Voilà trois ans que vous faîtes des mathématiques, vous avez été, lors de cette dernière année, au contact de structures abstraites (espaces affines, topologiques, sous-groupes distingués, …) que vous avez pu assimiler mais

Et après, on fait quoi ?

Version 1

Par GALLOIS Dimitri.

Préface de l'auteur

Ce texte est destiné aux étudiants sortant de L3 et plus particulièrement aux étudiants en Maths Fondamentales de Reims, qui constituent en fait ma classe. Ce texte est à évocation culturelle, vous ne trouverez nullement de définitions formelles, ni même une introduction à une quelconque théorie et donc je risque de fort décevoir les étudiants de Master. J'espère que ce document amusera les étudiants et y trouveront le même plaisir que j'ai quand j'ouvre mes bouquins, un plaisir comparable à celui de voir un film de science-fiction ; un monde étrange, nouveau, aux allures si éloigné du nôtre, et pourtant…c'est le nôtre.

Cependant, je fournirai après chaque explication des notions vues, une légère bibliographie ouvrant à l'introduction des notions. Je pense que la plupart de ces livres sont à la bibliothèque universitaire du campus Moulin de la Housse, Reims. Si, dans le cas le plus malheureux, ces livres ne se trouvent pas dans la bibliothèque concernée, et en état de conscience du prix d'un livre (qu'au final nous ne sommes pas sûr de lire), j'accepte volontiers de prêter ceux de ma collection (on prendra juste garde au fait que tous n'y sont pas). De plus ceux qui n'ont pas trop envie de faire des maths, mais plus d'améliorer leur culture, je conseille à la fin du texte une liste d'ouvrage aisé à lire (même pour des premières S pour certains).

Si, ne serait-ce qu'un étudiant, à la suite de cela, accepte de lire un des ouvrages que j'aurais conseillé, je considérerais cela comme une victoire pour moi ; car c'est très méconnu, mais les mathématiques sont passionnantes, elles sont pleine d'une histoire, parfois allant jusqu'à la légende, avec ces héros, ces débauches et ses succès incroyables. Et j'espère par là que cet étudiant poursuivra sa culture mathématique, non pas forcément avec rigueur, mais au moins avec intérêt, et qu'à son tour il la transmettra à ses élèves.

Ensemble, vivons-les.

Sommaire :

P.S. : Vous pouvez sauter une partie si elle ne vous inspire pas, ou si elle vous semble devenir trop complexe, mais certaines utilisent les précédentes.

La forme de l'univers.

Que vous ayez suivi MA504 ou MA505, vous n'avez vu qu'une infime partie de ce que représente le mot « Topologie » pour un mathématicien. En effet les premières topologies que vous ayez pu voir sont celles des espaces vectoriels normés, ou celle des espaces métriques. Mais il existe une théorie plus générale encore, englobant ces deux notions (elle était au programme l'année dernière) : la topologie générale.

En cours nous avons pu voir que les ouverts d'un espace métrique vérifiaient trois propriétés :

Une union quelconque d'ouverts est un ouvert ;

Une intersection FINIE d'ouverts est un ouvert ;

L'ensemble vide, et l'ensemble entier sont des ouverts.

Alors on a généralisé la notion d'espace métrique en espace topologique en transformant ces propriétés en axiomes :

On appelle espace topologique, un couple (X, T) où T est un ensemble de partie de X stable par union, par intersection finie et contenant X et $\emptyset$.

Schématiquement, on représente mentalement un ouvert comme une « boule » (légèrement tordue) privée de son contour (et on symbolise ça en mettant des pointillés). M.Bounzouina a alors démontré que tout espace métrique est un espace topologique. Globalement la topologie ignore les formes des objets, elle étudie juste « ce qui est au bord » et « ce qui est bel et bien à l'intérieur », et donc en définissant quelque chose « qui rapetisse » on vient l'air de rien d'écrire une notion de convergence. Car c'est ce qui se passe dans le cas très concret de l'espace des nombres réels (qui est un espace topologique car métrique, et c'est même un evn) ; on choisit quelque chose de plus en plus petit, défini grâce à la distance par une boule de rayon ε… (Voir la notion de filtre).

De même la topologie ignorant les formes voit les transformations fondamentales être des « transformations continues », et on appelle fonction continue une fonction dont la réciproque d'un ouvert est un ouvert. Visuellement, les fonctions continues ne peuvent pas couper en deux quelques choses puisqu'elles doivent laisser le « bord » intact. Je m'explique, une fonction continue transforme quelque chose en un morceau, en quelque chose en un morceau, pas deux. Ces morceaux (qu'on appelle composante connexe) sont très important pour le chapitre 9.

Voici un exercice intéressant à chercher pour ceux qui veulent :

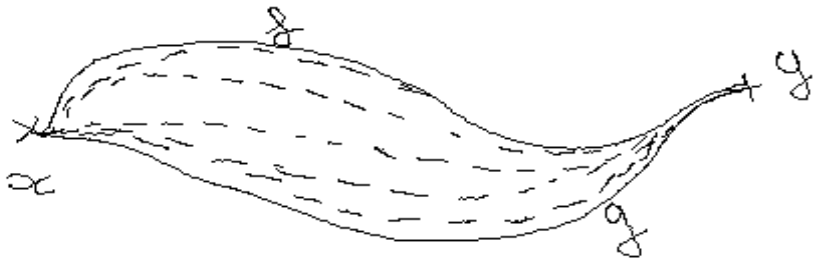
On appelle groupe topologique, un groupe munie d'une topologie rendant continue les applications produit $(x, y) \to xy$ et inverse $x \to x^{-1}$, montrer que la composante connexe de l'élément neutre est un sous-groupe du groupe.

Peu importe l'espace topologique, un grand intérêt des mathématiques est de travailler localement, car bien souvent on ne connait pas la forme générale d'un objet (comme une fonction vectorielle par exemple), et alors on travaille localement, c'est-à-dire près d'un point. On peut alors définir des notions comme localement compact, localement connexe, etc. Et alors $\mathbb{R}$ bien qu'il ne soit pas compact, il devient localement compact. Définir localement aura un grand intérêt par la suite (cf. Chapitre 9).

Au-delà des définitions locales, il existe aussi des définitions globales (compact, connexe, etc.) que vous avez déjà rencontré, disons que (globalement) connexe signifie « en un seul morceau » et compact signifie qu'une infinité (en fait dénombrable) d'éléments finissent par s'agglutiner autour d'un point (Théorème de Bolzano-Weiestrass). Nous allons maintenant introduire une notion fondamentale : la simple connexité. Vulgairement, une surface simplement connexe est une surface sans trou, comme $\mathbb{R}$. La simple connexité est invariante par les applications continues (car « les applications continues conservent les bords »).

Plus formellement, soient x et y deux points de X, X étant un espace topologique. On appelle chemin de x à y une application continue de $[0,1]$ dans X telle que $f(0) = x$ et $f(1) = y$.

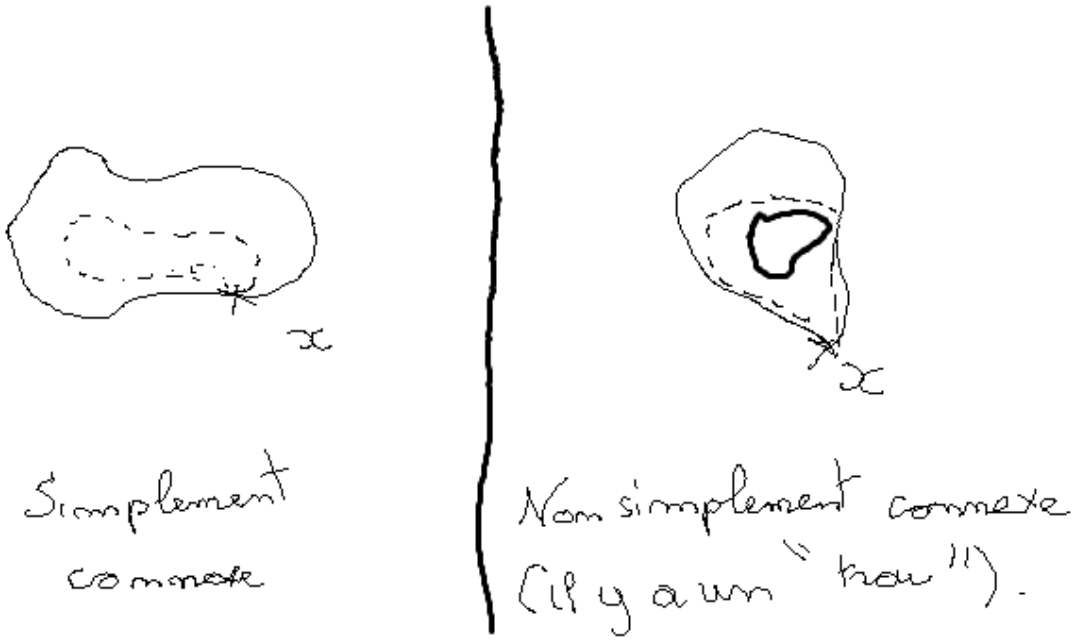
Pour deux chemins f et g, on dit que f et g sont homotopes si et seulement si on peut transformer de manière continue f en g. Par exemple, les deux chemins suivants sont homotopes :



Excusez l'horreur du dessin, mais j'ai utilisé une tablette graphique, et je ne suis pas très doué avec.

On peut voir sur le dessin la transformation de f à g.

Maintenant qu'on a l'homotopie, on dira qu'une surface est simplement connexe, si pour chaque point x, et pour chaque chemin f de x dans x, f est homotope à la fonction constante de valeur x.



On voit qu'à gauche le chemin se rétracte, et qu'à droite c'est impossible, c'est comme ça qu'on interprète l'existence d'un trou.

Par exemple : la sphère est simplement connexe contrairement au tore (un donut).

Allons plus loin : Soit f un chemin de x à y, et g un chemin de y à z, alors il existe un chemin de x à z et son expression est :

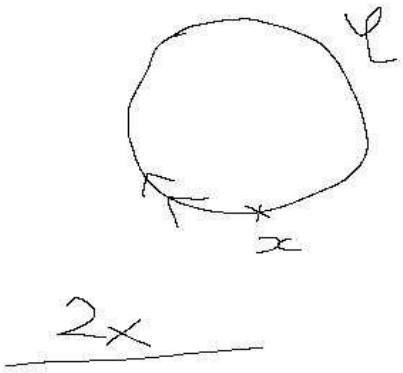
$$h: t \to \begin{cases} f(2t) \text{ si } t \in \left[0; \frac{1}{2}\right] \\ g(2t-1) \text{ si } t \in \left[\frac{1}{2}; 1\right] \end{cases}$$

(Ça on l'a vu en TD avec Mme Bruyant).

On a l'existence d'un chemin inverse : $t \to f(-t)$; d'un « espèce d'élément neutre (même s'il dépend des extrémités considérés) et on a l'associativité des chemins. Une telle structure n'est pas un groupe, la composition n'est même pas forcément définie, mais les similitudes avec les axiomes de groupes forcent les mathématiciens à appeler ça « groupoïde ». En réalité on travaille plutôt sur les classes d'équivalence construite par l'homotopie ; on appelle ce groupoïde, le groupoïde de Poincaré.

Lorsque nous nous restreignons aux chemins d'origine et d'extrémité commune (appelé alors lacet) le groupoïde devient un groupe. Appelé donc groupe de Poincaré. Une surface est simplement connexe si par définition ; en chaque point, le groupe de Poincaré est le groupe trivial réduit à son élément neutre.

Considérons le cercle, on voit clairement qu'il n'est pas simplement connexe (il y a un trou énorme !). Calculons son groupe fondamental. Soit x un point du cercle. Alors les lacets feront toujours le cercle, on pourrait croire qu'il n'y a que le lacet allant dans un sens, et celui dans l'autre mais non : le nombre de fois que le lacet fait le tour compte aussi. En effet la composé de ce chemin par lui-même ça compte aussi. On a donc le groupe engendré par un élément, et à chaque composition on forme un nouveau lacet, il est donc monogène infini. En gros c'est le groupe $\mathbb{Z}$ (à isomorphisme près).



Et l'opération sur $\mathbb{Z}$ est claire : si on effectue le tour une fois dans un sens, et deux fois dans l'autre, on vient d'effectuer l'opération $1 - 2 = -1$ qui veut dire qu'on vient de faire un tour en sens inverse.

Cette branche des mathématiques s'appelle la topologie algébrique (elle étudie la topologie à l'aide de l'algèbre), son point culminant est la conjecture de Poincaré ; un des sept problèmes du millénaire.

Note 1 : Les problèmes du millénaire sont sept problèmes qui ont été posés en 2000 par le Clay Mathematic Institute (dont le directeur est très sympa, et il parle très bien français) en analogie avec David Hilbert qui a posé en 1900 vingt-trois problèmes donnant un état général de la connaissance mathématique de l'époque. Evidemment certains problèmes sont restés entre les deux (comme l'hypothèse de Riemann). Chaque problème est récompensé du prix Clay, d'un million de dollars ; pour l'instant un seul à été résolu, la conjecture de Poincaré (précédemment cité) en 2002, qui concerne la simple connexité de certaines surfaces de dimension 3, l'énoncé était de montrer que ce serait toujours homéomorphe à la sphère.

Note 2 : Henri Poincaré (cousin du président Raymond) était un mathématicien du XIXème siècle, et sûrement le dernier à avoir eu une vue globale sur les mathématiques, car depuis elles se sont spécifiées, elles se sont diversifiées et complexifiées, et toutes les connaitre est devenu impossible…

Les sources pour poursuivre :

- Topologie et analyse pour la 3^{ième} année (Georges Skandalis)
- Eléments de topologie algébrique (Claude Godbillon) ; on remarquera une erreur légère sur la tranche, algébrique est muni d'un « s », erreur ou façon de voir les choses ? Le livre demande des connaissances déjà avancées en topologie, le seul cours de MA50X ne suffira pas (c'est pour ça que je préconise le livre de G.Skandalis, très bien fait, avec de beaux dessins pour comprendre et tout ça).

Quand l'infini devient un point

Cette section va étudier des notions de géométrie, elle est totalement indépendante de la précédente (et ce sera le cas pour quasiment toutes les sections en fait), plus précisément la notion d'espace projectif ; qui est resté pendant longtemps au programme de l'agrégation, mais Mme Barka m'a annoncé au rdv du tutorat que ça venait juste d'être retiré pour être remplacée par les distributions. Les prérequis sont plutôt faible par rapport à la section précédente : juste les connaissances d'algèbre linéaire contenu dans le programme de 1ère année, et comprendre ce que signifie « quotienter » par une relation d'équivalence.

Dans le pire des cas, en voici un rappel : Soit $\mathcal{R}$ une relation d'équivalence sur un ensemble E, on appelle classe de x l'ensemble des éléments équivalents à x, les classes forment une partition de E (leur union forme E, elles sont disjointes deux à deux, et toutes non vide) et on appelle quotient de E par $\mathcal{R}$ l'ensemble des classes de E pour la relation $\mathcal{R}$.

Globalement un espace projectif c'est définir comme des points les droites d'un espace vectoriel ou plus formellement : Soit E un espace vectoriel, on note E' l'espace $E\backslash\{0\}$; on définit une relation d'équivalence sur E' comme suit : deux vecteurs u et v sont équivalents si et seulement s'ils sont colinéaires, on note $\sim$ cette relation. Alors on appelle espace projectif (associé à E) l'espace quotient de E' par $\sim$ et on le note $P(E)$.

C'est-à-dire que les points à l'infini ont été ramenés « sur la sphère unité » et que dans une telle optique deux parallèles c'est deux droites qui se croisent à l'infini. Et d'ailleurs si E est de dimension finie $n+1$ alors $P(E)$ sera de dimension n, car on a contracté les droites en des points.

Plus formellement, on a une projection π de E sur $P(E)$, c'est la surjection canonique de E dans $P(E)$, on appelle alors sous-espace projectif de $P(E)$ tout ensemble de la forme $\pi(F\backslash\{0\})$ où F est un sous-espace vectoriel de E. Comme π vérifie $\pi(\bigcap_i F_i\backslash\{0\}) = \bigcap_i \pi(F_i\backslash\{0\})$ l'intersection de sous-espaces projectifs est un sous-espace projectif, on peut alors définir l'espace projectif engendré par une partie S de $P(E)$, c'est l'intersection de tous les sous-espaces projectifs contenant S et on le note $Proj(S)$.

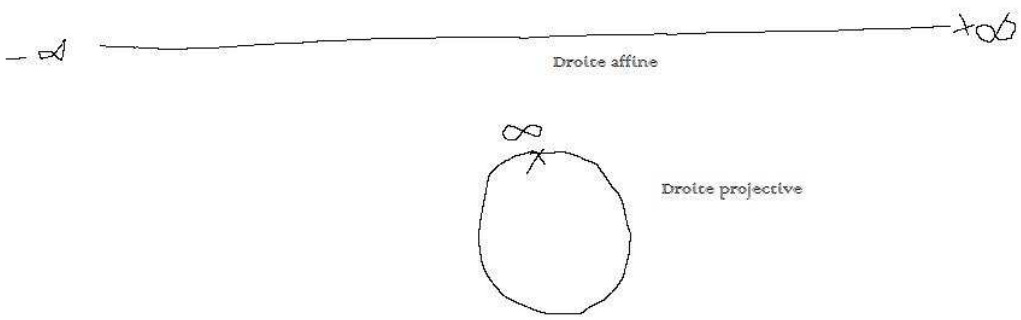
On introduit alors la notion de coordonnées homogènes : soit $(x_0, x_1, \dots, x_n)$ un vecteur x de $\mathbb{K}^n$, alors $(x_0, x_1, \dots, x_n)$ est appelé un système de coordonnées homogène de $\pi(x)$. On remarque que si $(x_0, x_1, \dots, x_n)$ est un système de coordonnées homogènes d'un point p de $P^n(\mathbb{K}) = P(\mathbb{K}^{n+1})$ alors le système $(\lambda x_0, \lambda x_1, \dots, \lambda x_n)$ en est un autre.

Si on considère l'espace E de dimension finie $n+1$ munie d'une base, alors on a $P(E) = P^n(\mathbb{K})$ avec un système de coordonnées homogènes $(x_0, x_1, \dots, x_n)$. On considère maintenant un hyperplan H (on suppose pour simplifier que l'équation de H est de la forme $x_0 = 0$, on peut toujours revenir à ce cas là) et on note $\overline{H}$ son image par π. L'application f de $A = P^n(\mathbb{K})\backslash\overline{H}$ qui a tout point de A de coordonnées homogènes $(x_0, x_1, \dots, x_n)$ associe $\left({x_1}/{x_0}, {x_2}/{x_0}, \dots, {x_n}/{x_0}\right)$ est une bijection (exercice pour ceux qui veulent) et si on veut on a même sa réciproque : $f^{-1} : (y_1, y_2, \dots, y_n) \to (1, y_1, y_2, \dots, y_n)$.

Pour conclure, on vient de plonger $P^n(\mathbb{K})\backslash\overline{H}$ dans un espace affine de dimension n, on peut dire que $P^n(\mathbb{K})$ est « grosso-modo » la réunion d'un espace affine (de dimension n) et d'un hyperplan (de $P^n(\mathbb{K})$, donc de dimension $n-1$). Les points de l'espace affine seront dit « à distance finie » et ceux de l'hyperplan « à l'infini ». Et ainsi on a une nouvelle vision des points à l'infini.

Pour clarifier les choses, on va voir des exemples particuliers. Commençons par la droite projective :

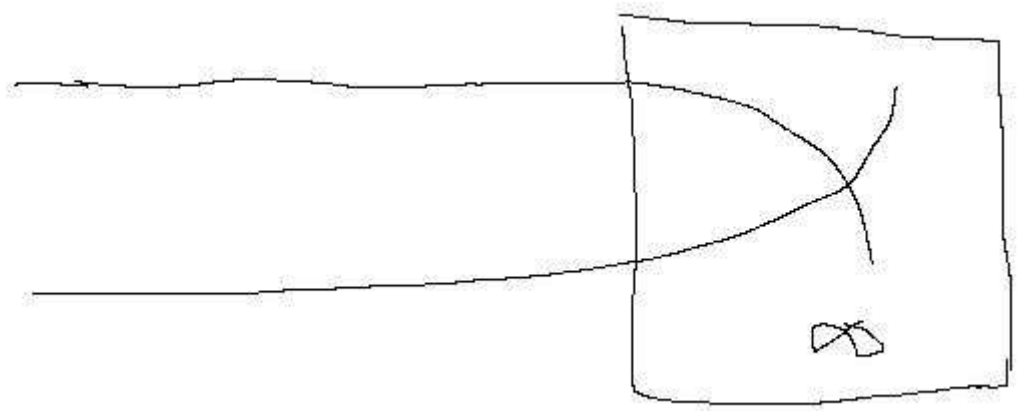
La droite projective est issue de $\mathbb{K}^2$, notons (x,y) les coordonnées des points de $\mathbb{K}^2$, choisissons comme « hyperplan » de $\mathbb{K}^2$ celui d'équation $y = 0$. Tous les points de l'hyperplan sont de la forme $(x,0)$, donc tous colinéaires entre eux, donc tous équivalents, ils correspondent donc à un seul et même point dans la droite projective, dont on peut prendre comme système de coordonnées homogènes (1,0). Les autres vecteurs de $\mathbb{K}^2$ de coordonnées (x,y) ont y non nul (car ils ne sont pas dans l'hyperplan d'équation $y = 0$) donc on peut choisir pour les points de la droite projective $(x/y\,,1) = (t\,,1)$ comme système de coordonnées homogènes. Les points $(t\,,1)$ s'identifie à une droite affine classique, et on a donc adjoint un point à l'infini $\infty = (1,0)$. De manière générale les espaces projectifs sont compacts et connexes, mais dans le cas de la droite projective, l'allure est triviale :



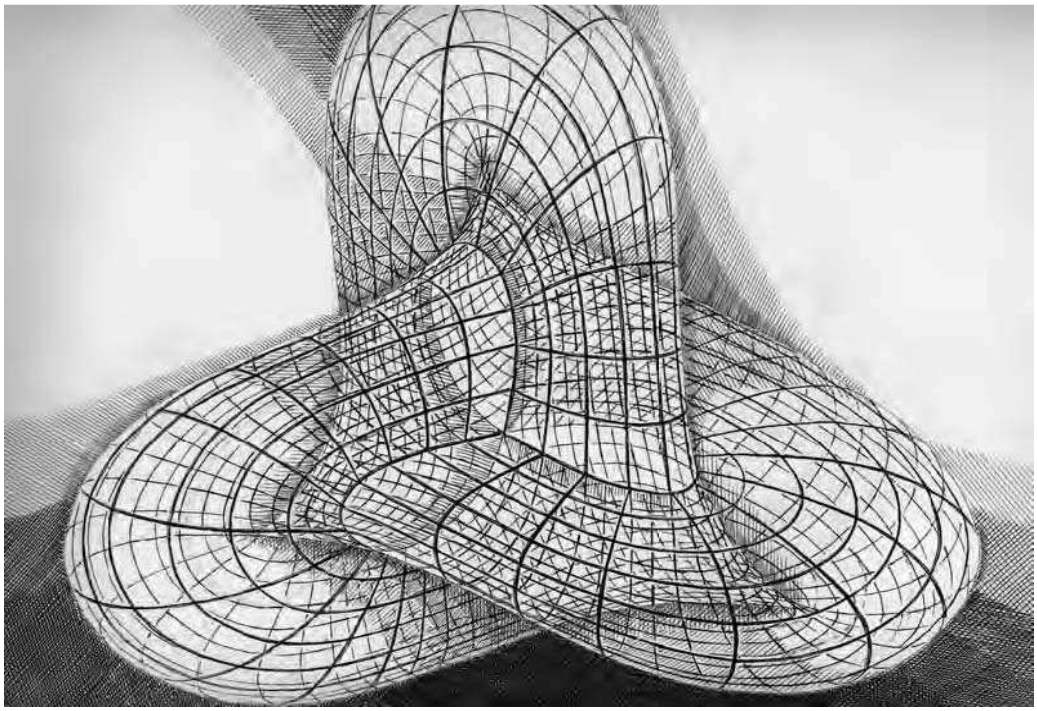
Ce passage de $\mathbb{R}$ à $P^1(\mathbb{R})$ s'appelle la compactification d'Alexandroff de $\mathbb{R}$. On peut assimiler l'espace projectif à un cercle (dans le cas où $\mathbb{K} = \mathbb{R}$) ou à une sphère (dans le cas où $\mathbb{K} = \mathbb{C}$).

