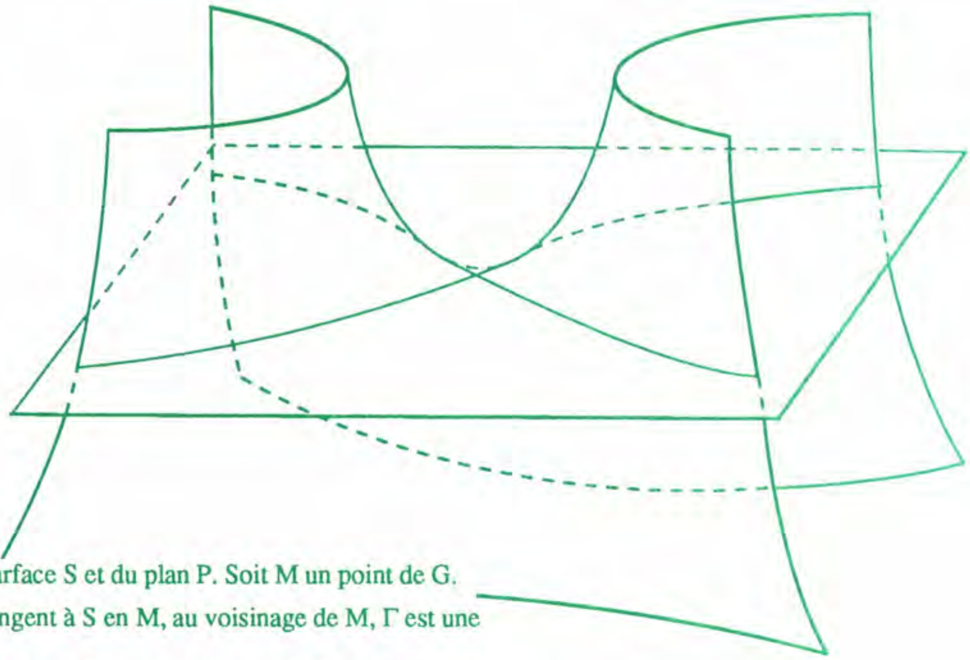
6 F

Université Nancy 1
Département de mathématiques



Soit Γ l'intersection de la surface S et du plan P. Soit M un point de G.

Si P n'est pas le plan tangent à S en M, au voisinage de M, Γ est une courbe régulière.

Si P est le plan tangent à S en M, et si la courbure de S en M est négative , alors au voisinage de M, Γ est une courbe qui a un point double en M.

S.M.2 M.P.I.

Cours de Mathématiques

Année 1989-90

Fascicule 2

Tirage 1992

© Edité et imprimé par l'**Institut de Recherche sur l'Enseignement des Mathématiques** - (Université de Nancy I - Faculté des Sciences) - B.P. 239 - 54506 VANDOEUVRE-les-NANCY Cedex
Dépôt légal : 2e trimestre 1991
n° de la publication : 2-85406-124-1
Responsable de la publication : Le Directeur de l'IREM, Bernard ANDRE

Avertissement

Ce cours a été conçu pour une section de seconde année de DEUG SM, à l'université Nancy I, en 1989-90. Il s'agit d'une section (dénommée Math-Physique-Informatique) à horaire de mathématiques renforcé (180 à 200h, options exclues). La majorité des étudiants concernés se destinent à une licence de mathématiques, quelques uns veulent entreprendre des études de physique, d'informatique, de mécanique, ou entrer dans une école d'ingénieurs. Il s'agit donc d'un public qui, n'ayant pas été habitué à l'abstraction au niveau du lycée, sera appelé à manipuler dès l'année suivante des notions purement théoriques. Notons que beaucoup d'étudiants suivent en outre une option probabilité, et une option algèbre et géométrie; ceci explique que certaines structures fondamentales - comme les groupes par exemple - n'apparaissent pas ici.

Table des matières du premier fascicule.

Séries de vecteurs et séries de fonctions

Au chapitre 10 nous avons défini une notion de suite convergente dans tout espace E muni d'une norme N. Nous pouvons nous demander si, comme dans le cas de $\mathbb{R}$, il est possible de démontrer la convergence d'une suite sans connaître sa limite.

La possibilité de prévoir la convergence d'une suite numérique sans en connaitre la limite nous a servi à démontrer des "théorèmes d'existence": théorème de la valeur intermédiaire (chapitre 8 §2) puis les théorèmes relatifs aux fonctions continues sur un compact (chapitre 10 §7, et §3). De la même façon la possibilité de prévoir "à priori" la convergence d'une suite de fonctions, nous servira à démontrer l'existence de (fonctions) solutions de certains problèmes, par exemple d'équations différentielles. Le processus sera toujours le même, on construit d'abord une suite de "solutions approchées", puis on démontre que cette suite a une limite qui est la solution (exacte) cherchée. Il en résulte, outre l'assurance que le problème posé a une solution, une suite qui en donne des approximations aussi bonnes que l'on veut. Dans la suite de ce cours on rencontrera plusieurs illustrations de cette technique (au §7 ci-dessous, puis au chapitre 12 §6, enfin au chapitre 20 §1).

Observons que, dans le cas des espaces vectoriels, il n'est pas question de généraliser le critère des suites croissantes majorées, puisque, E n'étant pas ordonné, la notion de sens de variation d'une suite d'éléments de E n'a aucune signification. C'est donc la condition de Cauchy que nous généraliserons.

§ 1 : Espaces complets

Nous dirons que $(u_n)_{n\in\mathbb{N}}$ est une suite de Cauchy pour la norme N si:

$$(\forall\varepsilon > 0)\ (\exists N_0)\ (\forall n)\ (\forall p)\ (n \geq N_0 \text{ et } p \geq N_0) \text{ implique } N(u_n-u_p) \leq \varepsilon.$$

Toute suite convergente pour N est une suite de Cauchy pour N (démonstration analogue à celle qui a été faite au chap.8 §2).*Mais il peut arriver que la réciproque ne soit pas vraie. Plus précisément:*

- *Si toute suite de Cauchy pour N, est convergente pour N, on dit que E est complet pour N.*
- *S'il existe des suites de Cauchy pour N, qui ne sont pas convergentes pour N, on dit que E n'est pas complet pour N.*

Exemples d'espaces complets

Soit K un sous-ensemble compact de $\mathbb{R}^n$, et soit $E = \mathscr{C}^0(K,\mathbb{R})$ (ou $\mathscr{C}^0(K,\mathbb{C})$). Munissons le de la norme uniforme $N_\infty(f) = \sup\limits_{x\in K}|f(x)|$, nous obtenons un espace complet. (la démonstration est au §2)

Tout espace de dimension finie est complet (quelle que soit la norme N choisie)

<u>Démonstration</u>: Soit $\{e_i\}$ une base de E, nous lui avons associé la norme $N_\infty(\sum\limits_i x^ie_i) = \sup\limits_i |x^i|$. Et nous savons qu'il existe $K(> 0)$ tel que $N_\infty \leq KN$. Si $(u_n)_{n\in\mathbb{N}}$ est une suite de Cauchy pour N, nous pouvons affirmer que:

$$(\forall\varepsilon)\ (\exists N),(\forall n)\ (\forall p)\ \ (n \geq N \text{ et } p \geq N) \text{ implique } N(u_n-u_p) \leq \frac{\varepsilon}{K}.$$

Ce qui entraîne $N_\infty(u_n-u_p) \leq \varepsilon$, et donc $(\forall i)\ |u_n^i-u_p^i| \leq \varepsilon$ (où $(u_n^i)_{n\in \mathbb{N}}$ est la i-ème suite coordonnée de $(u_n)_{n\in \mathbb{N}}$). Ainsi les suites coordonnées de $(u_n)_{n\in \mathbb{N}}$ sont de Cauchy; donc elles convergent (car $\mathbb{R}$ est complet). Donc $(u_n)_{n\in \mathbb{N}}$ converge pour toutes les normes sur E (chap.10 § 2). Ce qui termine la démonstration.

§ 2 : L'espace $\mathscr{C}^0(K,\mathbb{R})$ muni de la norme uniforme.

Soit K un sous-ensemble compact de $\mathbb{R}^n$ (muni d'une norme ν), et soit $E = \mathscr{C}^0(K,\mathbb{R})$ (ou $\mathscr{C}^0(K,\mathbb{C})$) l'espace des fonctions continues sur K.Posons $N_\infty(f) = \sup_{x\in K}|f(x)|$ (c'est parce que K est compact, que nous pouvons affirmer que, pour toute f continue, $\sup_{x\in K}|f(x)|$ existe). Nous obtenons une norme qui fait de E un espace complet.

La démonstration est un peu longue. Elle comporte deux étapes. Nous considérons une suite $(f_n)_{n\in \mathbb{N}}$ qui est de Cauchy pour N_∞.

<u>Etape 1</u> : Construction de la limite f. Pour tout x, $N_\infty(f_n-f_p) \leq \varepsilon$ implique $|f_n(x)-f_p(x)| \leq \varepsilon$; donc le fait que $(f_n)_{n\in \mathbb{N}}$ soit une suite de Cauchy pour N_∞, entraine que $(f_n(x))_{n\in \mathbb{N}}$ est une suite de Cauchy numérique, donc une suite convergente. Nous appellerons f(x) sa limite. Nous avons ainsi défini une application f de K dans $\mathbb{R}$.

Nous savons que: $(\forall\varepsilon>0)\ (\exists N_0)\ (\forall n)\ (\forall p)\ (n \geq N_0 \text{ et } p \geq N_0)$ implique $N_\infty(f_n - f_p) \leq \varepsilon$.

C'est à dire: $(\forall\varepsilon>0)\ (\exists N_0)\ (\forall n)\ (\forall p)\ (\forall x \in K)\ (n \geq N_0 \text{ et } p \geq N_0)$ implique $|f_n(x)-f_p(x)| \leq \varepsilon$.

Et puisque f(x) est la limite de $(f_p(x))_{p\in \mathbb{N}}$, nous aurons (en passant à la limite sur p):

(i) $(\forall\varepsilon > 0)\ (\exists N_0)\ (\forall n)\ (\forall x\in K)\ (n \geq N_0)$ implique $|f_n(x)-f(x)| \leq \varepsilon$.

Remarque: Si nous savions que f est continue, cette formule (i) nous permettrait d'affirmer que

$$(\forall\varepsilon>0)\ (\exists N_0)\ (\forall n)\ (n \geq N_0) \text{ implique } N_\infty(f_n-f) = \sup_{x\in K}|f_n(x)-f(x)| \leq \varepsilon.$$

Et nous aurions ainsi démontré que $(f_n)_{n\in \mathbb{N}}$ converge vers f pour N_∞. Il nous reste donc à démontrer que f est continue.

<u>Etape 2</u> : Montrons que la fonction f est continue. Choisissons x dans K et $\alpha>0$; pour montrer la continuité en x, il nous faut trouver des conditions sur h pour que $|f(x)-f(x+h)| \leq \alpha$.

Ecrivons: $|f(x)-f(x+h)| \leq |f(x)-f_n(x)| + |f_n(x)-f_n(x+h)| + |f_n(x+h)-f(x+h)|$.

En choisissant n de telle façon que $(\forall x)$, $|f_n(x)-f(x)| \leq \frac{\alpha}{3}$ (ce qui est possible d'après (i)), nous obtenons:

$$|f(x)-f(x+h)| \leq \frac{2\alpha}{3} + |f_n(x)-f_n(x+h)|.$$

Pour réaliser l'inégalité voulue, nous allons rendre $|f_n(x)-f_n(x+h)|$ inférieur ou égal à $\frac{\alpha}{3}$.

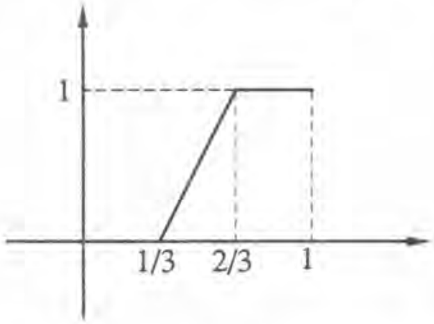
Puisque f_n est continue en x, il existe η tel que $\nu(h) \leq \eta$ implique $|f_n(x)-f(x+h)| \leq \frac{\alpha}{3}$. On en déduit que si $\nu(h) \leq \eta$, alors $|f(x)-f(x+h)| \leq \alpha$. Ce qui prouve la continuité de f, et - compte tenu de la remarque qui termine la 1ère étape - termine la démonstration.

§ 3 : Exemples d'espaces non complets

L'espace des fonctions continues sur [0,1] muni de la norme $N_1(f) = \int_0^1 |f(t)|dt$ *n'est pas complet. L'espace des fonctions continues sur [0,1] muni de la norme* N_2 *n'est pas complet.*

Négligeant de démontrer la première affirmation, donnons une preuve de la seconde. Pour cela nous allons exhiber une suite de fonctions qui est de Cauchy pour N_2 et n'a pas de limite pour N_2.

Soit f définie par $\begin{cases} f(t) = 0 \text{ si } t \leq 1/3 \\ f(t) = 3t-1 \text{ si } 1/3 \leq t \leq 2/3 \\ f(t) = 1 \text{ si } t \geq 2/3 \end{cases}$



Et soit f^n la puissance n-ième de f.

Exercice a: Calculer $\int_0^1 f^n(t)dt$. Montrer que, pour $p \geq 0$, $f^{n+p} \leq f^n$. En déduire $N_2(f^n - f^p)$ et $N_1(f^n - f^p)$. Montrer que $(f^n)_{n\in\mathbb{N}}$ est une suite de Cauchy pour N_2 et pour N_1.

Exercice b: Supposons que la suite $(f^n)_{n\in\mathbb{N}}$ converge, pour N_2, vers une limite φ (continue sur [0,1]) et montrons que nous obtenons une situation absurde.

1) Démontrer que $(\forall t)$ $\int_0^t (f^n-\varphi)^2(\theta)d\theta \leq \int_0^1 (f^n-\varphi)^2(\theta)d\theta$. En raisonnant dans $\mathscr{C}([0,t],\mathbb{R})$ muni de la norme N_2, en déduire que la suite $\left(\int_0^t (f^n)^2(t)dt\right)_{n\in\mathbb{N}}$ converge vers $\int_0^t \varphi^2(\theta)d\theta$

2) En déduire (pour tout t) la valeur de $\int_0^t \varphi^2(\theta)d\theta$. Et montrer que l'hypothèse que φ^2 est continue en $\frac{2}{3}$ est absurde.

<u>Analysons le vocabulaire</u>: Dans l'exemple précédent, on se rend facilement compte que la seule limite possible serait la fonction φ *définie par* $\varphi(t) = 0$ *si* $t < \frac{2}{3}$ *et* $\varphi(t) = 1$ *si* $t \geq \frac{2}{3}$ *(d'ailleurs* $\lim_n \varphi_n(t) = 1$ *si* $t < \frac{2}{3}$ *et* $\lim_n \varphi_n(t) = 1$ *si* $t \geq \frac{2}{3}$ *). Mais une telle fonction n'est pas continue. Pour que les* N_2*-suites de Cauchy convergent, il faudrait donc ajouter des points à* $\mathscr{C}([0,1],\mathbb{R})$ *; il faudrait le 'compléter'. On pourrait penser à ajouter toutes les fonctions f dont le carré est intégrable au sens de Riemann (ie: celles pour lesquelles on sait étendre la définition de* N_2*). Mais cela ne suffit pas. Il faut bâtir une nouvelle théorie de l'intégrale, ce que nous ne ferons pas dans le cadre de ce cours.*

§ 4 : Séries sommables dans un espace complet

Soit E un espace muni d'une norme N, nous pouvons y définir des séries sommables.

Séries sommables pour N:

Etant donnée une série $((u_n))_{n\in\mathbb{N}}$, *nous dirons qu'elle est sommable pour N, si la suite* $(v_n)_{n\in\mathbb{N}}$ *définie par* $v_n = \sum_{i=0}^{n} u_i$ *(ou éventuellement* $v_n = \sum_{i=1}^{n} u_i$, $\sum_{i=2}^{n} u_i$,...*) est convergente pour N.*

Si la série $((u_n))_{n\in\mathbb{N}}$ est sommable pour N, alors la suite $(u_n)_{n\in\mathbb{N}}$ converge vers 0 pour N (de même que si la série numérique $((u_n))_{n\in\mathbb{N}}$ est sommable, la suite $(u_n)_{n\in\mathbb{N}}$ converge vers 0). Car si la série est sommable les suites $\sum_{i\leq n} u_i$ et $\sum_{i\leq n+1} u_i$ ont même limite, donc leur différence tend vers 0.

Exercice c : En se plaçant dans $\mathscr{C}^0(K,\mathbb{R})$ muni de la norme uniforme, donner un exemple de suite $(f_n)_{n\in\mathbb{N}}$ qui converge vers 0 pour N, telle que la série $((f_n))_{n\in\mathbb{N}}$ ne soit pas sommable.

Si E n'est pas complet, il n'existe aucun procédé raisonnable permettant de montrer la sommabilité d'une série si nous ne pouvons pas en calculer explicitement la somme. Ce qui rend la théorie des séries à peu près sans intérêt. Si E est complet nous pourrons au contraire montrer la sommabilité d'une série en étudiant la série des normes, et utiliser ainsi les séries comme un procédé de définition.

Séries normalement sommables pour N.

Dans ce qui suit E est supposé complet pour la norme N.

Nous dirons que la série $((u_n))_{n\in\mathbb{N}}$ est normalement sommable pour N, si la série des normes $((N(u_n)))_{n\in\mathbb{N}}$ est sommable. Toute série normalement sommable pour N, est sommable pour N (parce que nous avons supposé E complet).

Démonstration: Elle consiste à montrer que les sommes partielles $v_n = \sum_{i=0}^{n} u_i$ forment une suite de Cauchy. Posons $w_n = \sum_{i=0}^{n} N(u_i)$, la suite $(w_n)_{n\in\mathbb{N}}$ est une suite numérique convergente, donc une suite de Cauchy. Nous avons les inégalités:

$$(\forall n)\ (\forall p>n) \qquad N(v_p-v_n) = N\Big(\sum_{i=n+1}^{p} u_i\Big) \leq \sum_{i=n+1}^{p} N(u_i) = w_p - w_n = |w_p - w_n|$$

Donc de: $(\forall\varepsilon)\ (\exists N)\ (\forall n)\ (\forall p)\ \ n \geq N$ et $p \geq N$ implique $|w_p - w_n| \leq \varepsilon$

nous pouvons conclure: $(\forall\varepsilon)\ (\exists N)\ (\forall n)\ (\forall p)\ \ n \geq N$ et $p \geq N$ implique $N(v_p - v_n) \leq \varepsilon$

Ce qui signifie que $(v_n)_{n\in\mathbb{N}}$ est une suite de Cauchy pour N.

Exercice e : Démontrer que la somme de deux séries sommables est sommable. Démontrer que la somme de deux séries normalement sommables est normalement sommable.

Ainsi pour les séries de vecteurs dans un espace complet, nous avons deux formes de sommabilité. La sommabilité simple, et la sommabilité normale, qui est l'analogue de la sommabilité absolue des séries numériques. Pour vérifier qu'une série de vecteurs $((u_n))_{n\in\mathbb{N}}$ est sommable nous aurons donc deux possibilités:

Si nous savons calculer sa somme S, nous calculerons (et, à défaut de pouvoir calculer, nous majorerons) les $N(S - \sum_{i=0}^{n} u_i)$. Mais cette méthode est un exercice d'école, car le cas vraiment intéressant est celui ou l'on ne connait pas S (puisqu'il s'agit précisément de le définir).

Si nous ne savons pas calculer S, nous calculerons (et, à défaut de pouvoir calculer, nous majorerons) les $N(u_i)$. Nous serons alors ramenés à une série numérique (Voir §6). Mais cette seconde méthode n'est valable que dans un espace complet; ce qui exclut certains espaces de dimension infinie.

Règles pour la convergence normale

Toutes les règles de sommabilité absolue que nous avons vues au chapitre 8, peuvent être recopiées. Voici les principales (sans démonstration).

En notant $((u_n))_{n\in\mathbb{N}}$ une série dans E et $((v_n))_{n\in\mathbb{N}}$ une série numérique positive:

Majoration:

Si $(\forall n)$ $N(u_n) \leq v_n$, et si $((v_n))_{n\in\mathbb{N}}$ est sommable, alors $((u_n))_{n\in\mathbb{N}}$ est normalement sommable ; et la norme du reste de $((u_n))_{n\in\mathbb{N}}$ est majorée par le reste du même ordre de $((v_n))_{n\in\mathbb{N}}$.

Utilisation de O:

Si $(N(u_n))_{n\in\mathbb{N}} = O((v_n))_{n\in\mathbb{N}}$, et si $((v_n))_{n\in\mathbb{N}}$ est sommable, alors $((u_n))_{n\in\mathbb{N}}$ est normalement sommable; et N(reste de $((u_n))_{n\in\mathbb{N}}$) = O(reste de $((v_n))_{n\in\mathbb{N}}$).

Utilisation de o:

Si $(N(u_n))_{n\in\mathbb{N}} = o((v_n))_{n\in\mathbb{N}}$, et si $((v_n))_{n\in\mathbb{N}}$ est sommable, alors $((u_n))_{n\in\mathbb{N}}$ est normalement sommable; et N(reste de $((u_n))_{n\in\mathbb{N}}$) = o(reste de $((v_n))_{n\in\mathbb{N}}$).

Utilisation de l'équivalence:

Si $(N(u_n))_{n\in\mathbb{N}}$ est équivalente à $(v_n)_{n\in\mathbb{N}}$, alors $((u_n))_{n\in\mathbb{N}}$ est normalement sommable si et seulement si $((v_n))_{n\in\mathbb{N}}$ (qui est positive!) est sommable.

Et N(reste de $((u_n))_{n\in\mathbb{N}}$) =O(reste de $((v_n))_{n\in\mathbb{N}}$).

Règle de d'Alembert: En comparant, grâce à ces règles, aux séries géométriques, on obtient:

Si $\left(\frac{N(u_{n+1})}{N(u_n)}\right)_{n\in\mathbb{N}}$ converge vers $\lambda < 1$, alors $((u_n))_{n\in\mathbb{N}}$ est normalement sommable, et le reste de $((u_n))_{n\in\mathbb{N}}$ est $O(\lambda^n)$.

Si $\frac{N(u_{n+1})}{N(u_n)}$ converge vers $\lambda > 1$ (ou vers 1 par valeurs supérieures), alors $((u_n))_{n\in\mathbb{N}}$ n'est pas sommable.

§ 5 : Le cas des espaces de dimension finie.

Les espaces de dimension finie sont complets; nous pouvons donc leur appliquer les résultats du paragraphe 4.

Soit E un espace de dimension finie, puisque toutes les normes sur E sont équivalentes, toute série sommable pour une norme N, est sommable pour toutes les autres normes.

Supposons donnée une base de E, nous savons qu'une suite est convergente si et seulement si ses suites coordonnées sont convergentes. *Donc une série est sommable si et seulement si ses séries coordonnées sont sommables*

De même $((u_n))_{n\in\mathbb{N}}$ est normalement sommable si et seulement si ses séries coordonnées $((u_n^i))_{n\in\mathbb{N}}$ sont absolument sommables. Car, si nous notons N_1 la norme définie par $N_1(x) = \sum_i |x^i|$, la série $((N_1(u_n)))_{n\in\mathbb{N}}$ est la somme des séries $((|u_n^i|))_{n\in\mathbb{N}}$. Comme il s'agit de séries positives $((N_1(u_n)))_{n\in\mathbb{N}}$ est sommable, si et seulement si les séries $((|u_n^i|))_{n\in\mathbb{N}}$ le sont.

En particulier pour une série de nombres complexes, la série des modules est sommable si et seulement si la série des parties réelles, et celle des parties imaginaires, sont absolument sommables.

§ 6 : Le cas des séries de fonctions uniformément sommables.

Soit une série de fonctions $((f_n))_{n\in\mathbb{N}}$ dans $\mathscr{C}([a,b],\mathbb{R})$. Comment savoir si elle est uniformément sommable ?

Fixons une valeur t. Pour toute fonction φ, nous avons $|\varphi(t)| \leq N_\infty(\varphi)$, donc si la série donnée est sommable pour N_∞ de somme g, la série numérique $((f_n(t)))_{n\in\mathbb{N}}$ est sommable, de somme g(t). Il en résulte que

** S'il existe des valeurs de t pour les quelles la série numérique $((f_n(t)))_{n\in\mathbb{N}}$ n'est pas sommable, la série donnée n'est pas sommable pour N_∞. Attention une série $((f_n))_{n\in\mathbb{N}}$ peut être sommable pour N_1 ou pour N_2, et que pour certaines valeurs de t, $((f_n(t)))_{n\in\mathbb{N}}$ ne soit pas sommable.*

** Si pour tout t, la série numérique $((f_n(t)))_{n\in\mathbb{N}}$ est sommable, et si on peut calculer sa somme, alors en notant cette somme g(t) on obtient la seule somme possible de la série donnée.*

** Si la fonction g n'est pas continue, elle ne peut être la somme, au sens de la norme uniforme, de la série donnée.*

** Si la fonction g ainsi définie est continue, il reste à calculer (ou, à défaut, à majorer) $N_\infty(g-\sum_{i\le n} f_i)$.*

** Si on n'est pas capable de calculer explicitement g(t), il est impossible de calculer ou de majorer $N_\infty(g-\sum_{i\le n} f_i)$. On peut alors calculer ou majorer les $N_\infty(f_n)$, pour essayer de montrer que la série numérique $((N_\infty(f_n)))_{n\in\mathbb{N}}$ est sommable.*

§ 7 : La méthode des approximations successives

Soit X une partie de l'espace normé E, et soit $f : X \to X$. On suppose qu'il existe $k < 1$ tel que (*) $(\forall x,y)\ N(f(x)-f(y)) \le k\, N(x-y)$. On notera que, si $f: [a,b] \to [a,b]$, à cause du théorème des accroissements finis, cette condition sera vérifiée par toute fonction dérivable dont la dérivée reste en valeur absolue au plus égale à k.

Supposons qu'il existe v dans X tel que $f(v) = v$. Alors pour tout x_0 dans X, la suite définie par la relation $(\forall n>0)\ x_n = f(x_{n-1})$ converge vers v.

En effet: $N(x_n-v) = N(f(x_{n-1}) - f(v)) \le k\, N(x_{n-1}-v)$; donc $(\forall n)\ N(x_n-v) \le k^n N(x_0-v)$. Et puisque $k<1$, ceci implique que la suite $(N(x_n-v))_{n\in\mathbb{N}}$ converge vers 0.

Nous disposons ainsi d'une méthode de calcul permettant de calculer les racines de certaines équations grace à des suites récurrentes.

Exemple: Soit à déterminer la solution v de l'équation $x^3+x-1 = 0$ (il y en a une et une seule, on s'en convaincra graphiquement). Cette équation s'écrit aussi $x = \dfrac{1}{x^2+1} = f(x)$.

Exercice d: Calculer f' et f''. Montrer que $(\forall t)\ |f'(t)| \le |f'(\frac{1}{\sqrt{3}})| = \frac{3\sqrt{3}}{8}$.

Puisque $|f'(t)|$ reste inférieure à $k = \frac{3\sqrt{3}}{8} < 1$, nous pouvons calculer v, en définissant une suite récurrente $(u_n)_{n\in\mathbb{N}}$ par $u_{n+1} = \dfrac{1}{1+(u_n)^2}$ (u_0 étant arbitrairement choisi).

Exercice d': Au moyen d'une calculatrice programmable, calculer cette racine à 10^{-6} près.

Mais nous pouvons modifier cette procédure de calcul pour en faire une démonstration de l'existence de v. Pour cela nous supposerons E complet et X fermé.

En effet: Partant de x_0 arbitraire dans X, nous définissons les x_n par $x_n = f(x_{n-1})$ (ce qui est possible puisque $f: X \to X$). Considérons la série définie par $u_n = x_n-x_{n-1}$

$$N(u_n) = N(x_n-x_{n-1}) = N(f(x_{n-1})-f(x_{n-2})) \le k\, N(x_{n-1}-x_{n-2}) \le k\, N(u_{n-1}).$$

(*) Une telle application est dite contractante.

Donc le rapport $\frac{N(u_n)}{N(u_{n-1})}$ reste au plus égal à $k < 1$; ce qui entraîne que la série numérique $((N(u_n)))_{n\in\mathbb{N}}$ est sommable, donc que la série $((u_n))_{n\in\mathbb{N}}$ est normalement sommable (règle de d'Alembert). Notons qu'on utilise ici le fait que E est complet.

Donc $(x_n)_{n\in\mathbb{N}}$ est une suite convergente, et sa limite v est dans X (car X est fermé). Puisque $N(f(v) - x_n) = N(f(v) - f(x_{n-1})) \leq k\, N(v - x_{n-1})$, la suite $(x_n)_{n\in\mathbb{N}}$ (qui converge vers v) converge aussi vers f(v). Donc f(v) = v.

Exercice e: Expliquer l'affirmation "sa limite est dans X car X est fermé".

Un exemple dans l'espace des matrices

Nous savons qu'il existe sur l'espace $\mathfrak{M}(n,n)$ des matrices carrées $n\times n$, une norme N et un nombre K (>0) tels que ($\forall$A et B) $N(AB) \leq K\, N(A)\, N(B)$ (Voir chap.10 Ex 24). Soit A une matrice, et l'application Φ: $\mathfrak{M}(n,n) \to \mathfrak{M}(n,n)$ définie par $\Phi(X) = Id+AX$. Alors

$$(\forall X)\ (\forall Y) \qquad N(\Phi(X)-\Phi(Y)) = N(A(X-Y)) \leq K\, N(A)N(X-Y).$$

Si $N(A)<1/K$, l'application Φ est contractante. Autrement dit nous pouvons appliquer le processus des approximations successives pour trouver V telle que $Id+AV = V$ (autrement dit pour trouver l'inverse de B=Id-A). Pour cela nous déterminerons la suite:

$U_0 = Id$

$U_1 = Id+AU_0 = Id+A.$

$U_2 = Id+AU_1 = Id+A+A^2.$

...

$U_n = Id+A+A^2+ \dots +A^n.$

Cette méthode de calcul de l'inverse d'une matrice B suppose malheureusement que celle-ci s'écrive Id–A avec A 'petite' (autrement dit que B soit voisine de Id). Elle ne peut donc pas être employée systématiquement.

Mais lorsque l'on inverse (avec une calculatrice) une matrice M (par exemple au moyen de la méthode du pivot) on trouve une matrice P telle que MP soit un peu différent de Id (à cause des erreurs d'arrondi de la machine). On écrit alors MP = 1-A, où A est 'petite'. Et on calcule alors comme ci-dessus $(1-A)^{-1}$, pour avoir avec une meilleure précision $M^{-1} = P(1-A)^{-1}$.

Un exemple dans un espace de fonctions

Cherchons à résoudre l'équation différentielle y' = y sur l'intervalle $[-\alpha,\alpha]$ (où l'on a choisi $0<\alpha<1$). Considérons l'espace $E = \mathscr{C}^0([-\alpha,\alpha],\mathbb{R})$ muni de la norme uniforme. Choisissons un nombre a_0, et à toute fonction f dans E, associons une fonction $\Phi(f)$ définie sur $[-\alpha,\alpha]$, en posant

$$\Phi(f): t \to a_0 + \int_0^t f(u)\, du$$

On a ainsi défini une application de E dans E.

Nous avons $\quad (\forall t) \quad \Phi(f)(t) - \Phi(g)(t) = \int_0^t [f(u) - g(u)]\, du$

$$(\forall t)\ \ |\Phi(f)(t) - \Phi(g)(t)| \leq \left| \int_0^t |f(u) - g(u)|\, du\right| \leq \left| \int_0^t N_\infty(f-g)\, du\right| \leq \alpha\, N_\infty(f-g)$$

Donc: $\qquad N_\infty(\Phi(f) - \Phi(g)) \leq \alpha\, N_\infty(f-g).$

Puisque $\alpha<1$, Φ est contractante. Donc la suite $(f_n)_{n\in\mathbb{N}}$ définie par $f_0 = 0$ et $(\forall n\geq 1)$ $f_n = \Phi(f_{n-1})$ est une suite convergente; et sa limite f est un "point fixe de Φ", c'est à dire une fonction f de E telle que $\Phi(f) = f$. C'est à dire que

$$f(t) = a_0 + \int_0^t f(u)\, du$$

On se convaincra facilement que cette formule entraine que f est dérivable de dérivée f. Autrement dit f est solution de l'équation y' = y. Par ailleurs on a $f(0) = \mathrm{Lim}\,(f_n(0)) = a_0$. En fait nous connaisssions déjà cette[(*)] solution, c'est $t \to a_0 e^t$. On démontre (par récurrence sur n) que $f_n(t) = a_0 \sum_{p=0}^{n} \frac{t^p}{p!}$. Donc

$$a_0 e^t = \lim_{n\to\infty} a_0 \sum_{p=0}^{n} \frac{t^p}{p!}.$$

§ 8 : Convergence des fonctions de classe C^1

Sur l'espace $\mathscr{C}^1([a,b],\mathbb{R})$ des fonctions de classe C^1 sur $[a,b]$ nous pouvons considérer les normes N_1, N_2, N_∞ définies pour l'espace des fonctions continues. Mais ces normes n'en font pas un espace complet. Pour N_1 et N_2 nous l'admettrons. Pour N_∞ voir l'exercice f.

Exercice f : Considérons $\mathscr{C}^1([a,b],\mathbb{R})$ (où $a<0<b$), et la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définie par $f_n(t) = \sqrt{t^2 + \frac{1}{n}}$. Montrer que, considérée dans $\mathscr{C}^0([a,b],\mathbb{R})$, cette suite a une limite pour N_∞. Quelle est cette limite ? En déduire que $\mathscr{C}^1([a,b],\mathbb{R})$ muni de la norme N_∞ n'est pas complet.

Ceci a pour conséquence que la norme uniforme est inadaptée pour construire des fonctions de classe C^1, au moyen des procédés de convergence a priori. Nous allons donc construire une nouvelle norme ν_∞ sur $\mathscr{C}^1([a,b],\mathbb{R})$, qui en fera un espace complet.

Choisissant x_0 quelconque dans $[a,b]$, nous poserons: $\qquad \nu_\infty(f) = |f(x_0)| + N_\infty(f')$.

Exercice g : Démontrer que $\nu_\infty(f) = 0$ implique $f = 0$. En déduire que ν_∞ est une norme.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy pour ν_∞. Puisque $(\forall n)\ (\forall p)\ N_\infty(f_n'-f_p') \le \nu_\infty(f_n - f_p)$, la suite des dérivées $(f_n'))_{n\in\mathbb{N}}$ est une suite de Cauchy pour N_∞; donc elle converge vers une fonction continue φ. De même la suite numérique $(f_n(x_0))_{n\in\mathbb{N}}$ est de Cauchy, donc converge vers un nombre α.

Posons $g(x) = \alpha + \int_{x_0}^{t} \varphi(\tau)\, d\tau$. Nous obtenons une fonction g de classe C^1 sur $[a,b]$. Et la suite $(\nu_\infty(f_n-g))_{n\in\mathbb{N}}$ est la somme de deux suites qui convergent vers 0, puisque

$$\nu_\infty(f_n-g) = |f_n(x_0)-\alpha| + N_\infty(f_n' - g') = |f_n(x_0)-\alpha| + N_\infty(f'_n-\varphi).$$

Donc $(f_n)_{n\in\mathbb{N}}$ converge vers g pour ν_∞.

