Méthodes – Calcul différentiel
✓ Règle de la chaîne

1 ✓ Comment faire un changement de coordonnées

Il ne s'agit pas ici de vous apprendre : comment faire un « bon » changement de variable (au sens : savoir lequel faire en vue de simplifier un problème). Ce n'est pas une compétence qu'on attend de vous, ni que vous pourriez avoir de toute façon. L'objectif est plutôt :
— comment ne pas se tromper en appliquant la règle de la chaîne ?
— comment l'écrire rigoureusement, et non « à la physicienne » ?

L'idée derrière la formule :

$$\frac{\partial(f\circ\varphi)}{\partial u}(u,v)=\frac{\partial x}{\partial u}(u,v)\frac{\partial f}{\partial x}(x(u,v),y(u,v))+\frac{\partial y}{\partial u}(u,v)\frac{\partial f}{\partial y}(x(u,v),y(u,v)),$$

est : la dérivée partielle par rapport à la première variable de $f\circ\varphi=f\circ(x,y)$ mesure les variations de $f\circ(x,y)$ quand cette variable varie, les autres étant fixées ; mais quand bouge la première variable, les variations de $(u,v)\mapsto x(u,v)$ et $(u,v)\mapsto y(u,v)$ sont certes selon la *première* variable, mais elles affectent les *deux* variables de f (vu qu'on calcule $f(x(u,v),y(u,v))$), donc la règle de la chaîne doit dépendre des variations par rapport aux DEUX variables de f, et des variations de x et y par rapport à la PREMIÈRE : on voit que le nom des variables n'intervient jamais vraiment dans ma réflexion, seule leur POSITION compte.

1.1 Comment l'écrire rigoureusement, et non « à la physicienne » ?

Le changement de variable sera en général formulé de manière non rigoureuse, sous la forme $(u,v)=(\diamondsuit,\heartsuit)$ (dépendant de x et y, « anciennes coordonnées ») ou $(x,y)=(\spadesuit,\clubsuit)$ (dépendant de u et v, « nouvelles coordonnées »). Je vais partir du principe, dans l'exposition ici, que l'objectif est d'exprimer les dérivées partielles par rapport aux « anciennes » variables, en fonction des dérivées partielles par rapport aux « nouvelles ». Ici, cela revient à retenir le changement de coordonnées $(x,y)=(\spadesuit,\clubsuit)$. Alors :

Utiliser la règle de la chaîne : du brouillon à la rédaction propre

Raisonnement au brouillon. On pose $(x,y)=(\spadesuit,\clubsuit)$ (expressions dépendant de u et v). Si l'on veut exprimer $\frac{\partial f}{\partial x}$ en fonction de $\frac{\partial F}{\partial u}$ et $\frac{\partial F}{\partial v}$:

1. On écrit improprement que $f(x,y)=F(u,v)$.
2. On écrit $(*)$: $\frac{\partial f}{\partial x}(x,y)=\frac{\partial F}{\partial x}(u,v)$ (on « applique $\frac{\partial}{\partial x}$ à l'égalité »).
3. On dit que pour faire apparaître des dérivées par rapport à u et v de F, on multiplie et divise par ce qu'il manque, à savoir ∂u et ∂v. On obtient $(\dagger)$: $\frac{\partial}{\partial x}=\frac{\partial u}{\partial x}\times\frac{\partial}{\partial u}+\frac{\partial v}{\partial x}\frac{\partial}{\partial v}$.
4. On injecte $(\dagger)$ dans $(*)$, et on obtient :

$$\frac{\partial f}{\partial x}(x,y)=\left(\frac{\partial u}{\partial x}\times\frac{\partial}{\partial u}+\frac{\partial v}{\partial x}\frac{\partial}{\partial v}\right)F(u,v)=\frac{\partial u}{\partial x}\times\frac{\partial F(u,v)}{\partial u}+\frac{\partial v}{\partial x}\times\frac{\partial F(u,v)}{\partial v}.$$

Rédaction rigoureuse.