Allons plus loin : le plan projectif. On commence de la même manière que la droite projective : on considère les coordonnées (x,y,z) de $\mathbb{K}^3$, et on choisit comme hyperplan H celui d'équation $z = 0$. Dans ce cas H est l'ensemble des points $(x,y,0)$ et en passant en projectif, on obtient une droite projective, D. Chaque vecteur de $\mathbb{K}^3$ en dehors de H voit $z \neq 0$, donc chaque point du plan projectif en dehors de D admet $(x/z, y/z, 1) = (t,s,1)$ comme système de coordonnées homogènes (qui s'identifie à $\mathbb{K}^2$). Le plan projectif appelle à regarder ses droites. On note $ax + by + cz = 0$ un plan de $\mathbb{K}^3$ (donc une droite du plan projectif), on a a,b,c non tous nuls (sinon c'est l'espace $\mathbb{K}^3$). On distingue deux cas : le premier avec $a = 0$, $b = 0$, donc c non nul, et alors on a comme équation $z = 0$. Et en « projectant » on obtient la droite D. Dans le deuxième cas, on a une droite $ax + by + cz = 0$. On peut calculer ses points à l'infini (i.e. l'intersection de son projeté avec D), à l'infini on a $z = 0$, ce sont donc des points de la forme $(x,y,0)$ et l'équation devient $ax + by = 0$, admettant une droite vectorielle comme solution, donc un unique point dans

l'espace projectif, et par exemple $(b\,,-a\,,0)$ sont ses coordonnées homogènes, on appellera direction de la droite ce point à l'infini ; si on regarde ses points à distance finie, on a $(x,y,1)$ comme coordonnées homogènes, soit l'équation $ax + by + c = 0$ (et on y voit bien une droite affine !). De plus si deux droites sont parallèles ($ax + by + c = 0$ et $ax + by + c' = 0$ en affine) dans le plan projectif, on a comme direction $(b,-a\,,0)$ pour les DEUX droites, autrement dit, elles admettent un unique point d'intersection à l'infini. On a enfin compris la notion de « croisement à l'infini ».



Un grand intérêt aussi du plan projectif c'est l'utilisation des coniques, mais je ne détaillerai pas cette notion. On peut représenter graphiquement un plan projectif (avec par exemple une surface de Boy, il y a d'autres modèles mais je ne connais que celui là). Voici la représentation graphique d'une surface de Boy :



Note 1 : Je me suis permis d'être un peu plus formel sur cette section que la précédente car premièrement j'avais peur de décourager le lecteur dès la première section, et de plus la seconde est indispensable pour l'agrégation, donc plus utile aux étudiants. On pourrait rajouter une troisième raison, c'est plus concret ; mais en mon sens, je trouve la topologie algébrique plus représentative quand on voit une surface de Boy (enfin ce n'est que mon avis).

Note 2 : J'ai galéré pour obtenir l'image de la surface de Boy car malheureusement je n'ai pas d'accès internet pour le moment. Je crois que ça a été la chose la plus difficile à faire de la section donc s'il vous plait, remontez légèrement la page pour voir l'image et applaudissez…

Lecture suggérée :

- Une introduction à la géométrie projective (Daniel Lehmann).

Par contre je n'ai pas du tout lu l'ouvrage, mais à ce que j'ai pu feuilleter l'auteur passe un temps relativement long sur le départ affine ; peut-être est ce plus concret ? Je l'ignore. En tout cas je ne conseille que cette lecture car je n'ai aucune référence en géométrie projective, à un point tel où je me demande d'où j'ai connu ça, peut être est ce Géométrie (Patrice Tauvel)…

Extensions de corps

Cette nouvelle section au programme de M1 sera vue, je pense, par la plupart des élèves, et j'espère par là introduire doucement le concept. Nous allons essentiellement étudier l'algèbre. Cette partie est indépendante des deux précédentes.

Nous allons commencer par un peu de théorie des corps. On prendra garde que contrairement au module MA602, les corps ne sont pas supposés commutatifs. Un corps est un anneau unitaire dont élément non nul est inversible. On commence par une notion fondamentale : soit A un anneau, on appelle caractéristique de A l'élément p définie par le plus petit entier non nul tel que $p.1 = 1 + 1 + \ldots + 1 = 0$. Plus formellement, l'homomorphisme de $\mathbb{Z}$ dans A défini par $k \to k.1$ admet pour noyau un idéal de $\mathbb{Z}$, donc de la forme $p\mathbb{Z}$ (c'est en fait un exercice de la feuille de TD numéro 2, mais comme nous en sommes pas encore là, je n'en parlerai pas), et c'est cet entier p que nous appelons caractéristique de A (on aura pris la convention $(0) = 0\mathbb{Z}$) et on note $p = car\, A$. On ajoutera que le sous-anneau d'un corps est toujours intègre.

Un premier théorème apparait : La caractéristique d'un corps est soit nulle, soit c'est un nombre premier.

Je donne exceptionnellement une preuve de ce résultat pour « amuser » le lecteur formel s'il s'ennuie : on considère de nouveau le noyau de l'homomorphisme ci-dessus, alors soit le noyau est (0) (injectivité) et donc la caractéristique est nulle, soit c'est un idéal de la forme $p\mathbb{Z}$ et donc l'image du morphisme est isomorphe à $\mathbb{Z}/p\mathbb{Z}$, comme l'image est intègre, $\mathbb{Z}/p\mathbb{Z}$ est intègre, donc p est premier. CQFD.

Pour aller plus loin, on appelle sous-corps premier d'un corps K donné, le plus petit sous-corps qu'il contient (i.e. l'intersection de tous ses sous-corps). Alors on a le théorème « fondamental » suivant :

Théorème : Soit K un corps, si $car\, K = 0$ alors K a son sous-corps premier isomorphe à $\mathbb{Q}$, et si $car\, K = p$ (p premier) alors le sous-corps premier de K est isomorphe à $\mathbb{Z}/p\mathbb{Z}$.

On a en corollaire que tout corps de caractéristique nulle est infini, ou par contraposée, tout corps fini à une caractéristique non nulle (donc première) ! En note sur les corps fini on a le théorème suivant :

Théorème (de Wedderburn) : Tout corps fini est commutatif.

Les préliminaires sur les corps étant faits (et ya rien de sexuel là !), on peut passer aux extensions de corps, mais avant cela quelques mots. Les extensions de corps permettent l'étude des éléments algébriques et transcendants (nous verrons ces notions plus loin) qui permettent de « préciser » l'irrationalité d'un nombre réel par exemple. C'est très important par exemple pour exprimer que π et e ne sont pas des expressions simples à calculer et que ce n'est pas pour rien qu'on s'acharne à les écrire avec des lettres. Une autre utilisation des extensions de corps est purement géométrique, elle permet de définir quelles sont les points constructibles à la règle et au compas (théorème de Wantzel).

On considère un corps K, une extension de corps (pour simplifier) est un corps L contenant le corps K (en réalité c'est la donnée d'un morphisme j de K vers un corps L), on écrira $L : K$ pour dire que L est une extension de corps de K. Par exemple, tout corps de caractéristique nulle est une extension du corps $\mathbb{Q}$, de même tout corps de caractéristique p est une extension de $\mathbb{Z}/p\mathbb{Z}$. Par exemple, $\mathbb{C}$ est une extension de corps de $\mathbb{R}$. Cependant une extension de corps ne précise pas les éléments qu'on a rajouté ; c'est pour cela qu'on étudie les extensions de corps obtenue par « adjonction » :

Soit $L : K$ une extension de corps, soit S une partie de L, alors $K \cup S$ est une partie de L, on peut trouver le plus petit corps contenant $K \cup S$, c'est l'intersection de tous les corps contenant ça. Et on note $K(S)$ ce corps. C'est une extension de corps et on appelle cette extension une extension par « adjonction » de S à K. Dans le cas particulier où $S = \{ a , b , c \dots \}$; on note $K(a, b, c, \dots)$ au lieu de $K(S)$.

Par exemple, dans la feuille de TD n°1 de Mme Benlolo on introduit $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(\sqrt{3})$. Ce sont des extensions de corps par adjonction. De même $\mathbb{R}(i) = \mathbb{C}$ est une extension par adjonction.

Une extension par adjonction est dîtes simple si on a adjoint un seul élément. Cependant certaines sont en apparence trompeuses (comme les filles…blague à part…), par exemple $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3})$ devient une extension simple. Les extensions simples offrent une étude au nombre adjoint, on introduit alors les notions suivantes, un nombre a de L est dit algébrique sur K s'il existe un polynôme P non constant à coefficient dans K tel que $P(a) = 0$. On dira aussi que l'extension $K(a)$ est algébrique. Un nombre qui n'est pas algébrique est dit transcendant. C'est le cas de π, démontré par Lidermann, autrement dit, il n'y a aucun polynome à coefficient dans $\mathbb{Q}$ annulant π. C'est pour cela que π ne puisse pas s'écrire avec des racines (et tout le merdier algébrique) mais toujours avec ln, des cos, etc. Hé oui, vous savez maintenant… Heureusement que vous n'avez pas posé la question au collège ; vous auriez fait peur au prof.

Le corps des fractions rationnelles d'un corps K, noté $K(X)$ est le corps des quotients de deux polynômes, c'est-à-dire dont les éléments sont de la forme P/Q avec P et Q deux polynôme (et évidemment Q non nul), vous avez un peu vu ce corps dans les décompositions en éléments simples. Plus précisemment, $K(X)$ est le corps des fractions de $K[X]$ mais qu'importe la signification…

On a le théorème : Si a est transcendant sur K alors $K(a)$ est isomorphe à $K(X)$. Je ne donnerai pas la preuve de ce théorème mais il est largement explicable : comme a n'annule aucun polynôme, le corps K(a) se forme par des polynômes 1, a, a^2, $a^2 + a$, etc. et donc on y voit bien son « isomorphisme » avec $K(X)$. Ainsi si a et b sont transcendants alors $K(a) \cong K(b)$. L'étude des extensions transcendantes n'apporte donc pas énormément d'intérêt. On va poursuivre sur les extensions algébriques (simple).

Soit a un élément algébrique sur un corps K. On introduit le morphisme $\begin{cases} K[X] \to K \\ P \to P(a) \end{cases}$. Considérons le noyau de ce morphisme, notons le Ker. Alors Ker est un idéal (car c'est un morphisme d'anneau) non nul car a est algébrique (i.e. il existe p tel que $p(a) = 0$) et propre (car le polynôme 1 n'annule pas a). Comme l'anneau $K[X]$ est principal (principal = tous ses idéaux sont engendré par un unique élément), on a $Ker = (p)$ avec un unique p unitaire sur $K[X]$. L'image du morphisme est intègre et son image est isomorphe à $K[X]/Ker = K[X]/(p)$, donc p est irréductible sur $K[X]$. On notera (par unicité de p) $p = I(a)$. Ce qui est remarquable c'est la récirpoque qui reste vraie : pour tout polynôme p unitaire et irréductible p, il existe une extension simple $K(a)$ tel que $p = I(a)$. On appelle degré d'extension de a le nombre $deg\ I(a)$.

Exemples : Considérons $\mathbb{C} = \mathbb{R}(i)$, on a $I(i) = X^2 + 1$ (il est bien unitaire, irreductible et vérifie $i^2 + 1 = 0$, par unicité c'est $I(i)$), on a $deg\ I(i) = 2$, donc i est de degré d'extension 2.

De même considérons $\mathbb{Q}(\sqrt[3]{2})$. On a $X^3 - 2$ qui annule $\sqrt[3]{2}$, il est unitaire, son irréductibilité ne sera pas montrée, et alors on a $\sqrt[3]{2}$ qui a un degré d'extension de 3. Ce résultat sera repris par la suite.

Maintenant que vous savez ce qu'est le degré d'une extension (algébrique simple) on étend au cas transcendant avec la convention que a a pour degré d'extension $+\infty$ si a est transcendant. Voyons le théorème de Wantzel : Un nombre a est constructible à la règle et au compas si son degré d'extension est une puissance de 2. Attention la réciproque est fausse !

Par exemple, π n'est pas constructible à la règle et au compas car π est transcendant, je renvoie au problème de la quadrature du cercle (que nous venons de montrer l'impossibilité ! Tant attendu depuis des millénaires !). De même, quelqu'un de l'antiquité grecque (je ne sais pas qui mais quelqu'un) a demandé après avoir vu son temple construit de construire (à la règle et au compas) un nouveau temple deux fois plus gros (dont le volume est double). Cette construction est impossible, si on suppose cela possible, alors on pourrait construire $\sqrt[3]{2}$ (arête de ce cube) or $\sqrt[3]{2}$ a un degré d'extension de 3, qui n'est certainement pas une puissance de 2. Pauvre architecte. Une minute de silence pour lui.

On va dans la foulée généraliser légèrement ce que nous venons de voir : Soit $L : K$ une extension de corps, on appelle degré d'extension et on note $(L : K)$ la dimension de L en tant que K-espace vectoriel. Par exemple $\mathbb{C} : \mathbb{R}$ est une extension de degré 2 (avec $(1, i)$ comme base possible). $\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}$ est une extension de degré 3 (avec $(1, \sqrt[3]{2}, \sqrt[3]{4})$ comme base possible). $\mathbb{R} : \mathbb{Q}$ est de degré d'extension infinie. On a alors :

Théorème : Si $L : K$ est une extension de la forme $K(a) : K$ alors le degré d'extension de $L : K$ est le degré d'extension de a.

Par exemple pour $\mathbb{C}$, on a $\mathbb{C} = \mathbb{R}(i)$ et i racine de $X^2 + 1$ (par définition de i), et on vient de montrer qu'importe la construction de i, on a aura tout nombre complexe de la forme $a + bi$.

Un autre théorème : Si $L : K$ et $K : J$ sont des extensions de corps, alors $L : J$ en est une et on a : $(L : J) = (L : K)(K : J)$.

Corollaire : $(\mathbb{R} : \mathbb{Q}) = +\infty$, car $(\mathbb{R} : \mathbb{Q}) = (\mathbb{R} : \mathbb{Q}(\pi))(\mathbb{Q}(\pi) : \mathbb{Q})$ or π est transcendant sur $\mathbb{Q}$, donc $(\mathbb{Q}(\pi), \mathbb{Q}) = +\infty$, et on a le resultat.

Note 1 : Ce qui va suivre peut vite devenir très compliqué, afin de ne pas perdre le lecteur je vais oublier le formalisme que j'avais commencé au début de cette section pour « adoucir » les notions, ce qui va suivre va traiter de la Théorie de Galois et des polynômes. Je vous assure que ça va en tout cas devenir nettement plus intéressant. De plus je ne suis pas sûr que ce qui va suivre est au programme de M1, je dirais plutôt M2 et c'est pour cela que ce n'est pas la peine de transformer quelque chose de sympa en un cours pur et dur.

Vous avez appris quand vous étiez des ados boutonneux à résoudre les équations polynômiales du second degré avec une méthode bien « algorithmique », on pourrait se demander s'il existe des résolutions pour les polynômes de degré supérieur, et avec un grand espoir une méthode générale pour tout degré. Il existe bel et bien une méthode pour les degrés 3 et 4, dans le cas de degré 4 la méthode (appelée méthode de Ferrari) est mémorisable, quant au cas des polynômes de degré 3 la méthode devient horrible et fait intervenir des nombres atroces dans un très long raisonnement (ceux qui ont eu le DM de M.Devie en sup traitant cette méthode confirmeront…) d'ailleurs le raisonnement est si horrible qu'il faut parfois être obligé de passer dans $\mathbb{C}$ pour trouver des racines réelles (c'est historiquement l'introduction des nombres complexes dans les mathématiques !) ; cette méthode est appelée méthode de Cardan. On pourrait alors se poser pour des méthodes en degré 5 et plus, cette question va occuper une majeure partie de l'algèbre du XVIIIème siècle. Ce sera Abel qui le premier va démontrer qu'il n'existe pas de méthode « par radicaux » pour le degré 5, cependant Abel, mathématicien norvégien à l' « origine » du prix Abel (pur équivalent du prix Nobel, contrairement à la médaille Fields qui demande d'avoir moins de 40 ans) meurt très jeune, 25 ans (d'une tuberculose) et les résultats ne se seront pas diffusés. Ce sera donc Evariste Galois qui mettra un terme à cette question en montrant l'impossibilité de résoudre les polynômes avec des racines pour les degrés au-delà de 5. Il utilise pour cela de fabuleuses remarques sur les symétries des racines et des raisonnements d'algèbre bien en avance sur son époque pour répondre au problème. Cependant lui aussi meurt jeune, à 21 ans après une histoire extrêmement complexe, particulièrement révolutionnaire il ira en prison ; il aime une fille qui le laisse tomber, il est totalement ignoré des mathématiciens contemporains (Poisson, Gauss, Fourier, …), il verra son père se suicider en même temps qu'un échec scolaire à son entrée en polytechnique. Il se fera renvoyer de l'école préparatoire (ENS aujourd'hui) ; et meurt d'un duel pour des raisons encore un peu sombre. Il sera (fort heureusement et vous l'aurez compris) reconnu en 1843 par Liouville. Voyons un peu comment on pourrait montrer ça.

Tout d'abord nous allons devoir définir ce que signifie « résoudre par radicaux », sans formalisme, résoudre par radicaux signifie que nous avons des racines (exemple $\sqrt{b^2 - 4ac}$) sur les coefficients pour exprimer les racines du polynôme. On part donc des polynômes unitaires à coefficients dans $\mathbb{Q}$ (en fait tout corps commutatif de caractéristique nulle) et on défini le fait de rajouter des « nombres à racines » par une extension issue du polynôme $X^n - a$ (a un élément de $\mathbb{Q}$). Exemple : $X^2 + 1$ adjoint i et $-i$, $X^2 - 1$ n'adjoint rien car 1 et -1 sont déjà dans $\mathbb{Q}$. Ces « nombres à racines » sont appelés radicaux.

On appelle extension radicielle de $\mathbb{Q}$ un corps K obtenu par une suite finie $\mathbb{Q} \subset K_1 \subset K_2 \ldots \subset K$ telle que : $K_{i+1} = K_i(a_i)$ avec a_i un radical de $\mathbb{Q}$.

On défini aussi le corps de décomposition d'un polynôme P non constant grosso-modo par l'adjonction de ses racines au corps $\mathbb{Q}$. C'est là qu'on va perdre tout notre formalisme, comme raccourci. Alors un polynôme résoluble par radicaux est un polynôme tel que son corps de décomposition soit inclus dans une extension radicielle. C'est ce qui exprime que les racines peuvent s'exprimer comme des radicaux des coefficients.

Revenons à un fait général : Soit $L : K$ une extension de corps, l'ensemble des automorphismes de L laissant fixe K s'appelle le groupe de Galois de l'extension $L : K$, on le note $Gal(L : K)$ et comme son nom l'indique, c'est un groupe. Dans le cas d'un polynome p sur K, on définit le groupe de Galois sur p comme le groupe de Galois de l'extension $E : K$ où E est le corps de décomposition de p (en gros l'adjonction de ses racines).

Maintenant que nous sommes revenu aux groupes, définissons ce que signifie un groupe résoluble, disons qu'un groupe résoluble c'est un groupe tel qu'en prenant successivement les sous-groupes dérivées (groupe engendré par les commutateurs : $xyx^{-1}y^{-1}$) on tombe au bout d'un moment sur le groupe trivial $\{e\}$. Autrement dit, il existe n tel que $D^n(G) = \{e\}$ avec $D^n(G) = D(D^{n-1}(G))$ et $D^0(G) = G$.

Le théorème important de Galois se situe là : Un polynôme est résoluble par radicaux, si et seulement si, son groupe de Galois est résoluble (oh que les mots sont bien choisis !).

Nous ajouterons que le groupe de Galois d'un polynôme de degré n est isomorphe à S_n, pour des raisons complexes issues des fonctions symétriques des racines dans un corps bien choisi. Ce truc un peu vague en fait amène à l'étude de résolubilité des groupes symétriques plutôt que ceux des groupes de Galois. Et alors comme tout groupe symétrique S_n n'est pas résoluble dès que $n \geq 5$, on a démontré l'impossibilité de résolutions par radicaux !

Note 2 : Evidemment comme vous venez de voir, j'ai fais main basse sur tout ce qui se passe réellement pour m'intéresser vraiment au minimum (et c'est même un peu abusé). Mais je ne voulais pas perdre le lecteur car en réalité c'est bien plus compliqué que l'image gentillette que j'ai pu donner. Mais l'idée incroyable à retenir c'est qu'un problème de résolution de polynôme a conduit à l'étude d'un groupe en lui-même (et laissant tomber le problème initial) : les groupes symétriques. D'ailleurs montrer leur non-résolution (pour $n \geq 5$) est plutôt simple, on a $D(S_n) = A_n$, et $D(A_n) = A_n$, on voit bien qu'on a une suite constante, on atteint jamais $\{e\}$.

Note 3 : Les groupes de Galois sont des objets fondamentaux en algèbre. On note $\overline{\mathbb{Q}}$ le corps des nombres algébriques sur $\mathbb{Q}$, le groupe de Galois $Gal(\overline{\mathbb{Q}} : \mathbb{Q})$ est un groupe qui fait l'objet de nombreuses recherches ; plus particulièrement on recherche ses représentations. Une représentation d'un groupe est un morphisme du groupe dans le groupe linéaire d'un espace vectoriel ; c'est un domaine de recherche actif. D'ailleurs, si je ne dis pas de bêtises, le centre de recherche de la fac de Reims travaille en partie sur les représentations des groupes (tout du moins l'équipe d'algèbre).

Livres sur le sujet :

- Algèbre corporelle (Antoine Chambert-Loir), c'est un livre plutôt agréable à lire, mais je ne sais pas s'il traite les chapitres dans l'ordre qu'il faudrait pour que ce soit simple.
- Extensions de corps et théorie de Galois (Josette Calais), livre très complet et avec un bon début didactique et de bons exemples. Entre les deux je conseille le second.

Que peut-on mesurer ?

Cette section sera un peu plus représentative que la précédente, et donc moins lourde. Elle concerne essentiellement le programme du module Intégration (MA60X) de L3, les gens qui suivent ce module s'ennuieront peut-être et si c'est le cas je leur conseille de passer à la prochaine partie directement. « Que peut-on mesurer ? » est toujours indépendante aux quatre parties précédentes à l'exception de la définition d'un espace topologique (et encore…). J'essaierai au maximum de ne pas me noyer dans des formalismes pour que même les élèves les plus faibles puissent en profiter ☺. Et c'est aussi pour refroidir les neurones après la partie précédente.

La première question qu'on peut se poser, c'est comment mesurer. Pour mesurer en fait on va choisir un ensemble de « morceaux » qu'on aura le droit de mesurer, et leur associer un nombre c'est les mesurer ; ce nombre peut être une longueur, une aire, un volume, une probabilité, etc. Voilà en gros ce que signifie mesurer. Nous allons donc commencer par définir les parties « qu'on pourra mesurer », i.e. les parties mesurables.

Soit X un ensemble. On définit les parties mesurables sur X comme étant les parties suivante : l'ensemble vide est mesurable, si une partie est mesurable alors son complémentaire est mesurable, et toute union dénombrable de parties mesurables est une partie mesurable. L'ensemble des parties mesurables s'appellent une tribu. Evidemment il y a plusieurs tribus sur un ensemble X. On a par exemple $P(X)$ (toutes les parties sont mesurables), on a aussi la tribu $\{\emptyset, X\}$, il y a de nombreux autres exemples. On appelle espace mesurable un couple (X, T) où X est un ensemble et T une tribu sur X.

Maintenant qu'on a une tribu (i.e. les parties qu'on peut mesurer) on défini la mesure de celles-ci par une application μ de T dans $\mathbb{R}_+$ telle que :

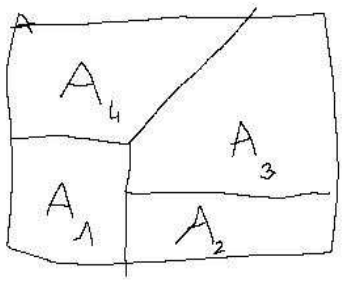
- $\mu(\emptyset) = 0$
- Pour toute famille au plus dénombrable de partie mesurable disjointes deux à deux on a $\mu(\bigcup A_i) = \sum(\mu(A_i))$. (On appelle ça la σ-additivité).

Le triplet (X, T, μ) est appelé espace mesuré. Dans le cas où $\mu(X) = 1$, le triplet alors appelé espace probabilisé, la mesure est dîtes mesure de probabilité et les parties mesurables sont appelées événements.

On a les propriétés simples suivantes : μ est croissante, si $\mu(X)$ est fini alors $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$.

Exemple : On peut définir une mesure (dîtes mesure de dénombrement) sur la tribu $(X, P(X))$ avec $\mu(A) = card\ A$ si A est fini, $+\infty$ sinon.

Une partie A est dite négligeable si elle est incluse dans une partie de mesure nulle. Cette partie n'est pas supposée mesurable.



On peut voir sur le dessin que nous avons la mesure (l'aire) de l'union des A_i égale à la somme des mesures des A_i.

On va introduire maintenant ce que signifie « fonction mesurable », en réalité ça va coïncider avec des intégrales (on n'a pas fait n'importe quoi non plus) ; On considère une fonction f d'un espace mesurable (A, T_A) dans un espace mesurable (B, T_B), on dit que f est mesurable si l'image réciproque de tout élément de T_B est un élément de T_A (on voit l'analogie avec la continuité dans les espaces topologiques), i.e. $\forall T \in T_B, f^{-1}(T) \in T_A \Leftrightarrow f$ mesurable.

Voilà maintenant les notions introduites, on va pouvoir discuter plus formellement des mesures plus classiques, dans des espaces qu'on connait bien. Tout d'abord il faut savoir que l'intersection de deux tribus est une tribu, donc on peut définir pour une partie S de $P(X)$ la tribu engendrée par S, c'est l'intersection de toutes les tribus contenant S. On se considère maintenant dans un espace topologique (X, T) (T désigne ici la topologie, ET NON la tribu), T est une partie de $P(X)$, la tribu engendrée par T est notée $B(X)$ et s'appelle la tribu borélienne de X, ses éléments sont appelés les boréliens. On considère une mesure μ sur cet espace, on peut « compléter » l'espace $(X, B(X), \mu)$ en un espace mesuré complet (qui voit toutes ses parties négligeables mesurables), qu'importe le procédé de construction, il existe toujours. L'espace ainsi complété s'écrira $(X, L(X), \bar{\mu})$ et $L(X)$ s'appelle tribu de Lebesgue. Mais dans la suite on travaillera sur les boréliens, plus facile.