On remarquera que la dérivée de g est φ, c'est à dire la limite (au sens de la norme uniforme N_∞) de la suite des fonctions dérivées des f_n. On remarquera aussi que $(\forall t)$:

$$\begin{aligned}
|f_n(t) - g(t)| &= |f(x_0) + \int_{x_0}^{t} f_n'(\tau)\, d\tau - \alpha - \int_{x_0}^{t} \varphi(\tau)\, d\tau| \\
&\le |f_n(x_0) - \alpha| + |\int_{x_0}^{t} [f_n'(\tau) - \varphi(\tau)]\, d\tau| \\
&\le |f_n(x_0) - \alpha| + |t - x_0|\, N_\infty(f_n' - \varphi) \\
&\le |f_n(x_0) - \alpha| + |b - a|\, N_\infty(f_n' - \varphi) \\
&\le \nu_\infty(f_n - g) + |b - a|\, \nu_\infty(f_n - g)
\end{aligned}$$

(*) Elle est unique, on le verra au chap.20

Et, puisque ces inagélités sont vraies quel que soit t, c'est que

$$N_\infty(f_n - g) \leq [1 + |b - a|]\, \nu_\infty(f_n - g)$$

Donc la suite $(f_n)_{n\in\mathbb{N}}$ converge vers g non seulement pour la norme ν_∞, mais aussi pour la norme N_∞. En particulier $(\forall t)$ la suite numérique $(f_n(t))_{n\in\mathbb{N}}$ converge vers g(t).

Nous retiendrons: *L'espace $\mathscr{C}^1([a,b],\mathbb{R})$ muni de la norme ν_∞ est complet, autrement dit toute suite $(f_n)_{n\in\mathbb{N}}$ qui est de Cauchy pour ν_∞, a une limite g (pour ν_∞). De plus g est aussi la limite de $(f_n)_{n\in\mathbb{N}}$ au sens de la norme uniforme N_∞.*

Exemple: Quelque soit t, la série $((\frac{\cos nt}{n^3}))_{n\in\mathbb{N}}$ est une série numérique sommable (car elle est majorée par la série $((\frac{1}{n^3}))_{n\in\mathbb{N}}$); notons g(t) sa somme. Nous allons montrer que cette somme est une application dérivable.

Pour démontrer que g est dérivable sur [a,b], plaçons nous dans $\mathscr{C}^1([a,b],\mathbb{R})$, muni de la norme ν_∞. Nous avons (en choisissant x_0=a)

$$\nu_\infty(\frac{\cos nt}{n^3}) = \frac{\cos a}{n^3} + N_\infty(\frac{-\sin nt}{n^2}) \leq \frac{1}{n^3} + \frac{1}{n^2}$$

Donc la série $((\nu_\infty(\frac{\cos nt}{n^3})))_{n\in\mathbb{N}}$ est sommable. Autrement dit la série de fonctions $((\frac{\cos nt}{n^3}))_{n\in\mathbb{N}}$ est normalement sommable pour ν_∞; donc sommable pour ν_∞ (puisque $\mathscr{C}^1([a,b],\mathbb{R})$, muni de la norme ν_∞ est complet). Sa somme au sens de ν_∞ est aussi sa somme au sens de N_∞; donc c'est la fonction g (qui est donc une fonction C^1).

Exercice h: On remplace x_0 par x_1; on obtient ainsi une nouvelle norme ν'_∞ sur $\mathscr{C}^1([a,b],\mathbb{R})$. En utilisant le théorème des accroissements finis, montrer qu'elle est équivalente à ν_∞.

§ 9 : La convergence simple des fonctions.

Sur l'espace $\mathscr{C}^0([a,b],\mathbb{R})$ nous avons trois normes $\mathcal{N}_1, \mathcal{N}_2$ et $\mathcal{N}_\infty$, donc trois façons de définir la convergence d'une suite de fonctions. Sur $\mathscr{C}^1([a,b],\mathbb{R})$ nous avons défini la norme ν_∞, donc une quatrième façon de définir la convergence d'une suite de fonctions. Lorsque l'on est en présence d'une série de fonctions, que l'on voudrait pouvoir considérer comme sommable, on se posera donc la question: pour quelles normes est elle sommable ? C'est la réponse à cette question, qui nous donnera les propriétés de la somme souhaitée; si elle est sommable pour $\mathcal{N}_\infty$, la somme est continue; si elle est sommable pour ν_∞, la somme est C^1; etc.

Ces considérations sur les normes sont une découverte récente (20-ème siècle). Pourtant on a considéré des séries de fonctions dés le début du 18-ème siècle. Comment ?

A l'origine, on considérait que la série $((u_n))_{n\in\mathbb{N}}$ est sommable si, pour toute valeur de la variable t, la série numérique $((u_n(t)))_{n\in\mathbb{N}}$ est sommable. C'est à dire qu'une suite $(\varphi_n)_{n\in\mathbb{N}}$ était dite convergente dés que, pour tout t, la suite numérique $(\varphi_n(t))_{n\in\mathbb{N}}$ était convergente. Lorsqu'il en est ainsi ont dit - à notre époque - que la suite converge simplement. On peut alors définir une fonction limite f, en posant $(\forall t)\, f(t) = \mathrm{Lim}\, \varphi_n(t)$. Malheureusement cette fonction limite peut avoir des propriétés fort désagréables. Par exemple:

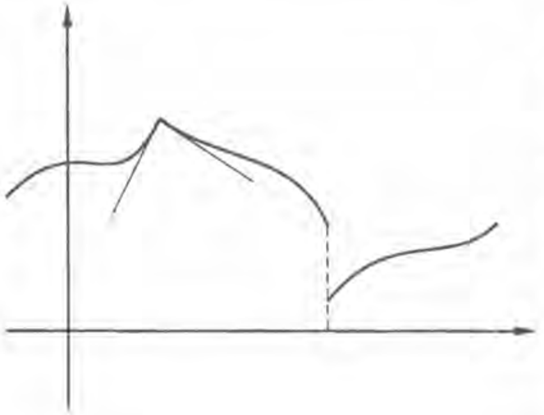
* *Il est possible qu'elle ne soit pas continue (par exemple sur $[0,\pi]$ la suite des fonctions $t \to \sin^n t$ converge simplement vers la fonction qui est nulle partout, sauf en $\pi/2$ où elle vaut 1), et encore moins dérivable.*
* *Il est possible que la suite des intégrales des φ_n ne converge pas vers l'intégrale de f (pour un exemple voir chapitre 14 §1)*
* *Il est possible que les φ_n ne convergent pas uniformément vers f.*

* *Il est possible que les φ_n ne convergent vers f ni pou $\mathcal{N}_1$, ni pour $\mathcal{N}_2$; mais inversement si les φ_n convergent vers f au sens de $\mathcal{N}_1$ ou de $\mathcal{N}_2$, il n'est pas certain qu'elles convergent simplement vers f.*

La convergence simple est donc un outil inadapté. Les outils efficaces sont les normes que l'on peut définir sur les espaces de fonctions. Nous en avons défini quatre, mais il en existe beaucoup d'autres. On se méfiera toutefois qu'en présence d'une suite ou d'une série de fonctions, c'est la convergence simple qu'il est le plus facile de vérifier; c'est pour cette raison que les mathématiciens du 18-ème siècle l'utilisaient; et il faut garder à l'esprit que l'on ne peut à peu près rien en déduire.

Remarque finale.

Jusqu'à maintenant vous n'avez jamais manipulé que des fonctions définies par des formules. Celles ci sont toutes continues, et dérivables sauf en quelques points ou la courbe représentative a une tangente verticale (par exemple: la fonction $\sqrt[3]{\ }$ à l'origine, ou la fonction Arccos en 1). Pour construire des fonctions non continues, ou non dérivables, vous disposiez d'un procédé: considérer une fonction définie par une certaine formule pout $t<t_0$, et par une autre formule pour $t>t_0$. Il y avait aussi la fonction caractéristique de $\mathbb{Q}$ (ou de $\mathbb{D}$). Tous ces exemples étaient bien artificiels.



Désormais nous disposons d'outils permettant de construire des fonctions comme limites de suites (ou sommes de séries) de fonctions. Ces fonctions sont parfois de classe C^∞ (celles que nous construirons au chapitre 12 par exemple). Mais des séries apparemment très simples pourront avoir pour sommes des fonctions fort bizarres. Notons par exemple qu'il est possible de construire des fonctions continues qui ne sont nulle part dérivables. Il est aussi possible de construire des fonctions qui sont dérivables, mais qui ne sont monotones sur aucun intervalle (il n'existe aucun intervalle sur lequel la dérivée garde un signe constant).

Ce que nous rencontrerons le plus souvent ce sont des fonctions dont nous serons incapables de dire si elles sont ou non continues (ou dérivables). Par exemple la somme de la série $((\frac{\sin n^3 t}{n^2}))_{n\geq 1}$ est continue (car elle est sommable pour la norme uniforme), mais les théorèmes dont nous disposons ne permettent pas de dire si elle est dérivable (elle ne l'est très vraisemblablement pas !).

Les graphiques de telles fonctions ne peuvent raisonnablement être tracés; ceci "donne du sens" à certains énoncés comme par exemple le théorème de la valeur intermédiaire. Pour les fonctions classiques dont on peut tracer la courbe représentative, il apparait comme une évidence graphique. C'est dans le cas d'une fonction comme celles que nous venons d'évoquer qu'on conçoit la nécessité de le démontrer. C'est d'ailleurs au début du XIX-ième siècle, lorsque l'on commença sérieusement à construire des limites de fonctions, qu'il perdit le nom de "principe de continuité" (principe que tout le monde considérait comme évident) pour devenir un théorème (première démonstration par Bolzano vers 1820).

Exercices sur le chapitre 11

* **Exercice 1** : Soit F le sous ensemble de $\mathscr{C}^0([0,1],\mathbb{R})$ formé des fonctions nulles en 0 et en 1. Montrer que F est un sous-espace vectoriel de $\mathscr{C}^0([0,1],\mathbb{R})$. On munit F de la norme N_∞, montrer qu'il est complet.

* **Exercice 2**: Soit F le sous-ensemble de $\mathscr{C}^0([0,1],\mathbb{R})$ formé des fonctions f telles que $\int_0^1 f(t)dt = 0$.

Montrer que F est un sous-espace vectoriel. On munit F de la norme uniforme; montrer qu'il est complet.

* **Exercice 3**:

a) Soit $((z_n))_{n\in\mathbb{N}}$ une série de nombres complexes normalement sommable. Soit k un entier au moins égal à 2. Montrer que la série $(((z_n)^k))_{n\in\mathbb{N}}$ est normalement sommable.

b) Soit $z_n = \dfrac{i^n}{\operatorname{Log} n}$. Montrer que $((z_n))_{n\geq 2}$ est sommable, mais pas normalement sommable. Montrer que $(((z_n)^2))_{n\geq 2}$ est sommable, mais pas normalement sommable. Montrer que $(((z_n)^4))_{n\geq 2}$ n'est pas sommable.

* **Exercice 4**: Soit $z\in\mathbb{C}$ (z fixé), la série $((\frac{1}{n^2+z^n}))_{n\geq 1}$ est elle sommable ? Est elle normalement sommable ? On distinguera deux cas suivant que le module de z est au plus égal à 1, ou supérieur à 1.

* **Exercice 5**: Soit $t\in\mathbb{R}$ (t fixé) , la série $((\frac{e^{itn}}{n}))_{n\geq 1}$ est elle sommable ? Est elle normalement sommable ?

* **Exercice 6**: Soit $z\in\mathbb{C}$ (z fixé), la série $((\frac{z(1-z^n)}{n^2}))_{n\geq 1}$ est elle sommable ? Est elle normalement sommable ? On distinguera deux cas suivant que le module de z est au plus égal à 1, ou supérieur à 1.

* **Exercice 7**: Soit f une fonction continue sur [a,b]. On pose $f_0 = f$ et $(\forall n\geq 1)$ $(\forall t)$ $f_n(t) = \frac{1}{2}\sin(f_{n-1}(t))$.

En utilisant l'inégalité $|\sin x| \leq |x|$, montrer par récurrence que $N_\infty(f_n) \leq \frac{1}{2^n} N_\infty(f_0)$. En déduire que la série $((f_n))_{n\in\mathbb{N}}$ est normalement sommable pour la norme uniforme.

* **Exercice 8**: Soit $f_n(x) = \dfrac{1}{1+(n+x)^2}$ des fonctions définies sur [0,1].

a) Calculer $N_\infty(f_n)$. Montrer que la série $((f_n))_{n\in\mathbb{N}}$ est normalement sommable pour la norme uniforme.

b) Calculer $N_\infty(f'_n)$. Montrer que la série $((f'_n))_{n\in\mathbb{N}}$ est normalement sommable pour la norme N_∞.

c) Montrer que la somme f de $((f_n))_{n\in\mathbb{N}}$ est une fonction de classe C^1.

d) Est-ce que f est de classe C^2 ?

* **Exercice 9**: Soit $f_n(t) = \frac{(-1)^n}{1+n+t}$ des fonctions définies sur $[0,1]$.

a) Calculer $N_\infty(f_n)$. Montrer que la série $((f_n))_{n\in\mathbb{N}}$ n'est pas normalement sommable pour N_∞.

b) On pose $g_n = f_{2n+1}+f_{2n}$. Montrer que $((g_n))_{n\in\mathbb{N}}$ est normalement sommable. Soit g sa somme.

c) On pose $F_n = \sum_{i=0}^{n} f_n$. Montrer que la suite $(F_{2n+1})_{n\in\mathbb{N}}$ converge uniformément vers g. Montrer que la suite $(F_{2n})_{n\in\mathbb{N}}$ converge uniformément vers g. En déduire que $((f_n))_{n\in\mathbb{N}}$ est sommable pour la norme uniforme.

** **Exercice 10**: Dans l'espace des matrices 2×2, on considère la norme $N_1\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \sup(|a|,|b|,|c|,|d|)$.

a) Montrer que $(\forall A,B)\ N_1(AB) \le 2\, N_1(A)N_1(B)$. En déduire que si la série (de matrices 2×2) $((M_n))_{n\in\mathbb{N}}$ est normalement sommable pour N_1 alors la série des carrés $(((M_n)^2))_{n\in\mathbb{N}}$ est normalement sommable pour N_1.

b) Montrer que si $N_1(A) < \frac{1}{2}$, alors la série $((A^n))_{n\in\mathbb{N}}$ est normalement sommable pour N_1.

c) Soit $A = \lambda\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ (où $\lambda > 0$). Calculer $A^2, A^3, A^4, \ldots, A^n, \ldots$ Calculer $N_1(A^n)$. Pour quelles valeurs de λ la série $((A^n))_{n\in\mathbb{N}}$ est elle sommable ? Pour quelles valeurs de λ est elle normalement sommable pour N_1 ? Calculer sa somme.

* **Exercice 11**: On pose, pour tout $x \ge 0$, $f_n(x) = x\, e^{-nx^2}$.

a) Montrer que $(\forall x)$ la série $((f_n(x)))_{n\in\mathbb{N}}$ est absolument sommable. Calculer sa somme $f(x)$.

b) Soit $[a,b]$ un intervalle compact ($b>a>0$). Tracer le graphique de f_n. En déduire $N_\infty(f_n|[a,b])$ (où N_∞ est la norme uniforme sur $[a,b]$). Montrer que $((f_n|[a,b]))_{n\in\mathbb{N}}$ est normalement sommable pour N_∞ ?

c) Soit $\varphi_n(t) = f(t) - \sum_{i=0}^{n} f_i(t)$. Calculer $\varphi_n(1/\sqrt{n})$ et $\operatorname{Lim}_n(\varphi_n(1/\sqrt{n}))$. En déduire que la série $((f_n|[0,1]))_{n\in\mathbb{N}}$ est sommable pour la norme uniforme, mais pas normalement sommable pour la norme uniforme.

* **Exercice 12**: Dans l'espace de la géométrie, on considère deux vecteurs $\vec{A}$ et $\vec{B}$.

On pose $\vec{X_0} = \vec{A}$ et $(\forall n\ge 1)\ \vec{X_n} = \vec{B} \wedge \vec{X_{n-1}}$. On se propose d'étudier la série de vecteurs $((\vec{X_n}))_{n\in\mathbb{N}}$.

a) On suppose $\|\vec{B}\| \le 1$; montrer que $((\vec{X_n}))_{n\in\mathbb{N}}$ est normalement sommable.

b) Montrer que, lorsque $\vec{A}$ et $\vec{B}$ sont perpendiculaires, et $\|\vec{B}\| \ge 1$, elle n'est pas sommable.

* **Exercice 13**: Soit f l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base naturelle est

$$M = \frac{1}{4}\begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

On choisit un vecteur $u_0(\alpha,\beta,\gamma)$ dans $\mathbb{R}^3$, et on définit une suite $(u_n)_{n\in\mathbb{N}}$ par la relation de récurrence $u_n = f(u_{n-1})$.

a) Pour quelles valeurs de u_0 la suite $(u_n)_{n\in\mathbb{N}}$ est elle convergente ? (on pourra faire un changement de base)

b) Pour quelles valeurs de u_0 la série $((u_n))_{n\in\mathbb{N}}$ est elle sommable ?

* **Exercice 14**: On considère dans $\mathscr{C}^0([0,1],\mathbb{R})$ la série de fonctions $((f_n))_{n\in\mathbb{N}}$ définie par $f_n(t) = \frac{t^n}{1+n^2t}$.

a) Montrer que, pour tout t la série numérique $((f_n(t)))_{n\in\mathbb{N}}$ est sommable.

b) Montrer que $N_\infty(f_n) = \frac{1}{1+n^2}$; en déduire que la somme est une fonction continue.

* **Exercice 15**: On considère dans $\mathscr{C}^0([0,1],\mathbb{R})$ la série de fonctions $((f_n))_{n\in\mathbb{N}}$ définie par $f_n(x) = x^{2n}(1-x)^n$. Montrer qu'elle est uniformément sommable et calculer sa somme.

** **Exercice 16**: On considère la série de fonctions $((\varphi_n))_{n\in\mathbb{N}}$ définie par $\varphi_n(t) = \dfrac{\sin nt}{1+n^2t^2}$.

a) Montrer que $(\forall t\in\mathbb{R})$ la série numérique $((\varphi_n(t)))_{n\in\mathbb{N}}$ est sommable.

b) Soit des nombres a et b tels que $0<a<b$. Montrer que la série $((\varphi_n|[a,b]))_{n\in\mathbb{N}}$ converge normalement pour la norme uniforme.

c) Calculer $\varphi_n(\frac{1}{n})$. En déduire que la série $((\varphi_n|[0,1]))_{n\in\mathbb{N}}$ n'est pas uniformément sommable.

* **Exercice 17**: On considère dans $\mathscr{C}^0([0,1],\mathbb{R})$ la série de fonctions $((f_n))_{n\in\mathbb{N}}$ définie par $f_n(t) = \dfrac{t}{1+n^3t}$.

a) Montrer que, pour tout t la série numérique $((f_n(t)))_{n\in\mathbb{N}}$ est sommable; on note $\varphi(t)$ sa somme.

b) Montrer que la fonction φ est continue.

* **Exercice 18**: On considère la série de fonctions sur $[0,1]$ définie par $u_n(t) = (\frac{t}{n+t})^n$.

a) Montrer que, pour tout t, la série numérique $((u_n(t)))_{n\geq 1}$ est sommable. On note S(t) sa somme.

b) Montrer que S est une fonction continue.

c) Montrer que S est une fonction de classe C^1.

** **Exercice 19**: On fixe un nombre A, et on considère la série de fonctions sur $[-A,A]$ définie par $\gamma_n(t) = \dfrac{t}{(1+t^2)^n}$.

a) Montrer que, pour tout t, la série numérique $((\gamma_n(t)))_{n\in\mathbb{N}}$ est sommable et calculer sa somme g(t).

b) La série $((\gamma_n))_{n\in\mathbb{N}}$ est elle sommable pour N_∞ ? Est elle normalement sommable pour N_∞ ?

c) Reprendre la question b en remplaçant l'intervalle $[-A,A]$ par un intervalle de la forme $[A,B]$ (où $0<A<B$).

** **Exercice 20**: Dans $\mathscr{C}^0([0,1],\mathbb{R})$ on considère la série $((f_n))_{n\in\mathbb{N}}$ définie par $f_n(t) = \sin \pi t^n$.

a) Montrer que $(\forall t)$ la série numérique $((f_n(t)))_{n\in\mathbb{N}}$ est sommable. On notera $\varphi(t)$ sa somme.

b) Montrer que la série $((f_n))_{n\in\mathbb{N}}$ n'est pas normalement sommable pour la norme uniforme. On pourra commencer par dessiner les graphiques de $f_1, f_2, f_3,\ldots$

c) Montrer que la série $((f_n))_{n\in\mathbb{N}}$ n'est pas sommable pour la norme uniforme.

* **Exercice 21**: Dans $\mathscr{C}^0([0,1],\mathbb{R})$ montrer que les séries suivantes sont normalement sommables pour N_∞.

a) $((f_n))_{n\in\mathbb{N}}$ définie par $f_n(t) = \sin\dfrac{t}{n^2}$.

b) $((g_n))_{n\in\mathbb{N}}$ définie par $g_n(t) = \dfrac{\sin nt}{n^2}$.

c) $((h_n))_{n\in\mathbb{N}}$ définie par $h_n(t) = e^{nt-n^2}$.

d) $((j_n))_{n\in\mathbb{N}}$ définie par $j_n(t) = \sin^n t$.

* **Exercice 22**: Dans $\mathscr{C}^0([0,1],\mathbb{R})$ les séries suivantes sont elles normalement sommables pour N_∞ ?

a) $((f_n))_{n\in\mathbb{N}}$ définie par $f_n(t) = t^n(1-t)$.

b) $((g_n))_{n\in\mathbb{N}}$ définie par $g_n(t) = \sin\pi(t^n - t^{n+2})$

c) $((h_n))_{n\in\mathbb{N}}$ définie par $h_n(t) = \cos(t+\frac{1}{n^2}) - \cos t$.

d) $((j_n))_{n\in\mathbb{N}}$ définie par $j_n(t) = \frac{\pi}{2} - \mathrm{Arctg}(t+n^2)$.

** **Exercice 23**: Dans $\mathscr{C}^0([0,1/2],\mathbb{R})$ on considére la série $((g_n))_{n\in\mathbb{N}}$ définie par $g_n(t) = 1 - \cos t^n$.

a) Montrer qu'elle est normalement sommable pour la norme uniforme.

b) Montrer que la série des dérivées $((g'_n))_{n\in\mathbb{N}}$ est normalement sommable pour la norme uniforme.

c) Montrer que la somme S de la série $((f_n))_{n\in\mathbb{N}}$ est une fonction de classe C^1.

* **Exercice 24**: Dans $\mathscr{C}^0([0,1],\mathbb{R})$ on considére la série $((g_n))_{n\in\mathbb{N}}$ définie par $g_n(t) = \frac{t^n(1-t)^n}{n^2}$.

a) Montrer qu'elle est normalement sommable pour la norme uniforme.

b) Montrer que la série des dérivées $((g'_n))_{n\in\mathbb{N}}$ est normalement sommable pour la norme uniforme.

c) Montrer que la somme S de la série $((f_n))_{n\in\mathbb{N}}$ est une fonction de classe C^1.

d) S est elle de classe C^2 ? C^3 ?...

** **Exercice 25**: On définit une suite de fonctions continues sur $[-a,+a]$ en choisissant f_0 quelconque, et en posant $(\forall n\geq 1)$ $f_n(x) = f_0(x) + \int_0^x f_{n-1}(t)dt$. On pose $g_n = f_{n+1} - f_n$. Etablir une formule donnant g_{n+1} à partir de g_n.

a) Supposons qu'il existe $k(\in\mathbb{N})$ et $A(>0)$ tels que $(\forall x\in[-a,+a])$ $|g_n(x)|\leq A|x|^k$. Démontrer que $(\forall x \in [-a,+a])$ $|g_{n+1}(x)| \leq A\frac{|x|^{k+1}}{k+1}$ (on distinguera 2 cas selon que x est positif ou négatif).

b) Montrer que $N_\infty(g_n) \leq \frac{a^{n-1}}{(n-1)!} N_\infty(g_1)$, puis que la série $((g_n))_{n\in\mathbb{N}}$ est normalement sommable pour N_∞.

c) Soit F la somme de cette série, montrer que $(\forall x)$ $F(x) = f_0(x) + \int_0^x F(t)\, dt$.

d) On suppose que f_0 est la fonction sinus; calculer F (on cherchera d'abord F')

* **Exercice 26**: Montrer que l'équation $x^3+2x^2+2x+2 = 0$ a une seule solution. Montrer qu'elle est équivalente à l'équation $x = -\frac{2x^2+2}{x^2+2}$. Résoudre cette dernière par approximations successives pour donner une valeur de la racine à 10^{-6} près.

** **Exercice 27**: On considére dans $\mathscr{C}^0([-\pi/2,\pi/2],\mathbb{R})$ la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définie par $f_n(t) = \sqrt{\frac{1}{n} + \sin^2 t}$.

1) Calculer pour tout t la limite f(t) de la suite numérique $(f_n(t))_{n\in\mathbb{N}}$.

2) La suite $(f_n)_{n\in\mathbb{N}}$ converge-t-elle uniformément vers f ?

** **Exercice 28**: On fixe un nombre A, et on considère la série de fonctions sur $[0,A]$ définie par $\gamma_n(t) = \frac{|t|^{3/2}}{(1+t)^n}$.

a) Montrer que, pour tout t, la série numérique $((\gamma_n(t)))_{n\in\mathbb{N}}$ est sommable et calculer sa somme g(t).

b) La série $((\gamma_n))_{n\in\mathbb{N}}$ est elle sommable pour N_∞ ? Est elle normalement sommable pour N_∞ ?

c) Est elle sommable pour N_2 ? pour N_1 ? (on calculera d'abord $\int_0^A \frac{dt}{(1+t)^n}$).

Outils fondamentaux du raisonnement

Quatrième partie : $\mathbb{N}$ et la récurrence.

§ 1 : Les axiomes de Peano et la récurrence.

Le nombre entier positif est le plus ancien de tous les objets mathématiques connus. C'est pourtant l'un de ceux dont les origines restèrent le plus longtemps mystérieuses. C'était un outil dont on se servait tout naturellement, mais personne ne s'était demandé quelles étaient les lois précises qui le régissaient. C'est Peano (1858-1932) qui, le premier, donna une description des propriétés fondamentales de $\mathbb{N}$. Les voici:

Tout élément de $\mathbb{N}$ a un suivant (autrement dit il existe une application $\sigma: \mathbb{N} \to \mathbb{N}$ qui à tout nombre n associe son suivant: celui que tout le monde appelle n+1)

Il existe un entier appelé zéro, qui n'est le suivant d'aucun autre.

Tout sous-ensemble de $\mathbb{N}$ qui contient 0, et qui, lorsqu'il contient un entier n, contient aussi son suivant, est $\mathbb{N}$ tout entier (autrement dit si $A \subset \mathbb{N}$, si A contient 0, et si $\sigma(A) \subset A$, alors $A = \mathbb{N}$).

Ce sont les axiomes de Peano; on démontre que toutes les propriétés familières des entiers en résultent. La troisième nous donne le classique axiome de récurrence: Pour toute propriété P

$$\left\{ \begin{array}{l} P(0) \\ (\forall n)\ P(n) \Rightarrow P(n+1) \end{array} \right\} \text{ implique } (\forall n)\ \ P(n)$$

Pour s'en convaincre, appliquer le troisième axiome à l'ensemble $A = \{n \text{ tels que } P(n)\}$.

Réduire ce que nous savons de $\mathbb{N}$ à ces trois propriétés est peut être une fantaisie dont l'intérèt est plus philosophique que mathématique. En revanche il est fondamental de remarquer que presque toutes les assertions $\mathcal{P}(n)$ portant sur un entier n, résultent d'une récurrence (ou de plusieurs récurrences successives). Quand on travaille sur les entiers le raisonnement par récurrence est l'outil essentiel.

On peut se demander comment on a pu faire des mathématiques pendant des siècles sans s'en apercevoir. La locution 'et ainsi de suite' est la forme première du raisonnement par récurrence; tant que les situations restent simples on a coutume de l'employer pour abréger et parceque cela donne des textes plus immédiatement compréhensibles. On prendra garde des limites d'une telle attitude:

Dès que les situations se compliquent, il devient esssentiel de poser correctement les récurrences, si l'on veut éviter les erreurs.

Les ordinateurs ne connaissent pas la commande 'et ainsi de suite'. Très souvent pour programmer un calcul on est obligé de faire apparaitre un argument récurrent, que l'on avait d'abord sous-entendu.

Quelques formes de récurrence:

Soit une propriété P(n); si on a vérifié P(0), et si P(n) implique P(n+1), alors $(\forall n)$ P(n). C'est la forme la plus simple du raisonnement par récurrence. On utilise surtout des variantes dont voici quelques unes.

Les récurrences "qui ne commencent pas à zéro". Soit à démontrer P(n) pour $n \geq 6$. Nous démontrerons par récurrence simple la propriété $P_1(q) = P(q+6)$. Autrement dit nous démontrerons:

D'abord $P_1(0)$ (c'est à dire $P(6)$).

Ensuite que $(\forall q\in \mathbb{N})$ $P_1(q)$ implique $P_1(q+1)$ (c'est à dire que $(\forall n\geq 6)$ $P(n)$ implique $P(n+1)$)

Les récurrences dont le pas utilise $(\forall i\leq n)$ $P(i)$. Supposons que l'on puisse vérifier $P(0)$, et que

$$(\alpha) \qquad ((\forall i\leq n)\ P(i)) \text{ implique } P(n+1).$$

Alors $(\forall n)$ $P(n)$; en effet: Considérons la propriété $Q(n) = \{(\forall i\leq n)\ P(i)\}$. On va démontrer $Q(n)$ par récurrence simple.

D'abord on vérifie $Q(0)$ (c'est à dire $P(0)$).

Ensuite on montre que $Q(n-1)$ implique $Q(n)$ (c'est (α)).

Exercice a: Un nombre premier est un nombre (au moins égal à 2) qui ne peut pas être écrit comme le produit de deux nombres au moins égaux à 2. Démontrer - par récurrence sur n - que tout nombre (au moins égal à 2) a un diviseur premier.

Les récurrences dont le pas utilise $P(n-1)$ et $P(n-2)$: Supposons qu'on sache démontrer que

$$(P(n-1) \text{ et } P(n-2)) \text{ implique } P(n).$$

Alors si l'on a vérifié $P(0)$ et $P(1)$, on peut affirmer que $(\forall n)$ $P(n)$.

C'est un cas particulier du précédent, puisque $((\forall i<n)\ P(i))$ implique $(P(n-1)$ et $P(n-2))$. Pour se ramener à une récurrence simple, il suffit de considérer la propriété (plus simple que celle que nous avions envisagée dans le cas précédent)

$$R(n) = (P(n) \text{ et } P(n-1))$$

Exercice b: Considérons une suite récurrente définie par $u_0 = 0$, $u_1 = 2$ et $(\forall n\geq 2)$ $u_n = u_{n-1} + 3u_{n-2}$. Démontrer que tous les termes de la suite sont pairs.

§ 2 : Les ensembles finis.

Pour démontrer qu'un ensemble E est fini on peut compter ses éléments, c'est à dire définir une bijection φ de E sur $\{1,\ldots,n\}$. Notons qu'une telle bijection définit une relation d'ordre total sur E (définie par: $e\leq e'$ si et seulement si $\varphi(e)\leq\varphi(e')$). Inversement si un ensemble fini est totalement ordonné, pour numéroter ses éléments, il suffit:

de donner le numéro 1 au plus petit de tous.

de donner le numéro 2 au plus petit de ceux qui restent.

etc.

En particulier en informatique, pour faire la liste de tous les éléments d'un ensemble fini, on doit d'abord décider dans quel ordre on les écrit.

Exercice c: Soit $\{a,b,c,d\}$ un ensemble à quatre éléments. Faire la liste de toutes les parties de E, en les écrivant dans l'ordre lexicographique (celui du dictionnaire).

Faire la liste des parties de E, en les rangeant par nombre d'éléments croissant, et en rangeant celles qui ont même nombre d'éléments dans l'ordre lexicographique.

Exercice d: Faire la liste des permutations de l'ensemble $\{1,2,3,4\}$ en les rangeant dans l'ordre lexicographique.

Mais il est souvent utile de pouvoir montrer qu'un ensemble est fini sans compter ses éléments. C'est Cantor qui, le premier, donna une caractérisation des ensembles finis, sans aucun recours à la notion d'entier. En voici trois variantes.

Un ensemble E est fini s'il est impossible de trouver une bijection de E sur une partie stricte (c'est à dire différente de E tout entier) de E.

Un ensemble E est fini si toute application injective de E dans E, est surjective.

Un ensemble E est fini si toute application surjective de E dans E, est injective.

Ajoutons que si E et F sont finis:

Card E < Card F $\Longleftrightarrow$ Il existe des injections de E dans F, et elles ne sont pas surjectives.

Card E > Card F $\Longleftrightarrow$ Il existe des surjections de E dans F, et elles ne sont pas injectives.*

Exercice e: Construire une application injective et non surjective de $\mathbb{Z}$ dans $\mathbb{Z}$. En construire une de $\mathbb{R}$ dans $\mathbb{R}$. Construire une application surjective et non injective de $\mathbb{Z}$ dans $\mathbb{Z}$. En construire une de $\mathbb{R}$ dans $\mathbb{R}$.

§ 3 : Calculs itératifs et calculs récursifs.

Dans ce qui précède les processus récurrents ont été employés pour faire des démonstrations. L'informatique les fait apparaitre sous un tout autre aspect. Pour un informaticien le problème fondamental est le suivant: Etant donnée une famille d'objets (de nombres, d'ensembles,...) définie par un processus récurrent, calculer un terme donné de cette suite. Il y a pour cela deux stratégies de calcul, que l'on va décrire sur un exemple.

Exemple 1: Calculer 50!, sachant que la suite n! est définie par $0! = 1$ et $(\forall n \geq 1)\ n! = n \times (n-1)!$.

Premier procédé. On se rend facilement compte qu'il va falloir calculer successivement 1!, 2!, 3!, Etc. On écrit donc l'algorithme:

$$\begin{cases} \text{Fact}(0) = 1 \\ \text{Pour n allant de 1 à 50 faire Fact(n)} = \text{n} \times \text{Fact(n-1)}. \end{cases}$$

C'est un type de calcul que vous pouvez programmer en Basic sur votre calculateur de poche (attention 50! est peut être déjà trop grand pour votre machine)

Second procédé: Les langages informatiques évolués (le Pascal par exemple) permettent de laisser la machine interpréter elle même la relation de récurrence. On écrit alors l'algorithme

$$\begin{cases} \begin{cases} \text{Si n = 0 alors Fact(n) = 1} \\ \text{Si n>0 alors Fact(n)} = \text{n} \times \text{Fact(n}-1) \end{cases} \\ \text{Calculer Fact(50).} \end{cases}$$

Et c'est la machine qui découvrira elle même que pour calculer 50!, il faut connaitre 49! ; elle calculera donc 49! ; elle découvrira elle même que, pour cela, il lui faut 48! ; etc.

Le premier calcul est dit "itératif"; le second est dit "récursif". Notons que dans le calcul itératif, c'est le programmeur qui a rangé les calculs dans l'ordre où on doit les effectuer. Dans le calcul récursif, cette partie du travail est faite par la machine. Ici il était tellement facile de mettre les calculs dans le bon ordre, qu'on voit mal l'intérêt qu'il peut y avoir à demander à la machine de le faire elle même. Ce n'est pas toujours aussi simple.

Exemple 2: Soit à écrire un algorithme qui permet de décomposer un nombre en facteurs premiers. Le raisonnement classique (il date de l'antiquité Grecque) qui permet de démontrer que tout nombre (≥ 2) a une décomposition en facteurs premiers, est une récurrence. C'est vrai pour 2. Si c'est vrai pour tous les nombres inférieurs à n, c'est vrai pour n; en effet, ou bien n est premier (et la décomposition est n), ou bien n n'est pas premier et il est le produit de deux nombres p et q (au moins égaux à 2, donc inférieurs à n), on obtient alors la décomposition de n en mettant cote à cote celles de p et de q. Pour transformer ceci en un algorithme, nous supposerons que nous disposons d'un sous-programme qui nous donne le

---

* Ces énoncés sont connus sous le nom de principe des tiroirs, parcequ'ils signifient que si l'on range n chaussettes dans p tiroirs, alors: si n<p il y a forcément au moins un tiroir vide; tandis que si n>p, il y a forcément deux chaussettes dans le même tiroir.

plus petit diviseur d'un nombre n (ce plus petit diviseur est un nombre premier). Nous écrirons alors l'algorithme récursif

$$\begin{cases} \text{Décomp}(1) = \emptyset \\ \text{Décomp}(n) = \text{Mettre bout à bout Plus Petit Div}(n) \text{ et Décomp}(n/\text{Plus Petit Div}(n)) \\ \text{Calculer Décomp}(444) \end{cases}$$

Pour calculer la décomposition de 444, la machine écrira $444 = 2\times222$, puis $222 = 2\times111$, puis $111 = 3\times37$, et 37 est premier. Puis, reprenant le calcul en sens contraire, $37 = 37$, $111 = 3\times37$, $222 = 2\times3\times37$, enfin $444 = 2\times2\times3\times37$. Pour faire un calcul itératif, il nous fallait décomposer d'avance le nombre n/Plus Petit Div(n). Mais nous ne le connaissions pas, nous savions seulement qu'il est au plus égal à $n-1$ (et même à $n/2$ si l'on se donne la peine de réfléchir un peu). Pour $n=444$, il fallait donc d'abord stocker les décompositions des 221 nombres compris entre 2 et 222. Dans ce cas le calcul récursif est de beaucoup le plus simple (c'est à dire celui qui demandera le moins de calculs).