1. On identifie les « variables qu'on dérive » (ci-dessus : u et v, par rapport à x). On leur donne donc des noms de *fonctions* pour qu'on puisse parler de dérivée : $u(x,y)=\cdots$ et $v(x,y)=\cdots$ (nécessite d'exprimer u et v en fonction de x et y en partant de la relation $(x,y)=(\spadesuit,\clubsuit)$).
2. Avec les notations rigoureuses, on *veut* poser F tel que : $f(x,y)=F(u(x,y),v(x,y))$ (et non $F(u,v)$). On pose $\varphi(x,y)=(u(x,y),v(x,y))$ pour abréger, de sorte qu'il soit clair qu'on *veut* poser F tel que : $f(x,y)=F\circ\varphi(x,y)$, et qu'il suffit de poser $F=f\circ\varphi^{-1}$ pour y parvenir.
3. On dit dans la copie qu'on peut utiliser la règle de la chaîne car l'application $\varphi:(x,y)\mapsto(u(x,y),v(x,y))$ est de classe C^1, et on réécrit l'égalité ci-dessus en remplaçant u, v par $u(x,y)$ et $v(x,y)$:

$$\frac{\partial f}{\partial x}(x,y)=\frac{\partial u}{\partial x}(x,y)\times\frac{\partial F}{\partial u}(u(x,y),v(x,y))+\frac{\partial v}{\partial x}(x,y)\times\frac{\partial F}{\partial v}(u(x,y),v(x,y)).$$

Le fait d'avoir une somme dans (†), au lieu de se contenter d'écrire : $\frac{\partial}{\partial x} = \frac{\partial u}{\partial u} \times \frac{\partial}{\partial x} = \frac{\partial u}{\partial x} \times \frac{\partial}{\partial u}$ ou $\frac{\partial}{\partial x} = \frac{\partial v}{\partial v} \times \frac{\partial}{\partial x} = \frac{\partial v}{\partial x} \times \frac{\partial}{\partial v}$, est contre-intuitif et injustifiable *a priori* dans notre « raisonnement au brouillon », du fait qu'il soit un raisonnement faux ; on peut à la rigueur se convaincre ainsi : comme x dépend de u et v, il serait bizarre de ne l'écrire qu'à l'aide de la dérivée par rapport à u ou par rapport à v, donc informellement on résout cette bizarrerie « en appliquant l'opération décrite ci-dessus avec ∂u ou ∂v séparément, puis en sommant pour tenir compte à la fois de u et v ».

Détaillons le contenu de l'encart ci-dessus :

1. Vous posez $x(u, v) = \spadesuit$ et $y(u, v) = \clubsuit$, c'est-à-dire : **vous remplacez les variables par des fonctions** (même si cela ne veut rien dire : tout est question de notations). Et vous donnez un nom à la fonction $(u, v) \mapsto (x(u, v), y(u, v))$; avec les notations ci-dessus, ce serait φ.

   Posez $V = \varphi(U)$, et si possible explicitez cet ensemble (si la donnée de départ est U bien entendu, sinon vous vous posez la question dans l'autre sens) : vous en avez besoin dans la résolution des équations aux dérivées partielles (en particulier il faut s'assurer que V est convexe).

2. Écrire désormais $f(u, v)$ au lieu de $f(x, y)$, et penser qu'ainsi nous avons bien fait le changement de variable, n'a AUCUN sens mathématique : **la fonction dans les « nouvelles variables » porte un autre nom, et s'obtient en composant f avec l'application de changement de coordonnées**. Je l'appelle F dans ce qui suit.

   Pour écrire le lien entre F et f et utiliser la règle de la chaîne, il s'agit *faussement* d'écrire une égalité entre $F(u, v)$ et $f(x, y)$. Autrement dit, on aimerait écrire **à la physicienne** :

   $$F(u, v) = f(x, y),$$

   chose qui n'est pas pertinente en mathématiques. Pour que ce soit rigoureux, vous remplacez $f(x, y)$ par $f(x(u, v), x(u, v))$, c'est-à-dire :

   $$\forall (u, v) \in V, \quad F(u, v) = f(x(u, v), y(u, v)) = f(\varphi(u, v)) = f \circ \varphi(u, v).$$

3. Si on veut les dérivées partielles de F (« nouvelles coordonnées ») en fonction de celles de f (« anciennes coordonnées »), alors il suffit d'appliquer la règle de la chaîne à $f \circ \varphi$, en vertu de l'égalité ci-dessus. **Sinon, vous devez calculer la bijection réciproque de** φ (c'est là qu'on voit qu'il dépend selon qu'on nous donne $(x, y) = (\spadesuit, \clubsuit)$ ou $(u, v) = (\diamondsuit, \heartsuit)$ au début).