Exercice (yen faut de temps en temps) : Soit X un espace métrique, on muni X de la tribu des boréliens $B(X)$. Montrer que toute partie dénombrable de X est mesurable (i.e. dans $B(X)$).

L'intérêt de fabriquer les boréliens est de pouvoir énoncer le théorème suivant :

Toute fonction continue est mesurable.

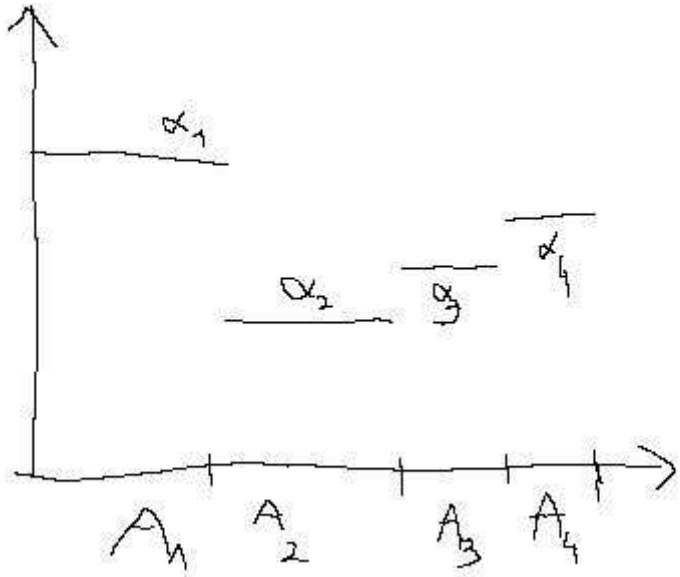
La preuve est « triviale » et laissée à celui qui s'ennuie ; on constate l'analogie très forte avec le théorème comme quoi toute fonction continue sur un segment est intégrable (au sens de Riemann). Ce n'est pas pour rien. Si on devait poursuivre l'analogie, on pourrait montrer que toute fonction qui admet un nombre de points de discontinuités dénombrable est mesurable, comme par exemple la fonction partie entière.

Dire qu'une fonction est mesurable est une chose, mais ce serait bien de la mesurer, comme on l'a fait pour les éléments d'une tribu. C'est là qu'on va introduire en fait la notion d'intégrale.

Cas particulier des fonctions simples.

On appelle fonction étagée une fonction qui prend un nombre fini de valeurs. Attention, il ne faut pas confondre étagée et en escalier, disons qu'étagée on fait les rectangles horizontalement et non verticalement.

On appelle fonction simple, une fonction étagée mesurable à valeur dans $\mathbb{R}_+$. Toutes les fonctions simples peuvent s'exprimer sous une forme $s = \sum \alpha_i \mathbf{1}_{A_i}$ avec les α_i tous différents deux à deux, les A_i tous disjoints deux à deux, l'union donnant l'espace de départ, et tous sont non vide, cette écriture est alors unique.



On définit l'intégrale d'une fonction simple comme le sens évident suivant : on mesure chaque morceau A_i (longueur d'un rectangle) et on multiplie par sa largeur, α_i, pour qu'on aie toute la fonction, on somme tous ces nombres, autrement dit :

$$\int s d\mu = \sum \alpha_i \mu(A_i)$$

Soit f une fonction positive, on peut approcher f par des fonctions simples (une suite croissante même) et donc on définit l'intégrale de f par borne supérieure des fonctions simples plus « petites » (inégalité de fonction), autrement dit :

$$\int f d\mu = sup_{s \leq f} \int s d\mu$$

On a de nombreuses propriétés, comme par exemple, la somme de deux fonctions mesurables l'est aussi, et son intégrale est la somme des intégrales de ses termes. On peut sortir les constantes positives (au pire on met une valeur absolue), etc. Cependant ce n'est pas encore l'intégrale voulue car celle-ci ne fonctionne que sur les fonctions positives. On passe au cas général en définissant l'intégrale de f comme ci : $\int f d\mu = \int \sup(f, 0)\, d\mu - \int \sup(-f, 0)\, d\mu$.

L'intégrale est définie et nous sommes retombées sur nos pattes, voyons maintenant ce que la théorie des mesures engendre. On considère la fonction caractéristique de $\mathbb{Q}$, celle qui vaut 1 sur $\mathbb{Q}$ et 0 ailleurs, cette fonction n'est pas intégrable au sens de Riemann (car $\mathbb{Q}$ n'est même pas un intervalle), pourtant elle l'est au sens de Lebesgue : Cette fonction prend 2 valeurs, c'est donc une fonction étagée, $\mathbb{Q}$ est dénombrable donc mesurable (cf. exercice) et on a pour tout rationnel $\mu(\{r\}) = 0$, $\mathbb{Q}$ étant dénombrable on a $\mu(\mathbb{Q}) = \sum \mu(\{r\}) = 0$. La fonction caractéristique de $\mathbb{Q}$ a donc pour intégrale (au sens de Lebesgue) $1 \times \mu(\mathbb{Q}) + 0 \times \mu(\mathbb{R}\backslash\mathbb{Q}) = 0$. Ce qui était impossible d'énoncer qu'avec l'intégrale de Riemann.

Cependant, l'intégrale de Lebesgue a ouvert de très nombreux problèmes « philosophiques » sur le choix des axiomes dans la théorie mathématique. En effet les mathématiques partent d'axiomes pour se définir et l'un d'eux fait tordre la pensée de façon brutale : l'axiome du choix.

L'axiome du choix énonce que pour un ensemble de parties quelconques on peut toujours choisir un élément dedans, autrement dit (plus formellement), tout produit cartésien d'ensemble non vide est non vide. Cet axiome à l'air certes évident, mais ouvre de nombreux problèmes conceptuels.

La possibilité ouverte par l'axiome du choix est de construire des ensembles non mesurables sur $\mathbb{R}$. Par exemple, les ensembles de Vitali. Qu'importe leur construction, ils ne sont pas mesurable (d'ailleurs ceux qui ont vu la construction TD ont du sentir la difficulté de construction de ces ensembles, sachez de plus que c'est la construction la plus simple connue). Leurs constructions utilisent dès le début l'axiome du choix (dans le choix des représentants des classes d'équivalences d'une certaine relation). Ce n'est pas mesurable…et alors ?

Ça ne semble être rien mais grâce aux parties non mesurables on peut démontrer le théorème suivant (qui vient heurter notre sens logique) :

Théorème de Banach-Tarski (vulgarisé) : Soit S la boule de $\mathbb{R}^3$, on peut découper S en un nombre fini de morceaux (je ne sais plus la valeur exacte) telle qu'en recollant les morceaux (sans les déformer : isométrie) on obtienne deux copies exactes de la boule S…

C'est à cause de résultats du genre que Lebesgue refusa la véracité de l'axiome du choix pour plus se tourner vers l'axiome de Solovay qui énonce que toute partie de $\mathbb{R}$ est mesurable et rendant faux le théorème de Banach-Tarski. Nous reparlerons de la notion d'axiome dans l'avant dernière partie.

Au-delà de la notion d'intégrale de Lebesgue, la théorie des mesures a trouvé sa place dans la théorie des probabilités : on appelle espace probabilisé un espace mesuré (X, T, μ) avec $\mu(X) = 1$. Les parties mesurables sont appelés « événements » et X est appelé univers. On donne un sens formel à la notion de probabilité vue avant. Et cette façon de voir les choses permet de mieux comprendre le paradoxe de Bertrand.

On veut répondre à la question « Soit D un disque, T un triangle équilatéral inscrit dans le cercle, quelle est la probabilité que la longueur d'une corde dépasse la longueur des cotés du triangle ? », Bertrand, utilisant trois méthodes différentes offrira trois réponses : $\frac{1}{2}$, $\frac{1}{3}$ et $\frac{1}{4}$. Bien sûr ces valeurs ne sont pas égales, le problème réside en fait que les trois méthodes changent à chaque fois l'espace mesurée (ou probabilisé)…

Lecture suggérée :

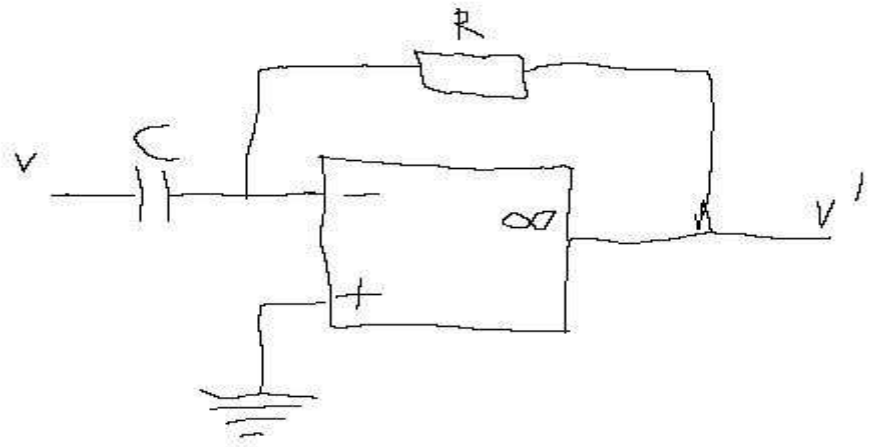
Je suis désolé je n'en vois pas… Vous pouvez toujours demander à un prof où trouver les références.

Quand l'oscilloscope dit vrai.

Cette section (et la suivante) sont des notions au programme de M1 Agreg, mais aussi des notions très profondes dans l'ingénierie, elles ouvrent des utilisations électroniques sans précédent. Cependant je tenterai au maximum de rester dans les mathématiques et me servir pour l'instant de la physique comme « historique », vous n'aurez pas à subir ce monde. Contrairement aux parties précédentes nous allons nous approcher de l'analyse pour quitter le formalisme algébriste, je pourrais m'enfoncer profondément dans le côté topologique de la théorie des distributions mais je vais plutôt en offrir un aspect purement « synthétique » comme l'aimait son inventeur, Laurent Schwartz.

Pourquoi les distributions sont-elles nées ?

Les électroniciens utilisent beaucoup de circuit pour transformer les signaux d'entrées. Un des circuits régulièrement utilisé est le suivant :



Le « gros carré » est un amplificateur opérationnel, j'ignore ce que c'est exactement (ma prof de physique de l'an dernier n'a pas voulu me décrire le composant), mais disons qu'il s'agissait pour moi d'un microprocesseur particulier. Bref, j'ignore pourquoi, mais ce circuit fait les choses suivantes :

Si on considère les tensions d'entrée et de sortie, mesurable avec un oscilloscope, on remarque les transformations :

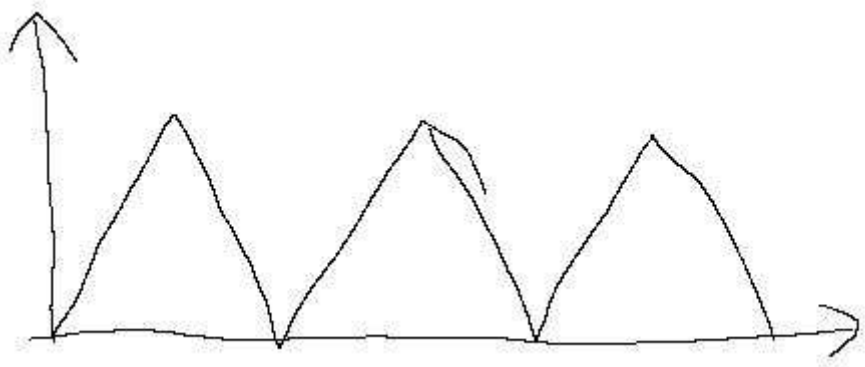
Une sinusoïde devient une autre sinusoïde translatée.

Un courant continu (constant) devient nul.

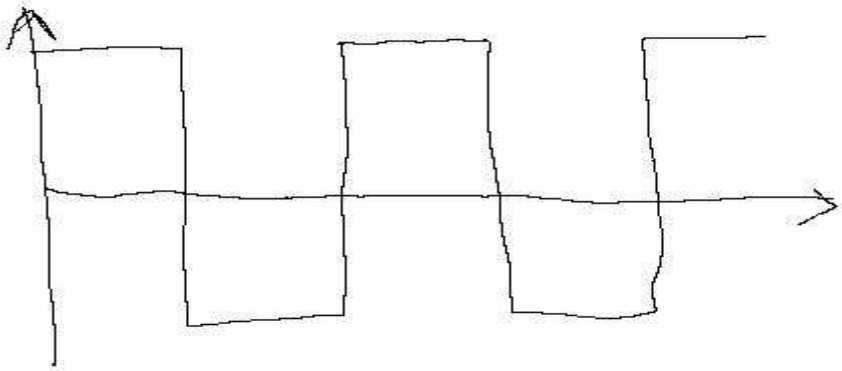
Les morceaux de droites obliques devenaient horizontaux.

Autrement dit le circuit « dérivait » ce qu'on avait.

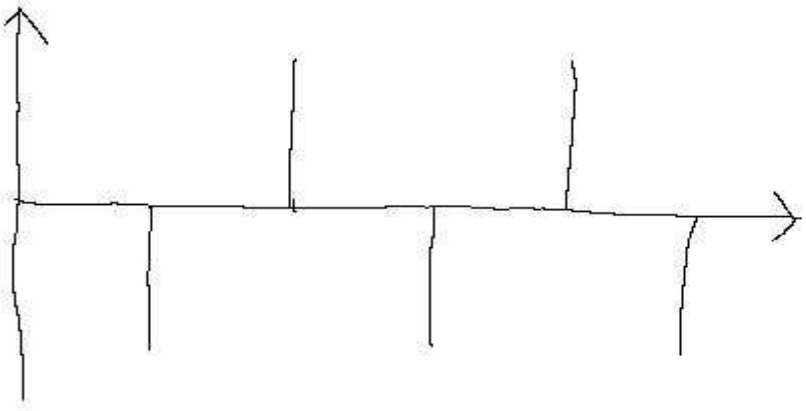
Le circuit dérivateur fonctionne très bien, malheureusement on remarque très vite le phénomène suivant : il dérive des fonctions non dérivables. Car bien souvent en électronique on utilise des tensions avec des signaux triangulaires, ou pire des échelons de tension (qui ne sont même pas continues), et « plus pire », il y a les signaux impulsions (qui parfois ne sont même plus des fonctions).



Fonction triangulaire



Des échelons de tension (fonction créneau)



Des impulsions.

Le signal triangulaire se dérive en signal créneau, qui lui-même se dérive en impulsions. Cette aberration mathématique n'a pas dérangé les ingénieurs qui l'utilisaient, certains physiciens ont tenté de donner un sens à ces expressions, mais c'était bien souvent fastidieux et parfois très peu rigoureux. On appelait à l'époque « Calcul symbolique » ces principes.

Pour mathématiser le fait, il a fallu décomposer ces « fonctions » élémentaires :

L'échelon unité (ou fonction de Heaviside) est définie par : $H: x \to \begin{cases} 1 \text{ si } x > 0 \\ 0 \text{ sinon} \end{cases}$.

L'impulsion (ou « fonction » de Dirac) est définie par : $\delta: x \to \begin{cases} 1 \text{ si } x = 0 \\ 0 \text{ sinon} \end{cases}$.

Ainsi on peut à partir de ces deux fonctions construire les signaux ci-dessus (hors triangulaire).

Ainsi les physiciens, par le « calcul symbolique » affirment que H se dérive en δ. Ce qui est inacceptable pour un mathématicien : H n'est même pas continue ! Encore le passage de triangulaire à créneau était « compréhensible » en « formant » ces points de discontinuités, mais affirmer qu'un point de discontinuité dérive en une impulsion est osé.

Laurent Schwartz résoudra le problème en « transformant » des fonctions en des « fonctionnelles ». Il considère une fonction qui ne prend pas ses valeurs dans $\mathbb{R}$, mais dans l'ensemble des fonctions à support compact (le support d'une fonction est l'adhérence des points où elle ne s'annule pas). Il pose alors $f(\varphi) = \int_{-\infty}^{+\infty} f(x)\varphi(x)dx$. Il dira que f est une distribution (et il la note en général T).

Bien-sûr ce que je fais là n'est pas exactement le sens prévu, car les distributions sont définies de façon légèrement plus compliquées, mais disons que l'intégrale existe car φ est à support compact, c'est donc une sommation sur un compact ; toujours définie pour des fonctions sommables sur tous les bornées (mesurables). Cette intégrale on la note $< T, \varphi >$ étant donné la linéarité à droite.

On ajoute une petite convention : $\delta(\varphi) = < \delta , \varphi > = \varphi(0)$.

En fait plus formellement, une distribution est une forme linéaire continue T de l'espace des fonctions compactes. Ce qui correspond bien avec la distribution pour une fonction ou la distribution issue de δ (distribution de Dirac). Il en existe bien d'autres mais dans le cas d'une FONCTION (avec de bonnes hypothèses) on considère la formule intégrale. On peut montrer que $\varphi \to < f, \varphi >$ et $\varphi \to < g, \varphi >$ sont égales si et seulement si f et g sont égales presque partout (i.e. égale sauf sur un ensemble négligeable, cf. section précédente). Les distributions définies depuis les fonctions sont appelées distributions régulières.

On définit des formules sur les distributions :

Soit S et T deux distributions, on note $S + T$ la somme définie par $\varphi \to < S, \varphi > + < T, \varphi >$.

Soit a une application de classe C^∞, on définie aT comme $\varphi \to < T, a\varphi >$.

On appelle dérivée de la distribution T noté T', la distribution $\varphi \to - < T , \varphi' >$.

On prendra garde au fait que T n'est pas TOUJOURS issue d'une fonction (c'est le cas de la distribution δ). C'est pourquoi nous avons du redéfinir les « transformations ». La dernière peut sembler étrange au premier abord.

Nous allons justifier cette « formule » :

Plaçons dans un cadre qu'on connait bien : les fonctions (aux bonnes hypothèses). Soit donc f une fonction, la distribution associée à f est la distribution $\varphi \to < f, \varphi > = \int_{-\infty}^{+\infty} f(x)\varphi(x)dx$
Nous avons :

$$< f', \varphi > = \int_{-\infty}^{+\infty} f'(x)\varphi(x)dx = \underbrace{[f(x)\varphi(x)]_{-\infty}^{+\infty}}_{=0\ (\varphi \text{ à support compact})} - \int_{-\infty}^{+\infty} f(x)\varphi'(x)dx = - < f, \varphi' >$$

Ainsi la transformation $< T', \varphi > = \ - < T, \varphi' >$ est justifiée.

Avec une telle transformation, toute les distributions deviennent indéfiniment dérivable avec :

$$< T^{(p)}, \varphi > = (-1)^p < T, \varphi^{(p)} >$$

Nous allons maintenant dériver la distribution issue de la fonction de Heaviside, que nous appellerons distribution de Heaviside, encore notée H.

$$< H', \varphi > = \ - < H, \varphi' > = \int_{-\infty}^{+\infty} H(x)\varphi'(x)dx = \int_{0}^{+\infty} \varphi'(x)dx = [\varphi(x)]_0^{+\infty} \underbrace{=}_{\varphi \text{ à support compact}} \varphi(0)$$

Ainsi, comme $\varphi(0) \ = \ \delta(\varphi) \ = < \delta, \varphi >$ on a montré la formule « $H' \ = \ \delta$ ».

Fort heureusement que ça se passe ainsi, on a enfin trouvé une méthode pour autoriser la physique à faire des calculs « faux », et ainsi comprendre les résultats de l'oscilloscope.

On pourrait aller plus loin en définissant un calcul de distribution pour des « pseudo-fonctions », c'est ce qui était d'ailleurs prévu à l'origine dans le document, mais après « recensement » ça se trouve être compliquée, alors j'en donne une version ultra vulgarisée (paix à l'âme de celui qui utilise les distributions et qui lit ça).

En gros (en très gros même), une intégrale qui n'existe pas sur un intervalle $]a, b[$ pourrait exister sur $]a + \varepsilon, b[$, et parfois on peut extraire un polynôme en $\frac{1}{x-a}$ devant telle que le reste soit alors sommable sur $]a, b[$ (sommable = l'intégrale existe). Et alors quand nous passons à l'intégrale sur $]a + \varepsilon, b[$ nous avons deux fonctions, dont l'une s'écrit en polynôme de $\frac{1}{\varepsilon}$ et l'autre ayant une limite finie quand ε tend vers 0. C'est ce deuxième morceaux (enfin sa valeur en 0) qui nous intéresse, c'est en gros, le morceau qui converge dans l'intégrale, et on l'appelle partie finie d'Hadamard, ou tout simplement partie finie et on note ce réel : $\text{Pf.} \int_a^b f(x)dx$.

Dès lors que la partie finie est définie on peut avoir une distribution T issue de ces fonctions en posant $< T, \varphi > = \ \text{Pf.} \int_a^b f(x)\varphi(x)dx$ (où $]a, b[$ contient en fait le support de φ).

Les distributions issues des fonctions (sommables sur tout bornés mesurables) sont appelées des « fonctions » (bien qu'elles n'en soient pas).

Les distributions issues des fonctions vu ci-dessus sont appelées « pseudo-fonctions ».

Les distributions ont leurs plus grandes applications en analyse de Fourier, en analyse harmonique et dans les EDP (équations aux dérivées partielles).

Un livre pour savoir :

- Théorie des distributions (Laurent Schwartz). La notion introduite n'est pas très conseillée pour les calculs, les applications sont énormément topologiques. La lecture est lourde, et sûrement difficile pour un L3 mais lisible quand on saute certains passages.

Savoir téléphoner

Cette section concerne les séries de Fourier, elles aussi très importantes dans le domaine de l'ingénierie et plus particulièrement dans le traitement du signal (et c'est au programme de l'Agreg). C'est une branche très ouverte dans les sciences en général, j'ai même entendu parler d'application en biologie. Mathématiquement elle présente un grand intérêt pour résoudre des EDP (en particulier les séries de Fourier ont été introduites par Fourier (ahah) dans son traité de la chaleur où il résout l'équation de la chaleur donnée par $\frac{\partial T}{\partial t} - k\Delta T = P$), et utilise beaucoup de notions un peu partout en mathématiques (algèbre, analyse, analyse fonctionnelle, géométrie (hilbertienne), théorie des mesures…). Pour ne pas noyer le lecteur dans cette branche, je vais tenter de plutôt introduire ce qu'est une série de Fourier plutôt que de définir le domaine d'application, on conclura sur le calcul de $\zeta(2)$ (j'essaie de faire cette section courte pour apprécier tout ce qui est mis en œuvre, et en plus parce que je n'aime pas les séries de Fourier). Là aussi ceux qui suivent le cours de complément de topologie vont s'ennuyer car c'est exactement ce qu'on fait au début du module (en moins rigoureux).

On va commencer par parler du tore (de dimension 1) : $\mathbb{T}$.

$\mathbb{T}$ est défini par $\mathbb{R}/2\pi\mathbb{Z}$, c'est-à-dire qu'on prend les intervalles d'amplitude 2π et on les regroupe pour choisir une classe de représentant (qui sera une espèce d'intervalle d'amplitude 2π à son tour), c'est-à-dire qu'on a une identification des fonctions 2π-périodique sur les fonctions à valeurs sur le tore (et l'intérêt est là). Le tore peut être imaginé par un cercle de périmètre 2π (d'où le 2π dans son expression).

Mais avant de travailler sur le tore (et donc les fonctions 2π-périodiques) nous allons faire quelques préliminaires (rappels ?) sur les espaces Hilbertiens, fondamentaux dans cette théorie.

On considère un corps commutatif $\mathbb{K}$ (ça peut-être $\mathbb{R}$ ou $\mathbb{C}$, bien souvent ce sera $\mathbb{C}$), soit σ un automorphisme de corps de $\mathbb{K}$, on note λ^σ au lieu de $\sigma(\lambda)$. On considère maintenant deux espaces vectoriels E et F sur le corps $\mathbb{K}$. Une application de E dans F est dîtes semi-linéaire si :

$$f(x + y) = f(x) + f(y) \text{ et } f(\lambda x) = \lambda^\sigma f(x).$$

Dans le cas où $\sigma = id$, nous revenons à une application linéaire, en pratique cet automorphisme σ est régulièrement la conjugaison dans $\mathbb{C}$. l'intérêt de cette notion est d'introduire l'équivalent des « formes bilinéaires », c'est-à-dire une forme sur E^2 linéaire à gauche et semi-linéaire à droite, on les appelle alors « formes sesquilinéaires ». Dès maintenant nous resterons dans $\mathbb{C}$, avec comme automorphisme la conjugaison.

On dit qu'une forme sesquilinéaire f est à symétrie hermitienne si $f(x, y) = \overline{f(y, x)}$.

On dit qu'une forme sesquilinéaire est positive si $f(x, x) \geq 0$, et définie positive si de plus $f(x, x) = 0 \Leftrightarrow x = 0$.

Une forme sesquilinéaire à symétrie hermitienne et définie positive est appelée une forme hermitienne. C'est l'équivalent du produit scalaire sur le corps $\mathbb{R}$. Pourquoi définir de manière aussi compliquée ? Et ben prenons comme forme hermitienne l'application $p : (a, b) -> a\bar{b}$, alors $p(z, z) = z\bar{z}$, donne une norme : $\sqrt{z\bar{z}}$, (ce succes réside dans le caratère involutif de l'automorphisme), tout comme pour le produit scalaire alors que partir d'une forme bilinéaire donnerai zz, qui n'est même pas toujours réel. On note $< x, y >$ au lieu de $f(x, y)$.

