflot dans une catégorie  $\mathcal{C}$ . On appellera *sélection* de  $\Phi$  un flot  $\Psi$  dans  $\mathcal{C}$  tel que  $\Psi = (I, Q, \Psi)$  et  $\Psi_{t,s} \in \Phi_{t,s}$  pour tout  $(s, t) \in \hat{I}$ . L'ensemble de toutes les sélections du multiflot  $\Phi$  sera noté  $\Gamma(\Phi)$ .  $\diamond$

On peut évidemment identifier l'ensemble  $\Gamma(\Phi)$  au sous-ensemble fermé du produit  $\prod_{(s,t) \in \hat{I}} \Phi_{t,s}$  dont les éléments sont les familles  $\Psi = \{\Psi_{t,s}\}_{(s,t) \in \hat{I}}$  telles que  $\Psi_{t,s} \circ \Psi_{s,r} = \Psi_{t,r}$  pour  $r, s, t \in I$  quelconques tels que  $r \leq s \leq t$ . Ainsi donc,  $\Gamma(\Phi)$  est muni d'une topologie naturelle, qu'on appellera la *topologie produit*.

Si  $J \subseteq I$  et  $\Psi = \{\Psi_{t,s}\}_{(s,t) \in \hat{I}} \in \Gamma(\Phi)$ , on notera  $\Psi \upharpoonright J$  la restriction de  $\Psi$  à  $J$ , c'est-à-dire l'élément  $\{\Psi_{t,s} : (s, t) \in \hat{J}\}$  de  $\Gamma(\Phi \upharpoonright J)$ .

**Proposition 4.5.5** Si  $\mathcal{C}$  est une catégorie  $m$ -topologique et  $\Phi = (I, Q, \Phi)$  est un flot dans  $M_c\mathcal{C}$  (c'est-à-dire un multiflot compact dans  $\mathcal{C}$ ), l'espace  $\Gamma(\Phi)$ , muni de la topologie produit, est compact et non vide. Si  $J$  est un sous-ensemble de  $I$ , l'application  $\Gamma(\Phi) \ni \Psi \rightarrow \Psi \upharpoonright J \in \Gamma(\Phi \upharpoonright J)$  est surjective.

PREUVE. La compacité de  $\Gamma(\Phi)$  est immédiate, puisque  $\Gamma(\Phi)$  est fermé dans le produit  $S = \prod_{(s,t) \in \hat{I}} \Phi_{t,s}$  et les  $\Phi_{t,s}$  sont compacts.

Démontrons maintenant que l'application

$$\Gamma(\Phi) \ni \Psi \rightarrow \Psi \upharpoonright J \in \Gamma(\Phi \upharpoonright J) \quad (43)$$

est surjective. Soit  $\Psi^J$  une sélection de  $\Phi \upharpoonright J$ . Il nous faut montrer qu'il existe une sélection  $\Psi$  de  $\Phi$  telle que  $\Psi \upharpoonright J = \Psi^J$ .

Si  $K \subseteq J$  et  $L \subseteq I$ , notons  $A(K, L)$  l'ensemble de tous les  $\Xi \in S$  vérifiant  $\Xi_{t,s} = \Psi^J_{t,s}$  pour tout  $(s, t) \in \hat{I} \cap (K \times K)$ , et  $\Xi_{t,s} \circ \Xi_{s,r} = \Xi_{t,r}$  pour  $r, s, t \in L$  quelconques tels que  $r \leq_I s \leq_I t$ . Il est clair que  $A(K, L)$  est toujours un sous-ensemble fermé de  $S$ .

Notre but est de démontrer que  $A(J, I) \neq \emptyset$ . Or, on a

$$A(J, I) = \bigcap_{L \in I_f, K \subseteq J \cap L} A(K, L),$$

où  $I_f$  désigne l'ensemble de tous les sous-ensembles finis de  $I$ . En outre, l'inclusion  $A(K_1 \cup K_2, L_1 \cup L_2) \subseteq A(K_1, L_1) \cap A(K_2, L_2)$  est toujours vérifiée. Il suffit donc de démontrer que  $A(K, L) \neq \emptyset$  si  $L \in I_f$  et  $K \subseteq J \cap L$ . Autrement dit, il suffit de démontrer la surjectivité en supposant que  $I$  soit fini. Puisque la surjectivité de l'application (43) est immédiate lorsque  $I$  est fini, notre conclusion est démontrée.

Montrons finalement que  $\Gamma(\Phi) \neq \emptyset$ . Si  $I = \emptyset$  on a  $\Gamma(\Phi) \neq \emptyset$ , puisque dans ce cas la famille vide appartient à  $\Gamma(\Phi)$ . Si  $I \neq \emptyset$ , on choisit un  $i \in I$  et on définit  $J = \{i\}$ . L'ensemble  $\Gamma(\Phi \upharpoonright J)$  est évidemment non vide. En vertu de la surjectivité de l'application (43), on a  $\Gamma(\Phi) \neq \emptyset$ , ce qui achève la démonstration.  $\diamond$

**Definition 4.5.6** Soit  $\mathcal{D}$  un calcul différentiel généralisé, et supposons que  $\Phi = (I, Q, \Phi)$  soit un flot dans  $\mathcal{C}^1MV$ . Soit  $\xi \in \text{Traj}(\Phi)$ . On dit que  $\Phi$  est *faiblement  $\mathcal{D}$ -différentiable le long de  $\xi$*  si  $\mathcal{D}(\Phi_{t,s}, \xi(s), \xi(t), Q_s) \neq \emptyset$  quel que soit  $(s, t) \in \hat{I}$ . On dit que  $\Phi$  est  *$\mathcal{D}$ -différentiable le long de  $\xi$*  s'il existe un multiflot compact linéaire  $\Lambda = (I, X, \Lambda)$  tel que  $X = X^{\Phi, t}$  et  $\Lambda_{t,s} \in \mathcal{D}(\Phi_{t,s}, \xi(s), \xi(t), Q_s)$  quel que soit  $(s, t) \in \hat{I}$ .  $\diamond$

