MATHEMATIQUES

Edition
PARIS-ONZE EDITION
Université de Paris-Sud

Cours de Topologie

Calcul différentiel Équations différentielles

Pour la licence MAF

Jean-Christophe Yoccoz

*Centre Scientifique
d'Orsay*

COURS DE TOPOLOGIE

CALCUL DIFFÉRENTIEL ÉQUATIONS DIFFÉRENTIELLES

POUR LA LICENCE MAF

Jean-Christophe YOCCOZ

Edition
Paris Onze édition

Directeur du Service Paris Onze édition
Christian ESKENAZI

N° Paris Onze édition *G 36*

ISBN 2–87800–037–4

1ère PARTIE.
RAPPELS DE TOPOLOGIE.

CHAPITRE I - Les concepts fondamentaux.

§. 1. Espaces vectoriels normés. Espaces métriques.

1.1. Espaces vectoriels normés.

Soit E un espace vectoriel.

DÉFINITION 1. Une **norme** sur E est une application : $x \mapsto \|x\|$ de E dans $\mathbf{R}^+$ qui vérifie :

(EN1) $\|x\| = 0 \Leftrightarrow x = 0$.

(EN2) $\|\lambda x\| = |\lambda| \; \|x\|$, pour $\lambda \in \mathbf{R}$, $x \in E$.

(EN3) $\|x + y\| \leq \|x\| + \|y\|$, pour $x, y \in E$.

EXEMPLE 1. La norme **euclidienne** sur $\mathbf{R}^n$ est définie par :

$$\|x\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$$

La paire $(E, \| \;\; \|)$ s'appelle un **espace vectoriel normé.**

1.2. Distances. Espaces métriques.

Soit X un ensemble.

DÉFINITION 2. Une **distance** sur X est une application d de $X \times X$ dans $\mathbf{R}^+$ vérifiant

(D1) $d(x, y) = 0 \Leftrightarrow x = y$.

(D2) $d(x, y) = d(y, x)$, pour $x, y \in X$.

(D3) $d(x, y) + d(y, z) \geq d(x, z)$, pour $x, y, z \in X$.

(La propriété (D3) est dite "inégalité du triangle").

Un **espace métrique** est une paire (X,d) formée d'un ensemble X et d'une distance d sur X.

Exemple fondamental.

Soit $(E,\| \ \|)$ un espace vectoriel normé. Pour $x,y \in E$, définissons :

$$d(x,y) = \|y-x\| \ .$$

L'application $d : E \times E \to \mathbf{R}^+$ est une distance sur E : en effet, (D1) résulte de (EN1), (D2) résulte de (EN2) (avec $\lambda = -1$), et (D3) résulte de (EN3).

Les espaces vectoriels normés sont donc des espaces métriques.

1.3. Sous-espaces d'un espace métrique.

Soit (X,d) un espace métrique, Y une partie de X. La restriction de l'application d à $Y \times Y \subset X \times X$ vérifie (de façon triviale) les propriétés (D1), (D2), (D3) : c'est donc une distance sur Y et la paire (Y,d) est un espace métrique.

En particulier, n'importe quelle partie de $\mathbf{R}^n$, muni de la norme euclidienne, ou plus généralement n'importe quelle partie d'un espace vectoriel normé $(E,\| \ \|)$ **peut être considérée comme un espace métrique.**

1.4. Distances et normes équivalentes.

DÉFINITION 3. Deux normes $\| \ \|_1$, $\| \ \|_2$ sur un espace vectoriel E sont **équivalentes** s'il existe une constante $C > 0$ telle qu'on ait :

$$\|x\|_2 \leq C\,\|x\|_1 \ ,$$
$$\|x\|_1 \leq C\,\|x\|_2 \ ,$$

pour tout $x \in E$.

DÉFINITION 3'. Deux distances d_1, d_2 sur un ensemble X sont **équivalentes** s'il existe une constante $C > 0$ telle qu'on ait :

$$d_2(x,y) \leq C\,d_1(x,y) \ ,$$
$$d_1(x,y) \leq C\,d_2(x,y) \ ,$$

pour tous $x,y \in X$.

Il est immédiat de voir que deux normes équivalentes sur un espace vectoriel définissent des distances équivalentes.

Toutes les propriétés des espaces métriques (et espaces vectoriels normés) que nous étudions dans la suite ne dépendent de la distance (ou de la norme) qu'à équivalence près.

EXEMPLE 2. Considérons les normes $\| \ \|_1$, $\| \ \|_\infty$ sur $\mathbf{R}^n$ définies par :

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \ ,$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \ .$$

(On vérifie facilement que les conditions (EN1), (EN2), (EN3) sont satisfaites par $\| \ \|_1$ et $\| \ \|_\infty$). Pour $x \in \mathbf{R}^n$, on a :

$$\|x\|_\infty \leq \|x\| \ ,$$

$$\|x\|_\infty \leq \|x\|_1 \ ,$$

$$\|x\| \leq \sqrt{n}\, \|x\|_\infty \ ,$$

$$\|x\|_1 \leq n\, \|x\|_\infty \ ,$$

donc les normes $\| \ \|$, $\| \ \|_1$, $\| \ \|_\infty$ sur $\mathbf{R}^n$ sont **équivalentes.**

REMARQUE. On verra un peu plus loin que deux normes quelconques sur un espace vectoriel **de dimension finie** sont équivalentes.

1.5. Produits d'espaces métriques.

Soient (X_1, d_1), $(X_2, d_2), \ldots, (X_n, d_n)$ des espaces métriques. Notons $X = X_1 \times X_2 \times \ldots \times X_n$. Pour deux éléments $x = (x_1, \ldots, x_n)$, $y = (y, \ldots, y_n)$ de X, posons :

$$D(x,y) = \max_{1 \leq i \leq n} d_i(x_i, y_i) \ .$$

On vérifie immédiatement que l'application $D : X \times X \to \mathbf{R}^+$ possède les propriétés (D1), (D2), (D3). La paire (X, D) est donc un espace métrique, dit espace produit des espaces métriques $(X_1, d_1), \ldots, (X_n, d_n)$.

REMARQUE 1. Soient $(E_1, \| \ \|_1), \ldots, (E_n, \| \ \|_n)$ des espaces vectoriels normés, $E = E_1 \times \ldots \times E_n$. Pour $x \in E$, posons (en écrivant $x = (x_1, \ldots, x_n)$)

$$\|x\| = \max_{1 \leq i \leq n} \|x_i\|_i \ .$$

On définit ainsi une norme sur E, et $(E, \| \ \|)$ est l'espace vectoriel normé produit de $(E_1, \| \ \|_1), \ldots, (E_n, \| \ \|_n)$.

REMARQUE 2. Si dans la remarque précédente, on prend, pour chaque $i \in \{1, \ldots, n\}$:

$$E_i = \mathbf{R} \ , \quad \|x\|_i = |x| \ ,$$

on obtient sur le produit $E = \mathbf{R}^n$ la norme $\| \ \|_\infty$ considérée plus haut.

REMARQUE 3. Soient $(X_1, d_1), \ldots, (X_n, d_n)$ des espaces métriques. Pour $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ appartenant à $X = X_1 \times \ldots \times X_n$, posons :

$$D_1(x,y) = \sum_{i=1}^{n} d_i(x_i, y_i) \ ,$$

$$D_2(x,y) = \left(\sum_{i=1}^{n} [d_1(x_i, y_i)]^2 \right)^{1/2} \ ;$$

Alors, D_1, D_2 sont des distances sur X **équivalentes** à la distance D considérée précédemment.

§. 2. Suites dans un espace métrique.

Soient (X, d) un espace métrique, $\underline{x} = (x_n)_{n \geq 0}$ une suite d'éléments de X, a un point de X.

DÉFINITION 4. On dit que a est **limite** de la suite $(x_n)_{n \geq 0}$, ou que la suite $(x_n)_{n \geq 0}$ **converge** vers a, si l'on a :

$$\lim_{n \to +\infty} d(a, x_n) = 0 \ .$$

On écrit alors $\lim\limits_{n \to +\infty} x_n = a$.

DÉFINITION 5. On dit que a est **point d'accumulation** de la suite $(x_n)_{n \geq 0}$ s'il existe une suite extraite $(x_{n_k})_{k \geq 0}$ convergeant vers a.

PROPOSITION 1. *Si la suite $(x_n)_{n \geq 0}$ converge vers a, alors a est l'unique point d'accumulation de la suite $(x_n)_{n \geq 0}$. En particulier, la limite d'une suite, si elle existe, est unique.*

DÉMONSTRATION : Supposons que $(x_n)_{n \geq 0}$ converge vers a. Il est clair que a est point d'accumulation de $(x_n)_{n \geq 0}$.

Soit b un point d'accumulation de $(x_n)_{n \geq 0}$, et $(x_{n_k})_{k \geq 0}$ une suite extraite convergeant vers b. Par définition, on a :

$$\lim_{n \to +\infty} d(x_n, a) = 0 \ ,$$

$$\lim_{k \to +\infty} d(x_{n_k}, b) = 0 \ .$$

D'après l'inégalité du triangle (D3), on a :

$$d(a, b) \leq d(a, x_{n_k}) + d(x_{n_k}, b) \ ,$$

pour tout $k \geq 0$. Faisant tendre k vers $+\infty$, on obtient $d(a, b) = 0$, d'où $a = b$ d'après (D1). □

REMARQUE. Une suite possédant un unique point d'accumulation ne converge pas nécessairement : considérer par exemple dans $\mathbf{R}$ la suite définie par $x_{2n} = n$, $x_{2n+1} = 0$, (pour $n \geq 0$), dont le seul point d'accumulation est 0.

EXEMPLE 3. Soient (X, d) un espace métrique, Y une partie de X, $(x_n)_{n\geq 0}$ une suite **dans Y**.

La suite $(x_n)_{n\geq 0}$ converge dans (Y, d) si et seulement si elle converge dans (X, d) et sa limite dans (X, d) appartient à Y.

Les points d'accumulation dans (Y, d) de la suite $(x_n)_{n\geq 0}$ sont les points d'accumulation dans (X, d) qui appartiennent à Y.

EXEMPLE 4. Soient $(X_1, d_1), \ldots, (X_k, d_k)$ des espaces métriques, et pour $i \in \{1, \ldots, k\}$, soit $(x_n^i)_{n\geq 0}$ une suite dans X_i. Pour $n \geq 0$, posons $x_n = (x_n^1, \ldots, x_n^k)$; on définit ainsi une suite $(x_n)_{n\geq 0}$ dans $X = X_1 \times \ldots \times X_k$.

La suite $(x_n)_{n\geq 0}$ converge dans (X, D) si et seulement si chacune des suites $(x_n^i)_{n\geq 0}$ (pour $i = 1, 2, \ldots, k$) converge dans (X_i, d_i). On a alors :

$$\lim_{n\to+\infty} x_n = \left(\lim_{n\to+\infty} x_n^1, \ldots, \lim_{n\to+\infty} x_n^k \right) .$$

REMARQUE. Si $a = (a^1, \ldots, a^k)$ est point d'accumulation de $(x_n)_{n\geq 0}$, alors, pour tout $i \in \{1, \ldots, k\}$, le point a^i est point d'accumulation de la suite $(x_n^i)_{n\geq 0}$ dans (X_i, d_i).

La réciproque n'est pas vraie : prenons $X_1 = X_2 = \mathbf{R}$, muni de la distance usuelle, et définissons, pour $n \geq 0$:

$$x_{2n} = (x_{2n}^1, x_{2n}^2) = (n, 0) \ ,$$
$$x_{2n+1} = (x_{2n+1}^1, x_{2n+1}^2) = (0, n) \ ;$$

on voit que 0 est point d'accumulation de la suite $(x_n^1)_{x\geq 0}$ et de la suite $(x_n^2)_{n\geq 0}$, mais $(0,0)$ n'est pas point d'accumulation de la suite $(x_n)_{n\geq 0}$ dans $\mathbf{R}^2$. (Cette suite ne possède pas de point d'accumulation).

§. 3. Boules, ouverts, fermés.

3.1. Définitions.

Soit (X, d) un espace métrique.

DÉFINITION 6. Soient $a \in X$, $r > 0$. La **boule ouverte** de centre a, de rayon r est la partie de X définie par :

$$B(a,r) = \left\{ x \in X \, , d(x,a) < r \right\} .$$

La **boule fermée** de centre a de rayon r est la partie de X définie par :

$$\overline{B}(a,r) = \left\{ x \in X \, , d(x,a) \leq r \right\} .$$

DÉFINITION 7. Une partie Y de X est **ouverte** si pour tout $y \in Y$ il existe $r > 0$ tel que $B(y,r) \subset Y$.

Une partie Y de X est **fermée** si son complémentaire $X - Y$ dans X est ouverte.

EXEMPLE 5. Une boule ouverte est une partie ouverte : soient $a \in X$, $r > 0$, $Y = B(a,r)$, $y \in Y$; par définition, on a $d(a,y) < r$, donc :

$$r' = r - d(a,y) > 0 .$$

La boule $B(y,r')$ est contenue dans Y : pour $z \in B(y,r')$, on a en effet

$$\begin{aligned} d(z,a) &\leq d(z,y) + d(y,a) \\ &< r' + d(y,a) = r . \end{aligned}$$

EXEMPLE 6. Une boule fermée est une partie fermée : soient $a \in X$, $r > 0$, $Y = \overline{B}(a,r)$; montrons que $X - Y$ est ouverte; soit $y \in X - Y$; on a donc :

$$r' = d(a,y) - r > 0 .$$

La boule ouverte $B(y,r')$ est contenue dans $X - Y$: en effet, pour $z \in B(y,r')$, on a :

$$\begin{aligned} d(a,y) &\leq d(a,z) + d(z,y) \\ &< d(a,z) + r' , \end{aligned}$$

donc

$$d(a,z) > d(a,y) - r' = r .$$

3.2. Unions, intersections de parties ouvertes, fermées.

Soit (X,d) un espace métrique.

PROPOSITION 2.

1) *Soit* $(U_i)_{i\in I}$ *une famille de parties ouvertes de* X. *L'union des* U_i *est une partie ouverte de* X.

2) *Soient* $U_1,\dots,U_k$ *des parties ouvertes de* X. *L'intersection* $\bigcap_{i=1}^{k} U_i$ *est une partie ouverte de* X.

DÉMONSTRATION :

1) Si $x \in \bigcup_{i\in I} U_i$, il existe i_0 tel que U_{i_0} contienne x ; comme U_{i_0} est ouverte, il existe $r > 0$ tel que $B(x,r) \subset U_{i_0}$. On a alors $B(x,r) \subset \bigcup_{i\in I} U_i$.

2) Si $x \in \bigcap_{i=1}^{k} U_i$, il existe, pour $i = 1,2,\dots,k$, un nombre $r_i > 0$ tel que $B(x,r_i) \subset U_i$. Posons $r = \min_{1\leq i\leq k} r_i > 0$. On a alors $B(x,r) \subset B(x,r_i) \subset U_i$ pour $1 \leq i \leq k$, d'où $B(x,r) \subset \bigcap_{i=1}^{k} U_i$.

Par passage aux complémentaires, on obtient :

PROPOSITION 2'.

1) *Soit* $(F_i)_{i\in I}$ *une famille de parties fermées de* X. *L'intersection des* F_i *est une partie fermée de* X.

2) *Soient* $F_1,\dots,F_k$ *des parties fermées de* X. *L'union* $\bigcup_{i=1}^{k} F_i$ *est une partie fermée de* X.

REMARQUE 1. La partie vide ϕ de X, et la partie totale X de X sont toujours des parties à la fois ouvertes et fermées de X.

REMARQUE 2. En général, l'intersection d'une famille **infinie** de parties ouvertes **n'est pas ouverte** : par exemple, dans $X = \mathbf{R}$, muni de la distance usuelle, considérons, pour $n \geq 1$:

$$U_n = \left]-\frac{1}{n},+\frac{1}{n}\right[= B\left(0,\frac{1}{n}\right) \quad ;$$

on a $\bigcap_{n\geq 1} U_n = \{0\}$ qui n'est pas ouverte.

DÉFINITION 8. Soit Y une partie de X. L'union des parties ouvertes de X contenues dans Y est la plus grande partie ouverte (de X) contenue dans Y et s'appelle **intérieur** de Y (dans X).

L'intersection des parties fermées de X contenant Y est la plus petite partie fermée (de X) contenant Y et s'appelle **adhérence** de Y (dans X).

On note $\mathrm{int}(Y)$ ou $\overset{\circ}{Y}$ l'intérieur de Y ; on note $\overline{Y}$ l'adhérence de Y.

On a immédiatement, pour $Y \subset X$:

$$\mathrm{int}\,(X - Y) = X - \overline{Y}$$
$$\overline{X - Y} = X - \mathrm{int}\,Y\ .$$

3.3. Caractérisation de l'adhérence et des parties fermées par les suites.

Soient (X, d) un espace métrique, et Y une partie de X.

PROPOSITION 3. *Pour un point z de X, les propriétés suivantes sont équivalentes :*

- **(i)** *z appartient à l'adhérence $\overline{Y}$ de Y ;*
- **(ii)** *pour tout $r > 0$, la boule $B(z, r)$ coupe Y ;*
- **(iii)** *il existe une suite $(y_n)_{n \geq 0}$ d'éléments de Y convergeant vers z (dans X).*

DÉMONSTRATION :

(iii)$\Rightarrow$(ii). Pour tout $r > 0$, on a $y_n \in Y \cap B(z, r)$ pour n assez grand.

(ii)$\Rightarrow$(iii). Pour $n \geq 1$, prendre $y_n \in Y \cap B\left(z, \dfrac{1}{n}\right)$.

(i)$\Rightarrow$(ii). Si $B(z,r) \cap Y = \phi$, alors $B(z, r)$ est une partie ouverte contenue dans $X - Y$, donc dans $\mathrm{int}\,(X - Y) = X - \overline{Y}$, et $z \notin \overline{Y}$.

(ii)$\Rightarrow$(i). Si $z \notin \overline{Y} = X - \mathrm{int}\,(X - Y)$, il existe $r > 0$ tel que $B(z, r) \subset \mathrm{int}\,(X - Y) \subset X - Y$. □

COROLLAIRE. *La partie Y de X est fermée si et seulement si la limite de toute suite $(y_n)_{n \geq 0}$ d'éléments de Y qui converge dans X appartient à Y.*

En effet, une partie est fermée si et seulement si elle est égale à son adhérence.

3.4. Sous-espaces.

Soient (X, d) un espace métrique, Y une partie de X, Z une partie de Y. On peut donc considérer Z comme une partie de l'espace métrique (X, d), mais aussi comme une partie du sous-espace métrique (Y, d).

Des définitions, il résulte facilement que :

1. Pour que Z soit une partie ouverte **dans Y**, il faut et il suffit qu'il existe une partie U de X ouverte dans X telle que $Z = U \cap Y$;

2. Pour que Z soit une partie fermée **dans Y** il faut et il suffit qu'il existe une partie F de X fermée dans X telle que $Z = F \cap Y$;

3. Si Y est une partie ouverte de X, alors Z est ouverte dans Y si et seulement si elle est ouverte dans X.

4. Si Y est une partie fermée de X, alors Z est une partie fermée dans Y si et seulement si c'est une partie fermée dans X.

Exercice. Démontrer ces assertions.

§. 4. Applications continues.

Soient (X_1, d_1), (X_2, d_2) deux espaces métriques et $f : X_1 \to X_2$ une application.

4.1. Continuité en un point.

DÉFINITION 9. Soit $a \in X_1$. On dit que f est **continue au point a** si pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que :

$$d_1(x, a) < \delta \Rightarrow d_2(f(x), f(a)) < \varepsilon .$$

En d'autres termes, f est continue au point a si l'image réciproque d'une boule ouverte de centre $f(a)$ contient toujours une boule ouverte de centre a.

PROPOSITION 4. *Soit $a \in X_1$. Les propriétés suivantes sont équivalentes :*

(i) *f est continue au point a ;*

(ii) *pour toute suite $(x_n)_{n\geq 0}$ convergeant vers a la suite $(f(x_n))_{n\geq 0}$ converge vers $f(a)$.*

DÉMONSTRATION :

(i)⇒(ii). Soit $(x_n)_{n\geq 0}$ une suite convergeant vers a. Soit $\varepsilon > 0$. Comme f est continue en a, il existe $\delta > 0$ tel que :

$$d_1(x,a) < \delta \Rightarrow d_2(f(x), f(a)) < \varepsilon \ .$$

Or il existe un entier n_0 tel que $d_1(x_n, a) < \delta$, pour $n \geq n_0$. On a donc $d_2(f(x_n), f(a)) < \varepsilon$ pour $n \geq n_0$. Comme $\varepsilon > 0$ est arbitraire, la suite $(f(x_n))_{n\geq 0}$ converge vers $f(a)$.

(ii)⇒(i). Supposons que f n'est pas continue en a ; il existe alors $\varepsilon_0 > 0$, et, pour $n \geq 1$, un point $x_n \in X_1$ vérifiant :

$$d_1(x_n, a) < \frac{1}{n} \quad \text{et} \quad d_2(f(x_n), f(a)) \geq \varepsilon_0 \ .$$

La suite $(x_n)_{n\geq 1}$ converge vers a, mais la suite $(f(x_n)_{n\geq 1}$ ne peut converger vers $f(a)$. □

4.2. Continuité.

DÉFINITION 10. L'application f est **continue** si elle est continue en tout point $a \in X_1$.

PROPOSITION 5. *Le propriétés suivantes sont équivalentes :*

(i) *f est continue;*

(ii) *l'image réciproque par f d'une partie ouverte de X_2 est une partie ouverte de X_1 ;*

(iii) *l'image réciproque par f d'une partie fermée de X_2 est une partie fermée de X_1.*

DÉMONSTRATION :

(ii) ⇔(iii). Soit Y une partie de X_2 ; on a :

$$f^{-1}(X_2 - Y) = X_1 - f^{-1}(Y) \ ,$$

et

$$Y \text{ ouverte } \Leftrightarrow X_2 - Y \text{ fermée} \ ,$$
$$f^{-1}(Y) \text{ ouverte } \Leftrightarrow X_1 - f^{-1}(Y) \text{ fermée} \ ,$$

d'où l'équivalence de (ii) et (iii).

(i)⇒(ii). Soient Y une partie ouverte de X_2, a un point de $f^{-1}(Y)$. Alors Y contient une boule ouverte centrée en $f(a)$; comme f est continue en a, $f^{-1}(Y)$ contient une boule ouverte centrée en a.

(ii)⇒(i). Soit $a \in X_1$ et B une boule ouverte de centre $f(a)$. Alors B est ouverte dans X_2, donc $f^{-1}(B)$ est ouverte dans X_1, et contient donc une boule ouverte de centre a.

REMARQUE. L'image par f d'une partie ouverte dans X_1 n'est pas nécessairement ouverte dans X_2. L'image par f d'une partie fermée dans X_1 n'est pas nécessairement fermée dans X_2. Par exemple, en prenant $X_1 = X_2 = \mathbf{R}$, muni de la distance usuelle, et l'application continue f définie par :

$$f(x) = \frac{x^2}{1+x^2} ,$$

l'image $f(\mathbf{R}) = [0,1[$ n'est ni ouverte ni fermée dans $\mathbf{R}$.

DÉFINITION 11. Soit $K > 0$. L'application f est **K-lipschitzienne** si on a, pour tous $x, y \in X_1$:

$$d_2(f(x), f(y)) \leq K\, d_1(x,y) .$$

PROPOSITION 6. *Si f est K-lipschitzienne, alors f est continue.*

DÉMONSTRATION : Il suffit de prendre $\delta = \varepsilon K^{-1}$ dans la définition de la continuité au point $a \in X_1$.

EXEMPLE 7. Soit $(E, \| \ \|)$ un espace vectoriel normé, qu'on munit de la distance associée $d(x,y) = \|x - y\|$.

Munissons $\mathbf{R}$ de la distance usuelle. L'application $x \mapsto \|x\|$ de E dans $\mathbf{R}$ est 1-lipschitzienne; en effet, d'après (EN3), on a, pour $x, y \in E$:

$$|\|x\| - \|y\|| \leq \|x - y\| .$$

EXEMPLE 7'. Soit (X,d) un espace métrique. On a muni en 1.5 le produit $X \times X$ de la distance :

$$D((x_1,x_2),(y_1,y_2)) = \text{Max}\,(d(x_1,y_1), d(x_2,y_2)) .$$

Munissons $\mathbf{R}$ de la distance usuelle, et considérons la distance d comme une application de l'espace métrique $(X \times X, D)$ dans $\mathbf{R}$. Elle est alors 2-lipschitzienne. En effet, pour (x_1,x_2) et (y_1,y_2) dans $X \times X$, on a, d'après (D3) :

$$\begin{aligned} |d(x_1,x_2) - d(y_1,y_2)| &\leq d(x_1,y_1) + d(x_2,y_2) \\ &\leq 2D((x_1,x_2),(y_1,y_2)) . \end{aligned}$$

4.3. Composition.

Soient (X_3, d_3) un troisième espace métrique, et $g : X_2 \to X_3$ une application.

PROPOSITION 7. *Soit $a \in X_1$. Si f est continue en a et g est continue en $f(a)$, alors $g \circ f$ est continue en a.*

DÉMONSTRATION : Si $(x_n)_{n\geq 0}$ converge vers a, alors $(f(x_n))_{n\geq 0}$ converge vers $f(a)$ (continuité de f en a) donc $(g(f(x_n)))_{n\geq 0}$ converge vers $g \circ f(a)$ (continuité de g en $f(a)$). Donc $g \circ f$ est continue en a. □

COROLLAIRE. *Si f et g sont continues, $g \circ f$ est continue.*

REMARQUE. Si f est K-lipschitzienne et g est K'-lipschitzienne (avec $K, K' > 0$), alors $g \circ f$ est KK'-lipschitzienne.

4.4. Homéomorphismes.

DÉFINITION 12. L'application f est un **homéomorphisme** si elle est continue, bijective, et l'application réciproque f^{-1} est continue.

L'application réciproque f^{-1} est alors aussi un homéomorphisme.
Si f et g sont des homéomorphismes, $g \circ f$ est un homéomorphisme.

Un homéomorphisme f transporte :

- parties ouvertes de X_1 en parties ouvertes de X_2
- parties fermées de X_1 en parties fermées de X_2
- intérieur d'une partie dans X_1 en intérieur de la partie image dans X_2
- adhérence d'une partie dans X_1 en adhérence de la partie image dans X_2
- suites convergentes dans X_1 en suites convergentes dans X_2
- limite d'une suite dans X_1 en limite de la suite image dans X_2
- points d'accumulation d'une suite dans X_1 en points d'accumulation de la suite image dans X_2.

EXEMPLE 8. Soient X un ensemble et d_1, d_2 deux distances équivalentes sur X. L'application identique $(X, d_1) \xrightarrow{i} (X, d_2)$ est un homéomorphisme : elle est bien sûr bijective, et, si C désigne la constante de la définition 3', alors i et i^{-1} sont C-lipschitziennes, donc continues.

En particulier, dans $\mathbf{R}^n$, les parties ouvertes, fermées et les notions d'adhérence, d'intérieur, de limite, de point d'accumulation **sont les mêmes** pour la norme euclidienne ou les normes $\| \ \|_1$, $\| \ \|_\infty$ (et en fait pour toutes les normes, car on verra que deux normes quelconques sur $\mathbf{R}^n$ sont équivalentes).

CHAPITRE II - Compacité.

§. 1. Espaces métriques compacts.

DÉFINITION 1. Un espace métrique (X, d) est **compact** si toute suite de points de X a au moins un point d'accumulation.

Une partie Y d'un espace métrique (X, d) est compacte si l'espace métrique (Y, d) est compact.

REMARQUE. La propriété, pour une partie Y d'un espace métrique (X, d), d'être compacte est une **propriété intrinsèque**, c.à.d. ne dépend que de l'espace métrique (Y, d), à la différence des propriétés d'être ouverte, ou fermée qui sont des propriétés extrinsèques, dépendant de l'espace métrique ambiant X.

PROPOSITION 1. *Soient (X_1, d_1), (X_2, d_2) des espaces métriques et $f : X_1 \to X_2$ une application continue. Si X_1 est compact, alors $f(X_1)$ est une partie compacte de X_2.*

DÉMONSTRATION : Soit $(y_n)_{n\geq 0}$ une suite dans $f(X_1)$; pour $n \geq 0$, il existe $x_n \in X_1$ tel que $f(x_n) = y_n$. Comme X_1 est compact, il existe une suite extraite $(x_{n_k})_{k\geq 0}$ qui converge vers une limite $a \in X_1$. Comme f est continue, la suite $(y_{n_k})_{k\geq 0} = (f(x_{n_k}))_{k\geq 0}$ converge vers $f(a) \in f(X_1)$.□

COROLLAIRE. *Soient (X_1, d_1), (X_2, d_2) deux espaces métriques et $f : X_1 \to X_2$ un homéomorphisme. Alors X_1 est compact si et seulement si X_2 est compact.*

PROPOSITION 2. *Soit Y une partie d'un espace métrique (X, d). Si Y est compact, Y est fermée dans X. Si X est compact et Y est fermée dans X, Y est compacte.*

DÉMONSTRATION : Supposons Y compacte. Une suite de points de Y, convergente dans X, a pour seul point d'accumulation dans X sa limite, celle-ci appartient donc à Y et Y est fermée (I 3.3, Corollaire).

Supposons X compact et Y fermée. Soit $(x_n)_{n\geq 0}$ une suite de points de Y. Comme X est compact, elle possède une suite extraite convergente dans X ; comme Y est fermée la limite de cette suite extraite appartient à Y (I 3.3, Corollaire).

PROPOSITION 3. *Soient $(X_1, d_1), \ldots, (X_k, d_k)$ des espaces métriques compacts. L'espace produit (X, D) est compact.*

DÉMONSTRATION :

1) $k = 2$. Soit $(x_n)_{n\geq 0} = (x_n^1, x_n^2)_{n\geq 0}$ une suite dans $X_1 \times X_2$. Comme X_1 est compact, il existe une suite extraite $(x_{n_k}^1)_{k\geq 0}$ convergente;

comme X_2 est compact, il existe une suite extraite $(x^2_{n_{k_\ell}})_{\ell\geq 0}$ de $(x^2_{n_k})_{k\geq 0}$ convergente dans X_2. La suite $(x_{n_{k_\ell}})_{\ell\geq 0}$ est alors convergente.

2) Cas général : par récurrence sur k. Si l'assertion de la proposition est valide pour $k-1$, alors $X_1\times\ldots\times X_{k-1}$ et X_k sont compacts, donc $X = X_1\times\ldots\times X_k$ est compact d'après 1). □

§. 2. Parties compactes de $\mathbf{R}^n$.

THÉORÈME 1 (BOLZANO-WEIERSTRASS). *L'intervalle* $[0,1]$ *est compact.*

DÉMONSTRATION : Elle dépend de la construction de $\mathbf{R}$; nous partons de l'idée intuitive que tout développement décimal illimité $0,\varepsilon_1\varepsilon_2\varepsilon_3\ldots$ (avec $\varepsilon_i\in\{0,1,\ldots,9\}$ représente un nombre réel $\in[0,1]$, et qu'inversement tout nombre réel $\in[0,1]$ est représentable par un tel développement (pour fixer les idées, nous choisissons, lorsque le nombre est décimal, le développement formé de 9 à partir d'un certain rang).

Soit $(x^0_n)_{n\geq 0}$ une suite de nombres réels appartenant à $[0,1]$. Notons $\varepsilon_i(x)$ la $i^{\text{ème}}$ décimale d'un nombre $x\in[0,1]$. Il existe $\theta_1\in\{0,1,\ldots,9\}$ et une suite extraite $(x^1_n)_{n\geq 0}$ de $(x^0_n)_{n\geq 0}$ telle qu'on ait $\varepsilon_1(x^1_n)=\theta_1$ pour $n\geq 0$.

On construit de même, par récurrence sur k, un entier $\theta_k\in\{0,1,\ldots,9\}$ et une suite $(x^k_n)_{n\geq 0}$ extraite de $(x^{k-1}_n)_{n\geq 0}$ telle qu'on ait $\varepsilon_k(x^k_n)=\theta_k$ pour $n\geq 0$.

Posons $y_k = x^k_k$ pour $k\geq 0$; la suite $(y_k)_{k\geq 0}$ est extraite de $(x^0_n)_{n\geq 0}$; on a, pour tout $i\geq 1$, $\varepsilon_i(y_k)=\theta_i$ pour $k\geq i$ (car alors y_k est l'un des x^i_n).

Soit y le nombre réel représenté par $0,\theta_1\theta_2\ldots$ On a

$$|y-y_k|\leq 10^{-k}\ ,\quad \forall k\geq 0$$

donc $(y_k)_{k\geq 0}$ converge vers y. □

DÉFINITION 2. Une partie Y d'un espace vectoriel normé $(E,\|\ \|)$ est **bornée** s'il existe $R>0$ tel qu'on ait $\|y\|\leq R$ pour tout $y\in Y$.

La propriété pour une partie d'être bornée ou non subsiste lorsqu'on remplace la norme de E par une norme équivalente.

PROPOSITION 4. *Les parties compactes de* $(\mathbf{R}^n,\|\ \|_\infty)$ *sont les parties fermées et bornées.*

DÉMONSTRATION : Soit Y une partie compacte de $(\mathbf{R}^n,\|\ \|_\infty)$. Alors Y est fermée (prop. 2). Si Y n'était pas bornée, il existerait une suite $(y_n)_{n\geq 0}$ dans Y vérifiant $\|y_n\|_\infty\geq n$ pour $n\geq 0$. Or une telle suite n'a pas de point d'accumulation.

Soit Y une partie fermée et bornée de $(\mathbf{R}^n,\|\ \|_\infty)$. Soit $R>0$ tel que $Y\subset\overline{B}(0,R)=[-R,+R]^n$. L'intervalle $[-R,+R]$ est homéomorphe à

$[0,1]$, donc compact (cor. de la prop. 1 et théorème 1). Par suite, le produit $[-R,+R]^n$ est compact (prop. 3). Comme Y est une partie fermée de $[-R,+R]^n$, Y est compacte (prop. 2). $\square$

COROLLAIRE 1. *Soit Y une partie compacte non vide de* $\mathbf{R}$. *Alors il existe a, b* dans $\mathbf{Y}$ *tels que $Y \subset [a,b]$.*

DÉMONSTRATION : Comme Y est bornée, elle possède une borne inférieure a et une borne supérieure b. Comme Y est fermée, ces bornes appartiennent à Y.

COROLLAIRE 2. *Soient (X,d) un espace métrique compact (non vide) et $f : X \to \mathbf{R}$ une application continue. Il existe a, b dans X tels qu'on ait $f(a) \leq f(x) \leq f(b)$ pour tout $x \in X$.*

("Une fonction continue sur un espace compact est bornée et atteint ses bornes").

Cela résulte de la proposition 1 et du corollaire 2.

La proposition 4 n'est pas vraie pour un espace vectoriel normé de dimension infinie : c'est le théorème de Riesz (voir problème).

PROPOSITION 5. *Toutes les normes sur $\mathbf{R}^n$ sont équivalentes.*

DÉMONSTRATION : L'équivalence entre normes est une relation d'équivalence (exerc.). Il suffit donc de montrer que tout norme $\| \ \|$ sur $\mathbf{R}^n$ est équivalente à $\| \ \|_\infty$.

Soit $S = \{x \in \mathbf{R}^n, \|x\|_\infty = 1\}$ la sphère unité de $(\mathbf{R}^n, \| \ \|_\infty)$. Dans $(\mathbf{R}^n, \| \ \|_\infty)$, S est bornée et fermée (car $\| \ \|_\infty$ est continue), donc compacte.

D'autre part, en notant $(e_1, \dots, e_n)$ la base canonique de $\mathbf{R}^n$, on a, pour $x = (x_1, \dots, x_n) \in \mathbf{R}^n$:

$$\|x\| \leq \sum_{i=1}^{n} |x_i| \, \|e_i\| \leq C_1 \, \|x\|_\infty \quad ,$$

avec $C_1 = \sum_{i=1}^{n} \|e_i\|$, d'où, pour $x, y \in \mathbf{R}^n$:

$$\Big| \|x\| - \|y\| \Big| \leq \|x - y\| \leq C_1 \, \|x - y\|_\infty \quad .$$

Ceci prouve que l'application $x \mapsto \|x\|$ de $(\mathbf{R}^n, \| \ \|_\infty)$ dans $\mathbf{R}$ est C_1-Lipschitzienne donc continue.

D'après le corollaire 2 ci-dessus, il existe $a \in S$ tel qu'on ait $\|x\| \geq \|a\| > 0$ pour $x \in S$. Pour tout x non nul dans $\mathbf{R}^n$, on a alors :

$$\|x\| = \|x\|_\infty \left\| \frac{x}{\|x\|_\infty} \right\| \geq \|a\| \, \|x\|_\infty \quad ,$$

donc les normes $\| \ \|$ et $\| \ \|_\infty$ sont équivalentes. $\square$

§. 3. Compacité et recouvrements.

DÉFINITION 3. Une famille $(A_i)_{i\in I}$ de parties d'un ensemble X est un **recouvrement** de X si $\bigcup_{i\in I} A_i = X$.

Un **sous-recouvrement** est une sous-famille $(A_j)_{j\in J}$, $J \subset I$ telle que $\bigcup_{j\in J} A_j = X$.

THÉORÈME 2 (LEBESGUE). *Soit $(U_i)_{i\in I}$ un recouvrement par parties* ***ouvertes*** *d'un espace métrique* ***compact*** *(X,d). Il existe alors $r > 0$ tel que toute boule ouverte de rayon r soit contenue dans l'un des U_i.*

DÉMONSTRATION : Sinon il existe une suite $(x_n)_{n\geq 1}$ telle que $B\left(x_n, \frac{1}{n}\right)$ pour $n \geq 1$ ne soit contenue dans aucun des U_i. Soit a un point d'accumulation de la suite $(x_n)_{n\geq 1}$. Il existe $i_0 \in I$ et $\varepsilon > 0$ tels que $B(a,\varepsilon) \subset U_{i_0}$. Il existe $n > 2\varepsilon^{-1}$ tel que $d(a,x_n) \leq \frac{\varepsilon}{2}$. On a alors

$$B\left(x_n, \frac{1}{n}\right) \subset B\left(x_n, \frac{\varepsilon}{2}\right) \subset B(a,\varepsilon) \subset U_{i_0} ,$$

une contradiction. □

PROPOSITION 6. *Soient (X,d) un espace métrique compact, et $r > 0$. Il existe un recouvrement* ***fini*** *de X par des boules ouvertes de rayon r.*

DÉMONSTRATION : Sinon, on pourrait construire une suite $(x_n)_{n\geq 0}$ telle qu'on ait $d(x_i,x_j) \geq r$ pour $i \neq j$. Une telle suite ne peut avoir de point d'accumulation.

THÉORÈME 3 (BOREL-LEBESGUE). *Pour qu'un espace métrique (X,d) soit compact, il faut et il suffit que tout recouvrement de X par parties* ***ouvertes*** *possède un sous-recouvrement* ***fini****.*

DÉMONSTRATION : Supposons (X,d) compact, et soit $(U_i)_{i\in I}$ un recouvrement par parties ouvertes de X.

Soit $r > 0$ satisfaisant la conclusion du théorème 2. D'après la prop. 6, il existe des boules ouvertes $B_1,\dots,B_k$ de rayon r telles que $X = \bigcup_{\ell=1}^{k} B_\ell$. D'après le théorème 2, pour tout $1 \leq \ell \leq k$, il existe $i_\ell \in I$ tel que $B_\ell \subset U_{i_\ell}$. On a alors

$$X = \bigcup_{\ell=1}^{k} U_{i_\ell} .$$

Inversement, si (X,d) n'est pas compact, il existe une suite $(x_n)_{n\geq 0}$ dans X sans point d'accumulation. Pour tout $x \in X$, il existe alors une boule ouverte B_x centrée en x ne contenant au plus qu'un nombre fini de valeurs de la suite $(x_n)_{n\geq 0}$. La famille $(B_x)_{x\in X}$ est un recouvrement de X par parties ouvertes, mais l'union de toute sous-famille finie ne peut contenir qu'un nombre fini de valeurs de la suite $(x_n)_{n\geq 0}$. □

COROLLAIRE. *Soit $(F_n)_{n\geq 0}$ une suite décroissante (pour l'inclusion) de parties fermées non vides d'un espace métrique compact (X,d). On a* $\bigcap_{n\geq 0} F_n \neq \phi$.

DÉMONSTRATION : Posons $U_n = X - F_n$. Si on a $\bigcap_{n\geq 0} F_n = \phi$, la famille croissante $(U_n)_{n\geq 0}$ est un recouvrement par parties ouvertes de X et il existe n_0 tel que $U_{n_0} = X$, c'est-à-dire $F_{n_0} = \phi$. □

§. 4. Uniforme continuité.

Soient (X_1, d_1), (X_2, d_2) des espaces métriques et $f : X_1 \to X_2$ une application.

DÉFINITION 4. L'application f est **uniformément continue** si pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que :

$$d_1(x,y) < \delta \Rightarrow d_2(f(x), f(y)) < \varepsilon .$$

Une application K-lipschitzienne est uniformément continue. Une application uniformément continue est continue.

THÉORÈME 4 (HEINE). *Si X_1 est compact et f est continue, alors f est uniformément continue.*

DÉMONSTRATION : Soit $\varepsilon > 0$; comme f est continue, il existe, pour tout $x \in X_1$, une boule ouverte B_x centrée en x dont l'image est contenue dans $B\left(f(x), \dfrac{\varepsilon}{2}\right)$. On a $\bigcup_{x\in X_1} B_x = X_1$, et X_1 est compact, donc il existe $\delta > 0$ tel que toute boule ouverte de rayon δ est contenue dans une boule B_x (théorème 2). Si $y, z \in X_1$ vérifient $d_1(y,z) < \delta$, on a donc $d_2(f(y), f(z)) < \varepsilon$. □

PROPOSITION 7. *Si X_1 est compact et f est bijective et continue, alors f est un homéomorphisme.*

DÉMONSTRATION : Il s'agit de vérifier que l'application réciproque f^{-1} est continue. Or, si Y est une partie fermée de X_1, Y est compacte car X_1 est compact (prop. 2) ; donc $(f^{-1})^{-1}(Y) = f(Y)$ est compacte (Prop. 1), et par conséquent fermée (prop. 2). □

CHAPITRE III - Connexité.

§. 1. Espaces métriques connexes.

