Calcul différentiel

« Dévore un homme et tu feras un bon repas. Épargne-le et il te nourrira toute ta vie. » – proverbe canin.

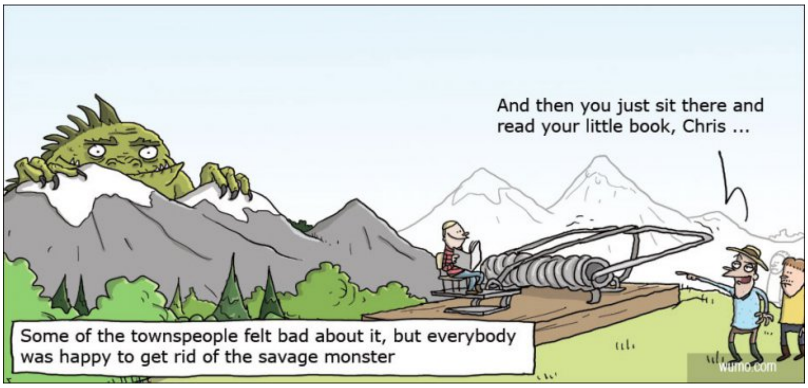
Table des matières

1 Fonctions de $\mathbb{R}$ dans $\mathbb{R}^p$ **2**
1.1 Dérivée en un point . . . 2
1.2 Propriétés opératoires . . . 2

2 Fonctions de $\mathbb{R}^p$ dans $\mathbb{R}$ **3**
2.1 Classe $\mathcal{C}^1$, différentielle . . . 4
2.2 Gradient . . . 5
2.3 Et de $\mathbb{R}^p$ dans $\mathbb{R}^q$? . . . 6
2.4 Composition (« règle de la chaîne ») . . . 6
2.5 Problèmes d'extrema . . . 7

3 Au delà du premier ordre **8**
3.1 Classe $\mathcal{C}^2$, Schwarz . . . 8
3.2 Hessienne ; retour aux extrema . . . 9
3.3 Des équations aux dérivées partielles . . . 10

4 Un peu de géométrie différentielle **10**
4.1 Qu'est-ce qu'une courbe du plan ? . . . 10
4.2 Au voisinage d'un point régulier . . . 11



1 Fonctions de $\mathbb{R}$ dans $\mathbb{R}^p$

On sait déjà que pour intégrer une fonction continue de $[a, b]$ dans $\mathbb{R}^p$, « on intègre composante par composante ». De même pour une fonction à valeurs dans un espace vectoriel de dimension finie (on intègre les coordonnées sur une base ; penser à $\mathbb{C}$ et $\mathrm{e}^{it} = \cos t + i \sin t$). Essentiellement, ça va être la même chose pour la dérivabilité/dérivation.

On gagne beaucoup à voir une application de $\mathbb{R}$ dans $\mathbb{R}^p$ comme une courbe paramétrée (« un point qui se déplace au cours du temps »). Les aspects cinématiques (vitesse, accélération) expliquent bien des choses, permettent de se construire une intuition.

1.1 Dérivée en un point

Définition 1 — *Dérivée*

Soit $f : I \to E$, avec E un espace vectoriel de dimension finie. On dit que f est dérivable en $t_0 \in I$ lorsqu'il existe $\overrightarrow{V} \in E$ tel que

$$f(t_0 + h) = f(t_0) + h\overrightarrow{V} + o(h)$$

c'est-à-dire :

$$\left\| f(t_0 + h) - \left(f(t_0) + h\overrightarrow{V} \right) \right\| = o(h)$$

Lorsque c'est le cas, $\overrightarrow{V}$ est alors unique, et on note alors $f'(t_0) = \overrightarrow{V}$ (mais voir ceci comme un *vecteur vitesse* est certainement une bonne idée...).

Cette définition n'est pas pratique pour vérifier qu'une fonction est effectivement dérivable. Heureusement on a la propriété suivante :

Proposition 1 — En pratique, c'est simple !

Si $f : I \to \mathbb{R}^n$ s'écrit $f = (f_1, ..., f_n)$, alors f est dérivable en t si et seulement si chacune des f_i l'est. Et lorsque c'est le cas, on a $f'(t) = (f_1'(t), ..., f_n'(t))$.
Plus généralement, si f est à valeurs dans E ayant pour base $(\overrightarrow{e_1}, ..., \overrightarrow{e_n})$ et que f s'écrit $f_1\overrightarrow{e_1} + \cdots + f_n\overrightarrow{e_n}$, alors f est dérivable si et seulement si les f_k le sont, et on a alors $f' = \sum\limits_{k=1}^{n} f_k'\overrightarrow{e_k}$.

Exemples :
- $E = \mathbb{R}^3$: l'application $t \mapsto (t^2, 1/t)$ est dérivable sur $\mathbb{R}^*$, de dérivée $t \mapsto (2t, -1/t^2)$.
- $E = \mathbb{C}$: l'application $t \mapsto t^2 + i \sin t$ est dérivable de dérivée $t \mapsto 2t + i \cos t$.
- $E = \mathcal{M}_2(\mathbb{C})$: l'application $t \mapsto \begin{pmatrix} \mathrm{e}^t & \sin t \\ t & -1 \end{pmatrix}$ est dérivable de dérivée $t \mapsto \begin{pmatrix} \mathrm{e}^t & \cos t \\ 1 & 0 \end{pmatrix}$.

On dispose évidemment des notions de fonction de classe $\mathcal{C}^k$:

Définition 2 — *Classe $\mathcal{C}^k$*

Soit $f : I \to \mathbb{R}^p$.
- On dit que f est de classe $\mathcal{C}^1$ lorsque f est dérivable en tout point de I et que l'application dérivée $f' : I \to \mathbb{R}^p$ est continue.
- Pour k entier strictement plus grand que 1, on dit que f est de classe $\mathcal{C}^k$ lorsque f est de classe $\mathcal{C}^1$, et que f' est de classe $\mathcal{C}^{k-1}$.
- f est dite de classe $\mathcal{C}^\infty$ lorsqu'elle est de classe $\mathcal{C}^k$ pour tout $k > 0$.

En pratique, $f = (f_1, ..., f_p)$ est de classe $\mathcal{C}^k$ si et seulement si les f_i le sont toutes.