**Remarque 4.5.7** Il existe des exemples de flots qui sont faiblement  $\mathcal{D}$ -différentiables le long d'une trajectoire sans être  $\mathcal{D}$ -différentiables le long de  $\xi$ .  $\diamond$

## 4.6 Variations

**Definition 4.6.1** Soit  $\mathcal{C}$  une catégorie et supposons que  $\Phi = (I, Q, \Phi)$  soit un  $\mathcal{C}$ -flot. Une *quasi-variation de  $\Phi$  dans  $\mathcal{C}$*  (ou  *$\mathcal{C}$ -quasi-variation de  $\Phi$* ) est un couple  $(C, \Psi)$ , tel que  $C$  est un cône convexe fermé dans un espace vectoriel réel  $E_C$  de dimension finie,  $\text{Int}_{E_C}(C) \neq \emptyset$ , et  $\Psi = \{\Psi^c\}_{c \in C}$  est une famille telle que (a)  $(I, Q, \Psi^c)$  est un  $\mathcal{C}$ -flot pour tout  $c \in C$ , et (b)  $\Psi^0 = \Phi$ . Le cône  $C$  est appelé le *cône des paramètres* de la quasi-variation  $\Psi$ . Une *variation de  $\Phi$  dans  $\mathcal{C}$*  (ou  *$\mathcal{C}$ -variation de  $\Phi$* ) est un germe<sup>16</sup> en 0 de quasi-variations ayant le même cône paramétrique. Autrement dit, on identifie deux quasi-variations  $(C_1, \Psi_1)$ ,  $(C_2, \Psi_2)$  de  $\Phi$ , si  $C_1 = C_2$  et il existe un voisinage relatif  $N$  de 0 dans  $C_1$  tel que  $\Psi_1^c = \Psi_2^c$  quel que soit  $c \in N$ , et une *variation de  $\Phi$*  est une classe d'équivalence de quasi-variations de  $\Phi$  par rapport à cette identification. Si  $(C, \Psi)$  est une quasi-variation de  $\Phi$ , on notera  $[C, \Psi]$  le germe de  $\Psi$  en  $c = 0$ , c'est-à-dire la variation  $\mathbf{V}$  telle que  $(C, \Psi) \in \mathbf{V}$ .  $\diamond$

Si  $\mathbf{V} = [C, \Psi]$  est une  $\mathcal{C}$ -variation du flot  $\Phi = (I, Q, \Phi)$ , et  $\mathcal{C}$  est une catégorie dont les objets sont des ensembles et les morphismes sont des applications multivoques, on notera  $\langle \Psi_{t,s} \rangle$ , pour  $(s, t) \in \hat{I}$ , l'application multivoque  $\mu$  de  $E_C \times Q_s$  dans  $Q_t$  telle que  $\mu(c, q) = \Psi_{t,s}^c(q)$  si  $c \in C$ ,  $q \in Q_s$ , et  $\mu(c, q) = \emptyset$  si  $c \in E_C \setminus C$ ,  $q \in Q_s$ .

<sup>16</sup>Considérant les familles  $\Psi$  comme des applications  $c \rightarrow \Psi^c$  de  $C$  dans la classe des  $\mathcal{C}$ -flots.

Si  $E$  est un espace vectoriel réel de dimension finie, nous noterons  $E \times \mathcal{E}\mathcal{V}_{\mathbb{R}}\mathcal{D}\mathcal{F}$  la catégorie dont les objets sont les produits  $E \times X$ , où  $X$  est un espace vectoriel réel de dimension finie, et les morphismes d'un objet  $E \times X$  dans un autre objet  $E \times Y$  sont les applications linéaires  $L : E \times X \rightarrow E \times Y$  telles que  $\pi_{1,E,Y} \circ L = \pi_{1,E,X}$ , où  $\pi_{1,U,V}$  désigne la projection  $U \times V \ni (u, v) \rightarrow u \in U$ . La catégorie  $E \times \mathcal{E}\mathcal{V}_{\mathbb{R}}\mathcal{D}\mathcal{F}$  est évidemment m-topologique.

Les éléments de  $\text{Hom}_{E \times \mathcal{E}\mathcal{V}_{\mathbb{R}}\mathcal{D}\mathcal{F}}(E \times X, E \times Y)$  sont les applications linéaires de la forme

$$E \times X \ni (c, x) \rightarrow (c, A \cdot c + B \cdot x) \in E \times Y, \quad (44)$$

où  $A \in L(E, Y)$  et  $B \in L(X, Y)$ . Nous noterons  $[[A, B]]$  l'application (44).

Si  $\Phi = (I, Q, \Phi)$  est un flot dans  $\mathbf{C}^1MV$ ,  $\xi \in \text{Traj}(\Phi)$ , et  $E$  est un espace vectoriel, nous noterons  $E \times X^{\Phi, \xi}$  la famille  $\{E \times X_t^{\Phi, \xi}\}_{t \in I}$ .

**Definition 4.6.2** Soit  $\Phi = (I, Q, \Phi)$  un flot dans la catégorie  $\mathbf{C}^1MV$ , et soit  $\xi \in \text{Traj}(\Phi)$ . Notons  $X = X^{\Phi, \xi}$ . Soit  $\mathbf{\Lambda} = (I, X, \Lambda)$  un multiflot compact linéaire sur  $(I, X)$ . On appelle *variation infinitésimale de  $\Phi$  associée à  $\mathbf{\Lambda}$  le long de  $\xi$*  un couple  $\mathbf{W} = (C, \Theta)$  tel que

1.  $C$  est un cône convexe fermé dans un espace un espace vectoriel réel  $E_C$  de dimension finie,
2.  $\text{Int}_{E_C}(C) \neq \emptyset$ ,
3.  $\Theta$  est une application de l'espace  $\Gamma(\mathbf{\Lambda})$  des sélections de  $\mathbf{\Lambda}$  dans l'ensemble des  $E \times \mathcal{E}\mathcal{V}_{\mathbb{R}}\mathcal{D}\mathcal{F}$ -multiflots compacts sur  $(I, E \times X^{\Phi, \xi})$ , telle que pour toute sélection  $L = \{L_{t,s} : (s, t) \in \hat{I}\} \in \Gamma(\mathbf{\Lambda})$  le multiflot  $\Theta(L)$  possède la propriété suivante : si  $(s, t) \in \hat{I}$  et  $W \in (\Theta(L))_{t,s}$ ,  $W$  est de la forme  $[[Z, L_{t,s}]]$  pour un  $Z \in L(E_C, X_t^{\Phi, \xi})$ ,
4. pour tout  $(s, t) \in \hat{I}$ , l'ensemble  $\{(L, Z) : [[Z, L_{t,s}]] \in (\Theta(L))_{t,s}\}$  est compact.  $\diamond$