Exemple 3: Définissons une suite récurrente par $u_0 = 2$, $u_1 = 1$ et $(\forall n\geq2)\ u_n = u_{n-1} + 2u_{n-2}$. Le calcul récursif de u_{50} mène à l'algorithme

$$\begin{cases} \begin{cases} \text{si } n = 0 \text{ alors } u(n) = 2 \\ \text{si } n = 1 \text{ alors } u(n) = 1 \\ \text{si } n\geq2 \text{ alors } u(n) = 2\,u(n-2) + u(n-1) \end{cases} \\ \text{Calculer } u(50) \end{cases}$$

Pour calculer u(50) la machine calculera d'abord u(48) (qu'elle mettra en mémoire), puis u(49). Mais lorsqu'elle effectuera ce second calcul, elle aura oublié tous les calculs intermédiaires qui lui ont permis de calculer u(48); elle ne sera même pas capable d'utiliser la valeur de u(48) qui se trouve quelque part dans la mémoire. Donc elle calculera 2 fois u(48).

Pour calculer u(48), elle calculera u(47) une première fois. Puis dans le calcul de u(49) elle calculera deux fois u(47), qui sera donc calculé trois fois en tout. On se persuadera facilement que u(46) sera calculé cinq fois, u(45) huit fois; etc. Il est clair que la méthode n'est pas très astucieuse.

Exercice f: Ecrire un programme calculant u(50) de façon itérative.

En fait la longueur du calcul est due au fait que la machine oublie de stocker les résultats intermédiaires dont elle aura besoin. Pour l'obliger à le faire il suffit de remplacer la suite scalaire u(n) par la suite vectorielle (v(n),u(n)) où $v(n) = u(n-1)$. On écrira alors l'algorithme récursif

$$\begin{cases} \begin{cases} \text{si } n = 1 \text{ alors } \begin{pmatrix} v(n) \\ u(n) \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ \text{si } n\geq1 \text{ alors } \begin{pmatrix} v(n) \\ u(n) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} v(n-1) \\ u(n-1) \end{pmatrix} \end{cases} \\ \text{Calculer } \begin{pmatrix} v(50) \\ u(50) \end{pmatrix} \end{cases}$$

Exercices sur le chapitre 11b

* **Exercice 1:** Montrer (par l'absurde), en utilisant l'une des trois formulations ci-dessus, que tout sous-ensemble d'un ensemble fini est fini.

Soit E un ensemble, et $\mathcal{P}(E)$ l'ensemble des parties de E; montrer que si $\mathcal{P}(E)$ est fini, alors E est fini; et réciproquement. On suppose que Card(E) = n, savez vous ce que vaut Card ($\mathcal{P}(E)$) ?

* **Exercice 2:** Démontrer par récurrence sur n que $\sum_{p=1}^{n} p^2 = \frac{n(n+1)(2n+1)}{6}$.

* **Exercice 3:** On se propose de calculer $B_n = \sum_{p=0}^{n} p^3$.

a) Montrer qu'il existe un polynôme P de degré 4 tel que $P(0) = 0$, $P(1) = B_1$, $P(2) = B_2$, $P(3) = B_3$ et $P(4) = B_4$.

b) Montrer par récurrence que $(\forall n)$ $B_n = P(n)$.

* **Exercice 4:** On note φ la fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui est nulle en 0, et qui en tout point $t \neq 0$ prend la valeur e^{-1/t^2}.

a) Soit P un polynôme, posons $f_P(t) = \varphi(t)\,P(1/t)$ (pour $t \neq 0$, et $f_P(0) = 0$). Montrer que f_P est dérivable en 0, de dérivée nulle. Montrer que f_P est dérivable en tout $t \neq 0$, et calculer $f'_P(t)$. Montrer que la fonction f'_P est de la forme f_Q, pour un polynôme Q que l'on précisera.

b) Démontrer par récurrence sur k que la fonction φ est de classe C^k.

c) Quelle est la série de Taylor de φ à l'origine ?

d) Calculer explicitement les dérivées de φ jusqu'à l'ordre 5.

* **Exercice 5:** On se propose de calculer les nombres $C_n^p = \frac{n!}{p!\,(n-p)!}$ au moyen d'une machine.

a) Pourquoi est il maladroit de calculer n!, p! et (n-p)! et de faire le quotient ?

b) On va utiliser le fait que $(\forall k)$ $C_k^0 = 1$, et $(\forall i<j)$ $C_j^i = C_{j-1}^{i-1} \times \frac{j}{i}$. Ecrire un algorithme itératif donnant C_{50}^{17}.

c) On écrit l'algorithme récursif suivant
$$\begin{cases} \text{si } p=0 \text{ alors } C_n^p = 1 \\ \text{si } p=n \text{ alors } C_n^p = 1 \\ \text{si non } C_n^p = C_{n-1}^{p-1} \times \frac{n}{p} \end{cases}$$

Démontrer qu'il permet de calculer tous les C_n^1. Démonter qu'il permet de calculer tous les C_n^2. Etc. Montrer par récurrence sur p qu'il permet de calculer tous les C_n^p.

** **Exercice 6:** Soit (pour $p \geq n$) $\mathcal{X}(p,n)$ l'ensemble des applications croissantes (au sens large) et surjectives de $\{1,\ldots,p\}$ dans $\{1,\ldots,n\}$. Soit $X(p,n) = \mathrm{Card}(\mathcal{X}(p,n))$. Notons que $(\forall n)$ $X(n,n) = X(n,1) = 1$.

a) Soit $\mathcal{Y}(p,n)$ le sous-ensemble de $\mathcal{X}(p,n)$ formé des applications qui envoient $p-1$ en n. Soit $\mathcal{Z}(p,n)$ le sous-ensemble formé de celles qui n'envoient pas $p-1$ en n. Montrer que $\mathrm{Card}\,\mathcal{Y}(p,n) = X(p-1,n)$ et que $\mathrm{Card}\,\mathcal{Z}(p,n) = X(p-1,n-1)$. Ecrire une relation récurrente permettant de calculer les $X(p,n)$.

b) Calculer à la main et récursivement $X(7,4)$. On notera tous les $X(p,n)$ que l'on doit déterminer en cours de calcul.

c) Ecrire un algorithme récursif permettant de calculer $X(7,4)$ au moyen d'une machine. Les $X(p,n)$ que la machine déterminera en cours de calcul sont les mêmes que ci-dessus. Mais certains seront calculés plusieurs fois. Les quels ? Combien de fois seront-ils calculés ? (cf: §3 Ex.3)

d) Ecrire un algorithme itératif permettant de calculer successivement tous les $X(2,n)$, tous les $X(3,n)$, tous les $X(4,n)$,...

Exercice 7: On note $\mathcal{R}(p,n)$ l'ensemble des applications de $\{1,\ldots,p\}$ dans $\{1,\ldots,n\}$. On note $\mathcal{S}(p,n)$ le sous-ensemble formé de celles qui sont surjectives.

a) Montrer que Card $\mathcal{S}(p,n) = n^p$. On fera une récurrence sur p.

b) Montrer que Card $\mathcal{S}(p,2) = 2^p - 2$. Montrer que Card $\mathcal{S}(p,3) = 3^p - 3\times 2^p + 3$. Que valent Card $\mathcal{S}(p,4)$, Card $\mathcal{S}(p,5)$,...

** **Exercice 8:** On définit $\varphi\colon \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ par $\varphi(p,q) = \dfrac{(p+q+1)^2+p-q-1}{2}$. On remarquera que

$$(p+q)^2 + (p+q) + 1 \le 2\varphi(p,q) < (p+q+1)^2 + (p+q+1).$$

a) Montrer que φ est une injection de $\mathbb{N}$ sur $\mathbb{N}$ (remarquer d'abord que $\varphi(p,q)$ permet de retrouver p+q)

b) Montrer que φ est surjectice (pour résoudre $\varphi(p,q) = n$, déterminer d'abord p+q).

c) Construire une bijection de $\mathbb{N}^3$ sur $\mathbb{N}^2$. Construire une bijection de $\mathbb{N}^3$ sur $\mathbb{N}$.

* **Exercice 9:** On dispose d'un sac de pommes de terre de 5kg, chaque tubercule pèse entre 1g et 500g, et il y en a 50. On dispose d'une balance précise au milligramme près. On se propose de montrer qu'en choisissant judicieusement des pommes de terre, et en en posant au moins une sur chaque plateau (mais toutes les pommes de terre du sac n'étant pas nécessairement utilisées), on peut équilibrer la balance.

a) Combien de tas peut on faire avec ces 50 pommes de terre. Montrer que deux de ces tas ont des poids qui diffèrent de moins de un milligramme.

b) Montrer qu'il existe deux tas sans élément commun, et dont les poids diffèrent de moins de un milligramme.

* **Exercice 10:** On pose $A_n = \displaystyle\sum_{k=1}^{n} \frac{1}{k}$. On veut montrer que $(\forall n\geq 2)$ A_n est le quotient d'un nombre impair par un nombre pair.

a) Montrer que si la propriété est vraie pour $n = 2p$, elle est vraie pour $n = 2p+1$.

b) Montrer que si la propriété est vraie pour p, elle est vraie pour $n = 2p$ (on écrira $A_n = \displaystyle\sum_{k=1}^{p} \frac{1}{2k} + \sum_{k=1}^{p} \frac{1}{2k-1}$

c) Conclure.

** **Exercice 11:** Soit $f\colon \mathbb{N} \to \mathbb{N}$. Soit $a \in \mathbb{N}$. On définit une suite $(u_n)_{n\in\mathbb{N}}$ par $u_0 = a$ et $(\forall n\geq 1)$ $u_n = f(u_{n-1})$.

a) Montrer que si $u_{n_0} = u_{n_0+1}$, alors $(\forall n>n_0)$ $u_n = u_{n+1}$.

b) Montrer que s'il existe n_0 et p (>0) tels que $u_{n_0} = u_{n_0+p}$, alors $(\forall k>0)$ $u_{n_0+k} = u_{n_0+p+k}$.

c) On pose $v_n = u_{2n} - u_n$. Montrer que, pour qu'un tel p existe, il est nécessaire et suffisant que l'un des v_n soit nul. Quelle relation existe-t-il alors entre n et p ?

d) Décrire un algorithme itératif permettant de calculer successivement les v_n. Décrire un algorithme itératif permettant de trouver le plus petit nombre p tel que, à partir d'un certain rang, $u_{n+p} = u_n$.

** **Exercice 12:** On sait que $C_n^2 = \dfrac{n(n-1)}{2}$; donc $C_2^2 = 1, C_3^2 = 3, \ldots$ On pose $C_1^2 = 0$. Pour tout $k\geq 2$, on définit un sous-ensemble T_k de $\mathbb{N}$, en posant $T_k = \{C_{k+1}^2 - C_i^2 \text{ où } 1\leq i<k\}$.

Pour tout $k\geq 1$, on définit un sous-ensemble S_k de $\mathbb{N}$ par $\begin{cases} S_1 = \{1\} \\ (\forall k\geq 2)\ S_k = \{2k-1\} \cup \{p+k \text{ où } p\in S_{k-1}\} \end{cases}$

En remarquant que $C_{k+1}^2 = C_k^2 + k$, montrer par récurrence que pour tout $k\geq 2$, $S_k = T_k$.

*** **Exercice 13:** Soit $x \notin \mathbb{Q}$, montrer que

$$(\forall y\in[-1,1])\ (\forall \varepsilon>0)\ (\exists n\in\mathbb{N}) \text{ tel que } |\sin(n\pi x) - y| \le \varepsilon.$$

Séries entières

Considérons une suite $(a_n)_{n\in\mathbb{N}}$ de nombres complexes; à cette suite et à tout nombre complexe z, nous pouvons associer la série de nombres complexes $((a_nz^n))_{n\in\mathbb{N}}$. Cette série est appelée la série entière de coefficients $(a_n)_{n\in\mathbb{N}}$.

Pour la commodité des démonstrations nous étudions seulement maintenant les séries entières. Pourtant historiquement la théorie des séries fut, au 18-ème siècle, essentiellement celle des séries entières. La raison en est la suivante. Considérons une fonction f qui est C^∞ au voisinage de 0, notons $f^{n)}$ ses dérivées successives en 0; nous pouvons lui associer la série entière $((\frac{f^{(n)}}{n!}z^n))_{n\in\mathbb{N}}$. C'est sa série de Taylor, dont les sommes partielles sont les polynômes de Taylor de f en 0. Nous démontrerons au §5 que - dans les bons cas, et en particulier pour toutes les fonctions familières aux mathématiciens de cette époque - cette série est sommable, et a pour somme la fonction f. La découverte de ces phénomènes date du début du 18-ème siècle (Taylor 1685-1727). On confondait alors le polynôme et la série (qui n'était qu'un polynôme à un nombre infini de termes), en négligeant complètement les problèmes de convergence. La mise au point de la théorie date du 19-ème siècle.

§ 1 : Convergence des séries entières

Nous allons d'abord étudier l'ensemble des z tels que $((a_nz^n))_{n\in\mathbb{N}}$ soit une série sommable. Cet ensemble est très variable: il peut être réduit à 0; il peut être $\mathbb{C}$ tout entier.

Exercice a: En utilisant la règle de d'Alembert, montrer que la série $((n!z^n))_{n\in\mathbb{N}}$ ne converge pour aucune valeur de z ($\neq 0$). Montrer que la série $((\frac{z^n}{n!}))_{n\in\mathbb{N}}$ converge pour toute valeur de z.

Le résultat fondamental est le suivant: *Il existe $\rho\in[0,\infty]$, tel que :*

α) pour $|z|<\rho$, la série $((a_nz^n))_{n\in\mathbb{N}}$ est absolument sommable.

β) pour $|z|>\rho$, la série $((a_nz^n))_{n\in\mathbb{N}}$ n'est pas sommable.

Si ρ est fini, pour $|z|=\rho$, il peut arriver que $((a_nz^n))_{n\in\mathbb{N}}$ soit sommable, il peut arriver que $((a_nz^n))_{n\in\mathbb{N}}$ ne soit pas sommable (cf. Ex c).

On dit que ρ est le rayon de convergence de la série entière $((a_nz^n))_{n\in\mathbb{N}}$.

Démonstration: Soit $X=\{x\in[0,\infty[$ tels que la suite $(|a_n|x^n)_{n\in\mathbb{N}}$ soit bornée$\}$, et soit ρ la borne supérieure de X (elle existe toujours, à condition de convenir que, lorsque X n'est pas majoré, elle est égale à $+\infty$).

* Si $|z|<\rho$, il existe x dans X tel que $|z|<x<\rho$. Puisque x est dans X, il existe M tel que $(\forall n)$ $|a_n|x^n<M$. Alors $(\forall n)$:

$$|a_nz^n|=|a_n|\,|z|^n=(|a_n|x^n)\left(\frac{|z|}{x}\right)^n\le M\left(\frac{|z|}{x}\right)^n .$$

Et puisque $\frac{|z|}{x}<1$, la série majorante $\left(\left(M\left(\frac{|z|}{x}\right)^n\right)\right)_{n\in\mathbb{N}}$ est sommable, donc la série $((a_nz^n))_{n\in\mathbb{N}}$ est normalement sommable.

* Si $|z|>\rho$, d'après la définition de ρ, la suite $(a_nz^n)_{n\in\mathbb{N}}$ n'est pas bornée, donc elle ne converge pas vers 0, donc $((a_nz^n))_{n\in\mathbb{N}}$ n'est pas sommable.

Exercice b: Quel est le rayon de convergence de la série $((z^n))_{n\in\mathbb{N}}$ (i.e.: $a_n = 1$ quel que soit n).

On sait que $((\frac{1}{n}))_{n\in\mathbb{N}}$ n'est pas sommable, et que $((\frac{(-1)^n}{n}))_{n\in\mathbb{N}}$ est sommable. En déduire le rayon de convergence de $((\frac{z^n}{n}))_{n\in\mathbb{N}}$.

Exercice c : On considère les séries $A = ((n\,z^n))_{n\in\mathbb{N}}$, $B = ((\frac{1}{n}z^n))_{n\in\mathbb{N}}$ et $C = ((\frac{1}{n^2}z^n))$.

Démontrer qu'elles ont toutes trois le même rayon de convergence $\rho = 1$. Démontrer que A n'est sommable pour aucune valeur de z telle que $|z| = 1$. Démontrer que B est sommable pour certains z tels que $|z| = 1$ ($z = -1,-i,+i$ par exemple) mais pas pour tous (utiliser le critère d'Abel). Démontrer que C est sommable pour tous les z tels que $|z| = 1$.

§ 2 : Les fonctions définies comme somme d'une série entière

Soit D le domaine de convergence de la série entière $((a_nz^n))_{n\in\mathbb{N}}$; on sait que c'est un disque ouvert augmenté éventuellement d'une partie de son bord. A tout Z dans D nous pouvons associer la somme de la série $((a_nZ^n))_{n\in\mathbb{N}}$. Nous obtenons ainsi une application f de D dans $\mathbb{C}$. Nous allons étudier les propriétés de cette fonction f.

Continuité de la somme: Nous allons montrer que f est continue sur le disque ouvert D' de rayon ρ. Pour cela nous utiliserons les méthodes du chapitre 11 §6. Il nous faut démontrer la convergence uniforme de la série de fonctions $((a_nz^n))_{n\in\mathbb{N}}$. Malheureusement elle ne converge pas uniformément sur D' (avec les définitions que nous avons données, cela n'aurait aucun sens puisque la norme N_∞ n'a été définie que pour des fonctions sur un compact K; nous pourrions prendre des définitions plus générales, mais alors ce serait faux).

Démonstration: Soit $0 < x < \rho$, notons D_x le disque fermé de centre l'origine et de rayon x, et considérons la série entière $((a_nz^n))_{n\in\mathbb{N}}$ comme une série de fonctions dans $\mathscr{C}(D_x,\mathbb{C})$. Choisissons r tel que $x < r < \rho$. Nous avons démontré au §1 qu'il existe M tel que $(\forall z)$ et $(\forall n)$ $|a_nz^n| \leq M(\frac{|z|}{r})^n$. Donc $(\forall n)$ $(\forall z\in D_x)$ (i.e. $|z|\leq x$) nous aurons $|a_nz^n| \leq M(\frac{x}{r})^n$. C'est à dire $(\forall n)$ $N_\infty(a_nz^n) \leq M(\frac{x}{r})^n$; et puisque $((M(\frac{x}{r})^n))_{n\in\mathbb{N}}$ est sommable, nous obtenons une série normalement sommable pour la norme uniforme.

La convergence uniforme sur D_x implique la continuité de $f|D_x$. Et ceci quel que soit $x < \rho$, donc f est continue sur la réunion des D_x, c'est à dire sur D'.

Exercice d : Soit f une fonction définie sur le disque ouvert D' de centre 0 et de rayon ρ. Montrer que si $f|D_x$ est continue quel que soit $x < \rho$, alors f est continue.

Par conséquent la somme d'une série entière $((a_nz^n))$ est une application continue sur le disque ouvert de centre 0 et de rayon ρ (c'est-à-dire sur l'ensemble D' des z tels que $|z|$ soit strictement inférieur au rayon de convergence)

Sommes de séries entiéres: La somme des séries entières $A = ((a_nz^n))_{n\in\mathbb{N}}$ et $B = ((b_nz^n))_{n\in\mathbb{N}}$ est la série $(A+B) = (((a_n+b_n)z^n))_{n\in\mathbb{N}}$. Puisque la somme de deux séries sommables est sommable, son domaine de convergence contient l'intersection des domaines de convergence de A et B; *autrement dit $\rho_{A+B} \geq inf(\rho_A,\rho_B)$.*

Si $\rho_A \neq \rho_B$, alors $\rho_{A+B} = inf(\rho_A,\rho_B)$. Supposons $\rho_A<\rho_B$, et considérons r tels que $\rho_A<r<\rho_B$. Puisque la somme d'une série sommable et d'une série non sommable est non sommable, la série $(((a_n+b_n)r^n))_{n\in\mathbb{N}}$ n'est pas sommable; donc $r > \rho_{A+B}$. Ce qui prouve que tout r supérieur à ρ_A, est aussi supérieur à ρ_{A+B}. Et ainsi $\rho_A \geq \rho_{A+B}$.

En prenant $A = ((z^n))_{n\in\mathbb{N}}$ et $B = (((\frac{1}{n!}-1)z^n))_{n\in\mathbb{N}}$, on constatera que, *lorsque* $\rho_A = \rho_B$, ρ_{A+B} *peut être beaucoup plus grand que* $\inf(\rho_A,\rho_B)$.

Bien entendu on a (pour tout z tel que les trois sommes existent)

$$\text{somme}(A+B)(z) = \text{somme}(A)(z) + \text{somme}(B)(z).$$

Exercice e: Soit f(z) la somme de $((nz^n))_{n\in\mathbb{N}}$. Calculer en fonction de f la somme de la série $((nz^{n+1}))_{n\in\mathbb{N}} = (((n-1)z^n))_{n\geq 1}$

En déduire que f(z) - zf(z) est la somme de la série géométrique $((z^n))_{n\in\mathbb{N}}$ et calculer f(z).

§ 3 : Série dérivée d'une série entière

On appelle série dérivée de $((a_nz^n))_{n\in\mathbb{N}}$, *la série* $((na_nz^{n-1}))_{n\in\mathbb{N}}$ (on a dérivé terme à terme). *Elle a même rayon de convergence que la série donnée (mais son comportement sur le bord du disque de convergence peut être différent* (voir Ex. c).

Démonstration: Notons ρ le rayon de convergence de $((a_nz^n))_{n\in\mathbb{N}}$, et ρ_1 celui de la série dérivée.

Montrons que $\rho \leq \rho_1$: Soit $r < \rho$, il existe M tel que $(\forall n)\ |a_n|\, r^n \leq M$. On en déduit que $(\forall z)$ $|n\, a_nz^{n-1}| \leq \frac{nM}{|z|}(\frac{|z|}{r})^n$. Pour $|z| < r$, la série positive $((\frac{nM}{|z|}(\frac{|z|}{r})^n))_{n\in\mathbb{N}}$ est sommable (règle de d'Alembert); donc la série dérivée $((na_nz^{n-1}))_{n\in\mathbb{N}}$ est aussi sommable. Ceci prouve que $\rho_1 \geq r$; et puisque ceci est vrai pour tout $r < \rho$, c'est que $\rho_1 \geq \rho$.

Montrons que $\rho_1 \leq \rho$: Pour $r < \rho_1$, la série $((na_nr^{n-1}))$ est absolument sommable. Et puisque $|a_nz^n| \leq |n\, a_nz^{n-1}|\times|z|$, la série $((a_nz^n))_{n\in\mathbb{N}}$ est sommable. Ceci est vrai pour tout $r < \rho_1$; donc $\rho_1 \leq \rho$.

Nous allons montrer que la somme g de la série dérivée $((na_nz^{n-1}))_{n\in\mathbb{N}}$ *est la dérivée de la somme f de* $((a_nz^n))_{n\in\mathbb{N}}$. Nous ferons deux démonstrations successives.

Première démonstration: Nous allons utiliser les méthodes du chapitre 11 §8.

Choisissons $r<\rho$, nous allons démontrer la dérivabilité de f sur l'intervalle [-r,r] (attention ici z devient une variable réelle). Autrement dit plaçons nous dans $\mathscr{C}^1([-r,r],\mathbb{C})$ muni de la norme ν_∞ (définie par $\nu_\infty(f) = |f(0)| + N_\infty(f'))$. Nous avons $\nu_\infty(a_nz^n) = |a_n\, 0^n| + N_\infty(na_nz^n) = N_\infty(na_nz^n)$.

Choisissons R tel que $r<R<\rho$. Puisque ρ est le rayon de convergence de $((na_nz^n))_{n\in\mathbb{N}}$, on sait qu'il existe M tel que, $\forall n$, $|na_n|\, R^n \leq M$. Il en résulte que $(\forall z\in[-r,r])$ $|na_nz^n| \leq M(\frac{r}{R})^n$; c'est à dire que $\nu_\infty(a_nz^n) = N_\infty(na_nz^n) \leq M(\frac{r}{R})^n$. La série numérique $((M(\frac{r}{R})^n))_{n\in\mathbb{N}}$ est sommable, puisque $r<R$. Donc la série de fonctions (sur [-r,r]) converge normalement pour la norme ν_∞. Ainsi sa somme est une fonction dérivable sur [-r,r], dont la dérivée est la somme de la série des dérivées. Ce qui termine la démonstration.

Seconde démonstration (plus difficile mais qui va nous donner un peu plus). Choisissons $r<\rho$. Il existe M tel que $(\forall n)$ $|a_n| \leq Mr^{-n}$. Choisissons z et h tels que $|z|+|h| < r$. On démontrera l'inégalité :

$$(\alpha) \qquad |f(z+h)-f(z)-hg(z)| \leq \frac{|h|^2Mr}{(r-|z|)^2(r-|z|-|h|)}$$

Restreignons f et g à l'intervalle $]-\rho,+\rho[$ de $\mathbb{R}$. Il est facile de montrer que (r et z étant fixés) la quantité $\frac{|h|^2Mr}{(r-|z|)^2(r-|z|-|h|)}$ est $o(h)$; donc cette inégalité prouve que f est une fonction dérivable, de dérivée g.

Mais pour z complexe nous pouvons aussi montrer que $\lim_{h\to 0}\left(\frac{|h|Mr}{(r-|z|)^2(r-|z|-|h|)}\right) = 0$. *Donc l'inégalité (α) prouve que (z étant complexe) le rapport* $\frac{f(z+h)-f(z)}{h}$ *a une limite quand h tend vers zéro, et que cette limite est g(z). Autrement dit l'application f:* $D(0,\rho) \to \mathbb{C}$ *est dérivable comme fonction de la variable complexe z.*

Démonstration de l'inégalité (α):

Posons $u_n = a_n\left[(z+h)^n - z^n - nhz^{n-1}\right]$ et $v_n = Mr^{-n}\left[(|z|+|h|)^n - |h|^n - n\,|h|\,|z|^{n-1}\right]$. Puisque $|a_n| \le Mr^{-n}$, on a (formule du binôme) $|u_n| \le Mr^{-n}\left|\sum_{i\ge 2} C_n^i\, z^{n-i}h^i\right|$; d'où $|u_n| \le Mr^{-n}\sum_{i\ge 2} C_n^i\,|z|^{n-i}\,|h|^i = v_n$.

La série $((v_n))_{n\in\mathbb{N}}$ est positive et sommable (car différence de séries sommables), appelons S sa somme. Elle est supérieure au module de la somme de $((u_n))_{n\in\mathbb{N}}$, c'est-à-dire à $|f(z+h)-f(z\text{-})-hg(z)|$. Il reste donc à calculer S. On a:

$$v_n = M\left(\frac{|z|+|h|}{r}\right)^n - M\left(\frac{|z|}{r}\right)^n - M\frac{|h|}{r}\,n\left(\frac{|z|}{r}\right)^{n-1}$$

D'où (sommes de séries géométriques et exercice e) :

$$S = M\frac{1}{1-\frac{|z|+|h|}{r}} - M\frac{1}{1-\frac{|z|}{r}} - M\frac{|h|}{r}\frac{1}{\left(1-\frac{|z|}{r}\right)^2}$$

Soit en réduisant : $$S = \frac{M|h|^2 r}{(r-|z|-|h|)(r-|z|)^2}\ .$$

Exemples : a) La série $\left(\left(\frac{x^n}{n}\right)\right)_{n\in\mathbb{N}}$ a pour série dérivée la série $((x^n))_{n\in\mathbb{N}}$. Elles ont pour sommes respectives $-\mathrm{Log}(1-x)$ et $\frac{1}{1-x}$ (rayon de convergence 1).

b) La série $\left(\left(\frac{x^n}{n!}\right)\right)$ est sa propre série dérivée, elle a pour somme e^x qui est sa propre dérivée (rayon de convergence infini).

c) La série entière $((u_nz^n))_{n\in\mathbb{N}}$ où $\left(u_{2n+1} = \frac{(-1)^n}{(2n+1)!}\right.$ et $\left.u_{2n}=0\right)$ a pour série dérivée $((v_n))_{n\in\mathbb{N}}$ où $\left(v_{2n} = \frac{(-1)^n}{(2n)!}\right.$ et $\left.v_{2n+1}=0\right)$; et réciproquement $((v_nz^n))_{n\in\mathbb{N}}$ a pour série dérivée $((-u_nz^n))_{n\in\mathbb{N}}$. Leurs sommes respectives sont $\sin x$ et $\cos x$ (Les rayons de convergence sont infinis).

§ 4 : Produits de séries entières

Soit $((u_n))_{n\in\mathbb{N}}$ et $((v_n))_{n\in\mathbb{N}}$ deux séries de nombres complexes normalement sommables, de sommes respectives U et V. Posons, pour tout n, $w_n = \sum_{i+j=n} u_iv_j$. Alors la série $((w_n))$ est normalement sommable, de somme UV.

Nous avons $U = \sum_i u_i$ et $V = \sum_j v_j$, donc $UV = \sum_{ij} u_iv_j$. C'est la distributivité de l'addition. Puis $\sum_{ij} u_iv_j = \sum_n \sum_{i+j=n} u_iv_j$. C'est l'associativité de l'addition. Malheureusement ces règles de calcul ne sont valables que pour des sommes finies, alors que les signes $\sum$ que nous utilisons ainsi désignent en fait la succession d'une somme finie et d'une limite. C'est cette limite qui rend incorrect le calcul simpliste que nous venons de faire, et la page de calcul qui suit est là pour montrer que le résultat est tout de même vrai lorsque les séries sont normalement sommables.

Posons $U_n = \sum_{i\le n} u_i$, $V_n = \sum_{i\le n} v_i$, et $W_n = \sum_{i\le n} w_i$. Il s'agit de montrer que la suite $(W_n)_{n\in\mathbb{N}}$ a une limite, et que cette limite est le produit des limites des suites $(U_n)_{n\in\mathbb{N}}$ et $(V_n)_{n\in\mathbb{N}}$. Posons

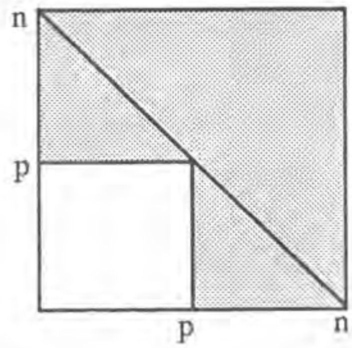
$$A_n = \sum_{i\le n} |u_i| \ , \quad B_n = \sum_{j\le n} |v_j| \quad \text{et} \quad C_n = \sum_{k\le n} \Big(\sum_{i+j=k} |u_i|\,|v_j| \Big).$$

Alors le produit U_nV_n (resp. A_nB_n) est la somme des u_iv_j (resp. $|u_iv_j|$) pour $0\le i\le n$ et $0\le j\le n$. Tandis que W_n (resp. C_n) est la somme des u_iv_j (resp. $|u_iv_j|$) pour $i+j \le n$. Donc $U_nV_n-W_n$ (resp. $A_nB_n - C_n$) est la somme des u_iv_j (resp. $|u_iv_j|$) pour $0 \le i \le n$, $0 \le j \le n$ et $i+j > n$. Il en résulte que:

$$|U_nV_n-W_n| \le A_nB_n - C_n \ .$$

(on est ici en présence de sommes finies, et le module de la somme est au plus égal à la somme des modules). Pour démontrer que la série $((w_n))_{n\in \mathbb{N}}$ est sommable de somme UV, il faut démontrer que la suite $(U_nV_n - W_n)_{n\in \mathbb{N}}$ converge vers 0, et pour cela nous démontrerons que la suite $(A_nB_n - C_n)_{n\in \mathbb{N}}$ converge vers 0.

Soit p la partie entière de $\frac{n}{2}$.



Alors $\alpha_n = A_nB_n - A_pB_p$ est la somme des $|u_iv_j|$ pour $0\le i\le n$, $0\le j\le n$, et (ou bien $i>p$, ou bien $j>p$). Donc α_n contient tous les $|u_iv_j|$ qui sont sommés dans $A_nB_n - C_n$, et ainsi $\alpha_n \ge A_nB_n - C_n \ge 0$.

Mais en écrivant $\alpha_n = [A_n - A_p]\, B_n + [B_n- B_p]\, A_p$, on voit que la suite $(\alpha_n)_{n\in \mathbb{N}}$ converge vers zéro c'est à dire vers $0\times\text{Lim } B_n + 0\times\text{Lim } A_p$. Donc $(A_nB_n - C_n)_{n\in \mathbb{N}}$ converge vers zéro.

Applications: 1) Si $((a_nz^n))_{n\in \mathbb{N}}$ et $((b_nz^n))_{n\in \mathbb{N}}$ sont deux séries entières de rayons de convergence ρ et ρ' et de sommes $f(z)$ et $g(z)$, et si l'on pose $c_n = \sum_{i+j=n} a_ib_j$, alors $((c_nz^n))_{n\in \mathbb{N}}$ a un rayon de convergence au moins égal à $\inf(\rho,\rho')$, et sa somme est égale à $f(z)g(z)$.

2) Puisque $\displaystyle\sum_{p+q=n} \frac{z^p}{p!}\frac{z'^q}{q!} = \frac{(z+z')^n}{n!}$, on a $\displaystyle\sum_{p\ge 0} \frac{z^p}{p!} \sum_{q\ge 0} \frac{z^q}{q!} = \sum_{n\ge 0} \frac{(z+z')^n}{n!}$. C'est à dire $e^z e^{z'} = e^{z+z'}$.

§ 5 : Développement en série entière des fonctions C^∞.

Considérons une application $f\colon U \to \mathbb{R}$, où U est un voisinage de 0 dans $\mathbb{R}$. Existe-t-il une série entière $((a_nz^n))_{n\in \mathbb{N}}$ de rayon de convergence non nul, dont la somme, restreinte au domaine réel, coïncide avec f au voisinage de 0 ?

Une condition nécessaire pour qu'il en soit ainsi est (cf § 3) que f soit indéfiniment dérivable au voisinage de zéro. Et (toujours d'après le § 3) on a alors $f^{(n)}(0) = n!\, a_n$. La série $\big(\big(\frac{f^{(n)}(0)}{n!} z^n\big)\big)_{n\in \mathbb{N}}$ est appelée la série de Taylor de f. Elle existe dès que f est indéfiniment dérivable.

Mais cette condition n'est pas suffisante. Il existe des fonctions indéfiniment dérivables, dont la série de Taylor a un rayon de convergence non nul, mais qui ne sont pas égales à la somme de cette série.
Exemple: Soit $f(x) = e^{-1/x^2}$ pour $x \ne 0$ et $f(0) = 0$. Un calcul classique montre que sa série de Taylor existe, et a tous ses coefficients nuls (c'est à dire que la fonction f est indéfiniment dérivable et que toutes ses dérivées sont nulles à l'origine; cf: chap.11b Ex.4). Elle a donc un rayon de convergence infini, mais sa somme est nulle, et donc différente de la fonction f.

Il existe aussi des fonctions indéfiniment dérivables, dont la série de Taylor a un rayon de convergence nul (elles sont plus difficiles à construire).