De même que $f'(t_0)$ sera vu comme le vecteur vitesse d'un point qui se déplace, $f''(t_0)$ représente *l'accélération* du point au temps t_0.

1.2 Propriétés opératoires

Essentiellement, tout ce qu'un physicien écrira naturellement (bref : qui est homogène et étend raisonnablement ce qu'on sait du calcul des dérivées pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$)... est vrai ! Pour les prouver (formellement, ou presque, mais ça vous donnera déjà le résultat et l'argument moral), écrivez des développements limités ! Vous pouvez commencer par reconstituer la preuve de la formule $(uv)' = u'v + uv'$, pour deux fonctions u et v dérivables de $\mathbb{R}$ dans $\mathbb{R}$.

Proposition 2 — Diverses propriétés – les fameux « théorèmes usuels »

> Soient $f, g : \mathbb{R} \to \mathbb{R}^p$, $\lambda : \mathbb{R} \to \mathbb{R}$ dérivables. Soient également $\alpha \in \mathbb{R}$, $u \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^q)$ et B une forme bilinéaire sur $\mathbb{R}^p$.
> - L'application $\alpha f + g$ est dérivable, avec $(\alpha f + g)' = \lambda f' + g'$.
> - L'application λf est dérivable, de dérivée $\lambda' f + \lambda f'$.
> - L'application $t \mapsto f(\lambda(t))$ est dérivable, de dérivée $t \mapsto \lambda'(t) f'(\lambda(t))$.
> - L'application $t \mapsto u(f(t))$ est dérivable, de dérivée $t \mapsto u(f'(t))$.
> - L'application $t \mapsto B(f(t), g(t))$ est dérivable, de dérivée $t \mapsto B(f'(t), g(t)) + B(f(t), g'(t))$.
>   *En particulier lorsque B est un produit scalaire, le déterminant dans une base donnée (en dimension 2) ou le produit vectoriel (en dimension 3).*
> - Si $\| \ \|$ est une norme euclidienne, alors l'application $t \mapsto \|f(t)\|^2$ est dérivable, de dérivée $t \mapsto 2\langle f(t) | f'(t) \rangle$.
>
> Évidemment, on pourra remplacer « dérivable » par « de classe $\mathcal{C}^k$ »...

Terminons cet annuaire en signalant que la bien connue (...) formule de Taylor avec reste intégral est encore vraie pour une fonction à valeurs vectorielles (en dimension finie, tout de même...) : il suffit d'appliquer la formule scalaire à chaque coordonnée.

Proposition 3 — Formule de Taylor avec reste intégral

> Soit f une fonction de classe $\mathcal{C}^{n+1}$ de I dans E, avec I un intervalle réel et E un espace de dimension finie. Soient également $a, b \in I$. On a alors :
>
> $$f(b) = f(a) + \sum_{k=1}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t) dt$$

La clef de la preuve est la formule d'intégration par parties, qui reste également vraie pour des fonctions à valeurs vectorielles.

Exercice 1. *Pourquoi, au fait ?*

2 Fonctions de $\mathbb{R}^p$ dans $\mathbb{R}$

Voici quelques exemples d'applications de $\mathbb{R}^p$ dans $\mathbb{R}$... ou plus généralement de E dans $\mathbb{R}$, avec E un espace vectoriel de dimension finie :

$$f_1 : (x, y) \in \mathbb{R}^2 \mapsto x^2 y - \cos(y) \mathrm{e}^x \qquad f_2 : M \in \mathcal{M}_n(\mathbb{R}) \mapsto \det(M) \qquad f_3 : M \in \mathbb{R}^2 \mapsto \Omega M^2$$

On notera que dans le dernier exemple, on a plus envie de voir les habitants de $\mathbb{R}^2$ comme des points plutôt que des vecteurs : c'est bien ce qu'on fera. Et bien entendu si $M, N \in \mathbb{R}^2$, on a envie d'écrire $N = M + \overrightarrow{MN}$, « avec $\overrightarrow{MN} = N - M$ » qu'on voit comme un vecteur...
Le but est d'obtenir un développement limité (à l'ordre 1 dans un premier temps) d'une fonction de $\mathbb{R}^p$ dans $\mathbb{R}$ au voisinage d'un point. Qu'est-ce que ça peut vouloir dire ? Comment exprimer un tel développement limité ?

Si on veut développer f au voisinage de a (au sens : pour tous les points proches de a, donc de la forme $a + h$ avec h un petit vecteur), on peut espérer quelque chose qui ressemble à ça :

$$f(a + h) = f(a) + \varphi(h) + o(\|h\|)$$

avec φ une forme linéaire (bien penser au fait que a est dans E (et peut être vu plutôt comme un point) et h également (mais qu'on verra plutôt comme un petit vecteur de déplacement). Mais on peut aussi vouloir développer f au voisinage de a dans une direction donnée, c'est-à-dire écrire quelque chose comme :

$$f(a + t\overrightarrow{d}) = f(a) + Kt + o(t)$$

(la variable étant maintenant un réel t, et le vecteur $\overrightarrow{d}$ porte la direction dans laquelle on regarde).

En pratique, les fonctions considérées ne seront pas définies forcément sur tout $\mathbb{R}^p$, mais seulement sur des ouverts de $\mathbb{R}^p$; passons dans un premier temps pour simplifier les définitions et énoncés...

2.1 Classe $\mathcal{C}^1$, différentielle

DÉFINITION 3 — *Fonctions de classe* $\mathcal{C}^1$

On suppose que f est une application de $\mathbb{R}^p$ dans $\mathbb{R}$.

— f possède une dérivée partielle selon la i-ème variable en $a = (a_1, ..., a_n)$ lorsque l'application

$$t \mapsto f(a_1, ..., a_{i-1}, t, a_{i+1}, ..., a_n)$$

est dérivable en a_i. Lorsque c'est le cas, cette dérivée est notée $\partial_i f(a)$.

— f est dite de classe $\mathcal{C}^1$ lorsqu'elle est dérivable par rapport à chacune de ses variables, et que les applications $\partial_i f$ (qui vont donc de E dans $\mathbb{R}$) sont toutes continues.

