Espaces vectoriels normés, premières notions de topologie.

Espace vectoriel normé.

Soit E un espace vectoriel. Soit une fonction $\|.\|:E \to R$ qui prend un vecteur de E et donne un réel. Une telle fonction est dite « norme » si elle valide trois critères :

- si $\|x\| = 0$, x est le vecteur nul (séparation)
- $\|\lambda x\| = |\lambda|\cdot\|x\|$ (homogénéité absolue)
- $\|x+y\| \leq \|x\|+\|y\|$ (inégalité triangulaire)

Si on munit E de cette norme, on obtient un <u>espace vectoriel normé</u>.

Exemples classiques.

Une norme est essentiellement une mesure de la longueur d'un vecteur. La norme $\|\cdot\|_p$ est définie par $\|x\|_p = (|x_1|^p+|x_2|^p+\ldots+|x_n|^p)^{1/p}$. La norme traditionnelle, dite Euclidienne, est le cas p = 2. Deux autres cas bien connus sont $\|x\|_1 = |x_1|+\ldots+|x_n|$, et $\|x\|_\infty = \max(|x_1|, \ldots, |x_n|)$, et donnent la « distance de Manhattan » et la « distance de Tchebychev ».

Espace métrique.

En partant d'un espace vectoriel normé, on peut rajouter la fonction distance définie par $d(x,y) = \|x-y\|$. Grâce à la définition de norme, la distance satisfait :

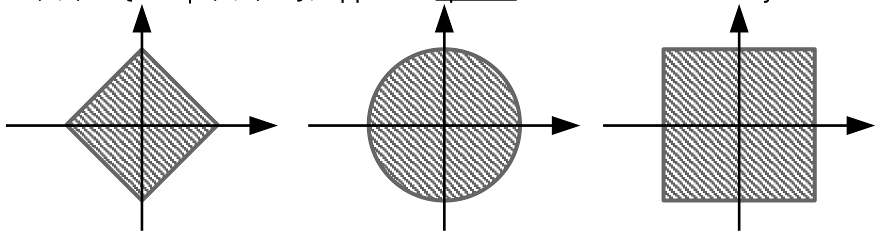
- $d(x,y) = d(y,x)$ pour tout x,y (symétrie)
- $d(x,y) = 0$ ssi $x = y$ (séparation)
- $d(x,y) \leq d(x,z)+d(z,y)$ pour tout x,y,z (inégalité triangulaire).

Toute fonction satisfaisant ces trois critères est une <u>distance</u> sur un ensemble X ; et (X,d) est alors un <u>espace métrique</u>.

Boules et sphères.

Soit (X,d) un espace métrique. Pour tout réel $r > 0$, on a :

- $B(a,r) = \{x\in X | d(a,x)<r\}$, appelé « <u>boule ouverte</u> de centre a et de rayon r ».
- $\overline{B}(a,r) = \{x\in X | d(a,x)\leq r\}$, appelé « <u>boule fermée</u> de centre a et de rayon r ».
- $S(a,r) = \{x\in X | d(a,x)=r\}$, appelé « <u>sphère</u> de centre a et de rayon r ».



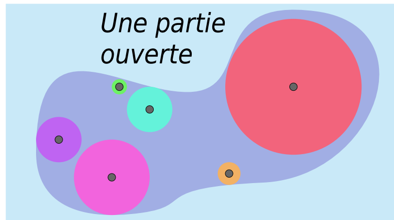
Exemples de boules (fermées) dans différentes normes : dans $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$.

Parties ouvertes, parties fermées.

Une partie U d'un espace métrique X est dite <u>*ouverte*</u> si pour tout $a\in U$, U contient une boule de centre a. Une partie F de X est dite <u>*fermée*</u> si son complémentaire (noté X\F) est une partie ouverte de X. Attention ! Ces deux notions ne sont pas mutuellement exclusives ; une partie peut être les deux (« ouverte-fermée », *clopen* en anglais), ou aucune des deux.

