

Chapitre 27 : Fonctions à deux variables.

MPSI Lycée Camille Jullian

21 juin 2024

Quel que soit le temps passé à faire des mathématiques, ce n'est jamais du temps perdu.

Cédric VILLANI.

*C'est un romain qui entre dans un bar en levant deux doigts :
« Je voudrais cinq bières s'il vous plait. »*

Pour ce dernier chapitre de l'année, nous allons aborder la notion de fonctions à deux variables, que vous généraliserez l'an prochain (aucune raison a priori de se limiter à deux variables!). Le but de ce chapitre est de développer des outils analytiques d'étude de ces fonctions, en adaptant ce que vous maîtrisez déjà sur les fonctions à deux variables, c'est-à-dire essentiellement les notions de continuité et de dérivabilité (et les calculs qui s'y rapportent comme les développements limités). Mais la simple définition (et l'interprétation) de ces outils devient nettement plus compliquée avec l'ajout de la deuxième variable, ce qui nous forcera à faire un rapide détour par l'étude de la topologie du plan avant de nous lancer dans les fonctions à deux variables proprement dites. On complètera également le chapitre avec quelques considérations hors-programme sur les calculs d'intégrales faisant intervenir des fonctions à deux (ou plus) variables, très utiles en physique.

Objectifs du chapitre :

- comprendre les bases du vocabulaire topologique : ouverts, fermés, et leur lien avec la notion de continuité d'une fonction.
- savoir manipuler correctement les dérivées de fonctions à deux variables.

1 Topologie dans le plan.

Ce court paragraphe consacré à la topologie de $\mathbb{R}^2$ reprendra simplement les notions déjà évoquées dans $\mathbb{R}$ en introduction du chapitre 9 (ouvert, voisinage, etc). Ces notions sont fondamentales si on veut définir correctement (et comprendre !) la notion de convergence (pour les suites dans le chapitre 9, mais nous avons repris la même définition pour définir ensuite les limites de fonctions, ce que nous referons ici). Il s'agit donc de généraliser notamment les notions d'intervalles ouverts ou fermés de $\mathbb{R}$, fondamentales dans les hypothèses de beaucoup de théorèmes d'analyse. Sauf mention contraire explicite, la distance utilisée sur $\mathbb{R}^2$ sera toujours la **distance euclidienne**, c'est-à-dire la distance définie par le produit scalaire canonique de $\mathbb{R}^2$: si $A(x, y)$ et $B(x', y')$ sont deux points du plan, alors $AB = \sqrt{(x - x')^2 + (y - y')^2}$ (coordonnées exprimées dans la base canonique).

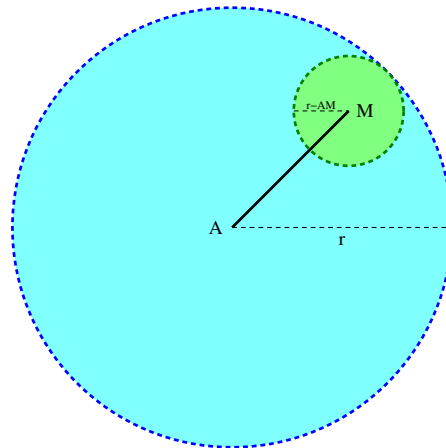
Définition 1. Soit $A \in \mathbb{R}^2$ et $r > 0$, on appelle **boule ouverte de centre A et de rayon r** l'ensemble $B(A, r) = \{M \in \mathbb{R}^2 \mid AM < r\}$. On appellera de même **boule fermée** de centre A et de rayon r l'ensemble $B_f(A, r) = \{M \in \mathbb{R}^2 \mid AM \leq r\}$.

Remarque 1. On pourrait étendre la notion de boule fermée en autorisant un rayon nul (la boule ne contient alors que le point A), mais pas celle de boule ouverte (qui deviendrait vide). Bien sûr, dans le cas du plan, on pourra aussi utiliser le terme de **disque** (ouvert ou fermé) plutôt que celui de boule, mais cette définition est en fait destinée à être généralisée à n'importe quel espace muni d'une distance (en particulier tous les espaces préhilbertiens).

Définition 2.

- Un sous-ensemble $V \subset \mathbb{R}^2$ est un **voisinage du point A** s'il contient la boule ouverte $B(A, r)$ pour un certain rayon $r > 0$.
- V est **ouvert** s'il est voisinage de chacun de ses points, autrement dit si $\forall A \in D, \exists r > 0, B(A, r) \subset D$, ou encore $\exists r > 0, \forall M \in \mathbb{R}^2, AM < r \Rightarrow M \in D$.
- Un sous-ensemble de $\mathbb{R}^2$ est **fermé** si et seulement s'il est le complémentaire d'un ouvert de $\mathbb{R}^2$.

Exemples : Une boule ouverte est toujours ouverte (encore heureux!). En effet, si $M \in B(a, r)$, alors, par définition de ce qu'est une boule ouverte, $AM < r$. En notant $\varepsilon = r - AM$, on a donc $\varepsilon > 0$, et $B(M, \varepsilon) \subset B(a, r)$: si $MN < \varepsilon$, alors l'inégalité triangulaire permet d'affirmer que $AN \leq AM + MN < AM + \varepsilon = r$.



Comme on pouvait s'en douter, une boule fermée est au contraire toujours fermée (démonstration similaire à la précédente). Attention par contre à ne pas croire que fermé est « le contraire d'ouvert ». Il existe des ensembles qui sont à la fois ouverts et fermés ($\mathbb{R}^2$ tout entier) et surtout énormément d'ensembles qui ne sont ni ouverts ni fermés (par exemple $\mathbb{Q}^2$, cf les calculs effectués plus bas sur cet exemple particulier).

Le demi-plan « ouvert » $D = \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$ est un ouvert. Si on remplace l'inégalité stricte par une inégalité large, on obtient un demi-plan fermé. De façon générale, un ensemble de points délimité par une droite (ou un cercle) sera ouvert s'il ne contient pas cette droite, fermé s'il la contient.

Remarque 2. Ces définitions semblent être intimement liées au choix de la distance euclidienne qu'on a effectué en début de paragraphe. Si on change cette distance pour une autre (par exemple en prenant un autre produit scalaire), on imagine a priori que les boules ne seront plus les mêmes (c'est effectivement le cas) et donc que les notions de voisinage et d'ouverts/fermés vont aussi changer. En

fait, pas du tout : vous verrez l'an prochain que, dans un espace vectoriel de dimension finie comme $\mathbb{R}^2$, toutes les normes sont **équivalentes**, ce qui revient exactement à dire qu'on peut toujours inclure une boule ouverte centrée en A définie à l'aide d'une certaine norme dans une boule ouverte centrée en A (mais de rayon différent) définie à l'aide d'une autre norme (je vous laisse méditer les exemples donnés au tout début du chapitre précédent à ce sujet). Ainsi, même si les boules ne sont pas les mêmes, les ensembles ouverts et fermés restent exactement les mêmes. On dit que la topologie de l'ensemble $\mathbb{R}^2$ est indépendante du choix de la norme.

Cette propriété n'est par contre pas du tout vérifiée dans un espace vectoriel de dimension infinie.

Prenons l'exemple de $\mathbb{R}[X]$, et des deux produits scalaires $\varphi(P, Q) = \int_0^1 P(t)Q(t) dt$ et $\psi(P, Q) =$

$\sum_{k=0}^{+\infty} a_k b_k$ (ce deuxième produit scalaire est simplement le produit scalaire canonique obtenu à partir des coefficients des deux polynômes, la somme n'est pas réellement infinie en pratique). Ces deux

produits scalaires définissent des notions de distance et de norme qui ne sont pas équivalentes : pour le second produit scalaire, tous les monômes X^n ont une norme égale à 1, donc la boule ouverte centrée en 0 et de rayon 1 ne contiendrait aucun de ces monômes. Mais pour le premier produit

scalaire, on a $\|X^n\| = \sqrt{\int_0^1 t^{2n} dt} = \frac{1}{\sqrt{2n+1}}$. Une boule ouverte centrée en 0, quel que soit son rayon, contiendrait donc nécessairement un de ces monômes, et ne pourrait jamais être incluse dans la boule ouverte de rayon 1 définie à l'aide de l'autre distance. Autrement dit, la boule ouverte centrée en 0 et de rayon 1 pour le produit scalaire canonique est un ouvert si on utilise la première norme, mais pas si on utilise la seconde (ce n'est plus un voisinage de 0).

Proposition 1. Une union quelconque d'ouverts de $\mathbb{R}^2$ est toujours un ouvert. Une intersection finie d'ouverts de E est toujours un ouvert.

Remarque 3. Symétriquement, toute intersection quelconque ou union finie de fermés est donc fermée.

Démonstration. C'est complètement trivial pour l'union : si un point A appartient à une union d'ouverts, alors il appartient à l'un des ouverts de cette union, et cet ouvert contient une boule ouverte centrée en A . L'union la contient également. Pour l'intersection, il faut se battre un peu plus : supposons que $A \in \bigcap_{1 \leq i \leq n} V_i$, avec V_i ouvert. Pour tout entier i , il existe alors un rayon $r_i > 0$ tel que $B(A, r_i) \subset V_i$. On pose alors simplement $r = \min_{1 \leq i \leq n} r_i$, et $B(A, r) \subset \bigcap V_i$. $\square$

Définition 3. Soit $V \subset \mathbb{R}^2$ et $A \in \mathbb{R}^2$.

- A est **intérieur** à V si V est un voisinage de A . On note $\overset{\circ}{V}$ l'ensemble des points intérieurs à V , appelé **intérieur** de l'ensemble V . L'intérieur de V est le plus grand ouvert contenu dans V .
- A est **adhérent** à V si, pour tout voisinage W de A , $V \cap W \neq \emptyset$. On note $\overline{V}$ l'ensemble des points adhérents à V , appelé **adhérence** de V . L'adhérence de V est le plus petit fermé contenant V .
- on dit que A est à la **frontière** de l'ensemble V si A est adhérent à V et à son complémentaire. On note ∂V l'ensemble des points à la frontière de V , appelé **frontière** de V . On a $\partial V = \overline{V} \setminus \overset{\circ}{V}$.