Concrètement, dire que f est dérivable par rapport à sa i-ème variable, c'est dire qu'en bloquant les autres, on obtient une « fonction » dérivable. On (en particulier les physiciens) utilise également d'autres notations quand $E = \mathbb{R}^p$:

— $\partial_i f(a)$ est souvent noté $\dfrac{\partial f}{\partial x_i}(a)$; Pour $p = 3$, on aura tendance à noter les variables x, y et z, ce qui fait qu'on parlera plutôt de $\dfrac{\partial f}{\partial x}(a)$, $\dfrac{\partial f}{\partial y}(a)$ et $\dfrac{\partial f}{\partial z}(a)$.

— Si tout le monde voit bien ce que signifie $\dfrac{\partial f}{\partial x}(3, 2, 1)$, la notation $\dfrac{\partial f}{\partial x}(x, y, z)$ devient plus délicate (il y a deux statuts différents pour les x...), et $\dfrac{\partial f}{\partial x}(z, x, y)$ devient carrément casse-gueule ! S'agit-il de la dérivée de f par rapport à sa première variable, qu'on évaluerait en (z, x, y), ou bien de la dérivée par rapport à sa première variable de l'application $(x, y, z) \mapsto f(z, x, y)$?

$$(\partial_1 f)(z, x, y) \neq \frac{\partial}{\partial x}\left(f(z, x, y)\right)$$

Accessoirement, le membre de droite de l'inégalité précédente n'est formellement pas bien défini (on doit dériver une fonction et non une expression).
Bref, ceux qui confondaient déjà $\left(f(x^2)\right)'$ et $f'(x^2)$ vont avoir du mal (...) ; et ceux qui avaient fini par distinguer ces deux choses... devront à nouveau être vigilants pour les fonctions de plusieurs variables.

L'intérêt principal de la classe $\mathcal{C}^1$ est qu'elle permet de faire des développements limités au voisinage des points ; et ce dans les deux sens évoqués plus haut : « globalement », ou dans une direction donnée.

THÉORÈME 1 — *Développements limités*

Soit f de classe $\mathcal{C}^1$ sur un ouvert Ω, et $a = (a_1, ..., a_n) \in \Omega$.

— Si $h = (h_1, ..., h_n)$, alors

$$f(a + h) = f(a) + \sum_{k=1}^{n} h_i \partial_k f(a) + o(\|h\|)$$

— Si $\overrightarrow{v} = (v_1, ..., v_n)$, alors :

$$f\left(a + t\overrightarrow{v}\right) = f(a) + \left(\sum_{k=1}^{n} v_i \partial_k f(a)\right) t + o(t)$$

REMARQUES :

- Ainsi, le caractère $\mathcal{C}^1$ implique le caractère $\mathcal{C}^0$... et c'est moins niais comme remarque qu'il n'y parait !
- L'énoncé a volontairement été donné pour une fonction définie sur un ouvert. Le point crucial est qu'on veut pouvoir avoir de vrais voisinages de a ; il est donc important que a soit intérieur à Ω... ce qui sera assuré dès que Ω est ouvert.
- Pour la dernière formule, on écrira aussi $f\left(a+t\overrightarrow{v}\right) = f(a) + tD_v f(a) + o(t)$ avec $D_v f(a) = \sum\limits_{k=1}^{n} v_i \partial_k f(a)$: c'est la « dérivée de f en a selon le vecteur $\overrightarrow{v}$ ».

DÉFINITION 4 — *Différentielle*

Pour f de classe $\mathcal{C}^1$ sur Ω et $a \in \Omega$, on note $df(a)$ la forme linéaire $h \mapsto \sum\limits_{k=1}^{p} \partial_k f(a) h_k$: $df(a)$ est la *différentielle de f en a*. On a alors :

$$f(a+h) = f(a) + df(a).h + o(\|h\|)$$

On note $df(a).h$ par abus de langage : c'est normalement $(df(a))(h)$...
Mais aussi :

$$f\left(a + t\overrightarrow{d}\right) = f(a) + \left(df(a).\overrightarrow{d}\right)t + o(t)$$

PROPOSITION 4 — Combinaisons linéaires et produits

Soit Ω un ouvert de $\mathbb{R}^n$. Soient également $f, g \in \mathcal{C}^1(\Omega, \mathbb{R})$, et $\lambda \in \mathbb{R}$.

- L'application $\Phi = \lambda f + g$ est de classe $\mathcal{C}^1$ sur Ω, avec pour tout $i \in [\![1,p]\!]$, $\partial_i \Phi = \lambda \partial_i f + \partial_i g$:
$$\forall a \in \Omega, \qquad d(\lambda f + g)(a) = \lambda df(a) + dg(a).$$
- L'application $\Psi = fg$ est de classe $\mathcal{C}^1$ sur Ω, avec pour tout $i \in [\![1,p]\!]$, $\partial_i \Psi = (\partial_i f)g + f\partial_i g$:
$$\forall a \in \Omega, \qquad d(fg)(a) = df(a)g(a) + f(a)dg(a).$$

2.2 Gradient

La partie linéaire du développement limité $f(a+h) = f(a) + df(a).h + o(\|h\|)$ est la forme linéaire $(h_1, ...h_n) \mapsto \sum\limits_{k=1}^{p} \partial_k f(a) h_k$. Comme toute forme linéaire en dimension finie, elle s'exprime sous la forme d'un produit scalaire :

$$df(a).h = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_p f(a) \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix}$$

DÉFINITION 5 — *Gradient*

Pour $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ et $a \in \Omega$, le *gradient de f en a* est le vecteur $\nabla f(a) = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_p f(a) \end{pmatrix}$.

On a alors les développements limités (naturels pour le physicien) :

$$f(a+h) = f(a) + \langle \nabla f(a) | h \rangle + o(\|h\|)$$

et

$$f\left(a + t\overrightarrow{v}\right) = f(a) + \langle \nabla f(a) | \overrightarrow{v} \rangle t + o(t)$$

Ainsi, quand $f \in \mathcal{C}^1(\Omega, \mathbb{R})$, ∇f est une application continue de Ω dans $\mathbb{R}^p$.

