Des jets aux infiniment petits: quand l'intuition se mue en rigueur.

Francis Borceux

Paris, le 1er décembre 2007

Le plan tangent à une surface — disons, une surface de l'espace réel usuel à trois dimensions — est quelque chose que nous pouvons définir facilement, de beaucoup de façons différentes d'ailleurs, en utilisant quelque part des dérivées. Rappelez-vous : la *dérivée* d'une fonction

$$f\colon \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto f(x)$$

c'est la fonction

$$\frac{df}{dx}\colon \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto \text{pente en } x \text{ de la tangente à } y = f(x) \ .$$

Pour raccourcir le langage, faisons une fois pour toute la convention que toutes les fonctions que je considère sont dérivables ; je ne le rappellerai plus.

On définit donc le plan tangent au moyen de dérivées ; mais la dérivée de quoi ? Ce que nous devons dériver c'est l'équation, ou les équations, qui décrivent la surface dans l'espace ambiant. Les équations qui nous disent où se trouve précisément la surface dans l'espace. Mais une sphère, ou un tore, ou un ellipsoïde, ce sont des surfaces qui existent par elles-mêmes, indépendamment de la manière dont on les plonge dans un espace à trois, quatre, ou cinq dimensions. Quand je vous parle du plan tangent à une sphère, vous comprenez ce que je veux dire : je n'ai pas besoin de vous donner des équations, de vous préciser la position de la sphère dans l'espace. Donc on devrait pouvoir parler de l'espace tangent à une surface, indépendamment de tout plongement de cette surface dans un espace à trois, ou quatre, ou cinq dimensions. Et c'est *a fortiori* vrai si l'on pense à des surfaces telles que le plan projectif : le plan usuel auquel on a ajouté des “points à l'infini”. Voilà une surface que plusieurs d'entre nous ont rencontrée ... en général sans même savoir que le plan projectif se plonge dans l'espace à six dimensions.

Les mathématiciens savent depuis pas mal de temps déjà comment définir une surface de manière intrinsèque, sans la regarder comme une partie d'un espace de dimension supérieure. Je me limiterai à l'approche la plus intuitive, même si elle est loin d'être la plus générale.

Aucun d'entre nous, je pense, n'a encore eu l'occasion d'embarquer dans un vaisseau spatial pour regarder la Terre de loin et faire de la géographie *de visu, en vraie grandeur*. Quand nous nous intéressons à recueillir une information concernant la surface de la Terre — la position d'une ville, le cours d'un fleuve, un itinéraire routier, l'altitude d'une montagne — nous ne montons pas dans une fusée pour aller embrasser tout cela d'un seul coup d'oeil, non : nous consultons *un atlas géographique*. C'est quoi, un atlas géographique ? C'est une collection de cartes locales, chacune représentant une partie de la surface de la Terre. Aujourd'hui je choisis de me simplifier la vie : je vais m'intéresser à une seule carte, pas à la façon dont on passe d'une carte à l'autre.

Et une carte de géographie, c'est quoi ? C'est la représentation, sur un morceau de plan, d'une surface qui dans la réalité, n'est pas plate du tout. Prenez une carte de l'*Institut Géographique National* : c'est une feuille toute plate, de la taille d'une table, mais qui — par exemple, grâce aux courbes de niveau — vous permet de déduire ce qu'est le relief sur une étendue de plusieurs kilomètres carrés. Les lignes de niveau vous permettent de repérer les collines, les vallées. L'écartement entre les lignes de niveau vous permet de juger d'un coup d'oeil si la pente est raide ou si elle est douce. Etc.

Mathématiquement, une "carte" d'un morceau de surface est la donnée d'un morceau du plan $\mathbb{R}^2$ et d'indications métriques qui permettent, à partir de cette représentation fatalement faussée qu'est la carte, de *recalculer* ce que sont les vraies longueurs, les vrais angles, les vraies directions sur la surface. J'ai bien dit *recalculer* — comme ont l'habitude de le faire les marins — pas simplement *mesurer à l'échelle* : car la carte est plate, alors que la surface de la Terre ne l'est pas. Plus généralement, une *surface de Riemann* est un atlas de cartes, c'est-à-dire de morceaux de $\mathbb{R}^2$, munis d'indications métriques permettant de calculer les vraies longueurs et les vrais angles sur la surface, ainsi que d'indications sur la manière dont des cartes voisines se chevauchent.

Mais revenons-en à nos bonnes vieilles surfaces plongées dans l'espace à trois dimensions. Des surfaces que nous voulons étudier à partir d'une carte bien plate, d'un morceau de $\mathbb{R}^2$ muni d'indications métriques. Le plan tangent en un point de la surface, c'est tout simplement l'espace de tous les vecteurs tangents en ce point. Et il reste à définir ce qu'est un vecteur tangent. Mais attention n'oublions pas que nous sommes à la recherche d'une définition intrinsèque, une définition qui ne fait pas référence au plongement particulier de la surface dans l'espace à trois dimensions. Une définition que nous pouvons donner à partir de notre carte toute plate et de ses indications métriques.

Le problème est qu'un vecteur tangent, cela ne se trouve pas à l'intérieur de la surface, cela se trouve en dehors de celle-ci, dans l'espace ambiant. Donc le vecteur tangent n'est pas représenté à l'intérieur de la carte. Mais un vecteur tangent à la surface, c'est en particulier un vecteur tangent à une courbe tracée sur la surface, une courbe qui elle, est bel et bien à l'intérieur de la surface. Donc une courbe que je peux représenter sur ma carte. Bien sûr, beaucoup de courbes différentes, tracées sur la surface, peuvent avoir le même vecteur tangent. Qu'à cela ne tienne : il suffit de faire ce que l'on appelle un *quotient*. Déclarons *équivalentes en un point* deux courbes tracées sur la surface et qui ont le même vecteur tangent en ce point. Pour préciser un vecteur tangent en un point, je peux tout aussi bien préciser quel est le “paquet” de toutes les courbes tracées sur la surface et admettant ce vecteur comme vecteur tangent. Donner l'ensemble de tous ces “paquets” de courbes, c'est équivalent à donner l'ensemble de tous les vecteurs tangents. Je peux donc définir le plan tangent à la surface comme étant l'ensemble de tous ces “paquets” de courbes, deux courbes étant dans le même “paquet” quand elles admettent le même vecteur tangent.