Pour un espace métrique X :

- la partie vide et X sont ouverts et fermés
- toute réunion d'ouverts est ouverte
- toute intersection finie d'ouverts est ouverte
- toute réunion finie de fermés est fermée
- toute intersection de fermés est fermée



Intérieur, adhérence, frontière …

Une partie contenant un point a est appelée « <u>voisinage</u> de a » si elle contient un boule de centre a. Un point x est dit <u>intérieur</u> à A si A est un voisinage du point, <u>adhérent</u> à A si tout voisinage de x rencontre A. L'ensemble des points intérieurs est appelé l'<u>intérieur</u>, noté $\mathring{A}$, (plus grand ouvert contenu dans A), et celui des points adhérents est appelé <u>adhérence</u>, notée $\bar{A}$, (plus petit fermé contenant A). Et la <u>frontière</u> de A est l'ensemble des points adhérents mais pas intérieurs à A ; on la note Fr(A) ou $\partial(A)$, égal à $\bar{A}\backslash\mathring{A}$.

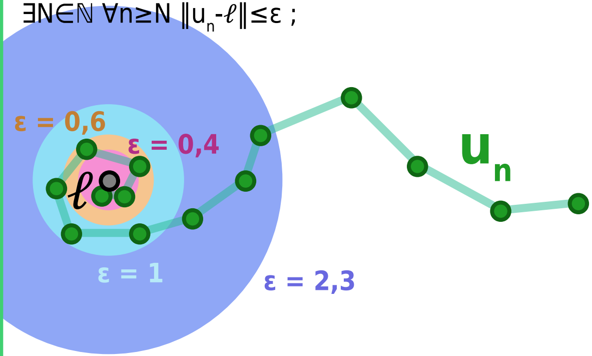
Topologie supplémentaire.

Une partie A de E est dite <u>bornée</u> si il existe M>0 tel que $\|x\|\leq M$ pour tout $x\in A$. Autrement dit, s'il existe une boule centrée à l'origine (de rayon M) contenant A.

Suites dans e.v.n.

Soit E un espace vectoriel normé, et (u_n) une suite à valeurs dans E. Alors :

- (u_n) est dite bornée si il existe M>0 tel que $\|u_n\| \leq M$ pour tout $n\in\mathbb{N}$;
- (u_n) est dite <u>convergente</u> vers $\ell\in E$ si $\forall\varepsilon>0$ $\exists N\in\mathbb{N}$ $\forall n\geq N$ $\|u_n-\ell\|\leq\varepsilon$;



Normes équivalentes.

Des normes N_1, N_2 sont dites équivalentes si il existe des constantes a,b>0 telles que $aN_1(x) \leq N_2(x) \leq bN_1(x)$ pour tout x.

Dimension finie.

Lorsque les espaces vectoriels concernés sont de dimension finie, on a alors que :

- toutes les normes sont équivalentes.
- une partie est compacte si et seulement si elle est fermée et bornée.

Caractérisations séquentielles.

Beaucoup des notions vues peuvent être définies en termes de suites ; ce que l'on appelle une caractérisation séquentielle.

- valeur d'adhérence : ℓ est dite <u>valeur d'adhérence</u> d'une suite (u_n) si $\forall\varepsilon>0$ $\forall N\in\mathbb{N}$ $\exists n\geq N$ $\|u_n-\ell\|\leq\varepsilon$. À comparer avec la définition de limite. Équivalent : si il existe une suite extraite de (u_n) convergente vers ℓ.
- un élément a de E appartient à l'adhérence de D si a est la limite d'une suites à éléments dans D. Et D est fermé si toute suite convergente à éléments dans D, a pour limite un élément de D.
- une partie A de E est dite « <u>dense</u> dans E » si tout ouvert non-vide de A rencontre E. La caractérisation séquentielle équivalente est: pour tout élément a de E, il y a une suite à éléments dans A qui converge vers a. De manière équivalente, si l'adhérence de A est E.
- une partie K de E est dite <u>compacte</u> si pour toute suite (u_n) à éléments dans K, il en existe une sous-suite extraite qui converge vers un élément de K. Toute partie compacte est fermée et bornée.

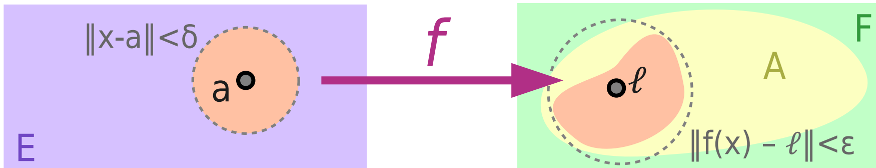
Continuité des fonctions.