Exemples : L'intérieur de $B_f(A, r)$ est $B(A, r)$. Réciproquement, $\overline{B(A, r)} = B_f(A, r)$. La frontière de $B(A, r)$ (ou de $B_f(A, r)$) est le **cercle** de rayon r centré en A .

L'ensemble $\mathbb{Q}^2$ a l'intéressante propriété suivante : son intérieur est vide ($\mathbb{Q}^2$ ne peut contenir aucune boule ouverte, puisqu'une telle boule ouverte contiendra toujours des points à coordonnées irrationnelles), mais son adhérence est égale à $\mathbb{R}^2$ (c'est une conséquence de la densité de $\mathbb{Q}$ dans $\mathbb{R}$: quel

que soit le point $A(x, y) \in \mathbb{R}^2$, on peut construire deux suites (x_n) et (y_n) de rationnels tels que $\lim x_n = x$ et $\lim y_n = y$, et les points $A_n(x_n, y_n)$ appartiennent alors à $\mathbb{Q}^2$ et peuvent être rendus aussi proches qu'on le souhaite du point A , ce qui prouve que $A \in \overline{\mathbb{Q}^2}$. Autrement dit, $\partial\mathbb{Q}^2 = \mathbb{R}^2$. Tout point du plan est à la frontière de $\mathbb{Q}^2$ (mais aussi à la frontière de son complémentaire!).

2 Fonctions à deux variables.

2.1 Aspect graphique.

Définition 4. Une **fonction à deux variables** est une application $f : \mathcal{D} \rightarrow \mathbb{R}$, où $\mathcal{D}$ est un sous-ensemble du plan $\mathbb{R}^2$ appelé domaine de définition de la fonction f .

Exemples : La fonction $f : (x, y) \mapsto x^3 + 2x^2y + xy^3 - 4y^2$ est une fonction à deux variables définie sur $\mathbb{R}^2$ tout entier. La fonction $g : (x, y) \mapsto \ln(x + y - 1)$ est une fonction définie sur l'ensemble des couples (x, y) vérifiant $x + y - 1 > 0$, qui se trouve être le demi-plan supérieur ouvert délimité par la droite d'équation $y = 1 - x$. La fonction $h : (x, y) \mapsto \sqrt{4 - x^2 - y^2}$ est définie sur la boule fermée $B_f(O, 2)$, en notant $O(0, 0)$ l'origine du repère canonique de $\mathbb{R}^2$.

Définition 5. Toute fonction obtenue comme combinaison linéaire des fonctions $(x, y) \mapsto x^n y^p$, avec $(n, p) \in \mathbb{N}^2$, est appelée **fonction polynômiale à deux variables**.

Remarque 4. Toutes les fonctions polynômiales sont définies sur $\mathbb{R}^2$ tout entier. Les **fonctions coordonnées** $(x, y) \mapsto x$ et $(x, y) \mapsto y$ sont des cas particuliers de fonctions polynômiales.

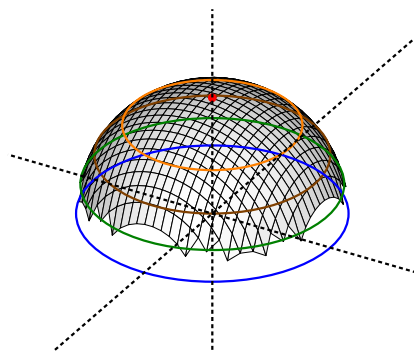
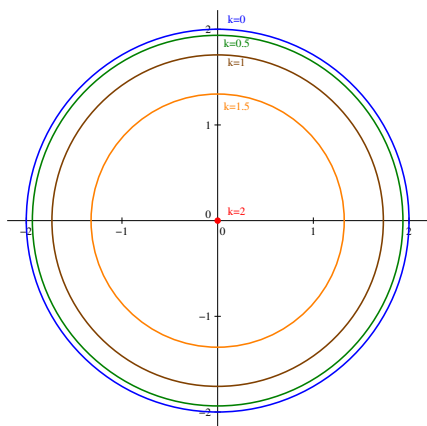
Définition 6. La **représentation graphique** d'une fonction à deux variables dans un repère $(O, \vec{i}, \vec{j}, \vec{k})$ de l'espace est la surface définie par l'ensemble des points $M(x, y, z)$ vérifiant $z = f(x, y)$.

Remarque 5. Il est très difficile en général de visualiser ce genre de représentations graphiques (et de les dessiner dans un plan), c'est pourquoi on en est souvent réduit à étudier des « morceaux » de cette surface représentative. Le plus simple pour cela est de fixer une valeur de x ou de y pour se ramener au cas d'une courbe représentative de fonction à une variable (c'est le principe des applications partielles qu'on définira un peu plus loin), mais on peut aussi fixer la valeur de la cote z des points qu'on cherche à représenter. C'est l'objet de la définition qui suit.

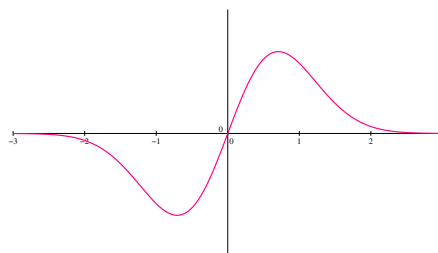
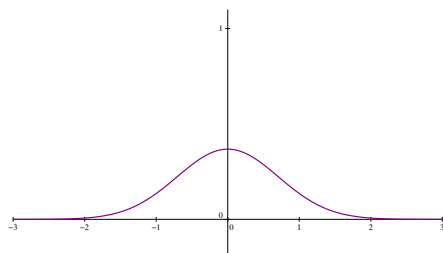
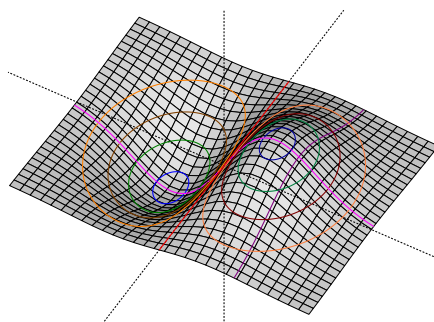
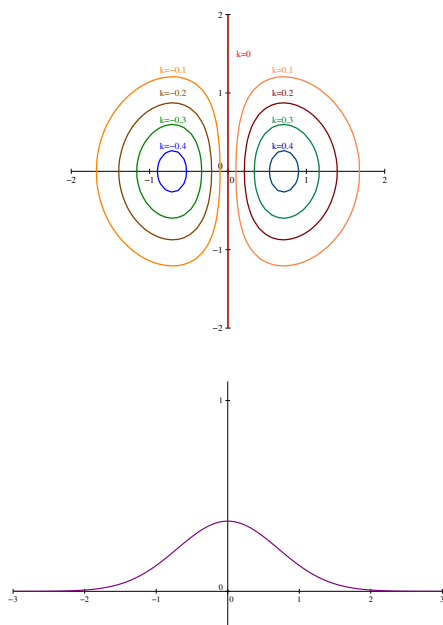
Définition 7. Soit $k \in \mathbb{R}$ et f une fonction de deux variables, la **ligne de niveau k** de la fonction f est l'ensemble des couples (x, y) vérifiant $f(x, y) = k$.

Remarque 6. Il s'agit donc de la coupe de la surface représentative de f par le plan « horizontal » d'équation $z = k$. La plupart du temps, une ligne de niveau n'est pas la courbe représentative d'une fonction à une variable.

Exemple : reprenons la fonction définie par $h(x, y) = \sqrt{4 - x^2 - y^2}$, ses lignes de niveaux sont définies par les équations $x^2 + y^2 = 4 - k^2$, si $k \geq 0$ (si $k < 0$, la ligne de niveau correspondante est vide). On reconnaît l'équation du cercle de centre 0 et de rayon $\sqrt{4 - k^2}$ si $k \leq 2$, un ensemble vide sinon. Ci-dessous on a représenté quelques unes de ces lignes de niveau (niveau 0 en bleu, niveau 2, réduite à un point, en rouge, les valeurs sont indiquées à côté des cercles), il ne reste plus qu'à relier mentalement ces courbes dans l'espace pour imaginer la surface représentative de la fonction f , dessinée à droite (oui, l'outil que j'utilise pour les tracés de surfaces dans l'espace gère assez mal le « bord » du domaine de définition, la surface se prolonge en fait jusqu'au cercle bleu). Ici, la surface correspondante est une demi-sphère centrée en O de rayon 2.



Un autre exemple avec les mêmes conventions, pour $f(x, y) = xe^{-x^2-y^2}$. On a d'abord représenté quelques lignes de niveau (les valeurs correspondantes sont indiquées à côté, comme pour l'exemple précédent), puis la surface complète avec les mêmes lignes de niveau. Enfin, on a ajouté cette fois-ci les représentations graphiques de fonctions partielles, ici les deux courbes des fonctions $y \mapsto f(1, y) = e^{-1-y^2}$ et $x \mapsto f(x, 0) = xe^{-x^2}$, également indiquées sur la surface complète (en violet et rose respectivement).



2.2 Continuité.

Définition 8. Soit f une fonction à deux variables définie sur un ensemble $D \subset \mathbb{R}^2$, et $a \in D$, alors f admet pour limite l en a si quel que soit le voisinage V de l , il existe un voisinage W de a tel que $f(W \cap D) \subset V$.

Remarque 7. Puisqu'un voisinage n'est autre qu'un ensemble contenant une boule ouverte centrée en a (ou un intervalle ouvert centré en l si on travaille dans $\mathbb{R}$), cette définition est équivalente à la définition plus classique suivante : $\forall \varepsilon > 0, \exists \eta > 0, \forall b \in D, d(a, b) < \eta \Rightarrow |f(b) - l| < \varepsilon$. On pourrait définir, comme dans le cas de fonctions à une variable, des limites infinies en un point a , mais nous n'en aurons pas vraiment besoin dans ce chapitre, on va donc s'en dispenser.