DÉFINITION 1. Un espace métrique (X, d) est **connexe** si les seules parties à la fois ouvertes et fermées dans X sont la partie vide ϕ et la partie totale X.

Une partie Y d'un espace métrique (X, d) est connexe si l'espace métrique (Y, d) est connexe.

La connexité, comme la compacité, est donc une propriété intrinsèque.

PROPOSITION 1. *Soit (X, d) un espace métrique non vide. Les propriétés suivantes sont équivalentes :*

(i) *(X, d) n'est pas connexe,*

(ii) *X est réunion de deux parties ouvertes disjointes non vides ;*

(iii) *X est réunion de deux parties fermées disjointes non vides ;*

(iv) *il existe une application continue non constante de X dans $\{0, 1\}$.*

DÉMONSTRATION : Exercice.

PROPOSITION 2. *Soient (X_1, d_1), (X_2, d_2) deux espaces métriques, et $f : X_1 \to X_2$ une application continue. Si X_1 est connexe, alors $f(X_1)$ est connexe.*

DÉMONSTRATION : Si $f(X_1)$ n'est pas connexe, il existe une application continue non constante g de $f(X_1)$ dans $\{0, 1\}$. L'application composée $g \circ f : X_1 \to \{0, 1\}$ est continue et non constante, donc X_1 n'est pas connexe. $\square$

COROLLAIRE. *Soient (X_1, d_1), (X_2, d_2) deux espaces métriques, et $f : X_1 \to X_2$ un homéomorphisme. Alors X_1 est connexe si et seulement si X_2 est connexe.*

PROPOSITION 3. *Soient (X, d) un espace métrique, et $(A_i)_{i \in I}$ une famille de parties connexes d'intersection non vide. Alors $\bigcup_{i \in I} A_i$ est connexe.*

DÉMONSTRATION : Une application continue : $\bigcup_{i \in I} A_i \to \{0, 1\}$ est constante sur chacun des A_i (prop. 1) donc sur $\bigcup_{i \in I} A_i$, car $\bigcap_{i \in I} A_i \neq \phi$. $\square$

PROPOSITION 4. *Soit Y une partie connexe d'un espace métrique (X,d). L'adhérence $\bar{Y}$ de Y dans X est connexe.*

DÉMONSTRATION : Soit $g : \bar{Y} \to \{0,1\}$ une application continue. Alors g est constante sur Y, et prend encore la même valeur sur $\bar{Y}$ d'après les propositions 3 et 4 du Chapitre I. □

REMARQUE. La même démonstration prouve que toute partie Z telle que $Y \subset Z \subset \bar{Y}$ est connexe.

PROPOSITION 5. *Soient $(X_1,d_1),\ldots,(X_k,d_k)$ des espaces métriques connexes. L'espace produit (X,D) est connexe.*

DÉMONSTRATION :

1) $k=2$. Fixons $x_0 \in X_1$. Pour tout $y \in X_2$, posons :

$$A_y = \Big(\{x_0\} \times X_2\Big) \cup \Big(X_1 \times \{y\}\Big) .$$

C'est une partie connexe de $X_1 \times X_2$ d'après la proposition 3. On a

$$\bigcap_{y \in X_2} A_y = \{x_0\} \times X_2 \ ,$$

$$\bigcup_{y \in X_2} A_y = X_1 \times X_2 \ ,$$

donc $X_1 \times X_2$ est connexe (proposition 3).

2) **Cas général** : cf. Chapitre II, prop. 3.

§. 2. Parties connexes de $\mathbf{R}$.

THÉORÈME 1. *Les parties connexes de $\mathbf{R}$ sont les intervalles (non vides).*

DÉMONSTRATION : Soit J une partie non vide de $\mathbf{R}$ qui n'est pas un intervalle. Il existe un point $c \in \mathbf{R} - J$, tel que $J \cap]-\infty, c[$ et $J \cap]c, +\infty[$ ne sont pas vides. Ce sont des parties ouvertes et disjointes de J dont J est l'union; donc J n'est pas connexe.

Soit J un intervalle compact non vide de $\mathbf{R}$: $J = [a,b]$; supposons que J soit union de deux parties fermées disjointes non vides F_0 et F_1, avec $a \in F_0$. Soit c la borne supérieure des nombres $d \in [a,b]$ tels que $[a,d] \subset F_0$. On a $[a,c[\subset F_0$, donc $[a,c] \subset F_0$ puisque F_0 est fermé. On a $c < b$ puisque F_1 n'est pas vide. Par définition de c, tout intervalle $[c,d]$, avec $c < d \leq b$, coupe F_1 (autrement $[a,d] \subset F_0$); comme F_1 est fermé, on a donc $c \in F_1$, et $F_0 \cap F_1 \neq \phi$. Donc J est connexe. Soient J un intervalle non vide de $\mathbf{R}$, $a \in J$; on a

$$J = \bigcup_{\substack{x \geq a \\ x \in J}} [a,x] \cup \bigcup_{\substack{x \leq a \\ x \in J}} [x,a] \ ,$$

donc J est connexe (proposition 3).

COROLLAIRE (THÉORÈME DES VALEURS INTERMÉDIAIRES). *Soit* (X,d) *un espace métrique* **connexe** *et* $f : X \to \mathbf{R}$ *une application continue. Alors* $f(X)$ *est un intervalle : en particulier, si* a, b *sont des points de* X *et* z *un nombre réel tels que* $f(a) \leq z \leq f(b)$, *il existe un point* c *de* X *tel que* $f(c) = z$.

DÉMONSTRATION : Proposition 2 et théorème 1.

§. 3. Connexité par arcs. Composantes connexes.

DÉFINITION 2. Soient (X,d) un espace métrique, et x, y deux points de X. Un **chemin de x à y** est une application **continue** $\gamma : [0,1] \to X$ telle que $\gamma(0) = x$, $\gamma(1) = y$.

DÉFINITION 3. Un espace métrique (X,d) est **connexe par arcs** si pour tous $x, y \in X$ il existe un chemin de x à y.

PROPOSITION 6. *Un espace métrique* (X,d) *connexe par arcs est connexe.*

DÉMONSTRATION : Soit $x_0 \in X$; pour tout $y \in X$, soit $\gamma_y : [0,1] \to X$ un chemin de x_0 à y ; posons $A_y = \gamma_y([0,1])$. D'après la proposition 2 et le théorème 1, A_y est connexe. On a :

$$\bigcup_{y \in X} A_y = X \ , \quad \bigcap_{y \in X} A_y \ni x_0 \ ;$$

d'après la proposition 3, X est connexe.

COROLLAIRE. *Une partie* **convexe** *de* $\mathbf{R}^n$ *est connexe.*

DÉMONSTRATION : En effet, elle est connexe par arcs.

(On rappelle qu'une partie C de $\mathbf{R}^n$ est convexe si pour tout $x, y \in C$, le segment joignant x à y est contenu dans C).

Les deux lemmes suivants serviront pour la proposition 7, et sont importants par eux-mêmes.

LEMME 1. *Soit* (X,d) *un espace métrique. La relation :*

" il existe un chemin de x à y "

est une relation d'équivalence.

DÉMONSTRATION :

Réflexivité. L'application $\gamma : [0,1] \to X$ constante égale à x est un chemin de x à x.

Symétrie. Si $\gamma : [0,1] \to X$ est un chemin de x à y, l'application $\overline{\gamma}(t) = \gamma(1-t)$ est un chemin de y à x.

Transitivité. Si $\gamma : [0,1] \to X$ est un chemin de x à y et $\gamma' : [0,1] \to X$ est un chemin de y à z, l'application γ'' définie par :

$$\gamma''(t) = \gamma(2t) \qquad 0 \leq t \leq \frac{1}{2}\,,$$
$$\gamma''(t) = \gamma'(2t-1) \quad \frac{1}{2} \leq t \leq 1$$

est un chemin de x à z.

LEMME 2. *Soient (X,d) un espace métrique* **connexe** *et $\mathcal{R}$ une relation d'équivalence sur X ; si chaque classe d'équivalence de $\mathcal{R}$ est* **ouverte** *dans X, alors il existe une seule classe (on a $x\mathcal{R}y$, $\forall x,y \in X$).*

DÉMONSTRATION : Soit C une classe d'équivalence; alors C est ouverte, non vide; $X - C$ est l'union des autres classes d'équivalences, donc est ouverte.

Comme X est connexe, on a $X - C = \phi$, d'où $X = C$.

PROPOSITION 7. *Une partie* **ouverte et connexe** *U de $\mathbf{R}^n$ est connexe par arcs.*

DÉMONSTRATION : Soit $\mathcal{R}$ la relation sur U :

" il existe un chemin de x à y (contenu dans U) " .

C'est une relation d'équivalence (lemme 1). Si $x \in U$, la classe de x contient toute boule de centre x contenue dans U. Les classes d'équivalence sont donc ouvertes. D'après le lemme 2, U est connexe par arcs. □

Introduisons finalement la notion de composante connexe. Soient (X,d) un espace métrique, a un point de X. L'union des parties **connexes** de X contenant a est une partie connexe (proposition 3) : la **composante connexe** de a dans X. C'est la plus grande partie connexe de X contenant a.

D'un point de vue un peu différent, considérons la relation sur X :

" il existe une partie connexe contenant x et y" .

C'est une relation d'équivalence (la transitivité résulte de la proposition 3), et la classe d'équivalence d'un point $a \in X$ est exactement la composante connexe de a dans X.

On a donc une **partition** de X en ses composantes connexes. Bien sûr, X est connexe si et seulement si X a une seule composante connexe.

Soit U une partie **ouverte** de $\mathbf{R}^n$. Soit $a \in U$. La composante connexe de a dans U contient toute boule de centre a contenue dans U.

Les composantes connexes de U sont donc des parties **ouvertes** de $\mathbf{R}^n$. On obtient donc, à partir de la proposition 7, le :

COROLLAIRE. *Soient U une partie* **ouverte** *de $\mathbf{R}^n$, x, y deux points de U. Alors il existe un chemin de x à y si et seulement si x et y appartiennent à la même composante connexe de U.*

§. 4. Compacts connexes.

PROPOSITION 8. *Soient (X, d) un espace métrique, et $(K_n)_{n\geq 0}$ une suite décroissante (pour l'inclusion) de parties **compactes, connexes, non vides** de X. Alors l'intersection $K = \bigcap_{n\geq 0} K_n$ est **compacte, connexe, non vide.***

DÉMONSTRATION : On peut supposer que $X = K_0$ (car compacité et connexité sont des propriétés intrinsèques). D'après le corollaire du Chapitre II, §.3, K est une partie compacte non vide de X. Supposons que K n'est pas connexe; il existe alors deux parties fermées (donc compactes) L_0, L_1 disjointes et non vides, dont K est l'union. L'application $(x, y) \mapsto d(x, y)$ de $L_0 \times L_1$ dans $\mathbf{R}$ est continue, et $L_0 \times L_1$ est compact, donc il existe $a \in L_0$, $b \in L_1$ tels que $d(x, y) \geq d(a, b) = \varepsilon > 0$, pour $x \in L_0$, $y \in L_1$. Posons :

$$U_0 = \bigcup_{x \in L_0} B\left(x, \frac{\varepsilon}{2}\right) ,$$

$$U_1 = \bigcup_{x \in L_1} B\left(x, \frac{\varepsilon}{2}\right) .$$

Alors U_0 et U_1 sont des parties ouvertes et disjointes dont l'union contient K. Posons :

$$K'_0 = K_0 - (U_0 \cup U_1) ,$$
$$K'_n = K_n \cap K'_0 , \qquad \text{pour } n \geq 1 .$$

Les K'_n forment une suite décroissante de parties compactes de K_0, dont l'intersection est vide.

D'après le corollaire du Chapitre II, §.3, il existe un entier p tel que $K'_p = \phi$. On a alors :

$$K_p = (K_p \cap U_0) \cup (K_p \cap U_1) ,$$

donc K_p n'est pas connexe. □

CHAPITRE IV - Espaces métriques complets et espaces de Banach.

§. 1. Suites de Cauchy. Espaces métriques complets.

Soit (X, d) un espace métrique.

DÉFINITION 1. Une suite $(x_n)_{n \geq 0}$ de points de X est une **suite de Cauchy** si

$$\lim_{\substack{m \to +\infty \\ n \to +\infty}} d(x_m, x_n) = 0 \ .$$

(En d'autre termes, pour tout $\varepsilon > 0$, il existe un entier n_0 tel qu'on ait $d(x_m, x_n) < \varepsilon$ lorsque m et n sont supérieurs à n_0).

PROPOSITION 1. *Toute suite convergente est une suite de Cauchy.*

DÉMONSTRATION : Exercice.

DÉFINITION 2. L'espace métrique (X, d) est **complet** si toute suite de Cauchy dans X est convergente.

Une partie Y de X est complète si l'espace métrique (Y, d) est complet.

La propriété pour une partie d'être complète est donc une propriété intrinsèque.

PROPOSITION 2. *Soit Y une partie de X. Si Y est complète, Y est fermée dans X. Si X est complet et Y est fermée dans X, Y est complète.*

DÉMONSTRATION : Supposons Y complète. Toute suite de points de Y, qui converge dans X, est une suite de Cauchy; elle doit donc converger dans Y ; par conséquent Y est fermée (Chapitre I, §. 3.3, Corollaire).

Supposons X complet et Y fermée; une suite de Cauchy de points de Y converge alors dans X (qui est complet) donc dans Y (qui est fermée dans X).

LEMME. *Une suite de Cauchy qui a un point d'accumulation est convergente.*

DÉMONSTRATION : Exercice.

PROPOSITION 3. *Un espace métrique compact est complet.*

DÉMONSTRATION : Cela résulte du lemme et de la définition d'un espace compact.

PROPOSITION 4. *Supposons que toutes les parties fermées et bornées de* (X,d) *soient* ***complètes****, alors* X *est complet.*

(Une partie Y d'un espace métrique (X,d) est bornée s'il existe $C > 0$ tel qu'on ait $d(x,y) < C$ pour $x, y \in Y$).

DÉMONSTRATION : Soit $(x_n)_{n\geq 0}$ une suite de Cauchy dans X. Posons :

$$Z = \left\{x_n, n \geq 0\right\} \ , \ Y = \bar{Z} \ ,$$

et montrons que Y est bornée. Il existe un entier n_0 tel qu'on ait $d(x_n, x_{n_0}) < 1$ pour $n \geq n_0$. Posons :

$$C_0 = \max_{0 \leq i \leq n_0} d(x_i, x_{n_0}) \ .$$

Soit y un point de Y ; la boule $B(y,1)$ rencontre Z (Chapitre I, § 3.3, proposition 3), donc il existe un entier n tel que $d(y, x_n) < 1$. On a par conséquent $d(y, x_{n_0}) < 2 + C_0$, et finalement $d(y_1, y_2) < 4 + 2C_0$ pour deux points quelconques y_1, y_2 de Y.

La partie Y est donc bornée; elle est aussi fermée, donc complète. La suite de Cauchy $(x_n)_{n\geq 0}$ est une suite de points de Y. Elle est donc convergente.

COROLLAIRE. *L'espace* $\mathbb{R}^n$ *(muni d'une norme quelconque) est complet.*

DÉMONSTRATION : En effet, les parties fermées et bornées sont compactes, donc complètes.

REMARQUE IMPORTANTE. Soient (X_1, d_1), (X_2, d_2) deux espaces métriques, et $f : X_1 \to X_2$ un homéomorphisme. Il est possible que X_1 soit complet sans que X_2 le soit, et vice versa. Par exemple, l'application $x \mapsto \operatorname{tg}(x)$ est un homéomorphisme de l'intervalle $I = \left]-\frac{\pi}{2}, +\frac{\pi}{2}\right[$ sur $\mathbb{R}$. Or $\mathbb{R}$ est complet (Corollaire), mais I qui n'est pas fermé dans $\mathbb{R}$ n'est pas complet (proposition 2).

Cependant, si on suppose de plus que f est **uniformément continue** et si X_2 est complet, alors X_1 est complet (exercice).

PROPOSITION 5. *Soient* $(X_1, d_1), \ldots, (X_k, d_k)$ *des espaces métriques complets. Alors l'espace produit* (X, D) *est complet.*

DÉMONSTRATION : Cela résulte du fait qu'une suite $(x_n)_{n\geq 0} = (x_n^1, \ldots, x_n^k)_{n\geq 0}$ dans X est convergente (resp. une suite de Cauchy) si et seulement si chacune des suites $(x_n^i)_{n\geq 0}$ est convergente (resp. une suite de Cauchy).

§. 2. Le théorème du point fixe.

DÉFINITION 3. Soit (X,d) un espace métrique. Une application $f: X \to X$ est **contractante** s'il existe $k \in \,]0,1[$ tel que f soit k-lipschitzienne, c'est-à-dire qu'on ait :

$$d(f(x), f(y)) \leq k\, d(x,y) \ , \quad \text{pour } x, y \in X \ .$$

DÉFINITION 4. Soient X un ensemble et $f: X \to X$ une application. Un **point fixe** de f est un point $a \in X$ tel que $f(a) = a$.

THÉORÈME 1. *Soient (X,d) un espace métrique **complet**, $f: X \to X$ une application **contractante** et x_0 un point de X. Soit $(x_n)_{n\geq 0}$ la suite vérifiant $x_{n+1} = f(x_n)$ pour $n \geq 0$.*

***Alors**, la suite $(x_n)_{n\geq 0}$ est convergente, et sa limite est **l'unique point fixe** de f.*

DÉMONSTRATION : Soit $k \in \,]0,1[$ tel que f soit k-lipschitzienne. Pour $n \geq 1$, on a :

$$\begin{aligned} d(x_n, x_{n+1}) &= d(f(x_{n-1}), f(x_n)) \\ &\leq k\, d(x_{n-1}, x_n) \ , \end{aligned}$$

donc

$$d(x_n, x_{n+1}) \leq k^n\, d(x_0, x_1) \ .$$

On en déduit, pour $m \geq n \geq 0$:

$$\begin{aligned} d(x_n, x_m) &\leq d(x_0, x_1) \sum_{\ell=n}^{m-1} k^\ell \\ &\leq k^n \ \frac{d(x_0, x_1)}{1-k} \ ; \end{aligned}$$

ceci prouve que $(x_n)_{n\geq 0}$ est une suite de Cauchy. Comme X est complet, elle est convergente; notons a sa limite. Comme f est continue, on a :

$$f(a) = \lim_{n\to+\infty} f(x_n) = \lim_{n\to+\infty} x_{n+1} = a \ ,$$

donc a est un point fixe de f. Finalement, soit b un point fixe de f. On a :

$$d(a,b) = d(f(a), f(b)) \leq k\, d(a,b) \ ,$$

avec $k \in \,]0,1[$, donc $d(a,b) = 0$ et $a = b$. □

Le théorème du point fixe a une **très grande importance** tant théorique que pratique. En effet, de très nombreux problèmes peuvent se présenter sous la forme de recherche de points fixes. Le théorème du point fixe est à la base de la démonstration des principaux théorèmes d'Analyse du cours : théorème d'inversion locale, théorème des fonctions implicites, théorème d'existence et d'unicité des solutions d'équations différentielles (théorème de Cauchy-Lipschitz).

D'un point de vue théorique, il faut surtout retenir l'existence et l'unicité du point fixe de l'application contractante f.

D'un point de vue pratique, la convergence de la suite $(x_n)_{n\geq 0}$ vers le point fixe permet de **déterminer** le point fixe avec une précision arbitraire; on a de plus un contrôle sur l'erreur commise; on a vu que :

$$d(x_m, x_n) \leq k^n (1-k)^{-1} d(x_0, x_1)$$

pour $m \geq n \geq 0$, et on a donc :

$$d(a, x_n) \leq k^n (1-k)^{-1} d(x_0, x_1) \ .$$

EXEMPLE. Etant donné un nombre réel x_0, on définit une suite $(x_n)_{n\geq 0}$ par la relation de récurrence :

$$x_{n+1} = x_n^2 - 100 + \sin(n) \ .$$

On va montrer qu'il existe une **unique** valeur de x_0 telle que la suite $(x_n)_{n\geq 0}$ soit **bornée** et **à valeurs positives**. On va aussi déterminer cette valeur à 10^{-3} près (avec une calculette).

1) Préliminaires. Montrons que si la suite $(x_n)_{n\geq 0}$ est bornée à valeurs positives, on a nécessairement $10 \leq x_n \leq 11$ pour tout $n \geq 0$.

Supposons qu'il existe un entier p tel que $x_p < 10$, on a alors $x_{p+1} = x_p^2 - 100 + \sin p < 1$, et (si $x_{p+1} \geq 0$), $x_{p+2} = x_{p+1}^2 - 100 + \sin(p+1) < -98$, contrairement à l'hypothèse que la suite est à valeurs positives.

Supposons qu'il existe un entier p tel que $x_p > 11$; on a alors :

$$\begin{aligned} x_{p+1} &= x_p^2 - 100 + \sin(p) \\ &> 11 x_p - 99 > 2x_p \ ; \end{aligned}$$

on a alors $x_{k+1} > 2x_k$ pour tout $k \geq p$, et la suite $(x_k)_{k\geq 0}$ n'est pas bornée.

2) L'espace métrique où on va appliquer le théorème du point fixe. Notons X l'ensemble des suites $(y_k)_{k\geq 0}$ à valeurs dans l'intervalle $[10, 11]$. Nous définissons une distance D sur X, la **distance uniforme** par :

$$D\Big((y_k)_{k\geq 0}, (z_k)_{k\geq 0}\Big) = \sup_{k\geq 0} |y_k - z_k| \ .$$

On verra au prochain paragraphe que l'espace des suites à valeurs dans un espace métrique compact (en l'occurence, $[10, 11]$), muni de la distance uniforme, est **complet**. Donc (X, D) est complet.

3) L'application contractante $F : X \to X$. Soit $y = (y_k)_{k\geq 0}$ un élément de X. Définissons une suite $z = F(y) = (z_k)_{k\geq 0}$ par la formule :

$$z_k = \sqrt{100 - \sin(k) + y_{k+1}}$$

(racine carrée **positive**). Nous vérifions successivement :

a) que $z = (z_k)_{k\geq 0}$ est un élément de X.

En effet, pour tout $k \geq 0$, on a $y_{k+1} \in [10, 11]$, donc

$$109 \leq 100 - \sin(k) + y_{k+1} \leq 112$$

et donc $z_k \in [10, 11]$.

On a donc bien défini une application F de X dans X.

b) que F est contractante (et on calcule la constante de contraction).

Soient $y = (y_k)_{k\geq 0}$, $y' = (y'_k)_{k\geq 0}$ deux éléments de X. Notons $z = F(y) = (z_k)_{k\geq 0}$ et $z' = F(y') = (z'_k)_{k\geq 0}$ leurs images par F. Pour $k \geq 0$, on a :

$$\begin{aligned}
|z_k - z'_k| &= \left|\sqrt{100 - \sin(k) + y_{k+1}} - \sqrt{100 - \sin(k) + y'_{k+1}}\right| \\
&= \frac{|y_{k+1} - y'_{k+1}|}{\sqrt{100 - \sin(k) + y_{k+1}} + \sqrt{100 - \sin(k) + y'_{k+1}}} \\
&\leq \frac{1}{20}\,|y_{k+1} - y'_{k+1}|\,,
\end{aligned}$$

car on a $y_{k+1} \geq 10$, $y'_{k+1} \geq 10$.

On a par conséquent :

$$D(z, z') = D(F(y), F(y')) \leq \frac{1}{20} D(y, y')$$

et l'application F est contractante.

4) Application du théorème du point fixe. L'espace métrique (X, d) est complet, et F est contractante, donc F a un **unique** point fixe $a = (a_k)_{k\geq 0}$; revenant à la définition de F, cela signifie qu'il existe une **unique** suite $(a_k)_{k\geq 0}$, à valeurs dans $[10, 11]$, qui vérifie :

$$a_k = \sqrt{100 - \sin(k) + a_{k+1}}\ , \quad \text{pour tout } k \geq 0$$

ou encore

$$a_{k+1} = a_k^2 - 100 + \sin(k)\ , \qquad \text{pour tout } k \geq 0\ .$$

Or, on a montré dans les préliminaires que si une suite $(x_k)_{k\geq 0}$ est à valeurs positives, bornée et vérifie la relation de récurrence :

$$x_{k+1} = x_k^2 - 100 + \sin(k)\ ,$$

cette suite est nécessairement à valeurs dans $[10, 11]$, en d'autre termes c'est un élément de X.

On a donc démontré l'existence et l'unicité du nombre réel x_0 qu'on avait annoncées : on a $x_0 = a_0$.

5) Détermination approchée de a_0. Soit $y^0 = (y^0_k)_{k\geq 0}$ l'élément de X tel que :

$$y^0_k = 10,5 \quad , \quad \text{pour tout } k \geq 0 \ .$$

Posons $y^1 = F(y^0)$. On a donc $y^1 = (y^1_k)_{k\geq 0}$, avec

$$y^1_k = \sqrt{100 - \sin(k) + 10,5} \ , \quad \text{pour tout } k \geq 0 \ .$$

Pour tout $k \geq 0$, on a :

$$10,46 < \sqrt{109,5} \leq y^1_k \leq \sqrt{111,5} < 10,56 \ .$$

On a donc

$$D(y^0, y^1) < 0.06 \ .$$

Posons $y^2 = F(y^1)$, $y^3 = F(y^2)$, etc...

On a vu plus haut que la distance de y^n au point fixe a de F vérifie :

$$D(a, y^n) < \left(\frac{1}{20}\right)^n \frac{1}{1 - \frac{1}{20}} D(y^0, y^1)$$

(où $\dfrac{1}{20}$ est le rapport de contraction de F). Pour que le second membre soit $< 10^{-3}$, il suffit de prendre $n = 2$; on aura alors, en écrivant $y^2 = (y^2_k)_{k\geq 0}$:

$$|a_0 - y^2_0| \leq D(a, y^2) < \frac{1}{380} 0.06 < 1,6 \ 10^{-4} \ .$$

Finalement on a :

$$y^2_0 = \sqrt{100 - \sin(0) + y^1_1} = \sqrt{100 + y^1_1} \ ,$$
$$y^1_1 = \sqrt{100 - \sin(1) + 10,5} \ ,$$

d'où

$$y^1_1 \approx 10,471797 \ ,$$
$$y^2_0 \approx 10,510556 \ ,$$
$$|a_0 - 10,5105| < 3.10^{-4} \ .$$

§. 3. Espaces d'applications continues et distance uniforme.

Soient (X_1, d_1), (X_2, d_2) deux espaces métriques. On note $\mathcal{C}(X_1, X_2)$ l'ensemble des applications continues de X_1 dans X_2, et $\mathcal{C}_b(X_1, X_2)$ le sous-ensemble des applications continues et bornées de X_1 dans X_2. (Une application $f : X_1 \to X_2$ est bornée si la partie $f(X_1)$ est bornée, c.à.d. s'il existe $C > 0$ tel qu'on ait $d_2(f(x), f(y)) < C$ pour tous $x, y \in X_1$).

REMARQUES.

1. Lorsque $X_1 = \mathsf{N}$ (ou Z) toute application de X_1 dans X_2 est continue; une telle application est simplement une suite à valeurs dans X_2. L'ensemble $\mathcal{C}(\mathsf{N}, X_2)$ est donc l'ensemble des suites à valeurs dans X_2, et $\mathcal{C}_b(\mathsf{N}, X_2)$ est le sous-ensemble des suites bornées.

2. Lorsque l'espace d'arrivée X_2 est borné, toute application continue est évidemment borné et on a donc $\mathcal{C}(X_1, X_2) = \mathcal{C}_b(X_1, X_2)$. C'est en particulier le cas lorsque X_2 est compact, puisqu'alors l'application continue $(x,y) \mapsto d_2(x,y)$ est bornée sur l'espace compact $X_2 \times X_2$.

3. Lorsque l'espace de départ X_1 est compact, on a encore $\mathcal{C}(X_1, X_2) = \mathcal{C}_b(X_1, X_2)$. Si $f : X_1 \to X_2$ est une application continue, la partie $f(X_1)$ est en effet compacte, donc bornée.

DÉFINITION 5. La formule

$$D(f,g) = \sup_{x \in X_1} d_2(f(x), g(x)) ,$$

pour $f, g \in \mathcal{C}_b(X_1, X_2)$ définit une distance sur l'ensemble $\mathcal{C}_b(X_1, X_2)$ dite **distance uniforme.** La convergence d'une suite (d'applications continues bornées) dans l'espace métrique $(\mathcal{C}_b(X_1, X_2)$, est dite **convergence uniforme.**

EXERCICE. Vérifier que la borne supérieure définissant $D(f,g)$ est finie; vérifier les axiomes $(D1)$, $(D2)$, $(D3)$ d'une distance pour D.

PROPOSITION 6. *Supposons X_2 **complet**. Alors, $(\mathcal{C}_b(X_1, X_2), D)$ est un espace métrique **complet**.*

DÉMONSTRATION : Soit $(f_n)_{n \geq 0}$ une suite de Cauchy dans $\mathcal{C}_b(X_1, X_2)$. Pour tout $x \in X_1$, on a :

$$(1) \qquad d_2(f_n(x), f_m(x)) \leq D(f_n, f_m) ,$$

donc $(f_n(x))_{n \geq 0}$ est une suite de Cauchy dans X_2, et elle converge vers une limite qu'on note $f(x)$.

Soit $\varepsilon > 0$. Il existe un entier n_0 tel qu'on ait $D(f_n, f_m) < \dfrac{\varepsilon}{3}$ pour $m, n \geq n_0$. En passant à la limite dans l'inégalité (1), on a donc, pour $x \in X_1$, $n \geq n_0$:

$$(2) \qquad d_2(f(x), f_n(x)) \leq \frac{\varepsilon}{3} .$$

Soit x_0 un point de X_1. Pour tout $x \in X_1$, on a :

$$d_2(f(x_0), f(x)) \leq d_2(f(x_0), f_{n_0}(x_0) + d_2(f_{n_0}(x_0), f_{n_0}(x)) \\ + d_2(f_{n_0}(x), f(x)) ,$$

d'où, d'après la relation (2) :

$$d_2(f(x_0), f(x)) \leq \frac{2\varepsilon}{3} + d_2(f_{n_0}(x_0), f_{n_0}(x)) \ . \tag{3}$$

L'application f_{n_0} est bornée; on déduit alors de la relation (3) que f est aussi bornée.

L'application f_{n_0} est continue; il existe donc $\delta > 0$ tel qu'on ait $d_2(f(x_0), f(x)) < \dfrac{\varepsilon}{3}$ pour $x \in B(x_0, \delta)$; d'après (3) on a :

$$d_2(f(x_0), f(x)) < \varepsilon \ ,$$

pour $x \in B(x_0, \delta)$. Ceci montre que f est continue en x_0.

On a montré que l'application f est un élément de $\mathcal{C}_b(X_1, X_2)$. La relation (2) implique que la suite $(f_n)_{n \geq 0}$ converge vers f. □

§. 4. Applications linéaires continues.

Soient $(E_1, \| \ \|_1)$ et $(E_2, \| \ \|_2)$ des espaces vectoriels normés.

PROPOSITION 7. *Soit* $u \in \mathcal{L}(E_1, E_2)$ *une application linéaire. Les propriétés suivantes sont équivalentes :*

(i) *u est lipschitzienne;*

(ii) *u est uniformément continue;*

(iii) *u est continue;*

(iv) *u est continue en 0 ;*

(v) *l'image par u d'une partie bornée est bornée;*

(vi *l'image par u de la boule unité $\overline{B}(0,1)$ est bornée;*

(vii) *l'image par u de la sphère unité est bornée.*

DÉMONSTRATION :

$$\begin{array}{lcl} (i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) & : & \text{déjà vu} \\ (v) \Rightarrow (vi) \Rightarrow (vii) & : & \text{évident} \ . \end{array}$$

(iv)$\Rightarrow$(v). Par définition de la continuité en 0, il existe $r > 0$ tel que $\|u(x)\|_2 < 1$ pour $\|x\|_1 < r$.

Si A est une partie bornée de E_1, il existe $R > 0$ tel que $A \subset B(0, R)$ et on a $u(A) \subset B\left(0, \dfrac{R}{r}\right)$.

(vii)$\Rightarrow$(i). Soit $C > 0$ tel que $\|u(x)\|_2 \leq C$ pour $\|x\|_1 = 1$.

Pour $x, y \in E_1$, distincts, on a :

$$\begin{aligned}\|u(x) - u(y)\|_2 &= \|u(x-y)\|_2 \\ &= \|x-y\|_1 \left\| u\left(\frac{x-y}{\|x-y\|_1}\right) \right\|_2 \\ &\leq C\,\|x-y\|_1 \ ,\end{aligned}$$

donc u est C-lipschitzienne. □

Les applications linéaires continues de E_1 dans E_2 forment un sous espace vectoriel de $\mathcal{L}(E_1, E_2)$ qu'on note $\mathcal{LC}(E_1, E_2)$. Pour $u \in \mathcal{LC}(E_1, E_2)$, la borne supérieure :

$$\|u\| = \sup_{\|x\|_1 = 1} \|u(x)\|_2$$

est finie (d'après (vii)) ; on a clairement :

$$\|u\| = \sup_{\bar{B}(0,1)} \|u(x)\|_2 = \sup_{x \neq 0} \frac{\|u(x)\|_2}{\|x\|_1} \ .$$

L'application u est $\|u\|$-lipschitzienne.

On vérifie facilement (exercice) que l'application $u \to \|u\|$ fait de $\mathcal{LC}(E, F)$ un espace vectoriel normé.

PROPOSITION 8. *Soient $(E_3, \| \ \|_3)$ un troisième espace vectoriel normé et $v : E_1 \times E_2 \to E_3$ une application* ***bilinéaire****. Les propriétés suivantes sont équivalentes :*

(i) *v est continue;*

(ii) *v est continue en 0 ;*

(iii) *il existe $C \geq 0$ tel qu'on ait :*

$$\|v(x_1, x_2)\|_3 \leq C\,\|x_1\|_1 \ \|x_2\|_2 \ ,$$

pour tous $x_1 \in E_1$, $x_2 \in E_2$.

DÉMONSTRATION : Exercice (cf. démonstration de la proposition 7).

Les applications bilinéaires continues de $E_1 \times E_2$ dans E_3 forment un espace vectoriel qu'on munit de la norme :

$$\begin{aligned}\|v\| &= \sup_{\|x_1\|_1 = 1,\, \|x_2\|_2 = 1} \|v(x_1, x_2)\|_3 \\ &= \sup_{x_1 \neq 0,\, x_2 \neq 0} \frac{\|v(x_1, x_2)\|_3}{\|x_1\|_1 \ \|x_2\|_2} \ .\end{aligned}$$

On aurait un résultat analogue pour des applications multilinéaires $E_1 \times E_2 \times \ldots \times E_k \to E_{k+1}$.

PROPOSITION 9. *Si E_1 est de* ***dimension finie****, toute application linéaire $u \in \mathcal{L}(E_1, E_2)$ est* ***continue****.*

REMARQUE. La propriété pour une application linéaire $u : E_1 \to E_2$ d'être continue ne change pas si on remplace les normes de E_1 ou E_2 par des normes équivalentes.

DÉMONSTRATION DE LA PROPOSITION 9 : D'après la remarque, on peut supposer que $E_1 = \mathbf{R}^n$, muni de la norme $\| \ \|_\infty$ (toutes les normes sont équivalentes sur $\mathbf{R}^n$). Soit $(e_i)_{1 \leq i \leq n}$ la base canonique. Pour $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$, on a :

$$\begin{aligned}
\|u(x)\|_2 &= \left\| u\left(\sum_1^n x_i e_i\right) \right\|_2 \\
&= \left\| \sum_1^n x_i u(e_i) \right\|_2 \\
&\leq \sum_1^n |x_i| \, \|u(e_i)\|_2 \\
&\leq \left(\sum_1^n \|u(e_i)\|_2 \right) \|x\|_\infty \, .
\end{aligned}$$

□

§. 5. Espaces de Banach.

DÉFINITION 6. Un **espace de Banach** est un espace vectoriel normé complet.

EXEMPLES.

1. Tout espace vectoriel normé **de dimension finie** est un espace de Banach.

2. Soit (X, d) un espace métrique. Notons $\mathcal{C}_b(X)$ l'espace vectoriel des fonctions (à valeurs réelles) continues et bornées sur X.

Pour $f \in \mathcal{C}_b(X)$, posons :

$$\|f\| = \sup_{x \in X} |f(x)| \, .$$

Ceci définit une norme sur $\mathcal{C}_b(X)$; la distance associée n'est autre que la **distance uniforme** introduite au §.3. Comme $\mathbf{R}$ est complet, $\mathcal{C}_b(X)$ est un espace de Banach d'après la proposition 6.

3. Lorsque X est compact, toute fonction continue sur X est bornée. L'espace vectoriel des fonctions continues sur X, noté $\mathcal{C}(X)$, est muni de la norme ci-dessus; c'est un espace de Banach.

4. Plus généralement, soient $(E, \| \ \|)$ un espace vectoriel normé, et (X, d) un espace métrique. L'espace vectoriel $\mathcal{C}_b(X, E)$ des applications continues et bornées de X dans E est muni de la norme :

$$|||f||| = \sup_{x \in X} \|f(x)\| \ ,$$

pour $f \in \mathcal{C}_b(X, E)$. **Si E est un espace de Banach** alors $\mathcal{C}_b(X, E)$ est aussi un espace de Banach.

PROPOSITION 10. *Soient $(E_1, \| \ \|_1)$, $(E_2 \| \ \|_2)$ deux espaces vectoriels normés. Si E_2 est un espace de Banach, alors $\mathcal{LC}(E_1, E_2)$ est un espace de Banach.*

DÉMONSTRATION : Soit $(u_n)_{n \geq 0}$ une suite de Cauchy dans $\mathcal{LC}(E_1, E_2)$. Notons pour abréger B_R la boule fermée de centre 0 de rayon R dans E_1.

D'après la proposition 7, pour tout $R > 0$, la suite des restrictions $(u_n/B_R)_{n \geq 0}$ à la boule B_R est une suite de Cauchy dans $\mathcal{C}_b(B_R, E_2)$. Comme E_2 est complet, $\mathcal{C}_b(B_R, E_2)$ l'est aussi (proposition 6), donc $(u_n)_{n \geq 0}$ converge uniformément sur B_R vers une application $u_R \in \mathcal{C}_b(B_R, E_2)$. Pour $R' > R$, $u_{R'}$ coïncide avec u_R sur B_R. L'application u qui coïncice avec u_R sur B_R (pour tout $R > 0$) est linéaire (exercice); elle est bornée sur B_1, donc continue (proposition 7). Comme $(u_n)_{n \geq 0}$ converge uniformément vers u sur B_1, la suite $(u_n)_{n \geq 0}$ converge aussi vers u dans $\mathcal{LC}(E_1, E_2)$ (par définition de la norme).

§. 6. Le théorème d'Ascoli.

Soient (X, d) un espace métrique **compact**, E un espace vectoriel normé **de dimension finie**, et A une partie de l'espace de Banach $\mathcal{C}(X, E)$.

DÉFINITION 7. La partie A est **équicontinue** si pour tout $\varepsilon > 0$, tout $x_0 \in X$, il existe $\delta > 0$ tel qu'on ait :

$$\forall x \in B(x_0, \delta), \quad \forall f \in A, \quad \|f(x) - f(x_0)\| < \varepsilon \ .$$

(En d'autres termes, "le même δ sert pour **toutes** les applications $f \in A$").

THÉORÈME 2 (ASCOLI). *Les propriétés suivantes sont équivalentes :*

(i) *l'adhérence de A dans $\mathcal{C}(X, E)$ est compacte;*

(ii) *A est équicontinue, et pour tout $x \in X$, l'ensemble $A(x) = \{f(x), f \in A\}$ est une partie bornée de E.*

DÉMONSTRATION :

(i)⇒(ii). Soit $x \in X$. L'application $f \mapsto f(x)$ de $\mathcal{C}(X,E)$ dans E est continue (exercice). Comme $\bar{A}$ est compact, son image par cette application l'est aussi, en particulier elle est bornée. Donc $A(x)$ est borné.

Soit $\varepsilon > 0$. D'après la proposition 6 du §. 3, Chapitre II, il existe un nombre fini d'applications $f_1, \ldots, f_k \in \bar{A}$ telles que :

$$A \subset \bigcup_{i=1}^{k} B\left(f_i, \frac{\varepsilon}{3}\right) .$$

Il existe un nombre $\delta > 0$ tel qu'on ait, pour $y \in B(x,\delta)$ et $1 \leq i \leq k$:

$$\|f_i(x) - f_i(y)\| < \frac{\varepsilon}{3} .$$

Soit $f \in A$; il existe un entier j tel que $f \in B\left(f_j, \frac{\varepsilon}{3}\right)$, et on a alors, pour $y \in B(x,\delta)$

$$\begin{aligned} \|f(x) - f(y)\| &\leq \|f(x) - f_j(x)\| + \|f_j(x) - f_j(y)\| \\ &\quad + \|f_j(y) - f(y)\| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon . \end{aligned}$$

Donc A est équicontinue.

(ii)⇒(i). La démonstration comporte plusieurs étapes.

1) Lemme. *A est **uniformément équicontinue**, c'est à dire que pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que :*

$$d(x,y) < \delta \Rightarrow \|f(x) - f(y)\| < \varepsilon ,$$

pour tout $f \in A$.

DÉMONSTRATION DU LEMME : Pour tout $x \in X$, il existe $\delta(x) > 0$ tel que :

$$d(x,y) < \delta(x) \Rightarrow \|f(x) - f(y)\| < \frac{\varepsilon}{2} , \quad \forall f \in A .$$

D'après le théorème de Lebesgue (Chapitre II, §.3, théorème 2), appliqué au recouvrement de X par les boules $B(x,\delta(x))$, il existe $\delta > 0$ tel que deux points $y, y' \in X$ vérifiant $d(y,y') < \delta$ dont contenus dans une même boule $B(x,\delta(x))$.

On a alors

$$\begin{aligned} \|f(y) - f(y')\| &\leq \|f(y) - f(x)\| + \|f(x) - f(y')\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon . \end{aligned}$$

□

2). Soient $\varepsilon > 0$, et $(f_n)_{n\geq 0}$ une suite d'éléments de A. Montrons qu'il existe une suite $(g_k)_{k\geq 0}$ extraite de $(f_n)_{n\geq 0}$ vérifiant

$$\|g_k - g_\ell\| < \varepsilon \ , \quad \forall k, \ell \geq 0 \ .$$

Soit $\delta > 0$ tel que :

$$\forall x, y \in X, \quad \forall f \in A, \quad d(x,y) < \delta \Rightarrow \|f(x) - f(y)\| \leq \frac{\varepsilon}{3} \ .$$

Soient $x_1, \ldots, x_m$ des points de X tels que :

$$X = \bigcup_{i=1}^{m} B(x_i, \delta)$$

(cf. Chapitre II, § 3, proposition 6).