Si  $\mathbf{W} = (C, \Theta)$  est une variation infinitésimale de  $\Phi$  associée à  $\mathbf{\Lambda}$  le long de  $\xi$ ,  $L \in \mathbf{\Lambda}$ , et  $(s, t) \in \hat{I}$ , nous définissons

$$Z_{t,s}^{\Theta}(L) \stackrel{\text{d\u00e9f}}{=} \left\{ Z \in L(E_C, X_t^{\Phi, \xi}) : [[Z, L_{t,s}]] \in (\Theta(L))_{t,s} \right\}.$$

Il ressort de cette formule que les ensembles  $Z_{t,s}^{\Theta}(L)$  vérifient l'identité

$$Z_{t,s}^{\Theta}(L) + L_{t,s} \circ Z_{s,r}^{\Theta}(L) = Z_{t,r}^{\Theta}(L)$$

pour  $r, s, t$  quelconques tels que  $r, s, t \in I$  et  $r \leq s \leq t$ .

On démontre alors par récurrence la proposition suivante :

**Proposition 4.6.3** *Étant donné un flot  $\Phi = (I, Q, \Phi)$  dans  $\mathbf{C}^1MV$ , une trajectoire  $\xi$  de  $\Phi$ , un multiflot compact linéaire  $\Lambda = (I, X^{\Phi, \xi}, \Lambda)$  sur  $(I, X^{\Phi, \xi})$ , une variation infinitésimale  $\mathbf{W} = (C, \Theta)$  de  $\Phi$  associée à  $\Lambda$  le long de  $\xi$ , un entier positif  $\nu$ , des temps  $t_0, \dots, t_\nu \in I$  tels que  $t_0 \leq t_1 \leq \dots \leq t_{\nu-1} \leq t_\nu$ , et une sélection  $L \in \Gamma(\Lambda)$ , l'identité suivante a lieu :*

$$Z_{t_\nu, t_0}^\Theta(L) = \sum_{k=1}^{\nu} L_{t_\nu, t_k} \circ Z_{t_k, t_{k-1}}^\Theta(L). \quad \diamond$$

Si  $\mathbf{W} = (C, \Theta)$  est une variation infinitésimale du  $\mathbf{C}^1MV$ -flot  $\Phi = (I, Q, \Phi)$  le long d'une trajectoire  $\xi$  de  $\Phi$ , associée au multiflot compact  $\Lambda = (I, X, \Lambda)$ , où  $X = X^{\Phi, \xi}$ , on notera  $\langle \mathbf{W}_{t,s} \rangle$ , pour  $(s, t) \in \hat{I}$ , l'ensemble de toutes les applications linéaires  $\pi_{2, E_C, X_t} \circ [[Z, L_{t,s}]] : E_C \times X_s \rightarrow X_t$ , pour  $L \in \Gamma(\Lambda)$ ,  $Z \in Z_{t,s}^\Theta(L)$ , où  $\pi_{2, U, V}$  désigne la projection  $U \times V \ni (u, v) \rightarrow v \in V$ . (Ainsi donc,  $\pi_{2, E_C, X_t} \circ [[Z, L_{t,s}]]$  est l'application  $E_C \times X_s \ni (c, x) \rightarrow Z \cdot c + L_{t,s} \cdot x \in X_t$ .) Il s'ensuit que  $\langle \mathbf{W}_{t,s} \rangle$  est un sous-ensemble compact de  $L(E_C \times X_s, X_t)$ .

**Definition 4.6.4** Soient  $\mathbf{W}^i = (C^i, \Theta^i)$ ,  $i = 1, 2$ , deux variations infinitésimales de  $\Phi$  associées à  $\Lambda$  le long de  $\xi$ , où  $\Phi = (I, Q, \Phi)$  est un flot dans  $\mathbf{C}^1MV$ ,  $\xi$  est une trajectoire de  $\Phi$ , et  $\Lambda = (I, X^{\Phi, \xi}, \Lambda)$  est un multiflot compact linéaire sur  $(I, X^{\Phi, \xi})$ . On appelle *variation composée* de  $\mathbf{W}^1$  et  $\mathbf{W}^2$  la variation infinitésimale

$$\mathbf{W}^1 \# \mathbf{W}^2 \stackrel{\text{déf}}{=} (C^1 \times C^2, \Theta^1 \# \Theta^2),$$

où, pour une sélection  $L \in \Gamma(\Lambda)$  quelconque,

$$((\Theta^1 \# \Theta^2)(L))_{t,s} \stackrel{\text{déf}}{=} \left\{ [[Z^1 \# Z^2, L_{t,s}]] : Z^1 \in Z_{t,s}^{\Theta^1}(L), Z^2 \in Z_{t,s}^{\Theta^2}(L) \right\},$$

et  $Z^1 \# Z^2$  est l'application linéaire

$$E_{C^1 \times C^2} \sim E_{C^1} \times E_{C^2} \ni (c^1, c^2) \rightarrow Z^1 \cdot c^1 + Z^2 \cdot c^2 \in X_t^{\Phi, \xi}. \quad \diamond$$