Un espace vectoriel complexe muni d'une forme hermitienne est appelé un espace préhilbertien (complexe). De plus si l'espace est complet pour ce produit scalaire, l'espace est dit espace de Hilbert.

Exemple : L'espace $\mathbb{R}^{\infty}$ des suites complexes de carré sommable, i.e. $\mathbb{R}^{\infty} = \{(a_n) \in \mathbb{R}^{\mathbb{N}}; \sum {a_n}^2 < +\infty\}$, est un espace de Hilbert (je ne le prouve pas, ça prendrait beaucoup de place).

Les espaces de Hilbert ont des familles de vecteurs importantes : les bases hilbertiennes. Une base hilbertienne est une famille de vecteurs orthogonales (c'est-à-dire que deux vecteurs distincts de la famille, u et v, vérifient $< u, v > = 0$), unitaire (c'est-à-dire que pour tout vecteur u, $< u, u > = 1$) et totale (c'est-à-dire que l'adhérence (i.e. on inclut le bord, on prend les limites des suites convergentes) de la partie engendrée par les vecteurs est H) ce qu'on écrit : $\overline{\text{Vect}\,(e_\iota)} = H$.

On remarquera que les espaces de Hilbert qui admettent une base hilbertienne n'admette pas forcement une base classique (et ce sera notre cas par la suite), disons que les bases hilbertiennes permettent de faire une « convergence ». Un espace de Hilbert admet TOUJOURS une base hilbertienne.

Soit (e_i) une famille de vecteur orthonormaux (orthogonaux + unitaires) de H (espace de Hilbert, en réalité on suppose de plus qu'il est séparable mais dans notre cas ce sera toujours réalisé), alors (e_i) est une base hilbertienne ssi $\forall x \in H$, $\sum_{n=0}^{+\infty} |\langle x, e_n \rangle|^2 = \|x\|^2 =< x, x >$. Cette dernière formule s'appelle la formule de Parseval.

Maintenant que vous êtes grand, on va pouvoir parler des séries de Fourier.

On considère l'espace des fonctions mesurables (c'est un peu plus compliquées, mais en « pratique » ce sera des fonctions continues par morceaux) définies sur le tore $\mathbb{T}$ (ou si vous préférez, $2\pi -$ périodique) et à valeur dans $\mathbb{C}$. On munit cet espace d'une forme hermitienne définie par :

$$< f, g > = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\overline{g(t)}dt$$

L'espace préhilbertien complexe obtenu est en réalité un espace de Hilbert.

On peut avoir une base hilbertienne : la famille des $e_n = e^{inx}$ pour $n \in \mathbb{Z}$.

Pour démontrer ce résultat : Le côté famille orthonormale est laissé au lecteur. La difficulté réside dans le coté « famille totale », mais la résolution de ce problème se trouve dans le théorème très puissant de Stone-Weierstrass (que je n'énonce pas), mais il existe une forme plus faible qui énonce qu'on peut approcher par des polynômes trigonométriques de façon uniforme, et ces polynômes trigonométriques sont justement ces e_n.

On appelle $n-$ ième coefficient de Fourier le nombre défini par $< f, e_n >$ que nous noterons $c_n(f)$.

La fonction $x \to c_n(f)e_n + c_{-n}(f)e_{-n}$ est appelée la n-ième Harmonique de f. La somme des harmoniques donne par théorème cette fonction f, autrement dit

$$f = \sum_{n\in\mathbb{Z}} c_n(f)e_n$$

Ainsi lorsque nous avons une fonction f 2π-périodique (un signal par exemple !), plutôt que d'envoyer une approximation de la fonction, on peut envoyer une liste de nombres (qui correspond aux c_n , mais en nombre fini), permettant de reconstituer le signal (de façon approximative) f avec la formule ci-dessus. L'intérêt dans la vie de tous les jours fait son apparition.

Comme vous êtes sage (enfin je pense qu'en lisant ça vous l'êtes) je vais vous donner la preuve attendue du résultat $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$.

On considère la fonction f définie par $f(x) = x$ sur le tore.

On considère sa série de Fourier : $f = \sum c_n(f)e_n$.

Calculons maintenant les $c_n(f)$:

On a pour n non nul, $c_n(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} xe^{inx}dx = i\frac{(-1)^n}{n}$, et $c_0(f) = 0$, donc $|\langle c_n(f), c_n(f)\rangle| = \frac{1}{n^2}$

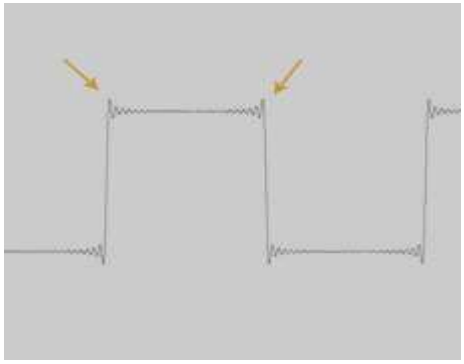
De plus, $< f, f > = \frac{\pi^2}{3}$.

Ainsi d'après l'égalité de Parseval, on a $\frac{\pi^2}{3} = \sum_{n\in\mathbb{Z}^*} \frac{1}{n^2}$

Et donc $\sum 1/n^2 = \frac{\pi^2}{6}$!

Je ne vais pas conclure cette section sur la preuve de $\zeta(2)$, j'aimerais encore parler de deux petites choses, la première concernant le phénomène de Gibbs. Le phénomène de Gibbs était le phénomène mis en œuvre dans le sujet math 1 de CCP l'année dernière, il concerne un problème de convergence point par point.

En effet si on considère un signal créneau 2π-périodique et qu'on étudie ses harmoniques, on arrive très vite à un phénomène du genre :



Les flèches montrent l'existence « d'oreilles » (qui vont grandir en ajoutant les harmoniques), et ces oreilles offrent quelques problèmes de convergence « en hauteur » (dite convergence normale). Ce phénomène n'existe pas dans le cas des fonctions continues C^1 par morceaux.

Bref je voulais juste parler de ce phénomène, mais le premier intérêt des séries de Fourier n'étaient pas d'approcher un signal comme j'en ai parlé tout le long mais bel et bien de résoudre l'équation de la chaleur, l'intérêt des séries de Fourier réside dans la facilité d'expression des coefficients de la dérivée : $c_n(f') = in\, c_n(f)$ ramenant l'étude de cette équation à quelque chose de souvent plus simple.

Note : Ce que je vais dire là n'a pas forcement à voir avec les séries de Fourier, mais je voudrais parler des EDP en général, comme vous avez pu le constater dans le sujet de partiel en Calcul Diff, les EDP peuvent se montrer simple à résoudre à condition d'en connaître une astuce ; nous sommes bien loin de méthodes générales comme le cas à une variable. L'étude des EDP est un très grand problème ouvert faisant intervenir nombreuses mathématiques allant de l'analyse complexe jusqu'à l'algèbre homologique (cohomologie de De Rham sur les formes différentielles), en passant par l'analyse de Fourier, les variétés différentielles, le calcul des variations, bref un problème large. Bien souvent les équations ne sont pas résolues, on en cherche plutôt l'EXISTENCE des solutions, et avec un peu de chance les étudier sans en connaître leurs expressions explicites, régularité etc. On peut voir par exemple l'équation de Laplace. Les EDP « ultimes » sont pour le moment les équations de Navier-Stokes en mécanique des fluides, elles forment un des problèmes du millénaire.

Lecture conseillée :

http://homepage.mac.com/noelle.pottier/page2/files/Chapitre_1.pdf

Hé oui, un PDF pour une fois, c'est pratique à avoir, je ne l'ai pas lu entièrement (à vrai dire je l'ai feuilleté seulement) il n'utilise pas d'espaces de Hilbert, donc il est largement utilisable à profit des étudiants qui n'ont pas M. Lévy-Bruhl en module de complément de « Topologie ».

La géométrie algébrique

On commence à entrer dans les sections abstraites, les sections qui vont suivre sont à peine au programme de Master (juste leurs introductions) et peuvent commencer à ouvrir des voies de recherche. Bien sûr ce que vous allez lire là est loin d'être suffisant, mais ça pourrait ouvrir votre curiosité et le livre que je conseillerai en bas, permet lui de s'introduire plus rigoureusement en géométrie algébrique (pas au point d'un thème de recherche, mais c'est un bon livre pour un mémoire de M2). Nous allons étudier la géométrie algébrique, c'est-à-dire que nous allons écrire un dictionnaire pour décrire algébriquement ce qui se passe en géométrie, c'est cette idée fondamentale qui a conduit les mathématiciens à poursuivre cette branche. Très ouvertes au cours du XXième siècle, en particulier par les japonais. C'est la mathématique Pokemon.

Introduction, nous allons démarrer avec les ensembles algébriques (certains les appelles variétés algébriques, mais la notion de variété utilise en fait la notion de faisceau, qui est très loin d'être évidente, nous en parlerons pas). Un ensemble algébrique, est grosso-modo, une courbe qui peut être décrit par une équation polynomiale, par exemple $X^2 + Y^2 = 1$ représente un cercle, c'est donc un ensemble algébrique. Nous verrons qu'on peut transformer ces courbes en idéal (d'un anneau !) et vice versa, ce qui constitue un premier lien géométrie – algèbre.

Plus formellement, soit S une partie de $K[X_1, X_2, \dots X_n]$ (espace des polynômes à n indeterminées, exemple $X^2 + Y^2 - 1$ est un polynôme à deux indeterminées sur le corps K), on considère l'ensemble $V(S)$ ensemble des zéros de tous les polynômes de S, c'est-à-dire $V(S) = \{x \in K^n \,;\, \forall P \in S, P(x) = 0\}$.

Dans le cas où S est un singleton $\{f\}$, $V(\{f\})$ est notée $V(f)$ et est appelée « hypersuface ».

Par exemple : $V(X^2 + Y^2 - 1)$ est le cercle. $V(1) = \emptyset, V(0) = K^n, V(\det - 1) = SL_n(K)$…

On peut montrer par exemple que l'application $S \to V(S)$ est décroissante.

De plus, l'idéal engendré par S, $< S >$, vérifie : $V(S) = V(< S >)$ (exercice pour le lecteur).

Un premier gros théorème apparait alors : C'est dû au théorème de Hilbert qui énonce que K étant noetherien on a $K[X_1, .. X_n]$ noetherien (qu'importe ce que cela signifie), mais disons que ça permet de mettre en évidence que tout ensemble algébrique $V(S)$ est une intersection finie d'hypersurfaces ! (Oh tiens ça rappelle un exo de la feuille de TD n° 1 du module MA502, Géométrie affine avec M.Ninet, qu'importe.)

On prendra garde par contre que l'application V n'est pas « injective », $V(X) = V(X^2)$.

C'est justement ce que va dire le gros but de ces quelques pages, parler d'un théorème aux abords « évidents » une courbe de degré d et une autre de degré d' se croisent en dd' points (par exemple : le cercle et une droite pas trop mal placée, se coupent en deux points). Bien sûr ce théorème sans contexte est faux : si déjà la courbe est tangente… c'est faux.

En réalité l'astuce va être de se placer dans le bon contexte, pour faire apparaitre le bon nombre de points il faut déjà ajouter des points qui n'existaient pas graphiquement, dans un corps algébriquement clos (tout polynôme de degré n admet exactement n racines), et donc la droite et le cercle se coupe presque toujours deux fois. Ensuite on compte les points multiples, par exemple, quand la droite est tangente au cercle, on le compte deux fois, car c'est une racine double. Et enfin une autre difficulté, deux droites parallèles ne se croisent pas, sauf qu'à la section « Quand l'infini devient un point », on a étudié le cas de croisement de deux droites parallèles, on va donc parler de courbes projectives, et ainsi une hyperbole et son asymptote se croiseront deux fois (à l'infini). Enfin le problème n'est pas résolu, par exemple, X et X^2 ont pour degré 1 et 2 mais ne se croisent pas deux fois, mais sur tout $X = 0$ qui admet une infinité de points, c'est ce qu'on appelle « des composantes communes », les expliciter est plus compliqué. Ainsi le théorème (appelé alors le théorème de Bezout) affirme que deux courbes projectives (planes) de degrés d et d', dans un corps algébriquement clos, et sans composante commune se croisent dd' fois.

Reprenons : On va préférer travailler avec les idéaux plutôt que les ensemble étant donné la remarque $V(S) = V(< S >)$. Et nous allons créer la notion duale (en gros, contraire) de V.

Soit V un ensemble de point de K^n, on considère $I(V)$ l'ensemble des polynômes s'annulant sur V, on peut montrer par exemple que c'est un idéal, on appelle alors $I(V)$ idéal de V.

I est elle aussi décroissante. De plus si V est algébrique on a $V(I(V)) = V$. Et en plus I est injective.

On prendra garde à ce fait là : $I \neq I(V(I))$, en effet prenons $I = < X^2 + Y^2 + 1 >$, et $K = \mathbb{R}$, on a :

$V(I) = \emptyset$, et donc $I(V(I)) = K[X, Y]$ qui n'est plus du tout I.

De plus I « oublie les puissances » : $I(V(< X >) = < X >$ et $I(V(< X^2 >)) = < X >$.

On prendra garde à ce piège : Que vaut $I(K^n)$? On est tenté de répondre 0 (idéal nul). Pourtant c'est faux, en effet, si on considère $K = \mathbb{Z}/2\mathbb{Z}$ alors le polynôme $X^2 + X$ s'annule sur tout K sans être nul. Pour rendre vrai le résultat il faut supposer de plus K infini.

Je voulais à l'origine ajouter une section « Arithmétique des anneaux » juste avant les extensions de corps pour bien partir avec cette partie, j'ai annulé le projet à cause du module de théorie des anneaux, je fais donc confiance au contenu du programme de L3 en espérant que les notions suivantes puissent être comprises, si dans le pire des cas vous voulez savoir tout de suite, ou à la fin du semestre nous n'avons pas vu les notions, vous pouvez venir me voir pour poser ces questions, ou encore voir Mme Benlolo qui s'occupe du module.

Nous allons maintenant parler d'irréductibilité (i.e. nous allons retirer les composantes communes) vous verrez que l'opération magique est justement de revenir en algèbre (comme nous venons de le commencer) pour s'enfoncer justement dans l'arithmétique.

Pour cela nous allons devoir faire un peu de topologie (décidemment, on se demande où est la géométrie là dedans), l'intersection d'ensembles algébriques est un ensemble algébrique, l'union finie aussi, ainsi les ensembles algébriques se comportent comme des fermés dans un espace topologique, ainsi on défini les ouverts comme les complémentaires des ensembles algébriques. C'est une topologie sur K^n, appelée topologie de Zariski.

On fera très très attention à cette horrible topologie : elle n'est pas du tout usuelle, elle n'est même pas métrique (impossible d'y définir une distance), et ses ouverts sont très gros (voyez vous-même, exemple d'ouvert : les complémentaires des hyperplans, ou le complémentaire d'un cercle), en fait ils sont si gros qu'on ne peut pas les « séparer », c'est-à-dire que deux ouverts quelconques se croisent toujours (on dit que la topologie n'est pas séparée), et la séparation est une chose essentielle pour prouver l'unicité d'une limite…

Maintenant revenons à nos « sans facteurs communs », on va commencer avec la topologie « de l'arithmétique » :

Soit un espace topologique X. Si pour tous ouverts U et V on a $U \cap V = \emptyset \implies U = \emptyset \text{ ou } V = \emptyset$, ou de manière équivalente tout ouvert non vide est dense, on dira que X est irréductible.

On remarque l'analogie avec l'intégrité, $ab = 0 \Rightarrow a = 0 \text{ ou } b = 0$. Ce n'est pas un hasard. De plus dans nos espaces « classiques », cette situation ne se produit jamais dès qu'il y a deux points, les composantes irréductibles sont les singletons, on pourra imaginer qu'en fait notre espace usuel est en quelque sorte un corps parmi les espaces topologiques.

Vous verrez que le mot « irréductible » n'est vraiment pas une « nouvelle » définition.

Soit A un anneau, un idéal I est dit premier si l'anneau quotient A/I est intègre. Cette définition est une analogie avec les nombres premiers, p est premier ssi $\mathbb{Z}/p\mathbb{Z}$ est intègre. (Attention ! premier ne signifie pas irréductible dans les anneaux, irréductible utilise une intersection, bien que tout idéal premier soit irréductible, la réciproque est fausse.)

On a le théorème suivant (que je ne montrerais pas) :

Un ensemble algébrique affine V est irréductible (pour la topologie de Zariski induite) si et seulement si $I(V)$ est premier (ou encore $\Gamma(V) = K[X_1, \dots, X_n]/I(V)$ intègre).

On vient de lier le mot « irréductible » des espaces topologiques, et « premier » dans les anneaux (je ne sais pas pourquoi on dit irréductible au lieu de premier en topologie alors que justement on les distingue dans les anneaux…).

Maintenant un autre théorème :

Soit V un ensemble algébrique affine, on peut l'écrire sous la forme $V = V_1 \cup V_2 \dots \cup V_n$, où les V_i sont irréductibles, avec $V_i \subsetneq V_j$ dès que $i \neq j$, de plus une telle écriture est unique à permutation près.

On remarque maintenant l'analogie CLAIRE (pas Petitjean hein !) avec l'arithmétique, particulièrement les anneaux factoriels, c'est encore une conséquence du théorème de Hilbert qui affirme que $K[X_1, \dots X_n]$ est noetherien.

Si quelqu'un n'a pas compris pourquoi on appelle ça géométrie algébrique, il peut s'arrêter de lire maintenant.

On dira que deux ensembles algébriques sont sans facteurs communs s'ils n'ont pas de composantes irréductibles communes.

Vous avez commencé à voir le rapprochement avec le théorème de Bezout, on a pu justifier une des assertions, la géométrie projective a déjà été décrite, le corps algébriquement clos c'est bon, il resterait à parler des multiplicités, mais ça devient vite compliqué à mon goût, donc nous nous arrêterons là pour le théorème de Bezout. Vous pourrez l'apprendre à votre petit frère ou sœur, pour qu'il puisse l'énoncé en classe de Spé Math.

Bref je ne m'arrête pas là pour autant :

Je vais expliquer le phénomène d'oubli des puissances grâce à un théorème :

(Attention les yeux) le Nullstellensatz (ou plus simplement théorème des zéros de Hilbert).

Il énonce que dans un corps algébriquement clos, on a : $I(V(I)) = \sqrt{I}$.

Où $\sqrt{I}$ désigne bel et bien le radical que vous avez vu sur les idéaux en TD ! De plus je pourrais parler un peu plus de cette preuve qui peut utiliser des extensions de corps (juste pour montrer les branchements des maths…) mais je ne le ferais pas…

C'est en fait ce qui explique la perte des puissances.

Note 1 : Emmy Noether était une demoiselle, c'est très rare une demoiselle qui faisait des maths de ce level à cette époque. Un anneau Noethérien est grosso-modo un anneau qui énonce qui si on divise tout le temps, on tombe sur une suite constante (toute suite croissante d'idéaux est stationnaire au bout d'un certain rang), $24 \rightarrow 12 \rightarrow 6 \rightarrow 3 \rightarrow 3 \rightarrow 1 \rightarrow 1 \ldots \rightarrow 1$.

Note 2 : Si vous n'avez pas compris du tout cette section, n'espérez pas lire la suivante, qui la poursuit en quelque sorte.

Note 3 : Cette partie, bien qu'elle soit dans mes préférées est volontairement courte, pourquoi ? Parce que si je la faisais plus longue, il y aurait beaucoup de notions à introduire pour y voir quelque chose d'intéressant. Ou alors j'aurais pu ajouter les lemmes, les preuves, mais ce serait noyer le lecteur dans la rigueur au point d'oublier le plus important. Et puis je le rappelle : c'est culturel, rien de plus.

Lectures conseillées :

- Géométrie algébrique (Daniel Perrin), un bon livre avec plein d'explications de ce qui se passe, mais très vite le livre peut sembler lourd (en particulier quand le lecteur va voir pour la première fois ce qu'est un espace annelé)
- http://www.math.ens.fr/~debarre/DEA99.pdf Quasiment la même chose, mais en PDF, donc gratuit.

Arithmétique des courbes

Vous avez du sentir dans la partie précédente l'abstraction se lever, non pas forcément le niveau, mais bel et bien l'abstraction. Il faut avouer que rapprocher les courbes par des idéaux est franchement nouveau pour un étudiant. Ces principes généraux seront mieux décris après, mais je vous jure que ça ira de pire en pire en poursuivant. Je vais essayer ici de décrire au mieux ce que représente l'arithmétique, malheureusement je ne pourrais pas en faire un exposé très long car je suis moi-même en train d'apprendre l'arithmétique, et plus précisément les formes modulaires.

Je vais approcher le problème historiquement.

Il était une fois, un monsieur casse-couille appelé Pierre de Fermat, qui n'étant seulement qu'amateur n'avait rien trouvé de mieux à faire que de lancer des « énigmes » d'arithmétique aux mathématiciens de l'époque juste pour dire « moi j'ai trouvé nanananère ». Les problèmes furent la plupart du temps tout de même résolus par les mathématiciens de l'époque, sauf un.

Avant même que Newton découvre la gravitation universelle, Fermat écrivit dans la marge de son édition bilingue d'Arithmetica de Diophante que décomposer une cube en somme de deux cubes, et une puissance quatrième en somme de deux puissances quatrièmes, etc. était impossible, mais que la marge était trop petite pour contenir cette « magnifique » preuve. Et puis il est mort, et là…gros blanc : elle où la preupreuve ?

Ce qu'il affirmait peut s'écrire aujourd'hui sous la forme :

$$\forall n \geq 3, \forall (x,y,z) \in \mathbb{N}^3, x^n + y^n = z^n \implies xyz = 0.$$

Nous allons voir entre nous, pourquoi $n \neq 2$.

Cherchons les solutions de l'équation $x^2 + y^2 = z^2$ dans $\mathbb{Z}$, ou de manière équivalente,

$(x/z)^2 + (y/z)^2 = 1$, ou encore les points rationnels de la courbe d'équation $X^2 + Y^2 = 1$.

On peut poser $X = \cos u$, et $Y = \sin u$, soit $\cos^2 u + \sin^2 u = 1$.

En posant $t = \operatorname{tg}\frac{u}{2}$ on a : $\cos u = \frac{1-t^2}{1+t^2}$ et $\sin u = \frac{2t}{1+t^2}$.

On a $\cos u$ et $\sin u$ rationnel ssi t l'est, car $t = \frac{\sin u}{1+\cos u}$.

Donc en posant $t = \frac{a}{b}$ on a : $(b^2 - a^2)^2 + (2ab)^2 = (b^2 + a^2)^2$ (il suffit de remonter dans les équations).

Ainsi, les solutions de l'équation $x^2 + y^2 = z^2$ sont :

$$x = a^2 - b^2,\ y = 2ab,\ z = a^2 + b^2, \text{où } a, b \in \mathbb{Z}.$$

Plus généralement, ce que nous avons réalisé là s'appelle une paramétrisation rationnelle, soit C une courbe plane d'équation $f(X,Y) = 0$, une paramétrisation rationnelle est un couple (a,b) où a et b sont deux fractions rationnelles tel que $f(a,b) = 0$. Les courbes qui admettent un tel paramétrage s'appelle des courbes unicursales. Par exemple les coniques sont des courbes unicursales, par contre, $x^n + y^n - 1$ $(n \geq 3)$ en n'est pas une, sinon on aurait des solutions du théorème de Fermat. Montrer qu'une courbe est unicursale ou non fait partie de la géométrie algébrique. Allons maintenant en arithmétique.

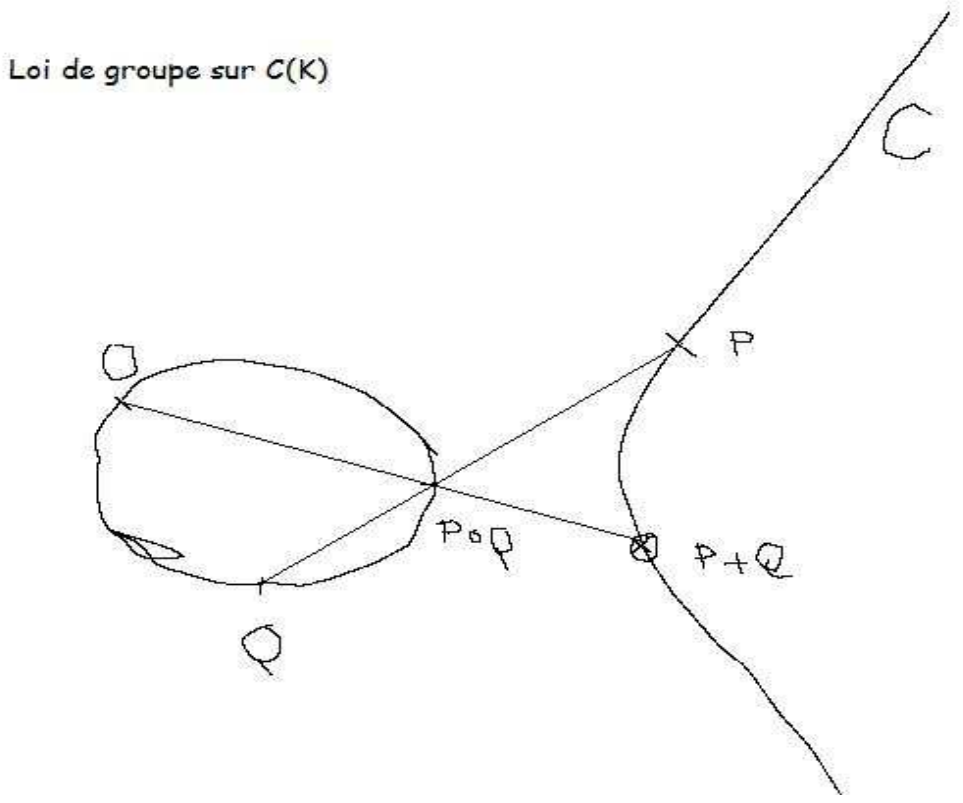
On considère le plan projectif P^2 $(= P^2(\mathbb{C}))$, un polynôme F est dit homogène de degré 3 si on a pour pour tout λ non nul : $F(\lambda X, \lambda Y, \lambda Z) = \lambda^3 F(X,Y,Z)$, une courbe solution de $F(X,Y,Z) = 0$ est appelée cubique (l'homogénéité de F est importante, car sinon écrire $F(X,Y,Z)$ n'a aucun sens, puisque ce sont des coordonnées homogènes).