Définition 9. Une fonction f définie sur D est **continue** en $a \in D$ si f admet pour limite $f(a)$ en a .

Une fonction f est continue sur D si elle est continue en tout point de D .

Exemples : les fonctions polynômiales sont définies et continues sur $\mathbb{R}^2$ tout entier.

Remarque 8. Pour les plus motivés, il existe en fait une définition nettement plus élégante et générale de la continuité. Si on appelle « ouvert de D » tous les ensembles qui sont intersections de D avec un ouvert de $\mathbb{R}^2$, alors f est continue sur D si et seulement si l'image réciproque de tout ouvert de $\mathbb{R}^2$ est un ouvert de D .

Théorème 1. Théorèmes généraux sur la continuité de fonctions à deux variables.

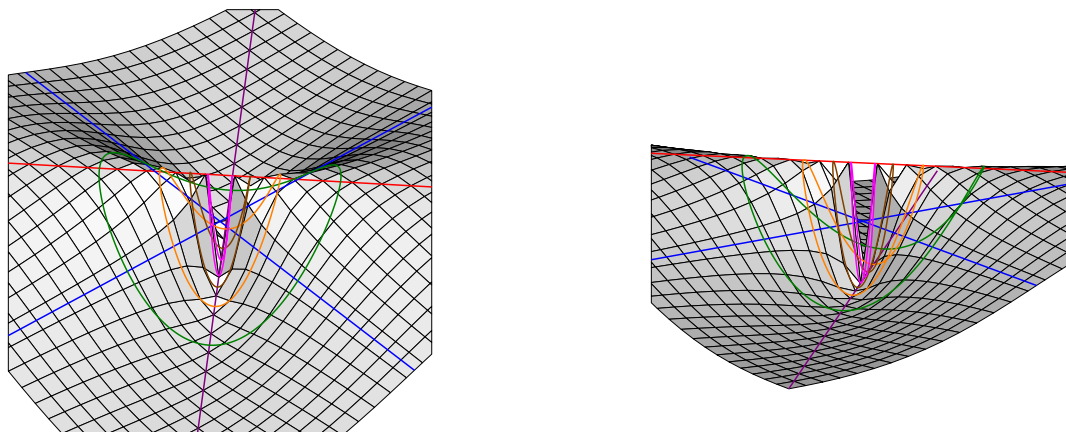
Toute somme, produit, quotient (si le dénominateur ne s'annule pas) de fonctions continues sur D reste continue sur D . La composée d'une fonction définie sur D et à valeurs dans $I \subset \mathbb{R}$ par une fonction (à une variable) continue sur I est continue sur D .

Exemple : La fonction $h : (x, y) \mapsto \sqrt{4 - x^2 - y^2}$ déjà évoquée plus haut est continue sur son ensemble de définition $D = B_f(O, 2)$ puisqu'elle est la composée d'une fonction polynômiale continue et positive sur $B_f(O, 2)$, et de la fonction racine carrée qui est continue sur $\mathbb{R}^+$.

Démonstration. Les preuves seraient en fait identiques à celles vues sur les fonctions à une variable, en remplaçant bien sûr toutes les valeurs absolues par des distances dans $\mathbb{R}^2$. $\square$

Exemple : à quoi peut bien ressembler une fonction qui n'est **pas** continue en un certain point de $\mathbb{R}^2$? Faut-il chercher des horreurs invraisemblables à bases de partie entières ou autres choses encore moins sympathiques ? On peut, bien sûr, mais il existe en fait des fonctions à deux variables définies de façon simple et qui ne sont pas continues partout où elles sont définies. Cela est notamment du au fait que les limites dans $\mathbb{R}^2$ sont plus compliquées à appréhender que celles dans $\mathbb{R}$, car on peut approcher un point du plan de plein de façon différentes (alors que sur une droite on ne peut s'approcher que par la gauche ou par la droite, la notion de limite à gauche ou à droite n'a d'ailleurs plus aucun sens dans $\mathbb{R}^2$). Posons par exemple $f(x, y) = \frac{2xy}{x^2 + y^2}$. La fonction f est définie et continue (par théorèmes généraux) sur $\mathbb{R}^2 \setminus \{(0, 0)\}$, mais admet-elle une limite en $(0, 0)$ qui permettrait de la prolonger par continuité à $\mathbb{R}^2$ tout entier ? Si on effectue le raisonnement naïf « je fais tendre x vers 0 à y fixé puis je fais tendre y vers 0 », on peut être tenté de penser que cette limite existe et va être nulle. En effet, à $y \neq 0$ fixé, $\lim_{x \rightarrow 0} \frac{2xy}{x^2 + y^2} = 0$. Mais quand on effectue ce raisonnement, on se rapproche en fait du point $(0, 0)$ « en suivant les axes » (on se déplace d'abord parallèlement à l'axe des abscisses, puis parallèlement à l'axe des ordonnées), ce qui est loin d'être la seule possibilité. Imaginons qu'à la place on se déplace suivant la diagonale d'équation $y = x$. Il faut alors calculer $f(x, x) = \frac{2x^2}{x^2 + x^2} = 1$, puis faire tendre x vers 0. Manifestement, on n'obtiendra pas une limite nulle. Ces deux calculs prouvent en fait qu'il ne peut pas y avoir de limite pour f en $(0, 0)$ (s'il y en avait une, elle serait unique). Allons un peu plus loin en faisant ce qu'on appelle techniquement un « passage en polaires » en exprimant f non pas en fonction des coordonnées cartésiennes x et y du point dont on cherche à calculer l'image, mais en fonction de ses coordonnées polaires r et θ . Ces dernières ne sont rien d'autre que l'équivalent du module et de l'argument des nombres complexes, on pose simplement $x = r \cos(\theta)$ et $y = r \sin(\theta)$ (ou, dans l'autre sens, $r = \sqrt{x^2 + y^2}$, et θ représente l'angle avec le premier axe, qu'on ne peut pas exprimer à l'aide d'une seule formule simple en fonction de x et de y). Faire tendre (x, y) vers $(0, 0)$ revient alors simplement à faire tendre la distance r vers 0. Or, $f(r, \theta) = \frac{2r \cos(\theta)r \sin(\theta)}{r^2} = \sin(2\theta)$. Cette valeur, qui ne dépend pas de r , n'a pas de limite quand r tend vers 0, elle peut se balader à loisir entre -1 et 1 selon l'angle avec lequel on approche de $(0, 0)$. Pour visualiser tout ça, on a dessiné ci-dessous la surface représentative de f , ainsi que les

deux premiers axes du repère sur lesquels f s'annule (en bleu), la diagonale $y = x$ où f est constante égale à 1 (en rouge), l'autre diagonale $y = -x$ où f est aussi constante égale à -1 (en violet), et les courbes obtenues en fixant la distance à l'origine du point (la valeur de r en coordonnées polaires) égale à 2, 1, $\frac{1}{2}$ ou $\frac{1}{5}$ (en vert, orange, marron et rose). Je vous mets même deux vues différentes pour essayer d'aider à voir ce qui se passe :



2.3 Dérivées partielles.

On ne peut pas étudier les variations d'une fonction à deux variables comme on le fait pour une fonction à une variable, puisque la simple notion de fonction croissante ou décroissante n'a pas d'équivalent quand on passe à deux variables. Il est cependant intéressant de calculer un analogue de la dérivée dans ce cadre, qui permet notamment de trouver les minima ou maxima de la fonction, comme c'est le cas pour une fonction à une variable. Le problème, comme dans le cas du calcul de limites évoqué plus haut, est qu'on peut très bien dériver suivant tout un tas de directions (après tout, un calcul de dérivée n'étant rien d'autre qu'un calcul de limite, c'est logique).

Définition 10. Soit f une fonction définie au voisinage de $a \in D$ et $u \in \mathbb{R}^2$, f est **dérivable en a dans la direction u** si le taux d'accroissement directionnel $\tau : h \mapsto \frac{f(a + hu) - f(a)}{h\|u\|}$ admet une limite finie quand h tend vers 0.

Remarque 9. Dans cette définition, on dérive en fait tout à fait classiquement une fonction à une seule variable, dont la courbe est obtenue en effectuant une coupe de la surface représentative de f par un plan vertical contenant le vecteur u . Le nombre obtenu représente tout aussi classiquement le coefficient directeur de la droite tangente à cette courbe dans le plan de coupe.

Pour obtenir une bonne vision d'ensemble de notre surface d'un point de vue dérivation, il faudrait donc théoriquement étudier la dérivabilité en a dans toutes les directions possibles (autrement dit prendre tous les choix possibles de vecteurs u de norme 1 par exemple), ce qui est pour le moins compliqué. En pratique, on se contentera de privilégier (très arbitrairement, mais ça suffira pour toutes les applications pratiques) deux de ces vecteurs, et on va prendre les plus simples : les vecteurs unitaires directeurs des deux axes. Cela revient à étudier les coupes de notre surface par des plans parallèles aux plans (xOz) et (yOz) (avec les notations traditionnelles pour les axes du repère).

Définition 11. Soit $f : (x, y) \mapsto f(x, y)$ une fonction à deux variables, les **applications partielles** associées à f au point $a = (x_0, y_0)$ sont les deux fonctions à une variable $f_{x, y_0} : x \mapsto f(x, y_0)$ et $f_{x_0, y} : y \mapsto f(x_0, y)$.

Remarque 10. Certaines des courbes tracées sur nos surfaces dans les paragraphes précédents étaient des courbes d'applications partielles (mais pas toutes).

Définition 12. Soit f une fonction à deux variables définie sur D , et $a = (x_0, y_0) \in D$, les **dérivées partielles** sont les dérivées de f en a suivant les directions données par les vecteurs $(1, 0)$ et $(0, 1)$. Elles sont notées $\frac{\partial f}{\partial x}(x_0, y_0)$ et $\frac{\partial f}{\partial y}(x_0, y_0)$. Ces valeurs correspondent aux dérivées (en x_0 et en y_0 respectivement) des fonctions partielles f_{x, y_0} et $f_{x_0, y}$.