**Definition 4.6.5** Soit  $\mathcal{D}$  un calcul différentiel généralisé. Soit  $\Phi = (I, Q, \Phi)$  un flot dans  $\mathbf{C}^1MV$ , et soit  $\xi$  une trajectoire de  $\Phi$  telle que  $\Phi$  est  $\mathcal{D}$ -différentiable le long de  $\xi$ . Soit  $\Lambda = (I, X^{\Phi, \xi}, \Lambda)$  une  $\mathcal{D}$ -différentielle de  $\Phi$  le long de  $\xi$ . Soit  $\mathbf{V} = [C, \Psi]$  une variation de  $\Phi$ , et soit  $\mathbf{W} = (C', \Theta)$  une variation infinitésimale de  $\Phi$  le long de  $\xi$  associée à  $\Lambda$ . Nous dirons que  $\mathbf{W}$  est une  *$\mathcal{D}$ -différentielle variationnelle de  $\mathbf{V}$  le long de  $\xi$  associée à  $\Lambda$*  si  $C' = C$ , et la relation

$$\langle \mathbf{W}_{t,s} \rangle \in \mathcal{D}(\langle \Psi_{t,s} \rangle, (0, \xi(s)), \xi(t), C \times Q_s)$$

est vérifiée pour tout  $(s, t) \in \hat{I}$ . \(\diamond\)

## 4.7 Le principe du maximum pour les flots et les différentielles généralisées

**Definition 4.7.1** Un système de flots est une triple  $\Sigma = (I, Q, \{\Phi^\eta\}_{\eta \in \mathcal{U}})$  où  $I$  est un ensemble totalement ordonné,  $Q = \{Q_t\}_{t \in I}$  est une famille paramétrisée de variétés de classe  $C^1$ ,  $\mathcal{U}$  est un ensemble, et, pour chaque  $\eta \in \mathcal{U}$ ,  $\Phi^\eta = (I, Q, \Phi^\eta)$  est un flot dans  $C^1MV$ .  $\diamond$

**Definition 4.7.2** Soit  $\Sigma = (I, Q, \{\Phi^\eta\}_{\eta \in \mathcal{U}})$  un système de flots. Soit  $\eta_* \in \mathcal{U}$ . On dira qu'une variation  $\mathbf{V} = [C, \Psi] = [C, \{\Psi^c\}_{c \in C}]$  du flot  $\Phi^{\eta_*}$  est une variation dans  $\Sigma$  s'il existe un voisinage  $N$  de 0 dans  $C$  tel que les flots  $\Psi^c$  appartiennent à l'ensemble  $\{\Phi^\eta : \eta \in \mathcal{U}\}$  pour tout  $c \in N$ .  $\diamond$

**Definition 4.7.3** Soit  $\Sigma = (I, Q, \{\Phi^\eta\}_{\eta \in \mathcal{U}})$  un système de flots. Soit  $\mathcal{D}$  un calcul différentiel généralisé. Soient  $\eta_* \in \mathcal{U}$ ,  $\xi_* \in \text{Traj}(\Phi^{\eta_*})$ . Supposons que  $\Phi^{\eta_*}$  soit  $\mathcal{D}$ -différentiable le long de  $\xi_*$ . Soit  $\mathbf{\Lambda} = (I, X, \Lambda)$  une  $\mathcal{D}$ -différentielle de  $\Phi^{\eta_*}$  le long de  $\xi_*$ . Soit  $\mathcal{W}$  un ensemble de variations infinitésimales de  $\Phi^{\eta_*}$  le long de  $\xi_*$  associées à  $\mathbf{\Lambda}$ . On dira que  $\mathcal{W}$  est admissible pour  $\Sigma$  si pour tout sous-ensemble fini  $\mathcal{W}^0$  de  $\mathcal{W}$  il existe une variation  $\mathbf{V}$  de  $\Phi^{\eta_*}$  dans  $\Sigma$  et une suite ordonnée  $(\mathbf{W}^1, \dots, \mathbf{W}^\nu)$  d'éléments de  $\mathcal{W}^0$ , contenant tous les éléments de  $\mathcal{W}^0$ , telle que la variation composée  $\mathbf{W}^1 \# \mathbf{W}^2 \# \dots \# \mathbf{W}^\nu$  est une différentielle variationnelle de  $\mathbf{V}$  le long de  $\xi_*$  associée à  $\mathbf{\Lambda}$ .  $\diamond$

Si  $\Sigma = (I, Q, \{\Phi^\eta\}_{\eta \in \mathcal{U}})$  est un système de flots,  $a \in I$ ,  $b \in I$ ,  $a \leq b$ , et  $q \in Q_a$ , nous noterons  $\mathcal{R}_{a,b}^\Sigma(q)$  l'ensemble des points  $\Sigma$ -accessibles à partir de  $q$  sur l'intervalle  $[a, b]$ , c'est-à-dire l'ensemble  $\bigcup_{\eta \in \mathcal{U}} \Phi_{b,a}^\eta(q)$ .

**Theorème 4.7.4** Soit  $\mathcal{D}$  un CDG possédant la propriété de l'application directionnellement ouverte. Soit  $\Sigma = (I, Q, \{\Phi^\eta\}_{\eta \in \mathcal{U}})$  un système de flots. Soient  $\eta_* \in \mathcal{U}$ ,  $\xi_* \in \text{Traj}(\Phi^{\eta_*})$ . Supposons que  $\Phi^{\eta_*}$  soit  $\mathcal{D}$ -différentiable le long de  $\xi_*$ . Soit  $\mathbf{\Lambda} = (I, X, \Lambda)$  une  $\mathcal{D}$ -différentielle de  $\Phi^{\eta_*}$  le long de  $\xi_*$ . Soit  $\mathcal{W}$  un ensemble admissible pour  $\Sigma$  de variations infinitésimales de  $\Phi^{\eta_*}$  le long de  $\xi_*$  associées à  $\mathbf{\Lambda}$ . Soient  $a, b \in I$  tels que  $a \leq_I b$ . Écrivons  $q_* = \xi_*(a)$ ,  $\tilde{q}_* = \xi_*(b)$ . Soit  $S$  un sous-ensemble de  $Q_b$  tel que  $\mathcal{R}_{a,b}^\Sigma(q_*) \cap S = \{\tilde{q}_*\}$ . Soit  $\mathcal{K}$  un multicône  $\mathcal{D}$ -tangent à  $S$  en  $\tilde{q}_*$ . Supposons qu'il existe un covecteur  $\mu \in T_{\tilde{q}_*}^* Q_b \setminus \{0\}$  tel que  $K \subseteq \{k : \mu(k) \geq 0\}$  et  $K \not\subseteq \{k : \mu(k) = 0\}$  pour tout  $K \in \mathcal{K}$ . Il existe alors une sélection  $L = \{L_{t,s}\}_{a \leq s \leq t \leq b}$  de la restriction  $\mathbf{\Lambda} \upharpoonright [a, b]$ , un covecteur  $\bar{\pi} \in T_{\xi_*(t)}^* Q_b \setminus \{0\}$ , et un cône  $K \in \mathcal{K}$ , tels que (a)  $\langle \bar{\pi}, k \rangle \geq 0$  pour tout  $k \in K$ , et (b) pour tout  $\mathbf{W} = (C, \Theta) \in \mathcal{W}$  il existe  $Z \in Z_{b,a}^\Theta$  tel que  $\langle \bar{\pi}, Z \cdot c \rangle \leq 0$  pour tout  $c \in C$ .