Nous allons démontrer que si f est indéfiniment dérivable sur U, et si il existe M et r tels que pour $|x| < r$, on ait, quel que soit $n \geq 0$, $|f^{(n)}(x)| \leq M\, r^{-n}(n!)$, alors la série de Taylor de f a un rayon de convergence au moins égal à r, et sa somme coïncide avec f sur $]-r,+r[$.

En effet posons $f(x) - \sum_{i\leq n-1} f^{(i)}(0)\frac{x^i}{i!} = g_n(x)$. Alors g_n est indéfiniment dérivable, et $g_n^{(i)}(0) = 0$ pour $i < n$, tandis que $g_n^{(n)}(x) = f^{(n)}(x)$. Donc $|g_n^{(n)}(x)| \leq M(n!)r^{-n}$. On en déduit que $(\forall x \in [-r,r])$ $|g_n(x)| \leq Mr^{-n}|x|^n$. Il en résulte que pour $|x| < r$, la suite $(|g_n(x)|)_{n\in\mathbb{N}}$ converge vers zéro; ce qui démontre d'une part que la série $\left(\left(f^{n}(0)\frac{x^n}{n!}\right)\right)_{n\in\mathbb{N}}$ est sommable, d'autre part que sa somme est $f(x)$.

Exercice f: On considère une fonction g définie sur $[-r,r]$ de classe C^∞ et telle que $(\forall x)\ |g^{(n)}(x)| \leq K|x^p|$, montrer que $(\forall x)\ |g^{(n-1)}(x)| \leq \frac{K}{p+1}|x^{p+1}|$. En déduire que si $(\forall x)\ |g^{(n)}(x)| \leq K$, alors $(\forall x)\ |g(x)| \leq \frac{K}{n!}|x^n|$.

On vérifie aisément que toutes les fonctions classiques exp, cos, sin, ... sont justiciables de ce théorème. C'est pourquoi à la fin du 18-ème siècle des mathématiciens comme Lagrange croyaient que toutes les fonctions étaient développables en série entière.

Extension au champ complexe des fonctions usuelles

La série $\left(\left(\frac{z^n}{n!}\right)\right)_{n\in\mathbb{N}}$ a un rayon de convergence infini; par définition, pour $z\in\mathbb{C}$, nous noterons $\exp z$ (ou e^z) la somme de cette série. Nous avons ainsi une application $\exp: \mathbb{C} \to \mathbb{C}$, dont la restriction à $\mathbb{R}$ est la fonction exponentielle ordinaire. Cette fonction vérifie la relation fonctionnelle $e^{z+u} = e^z e^u$. C'est la "bonne définition" de l'exponentielle complexe, celle qui fait apparaitre $z \to e^z$ comme un prolongement naturel de $t \to e^t$.

Il est facile de montrer que $e^{\bar{z}} = (\overline{e^z})$. Donc, pour t réel, $1 = e^{it-it} = e^{it}e^{-it} = |e^{it}|^2$; ce qui montre que e^{it} est un nombre complexe de module 1. Par ailleurs

$$e^{it} = \sum_{n\geq 0} \frac{(it)^n}{n!} = \sum_{p\geq 0} \frac{(-1)^p t^{2p}}{(2p)!} + i\sum_{q\geq 0} \frac{(-1)^q t^{2q+1}}{(2q+1)!} = \cos t + i \sin t.$$

On a donc $e^{x+iy} = e^x(\cos y + i \sin y)$, ce qui redonne la formule qui servait de définition lorsque vous étiez en terminale.

Les fonctions $\mathrm{ch}\, x = \frac{e^x+e^{-x}}{2}$ et $\mathrm{sh}\, x = \frac{e^x-e^{-x}}{2}$ s'étendent de façon tout à fait analogue en posant $\mathrm{ch}\, z = \frac{e^z+e^{-z}}{2}$ et $\mathrm{sh}\, z = \frac{e^z-e^{-z}}{2}$. De même, pour x réel on a $\cos x = \frac{e^{ix}+e^{-ix}}{2}$ et $\sin x = \frac{e^{ix}-e^{-ix}}{2i}$, donc on pose pour $z\in\mathbb{C}$

$$\cos z = \frac{e^{iz}+e^{-iz}}{2} \quad \text{et} \quad \sin z = \frac{e^{iz}-e^{-iz}}{2i}.$$

Ces deux fonctions sont aussi les sommes des séries entières $\left(\left(\frac{(-1)^p z^{2p}}{(2p)!}\right)\right)_{n\in\mathbb{N}}$ et $\left(\left(\frac{(-1)^q z^{2q}}{(2q+1)!}\right)\right)_{n\in\mathbb{N}}$.

Remarquons aussi les relations $\cos z = \mathrm{ch}(iz)$ et $\sin z = \frac{1}{i}\mathrm{sh}\, iz$.

§ 6 : Résolution de quelques équations différentielles

Nous allons utiliser les séries entières pour résoudre certaines équations différentielles. Il s'agit de mettre en œuvre le programme que nous avions fixé au début du chapitre 11. Pour résoudre un problème on construit une suite de fonctions qui en sont des solutions approchées, puis on montre que cette suite a une limite qui est la solution (exacte) du problème.

Nous donnerons simplement un exemple. Trouver une fonction f dérivable deux fois au voisinage de l'origine telle que l'on ait $(\forall t)\quad t^2f''(t) + tf'(t) + f(t) = \frac{1}{1-t}$.

Cherchons une série entière $((a_nz^n))_{n\in\mathbb{N}}$ dont la somme f (restreinte à l'intervalle $]-\rho,+\rho[$), vérifie cette équation. Nous voulons donc avoir

$$z^2\Big(\sum_{n\geq 0} n(n-1)a_nz^{n-2}\Big) + z\Big(\sum_{n\geq 0} n\, a_nz^{n-1}\Big) + \sum_{n\geq 0} a_nz^n = \sum_{n\geq 0} z^n$$

Cette égalité est une égalité entre les sommes supposées de deux séries entières. Nous allons exprimer que ces deux séries ont, pour tout n, le même terme en z^n. Ceci nous donnera les valeurs des a_n. Nous poserons ensuite les problèmes de convergence. Cette méthode qui consiste à oublier dans un premier temps la convergence des séries (c'est à dire en fait l'essentiel de la signification de ce que l'on écrit), est appelée la "phase formelle du calcul". L'expérience prouve que, lorsque l'on travaille avec des fonctions développables en série entière, on a intérêt à traiter d'abord le problème de façon formelle; encore faut-il donner ensuite un sens aux résultats que l'on a trouvés, en démontrant la convergence. C'est le processus que nous avons déjà mis en œuvre à propos des sommes, des produits ou des dérivées de séries entières.

Nous écrirons donc successivement:

Pour n = 0 $\quad 0 + 0 + a_0 = 1$

Pour n = 1 $\quad 0 + 1\times a_1 + a_1 = 1$

Pour n = 2 $\quad 2(2-1)\times a_2 + 2\times a_2 + a_2 = 1$

Pour n = 3 $\quad 3(3-1)\times a_3 + 3\times a_3 + a_3 = 1$

Pour n quelconque $\quad n(n-1)\, a_n + n\, a_n + a_n = 1$

D'où il résulte, pour tout n, $a_n(n^2+1) - 1 = 0$, donc $a_n = \frac{1}{n^2+1}$.

La série entière $\left(\left(\frac{z^n}{n^2+1}\right)\right)_{n\in\mathbb{N}}$ a pour rayon de convergence 1. Sa somme définit une application f: $]-1,+1[\ \to \mathbb{R}$, indéfiniment dérivable. Nous allons montrer que cette somme vérifie l'équation proposée sur l'intervalle $]1,1[$.

Pour tout t, posons $f_n(t) = \sum_{i\leq n} \frac{t^i}{i^2+1}$ et $\varphi_n(t) = \sum_{i\leq n} t^i$. Alors pour tout n nous avons

$$t^2f_n''(t) + tf_n'(t) + f_n(t) = \varphi_n(t)$$

Soit $0 < r < 1$, pour t dans $[-r,r]$ les suites de fonctions $(f_n)_{n\in\mathbb{N}}$, $(f_n')_{n\in\mathbb{N}}$ et $(f_n'')_{n\in\mathbb{N}}$ convergent uniformément vers f, f' et f''. Donc, en passant à la limite:

$$t^2f''(t) + t\,f'(t) + f(t) = \frac{1}{1-t}.$$

Ceci quelque soit $t\in[-r,r]$. Mais r est quelconque (et < 1) donc ceci est vrai pourvu que $|t|<1$.

Exercice g: Trouver toutes les fonctions développables en série entière au voisinage de 0, qui sont solutions de l'équation différentielle $t\,f''(t) - f(t) = 1$.

On note f_1 la solution dont la dérivée à l'origine vaut 1. Trouver un polynôme P tel que, sur l'intervalle $[-10,10]$, $N_\infty(f_1-P) \leq 10^{-4}$. Est-ce que, sur l'intervalle $[-10,10]$, $N_\infty(f_1' - P') \leq 10^{-4}$?

Exercices sur le chapitre 12

* **Exercice 1**: Déterminer les rayons de convergence des séries

$u_n = n^n z^n$ $\quad w_n = \frac{1}{n^n} z^n$ $\quad x_n = n\sqrt{n}\, z^n$ $\quad y_n = \frac{n^{n+1}}{n!} z^n$ $\quad a_n = (\text{Log } n!) z^n$

$b_n = \frac{n!}{n^n} z^n$ $\quad c_n = \text{Log } n\, z^n$ $\quad d_n = \frac{\text{sh } n}{\text{ch}^2 n} z^n$ $\quad g_n = \text{ch } n\, z^n$

$f_n = P(n)\, z^n$ (où P est un polynôme) $\quad h_n = \frac{1}{\text{ch } n} z^n$

* **Exercice 2**: Développer en série entière les fonctions suivantes. En précisant les rayons de convergence des séries trouvées.

$f(z) = \sin(\frac{\pi}{4} + z)$ $\quad g(z) = e^z \cos z$ $\quad k(z) = \frac{1}{1+z+z^2}$ $\quad m(z) = \text{Arctg } \frac{1+z}{1-z}$

* **Exercice 3**: Déterminer le rayon de convergence des séries suivantes:

$a = (((1+\frac{1}{n})^n z^n))_{n \geq 1}$ $\quad b = (((2^n-1) z^n))_{n \in \mathbb{N}}$

$c = ((\text{Arctg } (n^\alpha) z^n))_{n \in \mathbb{N}}$ $(\alpha \in \mathbb{R})$ $\quad d = ((\frac{2^n}{3^n+4^n} z^n))_{n \in \mathbb{N}}$

** **Exercice 4**: On considère la série entière $u = ((\frac{z^n}{n(n+1)}))_{n \geq 1}$

a) Quel est son rayon de convergence ρ ?

b) Montrer qu'elle est normalement sommable (pour la norme N_∞) sur le disque de centre l'origine et de rayon ρ.

c) Calculer sa somme f, en calculant d'abord la dérivée seconde de la fonction $x \to x\, f(x)$.

** **Exercice 5**: Déterminer les rayons de convergence et les sommes des séries suivantes

$u_n = n(n-1)(n-2) z^n$ $\quad v_n = \frac{1}{n(n-1)} z^n$ $\quad u_n = \cos(nt) z^n$ $\quad v_n = \sin(nt) z^n$.

** **Exercice 6**: Soit P un polynôme.

a) Calculer le rayon de convergence de la série $u = ((\frac{P(n)}{n!} z^n))_{n \geq 1}$.

b) On suppose que $P(n) = n^2+1$. Calculer la somme de u.

c) On suppose que $P(n) = n^5+n^4-n+2$. Calculer la somme de u.

* **Exercice 7**: Développer en série entière les fonctions (on précisera les rayons de convergence)

$f(x) = \frac{1}{x-2}$ $\quad g(x) = \frac{1}{x^2-4x+3}$ $\quad h(x) = \frac{1}{x^2+1}$

* **Exercice 8**: Quels sont les rayons de convergence des séries

$((\sqrt{n!}\ x^n))_{n \in \mathbb{N}}$ $\quad (((\frac{1}{n!} + n) x^n))_{n \in \mathbb{N}}$ (calculer la somme) $\quad (((n^3+n^2+1) x^n))_{n \in \mathbb{N}}$ (calculer la somme)

$((\frac{(2n)!\ x^n}{n^n}))_{n \in \mathbb{N}}$ $\quad (((\frac{1}{n} + n) x^n))_{n \in \mathbb{N}}$ (calculer la somme) $\quad ((\frac{n+1}{(2n)!} x^n))_{n \in \mathbb{N}}$ (calculer la somme)

$(((\frac{1}{n} + \frac{1}{(n^n)!}) x^n))_{n \in \mathbb{N}}$

* **Exercice 9**: Trouver un développement en série entière de la fonction $z \to \frac{1}{1-z+z^2}$. Quel est le rayon de convergence de la série trouvée ?

* **Exercice 10**: Déterminer le rayon de convergence des séries suivantes

$((\text{ch}\, n\, x^n))_{n\in \mathbb{N}}$ $((\log n\, x^n))_{n\in \mathbb{N}}$ $((\frac{\log n}{n!} x^n))_{n\in \mathbb{N}}$ $((\frac{1}{2+\cos n} x^n))_{n\in \mathbb{N}}$

** **Exercice 11**: Déterminer le rayon de convergence de la série entière $(((1+\cos n)\, x^n))_{n\in \mathbb{N}}$. Déterminer sa somme. On utilisera le fait que $\cos n + i \sin n = e^{in}$.

* **Exercice 12**: Trouver le rayon de convergence et la somme des séries entières

$(((n+1)n\, x^n))_{n\in \mathbb{N}}$ $(((n+1)\, x^n))_{n\in \mathbb{N}}$ $((n^2 x^n))_{n\in \mathbb{N}}$

* **Exercice 13**: Développer en série entière les fonctions suivantes:

a) $t \to \sin t \,\text{sh}\, t$. b) $t \to e^t \cos t$.

* **Exercice 14** : Trouver toutes les séries entières dont la somme $f(z)$ est solution de l'équation différentielle

$$(1+z)^2 f'' - zf' + f = 0.$$

Préciser le rayon de convergence de ces séries.

** **Exercice 15**: a) Montrer que $f(x) = (x+\sqrt{1+x^2})^m$ est solution de l'équation différentielle:

$$(1+x^2)\, y'' + xy' - m^2 y = 0$$

b) En déduire la série de Taylor de la fonction f.

* **Exercice 16**: Déterminer les séries entières dont les sommes sont solutions de l'équation différentielle $(1+x^2)\, y'' - 2y = x$. Quels sont leurs rayons de convergence ?

* **Exercice 17**: Déterminer les séries entières dont les sommes sont solutions de l'équation $x(2-x)\, y' + (1-x)\, y = 1$. Quel est leur rayon de convergence ?

** **Exercice 18**: On considére l'équation différentielle $y'' = xy$.

a) A quelles conditions la série entière $((a_n z^n))_{n\in \mathbb{N}}$ est elle - formellement - solution de cette équation ? Montrer que toutes ces séries ont un rayon de convergence infini.

b) Montrer que ($\forall \alpha$ et β) il existe une unique solution (développable en série entière) g de cette équation telle que $g(0) = \alpha$ et $g'(0) = \beta$.

c) On choisit $\alpha = \beta = 1$, calculer à 10^{-3} près la valeur de g en 1,5. Calculer $g'(1,5)$.

** **Exercice 19**: Soit $((a_n z^n))_{n\in \mathbb{N}}$ une série entière. On suppose que la suite $(a_n)_{n\in \mathbb{N}}$ est périodique, montrer que la série a un rayon de convergence égal à 1, et montrer que sa somme est une fraction rationnelle (que l'on précisera).

On suppose maintenant qu'il existe k et p tels que $(\forall n)\ a_{n+p} = ka_n$. Quel est le rayon de convergence de la série ? Quelle est sa somme ?

** **Exercice 20**: Soit E l'ensemble des séries entières dont le rayon est au moins égal à ρ ($\rho>0$). Montrer que E est un espace vectoriel. A toute série on associe sa série dérivée, on obtient ainsi une application linéaire de E dans lui même. Quelle est son image ? Quel est son noyau ?

Fonctions de classe C^1 sur $\mathbb{R}^2$ et sur $\mathbb{R}^3$

Pour étudier le comportement local des fonctions d'un espace vectoriel dans un autre, une première stratégie consiste à essayer de se ramener aux fonctions de $\mathbb{R}$ dans lui même. Pour cela il existe deux procédures, que nous avons déjà rencontrées au chapitre 10 à propos de la continuité.

Si $f: \mathcal{E} \to \mathcal{F}$, en prenant une base de $\mathcal{F}$, on peut essayer de remplacer f par ses fonctions coordonnées, qui sont à valeurs dans $\mathbb{R}$.

Si $f: \mathcal{E} \to \mathcal{F}$, en prenant une base de $\mathcal{E}$, on peut restreindre f à des sous-espaces de dimension 1 de $\mathcal{E}$. En identifiant ces sous-espaces à $\mathbb{R}$, on se ramène au cas de fonctions de $\mathbb{R}$ dans $\mathcal{F}$.

La généralisation de la notion de dérivabilité aux fonctions de $\mathcal{E}$ dans $\mathcal{F}$ que nous allons donner, utilise essentiellement le second de ces deux artifices. Ceci mène à définir la notion de dérivée partielle.

Mais la bonne notion, celle qui donne des théorèmes intéressants, n'est pas celle de 'fonction ayant des dérivées partielles', c'est celle de 'fonction ayant des dérivées partielles continues', ce que l'on appelle des fonctions de classe C^1.

Dans tout ce chapitre on considère des espaces vectoriels de dimension finie sur $\mathbb{R}$. Pour tout ce qui concerne la différentiabilité les espaces de dimension au moins égale à 2 ont tous des comportements assez analogues, mais, sur bien des points, assez différents des espaces de dimension 1 (pour lesquels on retrouve la théorie déjà connue des fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$). Tout ce qui suit est valable pour les espaces de dimension n quelconque mais, pour simplifier, les démonstrations (et bon nombre d'énoncés) sont écrites en dimension 2.

Dans tout ce qui suit E, F,... désignent des espaces normés de dimension n, p,... et U est un ouvert de E.

§ 1 : Dérivées partielles et applications de classe C^1.

Dérivées selon un vecteur.

Soit X un vecteur de E, et $f: U(\subset E) \to F$. Pour tout M_0 dans U, nous pouvons considérer l'application

$$f_X: t \to f(M_0+tX)$$

(qui est une fonction de $\mathbb{R}$ dans F définie pour t voisin de 0)

Si f_X est dérivable en 0, on dit que f a, au point M_0, une dérivée par rapport au vecteur X. Cette dérivée est un vecteur de F. Nous la noterons $D_Xf(M_0)$.

Une propriété évidente: Si X = 0, $D_Xf(M_0)$ existe, et est nul. Si $Y= \lambda X$, et si $D_Xf(M_0)$ existe, alors $D_Yf(M_0)$ existe et est égal à $\lambda D_Xf(M_0)$.

Dérivées partielles.

Fixons une base $\{e_1,...,e_n\}$ de E, nous appellerons j-ième dérivée partielle de f au point M_0 (par rapport à cette base) le vecteur $D_{e_j}f(M_0)$, que nous noterons souvent $\frac{\partial f}{\partial x_j}(M_0)$ (si les coordonnées dans la base $\{e_i\}$ sont notées $x_1,..x_n$).

Changeons de point de vue: Soit M un point de E, et soit $\{x_1,...,x_n\}$ ses coordonnées dans la base $\{e_1,...,e_n\}$, nous pouvons noter $f(x_1,...,x_n)$ au lieu de f(M). Ainsi f, qui était une fonction d'une variable vectorielle, apparait comme une fonction de n variables numériques. Avec ces nouvelles notations, et en notant $(x_1^0,...,x_n^0)$ les coordonnées de M_0, nous avons

$$f_{e_j}(t) = f(x_1^0,...,x_j^0+t,...,x_n^0)$$

Calculer la j-iéme dérivée partielle de f en M_0, c'est donc calculer la dérivée de la fonction de n variables par rapport à son j-ième argument (en faisant comme si tous les autres étaient des constantes).

Exemple: Soit f: $\mathbb{R}^3 \to \mathbb{R}$ qui à M associe sa distance à l'origine (pour la norme euclidienne naturelle de $\mathbb{R}^3$). Sa valeur en M(x,y,z) est donnée par la formule $f(x,y,z) = \sqrt{x^2+y^2+z^2}$. En tout point différent de l'origine, elle a des dérivées partielles, qui s'écrivent:

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

A l'origine elle n'a aucune dérivée partielle (car $f_{e_i}(t) = |t|$ n'est pas dérivable en 0).

Les fonctions de classe C^1 et leurs différentielles.

Si f: $U \to \mathbb{R}$ a en chaque point de U des dérivées partielles, celles ci définissent n fonctions $(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})$ de U dans F. Si ces fonctions sont continues, on dit que f est de classe C^1.

Soit f: $U \to F$ est de classe C^1, et soit M un point de U. *Il existe une unique application linéaire d_Mf de E dans F, qui prend la valeur $\frac{\partial f}{\partial x_i}(M)$ sur le vecteur e_i. Cette application linéaire est appelée la différentielle de f en M.*

Exemple 1: La différentielle de la fonction f de l'exemple précédent, au point M(1,–2,2) est l'application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}$, qui

$$* \text{ à } (1,0,0) \text{ associe } 1/3 = \frac{\partial f}{\partial x}(1,-2,2)$$
$$* \text{ à } (0,1,0) \text{ associe } -2/3 = \frac{\partial f}{\partial y}(1,-2,2)$$
$$* \text{ à } (0,0,1) \text{ associe } 2/3 = \frac{\partial f}{\partial z}(1,-2,2)$$

Sa matrice, dans les bases naturelles, est donc (1/3,–2/3,2/3).

Exemple 2: Soit f l'application de $\mathbb{R}^3$ dans $\mathbb{R}^2$ qui à tout M associe le couple formé de sa distance à l'origine et de sa distance au point (1;2;3). Ses dérivées partielles en M(x,y,z) sont donc des vecteurs de $\mathbb{R}^2$. Nous avons:

$$f(x,y,z) = (\sqrt{x^2+y^2+z^2}, \sqrt{(x-1)^2+(y-2)^2+(z-3)^2})$$
$$\frac{\partial f}{\partial x}(x,y,z) = (\frac{x}{\sqrt{x^2+y^2+z^2}}, \frac{x-1}{\sqrt{(x-1)^2+(y-2)^2+(z-3)^2}})$$
$$\frac{\partial f}{\partial y}(x,y,z) = (\frac{y}{\sqrt{x^2+y^2+z^2}}, \frac{y-2}{\sqrt{(x-1)^2+(y-2)^2+(z-3)^2}})$$

$$\frac{\partial f}{\partial z}(x,y,z) = \left(\frac{z}{\sqrt{x^2+y^2+z^2}}, \frac{z-3}{\sqrt{(x-1)^2+(y-2)^2+(z-3)^2}}\right)$$

Sa différentielle au point M(3;4;0) est donc l'application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^2$, dont la matrice dans les bases naturelles s'écrit:

$$\begin{pmatrix} 3/5 & 4/5 & 0 \\ 2/\sqrt{17} & 2/\sqrt{17} & -3/\sqrt{17} \end{pmatrix}$$

Remarque: Cette définiton des applications de classe C^1 fait intervenir de façon essentielle une base de E, et il n'est pas évident que si nous avions choisi une autre base nous aurions obtenu les mêmes applications de classe C^1, et la même différentielle. Ceci est vrai, mais ne peut être démontré dés maintenant (voir remarque à la fin du §2).

§ 2 : Développement à l'ordre 1 des fonctions de classe C^1.

Soit f: U → F est de classe C^1, et soit M un point de U. Il existe une fonction ε (définie sur E) continue et nulle en 0, telle que ($\forall h \in E$)

$$N_F(f(M+h) - f(M) - d_Mf(h)) \le N_E(h)\ \varepsilon(h)$$

Nous allons le démontrer pour une fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$. Nous utiliserons la norme(*) de $\mathbb{R}^2$ définie par $N(x,y) = \sup(|x|,|y|)$. Choisissons $M = (x_0,y_0)$. Notons D_1f et D_2f au lieu de $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$.

Il nous faut majorer:

$$A(h) = f(M+h) - f(M) - \varphi\, D_1f(M) - \psi\, D_2f(M) \qquad (\text{où } h = (\varphi,\psi))$$

La formule des accroissements finis(*), appliquée à $t \to f(x_0+\varphi,t)$ entre y_0 et $y_0+\psi$, nous donne:

$$\exists\theta_2\ (0 < \theta_2 < 1) \text{ tel que } \ f(x_0+\varphi,y_0+\psi) - f(x_0+\varphi,y_0) = \psi\, D_2f(x_0+\varphi,y_0+\theta_2\psi)$$

La même formule appliquée à $t \to f(t,y_0)$ entre x_0 et $x_0+\varphi$, nous donne:

$$\exists\theta_1\ (0 < \theta_1 < 1) \text{ tel que } \ f(x_0+\varphi,y_0) - f(x_0,y_0) = \varphi\, D_2f(x_0+\theta_1\varphi,y_0)$$

En sommant ces deux formules nous obtenons:

$$A(h) = [f(x_0+\varphi,y_0+\psi) - f(x_0+\varphi,y_0) - \psi\, D_2f(M)] + [f(x_0+\varphi,y_0) - f(x_0,y_0) - \varphi\, D_1f(M)]$$

$$A(h) = [\psi\, D_2f(x_0+\varphi,y_0+\theta_2\psi) - \psi\, D_2f(x_0,y_0)] + [\varphi\, D_1f(x_0+\theta_1\varphi,y_0) - \varphi\, D_1f(x_0,y_0)]$$

$$A(h) = \psi\, [D_2f(x_0+\varphi,y_0+\theta_2\psi) - D_2f(x_0,y_0)] + \varphi\, [D_1f(x_0+\theta_1\varphi,y_0) - D_1f(x_0,y_0)]$$

Soit, puisque $|\varphi| \le N_E(h)$ et $|\psi| \le N_E(h)$:

$$|A(h)| \le N_E(h)\,|D_2f(x_0+\varphi,y_0+\theta_2\psi) - D_2f(x_0,y_0)| + N_E(h)\,|D_1f(x_0+\theta_1\varphi,y_0) - D_1f(x_0,y_0)| = N_E(h)\ \varepsilon(h)$$

Mais puisque les fonctions D_1f et D_2f sont continues en M: "lorsque h tend vers 0", les quantités $D_2f(x_0+\varphi,y_0+\theta_2\psi) - D_2f(x_0,y_0)$ et $D_1f(x_0+\theta_1\varphi,y_0) - D_1f(x_0,y_0)$ "tendent vers 0". Et ainsi A(h) est majoré par une quantité de la forme $N_E(h)\varepsilon(h)$ où ε tend vers zéro lorsque h tend vers zéro (donc ε se prolonge en une fonction continue et nulle en 0).

Exercice a : Précisons ce dernier argument.

a) Posons $k = (\varphi,\theta_2\psi)$; Puisque θ_2 dépend de φ et ψ, nous écrirons plutôt $k = (\varphi,\theta_2(\varphi,\psi)\psi)$. Montrer que $N(k) \le N(h)$. Ecrire que D_2f est continue en M (($\forall\alpha>0$) ($\exists\eta>0$) ...).

En déduire que la fonction $(\varphi,\psi) \to |D_2f(x_0+\varphi,y_0+\theta_2(\varphi,\psi)\psi) - D_2f(x_0,y_0)|$ est continue en (0,0).

b) Démontrer de même que $(\varphi,\psi) \to |D_1f(x_0+\theta_1(\varphi)\varphi,y_0) - D_1f(x_0,y_0)|$ est continue en (0,0).

c) En déduire que la fonction ε est continue et nulle pour h = 0.

(*) On démontre que, puisque toutes les normes sur $\mathbb{R}^2$ sont équivalentes, le résultat ne dépend pas de la norme utilisée.

(*) Rappelons l'énoncé de cette formule : Si f : [a,b] → $\mathbb{R}$ est continue sur [a,b] et dérivable sur]a,b[, alors il existe un nombre θ compris entre 0 et 1 tel que $f(b) - f(a) = (b-a)\, f'(a+\theta(b-a))$.

Conséquence: *Si f est de classe C^1, pour tout point M et tout vecteur X, la dérivée $D_Xf(M)$ existe et est égale à $d_Mf(X)$.* Démonstration: Faisons h = tX dans la formule précédente, nous avons

$$N_F(f(M+tX) - f(M) - d_Mf(tX)) \leq N_E(tX)\,\varepsilon(tX) = |t|\, N_E(X)\,\varepsilon(tX)$$

Autrement dit, avec les notations du début du §1:

$$N_F(f_X(t) - f_X(0) - t\, d_Mf(X)) \leq |t|\, N_E(X)\,\varepsilon(tX)$$

Et puisque $\varepsilon(tX)$ tend vers zéro lorsque t tend vers zéro, $|t|\, N_E(X)\,\varepsilon(tX) = o(t)$, donc la fonction f_X (à valeurs vectorielles dans F), est dérivable en 0, de dérivée $d_Mf(X)$.

Ainsi la différentielle d_Mf apparait comme l'application qui à tout vecteur X associe D_Xf; en particulier une condition nécessaire (non suffisante !!) pour que f soit de classe C^1, est que

d'une part ($\forall X$ et $\forall M$) $D_Xf(M)$ existe

d'autre part ($\forall X$, $\forall Y$ et $\forall M$) $D_Xf(M) + D_Yf(M) = D_{X+Y}f(M)$.

La notion de fonction différentiable.

De façon générale on dit qu'une fonction $f: U \to F$ est différentiable en M, s'il existe une application linéaire λ de E dans F, et une application $\varepsilon: E \to \mathbb{R}$, continue et nulle en 0, telles que:

$$(\forall h) \quad N_F(f(M+h) - f(M) - \lambda(h)) \leq N_E(h)\,\varepsilon(h)$$

Nous venons de montrer que les fonctions de classe C^1 sont différentiables; il existe des fonctions qui sont différentiables et ne sont pas de classe C^1; nous n'en donnerons aucun exemple.

Exercice b: Considérons $f: \mathbb{R}^2 \to \mathbb{R}$ définie en coordonnées polaires par $f(\rho,\theta) = \rho \sin 3\theta$. Soit X le vecteur de coordonnées polaires (r,ω). Montrer que $D_Xf(0)$ existe, et donner sa valeur. Montrer que $X \to D_Xf(0)$ n'est pas une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}$, et que f ne peut donc pas être différentiable en 0.

Exercice c: Considérons l'application $f: \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = \dfrac{x^2y}{x^2+y^2}$ si $(x,y) \neq (0,0)$, et par $f(0,0) = 0$. Soit X le vecteur de coordonnées (a,b). Montrer que $D_Xf(0)$ existe, et donner sa valeur. Montrer que $X \to D_Xf(0)$ n'est pas une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}$, et que f ne peut donc pas être différentiable en 0.

Exercice d: Soit λ une application linéaire de E dans F. Montrer qu'elle est différentiable en tout point de E. Quelle est sa différentielle ?

Mêmes questions pour l'application affine $X \to A_0 + \lambda(X)$ (A_0 constante)

Nous savons qu'une fonction $f: \mathbb{R} \to \mathbb{R}$ est dérivable en x_0 (de dérivée α) si

$$f(x_0+h) = f(x_0) + h\,\alpha + o(h)$$

La définition des fonctions différentiables peut être considérée comme une généralisation de cette formule. La fonction $h \to h\alpha$ est remplacée par une application linéaire. Et le terme complémentaire $N_E(h)\,\varepsilon(h)$ exprime le $o(h)$. Ainsi la notion de différentiabilité se trouve ramenée à celle de développement limité d'ordre 1. La fonction affine:

$$h \to f(M) + d_Mf(h)$$

Est la partie polynômiale du développement à l'ordre 1 de f en M. On l'appelle l'application affine 'tangente à f en M'. C'est la fonction affine qui 'approche le mieux f au voisinage de M'.

On retiendra que, pour une fonction f de $\mathbb{R}$ dans $\mathbb{R}$, dérivabilité est synonyme de différentiabilité, et que la dérivée est l'unique coefficient de la matrice de la différentielle. Autrement dit la différentielle d_mf est l'application qui à h associe $hf'(m)$.

Les fonctions coordonnées d'une application de classe C^1.

Soit $\{\varphi_1,\dots,\varphi_n\}$ une base de F, et soit $f: U \to F$. Notons $f^1,\dots,f^n$ les fonctions coordonnées de f dans cette base. Alors f est de classe C^1, si et seulement si toutes les fonctions coordonnées f^j le sont. De plus les fonctions coordonnées de $\frac{\partial f}{\partial x_i}$ sont les $\frac{\partial f^j}{\partial x_i}$.

En effet lorsque nous parlons de dérivées partielles, nous parlons de dérivées d'une fonction d'une seule variable (puisque, lorsqu'on dérive, toutes les autres sont considérées comme des constantes). Et pour les fonctions d'une variable nous savons que les dérivées des fonctions coordonnées sont les fonctions coordonnées de la dérivée.

Les fonctions "définies par des formules".

Pour définir une fonction de classe C^1, il y a deux méthodes. D'une part écrire une formule. D'autre part fabriquer des limites de fonctions définies par des formules, comme nous avons appris à le faire au chapitre 11 pour les fonctions d'une variable. Dans le cadre de ce cours, ce second procédé ne sera pas utilisé pour les fonctions sur $\mathbb{R}^n$ ($n\geq 2$). Les fonctions 'de plusieurs variables' que nous rencontrerons seront données par des formules, c'est à dire obtenues à partir des fonctions classiques (exp, log, sin, cos,...) par composition, et par des sommes, des produits et des quotients. Ces fonctions sont toutes de classe C^1, avec les restrictions suivantes:

Les fonctions $\sqrt{t}$, $\sqrt[3]{t}$, t^k ($k \notin \mathbb{N}$) ne sont pas dérivables en 0, donc, lorsque f est de classe C^1, $\sqrt{f}$, $\sqrt[3]{f}$, f^k ($k \notin \mathbb{N}$) n'ont (en général) pas de dérivées partielles (et donc ne sont pas différentiables) aux points u tels que $f(u) = 0$.

Les fonctions Arcsin et Arccos ne sont pas dérivables en 1 et -1 (de même Argch n'est pas dérivable en 1), donc Arcsin (f), Arccos (f) n'est (en général) pas différentiable aux points u tels que $f(u) = \pm 1$ (de même Argch (f) n'est (en général) pas différentiable aux points u tels que $f(u) = 1$).

Ceci nous donne de nombreux exemples de fonctions dont nous pouvons affirmer qu'elles sont différentiables. Nous savons aussi calculer leurs dérivées partielles; ce qui nous donnera (la matrice de) leur différentielle.

Remarque: Considérons maintenant deux bases $\{e_i\}$ et $\{\varepsilon_j\}$ de E, il existe des scalaires α_j^i tels que ($\forall j$) $\varepsilon_j = \sum_j \alpha_j^i e_i$. Supposons que les dérivées partielles de f dans la base $\{e_i\}$ existent et sont continues.

La formule liant la différentielle aux dérivées partielles $\frac{\partial f}{\partial x_i}$, nous dit alors que les dérivées $D_M f(\varepsilon_j)$ existent et que (en notant x_i les coordonnées dans la base $\{e_i\}$, ξ_j les coordonnées dans la base $\{\varepsilon_j\}$):

$$\frac{\partial f}{\partial \xi_j}(M) = d_M f(\varepsilon_j) = d_M f\Big(\sum_j \alpha_j^i e_i\Big) = \sum_j \alpha_j^i\, d_M f(e_i) = \sum_j \alpha_j^i \frac{\partial f}{\partial x_i}(M)$$

Sachant ainsi l'effet du changement de base sur les dérivées partielles, nous pouvons constater que les $\frac{\partial f}{\partial \xi_j}$ existent également et sont continues (car combinaisons linéaires de fonctions continues); ainsi se trouve démontré que la notion de fonction de classe C^1 ne dépend pas de la base qui sert à la définir (cf: la remarque qui termine le §1).

§ 3 : Composition des applications de classe C^1.