Remarque 11. En pratique, pour calculer ces dérivées partielles (ce qu'on fera plutôt sur un ouvert que simplement en un point, comme dans le cas des dérivées de fonctions d'une seule variable), on dérive en considérant l'une des deux variables comme une constante (on dit qu'on dérive la fonction f par rapport à x ou y respectivement), mais chacune des deux dérivées partielles reste une fonction à deux variables.

Exemples : Reprenons l'exemple de la fonction $f : (x, y) \mapsto x^3 + 2x^2y + xy^3 - 4y^2$. On calcule $\frac{\partial f}{\partial x}(x, y) = 3x^2 + 4xy + y^3$ et $\frac{\partial f}{\partial y}(x, y) = 2x^2 + 3xy^2 - 8y$.

De même, la fonction $g : (x, y) \mapsto \ln(x + y - 1)$ a pour dérivées partielles $\frac{\partial g}{\partial x}(x, y) = \frac{1}{x + y - 1}$ et $\frac{\partial g}{\partial y}(x, y) = \frac{1}{x + y - 1}$.

Enfin, la fonction $h : (x, y) \mapsto \sqrt{4 - x^2 - y^2}$ vérifie $\frac{\partial h}{\partial x}(x, y) = \frac{-2x}{2\sqrt{4 - x^2 - y^2}} = -\frac{x}{\sqrt{4 - x^2 - y^2}}$ et $\frac{\partial h}{\partial y}(x, y) = -\frac{y}{\sqrt{4 - x^2 - y^2}}$.

Théorème 2. Théorèmes généraux sur la dérivabilité de fonctions à deux variables.

Toute somme, produit, quotient (si le dénominateur ne s'annule pas) de fonctions admettant des dérivées partielles en a admet des dérivées partielles en a . Les formules habituelles concernant ces opérations restent valables : $\frac{\partial(f + g)}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x}$, $\frac{\partial(\lambda f)}{\partial x} = \lambda \frac{\partial f}{\partial x}$, $\frac{\partial(fg)}{\partial x} = f \times \frac{\partial g}{\partial x} + g \times \frac{\partial f}{\partial x}$ et $\frac{\partial(\frac{1}{f})}{\partial x} = -\frac{1}{f^2} \times \frac{\partial f}{\partial x}$. Bien sûr, on a des formules identiques pour les dérivées partielles par rapport à y .

Démonstration. C'est évident, puisque les dérivées partielles ne sont en fait rien d'autre que des dérivées des fonctions partielles qui sont des fonctions à une variable. Tout le formulaire en découle immédiatement. On pourrait d'ailleurs exprimer de même les dérivées (partielles) de la composée d'une fonction à deux variables par une fonction à une variable. $\square$

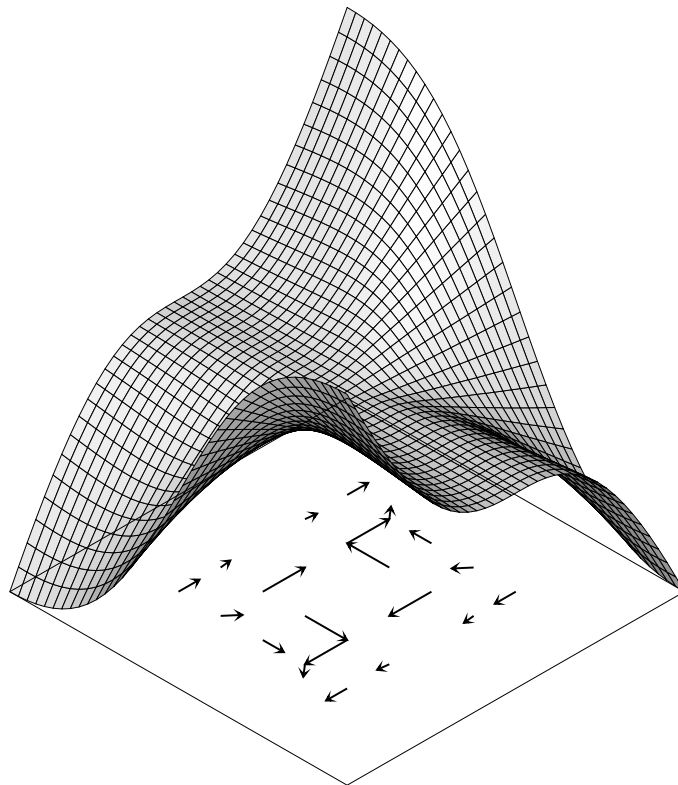
Remarque 12. Attention, l'existence de dérivées partielles ne garantit absolument pas la continuité de la fonction au point étudié, puisqu'on se contente de regarder suivant deux directions particulières. Ainsi, la fonction $f : (x, y) \mapsto \frac{2xy}{x^2 + y^2}$, prolongée en $(0, 0)$ par $f(0, 0) = 0$ n'est pas continue en $(0, 0)$ (cf exemple déjà donné un peu plus haut) mais elle admet des dérivées partielles en ce même point (qui sont nulles puisque les deux applications partielles en $(0, 0)$ sont nulles). Graphiquement, il existe bien des vecteurs tangents à la surface en $(0, 0)$ dans la direction des deux axes.

Définition 13. Soit f une fonction admettant des dérivées partielles en $a = (x_0, y_0)$, on appelle **gradient de f en a** le vecteur $\nabla f(a) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right)$.

Remarque 13. Pour les plus curieux, le drôle de symbole utilisé pour noter le gradient est appelé nabla. On a aussi le droit de noter plus classiquement ce vecteur $\overrightarrow{\text{grad}} f(a)$, mais c'est évidemment nettement moins classe.

Proposition 2. Le gradient de f en a est dirigé suivant la plus grande pente positive de la surface représentative de f au voisinage de a . De plus, le vecteur gradient est orthogonal à la ligne de niveau contenant le point $(a, f(a))$.

Démonstration. Ces résultats découlent des théorèmes que nous allons énoncer dans le paragraphe suivant (règle de la chaîne notamment). Contentons-nous pour l'instant d'un exemple visuel, où on a tracé la surface représentative de la fonction $f : (x, y) \mapsto \frac{x^3 y}{x^2 + y^2}$, ainsi que les gradients en certains points (les gradients sont les vecteurs représentés arbitrairement dans le plan $z = -2$, donc « en-dessous » de la surface représentative de f , le vecteur gradient donne la direction de la pente la plus forte au point de la surface situé juste au-dessus de lui. Par ailleurs, les normes des gradients ont été divisées par 4 pour que ce soit vaguement plus lisible, les pentes étant assez fortes sur cette surface).



□

Remarque 14. On dispose des mêmes formules pour le gradient que pour les dérivées partielles dont il est issu (gradient d'une somme, d'un produit, d'un quotient notamment).

2.4 Fonctions à deux variables de classe $\mathcal{C}^1$.

Définition 14. La fonction f est de classe $\mathcal{C}^1$ sur D si elle admet en tout point de D des dérivées partielles, et si ces dérivées partielles sont continues sur D .

Exemples : Les fonctions polynômiales sont de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. Plus généralement, toute fonction obtenue par opérations élémentaires à partir de polynômes et de fonctions à une variable usuelles connues pour être $\mathcal{C}^1$ sera de classe $\mathcal{C}^1$ sur son ensemble de définition. Par exemple, $h : (x, y) \mapsto \sqrt{4 - x^2 - y^2}$ sera certainement de classe $\mathcal{C}^1$ sur le disque **ouvert** $B(O, 2)$, mais n'admet pas de dérivée partielle à la frontière de ce disque (donc sur le cercle de centre O et de rayon 2).

Exemple : La fonction norme $u(x, y) \mapsto \|u\| = \sqrt{x^2 + y^2}$ est une fonction définie et continue sur $\mathbb{R}^2$, qui est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2 \setminus \{(0, 0)\}$, mais n'admet pas de dérivées partielles en $(0, 0)$. C'est bien sûr cohérent avec le fait que ses applications partielles en $(0, 0)$ coïncident avec la fonction valeur absolue, qui n'est pas dérivable en 0.

Définition 15. Soit f une fonction définie en $a(x_0, y_0)$, f est **différentiable** en a s'il existe une forme linéaire $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ telle que $f(x_0 + h, y_0 + k) \underset{(h,k) \rightarrow (0,0)}{=} f(x_0, y_0) + \varphi(h, k) + o(\|(h, k)\|)$. On dit également que f admet alors un **développement limité d'ordre 1 en a** , et l'application φ est appelée **différentielle de f au point a** et notée df_a .

Non, ne fuyez pas, ce n'est en fait pas si compliqué que ça à comprendre, puisqu'il s'agit exactement de la même chose que la dérivabilité d'une fonction à une variable, mais exprimée avec des applications linéaires pour rendre les choses plus faciles à manier. Prenons donc pour commencer l'exemple d'une fonction à une seule variable : $f(x) = 1 - 2x - x^2$. Cette fonction est dérivable en 0 et vérifie $f'(0) = 2$. On peut exprimer ceci à l'aide d'un développement limité à l'ordre 1 : $f(x) \underset{x \rightarrow 0}{=} 1 - 2x + o(x)$. C'est exactement le principe de la définition ci-dessus : on a négligé tous les termes « plus petits que x » pour ne conserver qu'une approximation d'ordre 1 de la fonction f au voisinage de 0 faisant intervenir l'application $\varphi : x \mapsto -2x$ qui est une forme linéaire sur $\mathbb{R}$. De façon formelle, on peut dire que f est dérivable en 0 s'il existe une telle forme linéaire φ telle que $f(x) \underset{x \rightarrow 0}{=} f(0) + \varphi(x) + o(x)$. Bien sûr, comme les formes linéaires de $\mathbb{R}$ dans $\mathbb{R}$ sont toutes de la forme $x \mapsto ax$, la seule information pertinente est la valeur du coefficient a , qui représente ce qu'on appelle habituellement nombre dérivé de f en 0. Mais on pourrait utiliser le même vocabulaire que pour les fonctions à deux variables et parler de différentielle de f en 0 pour désigner l'application φ .