PREUVE. Choisissons une norme  $\|\cdot\|$  sur  $T_{\tilde{q}_*}^* Q_b$ , et notons  $\mathbb{S}$  la sphère unité de  $T_{\tilde{q}_*}^* Q_b$ . Notons  $A$  l'espace  $\Gamma(\mathbf{\Lambda} \upharpoonright [a, b])$  de toutes les sélections de  $\mathbf{\Lambda} \upharpoonright [a, b]$ . Nous savons déjà (cf. prop. 4.6.3) que  $A$  est compact et  $A \neq \emptyset$ .

Si  $\tilde{\mathcal{W}}$  est un sous-ensemble de  $\mathcal{W}$ , notons  $B(\tilde{\mathcal{W}}^0)$  l'ensemble de tous les couples  $(\bar{\pi}, L)$  tels que (1)  $\bar{\pi} \in T_{\tilde{q}_*}^* Q_b$ , (2)  $\|\bar{\pi}\| = 1$ , (3) il existe  $K \in \mathcal{K}$  tel que

$\langle \bar{\pi}, k \rangle \geq 0$  pour tout  $k \in K$ , (4)  $L \in A$ , et (5) pour tout  $\mathbf{W} = (C, \Theta) \in \tilde{\mathcal{W}}$  il existe  $Z \in Z_{b,a}^\Theta$  tel que  $\langle \bar{\pi}, Z \cdot c \rangle \leq 0$  pour tout  $c \in C$ .

On vérifie aisément que  $B(\mathcal{W}^0) = \bigcap_{\tilde{\mathcal{W}} \subseteq \mathcal{W}, \tilde{\mathcal{W}} \text{ fini}} B(\tilde{\mathcal{W}}^0)$  et que  $B(\tilde{\mathcal{W}}^0)$  est un sous-ensemble compact de  $\mathbb{S} \times A$ . Il s'ensuit que, pour montrer que  $B(\mathcal{W}^0) \neq \emptyset$ , il suffit de démontrer que  $\bigcap_{\tilde{\mathcal{W}} \subseteq \mathcal{W}, \tilde{\mathcal{W}} \in F} B(\tilde{\mathcal{W}}) \neq \emptyset$  si  $F$  est un ensemble fini de sous-ensembles finis de  $\mathcal{W}$ . Or  $\bigcap_{\tilde{\mathcal{W}} \subseteq \mathcal{W}, \tilde{\mathcal{W}} \in F} B(\tilde{\mathcal{W}}) = B(\bigcup_{\tilde{\mathcal{W}} \in F} \tilde{\mathcal{W}})$ . Il suffit donc de montrer que  $B(\tilde{\mathcal{W}}) \neq \emptyset$  pour tout sous-ensemble fini  $\tilde{\mathcal{W}}$  de  $\mathcal{W}$ .

Soit  $\tilde{\mathcal{W}}$  un sous-ensemble fini de  $\mathcal{W}$ . Soit  $(\mathbf{W}^1, \dots, \mathbf{W}^\nu)$  une suite finie d'éléments de  $\tilde{\mathcal{W}}$  contenant tous les éléments de  $\tilde{\mathcal{W}}$ , telle qu'il existe une variation  $\mathbf{V}$  de  $\Phi^{\eta^*}$  dans  $\Sigma$  possédant la propriété que la variation composée  $\mathbf{W} = \mathbf{W}^1 \# \mathbf{W}^2 \# \dots \# \mathbf{W}^\nu$  est une différentielle variationnelle de  $\mathbf{V}$  le long de  $\xi_*$  associée à  $\mathbf{\Lambda}$ .

Écrivons  $\mathbf{W}^i = (C^i, \Theta^i)$ ,  $\mathbf{W} = (C, \Theta)$ . Il s'ensuit que  $C = C^1 \times \dots \times C^\nu$ , et que  $Z_{t,s}^{\Theta^i}(L) = \{Z^1 \# Z^2 \# \dots \# Z^\nu : Z^i \in Z_{t,s}^{\Theta^i}(L) \text{ pour } i = 1, \dots, \nu\}$  pour tout  $(s, t) \in \hat{I}$ ,  $L \in \Phi_{t,s}$ .

La variation  $\mathbf{V}$  est donc de la forme  $[C, \Psi]$ , où  $\Psi = \{\Psi^c\}_{c \in C}$ . Ainsi donc, on peut supposer que chaque  $\Psi^c$  appartienne à l'ensemble  $\{\Phi^\eta : \eta \in \mathcal{U}\}$ .

Soit  $M$  l'application multivoque  $\langle \Psi_{b,a} \rangle$ . Il s'ensuit que  $M$  est une application multivoque de  $E_C \times Q_a$  dans  $Q_b$ , et que  $\langle \mathbf{W}_{b,a} \rangle \in \mathcal{D}(M, (0, q_*), \tilde{q}_*, C \times Q_b)$ . Or,  $\langle \mathbf{W}_{b,a} \rangle$  est l'ensemble de toutes les applications linéaires  $\zeta_L^{Z^1, \dots, Z^\nu}$ , pour  $L \in A$ ,  $Z^i \in Z_{b,a}^{\Theta^i}(L)$ , où  $\zeta_L^{Z^1, \dots, Z^\nu}(c^1, \dots, c^\nu, v) \stackrel{\text{déf}}{=} L_{b,a} \cdot v + \sum_{i=1}^\nu Z^i \cdot c^i$ .