Soit E et F deux espaces vectoriels normés, A une partie de E et $f:A\to F$ une fonction. Soit $a\in\bar{A}$, on dit que f a pour <u>limite</u> ℓ en a si

$$\forall\varepsilon>0,\ \exists\delta>0,\ \forall x\in A,\ \|x-a\|<\delta \Rightarrow \|f(x) - \ell\|<\varepsilon.$$

De manière équivalente, si et seulement si pour toute suite (u_n) de A qui converge vers a, $(f(u_n))$ converge vers ℓ.

On dit que f est <u>continue</u> en a si f(x) tend vers f(a) quand $x\in E$ tend vers a. Et donc, si f est continue en tout point de X, f est continue. De plus, f est <u>uniformément continue</u> sur A si $\forall\varepsilon>0$, $\exists\delta>0$ tel que $\forall x,y\in A$, $\|x-y\|<\delta \Rightarrow \|f(x)-f(y)\|<\varepsilon$. Et f est <u>lipschitzienne</u> de rapport $k\in\mathbb{R}$ si $\forall x,y\in A$ $\|f(x)-f(y)\| \leq k\|x-y\|$.



Sur un compact.

Soit $f:K\to F$ une application continue où K est une partie compacte d'un espace vectoriel normé E. Alors f(K) est un compact de F.

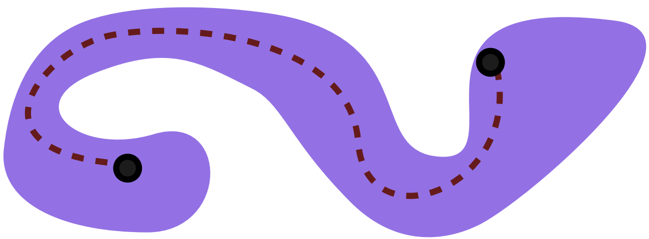
Théorème de Heine. Toute fonction continue sur un compact est uniformément continue.

Extrema d'une fonction continue.

Soit f une fonction à valeurs dans $\mathbb{R}$.
On dit que f admet un <u>minimum global</u> en $a\in U$ si, pour tout $x\in U$, $f(x)\geq f(a)$. On dit que f admet un <u>minimum local</u> en a si il existe $r>0$ tel que pour tout $x\in U\cap B(a,r)$, on a $f(x)\geq f(a)$. On a des définitions similaires pour maximum global et local.
Si D est une partie compacte d'un espace topologique, et f est une fonction continu, il existe des extrema globaux. En effet, l'image f(D) est alors une partie compacte de $\mathbb{R}$, et est donc bornée, et admet minimum et maximum !

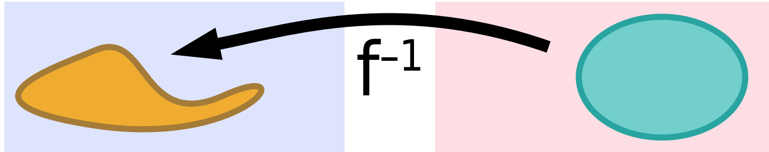
Partie connexe par arc.

Soit A une partie de E, et $x,y\in A$. Une application continue $f:[0,1]\to A$ telle que $f(0) = x$ et $f(1) = y$ est appelée <u>chemin continu</u> tracé dans A de x vers y. Une partie A de E est <u>connexe par arcs</u> si il y a un chemin continu de x vers y (tracé dans A) pour tout $x,y\in A$. Dans $\mathbb{R}$, les seules parties connexes par arcs sont les intervalles. Si $f:A\to F$ est continue et A connexe par arcs, alors f(A) est connexe par arcs !



Continuité : ouverts et fermés.

L'image réciproque d'un ouvert par une application continue est un ouvert. L'image réciproque d'un fermé par une application continue est un fermé. La condition "f est continue" est équivalente à "l'image réciproque par f de tout ouvert est un ouvert" et à "l'image réciproque par f de tout fermé est un fermé".



Suite de Cauchy.

Soit une suite (r_n) de points d'un espace métrique. C'est une <u>suite de Cauchy</u> si $\forall\varepsilon>0\ \exists N\in\mathbb{N}\ \forall p\geq N\ \forall q\geq N\ |r_p-r_q|<\varepsilon$. Autrement dit, si pour tout ε, i y a un rang à partir duquel tout les éléments tiennent dans un intervalle de taille ε. Dans un espace métrique, toute suite convergente est une suite de Cauchy. Mais la réciproque ne l'est pas toujours; un espace métrique où toute suite de Cauchy converge est dit <u>complet</u>. $\mathbb{R}$ et $\mathbb{C}$ sont complets; $\mathbb{Q}$ ne l'est pas. [0,1] est complet,]0,1[ne l'est pas.