Mais cela ne résoud en rien notre problème car si les courbes sont bien tracées sur la surface, la relation d'équivalence entre deux courbes — le fait d'avoir le même vecteur tangent — utilise explicitement ce vecteur tangent … la chose précisément que nous cherchions à définir intrinsèquement.

Mais cette nouvelle difficulté est facile à surmonter. Considérons une surface S dans l'espace à trois dimensions et une carte de cette surface. Il s'agit d'un morceau C du plan $\mathbb{R}^2$, et d'une fonction

$$f\colon C \longrightarrow S \subseteq \mathbb{R}^3, \quad (u,v) \mapsto f(u,v)$$

qui me dit quel point de la surface correspond à chaque point de la carte. Plus des données métriques que je ne précise pas ici. Pour supporter l'intuition, imaginez le cas de la sphère et pensez que u et v sont la “longitude” et la “latitude” du point correspondant sur la sphère.

Une courbe sur la surface peut bien sûr être décrite en dessinant cette courbe sur la carte, c'est-à-dire en donnant une fonction

$$g\colon]a,b[\longrightarrow C \subseteq \mathbb{R}^2, \quad t \mapsto g(t).$$

La courbe dans l'espace, la vraie courbe, sur la vraie surface, est alors simplement le composé $f \circ g$:

$$]a, b[\xrightarrow{\quad g \quad} C \xrightarrow{\quad f \quad} S.$$

Pour la facilité des notations, convenons que nous avons choisi les paramètres de sorte que $0 \in]a, b[$ et que $g(0) = (u_0, v_0)$ soit le point en lequel nous voulons calculer le plan tangent.

Vous ne serez guère étonnés si je vous dis que deux courbes tracées sur la surface sont tangentes si et seulement si leurs représentations sur la carte sont tangentes. Si ce n'était pas le cas, avouez que la carte serait d'une piètre qualité ! C'est bien le cas, rassurez-vous, et cela se démontre facilement à coup de dérivées partielles. Je vous fais grâce des détails.

Considérons donc une autre courbe tracée sur la surface et passant par le même point de coordonnées (u_0, v_0). Il n'y a bien sûr aucune restriction à écrire cette seconde courbe en fonction du même paramètre t :

$$]a, b[\xrightarrow{\quad h \quad} C \xrightarrow{\quad f \quad} S.$$

Cette seconde courbe sur la surface admet le même vecteur tangent que la première, précisément quand les deux courbes tracées sur la carte C

$$g \colon]a, b[\longrightarrow C, \quad h \colon]a, b[\longrightarrow C$$

admettent le même vecteur tangent. Et rappelons-nous qu'une tangente, cela se calcule au moyen d'une dérivée. Donc finalement les deux courbes $f \circ g$ et $f \circ h$ tracées sur la surface admettent le même vecteur tangent au point considéré si et seulement si

$$\frac{dg}{dt}(0) = \frac{dh}{dt}(0)$$

Cette fois c'est gagné : il s'agit là d'une condition exprimée uniquement dans la carte C, sans plus recourir à la fonction f qui nous explique comment la surface est positionnée dans l'espace. Deux courbes tracées sur la carte C par un point (u_0, v_0) sont donc décrétées *équivalentes* lorsqu'elles ont le même vecteur tangent dans cette carte, c'est-à-dire quand

$$\frac{dg}{dt}(0) = \frac{dh}{dt}(0).$$

Une classe d'équivalence pour cette relation d'équivalence, un "paquet de courbes" équivalentes en ce sens, c'est ce qu'on appelle un *jet*. L'ensemble de ces *jets*, c'est ce qu'on appelle l'*espace tangent* à la surface au point (u_0, v_0). Et on a ainsi défini l'espace tangent, à partir de la carte, sans se référer au plongement dans l'espace.

La notion de *jet* est due à *Charles Ehresmann*. Celle que j'ai définie ici est celle de *1-jet*, car elle n'utilise que des dérivées premières. Il y a plus généralement une notion de r*-jet*, utilisant des dérivées jusqu'à l'ordre r. Et bien sûr au lieu de courbes et de surfaces, c'est-à-dire de dimensions 1 et 2, la théorie des jets s'applique tout aussi bien en dimensions supérieures. Cette théorie des jets, due à *Ehresmann*, est aujourd'hui utilisée de manière universelle en géométrie différentielle, au point que beaucoup de gens ne savent plus d'où elle vient, tellement elle apparaît naturelle. La notion de *jet* a en particulier été utilisée par d'autres grands mathématiciens tels *René Thom* et *André Weil* pour étudier les singularités des variétés différentiables.

Et bien même si cela ne saute pas encore aux yeux, cette notion de *jet* nous amène aux portes d'une théorie des infiniment petits, suivant une idée de *Lawvere* et de *Kock*.

Pour comprendre cela, intéressons-nous au cas particulier d'une parabole tracée dans notre carte C de la surface :

$$y = x^2.$$

Pour reprendre les notations antérieures, cette parabole est la courbe

$$g\colon]a,b[\longrightarrow C, \quad t \mapsto (t, t^2).$$

Le vecteur tangent à cette parabole est donné par la dérivée de g

$$\frac{dg}{dt} = (1, 2t)$$

et sa valeur à l'origine est

$$\frac{dg}{dt}(0) = (1, 0).$$

Comme seconde courbe, considérons l'axe des x, qui est d'équation

$$y = 0.$$

Avec les notations antérieures, il s'agit de la courbe

$$h\colon]a,b[\longrightarrow C, \quad t \mapsto (t, 0).$$

Le vecteur tangent est donné par la dérivée de h

$$\frac{dh}{dt} = (1, 0).$$

C'est un vecteur constant, qui ne dépend plus de t, donc en particulier

$$\frac{dh}{dt}(0) = (1, 0).$$

Conclusion, la parabole et l'axe des x admettent le même vecteur tangent à l'origine

$$\frac{dg}{dt}(0) = (1, 0) = \frac{dh}{dt}.$$

La parabole et l'axe des x définissent le même *jet* à l'origine.