PROPRIÉTÉS : Soient f et g de classe $\mathcal{C}^1$ de Ω (ouvert de $\mathbb{R}^p$) dans $\mathbb{R}$; soit également $\lambda \in \mathbb{R}$:

- (on a déjà vu que) $\lambda f + g$ est de classe $\mathcal{C}^1$, et $\nabla(\lambda f + g) = \lambda \nabla f + \nabla g$;
- (on a déjà vu que) fg est de classe $\mathcal{C}^1$, et $\nabla(fg) = g\nabla f + f\nabla g$ (oui, c'est homogène !) ;

— si $f(a) \neq 0$ alors f ne s'annule pas au voisinage de a, et $1/f$ est de classe $\mathcal{C}^1$ sur ce voisinage, avec $\nabla(1/f) = -\dfrac{1}{f^2}\nabla f$.

EXEMPLES :

— En électrostatique, $\overrightarrow{E} = -\nabla V$, et la différence de potentiel entre M et $M + d\overrightarrow{\ell}$ est $dV = -\overrightarrow{E}.d\overrightarrow{\ell}$. *En particulier, si on se déplace avec un mouvement orthogonal au champ, le potentiel ne varie pas : on est sur une surface équipotentielle.*

— La loi de Fourier exprime le flux de chaleur à l'aide de la température : $\overrightarrow{j_Q} = -\lambda\nabla T$:

$$T(M + d\overrightarrow{\ell}) = T(M) - \frac{1}{\lambda}\langle\overrightarrow{j_Q}|d\vec{\ell}\rangle$$

En particulier, si on se déplace de façon orthogonale au flux de chaleur, la température ne bouge pas : on est sur une surface isotherme.

2.3 Et de $\mathbb{R}^p$ dans $\mathbb{R}^q$?

Si f est une telle application, alors on peut l'écrire $f = (f_1, ..., f_q)$, avec chaque f_i qui est une application de $\mathbb{R}^p$ dans $\mathbb{R}$. L'application f est alors dite de classe $\mathcal{C}^1$ (et plus tard $\mathcal{C}^n$) lorsque chaque f_i l'est.
C'est tout !

2.4 Composition (« règle de la chaîne »)

On peut vouloir composer des fonctions de $\mathbb{R}^p$ dans $\mathbb{R}^q$ et de $\mathbb{R}^q$ dans $\mathbb{R}^s$. Començons par le cas $p = s = 1$.

PROPOSITION 5 — Règle de la chaîne

Si $\gamma \in \mathcal{C}^1(I, \mathbb{R}^n)$ (I intervalle de $\mathbb{R}$) avec $\gamma(t) = (x_1(t), ..., x_n(t))$ et $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$, alors $g = f \circ \gamma : t \mapsto f(x_1(t), ..., x_n(t))$ est de classe $\mathcal{C}^1$, avec :

$$\forall t \in I, \qquad g'(t) = x_1'(t)\partial_1 f(\gamma(t)) + \cdots + x_n'(t)\partial_n f(\gamma(t)) = \sum_{k=1}^{n} x_k'(t)\partial_k f(\gamma(t))$$

Et si $\Phi = (\Phi_1, ..., \Phi_n) \in \mathcal{C}^1(\mathbb{R}^p, \mathbb{R}^n)$ (chaque Φ_i étant donc dans $\mathcal{C}^1(\mathbb{R}^p, \mathbb{R})$...), alors

$$\Psi = f \circ \Phi : x \in \mathbb{R}^p \mapsto f(\Phi_1(x), ..., \Phi_n(x))$$

est également de classe $\mathcal{C}^1$, avec :

$$\forall a \in \mathbb{R}^p \ \forall i \in [\![1, p]\!], \qquad \partial_i\Psi(a) = \sum_{k=1}^{n} \partial_i\Phi_k(a)\partial_k f(\Phi(a))$$

Ou encore, avec les notations à la physicienne :

$$\forall a \in \mathbb{R}^p \ \forall i \in [\![1, p]\!], \qquad \frac{\partial\Psi}{\partial x_i}(a) = \sum_{k=1}^{n} \frac{\partial\Phi_k}{\partial x_i}(a)\frac{\partial f}{\partial u_k}(\Phi(a)) = \langle\nabla f(\Phi(a))| \begin{pmatrix} \frac{\partial\Phi_1}{\partial x_i}(a) \\ \vdots \\ \frac{\partial\Phi_n}{\partial x_i}(a) \end{pmatrix} \rangle$$

EXEMPLES :

— Concrètement, pour $n = 2$: si $g(u, v) = f(X(u, v), Y(u, v))$, alors :

$$\partial_1 g(u, v) = \partial_1 X(u, v)\partial_1 f(X(u, v), Y(u, v)) + \partial_1 Y(u, v)\partial_2 f(X(u, v), Y(u, v)),$$

ce que les physiciens préféreront écrire :

$$\frac{\partial g}{\partial u}(u, v) = \frac{\partial X}{\partial u}(u, v)\frac{\partial f}{\partial x}(X(u, v), Y(u, v)) + \frac{\partial Y}{\partial u}(u, v)\frac{\partial f}{\partial y}(X(u, v), Y(u, v))$$

et même :
$$\frac{\partial g}{\partial u} = \frac{\partial X}{\partial u}\frac{\partial f}{\partial x} + \frac{\partial Y}{\partial u}\frac{\partial f}{\partial y}.$$

Si on décide d'utiliser la notation à la physicienne, on change le nom des variables pour f et g : (u,v) pour g et (x,y) pour f, sinon c'est une boucherie !