Par hypothèse, la suite $(f_n(x_1), \ldots, f_n(x_m))_{n\geq 0}$ dans E^m est bornée ; comme E^m est de dimension finie, elle a un point d'accumulation. On peut donc extraire de $(f_n)_{n\geq 0}$ une suite $(g_k)_{k\geq 0}$ telle qu'on ait, pour $k \geq 0$, $\ell \geq 0$, $1 \leq i \leq m$:

$$\|g_k(x_i) - g_\ell(x_i)\| \leq \frac{\varepsilon}{3} \ .$$

Soient $k \geq 0$, $\ell \geq 0$, $x \in X$; il existe un entier i tel que $x \in B(x_i, \delta)$ et on a :

$$\begin{aligned} \|g_k(x) - g_\ell(x)\| &\leq \|g_k(x_i) - g_k(x)\| + \|g_k(x_i) - g_\ell(x_i)\| \\ &\quad + \|g_\ell(x_i) - g_\ell(x)\| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \ . \end{aligned}$$

3) Processus diagonal. (Cf. Chapitre II, théorème de Bolzano-Weierstrass).

Montrons que de toute suite (f_n) d'éléments de A, on peut extraire une suite de Cauchy.

D'après la partie 2), on peut extraire une suite $(f_n^1)_{n\geq 0}$ satisfaisant :

$$\|f_k^1 - f_\ell^1\| < \frac{1}{2}, \forall k, \ell \geq 0 \ ,$$

par récurrence, on construit pour $j \geq 2$ une suite $(f_n^j)_{n\geq 0}$ extraite de $(f_n^{j-1})_{n\geq 0}$ satisfaisant :

$$\left\|f_k^j - f_\ell^j\right\| < \frac{1}{2^j} \ , \ \forall k, \ell \geq 0 \ .$$

Posons $g_n = f_n^n$ pour $n \geq 0$. Pour $m \geq n > 0$, on a

$$\|g_n - g_m\| < \frac{1}{2^n} \ ,$$

donc $(g_n)_{n\geq 0}$ est une suite de Cauchy, extraite de la suite $(f_n)_{n\geq 0}$.

4) Montrons que l'adhérence de A est compacte.

Soit $(f'_n)_{n\geq 0}$ une suite d'éléments de $\overline{A}$. Pour $n \geq 0$, soit f_n un élément de A tel que

$$(*) \qquad \|f_n - f'_n\| < 2^{-n} .$$

Soit $(f_{n_k})_{k\geq 0}$ une suite de Cauchy extraite de $(f_n)_{n\geq 0}$. Elle converge dans $\mathcal{C}_b(X, E)$ (proposition 6); comme $\overline{A}$ est fermé, sa limite appartient à $\overline{A}$. D'après $(*)$, la suite $(f'_{n_k})_{k\geq 0}$, extraite de $(f'_n)_{n\geq 0}$, a la même limite. □

2ème PARTIE.
CALCUL DIFFERENTIEL.

CHAPITRE V : La différentielle.

Dans ce chapitre, on désigne par E et F deux espaces vectoriels de **dimension finie**, de dimensions respectives m et n.

On notera en général $(e_1, \dots, e_m)$ une base de E, et $x_1, \dots, x_m$ les coordonnées d'un vecteur x de E dans cette base. (On a alors une identification de E à $\mathbf{R}^m$). De même, on notera $(f_1, \dots, f_n)$ une base de F et $y_1, \dots, y_n$ les coordonnées d'un vecteur y de F.

On munit E et F de normes, toutes deux notées $\| \ \|$; on rappelle que toutes les normes sur E (ou sur F) sont équivalentes. Pour cette raison, tout ce qui suit ne **dépend pas** des normes choisies sur E et F.

NOTATION. Soit (X, d) un espace métrique, E un espace vectoriel normé, x_0 un point de X, $\Phi : X \to E$ et $\varphi : X \to \mathbf{R}_+$ des applications. La propriété :

$$\text{“ } \exists r > 0,\ \exists c > 0 \text{ tq } \|\Phi(x)\| \leq C\varphi(x) \text{ pour } x \in B(x_0, r) \text{ ”}$$

se note $\Phi = O_{x_0}(\varphi)$ ou $\Phi = O(\varphi)$ lorsque la valeur de x_0 est claire d'après le contexte. La propriété :

$$\text{“ } \forall \varepsilon > 0,\ \exists r > 0 \text{ tq } \|\Phi(x)\| \leq \varepsilon\varphi(x) \text{ pour } x \in B(x_0, r) \text{ ”}$$

se note $\Phi = o_{x_0}(\varphi)$ ou $\phi = o(\varphi)$.

REMARQUE. Tout ce qui suit (à l'exception de ce qui utilise bases et coordonnées) se généralise au cas où E et F sont des espaces de Banach. La seule modification requise est d'exiger que toutes les applications linéaires rencontrées soient **continues**.

§. 1. La différentielle.

Soient U une partie **ouverte** de E, a un point de U, et $H : U \to F$ une application.

DÉFINITION 1. L'application H est **différentiable** au point a s'il existe une application **linéaire** $L : E \to F$ telle qu'on ait :

$$(*) \qquad H(x) - H(a) - L(x - a) = o_a(\|x - a\|) .$$

En d'autres termes, on a :

$$(*') \qquad \lim_{\substack{x \to a \\ x \neq a}} \frac{H(x) - H(a) - L(x - a)}{\|x - a\|} = 0 .$$

PROPOSITION 1. *Si H est différentiable en a, il existe une **unique** application linéaire L vérifiant la relation* $(*)$.

DÉMONSTRATION : Soient L_1, L_2 des applications linéaires vérifiant $(*)$ et u un vecteur non nul de E. En utilisant la relation $(*')$ pour $x = a+\varepsilon u$, on obtient :

$$\lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} \frac{L_1(\varepsilon u) - L_2(\varepsilon u)}{\|\varepsilon u\|} = 0 \ ;$$

or on a, pour $\varepsilon > 0$:

$$\frac{L_1(\varepsilon u) - L_2(\varepsilon u)}{\|\varepsilon u\|} = \|u\|^{-1} \left(L_1(u) - L_2(u)\right) \ ,$$

donc $L_1(u) = L_2(u)$ et $L_1 = L_2$.

DÉFINITION 2. Lorsque H est différentiable en a, l'unique application linéaire vérifiant $(*)$ s'appelle la **différentielle** de H en a. On la note $DH(a)$ ou D_aH.

PROPOSITION 2. *Si H est différentiable en a, alors H est continue en a.*

DÉMONSTRATION : Soit $r > 0$ tel que $B(a,r) \subset U$ et qu'on ait :

$$\|H(x) - H(a) - L(x-a)\| \leq \|x - a\|$$

pour $x \in B(a,r)$. D'après le chapitre IV, propositions 7 et 9, il existe $C > 0$ tel qu'on ait :

$$\|L(x)\| \leq C\,\|x\| \ , \quad \forall x \in E \ .$$

On a donc, pour $x \in B(a,r)$:

$$\|H(x) - H(a)\| \leq (C+1)\,\|x - a\| \ ,$$

ce qui implique que H est continue en a.

PROPOSITION 3. *Identifions F à $\mathbf{R}^n$ au moyen d'une base $(f_1,\dots,f_n)$ de F, et notons $H_i : U \to \mathbf{R}$ pour $i = 1,\dots,n$ les applications coordonnées de H. Pour que H soit différentiable en a il faut et il suffit que chaque H_i le soit, et on a alors, pour $v \in E$:*

$$D_aH(v) = \sum_{i=1}^{n} D_a H_i(v) f_i \ .$$

DÉMONSTRATION : Exercice.

§. 2. Différentielles et dérivées partielles.

Soient U une partie ouverte de E, a un point de U, $H : U \to F$ une application et E_1 un sous espace vectoriel de E.

Définissons :

$$U_1 = \left\{ v \in E_1 \ , \ a + v \in U \right\} .$$

Comme l'application $v \mapsto a + v$ de E_1 dans E est continue, U_1 est une partie ouverte de E_1.

DÉFINITION 3. On dit que H est **différentiable en a dans la direction de E₁** si l'application $v \mapsto H(a+v)$ de U_1 dans F est différentiable en 0. La différentielle en 0 de cette application est alors une application linéaire de E_1 dans F qu'on appelle **différentielle partielle** de H dans la direction de E_1.

EXEMPLE. Soit w un vecteur non nul de E. Prenons pour E_1 le sous-espace vectoriel (de dimension 1) engendré par w, qu'on identifie à $\mathbf{R}$ par l'isomorphisme : $t \mapsto tw$.

L'application H est différentiable en a dans la direction de E_1 (on dit encore, dans ce cas, **dérivable** dans la direction de w) si et seulement si la limite

$$\lim_{\substack{t \to 0 \\ t \neq 0}} \frac{H(a+tw) - H(a)}{t}$$

existe. On l'appelle alors la **dérivée** de H dans la direction de w. En notant v cette limite, la différentielle partielle de H dans la direction de E_1 est l'application linéaire $tw \mapsto tv$ $(t \in \mathbf{R})$ de E_1 dans F.

DÉFINITION 4. Soit $(e_1, \ldots, e_m)$ une base de E. Les dérivées de H dans la direction de $e_1, \ldots, e_m$ sont, quand elles existent, appelées **dérivées partielles** de l'application H (relatives à la base $(e_1, \ldots, e_m)$). La dérivée partielle dans la direction de e_i est notée $\partial_i H(a)$ ou $\dfrac{\partial}{\partial x_i} H(a)$.

PROPOSITION 4. *Si H est différentiable en a, alors H est différentiable en a dans la direction de E_1 et sa différentielle partielle dans la direction de E_1 est la restriction à E_1 de la différentielle de H en a.*

DÉMONSTRATION : Exercice.

COROLLAIRE. *Si H est différentiable en a, alors les dérivées partielles $\partial_i H(a)$, pour $1 \leq i \leq n$, sont définies et on a :*

$$\partial_i H(a) = D_a H(e_i) \ .$$

REMARQUE. Il est possible, lorsque $m \geq 2$, que H soit dérivable en a dans la direction de tout vecteur non nul $w \in E$ sans que H soit différentiable en a. Prenons par exemple $U = E = \mathbf{R}^2$, $a = (0,0)$, $F = \mathbf{R}$,

$$H(x,y) = \frac{x^2 y}{x^2 + y^2} \quad \text{si} \quad (x,y) \neq (0,0)$$
$$H(0,0) = 0 \ .$$

Soit $w \in \mathbf{R}^2$. Pour tout réel t, on a

$$H(tw) = tH(w)$$

donc H est dérivable en 0 dans la direction de w, de dérivée $H(w)$. $\big[$ En particulier, on a $\partial_1 H(0,0) = \partial_2 H(0,0) = 0\big]$. D'après la proposition 4, H n'est pas différentiable en $(0,0)$: sa différentielle en 0 serait l'application H, qui n'est pas linéaire.

DÉFINITION 5. Soient $(e_1, \ldots, e_m)$, $(f_1, \ldots, f_n)$ des bases de E, F respectivement. Notons $H_1, \ldots, H_n$ les applications coordonnées de H, et supposons que H est différentiable en a. La matrice

$$\Big(\partial_j H_i(a)\Big)_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq m,}}$$

a n lignes et m colonnes s'appelle **matrice jacobienne** de H en a. C'est la matrice de la différentielle de H en a relative aux bases de E et F considérées.

Lorsque $\mathbf{m} = \mathbf{n}$, son déterminant s'appelle jacobien de H en a.

§. 3. Propriétés élémentaires. Règles de calcul.

3.1. Applications affines.

Soit $u : E \to F$ une application linéaire, b un point de F, et $H : E \to F$ l'application affine définie par :

$$H(x) = u(x) + b \ .$$

En tout point a de E, l'application H est différentiable et sa différentielle est u.

3.2. Somme.

Soient H_1, H_2 des applications d'une partie ouverte U de E dans F.

Si H_1 et H_2 sont différentiables en un point a de U, alors $H_1 + H_2$ l'est aussi et on a :

$$D_a(H_1 + H_2) = D_a H_1 + D_a H_2 \ .$$

En termes de dérivées partielles relatives à une base de E, on a donc, pour $1 \leq i \leq m$:

$$\partial_i(H_1 + H_2) = \partial_i H_1 + \partial_i H_2 \ .$$

3.3. Composition.

Soit G un troisième espace vectoriel (normé) de dimension finie. Soient U une partie ouverte de E, H une application de U dans F, V une partie ouverte de F, K une application de V dans G.

PROPOSITION 5. *Soit a un point de U. On suppose que H est différentiable en a, que $H(a) \in V$, et que K est différentiable en $H(a)$. Alors, l'application $K \circ H$ est définie sur une boule ouverte centrée en a, est différentiable au point a, et sa différentielle en a est :*

$$D_a(K \circ H) = D_{H(a)}K \circ D_a H \ .$$

DÉMONSTRATION : L'application H est continue en a (proposition 2), donc il existe $r > 0$ tel que $H(B(a,r)) \subset V$. L'application $K \circ H$ est donc définie sur $B(a,r)$.

Posons $L_1 = D_a H$, $L_2 = D_{H(a)}K$, $C_1 = \|L_1\|$, $C_2 = \|L_2\|$ (où les normes dans les espaces vectoriels $L(E,F)$, $L(F,G)$ ont été définies au Chapitre IV, §. 4).

On a vu dans la démonstration de la proposition qu'il existe $r_0 \in]0,r[$ tel qu'on ait :

$$\|H(x) - H(a)\| \leq (1 + C_1)\,\|x - a\|$$

pour tout $x \in B(a, r_0)$.

Soit $\varepsilon > 0$; il existe $r_1 \in]0, r_0[$ tel qu'on ait, pour $\|x - a\| < r_1$, $\|y - H(a)\| < r_1$

$$\|H(x) - H(a) - L_1(x-a)\| \leq \frac{\varepsilon}{1 + C_1 + C_2}\,\|x - a\| \ ,$$

$$\|K(y) - K(H(a)) - L_2(y - H(a))\| \leq \frac{\varepsilon}{1 + C_1 + C_2}\,\|y - H(a)\| \ .$$

Soit $x \in B\left(a, \dfrac{r_1}{1 + C_1}\right)$. Posons $y = H(x)$; on a alors :

$$\|y - H(a)\| \leq (1 + C_1)\,\|x - a\| < r_1 \ .$$

Posons :

$$R_1 = H(x) - H(a) - L_1(x - a) \ ,$$
$$R_2 = K(y) - K(H(a)) - L_2(y - H(a)) \ .$$

On a :

$$\begin{aligned} K \circ H(x) &= K(y) \\ &= K \circ H(a) + L_2(y - H(a)) + R_2 \ , \end{aligned}$$

$$\begin{aligned} L_2(y - H(a)) &= L_2(L_1(x-a) + R_1) \\ &= L_2 \circ L_1(x - a) + L_2(R_1) \ , \end{aligned}$$

donc

$$K \circ H(x) = K \circ H(a) + L_2 \circ L_1(x - a) + R_2 + L_2(R_1) \ ,$$

avec

$$\|R_2\| \leq \frac{\varepsilon}{1+C_1+C_2} \|y - H(a)\| \leq \frac{\varepsilon(1+C_1)}{1+C_1+C_2} \|x-a\|$$

$$\|L_2(R_1)\| \leq C_2 \|R_1\| \leq \frac{\varepsilon C_2}{1+C_1+C_2} \|x-a\| \ ,$$

d'où finalement :

$$\|K \circ H(x) - K \circ H(a) - L_2 \circ L_1(x-a)\| \leq \varepsilon \|x-a\| \ ,$$

ce qui termine de prouver la proposition.

Soient $(e_1, \ldots, e_m)$ une base de E, $(f_1, \ldots, f_n)$ une base de F, $H_1, \ldots, H_n$ les applications coordonnées de H relatives à cette base.

En exprimant que la matrice jacobienne de $K \circ H$ en a est le produit de la matrice jacobienne de K en $H(a)$ et de la matrice jacobienne de H en a, on obtient, pour $1 \leq i \leq m$

$$\partial_i(K \circ H)(a) = \sum_{\ell=1}^{n} \partial_i H_\ell(a) \partial_\ell K(a) \ .$$

3.4. Produits.

Supposons que E soit le produit de deux espaces vectoriels E_1, E_2 et que H soit une application **bilinéaire** de $E_1 \times E_2$ dans F.

PROPOSITION 6. *L'application H est alors différentiable en tout point $a = (a_1, a_2)$ de E, et sa différentielle en a vérifie :*

$$D_a H(v_1, v_2) = H(v_1, a_2) + H(a_1, v_2) \ ,$$

pour tout $v = (v_1, v_2) \in E$.

DÉMONSTRATION : L'application H est continue (exercice) et il existe une constante $C > 0$ telle que :

$$\forall x_1 \in E_1, \quad \forall x_2 \in E_2, \qquad \|H(x_1, x_2)\| \leq C \|x_1\| \|x_2\| \ ,$$

(Chapitre IV, Proposition 8).

D'autre part, comme toutes les normes sur E sont équivalentes, il existe $C' > 0$ tel qu'on ait :

$$\forall x = (x_1, x_2) \in E, \quad \mathrm{Max}\,(\|x_1\|, \|x_2\|) \leq C' \|x\| \ .$$

On a alors, pour $a = (a_1, a_2)$, $v = (v_1, v_2)$ dans E :

$$\begin{aligned} H(a+v) &= H(a_1, a_2) + H(a_1, v_2) + H(v_1, a_2) + H(v_1, v_2) \\ &= H(a_1, a_2) + H(a_1, v_2) + H(v_1, a_2) + O(\|v_1\| \|v_2\|) \\ &= H(a_1, a_2) + H(a_1, v_2) + H(v_1, a_2) + O(\|v\|^2) \ , \end{aligned}$$

d'où la proposition.

En combinant les propositions 5 et 6, on obtient les règles de différentiation pour différentes sortes de "produits" :

1) Prenons $F = \mathbf{R}$; si H_1, H_2 sont deux applications à valeurs réelles, définies sur une partie ouverte U de E, et différentiables en un point a de U, alors le produit $H_1 H_2$ est différentiable en a et on a, pour $v \in E$:

$$D_a(H_1 H_2)(v) = H_1(a) D_a H_2(v) + H_2(a) D_a H_1(v) \ .$$

2) Soient U une partie ouverte de E, $H_1 : U \to F$ et $H_2 : U \to \mathbf{R}$ des applications, différentiables en un point $a \in U$. Alors $H_1 H_2$ est différentiable en a et :

$$D_a(H_2 H_1)(v) = D_a H_2(v) H_1(a) + H_2(a) D_a H_1(v) \ .$$

3) Si F est muni d'un produit scalaire $\langle \, , \rangle$, U est une partie ouverte de E, H_1, H_2 sont des applications de U dans F, différentiables en un point a de U, alors $\langle H_1, H_2 \rangle$ est différentiable en a et on a :

$$D_a \langle H_1, H_2 \rangle (v) = \langle D_a H_1(v), H_2(a) \rangle + \langle H_1(a), D_a H_2(v) \rangle \ .$$

4) Si U est une partie ouverte de E, H_1, H_2 des applications de U dans $\mathbf{R}^3$, différentiables en un point $a \in U$, le produit extérieur $H_1 \wedge H_2$ est différentiable en a, avec :

$$D_a(H_1 \wedge H_2)(v) = D_a H_1(v) \wedge H_2(a) + H_1(a) \wedge D_a H_2(v) \ .$$

Par rapport à une base $(e_1, \ldots, e_m)$ de E, on a des règles analogues pour les dérivées partielles : par exemple, dans le premier cas ci-dessus :

$$\partial_i(H_1 H_2)(a) = H_1(a) \partial_i H_2(a) + \partial_i H_1(a) H_2(a) \ ,$$

pour $1 \leq i \leq m$.

Remarque. De façon plus générale, si E est le produit de k sous espaces $E_1, \ldots, E_k$ et $H : E \to F$ est une application $k-$linéaire, l'application H est différentiable en tout point $a = (a_1, \ldots, a_k)$ de E et on a, pour $v = (v_1 \ldots v_k) \in E$:

$$\begin{aligned} D_a H(v) = {} & H(a_1, \ldots, a_{k-1}, v_k) + H(a_1, \ldots, v_{k-1}, a_k) \\ & + \ldots + H(a_1, v_2, \ldots, a_{k-1}, a_k) + H(v_1, \ldots, a_k) \ . \end{aligned}$$

On déduit par exemple de ceci et de la proposition 5 la formule suivante : soient U une partie ouverte de E et H_1, H_2, H_3 des applications de U dans $\mathbf{R}^3$ différentiables en un point $a \in U$. Alors l'application $K = \det(H_1, H_2, H_3)$ de U dans $\mathbf{R}$ est différentiable en a, et sa différentielle est donnée par :

$$\begin{aligned} D_a K(v) = {} & \det(D_a H_1(v), H_2(a), H_3(a)) + \det(H_1(a), D_a H_2(v), H_3(a)) \\ & + \det(H_1(a), H_2(a), D_a H_3(v)) \ . \end{aligned}$$

3.5. Inverses.

Soit G un espace vectoriel de dimension finie. Prenons $E = F = \mathcal{L}(G)$, $U = GL(G) = \{u \in \mathcal{L}(G), \det u \neq 0\}$. C'est une partie ouverte de E.

PROPOSITION 7. *L'application $H : u \mapsto u^{-1}$ de U dans $\mathcal{L}(G)$ est différentiable en tout point $a \in U$, et on a :*

$$D_a H(v) = -a^{-1} \circ v \circ a^{-1} \ .$$

DÉMONSTRATION : Soient $a, x \in U$; posons $v = x - a$. Il s'agit d'estimer la norme de l'élément :

$$w = x^{-1} - a^{-1} + a^{-1} \circ v \circ a^{-1} \ .$$

On a :

$$\begin{aligned} x \circ w \circ a &= a - x + x \circ a^{-1} \circ v \\ &= -v + (a+v) \circ a^{-1} \circ v \\ &= v \circ a^{-1} \circ v \ . \end{aligned}$$

On a donc $w = x^{-1} \circ v \circ a^{-1} \circ v \circ a^{-1}$, d'où :

$$\|w\| \leq \|a^{-1}\|^2 \, \|x^{-1}\| \, \|v\|^2 \ .$$

Or on a, d'après la définition de w :

$$\|x^{-1}\| \leq \|a^{-1}\| + \|w\| + \|a^{-1}\|^2 \ \|v\| \ .$$

On en déduit :

$$\|w\| \left(1 - \|a^{-1}\|^2 \|v\|^2\right) \leq \|a^{-1}\|^3 \|v\|^2 \left(1 + \|a^{-1}\| \, \|v\|\right) \ ,$$

d'où, pour $\|v\|$ assez petit :

$$\|w\| \leq 2 \, \|a^{-1}\|^3 \ \|v\|^2 \ .$$

La proposition en résulte.

Lorsque $G = \mathbf{R}$, en combinant les propositions 5 et 7, on obtient :

COROLLAIRE. *Soient U une partie ouverte d'un espace vectoriel de dimension finie et $H : U \to \mathbf{R}^*$ une application. Si H est différentiable en un point $a \in U$, l'application $\dfrac{1}{H}$ l'est aussi et on a, pour $v \in E$:*

$$D_a \left(\frac{1}{H}\right)(v) = -\frac{D_a H(v)}{(H(a))^2} \ .$$

En termes de dérivées partielles (relatives à une base $e_1, \ldots, e_m$ de E) :

$$\partial_i \left(\frac{1}{H}\right)(a) = -\frac{\partial_i H(a)}{(H(a))^2} \ .$$

CHAPITRE VI : Applications de classe C^1.

§. 1. Applications de classe C^1.

Soient E, F deux espaces vectoriels (normés) de dimensions finie, U une partie ouverte de E, et $H : U \to F$ une application.

DÉFINITION 1. On dit que H est **de classe C^1** (ou encore que H est **continûment différentiable**) si les deux conditions suivantes sont réalisées :

(i) H est différentiable en tout point de U,
(ii) L'application $x \mapsto DH(x)$ de U dans $\mathcal{L}(E,F)$ est continue.

REMARQUES.

1) Identifions F à $\mathbb{R}^n$ au moyen d'une base $(f_1, \ldots, f_n)$ de F, et notons $H_1, \ldots, H_n$ les applications coordonnées de H. D'après la proposition 3 du Chapitre V, pour que H soit de classe C^1, il faut et il suffit que chacune des applications H_i, $1 \leq i \leq n$, le soit.

2) La condition (ii) ci-dessus est équivalente à la condition :

(ii′) L'application : $(x,v) \mapsto D_x H(v)$ de $U \times E$ dans F est continue.

DÉMONSTRATION :

(ii)$\Rightarrow$(ii'). En effet l'application $(x,v) \mapsto D_x H(v)$ est composée de l'application

$$(x,v) \mapsto (D_x H, v)$$

de $U \times E$ dans $\mathcal{L}(E,F) \times E$, qui est continue par hypothèse, et de l'application :

$$(L,v) \to L(v)$$

de $\mathcal{L}(E,F) \times E$ dans F, qui est bilinéaire donc continue.

(ii')$\Rightarrow$(ii). Soient $(e_1, \ldots, e_m)$ une base de E, et $(f_1, \ldots, f_n)$ une base de F. Par hypothèse, l'application : $x \mapsto D_x H(e_i)$ est continue, donc le $i^{\text{ième}}$ vecteur colonne de la matrice jacobienne de H dépend continûment du point $x \in U$.

L'application $x \mapsto D_x H$ est donc continue.

[**Note** : l'équivalence de (ii) et (ii′) n'a plus lieu lorsque E est un espace (de Banach) de dimension infinie. On a alors seulement : (ii) $\Rightarrow$ (ii′).]

Soit $(e_1, \ldots, e_m)$ une base de E.

PROPOSITION 1. *Pour que H soit de classe C^1, il faut et il suffit que les deux conditions suivantes soient réalisées :*

(iii) *pour tout $x \in U$, pour tout $i \in \{1, \ldots, m\}$, l'application H admet en x une dérivée partielle $\partial_i H(x)$ dans la direction de e_i,*

(iv) *pour tout $i \in \{1, \ldots, m\}$, l'application $x \mapsto \partial_i H(x)$ de U dans F est continue.*

DÉMONSTRATION : Si H est de classe C^1, la condition (iii) est réalisée (Chapitre V, proposition 4) et on a $\partial_i H(x) = D_x H(e_i)$, donc la condition (iv) est réalisée (remarque 2 ci-dessus).

Inversement, supposons les conditions (iii) et (iv) vérifiées par H. D'après la remarque 2 ci-dessus et la proposition 4 du Chapitre V, il suffit de démontrer que H est différentiable en tout point $a \in U$, ou encore que chaque application coordonnée de H (par rapport à une base de F) est différentiable en tout point $a \in U$. Or ces applications coordonnées satisfont aux conditions (iii) et (iv) s'il en est de même pour H. On peut donc supposer $F = \mathbf{R}$.

Soient $a \in U$, et $\varepsilon > 0$. Il existe $\delta > 0$ tel que pour tout vecteur $v = \sum_1^m v_i e_i$ vérifiant :

(1) $\operatorname{Max} |v_i| < \delta$

on ait :

(2) $a + v \in U$,

(3) $\forall 1 \leq i \leq m$, $|\partial_i H(a+v) - \partial_i H(a)| < \varepsilon$.

Soit $v \in E$ vérifiant (1). Posons $v(0) = 0$ et

$$v(k) = \sum_{i=1}^{k} v_i e_i \ , \quad 1 \leq k \leq m \ .$$

On a alors $v(m) = v$, d'où :

$$H(a+v) - H(a) = \sum_{k=1}^{m} \Big(H(a + v(k-1) + v_k e_k) - H(a + v(k-1)) \Big) \ .$$

Soit $1 \leq k \leq m$. D'après (2), l'application $\Delta_k H : t \mapsto H(a + v(k-1) + te_k)$ est définie sur $]-\delta, +\delta[$, et d'après (iii) et (iv), elle est continûment différentiable sur cet intervalle. On a donc :

$$\Delta_k H(v_k) - \Delta_k H(0) = v_k \int_0^1 \partial_k H\big(a + v(k-1) + s v_k e_k\big)\, ds \ .$$

D'après la relation (3), on a, pour $0 \leq s \leq 1$:

$$\big|\partial_k H\big(a + v(k-1) + s v_k e_k\big) - \partial_k H(a)\big| < \varepsilon \ ,$$

et on en déduit :

$$|\Delta_k H(v_k) - \Delta_k H(0) - v_k \partial_k H(a)| \leq \varepsilon\, |v_k| \ ,$$

$$\left| H(a+v) - H(a) - \sum_{k=1}^{m} v_k \partial_k H(a) \right| \leq \varepsilon \sum_{k=1}^{n} |v_k| \ .$$

L'application H est donc différentiable en a.

Reprenons le §. 3 du chapitre précédent. Des différentes formules qui s'y trouvent, il ressort immédiatement que :

1) Une somme (finie) d'applications de classe C^1 est de classe C^1.

2) Soient G un troisième espace vectoriel de dimension finie, U une partie ouverte de E, V une partie ouverte de F, $H : U \to F$ et $K : V \to G$ des applications de classe C^1. Alors la composée $K \circ H : U \cap H^{-1}(V) \to G$ est de classe C^1.

3) Une application affine est de classe C^1. Une application bilinéaire est de classe C^1 (et plus généralement, une application multilinéaire).

4) Si $H_1, H_2 : U \to \mathbf{R}$ sont de classe C^1, le produit l'est aussi. De même pour les autres types de produits considérés en §. 3.4, Chapitre V.

En particulier, toute application **polynômiale** $H : E \to \mathbf{R}$ est de classe C^1.

5) Si $H : U \to \mathbf{R}^*$ est de classe C^1, l'application $\dfrac{1}{H}$ l'est aussi.

§. 2. Le théorème de la moyenne.

THÉORÈME. *Soient $I = [\alpha, \beta]$ un intervalle compact de $\mathbf{R}$, F un espace vectoriel normé (de dimension finie), et $\Phi : I \to F$, $\varphi : I \to \mathbf{R}$ deux applications. On suppose que Φ et φ sont dérivables en tout point de I, et qu'on a, pour tout $x \in I$:*

$$\|\Phi'(x)\| \leq \varphi'(x) \ .$$

On a alors

$$\|\Phi(\beta) - \Phi(\alpha)\| \leq \varphi(\beta) - \varphi(\alpha) \ .$$

REMARQUE. La notation $\Phi'(x)$ désigne la dérivée de Φ en x, c'est-à-dire le vecteur $D_x\Phi(1)$ de F. De même, $\varphi'(x)$ est le nombre réel $D_x\varphi(1)$. Au point α, on a :

$$\Phi'(\alpha) = \lim_{\substack{t>0 \\ t\to 0}} t^{-1}(\Phi(\alpha+t) - \Phi(\alpha))$$

(l'existence de la limite fait partie de l'hypothèse), et des expressions analogues pour $\Phi'(\beta)$, $\varphi'(\alpha)$, $\varphi'(\beta)$.

DÉMONSTRATION : Soit $\varepsilon > 0$. Considérons l'ensemble A des nombres $\gamma \in [\alpha,\beta]$ vérifiant :

$$\forall t \in [\alpha,\gamma] \ , \quad \|\Phi(t) - \Phi(\alpha)\| \leq \varphi(t) - \varphi(\alpha) + \varepsilon(t-\alpha) \ .$$

Nous allons montrer que $A = [\alpha,\beta]$. Remarquons d'abord que :

1) $\alpha \in A$;

2) si $\gamma \in A$, on a $[\alpha,\gamma] \subset A$;

3) A est fermé, car Φ et φ sont continues.

De ces trois propriétés il résulte qu'il existe un nombre $\theta \in [\alpha,\beta]$ tel que $A = [\alpha,\theta]$. Supposons qu'on ait $\theta < \beta$. Comme $\theta \in A$, on a :

$$\|\phi(\theta) - \phi(\alpha)\| \leq \varphi(\theta) - \varphi(\alpha) + \varepsilon(\theta - \alpha) \ .$$

Comme φ et Φ sont dérivables en θ, il existe $\delta > 0$ tel qu'on ait $[\theta,\theta+\delta] \subset [\alpha,\beta]$ et, pour $0 \leq t \leq \delta$:

$$\|\Phi(\theta+t) - \Phi(\theta) - t\Phi'(\theta)\| \leq \frac{\varepsilon}{2}t \ ,$$

$$|\varphi(\theta+t) - \varphi(\theta) - t\varphi'(\theta)| \leq \frac{\varepsilon}{2}t \ .$$

Pour $t \in [0,\delta]$, on a donc :

$$\begin{aligned} \|\Phi(\theta+t) - \Phi(\theta)\| &\leq t\,\|\Phi'(\theta)\| + \frac{\varepsilon}{2}t \\ &\leq t\varphi'(\theta) + \frac{\varepsilon}{2}t \\ &\leq \varphi(\theta+t) - \varphi(\theta) + \varepsilon t \ , \end{aligned}$$

$$\begin{aligned} \|\Phi(\theta+t) - \Phi(\alpha)\| &\leq \|\Phi(\theta+t) - \Phi(\theta)\| + \|\Phi(\theta) - \Phi(\alpha)\| \\ &\leq \varphi(\theta+t) - \varphi(\alpha) + \varepsilon(\theta+t-\alpha) \ . \end{aligned}$$

Ceci implique qu'on a $[\alpha,\theta+\delta] \subset A$, en contradiction avec la définition de θ. On a donc $[\alpha,\beta] = A$, d'où :

$$\|\phi(\beta) - \phi(\alpha)\| \leq \varphi(\beta) - \varphi(\alpha) + \varepsilon(\beta - \alpha) \ .$$

Comme ceci est vrai pour tout $\varepsilon > 0$, on conclut :

$$\|\phi(\beta) - \phi(\alpha)\| \leq \varphi(\beta) - \varphi(\alpha) \ .$$

On applique souvent le théorème précédent dans la situation suivante : E, F sont des espaces vectoriels normés de dimension finie, U est une partie ouverte de E, $H : U \to F$ est une application de classe C^1, a_0, a_1 sont deux points de U tels que le segment joignant a_0 à a_1 est contenu dans U. On prend alors $\alpha = 0$, $\beta = 1$, et on définit Φ par :

$$\Phi(t) = H(a_0 + t(a_1 - a_0)) \ , \quad 0 \leq t \leq 1 \ .$$

Posons, pour $0 \leq t \leq 1$, $a_t = a_0 + t(a_1 - a_0)$ (la définition est cohérente pour $t = 0$ et $t = 1$).

Pour $t \in [0,1]$, on a :

$$\Phi'(t) = D_{a_t}H(a_1 - a_0) \ .$$

COROLLAIRE 1. *Si $M \geq 0$ est une constante telle que*

$$\forall t \in [0,1] \ , \quad \|D_{a_t}H(a_1 - a_0)\| \leq M \ ,$$

alors on a $\|H(a_1) - H(a_0)\| \leq M$.

DÉMONSTRATION : Il suffit de prendre $\varphi(t) = Mt$.

COROLLAIRE 2. *Si $M \geq 0$ est une constante telle que*

$$\forall t \in [0,1] \ , \quad \|D_{a_t}H\| \leq M \ ,$$

alors on a $\|H(a_1) - H(a_0)\| \leq M \|a_1 - a_0\|$.

COROLLAIRE 3. *Si $M \geq 0$ est une constante telle que*

$$\forall t \in [0,1] \ , \quad \|D_{a_t}H - D_{a_0}H\| \leq M \ ,$$

alors on a $\|H(a_1) - H(a_0) - D_{a_0}H(a_1 - a_0)\| \leq M \|a_1 - a_0\|$.

DÉMONSTRATION : Il suffit d'appliquer le corollaire 2 à l'application $\widetilde{H}(x) = H(x) - D_{a_0}H(x)$. □

COROLLAIRE 4. *Supposons que U soit* **connexe**, *et que la différentielle de H en tout point de U soit nulle.*

Alors H est constante.

DÉMONSTRATION : La relation " $H(x) = H(y)$ " entre points de U est une relation d'équivalence. En prenant $M = 0$ dans le corollaire 2, on voit que chaque classe d'équivalence est ouverte. D'après le lemme 2 du §. 3 du Chapitre III, il n'y a qu'une seule classe d'équivalence, c'est-à-dire que H est constante.

COROLLAIRE 5. *Supposons que U soit* **convexe** *et qu'il existe une constante $M \geq 0$ telle qu'on ait $\|D_xH\| \leq M$ pour tout $x \in U$. Alors l'application H est M-lipschitzienne, c'est-à-dire qu'on a, pour $x, y \in U$:*

$$\|H(x) - H(y)\| \leq M \|x - y\| \ .$$

DÉMONSTRATION : Comme U est convexe, pour tous $x, y \in U$, le segment joignant x à y est contenu dans U et on peut appliquer le corollaire 2. □

§. 3. Suite d'applications de classe C^1.

Soient E, F des espaces vectoriels normés de dimension finie, U une partie ouverte de E, et $(H_n)_{n\geq 0}$ une suite d'applications **continues** de U dans F.

LEMME 1. *Les propriétés suivantes sont équivalentes :*

(i) *La suite $(H_n)_{n\geq 0}$* ***est localement uniformément bornée :*** *pour tout $x \in U$, il existe $r(x) > 0$ et $C(x) > 0$ tels que $B(x, r(x)) \subset U$ et $\|H_n(y)\| \leq C(x)$ pour tous $n \geq 0$, $y \in B(x, r(x))$.*

(ii) *La suite $(H_n)_{n\geq 0}$ est* ***uniformément bornée sur les parties compactes de*** U *: pour toute partie compacte K de U, il existe $C(K) > 0$ tel que $\|H_n(y)\| \leq C(K)$ pour tous $n \geq 0$, $y \in K$.*

DÉMONSTRATION :

(ii)$\Rightarrow$(i). Soit $x \in U$; il existe $r > 0$ tel que $B(x, r) \subset U$. Prenons $r(x) = \frac{1}{2}r$, $K = \overline{B}(x, r(x))$: on obtient (i) avec $C(x) = C(K)$.

(i)$\Rightarrow$(ii). Soit K une partie compacte de U. Pour tout $x \in K$, soient $r(x) > 0$, $C(x) > 0$ vérifiant (i). On a $K \subset \bigcup_{x\in K} B(x, r(x))$, donc, d'après le théorème de Borel-Lebesgue, il existe des points $x_1, \ldots, x_k$ de K tels que $K \subset \bigcup_{i=1}^{k} B(x_i, r(x_i))$. La condition (ii) est alors vérifiée en prenant $C(K) = \max_{1\leq i\leq k} C(x_i)$.

LEMME 2. *Soit $H : U \to F$ une application. Les propriétés suivantes sont équivalentes :*

(i) *La suite $(H_n)_{n\geq 0}$ converge* ***localement uniformément vers*** H *: pour tout $x \in U$, il existe $r(x) > 0$ tel que :*

$$\lim_{n\to+\infty} \Big(\sup_{\|y-x\|<r(x)} \|H_n(y) - H(y)\| \Big) = 0 .$$

(ii) *La suite $(H_n)_{n\geq 0}$ converge vers H* ***uniformément sur les parties compactes de*** U *: pour toute partie compacte K de u, on a :*

$$\lim_{n\to+\infty} \Big(\sup_{y\in K} \|H_n(y) - H(y)\| \Big) = 0 .$$

La démonstration du lemme 2, très similaire à celle du lemme 1, est laissée en exercice.

REMARQUE. La propriété (i) (ou (ii)) du lemme 2 implique que H est continue (proposition 6 du chapitre IV).

PROPOSITION 2. *On suppose que :*

(i) *U est* ***connexe****;*

(ii) *les applications H_n, pour $n \geq 0$, sont de classe C^1 ;*

(iii) *il existe un point $a \in U$, tel que la suite $(H_n(a))_{n\geq 0}$ soit bornée;*

(iv) *la suite $(DH_n)_{n\geq 0}$ d'applications de U dans $\mathcal{L}(E,F)$ est* ***localement uniformément bornée****.*

Alors, il existe une application continue H de U dans F et une suite extraite $(H_{n_k})_{k\geq 0}$ qui converge vers H ***uniformément sur les parties compactes*** *de U.*

DÉMONSTRATION : Elle comporte plusieurs étapes.

1. Soit $x \in U$. Montrons que la suite $(H_n(x))_{n\geq 0}$ est bornée.

Comme U est connexe, il existe des points $x_0 = a$, $x_1, \dots, x_k = x$ dans U tels que, pour $0 \leq i \leq k-1$, le segment S_i joignant x_i à x_{i+1} soit contenu dans U. (La démonstration de cette assertion est laissée en exercice : s'inspirer de la démonstration de la proposition 7 du Chapitre III).

L'union $K = \bigcup_{i=0}^{k-1} S_i$ est une partie compacte de U ; d'après le lemme 1, il existe donc $C(K) > 0$ tel qu'on ait, pour tous $n \geq 0$, $y \in K$:

$$\|D_y H_n\| \leq C(K) .$$

D'après le corollaire 2 du §. 2, on a alors, pour tout $n \geq 0$:

$$\begin{aligned}\|H_n(x)\| &\leq \|H_n(a)\| + \sum_{i=0}^{k-1} \|H_n(x_{i+1}) - H_n(x_i)\| \\ &\leq \|H_n(a)\| + C(K) \sum_{i=0}^{k-1} \|x_{i+1} - x_i\| ,\end{aligned}$$

donc la suite $(H_n(x))_{n\geq 0}$ est bornée.

2. Soit K une partie compacte de U. Montrons qu'il existe une suite extraite $(H_{n_k})_{k\geq 0}$ dont les restrictions à K forment une suite convergeant uniformément sur K.

La suite $(H_n)_{n\geq 0}$ est **équicontinue** : en effet, d'après l'hypothèse (iv) et le corollaire 2 du §. 2, pour tout $x \in U$ il existe $r(x) > 0$ et $C(x) > 0$ tels qu'on ait, pour tout $n \geq 0$, $y \in B(x, r(x))$:

$$\|H_n(y) - H_n(x)\| \leq C(x)\ \|y - x\| .$$

D'autre part, on a montré en 1. que pour tout $x \in U$ la suite $(H_n(x))_{n\geq 0}$ est bornée. Il résulte donc du théorème d'Ascoli que les restrictions à K des applications H_n forment une partie de $\mathcal{C}(K,F)$ dont l'adhérence est **compacte**. L'assertion annoncée en résulte.