   Pour calculer φ^{-1}, vous résolvez pour tout $(x, y) \in V$ l'équation $\varphi(u, v) = (x, y)$ d'inconnue $(x, y) \in V$, et vous appelez la solution $\varphi^{-1}(x, y)$. Alors, l'égalité $F = f \circ \varphi$ implique, après composition par φ^{-1} :

   $$\forall (x, y) \in U, \quad f(x, y) = F(\varphi^{-1}(x, y)),$$

   et vous appliquez la règle de la chaîne à cette égalité-là.

Si, à la place, vous partez de $(u, v) = (\diamondsuit, \heartsuit)$ (expression des nouvelles coordonnées en fonction des anciennes), voici les modifications à apporter : vous posez $u(x, y) = \diamondsuit$ et $v(x, y) = \heartsuit$ (**remplacement par des fonctions**). Ensuite, pour le lien entre F (fonction exprimée dans les nouvelles coordonnées) et f, il s'agit *faussement* d'écrire une égalité entre $F(u, v)$ et $f(x, y)$. Autrement, on aimerait écrire **à la physicienne** :

$$F(u, v) = f(x, y),$$

chose qui n'est pas pertinente en mathématiques. Pour que ce soit rigoureux, vous remplacez $F(u, v)$ par $F(u(x, y), v(x, y))$ (et vous posez $\psi : (x, y) \mapsto (u(x, y), v(x, y))$ par commodité), de sorte que :

$$F \circ \psi(x, y) = f(x, y),$$

et le tour est joué. Vous pouvez reprendre au point 3, en remplaçant φ^{-1} par ψ.

2 ✓ Comment résoudre une équation aux dérivées partielles

Supposons que l'on cherche toutes les applications $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^1 vérifiant une équation différentielle linéaire impliquant les dérivées partielles $\partial_i f$ et $\partial^2_{i,j} f$. Par exemple :

$$\forall (x,y) \in U,\ \frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial y}(x,y) = xy, \quad \text{et :} \quad \forall (x,y) \in U,\ x^2 \frac{\partial^2 f}{\partial x^2}(x,y) + 2xy \frac{\partial^2 f}{\partial x \partial y}(x,y) + y^2 \frac{\partial^2 f}{\partial y^2}(x,y) = 0.$$

L'objectif sera systématiquement de vous ramener à des équations où il n'intervient QU'UNE SEULE dérivée partielle :

$$\partial_i F = 0, \quad \text{ou } \partial_i F = g, \quad \text{ou } \partial^2_{i,j} F = 0, \quad \text{ou } \partial^2_{i,j} F = h,$$

après un changement de coordonnées appropriées (F est le nom de la fonction obtenue après ce changement de coordonnées). De la sorte, vous ignorez (pour ainsi dire) les autres variables, et vous avez là une équation différentielle tout à fait ordinaire, que vous savez résoudre en intégrant. J'en redis deux mots ci-dessous.

Pour y parvenir, on vous explicitera le changement de coordonnées. **Votre première tâche sera de renommer ce changement de coordonnées en lui donnant un nom de fonction**. Ainsi, si f est une fonction des variables x et y, dire : « on pose $(u,v) = (\diamondsuit, \heartsuit)$ » (dépendant de x et y) revient à poser :

$$f(x,y) = F \circ \psi(x,y), \quad \text{où : } \psi(x,y) = (\diamondsuit, \heartsuit) = (u(x,y), v(x,y)).$$

On pose $u(x,y) = \diamondsuit$ et $v(x,y) = \heartsuit$ pour mieux comprendre ce que représentent ces deux coordonnées de ψ. Je note aussi $V = \psi(U)$ pour la suite : c'est l'ensemble des valeurs prises par les « nouvelles coordonnées ».

L'application F représente « f dans les nouvelles coordonnées », tandis qu'on dira que f est exprimé dans les « anciennes coordonnées ». Ensuite :

1. On exprime les dérivées partielles de f en fonction de celles de F (les « anciennes » en fonction des « nouvelles »). D'un point de vue cognitif, vous vous en sortirez mieux si vous écrivez que vous dérivez F par rapport à u et v, même si en vérité le nom des variables n'a pas d'importance.
2. Dans l'équation aux dérivées partielles vérifiée par f, vous remplacez toutes les « anciennes » dérivées partielles (celles de f) par leurs expressions en fonction des « nouvelles ». Vous tomberez alors sur une équation de la forme :

$$\forall (x,y) \in U,\ \partial_i F(\psi(x,y)) = 0, \quad \text{ou} \quad \partial_i F(\psi(x,y)) = g(x,y).$$