— Si on fait un changement de variable linéaire : $g(u,v) = f(au+bv, cu+dv)$, alors (à la physicienne) :
$$\frac{\partial g}{\partial u} = a\frac{\partial f}{\partial x} + c\frac{\partial f}{\partial y} \qquad \text{et} \qquad \frac{\partial g}{\partial v} = b\frac{\partial f}{\partial x} + d\frac{\partial f}{\partial y}$$

— Dans un changement de variable polaire, on définit $g(r,\theta) = f(r\cos\theta, r\sin\theta)$ (c'est le fameux « on pose $x = r\cos\theta$ et $y = r\sin\theta$ ») ; on a alors :
$$\frac{\partial g}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \sin\theta\frac{\partial f}{\partial y} \qquad \text{et} \qquad \frac{\partial g}{\partial \theta} = -r\sin\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}$$

Le physicien écrira ce qui précède, ou bien :
$$\frac{\partial g}{\partial r}(r,\theta) = \cos\theta\frac{\partial f}{\partial x}(x,y) + \sin\theta\frac{\partial f}{\partial y}(x,y) \qquad \text{et} \qquad \frac{\partial g}{\partial \theta}(r,\theta) = -r\sin\theta\frac{\partial f}{\partial x}(x,y) + r\cos\theta\frac{\partial f}{\partial y}(x,y)$$

alors que le matheux (surtout s'il doit dériver une deuxième fois) écrira plutôt :
$$\frac{\partial g}{\partial r}(r,\theta) = \cos\theta\frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta\frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) \qquad \text{et} \qquad \frac{\partial g}{\partial \theta}(r,\theta) = \cdots$$

Exercice 2. *On suppose que f est une application de classe $\mathcal{C}^1$ de C dans $\mathbb{R}$, avec C un convexe de $\mathbb{R}^p$. On suppose de plus que ∇f est uniformément nul sur C.*
Montrer que f est constante.
On pourra s'intéresser, avec a et b fixés dans C, à l'application $t \mapsto f(a+t(b-a))$.

Exercice 3. *Soit $\varphi \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$. Justifier le caractère $\mathcal{C}^1$ et calculer les dérivées partielles des applications suivantes :*
$$f : (x,y) \mapsto \varphi(x+y); \qquad g : (x,y) \mapsto \varphi(x^2+y^2); \qquad h : (x,y) \mapsto \varphi(xy)$$

2.5 Problèmes d'extrema

Vous vous souvenez ce qu'est un extrémum local ou global, pour une fonction de $\mathbb{R}$ dans $\mathbb{R}$ (je ne parle pas de théorème, mais de définition) ?

— Oui bien sûr ? *Alors pour une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$, c'est pareil !*

— Non ? *Alors laissez tomber ce paragraphe : il est trop tard pour s'y mettre.*

On rappelle que si une fonction est continue sur un fermé borné (en dimension finie) et à valeurs réelles, alors elle possède un maximum et un minimum (globaux). Bien entendu, qui dit global dit local.
Une façon de rechercher les extrema globaux peut donc être, dans une phase d'analyse, de chercher les locaux ; puis dans la synthèse... voir s'ils sont mieux que cela.

Exercice 4. *Énoncer un/le théorème donnant une condition nécessaire/suffisante d'existence d'un extrémum local, pour une fonction dérivable de $\mathbb{R}$ dans $\mathbb{R}$.*

Exercice 5. *Chercher et trouver un contre-exemple à votre énoncé précédent. Réparer l'énoncé en question.*

Exercice 6. *Chercher et trouver un contre-exemple à l'énoncé réparé. Réparer à nouveau l'énoncé réparé.*

Exercice 7. *On suppose que $f : \mathbb{R}^2 \to \mathbb{R}$ possède un minimum local en $a = (a_1, a_2)$. Montrer :* $\partial_1 f(a_1, a_2) = 0$

Théorème 2 — *Condition* ***nécessaire*** *d'extremum*

Si f est de classe $\mathcal{C}^1$ de Ω (ouvert de $\mathbb{R}^p$ dans $\mathbb{R}$ et possède un extremum local en a, alors $df(a) = 0$ (ou encore : $\nabla f(a) = 0$).

Exercice 8. *La condition nécessaire citée plus haut est-elle suffisante ?*

DÉFINITION 6 — *Point critique*

> Pour une fonction f de classe $\mathcal{C}^1$ sur un ouvert Ω, on appelle *point critique* tout $a \in \Omega$ tel que $df(a) = 0$.
> *Certains désignent plutôt par point critique la valeur $f(a)$, attention...*

Les figures suivantes représentent les graphes de deux exemples typiques de fonctions possédant un point critique en 0.

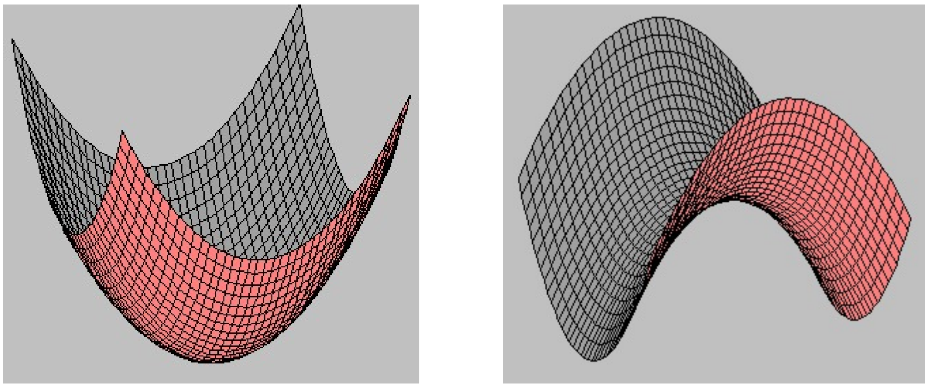


FIGURE 1 – Paraboloïde ($z = x^2 + y^2$) et paraboloïde hyperbolique ($z = x^2 - y^2$)

Tout est simple lorsqu'on peut séparer les dépendances en x et y dans $f(x, y)$!

Exercice 9. *Déterminer les points critiques et extrema locaux/globaux des applications suivantes :*

1. $(x, y) \in \mathbb{R}^2 \mapsto x^4 + x^5 + y^4$; $(x, y) \in \mathbb{R}^2 \mapsto x^2 + x^{19} - y^2$;
2. $(x, y) \in \mathbb{R}^2 \mapsto x^{17} + y^{17}$; $(x, y) \in \mathbb{R}^2 \mapsto x^{19} - y^{16} + y^{512}$;
3. $(x, y) \in \mathbb{R}^2 \mapsto \sin(x^2 + y^2)$.

(dans le dernier cas, on ne peut pas séparer les dépendances en x et y, mais...)

3 Au delà du premier ordre

On a vu plus haut que pour avoir un extremum en un point une condition *nécessaire* est que le gradient y soit nul. On va développer des outils nous fournissant une *condition suffisante*.