Passons maintenant au cas de la dimension 2. Notre différentielle est maintenant une application linéaire à deux variables, mais ce n'est pas pour autant beaucoup plus compliqué : on sait que $\mathcal{L}(\mathbb{R}^2, \mathbb{R})$ est un espace vectoriel de dimension 2, et on peut même en donner une base facilement ! Si on considère les deux formes linéaires $\varphi : (x, y) \mapsto x$ et $\psi : (x, y) \mapsto y$, toute forme linéaire sur $\mathbb{R}^2$ peut se décomposer comme une combinaison linéaire de φ et de ψ . Mieux, φ et ψ sont les différentielles en tout points de deux fonctions à deux variables particulièrement simples : les applications coordonnées $(x, y) \mapsto x$ et $(x, y) \mapsto y$ (qui coïncident avec leur différentielle puisqu'il s'agit déjà d'applications linéaires). Ces deux formes linéaires sont tellement essentielles qu'on leur donne un nom particulier : on les note dx et dy (oui, oui, c'est bien cette même notation qui est cachée dans les calculs d'intégrales). Si une fonction à deux variables f est différentiable en un point a , on devrait donc pouvoir écrire $df_a = \lambda dx + \mu dy$. Que représentent alors les constantes λ et μ ? Tout simplement une pente de tangente dans les directions données par les deux vecteurs unitaires du repère canonique, c'est-à-dire exactement la définition des dérivées partielles. D'où le théorème fondamental suivant :

Théorème 3. Si f est une fonction de classe $\mathcal{C}^1$ sur D , alors f est différentiable en tout point $a \in D$, et $df_a = \frac{\partial f}{\partial x}(a)dx + \frac{\partial f}{\partial y}(a)dy$.

Démonstration. La démonstration peut se faire à l'aide des outils d'analyse « à une variable » développés cette année, mais elle n'est pas au programme, nous nous en dispenserons donc. □

Remarque 15. Il existe d'autres façons d'écrire ce résultat qu'il est bon d'avoir également en tête : si f est de classe $\mathcal{C}^1$ en $a(x_0, y_0)$, alors

- $f(x_0 + h, y_0 + k) \underset{(h,k) \rightarrow (0,0)}{=} f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k + o(\|(h, k)\|)$
- $f(a + u) \underset{u \rightarrow (0,0)}{=} f(a) + \langle \nabla f(a), u \rangle + o(\|u\|)$

Exemple : Soit $f : (x, y) \mapsto x^3y - 2x^2y^2$. La fonction f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ en tant que fonction polynômiale. On calcule ici sans difficulté $\frac{\partial f}{\partial x}(x, y) = 3x^2y - 4xy^2$ et $\frac{\partial f}{\partial y}(x, y) = x^3 - 4x^2y$. Sans surprise, la différentielle de f en $(0, 0)$ est nulle puisque tous les termes qui composent f sont de degré 4 et donc largement négligeables par rapport aux valeurs de x et de y . Par contre, on peut constater que $\frac{\partial f}{\partial x}(1, 1) = -1$ et $\frac{\partial f}{\partial y}(1, 1) = -3$, donc que $df_{(1,1)} = -dx - 3dy$. Cela signifie que, dans un voisinage proche de $(1, 1)$, on aura l'approximation du premier ordre suivante : $f(x, y) \simeq -1 - (x - 1) - 3(y - 1)$.

Proposition 3. Si f est différentiable en a , alors f est continue en a .

Démonstration. C'est évident à partir de la formule du développement limité à l'ordre 1 en a . Ce sera en particulier le cas de toutes les fonctions de classe $\mathcal{C}^1$, mais la réciproque n'est pas vraie. Il est d'ailleurs maintenant temps de vous avouer que la notion de différentiabilité n'est pas officiellement à votre programme de première année. $\square$

Définition 16. Soit f une fonction de classe $\mathcal{C}^1$ en $a = (x_0, y_0)$, le plan d'équation $z = f(a) + \frac{\partial f}{\partial x}(a)(x - x_0) + \frac{\partial f}{\partial y}(a)(y - y_0)$ est appelé **plan tangent** à la surface représentative de la fonction f en a .

Remarque 16. Il s'agit bien évidemment de l'équivalent de la droite tangente à une courbe de fonction à une variable dérivable en un point. Le plan tangent en a est tout simplement le plan passant par le point de la surface $A(x_0, y_0, f(a))$, et admettant comme base de vecteurs directeurs $\left(\left(1, 0, \frac{\partial f}{\partial x}(a) \right), \left(0, 1, \frac{\partial f}{\partial y}(a) \right) \right)$.

Exemple : Soit $f : (x, y) \mapsto \sin(x + 2y)$, alors f est une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$, et $\frac{\partial f}{\partial x}(x, y) = \cos(x + 2y)$, $\frac{\partial f}{\partial y}(x, y) = 2\cos(x + 2y)$. On en déduit que l'équation du plan tangent à la surface représentative de f en son point $A(0, 0, 0)$ a pour équation $z = x + 2y$.

Théorème 4. Règle de la chaîne.

Soient x et y deux fonctions à une variable de classe $\mathcal{C}^1$ sur un intervalle I , et f une fonction à deux variables de classe $\mathcal{C}^1$ sur un domaine D tel que $\forall t \in I, (x(t), y(t)) \in D$, alors la fonction $t \mapsto F(t) = f(x(t), y(t))$ est une fonction de classe $\mathcal{C}^1$ sur I , et $\forall t \in I$, $F'(t) = x'(t)\frac{\partial f}{\partial x}(x(t), y(t)) + y'(t)\frac{\partial f}{\partial y}(x(t), y(t))$.

Démonstration. En notant $u : t \mapsto (x(t), y(t))$, on aura donc $F = f \circ u$, avec $u : \mathbb{R} \rightarrow \mathbb{R}^2$ et $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. On calcule alors, au voisinage d'un certain réel a , $F(a+h) = f(u(a+h)) = f(u(a) + (u(a+h) - u(a)))$. En posant $k = u(a+h) - u(a)$, le caractère $\mathcal{C}^1$ de l'application u assure que $\lim_{h \rightarrow 0} k = (0, 0)$. On peut donc écrire le développement limité à l'ordre 1 de la fonction à deux variables f

au voisinage de $u(a)$: $F(a+h) = f(u(a)) + \frac{\partial f}{\partial x}(u(a))(x(a+h) - x(a)) + \frac{\partial f}{\partial y}(u(a))(y(a+h) - y(a)) + o(h)$ (en effet, le o devrait être un $o(\|u(a+h) - u(a)\|) = o(\sqrt{(x(a+h) - x(a))^2 + (y(a+h) - y(a))^2}) = o(\sqrt{h^2 x'(a)^2 + h^2 y'(a)^2 + o(h^2)}) = o(h)$). Il ne reste plus qu'à redévelopper $x(a+h) - x(a)$ sous la forme $hx'(a) + o(h)$, et de même pour le deuxième terme, et à regrouper tous les $o(h)$ pour obtenir finalement $F(a+h) = F(a) + h \left(\frac{\partial f}{\partial x}(u(a))x'(a) + \frac{\partial f}{\partial y}(u(a))y'(a) \right) + o(h)$, ce qui prouve exactement que F est dérivable en a (elle y admet un développement limité à l'ordre 1), et donne la formule souhaitée pour $F'(a)$. $\square$

Exemple : soit $f : (x, y) \mapsto x^2 e^y$ et $F : t \mapsto f(t, \ln(t))$. La fonction F est de classe $\mathcal{C}^1$ sur $]0, +\infty[$, puisque les deux fonctions $t \mapsto t$ et $t \mapsto \ln(t)$ le sont, et que f est elle-même $\mathcal{C}^1$ sur $\mathbb{R}$. De plus, en notant $x(t) = t$ et $y(t) = \ln(t)$ (pour garder les notations de l'énoncé précédent), alors $x'(t) = 1$ et $y'(t) = \frac{1}{t}$. Enfin, $\frac{\partial f}{\partial x}(x, y) = 2xe^y$, donc $\frac{\partial f}{\partial x}(t, \ln(t)) = 2t^2$, et $\frac{\partial f}{\partial y}(x, y) = x^2 e^y$, donc $\frac{\partial f}{\partial y}(t, \ln(t)) = t^3$. La règle de la chaîne permet alors d'affirmer que la dérivée de la fonction F est donnée par $F'(t) = 1 \times 2t^2 + \frac{1}{t} \times t^3 = 3t^2$. En effet, cela n'a rien de surprenant puisque $F(t) = f(t, \ln(t)) = t^2 e^{\ln(t)} = t^3$.

Corollaire 1. Une fonction f de classe $\mathcal{C}^1$ sur $D \subset \mathbb{R}^2$ y admet une dérivée directionnelle suivant tous les vecteurs en tout point de D . De plus, si $a \in D$ et $u \in \mathbb{R}^2$ unitaire, alors la dérivée de f en a suivant $u = (\alpha, \beta)$ est égale à $\frac{\partial f}{\partial x}(a) \times \alpha + \frac{\partial f}{\partial y}(a) \times \beta$.