3. Pour tout $m \geq 1$, définisons :

$$K_m = \left\{x \in U \,:\, \|x\| \leq m \text{ et } B\left(x, \frac{1}{m}\right) \subset U\right\}.$$

On a $U = \bigcup_{m \geq 1} K_m$. La condition $B\left(x, \frac{1}{m}\right) \subset U$ est toujours vérifiée lorsque $U = E$, et signifie $d(x, E-u) \geq \frac{1}{m}$ sinon. Or l'application : $x \mapsto d(x, E-U)$ est continue (exercice). Pour tout $m \geq 1$, K_m est par conséquent fermée et bornée, donc **compacte**. Si K est une partie compacte de U, K est bornée et il existe x dans K tel que pour tout $y \in K$ on ait $d(y, E-U) \geq d(x, E-U)$. Il existe donc un entier $m = m(K)$ tel que $K \subset K_m$.

4. Processus diagonal.

D'après la partie 2., il existe une suite $(H_n^1)_{n\geq 0}$ extraite de $(H_n)_{n\geq 0}$ convergeant uniformément sur K_1. Or toute suite extraite de $(H_n)_{n\geq 0}$ vérifie encore les hypothèses de la proposition. On peut donc d'après la partie 2. construire par récurrence pour $m \geq 2$ une suite $(H_n^m)_{n\geq 0}$ extraite de $(H_n^{m-1})_{n\geq 0}$ qui converge uniformément sur K_m.

Posons, pour $n \geq 1$, $\tilde{H}_n = H_n^n$. La suite $(\tilde{H}_n)_{n\geq 0}$ est extraite de $(H_n)_{n\geq 0}$; pour tout $m \geq 1$, elle est uniformément convergente sur K_m ; d'après la partie 3., elle est donc uniformément convergente sur toute partie compacte K de U. Sa limite H est automatiquement continue (cf. Remarque précédant la proposition 2). Ceci termine la démonstration de la proposition 2.

PROPOSITION 3. *On suppose que :*

(i) *U est* ***connexe****;*

(ii) *les applications H_n, pour $n \geq 0$ sont de classe C^1 ;*

(iii) *il existe $a \in U$ tel que la suite $(H_n(a))_{n\geq 0}$ soit convergente;*

(iv) *il existe une application L de U dans $\mathcal{L}(E,F)$ telle que la suite $(DH_n)_{n\geq 0}$ converge* ***localement uniformément*** *vers L.*

Alors la suite $(H_n)_{n\geq 0}$ converge ***uniformément sur les parties compactes*** *de U. Sa limite H est de classe C^1 et on a $DH = L$.*

DÉMONSTRATION :

1. Soit $x \in U$. Il existe $r(x) > 0$ tel que $\bar{B}(x, r(x)) \subset U$ et que la suite $(DH_n)_{n\geq 0}$ converge vers L uniformément sur $B(x, r(x))$. Cela signifie que la suite $(C_n)_{n\geq 0}$ définie par :

$$C_n = \sup_{m\geq n} \sup_{\|y-x\|<r(x)} \|D_y H_n - D_y H_m\|$$

converge vers 0.

D'après le corollaire 2 du §. 2, on a, pour $y, y' \in B(x, r(x))$ et $m \geq n \geq 0$:

$$\text{(1)} \qquad \|[H_n(y') - H_m(y')] - [H_n(y) - H_m(y)]\| \leq C_n \|y - y'\| ,$$

donc la suite $(H_n(y))_{n\geq 0}$ est de Cauchy si et seulement si la suite $(H_n(y'))_{n\geq 0}$ est de Cauchy.

2. Soit A l'ensemble des points $y \in U$ tels que la suite $(H_n(y))_{n\geq 0}$ soit convergente. Montrons que $A = U$.

En effet, A n'est pas vide, car $a \in A$. La partie A est ouverte : si $x \in A$, $B(x, r(x)) \subset A$ d'après 1. La partie A est fermée (dans U) : si une suite $(x_n)_{n\geq 0}$ d'éléments de A converge vers une limite $x \in U$, on a $x_n \in B(x, r(x))$ pour n assez grand, donc $x \in A$ d'après 1.

Comme U est connexe, ceci implique $A = U$.

3. Notons $H(y)$ la limite de la suite $(H_n(y))_{n\geq 0}$, pour $y \in U$. Soit $x \in U$; en prenant $y' = x$ et en laissant tendre m vers $+\infty$ dans la relation (1), on obtient pour $y \in B(x, r(x))$, $n \geq 0$:

$$\|H_n(y) - H(y)\| \leq \|H_n(x) - H(x)\| + C_n \|y - x\| ,$$

donc la suite $(H_n)_{n\geq 0}$ converge uniformément sur $B(x, r(x))$. D'après le lemme 2, elle converge vers H uniformément sur les parties compactes de U.

4. Soit $x \in U$. Montrons que H est différentiable en x, avec $D_x H = L(x)$.

Soient $\varepsilon > 0$, n un entier tel que $C_n \leq \frac{\varepsilon}{3}$. On a alors :

$$\|D_x H_n - L(x)\| \leq \frac{\varepsilon}{3} ,$$

et

$$\|[H(y) - H(x)] - [H_n(y) - H_n(x)]\| \leq \frac{\varepsilon}{3} \|y - x\|$$

pour tout $y \in B(x, r(x))$, d'après la relation (1).

Comme H_n est différentiable en x, il existe $\delta \in]0, r(x)[$ tel qu'on ait, pour tout $y \in B(x, \delta)$:

$$\|H_n(y) - H_n(x) - D_x H_n(y - x)\| \leq \frac{\varepsilon}{3} \|y - x\| .$$

Pour $y \in B(x,\delta)$, on a donc :

$$\begin{aligned}
\Big\|H(y)-H(x) - L(x)(y-x)\Big\| \\
&\leq \Big\| [H(y)-H(x)] - [H_n(y)-H_n(x)] \Big\| \\
&\quad + \Big\|H_n(y) - H_n(x) - D_x H_n(y-x)\Big\| \\
&\quad + \Big\|D_x H_n(y-x) - L(x)(y-x)\Big\| \\
&\leq \frac{\varepsilon}{3}\,\|y-x\| + \frac{\varepsilon}{3}\,\|y-x\| + \frac{\varepsilon}{3}\,\|y-x\| \ ,
\end{aligned}$$

ce qui démontre l'assertion annoncée.

Or, d'après la Remarque précédant la proposition 2, l'application L est continue. On conclut que H est de classe C^1. □

CHAPITRE VII - Le théorème des fonctions implicites.

§. 1. Le théorème d'inversion locale.

Soient E, F des espaces vectoriels de <u>même</u> dimension finie, U une partie ouverte de E, V une partie ouverte de F et $H : U \to F$ une application.

DÉFINITION. On dit que H est un **difféomorphisme de classe C^1 de U sur V** si :

(i) H est une application de classe C^1 dans U ;
(ii) H est une bijection de U sur V ;
(iii) la bijection réciproque H^{-1} est une application de classe C^1.

Soit H un difféomorphisme de classe C^1 de U sur V. Comme toute application de classe C^1 est continue, H et H^{-1} sont continues, donc H est un homéomorphisme de U sur V.

Soit $y \in V$; comme $H \circ H^{-1} = \mathrm{id}$, il résulte de la formule pour la différentielle d'une composition (Chapitre V, §. 3.3, Proposition 5) qu'on a :

$$D_y(H^{-1}) = \left(D_{H^{-1}(y)} H\right)^{-1} .$$

LEMME. *Soit E un espace de Banach. Soit $u \in \mathcal{LC}(E)$. Si $\|u\| < 1$, $(\mathrm{id} - u)$ est inversible et son inverse $v = \sum_{n\geq 0} u^n$ appartient à $\mathcal{LC}(E)$.*

DÉMONSTRATION : Pour $m \geq 0$, posons :

$$v_m = \sum_{n=0}^{m} u^n .$$

On a $v_m \in \mathcal{LC}(E)$ et pour $m' > m$:

$$\|v_m - v_{m'}\| = \left\| \sum_{n=m+1}^{m'} u^n \right\| \leq \sum_{n=m+1}^{m'} \|u\|^n \leq (1 - \|u\|)^{-1} \|u\|^{m+1} ,$$

donc $(v_m)_{m\geq 0}$ est une suite de Cauchy dans $\mathcal{LC}(E)$. Or $\mathcal{LC}(E)$est un espace de Banach (Chapitre IV, §. 5, Proposition 10), donc $(v_m)_{m\geq 0}$ est convergente. Soit v la limite. Pour $m \geq 0$, on a $(\mathrm{id} - u) \circ v_m = v_m \circ (\mathrm{id} - u) = \mathrm{id} - u^{m+1}$; comme l'application $(f, g) \to (f \circ g)$ de $\mathcal{LC}(E) \times \mathcal{LC}(E)$ dans $\mathcal{LC}(E)$ est continue, on a par passage à la limite $(\mathrm{id} - u) \circ v = v \circ (\mathrm{id} - u) = \mathrm{id}$.

THÉORÈME (d'inversion locale). *Soient* E, F, U, H *comme ci-dessus, et* $a \in U$. *On suppose que la différentielle de* H *en* a *est une application linéaire* ***inversible*** *de* E *dans* F. *Alors il existe une partie ouverte* U_1 *de* U *contenant* a *et une partie ouverte* V_1 *de* F *contenant* $H(a)$ *telles que la restriction de* H *à* U_1 *soit* ***un difféomorphisme de classe*** $\mathbf{C}^1$ ***de*** $\mathbf{U}_1$ ***sur*** $\mathbf{V}_1$.

DÉMONSTRATION :

1. Soit $A : E \to F$ l'application affine :

$$A(x) = H(a) + D_a H(x-a) .$$

Elle est inversible puisque $D_a H$ l'est. Posons :

$$\widetilde{H}(x) = A^{-1} \circ H(x) , \quad \text{pour } x \in U ;$$

alors $\widetilde{H}$ est une application de classe C^1 de U dans E ; on a $\widetilde{H}(a) = a$ et $D_a\widetilde{H} = \mathrm{id}_E$. Posons, pour $x \in U$:

$$\widetilde{H}(x) = x + \varphi(x) .$$

Alors φ est une application de classe C^1 de U dans E, et on a $\varphi(a) = 0$, $D_a\varphi = 0$.

2. Soit $\delta > 0$ un nombre assez petit pour qu'on ait $\overline{B}(a,\delta) \subset U$ et :

$$\forall x \in \overline{B}(a,\delta) , \quad \|D_x\varphi\| \leq \frac{1}{2} .$$

D'après le théorème de la moyenne, on a, pour $x, x' \in \overline{B}(a,\delta)$:

$$\|\varphi(x) - \varphi(x')\| \leq \frac{1}{2} \|x - x'\| \tag{1}$$

et en particulier :

$$\|\varphi(x)\| \leq \frac{1}{2} \|x - a\| . \tag{2}$$

3. Soit $y \in B(a, \frac{\delta}{2})$. Pour $x \in \overline{B}(a,\delta)$, posons :

$$f_y(x) = y - \varphi(x) .$$

D'après (2), l'application f_y envoie $\overline{B}(a,\delta)$ dans $B(a,\delta)$, et d'après (1) elle est $\frac{1}{2}$–lipschitzienne, donc contractante.

Or $\overline{B}(a,\delta)$ est complet (c'est une partie fermée de E, qui est complet). D'après le théorème du point fixe, l'application f_y a un unique point fixe; notons le $K(y)$. On a $K(y) \in B(a,\delta)$.

4. Soit $\varepsilon > 0$. Comme $D_a\varphi = 0$, il existe $\delta' \in \left]0, \frac{\delta}{2}\right[$ tel qu'on ait, pour $y \in B(a, \delta')$:

$$\|\varphi(y)\| \leq \frac{\varepsilon}{2}\, \|y - a\| \ .$$

Soit $y \in B(a, \delta')$; posons $y_0 = y$ et définissons pour $n \geq 1$ le point y_n par $y_n = f_y(y_{n-1})$. On a :

$$\|y_1 - y_0\| = \|f_y(y) - y\| = \|\varphi(y)\| \leq \frac{\varepsilon}{2}\, \|y - a\|$$

$$\|y_{n+1} - y_n\| \leq \frac{\varepsilon}{2^n}\, \|y - a\| \ , \ \text{pour } n \geq 0$$

et $\lim\limits_{n \to +\infty} y_n = K(y)$, d'où :

$$\|K(y) - y\| \leq \varepsilon\, \|y - a\| \ .$$

On conclut que l'application $K : B\left(a, \frac{\delta}{2}\right) \to E$ est différentiable en a, sa différentielle y étant égale à l'identité.

5. Posons

$$U_1 = \widetilde{H}^{-1}\left(B\left(a, \frac{\delta}{2}\right)\right) \cap B(a, \delta) \ ,$$

$$V_1 = A\left(B\left(a, \frac{\delta}{2}\right)\right) \ .$$

Comme $\widetilde{H}$ est continue et A est un homéomorphisme, U_1 est une partie ouverte de U et V_1 une partie ouverte de F. On a $a \in U_1$ et $H(a) = A(a) \in V_1$.

On a $\widetilde{H}(U_1) \subset B\left(a, \frac{\delta}{2}\right)$, donc $H(U_1) \subset V_1$. Soient $y \in B\left(a, \frac{\delta}{2}\right)$, $z = A(y)$ et $x \in U_1$. On a :

$$H(x) = z \Leftrightarrow \widetilde{H}(x) = y \Leftrightarrow f_y(x) = x \ ,$$

avec, $x \in B(a, \delta)$, d'où

$$H(x) = z \Leftrightarrow x = K(y) \ ,$$

ce qui montre que H est une bijection de U_1 sur V_1 et que la bijection réciproque est $K \circ A^{-1}$.

6. Soient $z \in V_1$, $x = K \circ A^{-1}(z) \in U_1$. On a $\|D_x\varphi\| < 1$, donc d'après le lemme $D_x\widetilde{H}$, et par suite D_xH, sont des applications linéaires inversibles. D'après ce qui précède il existe une partie ouverte U_1' contenant x, une partie ouverte V_1' contenant z telles que la restriction de H à U_1' soit une bijection de U_1' sur V_1', et que la bijection réciproque K' soit différentiable en z. Donc K' est continue en z, et il existe $\delta' > 0$ tel que $B(z,\delta') \subset V_1 \cap V_1'$, $K'(B(z,\delta')) \subset U_1 \cap U_1'$. On a $K' = K \circ A^{-1}$ sur $B(z,\delta')$, donc $K \circ A^{-1}$ est différentiable en tout point de V_1. Finalement, pour $z \in V_1$, on doit avoir

$$D_z[K \circ A^{-1}] = \left[D_{K \circ A^{-1}(z)} H\right]^{-1}$$

(Chapitre V, §. 3.3), donc $K \circ A^{-1}$ est une application de classe C^1.

REMARQUE. Soient E, F, U, H comme ci-dessus. Supposons que pour tout $x \in U$ la différentielle $D_x H$ de x en U soit inversible. Si de plus **l'application H est injective** alors H est un difféomorphisme de classe C^1 de U sur $H(U)$. Mais l'hypothèse n'est pas suffisante pour assurer l'injectivité de H.

EXEMPLE. Prenons $E = F = \mathbf{R}^2$, $U = \{(x,y), y > 0\}$, $H(x,y) = (y\cos x, y \sin x)$ ("coordonnées polaires"). La matrice jacobienne de H en un point (x,y) est

$$\begin{pmatrix} -y\sin x & \cos x \\ y\cos x & \sin x \end{pmatrix}$$

et est donc inversible pour tout $(x,y) \in U$. Mais on a $H(x+2\pi, y) = H(x,y)$ pour tout $(x,y) \in U$, donc H n'est pas injective.

§. 2. Le théorème des fonctions implicites.

Soient E un espace vectoriel de dimension finie, et E_1, E_2 des sous-espaces vectoriels de E tels que $E = E_1 \oplus E_2$. Soient U une partie ouverte de E, et H une application de classe C^1 dans U à valeurs dans un espace vectoriel de dimension finie F. Soit (a_1, a_2) un point de E (avec $a_1 \in E_1$, $a_2 \in E_2$).

THÉORÈME (des fonctions implicites). *On suppose que E_2 et F ont **même dimension** et que la différentielle partielle dans la direction de E_2 de l'application H au point (a_1, a_2) est une application linéaire **inversible** de E_2 dans F.*

Il existe alors un ouvert U_1 de E_1 contenant a_1 et un ouvert U_2 de E_2 contenant a_2 tels que :

1) $U_1 \times U_2 \subset U$;
2) *pour tout x_1 dans U_1, l'équation :*

$$H(x_1, x_2) = H(a_1, a_2)$$

*a **une et une seule** solution x_2 **dans** U_2 ;*

3) *si on note $K(x_1)$ cette solution, l'application $K : U_1 \to U_2$ est **de classe** C^1 dans U_1.*

DÉMONSTRATION : Nous définissons une application $\widetilde{H}$ de U dans $E_1 \times F$ par la formule :

$$\widetilde{H}(x_1, x_2) = (x_1, H(x_1, x_2)) .$$

Cette application est de classe C^1 et sa différentielle au point (a_1, a_2) est donnée par :

$$D_{(a_1,a_2)}\widetilde{H}(v_1,v_2) = \big(v_1, D_{(a_1,a_2)}H(v_1,v_2)\big) \ .$$

On a, pour $w_1 \in E_1$, $w \in F$:

$$D_{(a_1,a_2)}\widetilde{H}(v_1,v_2) = (w_1,w) \Leftrightarrow \begin{cases} w_1 = v_1 \\ D_{(a_1,a_2)}H(v_1,v_2) = w \end{cases}$$

$$\Leftrightarrow \begin{cases} v_1 = w_1 \\ D_{(a_1,a_2)}H(0,v_2) = w - D_{(a_1,a_2)}H(w_1,0) \ . \end{cases}$$

Comme l'application $v_2 \mapsto D_{(a_1,a_2)}H(0,v_2)$ est supposée inversible, la différentielle $D_{(a_1,a_2)}\widetilde{H}$ est inversible.

D'après le théorème d'inversion locale, il existe un ouvert $\widetilde{U} \subset U$ contenant (a_1,a_2) et un ouvert $\widetilde{V} \subset E_1 \times F$ contenant $(a_1, H(a_1,a_2))$ tels que la restriction de $\widetilde{H}$ à $\widetilde{U}$ soit un difféomorphisme de classe C^1 de $\widetilde{U}$ sur $\widetilde{V}$.

Notons $\widetilde{K} : \widetilde{V} \to \widetilde{U}$ le difféomorphisme inverse. Quitte à diminuer $\widetilde{U}$, nous pouvons supposer que $\widetilde{U}$ est de la forme $\widetilde{U} = \widetilde{U}_1 \times U_2$, où $\widetilde{U}_1$ est un ouvert de E_1 contenant a_1 et U_2 un ouvert de E_2 contenant a_2.

Posons :

$$U_1 = \left\{x \in \widetilde{U}_1,\ (x_1, H(a_1,a_2)) \in \widetilde{V}\right\} \ .$$

C'est une partie ouverte de $\widetilde{U}_1$ contenant a_1. On a :

$$U_1 \times U_2 \subset \widetilde{U}_1 \times U_2 = \widetilde{U} \subset U \ ;$$

pour $x_1 \in U_1$, $x_2 \in U_2$, on a (en posant $x = (x_1,x_2)$) :

$$H(x_1,x_2) = H(a_1,a_2) \Leftrightarrow \widetilde{H}(x) = (x_1, H(a_1,a_2))$$
$$\Leftrightarrow x = \widetilde{K}\,(x_1, H(a_1,a_2)) \ ,$$

d'où l'assertion 2) de l'énoncé; si on écrit $\widetilde{K} = (K_1,K_2)$ avec $K_1 : \widetilde{V} \to \widetilde{U}_1$ et $K_2 : \widetilde{V} \to U_2$, on a, pour $x_1 \in U_1$:

$$K(x_1) = K_2\,(x_1, H(a_1,a_2)) \ ;$$

Comme $\widetilde{K}$ (et donc K_1, K_2) sont de classe C^1 dans $\widetilde{V}$, K est de classe C^1 dans U_1. □

Complément. Différentielle de l'application K.

Pour $x \in U$, notons $D_x^{(1)}H \in \mathcal{L}(E_1,F)$ la différentielle de H en x dans la direction de E_1 et $D_x^{(2)}H \in \mathcal{L}(E_2,F)$ la différentielle de H en x dans la direction de E_2. L'hypothèse du théorème est que $D_{(a_1,a_2)}^{(2)}H$ est inversible.

On a choisi U_1, U_2 suffisamment petits pour que la différentielle $D_{(x_1,x_2)}^{(2)}H \in \mathcal{L}(E_2,F)$ soit inversible lorsque $x_1 \in U_1$, $x_2 \in U_2$.

Dans U_1, on a l'identité :

$$H(x_1, K(x_1)) = H(a_1, a_2) \ ;$$

en différentiant au point x_1, on obtient, **en posant $\boldsymbol{x = (x_1, K(x_1))}$:**

$$D_x^{(1)} H + D_x^{(2)} H \circ D_{x_1} K = 0$$

(égalité dans $\mathcal{L}(E_1, F)$), d'où la formule suivante pour la différentielle de K en x_1 (en posant $x = (x_1, K(x_1))$) :

$$\boxed{D_{x_1} K = -(D_x^{(2)} H)^{-1} \circ D_x^{(1)} H}$$

(égalité dans $\mathcal{L}(E_1, E_2)$).

§. 3. Application 1 : surfaces dans $\mathbf{R}^3$.

3.1. Notons x_1, x_2, x_3 les coordonnées usuelles dans $\mathbf{R}^3$.

Soient U une partie ouverte de $\mathbf{R}^3$, H une application de classe C^1 de U dans $\mathbf{R}$.

Considérons la surface S d'équation $H = 0$, c'est-à-dire :

$$S = \left\{ x \in U,\ H(x) = 0 \right\} .$$

Soit $a = (a_1, a_2, a_3)$ un point de S. Nous **supposons** qu'on a

$$\boxed{\boxed{D_a H \neq 0}} \qquad (\text{dans}\, \mathcal{L}(\mathbf{R}^3, \mathbf{R}))$$

et nous allons **sous cette hypothèse** décrire S **au voisinage de** a (c'est-à-dire décrire $S \cap V$, où V est un ouvert suffisamment petit de $\mathbf{R}^3$ contenant a).

3.2. Description locale comme graphe.

La différentielle de H en a est la forme linéaire :

$$D_a H(x_1, x_2, x_3) = \partial_1 H(a) x_1 + \partial_2 H(a) x_2 + \partial_3 H(a) x_3 \ .$$

L'hypothèse est donc qu'**au moins un** des trois nombres $\partial_1 H(a)$, $\partial_2 H(a)$, $\partial_3 H(a)$ est non nul.

Supposons **par exemple** qu'on ait $\partial_3 H(a) \neq 0$. Nous allons appliquer le théorème des fonctions implicites avec :

$$E_1 = \left\{ x \in \mathbf{R}^3,\ x_3 = 0 \right\} ,$$

$$E_2 = \left\{ x \in \mathbf{R}^3,\ x_1 = x_2 = 0 \right\} , \quad F = \mathbf{R} \ .$$

(Nous identifions E_1 à $\mathbf{R}^2$ par l'application $(x_1, x_2, 0) \mapsto (x_1, x_2)$ et E_2 à $\mathbf{R}$ par l'application $(0, 0, x_3) \mapsto x_3$).

La différentielle de H en a dans la direction de E_2 est donnée par :

$$D_a^{(2)} H(t) = \partial_3 H(a) t$$

et est inversible puisque $\partial_3 H(a) \neq 0$.

On conclut qu'il existe :

- une partie ouverte U_{12} de $\mathbf{R}^2$ contenant (a_1, a_2),
- une partie ouverte U_3 de $\mathbf{R}$ contenant a_3,
- une application K de classe C^1 de U_{12} dans $\mathbf{R}$ telles que :

$$(U_{12} \times U_3) \cap S = \Big\{(x_1, x_2, K(x_1, x_2)), (x_1, x_2) \in U_{12}\Big\} .$$

En posant $V = U_{12} \times U_3$, on a donc décrit $S \cap V$ comme le **graphe** de l'application de classe C^1 $K : U_{12} \to \mathbf{R}$.

D'après le complément au théorème des fonctions implicites, les dérivées partielles de K en un point $x = (x_1, x_2) \in U_{12}$ sont données par :

$$\partial_1 K(x) = -\frac{\partial_1 H(x, K(x))}{\partial_3 H(x, K(x))} ,$$

$$\partial_2 K(x) = -\frac{\partial_2 H(x, K(x))}{\partial_3 H(x, K(x))} .$$

EXERCICE. Traiter de façon analogue le cas $\partial_2 H(a) \neq 0$.

3.3. Plan tangent.

La forme linéaire $D_a H$ n'étant pas (par hypothèse) identiquement nulle, son **noyau** est un plan (vectoriel) qu'on note $T_a S$ et appelle plan tangent (vectoriel) à S en a. L'équation de ce plan est donc $D_a H(x_1, x_2, x_3) = 0$.

Le plan affine parallèle à $T_a S$ et passant par a est appelé plan tangent (affine) à S en a. Son équation est

$$D_a H(x_1 - a_1, x_2 - a_2, x_3 - a_3) = 0 .$$

Nous allons décrire S au voisinage de a dans un repère de $\mathbf{R}^3$ adapté au point a. Notons (e_1, e_2, e_3) la base canonique de $\mathbf{R}^3$, et choisissons une base orthonormée (f_1, f_2, f_3) de $\mathbf{R}^3$ telle que $f_1, f_2 \in T_a S$. Notons $\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{x}_3)$ les coordonnées dans le repère d'origine a déterminé par (f_1, f_2, f_3) d'un point $x = (x_1, x_2, x_3)$ de $\mathbf{R}^3$. On a donc

$$x = a + R(\bar{x}) = \widehat{R}(\bar{x}) ,$$

où R est le déplacement (vectoriel) de $\mathbf{R}^3$ tel que

$$R(f_i) = e_i \ , \quad i = 1, 2, 3$$

et $\widehat{R}$ est le déplacement affine correspondant envoyant 0 sur a.

Dans les nouvelles coordonnées :

- le point a a pour coordonnées $(0, 0, 0)$
- la surface S a pour équation

$$H \circ \widehat{R}(\bar{x}) = 0$$

– les plans tangents vectoriels et affines coïncident et ont pour équation $\bar{x}_3 = 0$.

En d'autre termes, si on pose $\overline{H} = H \circ \widehat{R}$, on a :

$$\begin{aligned} \overline{H}(0,0,0) &= 0 \ , \\ \partial_1 \overline{H}(0,0,0) &= 0 \ , \\ \partial_2 \overline{H}(0,0,0) &= 0 \ , \\ \partial_3 \overline{H}(0,0,0) &\neq 0 \ . \end{aligned}$$

D'après 3.2, on peut donc écrire :

$$S \cap V = \left\{ \bar{x}, \bar{x}_3 = \overline{K}(\bar{x}_1, \bar{x}_2) \right\}$$

avec V un ouvert de $\mathbf{R}^3$ contenant a.

$\overline{K}$ une application de classe C^1, définie sur un ouvert de $\mathbf{R}^2$ contenant $(0,0)$, à valeurs dans $\mathbf{R}$ et vérifiant

$$\frac{\partial \overline{K}}{\partial \bar{x}_1}(0,0) = \frac{\partial \overline{K}}{\partial \bar{x}_2}(0,0) = 0 \ .$$

On retrouve la notion géométrique de plan tangent puisque la distance d'un point de $S \cap V$ à $\{\bar{x}_3 = 0\} = T_a S$ est donnée par :

$$|\bar{x}_3| = o(\|\bar{x}\|) = o(\|x - a\|) \ .$$

[Si H est de classe C^2 (cf. chapitre suivant), on obtient même $|\bar{x}_3| = O(\|x - a\|^2)$].

EXERCICE. En utilisant 3.2, montrer que le plan affine tangent est l'unique plan affine P satisfaisant

$$d(x, P) = o(\|x - a\|)$$

lorsque x tend vers a dans S.

3.4. Surfaces paramétrées.

Nous avons jusqu'à présent considéré le cas d'une surface définie par une équation $\{H = 0\}$. Considérons maintenant une surface (encore notée S) définie par un paramétrage : soient W une partie ouverte de $\mathbf{R}^2$, $\varphi = (\varphi_1, \varphi_2, \varphi_3)$ une application de classe C^1 de W dans $\mathbf{R}^3$, (s_0, t_0) un point de W (on note (s,t) les coordonnées dans $\mathbf{R}^2$).

Il s'agit maintenant, pour $\varepsilon > 0$ assez petit, de décrire l'image S_ε par φ de la boule de centre (s_0, t_0), de rayon ε.

Nous **supposons** que la différentielle $D_{(s_0,t_0)}\varphi \in \mathcal{L}(\mathbf{R}^2, \mathbf{R}^3)$ est **injective.**

Cela signifie que les vecteurs $\dfrac{\partial \varphi}{\partial s}(s_0, t_0)$ et $\dfrac{\partial \varphi}{\partial t}(s_0, t_0)$ sont **linéairement indépendants.**

Supposons **par exemple** qu'on ait

$$\det \begin{pmatrix} \dfrac{\partial\varphi_1}{\partial s} & \dfrac{\partial\varphi_1}{\partial t} \\ \dfrac{\partial\varphi_2}{\partial s} & \dfrac{\partial\varphi_2}{\partial t} \end{pmatrix} (s_0, t_0) \neq 0 \ .$$

L'application $\widetilde{\varphi} : (s,t) \mapsto (\varphi_1(s,t), \varphi_2(s,t))$ est de classe C^1, et sa différentielle en (s_0, t_0) est alors inversible. D'après le théorème d'inversion locale, il existe une boule ouverte B de centre (s_0, t_0), de rayon $\varepsilon_0 > 0$ telle que la restriction de $\widetilde{\varphi}$ à B soit un difféomorphisme de classe C^1 de B sur $\widetilde{\varphi}(B)$.

On a alors :

$$S_{\varepsilon_0} = \Big\{(x_1, x_2, x_3)\,,\, (x_1, x_2) \in \widetilde{\varphi}(B)\,,\, x_3 = \varphi_3 \circ \widetilde{\varphi}^{-1}(x_1, x_2)\Big\}$$

(représentation comme graphe, cf. 3.2).

L'équation $x_3 - \varphi_3 \circ \widetilde{\varphi}^{-1}(x_1, x_2) = 0$ obtenue pour S_{ε_0} nous permet de calculer le plan tangent (vectoriel) à S_{ε_0} en $a = \varphi(s_0, t_0)$; son équation est :

$$D_a H(x_1, x_2, x_3) = 0 \ ,$$

avec $H(x_1, x_2, x_3) = x_3 - \varphi_3 \circ \widetilde{\varphi}^{-1}(x_1, x_2)$.

On voit que :

$$D_a H\left(\frac{\partial\varphi}{\partial s}(s_0, t_0)\right) = \frac{\partial\varphi_3}{\partial s}(s_0, t_0) - \frac{\partial\varphi_3}{\partial s}(s_0, t_0) = 0 \ ,$$

$$D_a H\left(\frac{\partial\varphi}{\partial t}(s_0, t_0)\right) = 0 \ ,$$

ce qui nous permet de conclure que le plan vectoriel tangent $T_a S_{\varepsilon_0}$ est **engendré** par $\dfrac{\partial\varphi}{\partial s}(s_0, t_0)$ et $\dfrac{\partial\varphi}{\partial t}(s_0, t_0)$. En d'autres termes :

$$\boxed{T_a S_{\varepsilon_0} = \mathrm{Im}\ \left(D_{(s_0, t_0)}\varphi\right)}$$

§. 4. Application 2 : courbes dans $\mathbf{R}^3$.

4.1. Soient U une partie ouverte de $\mathbf{R}^3$, et $H = (H_1, H_2)$ une application de classe C^1 de U dans $\mathbf{R}^2$. Considérons la courbe C d'équation $H = 0$:

$$C = \Big\{(x_1, x_2, x_3) \in U,\ H_1(x_1, x_2, x_3) = H_2(x_1, x_2, x_3) = 0\Big\} \ .$$

Soit $a = (a_1, a_2, a_3)$ un point de C.

On **suppose** que la différentielle de H en a est **surjective**, et on veut décrire sous cette hypothèse la courbe C au voisinage de a.

4.2. Comme la différentielle $D_a H$ est surjective, son noyau est de dimension 1 et est appelé droite (vectorielle) tangente à C en a, notée $T_a C$. La droite affine parallèle à $T_a C$ passant par a est la droite (affine) tangente à C en a.

Supposons **par exemple** que $T_a C$ ne soit pas contenu dans le plan d'équation $x_3 = 0$.

Cela revient à dire que la restriction de $D_a H$ à ce plan est une application linéaire **inversible** de ce plan dans $\mathbf{R}^2$.

Nous appliquons le théorème des fonctions implicites à H au point a en prenant :

$$E_1 = \Big\{(0,0,x_3),\ x_3 \in \mathbf{R}\Big\}$$
$$E_2 = \Big\{(x_1,x_2,0),\ x_1,x_2 \in \mathbf{R}\Big\}\ ,\quad F = \mathbf{R}^2\ .$$

On conclut qu'il existe :

- un ouvert U_3 de $\mathbf{R}$ contenant a_3,
- un ouvert U_{12} de $\mathbf{R}^2$ contenant (a_1,a_2),
- une application de classe C^1 $K = (K_1,K_2)$ de U_3 dans $\mathbf{R}^2$

telles que :

$$(U_{12} \times U_3) \cap C = \Big\{ (K_1(x_3),K_2(x_3),x_3)\ ,\ x_3 \in U_3 \Big\}\ ,$$

c'est-à-dire qu'en posant $V = U_{12} \times U_3$, on a présenté $C \cap V$ comme le graphe de l'application de classe C^1 $K : U_3 \to \mathbf{R}^2$.

En dérivant par rapport à x_3 les relations $H_i\,(K_1(x_3),K_2(x_3),x_3) = 0$ $(i = 1,2)$ et en résolvant le système linéaire obtenu, on a :

$$K_1'(a_3) = -\frac{\partial_3 H_1 \partial_2 H_2 - \partial_3 H_2 \partial_2 H_1}{\partial_1 H_1 \partial_2 H_2 - \partial_1 H_2 \partial_2 H_1}(a)$$
$$K_2'(a_3) = -\frac{\partial_3 H_2 \partial_1 H_1 - \partial_3 H_1 \partial_1 H_2}{\partial_1 H_1 \partial_2 H_2 - \partial_1 H_2 \partial_2 H_1}(a)\ .$$

4.3. On peut aussi décrire C au voisinage de a dans un nouveau repère d'origine a associé à une base orthonormée (f_1,f_2,f_3) de $\mathbf{R}^3$ telle que f_3 soit un générateur de la droite vectorielle tangente à C en a.

En notant $\bar{x}_1,\bar{x}_2,\bar{x}_3$ les coordonnées dans ce nouveau repère, on a, comme en 3.3

$$x = \widehat{R}(\bar{x})\ ,$$

où $\widehat{R}$ est un déplacement affine de $\mathbf{R}^3$. L'équation de C dans ces coordonnées devient :

$$H \circ \widehat{R}(\bar{x}) = 0\ ,$$

c'est-à-dire

$$\overline{H}_1(\bar{x}_1,\bar{x}_2,\bar{x}_3) = \overline{H}_2(\bar{x}_1,\bar{x}_2,\bar{x}_3) = 0\ ,$$

où on a posé $\bar{H}_1 = H_1 \circ \hat{R}$, $\bar{H}_2 = H_2 \circ \hat{R}$. Vu le choix du nouveau repère, on a :

$$\bar{H}_1(0,0,0) = \bar{H}_2(0,0,0) = 0 \ ,$$
$$\partial_3 \bar{H}_1(0,0,0) = \partial_3 \bar{H}_2(0,0,0) = 0 \ ,$$

et la surjectivité de $D_a H$ se traduit par la condition

$$\partial_1 \bar{H}_1 \partial_2 \bar{H}_2 - \partial_1 \bar{H}_2 \partial_2 \bar{H}_1(0,0,0) \neq 0 \ .$$

Il existe alors un ouvert V de $\mathbf{R}^3$ contenant a tel qu'on puisse écrire :

$$C \cap V = \left\{ x \in \mathbf{R}^3, \bar{x}_1 = \bar{K}_1(\bar{x}_3), \bar{x}_2 = \bar{K}_2(\bar{x}_3) \right\}$$

où $\bar{K}_1$, $\bar{K}_2$ sont deux applications de classe C^1 définies sur un ouvert $\bar{U}_3$ de $\mathbf{R}$ contenant 0, et on a :

$$\bar{K}_1'(0) = \bar{K}_2'(0) = 0 \ .$$

On a donc :

$$\bar{x}_1 = o(|\bar{x}_3|), \bar{x}_2 = o(|\bar{x}_3|)$$

pour $x \in C \cap V$, d'où :

$$d(x-a, T_a C) = o(\|x-a\|) \ ,$$

caractérisation géométrique de la droite tangente.

4.4. Courbes paramétrées.

Supposons maintenant qu'au lieu d'être donnée par des équations $H_1 = H_2 = 0$, une courbe C soit obtenue comme l'image d'un paramétrage $s \mapsto \varphi(s) = (\varphi_1(s), \varphi_2(s), \varphi_3(s))$, où φ est une application de classe C^1, à valeurs dans $\mathbf{R}^3$, définie sur une partie ouverte W de $\mathbf{R}$.

Soit $s_0 \in W$; nous supposons que la différentielle $D_{s_0}\varphi$ est **injective** et voulons décrire, pour $\varepsilon > 0$ assez petit, l'image :

$$C_\varepsilon = \varphi(]s_0 - \varepsilon, s_0 + \varepsilon[) \ .$$

Dire que $D_{s_0}\varphi$ est injective revient à dire que la dérivée $\varphi'(s_0) = (\varphi_1'(s_0), \varphi_2'(s_0), \varphi_3'(s_0))$ est non nulle, c'est-à-dire qu'**au moins un** des trois nombres $\varphi_1'(s_0), \varphi_2'(s_0), \varphi_3'(s_0)$ est non nul.

Supposons **par exemple** qu'on ait $\varphi_3'(s_0) \neq 0$. Alors, d'après le théorème d'inversion locale, il existe $\varepsilon_0 > 0$ et un intervalle ouvert U_3 contenant $\varphi_3(s_0)$ tels que la restriction de φ_3 à $]s_0 - \varepsilon_0, s_0 + \varepsilon_0[$ soit un difféomorphisme de classe C^1 de cet intervalle sur U_3. On a donc :

$$C_{\varepsilon_0} = \left\{ \left(\varphi_1 \circ \varphi_3^{-1}(x_3), \varphi_2 \circ \varphi_3^{-1}(x_3), x_3\right), \ x_3 \in U_3 \right\}$$

représentation de C_{ε_0} comme graphe comme en 4.2.

On vérifie immédiatement que la droite tangente (vectorielle) à C_{ε_0} en $a = \varphi(s_0)$ est engendrée par $\varphi'(s_0)$, c'est-à-dire égale à l'image de la différentielle $D_{s_0}\varphi$.

§. 5. Autres exemples d'applications.

5.1. Zéros simples de polynômes.

Soit d un entier au moins égal à 1. Notons $\mathcal{P}_d$ l'espace vectoriel des polynômes à coefficients réels de degré au plus égal à d.

La dimension de $\mathcal{P}_d$ est $d+1$.

Soit $P_0 \in \mathcal{P}_d$; on dit qu'un nombre réel $\lambda_0 \in \mathbf{R}$ est un zéro simple de P_0 si on a :

$$P_0(\lambda_0) = 0 \ , \quad P_0'(\lambda_0) \neq 0 \ .$$

Considérons l'application H de $\mathcal{P}_d \times \mathbf{R}$ dans $\mathbf{R}$ définie par

$$H(P, \lambda) = P(\lambda) \ .$$

L'application H est de classe C^1 (elle est polynomiale par rapport à λ et aux coefficients de P), et ses différentielles partielles dans les directions de $\mathcal{P}_d$ et $\mathbf{R}$ sont respectivement, au point (P_0, λ_0) :

$$D^{(1)}_{(P_0,\lambda_0)} H(P) = P(\lambda_0)$$

(car l'application H est linéaire par rapport à P)

$$D^{(2)}_{(P_0,\lambda_0)} H(\lambda) = P_0'(\lambda_0)\lambda$$

(où P_0' est le polynôme de degré $\leq d-1$ dérivé de P_0).

Si λ_0 est un zéro simple de P_0, on a donc $H(P_0, \lambda_0) = 0$ et la différentielle partielle $D^{(2)}_{(P_0,\lambda_0)} H$ est inversible.

D'après le théorème des fonctions implicites, il existe alors une partie ouverte V de $\mathcal{P}_d$ contenant P_0, un nombre $\varepsilon > 0$, et une application de classe C^1 $\lambda : V \to \mathbf{R}$ tels que, pour $P \in V$, $\lambda(P)$ soit le seul zéro de P contenu dans l'intervalle $]\lambda_0 - \varepsilon, \lambda_0 + \varepsilon[$. On a en particulier $\lambda(P_0) = \lambda_0$. De plus, l'application $P \mapsto P'(\lambda(P))$ étant continue, on peut supposer V et ε assez petits pour qu'on ait $P'(\lambda(P)) \neq 0$ pour $P \in V$, ce qui exprime que $\lambda(P)$ est un zéro simple de P.

La différentielle en P_0 de l'application λ est donnée par :

$$D_{P_0}\lambda(Q) = -\frac{Q(\lambda_0)}{P_0'(\lambda_0)} \ .$$

5.2. Points fixes simples d'une application dépendant d'un paramètre.

Soient P, E des espaces vectoriels de dimension finie, U une partie ouverte de E, V une partie ouverte de P.

Considérons une application de classe C^1, notée H, de $V \times U$ dans E. On pense à $t \in V$ comme à un paramètre, et pour chaque valeur de $t \in V$, on note H_t l'application : $x \mapsto H(t, x)$ de U dans E.

Soient $t \in V$, $x \in U$. On dit que x est point fixe de H_t si $H_t(x) = x$, et que c'est un point fixe **simple** si de plus 1 n'est pas valeur propre de l'endomorphisme $D_x H_t \in \mathcal{L}(E)$.

Soient $t_0 \in V$, $x_0 \in U$; supposons que x_0 est un point fixe simple de H_{t_0}.