3.1 Classe $\mathcal{C}^2$, Schwarz

DÉFINITION 7 — *Classe $\mathcal{C}^2$*

> Une application $f : \mathbb{R}^p \to \mathbb{R}$ est dite de classe $\mathcal{C}^2$ lorsqu'elle est de classe $\mathcal{C}^1$, et que ses p dérivées partielles $\partial_i f$ (qui sont chacune des applications de $\mathbb{R}^p$ dans $\mathbb{R}$) sont elles même des applications de classe $\mathcal{C}^1$.

Exercice 10. *Montrer que $f : (x, y) \mapsto x \ln(1 + x^2 + y^2)$ est de classe $\mathcal{C}^2$. Calculer $\partial_2\partial_1 f(x, y)$ ainsi que $\partial_1\partial_2 f(x, y)$*

Vous savez quoi ? Ce n'est pas un miracle/accident !

THÉORÈME 3 — *Schwarz*

> Si f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^p$; alors pour tout $(i, j) \in [\![1, p]\!]^2$ et tout $a \in \mathbb{R}^p$: $\partial_i\partial_j f(a) = \partial_j\partial_i f(a)$.

Plutôt que de parler de $\partial_1\partial_2 f$ (resp. $\dfrac{\partial}{\partial x}\dfrac{\partial f}{\partial y}$), on notera plutôt $\partial^2_{1,2} f$ (resp. $\dfrac{\partial^2 f}{\partial x \partial y}$).
Terminons par une notion que vous avez déjà rencontrée quelques fois ailleurs qu'en maths !

DÉFINITION 8 — *Laplacien*

Soit f une application de classe $\mathcal{C}^2$ de $\mathbb{R}^2$ dans $\mathbb{R}$. Son laplacien, noté Δf est défini par :

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$

Et bien entendu, en dimension 3 : $\Delta f = \dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2} + \dfrac{\partial^2 f}{\partial z^2}.$

3.2 Hessienne ; retour aux extrema

Petits rappels du cours de première année (fonctions d'une variable réelle) et de deuxième année (matrices symétriques réelles, endomorphismes autoadjoints) :

- Si $f \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ alors $f(x_0 + t) = f(x_0) + f'(x_0)t + \frac{f''(x_0)}{2}t^2 + o(t^2)$ quand t tend vers 0.
- Pour que $f \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ possède un extremum en x_0 il est nécessaire d'avoir $f'(x_0) = 0$ (le prouver) mais ce n'est pas suffisant (dire pourquoi).
- Pour que $f \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$ possède un minimum en x_0 il est suffisant d'avoir $f'(x_0) = 0$ et $f''(x_0) > 0$ (le prouver) mais ce n'est pas nécessaire (dire pourquoi).
- Si $A \in \mathcal{S}_n(\mathbb{R})$ possède pour valeurs propres comptées avec leurs multiplicités et ordonnées $\lambda_1 \leqslant \cdots \leqslant \lambda_n$ alors pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$ on a $X^T AX \geqslant \lambda_1 X^T X$, ou encore : $\langle u(x)|x\rangle \geqslant \lambda_1 \|x\|^2$ avec les notations qu'on imagine.

La hessienne va jouer en dimension $n \geqslant 2$ le rôle de la dérivée seconde en dimensionn 1.

DÉFINITION 9 — *Hessienne*

Si $f \in \mathcal{C}^2(\mathbb{R}^p, \mathbb{R})$ et $a \in \mathbb{R}^p$, alors la hessienne de f en a est la matrice (symétrique)

$$H_f(a) = \left(\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)\right)_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{R}).$$

Exercice 11. *Calculer la hessienne en* 0 *des applications suivantes :*

- $(x, y) \mapsto xy + x^5$;
- $(x, y, z) \mapsto x^3y^3 + xyz + 3x + y - z$;
- $(x, y) \mapsto \cos(x + 2y)$;
- $(x, y) \mapsto \cos(2x + y)$.

THÉORÈME 4 — *Développements limités à l'ordre* 2

Si $f \in \mathcal{C}^2(\mathbb{R}^p, \mathbb{R})$ et $a \in \mathbb{R}^p$, alors :

$$f(a + h) = f(a) + \langle h|\nabla f(a)\rangle + h^T H_f(a)h + o(\|h\|^2)$$

lorsque $\|h\|$ tend vers 0.

On dispose maintenant d'un moyen permettant de détecter la plupart des extrema locaux (mais pas tous, comme pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$).

PROPOSITION 6 — Conditions nécessaires/suffisantes pour présenter un extremum en un point

Pour que $f \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R})$ possède un minimum en a :

- il est *nécessaire* d'avoir $\nabla f(a) = 0$ et $H_f(a) \in \mathcal{S}_n^+(\mathbb{R})$;
- il est *suffisant* d'avoir $\nabla f(a) = 0$ et $H_f(a) \in \mathcal{S}_n^{++}(\mathbb{R})$.

REMARQUES :

- Si $H_f(a)$ possède 0 comme valeur propre (« shit happens »), on ne peut rien dire.
- En pratique si $n = 2$, les deux valeurs propres sont non nulles et de même signe si et seulement si leur produit est strictement positif : il suffit donc de regarder le déterminant de la hessienne.
  - s'il est strictement négatif, alors la fonction présente un « point-col » en a : il n'y a ni minimum ni maximum ;
  - s'il est nul on ne peut rien dire ;
  - s'il est strictement positif alors les deux valeurs propres sont de même signe strict, donc il y a un minimum local (si elles sont positives) ou un maximum local (si elles sont négatives). Sans les calculer, le signe de la trace dit quel est le signe commnu de ces deux valeurs propres !

3.3 Des équations aux dérivées partielles

Il s'agit d'équations différentielles... faisant intervenir les dérivées partielles. Il n'y a ni résultat ni technique générale, mais quelques exemples significatifs à avoir compris. On réalise systématiquement une analyse-synthèse.

— *La base* : $\dfrac{\partial^2 f}{\partial x \partial y} = 0$.

  On fixe x_0 et on considère $y \mapsto \dfrac{\partial f}{\partial x}(x_0, y)$...
  On trouve finalement comme solutions les applications de la forme

$$(x, y) \mapsto \varphi_1(x) + \varphi_2(y)$$

  avec $\varphi_1, \varphi_2 \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$.