Soit  $\tilde{M}$  l'application multivoque de  $E_C$  dans  $Q_b$  donnée par  $\tilde{M}(c) = M(c, q_*)$ . Ainsi donc,  $\tilde{M} = M \circ \iota$ , où  $\iota$  est l'inclusion  $E_C \ni c \rightarrow (c, q_*) \in E_C \times Q_b$ . En vertu de la règle de différentiation des fonctions composées, on obtient

$$\Delta \in \mathcal{D}(\tilde{M}, 0, \tilde{q}_*, C),$$

où  $\Delta$  est l'ensemble de toutes les applications linéaires  $\tilde{\zeta}_L^{Z^1, \dots, Z^\nu} : E_C \rightarrow T_{\tilde{q}_*} Q_b$ , pour  $L \in A$ ,  $Z^i \in Z_{b,a}^{\Theta^i}(L)$ , et  $\tilde{\zeta}_L^{Z^1, \dots, Z^\nu}(c^1, \dots, c^\nu) \stackrel{\text{déf}}{=} \sum_{i=1}^\nu Z^i \cdot c^i$ .

L'ensemble  $\tilde{M}(C)$  étant inclus dans  $\mathcal{R}_{a,b}^\Sigma(q_*)$ , on déduit de la propriété de transversalité que les multicônes  $\Delta(C)$  et  $\mathcal{K}$  ne sont pas fortement transversaux. Il s'ensuit que ces multicônes satisfont à une des deux conditions suivantes :

- (a)  $\Delta(C)$  et  $\mathcal{K}$  ne sont pas transversaux,
- (b)  $\Delta(C)$  et  $\mathcal{K}$  sont transversaux mais pas fortement transversaux.

Considérons d'abord le cas (b). Soient  $D \in \Delta(C)$ ,  $K \in \mathcal{K}$ . Soit  $\bar{k} \in K$  tel que  $\mu(\bar{k}) > 0$ . En vertu de la transversalité de  $D$  et  $K$ , on peut écrire  $\bar{k} = d - k$ ,  $d \in D$ ,  $k \in K$ . Il s'ensuit que  $d = k + \bar{k}$ , d'où l'on déduit que  $d \in D \cap K$ . Or,  $\mu(d) = \mu(k) + \mu(\bar{k}) > 0$ . Ainsi donc, l'intersection  $D \cap K \cap \{v : \mu(v) > 0\}$  n'est pas vide. Nous avons donc montré que  $D \cap K \cap \{v : \mu(v) > 0\} \neq \emptyset$  pour tout  $D \in \Delta(C)$ ,  $K \in \mathcal{K}$ . La transversalité de  $\Delta(C)$  et  $\mathcal{K}$  implique donc que  $\Delta(C)$  et  $\mathcal{K}$  sont fortement transversaux, contrairement à (b).

In s'ensuit que le cas (b) est en fait impossible. Nous sommes donc ramenés au cas (a), ce qui veut dire qu'il existe  $D \in \Delta(C)$ ,  $K \in \mathcal{K}$ , et un covecteur  $\bar{\pi} \in T_{\bar{q}_b}^* Q_b \setminus \{0\}$  tels que  $\langle \bar{\pi}, d \rangle \leq 0$  pour tout  $d \in D$  et  $\langle \bar{\pi}, k \rangle \geq 0$  pour tout  $k \in K$ . On peut donc supposer que  $\|\bar{\pi}\| = 1$ .

Puisque  $D = \Delta(C)$ , il existe  $L \in A$ ,  $(Z^1, \dots, Z^\nu) \in Z_{b,a}^{\Theta^1}(L) \times \dots \times Z_{b,a}^{\Theta^\nu}(L)$ , tels que  $D = \tilde{\zeta}_L^{Z^1, \dots, Z^\nu}$ . Il en résulte que  $\langle \bar{\pi}, \sum_{i=1}^\nu Z^i \cdot c^i \rangle \leq 0$  pour tout  $(c^1, \dots, c^\nu) \in C^1 \times \dots \times C^\nu$ . On en déduit que  $\langle \bar{\pi}, Z^i \cdot c \rangle \leq 0$  pour tout  $c \in C^i$ , pour  $i = 1, \dots, \nu$ . Ceci implique que  $(\bar{\pi}, L) \in B(\mathcal{W})$ . On a donc  $B(\mathcal{W}) \neq \emptyset$ , ce qui achève notre démonstration.  $\diamond$

Reprenons les notations du Théorème 4.7.4. Une *variation infinitésimale ponctuelle* est une variation infinitésimale construite de la façon suivante : on choisit un cône convexe fermé  $C$  à intérieur non vide dans un espace vectoriel réel  $E_C$  de dimension finie, un temps  $\tau \in I$ , et un ensemble compact nonvide  $\mathbf{M}$  d'applications linéaires de  $E_C$  dans  $X_t$ . Nous définissons alors  $\Theta_{t,s}(L)$ , pour  $L \in \Gamma(\mathbf{A})$ , par la formule

$$\Theta_{t,s}(L) = \begin{cases} [[L_{t,\tau} \circ \mathbf{M}, L_{t,s}]] & \text{si } s \leq \tau < t, \\ [[0, L_{t,s}]] & \text{si } s > \tau \text{ ou } t \leq \tau. \end{cases}$$

La variation  $\mathbf{W} = (C, \Theta)$  obtenue de cette façon sera notée  $((C, \tau, \mathbf{M}))$ . Le temps  $\tau$  est appelé le *point d'application* de  $\mathbf{W}$ .