Espace de Banach.

Un <u>espace de Banach</u> est un espace vectoriel normé qui est complet (pour la distance issue de la norme). Ces espaces sont très utiles en analyse fonctionnelle. Un <u>espace de Hilbert</u> est un espace de Banach dont la norme vient d'un produit scalaire (ou hermitien) $\langle\cdot,\cdot\rangle$ par la formule $\|x\| = \sqrt{\langle x,x\rangle}$. Les espaces euclidiens $\mathbb{R}^n$ traditionnels sont de Hilbert, de même les espaces hermitiens $\mathbb{C}^n$ (avec pour norme $\langle x,y\rangle = x_1\overline{y}_1 + \ldots + x_n\overline{y}_n$) aussi ! Ils sont très utiles en physique. Dans chaque espace de Hilbert, il existe un analogue du théorème de Pythagore et de la loi des parallélogrammes !

Dérivée selon un vecteur.

Soit E un espace vectoriel normé de dimension finie p, F un e.v. de dimension finie n, et (e_1, ..., e_p) une base de E. Soit *U* un ouvert de E et f:*U*→F. Soit v un vecteur de E, et a∈U. La <u>dérivée de f en a suivant le vecteur v</u> est définie comme

$$D_v f(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t}$$

et n'existe bien sûr que si t↦f(a+tv) est dérivable en 0. Les <u>dérivées partielles</u> sont les dérivées de f en a selon les vecteurs de bases e_i. Elles ont une notation spéciale :

$$\frac{\partial f}{\partial x_i}(a) = \partial_i f(a) = \lim_{t \to 0} \frac{f(a+te_i) - f(a)}{t}$$

La fonction f = (f_1, ..., f_n) admet une dérivée par rapport à v en a si et seulement si tous les f_i en admettent ; et alors on a $D_v f(a) = (D_v f_1(a), ..., D_v f_n(a))$

Différentiabilité.

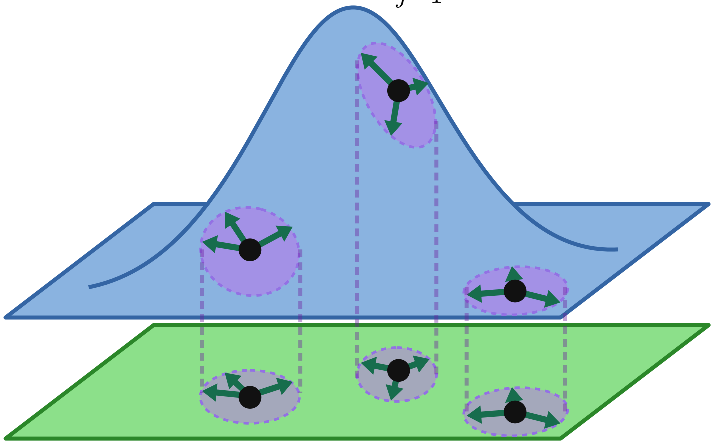
On peut utiliser la notion de différentielle et différentiabilité pour aborder les dérivées partielles et selon un vecteur. On dit que f est <u>différentiable</u> en a si il existe un voisinage V de 0 dans E, pour lequel on peut trouver une application linéaire L:E→F et une application ε:V→F vérifiant $\lim_{h \to 0} \varepsilon(h) = 0$, telles que f(a+h) = f(a)+L(h)+‖h‖ε(h) pour tout h∈V.
Si c'est le cas, L est unique est appelée <u>différentielle</u> de f en a ; on la note df_a. On retrouve la dérivée selon un vecteur : $D_v f(a) = df_a(v)$ (quand f est différentiable en a). De plus, vu que L est linéaire, il suffit de connaître les dérivées partielles pour connaître la dérivée selon n'importe que vecteur :

$$\mathrm{d}f_a(h) = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(a) h_i$$

Composition et opérations algébriques.