Une telle cubique est dite lisse si ses dérivées partielles ne s'annulent pas simultanément. Soit C une telle cubique, on note $C(K)$ l'ensemble des points « rationnels » de cette cubique (à coordonnées homogènes rationnelles).

On considère un point O de cette courbe. Soit d une droite de P^2, d'après le théorème de Bezout en géométrie algébrique, elle coupe C en trois points, soit P et Q deux des points, on note $R = P * Q$. On note alors $P + Q = O * (P * Q)$ et $-P = (O * O) * P$.

Il n'y a rien d'incroyable dirait-on, mais cette construction ouvre le théorème suivant :

Si $O \in C(K)$, alors $+$ est une loi de groupe pour $C(K)$.



En arithmétique, le polynôme F n'est pas « pris au hasard », c'est une cubique dite de Weierstrass, c'est-à-dire d'équation : $ZY^2 = X^3 + aXZ^2 + bZ^3$, avec $\Delta = 4a^3 - 27b^2 \neq 0$.

La condition $\Delta \neq 0$ permet d'obtenir une courbe lisse. On voit bien dans les monômes l'homogénéité du polynôme. De même en pratique, on repassera en affine $y^2 = x^3 + ax + b$, et on considérera comme point O, le point à l'infini de $C(K)$.

En général en arithmétique on étudie la courbe elliptique sur $\mathbb{Q}$, et dont le résultat principal est le théorème de Mordell-Weil. Cependant, étant très difficile (puisque je ne comprends pas tout, voire pas grand chose) ; on va plutôt parler des courbes elliptiques sur notre ami $\mathbb{C}$. Et plus précisemment du lien entre fonction elliptique et courbe elliptique puisque c'est intéressant.

Vous n'avez pas encore fait d'analyse complexe (ça viendra l'année prochaine ☺) on va alors introduire quelques petites notions ; une fonction holomorphe sur $\mathbb{C}$ est grosso-modo une fonction dérivable sur $\mathbb{C}$, contrairement à $\mathbb{R}$, une fonction dérivable est toujours indéfiniment dérivable, et différentiable ne signifie pas dérivable, il faut ajouter la condition de Cauchy. On a le théorème très important suivant :

Théorème de Liouville : Toute fonction holomorphe et bornée dans $\mathbb{C}$ est constante sur $\mathbb{C}$.

La preuve de ce théorème est très compliquée pour ceux qui n'ont jamais fais d'analyse complexe, pour ceux qui en ont déjà vu, disons qu'on applique la formule intégrale de Cauchy sur un cercle simple avec un rayon bien choisi pour entourer deux points, la différence des images (en module) est alors bornée par une expression en $1/r$, bref, ça n'intéresse personne là.

On défini ce qu'est un réseau Λ sur $\mathbb{C}$, disons qu'un réseau est un quadrillage (ou de façon plus formelle, un réseau est de la forme $\mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ avec ω_1 et ω_2 non $\mathbb{R}$-linéairement dépendant). On considère une fonction méromorphe (holomorphe sauf sur des points isolés où ça correspond au quotient de deux fonctions holomorphes formant des pôles) sur $\mathbb{C}$ et à valeur dans $\mathbb{C} \cup \{\infty\}$, s'il existe un réseau Λ telle que $f(z + \omega) = f(z)$ pour tous points du réseau Λ et pour tout z complexe.

Une telle fonction est appelée fonction elliptique.

On a le théorème suivant : Toute fonction elliptique entière (i.e. développable en série entière) est constante. La preuve est triviale : cette fonction est bornée sur l'adhérence d'un parallélogramme élémentaire du réseau Λ associé à la fonction, donc bornée dessus, donc bornée sur $\mathbb{C}$, donc constante.

Un autre théorème (encore de Liouville) dit qu'une fonction elliptique ne peut pas admettre un unique pôle d'ordre un (autrement dit, il n'y a pas qu'un terme de la forme $\frac{1}{z-\omega}$ dans son développement). Donc une fonction elliptique admet soit un pôle multiple, soit plusieurs pôles d'ordre 1, soit une combinaison des deux.

Ces théorèmes permettent de construire une fonction elliptique sur un réseau Λ, appelée fonction de Weierstrass (qu'importe le procédé, mais disons qu'elle vient d'une intégration d'une elliptique d'ordre 3) qui est définie par : $\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda \setminus \{0\}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{z^2} \right)$

On peut démontrer que les courbes elliptiques forment un corps, qui de plus est engendré par $\wp$ et $\wp'$: tout fonction elliptique f de réseau Λ, peut s'écrire $p(\wp) + \wp' s(\wp)$ où p et s sont des fractions rationnelles.

De plus les fonctions $\wp$ et $\wp'$ sont reliés par : $\wp' = 4\wp^3 - g_2\wp - g_3$.

Avec g_2 et g_3 deux nombres complexes issus du choix du réseau.

On pourrait se demander quel lien cette affreuse fonction pourrait avoir avec une courbe elliptique, même sur $\mathbb{C}$?

Pourtant c'est devant votre nez, là quelques lignes plus haut…vous ne voyez pas un cube ?

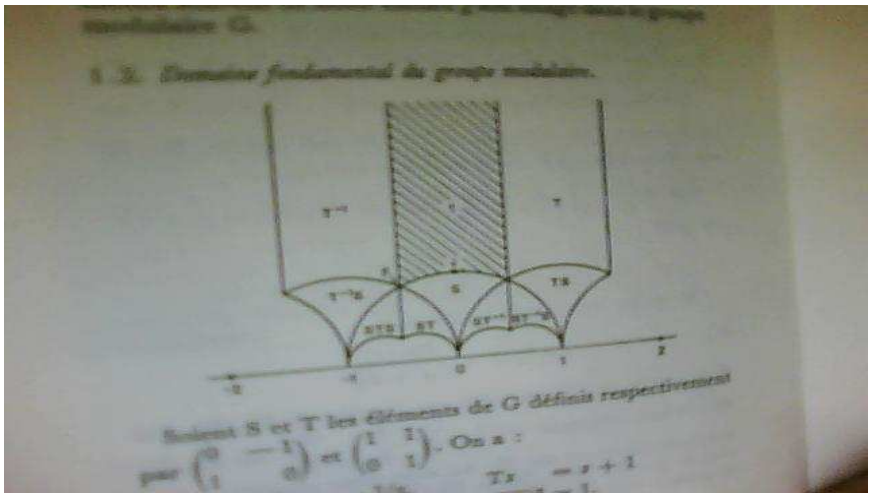
Autrement dit, l'application $z \to (\wp(z), \wp'(z), 1)$ étendue avec $\omega \to (0, 1, 0)$ est une bijection entre $\mathbb{C}/\Lambda$ (ça correspond à un tore complexe) et la cubique projective d'équation $TY^2 = 4X^3 - g_2XT^2 - g_3T^3$.

Voilà, vous avez tracé un lien entre fonction elliptique et courbe elliptique.

J'ajoute juste que pour toute courbe elliptique complexe on peut trouver un réseau tel que $a = g_2$ et $b = g_3$. Autrement dit on peut étudier les courbes elliptiques grâce à la fonction de Weierstrass.

Sachant que les courbes elliptiques sur $\mathbb{C}$ sont les courbes les plus simples à comprendre, je vous laisse mesurer l'étendue sur le corps $\mathbb{Q}$. C'est pour cela que j'arrête cette section ici, afin de ne pas trop perdre l'étudiant qui lit ça. Disons que grosso-modo, on peut ramener une solution hypothétique à l'étude d'une courbe elliptique (dite de Frey-Hellegouarch) dont on étudie ses éléments de torsions, reliés à leur tour par la hauteur de Néron-Tate d'une courbe elliptique et le théorème de Mordell-Weil, et ça peut nous emmener très loin…

Comme par exemple les formes modulaires (j'en parle car en ce moment je lis un cours dessus) où on peut trouver le superbe dessin :



Note 1 : Excusez la photo, je n'ai pas Internet, j'ai du photographier avec une webcam le dessin (parce que le dessiner, hum hum). Référence : Cours d'arithmétique de J-P Serre.

Note 2 : La conjecture de Fermat, démontrée par Wiles en 1994, a duré plus de 300 ans à presque tous les mathématiciens, c'est rare que des résultats durent si longtemps (cf. Conjecture de Kepler, 1998, et la conjecture de Goldbach, non résolue).

Note 3 : Ce dernier théorème de Fermat a été démontré en affaiblissant les hypothèses d'une conjecture bien plus forte, les conjectures en arithmétique sont extrêmement nombreuses : conjecture de Dénès, conjecture de Szpiro (pas le dragon xd), conjecture a,b,c, conjecture de Catalan, et enfin la conjecture de Birch Swinnerton-Dyer. Cette conjecture est un des problèmes du millénaire.

Note 4 : J'aurais pu aussi introduire un côté « théorie analytique des nombres » avec les recherches de densité etc. Ou même en faire une partie, mais je ne le ferais, je me contenterai de dire que ces études peuvent aussi mener à la théorie des nombres, elle aussi très ouverte, avec en particulier l'hypothèse de Riemann (un des 7 problèmes du millénaire, vu par certains par le saint-graal des maths) mais pour moi c'est la conjecture de Hodge, enfin ce n'est que mon point de vue.

Lectures sur le thème (pas facile de trouver des noms différents tout le temps) :

- Invitation aux mathématiques de Fermat-Wiles (Yves Hellegouarch), un incontournable si on aime l'arithmétique.
- Arithmétique (Marc Hindry), un beau livre.

Cependant je dois avouer que ces deux ouvrages ne sont pas faciles à lire et vont décourager beaucoup de personnes, donc j'ignore ce qui existe en arithmétique mais je conseille d'aller à la BU pour trouver des livres qui vous plaisent, je peux par exemple conseiller de lire des petites choses sur les corps finis, et si vous avez des questions dessus, vous pouvez aller voir Mme Barka (si je ne dis pas de bêtise, on m'a dit quand j'étais en MPSI qu'elle était spécialiste des corps finis).

J'ajoute donc à la fin, un livre qui parle de tout ça de façon très simplifiée (pour les lycéens) le livre de Simon Siggh « Le dernier théorème de Fermat », il est à la BU, au pire Adrien Didelet l'a.

Le calcul différentiel tordu

Comme vous avez pu le constater, généraliser les structures a beaucoup occupé les mathématiciens du début du XXème siècle (notamment avec le groupe Bourbaki), comme par exemple la notion de vecteurs, de limite, d'opération, mais aussi l'arithmétique etc. La question se pose alors, peut on généraliser l'analyse ? La question n'est plus soulevée car plus personne ne se la pose (à part moi xd, mais peut-être elle n'est même pas fondée), une généralisation possible a été de passer des espaces $\mathbb{R}^n$ aux espaces de Banach, où la notion de différentielle prend tout son sens. Bref je ne m'étale pas dessus car c'est le module précédent. Cependant les mathématiciens ont pu aller un peu plus loin : ils ont défini un calcul différentiel sur des espaces topologiques (pas trop méchants), mais ce n'est pour moi qu'un « transport » de calcul, et donc pas une grande généralisation.

Pourquoi faire ?

Si vous avez la chance de faire de la voile un jour (ou même du bateau en général, chance que je n'ai pas eu pour l'instant) vous constaterez tout de suite l'intérêt de faire de faire des calculs sur autre chose que $\mathbb{R}^n$: on aimerait étudier sa vitesse, évaluer son déplacement, pour le bateau, or celui ne se déplace pas selon un plan mais bel et bien une sphère. L'interêt est donc d'expliquer comment on peut faire du calcul sur ces surfaces.

On peut commencer par du concret : la géométrie différentielle des « sous-variétés de $\mathbb{R}^n$ ».

Par exemple, si nous choisissons de dessiner une courbe dans $\mathbb{R}^2$, on peut déjà étudier cette courbe grâce aux outils que vous connaissez bien, intégrale, différentielle, etc.

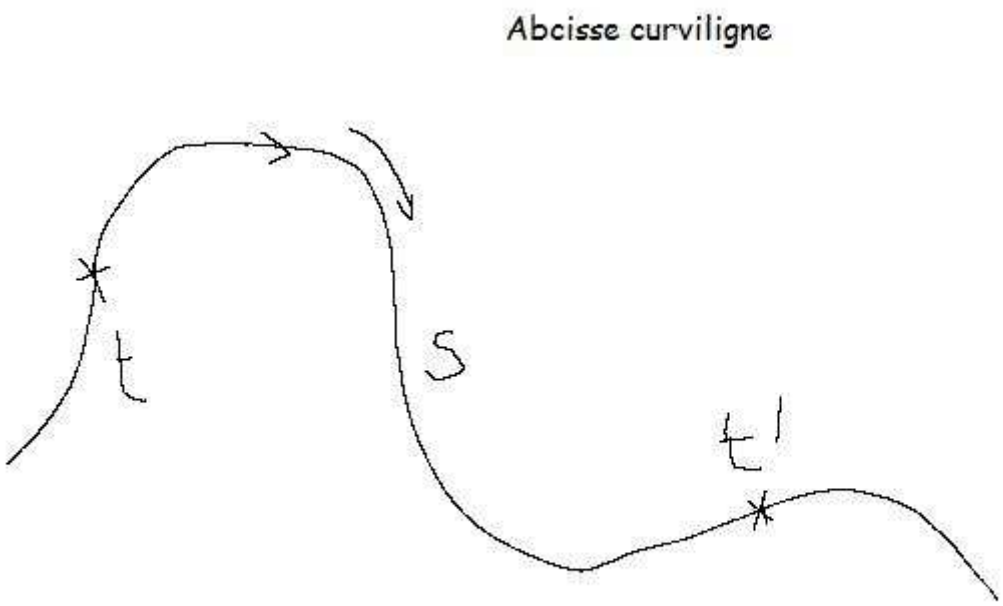
En effet, une courbe C est la donnée d'un couple (I, a) où I est un intervalle, et a une certaine application (en fait une immersion mais qu'importe) à valeur dans $\mathbb{R}^2$, appelée paramétrisation de C. On peut étudier les tangentes de la courbes, c'est-à-dire, les droites engendrées par les vecteurs $a'(t)$. On peut aussi calculer la longueur de la courbe entre deux instants t et t' avec :

$$L = \int_t^{t'} \|a'(u)\| du$$

Mais je ne suis pas là pour donner un cours de géométrie des courbes (vous pouvez revoir dans vos cours précédents en cas de doute), on va plutôt parler d'une notion importante : les paramétrisations directement équivalentes.

Disons qu'une paramétrisation directement équivalente correspond à « changer d'unité » pour la courbe, par exemple, pour un trajet en voiture (en supposant qu'elle ne s'arrête pas), on peut compter la position instantanée durant le trajet, ou encore la distance depuis le point de départ. Soit $C = (I, a)$ une courbe, C admet une paramétrisation (invariante par translation) directement équivalente à a telle que les vecteurs tangents soit unitaires, c'est donc une fonction s, appelée abscisse curviligne. Si on devait s'en donner une intuition, disons qu'il s'agisse de la distance parcourue depuis un point donné (et on comprend mieux l'invariance par translation).

Plus formellement deux paramétrisations sont dites directement équivalentes s'il existe une difféomorphisme strictement positif p tel que $p \circ a = b$. S'il devient négatif on reparcourt la courbe dans l'autre sens, s'il est nul, le point stagne alors que le temps avance, ce qui pose problème.



On peut calculer cette abscisse curviligne en fonction du paramètre t :

$$s(t) = \int_{t_0}^{t} \|a'(u)\| du$$

Par exemple :

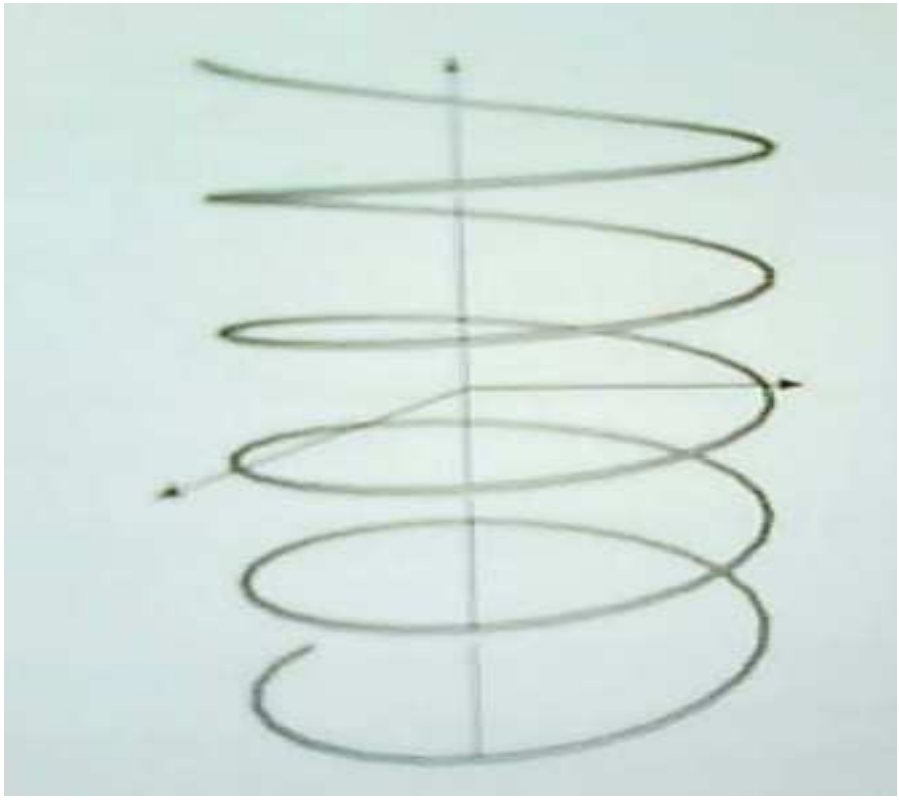
Considérons l'hélice circulaire, de rayon r et de montée p.

Elle est paramétrée par a : $t \to (r\cos t,\ r\sin t, pt)$

Son abscisse curviligne se calcule donc par : $s = t\sqrt{r^2 + p^2} = \frac{t}{c}$

On peut donc la paramétrer par $g: s \to (r\cos cs, r\sin cs, pcs)$

Cette paramétrisation permet par exemple de calculer la longueur de cette hélice : c'est la différente $s(t') - s(t)$.



Voilà à quoi peut « ressembler » de la géométrie différentielle sur les courbes de $\mathbb{R}^n$. On peut généraliser ce principe à n'importe quelle surface, avec une paramétrisation $\mathbb{R}^p \to \mathbb{R}^n$. Cette généralisation décrit en fait un cas particulier de ce qu'on appelle « variété différentielle » qui sont en fait les sous variétés de $\mathbb{R}^n$. Le but de cette section est de montrer comment on a pu généraliser ces concepts. Je ne m'attarde pas sur les courbes, ni même les surfaces afin de ne pas prendre trop de place dans le document, car là j'avoue que nous avons rien vu, mais je pense que toute la classe a déjà vue ces notions (abscisse curviligne, fonction angulaire, tangente, normale, courbure, …).

La généralisation du calcul différentiel a pris son essor grâce à deux remarques (mais quelles remarques !) :

- Le théorème d'inversion local (que je n'énonce pas).
- Le théorème d'invariance des ouverts.

Le théorème d'invariance des ouverts s'énonce ainsi : Si A est un ouvert de $\mathbb{R}^p$ et si B est un ouvert de $\mathbb{R}^n$, si de plus A et B sont homéomorphes alors $n = p$. La preuve de ce résultat utilise des notions très(?) compliquées de topologie algébrique (voir première section) et d'algèbre, on peut aussi partir du théorème de Jordan (peut-être est ce plus accessible ?) qui énonce qu'un ensemble S homéomorphe à l'hypersphère de dimension $n-1$ vérifie : $\mathbb{R}^n \backslash S$ a deux composantes connexes dont seulement l'une est bornée.

Sur la figure ci-dessous, on voit bien les deux composantes connexes, et celle en grise est clairement celle bornée. Une question intéressante est de se demander si cette partie est toujours homéomorphe à la boule unité (ouverte), la réponse est oui pour $n = 2$, non dans le cas général.

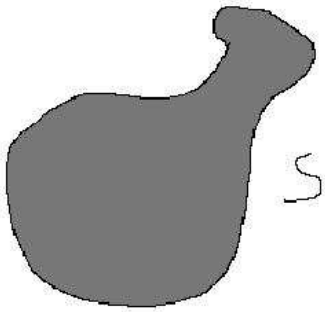
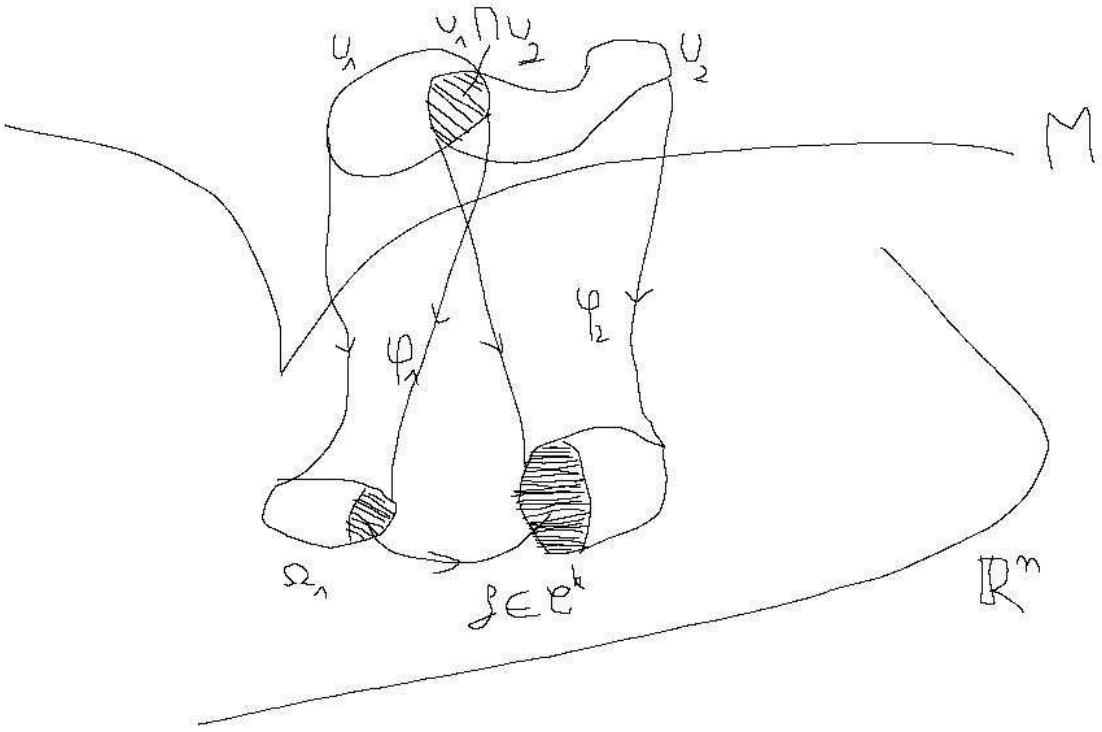


Illustration du théorème de Jordan

Grâce à ces deux théorèmes, on va définir le calcul différentiel sur un espace topologique quelconque. Disons grosso-modo, que l'espace topologique va être muni d'une « paramétrisation » comme on l'a fait pour les courbes.

On considère donc un espace topologique M. On introduit un ensemble A de couple (U_i, φ_i) où les U_i sont des ouverts de M et les φ_i des homéomorphismes de U_i dans un ouvert Ω_i de $\mathbb{R}^n$. Il peut arriver (et ce sera toujours le cas) que deux ouverts de A se croisent, alors leurs images dans $\mathbb{R}^n$ peuvent être disjointes, et donc sur l'intersection on choisit une application d'une certaine classe (disons k), appelée changement de carte, pour passer d'un ouvert de $\mathbb{R}^n$ à l'autre. Pour que ce soit plus clair voyons le dessin :



A est appelé atlas de la variété, les φ_i sont les cartes locales de M, les U_i les ouverts de cartes, l'entier n c'est la dimension de M et enfin k désigne sa classe, k peut valoir $+\infty$.

Ce que nous venons de faire, c'est transposer le calcul de $\mathbb{R}^n$ sur un espace topologique (en pratique on le suppose compact). On prendra garde au fait que M n'est pas toujours représentable dans $\mathbb{R}^3$ même pour des variétés de dimension 2. Pour effectuer des calculs sur M, on redescend sur $\mathbb{R}^n$, on effectue les calculs comme nous avons vu au module MA501, puis on remonte grâce à un φ_i^{-1}. C'est une « généralisation » du théorème d'inversion locale. On ajoutera que la dimension est bien définie à un unique entier grâce au théorème de l'invariance de l'ouvert.