Considérons l'application K de $V \times U$ dans E définie par :

$$K(t,x) = H(t,x) - x \ .$$

Comme H est de classe C^1, K l'est aussi.

Les différentielles en un point (t,x) de H et K dans les directions de $E_1 = P$ et de $E_2 = E$ sont reliées par les formules :

$$D^{(1)}_{(t,x)}K(s) = D^{(1)}_{(t,x)}H(s)$$
$$D^{(2)}_{(t,x)}K(y) = D^{(2)}_{(t,x)}H(y) - y \ .$$

Comme x_0 est point fixe simple de H_{t_0}, on a $K(t_0,x_0) = 0$ et la différentielle partielle $D^{(2)}_{(t_0,x_0)}K \in \mathcal{L}(E)$ est injective, donc inversible.

D'après le théorème des fonctions implicites, il existe donc :

- une partie ouverte V_1 de V contenant t_0 ;
- une partie ouverte U_1 de U contenant x_0 ;
- une application de classe C^1, notée X, de V_1 dans E,

telles que pour $t \in V_1$, $X(t)$ est l'unique point fixe de H_t appartenant à U_1. (On a en particulier $X(t_0) = x_0$). De plus, comme l'application $t \mapsto \det(D_{X(t)}H_t - \mathrm{id}_E)$ est continue, le point $X(t)$ est un point fixe **simple** de H_t si l'on choisit U_1, V_1 assez petits.

CHAPITRE VIII - Différentielles d'ordre supérieur.

§. 1. Différentielle seconde.

1.1. Soient E, F des espaces vectoriels de dimension finie, U une partie ouverte de E et H une application de classe C^1 de U dans F.

La différentielle DH de H est alors l'application $a \mapsto D_aH = DH(a)$ définie dans U et à valeurs dans $\mathcal{L}(E,F)$. Cette application est continue par définition des applications de classe C^1.

DÉFINITION 1. On dit que H est de classe C^2 si l'application DH de U dans $\mathcal{L}(E,F)$ est de classe C^1.

La différentielle $D(DH)$ est alors une application continue de U dans $\mathcal{L}(E,\mathcal{L}(E,F))$.

LEMME. *Notons $\mathcal{L}^2(E,F)$ l'espace vectoriel des applications bilinéaires de $E \times E$ dans F.*

Pour $T \in \mathcal{L}(E,\mathcal{L}(E,F))$, on définit une application $\widehat{T}$ de $E \times E$ dans F par la formule $\widehat{T}(x,y) = [T(x)](y)$.

Alors, on a $\widehat{T} \in \mathcal{L}^2(E,F)$ et l'application $T \to \widehat{T}$ est un isomorphisme linéaire, dit **canonique**, *qui permet d'identifier $\mathcal{L}(E,\mathcal{L}(E,F))$ et $\mathcal{L}^2(E,F)$.*

DÉMONSTRATION : La bilinéarité de $\widehat{T}$ est évidente, ainsi que la linéarité de l'application $T \to \widehat{T}$.

Pour $S \in \mathcal{L}^2(E,F)$, $x \in E$, définissons une application S_x de E dans F par :

$$S_x(y) = S(x,y) \ .$$

On a $S_x \in \mathcal{L}(E,F)$ pour tout $x \in E$. De plus, l'application $\check{S} : x \mapsto S_x$ de E dans $\mathcal{L}(E,F)$ est linéaire. Finalement, l'application $S \mapsto \check{S}$ de $\mathcal{L}^2(E,F)$ dans $\mathcal{L}(\mathcal{L}(E,F))$ est clairement l'inverse de l'application $T \mapsto \widehat{T}$. □

DÉFINITION 2. Soit $H : U \to F$ une application de classe C^2. La différentielle $D_a(DH)$ de l'application DH en un point $a \in U$, considérée comme élément de $\mathcal{L}^2(E,F)$ suivant le lemme, sera notée D_a^2H et appelée différentielle seconde de H au point a.

Compte tenu de la proposition 3 du Chapitre V et de la proposition 1 du Chapitre VI, on obtient facilement la proposition suivante, dont la démonstration est laissée en exercice.

PROPOSITION 1. *Soient* $(e_1,\dots,e_m)$ *une base de* E, $(f_1,\dots,f_n)$ *une base de* F, H *une application de* U *dans* F, $H_1,\dots,H_n$ *ses applications coordonnées.*

Pour que H *soit de classe* C^2 *il faut et il suffit que, pour tous* $1 \le i,j \le m$, $1 \le k \le n$:

- *la dérivée partielle* $\partial_j H_k(x)$ *existe pour tout* $x \in U$;
- *la dérivée partielle* $\partial_i(\partial_j H_k)(x)$ *existe pour tout* $x \in U$;
- *l'application* $x \mapsto \partial_i(\partial_j H_k)(x)$ *est continue dans* U.

En termes de coordonnées, la différentielle seconde de H *en un point* $a \in U$ *est donnée par :*

$$D_a^2 H\left(\sum_1^m v_i e_i, \sum_1^m v'_j e_j\right) = \sum_{k=1}^n \left(\sum_{i,j=1}^m \partial_i(\partial_j H_k)(a) v_i v'_j\right) f_k \ .$$

1.2. Le lemme de Schwartz.

Soit H une application de classe C^2 de U dans F.

THÉORÈME 1 (lemme de Schwartz). *Pour tout* $a \in U$, *la différentielle seconde* $D_a^2 H \in \mathcal{L}^2(E,F)$ *est une application bilinéaire* ***symétrique***.

En d'autres termes, si $(e_1,\dots,e_m)$ *est une base de* E, *on a, pour tous* $1 \le i,j \le m$, $a \in U$:

$$\partial_i(\partial_j H)(a) = \partial_j(\partial_i H)(a) \ .$$

DÉMONSTRATION : D'après la formule pour la différentielle seconde ci-dessus, il suffit de prouver que si H est une application de classe C^2, à valeurs réelles, définie dans un ouvert U de $\mathbf{R}^2$ contenant 0, on a $\partial_1(\partial_2 H)(0) = \partial_2(\partial_1 H)(0)$.

Soit H une telle application ; soit $\delta > 0$ tel que $[-\delta,+\delta] \times [-\delta,+\delta] \subset U$.

Pour $|u| \le \delta$, posons :

$$\Delta H(u) = H(u,u) - H(u,0) - H(0,u) + H(0,0) \ .$$

On a :

$$H(u,u) - H(u,0) = u\int_0^1 \partial_2 H(u,tu)\,dt \ ,$$

$$H(0,u) - H(0,0) = u\int_0^1 \partial_2 H(0,tu)\,dt \ ,$$

$$\partial_2 H(u,tu) - \partial_2 H(0,tu) = u\int_0^1 \partial_1(\partial_2 H)(su,tu)\,ds \ ,$$

d'où

$$\Delta H(u) = u^2 \int_0^1 \int_0^1 \partial_1(\partial_2 H)(su,tu)\,ds\,dt \ .$$

De même, en échangeant le rôle des coordonnées :

$$\Delta H(u) = u^2 \int_0^1 \int_0^1 \partial_2(\partial_1 H)(su,tu)\,dt\,ds \ ,$$

d'où on conclut :

$$\lim_{\substack{u\to 0\\ u\neq 0}} \frac{1}{u^2}\Delta H(u) = \partial_1(\partial_2 H)(0,0) = \partial_2(\partial_1 H)(0,0) \ .$$

§. 2. Différentielles d'ordre supérieur.

2.1. On définit de manière récurrente la notion d'application de classe C^k, pour tout entier $k \geq 2$.

DÉFINITION 3. Soient E, F des espaces vectoriels de dimension finie, U une partie ouverte de E, H une application de U dans F, k un entier au moins égal à 2. On dit que H est de classe C^k si H est de classe C^1, et si l'application différentielle DH de U dans $\mathcal{L}(E,F)$ est de classe C^{k-1}.

REMARQUE. On appelle parfois les applications continues applications de classe C^0.

DÉFINITION 4. On dit que H est de classe C^∞ si H est de classe C^k pour tout entier $k \geq 1$.

2.2. Pour $k \geq 1$, notons $\mathcal{L}^k(E,F)$ l'espace vectoriel des applications $k-$linéaires de E^k dans F. Comme dans le lemme de 1.1, on a une identification canonique entre les espaces $\mathcal{L}(E,\mathcal{L}^{k-1}(E,F))$ et $\mathcal{L}^k(E,F)$.

Ceci permet, pour une application H de classe C^k dans un ouvert U de E, à valeurs dans F, de considérer la différentielle d'ordre k de H en un point $a \in U$, qu'on note $D^k_a H$, comme un élément de $\mathcal{L}^k(E,F)$. L'application $D^k H$ est donc une application de U dans $\mathcal{L}^k(E,F)$.

2.3. Pour qu'une application H, de U dans F soit de classe C^k, il faut et il suffit (étant données des bases $(e_1,\ldots,e_m)$ de E et $(f_1,\ldots,f_n)$ de F) que pour tout $1 \leq s \leq n$, tout $1 \leq \ell \leq k$, tous entiers $i_1,\ldots,i_\ell$ dans $[1,m]$, la dérivée partielle $\partial_{i_1}(\partial_{i_2}(\ldots\partial_{i_\ell} H_s))$ existe en tout point x de U et dépende continûment du point x.

2.4. Soient k un entier au moins égal à 2, H une application de classe C^k de U dans F.

Soient $(e_1,\ldots,e_m)$ une base de E, ℓ un entier au plus égal à k, $i_1,\ldots,i_\ell$ des entiers dans $[1,m]$.

Pour toute permutation σ de $\{1,\ldots,\ell\}$ et tout point a de U, on a :

$$\partial_{i_1}(\partial_{i_2}\ldots(\partial_{i_\ell}H)(a)) = \partial_{i_{\sigma(1)}}\left(\partial_{i_{\sigma(2)}}\cdots\left(\partial_{i_{\sigma(\ell)}}H\right)(a)\right) .$$

Cela résulte du lemme de Schwartz, car toute permutation est un produit de transpositions.

Cela signifie que la différentielle $D_a^\ell H$ d'ordre ℓ de H en a est une application ℓ-linéaire **symétrique** de E^ℓ dans F.

REMARQUE. Une application ℓ-linéaire **symétrique** Q de E^ℓ dans F est entièrement déterminée par l'application $x \mapsto Q(x,x,\ldots,x)$.

§. 3. Propriété des applications de classe C^k.

Soient E, F des espaces vectoriels de dimension finie, U une partie ouverte de E, k un entier au moins égal à 1.

3.1. Si H_1, H_2 sont des applications de classe C^k de U dans F, alors $H_1 + H_2$ est aussi de classe C^k et on a :

$$D^k(H_1 + H_2) = D^k H_1 + D^k H_2 .$$

3.2. Une application affine $H(x) = u(x) + b$ $(u \in \mathcal{L}(E,F),\ b \in F)$ est de classe C^∞ et on a

$$D_x H = u \ , \quad \forall x \in E$$
$$D_x^2 H = 0 \ , \quad \forall x \in E .$$

(Inversement, une application H de classe C^2 de E dans F telle que $D_x^2 H = 0$ pour tout $x \in E$ est affine : exercice).

3.3. Supposons que $E = E_1 \times E_2$, et que H soit une application bilinéaire de $E_1 \times E_2$ dans F. On a vu que

$$D_{(a_1,a_2)}H(x_1,x_2) = H(a_1,x_2) + H(x_1,a_2) ,$$

par conséquent DH est linéaire et on a :

$$D^2_{(a_1,a_2)}H((x_1,x_2),(y_1,y_2)) = H(y_1,x_2) + H(x_1,y_2) ,$$
$$D^3_{(a_1,a_2)}H = 0 ,$$

donc H est de classe C^∞.

3.4. Plus généralement, toute application polynômiale de E dans F est de classe C^∞.

[On dit qu'une application H de E dans F est polynomiale si étant données des bases $(e_1, \ldots, e_m)$ de E et $(f_1, \ldots, f_n)$ de F, chaque composante H_ℓ de H $(1 \leq \ell \leq n)$ s'exprime comme un polynôme de m variables) en fonction des coordonnées $x_1, \ldots, x_m$ du point $x \in E$].

3.5. Soient G un troisième espace vectoriel de dimension finie, et V une partie ouverte de F. Soient H une application de classe C^k de U dans V, et K une application de classe C^k de V dans G. Alors $K \circ H$ est de classe C^k dans U.

On montre ceci par récurrence sur k, le cas $k = 1$ ayant déjà été traité. On a :

$$D_x(K \circ H) = D_{H(x)}K \circ D_x H \ .$$

Les applications $x \mapsto D_x H$ et $y \mapsto D_y K$ sont de classe C^{k-1} dans U, V respectivement. Par hypothèse de récurrence, l'application $x \mapsto D_{H(x)}K$ est de classe C^{k-1} dans U. Finalement, l'application $(u, v) \mapsto v \circ u$ de $\mathcal{L}(E, F) \times \mathcal{L}(F, G)$ dans $\mathcal{L}(E, G)$ est bilinéaire donc de classe C^∞. On conclut que l'application $x \mapsto D_x(K \circ H)$ est de classe C^{k-1}, donc que $K \circ H$ est de classe C^k.

Il existe une formule pour $D^k(K \circ H)$ (formule de Faa-di-Bruno), mais celle-ci est assez compliquée.

Exercice. Calculer $D^2(K \circ H)$.

3.6. Si $H_1 : U \to \mathbb{R}$ et $H_2 : U \to F$ sont de classe C^k, alors $H_1 H_2$ est de classe C^k.

Cela résulte de 3.4 et 3.5 (le produit est une application bilinéaire).

On a un résultat analogue pour d'autres espèces de produits (produit scalaire, produit vectoriel dans $\mathbb{R}^3, \ldots$).

3.7. Supposons que H soit un difféomorphisme de classe C^1 de U sur une partie ouverte V de F (E et F ont alors même dimension) et que H soit une application de classe C^k.

La bijection réciproque $H^{-1} : V \to U$ est alors de classe C^k. En effet, pour $y \in V$, on a :

$$D_y H^{-1} = \left[D_{H^{-1}(y)} H\right]^{-1} \ .$$

Soit $\mathcal{U}$ la partie ouverte de $\mathcal{L}(E, F)$ formée des applications inversibles. L'application $u \mapsto u^{-1}$ de $\mathcal{U}$ dans $\mathcal{L}(F, E)$ est de classe C^∞ (exercice, à partir du Chapitre V, §.3.5). On montre alors par récurrence sur ℓ que H^{-1} est de classe C^ℓ pour $1 \leq \ell \leq k$: on sait déjà que H^{-1} est de classe C^1 ; si H^{-1} est de classe C^ℓ, avec $\ell < k$, alors l'application

$y \mapsto \left(D_{H^{-1}(y)}H\right)^{-1}$ est composée d'une application de classe C^ℓ, d'une application de classe C^{k-1} et d'une application de classe C^∞. Elle est donc de classe C^ℓ et H^{-1} est de classe $C^{\ell+1}$.

Dans la situation précédente, on dit que H est un difféomorphisme de classe C^k de U sur V.

3.8. Reprenons les notations du théorème des fonctions implicites (Chapitre VII, §.2).

Si H est une application satisfaisant aux hypothèses de ce théorème, et si de plus H est une application de classe C^k, alors l'application K obtenue dans la conclusion du théorème est aussi de classe C^k.

Cela résulte de la formule pour la différentielle de K, par une démonstration extrêmement semblable à celle de 3.7 (exercice).

§. 4. La formule de Taylor.

Soient E, F des espaces vectoriels de dimension finie, U une partie ouverte de E, k un entier au moins égal à 1.

Soient a_0, a_1 des points de U. On **suppose** que le segment $[a_0, a_1]$ est contenu dans U, c'est-à-dire que pour tout $t \in [0,1]$ le point $a_t = a_0 + t(a_1 - a_0)$ appartient à U. On pose $v = a_1 - a_0$.

THÉORÈME 2 (formule de Taylor). *Soit $H : U \to F$ une application de classe C^k. On a :*

$$H(a_1) = H(a_0) + D_{a_0}H(v) + \frac{1}{2!}D^2_{a_0}H(v,v) + \dots$$
$$+ \frac{1}{(k-1)!}D^{k-1}_{a_0}H(v,\dots,v) + \frac{1}{(k-1)!}\int_0^1 (1-t)^{k-1} D^k_{a_t}H(v,\dots,v)\,dt \;.$$

DÉMONSTRATION : Il suffit de démontrer la formule précédente pour chacune des composantes (dans une base de F) de l'application H. On peut donc supposer $F = \mathbb{R}$.

Pour $0 \leq t \leq 1$, posons :

$$h(t) = H(a_0 + tv) \;.$$

L'application h est de classe C^k et on a :

$$h'(t) = D_{a_t}H(v) \;,$$
$$h''(t) = D^2_{a_t}H(v,v) \;,$$
$$h^{(k)}(t) = D^k_{a_t}H(v,\dots,v) \;,$$

or la formule de Taylor pour h à l'ordre k s'écrit :

$$h(1) = h(0) + h'(0) + \frac{h''(0)}{2!} + \dots + \frac{h^{(k-1)}(0)}{(k-1)!} + \frac{1}{(k-1)!}\int_0^1 (1-t)^{k-1} h^{(k)}(t)\,dt$$

d'où par substitution la formule du théorème. □

COROLLAIRE (Développement de Taylor à l'ordre k).

Soient $H : U \to F$ une application de classe C^k et $a \in U$. On a :

$$H(a+v) = H(a) + D_aH(v) + \frac{1}{2!}D_a^2H(v,v) + \dots$$
$$+ \frac{1}{k!}D_a^kH(v,\dots,v) + o(\|v\|^k) \ .$$

DÉMONSTRATION : Il s'agit de voir qu'on a :

$$\frac{1}{(k-1)!}\int_0^1 (1-t)^{k-1} D_{a+tv}^kH(v,v,\dots,v) = \frac{1}{k!}D_a^kH(v,\dots,v) + o(\|v\|^k) \ .$$

Soit $\varepsilon > 0$. Comme l'application $x \mapsto D_x^kH$ de U dans $\mathcal{L}^k(E,F)$ est continue, il existe $\delta > 0$ tel qu'on ait, pour $\|v\| < \delta$:

$$\left\|D_{a+v}^kH - D_a^kH\right\| \leq \varepsilon \ .$$

On aura alors, pour $\|v\| < \delta$, $t \in [0,1]$:

$$\left\|D_{a+tv}^kH(v,\dots,v) - D_a^kH(v,\dots,v)\right\| \leq \varepsilon\,\|v\|^k \ .$$

Comme on a par ailleurs $\displaystyle\int_0^1 (1-t)^{k-1}\,dt = \frac{1}{k}$, on obtient finalement, pour $\|v\| < \delta$:

$$\left\|\frac{1}{(k-1)!}\int_0^1 (1-t)^{k-1} D_{a+tv}^kH(v,\dots,v)\,dt - \frac{1}{k!}D_a^kH(v\dots v)\right\| \leq \frac{\varepsilon}{k!}\,\|v\|^k \ ,$$

d'où le corollaire.

CHAPITRE IX - Points critiques et extrema.

On s'intéresse dans ce chapitre à la situation suivante : Soient E un espace vectoriel de dimension finie, U une partie ouverte de E, H une application de classe C^1 de U <u>**dans R**</u>. Soit $a \in U$. On se propose de décrire l'allure de la fonction H au voisinage de a.

§. 1. Points réguliers et critiques.

1.1. La différentielle de H au point a est une forme linéaire de E dans $\mathbf{R}$. Elle est identiquement nulle ou surjective.

DÉFINITION 1. On dit que a est **point régulier** de H si $D_a H \neq 0$, **point critique** dans le cas contraire.

En termes de dérivées partielles par rapport à une base $(e_1, \ldots, e_m)$ de E, le point a est critique si on a $\partial_1 H(a) = \partial_2 H(a) = \ldots = \partial_m H(a) = 0$, régulier s'il existe $1 \leq i \leq m$ tel que $\partial_i H(a) \neq 0$.

1.2. Allure de H aux points réguliers.

Soit $a \in U$ un point régulier de H.

Par définition de la différentielle, on a, pour v asez petit :

$$H(a+v) = H(a) + D_a H(v) + o(\|v\|) .$$

On peut, en se plaçant dans un système de coordonnées adéquat, simplifier cette expression de H.

Comme a est régulier, le noyau de la forme linéaire $D_a H$ est un sous-espace de dimension $n-1$.

On peut donc choisir une base de E $(e_1, \ldots, e_m)$ telle que $(e_2, \ldots, e_m)$ soit une base de $\operatorname{Ker} D_a H$, et qu'on ait $D_a H(e_1) = 1$. En d'autres termes, dans cette base on a :

$$\partial_1 H(a) = 1 \ , \quad \partial_2 H(a) = \ldots \partial_m H(a) = 0 \ ,$$

et la formule ci-dessus s'écrit :

$$H(a+v) = H(a) + v_1 + o(\|v\|)$$

(avec $v = \sum_{i=1}^{m} v_i e_i$).

Un changement de coordonnées **non linéaire** permet de simplifier encore l'expression de H. Considérons en effet l'application L de U dans E donnée (dans la base précédente) par :

$$L(x_1, \ldots, x_m) = (H(x) - H(a) + a_1, x_2, \ldots, x_m) \ .$$

On a $L(a) = a$ et L est une application de classe C^1 dans U, vérifiant $D_a L = \mathrm{id}_E$. D'après le théorème d'inversion locale, il existe des parties ouvertes V_1, V_2 de E contenant a, avec $V_1 \subset U$, telles que la restriction de L à V_1 soit un difféomorphisme de classe C^1 de V_1 sur V_2.

On considère L comme définissant un **changement de coordonnées** (non linéaire) dans l'ouvert V_1 : les anciennes coordonnées $(x_1, \ldots, x_m)$ et les nouvelles coordonnées $(y_1, \ldots, y_m)$ d'un point $x \in V_1$ sont reliées par les formules

$$y_1 = H(x) - H(a) + a_1\,,\ y_2 = x_2, \ldots, y_m = x_m\ ,$$

et dire que L est un difféomorphisme de classe C^1 revient à dire qu'on peut aussi exprimer les anciennes coordonnées $(x_1, \ldots, x_m)$ en fonction des nouvelles $(y_1, \ldots, y_m)$ par des applications de classe C^1.

Dans les nouvelles coordonnées $(y_1, \ldots, y_m)$, on a évidemment, pour $y = (y_1, \ldots, y_m) \in V_2$:

$$H(y) = H(a) + y_1 - a_1\ ,$$

c'est-à-dire qu'on a (au prix d'un changement non linéaire de coordonnées) éliminé le reste $o(\|y - a\|)$ dans les expressions précédentes.

§. 2. Matrice hessienne.

2.1. Nous supposons maintenant que H est de classe C^2, et que $a \in U$ est un point **critique** de H.

On a donc $D_a H = 0$, donc

$$H(a + v) = H(a) + o(\|v\|)\ ,$$

une expression qui n'est pas suffisante pour décrire H au voisinage de a.

Le développement de Taylor à l'ordre 2 s'écrit :

$$H(a + v) = H(a) + \frac{1}{2} D_a^2 H(v, v) + o(\|v\|^2)\ .$$

La forme bilinéaire **symétrique** $(v, w) \mapsto D_a^2 H(v, w)$ (ou la forme quadratique associée $v \mapsto D_a^2 H(v, v)$) s'appelle forme **hessienne** de H au point critique a.

Dans une base $(e_1, \ldots, e_m)$ de E, la matrice **symétrique** associée à la forme hessienne est :

$$\Big(D_a^2 H(e_i, e_j)\Big)_{1 \leq i,j \leq m} = \Big(\partial_i \partial_j H(a)\Big)_{1 \leq i,j \leq m}$$

et s'appelle **matrice hessienne** de H en a.

2.2. Rappels sur les formes quadratiques.

Soit Q une forme quadratique dans un espace de dimension finie E, B la forme bilinéaire symétrique associée à Q. On a donc

$$\begin{aligned} Q(x) &= B(x,x) \\ B(x,y) &= \frac{1}{2}\,[Q(x+y)-Q(x)-Q(y)] \\ &= \frac{1}{4}\,[Q(x+y)-Q(x-y)]\ . \end{aligned}$$

Le **noyau de B** (ou Q) est par définition l'ensemble des $x \in E$ tels que $B(x,y) = 0$ pour tout $y \in E$. C'est un sous-espace vectoriel de E. On dit que B (ou Q) est **non dégénérée** si son noyau est $\{0\}$.

On dit que Q (ou B) est **positive** si $Q(x) \geq 0$ pour tout $x \in E$, **négative** si $Q(x) \leq 0$ pour tout $x \in E$, **indéfinie** si Q prend des valeurs strictement positives et strictement négatives.

Une forme non dégénérée et positive est dite **définie positive**; cela revient à dire qu'on a $Q(v) > 0$ pour tout $v \in E$, $v \neq 0$.

De même une forme non dégénérée et négative est dite **définie négative.**

EXEMPLE. Dans $\mathbb{R}^2$ (coordonnées en x_1, x_2)

$Q_0(x_1,x_2) = x_1^2 + x_2^2$	est définie positive,
$Q_1(x_1,x_2) = -x_1^2 - 2x_2^2$	est définie négative,
$Q_2(x_1,x_2) = x_1^2 - x_2^2$	est non dégénérée mais indéfinie,
$Q_3(x_1,x_2) = x_1^2$	est positive mais dégénérée.

Soit $(e_1,\dots,e_m)$ une base de E. La matrice associée à B (ou Q) est

$$M = \Big(B(e_i,e_j)\Big)_{1\leq i,j\leq m}$$

Cette matrice est **symétrique** et on a pour des vecteurs $x = \sum x_i e_i$, $y = \sum y_i e_i$:

$$B(x,y) = {}^t x M y = {}^t y M x$$

(où x, y dans la formule sont des vecteurs colonnes, et ${}^t x, {}^t y$ sont les vecteurs lignes correspondants).

Les valeurs propres d'une matrice symétrique sont **toutes réelles.** On a :

Q non dégénérée	$\Leftrightarrow$	0 n'est pas valeur propre de $M \Leftrightarrow \det M \neq 0$
Q positive	$\Leftrightarrow$	les valeurs propres de M sont positives ou nulles
Q négative	$\Leftrightarrow$	les valeurs propres de M sont négatives ou nulles
Q définie positive	$\Leftrightarrow$	les valeurs propres de M sont strictement positives
Q définie négative	$\Leftrightarrow$	les valeurs propres de M sont strictement négatives
Q indéfinie	$\Leftrightarrow$	M a des valeurs propres strictement positives et des valeurs propres strictement négatives.

Notons p le nombre de valeurs propres strictement positives de M (comptées avec multiplicité), q le nombre de valeurs propres strictement négatives de M. Le triplet $(p, q, m-p-q)$ s'appelle la **signature** de la forme quadratique Q. Il ne dépend pas de la base de E considérée. En effet :

– l'entier p est la dimension maximale d'un sous-espace E_1 de E sur lequel la forme Q soit définie positive;

– l'entier q est la dimension maximale d'un sous-espace E_2 de E sur lequel la forme Q soit définie négative;

– l'entier $m-p-q$ est la dimension du noyau de B (ou Q).

Il existe une base $(e_1, \ldots, e_m)$ de E telle que, si l'on note $x_1, \ldots, x_m$ les coordonnées associées, Q s'exprime sous la forme :

$$Q(x) = \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{p+q} x_i^2 ,$$

le noyau de Q étant alors engendré par $e_{p+q+1}, \ldots, e_m$.

2.3. Points critiques non dégénérés.

DÉFINITION 2. Le point critique a de H est **non dégénéré** (ou **de Morse**) si la forme hessienne $D_a^2 H$ est non dégénérée.

Par rapport à une base $(e_1, \ldots, e_m)$ de E, le point critique a de H est donc non dégénéré si et seulement si on a :

$$\det\Big(\partial_i \partial_j H(a)\Big)_{1 \leq i,j \leq m} \neq 0 \ .$$

On étudie au §.3 quelques relations entre la signature de la forme hessienne en a et l'allure de H au voisinage de a. On ne peut s'empêcher de mentionner le résultat fondamental (hors programme) suivant :

THÉORÈME (Lemme de Morse). *Soient $H : U \to \mathbf{R}$ une fonction de classe C^2. Soit $a \in U$ un point critique* ***non dégénéré*** *de H. Notons $(p, q, 0)$ la signature de la forme hessienne de H en a (on a $p + q = \dim E = m$). Il existe alors, dans un ouvert V_1 contenant a, un système de coordonnées (non linéaire) $y_1, \ldots, y_m$ qui se déduit de l'ancien système par un difféomorphisme de classe C^1, et dans lequel H s'exprime comme suit :*

$$H(y) = H(a) + \sum_{i=1}^{p} y_i^2 - \sum_{i=p+1}^{p+q} y_i^2 \ .$$

Ce résultat est à rapprocher de celui discuté en §.1.2 : en un point critique non dégénéré, on peut, par un changement de coordonnées adéquat, supprimer le reste dans le développement de Taylor à l'ordre 2.

§. 3. Extrema.

3.1. DEFINITION 3. Un point $a \in U$ est un **maximum local** [resp. **minimum local**] de la fonction $H : U \to \mathbf{R}$ s'il existe une partie ouverte V contenant a telle qu'on ait :

$$H(x) \leq H(a) \qquad \text{pour } x \in V$$
$$[\text{resp. } H(x) \geq H(a) \qquad \text{pour } x \in V] .$$

Si on a la relation plus forte :

$$H(x) < H(a) \quad \text{pour } x \in V, \quad x \neq a$$
$$[\text{resp. } H(x) > H(a) \quad \text{pour } x \in V, \quad x \neq a] ,$$

on dit que a est un maximum local **strict** [resp. minimum local **strict**] de H.

On dit que a est un **extremum local** de H si c'est un maximum ou minimum local, un **extremum local strict** si c'est un maximum ou minimum local strict.

3.2. La proposition suivante donne des conditions **nécessaires** pour qu'un point $a \in U$ soit un extremum local.

PROPOSITION 1. *Soient $H : U \to \mathbf{R}$ une fonction de classe C^2 et a un extremum local de H.*

*Alors a est un **point critique** de H. De plus la forme hessienne de H en a est **positive** si a est un minimum local, **négative** si a est un maximum local.*

COROLLAIRE. *Si a est un point **régulier** de H, ou un point critique dont la forme hessienne est **indéfinie**, alors a n'est pas un extremum local de H.*

DÉMONSTRATION DE LA PROPOSITION : Supposons que a soit un point régulier de H.

Il existe alors $v \in E$ tel que $D_aH(v) > 0$. Le développement de Taylor à l'ordre 1 montre qu'on a, pour $t > 0$ suffisamment petit :

$$H(a + tv) = H(a) + tD_aH(v) + o(t) > H(a) ,$$
$$H(a - tv) = H(a) - tD_aH(v) + o(t) < H(a) ,$$

donc a n'est pas un extremum local de H.

Supposons maintenant que a soit un point critique de H, mais que la forme hessienne D_a^2H de H en a ne soit pas positive. Il existe donc un vecteur $v \in E$ tel que

$$D_a^2H(v,v) < 0 .$$

Pour $|t|$ suffisamment petit, on a, d'après le développement de Taylor à l'ordre 2 :

$$H(a + tv) = H(a) + \frac{t^2}{2} D_aH(v,v) + o(t^2) < H(a)$$

$(t \neq 0)$, donc a n'est pas un minimum local de H.

De même, on montre que si la hessienne n'est pas négative, a n'est pas un maximum local de H. □

3.3. La proposition suivante donne maintenant des conditions **suffisantes** pour qu'un point soit extremum local (strict).

PROPOSITION 2. *Soient $H : U \to \mathbf{R}$ une fonction de classe C^2 et $a \in U$ un point critique de H.*

*Si la forme hessienne de H en a est **définie positive**, alors a est un **minimum local strict** de H. Si la forme hessienne de H en a est **définie négative**, alors a est un **maximum local strict** de H.*

DÉMONSTRATION : Supposons que la forme hessienne D_a^2H de H en a soit définie positive. Munissons E d'une norme $\| \ \|$ et considérons la sphère $S = \{x, \|x\| = 1\}$. Alors S est compacte, et on a $D_a^2H(v,v) > 0$ pour $v \in S$; donc il existe $\Delta > 0$ tel qu'on ait $D_a^2H(v,v) \geq \Delta$ pour $v \in S$. On a alors $D_a^2H(v,v) \geq \Delta \ \|v\|^2$ pour tout $v \in E$, donc on obtient, dans le développement de Taylor à l'ordre 2 :

$$\begin{aligned} H(a+v) &= H(a) + D_a^2H(v,v) + o(\|v\|^2) \\ &\geq H(a) + \Delta \ \|v\|^2 + o(\|v\|^2) \ ; \end{aligned}$$

on conclut que a est un minimum local strict de H.

Le cas où D_a^2H est définie négative se traite de manière analogue.

3.4. Recherche d'extrema.

Soit $H : U \to \mathbf{R}$ une application de classe C^2. On recherche les extrema locaux de H, et on cherche à déterminer leur type.

On commence par déterminer l'ensemble $\sum$ des **points critiques** de H, c'est-à-dire qu'on résout le système d'équations :

$$\partial_1 H(x) = \partial_2 H(x) = \ldots = \partial_m H(x) = 0$$

(dans une base $(e_1, \ldots, e_m)$ de E).

On sait en effet que tout extremum local appartient à $\sum$ (proposition 3.2).

Pour chaque point a de $\sum$, on calcule la matrice hessienne $(\partial_i \partial_j H(a))_{1 \leq i,j \leq m}$ et la forme hessienne D_a^2H correspondante.

Si la forme hessienne est **indéfinie**, a n'est pas extremum local de H.

Si la forme hessienne est **définie positive**, a est un minimum local strict de H.

Si la forme hessienne est **définie négative**, a est maximum local strict de H.

Si la forme hessienne est **positive**, ou **négative**, mais **dégénérée**, seule une étude plus fine peut permettre de conclure.

EXEMPLE. Considérons les fonctions

$$H_+(x,y) = x^2 + y^4 \ ,$$
$$H_0(x,y) = x^2 \ ,$$
$$H_-(x,y) = x^2 - y^4 \ .$$

Le point $(0,0)$ est un point critique de H_+, H_0 et H_-. Les formes hessiennes de H_+, H_0 et H_- sont toutes trois égales à :

$$Q(x,y) = x^2 \ ,$$

qui est une forme quadratique **positive** mais **dégénérée.** On voit facilement que $(0,0)$:

- est un minimum local strict de H_+,
- est un minimum local, non strict, de H_0,
- n'est pas un extremum local de H_-.

§. 4. Extrema liés.

4.1. On considère dans ce paragraphe la situation suivante. Soient E un espace vectoriel, de dimension finie m, et U une partie ouverte de E.

Soient $F_1, \ldots, F_k$ des fonctions de classe C^2 de U dans $\mathbf{R}$. On note F l'application de classe C^2 de U dans $\mathbf{R}^k$ dont les composantes sont $F_1, \ldots, F_k$.

Soit M la partie de U d'équation $F = 0$, c'est-à-dire :

$$M = \Big\{ x \in U \, , \ F_1(x) = F_2(x) = \ldots = F_k(x) = 0 \Big\} \ .$$

Etant donnée une application H de classe C^2 de U dans $\mathbf{R}$, on cherche à déterminer les extrema locaux de la fonction $x \mapsto H(x)$ **lorsque x est astreint à varier dans M.**

DÉFINITION 4. Un point $a \in M$ est un **minimum local de H sur M** s'il existe un ouvert $V \subset U$ contenant a tel que

$$H(x) \geq H(a) \ , \ \text{pour } x \in V \cap M \ .$$

C'est un **minimum local strict de H sur M** si on a la relation plus forte :

$$H(x) > H(a) \ , \ \text{pour } x \in V \cap M, \ x \neq a \ .$$

On définit de façon similaire un **maximum local**, et un **maximum local strict** de H sur M.

4.2. Points réguliers de M.

Les techniques de calcul différentiel introduites jusqu'ici ne permettent d'analyser la fonction H qu'au voisinage des points **réguliers** de M, définis ci-après.

PROPOSITION 3. *Soit $a \in M$. Les propriétés suivantes sont équivalentes :*

(i) *La différentielle D_aF est surjective;*

(ii) *Les formes linéaires $D_aF_1, \ldots, D_aF_k$ sont linéairement indépendantes;*

(iii) *la dimension de $\mathrm{Ker}\,(D_aF) = \bigcap_{i=1}^{k} \mathrm{Ker}\,(D_aF_i)$ est $m - k$ (ou $m = \dim E$).*

*Si l'une (donc les trois) de ces conditions est satisfaite, on dit que a est un point **régulier** de M et on appelle $\mathrm{Ker}\,(D_aF) = \bigcap_{i=1}^{k} \mathrm{Ker}\,(D_aF_i)$ le **sous-espace** (vectoriel) **tangent** à M en a ; on le note T_aM.*

*Les formes linéaires $D_aF_1, \ldots, D_aF_k$ forment alors une **base** de l'espace des formes linéaires de E dans $\mathbf{R}$ qui s'annulent sur T_aM.*

DÉMONSTRATION : On a $F = (F_1, \ldots, F_k)$, donc

$$D_aF = (D_aF_1, \ldots, D_aF_k) \text{ et } \mathrm{Ker}\,(D_aF) = \bigcap_{i=1}^{k} \mathrm{Ker}\,(D_aF_i) .$$

Une application linéaire $u \in \mathcal{L}(E, \mathbf{R}^k)$ est surjective si et seulement si son noyau $\mathrm{Ker}\,u$ est de dimension $m - k$. Donc (i) est équivalent à (iii). Si D_aF n'est pas surjective, il existe des nombres réels $\lambda_1, \ldots, \lambda_k$ **non tous nuls** tels que l'image de D_aF soit contenue dans l'hyperplan de $\mathbf{R}^k$ d'équation $\sum_{i=1}^{k} \lambda_i x_i = 0$. Vu la formule de D_aF, cela signifie qu'on a, pour tout $v \in E$:

$$\sum_{i=1}^{k} \lambda_i D_aF_i(v) = 0 ,$$

d'où

$$\sum_{i=1}^{k} \lambda_i D_aF_i = 0$$

et les formes linéaires $D_a F_1, \dots, D_a F_k$ ne sont pas linéairement indépendantes.

Inversement, si ces formes ne sont pas linéairement indépendantes, il existe $\lambda_1, \dots, \lambda_k$ non tous nuls tels que $\sum_{i=1}^{k} \lambda_i D_a F_i = 0$, et l'image de $D_a F$ est contenue dans l'hyperplan d'équation $\sum_{i=1}^{k} \lambda_i x_i = 0$. On a donc (i)$\Leftrightarrow$(ii), ce qui termine la démonstration de l'équivalence des trois conditions.

Supposons que a est un point régulier de M. On a donc

$$\dim (T_a M) = \dim (\operatorname{Ker} D_a F) = m - k\ .$$

Soit E' le dual de E, c'est-à-dire l'espace vectoriel des formes linéaires de E dans $\mathbf{R}$. Si E_1 est un sous-espace vectoriel de E, on note E_1^0, et on appelle orthogonal de E_1 dans E', le sous-espace vectoriel de E' formé des formes linéaires de E dans $\mathbf{R}$ qui s'annulent sur E_1.

On rappelle que la dimension de E_1^0 est alors égale à $\dim E - \dim E_1$.

Prenons ici $E_1 = T_a M = \operatorname{Ker}(D_a F) = \bigcap_{i=1}^{k} \operatorname{Ker}(D_a F_i)$. On a $\dim E_1 = m - k$, donc $\dim E_1^0 = k$.

Or les formes linéaires $D_a F_1, \dots, D_a F_k$ sont linéairement indépendantes, et appartiennent à E_1^0. Elles forment donc une base de E_1^0. □

PROPOSITION 4. *Les points réguliers de M forment une partie ouverte de M.*

DÉMONSTRATION : L'application $x \mapsto D_x F$ de M dans $\mathcal{L}(E, \mathbf{R}^k)$ est continue puisque F est de classe C^2.

La proposition résulte donc de la caractérisation (i) des points réguliers et du lemme suivant :

LEMME. *Soient E, F des espaces vectoriels de dimension finie. Dans $\mathcal{L}(E, F)$, l'ensemble des applications linéaires surjectives est ouvert.*

DÉMONSTRATION : Posons $m = \dim E$, $k = \dim F$. Si $k > m$ aucune application linéaire de E dans F n'est surjective. Si $k \leq m$, une matrice à k lignes et m colonnes représente (dans des bases fixées de E et F) une application linéaire surjective si et seulement si l'un des $k \times k$ mineurs est non nul. D'où le lemme.

4.3. Points critiques relatifs de H sur M.

DÉFINITION 5. Soit a un point régulier de M. On dit que a est un **point critique relatif de H sur M** si le noyau de $D_a H$ contient le sous-espace tangent $T_a M$ en a à M.

Soit a un point critique relatif de H sur M. Comme a est régulier, les formes linéaires $D_aF_1, \ldots, D_aF_k$ sont linéairement indépendantes. Comme D_aH s'annule sur T_aM, il existe, d'après la proposition 3, un **unique** k-uplet $(\lambda_1, \ldots, \lambda_k)$ tel que l'on ait :

$$D_aH = \sum_{i=1}^{k} \lambda_i D_a F_i \; .$$

Les nombres réels $\lambda_1, \ldots, \lambda_k$ déterminés par cette relation s'appellent les **multiplicateurs de Lagrange** de H au point critique relatif a.

DÉFINITION 6. Soit a un point critique relatif de H sur M. La restriction au sous-espace tangent T_aM de la forme quadratique $D_a^2H - \sum_{i=1}^{k} \lambda_i D_a^2 F_i$, considérée comme forme quadratique sur T_aM, s'appelle **forme hessienne relative de H sur M**.