On a des propriétés similaires à la dérivée traditionnelle. Si f:U→F et g:U→F sont différentiables en a∈U, alors toute combinaison linéaire λf+μg l'est aussi et $d(\lambda f+\mu g)_a = \lambda df_a + \mu dg_a$. La différentiation est linéaire ! Si f:U→V et g:V→G (où U est un ouvert d'un e.v.n. E et V un ouvert d'un e.v. F), f est différentiable en a et g est différentiable en f(a), alors g∘f est différentiable en a et $d(g \circ f)_a = dg_{f(a)} \circ df_a$. Dans le formalisme des dérivées partielles, les dérivées partielles sont données par :

$$\frac{\partial g \circ f}{\partial x_i}(a) = \sum_{j=1}^{n} \frac{\partial g}{\partial y_j}(f(a)) \frac{\partial f_j}{\partial x_i}(a)$$



Exemple d'une fonction f:ℝ²→ℝ, avec plusieurs points en lesquels f est différentiable et des voisinages sur lesquels on peut donc écrire f(a)+L(h)+‖h‖ε(h) pour chaque vecteur h.

Classe d'une fonction.

On dit que f est de classe C^1 si f est différentiable sur U et si et si l'application a↦df_a est continue. De manière équivalente, si toutes les dérivées partielles de f existent sur U et sont continues sur U.

Classe (encore).

Pour k≥2, les <u>dérivées partielles d'ordre k</u> de f sont définies par récurrence comme les dérivées partielles des dérivées partielles d'ordre k-1 de f. Notamment, une dérivée partielle d'ordre 2 s'écrit comme

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$$

Et plus généralement :

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}} = \frac{\partial}{\partial x_{i_1}}\left(\frac{\partial}{\partial x_{i_2}}\left(\cdots\left(\frac{\partial f}{\partial x_{i_k}}\right)\right)\right)$$

Et donc, f:U→F est de <u>classe</u> C^k (k≥1) sur U si elle admet toutes les dérivées partielles possibles jusqu'à l'ordre k et qu'elles sont toutes continues sur U.

Théorème de Schwarz.

Si f est classe C^2 sur U, alors pour tout a∈U :

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$$

Plan tangent.

Soit f:U⊂ℝ²→ℝ différentiable en U. On peut alors définir une surface X = {(x,y,z)∈ℝ³; (x,y)∈U, z=f(x,y)}; soit donc (x_0, y_0, z_0) un point de X. Alors le <u>plan tangent</u> à X en ce point est défini par l'équation

$$z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

Plus généralement, c'est l'ensemble des vecteurs tangents; un vecteur est <u>tangent</u> à X en un point x si il existe un arc γ (défini sur un intervalle]-ε,ε[avec ε>0) dans X dérivable et tel que γ(0) = x et γ'(0) = v.

Notations.

Par la suite, nous allons utiliser la notation suivante : pour un n-uplet α = (α_1, ..., α_n), on note |α| = α_1+...+α_n. Pour **x** = (x_1, ... x_n), on note $\mathbf{x}^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ et α! = α_1! ... α_n! Et on a aussi

$$\frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}$$

Formule de Taylor-Young.

Soit f:U⊂ℝⁿ→ℝ une fonction de classe C^k sur U, et a un point de U. On définit le k-ième <u>polynôme de Taylor</u> en a comme :

$$\mathrm{P}^k_{f,\mathbf{a}}(\mathbf{x}) = \sum_{m=0}^{k} \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = m}} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha}(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a})^\alpha$$

Alors ce polynôme a les mêmes dérivées partielles d'ordres inférieurs à k en a que f; et on a le développement suivant :

$$f(\mathbf{x}) = P^k_{f,\mathbf{a}}(\mathbf{x}) + o\left(\|\mathbf{x} - \mathbf{a}\|^k\right)$$

Et c'est le seul polynôme de degré k (ou inférieur) qui valide ces deux conditions ! Sont utiles deux cas particuliers : k = 1 et k = 2. On a :

$$P^1_{f,\mathbf{a}}(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a}) h_i$$

$$P^2_{f,\mathbf{a}}(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a}) h_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) h_i h_j$$

Appliquée sur une fonction f:ℝ²→ℝ, le premier polynôme de Taylor nous permet de retrouver la formule du plan tangent :

$$f(\mathbf{x}) = f(\mathbf{a}) + \frac{\partial f}{\partial x}(\mathbf{a})(x - x_0) + \frac{\partial f}{\partial y}(\mathbf{a})(y - y_0) + o(\|\mathbf{x} - \mathbf{a}\|)$$

Et, vu que z = f(x,y) et z_0 = f(x_0,y_0), on retrouve la formule.

Réduction de Gauss.