Démonstration. Il suffit d'appliquer la règle de la chaîne à la fonction $F : t \mapsto f(a + tu) = f(x_0 + t\alpha, y_0 + t\beta)$ pour obtenir sa dérivée en 0, qui est par définition égale à la dérivée directionnelle de f en a suivant u . $\square$

Remarque 17. C'est également la règle de la chaîne qui permet de démontrer que les vecteurs gradients sont toujours orthogonaux aux lignes de niveau : si on admet que la ligne de niveau k peut être parcourue comme une certaine courbe paramétrée, $f(a) = k \Leftrightarrow a = (x(t), y(t))$, avec x et y deux fonctions de classe $\mathcal{C}^1$ sur un certain intervalle $I \subset \mathbb{R}$ (cette hypothèse n'a d'ailleurs rien d'évident, mais c'est le cas des quelques lignes de niveau dont on a pu calculer des équations jusqu'ici). On peut alors dire que $f(x(t), y(t)) = k$ sur tout l'intervalle I , donc, en notant comme précédemment $F(t) = f(x(t), y(t))$, la fonction F a une dérivée nulle. La règle de la chaîne affirme alors que $\frac{\partial f}{\partial x}(x(t), y(t)) \times x'(t) + \frac{\partial f}{\partial y}(x(t), y(t)) \times y'(t) = 0$. Autrement dit, le vecteur gradient au point $a = (x(t), y(t))$ est orthogonal au vecteur $(x'(t), y'(t))$, qui est un vecteur tangent à la ligne de niveau parcourue par les points de coordonnées $(x(t), y(t))$. C'est exactement cela que signifie « l'orthogonalité du gradient par rapport aux lignes de niveau ».

Théorème 5. Règle de la chaîne, deuxième version.

Soient u et v deux fonctions de classe $\mathcal{C}^1$ sur $U \subset \mathbb{R}^2$, telles que $\forall (x, y) \in U$, $(u(x, y), v(x, y)) \in V$, et f une fonction de classe $\mathcal{C}^1$ sur V , alors la fonction F définie sur U par $F(x, y) = f(u(x, y), v(x, y))$ est de classe $\mathcal{C}^1$ sur U , et

$$\begin{aligned} \bullet \quad \frac{\partial F}{\partial x} &= \frac{\partial f}{\partial u} \times \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \times \frac{\partial v}{\partial x} \\ \bullet \quad \frac{\partial F}{\partial y} &= \frac{\partial f}{\partial u} \times \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \times \frac{\partial v}{\partial y} \end{aligned}$$

Exemple : pour illustrer cette nouvelle version du théorème, on va prendre un cas très classique de changement de variable, le passage en coordonnées polaires. On souhaiterait exprimer les dérivées partielles d'une fonction à deux variables par rapport aux coordonnées polaires r et θ en fonction des dérivées partielles « classiques » par rapport à x et y . Le problème étant qu'il est plus facile de faire le calcul dans l'autre sens, en posant $x = r \cos(\theta)$ et $y = r \sin(\theta)$. Soit donc une fonction f à deux variables telle que $f(x, y) = g(r, \theta)$. On peut appliquer les formules du théorème précédent pour obtenir $\frac{\partial g}{\partial r}(r, \theta) = \frac{\partial f}{\partial x} \times \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \times \frac{\partial y}{\partial r} = \cos(\theta) \frac{\partial f}{\partial x}(r \cos(\theta), r \sin(\theta)) + \sin(\theta) \frac{\partial f}{\partial y}(r \cos(\theta), r \sin(\theta))$, et $\frac{\partial g}{\partial \theta}(r, \theta) = \frac{\partial f}{\partial x} \times \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \times \frac{\partial y}{\partial \theta} = -r \sin(\theta) \frac{\partial f}{\partial x}(r \cos(\theta), r \sin(\theta)) + r \cos(\theta) \frac{\partial f}{\partial y}(r \cos(\theta), r \sin(\theta))$.

Pour obtenir les formules dans l'autre sens, il faut en fait résoudre un système de deux équations à deux inconnues, ce qu'on peut ici faire très rapidement à l'aide de combinaisons linéaires astucieuses : en multipliant la première équation par $r \sin(\theta)$ et la deuxième par $\cos(\theta)$, leur somme donne $r \sin(\theta) \frac{\partial g}{\partial r} + \cos(\theta) \frac{\partial g}{\partial \theta} = \frac{\partial f}{\partial y}(r \sin^2(\theta) + r \cos^2(\theta))$, donc $\frac{\partial f}{\partial y} = \sin(\theta) \frac{\partial g}{\partial r} + \frac{\cos(\theta)}{r} \frac{\partial g}{\partial \theta}$. De même, en les multipliant par $r \cos(\theta)$ et par $-\sin(\theta)$, on aura $\frac{\partial f}{\partial x} = \cos(\theta) \frac{\partial g}{\partial r} - \frac{\sin(\theta)}{r} \frac{\partial g}{\partial \theta}$.

Ce genre de calcul a des applications classiques dans la résolution des équations aux dérivées partielles, qui ne sont rien d'autre que des équations différentielles dont les inconnues sont des fonctions à deux variables. Imaginons par exemple qu'on veuille résoudre sur $\mathbb{R}^2 \setminus \{(0, 0)\}$ l'équation $y \frac{\partial f}{\partial x} = x \frac{\partial f}{\partial y}$. En posant $f(x, y) = g(r, \theta)$ comme précédemment, on peut transformer notre équation en équation équivalente en coordonnées polaires : $r \sin(\theta) \cos(\theta) \frac{\partial g}{\partial r} - \sin^2(\theta) \frac{\partial g}{\partial \theta} = r \cos(\theta) \sin(\theta) \frac{\partial g}{\partial r} + \cos^2(\theta) \frac{\partial g}{\partial \theta}$, ce qui se simplifie en $\frac{\partial g}{\partial \theta} = 0$. Autrement dit, la fonction à deux variables g est constante à r fixé, ce qui revient à dire que $g(r, \theta) = \varphi(r)$, où φ est une fonction de classe $\mathcal{C}^1$ sur $]0, +\infty[$. Autrement dit, il existe une fonction de classe $\mathcal{C}^1$ sur $]0, +\infty[$ telle que $f(x, y) = \varphi(\sqrt{x^2 + y^2})$. Il faudrait faire un calcul pour vérifier que la réciproque est vraie, ce qui est bien le cas.

2.5 Extrema de fonctions à deux variables.

Définition 17. Un **point critique** pour une fonction f à deux variables est un point $a = (x_0, y_0)$ vérifiant $\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0$.

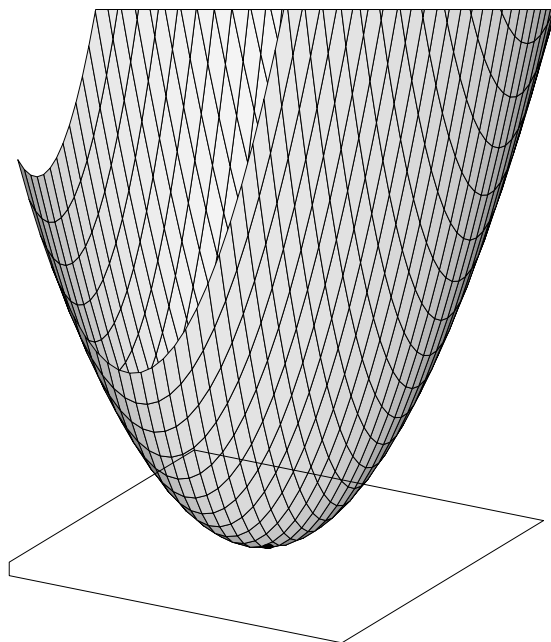
Théorème 6. Si une fonction à deux variables f définie sur un ouvert de $\mathbb{R}^2$ y admet un minimum ou un maximum local atteint en un point a , alors ce point est un point critique de f .

Remarque 18. Attention, comme dans le cas des fonctions à une variable, la réciproque n'a aucune raison d'être vraie. Elle le sera même « encore moins souvent », puisqu'un point correspondant par exemple à un maximum de la première fonction partielle, mais à un minimum de la deuxième fonction partielle, ne pourra jamais être un extremum local de la fonction f .

Démonstration. Si a correspond par exemple à un minimum local de f alors il correspond aussi à un minimum local des deux fonctions partielles associées au point a , ce qui assure l'annulation des deux dérivées partielles correspondantes. $\square$

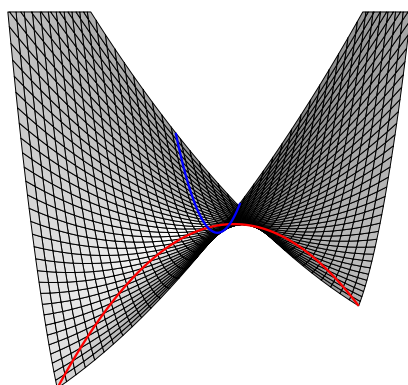
Exemple : soit $f : (x, y) \mapsto x^2 - 3x + xy + y^2$. La fonction f est polynômiale, donc de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. De plus, $\frac{\partial f}{\partial x}(x, y) = 2x - 3 + y$ et $\frac{\partial f}{\partial y}(x, y) = x + 2y$. Les points critiques pour la fonction f sont donc obtenus en résolvant le système $\begin{cases} 2x + y = 3 \\ x + 2y = 0 \end{cases}$. On obtient pratiquement en additionnant les deux équations et en les soustrayant $x + y = 1$ et $x - y = 3$, donc on déduit facilement $x = 2$ et $y = -1$, donc le seul point critique de f est $a = (2, -1)$.