4.4. Extrema de H sur M.

PROPOSITION 5. *Soit a un point régulier de M.*

(1) *Si a est un minimum local (resp. maximum local) de H sur M, alors a est un point critique relatif de H et la hessienne relative est positive (resp. négative).*

(2) *Si a est un point critique relatif de H sur M, de hessienne relative définie positive (resp. définie négative) alors a est un minimum local strict (resp. maximum local strict) de H sur M.*

DÉMONSTRATION : Translatons l'origine des coordonnées, de façon à avoir $a = 0$. Choisissons un supplémentaire E_2 de $E_1 = T_aM$ et identifions E à $E_1 \times E_2$. On a $\dim E_1 = k$ et $E_2 \cap \operatorname{Ker} D_aF = \{0\}$, donc la différentielle de F dans la direction de E_2 est inversible. D'après le théorème des fonctions implicites, il existe des ouverts V_1, V_2 de E_1, E_2 respectivement, contenant 0, et une application K de classe C^2 de V_1 dans E_2 tels que :

$$M \cap (V_1 \times V_2) = \Big\{(x, K(x)), x \in V_1\Big\} \;,$$
$$K(0_{E_1}) = O_{E_2} \;,$$
$$D_0K = 0 \; .$$

Posons $L(x) = (x, K(x)) \in M$, pour $x \in V_1$. Alors L est un homéomorphisme de V_1 sur $M \cap (V_1 \times V_2)$, donc 0 est un minimum local (resp. maximum local, resp. minimum local strict, resp. maximum local strict) de H sur M si et seulement si 0 est un minimum local (resp. maximum local, resp. minimum local strict, resp. maximum local strict) de $\widetilde{H} = H \circ L$. Or on a :

$$D_0\widetilde{H} = D_0H \circ D_0L = D_0H \Big/ E_1 \;,$$

(D_0L est l'injection de E_1 dans $E_1 \times E_2 = E$) donc si 0 est extremum local de H sur M, alors 0 est point critique relatif de H sur M.

Supposons maintenant que H est un point critique relatif de H sur M. Notons $\lambda_1, \ldots, \lambda_k$ les multiplicateurs de Lagrange correspondants. Posons :

$$H_1 = H - \sum \lambda_i F_i ,$$
$$\widetilde{H}_1 = H_1 \circ L .$$

Les fonctions H et H_1 coïncident dans M, et la hessienne relative de H en 0 est $D_0^2 H_1 \big/ E_1$. On a, pour $x \in V_1$:

$$D_x \widetilde{H}_1 = D_{L(x)} H_1 \circ D_x L$$

d'où (comme $D_0 H_1 = 0$) :

$$D_0^2 \widetilde{H}_1 = D_0^2 H_1 \circ (D_0 L, D_0 L) ,$$

c'est-à-dire que la hessienne relative de H en 0 est précisément $D_0^2 \widetilde{H}_1$. Comme on a vu plus haut qu'il revient au même d'étudier H_1 dans M ou $\widetilde{H}_1$ dans V_1, et que H et H_1 coïncident dans M, les conclusions de la proposition résultent des propositions 1 et 2, appliquées à $\widetilde{H}_1$.

4.5. Exemple.

Dans $\mathbf{R}^2$, soient C_1 le cercle d'équation $x^2 + y^2 = 1$, C_2 le cercle d'équation $(x - c)^2 + y^2 = R^2$. On suppose que $c > 0$ et $R > 1 + c$ (de sorte que C_1 est intérieur à C_2) et on cherche à déterminer les extrema locaux du carré de la distance euclidienne entre un point de C_1 et un point de C_2.

On définit donc :

$$M = C_1 \times C_2 \subset \mathbf{R}^2 \times \mathbf{R}^2 ,$$

on note (x_1, y_1, x_2, y_2) les coordonnées de $\mathbf{R}^2 \times \mathbf{R}^2$, de sorte que l'équation de M est :

$$\begin{cases} F_1(x_1, y_1, x_2, y_2) = x_1^2 + y_1^2 - 1 = 0 \\ F_2(x_1, y_1, x_2, y_2) = (x_2 - c)^2 + y_2^2 - R^2 = 0 . \end{cases}$$

On cherche les points critiques et les extrema locaux sur M de la fonction :

$$H(x_1, x_2, y_1, y_2) = (x_1 - x_2)^2 + (y_1 - y_2)^2 .$$

En un point $z = (a_1, b_1, a_2, b_2)$ de $C_1 \times C_2$, on a :

$$D_z F_1(x_1, y_1, x_2, y_2) = 2a_1 x_1 + 2b_1 y_1 ,$$
$$D_z F_2(x_1, y_1, x_2, y_2) = 2(a_2 - c)x_2 + 2b_2 y_2 ,$$

donc z est régulier et les vecteurs $e_1(z) = (-b_1, a_1, 0, 0)$, $e_2(z) = (0, 0, -b_2, a_2 - c)$ forment une base de $T_z M$. On a

$$D_z H(x_1, y_1, x_2, y_2) = 2(a_1 - a_2)(x_1 - x_2) + 2(b_1 - b_2)(y_1 - y_2)$$
$$D_z H(e_1(z)) = 2(a_2 b_1 - a_1 b_2)$$
$$D_z H(e_2(z)) = -2(a_2 b_1 - a_1 b_2) + 2c(b_1 - b_2) .$$

Le point z est donc point critique relatif de H sur M si et seulement si (comme $c \neq 0$)

$$\begin{cases} b_1 - b_2 = 0 \\ a_1 b_2 - a_2 b_1 = 0 \end{cases}$$

c'est-à-dire

$$b_1 = b_2 = 0 \quad \text{ou} \quad b_1 = b_2 \ , \quad a_1 = a_2 \ .$$

Comme $C_1 \cap C_2 = \phi$ (car $R > 1 + c$), le second cas est impossible et on obtient 4 points critiques relatifs :

$$a_1 = \pm 1 \ , \quad a_2 = c \pm R \ , \quad b_1 = b_2 = 0 \ .$$

En ces points critiques, on a $T_z M = \{x_1 = x_2 = 0\}$

$$D_z H = \frac{a_1 - a_2}{a_1} D_z F_1 + \frac{a_2 - a_1}{a_2 - c} D_z F_2$$

et la hessienne relative de H en ces points est :

$$Q_z(y_1, y_2) = 2\left[(y_1 - y_2)^2 - \frac{a_1 - a_2}{a_1} y_1^2 - \frac{a_2 - a_1}{a_2 - c} y_2^2\right] .$$

1. $a_1 = -1$, $a_2 = c + R$.

$$Q(y_1, y_2) = 2\left[(y_1 - y_2)^2 - (1 + c + R) y_1^2 - \frac{c + R + 1}{R} y_2^2\right]$$

de matrice

$$2\begin{pmatrix} -c - R & -1 \\ -1 & -\dfrac{1+c}{R} \end{pmatrix}$$

donc définie négative.

C'est un maximum local strict.

2. $a_1 = -1$, $a_2 = c - R$.

$$Q(y_1, y_2) = 2\left[(y_1 - y_2)^2 - (R + 1 - c) y_1^2 - \frac{R - 1 - c}{R} y_2^2\right]$$

de matrice

$$2\begin{pmatrix} R - c & -1 \\ -1 & \dfrac{1+c}{R} \end{pmatrix}$$

donc définie positive.

C'est un minimum local strict.

3. $a_1 = +1$, $a_2 = c + R$.

$$Q(y_1, y_2) = 2\left[(y_1 - y_2)^2 - (1 - c - R)y_1^2 - \frac{c + R - 1}{R} y_2^2\right]$$

de matrice

$$2\begin{pmatrix} c + R & -1 \\ -1 & \dfrac{1-c}{R} \end{pmatrix}$$

non dégénérée indéfinie : ce n'est pas un extremum local.

4. $a_1 = +1$, $a_2 = c - R$

$$Q(y_1, y_2) = 2\left[(y_1 - y_2)^2 - (1 - c + R)y_1^2 - \frac{1 + R - c}{R} y_2^2\right]$$

de matrice

$$2\begin{pmatrix} c - R & -1 \\ -1 & \dfrac{c-1}{R} \end{pmatrix}$$

non dégénérée indéfinie : ce n'est pas un extremum local.

REMARQUE. $M = C_1 \times C_2$ est une partie compacte de $\mathbf{R}^2 \times \mathbf{R}^2$, et H est continue, donc H est bornée sur M et y atteint son minimum et son maximum (absolus). Les points où les bornes sont atteintes sont des extrema locaux. Donc le seul point où H atteint son minimum (sur M) est :

$$a_1 = -1, \quad a_2 = c - R, \quad b_1 = b_2 = 0, \quad H = (R - 1 - c)^2$$

et le seul point où le maximum est atteint est :

$$a_1 = -1, \quad a_2 = c + R, \quad b_1 = b_2 = 0, \quad H = (R + 1 + c)^2 \ .$$

3ème PARTIE.
EQUATIONS DIFFERENTIELLES.

CHAPITRE X : Existence de solutions.

§. 1. Généralités.

1.1. Equations différentielles.

Une équation différentielle se présente sous la forme générale suivante :

$$(*) \qquad \widetilde{H}\Big(t, y(t), y'(t), \dots, y^{(k)}(t)\Big) = 0$$

où :

– E est un espace vectoriel de dimension finie;

– I est un intervalle ouvert de $\mathbf{R}$, U est une partie ouverte de E^{k+1} ;

– $\widetilde{H}$ est une application donnée, en général au moins de classe C^1, définie sur $I \times U$, et à valeurs dans E.

Une **solution** de l'équation $(*)$ est une application $y : J \to E$ telle que :

– J est un sous-intervalle de I ;

– y est de classe C^k dans l'intervalle J ;

– pour tout point $t \in J$, le point $\Big(y(t), \dots, y^{(k)}(t)\Big) \in U$ et $(*)$ est vérifiée.

1.2. Exemple.

Soient n corps dans l'espace, de masses $m_1, \dots, m_n$ de positions $y_1, \dots, y_n$ soumis à la gravitation universelle. La loi de Newton s'écrit :

$$y_i'' = -\sum_{j \neq i} \frac{m_j(y_i - y_j)}{\|y_i - y_j\|^3} \ , \quad 1 \leq i \leq n \ .$$

C'est une équation différentielle du type $(*)$, où :

– $k = 2$, $I = \mathbf{R}$, $E = (\mathbf{R}^3)^n$;

– U est la partie ouverte de $E \times E \times E$ formée des points $(y_1, \dots, y_n, y_1', \dots, y_n', y_1'', \dots, y_n'') \in (\mathbf{R}^3)^3$ tels que $y_i \neq y_j$ pour $i \neq j$;

— $\widetilde{H}$ est l'application (de classe C^∞) de $\mathbf{R} \times U$ dans $(\mathbf{R}^3)^n = E$ dont la $i^{\text{ème}}$ composante $\widetilde{H}_i : \mathbf{R} \times U \to \mathbf{R}^3$ est donnée par :

$$\widetilde{H}_i(t, y_1, \ldots, y_n'') = y_i'' + \sum_{j \neq i} \frac{m_j(y_i - y_j)}{\|y_i - y_j\|^3} .$$

L'étude des solutions de cette équation différentielle constitue la Mécanique céleste. Cette étude est facile pour $n = 2$ (Lois de Kepler), extrêmement difficile pour $n \geq 3$ où beaucoup de questions centrales restent non résolues. La Mécanique céleste a été historiquement une motivation puissante dans le développement de la théorie générale des équations différentielles.

1.3. Terminologie.

On ne s'intéresse dans la suite qu'à des équations différentielles **explicites** de la forme :

$$(**) \qquad y^{(k)}(t) = H\Big(t, y(t), \ldots, y^{(k-1)}(t)\Big)$$

où E est un espace vectoriel de dimension finie, I est un intervalle ouvert de $\mathbf{R}$, U est une partie ouverte de E^k et H est une application, en général de classe C^1, de $I \times U$ dans E.

REMARQUE. D'après le théorème des fonctions implicites, on peut, au moins localement, se ramener de la forme $(*)$ à la forme $(**)$ lorsque la différentielle partielle de $\widetilde{H}$ dans la direction de $y^{(k)}$ est inversible.

L'entier $k \geq 1$ s'appelle l'**ordre** de l'équation différentielle $(**)$. Par analogie avec la physique, la variable $t \in I$ s'appelle le **temps**. L'équation $(**)$ est dite **autonome** si l'application H ne dépend pas du temps (on a alors $I = \mathbf{R}$).

Une équation différentielle autonome d'ordre 1 se présente donc sous la forme :

$$(***) \qquad y' = H(y) ,$$

et on dit qu'elle est définie par le **champ de vecteurs** $y \mapsto H(y)$ (défini dans U).

La raison pour cette terminologie est la suivante : soit $y : J \to U$ une solution de $(***)$ (où J est un intervalle de $\mathbf{R}$); l'image $y(J) = \Gamma$ est donc une courbe dans U paramétrée par y ; en un point $y(t_0)$ tel que $H(y(t_0)) \neq 0$, la courbe Γ admet une tangente dont la direction est précisément engendrée par $H(y(t_0))$. (On verra que soit y est constante dans J, et alors Γ est un point, soit y' ne s'annule pas dans J, et alors Γ admet une tangente en chacun de ses points).

1.4. Manipulations formelles.

Les manipulations qui suivent ne représentent pas un progrès vers la résolution de $(**)$, mais permettent simplement de présenter la théorie sous un aspect plus unifié.

1.4.1. Réduction à l'ordre 1.

Soit $(**)$ une équation différentielle.

Posons $E_1 = E^k$, $Y = (y, y', \dots, y^{(k-1)})$. L'équation différentielle $(**)$ est alors équivalente à l'équation :

$$(**)' \qquad Y' = \widehat{H}(t, Y) \ ,$$

où $\widehat{H} : I \times U \to E_1$ est l'application :

$$\widehat{H}(t, y_0, \dots, y_{k-1}) = (y_1, \dots, y_{k-1}, H(t, y_0, \dots, y_{k-1})) \ .$$

L'équation $(**)'$ est d'ordre 1 : on a perdu en dimension ($\dim E_1 = k \dim E$) ce qu'on a gagné en ordre.

1.4.2. Réduction aux champs de vecteurs.

A partir de l'équation $(**)'$, posons :

$$\begin{aligned} E_2 &= \mathbf{R} \times E_1 \ , \\ V &= I \times U \ , \\ Z &= (t, Y) \in V \ , \\ \overline{H}(Z) &= (1, \widehat{H}(Z)) \ . \end{aligned}$$

L'équation $(**)'$ est équivalente à :

$$Z' = \overline{H}(Z) \ ,$$

associée au champ de vecteurs dans V définni par $\overline{H}$.

§. 2. Le théorème de Cauchy-Lipschitz.

2.1. Soit

$$(**) \qquad y' = H(t, y)$$

une équation différentielle d'ordre 1, où :

– E est un espace vectoriel de dimension finie;

– I est un intervalle ouvert de $\mathbf{R}$, U est une partie ouverte de E ;

– H est une application de classe C^1 de $I \times U$ dans E.

Une **condition initiale** pour cette équation différentielle est la donnée d'un point (t_0, y_0) dans $I \times U$.

Une **solution** de $(**)$ de **condition initiale** $(\mathbf{t_0}, \mathbf{y_0})$ est une application y de classe C^1, définie sur un sous-intervalle J de I contenant t_0, à valeurs dans U, et satisfaisant :

$$\begin{cases} y(t_0) = y_0 \ , \\ y'(t) = H(t, y(t)) \ , \ \forall t \in J \ . \end{cases}$$

THÉORÈME 1 (Cauchy-Lipschitz). *Soit* (t_0, y_0) *une condition initiale pour l'équation* $(**)$.

A) *Existence.*

Il existe $\varepsilon_0 > 0$ *tel que l'équation* $(**)$ *admette une solution* y *de condition initiale* (t_0, y_0) *définie sur* $J_0 = [t_0 - \varepsilon_0, t_0 + \varepsilon_0]$.

B) *Unicité.*

Si J *est un sous-intervalle de* J_0 *contenant* t_0 *et* $\tilde{y}$ *est une solution de* $(**)$, *de condition initiale* (t_0, y_0), *définie sur* J, **alors** $\tilde{y}$ *coïncide avec la restriction de* y *à* J.

2.2. Démonstration.

1. Munissons E d'une norme et choisissons $\delta > 0$ tel qu'on ait :

$$[t_0 - \delta, t_0 + \delta] \subset I \ , \tag{1}$$

$$\overline{B}(y_0, \delta) \subset U \ . \tag{2}$$

Comme H est continue sur la partie compacte $L_0 = [t_0 - \delta, t_0 + \delta] \times \overline{B}(y_0, \delta)$ de $I \times U$, il existe une constante $M > 0$ telle qu'on ait :

$$\|H(t, y)\| \leq M \ , \quad \text{pour } (t, y) \in L_0 \ . \tag{3}$$

Pour $(t, y) \in I \times U$, notons $D^{(2)}H$ la différentielle partielle de H dans la direction de E. L'application : $(t, y) \mapsto D^{(2)}_{(t,y)}H$ de $I \times U$ dans $\mathcal{L}(E)$ est continue (puisque H est de classe C^1), et par conséquent il existe $K > 0$ tel qu'on ait :

$$\left\| D^{(2)}_{(t,y)}H \right\| \leq K \ , \quad \text{pour } (t, y) \in L_0 \ . \tag{4}$$

Comme la boule $\overline{B}(y_0, \delta)$ est convexe, on a, d'après (4) et le théorème de la moyenne, pour $t \in [t_0 - \delta, t_0 + \delta]$ et $y_1, y_2 \in \overline{B}(y_0, \delta)$:

$$\|H(t, y_1) - H(t, y_2)\| \leq K \, \|y_1 - y_2\| \ . \tag{5}$$

2. On pose $\varepsilon_0 = \min\left(\delta, \dfrac{\delta}{M}, \dfrac{1}{2K}\right)$, $J_0 = [t_0 - \varepsilon_0, t_0 + \varepsilon_0]$. Soit J_1 un intervalle contenu dans J_0 et contenant t_0.

LEMME. *Pour qu'une application continue* $\tilde{y}$ *de* J_1 *dans* U *soit solution de* $(**)$, *de condition initiale* (t_0, y_0), *il faut et il suffit qu'on ait, pour tout* $t \in J_1$:

$$\tilde{y}(t) = y_0 + \int_{t_0}^{t} H(s, \tilde{y}(s))\, ds \ . \tag{6}$$

DÉMONSTRATION : La nécessité de (6) s'obtient en intégrant entre t_0 et $t \in J_1$ l'équation $(**)$.

Inversement, si l'application continue $\tilde{y} : J_1 \to U$ vérifie (6), alors le membre de droite de cette relation est dérivable en tout point $t \in J_1$, de dérivée $H(t, \tilde{y}(t))$; donc $\tilde{y}$ est de classe C^1 et est une solution de $(**)$ définie sur J_1. Par ailleurs, la relation (6), pour $t = t_0$, donne $\tilde{y}(t_0) = y_0$.

3. Munissons l'espace $X = \mathcal{C}(J_1, \overline{B}(y_0, \delta))$ des applications continues de J_1 dans $\overline{B}(y_0, \delta)$ de la distance uniforme. L'espace métrique X est alors **complet** (Chapitre IV, §.3, Proposition 6). Pour $z \in X$, définissons une application $\mathcal{F}(z)$ de J_1 dans E par :

$$\mathcal{F}(z)(t) = y_0 + \int_{t_0}^{t} H(s, z(s))\, ds \ .$$

L'application $\mathcal{F}(z)$ est continue. D'après (3), on a, pour $t \in J_1$:

$$\|\mathcal{F}(z)(t) - y_0\| \leq M\,|t - t_0| \leq \varepsilon_0 M \leq \delta \ , \tag{7}$$

donc $\mathcal{F}(z)$ prend ses valeurs dans $\overline{B}(y_0, \delta)$ et c'est donc un élément de X.

On a ainsi défini une application $\mathcal{F} : z \mapsto \mathcal{F}(z)$ de X dans X. Notons D la distance uniforme de X ; pour $z, \tilde{z} \in X$, on a :

$$\begin{aligned} D(\mathcal{F}(z), \mathcal{F}(\tilde{z})) &= \sup_{t \in J_1} \|\mathcal{F}(z)(t) - \mathcal{F}(\tilde{z})(t)\| \\ &= \sup_{t \in J_1} \left\| \int_{t_0}^{t} \Big[H(s, z(s)) - H(s, \tilde{z}(s)) \Big]\, ds \right\| \ . \end{aligned}$$

Or, pour $s \in J_1$, on a, d'après la relation (5) :

$$\begin{aligned} \|H(s, z(s)) - H(s, \tilde{z}(s))\| &\leq K\, \|z(s) - \tilde{z}(s)\| \\ &\leq K\, D(z, \tilde{z}) \ , \end{aligned}$$

et on obtient donc :

$$\begin{aligned} D(\mathcal{F}(z), \mathcal{F}(\tilde{z})) &\leq \sup_{t \in J_1} |t - t_0|\, K\, D(z, \tilde{z}) \\ &\leq \varepsilon_0 K\, D(z, \tilde{z}) \leq \frac{1}{2} D(z, \tilde{z}) \ . \end{aligned}$$

L'application $\mathcal{F} : X \to X$ est donc $\frac{1}{2}$-Lipschitzienne, et en particulier contractante. Comme X est complet, elle possède dans X un **unique point fixe**.

4. Existence. Prenons $J_1 = J_0$. La partie 3 montre qu'il existe une application continue y de J_0 dans $\overline{B}(y_0, \delta)$ vérifiant la relation (6). Le lemme de la partie 2 montre que y est une solution de $(**)$, de condition initiale (t_0, y_0).

Unicité. Soient J un sous-intervalle de J_0, contenant t_0, et $\tilde{y}$ une solution de $(**)$, de condition initiale (t_0, y_0), définie sur J. L'application $\hat{y}$ satisfait donc (6) pour tout $t \in J$.

Soit J_1 la composante connexe de t_0 dans $\tilde{y}^{-1}(\bar{B}(y_0,\delta))$. D'après la partie 3, on a $\tilde{y}(t) = y(t)$ pour $t \in J_1$. Or d'après la relation (7), on a :

$$\|\tilde{y}(t) - y_0\| = \|y(t) - y_0\| < \delta$$

pour $|t - t_0| < \varepsilon_0$. On a donc $J = J_1$, et $\tilde{y}$ coïncide avec la restriction de y à J.

2.3. Remarque. L'hypothèse que H est de classe C^1 n'est intervenue dans la démonstration que pour obtenir dans la partie 1 des constantes δ, M, K vérifiant les relations (1), (2), (3), (5). Le théorème de Cauchy-Lipschitz (et sa démonstration) est encore valide lorsqu'on suppose seulement que :

– l'application $H : I \times U \to E$ est continue.

– l'application H est localement uniformément lipschitzienne par rapport à la variable y, c'est-à-dire que pour tout $(t_0, y_0) \in I \times U$, il existe $\delta > 0$, $K > 0$ tels que les relations (1), (2), (5) soient satisfaites.

2.4. Complément sur l'unicité.

Soient y, $\tilde{y}$ deux solutions de l'équation différentielle (**), définies respectivement sur des sous-intervalles J, $\tilde{J}$ de I.

PROPOSITION 1. *S'il existe un temps $t_0 \in J \cap \tilde{J}$ tel que $y(t_0) = \tilde{y}(t_0)$, alors les solutions y et $\tilde{y}$ coïncident sur $J \cap \tilde{J}$ et définissent ensemble une solution de (**) sur $J \cup \tilde{J}$.*

DÉMONSTRATION : L'ensemble E des points $t \in J \cap \tilde{J}$ tels que $y(t) = \tilde{y}(t)$ est non vide par hypothèse. Il est fermé dans $J \cap \tilde{J}$ puisque y et $\tilde{y}$ sont continues. Il est ouvert dans $J \cap \tilde{J}$ d'après l'unicité dans le théorème de Cauchy-Lipschitz. On a donc $E = J \cap \tilde{J}$, puisque $J \cap \tilde{J}$ est un intervalle.

Finalement, il est clair que l'application $\hat{y}$ définie par :

$$\begin{aligned} \hat{y}(t) &= y(t) \ , \quad t \in J \\ &= \tilde{y}(t) \ , \quad t \in \tilde{J} \end{aligned}$$

est une solution de (**) définie sur $J \cup \tilde{J}$.

2.5. Complément sur l'existence.

Soit $\mathcal{K}$ une partie **compacte** de $I \times U$.

PROPOSITION 2. *Il existe un nombre $\varepsilon(\mathcal{K}) > 0$ tel que, pour toute condition initiale $(t_0, y_0) \in \mathcal{K}$, l'équation (**) admette une solution de condition initiale (t_0, y_0) définie sur $[t_0 - \varepsilon(\mathcal{K}), t_0 + \varepsilon(\mathcal{K})]$.*

DÉMONSTRATION : Pour tout $(t, y) \in \mathcal{K}$, soient $\delta(t,y)$, $M(t,y)$, $K(t,y)$ des nombres strictement positifs tels qu'on ait :

$$[t - \delta(t,y) \ , \ t + \delta(t,y)] \subset I \ ,$$

$$\overline{B}(y,\delta(t,y)) \subset U \ ,$$
$$\|H(\tilde{t},\tilde{y})\| \leq M(t,y) \ , \quad \text{pour} \ |t-\tilde{t}| \leq \delta(t,y) \ ,$$
$$\|y-\tilde{y}\| \leq \delta(t,y) \ ;$$
$$\|H(\tilde{t},\tilde{y}) - H(\tilde{t},\tilde{\tilde{y}})\| \leq K(t,y)\,\|\tilde{y}-\tilde{\tilde{y}}\| \ ,$$

pour

$$|t-\tilde{t}| \leq \delta(t,y) \ , \quad \|y-\tilde{y}\| \leq \delta(t,y) \ ,$$
$$\|y-\tilde{\tilde{y}}\| \leq \delta(t,y) \ .$$

(Cf. relations (1), (2), (3), (5) de 2.2).

Posons :

$$V(t,y) = \left]t - \frac{1}{2}\delta(t,y), t + \frac{1}{2}\delta(t,y)\right[\times B\left(y, \frac{1}{2}\delta(t,y)\right) \ .$$

Alors $(V(t,y))_{(t,y)\in\mathcal{K}}$ est un recouvrement de $\mathcal{K}$ par parties ouvertes; d'après le théorème de Borel-Lebesgue, il existe donc des points $(t_1,y_1),\ldots,(t_\ell,y_\ell)$ de $\mathcal{K}$ tels que :

$$\mathcal{K} \subset \bigcup_{i=1}^{\ell} V(t_i,y_i) \ .$$

Posons :

$$\overline{M} = \max_{1\leq i\leq \ell} M(t_i,y_i)$$
$$\overline{K} = \max_{1\leq i\leq \ell} K(t_i,y_i)$$
$$\bar{\delta} = \frac{1}{2}\min_{1\leq i\leq \ell} \delta(t_i,y_i)$$
$$\varepsilon = \varepsilon(\mathcal{K}) = \min\left(\bar{\delta}, \frac{\bar{\delta}}{\overline{M}}, \frac{1}{2\overline{K}}\right) \ .$$

Soit $(t_0,y_0) \in \mathcal{K}$. Il existe $i \in \{1,\ldots,\ell\}$ tel que $(t_0,y_0) \in V(t_i,y_i)$. On a alors :

$$[t_0-\bar{\delta}, t_0+\bar{\delta}] \subset [t_i - \delta(t_i,y_i), t_i + \delta(t_i,y_i)] \ ,$$
$$\overline{B}(y_0,\bar{\delta}) \subset \overline{B}(y_i,\delta(t_i,y_i)) \ ,$$

donc

$$\|H(t,y)\| \leq \overline{M} \ \text{pour} \ |t-t_0| \leq \bar{\delta} \ , \ \|y-y_0\| \leq \bar{\delta} \ ,$$
$$\|H(t,y) - H(t,\tilde{y})\| \leq \overline{K}\,\|y-\tilde{y}\|$$

pour

$$|t-t_0| \leq \bar{\delta} \ ,$$
$$\|y-y_0\| \leq \bar{\delta} \ ,$$
$$\|\tilde{y}-y_0\| \leq \bar{\delta} \ .$$

On peut donc appliquer le reste de la démonstration de 2.2 et conclure qu'il existe une solution de (**) de condition initiale (t_0,y_0) définie sur $[t_0-\varepsilon(\mathcal{K}), t_0+\varepsilon(\mathcal{K})]$. □

§. 3. Solutions maximales.

3.1. On considère toujours l'équation différentielle

$$(**) \qquad y' = H(t,y)$$

pour laquelle on garde les mêmes notations que précédemment.

Soit (t_0, y_0) une condition initiale.

THÉORÈME 2.

A) *Il existe un intervalle* ***ouvert*** J *contenant* t_0, *contenu dans* I, *et une solution* y *de* $(**)$, *de condition initiale* (t_0, y_0), *définie sur* J *qui est* ***maximale*** *dans le sens suivant : Si* $\widetilde{y}$ *est une autre solution de condition initiale* (t_0, y_0), *définie sur un intervalle* $\widetilde{J}$, *alors* $\widetilde{J} \subset J$ *et* $\widetilde{y}$ *coïncide avec la restriction de* y *à* $\widetilde{J}$.

B) ***La solution maximale*** y ***sort de tout compact*** *de* $I \times U$, *c'est-à-dire que pour toute partie compacte* $\mathcal{K}$ *de* $I \times U$, *il existe un intervalle compact* $[a,b] \subset J$ *tel que* $(t, y(t)) \notin \mathcal{K}$ *si* $t \in J - [a,b]$.

DÉMONSTRATION : Soit $\mathcal{J}$ l'ensemble des paires $(\widetilde{J}, \widetilde{y})$ telles que :

– $\widetilde{J}$ est un sous-intervalle de I, contenant t_0 ;

– $\widetilde{y}$ est une solution de $(**)$, de condition initiale (t_0, y_0), définie sur $\widetilde{J}$.

Posons $J = \bigcup_{(\widetilde{J},\widetilde{y}) \in \mathcal{J}} \widetilde{J}$. Alors J est un intervalle contenu dans I, contenant t_0.

On définit une application $y : J \to U$ de la façon suivante : pour $t \in J$, soit $(\widetilde{J}, \widetilde{y}) \in \mathcal{J}$ tel que $t \in \widetilde{J}$; on pose $y(t) = \widetilde{y}(t)$. Cette définition ne dépend pas de l'élément $(\widetilde{J}, \widetilde{y})$ considéré : si $(\widetilde{J}_1, \widetilde{y}_1)$ est un autre élément de $\mathcal{J}$ tel que $t \in \widetilde{J}_1$, les solutions $\widetilde{y}$ et $\widetilde{y}_1$ de $(**)$ coïncident sur $\widetilde{J} \cap \widetilde{J}_1$, donc en t, d'après la proposition 1.

L'application y est clairement une solution de $(**)$ de condition initiale (t_0, y_0), définie sur J.

Montrons que J est ouvert, c'est-à-dire que les bornes supérieures et inférieures de J (prises dans $\mathbf{R} \cup \{-\infty, +\infty\}$) n'appartiennent pas à J.

Si par exemple la borne supérieure T de J appartenait à J, il existerait d'après le théorème 1 un nombre $\varepsilon > 0$ et une solution $\widehat{y}$ de $(**)$ définie sur $[T - \varepsilon, T + \varepsilon]$ tels que $\widehat{y}(T) = y(T)$.

L'application qui coïncide avec y sur J et avec $\widehat{y}$ sur $[T, T+\varepsilon]$ serait une solution de $(**)$ de condition initiale (t_0, y_0) définie sur un intervalle contenant strictement J : ceci contredit la définition de J.

On démontre de même que la borne inférieure de J ne peut appartenir à J, et l'intervalle J est donc ouvert.

Soit $\mathcal{K}$ une partie compacte de $I \times U$. Soit $\varepsilon(\mathcal{K})$ un nombre strictement positif satisfaisant les conclusions de la proposition 2. Posons :

$$\bar{a} = \inf \mathcal{K}_1 \ , \quad \bar{b} = \sup \mathcal{K}_1 \ ,$$

où $\mathcal{K}_1$ est la partie compacte de I, image de $\mathcal{K}$ par la projection $(t,y) \mapsto t$ de $I \times U$ sur I. Posons aussi :

$$\begin{aligned} \bar{\bar{a}} &= \inf J \in \mathbf{R} \cup \{-\infty\} \ , \\ \bar{\bar{b}} &= \sup J \in \mathbf{R} \cup \{+\infty\} \ , \\ a &= \max(\bar{a}, \bar{\bar{a}} + \varepsilon(\mathcal{K})) \in \mathbf{R} \ , \\ b &= \min(\bar{b}, \bar{\bar{b}} - \varepsilon(\mathcal{K})) \in \mathbf{R} \ . \end{aligned}$$

Soit $t \in J$ tel que $(t, y(t)) \in \mathcal{K}$. On a donc $t \in [\bar{a}, \bar{b}]$. D'après la proposition 2, il existe une solution $\widehat{y}$ de $(**)$ définie sur $[t-\varepsilon(\mathcal{K}), t+\varepsilon(\mathcal{K})]$ vérifiant $\widehat{y}(t) = y(t)$. D'après la maximalité de (J, y), on doit avoir $[t-\varepsilon(\mathcal{K}), t+\varepsilon(\mathcal{K})] \subset J$, donc $t \geq \bar{\bar{a}} + \varepsilon(\mathcal{K})$, $t \leq \bar{\bar{b}} - \varepsilon(\mathcal{K})$ et finalement $t \in [a, b]$. $\square$

3.2. Remarque. La partie B) du théorème 2 n'a d'intérêt que lorsque J est **strictement contenu** dans I.

Supposons par exemple que la borne supérieure $\bar{\bar{b}}$ de J soit finie et **appartienne à I.** Soit $\mathcal{K}_2$ une partie **compacte de U.** Alors $[t_0, \bar{\bar{b}}] \times \mathcal{K}_2$ est une partie compacte de $I \times U$. D'après la partie B) du théorème 2, il existe un temps $b \in \left[t_0, \bar{\bar{b}}\right[$ tel que pour tout temps $t \in \left]b, \bar{\bar{b}}\right[$ on ait $y(t) \notin \mathcal{K}_2$. En d'autres termes, **si** la borne supérieure du domaine de définition de la solution maximale appartient à I, **alors** cette solution sort de **tout compact de U** lorsque le temps s'approche de cette borne supérieure.

3.3. Exemple fondamental.

Supposons qu'on ait $I = \mathbf{R}$, $U = E$. Soit $(t_0, y_0) \in \mathbf{R} \times E$ une condition initiale. Notons y la solution maximale de $(**)$ de condition initiale (t_0, y_0), J son domaine de définition. On a la dichotomie :

- Soit $[t_0, +\infty) \subset J$, c'est-à-dire que la solution maximale est définie pour tout temps $t \geq t_0$;
- Soit $\sup J = \bar{\bar{b}} < +\infty$, et on a alors :

$$\lim_{t \to \bar{\bar{b}}} \|y(t)\| = +\infty \ .$$

(Dans le premier cas, on ne peut rien dire, sans hypothèses supplémentaires, sur le comportement de $y(t)$ lorsque t tend vers $+\infty$).

De même :

- Soit $(-\infty, t_0] \subset J$, c'est-à-dire que la solution maximale est définie pour tout temps $t \leq t_0$;

– Soit $\inf J = \bar{\bar{a}} > -\infty$, et on a alors :

$$\lim_{t \to \bar{\bar{a}}} \|y(t)\| = +\infty \ .$$

§. 4. Equations différentielles d'ordre k.

4.1. Le cas d'une équation différentielle

$$y^{(k)} = H\left(t, y, \ldots, y^{(k-1)}\right)$$

d'ordre $k \geq 2$ se ramène à celui d'une équation d'ordre 1, comme expliqué en 1.4.1, par l'introduction de la variable

$$Y = \left(y, y', \ldots, y^{(k-1)}\right) \in E^k \ .$$

La donnée d'une condition initiale (t_0, Y_0) est donc la donnée :

- d'un temps $t_0 \in I$
- des valeurs $y(t_0)$, $y'(t_0), \ldots, y^{(k-1)}(t_0)$ de la solution cherchée et de ses dérivées jusqu'à l'ordre $(k-1)$ au temps t_0.

Ceci étant, tous les résultats des paragraphes 2 et 3 s'appliquent à l'équation différentielle précédente.

4.2. Exemple.

Reprenons l'exemple 1.2 de la Mécanique céleste.

En posant $y = (y_1, \ldots, y_n) \in (\mathbf{R}^3)^n$, on a une équation différentielle du second ordre :

$$y'' = H(t, y, y')$$

où l'application H ne dépend en fait que de y et sa $i^{\text{ème}}$ composante H_i est :

$$H_i(t, y, y') = -\sum_{j \neq i} m_j \ \frac{y_i - y_j}{\|y_i - y_j\|^3} \ .$$

L'application H est de classe C^∞ dans l'ouvert $\mathbf{R} \times V \times (\mathbf{R}^3)^n$, où :

$$V = \left\{ y \in (\mathbf{R}^3)^n \,\middle|\, \forall i \neq j \ , \ y_i \neq y_j \right\} \ .$$

Dans le cadre des §.2.3, on a donc $I = \mathbf{R}$, $U = V \times (\mathbf{R}^3)^n$. Se donner une condition initiale revient à se donner un temps t_0, les positions $y_1(t_0), \ldots, y_n(t_0)$ **distinctes** des masses au temps t_0 et leurs vitesses $y'_1(t_0), \ldots, y'_n(t_0)$.

Pour tout $R > 0$, définissons :

$$K_R = \Big\{ (y_1, \ldots, y_n, \ y'_1, \ldots, y'_n) \in (\mathbf{R}^3)^n \times (\mathbf{R}^3)^n \Big|$$
$$\sum_{i=1}^n \|y_i\| + \sum_{i=1}^n \|y'_i\| + \sum_{i<j} \|y_i - y_j\|^{-1} \leq R \Big\} \ .$$

C'est une partie compacte de $U = V \times (\mathbf{R}^3)^n$.

Inversement, pour toute partie compacte $\mathcal{K}$ de U il existe $R > 0$ tel que $\mathcal{K} \subset K_R$.

Soit $(t_0, y(t_0), y'(t_0))$ une condition initiale, y la solution maximale correspondante. Alors

- soit y est définie pour tout temps $t \geq t_0$;
- soit la borne supérieure t_1 du domaine de définition de y est finie et on a :

$$\lim_{t \to t_1} \left(\sum_{i=1}^{n} \|y_i(t)\| + \sum_{i=1}^{n} \|y_i'(t)\| + \sum_{i<j} \|y_i(t) - y_j(t)\|^{-1} \right) = +\infty \ .$$

[En fait on peut montrer qu'on a plus précisément :

$$\lim_{t \to t_1} \sum_{i=1}^{n} \|y_i'(t)\| = +\infty$$

$$\lim_{t \to t_1} \sum_{i<j}^{n} \|y_i(t) - y_j(t)\|^{-1} = +\infty \ ] \ .$$

CHAPITRE XI : Equations différentielles linéaires et comparaison des solutions.

§. 1. Rappels sur les équations linéaires du premier ordre sur $\mathbf{R}$.

1.1. Equations sans second membre.

Soient I un intervalle ouvert, et λ une fonction continue de I dans $\mathbf{R}$.

Considérons l'équation différentielle :

$$y' = \lambda(t)y \ .$$

Soit $(t_0, y_0) \in I \times \mathbf{R}$ une condition initiale pour cette équation différentielle. La solution maximale correspondante est alors définie sur l'intervalle I tout entier et est explicitement donnée par :

$$y(t) = y_0 \ \exp\left(\int_{t_0}^{t} \lambda(s)\,ds\right) \ .$$

1.2. Equations avec second membre.

Soient I, λ comme ci-dessus et μ une autre fonction continue de I dans $\mathbf{R}$.

Considérons l'équation différentielle :

$$y' = \lambda(t)y + \mu(t) \ .$$

Soit (t_0, y_0) une condition initiale pour cette équation différentielle. Là encore, la solution maximale correspondante est définie sur l'intervalle I tout entier et est déterminée par une formule explicite : on recherche cette solution par la méthode de **variation des constantes**, c'est-à-dire qu'on l'écrit sous la forme :

$$\begin{aligned} y(t) &= z(t) \ \exp\left(\int_{t_0}^{t} \lambda(s)\,ds\right) \\ &= z(t)\varphi(t) \ ; \end{aligned}$$

on a :

$$\begin{aligned} \varphi'(t) &= \lambda(t)\varphi(t) \ ; \\ \varphi(t_0) &= 1 \end{aligned}$$

de sorte qu'on doit avoir :

$$\begin{aligned} z(t_0) &= y_0 \ , \\ z'(t)\varphi(t) + z(t)\varphi'(t) &= \lambda(t)z(t)\varphi(t) + \mu(t) \ , \end{aligned}$$

ou encore

$$z'(t) = \frac{\mu(t)}{\varphi(t)}\,, \quad z(t_0) = y_0$$

et finalement :

$$z(t) = y_0 + \int_{t_0}^{t} \mu(s)\ \exp\left(-\int_{t_0}^{s} \lambda(u)du\right) ds\ .$$

1.3. Remarque. On a seulement supposé que λ et μ étaient continues, de sorte que les hypothèses du théorème d'existence et d'unicité de Cauchy-Lipschitz ne sont pas nécessairement vérifiées. Cependant, la remarque du Chapitre X, §.2.3 s'applique ici, de sorte que les **conclusions** du théorème de Cauchy-Lipschitz sont encore valables pour les équations différentielles considérées en 1.1 et 1.2.

§. 2. Le lemme de Grönwall.

2.1. Enoncé.

Soit y une application de classe C^1 définie sur un intervalle ouvert I de $\mathbf{R}$ et à valeurs dans un espace vectoriel normé de dimension finie E.

Soit $t_0 \in I$. On suppose que pour $t \geq t_0$, $t \in I$, l'application y vérifie :

$$\|y'(t)\| \leq \lambda(t)\,\|y(t)\| + \mu(t)\ ,$$

où λ et μ sont deux fonctions continues définies sur I, **la fonction λ étant à valeurs positives ou nulles.**

THÉORÈME 1 (Lemme de Grönwall). *Sous les hypothèses précédentes, soit φ une solution, définie sur I, de l'équation différentielle :*

$$\varphi'(t) = \lambda(t)\varphi(t) + \mu(t)$$

et satisfaisant de plus :

$$\|y(t_0)\| \leq \varphi(t_0)\ .$$

Alors on a, pour tout $t \in I$, $t \geq t_0$:

$$\|y(t)\| \leq \varphi(t)\ .$$

REMARQUE. Lorsque la fonction λ est identiquement nulle, l'énoncé précédent se réduit au théorème de la moyenne. Inversement, on va démontrer le théorème à l'aide du théorème de la moyenne.