C'est un algorithme qui permet d'écrire toute forme quadratique comme une combinaison linéaire de carrés (de formes linéaires indépendantes linéairement). La méthode est, pour une forme quadratique

$$Q(x) = \sum_{i=1}^{n} a_{i,i}x_i^2 + \sum_{1\leq i,j\leq n} a_{i,j}x_ix_j$$

- si un des $a_{i,i}$ est non-nul : pour simplifier la suite, on suppose que c'est $a_{1,1}$, que l'on note alors a. On a

$Q(x) = ax_1^2+x_1B(x_2, …, x_n)+C(x_2, …, x_n)$

(B est linéaire et C quadratique) ; on peut alors compléter le carré pour avoir

$$Q(x) = a\left(x_1 + \frac{B(x_2,\dots,x_n)}{2a}\right)^2 + C(x_2,\dots,x_n) - \frac{B(x_2,\dots,x_n)^2}{4a}$$

On a un multiple d'un carré, plus une forme quadratique avec une variable de moins; on applique la méthode dessus (récursivement).

- tout les $a_{i,i}$ sont nuls; si tout les $a_{i,j}$ sont nuls, il n'y a rien à faire. Supposons donc qu'un des $a_{i,j}$, disons $a_{1,2}$ pour simplifier, soit non-nul; notons le a. On a alors

$Q(x) = ax_1x_2+x_1B(x_3, …, x_n)+x_2C(x_3, …, x_n)+D(x_3, …, x_n)$

(B, C linéaires, D quadratique); on factorise pour avoir $a(x_1+C/a)(x_2+B/a)+(D–BC/a)$; en utilisant l'identité uv = 1/4((u+v)²-(u-v)²), on peut écrire

$$Q(x) = \frac{a}{4}\left(x_1 + x_2 + \frac{B+C}{a}\right)^2 - \frac{a}{4}\left(x_1 - x_2 + \frac{C-B}{a}\right)^2 + \left(D - \frac{BC}{a}\right)$$

On a deux multiples de carré, plus une forme quadratique (D-BC/a) avec deux variables de moins; on applique la méthode dessus. L'algorithme fonctionne par récursion : on forme un ou deux carrés, on réapplique l'algorithme sur le reste, jusqu'à ce qu'il ne reste plus de termes.

Forme linéaire, quadratique.

Un polynôme homogène est un polynôme dont tout les monômes sont de même degré. Quand ce degré est 1, on dit que le polynôme est une forme linéaire ; pour un degré de 2, on parle de forme quadratique. Pour deux variables, une forme quadratique est de la forme Q(x,y) = ax²+bxy+cy², pour trois variables Q(x,y,z) = ax²+by²+cz²+dxy+exz+fyz, et ainsi de suite. Une forme quadratique Q est dite définie positive si Q(x) > 0 pour tout x, positive si Q(x) ≥ 0 pour tout x, et ainsi de suite. Si la réduction de Gauss donne autant de carrés précédés de coefficients non-nuls qu'il n'y a de variables, la forme quadratique est dite non-dégénérée.

Position du plan tangent.

On peut toujours écrire

$$P^2_{f,\mathbf{a}}(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + D_\mathbf{a}f(\mathbf{h}) + \frac{1}{2}Q_\mathbf{a}f(\mathbf{h})$$

où $D_\mathbf{a}f$ est une forme linéaire et $Q_\mathbf{a}f$ une forme quadratique. La forme linéaire correspond au plan tangent en a ; on étudie la forme quadratique. Le plan est (localement) au-dessus pour $Q_\mathbf{a}f$ définie négative, (localement) en dessous pour $Q_\mathbf{a}f$ définie positive; et ainsi de suite.

Points critiques.

Le terme linéaire $D_\mathbf{a}f(\mathbf{x}-\mathbf{a})$ est nul si a est un point critique de f, c'est le signe de la forme quadratique qui nous indique le type. Si $Q_\mathbf{a}f$ est définie négative, alors f admet un maximum en a ; si $Q_\mathbf{a}f$ est définie positive, alors f admet un minimum en a ; sinon, si $Q_\mathbf{a}f$ est non-dégénérée mais ni définie positive ni définie négative, alors f admet un point de selle en a.

Fonctions polynômiales.

Une fonction polynômiale sur $\mathbb{R}^n$ est un polynôme en les coordonnées de l'espace. Ces fonctions sont continues et différentiables. En fait, elles sont différentiables une infinité de fois ; et donc de classe C^∞.