Intuitivement, deux courbes g et h donnant lieu au même jet en un point, c'est-à-dire ayant le même vecteur tangent en un point, donc la même tangente en ce point, sont de plus en plus indiscernables lorsque l'on s'approche de plus en plus du point considéré. En d'autres termes, si t est un réel de plus en plus petit, les deux morceaux de courbes déterminés par g et h sur l'intervalle $]-t, +t[$ sont de plus en plus indiscernables. Et si cela avait un sens, et bien on pourrait dire que pour une valeur de t infiniment petite, les deux courbes g et h coincident.

Et bien pendant un moment, faisons semblant de croire que des nombres réels infiniment petits, cela existe. Faisons semblant de croire que nos deux fonctions g et h qui ont le même vecteur tangent à l'origine sont égales sur tous les nombres réels infiniment petits. Alors pour tout nombre réel t infiniment petit

$$(t, t^2) = (t, 0).$$

Conclusion, si nous prenions au sérieux l'existence de nombres réels infiniment petits, notre raisonnement nous conduirait à admettre que $t^2 = 0$ pour tout nombre réel t infiniment petit.

Mais tout cela n'était qu'un jeu bien sûr, car nous savons tous que le seul nombre réel t tel que $t^2 = 0$, c'est $t = 0$. Ce n'est certainement pas un mathématicien qui va vous dire le contraire.

Pourtant c'est bien là l'idée de la théorie des infiniment petits de *Lawvere* et *Kock*. Un nombre petit a un carré encore plus petit : un millième au carré, cela ne fait plus qu'un millionième. Un nombre très très petit a un carré presque négligeable. Et bien un nombre est décrété infiniment petit quand son carré devient tout bonnement nul. Existe-t-il de tels nombres, que bien sûr dans la suite j'éviterai d'encore appeler *nombres réels*? Pourquoi pas!

Bien sûr votre religion vous interdit de penser qu'il puisse exister des nombres autres que 0 dont le carré soit nul. C'est bien dommage : vous devriez changer de religion. Rappelez-vous, à une certaine époque, le même genre d'attitude interdisait aux mathématiciens de penser qu'un nombre négatif puisse avoir une racine carrée. Et pourtant Dieu sait — le dieu que vous voulez, puisque je vous ai dit de changer de religion — et pourtant Dieu sait si les nombres complexes sont importants en mathématique et en physique. Mais puisque nous sommes d'accord aujourd'hui de considérer qu'un nombre négatif puisse avoir une racine carrée — racine carrée qui n'est plus un nombre réel, mais un nombre complexe — posons-nous sérieusement la question de savoir s'il n'existerait pas une autre notion de nombre, plus générale que celle de nombre réel, pour laquelle un nombre non nul pourrait avoir un carré nul.

Oh, des objets de carré nul, il en existe beaucoup en mathématique. Par exemple, prenez la théorie des matrices. Il existe beaucoup de matrices non nulles dont le carré est nul. Par exemple

$$\begin{pmatrix} 0 & r \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & r \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Mais voici un autre exemple, beaucoup plus significatif pour le problème qui nous intéresse. Vous savez tous ce que sont des polynomes, comment on les additionne, on les soustrait, on les multiplie. Par exemple

$$\begin{aligned}
&(aX+b)+(cX^2+dX+e) = cX^2+(a+d)X+(b+e);\\
&(aX+b)-(cX^2+dX+e) = -cX^2+(a-d)X+(b-e);\\
&(aX+b)\times(cX^2+dX+e) = (ac)X^3+(bc+ad)X^2+(ae+bd)X+(be).
\end{aligned}$$

Les polynomes constitutent ce que l'on appelle un *anneau* : un ensemble muni d'une addition, d'une soustraction et d'une multiplication, qui satisfont à toutes les égalités arithmétiques usuelles de ces opérations.

Et bien maintenant faisons ce que l'on appelle le *quotient* de l'anneau des polynomes par l'équation

$$X^2 = 0.$$

Cela veut dire quoi ? Et bien cela veut dire qu'à partir de maintenant je remplace systématiquement X^2 par zéro. Absurde me direz-vous. Peut-être, mais pas tant que cela. Faisons une comparaison.

Imaginez que vous ayez une petite calculatrice de poche, reçue en cadeau dans une boîte de poudre à lessiver, une calculatrice qui travaille uniquement avec huit décimales. Que vous le veuillez ou non, pour votre calculatrice, un millionième au carré, c'est égal à zéro :

$$(0,00000100)^2 = 0,00000000.$$

Et bien sûr aussi,

$$\pi = 3,14159265$$

pas une décimale de plus ! Pourtant votre calculatrice fait des calculs, de manière parfaitement logique : son arithmétique répond à des règles précises, est tout ce qu'il y a de plus cohérent. Et de plus utile, avouez-le. Pourtant cette arithmétique de la calculatrice n'obéit pas aux règles usuelles de la vraie arithmétique des vrais nombres réels. Puisque par exemple, un millionième au carré est égal à zéro !

Et bien faisons maitenant une chose analogue avec les polynomes, en identifiant cette fois X^2 à 0. A fortiori

$$X^3{=}X^2X{=}0X{=}0,\ X^4{=}X^2X^2{=}0X^2{=}0,\ X^5{=}X^2X^3{=}0X^3{=}0,\ \ldots.$$

Tout ce qu'il nous reste, ce sont les polynomes du premier degré. Notre "calculatrice" à polynomes travaille donc uniquement dans l'ensemble

$$A = \{aX + b | a, b \in \mathbb{R}\}$$

des polynomes du premier degré. Vous serez je suppose d'accord avec moi pour dire qu'il est parfaitement légitime, même si un peu curieux, de vouloir se restreindre à travailler uniquement avec les polynomes du premier degré. Bien sûr si on additionne ou si on soustrait des polynomes du premier degré, on trouve toujours des polynomes du premier degré. Mais qu'en est-il de la multiplication ?

$$(aX + b) \times (cX + d) = (ac)X^2 + (ad + bc)X + (bd).$$

Le résultat n'est plus un polynome du premier degré ! Mais bien sûr puisque nous avons choisi d'identifier X^2 à 0, il parait cohérent de considérer les opérations suivantes sur notre ensemble A des polynomes du premier degré :

$$\begin{aligned}(aX + b) + (cX + d) &= (a + c)X + (b + d);\\ (aX + b) - (cX + d) &= (a - c)X + (b - d);\\ (aX + b) \times (cX + d) &= (ad + bc)X + (bd).\end{aligned}$$

Voilà, c'est tout. Imposer l'équation $X^2 = 0$ cela revient à dire qu'au lieu de travailler avec tous les polynomes, on se restreint à ne considérer que des polynomes du premier degré. Et sur les polynomes du premier degré, on définit les trois opérations ci-dessus, parfaitement légitimes, même si la troisième est tout-à-fait inhabituelle.