On peut aussi définir ce qu'on appelle des variétés à bords, avec des ouverts homéomorphes à une partie de $\mathbb{R}^n$ de la forme $\mathbb{R}^{n-1} \times \mathbb{R}_+$, ces ouverts constituent « le bord » de la variété ; nous en parlerons pas, et toute nos variétés seront sans bords.

Par exemple la sphère de dimension 2, S^2 est un espace topologique, on peut y définir une structure de variété différentielle grâce aux projections stéréographiques par rapport au pôle nord et au pôle sud. Un théorème énonce que les atlas possibles de la sphère admettent toujours au moins DEUX cartes (maintenant regardez une carte géographique, et montrez moi où se trouve le pôle nord, et où se trouve le pôle sud).

Soient M et N deux variétés différentielles (variété de classe au moins 1), f une application de M dans N. On dit que f est différentiable en x, si pour une carte (et son ouvert de carte contenant x) donnée (U_i, φ_i) l'application $f \circ \varphi_i^{-1}$ est différentiable en $\varphi_i(x)$. C'est une première transposition du calcul différentiel sur un espace topologique. Evidemment cette définition n'a un sens que si en changeant de carte contenant x on conserve la différentiabilité de f ; ce qui est le cas puisque les changements de cartes sont des difféomorphismes, par composition on a conservé les difféomorphismes (voilà l'intérêt des changements de cartes). Cette définition n'est pas exacte car on ne précise pas bien le domaine d'arrivée, mais disons qu'on revient sur $\mathbb{R}^m$ grâce encore aux cartes de l'arrivée.

On obtient des théorèmes usuels comme : la composée de deux applications différentiables l'est aussi. Ou encore un théorème d'inversion locale pour les variétés.

Allons plus loin : Soit M une variété différentielle. Soit x un point de M. Définir un vecteur tangent en x de la variété M est souvent difficile, la façon la plus répandue (et peut être la plus simple) est de prendre des courbes sur la variété (des applications d'un intervalle dans M) passant par x, et de revenir sur les intervalles pour effectuer les calculs, si les dérivées coïncident on parle de tangence des deux courbes. Cette tangence est une relation d'équivalence, dont les classes d'équivalences forment les vecteurs tangents. On peut démontrer que ces vecteurs tangents forment un espace vectoriel, dit espace tangent de la variété M en x.

A tout vecteur x, on associe son espace tangent T_xM, on introduit alors l'ensemble :

$$TM = \bigcup_{x \in M} \{x\} \times T_xM$$

Le couple (TM, p) où p : $\begin{cases} TM \to M \\ (x, v) \to x \end{cases}$ est appelé le fibré vectoriel de M.

Nous avons ainsi transformé le calcul de $\mathbb{R}^n$ en un calcul différentiel sur les espaces topologiques. Il existe une autre définition pour définir la notion de variété différentielle, en réalité cette définition fonctionne aussi en géométrie algébrique et permet de définir les variétés générales, on utilise pour cela le principe de faisceau qu'on ne discutera pas ici.

Par contre, nous verrons dans la dernière partie la notion de préfaisceau sur un espace topologique. En réalité vous en avez déjà vu un en MA601, mais qu'importe, le sujet n'est pas là.

Les variétés différentielles ouvrent la porte à l'analyse de Lie ; et ce à cause des dérivations $((fg)' = f'g + g'f)$. Le sujet des groupes de Lie est très interressant car il permet une étude à la fois algébrique et à la fois analytique de ce qu'on peut rencontrer, par exemple $GL_n(\mathbb{R})$ est un groupe de Lie. Cette notion poursuit l'idée des groupes topologiques.

J'ai énoncé tout à l'heure qu'en mon sens il n'y avait pas de vraie généralisation du calcul différentiel, j'ai prononcé ces mots suite à la remarque que vous pouvez comprendre vous-même : « On transpose le calcul de $\mathbb{R}^n$ sur un espace topologique ». J'ignore encore si le principe généralise réellement les calculs et si non, si le calcul est généralisable. A méditer.

Note 1 : Comme vous aurez pu le constater, cette section est « intuitive » par rapport aux sections précédentes. Son placement dans les derniers chapitres s'explique par le fait que je voulais à l'origine introduire les notions de champs de vecteurs, de groupe de Lie, bref aller bien plus loin que le fibré tangent ; mais je m'arrête là pour ne pas devenir longuet.

Note 2 : Je remercie M.Bounzouina avec qui nous avons cherché à comprendre pourquoi la classe d'une variété était toujours définie, problème qu'il a résolu en plongeant les variétés dans un espace R^n. Par contre j'avoue que sa « preuve » (orale) était fausse car certaines variétés ne peuvent justement pas être plongées, mais ce n'est pas très usuel.

Note 3 : Le groupe Bourbaki est un groupe de mathématiciens du début du XXème siècle se réunissant à l'ENS Ulm où ils projetèrent d'écrire l'ensemble des connaissances mathématiques de leur époque (à l'image des éléments d'Euclide) ; les livres devinrent énormes, et le style Bourbakiste se trouve encore chez de nombreux auteurs, le défaut est de traiter le cas le plus général d'abord pour ensuite s'intéresser aux exemples.

Pour en savoir plus :

- Géométrie et Topologie différentielles (Jean-Yves le Dimet), un livre très propre, avec de nombreux exemples, très agréable à lire, il manque un peu de généralité mais ça n'avait pas l'air d'être le but du livre ; je le conseille vivement.
- Introduction aux variétés différentielles (Jacques Lafontaine), un livre plus complet, allant un peu plus vite que le précédent, je le conseille à celui qui a déjà lu le précédent.

La dimension

Cette section n'a pas de but particulier, je compte y mettre plusieurs notions sans vraiment les rapprocher entre elles, elle concerne la notion de dimension en mathématiques. Son existence est purement synthétique par rapport aux sections précédentes, et la question de dimension est une question rémanente chez les gens de « tous les jours ». Je vais suivre durant ce texte une liste d'idée chainée que j'ai pu voir lors d'une conférence par le président du Clay Mathematic Institute (celui qui remet les 7 prix du millénaire). J'utiliserai aussi cette section pour en profiter pour parler « un peu » et détendre les lectures parfois un peu lourde des chapitres précédents. J'ajouterai juste que ce que je vais dire là ne figure dans absolument aucun cours, c'est un agencement que je garde en mon côté artistique (très mauvais d'ailleurs quand on voit les dessins réalisés sous paint, à la souris).

Bref, comme vous l'aurez compris, on va parler ici de la notion de dimension.

La dimension est quelque chose de très difficile à définir à cause de l'existence d'espace extrêmement compliqué. Dans le cas de quelque chose de régulier et simple comme un espace vectoriel, la notion de dimension est très simple à définir : c'est le cardinal d'une base.

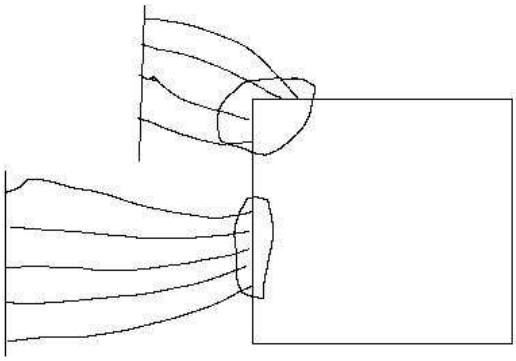
Globalement, définir une dimension c'est tenter de trouver un invariant par transformation qui permet de décrire en quelque sorte la « taille » de l'espace. Par exemple, dans le cas des espaces vectoriels, la dimension est invariante par les automorphismes.

Supposons maintenant que nous allions en topologie, on considère $\mathbb{R}$, et on définit $\mathbb{R}^2$ par $\mathbb{R} \times \mathbb{R}$. Supposons que l'on définisse une dimension « topologique », on a l'intuition humaine de penser que celle de $\mathbb{R}^2$ et celle de $\mathbb{R}$ diffère. C'est ce qui va arriver, peu importe la définition, il faut qu'elle soit invariante par les homéomorphismes. Supposons que $\mathbb{R}^2$ et $\mathbb{R}$ soient homéomorphes, alors si on retire un point à $\mathbb{R}$ (par exemple 0) nous obtenons deux composantes connexes alors que si je retire un point à $\mathbb{R}^2$, l'ensemble reste connexe, on a mis en évidence que les deux espaces ne sont pas homéomorphes, ça devient probable qu'ils n'ont pas la même dimension. De même, supposons $\mathbb{R}^2$ et $\mathbb{R}^n$ (pour $n \geq 3$) homéomorphes, en retirant un point à $\mathbb{R}^2$ on obtient un espace non simplement connexe, alors que si on le retire dans $\mathbb{R}^n$, on conserve la simple connexité (cf. section 1). Pour montrer de façon plus générale que $\mathbb{R}^n$ et $\mathbb{R}^m$ ne sont pas homéomorphes on utilise la théorie de l'homologie, bien trop compliquée pour être exposée. En posant dim $\mathbb{R}^n = n$, nous ne risquons pas de tomber sur des contradictions étant donné les définitions au dessus. Bien qu'on ait beaucoup d'espace topologique, il nous manque encore beaucoup d'espace, même simple, comme par exemple la sphère ou le cercle.

On pourrait traiter le cas du cercle de deux façons, la première utilisant la géométrie projective vue à la section 2, en effet on a parlé d'homéomorphisme du cercle avec la droite projective $P^1(\mathbb{R})$. On serait alors tenté de définir la dimension d'un espace topologique comme celle de l'espace affine où il est issu moins un. Et ainsi on pourrait dire que le cercle est de dimension 1.

La seconde méthode consiste à utiliser les variétés topologiques (cf. section précédente), on peut paramétrer le cercle par $t \to (\cos t, \sin t)$. La présence d'une unique variable indique que cette variété topologique (c'est une variété puisque nous y avons mis un atlas, d'une carte) est de dimension 1. On pourrait alors parler de la dimension comme le « nombre de variable » nécessaire pour parler de cette structure, en fait la dimension de l'espace $\mathbb{R}^n$, comme dit avant.

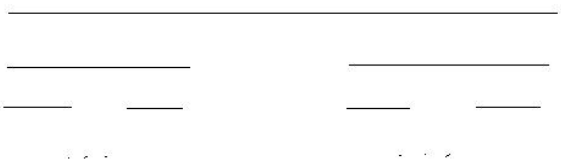
Cette méthode possède deux avantages, premièrement on peut voir très simplement la dimension d'une variété : Prenons par exemple un carré, on voit que localement c'est l'image d'un segment par une application continue. Voir figure ci-dessous :



On y voit bien d'ailleurs l'atlas qui se dessine sur le carré, c'est une variété topologique (de classe 0).

L'autre avantage de cette méthode est de pouvoir parler de dimension d'espace nettement plus compliqué comme le ruban de Moebius (surface obtenue à partir d'une bandelette de papier qu'on a recollé en inversant les deux bords, ce procédé est très courant en topologie algébrique).

Malheureusement cette définition ne va toujours pas, comment pourrait-on parler de la dimension de $\mathbb{Q}$? Ou même d'un espace topologique en général. Par exemple certains espaces très compliqués vont avoir du mal à admettre une dimension, c'est le cas par exemple de l'ensemble triadique de Cantor :



De plus certaines variétés algébriques (disons ensembles algébriques) n'admettent pas de structure de variété topologique ; donc y définir une dimension devient impossible. Considérons par exemple $V(XY)$ qui se représente par une croix dans $\mathbb{R}^2$. Ce n'est certainement pas une variété topologique, si on retire le point d'intersection, alors localement cette surface admet quatre composantes connexes, contre deux si on en retire un ailleurs. Le cas des variétés algébriques peut être résolu grâce à une dimension un peu étrange, la dimension d'un anneau.

On considère un anneau A, on rappelle qu'un idéal premier de I est un idéal de A tel que A/I est intègre. On appelle dimension de Krull la longueur maximale d'une chaine strictement croissante d'idéaux premiers de A ; cette chaine peut éventuellement être infinie. Par exemple dans $\mathbb{Z}$, tout idéal premier étant maximal on a dim $\mathbb{Z} = 1$. On définit alors la dimension (dite de Krull) d'un ensemble algébrique comme la dimension de Krull de l'anneau $\Gamma(V) = K[X_1, \dots, X_n]/I(V)$ et elle correspond à l'intuition qu'on a des dimensions des ensembles algébriques.

Cette définition ouvre à une définition topologique de la dimension (encore appelée dimension de Krull) malheureusement cette définition n'est adaptée qu'aux topologies de type Zariski, et laisse donc des « gros espaces » comme $\mathbb{R}^n$ de dimension nulle (allez expliquer ça à un physicien maintenant).

Cependant il existe une notion de dimension pour les espaces topologiques quelconques, en fait trois. J'en introduis une qui se construit récursivement (parce qu'un exemple est simple) la voici :

Il nous faut introduire la notion de base topologique.

Soit (X, T) un espace topologique, on appelle base topologique de X, un sous-ensemble B de T tel que tout élément de T (ouvert de X) s'écrive comme réunion d'élément de B.

Les boules ouvertes d'un espace métrique forment toujours une base topologique.

Nous allons maintenant définir la dimension d'un espace topologique (attention ça va être un peu bizarre) :

On le fait « récursivement » de cette manière :

Si pour tout voisinage U d'un point, il existe un voisinage V de bord (ou frontière je ne sais plus comment M Bounzouina l'a appelé) vide, alors l'espace est dit de dimension 0.

On dit alors qu'il est de dimension au plus n, si le bord de V est alors de dimension au plus $n - 1$.

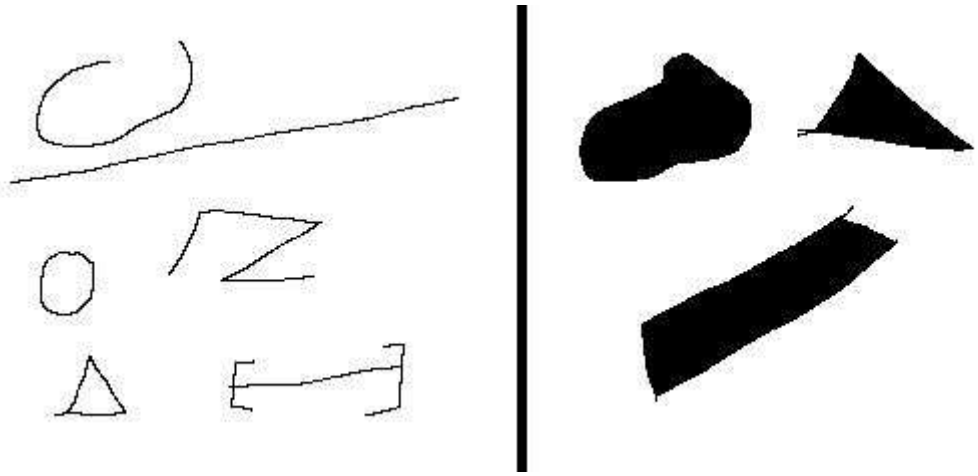
Un espace est dit de dimension n, s'il est au plus de dimension n et s'il n'est pas au plus de dimension $n - 1$. De dimension infinie s'il n'est pas de dimension au plus n pour tout n.

La dimension ainsi définie nous va bien :

La dimension est invariante par homéomorphisme, de plus, la dimension de $\mathbb{R}^n$ est n (la preuve utilise l'homologie en topologie algébrique). Elle permet le calcul de dimension d'espace plus compliquée comme $\mathbb{Q}^2$.

Soit (x, y) un élément de $\mathbb{Q}^2$, alors pour tout voisinage de (x, y) il existe un rectangle contenu dans le voisinage et dont les coordonnées soient irrationnelles, sont bord dans $\mathbb{Q}^2$ est donc vide. Ainsi $\mathbb{Q}^2$ est un espace topologique de dimension nulle (pour la distance induite de celle de $\mathbb{R}^2$).

Je vais donner sur $\mathbb{R}^2$ quelques exemples d'espaces de dimensions 1 et 2 :



A gauche ce sont des espaces de dimension 1, à droite de dimension 2, la bande peut être infinie.

Pourquoi cette notion de dimension n'est elle pas invariante par les applications continues ?

Ce n'est pas un défaut de cette définition de dimension, mais un défaut conceptuel, on se représente toujours l'image d'un segment par une application continue étant une espèce de courbe non coupée en morceau, pourtant ça peut devenir plus compliqué, on peut construire une application continue allant du segment unité dans le carré unité et étant surjective. (Je ne fournis pas de dessin car je suis incapable de dessiner ça à la souris, vous pouvez aller voir « courbe de Hilbert » sur Internet, mais j'ignore où précisément).

Cette dimension appelée dimension de Menger-Urysohn (ou encore petite dimension inductive) semble bien correspondre, cependant d'autres personnes ont donné une définition de la dimension, on citera par exemple la grande dimension inductive (ou encore dimension de Brouwer-Čech). En mon sens, je préfère la définition la plus moderne d'espace topologique : la dimension de Čech-Lebesgue, qui est en quelque sorte entre les deux précédentes. Ces trois dimensions coïncident dans nos espaces usuels, les espaces métrisables séparables. Pourtant il existe par exemple un espace topologique (et même normal séparable) vérifiant $\mathrm{ind}(X) = 0$, $\dim(X) = 1$, $\mathrm{Ind}(X) = 2$, ind désigne la petite dimension inductive, Ind la grande dimension inductive. L'avantage certain de cette dernière définition de dimension est qu'elle n'utilise pas de principe récursif.

Note 1 : L'unique référence que j'ai de cette notion est « dimension topologique et système dynamique » de Michel Coornaert, qui se trouve être aveugle (pauvre monsieur) et son cours fut rédigé au tableau par un de mes amis Stéphane Laurent, probabiliste à l'université de Strasbourg. Je voulais lui faire ce petit clin d'œil…

La dimension de Čech-Lebesgue consiste à recouvrir un minimum de fois l'espace pour calculer sa dimension, plus précisément, le nombre de croisement entre les ouverts indique la dimension de l'espace. La définition formelle étant légèrement tordue, je ne suis pas sûr de pouvoir être bien clair dans ce que je dis, mais je vais tenter le coup :

Imaginez un espace topologique, on choisit de le recouvrir par des ouverts en nombre fini (on pourrait commencer par X), alors depuis ce recouvrement on définit des ouverts inclus dans les premiers qui recouvrent de nouveau cet espace afin de minimiser le nombre de « débordement ».

On choisit alors le plus grand « débordement possible » parmi les recouvrements finis initiaux, c'est ce nombre qu'on appelle dimension de l'espace.

La définition formelle étant :

$$\dim X = \sup_{\alpha} \min_{\beta \succ \alpha} \sup_{x \in X} (\mathrm{card}\{i \in I; x \in A_i\} - 1)$$

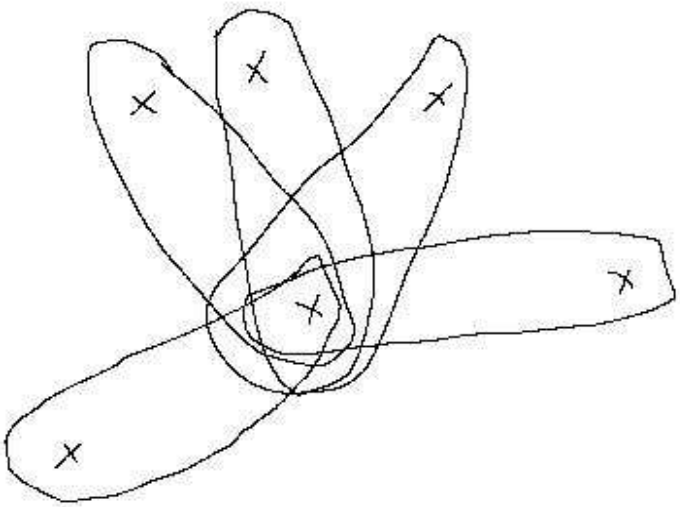
où α parcourt les recouvrements ouvert fini et $\succ$ signifie "être plus fin"

Vous comprenez pourquoi ça m'est difficile d'expliquer là ?

Par exemple pour un espace discret, peu importe le recouvrement fini d'ouvert du départ, les singletons (qui sont ouverts) recouvre tout l'espace. Comme ils ne se touchent pas, il n'y aucun débordement, donc on obtient toujours 0 (et ce peu importante le recouvrement initial), ainsi la plus grande valeur possible est 0, donc $\dim X = 0$.

Un autre exemple : on considère un ensemble X fini contenant au moins deux éléments. L'un des éléments est noté x, et on définit une topologie sur X en prenant comme ouvert soit l'ensemble vide, soit les ouverts contenant x. Calculons sa dimension : Recouvrons X par des ouverts (en nombre fini), on peut minimiser les débordements en prenant par exemple les ouverts de la forme $\{x, y\}$ avec $y \neq x$, ainsi les ouverts ne se croisent qu'en x, par contre on ne peut pas réduire ce débordement « en x » puisque tous les ouverts contiennent x, de plus, c'est bel et bien un recouvrement, on parcourt tous les y, donc on a débordé n fois sur x. On ne peut pas faire plus, c'est donc la dimension de cet espace topologique.

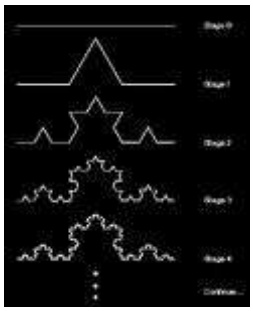
Le recouvrement parfait : plus "fin" que tous les autres.



Nous avons enfin exhibé une méthode pour calculer les dimensions de presque n'importe quel espace, cependant je vais poursuivre sur le cas très particulier de la dimension de Hausdorff.

La dimension de Hausdorff permet de calculer des dimensions nettement plus compliquées en particulier cette des fractales.

Par exemple, si on considère le flocon de Von Koch (voir figure plus basse), on peut voir que c'est « un peu plus épais » qu'une droite, sans être pour autant un plan. Cette notion de dimension vient en fait parfaire la remarque que nous avons dis sur la courbe de Hilbert.



On définit par fractale simple un ensemble qui se répète dans lui-même, plus formellement :

Un ensemble F d'un espace euclidien est appelée fractale simple s'il existe une partition de cet ensemble en partie semblable à F i.e. il existe une similitude pour revenir à F. C'est le cas par exemple de la courbe de Von Koch ci-dessus.

Soit maintenant E une partie bornée d'un espace métrique et $a > 0$, on définit l'a-mesure de E comme : $m_a(E) = \lim_{\varepsilon \to 0}(\inf \sum \delta_i^a)$ où l'inf. est pris sur tous les recouvrements par des boules de diamètre $\delta_i < \varepsilon$.

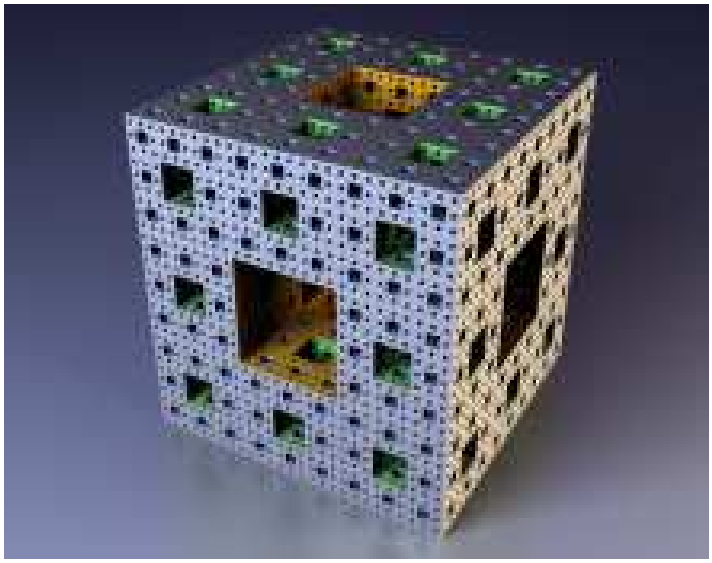
Besicovitch montrera alors qu'il existe un réel u tel que si $a > u$ alors $m_a(E) = 0$ et si $a < u$ alors $m_a(E) = +\infty$, c'est ce réel u qu'on appelle dimension de Hausdorff. C'est en ce sens qu'elle est plus compliquée : ce n'est pas forcement un entier.

On définit par ensemble fractal un ensemble bornée telle que sa dimension de Hausdorff dépasse sa dimension topologique. En réalité la dimension de Hausdorff précise la dimension de l'ensemble au-delà de sa dimension topologique (peu importe laquelle, les trois coïncident ici).