DÉMONSTRATION : Supposons d'abord qu'on ait :

$$\|y(t_0)\| < \varphi(t_0) ,$$

et montrons qu'on a alors, pour $t \in I$, $t \geq t_0$:

$$\|y(t)\| < \varphi(t) .$$

Dans le cas contraire, il existe $t_1 \in I$, $t_1 > t_0$ tel qu'on ait :

$$\begin{aligned} &\|y(t_1)\| = \varphi(t_1) \\ &\|y(t)\| < \varphi(t) , \quad \forall t \in [t_0, t_1[\ . \end{aligned}$$

On a alors, pour $t \in [t_0, t_1]$:

$$\begin{aligned} \|y'(t)\| &\leq \lambda(t) \, \|y(t)\| + \mu(t) \\ &\leq \lambda(t)\varphi(t) + \mu(t) = \varphi'(t) , \end{aligned}$$

d'où, par le théorème de la moyenne :

$$\|y(t_1) - y(t_0)\| \leq \varphi(t_1) - \varphi(t_0) ,$$

ce qui implique $\|y(t_1)\| < \varphi(t_1)$, une contradiction.

Pour traiter le cas restant $\|y(t_0)\| = \varphi(t_0)$, notons, pour $\varepsilon \in \mathbf{R}$, φ_ε la solution définie sur I de l'équation différentielle

$$\varphi'_\varepsilon(t) = \lambda(t)\varphi_\varepsilon(t) + \mu(t)$$

telle que $\varphi_\varepsilon(t_0) = \|y(t_0)\| + \varepsilon$.

D'après 1.2, l'application $(\varepsilon, t) \mapsto \varphi_\varepsilon(t)$ de $\mathbf{R} \times I$ dans $\mathbf{R}$ est continue. D'après la première partie de la démonstration, on a, pour $t \geq t_0$, $t \in I$ et $\varepsilon > 0$:

$$\|y(t)\| < \varphi_\varepsilon(t) .$$

On conclut donc qu'on a, pour $t \in I$, $t \geq t_0$:

$$\|y(t)\| \leq \varphi_0(t) .$$

2.2. Temps antérieurs à t_0.

L'énoncé précédent a permis de contrôler la norme de l'application y pour des temps postérieurs au temps t_0.

En changeant le sens du temps, on va obtenir un contrôle similaire pour les temps antérieurs à t_0.

Supposons donc que l'application y vérifie, pour $t \in I$, $t \leq t_0$:

$$\|y'(t)\| \leq \lambda(t) \, \|y(t)\| + \mu(t) ,$$

où λ et μ sont des fonctions continues définies sur I, la fonction λ étant à **valeurs positives ou nulles.**

Posons $\widetilde{\lambda}(t) = \lambda(-t)$, $\widetilde{\mu}(t) = \mu(-t)$, $\widetilde{y}(t) = y(-t)$. On a alors $\widetilde{y}'(t) = -y'(-t)$, donc

$$\|\widetilde{y}'(t)\| \leq \widetilde{\lambda}(t) \, \|\widetilde{y}(t)\| + \widetilde{\mu}(t) ,$$

pour $t \geq -t_0$ tel que $-t \in I$.

D'après le lemme de Grōnwall, on aura :

$$\|\tilde{y}(t)\| \leq \tilde{\varphi}(t) ,$$

pour $t \geq -t_0$ tel que $-t \in I$, à condition que $\tilde{\varphi}$ soit une solution de l'équation

$$\tilde{\varphi}'(t) = \tilde{\lambda}(t)\tilde{\varphi}(t) + \tilde{\mu}(t)$$

satisfaisant $\tilde{\varphi}(-t_0) \geq \|\tilde{y}(-t_0)\|$.

En posant $\varphi(t) = \tilde{\varphi}(-t)$, cela revient à dire que, si φ est une solution (définie dans I) de l'équation différentielle :

$$\varphi'(t) = -\lambda(t)\varphi(t) - \mu(t)$$

satisfaisant de plus :

$$\varphi(t_0) \geq \|y(t_0)\| ,$$

alors on a, pour $t \in I$, $t \leq t_0$:

$$\|y(t)\| \leq \varphi(t) .$$

2.3. Exemple fondamental.

Soit y une application de classe C^1 définie sur un intervalle ouvert I de $\mathbf{R}$ et à valeurs dans un espace vectoriel de dimension finie E.

Supposons qu'il existe un nombre réel $\lambda \geq 0$ tel qu'on ait, pour tout $t \in I$:

$$\|y'(t)\| \leq \lambda \|y(t)\| .$$

Soit $t_0 \in I$. On a alors, pour $t \in I$:

$$\|y(t)\| \leq \|y(t_0)\| \; e^{\lambda|t-t_0|} .$$

En effet, pour $t \geq t_0$, on compare y à la solution

$$\varphi_+(t) = \|y(t_0)\| \; e^{\lambda(t-t_0)}$$

de l'équation

$$\varphi_+'(t) = \lambda\varphi_+(t) ,$$

tandis que pour $t \leq t_0$, on compare y à la solution

$$\varphi_-(t) = \|y(t_0)\| \; e^{-\lambda(t-t_0)}$$

de l'équation

$$\varphi_-'(t) = -\lambda\varphi_-(t) .$$

2.4. Conséquence pour les solutions maximales.

Soient I un intervalle ouvert de $\mathbf{R}$ et E un espace vectoriel normé de dimension finie.

On considère une équation différentielle :

$$(**) \qquad y' = H(t,y)$$

où H est une application de classe C^1 de $I \times E$ dans E (il suffit en fait de supposer que H vérifie les conditions de la remarque du Chapitre X, §.2.3).

On **suppose** qu'il existe deux fonctions continues λ, μ définies sur I telles qu'on ait, pour tout $(t,y) \in I \times E$:

$$\|H(t,y)\| \leq \lambda(t)\,\|y\| + \mu(t)$$

(les fonctions λ et μ sont donc nécessairement à valeurs positives).

PROPOSITION 1. *Sous les hypothèses précédentes, toute solution maximale de l'équation différentielle* $(**)$ *est définie sur l'intervalle* I *tout entier.*

DÉMONSTRATION : Soient y une solution maximale de l'équation $(**)$, J son domaine de définition, $t_0 \in J$. Soit φ la solution, définie dans I, de l'équation

$$\varphi'(t) = \lambda(t)\varphi(t) + \mu(t)$$

telle que $\varphi(t_0) = \|y(t_0)\|$. Pour $t \in J$, on a :

$$\begin{aligned}\|y'(t)\| &= \|H(t,y(t))\| \\ &\leq \lambda(t)\,\|y(t)\| + \mu(t)\ ,\end{aligned}$$

donc, d'après le lemme de Grönwall, on a :

$$\|y(t)\| \leq \varphi(t)$$

pour $t \geq t_0$, $t \in J$. Si la borne supérieure t_1 de J est finie et appartient à I, l'application φ est bornée sur l'intervalle compact $[t_0, t_1]$, et il en est donc de même pour y. Mais ceci contredit le Théorème 2, partie B), §.3.1 du Chapitre X (cf. aussi Remarque §.3.2).

Par conséquent la borne supérieure de J est ausi celle de I. De même la borne inférieure de J est aussi celle de I. On a donc $J = I$. □

REMARQUE. On peut bien sûr aussi se servir du lemme de Grönwall pour contrôler la norme des solutions maximales sur l'intervalle I tout entier.

§. 3. Equations linéaires du premier ordre en dimension quelconque.

3.1. Equations sans second membre.

Soient I un intervalle ouvert de $\mathbf{R}$ et E un espace vectoriel (normé) de dimension finie.

On considère l'équation différentielle :

$$(*) \qquad y' = A(t)y \ ,$$

où A est une application continue de I dans $\mathcal{L}(E)$, et on cherche des solutions y à valeurs dans E.

Les hypothèses de la remarque du Chapitre X, §.2.3 sont satisfaites, donc il en est de même pour les conclusions du théorème de Cauchy-Lipschitz.

PROPOSITION 2.

1) *Toute solution maximale de l'équation différentielle* $(*)$ *est définie sur l'intervalle* I *tout entier.*
2) *Les solutions maximales de* $(*)$ *forment un sous-espace vectoriel* S *de l'espace* $C^1(I,E)$ *des applications de classe* C^1 *de* I *dans* E.
3) *La dimension de* S *est égale à celle de* E. *Plus précisément, pour tout* $t_0 \in I$, *l'application* $y \mapsto y(t_0)$ *est un isomorphisme de* S *sur* E.

DÉMONSTRATION : La partie 1) résulte de la Proposition 1 (en prenant $\mu(t) \equiv 0$, $\lambda(t) = \|A(t)\|$).

Toute solution maximale de $(*)$ est donc un élément de $C^1(I,E)$. Il est clair que si y_1, y_2 sont des solutions (maximales) de $(*)$ et $\lambda_1, \lambda_2 \in \mathbf{R}$, alors $\lambda_1 y_1 + \lambda_2 y_2$ est encore une solution (maximale) de $(*)$; d'où la partie 2).

Finalement, l'application $y \mapsto y(t_0)$ de S dans E est évidemment linéaire, et elle est injective et surjective d'après l'unicité et l'existence dans le théorème de Cauchy-Lipschitz. □

Munissons $\mathcal{L}(E)$ de la norme d'applications linéaires ; on a alors, pour tout $y \in S$, $t \in I$

$$\|y'(t)\| \leq \|A(t)\| \, \|y(t)\| \ .$$

Pour tout $t_0 \in I$, nous obtenons alors, par le lemme de Grönwall :

$$\|y(t)\| \leq \|y(t_0)\| \ \exp\left(\left|\int_{t_0}^{t} \|A(s)\| \, ds\right|\right) \ , \quad \forall t \in I \ .$$

3.2. Equations avec second membre.

Soient I, E, A comme précédemment. Nous considérons maintenant l'équation différentielle :

$$(**) \qquad y' = A(t)y + b(t) ,$$

où b est une application continue de I dans E.

L'équation $(**)$ vérifie à nouveau les hypothèses de la remarque du Chapitre X, §.2.3, donc les conclusions du théorème de Cauchy-Lipschitz sont vérifiées.

PROPOSITION 3.

1) *Toute solution maximale de* $(**)$ *est définie sur l'intervalle* I *tout entier.*
2) *Les solutions maximales de* $(**)$ *forment un sous-espace* ***affine*** $\widetilde{S}$ *de* $C^1(I,E)$*, dont la direction est le sous-espace vectoriel* S *des solutions de* $(*)$*. En particulier, la dimension de* $\widetilde{S}$ *est égale à celle de* E*.*

DÉMONSTRATION : La partie 1) est de nouveau conséquence de la proposition 1 (en prenant $\lambda(t) = \|A(t)\|$, $\mu(t) = \|b(t)\|$).

Soit y_0 une solution maximale de $(**)$. Une vérification immédiate montre qu'une application $y \in C^1(I,E)$ est solution (maximale) de $(**)$ si et seulement si l'application $y - y_0$ est solution de $(*)$, c'est-à-dire appartient à S. Ceci démontre la partie 2) de la proposition. □

On peut utiliser, de même qu'en 3.1, le lemme de Grönwall pour limiter la croissance des solutions maximales de $(**)$.

3.3. Equations à coefficients constants.

On se propose d'étudier plus précisément l'équation différentielle linéaire du premier ordre :

$$(*) \qquad y' = Ay ,$$

où A est un élément de $\mathcal{L}(E)$ **indépendant du temps** t.

Nous nous contenterons d'étudier cette équation sous l'hypothèse supplémentaire que l'endomorphisme A de E est **diagonalisable sur** $\mathbf{C}$, ce qui est en particulier le cas lorsque toutes les racines (réelles ou complexes) du polynôme caractéristique de A sont simples.

Il existe alors une décomposition :

$$E = \left(\bigoplus_{i=1}^{r} E_i\right) \oplus \left(\bigoplus_{j=1}^{s} E'_j\right)$$

ayant les propriétés suivantes :

– pour $1 \le i \le r$, chaque sous-espace E_i est de dimension 1, invariant par A, et il existe une valeur propre **réelle** λ_i de A telle qu'on ait $Ax = \lambda_i x$ pour $x \in E_i$;

– Pour $1 \le j \le s$, chaque sous-espace E'_j est de dimension 2, invariant par A ; il existe une paire μ_j, $\bar{\mu}_j$, avec $\operatorname{Im} \mu_j > 0$, de valeurs propres **complexes conjuguées** de A, et une identification de E'_j à $\mathbf{C}$ telles qu'on ait $Az = \mu_j z$ pour $z \in E'_j$.

En identifiant E à $\mathbf{R}^r \times \mathbf{C}^s$ (où on note les coordonnées $(x_1, \ldots, x_r, z_1, \ldots, z_s)$), l'équation différentielle considérée devient :

$$\begin{cases} x'_i = \lambda_i x_i \ , & 1 \le i \le r \\ z'_j = \mu_j z_j \ , & 1 \le j \le s \ . \end{cases}$$

On peut alors résoudre explicitement :

$$\begin{cases} x_i(t) = e^{\lambda_i t} \ x_i(0) \ , & 1 \le i \le r \\ z_j(t) = e^{\mu_j t} \ z_j(0) \ , & 1 \le j \le s \ . \end{cases}$$

L'allure de la courbe (paramétrée par t) décrite par $z_j(t)$ est plus précisément décrite par :

$$\begin{cases} |z_j(t)| & = e^{(\operatorname{Re} \mu_j)t} \, |z_j(0)| \\ \operatorname{Arg} z_j(t) & = t(\operatorname{Im} \mu_j) + \operatorname{Arg} z_j(0) \ . \end{cases}$$

Lorsque $z_j(0)$ est non nul, c'est une spirale logarithmique si $\operatorname{Re} \mu_j \neq 0$, un cercle (parcouru périodiquement avec une période $2\pi i \big/ \mu_j$) si $\operatorname{Re} \mu_j = 0$.

En regroupant certains des sous-espaces E_i, E'_j on obtient la décomposition fondamentale de E :

$$E = E^- \oplus E^0 \oplus E^+$$

E^- est la somme directe des sous-espaces E_i et E'_j correspond à des valeurs propres de partie réelle **strictement négative**; on l'appelle le sous-espace **stable** du champ de vecteurs (linéaire) considéré et c'est exactement l'ensemble des solutions $y(t)$ telles que :

$$\lim_{t \to +\infty} \|y(t)\| = 0 \ .$$

E^0 est la somme directe des sous-espaces E_i et E'_j correspondant à des valeurs propres **nulles ou imaginaires pures**; on l'appelle le sous-espace **central** du champ de vecteurs considéré; c'est exactement l'ensemble des solutions $y(t)$ telles que :

$$\sup_{t \in \mathbb{R}} \|y(t)\| < +\infty \ .$$

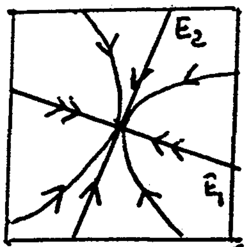
E^+ est la somme directe des sous-espaces E_i ou E'_j correspondant à des valeurs propres de partie réelle strictement positive; on l'appelle le sous-espace **instable** du champ de vecteurs considéré; c'est exactement l'ensemble des solutions $y(t)$ telles que :

$$\lim_{t \to -\infty} \|y(t)\| = 0 \ .$$

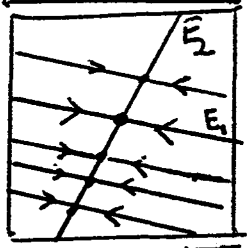
EXEMPLE. Décrivons l'allure des solutions lorsque $\dim E = 2$ et les valeurs propres de A sont **simples**.

1. A a deux valeurs propres réelles $\lambda_1 < \lambda_2$.

1.a).
$\lambda_1 < \lambda_2 < 0$
$E^- = E$
$E^0 = E^+ = \{0\}$.

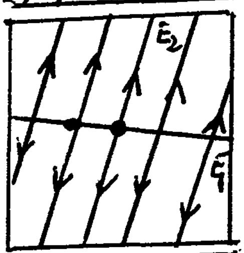


1.b).
$\lambda_1 < \lambda_2 = 0$
$E^- = E_1$
$E^0 = E_2$
$E^+ = \{0\}$.

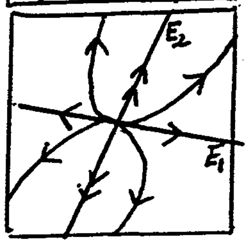


1.c).
$\lambda_1 < 0 < \lambda_2$
$E^- = E_1$
$E^0 = \{0\}$
$E^+ = E_2$.

1.d).
$\lambda_1 = 0 < \lambda_2$
$E^- = \{0\}$
$E^0 = E_1$
$E^+ = E_2$.

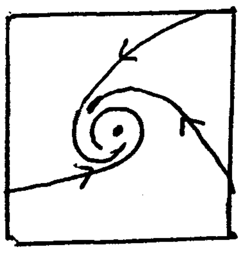


1.e).
$0 < \lambda_1 < \lambda_2$
$E^- = E^0 = \{0\}$
$E^+ = E$.



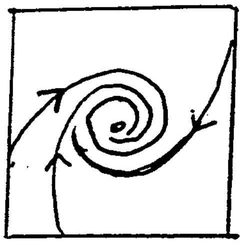
2. A a deux valeurs propres complexes conjuguées $\mu, \bar{\mu}$, et dans une identification de E à $\mathbb{C}$ qui **préserve l'orientation**, A s'identifie à la multiplication par μ.

2.a).
$\mathrm{Re}\,\mu < 0$
$\mathrm{Im}\,\mu > 0$
$E^- = E$, $E^0 = E^+ = \{0\}$.



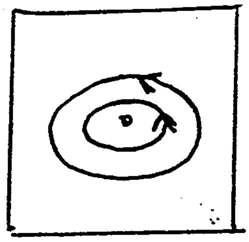
2.b).

$\operatorname{Re}\mu < 0$
$\operatorname{Im}\mu < 0$
$E^- = E,\ E^0 = E^+ = \{0\}.$



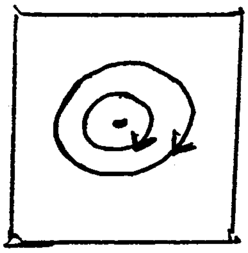
2.c).

$\operatorname{Re}\mu = 0$
$\operatorname{Im}\mu > 0$
$E^- = E^+ = \{0\},\ E^0 = E.$



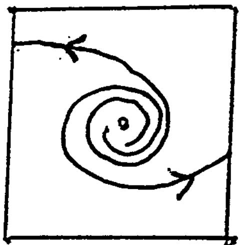
2.d).

$\operatorname{Re}\mu = 0$
$\operatorname{Im}\mu < 0$
$E^- = E^+ = \{0\},\ E^0 = E.$



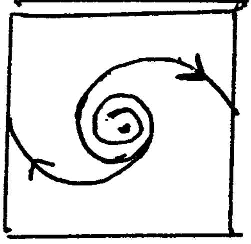
2.e).

$\operatorname{Re}\mu > 0$
$\operatorname{Im}\mu > 0$
$E^- = E^0 = \{0\},\ E^+ = E.$



2.f).

$\operatorname{Re}\mu > 0$
$\operatorname{Im}\mu < 0$
$E^- = E^0 = \{0\},\ E^+ = E.$



§. 4. Dépendance des conditions initiales. L'équation aux variations.

4.1. On considère dans ce paragraphe une équation différentielle

$$(*) \qquad y' = H(t,y)$$

où :

– I est un intervalle ouvert de $\mathbf{R}$

– U est une partie ouverte d'un espace vectoriel normé E de dimension finie

– H est une application de classe C^1 de $I \times U$ dans E.

On se fixe une fois pour toutes un temps initial $t_0 \in I$.

Etant donné un point $y \in U$; on note $J(y)$ l'intervalle de définition de la solution maximale de $(*)$ de condition initiale (t_0, y), et on note

$t \to \Phi(t,y)$ cette solution maximale. L'application Φ est donc définie sur la partie :

$$V = \bigcup_{y\in U} J(y) \times \{y\}$$

de $I \times U$. Le point $\Phi(t,y)$ représente donc la position à l'instant t de la solution qui était à l'instant initial t_0 au point y. On a, pour $(t,y) \in V$:

$$\frac{\partial}{\partial t}\Phi(t,y) = H(t,\Phi(t,y)) \ .$$

Pour $(t,y) \in I \times U$, on note $D^{(2)}H(t,y)$ la différentielle partielle de l'application H dans la direction de E.

4.2.

THÉORÈME 2.

1) *V est une partie ouverte de $\mathbf{R} \times E$;*

2) *L'application Φ de V dans E est de classe C^1 ;*

3) *Notons $D^{(2)}\Phi$ la différentielle partielle de Φ dans la direction de E. Soient $y_0 \in U$, $v_0 \in E$; posons $v(t) = D^{(2)}\Phi(y_0,t)(v_0)$ pour $t \in J(y_0)$. Alors l'application $t \mapsto v(t)$ de $J(y_0)$ dans E est la solution maximale de condition initiale (t_0,v_0) de l'équation différentielle **linéaire** (dite **équation aux variations**) :*

$$v' = D^{(2)}H(t,\Phi(t,y_0))v.$$

L'équation aux variations est du type considéré en 3.1, dont les propriétés sont données dans la proposition 2. Rappelons que ses solutions maximales sont définies sur l'intervalle $J(y_0)$ tout entier. Pour chaque $t \in J(y_0)$, l'application $v \mapsto v(t)$ est un isomorphisme de l'espace vectoriel S des solutions de l'équation aux variations sur E. On en déduit que l'endomorphisme $D^{(2)}\Phi(t,y_0) \in \mathcal{L}(E)$ est toujours ***inversible***.

4.3. Démonstration. (esquissée).

1. Soient $y_0 \in U$, $t_1 \in J(y_0)$. Notons L l'intervalle compact d'extrémités t_0 et t_1. Nous supposerons d'abord qu'on a, pour tout $t \in L$:

$$H(t,y_0) = 0 \ ,$$
$$D^{(2)}H(t,y_0) = 0 \ .$$

On a alors

$$\Phi(t,y_0) = y_0 \ , \quad \forall t \in L \ .$$

Posons $\ell = |t_1 - t_0|$, et fixons $\varepsilon > 0$. Comme $D^{(2)}H$ est continue et L est compact, il existe par le théorème de Borel-Lebesgue un nombre $\delta > 0$ tel qu'on ait :

$$\left\|D^{(2)}H(t,y)\right\| \le \varepsilon$$

pour $(t,y) \in L \times \overline{B}(y_0,\delta)$. On a alors, par le théorème de la moyenne :

$$(1) \qquad \|H(t,y_1) - H(t,y_2)\| \leq \varepsilon\ \|y_1 - y_2\| \ ,$$

pour $t \in L$, $y_1, y_2 \in \overline{B}(y_0,\delta)$.

Munissons l'espace $X = \mathcal{C}(L, \overline{B}(y_0,\delta))$ des applications continues de L dans $\overline{B}(y_0,\delta)$ de la distance uniforme. Pour tout $y \in B\left(y_0, \dfrac{\delta}{2}\right)$ et toute application $z \in X$, définissons une application $\mathcal{F}_y(z)$ de L dans E par :

$$\mathcal{F}_y(z)(t) = y + \int_{t_0}^{t} H(s, z(s))\, ds \ .$$

D'après la relation (1), on a, pour $s \in L$:

$$(2) \qquad \|H(s,z(s))\| \leq \varepsilon\ \|z(s) - y_0\| \leq \varepsilon\delta \ .$$

Supposons $\varepsilon < \dfrac{1}{2\ell}$**.** On a alors, pour $t \in L$:

$$\begin{aligned}\|\mathcal{F}_y(z)(t) - y_0\| &\leq \|y - y_0\| + |t - t_0|\ \varepsilon\delta \\ &\leq \frac{\delta}{2} + \frac{\delta}{2} = \delta \ .\end{aligned}$$

D'autre part, $\mathcal{F}_y(z)$ est clairement continue dans L. C'est donc un élément de X et on a défini une application $\mathcal{F}_y$ de X dans X.

Pour $z, \widetilde{z} \in X$, on a, d'après la relation (1) :

$$\|\mathcal{F}_y(z) - \mathcal{F}_y(\widetilde{z})\| \leq \varepsilon\ell\ \|z - \widetilde{z}\| \ ,$$

donc $\mathcal{F}_y$ est une application contractante de l'espace métrique complet X dans lui même. Son unique point fixe Z_y est l'application $t \mapsto \Phi(t,y)$ de L dans $\overline{B}(y_0,\delta)$ solution de $(*)$ de condition initiale (t_0, y) (cf. Chapitre X, §.2.2, Lemme). Soit $z_0 \in X$ l'application constante de valeur y. On a :

$$\lim_{n\to+\infty} \mathcal{F}_y^n(z_0) = Z_y \ ,$$

et plus précisément, comme l'application $\mathcal{F}_y$ est $\varepsilon\ell-$lipschitzienne :

$$\left\|\mathcal{F}_y^n(z_0) - Z_y\right\| \leq \frac{(\varepsilon\ell)^n}{1-\varepsilon\ell}\ \|\mathcal{F}_y(z_0) - z_0\| \ .$$

D'après la relation (2) ci-dessus, on a :

$$\|\mathcal{F}_y(z_0) - z_0\| \leq \varepsilon\ell\ \|y - y_0\| \ ,$$

et on conclut qu'on a, pour tout $t \in L$, $y \in B\left(y_0, \dfrac{\delta}{2}\right)$:

$$\|\Phi(t,y) - y\| \leq \frac{\varepsilon\ell}{1-\varepsilon\ell}\ \|y - y_0\| \ .$$

On a donc démontré, dans le cas particulier considéré, que

– V contient un "tube" de la forme $L \times B\left(y_0, \frac{\delta}{2}\right)$.

– Pour tout $t \in L$, l'application $y \mapsto \Phi(t,y)$ de $B\left(y_0, \frac{\delta}{2}\right)$ dans E est différentiable au point y_0, sa différentielle étant l'application identique de E dans E.

2. Passons au cas général.

Soient de nouveau $y_0 \in U$, $t_1 \in J(y_0)$. Pour $t \in J(y_0)$, posons $A(t) = D^{(2)}H(t, \Phi(t, y_0))$ et considérons l'équation aux variations :

$$\text{(V)} \qquad v' = A(t)v \ ,$$

pour une application v de $J(y_0)$ dans E. D'après la proposition 2 de §.3.1, il existe une application de classe C^1, notée B, de $J(y_0)$ dans $GL(E)$, telle que toute solution maximale v de (V) satisfait :

$$v(t) = B(t)v(t_0) \ , \quad \forall t \in J(y_0) \ .$$

On a donc :

$$\text{(3)} \qquad B(t_0) = \mathrm{id}_E \ ,$$

$$\text{(4)} \qquad B'(t) = A(t)B(t) \ , \quad \forall t \in J(y_0) \ .$$

Il existe $\rho > 0$ et un intervalle ouvert J contenant t_0 et t_1 tels que la formule :

$$\text{(5)} \qquad K(t,w) = H(t, \Phi(t,y_0) + B(t)w) - H(t, \Phi(t,y_0)) - A(t)B(t)w$$

définisse une application K de $J \times B(0,\rho)$ dans E.

Nous appliquons maintenant à l'équation différentielle

$$(*)' \qquad w' = K(t,w)$$

la première partie de la démonstration. On a en effet :

$$K(t,0) = 0 \ , \qquad \forall t \in J \ ,$$
$$D^{(2)}K(t,0) = 0 \ , \qquad \forall t \in J \ .$$

[**Remarque** : L'application $t \mapsto A(t)$ est a priori seulement continue, donc K n'est pas nécessairement de classe C^1. Mais on observera qu'on a utilisé dans la première partie de la démonstration uniquement le fait que H est différentiable **dans la direction de E**, et que la différentielle partielle correspondante est continue dans $I \times U$. Ces hypothèses sont vérifiées par K dans $J \times B(0,\rho)$. Elles sont d'ailleurs suffisantes pour garantir la validité du théorème 2].

Soit L un intervalle compact contenu dans J, et contenant t_0 et t_1 dans son intérieur. D'après la partie 1, il existe $\rho > \delta > 0$ et une application Ψ de $L \times B(0,\delta)$ dans E tels que :

– pour tout $w_0 \in B(0,\delta)$, l'application $t \mapsto \Psi(t,w_0)$ est solution de $(*)'$, définie sur L, de condition initiale (t_0,w_0) ;

– pour tout $t \in L$, l'application $w \mapsto \Psi(t,w)$ est différentiable en 0, sa différentielle étant l'application identique de E.

Un calcul immédiat montre alors que l'application $t \mapsto \Phi(t,y_0) + B(t)\Psi(t,w)$ est une solution définie sur L de l'équation $(*)$, de condition initiale (t_0,y_0+w). On a donc, pour $w \in B(0,\delta)$, $t \in L$:

$$\Phi(t,y_0+w) = \Phi(t,y_0) + B(t)\Psi(t,w) \ .$$

On conclut que le domaine de définition V de Φ contient $L \times B(y_0,\delta)$. Comme y_0 et t_1 étaient arbitraires, cela montre que V est ouvert.

On conclut aussi que Φ possède au point (t_1,y_0) une différentielle partielle dans la direction de E, égale à $B(t_1)$. On a donc démontré les conclusions 1 et 3 du théorème.

On laisse au lecteur consciencieux le soin de vérifier que les différentielles partielles

$$\frac{\partial}{\partial t}\Phi(t,y) = H(t,\Phi(t,y)) \quad \text{et} \quad D^{(2)}\Phi(t,y)$$

sont continues dans V, ce qui termine la démonstration du théorème.

4.4. On trouvera au chapitre suivant des applications importantes du théorème 2.

Le théorème 2 s'applique en particulier lorsqu'on considère une équation différentielle dépendant (différentiablement) de paramètres; considérons en effet une famille :

$$y' = H(t,y,p)$$

d'équations différentielles dépendant d'un paramètre p variant dans un ouvert W d'un espace vectoriel de dimension finie P. On garde les notations de 4.1, et on suppose que H est de classe C^1 dans $I \times U \times W$.

La famille précédente peut être considérée comme une unique équation différentielle dans l'espace produit $E \times P$:

$$\begin{cases} y' = H(t,y,p) \\ p' = 0 \end{cases}$$

à laquelle on applique le théorème 2.

Soit $t_0 \in I$ un instant initial. Pour $y_0 \in U$, $p_0 \in W$, notons $t \mapsto \Phi(t,y_0,p_0)$ la solution maximale de

$$(*_{p_0}) \qquad y' = H(t,y,p_0)$$

de condition initiale (t_0,y_0), et notons $J(y_0,p_0)$ son domaine de définition. Le domaine de Φ est donc :

$$V = \Big\{(t,y,p) \in I \times U \times W, \quad t \in J(y,p)\Big\} \ .$$

On a alors, d'après le théorème 2, les conclusions suivantes :

1. V est ouvert dans $\mathbf{R} \times E \times P$, et Φ est de classe C^1 dans V.

2. Notons $D^{(2)}H$, $D^{(3)}H$ les différentielles partielles de H dans les directions de E et P, $D^{(3)}\Phi$ la différentielle partielle de Φ dans la direction de P.

Soient $y_0 \in U$, $p_0 \in W$, $q_0 \in P$; posons, pour $t \in J(y_0)$:

$$w(t) = D^{(3)}\Phi(t, y_0, p_0)(q_0) \in E ,$$

$$A(t) = D^{(2)}H\Big(t, \Phi(t, y_0, p_0), p_0\Big) \in \mathcal{L}(E) ,$$

$$b(t) = D^{(3)}H\Big(t, \Phi(t, y_0, p_0), p_0\Big)(q_0) \in E .$$

Alors, l'application w est la solution maximale, de condition initiale $(t_0, 0)$ de l'équation linéaire avec second membre :

$$w'(t) = A(t)\,w(t) + b(t) ,$$

(équation aux variations du paramètre).

CHAPITRE XII : Champs de vecteurs.

§. 1. La propriété de flot.

1.1. On considère dans ce chapitre une équation différentielle **autonome** du 1er ordre :

$$(*) \qquad y' = H(y)$$

définie par un **champ de vecteurs** H de classe C^1 dans un ouvert U de l'espace vectoriel de dimension finie E.

La propriété suivante est évidente mais **fondamentale** :

PROPOSITION 1. *Pour $t_0 \in \mathbf{R}$, notons T_{t_0} la translation $t \mapsto t + t_0$ de $\mathbf{R}$. Soit $t_0 \in \mathbf{R}$. Si y est une solution de $(*)$, définie sur un intervalle J, alors l'application $t \mapsto y(t - t_0)$ définie sur l'intervalle $T_{t_0}(J)$, est aussi une solution de $(*)$.*

DÉMONSTRATION : Pour $t \in T_{t_0}(J)$, posons $z(t) = y(t - t_0)$. Comme y est de classe C^1 dans J, z est de classe C^1 dans $T_{t_0}(J)$ et on a :

$$z'(t) = y'(t - t_0) = H(y(t - t_0)) = H(z(t)) \ .$$

□

COROLLAIRE. *Soient $t_0 \in \mathbf{R}$, $y_0 \in U$. Notons y la solution maximale de $(*)$ de condition initiale $(0, y_0)$, et J le domaine de définition de y. La solution maximale de $(*)$ de condition initiale (t_0, y_0) est alors l'application $t \mapsto y(t - t_0)$, et son domaine de définition est l'intervalle $T_{t_0}(J)$.*

DÉMONSTRATION : Soient $\bar{z}$ la solution maximale de $(*)$ de condition initiale (t_0, y_0), J' son domaine de définition. D'après la proposition 1, on a $z(t) = y(t - t_0)$ pour $t \in T_{t_0}(J)$ et $T_{t_0}(J) \subset J'$. Si cette inclusion était stricte, l'application $t \mapsto z(t + t_0)$ serait une solution de $(*)$ de condition initiale $(0, y_0)$ définie sur un intervalle $T_{-t_0}(J')$ contenant strictement J. □

1.2. Flot de l'équation.

Nous allons reprendre, pour l'équation $(*)$ définie par le champ de vecteurs H, certaines des conclusions du théorème 2 du Chapitre XI, §.4.

Nous choisissons une fois pour toutes comme instant initial $t_0 = 0$.

Pour $y_0 \in U$, nous notons $t \to \Phi(t, y_0)$ la solution maximale de $(*)$ de condition initiale $(0, y_0)$, et notons $J(y_0)$ son intervalle de définition. Le domaine de définition de Φ est donc :

$$V = \Big\{(t,y) \in \mathbf{R} \times U,\ t \in J(y)\Big\} .$$

Pour $t \in \mathbf{R}$ fixé, nous notons Φ_t l'application $y \mapsto \Phi(t,y)$. Le domaine de définition de Φ_t est donc :

$$U_t = \Big\{y \in U,\ t \in J(y)\Big\}$$

c'est-à-dire l'ensemble des positions initiales (à l'instant $t_0 = 0$) pour lesquelles existe une solution de $(*)$ définie sur $[0,t]$ (pour $t \geq 0$) ou $[t,0]$ (pour $t \leq 0$). On peut écrire :

$$V = \Big\{(t,y) \in \mathbf{R} \times U,\ y \in U_t\Big\} .$$

THÉORÈME 1.

1. *V est ouvert dans $\mathbf{R} \times E$, et Φ est de classe C^1 dans V. Pour tout réel t, U_t est ouvert dans E.*

2. *On a $U_0 = U$, et $\Phi_0 = \mathrm{id}_U$.*

3. *Soit $t \in \mathbf{R}$. Alors Φ_t est un difféomorphisme de classe C^1 de U_t sur U_{-t}, dont l'inverse est Φ_{-t}.*

4. *Soient $s,t \in \mathbf{R}$. Dans la partie ouverte $\Phi_t^{-1}(U_s)$ de U, on a :*

$$\Phi_{s+t} = \Phi_s \circ \Phi_t .$$

DÉMONSTRATION : La partie 1 résulte du théorème 2 du Chapitre XI, §.4.2, et la partie 2 est évidente.

Soit $t \in \mathbf{R}$; d'après la partie 1 l'application Φ_t est de classe C^1 dans l'ouvert U_t.

Soit $y \in U_t$; d'après la proposition 1, on a :

$$J(\Phi_t(y)) = T_{-t}(J(y)) , \tag{1}$$

$$\Phi_s(\Phi_t(y)) = \Phi_{s+t}(y) , \ \forall s \in J(\Phi_t(y)) . \tag{2}$$

Comme $0 \in J(y)$, on a $-t \in J(\Phi_t(y))$, c'est-à-dire que $\Phi_t(y) \in U_{-t}$, et

$$\Phi_{-t}(\Phi_t(y)) = \Phi_0(y) = y .$$

En remplaçant t par $-t$, on voit que Φ_{-t} est une application de classe C^1 de U_{-t} dans U_t, vérifiant $\Phi_t(\Phi_{-t}(\tilde{y})) = \tilde{y}$ pour $\tilde{y} \in U_{-t}$. Donc Φ_t est un difféomorphisme de classe C^1 de U_t sur U_{-t} et son inverse est Φ_{-t}.

Soit $s \in \mathbf{R}$, tel que $\Phi_t(y) \in U_s$. On a donc $s \in J(\Phi_t(y))$, d'où la partie 4 d'après la relation (2).

1.3. Champs de vecteurs complets.

On dit que le champ de vecteurs défini par H est **complet** si toute solution maximale de l'équation différentielle $(*)$ est définie sur la droite réelle toute entière.

Supposons qu'il en est ainsi. On a alors :

$$V = \mathbf{R} \times U \ ,$$
$$U_t = U\,, \quad \forall t \in \mathbf{R}\ .$$

Par ailleurs, notons $\mathrm{Diff}^1(U)$ l'ensemble des difféomorphismes de classe C^1 de U sur U.

Muni de la composition, $\mathrm{Diff}^1(U)$ est un groupe.

La partie 3 du théorème 1 montre que, pour tout réel t, l'application Φ_t appartient à $\mathrm{Diff}^1(U)$.

Les parties 2, 3 et 4 du théorème 1 montrent alors que l'application $t \mapsto \Phi_t$ de $\mathbf{R}$ dans $\mathrm{Diff}^1(U)$ est un homomorphisme du groupe (additif) $\mathbf{R}$ dans $\mathrm{Diff}^1(U)$.

§. 2. Singularités.

2.1. DÉFINITION 1. On dit qu'un point $a \in U$ est une **singularité** (ou un **point singulier**) du champ de vecteurs défini par H si on a $H(a) = 0$.

Soit $a \in U$. Si a est une singularité du champ de vecteurs défini par H, la solution maximale de $(*)$ de condition initiale $(0, a)$ est définie sur tout $\mathbf{R}$ et constamment égale à a. On a donc $a \in U_t$ et $\Phi_t(a) = a$ pour tout réel t.

Si au contraire a n'est pas une singularité du champ de vecteurs, le point $\Phi_t(a)$ n'est pour aucune valeur de $t \in J(a)$ une singularité. La courbe décrite par $\Phi_t(a)$, lorsque t décrit $J(a)$, est appelée **orbite** ou **trajectoire** du point a ; elle admet en chacun de ses points une tangente engendrée par la valeur de H en ce point.

Lorsque l'équation différentielle décrit un phénomène physique, une singularité correspond à la notion physique d'**équilibre**.

PROPOSITION 2. *Soit $y \in U$. Supposons qu'on ait $y \in U_t$ pour tout $t \geq 0$ et qu'il existe $a \in U$ tel que :*

$$\lim_{t \to +\infty} \Phi_t(y) = a\ .$$

Alors a est une singularité du champ de vecteurs défini par H. De même, si on suppose qu'on a $y \in U_t$ pour tout $t \leq 0$ et qu'il existe $a' \in U$ tel que :

$$\lim_{t \to -\infty} \Phi_t(y) = a'\ ,$$

alors a' est une singularité du champ de vecteurs.

DÉFINITION 2. Sous les hypothèses de la proposition, on dit que y appartient à l'ensemble **stable** de la singularité a, ou à l'ensemble **instable** de la singularité a'.

DÉMONSTRATION : Supposons que $y \in U_t$ pour tout $t \geq 0$ et qu'on ait :

$$\lim_{t \to +\infty} \Phi_t(y) = a \ .$$

Soit $s \in \mathbf{R}$. Pour $t \geq \max(0, -s)$ on a :

$$\Phi_s(\Phi_t(y)) = \Phi_{s+t}(y) \ .$$

Pour $|s|$ assez petit, on a $a \in U_s$ et Φ_s est continue donc

$$\Phi_s(a) = \lim_{t \to +\infty} \Phi_s(\Phi_t(y)) = \lim_{t \to +\infty} \Phi_{s+t}(y) = a \ ,$$

donc

$$H(a) = \frac{d}{ds} \Phi_s(a) \big|_{s=0} = 0 \ .$$

L'autre cas se traite de façon analogue.

2.2. Puits et Sources.

Soit a une singularité du champ de vecteurs défini par H. L'**équation aux variations** pour la solution maximale constamment égale à a s'écrit :

$$v' = D_a H(v) \ .$$

C'est une équation différentielle linéaire à **coefficients constants** (cf. Chapitre XI, §.3.4), dite **équation linéarisée** du champ de vecteurs à la singularité a.

DÉFINITION 3. On dit que la singularité a est un **puits** (resp. une **source**) si toutes les valeurs propres (réelles ou complexes) de $D_a H \in \mathcal{L}(E)$ ont une partie réelle **strictement négative** (resp. **strictement positive**).

DÉFINITION 4. On dit que la singularité a est **attractive** (resp. **répulsive**) si l'ensemble stable $W^s(a)$ de a (resp. l'ensemble instable $W^u(a)$ de a) contient une boule ouverte centrée en a.

PROPOSITION 3. *Soit a une singularité du champ de vecteurs. Si a est attractive (resp. répulsive) alors $W^s(a)$ (resp. $W^u(a)$) est une partie ouverte de U.*

DÉMONSTRATION : Commençons par un lemme.

LEMME. *Soient a une singularité du champ de vecteurs, $W^s(a)$ son ensemble stable, $W^u(a)$ son ensemble instable, $\delta > 0$ tel que $B(a,\delta) \subset U$. On a alors :*

$$W^s(a) = \bigcup_{t \geq 0} \Phi_t^{-1}\Big(W^s(a) \cap B(a,\delta)\Big)$$

$$W^u(a) = \bigcup_{t \leq 0} \Phi_t^{-1}\Big(W^u(a) \cap B(a,\delta)\Big) .$$

DÉMONSTRATION DU LEMME : Soit $y \in W^s(a)$. On a $\lim\limits_{t \to +\infty} \Phi_t(y) = a$, donc il existe $t_0 \geq 0$ tel que $z = \Phi_{t_0}(y) \in B(a,\delta)$. On a $\Phi_s(z) = \Phi_{s+t_0}(y)$ pour $s \geq 0$, donc $z \in W^s(a)$ et $y \in \Phi_{t_0}^{-1}\Big(W^s(a) \cap B(a,\delta)\Big)$. Inversement, soient $t_0 \geq 0$ et $y \in U$ tels que $y \in \Phi_{t_0}^{-1}\Big(W^s(a) \cap B(a,\delta)\Big)$. On a donc $\Phi_{t_0}(y) = z \in W^s(a) \cap B(a,\delta)$, donc $y \in U_t$ pour tout $t \geq 0$ et $\Phi_t(y) = \Phi_{t-t_0}(z)$ pour $t \geq 0$ d'où $y \in W^s(a)$.