Un lemme préliminaire

Soit X et Y deux espaces normés de dimension finie (notons N_X et N_Y leurs normes), et soit φ une application linéaire de X dans Y; il existe un nombre K, tel que $(\forall x\in X)\ N_Y(\varphi(x)) \le K\, N_X(x)$.

Pour nous en convaincre choisissons des bases $\{e_i\}$ et $\{f_j\}$ de X et de Y, et définissons de nouvelles normes par $\nu_X(\sum x_i e_i) = \sup(|x_i|)$ et $\nu_Y(\sum y_j f_j) = \sup(|y_j|)$. Ces nouvelles normes sont équivalentes aux anciennes; il existe donc des nombres (strictement positifs) α,β,γ et δ, tels que

$$(\forall x\in X) \qquad \alpha\, N_X(x) \le \nu_X(x) \le \beta\, N_X(x)$$
$$(\forall y\in Y) \qquad \gamma\, N_Y(y) \le \nu_Y(y) \le \delta\, N_Y(y)$$

Soit k un majorant des valeurs absolues des coefficients de la matrice de φ dans ces bases. Il est clair que toutes les coordonnées de $\varphi(x)$ sont au plus égales à $n\times k\times \nu_X(x)$ (où n = dim X). Par conséquent $\nu_Y(\varphi(x)) \le n\times k\times \nu_X(x)$.

Nous aurons donc, pour tout x:

$$N_Y(\varphi(x)) \le \frac{1}{\gamma}\nu_Y(\varphi(x)) \le \frac{1}{\gamma}\times n\times k\times \nu_X(x) \le \frac{1}{\gamma} n\times k\times \beta\, N_X(x)$$

Et il suffit de poser $K = \frac{1}{\gamma}\times n\times k\times\beta$ pour avoir le résultat annoncé.

Continuité des applications différentiables.

Soit f différentiable en M, et soit K tel que $(\forall x)\ N_F(d_M f(x)) \le K\, N_E(x)$. Puisque dans l'inégalité:

$$(\forall h) \qquad N_F(f(M+h) - f(M) - d_M f(h)) \le N_E(h)\,\varepsilon(h)$$

la fonction ε est continue et nulle en 0, nous pouvons trouver α (>0) tel que pour $N_E(h) \le \alpha$, $\varepsilon(h)$ soit au plus égal à 1. Alors pour $N_E(h) \le \alpha$, nous aurons (inégalité du triangle)

$$N_F(f(M+h) - f(M)) \le N_F(d_M f(h)) + N_E(h)$$
$$N_F(f(M+h) - f(M)) \le (K+1)\, N_E(h) \qquad \text{(K donné par le lemme ci-dessus)}$$

En particulier si la suite $(h_n)_{n\in\mathbb{N}}$ converge vers 0 (c'est à dire si la suite $(M+h_n)_{n\in\mathbb{N}}$ converge vers M), la suite $(N_F(f(M+h_n) - f(M)))_{n\in\mathbb{N}}$ converge aussi vers 0 (autrement dit la suite $(f(M+h_n))_{n\in\mathbb{N}}$ converge vers f(M). Ceci signifie (cf: chap 10 §4) que f est continue en M.

Nous retiendrons que toute fonction différentiable en M (en particulier toute fonction de classe C^1) est continue.

En fait l'inégalité que nous venons de démontrer nous donne un résultat un peu plus précis que la continuité puisqu'elle montre que $(N_F(f(M+h_n) - f(M)))_{n\in\mathbb{N}} = O(N_E(h_n))_{n\in\mathbb{N}}$. *Autrement dit la suite* $(N_F(f(M+h_n) - f(M)))_{n\in\mathbb{N}}$ *converge vers zéro 'au moins aussi vite que'* $(N_E(h_n))_{n\in\mathbb{N}}$.

Composées d'applications différentiables.

La composée d'une application f différentiable en M et d'une application g différentiable en f(M), est différentiable en M; et la différentielle de la composée est donnée par la formule:

$$d_M(g\circ f) = d_{f(M)}g\circ d_M f.$$

Démonstration: Soit E, F et G trois espaces normés de dimension finie. Supposons f: $U(\subset E) \to F$, $f(U)\subset V\subset F$, et g: $V \to G$. Il nous faut majorer:

$$A(h) = g\circ f(M+h) - g\circ f(M) - d_{f(M)}g\circ d_M f(h)$$

Soit $$A(h) = g\circ f(M+h) - g\circ f(M) - d_{f(M)}g[f(M+h)-f(M)] + d_{f(M)}g[f(M+h)-f(M)-d_M f(h)]$$

*Majoration de $B(h) = d_{f(M)}g[f(M+h)-f(M)-d_Mf(h)]$. Puisque f est différentiable en M, il existe une fonction η nulle et continue en 0, telle que $N_F(f(M+h)-f(M)-d_Mf(h)) \leq N_E(h)\, \eta(h)$. Et le lemme ci dessus nous dit qu'il existe K_1 tel que:

$$N_G(B(h)) \leq K_1\ N_F[f(M+h)-f(M)-d_Mf(h)] \leq K_1\ N_E(h)\, \eta(h)$$

*Majoration de $k(h) = f(M+h)-f(M)$. Nous venons de voir que $N_F(k(h)-d_Mf(h)) \leq N_E(h)\, \eta(h)$, où η est continue est nulle en 0. Pour h assez petit, $\eta(h) \leq 1$, et alors (inégalité du triangle):

$$N_F(k(h)) \leq N_F(d_Mf(h)) + N_E(h)\times 1$$

En appliquant le lemme ci-dessus nous obtenons un nombre K tel que $N_F(d_Mf(h)) \leq K\ N_E(h)$; donc

$$N_F(k(h)) \leq N_E(h)\times(1+K)$$

*Majoration de $C(h) = g\circ f(M+h) - g\circ f(M) - d_{f(M)}g[f(M+h)-f(M)]$. Puisque g est differentiable en f(M) il existe une fonction η' (définie au voisinage de 0 dans F) continue et nulle en 0, telle que:

$$N_G(C(h)) = N_G\{g(f(M)+k(h)) - g(f(M)) - d_{f(M)}g[f(M)+k(h)-f(M)]\} \leq N_F(k(h))\, \eta'(k(h))$$

$$N_G(C(h)) \leq N_E(h)\times(1+K)\, \eta'(k(h))$$

Posons $\eta''(h) = \eta'(k(h))$; nous avons:

$$N_G(C(h)) \leq (K+1)\, N_E(h)\, \eta''(h)$$

Nous avons $\eta''(0) = 0$; et, puisque f est continue en M, et que η est une fonction continue en 0, η'' est une fonction de h qui est continue en 0.

Les majorations de B et de C ainsi trouvées nous permettent d'affirmer que:

$$N_G(A(h)) \leq N_E(h)\ \{K_1\, \eta(h) + (K+1)\, \eta''(h)\}$$

Où $\{K_1\, \eta(h) + (K+1)\, \eta''(h)\}$ est continue et nulle en 0. Ce qui prouve que $g\circ f$ est différentiable en M.

Composition des applications de classe C^1

La composée de deux fonctions de classe C^1, est une fonction de classe C^1.

En effet si f et g sont de classe C^1, elles sont différentiables (en chaque point); donc leur composée est différentiable (en chaque point), et elle a des dérivées partielles. Il reste à montrer que ces dérivées sont continues. Pour cela nous allons les calculer; elles sont données par la formule $d_M(g\circ f) = d_{f(M)}g\circ d_Mf$, qu'il nous faut expliciter.

Des bases de $\{\varepsilon_i\}$ de E, $\{\varphi_j\}$ de F et $\{\gamma_k\}$ de G étant choisies, nous aurons

$$\frac{\partial(g\circ f)}{\partial x_i}(M) = d_M(g\circ f)(\varepsilon_i) = d_{f(M)}g\circ d_Mf(\varepsilon_i) = d_{f(M)}g(d_Mf(\varepsilon_i)).$$

Dans cette formule $d_Mf(\varepsilon_i)$ est la i-ème dérivée partielle de f, qui (en notant f^j les fonctions coordonnées de f) est égale à $\sum_j \frac{\partial f^j}{\partial x_i}(M)\, \varphi_j$. Donc:

$$\frac{\partial(g\circ f)}{\partial x_i}(M) = d_{f(M)}g\Big(\sum_j \frac{\partial f^j}{\partial x_i}(M)\, \varphi_j\Big) = \sum_j \frac{\partial f^j}{\partial x_i}(M)\ d_{f(M)}g(\varphi_j) = \sum_j \frac{\partial f^j}{\partial x_i}(M)\ \frac{\partial g}{\partial y_j}(f(M))$$

Soit $$\frac{\partial(g\circ f)}{\partial x_i}(M) = \sum_j \frac{\partial f^j}{\partial x_i}(M)\ \frac{\partial g}{\partial y_j}(f(M))$$

Telle est l'écriture développée en termes de dérivées partielles, de la formule de composition des différentielles. Notons que les $\frac{\partial f^j}{\partial x_i}(M)$ sont des scalaires, et les $\frac{\partial g}{\partial y_j}(f(M))$ des vecteurs de G, et que la somme a p (=dim F) termes.

Et puisque $M \to f(M)$, $M \to \dfrac{\partial f^j}{\partial x_i}(M)$ et $P \to \dfrac{\partial g}{\partial y_j}(P)$ sont continues, il est maintenant clair que les dérivées partielles de $g \circ f$ sont des fonctions continues de M; ce que nous cherchions à démontrer.

Remarque: La formule $\dfrac{\partial (g \circ f)}{\partial x_i}(M) = \displaystyle\sum_j \dfrac{\partial f^j}{\partial x_i}(M)\ \dfrac{\partial g}{\partial y_j}(f(M))$ *est la forme développée du produit de matrices correspondant à* $d_{f(M)}g \circ d_M f$. *C'est aussi la forme ancienne de ce produit; celle que l'on utilisait avant que l'usage des matrices soit banalisé. Elle est encore fréquemment utilisée, dans les cours de physique par exemple. Il faut se souvenir que, dès que les situations se compliquent, la notation matricielle se révèle plus commode.*

§ 4 : Les propriétés algébriques des applications de classe C^1.

Somme et produit par un scalaire

Soit $U \subset E$ *et* $f: U \to F$, $g: U \to F$ *et* $\lambda: U \to \mathbb{R}$ *trois applications de classe* C^1, *alors l'application* $f+\lambda g: U \to F$ *est de classe* C^1 *et :*

$$d_M(f+\lambda g) = d_M f + d_M\lambda \times g(M) + \lambda(M) \times d_M g$$

(C'est à dire, pour tout h dans E : $d_M(f+\lambda g)(h) = d_M f(h) + d_M\lambda(h) g(M) + \lambda(M) d_M g(h)$*).*

Démonstration: Pour tout vecteur X de E, nous avons

$$(f+\lambda g)(M+tX) = f(M+tX) + \lambda(M+tX)\ g(M+tX)$$

Puisque f, g et λ sont de classe C^1, on peut dériver par rapport à t pout t=0; ce qui donne:

$$(\forall X) \qquad D_X(f+\lambda g)(M) = D_X f(M) + D_X\lambda(M) \times g(M) + \lambda(M) \times D_X g(M)$$

En appliquant cette formule au cas où X est l'un des vecteurs de base (ie: de la base choisie dans E), on voit que si f, g et λ ont des dérivées partielles continues, alors $f+\lambda g$ a aussi des dérivées partielles continues, autrement dit est de classe C^1. Mais cette même formule s'écrit aussi:

$$d_M(f+\lambda g)(X) = d_M f(X) + d_M\lambda(X) \times g(M) + \lambda(M) \times d_M g(X)$$

Ce qui est, aux notations près, la formule annoncée pour la différentielle.

Produits scalaires

Supposons que F soit muni d'un produit scalaire; alors si $f: U \to F$ *et* $g: U \to F$ *sont de classe* C^1, *l'application* $\langle f,g\rangle$ *de U dans* $\mathbb{R}$ *qui à u associe* $\langle f(u),g(u)\rangle$, *est de clase* C^1. *Et sa différentielle en M est l'application*

$$d_M\langle f,g\rangle : h \to \langle d_M f(h), g(a)\rangle + \langle f(a), d_M g(h)\rangle$$

Démonstration: Pour tout vecteur X, nous avons $\langle f,g\rangle(M+tX) = \langle f(M+tX), g(M+tX)\rangle$. En dérivant par rapport à t pour t = 0 (ce qui est possible puisque $t \to f(M+tX)$ et $t \to g(M+tX)$ sont de classe C^1), nous aurons:

$$D_X\langle f,g\rangle(M) = \langle D_X f(M), g(M)\rangle + \langle f(M), D_X g(M)\rangle$$

En prenant pour X les vecteurs d'une base de E, cette formule montre que $\langle f,g\rangle$ est de classe C^1, et elle s'écrit:

$$d_M\langle f,g\rangle : X \to \langle d_M f(X), g(M)\rangle + \langle f(M), d_M g(X)\rangle$$

Ce qui est, aux notations près, le résultat annoncé.

Cas où l'espace d'arrivée F est muni d'un produit

Supposons que F soit muni d'un produit (on notera $\alpha.\beta$ le produit de α et β) et d'une norme tels que:

$$(\exists k)(\forall\alpha)(\forall\beta)\ N_F(\alpha.\beta) \le k\, N_F(\alpha)\, N_F(\beta).$$

Si $f: U \to F$ et $g: U \to F$ (où $U \subset E$) sont de classe C^1, alors la fonction produit $f.g$ (qui à u associe $f(u).g(u)$) est de classe C^1. Sa différentielle est donnée par :

$$d_M f.g(h) = d_M f(h).g(M) + f(M).d_M g(h).$$

En particulier: Si f et g sont des fonctions de classe C^1 à valeurs complexes, la fonction produit est de classe C^1. Si f et g sont des fonctions de classe C^1 à valeurs dans l'espace E_3 des vecteurs de la géométrie, la fonction $f \wedge g$ (qui à u associe le produit vectoriel $f(u)\wedge g(u)$) est de classe C^1. Si f et g sont des fonctions de classe C^1 à valeurs matricielles, alors la fonction produit fg est aussi de classe C^1.

La démonstration est analogue aux deux précédentes, c'est à dire que l'on se ramène au cas des fonctions sur $\mathbb{R}$ en calculant les dérivées partielles.

§ 5 : Le théorème des accroissements finis

Pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ nous connaissons

a) la formule des accroissements finis

b) l'inégalité des accroissements finis

Ces deux énoncés permettent d'évaluer le taux d'accroissement de la fonction à partir des valeurs de la dérivée. Dans le cas des fonctions sur un espace de dimension plus grande que 1, on ne peut espérer avoir un résultat aussi précis que la formule des accroissements finis; on va seulement démontrer l'inégalité suivante, connue sous le nom de théorème des accroissements finis.

Soit f une application différentiable d'un ouvert U de E dans F. On suppose que N_F est une norme euclidienne (c'est à dire une norme associée à un produit scalaire) sur F, tandis que N_E est une norme quelconque sur E. On suppose que K est un nombre tel que ($\forall u \in U$) ($\forall e \in E$) :

$$N_F(d_u f(e)) \le K\, N_E(e)$$

(Noter que le lemme du §2 nous dit que, pour tout u, il existe un tel K; notre hypothèse signifie seulement qu'il en existe un qui est valable pour tous les u)

Alors pour tout couple (A,B) de points de U tels que le segment [AB] soit dans U, on a:

$$N_F[f(B)-f(A)] \le K\, N_E(B-A)$$

Démonstration: Notons <.,.> le produit scalaire qui définit la norme N_F. Pour tout t dans [0,1], posons $\varphi(t) = f(A + t(B-A))$ et $\psi(t) = \langle f(B)-f(A),\varphi(t)\rangle$. Nous savons (cf: §4) que ψ est différentiable (c'est à dire dérivable puisque c'est une fonction réelle d'une variable réelle). Et on a, pour $h \in \mathbb{R}$:

$$d_t\Psi(h) = \langle f(B)-f(A), d_{A+t(B-A)}f(h(B-A))\rangle$$

C'est à dire $\quad \dfrac{d\psi}{dt}(t) = \langle f(B)-f(A), d_{A+t(B-A)}f(B-A)\rangle$

D'où: $\quad \left|\dfrac{d\psi}{dt}(t)\right| \le N_F[f(B)-f(A)]\ N_F[d_{A+t(B-A)}f(B-A)]$

Ou encore: $\quad \left|\dfrac{d\psi}{dt}(t)\right| \le N_F[f(B)-f(A)]\ K\, N_E(B-A)$

La fonction ψ étant continue sur [0,1] et dérivable sur]0,1[, nous pouvons lui appliquer l'inégalité des accroissements finis pour les fonction de $\mathbb{R}$ dans $\mathbb{R}$. Donc:

$$|\psi(1) - \psi(0)| \le (1-0) \times \left[N_F[f(B)-f(A)]\ K\, N_E(B-A)\right]$$

Soit : $|\langle f(B)-f(A),f(B)\rangle - \langle f(B)-f(A),f(A)\rangle| \leq N_F[f(B)-f(A)]\ K\, N_E(B-A)$

Ou encore: $\{N_F[f(B)-f(A)]\}^2 \leq N_F[f(B)-f(A)]\ K\, N_E(B-A)$

Donc $N_F[f(B)-f(A)] \leq K\, N_E(B-A)$

Ce que l'on cherchait à démontrer.

Cas des applications de classe C^1

Si f est de classe C^1 sur U, l'application df: $U \rightarrow \mathcal{L}(E,F)$ est continue; autrement dit, dans des bases quelconques de E et F, les coefficients de la matrice (m_{ij}) de df sont des fonctions continues. Donc (chap 10 §7) pour tout compact A contenu dans U, il existe un nombre K_A tel que $(\forall u \in A)$ $(\forall i,j)$ on ait $|m_{ij}(u)| \leq K_A$. D'après le lemme du §2 (ou plutôt sa démonstration), pour tout u dans A nous aurons alors:

$$N_F(d_u f(h)) \leq \lambda\, K_A\, N_E(h)$$

où λ est un nombre qui dépend des normes N_E et N_F choisies. *Ainsi pour une fonction de classe C^1, les hypothèses du théorème des acroissements finis ne sont peut être pas vérifiées sur U tout entier, mais elles le sont sur tout compact contenu dans U. En particulier pour tout point u de U, il existe une boule centrée en u, et sur laquelle on peut appliquer le théorème des accroissements finis (pour une constante K bien choisie et qui dépend de cette boule).*

Exercice k: Montrer que l'application $z \rightarrow e^z$ $(\mathbb{C} \rightarrow \mathbb{C})$ est de classe C^1. Puis trouver un nombre K_R tel que, quel que soient z et z' de module au plus égal à R, on ait $|e^z - e^{z'}| \leq K_R\, |z - z'|$.

§ 6 : Applications de classe $C^2, C^3, \ldots, C^\infty$.

Soit f une application de classe C^1 d'un ouvert U de E dans F. La correspondance qui à tout u de U associe la différentielle de f en u, est une application de U dans l'espace vectoriel $\mathcal{L}(E,F)$. Si cette application df: $U \rightarrow \mathcal{L}(E,F)$ est de classe C^1, nous dirons que f est de classe C^2 (remarquons que ceci a un sens puisque $\mathcal{L}(E,F)$ est de dimension finie). Des bases de E et F étant choisies, les fonctions coordonnées de df (i.e. les coefficients de la matrice) sont les fonctions coordonnées des dérivées partielles (premières) de f. Pour que f soit de classe C^2 (c'est à dire que df soit de classe C^1), il suffit que ces dérivées partielles premières aient elles mêmes des dérivées partielles continues, c'est à dire que f ait des dérivées partielles secondes continues.

On définit de même les applications de classe C^3: ce sont les applications f telles que df soit de classe C^2. Et ainsi de suite. Lorsque f est de classe C^r quel que soit r dans $\mathbb{N}$, on dit que f est de classe C^∞.

Dans la plupart des exercices, nous utiliserons des fonctions définies par des formules; c'est à dire des fonctions obtenues à partir des fonctions classiques (exp, log, sin,cos,...) par composition, et par des sommes, des produits et des quotients. Ces fonctions sont toutes de classe C^∞,avec les restrictions suivantes (déjà vues au § 2)

Les fonctions $\sqrt{t}$, $\sqrt[3]{t}$, t^k $(k \notin \mathbb{N})$ ne sont pas dérivables en 0, donc, lorsque f est de classe C^∞, $\sqrt{f}$, $\sqrt[3]{f}$, f^k $(k \notin \mathbb{N})$ n'ont (en général) pas de dérivées partielles (et donc ne sont pas C^∞) aux points u tels que $f(u) = 0$.

Les fonctions Arcsin et Arccos ne sont pas dérivables en 1 et -1 (de même Argch n'est pas dérivable en 1), donc Arcsin (f), Arccos (f) n'est (en général) pas C^∞ aux points u tels que $f(u) = \pm 1$ (de même Argch (f) n'est (en général) pas C^∞ aux points u tels que $f(u) = 1$).

Le lemme de symétrie de Schwarz

Rappelons que pour toute fonction f de classe C^2, et tout couple (i,j), on a $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$. *Donc les dérivées d'ordre r d'une fonction f de classe C^r ne dépendent pas de l'ordre dans lequel s'effectuent les dérivations.*

Complément : Champs de gradient.

Gradient d'une fonction.

Considérons une fonction différentiable f définie sur un ouvert U d'un espace E muni d'un produit scalaire (noté <.,.>) et à valeurs dans $\mathbb{R}$. Sa différentielle en un point M est une forme linéaire sur E. *Or il existe un unique vecteur V de E, tel que, quel que soit h:*

$$d_M f(h) = \langle V,h \rangle$$

Pour calculer V, il suffit de calculer ses coordonnées dans une base orthonormée $\{e_1,\ldots,e_n\}$; elles sont données par $V_i = \langle V,e_i \rangle = d_M f(e_i) = \frac{\partial f}{\partial x_i}(M)$. *Ce vecteur est appelé le gradient de f.*

Dans toute base orthonormée les coordonnées du gradient de f sont donc les dérivées partielles de f au point considéré. On prendra garde que ceci devient faux lorsque la base n'est plus orthonormée.

Nous avons ainsi associé à tout M dans U un vecteur $\mathrm{grad}_M f$ de E, autrement dit nous avons défini un champ de vecteurs sur U. Si f est de classe C^1, ce champ est continu.

Considérons maintenant une courbe de classe C^1, c'est à dire une application φ de classe C^1 de [a,b] dans E; et calculons l'intégrale du gradient de f le long de cette courbe.

$$\int_a^b \langle \mathrm{grad}_{\varphi(t)} f, \frac{d\varphi}{dt}(t) \rangle\, dt = \int_a^b d_{\varphi(t)} f(\frac{d\varphi}{dt}(t))\, dt = \int_a^b \frac{d}{dt}[f(\varphi(t))]\, dt = f(\varphi(b)) - f(\varphi(a))$$

Autrement dit cette intégrale ne dépend pas de la courbe mais uniquement de ses extrémités.

Ici encore on est en présence d'anciennes notations, datant d'une époque où les phénomènes linéaires n'étaient pas bien compris. La construction (au moyen du produit scalaire) de $\mathrm{grad}_M f$ à partir de $d_M f$ n'apporte rien. C'est une simple traduction permettant d'employer le langage des vecteurs plutôt que celui des formes linéaires.

Caractérisation des champs de gradient.

Soit X un champ de vecteurs de classe C^1 sur E. Cherchons des conditions pour que X soit le champ de gradient d'une fonction. Pour cela choisissons une base orthonormée $\{e_i\}$ de E, et notons $(X_i)_{i=1,\ldots,n}$ les coordonnées de X dans cette base. Si X est le champ de gradient d'une fonction f, alors $(\forall i)\ X_i = \frac{\partial f}{\partial x_i}$; par conséquent (d'après la formule de symétrie de Schwarz)

$$\frac{\partial X_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial X_j}{\partial x_i}$$

Ainsi une condition nécessaire pour que X soit un champ de gradient, est que ($\forall i \neq j$) $\frac{\partial X_i}{\partial x_j} - \frac{\partial X_j}{\partial x_i} = 0$.

Cette condition est localement suffisante; ce qui signifie que, si elle est vérifiée, pour tout point M il existe un ouvert U contenant M, et une fonction f: U $\to$ $\mathbb{R}$ tels que (en tout point P de U) on ait $X(P) = \mathrm{grad}_P f$. Nous l'admettrons.

<u>Exemple</u>: Soit le champ X défini sur le complémentaire de l'origine dans $\mathbb{R}^2$, par $X = (\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2})$.

On vérifie aisément que $\frac{\partial}{\partial y}(\frac{-y}{x^2+y^2}) = \frac{\partial}{\partial x}(\frac{x}{x^2+y^2}) = \frac{y^2-x^2}{(x^2+y^2)^2}$. Donc notre condition nécessaire est vérifiée.

Sur l'ouvert $U_1 = \{(x,y)$ tels que $x>0\}$, X est le gradient de la fonction $\mathrm{Arctg}\,(\frac{y}{x})$.

Sur l'ouvert $U_2 = \{(x,y)$ tels que $y>0\}$, X est le gradient de la fonction $-\mathrm{Arctg}\,(\frac{x}{y})$

Sur l'ouvert $U_3 = \{(x,y)$ tels que $x<0\}$, X est le gradient de la fonction $\mathrm{Arctg}\,(\frac{y}{x})$

Sur l'ouvert $U_4 = \{(x,y)$ tels que $y<0\}$, X est le gradient de la fonction $-\mathrm{Arctg}\,(\frac{x}{y})$.

Mais il n'existe aucune fonction définie sur le complémentaire de l'origine (c'est à dire définie sur l'ensemble de définition de X tout entier) et dont le gradient est f. En fait une fonction dont le gradient est X est nécessairement de la forme M $\rightarrow$ Constante + Mesure de l'angle (Ox,OM); elle ne peut donc pas être définie dans le complémentaire de l'origine, puisque, lorsqu'on fait le tour de l'origine une mesure de l'angle (Ox,OM) varie de 2π.

<u>Cas particuliers</u>: En dimension 2, notre condition se réduit à $\frac{\partial X_1}{\partial x_2} - \frac{\partial X_2}{\partial x_1} = 0$.

En dimension 3, nous devrons écrire $\frac{\partial X_1}{\partial x_2} - \frac{\partial X_2}{\partial x_1} = 0$, $\frac{\partial X_2}{\partial x_3} - \frac{\partial X_3}{\partial x_2} = 0$ et $\frac{\partial X_3}{\partial x_1} - \frac{\partial X_1}{\partial x_3} = 0$.

Nous sommes donc amenés à écrire le champ de vecteurs

$$Y = (\frac{\partial X_2}{\partial x_3} - \frac{\partial X_3}{\partial x_2})\, e_1 + (\frac{\partial X_3}{\partial x_1} - \frac{\partial X_1}{\partial x_3})\, e_2 + (\frac{\partial X_1}{\partial x_2} - \frac{\partial X_2}{\partial x_1})\, e_3$$

La nullité de Y est une condition nécessaire et localement suffisante pour que X soit un champ de gradient.

Mais ce champ Y dépend de la base $\{e_i\}$. Rappelons d'abord que cette base doit être orthonormée pour que notre calcul ait un sens. Mais notons que si nous remplaçons $\{e_i\}$ par $\{-e_i\}$ (ce qui change l'orientation de la base), les X_i sont changés en leurs opposés; les $\frac{\partial X_i}{\partial x_j}$ ne sont pas changés, donc Y est changé en son opposé. On démontre que (si l'on oriente E) et si l'on fait ce calcul dans deux bases orthonormées directes, on trouve le même champ Y. Ce champ Y (calculé en base orthonormée directe) est appelé le rotationel du champ X.

Exercices sur le chapitre 13

* **Exercice 1:** On considère l'application f de $\mathbb{R}^2$ dans $\mathbb{R}^2$ qui à (x,y) associe (cos (x+y), sin (x-y)). Elle est de classe C^1. Ecrire la matrice de sa différentielle au point M(x,y). Quels sont les points M tels que d_Mf soit inversible ?

* **Exercice 2:** On considère l'application f de $\mathbb{R}^3$ dans $\mathbb{R}^2$ qui à (x,y,z) associe $(x^2+y^2+z^2, x-y+2z, xy+yz+zx)$. Elle est de classe C^1. Ecrire la matrice de sa différentielle au point M(x,y,z).

* **Exercice 3:** Soit f l'application de $\mathbb{C}^*$ dans $\mathbb{C}$ qui à z associe $z + \frac{1}{z}$. Montrer que, considérée comme une application de $(\mathbb{R}^2)^*$ dans $\mathbb{R}^2$, elle est différentiable, et calculer d_zf.

* **Exercice 4:** Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = \frac{xy^3}{x^2+y^2}$ pour $(x,y) \neq (0,0)$, et $f(0,0) = 0$. Calculer $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$. Montrer que $(\exists K)$ tel que $(\forall(x,y))$ $|\frac{\partial f}{\partial x}(x,y)| \leq K\sqrt{x^2+y^2}$ et $|\frac{\partial f}{\partial y}(x,y)| \leq K\sqrt{x^2+y^2}$. Montrer que f est de classe C^1.

** **Exercice 5:** Soit $f: \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = \frac{x^4 - y^2 - x^2}{x^4+y^2+x^2}$ pour $(x,y) \neq (0,0)$, et $f(0,0) = 0$. Est elle continue à l'origine? Est elle de classe C^1?

Mêmes questions pour g définie par $g(x,y) = xf(x,y)$.

* **Exercice 6:** On considére $\mathbb{R}^3$ muni de sa norme euclidienne naturelle. On définit $f: (\mathbb{R}^3)^* \to \mathbb{R}^3$ par $f(M) = \frac{M}{\|M\|}$. Montrer que f est de classe C^1. Quelle est sa différentielle au point M(x,y,z)?

** **Exercice 7:** Soit E l'espace des matrices réelles 3×3. Soit F l'application qui à toute matrice associe son déterminant. Montrer qu'elle est de classe C^1; calculer sa différentielle en M_0, où

$$M_0 = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

(On pourra utiliser les développements du déterminant par rapport à ses lignes ou à ses colonnes).

*** **Exercice 8:** Soit E l'espace des matrices réelles 3×3. Soit Φ l'application de E dans E, qui à toute matrice associe son inverse. Montrer qu'elle est de classe C^1 (on pourra utiliser les formules qui donnent l'inverse de M au moyen des déterminants, et tenir compte de l'exercice 7). Démontrer les formules:

$$(M+H)^{-1} = M^{-1} - (M+H)^{-1}HM^{-1}$$
$$(M+H)^{-1} = M^{-1} - M^{-1}HM^{-1} + (M+H)^{-1}HM^{-1}HM^{-1}$$

On rappelle qu'il existe sur E une norme N telle que ($\forall$A et B) $N(AB) \leq 3N(A)N(B)$. En utilisant le fait que Φ est de classe C^1 (et en utilisant le théorème des accroisssements finis), montrer qu'il existe ε (>0) tel que

$$N(H) \leq \varepsilon \text{ implique } N((M+H)^{-1} - M^{-1}) \leq \frac{1}{2}N(M^{-1}), \text{ donc } N((M+H)^{-1}) \leq \frac{3}{2}N(M^{-1})$$

En déduire une expression simple de la différentielle de Φ.

* **Exercice 9**: Soit f: $\mathbb{R}^3 \to \mathbb{R}^2$ de classe C^1; on considére l'application g: $\mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$g(\omega,\varphi) = f(\cos\omega + \sin\varphi, \sin\omega + \cos\varphi, e^{\omega-\varphi}) .$$

Alors g est aussi de classe C^1. Soit M(1,1,1), on suppose que d_Mf a pour matrice dans les bases naturelles de $\mathbb{R}^3$ et $\mathbb{R}^2$ $A = \begin{pmatrix} 1 & 3 & 4 \\ 2 & -1 & 3 \end{pmatrix}$ Déterminer la matrice de d_Pg, où $P=(\pi/2,\pi/2)$.

* **Exercice 10**: Soit E l'espace des matrices n×n, on considère l'application de E dans lui même qui à M associe M^2. Montrer que cette application est différentiable, et calculer sa différentielle en A (Attention: si A n'est pas une matrice scalaire, cette différentielle n'est pas l'application $H \to 2MH$).

Calculer de même la différentielle en A de l'application qui à toute M associe $M^5 - 3M^4 + M$.

* **Exercice 11**: a) Soit $f:\mathbb{C} \to \mathbb{C}$ définie par $f(z) = z^3 + 2z$. Montrer que z est différentiable, et que sa différentielle au point z est l'application $h \to (3z^2 + 2)h$.

b) On choisit pour norme sur $\mathbb{C}$ le module. Trouver un nombre M tel que ($\forall$z de module au plus égal à 2) ($\forall h\in \mathbb{C}$) $|3z^2 + 2| \leq M$.

En déduire un nombre k tel que ($\forall$z et z' de module au plus égal à 2) $|f(z)-f(z')| \leq k|z-z'|$.

* **Exercice 12**: Soit f une fonction de classe C^2 de $\mathbb{R}^3$ dans $\mathbb{R}$. Calculer, en fonction des dérivées partielles premières et seconde de f, la dérivée première et la dérivée seconde de $g(t) = f(\cos t, \sin t, t)$.

** **Exercice 13:** La fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ qui est définie par $f(x,y) = \dfrac{x^2y^3}{x^4+y^4}$ pour $(x,y) \neq (0,0)$, et par $f(0,0) = 0$, est elle de classe C^1 ?

** **Exercice 14**: Soit f: $\mathbb{R}^2 \to \mathbb{R}$ de classe C^1. Soit g: $\mathbb{R}^2 \to \mathbb{R}^2$ définie par $g(x,y) = (2x+y,x-3y)$

a) Calculer (en fonction des dérivées partielles de f) les dérivées partielles d'ordre 1 et d'ordre 2 de f∘g au point (x_0,y_0).

b) Quelles sont les fonctions h: $\mathbb{R}^2 \to \mathbb{R}$ telles que $2\dfrac{\partial^2 h}{\partial x^2} - 5\dfrac{\partial^2 h}{\partial x \partial y} - 3\dfrac{\partial^2 h}{\partial y^2} = 0$?

** **Exercice 15:** Soit f: $\mathbb{R}^2 \to \mathbb{R}$ de classe C^1. Soit g: $\mathbb{R}^2 \to \mathbb{R}^2$ définie par $g(x,y) = (x+y,x-y)$

a) Calculer (en fonction des dérivées partielles de f) les dérivées partielles d'ordre 1 et d'ordre 2 de f∘g au point (x_0,y_0).

b) Quelles sont les fonctions h: $\mathbb{R}^2 \to \mathbb{R}$ telles que $\dfrac{\partial^2 h}{\partial X^2} - 2\dfrac{\partial^2 h}{\partial X \partial Y} + \dfrac{\partial^2 h}{\partial Y^2} = 0$?

* **Exercice 16:** On considère l'application G de $\mathbb{C}$ dans lui même qui à z associe $z^3 + z + 1$. Montrer qu'elle est de classe C^1. Montrer que ($\forall z_0$) sa différentielle en z_0 est une similitude directe de $\mathbb{C}$.

** **Exercice 17**: Soit U le complémentaire de l'origine dans $\mathbb{C}$. Montrer que U est un ouvert dans $\mathbb{C}$. Soit G: $U \to U$ définie par $G(z) = \dfrac{1}{\bar{z}}$. Montrer que G est de classe C^1. Montrer que ($\forall z_0$) sa différentielle en z_0 est une similitude inverse.

** **Exercice 18**: Pour tout nombre complexe z différent de 2, on pose $F(z) = \frac{z+2}{z-2}$.

a) Montrer que l'on a ainsi défini une bijection F de l'ouvert $U = \complement_{\mathbb{C}}\{2\}$ sur un ouvert V de $\mathbb{C}$ que l'on précisera.

b) Montrer que F^{-1} et F sont de classe C^1. Calculer la différentielle de F^{-1} au point $F(x_0,y_0)$. Montrer que c'est une similitude directe.