On calcule alors $f(2, -1) = 4 - 6 - 2 + 1 = -3$, qui est potentiellement un extremum local de f . Pour vérifier que c'est bien le cas, un seul moyen à notre disposition : calculer $f(2+h, -1+k) - f(2, -1) = (2+h)^2 - 3(2+h) + (2+h)(k-1) + (k-1)^2 + 3 = 4 + 4h + h^2 - 6 - 3h + 2k - 2 + hk - h + k^2 - 2k + 1 + 3 = h^2 + k^2 + hk = \left(h + \frac{k}{2}\right)^2 + \frac{3}{4}k^2 \geq 0$, ce qui prouve que $f(2+h, -1+k) \geq f(2, -1)$ quelles que soient les valeurs de h et k . Non seulement a correspond à un minimum local pour f , mais il s'agit même d'un minimum global. C'est bien sûr du au fait que la surface représentative de f est très particulière, il s'agit d'un parabolôïde représenté ci-dessous, avec le plan tangent horizontal d'équation $z = -3$ obtenu au point a :



Exemple 2 : soit $f(x, y) = x^2 + y^2 + 4xy - 2$, la fonction f est à nouveau polynômiale donc de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. On calcule donc $\frac{\partial f}{\partial x}(x, y) = 2x + 4y$ et $\frac{\partial f}{\partial y}(x, y) = 2y + 4x$. L'annulation des deux dérivées partielles impose les deux conditions $y = -2x$ et $x = -2y$, qui ne peuvent se produire que si $x = y = 0$. L'origine est donc le seul point critique de la fonction f . De plus, $f(0, 0) = -2$,

mais $f(x, 0) = x^2 - 2 \geq -2$, et $f(x, -x) = 2x^2 - 4x^2 - 2 = -2 - 2x^2 \leq 0$ (et strictement inférieur pour des points situés aussi proches de $(0, 0)$ qu'on le veut), ce qui prouve que la fonction ne peut pas admettre d'extremum local (ni a fortiori global) en $(0, 0)$. La fonction n'a donc pas du tout d'extremum local. Le point $(0, 0)$ correspond en fait à ce qu'on appelle parfois un « point col » ou « point selle » pour la surface représentative de f : le point correspond à un minimum si on coupe la surface suivant certains plans verticaux, mais à un maximum suivant d'autres. La figure ci-dessous permet de visualiser ce phénomène, on y a aussi représenté les coupes à $y = 0$ (en bleu) et suivant $y = -x$ (en rouge) utilisées pour le calcul :



3 Outils supplémentaires de calcul intégral.

3.1 Intégrales multiples.

Nous avons appris, il y a quelques chapitres de ça, à calculer des intégrales d'une fonction à une variable, en les définissant à l'aide de la convergence d'intégrales de fonctions en escalier. Nous avons même vu plus précisément que cette technique de calcul, appelée intégrale de Riemann, donnait une valeur à toute intégrale de fonction continue par morceaux sur un segment, et que cette intégrale était liée à la notion de calcul de primitive. Peut-on adapter ces calculs pour produire une théorie de l'intégration permettant de calculer des intégrales de fonctions à deux variables, c'est-à-dire calculant le volume situé entre le plan de coordonnées (xOy) et la surface représentative d'une telle fonction ? Oui, mais des difficultés techniques vont voir le jour, notamment concernant le domaine d'intégration : dans le cas d'une fonction à une variable, le calcul d'intégrale se fait sur un segment (ou, vous le verrez l'an prochain, sur un intervalle éventuellement non borné). Mais dans le plan, on aimerait pouvoir calculer des intégrales sur des domaines plus variés que de simples produits d'intervalles de $\mathbb{R}$ (qui donneraient des rectangles). C'est possible, mais on va tout de même commencer par traiter le cas plus simple de rectangles. Dans toute cette dernière partie de cours, qui est complètement hors-programme, on donnera souvent des définitions souples et on démontrera peu de théorèmes (voire même pas du tout), le but étant simplement de donner une idée du fonctionnement des concepts évoqués. Dans la définition qui suit, par exemple, on se gardera bien d'expliquer ce que signifie « continue par morceaux » pour une fonction à deux variables.

Définition 18. Soit f une fonction à deux variables continue par morceaux sur l'ensemble $D = [a, b] \times [c, d]$. Les sommes de Riemann associées à la fonction f sont les sommes $R_{n,p} = \sum_{i=1}^n \sum_{j=1}^p f(x_i, y_j) \times \frac{(b-a)(d-c)}{np}$, où $x_i = a + i \times \frac{b-a}{n}$, et $y_j = c + j \times \frac{d-c}{p}$.

Théorème 7. Les sommes de Riemann associées à la fonction f convergent quand n et p tendent tous les deux vers $+\infty$ vers une valeur appelée intégrale double de f sur le rectangle $[a, b] \times [c, d]$, et notée $\iint_{[a,b] \times [c,d]} f(x, y) \, dx \, dy$.

Remarque 19. Les intégrales doubles vérifient les propriétés habituelles de l'intégration : linéarité, positivité, relation de Chasles.

Théorème 8. Théorème de Fubini.

$$\iint_{[a,b] \times [c,d]} f(x, y) \, dx \, dy = \int_c^d \left(\int_a^b f(x, y) \, dx \right) dy = \int_a^b \left(\int_c^d f(x, y) \, dy \right) dx.$$

Exemple : en pratique, le théorème de Fubini permet de calculer des intégrales doubles à l'aide de deux intégrations usuelles successives, en faisant toutefois attention au fait qu'on doit intégrer « suivant la bonne variable ».

Ainsi, $\int_0^1 \int_{-1}^1 (x^2 y - 1) \, dx \, dy = \int_0^1 \left[\frac{x^3 y}{3} - x \right]_{-1}^1 dy = \int_0^1 \frac{2y}{3} - 2 \, dy = \left[\frac{y^2}{3} - 2y \right]_0^1 = -\frac{5}{3}$. On peut bien sûr aussi calculer l'intégrale dans l'autre sens : $\int_{-1}^1 \int_0^1 (x^2 y - 1) \, dy \, dx = \int_{-1}^1 \left[\frac{x^2 y^2}{2} - y \right]_0^1 dx = \int_{-1}^1 \frac{x^2}{2} - 1 \, dx = \left[\frac{x^3}{6} - x \right]_{-1}^1 = -\frac{5}{3}$.

Définition 19. Si f est une fonction à deux variables continue par morceaux sur un domaine D inclus dans un rectangle $[a, b] \times [c, d]$, l'intégrale de f sur D est l'intégrale de la fonction g définie sur $[a, b] \times [c, d]$ par $g(x) = f(x)$ si $x \in D$, et $g(x) = 0$ sinon. On note cette intégrale $\iint_D f(x, y) \, dx \, dy$.

Cette définition est bien entendu une arnaque totale, on complète simplement la fonction f par la fonction nulle pour ne pas changer la valeur de l'intégrale.

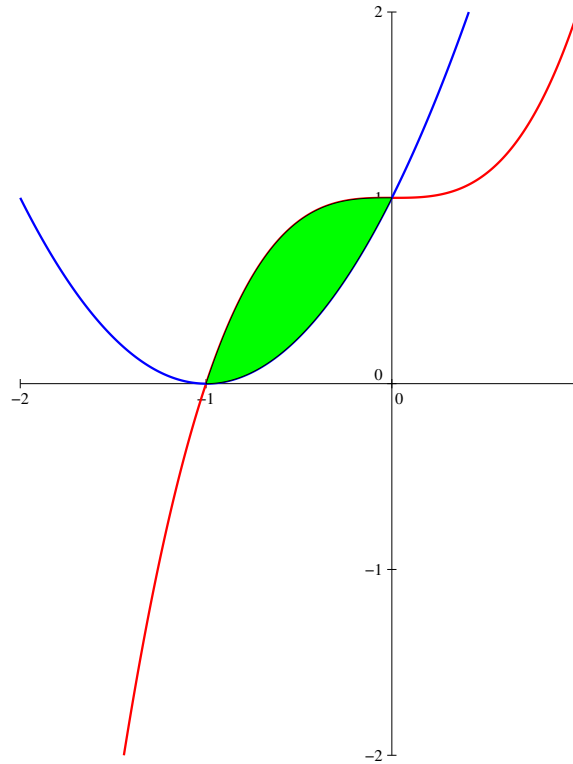
Théorème 9. Théorème de Fubini, deuxième version.

Si le domaine d'intégration D peut être décrit sous la forme $D = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b \text{ et } c(x) \leq y \leq d(x)\}$, où c et d sont des fonctions continues sur le segment $[a, b]$, alors

$$\iint_D f(x, y) \, dx \, dy = \int_a^b \int_{c(x)}^{d(x)} f(x, y) \, dy \, dx.$$

Remarque 20. On a bien sûr une formule symétrique si on décide de décrire le domaine D en faisant varier y entre c et d , puis x entre $a(y)$ et $b(y)$.

Exemple 1 : On souhaite calculer l'aire de la surface du plan comprise entre les deux courbes d'équation $f(x) = (x + 1)^2$ et $g(x) = x^3 + 1$ (cf ci-dessous). Il est facile de constater que ces deux courbes se coupent en exactement deux points, de coordonnées $(-1, 0)$ et $(0, 1)$. L'aire recherchée est alors simplement l'intégrale de la fonction constante égale à 1 sur le domaine $D = \{(x, y) \in \mathbb{R}^2 \mid x \in [-1, 0], (x + 1)^2 \leq y \leq x^3 + 1\}$.



Cette intégrale se calcule facilement : $\int_{-1}^0 \int_{x^2+2x+1}^{x^3+1} 1 \, dy \, dx = \int_{-1}^0 x^3 - x^2 - 2x \, dx = \left[\frac{x^4}{4} - \frac{x^3}{3} - x^2 \right]_{-1}^0 = -\frac{1}{4} - \frac{1}{3} + 1 = \frac{5}{12}$.

Exemple 2 : essayons d'appliquer cette méthode pour calculer le volume V de la boule unité (centré en O , rayon 1) dans l'espace. On calculera plutôt en pratique le volume d'une demi-boule (qu'on multipliera donc brillamment par 2 pour obtenir la valeur de V), qui correspond à l'intégrale double de la fonction $f : (x, y) \mapsto \sqrt{1 - x^2 - y^2}$ calculée sur un domaine D qui n'est autre que le cercle trigonométrique, qu'on peut lui même décrire par les inégalités $-1 \leq x \leq 1$ et $-\sqrt{1 - x^2} \leq y \leq \sqrt{1 - x^2}$. Autrement dit, $V = 2 \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1 - x^2 - y^2} \, dy \, dx = 2 \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1 - x^2} \sqrt{1 - \frac{y^2}{1 - x^2}} \, dy \, dx$.