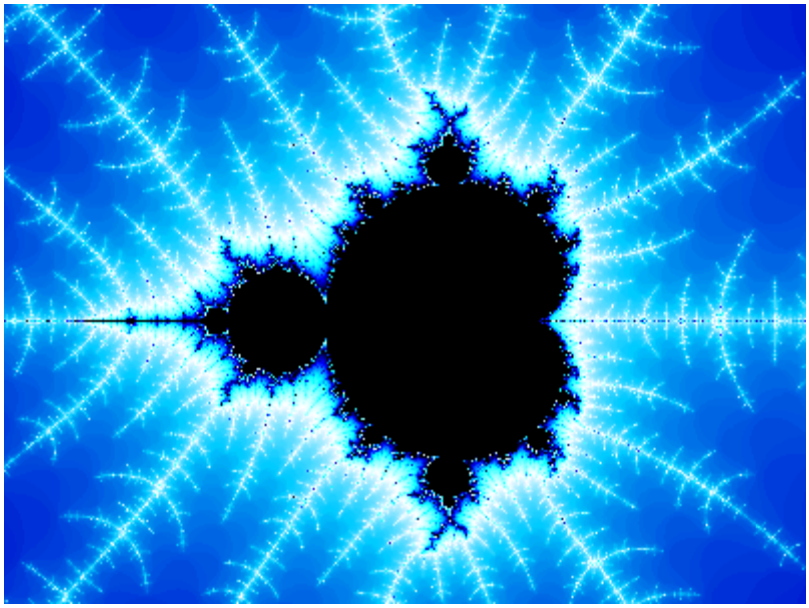
Une remarque fabuleuse de ces dimensions est leur facilité de calcul dans quelques cas, soit F une fractale simple, on dit qu'elle sera régulière si on peut trouver une partition finie telle que toutes les similitudes aient le même rapport r, notons n le nombre d'éléments de la partition. Alors un théorème fondamental énonce que la dimension de Hausdorff de ces espaces vaut exactement $-\log n / \log r$.

Par exemple, dans le cas de la courbe de Von Koch, à chaque itération on partition en 4 éléments qui ont une longueur valant le tiers de la longueur d'origine, donc sa dimension de Hausdorff vaut : $ln\ 4 / ln\ 3$ ce qui correspond à environ $1{,}26$ c'est ce qui explique son placement entre droite et plan.

Comme exercice, je peux donner à calculer la dimension de Hausdorff de l'éponge de Menger.



On ajoutera finalement que certaines fractales ne sont pas simple et font l'objet de beaucoup d'étude, notamment les ensembles de Julia, ou encore le bonhomme de Mandelbrot.



Je conclurai cette section sur ce beau dessin…

Lecture sur le sujet :

- Dimension topologique et système dynamique (Michel Coornaert), par contre je vous souhaite bonne chance pour le trouver, il est en collection SMF, de plus seul le chapitre 1 concerne vraiment la dimension topologique de façon générale. Le livre bien qu'accessible après avoir lu le livre de Skandalis (cf section 1) demande à partir d'un moment donné des connaissances relevées en système dynamique (le livre étudie la dimension moyenne de Gromov).
- Je pense par contre que vous trouverez beaucoup de référence sur les fractales. Je pense au livre de Benoit Mandelbrot par exemple.

L'origine logique des mathématiques

Cette avant dernière section sera introduite pour un titre purement culturel car bien loin des enseignements scolaire classiques. Nous parlerons principalement la théorie des ensembles et nous introduirons cet aspect de façon « historique ». Ce que vous allez lire semblera abstrait car les bases pour vous y attacher seront peu nombreuses ; je n'introduirais pas le concept de logique pour ne pas trop m'encombrer et je pense qu'il y a un module de logique dans les années précédentes. En dernière remarque avant de commencer, je tiens à remercier Benoit (Dejoncheere) pour avoir eu l'idée d'introduire cette section mais je ne suivrais pas son idée d'introduire directement toutes les notions, et me permettrait de mettre cette section à titre culturel.

Pour commencer, nous allons faire un bon en arrière de deux millénaires, à l'époque de la « naissance » des mathématiques et de sa diffusion dans le monde entier. C'est-à-dire à l'époque d'Euclide environ 300 ans avant notre ère.

Euclide eu l'idée de rédiger un ouvrage portant sur l'ensemble des connaissances mathématiques de son époque. Cet ouvrage n'est pas seulement une synthèse, mais il est le plus ancien ouvrage qui expose une façon de démontrer les choses. C'est-à-dire qu'Euclide part d'un ensemble d'axiomes (qu'il considère comme être le plus bas possible) et de postulats avec lesquels il démontre tous les théorèmes connu de l'époque (comme par exemple le théorème de Pythagore). Ce livre est un monument mathématique mais aussi historique : écrit en 13 volumes il fut l'un des premiers imprimés en Europe ; c'est aussi le livre le plus édité (après la Bible) au monde. Le livre est remarquablement bien écrit, à un tel point qu'il est encore lisible par les étudiants actuels (à condition qu'il soit traduit), de plus il a été au programme scolaire pendant plusieurs siècles. Pour mesurer la portée de ce livre, en voici un axiome : « Si deux nombres sont égaux, alors ajouter le même nombre à chacun d'eux formera deux nombres encore égaux ».

Deux-mille ans plus tard, un groupe de mathématiciens français eurent l'idée de produire un équivalent des éléments mais pour les mathématiques modernes (en réalité on différencie mathématiques modernes et mathématiques contemporaines mais ces différences sont d'ordre quasi politique et je n'en détaillerai pas le débat) et cet ouvrage s'appellera « Eléments des mathématiques » et deviendra le deuxième ultime ouvrage des mathématiques. Le groupe est nommé « Nicolas Bourbaki » (le choix du nom n'est pas encore totalement déterminé) et représente un mathématicien imaginaire qui travaille dans un pays imaginaire. Bref ce groupe très secret avait des règles strictes : à 50 ans, chaque membre devait céder sa place à quelqu'un de plus jeune. Le groupe Bourbaki est aujourd'hui disséminé mais existe encore, les membres actuels gardant leurs noms secrets ; on peut néanmoins les « repérer » par leur style très particulier.

Eléments des mathématiques se compose de 11 livres, le premier étant théorie des ensembles, et le dernier l'histoire des mathématiques même si chaque livre est divisé en chapitre et les éditeurs n'en publient que quelques une à chaque fois, faisant de ce traité une œuvre colossale dont le niveau est très disparate. Cependant l'ouvrage est difficilement accessible à un étudiant et son manque d'exemple le condamnera. L'œuvre est aujourd'hui inachevée et n'est plus publiée voire presque totalement vendue ; il fallut attendre 2006 pour qu'il soit réédité. Pour info : certains chapitres de cet ouvrage sont à la BU.

Nous allons parlons du premier livre : la théorie des ensembles.

Avant toute chose, nous allons parler de la philosophie bourbakiste : elle concerne la généralisation des structures (c'est ce que nous verrons au chapitre suivant), et la recherche d'un système d'axiome relativement fin pour tout énoncer. Je vais donner quelques façons de voir les choses pour un bourbakiste.

On considère deux ensembles E et F, une relation R sur E dans F est un sous-ensemble de $E \times F$ (non vide). On note alors xRy au lieu de $(x, y) \in R$.

Dans le cas où $E = F$:

On dit qu'une relation est réflexive si : $\forall x \in E, xRx$

Symétrique si : $\forall (x, y) \in E^2, xRy \Leftrightarrow yRx$

Asymétrique si elle n'est pas symétrique.

Antisymétrique si : $\forall (x, y) \in E^2, xRy$ et $yRx \implies x = y$

Transitive si : $\forall (x, y, z) \in E^3, xRy$ et $yRz \implies xRz$

Dans le cas général :

Surjective à gauche si : $\forall x \in E, \exists y \in F$; xRy

Injective à droite si : $\forall (x, y, z) \in E^3, xRy$ et $xRz \implies y = z$

Une relation binaire réflexive, symétrique et transitive est appelée relation d'équivalence.

Une relation binaire réflexive, antisymétrique et transitive est appelée relation d'ordre.

Une relation surjective à gauche et injective à droite est appelée application.

Une relation injective à droite est appelée fonction.

Dans ces deux derniers cas, on note $y = R(x)$ au lieu de $(x, y) \in R$.

Il existe encore de nombreuses relations prédominantes en mathématiques comme les relations de préordres, mais n'étant pas un grand connaisseur je ne peux pas exposer longtemps dessus.

Toutes ces jolies définitions, ont beau être aussi général que l'on veut, elles utilisent toute une notion commune : la notion d'ensemble. Dans la folie de recherche des définitions, on pourrait se poser la question « qu'est ce qu'un ensemble ? ».

Cette question restera longtemps essentielle en mathématique et fut initiée par Georg Cantor (qui deviendra fou) et il faut se « mettre d'accord » pour le choix des axiomes de la construction de la notion d'ensemble. Ces choix d'axiomes nous en avons déjà parlé à la partie 4 au niveau du paradoxe de Banach-Tarski. Il faut être très minutieux quand on pose un axiome : pendant pas loin de 2000 ans l'un des postulats d'Euclide sera admis par tous comme « évident » pourtant aujourd'hui beaucoup d'étudiants savent qu'il existe des géométries le rendant faux et il ne s'agit pas seulement de délires mathématiques car ces géométries sont à la base de la relativité restreinte.

Plus précisément il s'agit du 5ième postulat d'Euclide énonçant : « Pour tout droite du plan et tout point extérieur à celle-ci il existe une et une seule droite parallèle à la droite donnée passant par ce point. » ; pour un exemple de géométrie non-euclidienne (géométrie où le 5ième postulat d'Euclide est faux) on peut imaginer une sphère comme l'espace étudié et les droites étant les grands cercles (ceux ayant le même rayon que la sphère), on peut voir que dans une telle optique, deux droites quelconques se croisent toujours. Bref cette discussion va ouvrir une grande méfiance quant au choix des axiomes pour définir ce que nous appelons « ensemble ». Dans la suite nous exposerons la théorie de Zermelo-Fraenkel, théorie la plus usée parmi les mathématiciens actuels.

Comme il faut quand même parler de quelque chose, on utilise ce que nous appelons une « collection ». J'ignore comment on définit actuellement une collection, ni même s'il existe une définition mathématique de ce mot, tout ce que j'en sais c'est qu'il faut bien un point de départ et c'est ce point qui n'est pas mathématisable et ouvre les portes à la philosophie, en particulier on peut méditer sur la question : « Existe-t-il quelque chose ? ». Gödel d'ailleurs appellera ce premier axiome (base de toutes les mathématiques) « Dieu » et tentera d'en démontrer l'existence mathématique, puis physique ; la folie l'emportera.

En théorie des ensembles on travaille avec une collection particulière dite « univers » et dont les objets seront précisément les ensembles. Disons que nous sommes obligés de définir la notion de « collection » à cause du théorème suivant :

Supposons qu'il existe un ensemble contenant tous les ensembles, alors cet ensemble se contient lui-même, et comme il contient aussi l'ensemble vide, il n'y a donc pas de bijection sur lui-même, ce qui est contradictoire. Pour ma note personnelle : je n'aime pas beaucoup cette preuve, car la théorie de Zermelo-Fraenkel (que toute la terre abrège en ZF) permet de montrer qu'un ensemble ne se contient pas lui-même ; depuis justement cette notion de collection (on tourne en rond). Fort heureusement le « paradoxe de Russel » est levé grâce à une preuve utilisant cette notion de collection et ne s'autoréférant pas. Le paradoxe de Russel s'énonce bien différemment en philosophie, vous pouvez le consulter si vous le désirez.

Maintenant nous allons faire un peu de math (il faut de temps en temps non ?) pour énoncer les axiomes fondateurs des ensembles, on peut donner comme définition d'ensemble les objets de l'univers (au sens philosophique). C'est-à-dire énoncer la théorie ZF.

ZF consiste en la donnée de quatre voire cinq axiomes qui vont être difficile à mettre en défaut (contrairement au postulat d'Euclide) :

Le premier axiome est l'axiome d'extensionalité. Les axiomes vont tous avoir une forme horrible à lire, mais à chaque fois (ou presque) je donnerai une phrase pour le comprendre.

Bref, axiome d'extensionalité (E) : $\forall x \forall y (\forall z (z \in x \Leftrightarrow z \in y) \Longrightarrow x = y$

qui se lit : « Deux ensembles sont égaux ssi ils contiennent les mêmes éléments ».

L'axiome de la réunion (S) : $\forall x \exists y \forall z (z \in y \Leftrightarrow \exists t (t \in x \text{ et } z \in t))$

qui se lit : « Soit A un ensemble, il existe un ensemble B dont les éléments sont les éléments des éléments de A » ou plus simplement encore par : « Une réunion d'ensemble est un ensemble ».

L'axiome de l'ensemble des parties (PE) : $\forall x \exists y \forall z (z \in y \Leftrightarrow z \subset x)$

qui se lit : « les sous-ensembles d'un ensemble forment un ensemble ». On ajoutera qu'il y a bien sûr une définition du symbole d'inclusion : $\forall x \forall z (x \subset z \Leftrightarrow \forall t (t \in x \Longrightarrow t \in z))$

Et enfin le dernier axiome : le schéma de remplacement.

C'est le seul que je me permets de ne pas énoncer de façon exacte ni même de façon imagée car il est uniquement d'une nature « technique ». Il permet par exemple de démontrer l'existence d'un ensemble vide, et l'axiome E permet d'en démontrer l'unicité (comme quoi nous n'obtenons pas n'importe quoi !).

On ajoute parfois l'axiome de l'infini (AF) : nous en reparlerons. Disons juste sa métaphore associée « l'infini existe ».

On peut à partir de cette liste d'axiome s'autoriser ce qui est possible en général pour les ensembles : définir une différence, une intersection mais aussi un produit cartésien, cependant il faudra parler de l'axiome du choix afin de ne pas dire « trop » de bêtises trop vite.

Kurt Gödel travaillera beaucoup sur la notion de « consistance » des axiomes c'est-à-dire qu'en considérant un système d'axiomes donné, peut on aboutir sur une contradiction ? C'est-à-dire démontrer que $0 \neq 0$ est un théorème. Je n'expose pas la théorie de la démonstration, mais disons qu'une démonstration est un enchainement de résultat des axiomes du système, le dernier résultat est appelé théorème.

Plus précisément cette notion de contradiction est relative : on admet un système d'axiome, en ajoutant un axiome est que la théorie peut devenir contradictoire, en clair est-ce que $0 \neq 0$ est un théorème de ces axiomes ?

Si on retourne quelques sections en arrière vous trouverez un axiome surprenant : l'axiome du choix. Cet axiome (qui admet de bien multiples formulations) a posé de nombreux problèmes aux mathématiciens : le théorème de Banach-Tarski (cf section 4). Au point de remettre en cause l'évidence de cet axiome, cependant Gödel démontrera (en 1938 je crois) la consistance relative de cet axiome par rapport à la théorie : on ne pourra pas se contredire avec cet axiome (contrairement à quelques vidéos réalisées par des fanatiques religieux persuadés du mensonge des mathématiques et que 1 = 2, bref remarque à part). Le hic, c'est que si on suppose l'axiome du choix faux, on ne contredit toujours pas la théorie ZF. C'est donc au choix du mathématicien.

Il existe de nombreux axiomes qui sont ajoutés à la théorie ZF quand on veut montrer quelque théorème particulier, j'ai déjà cité notamment l'axiome de l'infini. Je vais en citer d'autres car ils sont essentiels en mathématique (pourquoi ne les a-t-on pas inclus dans la théorie ZF alors ? Tout simplement parce qu'historiquement, la théorie ZF n'a été « inventée » uniquement dans le but de définir correctement une théorie des ENSEMBLES, et non un besoin de construction de toutes les mathématiques…). Ces axiomes concernent plutôt la construction de structure bien plus avancée que les ensembles et les collections, comme par exemple le semi-groupe usuel N et c'est pour ça qu'avant de les annoncer, je dois m'attarder légèrement sur la théorie des ordinaux.

Deux ensembles sont dit équipotents s'il existe une bijection de l'un dans l'autre, en pratique pour trouver des bijections on peut soit la construire explicitement, soit considérer des structures (en fait abélienne, voir la section d'après), soit utiliser le théorème de Cantor-Bernstein : s'il existe une injection de A dans B, et une injection de B dans A alors il existe une bijection de A dans B. Cette relation d'équipotence est une relation d'équivalence et nous allons choisir un bon système de représentant (on ne prend pas n'importe quoi) : les ordinaux.

Définition d'un ordinal : C'est un ensemble bien ordonné pour la relation d'appartenance et transitif, c'est-à-dire, que toute partie (non vide) d'un ordinal admet un minimum, et tout élément de cet ensemble est inclus dans l'ensemble. C'est une définition un peu compliquée, mais quelques exemples rafraichiront ça :

Les ensembles suivants sont des ordinaux : $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, … contrairement à $\{\emptyset, \{\{\emptyset\}\}$.

Il y a de nombreuses propriétés sur les ordinaux comme par exemple : les éléments d'un ordinal sont des ordinaux, etc. Une essentielle : un ordinal n'appartient pas à lui-même.

Soient O et W deux ordinaux, alors on a l'une des ces trois assertions :

$O = W, O \in W \text{ ou } W \in O.$

Soit O un ordinal, alors le plus petit des ordinaux strictement plus grand se note O' et vaut exactement $O \cup \{O\}$, on l'appelle le successeur de O.

On a le théorème très important suivant : tout ensemble est équipotent à un ordinal. Le plus grand de ces ordinaux est appelé le cardinal de l'ensemble, et on note $\#E$ le cardinal de E.

Théorème de Cantor : $\#E < \#P(E)$.

On peut sommer des cardinaux considérant deux ensembles disjoints A et B, $\#A + \#B = \#(A \cup B)$. On peut aussi définir le produit depuis le produit cartésien, et la puissance depuis les applications de B vers A.

Voilà vous connaissez la notion d'ordinal, on peut maintenant revenir aux axiomes dont je vous ai parlé plus haut : un ordinal est dit fini si toute ordinal inclus non vide admet un prédécesseur (un prédécesseur de W est un ordinal O tel que $O' = W$, mais normalement vous l'aviez deviné). En réalité vous connaissez bien (et ce depuis très longtemps) les ordinaux finis, mais sous un autre nom : les entiers naturels. D'ailleurs en pratique, le plus petit ordinal on le note 0, son successeur on le note 1, puis on note $2 = 1'$, etc…

Axiome de l'infini : il existe un ordinal infini.

Si on suppose cet axiome vrai, alors on peut créer une notion d'ordinal limite (que je ne précise pas pour n'embrouiller personne) qui contient tous les ordinaux qui lui sont plus petit. Par exemple, dans les ordinaux infinis, il y a le plus petit des ordinaux limites, c'est celui qui contient tous les ordinaux finis, on le note $\mathbb{N}$, et on note $\aleph_0$ son cardinal. Et ton ensemble de cardinal $\aleph_0$ est dit dénombrable (on commence enfin à retomber sur nos pattes : ouf !).

Maintenant énonçons d'autres axiomes :

Un cardinal C est dit inaccessible si pour tout cardinal A plus petit que C on a $2^C < A$. Ca signifie métaphoriquement qu'il n'existe pas un énorme ensemble dont ses parties soient elles-même plus grosse que tous les cardinaux inférieurs. (Personnellement je symbolise dans ma tête un cardinal inaccessible comme un enorme trou…).

L'axiome d'accessibilité : tout cardinal est accessible (i.e. il n'y a pas de cardinaux inaccessibles).

La théorie ZF + l'axiome d'accessibilité est consistante (relativement à ZF).

L'hypothèse du continu : $\aleph_1 = 2^{\aleph_0}$.

$\aleph_1$ désigne un certain ordinal que je n'explique pas (trop long à expliquer) mais on peut se le représenter comme étant le cardinal de $\mathbb{R}$, d'où le nom « hypothèse du continu ». On a longtemps cru que l'hypothèse du continu était un théorème démontrable, c'était même un des 23 problèmes de Hilbert. Maintenant nous lui avons donné sa bonne place d'axiome…

Il existe un dernier axiome usuel, l'axiome de fondation qui énonce en grosso-modo qu'on ne peut pas construire une suite décroissante pour la relation d'appartenance. On encore qu'il n'existe pas de suite strictement décroissante d'entiers naturels, cette remarque est incroyablement féconde : elle permet de prouver des résultats par « descente infinie », j'ignore exactement où se situe le programme scolaire de L3 et avant (que ce soit prépa ou fac) donc j'ignore si vous avez déjà fais une descente infinie ; il y a juste certains qui ont fait l'option informatique qui l'ont déjà pratiqué en info en prépa (mais si, souvenez vous…). On se sert aussi beaucoup de l'axiome de fondation en théorie des graphes pour des fermetures transitives, etc. En réalité cet axiome est grosso-modo THE axiome pratique, mais presque tous le pratiquent sans le savoir… ☹.

C'est sur ça que je terminerais la section : je pense que le lecteur en a suffisemment vu sur l'origine des mathématiques, et je pense qu'il s'accordera avec moi sur le choix du mot « orgine ». On aurait pu s'enfoncer bien plus loin : il y a tant à dire, comme les théorèmes de Gödel, l'informatique théorique, la théorie des graphes, bref ces « petites maths » interressantes qui font d'ailleurs la passion de Benoit (quant à moi je préfère la partie suivante, mais c'est qu'une question de goût.

Lecture suggerée :

Théorie des ensembles (Jean-Louis Krivine). Je vous préviens d'avance, l'ouvrage (et tout ouvrage de théorie des ensembles) est indigeste, donc si vous ne vous sentez pas à la hauteur, passez à autre chose (et je ne donne pas souvent ce conseil). Ma copine m'a dit avoir trouvé un cours de théorie des ensembles sur Internet en tapant « Conférences de mathématiques » dans Google vidéo. Je ne l'ai pas visionnée, mais ça peut peut-être valloir le coup.

Un résumé du livre se trouve en powerpoint ici :

http://perso.ens-lyon.fr/pierre.lescanne/ENSEIGNEMENT/LOGIQUE/04-05/slides-ensembles.pdf

(Je trouve ça d'ailleurs étonant que certains auteurs font des pâles copies d'autres cours… mais bon au moins le pdf est sûr d'être gratuit).

Les mathématiques absolues…

Vous voici enfin dans la toute dernière section de mon texte. Ce que nous allons voir maintenant est ce qu'on appelle la « théorie des catégories », je vais utiliser des connaissances incluses dans toutes les parties précédentes (sauf les parties 5 et 6), cette théorie représente une synthèse de l'aboutissement mathématiques, elle n'aurait peut-être pas pu voir le jour sans les travaux exceptionnels des Bourbakistes. Je ne montrerais que l'entrée en la matière afin que tout lecteur ayant lu ce qui précède puisse comprendre.

Pourquoi les mathématiques « absolues » ?

Comme on peut le voir sur la « couverture », vous avez été au contact de structure toujours plus abstraite mais qu'importe les structures étudiées il existe des analogies entre elles, comme par exemple l'existence de morphisme conservant ces structures, parfois l'utilisation d'un noyau pour prouver l'injectivité, etc. On peut se demander alors s'il existe une généralisation de tous ces concepts ?

La réponse est : oui !

La notion clé est l'utilisation de ce qu'on appelle une « catégorie ». En voici une définition :

On appelle catégorie ce qui contient :

Une classe Ob appelée classe des objets, les éléments sont tout simplement appelés « objet » ;

Et pour tout couple d'objet (A, B), d'un ensemble $Hom(A, B)$ dont les éléments sont appelés « morphismes » de la catégorie ;

Et pour tout triplet d'objets (A, B, C), on peut associer une loi de composition :

$Hom(A, B) \times Hom(B, C) \to Hom(A, C),\ (f, g) \to g \circ f.$

De plus on ajoute les axiomes suivants :

Pour des objets bien choisis, $h \circ (g \circ f) = (h \circ g) \circ f$ (Associativité)

Pour tout objet A, il existe un morphisme id_A tel que, pour tout objet B et tout morphisme f de A vers B et tout morphisme g de B vers A on ait : $id_A\ o\ f = f$ et $g\ o\ id_A = g$.

Note 1 : La structure de classe est définie grâce à la théorie des ensembles vue plus haut, c'est nécessaire car nous pourront choisir la catégorie des « ensembles », qui ne sera pas un ensemble.

On peut déjà montrer si on le désire que pour tout objet A, il y a une unique identité id_A (la preuve est analogue que l'élément neutre d'un groupe).

Cette définition au premier abord semble très abstraite, c'est pour cela que je vais donner un nombre d'exemples qui je l'espère vont éclairer votre lanterne, on ajoutera ensuite des exemples originaux que vous ne pouviez pas trop deviner pour montrer la diffusion de cette branche.

On note $A \to B$ pour dire qu'il s'agit d'un morphisme de A vers B.

Premier exemple :

La catégorie des ensembles notée **Ens** admet les ensembles pour objet et les applications pour morphismes.

Deuxième exemple :

La catégorie des groupes notée **Gr** admet les groupes pour objet et les homomorphismes de groupes pour morphismes.

Troisième exemple :

La catégorie des espaces vectoriels notée **Ev** admet les espaces vectoriels pour objet et les applications linéaires pour morphismes.

Quatrième exemple :

La catégorie des espaces topologiques notée **Top** admet les espaces topologiques pour objets et les applications continues pour morphismes.

Cinquième exemple (exotique) :

Soit M un monoïde, on peut définir sur M une catégorie ayant pour objet M et les morphismes les éléments de M (la composition de deux morphismes est alors le produit de deux éléments).

Sixième exemple (exotique) :

La catégorie des matrices **Mat** où les objets sont les entiers naturels et les morphismes $m \to n$ les matrices à n lignes et m colonnes, la composition devenant le produit matriciel.

Septième exemple (exotique) :

Soit $(A, \leq)$ un ensemble ordonné, on peut définir une catégorie ayant pour objet les éléments de A et l'existence d'un morphisme $a \to b$ si $a \leq b$, la transitivité de $\leq$ permet de trouver la composition.

J'espère que ces exemples vous ont permis de vous rendre compte de ce que représente la notion de catégorie, vous voyez elle n'est pas si méchante. Le plus incroyable de la théorie des catégories est de pouvoir étudier des structures sans vraiment savoir à quoi elles peuvent ressembler, ni même connaitre les objets qui la compose.