** **Exercice 19:** Soit f une apllication de classe C^2 de $\mathbb{R}^2$ dans $\mathbb{R}$. On suppose que

$$\frac{\partial f}{\partial X}(0,0) = 1 \qquad \frac{\partial f}{\partial Y}(0,0) = 2 \qquad \frac{\partial^2 f}{\partial X^2}(0,0) = 1 \qquad \frac{\partial^2 f}{\partial X \partial Y}(0,0) = 2 \qquad \frac{\partial^2 f}{\partial Y^2}(0,0) = -1$$

Soit g: $\mathbb{R}^2 \to \mathbb{R}^2$ définie par $g(x,y) = (\text{sh } x \cos y, \text{ch } x \sin y)$.

a) Calculer les dérivées partielles premières de $f \circ g$ en (0,0).

b) Calculer les dérivées partielles secondes de $f \circ g$ en (0,0).

* **Exercice 20:** Soit f: $\mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x,y) = (\cos x \sin y, \sin x \cos y)$.

a) Calculer la différentielle de f au point (x,y).

b) Démontrer que ($\forall$ A et B dans $\mathbb{R}^2$) $\|f(A) - f(B)\| \leq 2\|A - B\|$. Où $\|X\|$ désigne la norme euclidienne naturelle du vecteur X.

*** **Exercice 21:** Soit f: $\mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x,y) = (x^2y + y^2 + x, x^3y^2 + y^4 + y)$. On note $\|A\|$ la norme euclidienne naturelle du vecteur A.

a) On pose $f = \text{Id} + \varphi$. Calculer $d_{(x,y)}\varphi$.

b) Montrer que $\|A\| \leq 1/10$ implique que $(\forall X)$ $\|d_A\varphi(X)\| \leq \frac{1}{2}\|X\|$.

c) En déduire que pour $\|A\| \leq 1/10$ et $\|B\| \leq 1/10$, on a $\|\varphi(A) - \varphi(B)\| \leq \frac{1}{2}\|A - B\|$

d) Montrer que la restriction de f à la boule de centre l'origine et de rayon 1/10, est injective.

Limites sous le signe intégrale

§ 1 : Intégrale de la limite d'une suite de fonctions continues sur [a,b].

Considérons une suite de fonctions $f_n(t)$ continues sur $[a,b]$. Supposons que $(\forall t)$ la suite $(f_n(t))_{n\in\mathbb{N}}$ converge vers un certain nombre $f(t)$; c'est à dire que la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f. Supposons que f soit une fonction intégrable, pouvons nous affirmer que $\int_a^b f(t)\,dt$ est la limite de la suite $\left(\int_a^b f_n(t)\,dt\right)_{n\in\mathbb{N}}$?

Sous les hypothèses que nous avons faites, il se peut que ce soit faux.

Exemple: Posons $f_n(t) = \dfrac{n^2t}{1+n^3t^3}$ sur $[0,1]$. Dans ce cas f est la fonction nulle (autrement dit, pour tout t dans $[0,1]$, $\lim\limits_{n\to\infty} f_n(t) = 0$). Pourtant

$$\int_0^1 f_n(t)\,dt = \int_0^1 \frac{n^2t}{1+n^3t^3}\,dt = \int_0^n \frac{u}{1+u^3}\,du \qquad (u = nt)$$

Donc la suite $\left(\int_0^1 f_n(t)\,dt\right)_{n\in\mathbb{N}}$ est croissante et tend vers $\int_0^\infty \frac{u}{1+u^3}\,du$ $\left(\neq \int_0^1 f(t)\,dt = 0\right)$.

Si la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f pour la norme N_∞, $\lim\limits_{n\to\infty} \int_a^b f_n(t)\,dt = \int_a^b f(t)\,dt$.

Pour s'en convaincre, on écrit les inégalités :

$$\left| \int_a^b f_n(t)\,dt - \int_a^b f(t)\,dt \right| \leq \int_a^b |f_n(t) - f(t)|\,dt \leq \int_a^b N_\infty(f_n-f)\,dt = (b-a)\,N_\infty(f_n-f)$$

Exercice a: a) On suppose que la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f pour la norme N_1; montrer que la suite $\left(\int_a^b f_n(t)\,dt\right)_{n\in\mathbb{N}}$ converge vers $\int_a^b f(t)\,dt$

b) On suppose que la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f pour la norme N_2; en considérant $\int_a^b f(t) - f_n(t)\,dt$ comme le produit scalaire de $f - f_n$ et de la constante 1, et en appliquant l'inégalité de Schwarz, montrer que la suite $\left(\int_a^b f_n(t)\,dt\right)_{n\in\mathbb{N}}$ converge vers $\int_a^b f(t)\,dt$

§ 2 : Intégrale de la limite d'une suite de fonctions continues sur]a,b[.

Nous travaillons ici sur un intervalle ouvert]a,b[; les intégrales considérées sont donc des intégrales au sens du chapitre 9, c'est à dire des intégrales absolument convergentes en a et en b.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions continues et intégables sur]a,b[. On suppose qu'il existe une fonction continue f telle que ($\forall t \in]a,b[$) la suite $(f_n(t))_{n\in\mathbb{N}}$ converge vers f(t). La fonction f est elle intégrable sur]a,b[? Son intégrale est elle la limite de la suite ($\int_a^b f_n(t)\,dt)_{n\in\mathbb{N}}$?

Sous ces hypothèses il est possible que les réponses à ces deux questions soient négatives; comme en témoignent les exemples suivants :

Exemple 1: Soit $]a,b[=]-\infty,+\infty[$, et soit $f_n(t) = \frac{1}{1+(n-t)^2}$. Alors ($\forall t$) la suite $(f_n(t))_{n\in\mathbb{N}}$ converge vers 0.

Ici la fonction f est donc intégrable, d'intégrale nulle; pourtant $\int_{-\infty}^{+\infty} \frac{1}{1+(n-t)^2}\,dt = \int_{-\infty}^{+\infty} \frac{1}{1+u^2}\,du = \pi$ (où l'on a posé t-n = u), donc la suite ($\int_a^b f_n(t)\,dt)_{n\in\mathbb{N}}$ converge mais n'a pas pour limite l'intégrale de f.

Exemple 2: Soit]a,b[=]0,1[, et soit $f_n(t) = \frac{n+n^2t^3}{1+n^2t^2}$. Alors ($\forall t\in]0,1[$) la suite $(f_n(t))_{n\in\mathbb{N}}$ converge vers t. La fonction f est donc la fonction $t \to t$ qui est intégrable, d'intégrale 1/2; pourtant

$$\int_0^1 \frac{n+n^2t^3}{1+n^2t^2}\,dt = \int_0^n \frac{n+\frac{u^3}{n}}{1+u^2}\,\frac{du}{n} = (\text{Arctg } n) + \frac{1}{2} - \frac{1}{2n^2}\text{Log}\,(1+n^2) \qquad (u=nt)$$

Donc la suite ($\int_a^b f_n(t)\,dt)_{n\in\mathbb{N}}$ converge vers $\frac{1}{2}+\frac{\pi}{2}$.

Exemple 3 : Soit]a,b[=]0,1[, et soit $f_n(t) = (1-2t)\,\frac{1-(1-2t)^{2n}}{t\,(1-t)}$. Alors ($\forall t$) la suite $(f_n(t))_{n\in\mathbb{N}}$ converge vers $\frac{1-2t}{t\,(1-t)}$, qui n'est pas intégrable sur]0,1[(car son intégrale ne converge ni en 0 ni en 1). Tandis que toutes les fonctions f_n sont intégrables (car elles ont une limite finie en 0 et en 1); et de plus ($\forall n$) $f_n(t) = -\,f_n(1-t)$, donc $\int_0^1 f_n(t)\,dt = 0$; donc la suite ($\int_0^1 f_n(t)\,dt)_{n\in\mathbb{N}}$ converge vers 0.

Le théorème de la convergence majorée. Pour pouvoir affirmer que la fonction limite f est intégrable, et que la suite ($\int_a^b f_n(t)\,dt)_{n\in\mathbb{N}}$ converge vers $\int_a^b f(t)\,dt$ nous devons donc faire des hypothèses supplémentaires. Nous allons démontrer le résultat suivant connu sous le nom de théorème de la convergence majorée. Les fonctions f_n étant toujours continues et intégrables sur]a,b[)

Si il existe une fonction continue positive g (absolument) intégrable sur]a,b[telle que ($\forall n$) et ($\forall t$) $|f_n(t)| \leq g(t)$.

Et si pour tout intervalle compact $I = [\alpha,\beta]$ contenu dans]a,b[, la suite $(f_n|I)_{n\in\mathbb{N}}$ converge uniformément vers f|I (c'est à dire si la suite ($\sup_{\alpha\leq t\leq\beta}(f_n(t)-f(t)))_{n\in\mathbb{N}}$ converge vers 0)

Alors la fonction f est intégrable sur]a,b[et la suite ($\int_a^b f_n(t)\,dt)_{n\in\mathbb{N}}$ converge vers $\int_a^b f(t)\,dt$.

Démonstration: Puisque sur tout intervalle $[\alpha,\beta]$ les f_n convergent uniformément vers f, on peut affirmer que f est continue. De plus $(\forall t)$ la suite $(f_n(t))_{n\in\mathbb{N}}$ converge vers f(t). Donc l'inégalité $(\forall t)$ $|f_n(t)| \le g(t)$, nous permet de conclure que $(\forall t)$ $|f(t)| \le g(t)$. Et ainsi la fonction $|f|$ est majorée par une fonction positive intégrable, donc f est absolument intégrable.

Pour tout intervalle $[\alpha,\beta]$ $(a<\alpha<\beta<b)$, nous pouvons écrire:

$$\left|\int_a^b f(t)\,dt - \int_a^b f_n(t)\,dt\right| \le \left|\int_a^\alpha f(t)\,dt - \int_a^\alpha f_n(t)\,dt\right| + \left|\int_\alpha^\beta f(t)\,dt - \int_\alpha^\beta f_n(t)\,dt\right| + \left|\int_\beta^b f(t)\,dt - \int_\beta^b f_n(t)\,dt\right|$$

$$\left|\int_a^b f(t)\,dt - \int_a^b f_n(t)\,dt\right| \le \left|\int_a^\alpha f(t)\,dt\right| + \left|\int_a^\alpha f_n(t)\,dt\right| + \left|\int_\alpha^\beta f(t)\,dt - \int_\alpha^\beta f_n(t)\,dt\right| + \left|\int_\beta^b f(t)\,dt\right| + \left|\int_\beta^b f_n(t)\,dt\right|$$

$$\left|\int_a^b f(t)\,dt - \int_a^b f_n(t)\,dt\right| \le \int_a^\alpha g(t)\,dt + \int_a^\alpha g(t)\,dt + \left|\int_\alpha^\beta f(t)\,dt - \int_\alpha^\beta f_n(t)\,dt\right| + \int_\beta^b g(t)\,dt + \int_\beta^b g(t)\,dt$$

$$\left|\int_a^b f(t)\,dt - \int_a^b f_n(t)\,dt\right| \le 2\int_a^\alpha g(t)\,dt + 2\int_\beta^b g(t)\,dt + \left|\int_\alpha^\beta f(t)\,dt - \int_\alpha^\beta f_n(t)\,dt\right|$$

Nous devons montrer que, quel que soit ε, pour n suffisamment grand, $\left|\int_a^b f(t)\,dt - \int_a^b f_n(t)\,dt\right|$ est inférieur à ε. Pour cela choisissons <u>d'abord</u> α et β tels que $\int_a^\alpha g(t)\,dt \le \varepsilon/8$ et $\int_\beta^b g(t)\,dt \le \varepsilon/8$.

L'intervalle $I = [\alpha,\beta]$ étant ainsi fixé, nous pouvons <u>ensuite</u> trouver un entier N tel que $n\ge N$ implique $N_\infty(f_n|I - f|I) \le \varepsilon/2(\beta-\alpha)$ (c'est possible puisque la suite $(f_n|I)_{n\in\mathbb{N}}$ converge uniformément vers $f|I$). Pour $n\ge N$ on aura alors:

$$\left|\int_a^b f(t)\,dt - \int_a^b f_n(t)\,dt\right| \le 2\,\varepsilon/8 + 2\varepsilon/8 + \int_\alpha^\beta N_\infty(f_n|I - f|I)\,dt \le \varepsilon/2 + \int_\alpha^\beta \varepsilon/2(\beta-\alpha)\,dt = \varepsilon$$

Ce qui termine la démonstration.

<u>Exemple 4</u> : Les fonctions f_n définies par $f_n(t) = \dfrac{\sin\frac{t}{n}}{1+t^2}$ sont intégrables sur $]-\infty,+\infty[$, puisqu'elles sont continues et que $(\forall n)$ et $(\forall t)$ $|f_n(t)| \le \dfrac{1}{1+t^2} = g(t)$.

Par ailleurs, quel que soit A (A>0) $\sup_{-A\le t\le A}|f_n(t)| \le \sup_{-A\le t\le A}\left|\sin\frac{t}{n}\right| \le \sup_{-A\le t\le A}\left|\frac{t}{n}\right| = \frac{A}{n}$; donc la suite $(f_n|[-A,A])_{n\in\mathbb{N}}$ converge uniformément vers 0. Donc pour tout I fermé borné $(f_n|I)_{n\in\mathbb{N}}$ converge uniformément vers 0.

Nous pouvons en conclure que la suite des intégrales $\displaystyle\int_{-\infty}^{+\infty} \frac{\sin\frac{t}{n}}{1+t^2}\,dt$ converge vers 0.

Application du théorème de la convergence majorée aux intégrales sur [a,b]

<u>Exemple 5</u>: Considérons dans $\mathscr{C}^0([0,\pi/2])$, la suite définie par $f_n(t) = \sin^n t$. Elle converge simplement vers la fonction f qui vaut 0 pour $t\ne\pi/2$, et 1 pour $t=\pi/2$. Cette fonction f est intégrable. Elle n'est pas continue, donc la suite ne peut pas converger uniformément (cf chap.11 §6). Pourtant la suite numérique $\left(\int_0^{\pi/2} \sin^n t\,dt\right)_{n\in\mathbb{N}}$ converge vers $\int_0^{\pi/2} f(t)\,dt = 0$. Nous l'avons déja démontré en cours d'année (Chap.7 Ex 12). Nous allons voir que ceci résulte du théorème de la convergence majorée.

Nous pouvons considérer toutes ces intégrales comme des intégrales généralisées sur l'intervalle ouvert $]0,\pi/2[$. La fonction constante égale à 1 est une fonction majorante (c'est à dire que $(\forall t)$ $(\forall n)$ $|\sin^n t|\leq 1$) intégrable sur $]a,b[$. Par ailleurs si $[\alpha,\beta] \subset]0,\pi/2[$, la suite des $\sin^n(t)$ converge uniformément vers 0 sur $[\alpha,\beta]$ (Puisque, si $\alpha<\beta$, $N_\infty^{[\alpha,\beta]}(\sin^n t) = \sin^n\beta$, est une suite qui converge vers 0).

Donc le théorème de la convergence majorée, nous dit que l'intégrale de la fonction limite f est la limite des intégrales.

Très souvent, lorsqu'une suite de fonctions sur [a,b] converge simplement vers une fonction f non continue aux extrémités de l'intervalle, la convergence de la suite des intégrales vers l'intégrale de f, résulte ainsi du théorème de la convergence majorée appliqué à l'intervalle ouvert. Les intégrales sont alors considérées comme des intégrales généralisées sur cet intervalle ouvert.

La convergence uniforme sur tout compact.

Au chapitre 10 nous n'avons défini la convergence uniforme que pour les fonctions sur un intervalle fermé borné. En effet si I est un intervalle soit non borné, soit ouvert en l'une au moins de ses extrémités, les fonctions continues sur I ne sont pas toutes bornées, et la formule $\mathcal{N}_\infty(f) = \sup_{t\in I} |f(t)|$ *n'a donc aucun sens. Nous aurions pu tourner cette difficulté en nous plaçant dans l'espace* $\mathcal{C}^0_b(I,\mathbb{R})$ *des fonctions continues et bornées sur I . Nous ne l'avons pas fait parceque:*

** D'une part de nombreuses suites très simples de fonctions bornées ne sont pas convergentes pour cette norme. Dans l'exemple 1 ci-dessus,* $\sup_{t\in]-\infty,+\infty[} |f_n(t)| = 1$ $(\forall n)$*; et la suite* $(f_n)_{n\in\mathbb{N}}$ *ne converge donc pas vers 0 pour la norme uniforme sur* $\mathcal{C}^0_b(]-\infty,+\infty[,\mathbb{R})$*. Dans l'exemple 5 nous avons* $\sup_{t\in]0,\pi/2[} |\sin^n t| = 1$*, donc la suite* $(\sin^n t)_{n\in\mathbb{N}}$ *ne converge pas vers 0 pour la norme uniforme sur* $\mathcal{C}^0_b(]0,\pi/2[,\mathbb{R})$*.*

** D'autre part la convergence pour cette norme n'implique pas la convergence des intégrales vers l'intégrale de la limite. En effet plaçons nous sur* $]-\infty,+\infty[$*, et considérons les fonctions* $f_n(t) = \frac{n}{n^2+t^2}$ *. Nous avons* $\sup_{t\in]-\infty,+\infty[} |f_n(t)| = f_n(0) = \frac{1}{n}$ *; donc la suite* $(f_n)_{n\in\mathbb{N}}$ *converge vers 0 uniformément sur* $\mathbb{R}$ *tout entier.*

Pourtant $(\forall n)$ $\int_{-\infty}^{+\infty} f_n(t)\, dt = \pi$*, nous donne une suite ne convergeant pas vers 0.*

Pour les fonctions sur un intervalle I non borné ou non fermé la 'bonne' notion de convergence est celle qui est décrite dans le théorème de le convergence majorée: pour tout intervalle $[\alpha,\beta]$ *la suite* $(f_n|[\alpha,\beta])_{n\in\mathbb{N}}$ *converge uniformément vers* $f|[\alpha,\beta]$*. Lorsqu'il en est ainsi on dit que la suite* $(f_n)_{n\in\mathbb{N}}$ *converge vers f 'uniformément sur tout compact'.*

Exrcice b: Etudier si les suites de fonctions des exemples 1 à 5 ci-dessus convergent "uniformément sur tout compact".

§ 3 : Continuité et dérivabilité d'une fonction définie par une intégrale sur [a,b]

Continuité sous le signe intégrale

Considérons une fonction $f(t,u)$ *définie sur* $[a,b]\times[c,d]$*, et supposons que* $(\forall u\in[c,d])$ *l'intégrale* $\int_a^b f(t,u)\ dt$ *existe. Que pouvons nous dire de la fonction* $\varphi: u \to \int_a^b f(t,u)\ dt$ *?*

Sous les hypothèses que nous avons faites φ *peut ne pas être continue.*

Exemple: Soit $f(t,u) = \dfrac{t}{(u^2+t^2)\,\mathrm{Log}\,u}$ pour $u \neq 0$ et $f(t,0) = 0 = \lim\limits_{u\to 0} f(t,u)$. Nous avons pour $u\neq 0$

$$\varphi(u) = \int_0^1 f(t,u)\ dt = \int_0^1 \frac{t}{(u^2+t^2)\,\mathrm{Log}\,u}\,dt = \int_0^{1/u} \frac{v}{(1+v^2)\,\mathrm{Log}\,u}\,dv = \frac{1}{2\log u}\,\mathrm{Log}\,(1+\frac{1}{u^2})$$

$(v=t/u)$

Et ainsi $\varphi(u)$ tend vers -1 lorsque u tend vers 0, alors que $\varphi(0) = 0$. Ceci est dû à la non continuité de la fonction f en (0,0).

Si la fonction f: $[a,b]\times[c,d] \to \mathbb{R}$ est continue, alors la fonction φ: $[c,d] \to \mathbb{R}$ est continue.

Pour nous en convaincre nous considèrerons φ comme l'application composée:

$$[c,d] \xrightarrow{\Phi} \mathscr{C}^0([a,b],\mathbb{R}) \xrightarrow{\mathcal{I}} \mathbb{R}$$

Où Φ est l'application qui à u asssocie l'application f_u: $t \to f(t,u)$; et où $\mathcal{I}$ est l'application qui à toute fonction associe son intégrale sur [a,b]. Nous munissons $\mathscr{C}^0([a,b],\mathbb{R})$ de la norme uniforme N_∞. Nous savons (§1) que l'application $\mathcal{I}$ est continue; ceci signifie que, lorsque $(g_n)_{n\in\mathbb{N}}$ converge vers g pour la norme N_∞, alors $(\int_a^b g_n(t)\ dt)_{n\in\mathbb{N}}$ converge vers $\int_a^b g(t)\ dt$. La continuité de φ résultera donc de la continuité de Φ.

Pour montrer que Φ est continue, nous utiliserons le fait que f: $[a,b]\times[c,d] \to \mathbb{R}$ est uniformément continue (cf: Chapitre 10 §8); c'est à dire que (en désignant par $\| \ \|$ une norme sur $\mathbb{R}^2$, par exemple $\|(x,y)\| = \sup(|x|,|y|)$):

$$(\forall\varepsilon>0)\ (\exists\eta>0)\ (\forall(t,u))\ (\forall(t',v))\ (\|(t-t',u-v)\| \leq \eta) \text{ implique } |f(t,u) - f(t',v)| \leq \varepsilon$$

Dans cette formule nous pouvons prendre $t = t'$, nous obtenons alors:

$$(\forall\varepsilon>0)\ (\exists\eta>0)\ (\forall u)\ (\forall v)\ (\forall t)\ (|u-v| = \|(t-t,u-v)\| \leq \eta) \text{ implique } |f(t,u) - f(t,v)| \leq \varepsilon$$

Donc: $(\forall\varepsilon>0)\ (\exists\eta>0)\ (\forall u)\ (\forall v)\ (|u-v| \leq \eta)$ implique $\sup\limits_{t\in[a,b]} |f(t,u) - f(t,v)| \leq \varepsilon$

C'est à dire $(\forall\varepsilon>0)\ (\exists\eta>0)\ (\forall u)\ (\forall v)\ (|u-v| \leq \eta)$ implique $N_\infty(f_u - f_v) \leq \varepsilon$

Ce qui exprime la continuité uniforme de Φ.

Exercice b: Comme chaque fois que l'on utilise la continuité uniforme, on peut, dans le cas des fonctions de classe C^1, avoir un résultat un peu plus précis grace au théorème des accroissements finis. Supposons donc que f est C^1, et que M est un nombre tel que $(\forall(h,k) \in \mathbb{R}^2)\ (\forall(t,u) \in [a,b]\times[c,d])$: $|d_{(t,u)}(h,k)| \leq M\sqrt{h^2+k^2}$.

En utilisant le théorème des accroissements finis, démontrer que $(\forall u)\ (\forall v)\ (\forall t)\ |f(t,u)-f(t,v)| \leq M|u-v|$; en déduire que $(\forall u)\ (\forall v)\ N_\infty(f_u-f_v) \leq M|u-v|$. En déduire un nombre K tel que $(\forall u)\ (\forall v)\ |\varphi(u)-\varphi(v)| \leq K|u-v|$.

Dérivation sous le signe intégrale

Si la fonction $(t,u) \to f(t,u)$ est de classe C^1 sur $[a,b]\times[c,d]$, alors φ: $[c,d] \to \mathbb{R}$ est de classe C^1, et sa dérivée est donnée par la formule $\varphi'(u) = \int_a^b \frac{\partial f}{\partial u}(t,u)\ dt$

Démonstration: Choisissons un u, et démontrons la dérivabilité de φ en u. Nous devons, pour tout ε (>0), trouver des conditions permettant d'affirmer que:

$$A = \int_a^b f(t,u+h)\ dt - \int_a^b f(t,u)\ dt - \int_a^b \frac{\partial f}{\partial u}(t,u)\,h\ dt$$

est, en valeur absolue, inféfrieur à $\varepsilon|h|$. Posons:

$$\omega(t,h) = f(t,u+h) - f(t,u) - \frac{\partial f}{\partial u}(t,u)\,h$$

Puisque $A = \int_a^b \omega(t,h)\,dt$, l'inégalité $|A| \leq \varepsilon|h|$ résultera de l'inégalité $(\forall t)\ |\omega(t,h)| \leq \frac{\varepsilon}{b-a}|h|$.

Puisque la fonction f est de classe C^1, sa dérivée partielle $\frac{\partial f}{\partial u}$ est continue (donc uniformément continue, puisque nous travaillons sur un sous-ensemble compact du plan), il existe donc un nombre α (>0) tel que $(\forall t)\ (\forall h)\ (|h| \leq \alpha)$ implique $|\frac{\partial f}{\partial u}(t,u+h) - \frac{\partial f}{\partial u}(t,u)| \leq \frac{\varepsilon}{b-a}$.

Exercice c : Expliquer pourquoi la continuité uniforme de D_u sur $[a,b]\times[c,d]$ permet d'affirmer qu'il existe $\alpha(>0)$ tel que $(\forall t)$ $(\forall h)\ (|h| \leq \alpha)$ implique $|\frac{\partial f}{\partial u}(t,u+h) - \frac{\partial f}{\partial u}(t,u)| \leq \frac{\varepsilon}{b-a}$.

Mais $\frac{\partial f}{\partial u}(t,u+h) - \frac{\partial f}{\partial u}(t,u) = \frac{\partial}{\partial h}(\omega(t,h))$, donc, d'après le théorème des accroissements finis:

$$(|h| \leq \alpha) \text{ implique } |\omega(t,h)| = |\omega(t,h) - \omega(t,0)| \leq \frac{\varepsilon}{b-a}|h|$$

Ce qui (implique $|A| \leq \varepsilon|h|$ et) termine la démonstration.

§ 4 : Continuité et dérivabilité d'une fonction définie par une intégrale sur]a,b[.

Considérons maintenant un intervalle]a,b[(borné ou non) et une fonction f de $]a,b[\times[c,d]$ dans $\mathbb{R}$. Il ne suffit plus que la fonction f soit continue, pour que φ soit continue, il ne suffit plus que f soit C^1, pour que φ soit C^1. Nous allons être obligés d'introduire des conditions de majoration (analogues à celle qui a été utilisée au §2). Plus précisément:

Soit $f:]a,b[\times[c,d] \to \mathbb{R}$ une fonction continue.

S'il existe une fonction $g:]a,b[\to \mathbb{R}^+$ intégrable, telle que $(\forall t)\ (\forall u)\ |f(t,u)| \leq g(t)$, alors, pour tout u, $\varphi(u) = \int_a^b f(t,u)\,dt$ existe (chap.9), et $\varphi: [c,d] \to \mathbb{R}$ est continue.

Si de plus la fonction $(t,u) \to f(t,u)$ est de classe C^1 sur $]a,b[\times[c,d]$, et s'il existe une fonction positive intégrable $h:]a,b[\to [0,\infty[$, telle que $(\forall t)\ (\forall u)\ |\frac{\partial f}{\partial u}(t,u)| \leq h(t)$, alors $\varphi: [c,d] \to \mathbb{R}$ est de classe C^1, et sa dérivée est donnée par la formule $\varphi'(u) = \int_a^b \frac{\partial f}{\partial u}(t,u)\,dt$

<u>Démonstration de la continuité</u>: Considérons une suite d'intervalles $[\alpha_n,\beta_n]$ tels que $(\alpha_n)_{n\in\mathbb{N}}$ soit décroissante et $\text{Lim}\ \alpha_n = a$ et $(\beta_n)_{n\in\mathbb{N}}$ croissante et $\text{Lim}\ \beta_n = b$. Et posons $\varphi_n(u) = \int_{\alpha_n}^{\beta_n} f(t,u)\,dt$. Sur l'intervalle $[\alpha_n,\beta_n]$ nous pouvons appliquer les résultats du §3, les fonctions φ_n sont donc continues. Par ailleurs $(\forall u)$, pour p>n:

$$|\varphi_p(u) - \varphi_n(u)| \leq \left|\int_{\alpha_p}^{\alpha_n} f(t,u)\,dt\right| + \left|\int_{\beta_n}^{\beta_p} f(t,u)\,dt\right| \leq \int_{\alpha_p}^{\alpha_n} |f(t,u)|\,dt + \int_{\beta_n}^{\beta_p} |f(t,u)|\,dt$$

$$|\varphi_p(u) - \varphi_n(u)| \leq \int_{\alpha_p}^{\alpha_n} g(t)\,dt + \int_{\beta_n}^{\beta_p} g(t)\,dt$$

Donc $$N_\infty(\varphi_p - \varphi_n) \leq \int_{\alpha_p}^{\alpha_n} g(t)\,dt + \int_{\beta_n}^{\beta_p} g(t)\,dt \qquad (N_\infty = \text{norme uniforme sur } [c,d])$$

Puisque l'intégrale de g converge en a et en b, quel que soit $\varepsilon>0$, il existe des nombres a_1 et b_1 tels que $\int_a^{a_1} g(t)\,dt \leq \varepsilon$ et $\int_{b_1}^{b} g(t)\,dt \leq \varepsilon$. Puisque les suite $(\alpha_n)_{n\in\mathbb{N}}$ et $(\beta_n)_{n\in\mathbb{N}}$ convergent respectivement vers a et b, il existe N tel que $n\geq N$ implique $\alpha_n \in\,]a,a_1]$ et $\beta_n\in [b_1,b[$. Il en résulte que si n et p sont supérieurs à N (et p>n):

$$N_\infty(\varphi_p-\varphi_n) \leq \int_{\alpha_p}^{\alpha_n} g(t)\,dt + \int_{\beta_n}^{\beta_p} g(t)\,dt \leq \int_a^{a_1} g(t)\,dt + \int_{b_1}^{b} g(t)\,dt \qquad \leq 2\varepsilon$$

Nous avons ainsi démontré que la suite $(\varphi_n)_{n\in\mathbb{N}}$ est une suite de Cauchy pour la norme uniforme. Puisque $\mathscr{C}^0([c,d],\mathbb{R})$ muni de la norme uniforme est complet, la suite $(\varphi_n)_{n\in\mathbb{N}}$ converge uniformément, vers une fonction continue. Et, puisque $(\forall u)$ $\operatorname{Lim}\varphi_n(u) = \varphi(u)$, cette limite est φ. Donc φ est continue.

Exercice d: Expliquer pourquoi il existe a_1 et b_1 tels que $\int_a^{a_1} g(t)\,dt \leq \varepsilon$ et $\int_{b_1}^{b} g(t)\,dt \leq \varepsilon$.

<u>Démonstration de la dérivabilité</u>: Nous considèrerons les mêmes fonctions φ_n; puisque f est C^1 - et à cause du §3 - elles sont C^1. Nous démontrerons que la suite $(\varphi_n)_{n\in\mathbb{N}}$ est convergente pour la norme ν_∞ sur $\mathscr{C}^1([c,d],\mathbb{R})$ (ce qui prouvera que sa limite φ est C^1; voir chapitre 11 §6), et pour cela nous démontrerons que c'est une suite de Cauchy pour ν_∞. Nous allons commencer par majorer $N_\infty(\frac{d\varphi_p}{du}-\frac{d\varphi_n}{du})$, de façon tout à fait analogue à ce que nous avons fait ci-dessus. Quel que soit u (si p>n):

$$\left|\frac{d\varphi_p}{du}(u)-\frac{d\varphi_n}{du}(u)\right| = \left|\int_{\alpha_p}^{\alpha_n}\frac{\partial f}{\partial u}(t,u)\,dt + \int_{\beta_n}^{\beta_p}\frac{\partial f}{\partial u}(t,u)\,dt\right| \leq \left|\int_{\alpha_p}^{\alpha_n}\frac{\partial f}{\partial u}(t,u)\,dt\right| + \left|\int_{\beta_n}^{\beta_p}\frac{\partial f}{\partial u}(t,u)\,dt\right|$$

$$\left|\frac{d\varphi_p}{du}(u)-\frac{d\varphi_n}{du}(u)\right| \leq \int_{\alpha_p}^{\alpha_n}\left|\frac{\partial f}{\partial u}(t,u)\right|dt + \int_{\beta_n}^{\beta_p}\left|\frac{\partial f}{\partial u}(t,u)\right|dt \leq \int_{\alpha_p}^{\alpha_n} h(t)\,dt + \int_{\beta_n}^{\beta_p} h(t)\,dt$$

Donc $N_\infty(\frac{d\varphi_p}{du}-\frac{d\varphi_n}{du}) \leq \int_{\alpha_p}^{\alpha_n} h(t)\,dt + \int_{\beta_n}^{\beta_p} h(t)\,dt$

Et en tenant compte de la majoration faite dans la démonstration de la continuité, nous obtenons

$$\nu_\infty(\varphi_p-\varphi_n) \leq \int_{\alpha_p}^{\alpha_n} h(t)\,dt + \int_{\beta_n}^{\beta_p} h(t)\,dt + \int_{\alpha_p}^{\alpha_n} g(t)\,dt + \int_{\beta_n}^{\beta_p} g(t)\,dt$$

Exercice e: En s'inspirant de ce qui a été fait dans la démonstration sur la continuité, en déduire que $(\varphi_n)_{n\in\mathbb{N}}$ est une suite de Cauchy pour la norme ν_∞. Et conclure.

<u>Exemple</u>: Posons $\varphi(u) = \int_{-\infty}^{+\infty}\frac{\cos ut}{1+t^4}\,dt$. Nous intégrons une fonction C^∞. Donc pour montrer que φ est continue il suffit de remarquer que $(\forall t)$ $(\forall u)$ $\left|\frac{\cos ut}{1+t^4}\right| \leq \frac{1}{1+t^4} = g(t)$ (la fonction g étant clairement positive et absolument intégrable sur $]-\infty,+\infty[$).

Pour démontrer que φ est de classe C^1, et que $\frac{\partial\varphi}{\partial u}(u) = \int_{-\infty}^{+\infty}\frac{\partial\varphi}{\partial u}(\frac{\cos ut}{1+t^4})\,dt$, il suffit de remarquer que $\left|\frac{\partial\varphi}{\partial u}(\frac{\cos ut}{1+t^4})\right| = \left|\frac{-t\sin tu}{1+t^4}\right| \leq \frac{|t|}{1+t^4} = h(t)$ (qui est évidemment une fonction positive intégrable sur $]-\infty,+\infty[$).

Exercices sur le chapitre 14

* **Exercice 1**: On pose $u_n = \int_0^1 \text{Log}(1+t^n) \sin t \, dt$. Montrer que la suite $(u_n)_{n\in\mathbb{N}}$ converge vers 0.

* **Exercice 2**: On pose $u_n = \int_0^1 \text{Log}(1+\sqrt[n]{t}) \sin t \, dt$. Montrer que la suite $(u_n)_{n\in\mathbb{N}}$ est convergente et calculer sa limite.

* **Exercice 3**: On pose $u_n = \int_1^\infty \text{Log}\left(\frac{1+t^{n+1}}{t^n}\right) \frac{dt}{t^2}$. Montrer que la suite $(u_n)_{n\in\mathbb{N}}$ est convergente et calculer sa limite.

* **Exercice 4**: On pose $u_n = \int_0^1 \text{Log}(\cos t^n) \, dt$. Montrer que la suite $(u_n)_{n\in\mathbb{N}}$ est convergente, et calculer sa limite.

** **Exercice 5**: Peut-on écrire $\int_0^1 \frac{\cos \pi t}{1+t^2} dt = \sum_{n=0}^{\infty} (-1)^n \int_0^1 t^{2n} \cos \pi t \, dt$?

** **Exercice 6**: Démontrer que la suite $(u_n)_{n\in\mathbb{N}}$ définie par $u_n = \int_{-1}^1 (1-t^2)^n \, dt$ converge vers 0.

*** **Exercice 7**: On pose $u(x) = \int_0^{\pi/2} \text{Log}(x+\cos tx) \, dt$.

a) Montrer que $u(x)$ existe pour tout $x \geq 0$.

b) Montrer que f est de classe C^1 pour $x>0$, et donner une expression de sa dérivée.

** **Exercice 8**: Montrer que la fonction définie (pour $x>0$) par $f(x) = \int_0^\infty e^{-t} \text{Log}(x+t) \, dt$ est de classe C^1, et donner une expression intégrale de sa dérivée.

* **Exercice 9**: Démontrer que la fonction f définie par $f(x) = \int_1^\infty e^{-t} \cos(t^2 x) \, dt$ est de classe C^1. Puis montrer que f est de classe C^2, et donner une expression intégrale de sa dérivée seconde.