Lorsque les variations  $\mathbf{W} \in \mathcal{W}$  sont ponctuelles, le théorème 4.7.5 prend la forme du « principe du maximum » :

**Théorème 4.7.5** *Sous les hypothèses du théorème 4.7.5, supposons en outre que toutes les variations  $\mathbf{W} \in \mathcal{W}$  soient ponctuelles, et que tous leur temps d'application appartiennent à l'intervalle  $[a, b]$ . Il existe alors une sélection  $L = \{L_{t,s}\}_{a \leq s \leq t \leq b}$  de la restriction  $\mathbf{A}[a, b]$ , un covecteur  $\bar{\pi} \in T_{\xi_*(t)}^* Q_b$ , et un cône  $K \in \mathcal{K}$ , tels que (a)  $\bar{\pi} \neq 0$ , (b)  $\langle \bar{\pi}, k \rangle \geq 0$  pour tout  $k \in K$ , et (c) le champ de covecteurs  $[a, b] \ni t \rightarrow \pi(t) \stackrel{\text{d'ef}}{=} \bar{\pi} \circ L_{b,t} \in T_{\xi_*(t)}^* Q_t$  (dit « champ adjoint ») vérifie la condition suivante : quelle que soit la variation  $\mathbf{W} = ((C, \tau, \mathbf{M})) \in \mathcal{W}$ , il existe  $M \in \mathbf{M}$  tel que  $\langle \pi(t), M \cdot c \rangle \leq 0$  pour tout  $c \in C$ .  $\diamond$*

## 4.8 Quotients différentiels généralisés

L'idée du « quotient différentiel généralisé » (abrév.: QDG) est un développement logique d'une caractérisation de la différentielle classique proposée en 1984 par Botsko et Gosser, cf. [2]. Botsko et Gosser remarquèrent que, si  $F$  est une application continue<sup>17</sup> de  $\mathbb{R}^n$  dans  $\mathbb{R}^m$  telle que  $F(0) = 0$ , et  $L \in \mathbb{R}^{m \times n}$ , alors  $L = DF(0)$  si et seulement si  $F$  admet une factorisation  $F(x) = G(x) \cdot x$  au voisinage de 0, où  $G$  est une fonction continue à valeurs dans  $\mathbb{R}^{m \times n}$  telle que

<sup>17</sup>Sans supposer que  $F$  soit continue, l'énoncé suivant est vrai :  $L = DF(0)$  si et seulement si  $F$  admet une factorisation  $F(x) = G(x) \cdot x$  au voisinage de 0, où  $G$  est une fonction continue en 0 à valeurs dans  $\mathbb{R}^{m \times n}$  telle que  $G(0) = L$ .

$G(0) = L$ . Dans notre généralisation, on permettra à  $F$  et  $G$  d'être multivoques, et l'égalité  $G(x) \cdot x = F(x)$  sera remplacée par l'inclusion  $G(x) \cdot x \subseteq F(x)$ . Pour compléter la définition, il faut trouver les conditions techniques précises qui remplaceront la continuité de  $G$  dans le cas multivoque, On est donc mené à se demander d'abord pourquoi la continuité est-elle importante dans le cas univoque. Une réflexion montre que *la continuité joue un rôle décisif parce que pour les applications continues on connaît des bons résultats d'existence de point fixes, notamment les théorèmes de Brouwer et de Schauder*. Ainsi donc, dans le cas multivoque on aura besoin d'une condition qui permette de démontrer des théorèmes de point fixe. On aboutit ainsi naturellement aux conditions du théorème de Kakutani, selon lesquelles  $G$  doit être à valeurs convexes non vides, et le graphe de la restriction de  $G$  à chaque ensemble compact doit être compact. Or, cette première tentative se révèle insuffisante, parce que la condition de convexité des valeurs de  $G$  n'est pas invariante par les transformations continues non linéaires. Une réflexion un peu plus approfondie montre que la bonne condition est la « régularité » de  $G$ .

**Definition 4.8.1** Soient  $X, Y$  des espaces métriques, et soit  $F$  une application multivoque de  $X$  dans  $Y$ . On dira que  $F$  est *régulière* si pour tout sous-ensemble compact  $K$  de  $X$  les deux conditions suivantes sont vérifiées : (a) l'ensemble  $\text{Gr}(F) \cap (K \times Y)$  est compact, et (b) il existe une suite d'applications continues  $h_j : K \rightarrow Y$  telle que « les  $h_j$  convergent vers  $F \upharpoonright K$  » dans le sens que, si  $U$  est un voisinage quelconque de  $\text{Gr}(F) \cap (K \times Y)$  dans  $K \times Y$ , alors il existe  $j_U$  tel que  $\text{Gr}(h_j) \subseteq U$  pour tout  $j \geq j_U$ .  $\diamond$

**Definition 4.8.2** Soient  $n, m \in \mathbb{Z}_+$ ,  $(F; S) \in \mathcal{A}_{n,m}$ . Soit  $\Lambda$  un sous-ensemble compact de  $\mathbb{R}^{m \times n}$ . On dira que  $\Lambda$  est un *quotient différentiel généralisé de  $F$  dans la direction de  $S$* , et on écrira  $\Lambda \in \mathcal{D}_{QDG}(F; S)$ , si pour tout voisinage  $W$  de  $\Lambda$  dans  $\mathbb{R}^{m \times n}$  il existe un voisinage  $U$  de 0 dans  $\mathbb{R}^m$  tel que  $U \cap S$  est compact, et une application multivoque régulière  $G$  de  $U \cap S$  dans  $W$  telle que  $G(x) \cdot x \subseteq F(x)$  quel que soit  $x \in U \cap S$ .  $\diamond$

Nous énonçons sans démonstration le résultat suivant, qu'on prouve par un argument de point fixe.