La démonstration pour $W^u(a)$ est similaire. □

La proposition est conséquence immédiate du lemme : si a est attractive on peut choisir $\delta > 0$ de façon que $W^s(a) \supset B(a,\delta)$; on a alors :

$$W^s(a) = \bigcup_{t \geq 0} \Phi_t^{-1}(B(a,\delta)) ,$$

et $W^s(a)$ est une partie ouverte de U. De même si a est répulsive. □

REMARQUE. Considérons l'équation différentielle

$$(\check{*}) \qquad y' = -H(y) .$$

Notons $(\check{\Phi}_t)_{t \in \mathbb{R}}$ le flot de cette équation, $\check{U}_t$ le domaine de $\check{\Phi}_t$. Si y est une solution de $(*)$, définie sur un intervalle J, alors l'application $\check{y}(t) = y(-t)$, définie sur $-J$, est solution de $(\check{*})$. On a donc, pour tout réel t :

$$\check{U}_t = U_{-t} , \quad \check{\Phi}_t = \Phi_{-t} .$$

Le passage de $(*)$ à $(\check{*})$ (appelé changement de sens du temps) transforme singularités de $(*)$ en singularités de $(\check{*})$, puits de $(*)$ en sources de $(\check{*})$ (et vice versa), singularités attractives de $(*)$ en singularités répulsives de $(\check{*})$ (et vice versa), ensemble stable d'une singularité pour $(*)$ en ensemble instable pour $(\check{*})$ (et vice versa).

THÉORÈME 2. *Un puits est une singularité attractive. Une source est une singularité répulsive.*

DÉMONSTRATION : Bien que le théorème soit valable sans restrictions, nous n'allons indiquer la démonstration que dans le cas où l'endomorphisme $D_a H \in \mathcal{L}(E)$ (où a est le puits considéré) est diagonalisable sur $\mathbf{C}$. Nous traitons le cas des puits, le cas des sources s'en déduisant en changeant le sens du temps.

Il existe donc par hypothèse des nombres réels strictement négatifs $\lambda_1, \ldots, \lambda_r$ des nombres complexes $\mu_1, \ldots, \mu_s$ de partie réelle strictement négative, et une identification de E et $\mathbf{R}^r \times \mathbf{C}^s$ tels que, si l'on note $(x_1, \ldots, x_r, z_1, \ldots, z_s)$ les coordonnées dans $\mathbf{R}^r \times \mathbf{C}^s$, on ait :

$$D_a H(x_1, \ldots, x_r, z_1, \ldots, z_s) = (\lambda_1 x_1, \ldots, \lambda_r x_r, \mu_1 z_1, \ldots, \mu_s z_s) \ .$$

Munissons $\mathbf{R}^r \times \mathbf{C}^s$ de la norme euclidienne : pour $y = (x_1, \ldots, z_s)$

$$\|y\|^2 = \sum_{i=1}^{r} x_i^2 + \sum_{j=1}^{s} |z_j|^2 \ .$$

On note $\langle \ , \ \rangle$ le produit scalaire associé. Soit $y \in \mathbf{R}^r \times \mathbf{C}^s$. On a :

$$\begin{aligned} \langle y, D_a H(y) \rangle &= \sum_{i=1}^{r} \lambda_i x_i^2 + \sum_{j=1}^{s} \operatorname{Re} \mu_j \, |z_j|^2 \\ &\leq c_0 \, \|y\|^2 \ , \end{aligned}$$

où on a désigné par c_0 le nombre :

$$c_0 = \max_{\substack{1 \leq i \leq r \\ 1 \leq j \leq s}} (\lambda_i, \operatorname{Re} \mu_j) < 0 \ .$$

Soit c un nombre réel dans $]c_0, 0[$. On a $H(a) = 0$ donc il existe un nombre $\delta > 0$ tel que pour $\|y\| < \delta$ on ait $y \in U$ et :

$$\|H(a+y) - D_a H(y)\| \leq (c - c_0) \, \|y\| \ ,$$

d'où on déduit, pour $\|y\| < \delta$

$$\begin{aligned} \langle H(a+y), y \rangle &= \langle D_a H(y), y \rangle + \langle H(a+y) - D_a H(y), y \rangle \\ &\leq c_0 \, \|y\|^2 + (c - c_0) \, \|y\|^2 = c \, \|y\|^2 \ . \end{aligned}$$

Considérons le champ de vecteurs défini par H dans l'ouvert $\widetilde{U} = B(a, \delta)$. Soient $y_0 \in \widetilde{U}$, y la solution maximale (dans $\widetilde{U}$) de condition initiale $(0, y_0)$, J le domaine de définition de y. Pour $t \in J$, posons :

$$\varphi(t) = e^{-2ct} \, \|y(t) - a\|^2 \ .$$

L'application φ est de classe C^1 dans J et on a, pour $t \in J$:

$$\begin{aligned} \varphi'(t) &= 2 \ e^{-2ct} \langle y'(t), y(t) - a \rangle - 2c \ e^{-2ct} \, \|y(t) - a\|^2 \\ &= 2 \ e^{-2ct} \Big[\langle H(y(t)), y(t) - a \rangle - c \, \|y(t) - a\|^2 \Big] \\ &\leq 0 \ , \quad \text{puisque} \ \ \|y(t) - a\| < \delta \ . \end{aligned}$$

On a donc, pour tout $t \geq 0$ dans J :

$$\varphi(t) \leq \varphi(0) = \|y_0 - a\|^2 \ ,$$

c'est-à-dire, pour $t \geq 0$, $t \in J$:

$$\|y(t) - a\| \leq e^{ct} \, \|y_0 - a\| \ .$$

D'après le théorème 2 du Chapitre X, ceci n'est possible (comme $c < 0$) que si $J \supset \mathbf{R}^+$. On a donc $y_0 \in U_t$ pour tout $t \geq 0$, $\Phi_t(y_0) \in B(a,\delta) = \widetilde{U}$ pour tout $t \geq 0$ et

$$\|\Phi_t(y_0) - a\| \leq e^{ct}\,\|y_0 - a\|$$

pour tout $t \geq 0$, ce qui montre bien que $y_0 \in W^s(a)$. On a donc $B(a,\delta) \subset W^s(a)$. □

REMARQUE. On a démontré (dans le cas particulier où D_aH est diagonalisable sur $\mathbf{C}$) une estimation :

$$\|\Phi_t(y) - a\| \leq e^{ct}\,\|y - a\| \ ,$$

valable pour $\|y - a\| < \delta$, une certaine constante $c < 0$ et tout temps $t \geq 0$. Une telle estimation est encore valable (pour $c < 0$, $\delta > 0$ adéquats) sans l'hypothèse de diagonalisabilité.

Inversement, on peut montrer que si un point $a \in U$ possède la propriété suivante : il existe $c < 0$, $\delta > 0$, $C > 0$ tels que pour tout $t \geq 0$, $y \in B(a,\delta)$ on ait $y \in U_t$ et

$$\|\Phi_t(y) - a\| \leq C\ e^{ct}\,\|y - a\| \ ,$$

alors a est un **puits**.

2.3. Selles en dimension 2.

DÉFINITION 5. On dit que la singularité a est une selle si l'endomorphisme $D_aH \in \mathcal{L}(E)$ possède au moins une valeur propre de partie réelle strictement positive, au moins une valeur propre de partie réelle strictement négative, et pas de valeurs propres nulles ou imaginaires pures.

Nous supposons dans la suite de ce numéro que E est de **dimension 2**. Soit a une selle du champ de vecteurs défini par H.

L'endomorphisme $D_aH \in \mathcal{L}(E)$ possède alors une valeur propre réelle λ_u strictement positive, et une valeur propre réelle λ_s strictement négative. Notons E_u, E_s les sous-espaces propres (de dimension 1), associés à λ_u, λ_s.

Nous souhaitons étudier les orbites du flot au voisinage du point a. Pour ceci, nous choisissons un repère dans E d'origine a, et d'axes parallèles à E_u, E_s. Nous notons x_u, x_s les coordonnées dans ce repère, de sorte que la matrice de l'endomorphisme D_aH dans cette base est $\begin{pmatrix} \lambda_u & 0 \\ 0 & \lambda_s \end{pmatrix}$.

Munissons E de la norme $\|x\| = \sup(|x_s|, |x_u|)$, et choisissons $\delta > 0$ de façon que le carré $R = \overline{B}(a,\delta)$ soit contenu dans U et qu'on ait, pour $x \in R$:

$$\|D_xH - D_aH\| < \frac{c}{10} \tag{1}$$

où on a posé $c = \min(\lambda_u, -\lambda_s) > 0$.

Soit T_1 le triangle $\left\{x \in R, |x_s| \leq x_u\right\}$. Posons :

$$\partial_1 R = \left\{x \in R, x_u = \delta\right\}$$

$$\partial^+ T_1 = \left\{x \in R, x_s = x_u > 0\right\}, \ \partial^- T_1 = \left\{x \in R, x_u = -x_s > 0\right\} .$$

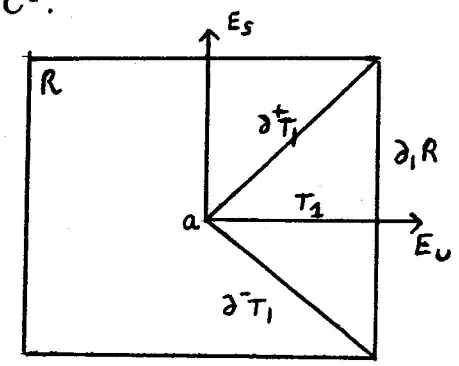
LEMME 1. *Soit b un point de T_1 distinct de a ; notons $A(b)$ la composante connexe de 0 dans l'ensemble $\left\{t \in \mathbf{R}, \Phi_t(b) \in T_1\right\}$.*

1) *La borne supérieure $t^+ = t^+(b)$ de $A(b)$ est finie, et on a $\Phi_{t^+}(b) \in \partial_1 R$;*

2) *la borne inférieure $t^- = t^-(b)$ de $A(b)$ est soit $-\infty$, auquel cas $b \in W^u(a)$, soit finie, auquel cas $\Phi_{t^-}(b) \in \partial^+ T_1 \cup \partial^- T_1$;*

3) *les applications $b \mapsto \Phi_{t^+}(b)$, $b \mapsto \Phi_{t^-}(b)$ sont continues dans $T_1 - \{a\}$; la restriction de $b \mapsto \Phi_{t^+}(b)$ à $\partial^+ T_1 \cup \partial^- T_1$ est injective;*

4) *la courbe paramétrée $t \mapsto \Phi_t(b)$, $t \in A(b)$ est le graphe d'une application de classe C^1.*



DÉMONSTRATION : Dans T_1, on a, d'après (1) et le théorème de la moyenne :

$$|H_u(x) - \lambda_u x_u| \leq \frac{c}{10} x_u ,$$

d'où

$$H_u(x) \geq \frac{9}{10} \lambda_u x_u . \tag{2}$$

Posons $\Phi_t(b) = (x_u(t), x_s(t))$ pour $t \in A(b)$. On a donc $x'_u(t) \geq \frac{9}{10} \lambda_u x_u(t) > 0$ pour $t \in A(b)$, d'où la quatrième assertion du lemme. On a $x_u(t) \geq b_u \left(1 + \frac{9}{10} \lambda_u t\right)$ pour $t \geq 0$, donc t^+ est fini. De même, si $t^- = -\infty$, on doit avoir $b \in W^u(a)$.

Soit $x \in \partial^+ T_1$; d'après (1) et le théorème de la moyenne on a $H_u(x) > 0$, $H_s(x) < 0$, donc il existe $\varepsilon > 0$ tel que $\Phi_t(x) \notin T_1$ pour $t \in [-\varepsilon, 0[$. De même si $x \in \partial^- T_1$. Le point $\Phi_{t^+}(b)$ ne peut donc (sauf si $b = (\delta,\delta)$ ou $b = (\delta,-\delta)$) appartenir à $\partial^+ T_1$ ou $\partial^- T_1$; comme il appartient au bord de T_1, il appartient à $\partial_1 R$.

Soit $x \in \partial_1 R$; on a $H_u(x) > 0$ donc il existe $\varepsilon > 0$ tel que $\Phi_t(x) \notin T_1$ pour $t \in]0,\varepsilon]$. Le point $\Phi_{t^-}(b)$ (si t^- est fini) appartient donc à $\partial^+ T_1 \cup \partial^- T_1$.

Les considérations précédentes montrent aussi que le point $\Phi_t(b)$ appartient à l'intérieur de T_1 lorsque t appartient à l'intérieur de $A(b)$.

L'injectivité de $b \mapsto \Phi_{t^+}(b)$ sur $\partial^+ T_1 \cup \partial^- T_1$ est immédiate.

La continuité des applications $b \mapsto \Phi_{t^+}(b)$, $b \mapsto \Phi_{t^-}(b)$ est laissée en exercice (utiliser le théorème des fonctions implicites). □

LEMME 2. *Soient $\delta_1 \in]0,\delta[$, $-\delta_1 \le y_1 < y_2 \le \delta_1$; notons φ_1, φ_2 les applications de classe C^1 de $[\delta_1,\delta] \subset E_u$ dans E_s telles que pour $i = 1$ ou 2 :*

$$\left\{\Phi_t(\delta_1, y_i), 0 \le t \le t^+(\delta_1, y_i)\right\} = \left\{(z, \varphi_i(z)), z \in [\delta_1, \delta]\right\} .$$

Alors l'application $\varphi_2 - \varphi_1$ est strictement positive et strictement décroissante dans $[\delta_1, \delta]$.

DÉMONSTRATION : On a $\varphi_2(\delta_1) = y_2 > y_1 = \varphi_1(\delta_1)$; comme $\varphi_2 - \varphi_1$ ne peut s'annuler sans être identiquement nulle (unicité des solutions), $\varphi_2 - \varphi_1$ est strictement positive dans $[\delta_1, \delta]$ (théorème des valeurs intermédiaires).

Soit $z \in [\delta_1, \delta]$; on a, pour $i = 1, 2$:

$$\varphi_i'(z) = \frac{H_s(z, \varphi_i(z))}{H_u(z, \varphi_i(z))} .$$

Pour $w \in [-z, z]$, on a, d'après (1) et le théorème de la moyenne :

$$\begin{aligned} \partial_s H_s(z,w) &\le \frac{9}{10}\lambda_s , \\ |H_s(z,w)| &\le \frac{11}{10}\,|\lambda_s|\, z , \\ |\partial_s H_u(z,w)| &\le \frac{1}{10}\lambda_u , \\ H_u(z,w) &\ge \frac{9}{10}\lambda_u z , \end{aligned}$$

d'où

$$H_u(z,w)\,\partial_s H_s(z,w) - H_s(z,w)\,\partial_s H_u(z,w) \le \frac{7}{10}\lambda_s \lambda_u z < 0 ,$$

l'application $w \mapsto H_s(z,w)\big/H_u(z,w)$ est donc strictement décroissante dans $[-z, +z]$, et on a :

$$\varphi_1'(z) > \varphi_2'(z) .$$

□

LEMME 3. *Il existe une application φ de classe C^1 de $[0,\delta] \subset E_u$ dans E_s possédant les propriétés suivantes :*

(i) $\varphi(0) = \varphi'(0) = 0$;

(ii) *le graphe de φ est contenu dans T_1, et ne coupe pas $\partial^+ T_1 \cup \partial^- T_1$;*

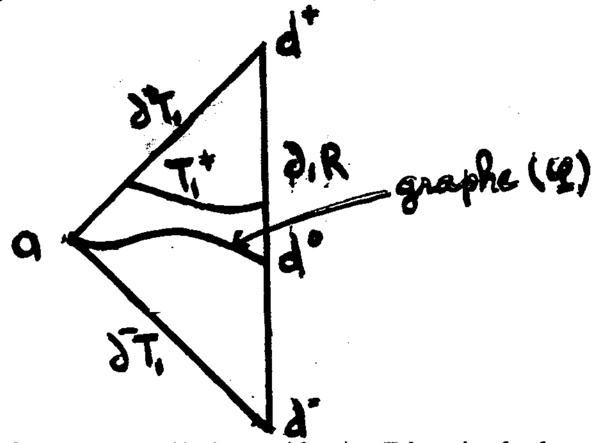
(iii) *posons $d^+ = (\delta,\delta)$, $d^- = (\delta,-\delta)$, $d^0 = (\delta,\varphi(\delta))$, $T_1^+ = \left\{x \in T_1, x_s > \varphi(x_u)\right\}$, $T_1^- = \left\{x \in T_1, x_s < \varphi(x_u)\right\}$.*

Si $b \in T_1^+$, on a $\Phi_{t^+}(b) \in \left]d^0,d^+\right]$, $t^- > -\infty$ et $\Phi_{t^-}(b) \in \partial^+ T_1$.

Si $b \in T_1^-$, on a $\Phi_{t^-}(b) \in \left]d^0,d^-\right]$, $t^- > -\infty$ et $\Phi_{t^-}(b) \in \partial^- T_1$.

Si $b = (b_u,\varphi(b_u))$ avec $b_u > 0$, on a $\Phi_{t^+}(b) = d^0$ et $t^- = -\infty$, $b \in W^u(a)$.

DÉMONSTRATION :



Pour $w \in [-\delta,+\delta]$, posons $d(w) = (\delta,w)$. D'après le lemme 1, il existe des réels w^-, w^+, avec $-\delta < w^- \le w^+ < \delta$ tels que l'application $b \mapsto \Phi_{t^+(b)}(b)$ soit un homéomorphisme de $\partial^+ T_1$ sur $[d(\delta), d(w^+)[$ et un homémorphisme de $\partial^- T_1$ sur $[d(-\delta), d(w^-)[$. Toujours d'après le lemme 1, il existe des applications φ^+, φ^- de classe C^1 de $]0,\delta] \subset E_u$ dans E_s ayant les propriétés suivantes : les graphes de φ^+ et φ^- sont contenus dans T_1 et ne rencontrent pas $\partial^+ T_1$ ni $\partial^- T_1$, le graphe de φ^+ est $\left\{\Phi_t(d(w^+)), t \le 0\right\}$ et celui de φ^- est $\left\{\Phi_t(d(w^-)), t \le 0\right\}$. On a donc :

$$\lim_{z\to 0} \varphi^+(z) = \lim_{z \to 0} \varphi^-(z) = 0 \ .$$

Or, d'après le lemme 2 la fonction $\varphi^+ - \varphi^-$ est strictement décroissante et strictement positive dans $]0,\delta]$ si $w^+ > w^-$. Ceci est impossible et on a donc $w^+ = w^-$, $\varphi^+ = \varphi^-$. Posons :

$$\begin{aligned} &d^0 = d(w^+) = d(w^-) \ , \\ &\varphi(z) = \varphi^+(z) = \varphi^-(z) \ , \quad 0 < z \le \delta \ , \\ &\varphi(0) = 0 \ . \end{aligned}$$

La fonction φ est continue dans $[0,\delta]$ et satisfait (ii). La propriété (iii) résulte de la définition de φ (si $b = (b_u,\varphi(b_u))$, avec $b_u > 0$) et de l'observation suivante : T_1^+ et T_1^- sont les deux composantes connexes de $T_1 -$graphe (φ) ; si $b \in T_1^+$, on a $\Phi_t(b) \in T_1 -$graphe (φ) pour tout $t \in A(b)$ (par unicité des solutions) donc $\Phi_t(b) \in T_1^+$ pour tout $t \in A(b)$ (car $\{\Phi_t(b), t \in A(b)\}$ est connexe).

Il reste à vérifier que φ est de classe C^1 dans $[0,\delta]$, et qu'on a $\varphi'(0) = 0$. On sait déjà que φ est de classe C^1 dans $]0,\delta]$. Soit $\varepsilon > 0$.

Il existe $0 < \delta_1 = \delta_1(\varepsilon) < \delta$ tel qu'on ait, pour $\|x\| \leq \delta_1$:

$$\|D_x H - D_a H\| < \varepsilon \ . \tag{3}$$

Montrons qu'on a $|\varphi(z)| \leq \frac{\varepsilon}{|\lambda_s|} z$ pour $z \in [0,\delta_1]$.

Sinon, il existerait en effet un intervalle $[\delta_2, \delta_3[$ tel qu'on ait :

$$\begin{aligned} &0 \leq \delta_2 < \delta_3 \leq \delta_1 \\ &|\varphi(z)| > \frac{\varepsilon}{|\lambda_s|} z \ , \quad \forall z \in]\delta_2, \delta_3[\ , \\ &|\varphi(\delta_2)| = \frac{\varepsilon}{|\lambda_s|} \delta_2 \ . \end{aligned}$$

Or d'après la relation (3) et le théorème de la moyenne, on a, pour $|x_s| \leq x_u \leq \delta_1$, $|x_s| \geq \frac{\varepsilon}{|\lambda_s|} x_u$:

$$|H_s(x) - \lambda_s x_s| \leq \varepsilon \, \|x\| \leq |\lambda_s x_s| \ .$$

La fonction $|\varphi|$ serait donc décroissante dans $[\delta_2, \delta_3]$, ce qui est impossible puisqu'on a :

$$|\varphi(\delta_2)| = \frac{\varepsilon}{|\lambda_s|} \delta_2 < \frac{\varepsilon}{|\lambda_s|} \delta_3 \leq |\varphi(\delta_3)| \ .$$

On a donc $|\varphi(z)| \leq \frac{\varepsilon}{|\lambda_s|} z$ pour $z \in [0,\delta_1]$. Ceci implique, d'après (3) et le théorème de la moyenne, pour $z \in]0,\delta_1]$:

$$\begin{aligned} |H_s(z,\varphi(z))| &\leq |\lambda_s|\,|\varphi(z)| + \varepsilon z \leq 2\varepsilon z \ , \\ H_u(z,\varphi(z)) &\geq \lambda_u z - \varepsilon z \ , \end{aligned}$$

donc (si on a choisi $\varepsilon < \lambda_u$)

$$|\varphi'(z)| = \left| \frac{H_s(z,\varphi(z))}{H_u(z,\varphi(z))} \right| \leq \frac{2\varepsilon}{\lambda_u - \varepsilon} \ .$$

Comme $\varepsilon > 0$ est arbitraire, on a bien

$$\lim_{z \to 0} \frac{\varphi(z)}{z} = 0 \ , \quad \lim_{z \to 0} \varphi'(z) = 0 \ ,$$

ce qui termine la démonstration du lemme. □

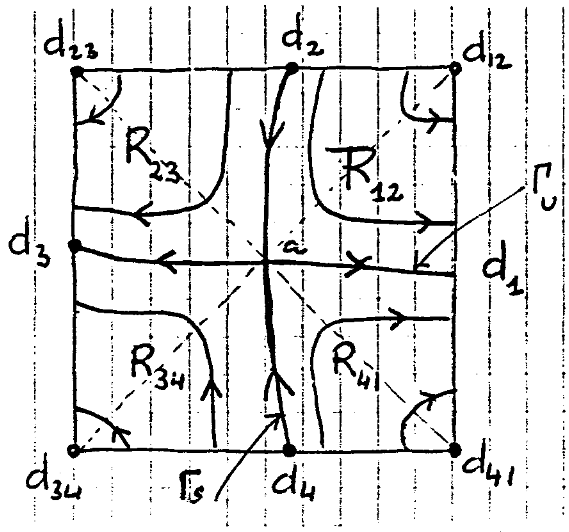
Les lemmes 1, 2 et 3 décrivent les trajectoires du champ de vecteurs dans le triangle T_1.

On peut faire une étude complètement analogue dans le triangle $T_3 = \left\{x \in R, |x_s| \leq -x_u\right\}$ et (en changeant le sens du temps) dans les triangles

$$T_2 = \left\{x \in R, |x_u| \leq x_s\right\} \;,\; T_4 = \left\{x \in R, |x_u| \leq -x_s\right\} .$$

En rassemblant les résultats obtenus, on arrive à une description de toutes les trajectoires du champ de vecteurs dans R.

On prendra garde que les notations dans le théorème ci-dessous sont **différentes** de celles utilisées précédemment.



$$d_{12} = (\delta, \delta), \; d_{23} = (-\delta, \delta)$$
$$d_{34} = (-\delta, -\delta), \; d_{41} = (\delta, -\delta)$$
$$d_1 = (\delta, (\varphi_u(\delta)), \; d_3 = (-\delta, (\varphi_u(-\delta))$$
$$d_2 = (\varphi_s(\delta), \delta), \; d_4 = (\varphi_s(-\delta), -\delta)$$

THÉORÈME 3. *Il existe une courbe Γ_u, graphe d'une application φ_u de classe C^1 de $[-\delta, +\delta] \subset E_u$ dans E_s, et une courbe Γ_s, graphe d'une application de classe C^1 de $[-\delta, +\delta] \subset E_s$ dans E_u qui possèdent les propriétés suivantes (où les points d_i, d_{ij} sont définis ci-dessus) :*

a) *la courbe Γ_u est contenue dans $W^u(a)$, et plus précisément égale à :*

$$\Gamma_u = \left\{\Phi_t(d_1), t \leq 0\right\} \cup \left\{\Phi_t(d_3), t \leq 0\right\} \cup \{a\} .$$

b) *La courbe Γ_s est contenue dans $W^s(a)$, et plus précisément égale à :*

$$\Gamma_s = \left\{\Phi_t(d_2), t \geq 0\right\} \cup \left\{\Phi_t(d_4), t \geq 0\right\} \cup \{a\} .$$

c) *La courbe Γ_u est contenue dans $T_1 \cup T_3$, et est tangente à E_u au point a ; la courbe Γ_s est contenue dans $T_2 \cup T_4$, et est tangente à E_s au point a.*

d) *La partie* $R - (\Gamma_s \cup \Gamma_u)$ *de* R *a 4 composantes connexes, notées* R_{12}, R_{23}, R_{34}, R_{41}. *Pour un point* $b \in R_{12}$, *la composante connexe de* 0 *dans l'ensemble* $\{t, \Phi_t(b) \in R\}$ *est un intervalle compact* $[t_-, t_+]$ *et on a :*

$$\begin{aligned} \Phi_t(b) &\in R_{12} \qquad \text{pour } t_- \leq t \leq t_+ \ , \\ \Phi_{t_-}(b) &\in]d_2, d_{12}] \ , \\ \Phi_{t_+}(b) &\in]d_1, d_{12}] \ . \end{aligned}$$

Un énoncé similaire vaut pour R_{23}, R_{34} *et* R_{41}.

REMARQUE. L'ensemble instable $W^u(a)$ de la singularité a est égal à :

$$W^u(a) = \left\{\Phi_t(d_1), d_1 \in U_t\right\} \cup \left\{\Phi_t(d_3), d_3 \in U_t\right\} \cup \{a\} \ .$$

En effet, il est clair qu'on a :

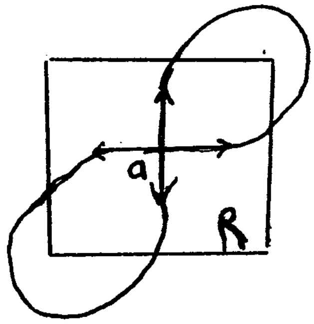
$$\begin{aligned} &a \in W^u(a) \\ &\left\{\Phi_t(d_1), d_1 \in U_t\right\} \subset W^u(a) \\ &\left\{\Phi_t(d_3), d_3 \in U_t\right\} \subset W^u(a) \ . \end{aligned}$$

Inversement, si $x \in W^u(a)$, il existe t_0 tel qu'on ait $\Phi_t(x) \in R$ pour tout $t \leq t_0$. On a alors, d'après le théorème 3, $\Phi_{t_0}(x) \in \Gamma_u$ d'où l'inclusion réciproque.

De façon analogue, on a :

$$W^s(a) = \left\{\Phi_t(d_2), d_2 \in U_t\right\} \cup \left\{\Phi_t(d_4), d_4 \in U_t\right\} \cup \{a\} \ .$$

Il est possible que l'intersection $W^u(a) \cap R$ contienne **strictement** Γ_u, et même qu'on ait $W^u(a) = W^s(a)$.



§. 3. Fonctions de Liapunov.

3.1. Champs gradients.

On suppose dans ce numéro que l'espace vectoriel E est muni d'un produit scalaire $\langle , \rangle$.

DÉFINITION 6. Soit $G : U \to \mathbf{R}$ une fonction de classe C^1. Le **champ de vecteurs gradient** de G est défini par l'unique application $H : U \to E$ qui vérifie, pour tout $y \in U$, $v \in E$:

$$\langle H(y), v\rangle = D_y G(v) .$$

Pour chaque point $y \in U$, l'existence et l'unicité du vecteur $H(y)$ satisfaisant la relation ci-dessus (pour tout $v \in E$) est garantie par le fait que $\langle\ ,\ \rangle$ soit non dégénéré.

Soit $(e_1, \ldots, e_m)$ une base orthonormée de E. Les composantes de H dans cette base sont données par :

$$H_i(y) = \partial_i G(y) , \quad 1 \leq i \leq m, \ y \in U .$$

On suppose que $G : U \to \mathbf{R}$ est une application de classe C^2. Le champ gradient H de G est alors de classe C^1. On considère l'équation différentielle :

$$(*) \qquad y' = +H(y) .$$

PROPOSITION 4.

1) *Soit $y : J \to U$ une solution de $(*)$; pour $t \in J$, on a :*

$$\frac{d}{dt} G(y(t)) = \|H(y(t))\|^2 \geq 0 .$$

2) *Un point $a \in U$ est une singularité de H si et seulement si c'est un point critique de G. Dans ce cas, a est un puits si et seulement si la forme hessienne $D_a^2 G$ est définie négative, a est une source si et seulement si la forme hessienne $D_a^2 G$ est définie positive, c'est une selle si et seulement si la forme hessienne $D_a^2 G$ est non dégénérée et indéfinie.*

DÉMONSTRATION : C'est un calcul facile, laissé en exercice.

REMARQUE. Il est d'usage en physique (en pensant à la pesanteur et à G comme la fonction hauteur) de considérer plutôt l'équation différentielle

$$y' = -H(y) .$$

Les puits de cette équation correspondent alors aux minima locaux non dégénérés de G, ce qui correspond mieux à la notion intuitive de "puits".

3.2. Fonctions de Liapunov. Champs de type gradient.

Peu de champs de vecteurs sont les champs gradients d'une application. On est ainsi amené à généraliser la notion de champ gradient.

Soit une équation différentielle :

$$(*) \qquad y' = H(y)$$

définie par un champ de vecteurs H de classe C^1 dans un ouvert U de E.

DÉFINITION 7. Une application G de classe C^1 de U dans $\mathbf{R}$ est une **fonction de Liapunov** pour le champ de vecteurs H si

1) les singularités de H sont les points critiques de G ;

2) en tout point $y \in U$ qui n'est pas une singularité de H, on a $D_yG(H(y)) > 0$.

EXEMPLES.

1) D'après la proposition 4, 1), le champ gradient d'une application G admet G pour fonction de Liapunov.

2) Soient H un champ de vecteurs admettant une fonction de Liapunov G, et $\lambda : U \to \mathbf{R}$ une fonction de classe C^1 à valeurs strictement positives. Alors G est aussi une fonction de Liapunov pour le champ λH.

3) Soient $G : U \to \mathbf{R}$ une application de classe C^1, H son champ gradient (relativement à un produit scalaire sur E, noté $\langle\ ,\ \rangle$). Si H_1 est un champ de vecteurs dans U admettant les mêmes singularités que H et vérifiant $\langle H(y), H_1(y)\rangle > 0$ pour tout point y qui n'est pas une singularité de H, alors G est une fonction de Liapunov pour H_1.

Inversement, soit H_1 un champ admettant une fonction de Liapunov G. Munissons E d'un produit scalaire $\langle\ ,\ \rangle$, et notons H le champ gradient de G. Les singularités de H coïncident avec celles de H_1, et on a $\langle H(y), H_1(y)\rangle > 0$ en tout point y qui n'est pas une singularité de H.

PROPOSITION 5. *Soit H un champ de vecteurs de classe C^1 admettant une fonction de Liapunov G. Soit $y : J \to U$ une solution non constante de l'équation différentielle :*

$$y' = H(y)\ .$$

Alors l'application $t \to G(y(t))$ est strictement croissante dans J.

DÉMONSTRATION : Comme y est non constante, $y(t)$ n'est pour aucune valeur de $t \in J$ une singularité de H et on a :

$$\frac{d}{dt}G(y(t)) = D_{y(t)}G(y'(t)) = D_{y(t)}G(H(y(t))) > 0\ .$$

□

Un champ de vecteurs H admettant une fonction de Liapunov est dit de type gradient.

On considère dans la fin de ce paragraphe un champ de vecteurs H de classe C^1 admettant une fonction de Liapunov G dans son domaine de définition U. On suppose que H n'a que des **singularités isolées**, c'est-à-dire que pour toute singularité a de H, il existe une boule $B(a, \varepsilon)$ ne contenant pas d'autre singularité de H.

THÉORÈME 4. *Considérons une solution maximale* $y : J \to U$ *de l'équation différentielle :*

$$(*) \qquad y' = H(y) \ .$$

Lorsque le temps t *se rapproche de la borne supérieure de* J, *on a l'alternative suivante :*

1) *ou bien la solution* y *sort de tout compact de* U, *c'est-à-dire que pour toute partie compacte* K *de* U *il existe* $t_0 = t_0(K) \in J$ *tel que* $y(t) \notin K$ *pour* $t \in [t_0, \sup J[$. *C'est la seule possibilité lorsque* $\sup J$ *est fini;*

2) *ou bien* $\sup J$ *est infini et il existe une singularité* a *de* H *telle que* $\lim_{t \to +\infty} y(t) = a$. *La trajectoire considérée est alors contenue dans la variété stable du point* a.

REMARQUES.

1) On a un énoncé similaire lorsque t se rapproche de la borne inférieure de J.

2) L'assertion du théorème lorsque $\sup J$ est fini a été démontrée auparavant (Chapitre X, Théorème 2). On supposera donc dans la démonstration que $\sup J = +\infty$.

DÉMONSTRATION : Munissons E d'une norme. Notons S l'ensemble (fermé) des singularités de H. Pour $1 > r > 0$, posons :

$$K_r = \left\{ x \in U, \ \ \|x\| \le r^{-1}, \ d(x, E - U) \ge r \right\} ,$$

$$L_r = K_r - \bigcup_{a \in S} B(a, r) \ .$$

Ce sont des parties compactes de U.

L'application $t \mapsto G(y(t))$ de J dans $\mathbf{R}$ est croissante et admet donc une limite (dans $\mathbf{R} \cup \{+\infty\}$) lorsque t tend vers $+\infty$. Si on a :

$$\lim_{t \to +\infty} G(y(t)) = +\infty \ ,$$

alors le premier terme de l'alternative du théorème est réalisé : pour toute partie compacte K de U, il suffit de choisir t_0 tel que $G(y(t_0)) > \sup_{x \in K} G(x)$.

On suppose dorénavant que :

$$\lim_{t \to +\infty} G(y(t)) < +\infty \ .$$

LEMME. *Pour tout* $1 > r > 0$, *il existe* $t_0 = t_0(r)$ *tel que* $y(t) \notin L_r$ *pour* $t \ge t_0$.

DÉMONSTRATION : Soit $1 > r > 0$. On a :

$$(1) \qquad d(L_r, E - L_{r/2}) \ge \frac{r}{2} \ .$$

Comme $L_{r/2}$ est compact et ne contient pas de singularités de H, il existe des constantes $C_1, C_2 > 0$ telles qu'on ait, pour tout $y \in L_{r/2}$:

$$D_y G(H(y)) \geq C_1 \ , \tag{2}$$

$$\|H(y)\| \leq C_2 \ . \tag{3}$$

Supposons que pour un temps $t_1 \in J$ on ait $y(t_1) \in L_r$. D'après les relations (1), (3) et le théorème de la moyenne, on a $y(t) \in L_{r/2}$ pour $|t - t_1| \leq \dfrac{r}{2C_2}$. D'après la relation (2), on a, pour $|t - t_1| \leq \dfrac{r}{2C_2}$:

$$\frac{d}{dt} G(y(t)) = D_{y(t)} G(H(y(t))) \geq C_1 \ ,$$

d'où

$$G\left(y\left(t_1 + \frac{r}{2C_2}\right)\right) - G\left(y\left(t_1 - \frac{r}{2C_2}\right)\right) \geq r\frac{C_1}{C_2} \ .$$

Si la conclusion du lemme n'était pas vérifiée, il existerait une suite $(t_n)_{n\geq 1}$ de temps satisfaisant :

$$t_{n+1} - t_n \geq \frac{r}{C_2} \ , \quad \forall n \geq 1 \ ,$$

$$t_n \in J \ , \ y(t_n) \in L_r \ , \quad \forall n \geq 1 \ .$$

On aurait alors pour $n \geq 1$:

$$\begin{aligned} G\left(y\left(t_{n+1} + \frac{r}{2C_2}\right)\right) &\geq \frac{C_1}{C_2} r + G\left(y\left(t_{n+1} - \frac{r}{2C_2}\right)\right) \\ &\geq \frac{C_1}{C_2} r + G\left(y\left(t_n + \frac{r}{2C_2}\right)\right) \ , \end{aligned}$$

d'où $\lim\limits_{t \to +\infty} G(y(t)) = +\infty$. $\square$

Supposons désormais que le premier terme de l'alternative du théorème ne soit pas réalisé. Il existe donc une partie compacte K de U et une suite $(t_n)_{n\geq 1}$ tendant vers $+\infty$ telles que $y(t_i) \in K$ pour tout $i \geq 1$. Il existe $r_0 > 0$ tel que $K \subset K_{r_0}$ (exercice). La partie $S \cap K_{r_0/2}$ est fermée et bornée, donc compacte, et ses points sont isolés. Elle est donc finie, et nous notons $a_1, \ldots, a_N$ ses éléments.

Soit $r_1 \in \left]0, \dfrac{r_0}{4}\right[$ assez petit pour que les boules $(B(a_j, r_1))_{1\leq j \leq N}$ soient disjointes. On a alors $\overline{B}(a_j, r_1) \subset K_{r_1}$, donc les boules $B(a_j, r_1)$, pour $1 \leq j \leq N$, sont des composantes connexes de $E - L_{r_1}$. D'après le lemme, il existe un temps t_0 tel que $y(t) \notin L_{r_1}$ pour $t \geq t_0$. Soit $i \geq 1$ tel que $t_i \geq t_0$. On a :

$$y(t_i) \in K_{r_0} \subset K_{r_1}$$

$$y(t_i) \notin L_{r_1} \ ,$$

donc $y(t_i) \in \bigcup_{a\in S} B(a,r_1)$. Mais, si $a \in S-K_{r_0/2}$, on a $B(a,r_1)\cap K_{r_0} = \phi$. Donc on a :

$$y(t_i) \in \bigcup_{j=1}^{N} B(a_j,r_1) \ .$$

Soit j tel que $y(t_i) \in B(a_j,r_1)$. L'image A par l'application y de l'intervalle $[t_0,+\infty[$ est connexe et contenue dans $E-L_{r_1}$, donc contenue dans une composante connexe de $E-L_{r_1}$. Comme $t_i \geq t_0$, $y(t_i) \in B(a_j,r_1)$ et $B(a_j,r_1)$ est une composante connexe de $E-L_{r_1}$, on a $A \subset B(a_j,r_1)$.

Soit $r \in]0,r_1]$; en appliquant à nouveau le lemme, on voit qu'il existe $t(r) \geq t_0$ tel que $y(t) \notin L_r$ pour $t \geq t(r)$. Mais on a :

$$B(a_j,r_1)\cap(E-L_r) = B(a_j,r) \ ,$$

donc $y(t) \in B(a_j,r)$ pour $t \geq t(r)$. On a donc :

$$\lim_{t\to+\infty} y(t) = a_j \ .$$

□

On est souvent en mesure, dans des situations particulières, d'éliminer le premier terme de l'alternative.

COROLLAIRE 1. *Supposons que la fonction de Liapunov G ait la propriété suivante : pour tout $\lambda \in \mathbf{R}$, $G^{-1}([\lambda,+\infty))$ est une partie compacte de U. Alors, on a $U_t = U$ pour tout $t \geq 0$, et tout point de U appartient à l'ensemble stable d'une singularité de H.*

DÉMONSTRATION : Exercice.

COROLLAIRE 2. *Soit a un maximum local strict de l'application G. Alors a est une singularité attractive du champ de vecteurs H. De même, un minimum local strict de l'application G est une singularité répulsive du champ de vecteurs H.*

DÉMONSTRATION : Il suffit de démontrer la première assertion, l'autre s'en déduisant en changeant le sens du temps et en remplaçant G par $-G$.

Soit $r_0 > 0$ tel qu'on ait :

(i) $\overline{B}(a,r_0) \subset U$,

(ii) $G(y) < G(a)$, pour $\|y-a\| = r_0$,

(iii) a est la seule singularité de H contenue dans $\overline{B}(a,r_0)$.

Soit $r_1 \in]0,r_0[$ tel qu'on ait :

$$\inf_{y\in B(a,r_1)} G(y) > \sup_{\|y-a\|=r_0} G(y) \ .$$

Soit $y_0 \in B(a,r_1)$. Montrons qu'on a $y_0 \in U_t$ pour tout $t \geq 0$ et $\lim_{t\to+\infty} \Phi_t(y_0) = a$.

Pour tout $t \geq 0$ tel que $y_0 \in U_t$, on a :

$$G(\Phi_t(y_0)) \geq G(y_0) > \sup_{\|y-a\|=r_0} G(y) \ ,$$

donc $\|\Phi_t(y_0) - a\| \neq r_0$. Comme l'application $t \to \|\Phi_t(y_0) - a\|$ est continue et prend une valeur strictement inférieure à r_0, on doit donc avoir $\Phi_t(y_0) \in B(a,r_0)$ pour tout $t \geq 0$ tel que $y_0 \in U_t$. D'après le théorème 4, ceci n'est possible que si l'on a $y_0 \in U_t$ pour tout $t \geq 0$ et

$$\lim_{t\to+\infty} \Phi_t(y_0) = a \ .$$

La boule $B(a,r_1)$ est donc contenue dans l'ensemble stable de la singularité a, et celle-ci est donc attractive.

Dépot Légal : 03.94

ISBN 2–87800–037–4
Paris Onze édition *G 36*