On effectue alors le changement de variable $\frac{y}{\sqrt{1 - x^2}} = \sin(t)$ (on peut, la valeur étant toujours comprise entre -1 et 1 par construction). Les bornes $y = -\sqrt{1 - x^2}$ et $y = \sqrt{1 - x^2}$ deviennent alors $\sin(t) = \pm 1$, soit $t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, et $dy = \sqrt{1 - x^2} \cos(t) \, dt$. De plus, avec le changement de variable effectué, $\sqrt{1 - \frac{y^2}{1 - x^2}} = \sqrt{1 - \sin^2(t)} = \cos(t)$ (formule toujours valable sur l'intervalle dans lequel varie t). On en déduit que $V = 2 \int_{-1}^1 (1 - x^2) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(t) \, dt = 2 \int_{-1}^1 (1 - x^2) \left[\frac{t}{2} + \frac{\sin(2t)}{4} \right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} dx =$

$\pi \int_{-1}^1 (1 - x^2) \, dx = \pi \left[x - \frac{x^3}{3} \right]_{-1}^1 = \frac{4\pi}{3}$.

Théorème 10. Changement de variables : passage en coordonnées polaires.

Si $f(x, y) = g(r, \theta)$, alors $\iint_D f(x, y) \, dx \, dy = \iint_D r g(r, \theta) \, dr \, d\theta$.

Bien sûr, pour que cette formule soit exacte, il faut que le domaine D décrit par les deux intégrales soit le même. Certains domaines sont plus faciles à exprimer en coordonnées polaires qu'en coordonnées cartésiennes, par exemple le cercle du calcul de volume de boule effectué plus haut, qui se décrit simplement avec les inégalités $0 \leq r \leq 1$ et $0 \leq \theta \leq 2\pi$. Comme $\sqrt{1-x^2-y^2} = \sqrt{1-r^2}$, on peut donc recalculer plus facilement ce volume en posant $V = \int_0^{2\pi} \int_0^1 2r\sqrt{1-r^2} dr d\theta = 2\pi \times \left[-\frac{2}{3}(1-r^2)^{\frac{3}{2}} \right]_0^1 = \frac{4\pi}{3}$.

Exemple : pour pouvoir donner un exemple extrêmement classique, on va tricher un peu en admettant que les résultats énoncés ci-dessus restent valables pour des intégrales impropres, à borne infinie. On souhaite calculer l'intégrale de Gauss $I = \int_0^{+\infty} e^{-t^2} dt$. Comme $e^{-x^2-y^2} = e^{-x^2} \times e^{-y^2}$, le théorème de Fubini permet d'écrire que $\int_0^{+\infty} \int_0^{+\infty} e^{-x^2-y^2} dx dy = \int_0^{+\infty} e^{-x^2} dx \times \int_0^{+\infty} e^{-y^2} dy = I^2$. Or, l'intégrale double précédente se calcule fort bien avec un passage en coordonnées polaires : le domaine $[0, +\infty[^2$ est décrit en coordonnées par $r \in [0, +\infty[$ et $\theta \in [0, \frac{\pi}{2}]$, donc $I^2 = \int_0^{\frac{\pi}{2}} \int_0^{+\infty} r e^{-r^2} dr d\theta = \frac{\pi}{2} \times \left[-\frac{e^{-r^2}}{2} \right]_0^{+\infty} = \frac{\pi}{4}$. On en déduit immédiatement que $I = \frac{\sqrt{\pi}}{2}$.

Bien sûr, vous allez me demander pourquoi ce facteur r apparaît lorsqu'on effectue le passage en polaires. Eh bien, il s'agit d'une sombre histoire de Jacobien que je ne peux pas vraiment vous expliquer en détail sans déborder trop violemment du programme officiel. Vous en apprendrez plus à ce sujet l'an prochain. Sachez en tout cas que toutes les techniques vues ici peuvent se généraliser à des calculs d'intégrales triples de fonctions dépendant de trois variables (les changements de variables sont alors encore plus compliqués). Un simple exemple de tel calcul : on veut calculer l'intégrale I de la fonction à trois variables $(x, y, z) \mapsto x$ sur le domaine de l'espace constitué des triplets (x, y, z) de réels positifs vérifiant $x + y + z \leq 1$. On a donc $I = \int_0^1 \int_0^{1-z} \int_0^{1-z-y} x dx dy dz = \frac{1}{2} \int_0^1 \int_0^{1-z} (1-z-y)^2 dy dz = -\frac{1}{6} \int_0^1 [(1-z-y)^3]_0^{1-z} dz = \frac{1}{6} \int_0^1 (1-z)^3 dz = \frac{1}{24} [(1-z)^4]_0^1 = \frac{1}{24}$.

3.2 Intégrales curvilignes.

Définition 20. Une **forme différentielle** ω sur $\mathbb{R}^2$ est une expression de la forme $P(x, y) dx + Q(x, y) dy$ (où P et Q sont deux fonctions à deux variables qui pour nous seront toujours continues). Une forme différentielle est **exacte** si elle correspond à la différentielle d'une fonction F , autrement dit si $\omega = dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy$.

Définition 21. Si $u : [a, b] \rightarrow \mathbb{R}^2$ est une fonction paramétrant une courbe du plan usuellement notée C , l'**intégrale curviligne** de la forme différentielle ω le long de la courbe C est l'intégrale $\oint_C \omega = \int_a^b \omega(u(t)) \cdot u'(t) dt$. Autrement dit, en notant $u(t) = (x(t), y(t))$, on a $\oint_C \omega = \int_a^b P(x(t), y(t)) x'(t) + Q(x(t), y(t)) y'(t) dt$.

Remarque 21. L'intégrale curviligne qu'on vient de définir ne dépend en fait pas du paramétrage choisi pour la courbe C : si une autre fonction v décrit la même courbe plane que u , avec le même point de départ et le même point d'arrivée, alors le calcul de l'intégrale curviligne correspondante donnera le même résultat.

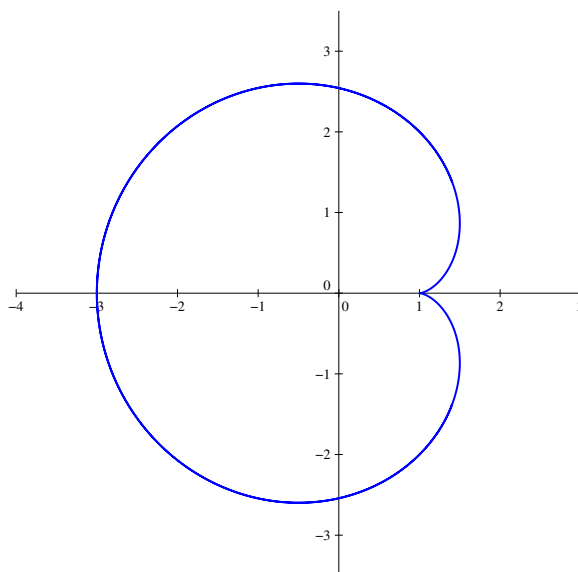
Proposition 4. Si ω est une forme exacte égale à la différentielle de la fonction F , alors $\oint_C \omega = F(u(b)) - F(u(a))$.

Autrement dit, l'intégrale d'une forme exacte le long d'une courbe ne dépend que du point de départ et du point d'arrivée de cette courbe, mais pas de la courbe elle-même. En particulier, l'intégrale curviligne d'une forme différentielle exacte le long d'une courbe fermée (points de départ et d'arrivée identiques) est toujours nulle.

Théorème 11. Formule de Green-Riemann.

Si C est une courbe fermée du plan délimitant un domaine D , de façon à ce que D soit situé « à gauche » de C (autrement dit, on parcourt le bord du domaine D dans le sens trigonométrique), et $\omega = P(x, y) dx + Q(x, y) dy$ une forme différentielle, alors $\oint_C \omega = \iint_D \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dx dy$.

La formule de Green-Riemann est une sorte de formule magique permettant de transformer mystérieusement une intégrale double en intégrale simple, particulièrement efficace notamment pour des calculs d'aires dans le plan. Mais pour pouvoir l'exploiter, il faut maîtriser le paramétrage des courbes planes, chose que vous n'étudiez plus vraiment en classes préparatoires. Un exemple quand même : on peut paramétrer le cercle trigonométrique C par la fonction $u : [0, 2\pi] \rightarrow \mathbb{R}^2$ définie par $u(t) = (\cos(t), \sin(t))$. Terminons ce chapitre avec un calcul rigolo faisant intervenir la formule de Green-Riemann : la courbe paramétrée par la fonction $u : \begin{cases} [0, 2\pi] & \rightarrow & \mathbb{R}^2 \\ t & \mapsto & (2 \cos(t) - \cos(2t), 2 \sin(t) - \sin(2t)) \end{cases}$ est la superbe cardioïde représentée ci-dessous :



En notant C cette cardioïde, et D le domaine qu'elle délimite dans le plan, on voudrait calculer l'aire de D . On sait que cette aire A peut s'exprimer sous la forme $A = \iint_D 1 dx dy$. On aimerait

appliquer la formule de Green-Riemann pour se ramener à quelque chose de plus facile, mais il faut pour cela trouver une forme différentielle $\omega = P \, dx + Q \, dy$ vérifiant $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$. Ce n'est pas très dur, il suffit de poser $P(x, y) = -y$ et $Q(x, y) = 0$ (il y a d'autres choix possibles). On peut alors affirmer que $A = -\oint_C y \, dx = \int_0^{2\pi} (2 \sin(t) - \sin(2t))(2 \sin(t) - 2 \sin(2t)) \, dt$ (on a simplement effectué le produit $-y(t)x'(t)$), donc $A = 2 \int_0^{2\pi} 2 \sin^2(t) - 3 \sin(t) \sin(2t) + \sin^2(2t) \, dt = 2 \int_0^{2\pi} 1 - \cos(2t) - 6 \sin^2(t) \cos(t) + \frac{1 - \cos(4t)}{2} \, dt = 2 \left[\frac{3}{2}t - \frac{\sin(2t)}{2} - 2 \sin^3(t) - \frac{\sin(4t)}{8} \right]_0^{2\pi} = 6\pi$.