**Proposition 4.8.3** *La théorie  $\mathcal{D}_{QDG}$  est un calcul différentiel généralisé possédant la propriété de l'application directionnellement ouverte.*  $\diamond$

**Remarque 4.8.4** La version du théorème 4.7.5 associée aux QDGs contient toutes les autres versions du principe du maximum énoncées dans ce travail. À présent, il reste encore des résultats que nous ne savons pas démontrer à partir de notre théorème général. Plus précisément, nous ne savons pas démontrer les résultats pour les inclusions différentielles que les mathématiciens de l'école d'« analyse non lisse » appellent « versions intrinsèques » (cf. par exemple Ioffe [12]). En outre, nous ne savons pas démontrer les résultats où dans la condition de transversalité intervient un « cône normal » qui n'est pas le polaire d'un de nos cônes ou multicônes tangents.  $\diamond$

**Remarque 4.8.5** La théorie des QDGs n'est pas comparable avec celles des conteneurs de dérivées ou des semi-différentielles. Par exemple, la fonction  $x \rightarrow x \sin \frac{1}{x}$  n'admet ni une semi-différentielle ni une multi-différentielle en 0, mais l'intervalle  $[-1, 1]$  est évidemment un QDG. Inversement, on peut construire des fonctions lipschitziennes admettant des conteneurs de dérivées qui ne sont pas des QDGs. Récemment (en Janvier 2000) nous avons trouvé un CDG encore plus général (celui des « path-integral generalized differentials »), qui contient tous les autres et possède la propriété de l'application directionnellement ouverte. Il est donc très probable que **la** « vraie théorie unifiée » des dérivées généralisées existe quelque part dans l'univers des idées mathématiques. Par contre, il ne nous semble pas évident à présent (en mai 2000) qu'une telle conclusion optimiste soit également justifiée pour les versions du principe du maximum. Ceci explique pourquoi notre travail se termine ici, sans une vraie « conclusion ».  $\diamond$

## Bibliographie

- [1] Bolza, O., *Über der 'Anormalen Fall' beim Lagrangeschen und Mayerschen Problem mit gemischten Bedingungen und variablen Endpunkten*. Math. Ann. **LXXIV**, 1913, pp. 430–446.
- [2] Botsko, M. W., and R. A. Gosser, *On the differentiability of functions of several variables*. Amer. Math. Monthly **92**, No. 9, 1985, pp. 663–665.
- [3] Bressan, A., *Directionally continuous selections and differential inclusions*. Funk. Ekvac. **31**, 1988, pp. 459–470.
- [4] Clarke, F. H., *The maximum principle under minimal hypotheses*. S.I.A.M. J. Control Optim. **14**, 1976, pp. 1078–1091.
- [5] Clarke, F.H., *Optimization and Nonsmooth Analysis*. Wiley, New York, 1983.
- [6] Fryszkowski, A., and T. Rzeżuchowski, *Continuous version of the Filippov-Ważewski relaxation theorem*. J. Diff. Equations **94**, 1993, pp. 254–265.
- [7] Gandz, S., *Die Harpedonaptai oder Seilspanner*. Quellen und Studien Gesch. der Math. **B 1**, 1930, pp. 255–277.
- [8] Halkin, H., *Necessary conditions in mathematical programming and optimal control theory*. In *Optimal control theory and its applications, Part I*, Lect. Notes Econ. Math. Systems **105**, Springer, Berlin, 1974, pp. 113–165.
- [9] Halkin, H., *Interior mapping theorem with set-valued derivatives*. Journal d'Analyse Mathématique **30**, 1976, pp. 200–207.

- [10] Halkin, H., *Necessary conditions for optimal control problems with differentiable or nondifferentiable data*. In *Mathematical Control Theory*, Lect. Notes Math. **680**, Springer-Verlag, Berlin, 1978, pp. 77–118.
- [11] Heath, T. L., *A Manual of Greek Mathematics*. Clarendon Press, 1931.
- [12] Ioffe, A., *Euler-Lagrange and Hamiltonian formalisms for dynamic optimization*. *Trans. Amer. Math. Society* **349**, 1997, pp. 2871–2900.
- [13] Kaskosz, B., and S. Lojasiewicz Jr., *A Maximum Principle for generalized control systems*. *Nonlinear Anal. TMA*, **9**, 1985, pp. 109–130.
- [14] Pontryagin, L.S., V.G. Boltyanskii, R.V. Gamkrelidze, & E.F. Mischenko, *The Mathematical Theory of Optimal Processes*. Wiley, New York, 1962.
- [15] Sussmann, H. J., *A strong version of the Lojasiewicz maximum principle*. In *Optimal Control of Differential Equations*, N. H. Pavel Ed., Marcel Dekker, New York, 1994, pp. 293–309.
- [16] Sussmann, H. J., *A strong version of the Maximum Principle under weak hypotheses*. In *Proc. 33rd IEEE Conference on Decision and Control, Orlando, FL, 1994*, IEEE publications, New York, 1994, pp. 1950–1956.
- [17] Sussmann, H. J., *An introduction to the coordinate-free maximum principle*. In *Geometry of Feedback and Optimal Control*, B. Jakubczyk and W. Respondek Eds., Marcel Dekker, New York, 1997, pp. 463–557.
- [18] Sussmann, H. J., *A strong maximum principle for systems of differential inclusions*. In *Proc. 35th IEEE Conference on Decision and Control, Kobe, Japan, Dec. 1996*, IEEE publications, 1996, pp. 1809–1814.
- [19] Sussmann, H.J., *Multidifferential calculus: chain rule, open mapping & transversal intersection theorems*. In *Optimal Control: Theory, Algorithms, and Applications*, W.W. Hager & P.M. Pardalos Eds., Kluwer, 1998, pp. 436–487.
- [20] Sussmann, H. J., *Geometry and optimal control*. In *Mathematical Control Theory*, J. Baillieul and J. C. Willems, Eds., Springer-Verlag, New York, 1998, pp. 140–198.
- [21] Sussmann, H. J., *Transversality conditions and a strong maximum principle for systems of differential inclusions*. In *Proc. 37th IEEE Conf. Decision and Control, Tampa, FL, Dec. 1998*. IEEE publications, 1998, pp. 1–6.
- [22] Tougeron, J.-C., *Idéaux de fonctions différentiables*. *Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 71*. Springer-Verlag, 1972.
- [23] Warga, J., *Optimization and controllability without differentiability assumptions*. *SIAM J. Control and Optimization*, **21**, 1983, pp. 837–855.
- [24] Ważewski, T., *Sur l'évaluation du domaine d'existence des fonctions implicites réelles ou complexes*. *Ann. Soc. Polon. Math.* **20**, 1947, pp. 81–120.