Et bien figurez-vous que contrairement à l'arithmétique de la calculatrice, ces nouvelles opérations continuent à vérifier toutes les égalités arithmétiques auxquelles vous êtes habitués, du genre

$$p(X) \times \big(q(X) + r(X)\big) = \big(p(X) \times q(X)\big) + \big(p(X) \times r(X)\big)$$

pour trois polynomes du premier degré ; etc. On a toujours ce que l'on appelle un *anneau* : un ensemble A muni d'une addition, d'une soustraction et d'une multiplication qui vérifient toutes les égalités arithmétiques usuelles de ces opérations. Mais cette fois, il y a bel et bien dans A des polynomes non nuls dont le carré est nul : à commencer par le polynome $p(X) = X$, qui est un vrai polynome non nul du premier degré, mais qui est tel que

$$p(X) \times p(X) = (1X + 0) \times (1X + 0) = 0.$$

Vous voyez : imposer l'égalité $X^2 = 0$ dans l'anneau des polynomes, ce n'est pas de la fumisterie. Cela revient tout simplement à construire à partir de l'anneau des polynomes un autre anneau A dans lequel cette équation $X^2 = 0$ est maintenant satisfaite. Mais la multiplication de ce nouvel anneau A — l'anneau des polynomes du premier degré — est cette fois une multiplication inhabituelle obtenue en combinant la multiplication usuelle et l'équation $X^2 = 0$ que l'on veut imposer. C'est ce que l'on appelle un *quotient algébrique* : le quotient de l'anneau des polynomes par l'équation $X^2 = 0$. Le processus du quotient par une équation est une opération fondamentale de l'algèbre moderne.

Quand on veut manipuler des racines carrées de nombres négatifs, on commence bien entendu par remplacer les nombres réels par les nombres complexes. Et bien la théorie des infiniment petits de *Lawvere* et *Kock* commence elle aussi par remplacer les nombres réels $\mathbb{R}$ par un certain anneau $\mathbf{R}$ contenant tous les nombres réels usuels et à l'intérieur duquel il est bien entendu parfaitement légitime de considérer

$$\mathbf{D} = \{r \in \mathbf{R} | r^2 = 0\}$$

l'ensemble des éléments de carré nul. Jusque là, rien d'extraordinaire. Et rien ne nous empêche, si cela nous amuse, d'appeler les éléments de $\mathbf{D}$ les *infiniment petits de l'anneau* $\mathbf{R}$. Mais *Lawvere* et *Kock* vont plus loin : ils prennent au sérieux l'idée que sur les infiniment petits, une fonction devient indiscernable de sa tangente, c'est-à-dire, devient une fonction du premier degré. Ils introduisent dès lors l'axiome :

Toute fonction $f : \mathbf{D} \longrightarrow \mathbf{R}$ s'écrit de manière unique
sous la forme $f(t) = at + b$, avec $a, b \in \mathbf{R}$.

Bien sûr faisant $t = 0$ on trouve $b = f(0)$. D'autre part s'il y a dans ce contexte une bonne notion de dérivée, on doit avoir

$$\frac{df}{dt}(0) = \frac{d(at + b)}{dt}(0) = a.$$

Lawvere et *Kock* prennent tout simplement cela comme définition : avec les notations de leur axiome, il définissent la dérivée de f par

$$\frac{df}{dt}(0) = a.$$

Voilà, c'est tout. La *géométrie différentielle synthétique*, dont l'étude a été initiée par *Lawvere* et *Kock*, c'est la géométrie différentielle obtenue en remplaçant les nombres réels par un anneau $\mathbf{R}$ satisfaisant à l'axiome ci-dessus, et la notion usuelle de dérivée par celle évoquée ci-avant.

Mais j'entends bien vos légitimes objections. Tout d'abord, un tel anneau $\mathbf{R}$, est-ce que cela existe ? Et même si un tel anneau existe, que diable tout cela a-t-il à voir avec la géométrie au sens usuel du terme ? Et bien pour vous mettre d'emblée … on ne peut plus mal à l'aise, observez simplement qu'un tel anneau, cela ne peut pas exister ! Du moins, un tel anneau avec de vrais infiniment petits non nuls. En effet considérez la fonction

$$f(t) = \begin{cases} 0 & \text{si } t = 0 \\ 1 & \text{si } t \neq 0 \end{cases}$$

L'axiome de *Kock–Lawvere* force cette fonction à être du premier degré sur les infiniment petits :

$$f(t) = at + b, \quad t \in \mathbf{D}.$$

On a alors

$$0 = f(0) = a0 + b = b.$$

Et s'il existe un infiniment petit non nul t, on obtient

$$1 = f(t) = at + b = at.$$

Mais $2t$ est encore infiniment petit puisque

$$(2t)^2 = 4t^2 = 4 \cdot 0 = 0.$$

Donc

$$1 = f(t + t) = a(t + t) = at + at = 1 + 1$$

et en simplifiant par 1, on trouve $0 = 1$! C'est raté ! Il ne peut exister aucun infiniment petit non nul t.

L'axiome de *Kock–Lawvere* est-il donc vide de sens ? Pas exactement. Car observez que pour donner le contre-exemple trivial ci-dessus, nous avons utilisé la règle du *tiers-exclu* : un nombre est nul, ou il ne l'est pas. Plus généralement : *une affirmation est vraie ou elle est fausse.* Le principe du *tiers-exclu* est certainement une règle de base de la logique classique, une logique qui — il faut bien le reconnaître — est finalement assez éloignée de la logique de *monsieur tout le monde.* Si je vous affirme : *Les Français font la grève*, vous n'allez pas me dire que j'ai raison — ou que j'ai tort. Vous me direz que certains jours à certains endroits j'ai raison, alors que d'autres jours à d'autres endroits j'ai tort. La valeur de vérité de la phrase *Les Français font la grève*, ce n'est pas *vrai* ou *faux*, c'est plus subtil que cela.