— *Équation de transport* : $\dfrac{\partial f}{\partial t} = K\dfrac{\partial f}{\partial x}\cdot$ On cherche la valeur de $f(x, t)$, et on va « poser $(u, v) = (x + Kt, t)$ », c'est-à-dire définir g telle que $g(u, v) = f(x, t)$ (avec les relations précédentes), soit encore : définir $g(u, v) = f(u - Kv, v)$. Les solutions de l'équation initiale sont alors les $(x, t) \mapsto \varphi(x + Kt)$, avec $\varphi \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$.

— *Équation de propagation* (ou de d'Alembert) : $\dfrac{\partial^2 f}{\partial x^2} = \dfrac{1}{c^2}\dfrac{\partial^2 f}{\partial t^2}\cdot$ « On pose $(u, v) = (x - ct, x + ct)$ »... et finalement les solutions sont les applications $(x, t) \mapsto \varphi_1(x - ct) + \varphi_2(x + ct)$, avec φ_1 et φ_2 dans $\mathcal{C}^2(\mathbb{R}, \mathbb{R})$.

— L'*équation de Fourier*, décrivant la propagation de la chaleur, est : $\dfrac{\partial^2 f}{\partial x^2} = K\dfrac{\partial f}{\partial t}\cdot$ Sa résolution nécessite des outils ad hoc... dont vous ne disposez pas !

Exercice 12. *En réalisant un changement de variable polaire, résoudre sur* $\mathbb{R}^2 \setminus \{0\}$ *:*

$$x\frac{\partial f}{\partial x}(x, y) + y\frac{\partial f}{\partial y}(x, y) = \sqrt{x^2 + y^2}.$$

Quelles solutions sont prolongeables sur $\mathbb{R}^2$ *?*

Exercice 13. *Résoudre l'équation de Navier-Stokes :*

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{1}{\rho}\nabla p + \gamma\nabla^2\mathbf{u} + \frac{1}{\rho}\mathbf{F}$$

4 Un peu de géométrie différentielle

Attention : il y a peu de dessins (essentiels à la compréhension de cette partie de cours) dans ce poly... vous voudrez donc bien les ajouter live !

4.1 Qu'est-ce qu'une courbe du plan ?

Trois points de vue sont en concurrence :
— graphe d'une application $g : I \to \mathbb{R}$ (l'ensemble des $(x, g(x))$, pour x décrivant I) ;
— courbe paramétrée (l'ensemble des $(f_1(t), f_2(t))$, pour t décrivant I, avec $f_1, f_2 : I \to \mathbb{R}$;
— courbe définie par une équation (ensemble des (x, y) tels que $f(x, y) = 0$).

Exemple : Le cercle trigonométrique !
— c'est la réunion des graphes des applications $x \in [-1, 1] \mapsto \pm\sqrt{1 - x^2}$;
— c'est la courbe paramétrée $\begin{cases} x = \cos t \\ y = \sin t \end{cases}$;
— c'est la courbe d'équation $x^2 + y^2 = 1$.

4.2 Au voisinage d'un point régulier

DÉFINITION 10 — *Point régulier*

> Lorsque $\nabla f(x_0, y_0) \neq \overrightarrow{0}$, on dit que (x_0, y_0) est un point régulier de la courbe d'équation $f(x, y) = f(x_0, y_0)$.

THÉORÈME 5 — *Au voisinage d'un point régulier*

> Si $\nabla f(x_0, y_0) \neq \overrightarrow{0}$ et que $\mathcal{C}$ désigne l'ensemble (*a.k.a. : la* ligne de niveau) d'équation $f(x, y) = f(x_0, y_0)$, alors :
> — $\mathcal{C}$ possède une paramétrisation locale $(x, y) = (f_1(t), f_2(t))$ avec f_1 et f_2 de classe $\mathcal{C}^1$ (on peut alors parler de courbe $\mathcal{C}^1$) ;
> — la tangente à $\mathcal{C}$ en $M_0 = (x_0, y_0)$ est orthogonale à $\nabla f(M_0)$; elle a donc pour équation $\langle \nabla f_{M_0} | \overrightarrow{M_0 M} \rangle = 0$, c'est-à-dire
>
> $$(x - x_0)\frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0)\frac{\partial f}{\partial y}(x_0, y_0) = 0$$

EXEMPLES :

— Pour le cercle d'équation $x^2 + y^2 = 1$, la tangente en (x_0, y_0) a donc pour équation $2x_0(x - x_0) + 2y_0(y - y_0) = 0$, c'est-à-dire $x_0 x + y_0 y = 1$. Il est assez raisonnable que le vecteur $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ soit normal à la tangente, non ?

— Dans le cas d'une courbe paramétrée $y = g(x)$, on peut prendre le point de vue « équationnel » $f(x, y) = 0$ en posant $f(x, y) = y - g(x)$. Tous les points de la courbe sont alors réguliers, puisque $\nabla f = \begin{pmatrix} -g'(x) \\ 1 \end{pmatrix}$. On retrouve par ailleurs le fait que la tangente, évidemment (?) dirigée par $\begin{pmatrix} 1 \\ g'(x_0) \end{pmatrix}$ (vous ne trouvez pas ça évident ? Alors reprenez votre cours de terminale sur la dérivation et les tangentes...) est bien orthogonale à $\nabla f(x_0, y_0)$.

Exercice 14. *Comment un ordinateur peut-il raisonnablement faire pour tracer :*

— *Le graphe d'une application $f : \mathbb{R} \to \mathbb{R}$?*
— *Une courbe paramétrée donnée par $f(t) = (x(t), y(t))$?*
— *Une courbe d'équation implicite $F(x, y) = 0$?*

4.3 Gradient et ligne de niveau

Les lignes de niveau d'une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ sont les ensembles

$$\mathcal{E}_K = \{M \in \mathbb{R}^2;\ f(M) = K\}$$

On a vu qu'en des points réguliers de f, les lignes de niveau sont de braves courbes paramétrées. On peut avoir des allures différentes soit globalement avec $x^2 + y^2 = 0$ ou $x^2 + y^2 = -1$... soit localement avec $x^2 - y^2 = 0$ (au voisinage de $(0, 0)$, cette ligne de niveau ne ressemble pas trop à une courbe...)