** **Exercice 10**: Soit la fonction définie pour $x>0$ par $f(x) = \int_0^1 \frac{\text{Log}(1+xt)}{1+t} dt$.

Montrer que f est C^1 et que, pour $x \neq 1$, $f'(x) = \frac{\text{Log } 2}{x-1} + \frac{1}{x(1-x)} \text{Log}(1+x)$.

** Exercice 11: Calculer $f(y) = \int_a^b \frac{dx}{y+x^2}$ $(y>0)$.

a) Calculer de deux façons la dérivée de f, en déduire $\int_a^b \frac{dx}{(1+x^2)^2}$.

b) Calculer de même $\int_a^b \frac{dx}{(1+x^2)^3}$, $\int_a^b \frac{dx}{(1+x^2)^4}$,...

** Exercice 12: Soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \frac{1}{\pi} \int_0^\pi \cos(x \sin t)\, dt$.

a) Montrer que f est de classe C^2.

b) Montrer que f est solution de l'équation différentielle $xy'' + y' + xy = 0$.

c) Montrer que la série entière $\left(\left(\frac{(-1)^n z^{2n}}{[2^n\, n!]^2}\right)\right)_{n\geq 0}$ a un rayon de convergence infini. Montrer que sa somme est solution de la même équation différentielle. Montrer que sa somme est f.

** Exercice 13: On définit une application $f : \mathbb{R}^2 \to \mathbb{R}$ par $f(x,y) = \int_{-y}^0 \frac{\sin xt}{1+t^2}\, dt$. Montrer que f est de classe C^1 et calculer sa différentielle au point (x,y).

** Exercice 14: Justifier l'égalité $\int_0^1 \frac{t\, dt}{1+t^3} = \sum_{n\geq 0} (-1)^n \int_0^1 t^{3n+1}\, dt$.

En déduire la somme de la série $\left(\left(\frac{(-1)^n}{3n+2}\right)\right)_{n\geq 0}$.

* Exercice 15: Soit $\varphi(x) = \int_{-\infty}^{+\infty} \frac{\sin t}{x^2+t^2}\, dt$. Démontrer que $\varphi(x)$ existe pour tout $x\neq 0$, et démontrer que φ est une fonction continue et dérivable sur $\mathbb{R}$ privé de 0.

** Exercice 16: On pose $u_n = \int_1^\infty \frac{\cos t}{t^{2n}}\, dt$. Démontrer que u_n existe pour tout $n\geq 1$. Démontrer que la série $(((-1)^n u_n))_{n\geq 1}$ est sommable et exprimer sa somme comme une intégrale.

** **Exercice 17:** On définit une fonction u par $u(x) = \int_0^{\infty} \frac{\cos xt}{1+t^2}\,dt$

a) Montrer que u est définie et continue sur $\mathbb{R}$ tout entier.

b) Montrer que (pour $x\neq 0$) $u(x) = \frac{1}{x}\int_0^{\infty} \frac{2t}{(1+t^2)^2}\sin xt\,dt$. En déduire que u est dérivable (pour $x\neq 0$).

c) Montrer que $u'(x) = \int_0^{\infty} \frac{-t}{1+t^2}\sin xt\,dt$ (on n'oubliera pas de justifier l'existence de cette dernière intégrale).

d) Montrer que (pour $x\neq 0$) $u(x) - u(0) = -x\int_0^{a} \frac{4\sin^2 v}{x^2+4v^2}\,dv$ (où $a = +\infty$ si $x>0$, et $a = -\infty$ si $x<0$) En déduire que u a une dérivée à droite en 0, mais n'est pas dérivable en 0.

(Extrait d'un partiel de 1982)

** **Exercice 18:** Soit g une fonction continue et bornée sur $\mathbb{R}$. On pose $G(x) = \int_{-\infty}^{+\infty} e^{-t^2} g(t-x)\,dt$.

a) Démontrer que G(x) existe quelque soit x. Démontrer que $G(x) = \int_{-\infty}^{+\infty} e^{-(t-x)^2} g(t)\,dt$.

b) En déduire que G est de classe C^∞, et donner une expression de sa dérivée d'ordre n.

*** **Exercice 19:** On pose $G(x) = \int_0^{\pi/2} \mathrm{Log}\,(\cos^2 t + x^2\sin^2 t)\,dt$. Calculer G(1).

Montrer que G est de classe C^1; calculer explicitement G'. En déduire G.

** **Exercice 20**: On considère dans $\mathscr{C}^0([0,1],\mathbb{R})$ la série $((f_n))_{n\in\mathbb{N}}$ définie par $f_n(t) = \frac{t}{1+t^2n^2}$.

a) Montrer que, quel que soit t, la série $((f_n(t)))_{n\in\mathbb{N}}$ est sommable. On note F sa somme.

b) Montrer que $((f_n))_{n\in\mathbb{N}}$ est uniformément sommable sur tout intervalle de la forme $[\varepsilon,1]$ ($\varepsilon>0$).

c) Montrer que F est continue sur $]0,1]$. Montrer que pour $t = \frac{1}{n}$, la série $((f_n(t)))_{n\in\mathbb{N}}$ possède au moins n termes supérieurs à $\frac{1}{2n}$; en déduire que F n'est pas continue en 0.

d) Soit (pour $0<a<1$) $I_a = \int_a^1 F(t)\,dt$. Donner une série dont la somme est I_a.

e) Montrer que $I = \int_0^1 F(t)\,dt$ existe, et donner une série de somme I.

Séries de Fourier

La lecture de ce qui suit suppose la connaissance des notions de produit scalaire, produit hermitien, systèmes orthonormés,... introduits au chapitre 4; le lecteur est prié de s'y reporter.

Dans tout ce chapitre, T est un nombre strictement positif fixé une fois pour toutes, et souvent égal à 2π. On note $\mathscr{C}^0_T(\mathbb{R},\mathbb{R})$ (resp. $\mathscr{C}^0_T(\mathbb{R},\mathbb{C})$) l'espace des fonctions continues sur $\mathbb{R}$, admettant T pour période, à valeurs dans $\mathbb{R}$ (resp. dans $\mathbb{C}$).

Choisissons un nombre strictement positif K, nous définirons un produit scalaire sur l'espace vectoriel $\mathscr{C}^0_T(\mathbb{R},\mathbb{R})$ en posant:

$$\langle f,g\rangle = K\int_0^T f(t)\,g(t)\,dt$$

Nous définirons un produit hermitien sur $\mathscr{C}^0_T(\mathbb{R},\mathbb{C})$ en posant:

$$(f\,|\,g) = K\int_0^T f(t)\,\overline{g(t)}\,dt$$

Exercice a: Montrer que l'on a ainsi défini un produit scalaire et un produit hermitien, la seule difficulté est de montrer que si $(f\,|\,f) = 0$, alors f est nulle. Pour cela on s'inspirera de l'exercice b du chapitre 4.

§ 1 : Polynômes trigonométriques.

Dans ce paragraphe T est égal à 2π.

On appellera polynômes trigonométriques complexes de degré n, les combinaisons linéaires (à coefficients complexes) de $e^{-int}, e^{-i(n-1)t}, \dots, e^{-it}, 1, e^{it}, \dots e^{i(n-1)t}, e^{int}$. Ils forment un sous-espace vectoriel de $\mathscr{C}^0_{2\pi}(\mathbb{R},\mathbb{C})$; nous le noterons $E^n_{\mathbb{C}}$.

L'espace $E^n_{\mathbb{C}}$ est de dimension 2n+1. Pour s'en convaincre il suffit de montrer que les 2n+1 vecteurs $e^{-int}, e^{-i(n-1)t}, \dots, e^{-it}, 1, e^{it}, \dots e^{i(n-1)t}, e^{int}$ sont indépendants. Pour cela on démontre qu'ils sont deux à deux orthogonaux.

$$(e^{ipt}\,|\,e^{iqt}) = K\int_0^{2\pi} e^{ipt}\,e^{-iqt}\,dt = K\int_0^{2\pi} e^{i(p-q)t}\,dt = \begin{cases} 0 \text{ si } p\neq q \\ 2\pi K \text{ si } p=q \end{cases}$$

On appelle polynômes trigonométriques à coefficients réels, les fonctions de $\mathscr{C}^0_{2\pi}(\mathbb{R},\mathbb{R})$ qui, considérées comme des fonctions à valeurs complexes, sont des polynômes trigonométriques à coefficients complexes; ils forment un sous-espace $E^n_{\mathbb{R}}$ de $\mathscr{C}^0_{2\pi}(\mathbb{R},\mathbb{R})$. Les fonctions 1, cos t, sin t, cos 2t, sin 2t,...,cos nt, sin nt forment une base de $E^n_{\mathbb{R}}$, qui est donc de dimension 2n+1 (tout comme $E^n_{\mathbb{C}}$).

En effet: Il est clair que 1, cos t, sin t, cos 2t, sin 2t,…,cos nt, sin nt sont dans $E^n_{\mathbb{R}}$, car cos αt et sin αt s'expriment en fonction de $e^{i\alpha t}$ et $e^{-i\alpha t}$. Ils sont indépendants car l'égalité

$$(\forall t) \quad \lambda_0 + \sum_{k=1}^{n} \lambda_k \cos kt + \mu_k \sin kt = 0$$

implique

$$(\forall t) \quad \lambda_0 + \sum_{k=1}^{n} \frac{1}{2}(\lambda_k - i\mu_k)\, e^{ikt} + \frac{1}{2}(\lambda_k + i\mu_k)\, e^{-ikt} = 0$$

D'où $\lambda_0 = 0$ et $(\forall k)$ $\lambda_k - i\mu_k = \lambda_k + i\mu_k = 0$ (puisque les e^{ikt} forment une base de $E^n_{\mathbb{C}}$); ce qui entraine bien $(\forall k)$ $\lambda_k = \mu_k = 0$.

Il nous reste à démontrer que ces 2n+1 fonctions engendrent $E^n_{\mathbb{R}}$. Nous ferons une récurrence sur n. C'est clair pour n = 0; supposons que ce soit vrai pour n-1. Soit P un élément de $E^n_{\mathbb{R}}$; il s'écrit $P(t) = Q(t) + a\, e^{int} + b\, e^{-int}$, où Q est dans $E^{n-1}_{\mathbb{R}}$ (donc combinaison linéaire de 1,cos t, sin t,…cos (n-1)t, sin (n-1)t d'après l'hypothèse de récurrence). Démontrons que $a\, e^{int} + b\, e^{-int}$ (qui est une fonction réelle, a et b étant complexes) est combinaison linéaire (à coefficients réels) de cos nt et sin nt. Nous avons

$$a\, e^{int} + b\, e^{-int} = (a+b) \cos nt + (a-b)\, i \sin nt$$

En faisant t = 0, puis $t = \pi/2n$, on voit que a+b et (a-b)i sont réels; d'où le résultat.

En fait {1, cos t, sin t, cos 2t, sin 2t,…,cos nt, sin nt} est une base orthogonale de $E^n_{\mathbb{R}}$. De plus $\langle 1,1\rangle = 2\pi K$ et $(\forall k>0)$ $\langle \cos kt, \cos kt\rangle = \langle \sin kt, \sin kt\rangle = \pi K$.

Dans toute la suite on choisit $K = \frac{1}{2\pi}$; ainsi la base $\{1, \sqrt{2}\cos t, \sqrt{2}\sin t, \ldots, \sqrt{2}\cos nt, \sqrt{2}\sin nt\}$ est orthonormée dans $E^n_{\mathbb{R}}$, et $\{e^{-int}, \ldots, 1, \ldots, e^{int}\}$ orthonormée dans $E^n_{\mathbb{C}}$.

Exercice b: Calculer les produits scalaires deux à deux de ces 2n+1 fonctions., et justifier ainsi ces affirmations.

§ 2 : Séries de Fourier.

Soit f dans $\mathscr{C}^0_{2\pi}(\mathbb{R},\mathbb{C})$, nous appellerons coefficients de Fourier (complexes) de f les produits hermitiens $\alpha_k(f) = (f \mid e^{ikt})$ $(\forall k \in \mathbb{Z})$. On appelle série de Fourier de f la série de fonctions définie par $u_0(t) = \alpha_0$ et $(\forall k>0)$ $u_k(t) = \alpha_k\, e^{ikt} + \alpha_{-k}\, e^{-ikt}$.

Soit f dans $\mathscr{C}^0_{2\pi}(\mathbb{R},\mathbb{R})$, nous appellerons coefficients de Fourier (réels) de f les produits scalaires $a_0(f) = \langle f,1\rangle$ et $(\forall k>0)$ $a_k(f) = \langle f, \sqrt{2}\cos kt\rangle$ et $b_k(f) = \langle f, \sqrt{2}\sin kt\rangle$. On appelle série de Fourier de f la série de fonctions définie par $u_0(t) = a_0$ et $(\forall k>0)$ $u_k(t) = a_k\sqrt{2}\cos kt + b_k\sqrt{2}\sin kt$.

Exercice c: Soit f une fonction dans $\mathscr{C}^0_{2\pi}(\mathbb{R},\mathbb{R})$, montrer que $\alpha_0 = a_0$, et que, quel que soitk, $a_k = \frac{1}{\sqrt{2}}(\alpha_k + \alpha_{-k})$ et $b_k = \frac{1}{i\sqrt{2}}(\alpha_k - \alpha_{-k})$. Vérifier que la série de Fourier réelle et la série de Fourier complexe de f coïncident.

Nous noterons $f_n = \sum_{k \le n} u_k$ les sommes partielles de la série de Fourier de f. Pour tout n, la fonction f_n est la projection orthogonale de f sur l'espace $E^n_{\mathbb{R}}$ (ou sur $E^n_{\mathbb{C}}$ dans le cas complexe). Pour s'en convaincre il suffit de remarquer que $f_n = \sum_i \langle f,e_i\rangle e_i$, où $\{e_i\}$ est une base orthonormée de l'espace sur lequel on projette (cf: chap.4 §4).

Variantes: Considérons le cas d'une période T différente de 2π. Nous remplacerons les fonctions $t \to e^{int}$ par les fonctions $t \to e^{2i\pi nt/T}$, qui sont de période T/n. L'espace $E^n_{\mathbb{C},T}$ sera le sous-espace vectoriel de $\mathscr{C}^0_T(\mathbb{R},\mathbb{C})$ engendré par les fonctions $t \to e^{2i\pi kt/T}$ pour $-n \le k \le n$. Ces fonctions en forment une base orthogonale. Ces vecteurs ont tous pour norme $\|\sqrt{KT}\|$; pour obtenir une base orthonormée nous prendrons donc les fonctions ε_i: $t \to \frac{1}{\sqrt{KT}} e^{2i\pi nt/T}$. Les produits hermitiens de f et de ces fonctions sont les coefficients de Fourier (complexes) de f; et les projections de f sur les $E^n_{\mathbb{C},T}$ sont les sommes partielles de la série de Fourier de f.

Nous remplacerons $t \to \cos kt$ et $t \to \sin kt$ par les fonctions $t \to \cos 2\pi kt/T$ et $t \to \sin 2\pi kt/T$. Ces fonctions (pour $k \le n$) engendrent l'espace $E^n_{\mathbb{R},T}$, et en forment une base orthogonale. Pour avoir une base orthonormée nous les remplacerons par φ_0: $t \to \frac{1}{\sqrt{KT}}$, et (pour $k>0$) φ_k: $t \to \frac{1}{\sqrt{KT/2}} \cos 2\pi kt/T$ et φ_k: $t \to \frac{1}{\sqrt{KT/2}} \sin 2\pi kt/T$. Les coefficients de Fourier réels de f sont les produits scalaires de f avec les éléments de cette base orthonormée, et la série de Fourier réelle de f a pour sommes partielles les projections de f sur les sous-espaces $E^n_{\mathbb{R},T}$. Les séries de Fourier réelle et complexe d'une fonction réelle sont identiques.

Exercice d: Ecrire les détail des constructions esquissées dans ce paragraphe, en faisant les calculs qui ont été omis.

Ainsi à toute fonction continue de période T nous avons associé une série. Celle ci est combinaison linéaire des fonctions périodiques que nous connaissions: les fonctions cos et sin. Ceci n'a d'intérêt que si nous pouvons montrer que cette série est sommable de somme f (i.e.: si la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f). Mais nous sommes dans un espace de fonctions, il nous faut donc, avant de parler de convergence, dire quelle norme nous utilisons. A priori nous pouvons choisir une norme arbitraire, mais deux d'entre elles sont particulièrement intéressantes:

La norme associée au produit scalaire (ou hermitien) qui nous a servi à définir la série de Fourier. Nous la noterons $\mathcal{N}_2$.

La norme uniforme $\mathcal{N}_\infty$ définie par $\mathcal{N}_\infty(f) = \sup_{t\in\mathbb{R}} (|f(t)|) = \sup_{t\in[0,T]} (|f(t)|)$. Notons que cette norme existe parceque toute fonction continue périodique sur $\mathbb{R}$ est bornée.

Nous allons envisager successivement les problèmes de convergence sous ces deux aspects. Et nous verrons qu'ils se présentent de façon tout à fait différente.

§ 3 : La convergence des séries de Fourier pour la norme quadratique.

La série de Fourier d'une fonction continue f est sommable pour la norme quadratique N_2, et sa somme est f. Autrement dit, au sens de la norme N_2, la suite $(f_n)_{n\in\mathbb{N}}$ a pour limite f. La démonstration est faite dans le complément 1.

Attention ceci n'implique pas que, pour une valeur t_0 donnée de la variable, la suite $(f_n(t_0))_{n\in\mathbb{N}}$ converge vers $f(t_0)$ (chap 10 Fig.page 121), nous y reviendrons au §4.

Les normes quadratiques des termes u_k de la série de Fourier sont données par

$N_2(u_0) = |a_0|$ $\qquad$ ($N_2(u_0) = |\alpha_0|$ dans le cas complexe).

$N_2(u_k) = \sqrt{(a_k)^2+(b_k)^2}$ $\qquad$ ($N_2(u_k) = \sqrt{|\alpha_k|^2+|\alpha_{-k}|^2}$ dans le cas complexe).

La série des carrés de ces normes est sommable de somme $[N_2(f)]^2$. *Autrement dit*

$$N_2(f)^2 = \sum_{i=0}^{\infty} N_2(u_i)^2 = (a_0)^2 + \sum_{k=1}^{\infty} (a_k)^2+(b_k)^2 = (\alpha_0)^2 + \sum_{k=1}^{\infty} |\alpha_k|^2+|\alpha_{-k}|^2.$$

Ce résultat est connu sous le nom de formule de Parseval. Il implique que les suites $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n>0}$, $(\alpha_n)_{n\in\mathbb{N}}$ *et* $(\alpha_{-n})_{n\in\mathbb{N}}$ *convergent vers 0.*

Pour nous en convaincre remarquons que $\sum_{i=0}^{n} N_2(u_i)^2$ est le carré de la norme de f_n (puisque $f_n = \sum_{i\leq n} u_i$ et que les u_i sont deux à deux orthogonaux). Et puisque f_n est la projection orthogonale de f sur l'espace des polynômes trigonométriques de degré n, nous avons ($f_n \perp (f-f_n)$ donc):

$$N_2(f)^2 = N_2(f-f_n)^2 + N_2(f_n)^2 = N_2(f-f_n)^2 + \sum_{i=0}^{n} N_2(u_i)^2 .$$

Puisque $(f_n)_{n\in\mathbb{N}}$ converge vers f pour la norme N_2, la suite $(N_2(f-f_n))_{n\in\mathbb{N}}$ converge vers 0. D'où le résultat annoncé.

<u>Exemple</u>: Soit f la fonction de période 2π définie par $(\forall t\in[0,2\pi])\ f(t) = t(2\pi-t)$. Ses coefficients de Fourier réels sont $a_0 = \frac{2\pi^2}{3}$, et $(\forall k>0)\ a_k = \frac{-2\sqrt{2}}{k^2}$ et $b_k = 0$. Donc $u_0(t) = \frac{2\pi^2}{3}$ (et $N_2(u_0) = \frac{2\pi^2}{3}$) et $(\forall k>0)\ u_k(t) = \frac{-4}{k^2}\cos kt$ (et $N_2(u_k) = \frac{2\sqrt{2}}{k^2}$). Tous ces calculs sont faciles, le lecteur est invité à les vérifier. La formule de Parseval s'écrit donc

$$\frac{8}{15}\pi^4 = [N_2(f)]^2 = \frac{4\pi^4}{9} + \sum_{1}^{\infty} \frac{8}{k^4}$$

Ce qui nous donne la somme de la série $((\frac{1}{k^4}))_{k\geq 1}$.

§ 4 : Convergence ponctuelle et convergence uniforme.

Conditions pour la convergence ponctuelle des séries de Fourier.

Une série de Fourier réelle est la somme de deux séries trigonométriques. *Nous savons que les suites* $(a_n)_{n\in\mathbb{N}}$ *et* $(b_n)_{n>0}$ *convergent vers 0. Si elles sont décroissantes le critère d'Abel nous dit que la série numérique* $((u_n(t)))_{n\in\mathbb{N}}$ *est sommable pour* $t\neq 0$. *Mais rien ne nous permet d'affirmer à priori que sa somme est* $f(t)$.

On peut montrer (et nous admettrons) *que si en un point* t_0 *la fonction f a une dérivée à droite et une dérivée à gauche, alors la série numérique* $((u_n(t)))_{n\in\mathbb{N}}$ *est sommable de somme* $f(t)$.

Ainsi dans l'exemple ci-dessus, nous avons pour tout t dans $[0,2\pi]$

$$(\forall t) \qquad t(2\pi-t) = \frac{2\pi^2}{3} + \sum_{1}^{\infty} \frac{-4}{k^2}\cos kt$$

Conditions pour la convergence uniforme des séries de Fourier.

Soit f de classe C^1 de période 2π, alors sa dérivée f ' est aussi périodique de période 2π, et ses coefficients de Fourier (réels) sont donnés par les formules

$$(\forall p \geq 1) \qquad a_p(f') = +\,p\,b_p(f) \qquad b_b(f') = -\,p\,a_p(f)$$

Exercice e: Démontrer ces formules au moyen d'une intégration par parties.

Nous savons que les coefficients de Fourier de f' forment des suites qui convergent vers 0. Il en résulte que les suites $(a_p(f))_{p\in\mathbb{N}}$ et $(b_p(f))_{p\geq 1}$ sont $\mathrm{o}(1/p)$. En itérant le procédé, nous démontrerons que, si f est de classe C^2, les suites $(a_p(f))_{p\in\mathbb{N}}$ et $(b_p(f))_{p\geq 1}$ sont $\mathrm{o}(1/p^2)$, ce qui implique que la série $((|a_p(f)| + |b_p(f)|))_{p\geq 1}$ est sommable. Et puisque, pour tout t, $|u_p(t)| \leq |a_p(f)| + |b_p(f)|$, nous pouvons affirmer que $N_\infty(u_p) \leq |a_p(f)| + |b_p(f)|$. Ainsi lorsque f est de classe C^2 la série $((u_p))_{p\in\mathbb{N}}$ est sommable pour la norme uniforme. On démontre que sa somme est f.

Exercice f: Se reporter au chap.11 (§6), et montrer que si la série de Fourier de f converge uniformément, et a pour somme g, alors elle converge vers g pour la norme N_2. En déduire que g = f.

On peut démontrer (nous ne le ferons pas) *que la seule hypothèse que f est de classe C^1 permet d'affirmer que sa série de Fourier est uniformément sommable.*

§ 5 : Séries de Fourier de quelques fonctions non continues.

Dans ce qui précède nous nous sommes limités à l'espace $\mathscr{C}^0_T(\mathbb{R},\mathbb{R})$ des fonctions continues. Cette hypothèse de continuité est souvent trop restrictive. *Nous dirons qu'une fonction périodique f est "continue avec sauts" si elle est définie et continue sauf en un nombre fini de points pour chaque période, et si en chacun des points où elle n'est pas continue, elle est continue à droite et a une limite à gauche.*

Exercice g: Démontrer qu'une telle fonction est bornée.

Nous noterons $\mathscr{C}^s_T(\mathbb{R},\mathbb{R})$ l'espace des fonctions (à valeurs réelles) "continues avec sauts" qui admettent 2π pour période. La définition du produit scalaire de $\mathscr{C}^0_T(\mathbb{R},\mathbb{R})$ s'étend à $\mathscr{C}^s_T(\mathbb{R},\mathbb{R})$ (puisque les fonctions bornées continues avec sauts sont intégrables), et dans $\mathscr{C}^s_T(\mathbb{R},\mathbb{R})$ nous pouvons considérer la suite $(E^n_{\mathbb{R}})_{n\in\mathbb{N}}$ des espaces de polynômes trigonométriques.

Nous pouvons donc reprendre pour les fonctions f de $\mathscr{C}^s_T(\mathbb{R},\mathbb{R})$ toutes les contructions du §2 ci-dessus. Les coefficients de Fourier réels de f seront définis par:

$$a_0(f) = <f,1> = \frac{1}{2\pi}\int_0^{2\pi} f(t)\,dt$$

$$a_1(f) = <f,\sqrt{2}\cos t> = \frac{1}{\pi\sqrt{2}}\int_0^{2\pi} f(t)\cos t\,dt \qquad b_1(f) = <f,\sqrt{2}\sin t> = \frac{1}{\pi\sqrt{2}}\int_0^{2\pi} f(t)\sin t\,dt$$

$$a_2(f) = <f,\sqrt{2}\cos 2t> = \frac{1}{\pi\sqrt{2}}\int_0^{2\pi} f(t)\cos 2t\,dt \qquad b_2(f) = <f,\sqrt{2}\sin 2t> = \frac{1}{\pi\sqrt{2}}\int_0^{2\pi} f(t)\sin 2t\,dt$$

......

$$a_k(f) = <f,\sqrt{2}\cos kt> = \frac{1}{\pi\sqrt{2}}\int_0^{2\pi} f(t)\cos kt\,dt \qquad b_k(f) = <f,\sqrt{2}\sin kt> = \frac{1}{\pi\sqrt{2}}\int_0^{2\pi} f(t)\sin kt\,dt$$

La série de Fourier de f est définie par

$$u_0 = a_0 \qquad \text{et pour } k\geq 1 \quad u_k(t) = a_k\sqrt{2}\cos kt + b_k\sqrt{2}\sin kt$$

Elle converge pour la norme N_2 vers la fonction f (donc la formule de Parseval reste vraie). La démonstration est dans le complément 1.

Exemple : Soit f la fonction de période 2π qui est définie par $(\forall x \in [0,2\pi[)\ f(x) = x$. Elle est "continue avec sauts"; les sauts étant aux points de la forme $2k\pi$. Ses coefficients de Fourier sont:

$$a_0 = \langle f,1\rangle = \frac{1}{2\pi}\int_0^{2\pi} t\,dt = \pi$$

$$a_n = \langle f,\sqrt{2}\cos nt\rangle = \frac{1}{\pi\sqrt{2}}\int_0^{2\pi} t\ \cos nt\,dt = 0 \qquad b_n = \langle f,\sqrt{2}\sin nt\rangle = \frac{1}{\pi\sqrt{2}}\int_0^{2\pi} t\ \sin nt\,dt = \frac{-\sqrt{2}}{n}$$

Sa série de Fourier est donc $((\pi, -2\sin t, \frac{-2}{2}\sin 2t, \ldots, \frac{-2}{n}\sin nt, \ldots))$. Elle est sommable au sens de la norme N_2, de somme f. La formule de Parseval s'écrit: $\pi^2 + \sum_{n\geq 1} \frac{2}{n^2} = (N(f))^2 = \frac{4\pi^2}{3}$.

Nous admettrons que, comme dans le cas des fonctions continues:

** Si en un point t_0 la fonction f est dérivable, alors la série numérique $((u_n(t_0)))_{n\in\mathbb{N}}$ est sommable de somme $f(t_0)$ (comme dans le cas des fonctions continues)*

** Si t_0 est un point de discontinuité, et si en ce point f a une dérivée à droite et une dérivée à gauche, alors la série numérique $((u_n(t_0)))_{n\in\mathbb{N}}$ est sommable de somme $\frac{1}{2}[\lim_{t\to t_0^+} f(t) + \lim_{t\to t_0^-} f(t)]$.*

Exercice g: Appliquer ces résultats à l'exemple ci-dessus; en déduire la somme de la série $((\frac{\sin kt}{k}))_{k\geq 1}$.

§ 6 : Familles de polynômes orthogonaux.

Résumons d'abord les constructions qui nous ont permis de définir les séries de Fourier. Nous avons un espace de fonctions E muni d'un produit scalaire. Dans cet espace nous avons une suite $(E_n)_{n\in\mathbb{N}}$ de sous-espaces de dimension finie. Cette suite est croissante (i.e. $(\forall n)\ E_n \subset E_{n+1}$). Alors à toute fonction f nous pouvons associer la suite $(f_n)_{n\in\mathbb{N}}$ formée des projections de f sur les E_n (ou, ce qui est équivalent, la série $((u_n))_{n\in\mathbb{N}}$ définie par $u_0 = f_0$ et (pour $n>0$) $u_n = f_n - f_{n-1}$).

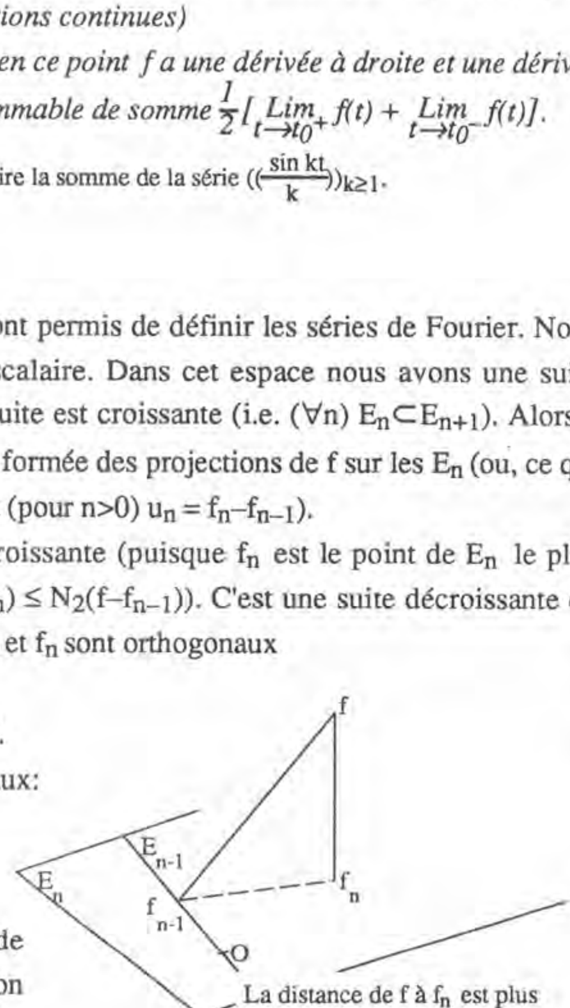
Il est clair que la suite $(N_2(f-f_n))_{n\in\mathbb{N}}$ est décroissante (puisque f_n est le point de E_n le plus proche de f, et puisque f_{n-1} est dans E_n, on a $N_2(f-f_n) \leq N_2(f-f_{n-1})$). C'est une suite décroissante de nombres positifs, donc elle a une limite ℓ. Puisque $f-f_n$ et f_n sont orthogonaux

$$[N_2(f)]^2 = [N_2(f_n)]^2 + [N_2(f-f_n)]^2$$

donc la suite $(N_2(f_n)^2)_{n\in\mathbb{N}}$ converge vers $N_2(f)^2 - \ell^2$.

Mais puisque les u_n sont deux à deux orthogonaux:

$$[N_2(f_n)]^2 = \sum_{i=0}^{n} [N_2(u_i)]^2$$



La distance de f à f_n est plus petite que la distance de f à f_{n-1}

Et ainsi la série $(((N_2(u_n)^2))_{n\in\mathbb{N}}$ est sommable de somme $N_2(f)^2 - \ell^2$. Dans le cas des séries de Fourier, on peut montrer que ℓ est nulle, ce qui nous donne la formule de Parseval.

Nous pouvons appliquer ce shéma de construction dans bien d'autres cas. Prenons par exemple l'espace $\mathscr{C}^0([a,b],\mathbb{R})$ des fonctions continues sur [a,b]. Nous le munirons d'un produit scalaire par la formule

$$\langle f,g\rangle = \int_a^b \varphi(t)\, f(t)\, g(t)\, dt$$

où φ est une fonction continue positive qui s'annule en au plus un nombre fini de points de [a,b] (voir chap.4, ex 15).

Soit E_n le sous-espace formé des polynômes de degré au plus n. Nous obtenons une suite croissante de sous-espaces, et nous pouvons répéter notre construction. Pour construire la suite $(f_n)_{n\in\mathbb{N}}$ nous construirons, par le procédé de Schmidt, une famille orthonomée $\{P_0, P_1, \ldots, P_n, \ldots\}$ telle que $(\forall n)$ $\{P_0, P_1, \ldots, P_n\}$ soit une base de E_n. Noter que P_n est de degré exactement n (il est dans E_n, donc il est de degré au plus n; et il n'est pas dans E_{n-1}, donc il est de degré supérieur à n–1). Alors nous aurons $u_n = \langle f,P_n\rangle P_n$ et $f_n = \sum_{i=0}^{n} \langle f,P_i\rangle P_i$.

Comme dans le cas des séries de Fourier la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f pour la norme N_2 (autrement dit $\ell = 0$). En effet pour toute fonction f continue sur [a,b] , et pour tout ε (>0) il existe un polynôme Q tel que $N_\infty(f-Q) \leq \varepsilon \left[\int_a^b \varphi(t)\, dt\right]^{-1/2}$ (c'est le théorème de Weierstrass dont une démonstration se trouve dans le complément 1). Soit $n = d°Q$, alors (puisque P_n est le point de E_n le plus proche de f au sens de N_2) $N_2(f-P_n) \leq N_2(f-Q) \leq \left[\int_a^b \varphi(t)\, [N_\infty(f-Q)]^2\, dt\right]^{1/2} = \varepsilon$. Nous avons ainsi démontré que $(\forall\varepsilon)$ il existe un terme de la suite (décroissante positive) $(N_2(f-P_n))_{n\in\mathbb{N}}$ qui est inférieur à ε; ce qui prouve que cette suite converge vers 0.

Nous avons donc une formule de Parseval, qui s'écrit:

$$[N_2(f)]^2 = \sum_{n=0}^{\infty} [N_2(u_n)]^2 = \sum_{n=0}^{\infty} [\langle f,P_n\rangle]^2$$

Ces constructions ont de multiples applications (résolutions d'équations différentielles, intégration numérique,...) qu'il serait trop long de développer ici. Notons qu'il y a autant de familles de polynômes orthogonaux (sur [a,b]) que de produits scalaires (voir Ex.12 et 13).

Complément 1 : La convergence des séries de Fourier pour la norme N_2.

Nous allons démontrer que la série de Fourier d'une fonction f continue avec sauts, est sommable pour la norme N_2, et a pour somme f. Notre démonstration reposera sur le théorème suivant:

Théorème de Weierstrass: *Quelle que soit g continue sur [a,b], et quel que soit $\varepsilon > 0$, il existe un polynôme Q tel que* $N_\infty(g-Q) = \sup_{x\in[a,b]} |g(x) - Q(x)| \leq \varepsilon$.

Préliminaires à la démonstration: Nous nous placerons en fait sur le segment [0,1]. Voici comment on démontre le cas général à partir du cas des fonctions définies sur [0,1]. Si g: $[a,b] \to \mathbb{R}$, posons $f(t) = g(a+t(b-a))$; f est définie sur [0,1]. Si $N_\infty^{[0,1]}(f-P) \leq \varepsilon$, nous poserons $Q(x) = P\left(\frac{x-a}{b-a}\right)$; et nous aurons $N_\infty^{[a,b]}(g-Q) \leq \varepsilon$.

Pour approcher f: $[0,1] \to \mathbb{R}$, posons $P_n^f(t) = \sum_{p=0}^{n} C_n^p\, t^p(1-t)^{n-p} f\left(\frac{p}{n}\right)$. On va montrer que P_n convient pourvu que n soit grand. Regardons d'abord trois cas particuliers.