Il est bien connu qu'à l'intérieur des mathématiques les plus classiques qui soient, il existe des modèles intuitionistes de la théorie des ensembles, des modèles où le principe du tiers-exclu ne s'applique pas. Parmi ces modèles on trouve notamment ce que l'on appelle les *topos* : par exemple, les topos de faisceaux. Un objet d'un tel univers, d'un tel topos, n'est plus quelque chose de rigide, mais — par exemple — est un objet mathématique qui se déforme de manière continue lorsque l'on parcourt les points d'un certain espace de paramètres : un paramètre qui peut être le temps, ou n'importe quoi d'autre. Dans une telle structure, vous pouvez très bien avoir un élément x dont le carré n'est pas nul à l'instant t, mais dont le carré devient nul à un instant ultérieur. Et dans un tel univers intuitioniste, il peut cette fois très bien exister des anneaux satisfaisant à l'axiome de *Kock–Lawvere.*

L'exemple le plus connu de topos est celui du topos des faisceaux sur un espace topologique. Un espace topologique, c'est tout simplement un ensemble dans lequel on dispose d'une bonne notion de partie ouverte et de partie fermée. Prenons l'exemple le mieux connu, le plan $\mathbb{R}^2$ usuel. Une partie est fermée quand elle contient tous les points de sa frontière ; elle est ouverte quand elle ne contient aucun point de sa frontière. Et bien un faisceau sur le plan, vu comme espace topologique, c'est un *ensemble qui évolue en fonction d'un paramètre*, un paramètre qui prend comme valeurs possibles *tous les ouverts du plan.*

Pour comprendre, voyons un exemple concret de faisceau sur le plan. Intéressons nous à nouveau aux surfaces : les surfaces d'équation

$$Z = f(X, Y).$$

Certaines de ces surfaces sont définies pour toutes les valeurs de X et de Y, comme par exemple le paraboloide

$$Z = X^2 + Y^2.$$

Mais d'autres ne sont pas définies partout, comme par exemple

$$Z = \frac{1}{X} + \frac{1}{Y}$$

qui n'est définie que pour $X \neq 0$ et $Y \neq 0$. Cela montre que toutes les surfaces qui nous intéressent s'organisent naturellement en un *faisceau* F

$$F = \big(F(U)\big)_{U\ \in \mathcal{O}}$$

où

- $\mathcal{O}$ est l'ensemble de tous les ouverts du plan ;
- $F(U)$ est l'ensemble de toutes les surfaces définies sur l'ouvert U.

Bien sûr un tel faisceau possède une structure interne assez naturelle :

- si une surface est définie sur un ouvert U, elle est a fortiori définie sur n'importe quel ouvert plus petit ;
- et si on a une surface définie sur l'ouvert U et une autre définie sur l'ouvert V, et si ces deux surfaces coincident sur l'intersection $U \cap V$, on peut les "recoller" pour obtenir une surface définie sur la réunion $U \cup V$ des deux ouverts.

Ce sont là précisément les deux axiomes qui définissent un faisceau.

Tout topos possède, comme je l'ai dit, une logique intrinsèque — de nature intuitioniste, c'est-à-dire sans la règle du *tiers-exclu* — une logique dans laquelle on peut interpréter toutes les formules de la théorie des ensembles. Dans le cas particulier du plan, on peut démontrer que la logique intrinsèque du topos de faisceaux se réduit à une expression particulièrement simple : la *logique locale*. En voici la description :

> *Une formule φ de la théorie des ensembles est* localement valide *dans notre faisceau F si et seulement si, pour chaque point du plan, elle est valide au sens classique du terme sur tout un voisinage de ce point.*

Et cela change tout ! Cela fournit une logique toute différente de la logique classique ! Voyons le sur deux cas particuliers.

Tout d'abord, quand on a une surface partout définie

$$Z = f(X, Y)$$

il est bien clair qu'en chaque point (X, Y) du plan

$$f(X, Y) = 0 \quad \text{ou} \quad f(X, Y) \neq 0.$$

C'est le tiers-exclu ! Et bien cela est faux dans la logique locale ! En effet, pour que cela soit vrai dans notre logique locale, il faudrait qu'étant donné un point (X_0, Y_0) quelconque du plan, l'une des deux conditions suivantes soit satisfaite :

- $f(X,Y) = 0$ pour tous les points (X,Y) dans un voisinage de (X_0,Y_0) ;

ou

- $f(X,Y) \neq 0$ pour tous les points (X,Y) dans un voisinage de (X_0,Y_0).

La première condition cause un réel problème : $f(X_0,Y_0)$ peut très bien être nul sans que ce soit le cas pour aucun autre point (X,Y) voisin. Par exemple dans le cas du paraboloide

$$Z = X^2 + Y^2$$

$f(X,Y)$ s'annule en $(0,0)$, mais en aucun autre point au voisinage de $(0,0)$. La formule

$$f(X,Y) = 0 \text{ ou } f(X,Y) \neq 0,$$

vraie classiquement, n'est pas valable dans la logique locale. La logique locale ne se contente pas du fait qu'une des deux propriétés soit vraie en chaque point du plan, elle exige que quand une propriété est vraie en un point, elle soit vraie également tout autour de ce point.

Voyons un autre exemple, offrant des conclusions opposées.

Il existe un entier $n \in \mathbb{N}$ tel que la fonction f est bornée par n.

C'est évidemment faux classiquement : prenez à nouveau le paraboloide

$$Z = X^2 + Y^2.$$

Quand X et Y deviennent grands, la fonction $f(X,Y) = X^2 + Y^2$ peut prendre des valeurs aussi grandes que l'on veut, donc n'est bornée par aucun entier. Mais sur un petit voisinage de chaque point (X_0,Y_0), la fonction $f(X,Y)$ est bornée, donc est bel et bien bornée par un certain entier n. L'affirmation ci-dessus est dès lors vraie au sens de la logique locale, alors qu'elle était fausse classiquement. Dans ce second exemple, la logique locale est *plus souple* que la logique classique, car elle nous autorise à choisir un autre entier n chaque fois que nous changeons de point (X_0,Y_0).