On définit tout d'abord la notion de catégorie duale d'une catégorie, c'est une catégorie ayant les même objets mais dont les morphismes sont ceux en ayant « renversé » les flèches, ainsi $A \to B$ est un morphisme de la catégorie duale si $B \to A$ est un morphisme de la catégorie. Si on note C la catégorie, on note C^* la catégorie duale. De plus la composition est prise dans l'autre sens.

L'intérêt de cette notion va être de pouvoir définir rapidement les notions « duales » comme le dual d'un produit qui est une somme amalgamée (nous verrons ça plus loin, pas de panique !).

Première étude : l'étude des morphismes.

Nous allons généraliser la notion d'injection et de surjection dans les catégories. On restera donc dans une catégorie C donnée jusqu'à ce que je dispense cette condition.

Soit $f : X \to Y$ un morphisme entre deux objets X et Y.

On appelle section de f un morphisme $s : Y \to X$ tel que $s \circ f = id_X$.

On appelle rétraction de f un morphisme $r : Y \to X$ tel que $f \circ r = id_Y$.

La rétraction est la notion duale de celle de section. Si f possède une retraction r et une section s alors $r = s$, et on dit que f est un isomorphisme et que X et Y sont isomorphes.

Plus généralement, on appelle monomorphisme un morphisme vérifiant : $f \circ g = f \circ h \implies g = h$.

De même on appelle épimorphisme un morphisme vérifiant : $g \circ f = h \circ f \implies g = h$.

Ces deux notions généralisent les notions d'injection et de surjection en remarquant que dans **Ens** par exemple, les monomorphismes sont les injections et les épimorphismes des surjections. Ce sera aussi le cas dans les catégories **Top**, **Gr**, **Ann** (des anneaux), bref ce qu'il y a d'usuel.

On remarquera que nous n'avons pas définis un isomorphisme comme un monomorphisme étant un épimophisme, car cette remarque est déjà fausse dans la catégorie **Ann**.

De même dans la catégorie **Top**, il y a des bijections, continues dans les deux sens qui ne sont pas des homéomorphismes (je vous rassure : c'est impossible dans les espaces métriques).

On retrouve nos propriétés usuelles : si $g \circ f$ est un monomorphisme alors f l'est aussi (analogie avec les injections) ; la composée de deux monomorphismes est un monomorphisme. Un morphisme possédant une rétraction est un monomorphisme.

Ces résultats dans la catégorie duale signifient par exemple que si $f \circ g$ est un épimorphisme alors f en est un. Voilà l'intérêt des catégories duales, c'est de pouvoir énoncer des énoncés qui se ressemblent qu'une seule fois.

Deuxième étude : les objets.

Certains objets sont fondamentaux en théorie des catégories. C'est le cas par exemple des objets initiaux, finaux et nuls.

Un objet X d'une catégorie C est dit final si pour tout objet Y il existe un unique morphisme de $Y \to X$.

De même un objet initial est un objet X tel que pour tout objet Y il existe un unique morphisme $X \to Y$.

Par exemple : Dans la catégorie **Ens**, $\{0\}$ est final (en effet, ya une unique application valant toujours 0 !) et $\emptyset$ est initial. Dans la catégorie **Ann**, $\mathbb{Z}$ est initial. Dans la catégorie **Gr** des groupes, un groupe

réduit à son élément neutre est initial et final. Dans la catégorie $(A, \leq)$ un objet final est un plus grand élément, un objet initial un plus petit élément.

Théorème : Deux objets finaux sont isomorphes. Deux objets initiaux sont isomorphes.

On appelle objet nul un objet à la fois initial et final, on note toujours 0 un objet nul.

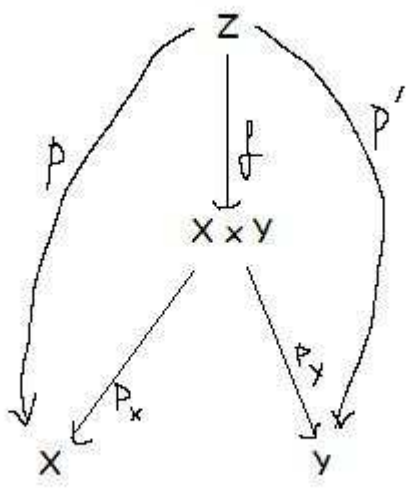
Par exemple, **Gr** admet un objet nul, **Ens** par contre non, car si c'était le cas, $\emptyset$ et $\{0\}$ serait en bijection. De même dans la catégorie des **Ev** l'espace vectoriel nul est un objet nul.

Nous allons maintenant généraliser des notions que vous connaissez bien : les produits, puis les sommes et enfin la notion de noyau.

Commençons donc par le produit, la définition vous sera surprenante, mais en réalité même dans des espaces usuels on ne peut pas toujours définir simplement le produit, c'est le cas par exemple de la topologie produit de deux topologies. De plus la définition que je vais donner est aussi une caractérisation pour les autres produits comme le produit cartésien d'ensemble, le produit de groupe, d'espace vectoriel, etc.

Soit X et Y deux objets ; on dit qu'un objet $X \times Y$ est un produit de X et Y s'il existe deux morphismes (appelés projections) $p_X : X \times Y \to X$ et $p_Y : X \times Y \to Y$, tel que s'il existe un autre triplet (Z, p, p') avec Z un objet ; $p : Z \to X$ et $p' : Z \to Y$ alors il existe un unique morphisme $f : X \times Y \to Z$ tel que $p = p_X \, o \, f$ et $p' = p_Y \, o \, f$.

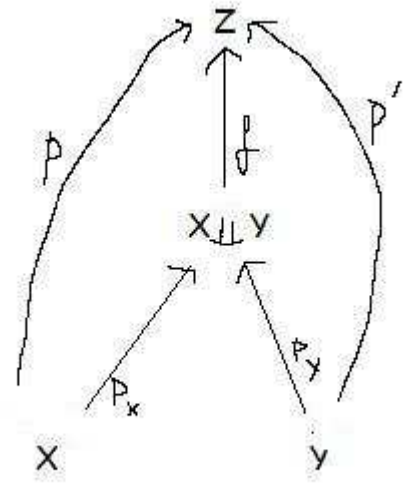
Ce qu'on « symbolise » par le diagramme :



Apprendre à lire un diagramme :

On dit que le diagramme est commutatif si entre deux objets quelconques A et B, si on a deux chemins allant de A vers B, alors les composées des applications formant les arcs selon les deux chemins sont égales, c'est ce qui exprime par exemple au dessus que $p_X \, o \, f = p$.

On définit de la même manière la somme (en fait elle est dite somme amalgamée, mais c'est chiant à dire) en renversant les flèches :



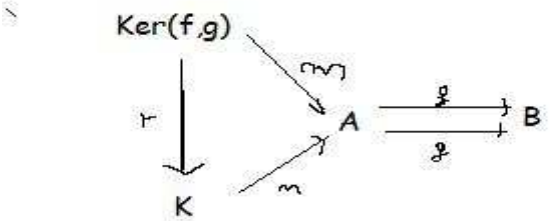
Par contre, afin de ne pas confondre, les morphismes p_X et p_Y sont appelés injections. On remarquera la notation « union disjointe » ; ce n'est pas un choix aléatoire : dans la catégorie des ensembles la somme de deux ensembles est leur union disjointe.

Exemple : Dans la catégorie des **Ev**, une somme de deux espaces vectoriels est leur somme directe.

Afin une dernière notion que vous connaissez un peu : la notion de noyau. De même cette définition vous semblera un peu bizarre, mais il se trouvera être nécessaire de définir ça ainsi à cause des catégories « bizarres » qui peut existe.

Soient f et g deux morphismes de A vers B. On appelle noyau de f et g, un couple $(Ker(f,g)\,,m)$ où $Ker(f,g)$ est un objet et m un morphisme $Ker(f,g) \to A$ avec $f \circ m = g \circ m$ tel que pour un autre couple $(K\,,n)$ ayant les mêmes propriétés il existe un unique morphisme $r : Ker(f,g) \to K$ tel que $n \circ r = m$.

On représente ceci par le diagramme commutatif suivant :



En raccourcit on parle du noyau pour désigner soit l'objet $Ker(f,g)$, soit le morphisme m, c'est selon le contexte. Le morphisme m est parfois appelé inclusion et c'est toujours un monomorphisme.

Par exemple : le noyau de deux morphismes f et g dans la catégorie **Ev** est $Ker(f - g)$. De même le noyau de deux applications dans la catégorie Ens est l'ensemble $\{f(x) = g(x), x \in A\}$.

Toutes les notions que nous avons introduites, noyau, produit, somme, ont la propriété remarquable d'être toujours le même à un isomorphisme près : deux noyaux des mêmes morphismes sont isomorphes (par exemple).

Vous aurez aussi constaté la présence de la phrase « il existe un unique morphisme tel que », cette phrase qualifie l'objet d'une propriété dite « propriété universelle ».

Voilà comment fonctionne globalement la théorie des catégories…

Une des principales notions des catégories est de construire des « variétés », peu importe comment elles sont. C'est ce qu'on appelle de façon générale la « théorie des faisceaux », j'introduirais juste le principe de faisceau (et vous verrez, ce sera déjà pas mal abstrait pour des gens qui n'ont pas l'habitude) et nous continuerons sur les exemples des variétés différentielles, et algébriques. Avant d'introduire ces notions nous allons devoir amener la notion de foncteur (pour moi l'un des plus objet mathématique, en réalité le foncteur H de l'algèbre homologique, mais ce n'est pas le sujet).

Un foncteur c'est (de façon très simplifiée) une « fonction » entre des catégories. De manière plus formelle :

On considère deux catégories C et C', un foncteur F de C dans C' est une opération qui à un objet X de C associe un objet $F(X)$ de C' ; qui a un morphisme $f : X \to Y$ de C associe un morphisme $F(f) : F(X) \to F(Y)$. De plus le foncteur doit vérifier $F(id_X) = id_{F(X)}$ et l'une des deux conditions suivantes :

i) $F(f \circ g) = F(f) \circ F(g)$
ii) $F(f \circ g) = F(g) \circ F(f)$

Dans la condition $i)$, le foncteur est dit « covariant », dans la seconde condition, il est dit « contravariant ». Bien souvent nous avons à faire à des foncteurs contravariants mais ce n'est pas toujours le cas. Parmi les foncteurs fondamentaux, citons les plus simples :

Exemple 1 : le foncteur identité, qui vérifie « F$(X) = X$ et $F(f) = f$ »

Exemple 2 : le foncteur oubli, par exemple **Gr** → **Ens**, qui fait « oublier » la structure pour revenir à un ensemble.

Exemple 3 : le foncteur inclusion, par exemple **Ab** → **Gr**.

Exemple 4 : On peut considérer le foncteur discrétisation **Ens** → **Top** qui à un ensemble associe l'espace topologique discret associé. De manière analogue pour le foncteur grossierisation.

Exemple 5 : Il existe aussi des foncteurs Hom que je ne définie pas mais qui sont essentiel en théorie des catégories.

Il est possible aussi de définir des « transformations de foncteurs », en fait appelées « transformations naturelles ». Je ne les définies pas car ça peut vite devenir trop abstrait pour des gens qui n'ont même pas encore vu ne serait ce que la topologie ; mais disons qu'il est possible de définir des foncteurs « inversibles », dit alors « équivalence de catégories ». On peut en voir un qui semble indiquer pourquoi on en parle : les catégories **Mat** (des matrices) et **Evf** (des espaces vectoriels finis) sont équivalentes. Bref il y a tant à dire sur les foncteurs comme la fidélité, la surjectivité etc. mais c'est une section, pas un cours complet.

D'autres exemples de foncteurs :

A la section 1, nous avons parlé d'une structure de groupe construite depuis un espace topologique connexe par arc (en particulier pour vérifier la simple connexité), cette opération est un foncteur, appelé foncteur de Poincaré.

A la section 7, nous avons défini pour tout ensemble algébrique une algèbre G(V), cette opération est encore un foncteur.

Comme quoi des foncteurs, il y en a !

Bref nous allons maintenant nous intéresser à un type de foncteur très particulier : les faisceaux.

Pour simplifier ce que nous allons voir, partons de ce qui a été fait en cours de MA601 avec M.Raoux.

On considère une équation différentielle pour une fonction $\mathbb{R} \to \mathbb{R}$. On suppose que cette équation vérifie le théorème de Cauchy-Lipchitz. Si nous résolvons cette équation sur un intervalle I, puis sur un intervalle J tel que $I \cap J$ soit non vide, alors il existe une unique solution qui coïncide sur $I \cup J$ tel que sa restriction à I soit la solution sur I et sa restriction à J soit la solution sur J. On appelle ce principe un « recollement » des solutions.

Nous pouvons voir dans le paragraphe précédent deux nouveaux mots : restriction et recollement.

Ce sont ces mots qui déterminent la définition d'un faisceau. Disons vulgairement qu'un faisceau sur un espace topologique X permet d'associer à un ouvert donné une fonction, telle que la restriction soit possible à un ouvert plus petit, et que sur deux ouverts qui se croisent il y ait une unique fonction telle que restreinte à chaque ouvert on trouve les fonctions sur les ouverts de départ.

De façon plus mathématique :

Un espace topologique X peut être vu comme une catégorie : ses objets sont les ouverts et ses morphismes sont les « relations d'inclusions » : $U \to V$ ssi $U \subset V$.

On appelle préfaisceau sur un espace topologique X un foncteur contravariant de la catégorie X dans une catégorie C. Les morphismes images sont appelés restrictions.

Je ne donnerai pas la définition la plus générale de faisceau car elle utilise les suites exactes (que j'introduis après) et elle se trouve être très abstraite. Je vais donc donner la définition d'un faisceau d'anneau (car en pratique ce sont ces faisceaux qu'on voit le plus souvent, vous verrez pourquoi).

On considère un préfaisceau F sur un espace topologique X à valeur dans **Ann**.

On dit que F est un faisceau (d'anneau) si et seulement si :

Pour un ouvert U donné de X, on recouvre U par des ouverts (pour simplifier deux seulement) V et W et pour tout élément sV de $F(V)$ et sW de $F(W)$ telle que $(F(V) \to F(V \cap W))(sV) = (F(W) \to F(V \cap W))(sW)$ alors il existe un unique s de $F(U)$ tel que $(F(U) \to F(V))(s) = sV$ et $(F(U) \to F(W))(s) = sW$.

La définition est déjà lourde, mais disons pour « mentaliser » l'idée, qu'elle exprime le principe de recollement. Je vous laisser mesurer le cas général.

Voyons quelques exemples :

On peut définir grosso-modo une variété comme la donné d'un espace topologique X (avec une base d'ouvert B), on y définit les bonnes fonctions sur cette base d'ouvert ou fonction régulière, on peut les engendrer par le principe de recollement à tout ouvert, et les fonctions régulières forment un anneau (faisceau d'anneau !), pour que si f et g sont de bonnes fonctions, alors $f + g$ en est une, de même pour le produit.

Les bonnes fonctions sont par exemple les fonctions différentiables ou polynômiales.

Dans le premier cas c'est ce qu'on appellera une variété différentiable, dans le second une variété algébrique. C'est ce qui généralise les sections 7 et 9.

Je ne donne pas la définition exacte car un peu lourde à énoncer parce qu'il faut introduire de nombreuses notations, surtout dans le cas algébrique.

Maintenant vous pouvez considérer l'intérêt des faisceaux.

Nous allons quitter ces notions pour faire quelques pas en arrière vers la notion de catégories et y définir le terme « catégorie abélienne ». Ce sont les catégories qui permettent l'étude de ce qu'on nomme « l'algèbre homologique ».

Soit C une catégorie.

Supposons qu'elle admette un objet 0, un morphisme $A \to B$ est dit nul s'il existe $i : A \to 0$ et $j : 0 \to B$ tel que : $f = j \circ i$. C'est-à-dire qu'on « envoie » tous les éléments de A sur l'objet nul, puis depuis l'objet nul on revient dans B, par exemple dans **Gr**, l'objet nul étant le groupe réduit à son élément neutre, les morphismes nuls sont les morphismes qui envoie tous les éléments de A sur l'élément neutre de B.

Grâce aux morphismes nuls, on peut définir ce qu'on appelle un VRAI noyau. Soit f un morphisme un noyau de f est un noyau des morphismes f et 0 (morphisme nul). Par unicité à isomorphisme près, on le notera $Ker\ f$.

On reprend une catégorie C, on dit qu'elle est pré-additive si on peut munir chaque $Hom(A, B)$ (étant des ensembles !) d'une structure de groupe vérifiant les bonnes distributivités avec la composition.

Soit C une catégorie pré-additive, elle est dite abélienne si elle vérifie les axiomes supplémentaires suivants :

1) C possède un objet nul.
2) Les produits et les sommes amalgamées existent.
3) Tous les morphismes ont un noyau (et un conoyau, dual d'un noyau).
4) Tous les monomorphismes sont des noyaux, et tous les épimorphismes des conoyaux.

En pratique on considère grandement deux catégories abéliennes : la catégorie **Ab** des groupes abéliens et la catégorie **Mod** des modules sur un anneau (un module c'est la notion étendue d'espace vectoriel pour des anneaux). Dans le cas **Ab** précisons les axiomes :

L'objet nul est le groupe trivial $\{e\}$.

Les produits (et les sommes) sont définis au sens où le comprend, comme en théorie des groupes MA502 avec M Alev

Les noyaux sont les couples $(Ker\, f, i)$ où $Ker\, f$ est le noyau tel que nous le connaissons et i la simple inclusion.

Le dernier point est un poil plus délicat : chaque monomorphisme est le noyau de $\pi : Y \to Y/f(X)$. Je n'exprime pas le cas des épimorphismes.

Dans une catégorie abélienne on retrouve des résultats connus :

f est un monomorphisme ssi son noyau est l'objet nul, i.e. $Ker\, f = 0$. (On a bien généralisé l'injection).

f est un isomorphisme ssi c'est un épimophisme ET un monomorphisme (hé oui c'est faux dans le cas général : la catégorie **Ann**).

Mais SURTOUT nous avons la proposition très puissante suivante :

Factorisation par l'image : Soit $f : A \to B$ un morphisme, il existe un monomorphisme $m : C \to B$ et un épimorphisme $e : A \to C$ tel que $f = m \circ e$. Et cette factorisation vérifie « la propriété universelle ».

Le morphisme m est appelé image de f (en réalité il existe une définition plus générale, mais dans une catégorie abélienne, ça devient une propriété, donc pour ne pas compliqué je l'ai défini seulement maintenant). Du coup l'image existe toujours (ce qui n'est pas le cas en général !).

Pour simplifier les notations, si nous avons un monomorphisme $f : A \to B$, son conoyau est noté B/A.

En fait la factorisation par l'image précise comment on obtient m et e, disons que ça permettra d'avoir le premier théorème d'isomorphisme par exemple que le noyau de $M \to M/Ker\, f$ est $Ker\, f$.

Nous nous approchons petit à petit de la notion d'algèbre homologique. La dénomination algèbre est justifiée puisque nous travaillerons toujours sur une catégorie abélienne (dans votre tête **Ab**) qui admet de bons objets qu'on aime bien comme des quotients, des noyaux, etc.

La notion suivante à voir est celle de suite exacte, très importante en algèbre homologique.

Soit C une catégorie abélienne.

Soient M, M' et M'' trois objets de cette catégorie avec des morphismes :

$M' \underset{f}{\underbrace{\rightarrow}} M \underset{g}{\underbrace{\rightarrow}} M''$. (Normalement il n'y a pas l'accolade, mais j'utilise Word et non du Latex, je promets cependant d'apprendre bientôt le Latex pour faire des choses correctes !).

Cette « suite » est dite exacte si et seulement si $Im\ g\ =\ Ker\ f$. (On généraliser la notion mais on s'en fout).

Pour une suite infinie : … $M_1 \underset{f_1}{\underbrace{\rightarrow}} M_2 \underset{f_2}{\underbrace{\rightarrow}} M_3$ …

On dit qu'elle est exacte si elle exacte pour chaque paire de morphismes consécutifs.

On a par exemple :

$f : A \mathrel{-}> B$ est un monomorphisme ssi la suite $0 \rightarrow A \underset{f}{\underbrace{\rightarrow}} B$ est exacte. (Et la notation duale pour le cas épimorphisme).

De même, soit une suite $A \underset{g}{\underbrace{\rightarrow}} B \underset{h}{\underbrace{\rightarrow}} C$. Soit (e, m) et (f, n) les factorisations par image de g et h. Alors la suite "g, h" est exacte si et seulement si "m, f" en est une.

De manière plus générale on appelle complexe une suite C de la forme :

$$\ldots \rightarrow M_{i+1} \underset{d_{i+1}}{\underbrace{\rightarrow}} M_i \underset{d_i}{\underbrace{\rightarrow}} M_{i-1} \rightarrow \ldots$$

Où pour tout i, $d_i \circ d_{i+1} = 0$. Les d_i sont appelées les différentielles du complexe.

Toute suite exacte est un complexe car les compositions s'annulent.

Dans un complexe donné on définit le i-ième objet d'homologie comme le quotient $Ker\ d_i\ /\ Im\ d_{i+1}$ et on le note $H_i(C)$. Lorsqu'une suite est exacte en M_i alors $H_i(C)$ est l'objet nul.

C'est là le principe de l'algèbre homologique !

Le but est d'étudier la suite des H_i permettant de trouver les exactitudes de la suite de départ et donc de comprendre comment est formé un certain objet (on utilise des résolutions mais peu importe). Cette mesure de l'exactitude permet de construire de nombreux invariants comme la dimension. Ces procédés homologique (et de façon dual, cohomologique) permettent de démontrer énormément de chose pour des grosses structures comme un non homéomorphisme entre $\mathbb{R}^n$ et $\mathbb{R}^m$ grâce aux rétractions, qu'importe.

C’est sur ce sujet que je termine cette section devenue bien longue (c’est mon thème « préféré »).

Je vous laisse méditer sur la diffusion de cette branche par un rappel :

On a utilisé la section 1 pour voir le foncteur de Poincaré, la section 10 pour l’invariant de la dimension, les sections 7 et 9 pour construire la notion de variété, la section 11 pour définir correctement la notion de catégorie.

Lecture conseillée : Algèbre (Serge Lang). PS : bonne chance…

Ouvrages généraux.

- L'atlas des mathématiques. L'avantage de ce livre c'est qu'il est jonché (une page sur deux) de schémas plein de couleurs, aidant grandement à des representations mentales. Défaut : ce n'est en aucun cas un cours.
- Le dictionnaire des mathématiques. Très bien réalisé, mais c'est comme son nom l'indique, un dictionnaire, donc pas très instructif, c'est juste une bonne référence sous la main, bien que ses connaissances soient trop culturelles elles sont trop vite limitées.
- Les maths pour l'agreg. Un incontournable, mais un peu cher.

Maintenant les ouvrages vulgarisés, vous n'apprendrez rien pouvant servir en cours, mais c'est ceux que je conseille le plus si vous n'êtes pas passionné par les mathématiques.

- Le dernier théorème de Fermat.
- Les 7 problèmes du millénaire. (C'est ce livre qui m'a lancé dans les mathématiques définitivement, ça a aussi l'air d'être le cas d'Adrien Didelet, juste un petit coup de pouce de ma part la lecture de ce livre et il est parti…).

CONCLUSION

Vous voilà arrivé en fin du texte, si vous avez tout lu je vous félicite grandement. J'ai pris plaisir à le rédiger pour ma classe mais je tiens juste à ajouter une petite note : j'avais fini le document lorsque j'ai mal enregistré le document et j'ai donc perdu 3h de données, ecoeuré il m'a fallut une semaine pour m'y remettre, ce qui explique les deux semaines de retard par rapport à ma « promesse » vis-à-vis l'attente de certains élèves. J'avais aussi fais une page de remerciement au début que j'ai au final effacée faute de bonne écriture. Je remercie donc ici les élèves qui prennent plaisir à lire le document et les professeurs qui m'ont motivé implicitement dans cette direction.

J'espère que ce texte fut à la fois clair et précis pour le lecteur, j'avoue ne pas avoir fais du mieux que je pouvais, car de manière générale je m'y prend bien autrement, mais les élèves à qui ce texte étaient destinés ont des connaissances mathématiques et j'en ai profité pour innover un peu ma façon d'écrire. D'ailleurs je préfère beaucoup en parler à l'oral, car les mathématiques sont truffées d'histoires croustillantes, d'où les deux derniers livres dans la liste des ouvrages généraux. J'aimerais tant que les mathématiques soient plus ouvertes au public. D'ailleurs un petit mot dessus, il serait tant d'annoncer à ceux qui propose les programmes que les mathématiques interressent les gens pour leurs histoires et non leur utilité, alors qu'en L, les élèves les plus démotivés des maths doivent apprendre les maths les plus barbantes qui soient… Et je donnerai beaucoup pour que les lecteurs de ce texte, normalement futurs profs, donnent à leurs élèves une vision moins scolaire des mathématiques et consacrent, même un temps minime, à conter quelques histoires incroyables des mathématiques… Je vous assure qu'avant ça, moi-même, je n'aimais pas les mathématiques, au point de vouloir faire L encore en $4^{\text{ième}}$.

Si vous avez une question, vous pouvez bien sûr me la poser en face à face pour avoir un éclaircissement ou une remarque, ou même juste venir me parler…

Je tiens aussi à m'excuser à certaines personnes de ne pas avoir les domaines qui les passionnent comme par exemple les probabilités (j'aurais pu parler des chaînes de Markov ou encore des percolations) mais je n'avais pas envie de mettre 36000 choses et puis c'est MA selection, si elle ne vous plait pas, vous n'avez qu'à rédiger votre propre document…

Voilà, c'est tout ce que j'avais à dire.