3. **Vous vérifiez que ψ est bijective**. Et, d'ailleurs, vous calculez sa réciproque $\psi^{-1} = \varphi$ explicitement, en résolvant pour tout $(u,v) \in \psi(U)$ l'équation $\psi(x,y) = (u,v)$ d'inconnue $(x,y) \in U$, et en appelant la solution $\varphi^{-1}(u,v)$: vous en aurez besoin pour plus tard. Pour l'heure, vous vous contentez d'en déduire que vous pouvez composer par $\psi^{-1} = \varphi$ l'égalité ci-dessus, pour en déduire :

$$\partial_i F = 0, \quad \text{ou} \quad \partial_i F = g \circ \varphi.$$

L'équation est valable sur V.

4. **On vérifie que V est convexe**, afin de pouvoir utiliser nos résultats sur les fonctions de classe C^1 constantes ou qui ne dépendent pas d'une variable.
5. Ensuite, cela dépend du cas de figure où l'on est :
   (a) Dans le cas où $\partial_1 F = 0$, cela signifie que F ne dépend pas de la première variable : il existe une application $a : I \to \mathbb{R}$ de classe C^1 telle que $F(u,v) = a(v)$ pour tout $(u,v) \in V$.
   (b) Dans le second cas, pour tout v on trouve une primitive $H_v : I \to \mathbb{R}$ de l'application $u \mapsto h_2(u,v)$, et dans ce cas $F(u,v) = H_v(u) + K(v)$, où $K : J \to \mathbb{R}$ est une application de classe C^1 (on l'interprète comme la « constante d'intégration » ; vu qu'on a dérivé selon la première variable, en intégrant on obtient une constante de la variable u, donc elle dépend uniquement de v).

   Évidemment, si on a à la place : $\partial_2 F(\psi(x,y)) = 0$, ou $\partial_2 F = h_1 \circ \varphi$, on reprend le même raisonnement, en inversant les rôles de u et v.

6. On en déduit f *via* l'égalité : $\forall(x,y)\in\mathbb{R}$, $f(x,y)=F\circ\psi(x,y)$, et les expressions trouvées ci-dessus ; on vérifie réciproquement que toutes les fonctions trouvées conviennent.

Remarque. Parfois, vous n'avez pas besoin d'exprimer les « anciennes » dérivées partielles en fonction des « nouvelles ». Ce sont des « coups de chance » où, en fait, le membre de gauche de l'équation différentielle est DÉJÀ la dérivée partielle de F par rapport à une variable. Vous rencontrez plus particulièrement cette situation lorsque vous avez un changement en coordonnées polaires $(r,\theta)\mapsto(r\cos(\theta),r\sin(\theta))$ à effectuer. Dans ce cas, pour reconnaître les dérivées par rapport à r et θ rien qu'à la vue de l'équation, il vaut mieux les avoir en tête en permanence (ici écrites sous forme non rigoureuse) :

$$\frac{\partial F}{\partial r}=\cos(\theta)\frac{\partial f}{\partial x}+\sin(\theta)\frac{\partial f}{\partial y},\quad \frac{\partial F}{\partial \theta}=-r\sin(\theta)\frac{\partial f}{\partial x}+r\cos(\theta)\frac{\partial f}{\partial y}.$$

Comme vous les utilisez aussi en Physique, j'ose espérer que ce n'est pas trop vous demander.

2.1 Et en cas de dérivées partielles secondes d'ordre 2 ?

Cela ne change rien. Il faut simplement intégrer deux fois l'équation :

$$\partial^2_{i,j}F=0,\quad\text{ou}\quad \partial^2_{i,j}F=h.$$

Vous choisissez si vous intégrez par rapport à la i^{e} variable, ou la j^{e}. Peu importe. Si vous commencez par intégrer par rapport à u, alors vous trouvez une primitive de h comme si c'était une fonction d'une seule variable (en l'occurrence u), et vous traitez v comme une constante. Par exemple, si h est l'application $(u,v)\mapsto uv+v$, alors en intégrant par rapport à u vous obtenez l'application $(u,v)\mapsto\frac{u^2v}{2}+uv+c(v)$, *avec la constante qui dépend de* v. Et vous recommencez, en intégrant cette fois par rapport à v. Il faut simplement donner un nom à une primitive de $v\mapsto c(v)$, puisque ce n'est pas une fonction explicite.

Table des matières