Voilà esquisée, à travers quelques exemples concrets, une très brève évocation de ce qu'est la logique intrinsèque d'un topos, une logique qui — en particulier — n'admet pas la loi du *tiers-exclu*. Mais je l'ai déjà dit : un topos, muni de sa logique intrinsèque, c'est un modèle *intuitioniste* de la théorie des ensembles. Ce n'est pas trivial à démontrer, bien entendu. Mais le résultat est magnifique ! Car il signifie qu'aussi compliqué que puisse être le topos particulier qui nous intéresse, peu importe : tout théorème que je peux démontrer dans les ensembles, en n'utilisant que les règles de la logique intuitioniste — c'est-à-dire sans le *tiers-exclu* — et bien ce théorème est automatiquement valable dans tout topos, aussi compliqué soit-il.

C'est la loi du *tiers-exclu* qui vidait de son sens l'axiome de *Kock–Lawvere* ! Mais elle n'est plus valable dans un topos. Et bien je vais maintenant tenter d'esquisser la construction d'un topos dans lequel il existe un anneau $\mathbf{R}$ qui satisfait à l'axiome de *Kock–Lawvere*, quand on interprète cet axiome dans la logique intrinsèque du topos.

Un anneau A, c'est comme nous l'avons vu un ensemble muni d'une addition, d'une soustraction et d'une multiplication. Mais alors, chaque fois que l'on prend un polynome à coefficients entiers, par exemple

$$p(X) = 2X^3 - 3X^2 + X - 2$$

on peut construire une fonction correspondante sur l'anneau A, fonction que je noterai encore p

$$p\colon A \longrightarrow A, \quad a \mapsto p(a) = aaa + aaa - aa - aa - aa + a - 1 - 1$$

puisque cela ne requiert que des additions, des soustractions et des multiplications. Plus généralement, si

$$p(X_1, \ldots, X_n)$$

est un polynome à plusieurs variables à coefficients entiers, on obtient une fonction correspondante

$$p\colon A^n \longrightarrow A; \quad (a_1, \ldots, a_n) \mapsto p(a_1, \ldots, a_n)$$

sur l'anneau.

Et bien de manière analogue, une $\mathcal{C}^\infty$-*algèbre* est par définition un ensemble A muni, pour toute fonction infiniment dérivable

$$f\colon \mathbb{R}^n \longrightarrow \mathbb{R}, \quad (r_1, \ldots, r_n) \mapsto f(r_1, \ldots, r_n)$$

d'une opération correspondante, que pour la facilité on continue à noter de la même manière

$$f\colon A^n \longrightarrow A, \quad (a_1, \ldots, a_n) \mapsto f(a_1, \ldots, a_n).$$

Les axiomes entre toutes ces opérations définies sur A sont toutes les égalités valides entre les fonctions correspondantes sur les réels.

Voyons un exemple concret, montrant tout l'intérêt de cette notion en géométrie : toute surface S au sens usuel donne lieu à une $\mathcal{C}^\infty$ algèbre, à savoir, l'ensemble des fonctions infiniment dérivables de S dans $\mathbb{R}$

$$A = \mathcal{C}^\infty(S, \mathbb{R}).$$

C'est bien une $\mathcal{C}^\infty$-algèbre, car toute fonction infiniment dérivable

$$f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$$

induit par composition une opération

$$\mathcal{C}^\infty(S,\mathbb{R})^n \longrightarrow \mathcal{C}^\infty(S,\mathbb{R}), \quad (h_1,\ldots,h_n) \mapsto f(h_1,\ldots,h_n)$$

où bien entendu

$$f(h_1,\ldots,h_n)\colon S \longrightarrow \mathbb{R}, \quad x \mapsto f\big(h_1(x),\ldots,h_n(x)\big).$$

L'argument qui précède n'a évidemment rien à voir avec le fait que S soit de dimension 2. Vous pouvez tout aussi bien prendre une *variété différentiable* S de dimension quelconque. Ou encore remplacer S par $\mathbb{R}$, tout simplement.

Parmi toutes les $\mathcal{C}^\infty$-algèbres, il y a celles de *présentation finie* : celles que l'on peut décrire au moyen d'un nombre fini de générateurs et un nombre fini d'équations. Qu'est-ce que cela veut dire ?

Faisons le parallèle avec les anneaux. Je dirai que trois éléments a, b, c d'un anneau A engendrent cet anneau, si tout élément de A peut se reconstruire à partir de a, b, c au moyen des opérations d'anneau, c'est-à-dire par additions, soustractions et multiplications. Par exemple si

$$A = \mathbb{R}[X,Y,Z]$$

est l'anneau des polynomes à trois variables X, Y, Z, les trois polynomes

$$q(X,Y,Z) = X, \quad r(X,Y,Z) = Y, \quad s(X,Y,Z) = Z$$

engendrent l'anneau A. Car par exemple le polynome

$$p(X,Y,Z) = 2XY^3 - 3XYZ$$

peut s'écrire

$$p(X,Y,Z) = XYYY + XYYY - XYZ - XYZ - XYZ$$

c'est-à-dire une combinaison des trois polynomes q, r, s au moyen des opérations d'anneau. Et bien sûr après cela on peut forcer des équations, comme nous l'avons fait tout-à-l'heure, pour l'équation $X^2 = 0$.

Je fais donc de même avec les $\mathcal{C}^\infty$-algèbres : je m'intéresse à celles que je peux reconstruire entièrement à partir d'un nombre fini de générateurs et un nombre fini d'équations, le tout au moyen des opérations de $\mathcal{C}^\infty$-algèbres.

Allez, juste une phrase à l'intention de ceux qui connaissent bien les topos. En notant $\mathcal{A}$ la catégorie des $\mathcal{C}^\infty$-algèbres de présentation finie, on définit sur la catégorie duale $\mathcal{A}^{\mathsf{op}}$ une topologie au sens de Grothendieck. Et l'on choisit comme univers de travail, le topos de faisceaux correspondant.