EXEMPLES :

— Les isobares (resp. isothermes) représentent sur une carte météo (ou en thermo les ensembles de points de même pression (resp. température.
— En électrostatique, les équipotentielles sont les ensembles de points ayant le même potentiel.

Ce qu'on retiendra : les lignes de niveau ont leurs tangentes orthogonales aux gradients du champ scalaire associée (c'est le théorème 4) ; le gradient de f montre dans quelle direction aller pour augmenter la valeur de f !

Exercice 15. *Représenter différentes lignes de niveau pour les applications suivantes :*

$$f_1 : (x, y) \mapsto x^2 + y^2; \qquad f_2 : (x, y) \mapsto xy; \qquad f_3 : (x, y) \mapsto x^2 - y^2$$

Représenter également quelques gradients...

Petite question pour terminer ce paragraphe : à votre avis, comment fait un logiciel dédié pour représenter les lignes de niveau d'une application donnée ?

4.4 Qu'est-ce qu'une surface de $\mathbb{R}^3$?

Trois points de vue en concurrence :
- « graphe » d'une application $f : I \times J \to \mathbb{R}$ (l'ensemble des $(x, y, f(x, y))$, pour $(x, y) \in I \times J$) ;
- surface paramétrée (l'ensemble des $(f_1(t), f_2(t), f_3(t))$, pour $t \in I$, avec $f_1, f_2, f_3 : I \to \mathbb{R}$;
- surface définie par une équation (ensemble des (x, y, z) tels que $f(x, y, z) = 0$).

Exemple : La sphère unité !
- c'est la réunion des graphes des applications $(x, y) \mapsto \pm\sqrt{1 - (x^2 + y^2)}$;
- c'est la surface de paramétrisation $\begin{cases} z = \cos\lambda \\ x = \sin\lambda\cos\theta \\ y = \sin\lambda\sin\theta \end{cases}$
- c'est la courbe d'équation $x^2 + y^2 + z^2 = 1$.

Les surfaces d'équation $x^2 \pm y^2 \pm z^2 = \pm 1$ fourniront des tas d'exemples simples et de bon goût !

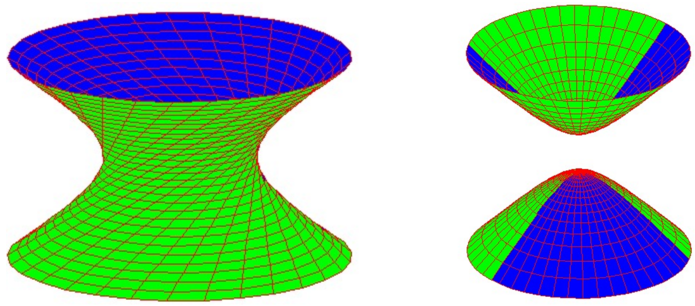


Figure 2 – Hyperboloïde à une nappe ($x^2 + y^2 - z^2 = 1$) et deux nappes ($x^2 - y^2 - z^2 = 1$)

Définition 11 — *Point régulier*

Lorsque $\nabla f(x_0, y_0, z_0) \neq \overrightarrow{0}$, on dit que (x_0, y_0, z_0) est un point régulier de la surface d'équation $f(x, y, z) = f(x_0, y_0, z_0)$.

4.5 Au voisinage d'un point régulier

Théorème 6 — *Au voisinage d'un point régulier...*

Lorsque $\nabla f(x_0, y_0, z_0) \neq \overrightarrow{0}$ et que $\mathcal{S}$ désigne l'ensemble (*a.k.a. : la* surface de niveau) d'équation $f(x, y, z) = f(x_0, y_0, z_0)$, alors :
- $\mathcal{S}$ possède une paramétrisation locale $(x, y, z) = (f_1(u, v), f_2(u, v), f_3(u, v))$ avec f_1, f_2 et f_3 de classe $\mathcal{C}^1$ (on peut alors parler de surface $\mathcal{C}^1$) ;
- le plan tangent à $\mathcal{S}$ en $M_0 = (x_0, y_0, z_0)$ possède pour vecteur normal $\nabla f(M_0)$; il a donc pour équation $\langle \nabla f_{M_0} | \overrightarrow{M_0M} \rangle = 0$, c'est-à-dire

$$(x - x_0)\frac{\partial f}{\partial x}(x_0, y_0, z_0) + (y - y_0)\frac{\partial f}{\partial y}(x_0, y_0, z_0) + (z - z_0)\frac{\partial f}{\partial z}(x_0, y_0, z_0) = 0$$

Par exemple, le plan tangent à la sphère d'équation $x^2 + y^2 + z^2 = R^2$ en (x_0, y_0, z_0) a pour équation (après une demi-ligne de calcul) : $xx_0 + yy_0 + zz_0 = R^2$.

On retrouve des résultats similaires aux lignes de niveau pour les surfaces de niveau : le gradient est orthogonal à ces surfaces, et indique la direction à suivre pour grimper les niveaux !

On peut vouloir dessiner une courbe sur une surface ! Il s'agit d'un ensemble de points de la forme $\{\gamma(t)\,,\ t \in I\}$ avec $\gamma \in \mathcal{C}^1(I, \mathbb{R}^3)$ telle que pour tout $t \in I$, $\Phi(\gamma(t)) = 0$, où $\Phi \in \mathcal{C}^1(\mathbb{R}^3, \mathbb{R})$ a son gradient qui ne s'annule pas (au moins localement).

Exemples :
- $\gamma : t \mapsto \left(\frac{1}{2}\cos t, \frac{1}{2}\sin t, \frac{\sqrt{3}}{2}\right)$ dessine un cercle sur une sphère.

— $\gamma : t \mapsto (\cos t \sin(10t), \sin t \sin(10t), \cos(10t))$ dessine un bidule sur une sphère.
La seule chose à savoir à ce sujet est que quand on trace une courbe sur une surface, les vecteurs tangents à cette courbe sont dans le plan tangent à la surface.

Exercice 16. *Énoncer plus formellement puis prouver le résultat précédent.*

On terminer par un exercice classique... jadis ! On pourra se reporter à la figure 2 pour visualiser le phénomène

Exercice 17. *Montrer que l'hyperboloïde à une nappe d'équation $x^2 + y^2 - z^2 = 1$ est une réunion de droites.*
On parle de surface réglée.