Mon Dieu, qu'est-ce que tout cela peut bien vouloir dire ! Et bien cela signifie que je vais travailler dans un univers $\mathbf{E}$ non plus d'ensembles rigides, mais de familles d'ensembles $X = (X_A)_{A\in\mathcal{A}}$ qui varient en fonction d'un paramètre A. Et cette fois l'ensemble $\mathcal{A}$ des paramètres n'est plus un ensemble d'ouverts comme tout-à-l'heure, mais est l'ensemble de toutes les $\mathcal{C}^\infty$-algèbres de présentation finie. Un "ensemble" de mon nouvel univers $\mathbf{E}$, c'est maintenant une famille d'ensembles indexée par toutes les $\mathcal{C}^\infty$-algèbres de présentation finie. Et bien sûr des propriétés de restriction et de recollement entre tout cela, propriétés de "faisceau" dont je vous fais grâce. En d'autres termes, ce sont les $\mathcal{C}^\infty$-algèbres de présentation finie qui jouent maintenant le rôle que les ouverts du plan jouaient dans notre exemple du faisceau des surfaces. Ce nouvel univers $\mathbf{E}$ est ce que l'on appelle un *topos de Grothendieck*, dont les objets sont encore appelés des *faisceaux*.

Parmi tous ces "faisceaux" il y a celui on ne peut plus banal défini par

$$\mathbf{R} = (X_A)_{A\in\mathcal{A}}, \quad X_A = A \text{ pour tout } A \in \mathcal{A}.$$

C'est un anneau car l'addition, la soustraction et la multiplication

$$+,\ -,\ \times\colon \mathbb{R}\times\mathbb{R} \longrightarrow \mathbb{R}$$

sont des fonctions infiniment dérivables, donc sont des opérations

$$+,\ -,\ \times\colon A\times A \longrightarrow A$$

de chaque $\mathcal{C}^\infty$-algèbre.

Intéressons-nous alors aux éléments de carré nul de cet anneau $\mathbf{R}$.

$$\mathbf{D} = \{r \in \mathbf{R} | r^2 = 0\}.$$

Les calculer n'est pas immédiat. Il faut pour cela observer que pour toute $\mathcal{C}^\infty$-algèbre A, on a toujours l'isomorphisme

$$A \cong \mathsf{Hom}\big(\mathcal{C}^\infty(\mathbb{R},\mathbb{R}), A\big)$$

où Hom désigne l'ensemble des homomorphismes de $\mathcal{C}^\infty$-algèbres. Cet isomorphisme est immédiat à établir :

- étant donné $f\colon \mathcal{C}^\infty(\mathbb{R},\mathbb{R}) \longrightarrow A$, on lui associe $f(\mathsf{id}_\mathbb{R}) \in A$;
- réciproquement étant donné $a \in A$, on lui associe l'homomorphisme

$$\mathcal{C}^\infty(\mathbb{R},\mathbb{R}) \longrightarrow A, \quad h \mapsto h(a);$$

ce qui a un sens puisque h est l'une des opérations qui existent sur A. Les algébristes auront reconnu que je viens de montrer que $\mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ est la $\mathcal{C}^\infty$-algèbre libre à un générateur. Mais peu importe.

Nous savons maintenant que notre anneau $\mathbf{R}$ peut de manière équivalente s'écrire

$$\mathbf{R} = \Big(\mathsf{Hom}\big(\mathcal{C}^\infty(\mathbb{R},\mathbb{R}), A\big)\Big)_{A\in\mathcal{A}}.$$

Reposons-nous alors la question : quand est-ce que l'élément $a \in A$ correspondant à l'homomorphisme

$$f\colon \mathcal{C}^\infty(\mathbb{R},\mathbb{R}) \longrightarrow A, \quad h \mapsto h(a)$$

est de carré nul ? Et bien pour le découvrir considérons la fonction identité

$$\mathsf{id}_\mathbb{R}\colon \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto x$$

et son carré

$$\mathsf{id}_\mathbb{R}^2\colon \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto x^2.$$

Cette fonction "élever au carré" est infiniment dérivable, donc est en particulier une opération de la théorie des $\mathcal{C}^\infty$-algèbres. Et puisque f est un homomorphisme, il respecte toutes les opérations de la théorie. Donc

$$f(\mathsf{id}_\mathbb{R}^2) = \big(f(\mathsf{id}_\mathbb{R})\big)^2 = a^2.$$

Il en résulte aussitôt que

$$a^2 = 0 \ \text{ si et seulement si } \ f(\mathsf{id}_\mathbb{R}^2) = 0.$$

Et bien sûr, des homomorphismes non nuls mais qui envoyent cependant la fonction x^2 sur 0, il y en a en général beaucoup. Dans un autre contexte, nous avons observé tout-à-l'heure que par exemple l'application

$$\delta\colon \mathcal{C}^\infty(\mathbb{R},\mathbb{R}) \longrightarrow \mathbb{R}, \quad g \mapsto \frac{dg}{dt}(0)$$

est telle que

$$\delta(t^2) = \frac{dt^2}{dt}(0) = 0.$$

Pour la petite histoire, j'ajouterai que la théorie des quotients algébriques nous permet de conclure que f est de carré nul si et seulement si il se factorise à travers le quotient

$$\mathcal{C}^\infty(\mathbb{R},\mathbb{R}) \longrightarrow\!\!\!\!\!\rightarrow \mathcal{C}^\infty(\mathbb{R},\mathbb{R})/\langle x^2 = 0\rangle,$$

quotient au même titre que celui que nous avons considéré tout-à-l'heure pour les polynomes. Il en résulte que le faisceau des éléments infiniment petits de l'anneau $\mathbf{R}$

$$\mathbf{D} = \{r \in \mathbf{R} | r^2 = 0\} \subseteq \mathbf{R}$$

peut équivalemment s'écrire

$$\mathbf{D} = \Big(\mathsf{Hom}\big(\ \mathcal{C}^\infty(\mathbb{R},\mathbb{R})/\langle x^2 = 0\rangle\ ,\ A\ \big)\Big)_{A\in\mathcal{A}}.$$

Ainsi le quotient tout-à-fait classique de l'anneau $\mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ par l'équation $x^2 = 0$ est en fin de compte l'opération classique qui fournit, dans la logique intrinsèque de notre topos $\mathbf{E}$ de faisceaux, les éléments de carré nul de l'anneau $\mathbf{R}$, les éléments que nous avons choisi d'appeler *infiniment petits*.

Et bien sachez que cet anneau $\mathbf{R}$ que je viens de décrire satisfait à l'axiome de *Kock–Lawvere*, pour autant bien sûr que l'on interprète cet axiome dans la logique intrinsèque du topos $\mathbf{E}$ dont j'ai esquissé la construction. Cela demande évidemment un peu de travail à le démontrer. Mais puisque les topos sont des modèles intuitionistes de la théorie des ensembles, cela prouve en tout cas qu'en théorie intuitioniste des ensembles, il existe bel et bien des anneaux satisfaisant à l'axiome de *Kock–Lawvere*. Nous venons précisément d'en construire un !

Mais il y a bien mieux ! En présence de l'axiome de *Kock–Lawvere*, on peut développer toute la géométrie différentielle d'une manière on ne peut plus intuitive. Une *surface* est maintenant définie comme un objet qui — localement — est isomorphe à $\mathbf{D}^2$. La signification exacte de cette phrasse doit bien entendu être précisée, mais elle force en tout cas le fait qu'une surface est tout simplement un objet mathématique qui, à distance infinitésimale, est isomorphe à un morceau infinitésimal du plan $\mathbf{R}^2$. Et bien sûr cette approche se généralise aussitôt aux variétés différentiables de dimension quelconque, autre que 2. On constate alors que la plupart des résultats de géométrie différentielle classique restent valables dans ce nouveau contexte, mais cette fois, avec des preuves en général particulièrement simples et intuitives, en termes d'éléments infiniment petits.

Très bien me direz-vous : s'il y a des gens que cela amuse de s'intéresser à cette nouvelle forme de géométrie, pourquoi pas. Cette géométrie est peut-être plus simple et plus intuitive, mais moi ce qui m'intéresse, c'est d'étudier les vraies surfaces de tout le monde, pas ces concepts bizarres sortis tout droit de quelques élucubrations surréalistes. Et bien vous avez tort ! Car tous les résultats que l'on peut démontrer en théorie intuitioniste des ensembles — c'est-à-dire sans utiliser le *tiers-exclu* — tous les résultats donc que l'on peut démontrer ainsi à partir de l'axiome de *Kock–Lawvere* sont *de facto* valables en géométrie différentielle classique. La géométrie différentielle synthétique, c'est *aussi* une technique de démonstration nouvelle, particulièrement intuitive, pour établir des théorèmes de géométrie différentielle classique.

Pourquoi cela ? Et bien parce qu'il existe un plongement des "bonnes" (= paracompactes, de classe $\mathcal{C}^\infty$) surfaces de tout le monde dans les surfaces au sens de *Kock–Lawvere*, considérées dans le topos $\mathbf{E}$ de faisceaux que j'ai esquissé ci-avant. Et quand je dis "surface", cela peut bien entendu être une variété de dimension quelconque, pas nécessairement 2. Rappelons-nous que pour une surface S,

$$\mathcal{C}^\infty(S, \mathbb{R})$$

est une $\mathcal{C}^\infty$-algèbre. Le plongement en cause de la surface usuelle S dans notre topos $\mathbf{E}$ est alors donné par le faisceau

$$\varphi(S) = \Big(\mathsf{Hom}\big(\ \mathcal{C}^\infty(S, \mathbb{R})\ ,\ A\ \big)\Big)_{A \in \mathcal{A}}$$

où Hom désigne à nouveau l'ensemble des homomorphismes de $\mathcal{C}^\infty$-algèbres. Et on vérifie que dans notre topos $\mathbf{E}$, ce faisceau $\varphi(S)$ est bien une surface au sens de la géométrie de *Kock–Lawvere*. Or il se trouve que le plongement φ préserve et reflète toutes les notions habituelles concernant les surfaces. Mais alors, quand un théorème concernant ces notions peut être prouvé en théorie intuitioniste des ensembles à partir de l'axiome de *Kock–Lawvere*, ce théorème est en particulier vrai dans tout topos, donc est vrai en particulier dans notre topos $\mathbf{E}$, pour l'anneau $\mathbf{R}$ considéré plus haut, et pour la surface $\varphi(S)$. Et via puisque le plongement φ reflète les propriétés des surfaces, le théorème se trouve démontré pour les surfaces usuelles.

Donc conclusion : en changeant de logique, simplement en excluant le *tiers-exclu* des raisonnements — mais tout en continuant à travailler dans les ensembles, j'insiste — on peut en toute rigueur utiliser des arguments en termes d'infiniment petits pour démontrer des théorèmes tout-à-fait classiques, à propos des surfaces les plus classiques qui soient, des surfaces qui vivent dans l'univers mathématique usuel où les infiniment petits n'ont pas droit de cité.

Références

[1] **Francis Borceux**, *Handbook of categorical algebra **3** : Categories of sheaves*, Encyclopedia of Mathematics and its Applications **53**, Cambridge University Press, 1994

[2] **Eduardo Dubuc**, *Sur les modèles de la géométrie différentielle synthétique*, Cahiers de Topologie et Géométrie différentielle **20**, 1979, 231–279

[3] **Charles Ehresmann**, *Les prolongements d'une variété différentiable*, Comptes Rendus de l'Académie des Sciences de Paris **233**, 1951, 598–600.

[4] **Anders Kock**, *Synthetic differential geometry*, London Mathematical Society Lecture Notes Series **51**, Cambridge University Press, 1951. Second edition, Cambridge University Press, 1996

[5] **René Lavendhomme**, *Basic Concepts of Synthetic Differential Geometry,* Kluwer Texts in the Mathematical Sciences **13**, Kluwer Academic Publishers, Dordrecht, 1996, xvi + 320pp.

[6] **René Lavendhomme**, *Lieux du Sujet*, collection "Champ Freudien", Seuil, Paris, 2001, 356 pp.

[7] **F.William Lawvere**, *Outline of Synthetic Differential Geometry*, 14 pp. manuscrit non publié, 1998, disponible à l'adresse `http://www.acsu.buffalo.edu/~wlawvere/downloadlist.html`

[8] **Ieke Moerdijk and Gonzalo Reyes**, *Models for smooth infinitesimal analysis*, Springer, 1991

[9] **André Weil**, *Théorie des points proches sur les variéts différentiables*, Conférence de Géométrie Différentielle, Strasbourg, 1953. Réédité dans les *Oeuvres complètes*, Volume **II**, Springer, 1979, 103–109

francis.borceux@uclouvain.be