1. Si $f(t) = 1$, on a clairement $P_n^f(t) = 1$

2. Si $f(t) = t$, on a $P_n^f(t) = \sum_{p=0}^{n} C_n^p t^p(1-t)^{n-p} \frac{p}{n} = \sum_{p=1}^{n} \frac{(n-1)!}{(p-1)!(n-p)!} t^{p-1}(1-t)^{n-p} t$

d'où $P_n^f(t) = \sum_{q=0}^{n-1} \frac{(n-1)!}{q!(n-q-1)!} t^q(1-t)^{n-q-1} t = t.$

3. Si $f(t) = t^2$, on a $P_n^f(t) = \sum_{p=1}^{n} C_n^p t^p(1-t)^{n-p} \frac{p^2}{n^2} = \sum_{p=1}^{n} C_n^p \left[\frac{p(p-1)}{n(n-1)} + \frac{p(n-p)}{n^2(n-1)} \right] t^p(1-t)^{n-p}$

d'où $P_n^f(t) = \sum_{p=2}^{n} \frac{(n-2)!}{(p-2)!(n-p)!} t^p(1-t)^{n-p} + \sum_{p=1}^{n-1} \frac{1}{n} \frac{(n-2)!}{(p-1)!(n-p-1)!} t^p(1-t)^{n-p}$

$$P_n^f(t) = \sum_{q=0}^{n-2} C_{n-2}^q t^q(1-t)^{n-2-q} t^2 + \sum_{q=0}^{n-2} \frac{1}{n} C_{n-2}^q t^q(1-t)^{n-q-2} t(1-t) = t^2 + \frac{1}{n} t(1-t).$$

<u>Démonstration</u>: Posons $\sup_{t\in[0,1]} |f(t)| = N_\infty(f)$. Puisque f est continue sur $[0,1]$, elle est uniformément continue; c'est-à-dire que $\forall \varepsilon > 0, \exists \alpha > 0$ tel que $|t-t'| \leq \alpha$ entraîne $|f(t)-f(t')| \leq \frac{\varepsilon}{2}$. On a alors :

$$f(t)-P_n^f(t) = \left[\sum_{p=0}^{n} C_n^p t^p(1-t)^{n-p}\right] f(t) - \sum_{p=0}^{n} C_n^p t^p(1-t)^{n-p} f(\tfrac{p}{n}) = \sum_{p=0}^{n} C_n^p t^p(1-t)^{n-p} (f(t)-f(\tfrac{p}{n}))$$

$$= \sum_{|t-\frac{p}{n}|\leq\alpha} C_n^p t^p(1-t)^{n-p} (f(t)-f(\tfrac{p}{n})) + \sum_{|t-\frac{p}{n}|>\alpha} C_n^p t^p(1-t)^{n-p} (f(t)-f(\tfrac{p}{n}))$$

d'où $$|f(t)-P_n^f(t)| \leq \sum_{|t-\frac{p}{n}|\leq\alpha} C_n^p t^p(1-t)^{n-p}\frac{\varepsilon}{2} + \sum_{|t-\frac{p}{n}|>\alpha} C_n^p t^p(1-t)^{n-p}\, 2\, N_\infty(f)$$

d'où $$|f(t)-P_n^f(t)| \leq \sum_{p=0}^{n} C_n^p t^p(1-t)^{n-p}\frac{\varepsilon}{2} + \sum_{|t-\frac{p}{n}|>\alpha} C_n^p t^p(1-t)^{n-p}\, 2\, N_\infty(f)$$

d'où $$|f(t)-P_n^f(t)| \leq \frac{\varepsilon}{2} + 2N_\infty(f) \sum_{|t-\frac{p}{n}|>\alpha} C_n^p t^p(1-t)^{n-p} \frac{(t-\frac{p}{n})^2}{\alpha^2} \qquad \left(\text{car } \frac{(t-\frac{p}{n})^2}{\alpha^2} > 1\right)$$

d'où $$|f(t)-P_n^f(t)| \leq \frac{\varepsilon}{2} + \frac{2N_\infty(f)}{\alpha^2} \sum_{p=0}^{n} C_n^p t^p(1-t)^{n-p}(t^2-2t\frac{p}{n}+\frac{p^2}{n^2})$$

d'où, d'après les exemples ci-dessus :

$$|f(t)-P_n^f(t)| \leq \frac{\varepsilon}{2} + \frac{2N_\infty(f)}{\alpha^2} \left[t^2-2t\times t+(t^2+\frac{1}{n}t(1-t))\right]$$

$$|f(t)-P_n^f(t)| \leq \frac{\varepsilon}{2} + \frac{2N_\infty(f)}{\alpha^2} \frac{1}{n} t(1-t) \leq \frac{\varepsilon}{2} + \frac{N_\infty(f)}{2n\alpha^2} \quad (\text{car pour } t\in[0,1]:\ |t(1-t)| \leq \frac{1}{4}).$$

Il est clair que, pour n assez grand, $\frac{N_\infty(f)}{2n\alpha^2} \le \frac{\varepsilon}{2}$, donc $|f(t)-P^f_n(t)| \le \varepsilon$. Et ceci quel que soit t, donc nous avons construit P_n tel que $N_\infty(f-P_n) \le \varepsilon$.

Convergence de la série de Fourier d'une fonction continue.

Remarquons d'abord que la suite $(N_2(f-f_n))_{n\in\mathbb{N}}$ est décroissante et positive (cf. début du §6). Pour démontrer qu'elle converge vers 0, nous allons donc montrer que $(\forall\varepsilon>0)$ il existe un n tel que $N_2(f-f_n) \le \varepsilon$. Pour cela nous montrerons qu'il existe un n et un élément g de E_n, tels que $N_2(f-g) \le \varepsilon$. Nous traiterons d'abord le cas des fonctions paires, puis celui des fonctions impaires (ce qui suffit puisque toute fonction est somme d'une fonction impaire et d'une fonction paire).

Cas où f est paire: Considérons la fonction $F = f\circ\text{Arccos}$. Elle est continue sur $[-1,1]$. Appliquons lui le théorème de Weierstrass. Quel que soit $\varepsilon>0$, il existe un polynôme P tel que $\sup_{x\in[-1,1]} |F(x) - P(x)| \le \varepsilon$.

Pour $t\in[0,\pi]$, $F(\cos(t)) = f(t)$ (par définition de F); cette égalité reste vraie sur $[-\pi,0]$ puisque les deux fonctions sont paires; elle est ainsi vraie sur une période, donc vraie pour tout t). La fonction $P\circ\cos$ est un polynôme trigonométrique. Et nous avons:

$$\sup_{t\in\mathbb{R}} |f(t) - P(\cos t)| = \sup_{x\in[-1,1]} |F(x) - P(x)| \le \varepsilon$$

Cas où f est impaire: Le nombre ε étant donné, nous appliquerons d'abord le théorème de Weierstrass, à la fonction $f|[-\pi,\pi]: [-\pi,\pi] \to \mathbb{R}$. Nous obtiendrons un polynôme Q tel que $\sup_{t\in[-\pi,\pi]} |f(t) - Q(t)| \le \varepsilon/3$.

Ce polynôme n'a aucune raison d'être impair, nous le remplacerons donc par le polynôme Q_1 défini par $Q_1(t) = \frac{1}{2}[Q(t) - Q(-t)]$. Pour tout t nous aurons

$$|f(t) - Q_1(t)| = \frac{1}{2}|[f(t) - Q(t)] - [f(-t) - Q(-t)]| \le \frac{1}{2}|f(t) - Q(t)| + \frac{1}{2}|f(-t) - Q(-t)]| \le \varepsilon/3.$$

C'est à dire $\sup_{t\in[-\pi,\pi]} |f(t) - Q_1(t)| \le \varepsilon/3$.

Nous poserons alors $Q_2(t) = Q_1(t) - \frac{t}{\pi} Q_1(\pi)$, de façon à obtenir un polynôme nul en 0 et en π.

Nous pourrons écrire $Q_2(t) = \sin t\, Q_3(t)$, où Q_3 est une fonction continue et paire sur $[-\pi,\pi]$; et nous allons appliquer à Q_3 ce que nous avons déjà démontré pour les fonctions paires. Il existe donc un polynôme P tel que $\sup_{t\in[0,\pi]} |Q_3(t) - P(\cos t)| \le \varepsilon/3$.

Exercice h: Justifier le fait que Q_3 est une fonction continue et paire.

Puisque $\sup_{t\in[-\pi,\pi]} |Q_1(t) - Q_2(t)| \le \varepsilon/3$, nous aurons $\sup_{t\in[-\pi,\pi]} |f(t) - Q_2(t)| \le 2\varepsilon/3$.

Puisque la fonction sinus est à valeurs dans $[-1,1]$

$$\sup_{t\in[0,\pi]} |Q_2(t) - P(\cos t)\sin t| = \sup_{t\in[0,\pi]} |Q_3(t)\sin t - P(\cos t)\sin t| \le \sup_{t\in[0,\pi]} |Q_3(t) - P(\cos t)| \le \varepsilon/3.$$

D'où (puisque les fonctions f et $(t \to P(\cos t)\sin t)$ sont impaires et de période 2π)

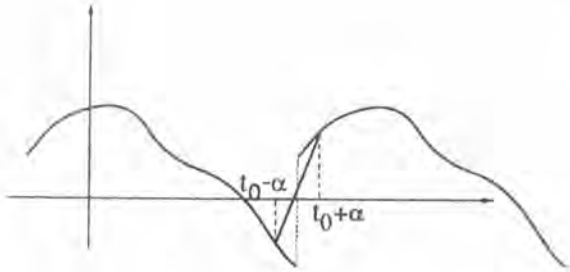
$$\sup_{t\in\mathbb{R}} |f(t) - P(\cos t)\sin t| = \sup_{t\in[-\pi,\pi]} |f(t) - P(\cos t)\sin t|$$
$$\le \sup_{t\in[-\pi,\pi]} |f(t) - Q_2(t)| + \sup_{t\in[0,\pi]} |Q_2(t) - P(\cos t)\sin t| \le \varepsilon.$$

Convergence de la série de Fourier d'une fonction continue avec sauts.

Pour simplifier l'écriture nous supposerons que f n'a qu'un seul point de discontinuité, noté t_0. Choisissons α $(\pi>\alpha>0)$, et définissons une fonction f_1 par

$$\begin{cases} \text{si } t \text{ appartient à un des intervalles } [t_0+\alpha+2k\pi, t_0-\alpha+(2k+1)\pi] \text{ alors } f_1(t) = f(t) \\ \text{sur chacun des } [t_0-\alpha+2k\pi, t_0+\alpha+2k\pi] \ f_1 \text{ est affine et coïncide avec f aux extrémités} \end{cases}$$

Cette fonction est continue, et (si M est un majorant des valeurs de $|f|$)



$$[N_2(f-f_1)]^2 = \int_{t_0-\alpha}^{t_0+\alpha} [f(t) - f_1(t)]^2\,dt$$

$$\leq \int_{t_0-\alpha}^{t_0+\alpha} 4M^2\,dt = 8M^2\alpha$$

Il en résulte que $(\forall\varepsilon>0)$ pour α assez petit, on a $N_2(f-f_1) \leq \varepsilon/2$. Mais puisque f_1 est continue on pourra trouver un polynôme trigonométrique G tel que $N_2(f_1-g) \leq \varepsilon/2$, et alors on aura $N_2(f-g) \leq \varepsilon$.

Complément 2 : L'équation des cordes vibrantes.

Ou comment les séries trigonométriques permettent de comprendre le fonctionnement d'une corde de guitare.

Partons d'un problème de mécanique. Une corde est fixée en ses deux extrémités A et B (Nous poserons AB = ℓ), elle est tendue entre ces deux points. Au repos elle a une position d'équilibre que nous supposerons être la ligne droite AB, ce qui signifie que nous faisons abstraction de la pesanteur. Si nous écartons la corde de cet équilibre, et si nous la lachons elle se met à vibrer. Notons que c'est exactement le geste du guitariste; dans un piano on se sert d'un marteau pour écarter la corde de sa position d'équilibre, puis on la laisse vibrer, le problème est donc mathématiquement à peu près identique; un violoniste frotte la corde avec un archet, celle ci n'est donc jamais soumise exactement aux lois que nous allons décrire, le mouvement est continuellement entretenu.

La mise en équation: Choisissons un système d'axes orthonormés (A,x,y,z) tel que l'axe Ox soit confondu avec AB. Décrire le mouvement c'est pour tout point M_x de la corde (occupant à l'équilibre la position (x,0,0)), dire sa position à l'instant t. Nous supposerons que cette position est de la forme (x,y(x,t),z(x,t)), c'est à dire que M_x reste dans un plan perpendiculaire à AB.

Il nous faut donc calculer les fonction y(x,t) et z(x,t), c'est à dire une fonction de $\mathbb{R}^2(x,t)$ (définie pour $x_A=0<x<x_B=\ell$, et $t>0$) dans $\mathbb{R}^2(y,z)$. Nous supposerons bien entendu que cette fonction est différentiable, ou plus précisément de classe C^k, avec k assez grand pour que nous puissions faire les calculs qui suivent.

Soit un petit morceau MM' de la corde; on note x l'abscisse de M et x+dx celle de M', donc la longueur MM' est égale à dx. Sa masse est μdx (ou l'on note μ la masse de l'unité de longueur de la corde, celle-ci étant bien entendu supposée homogène).

Les forces appliquées, à l'instant t, à ce morceau de corde sont au nombre de deux, l'une $\vec{F}_M$ est appliquée en M , l'autre $\vec{F}_{M'}$ en M'. Elles ont même intensité τ (c'est la tension de la corde), elles sont tangentes en M et M' à la corde, et sont "de sens contraire" (c'est à dire que si l'arc MM' était un morceau de droite, leur somme serait nulle). Donc $\vec{F}_M = \tau\,\vec{T}_M$, où $\vec{T}_M$ est un vecteur unitaire tangent à la courbe $x \to (x,y(x,t),z(x,t))$.Donc (en notant $\frac{\partial y}{\partial x}(x,t)$ et $\frac{\partial z}{\partial x}(x,t)$ les dérivées partielles de y et z par rapport à x) les coordonnées de $\vec{F_M}$ sont:

$$-\vec{F}_M = \tau\, \vec{T}_M = \tau \left(\frac{1}{\sqrt{1+(\frac{\partial y}{\partial x}(x,t))^2+(\frac{\partial z}{\partial x}(x,t))^2}} ; \frac{\frac{\partial y}{\partial x}(x,t)}{\sqrt{1+(\frac{\partial y}{\partial x}(x,t))^2+(\frac{\partial z}{\partial x}(x,t))^2}} ; \frac{\frac{\partial z}{\partial x}(x,t)}{\sqrt{1+(\frac{\partial y}{\partial x}(x,t))^2+(\frac{\partial z}{\partial x}(x,t))^2}} \right)$$

Remplaçons dans cette formule x par x+dx, nous aurons $\vec{F_{M'}}$. Et il nous faut calculer $\vec{F_{M'}} + \vec{F_M}$.

Nous supposerons que $\frac{\partial y}{\partial x}(x,t)$ et $\frac{\partial z}{\partial x}(x,t)$ sont assez petits pour que l'on puisse assimiler les dénominateurs à 1 (ce qui signifie que l'angle de $\vec{T}_M$ et de la droite AB reste petit). Alors les coordonnées de $\vec{F_{M'}} + \vec{F_M}$ sont :

$$(0 ; \tau [\frac{\partial y}{\partial x}(x+dx,t) - \frac{\partial y}{\partial x}(x,t)] ; \tau [\frac{\partial z}{\partial x}(x+dx,t) - \frac{\partial z}{\partial x}(x,t)])$$

Puisque dx est petit nous les assimilerons au premier terme de leur développement limité (en dx), c'est à dire à:

$$(0 ; \tau \frac{\partial^2 y}{\partial x^2}(x,t)\, dx ; \tau \frac{\partial^2 z}{\partial x^2}(x,t)\, dx)$$

Puisque les coordonnées de l'accélération de M sont $(0; \frac{\partial^2 y}{\partial t^2}(x,t) ; \frac{\partial^2 z}{\partial t^2}(x,t))$, l'équation fondamentale de la dynamique s'écrit alors:

$$\frac{\partial^2 y}{\partial t^2}(x,t)\, \mu = \tau \frac{\partial^2 y}{\partial x^2}(x,t) \quad \text{et} \quad \frac{\partial^2 z}{\partial t^2}(x,t)\, \mu = \tau \frac{\partial^2 z}{\partial x^2}(x,t)$$

Les deux fonctions coordonnées y et z sont donc régies par la même équation aux dérivées partielles; nous poserons $v^2 = \frac{\tau}{\mu}$ et nous chercherons les fonctions u (de x et de t) telles que:

$$(1) \qquad \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2}(x,t) = \frac{\partial^2 u}{\partial x^2}(x,t)$$

En n'oubliant pas que les solutions qui nous intéressent doivent vérifier les "conditions initiales" suivantes:

α) A l'instant t=0, on lache la corde dans une certaine position. Autrement dit $(\forall x)$ $u(x,0) = u_0(x)$, où u_0 est une fonction donnée. De plus à l'instant t=0, la vitesse de tous les points de la corde est nulle, ce qui s'exprime par $(\forall x)\ \frac{\partial u}{\partial t}(x,0) = 0$

β) Quelque soit t, $u(0,t) = u(\ell,t) = 0$ (ce qui exprime le fait que les deux extrémités de la corde sont fixes).

En fait seules les valeurs correspondant à $0 < x < \ell$ et $t>0$ nous intéressent.

Etude mathématique de l'équation des cordes vibrantes.

A) Etude des solutions stationnaires: Nous dirons qu'une solution de l'équation (1) est stationnaire si elle s'écrit comme un produit $u(x,t) = f(x)\, g(t)$, où f et g sont de classe C^2.

Exercice i: 1) Considérons 4 fonctions $\alpha,\beta,\gamma,\delta$ définies sur $\mathbb{R}$ et non identiquement nulles, on suppose que $(\forall x)(\forall t)$ $\alpha(x)\beta(t) = \gamma(x)\delta(t)$, montrer qu'il existe un nombre λ tel que $\alpha = \lambda\gamma$ et $\beta = \lambda\delta$. En déduire que, si $u(x,t) = f(x)\, g(t)$ est une solution stationnaire (non nulle), il existe un nombre λ tel que $(\forall x)(\forall t)$ $f''(x) = \lambda f(x)$ et $g''(t) = \lambda v^2 g(t)$.

2) Déterminer les valeurs possibles de f et g si λ est nul; démontrer alors que la condition initale β ne peut pas être vérifiée (sauf $f = 0$ donc $u = 0$).

3) Démontrer de même que λ ne peut pas être strictement positif.

4) Déterminer les valeurs possibles de f et g si λ est strictement négatif. En utilisant la condition initiale β montrer que les seules valeurs possibles de λ sont $-k^2\frac{\pi^2}{\ell^2}$, où k est un entier (positif). En déduire que les seules solutions stationnaires sont les fonctions $(t,u) \to \ell \sin(\frac{k\pi}{\ell}x)\sin(\frac{k\pi}{\ell}v(t-t_0))$.

B) <u>Démonstration de l'existence d'une solution</u> : Sur l'espace E des fonctions de période 2ℓ on met le produit scalaire:

$$<\alpha,\beta> = \int_{-\ell}^{\ell} \alpha(x)\,\beta(x)\,dx$$

Exercice j: Trouver des scalaires k_0 et k_1 tels que les fonctions $\varepsilon_0: x \to k_0$ et $(\forall p = 1,2,..,n,\ldots)$ $\varepsilon_p: x \to k_1 \cos\frac{\pi}{\ell}px$ et $\varphi_p: x \to k_1 \sin\frac{\pi}{\ell}px$ forment un système orthonormé dans E.

Nous supposerons que u_0 est une fonction de classe C^2 (définie sur $[0,\ell]$). Nous définissons une fonction $\overline{u}_0$ impaire de période 2ℓ, en posant

$$\begin{cases} \text{Si } 0 \le x \le \ell : \overline{u}_0(x) = u_0(x) \\ \text{Si } -\ell \le x \le 0 : \overline{u}_0(x) = -u_0(-x) \end{cases}$$

On pose $a_p = <\overline{u}_0,\varphi_p>$ (on notera que $(\forall p)$ $<\overline{u}_0,\varepsilon_p> = 0$); les a_p sont les coefficients de Fourier de $\overline{u}_0$. Et on considère la série de fonctions $((U_p))_{p\in\mathbb{N}}$ définie par

$$U_p(x,t) = a_p\,\varphi_p(x)\cos(\frac{\pi v}{\ell}pt)$$

Exercice k: 1) On suppose que $a_p = o(1/p^2)$. Démontrer que pour toute valeur de (x,t), la série numérique $((U_p(x,t)))_{p\ge1}$ est sommable. On note $U(x,t)$ sa somme.

2) Soit R un rectangle dans le plan des (x,t). On note F l'espace des fonctions continues sur R, et on le munit de la norme uniforme. On note $U_p|R$ la restriction de U_p à R. Majorer la norme de $U_p|R$. Montrer que, si $a_p = o(1/p^2)$, la série $((U_p|R))_{p\ge1}$ est uniformément sommable. Qu'en résulte-t-il pour la fonction U ?

Exercice l: 1) Soit $[a,b]$ un intervalle de $\mathbb{R}$, on note G l'espace des fonctions de classe C^1 sur $[a,b]$, et on le munit de la norme $v_\infty(f) = |f(a)| + N_\infty(f')$. On fixe t et on définit des fonctions V_p dans G en posant $V_p(x) = U_p(x,t)$. Majorer $v_\infty(V_p)$. Montrer que, si $a_p = o(1/p^3)$, la série $((V_p))_{p\ge1}$ est sommable pour v_∞, et que U a en chaque point une dérivée partielle par rapport à x.

2) On suppose que $a_p = o(1/p^3)$, en s'inspirant de l'exercice k, montrer que la dérivée partielle $\frac{\partial U}{\partial x}$ est continue.

Ainsi sous l'hypothèse que $a_p = o(1/p^3)$, la fonction U a une dérivée partielle en x qui est continue. Cette dérivée est obtenue en dérivant terme à terme la série qui donne U. On démontrerait de même que U a une dérivée partielle en t qui est continue et que l'on obtient en dérivant la série terme à terme. Si $a_p = o(1/p^4)$, on démontre de même que la fonction U est de classe C^2, et que

$$\frac{\partial^2 U}{\partial x^2} = \sum_{p\ge1}\frac{\partial^2 U_p}{\partial x^2} \quad \text{et} \quad \frac{\partial^2 U}{\partial t^2} = \sum_{p\ge1}\frac{\partial^2 U_p}{\partial t^2}$$

Puisque les U_i sont des solutions de l'équation, il en résulte immédiatement que U est une solution de l'équation. On vérifie facilement que $(\forall t)$ $U(0,t) = U(\ell,t) = 0$, et $(\forall x)$ $U(x,0) = \overline{u}_0(x)$ et $\frac{\partial U}{\partial t}(x,0) = 0$).

Nous avons donc construit une solution correspondant à la donnée initiale que nous avions fixée.

Notre hypothèse principale est que les coefficients de Fourier de $\bar{u}_0$ tendent suffisamment vite vers 0. Cette hypothèse est vérifiée dés que $\bar{u}_0$ est de classe C^4.

Exercice m: Etablir une relation entre $<\bar{u}_0,\varphi_p>$ et $<\bar{u}_0',\varepsilon_p>$. Etablir une relation entre $<\bar{u}_0,\varepsilon_p>$ et $<\bar{u}_0',\varphi_p>$. En déduire que si $\bar{u}_0$ est de classe C^4, $a_p = o(1/p^4)$.

<u>Commentaires</u>: Pour être certain que la solution ainsi construite est physiquement valable, il resterait à montrer que c'est la seule solution vérifiant les conditions initiales; nous ne le ferons pas. Cette théorie date de la fin du 18-ième siécle, elle est pour beaucoup dans l'introduction des séries de Fourier. Elle montre que la vibration est somme d'une vibration sinusoïdale fondamentale (c'est le premier terme de la série) de fréquence $\frac{v}{2\ell}$, appelée fréquence fondamentale, et de vibrations sinusoïdales (en nombre infini) dont les fréquences sont les multiples de la fréquence fondamentale; ce sont les harmoniques. Musicalement la fréquence fondamentale détermine la hauteur du son, les harmoniques définissent le timbre. Actuellement on est capable, par des procédés électroniques, de déterminer les intensités des harmoniques d'un son musical (c'est à dire sa série de Fourier, ou les coefficients a_p ci-dessus). On peut alors reconstituer par addition de sons sinusoïdaux, un son ayant même série de Fourier (jusqu'à un certain rang, puisqu'il n'est pas question de traiter expérimentalement des sommes infinies). Et l'oreille reconnait alors le timbre du son initial. Autrement dit la série de Fourier a été introduite pour les besoins du calcul, mais en fait c'est elle qu'on entend.

<u>Critique de la mise en équation</u>:

La mise en équation est parsemée d'approximations diverses: "nous supposerons que M_x reste dans un plan parallèle au plan Oyz", "puisque dx est petit nous les assimilerons au premier terme de leur développement limité",... L'équation du mouvement est beaucoup plus compliquée que celle que nous avons écrite; nous n'avons pas tenu compte des déplacements de la corde dans la direction de Ox; nous n'avons pas tenu compte des phénomènes d'élasticité...Elle pourrait s'écrire shématiquement

$$\frac{\partial^2 u}{\partial t^2}(x,t) = \mathcal{F}\left(t,x,u(x,t),\frac{\partial u}{\partial t}(x,t),\frac{\partial u}{\partial x}(x,t),\frac{\partial^2 u}{\partial x^2}(x,t),\dots\right)$$

A chaque instant l'accélération du point M_x dépent de x, de t, et des dérivées partielles de u. Une telle formule serait évidemment plus difficile à établir, mais surtout elle conduirait à une équation aux dérivées partielles que nous ne saurions absolument pas résoudre. La fonction $\mathcal{F}$ est une "bonne fonction", c'est à dire une fonction différentiable; pour obtenir l'équation $\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$, nous l'avons remplacée par sa différentielle pour $u = 0$; c'est ce qu'on appelle "linéariser l'équation". On le constate par exemple lorsque dans la formule donnant la différence des forces, on ne tient compte que du développement à l'ordre 1.

C'est une attitude classique en analyse de remplacer ainsi un problème que l'on ne sait pas résoudre, par un problème voisin (très souvent obtenu en remplaçant une fonction par sa différentielle - c'est la linéarisation - ou par son développement d'ordre 2) . Mais peut-on se permettre de remplacer ainsi un problème par un problème voisin ? Deux attitudes sont possibles.*

Cette théorie fut inventée à la fin du XVIII-ème siècle; elle permit de donner une explication aux règles de l'accord des instruments à corde (et aussi des instruments à vent, car la vibration d'un colonne d'air obéit à la même équation) dont l'étude empirique remonte à l'antiquité (Pythagore). Sur le plan de la compréhension mécanique du phénomène, l'avancée était telle qu'aucun expérimentateur ne pouvait mettre en doute sa validité.

* Voir aussi le complément du chapitre 20.

C'est la conformité aux lois préétablies de l'harmonie qui servait en quelque sorte de justification ultime. C'est le type même de l'explication physique; elle permet de donner une cohérence à des phénomènes expérimentaux connus. Par exemple si deux cordes (de longueurs, de natures, de tensions différentes) donnent la même note, et si on les raccourcit dans les mêmes proportions, elles donnent encore la même note; si on les raccourcit de moitié on obtient l'octave de la note initiale; si elles donnaient le même do, et si on les raccourcit de 1/3, on obtient le même sol.

Sur un plan purement mathématique cette approximation est inacceptable. Il nous faut alors poser le problème suivant: Les équations voisines (en un sens qu'il nous faudrait préciser) de $\frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$ *, ont elles des solutions qui ressemblent à celles que nous venons de trouver. Par exemple est ce qu'elles sont périodiques en t ? Est ce que la formule donnant leur période est voisine de celle que nous avons trouvée ?... Il est bien entendu totalement exclu de résoudre de tels problèmes avec les moyens dont nous disposons.*

Exercices sur le chapitre 15

* **Exercice 1** : a) Soit f la fonction de période 2π qui vaut t^2 pour $-\pi \leq t \leq \pi$. Calculer ses coefficients de Fourier.

b) Pourquoi la série de Fourier de f converge-t-elle vers 0 au point t = 0? Calculer la somme de $((\frac{(-1)^n}{n^2}))_{n>0}$.

c) Utiliser le fait que la série de Fourier converge pour la norme N_2, pour calculer la somme de la série $((\frac{1}{n^4}))_{n>0}$

** **Exercice 2** : Soit f la fonction de période 2π définie par $f(t) = |\sin t|$.

a) Calculer ses coefficients de Fourier

b) calculer $\sum\limits_{k=1}^{\infty} \frac{1}{4k^2-1}$, $\sum\limits_{k=1}^{\infty} \frac{(-1)^k}{4k^2-1}$ et $\sum\limits_{k=1}^{\infty} (\frac{1}{4k^2-1})^2$

* **Exercice 3** : Soit x(t) à valeurs réelles, de classe C^1, qui admet 2π pour période et telle que $\int_0^{2\pi} x(t)\, dt = 0$. Montrer que $\int_0^{2\pi} [x(t)]^2\, dt \leq \int_0^{2\pi} [x'(t)]^2\, dt$. Pour quelles fonctions x a-t-on égalité ?

* **Exercice 4** : Soit k un nombre réel, déterminer les fonctions x de classe C^2 qui admettent 2π pour période et telles que $(\forall t)$ $x''(t) + k\, x(t) = \sin 2t$. On cherchera d'abord les coefficients de Fourier de x.

** **Exercice 5** : On désigne par f_a la fonction de période 2π qui est égale à t-a en tout point de l'intervalle $[a, 2\pi+a]$.

a) Calculer les coeficients de Fourier de f_0. En déduire ceux de f_a.

b) Soit $(\forall n \geq 0)$ $\alpha_n = \frac{n}{n^2+1}$ et $(\forall n>0)$ $\beta_n = 0$. Montrer que ce sont les coefficients de Fourier d'une fonction g qui est la somme d'une fonction C^1 et d'une fonction de la forme λf_0.

*** **Exercice 6:** Soit f la fonction périodique de période 1 telle que sur [0,1] on ait $f(x) = x(1-x)$.

a) Calculer les coefficients de Fourier de f, et les termes u_n de la série de Fourier de f.

b) Soit $U_n(t) = \int_0^t u_n(\tau)\, d\tau$. Montrer que les U_n $(n \geq 1)$ sont les termes de la série de Fourier de la fonction F de période 1 qui sur [0,1] est définie par $F(t) = \int_0^t (f(\tau) - u_0)\, d\tau$.

c) En itérant cette intégration terme à terme, calculer les sommes des séries $((\frac{(-1)^n}{n^4}))_{n\geq 1}$ et $((\frac{1}{n^4}))_{n\geq 1}$.

d) Au moyen de la formule de Parseval, calculer les sommes des séries $((\frac{1}{n^6}))_{n\geq 1}$ et $((\frac{1}{n^8}))_{n\geq 1}$.

** **Exercice 7:** Soit f la fonction de période 2π qui vaut $\cos \frac{x}{2}$ pour $|x| \leq \pi$.

a) Calculer ses coefficients de Fourier, et les termes u_n de sa série de Fourier.

b) Que valent les sommes $\sum\limits_{n=0}^{\infty} u_n(0)$, $\sum\limits_{n=0}^{\infty} u_n(\pi)$, $\sum\limits_{n=0}^{\infty} u_n(2\pi/3)$?

c) Posons $U_n(t) = \int_0^t u_n(\tau)\,d\tau$. Montrer que la série $((U_n))_{n\geq 1}$ est sommable pour la norme N_2 ? Quelle est sa somme ?

* **Exercice 8**: Notons a_i et b_i les coefficients de Fourier de f. Quels sont les coefficients de Fourier de $f_a : t \to f(t+a)$? Quels sont les coefficients de Fourier de $\varphi_n : t \to f(t/n)$ $(n \in \mathbb{N})$?

* **Exercice 9**: a) Soit f la fonction de période 2π qui sur $[-\pi,\pi[$ est définie par $f(t) = e^{at}$. Calculer ses coefficients de Fourier complexes.

b) Calculer les coefficients de Fourier réels de $g: t \to f(t) + f(-t)$ et de $h: t \to f(t) - f(-t)$

c) En déduire les sommes des séries $((\frac{1}{1+n^2}))_{n\in\mathbb{N}}$ et $((\frac{(-1)^n}{1+n^2}))_{n\in\mathbb{N}}$.

** **Exercice 10**: On considère la fonction f de période 2π qui sur $[-\pi,\pi]$ est définie par $f(t) = t(\pi^2-t^2)$.

a) Calculer ses coefficients de Fourier.

b) Montrer que sa série de Fourier $((u_n))_{n\in\mathbb{N}}$ converge normalement pour N_∞.

c) Ecrire la formule de Parseval et en déduire la somme de la série $((\frac{1}{p^6}))_{p\geq 1}$.

d) Montrer que la série $((u'_n))_{n\in\mathbb{N}}$ converge normalement pour N_∞. Quelle est sa somme ?

e) Calculer les sommes des séries $((\frac{1}{p^2}))_{p\geq 1}((\frac{(-1)^n}{p^2}))_{p\geq 1}$.

*** **Exercice 11**: Soit f une fonction de période 2π. On se demande si l'équation différentielle $y'' + y = f$ a une solution de période 2π.

a) On suppose qu'il existe une telle solution g, calculer les coefficients de Fourier de g" en fonction de ceux de g. En déduire que les coefficients $a_1(f)$ et $b_1(f)$ sont nécessairement nuls. Dans la suite on admet que f vérifie cette condition nécessaire.

b) Soit f la fonction paire de période 2π définie par $f(t) = t$ si $t\in[0,\pi/2]$ et $f(t) = \pi-t$ si $t\in[\pi/2,\pi]$. Calculer sa série de Fourier.

c) Montrer qu'il existe une seule fonction de période 2π dont la série de Fourier possède les propriétés souhaitées. Calculer sa série de Fourier.

d) Montrer que cette fonction est de classe C^2, et est une solution du problème posé.

** **Exercice 12**: On note E l'espace des fonctions continues sur $[-1,1]$, muni du produit scalaire $\langle f,g\rangle = \int_{-1}^{1} \sqrt{1-t^2}\, f(t)\, g(t)\, dt$.

a) Montrer que les fonctions $P_n : x \to \frac{\sin((n+1)\text{Arccos}\, x)}{\sqrt{1-x^2}}$ et $Q_n : x \to \cos((n+1)\text{Arccos}\, x))$ sont des fonctions polynomiales de degré n. Calculer $P_0, P_1, \ldots, P_5$. Les P_n sont appelés les polynômes de Tchebichev.

b) Montrer que ces polynômes forment une base orthonormée de E.

** **Exercice 13**: On note E l'espace des fonctions continues sur $[-1,1]$, et on le munit du produit scalaire $\langle f,g\rangle = \int_{-1}^{1} f(t)\, g(t)\, dt$

a) On pose $P_n(t) = \frac{d^n}{dt^n}(x^n(1-x)^n)$. Montrer que la famille des polynômes P_n est orthogonale dans E (pour calculer $\langle P_n,P_p\rangle$, on fera des intégrations par parties). Les P_n sont appelés les polynômes de Legendre.

b) Calculer la norme de P_1, P_2, P_3, P_4, P_5.

Etude locale des fonctions de classe C^∞ de $\mathbb{R}^2$ dans $\mathbb{R}$.

§ 1 : Qu'est ce qu'une courbe ?

Considérons un plan $\mathcal{P}$ muni d'un repère orthonormé $(O,\vec{i},\vec{j})$. La droite parallèle au vecteur $\vec{v}\,(1;2)$ et passant par $M(1;3)$ peut être considérée:

** soit comme l'ensemble des points $M(x,y)$ tels que $2x - y = -1$.*

** soit comme l'image de $\mathbb{R}$ par l'application $t \to (1-t\,;\,3-2t)$.*

De même le cercle de centre $\Omega(1,2)$ et de rayon 1, peut être défini :

** soit comme l'ensemble des points $M(x,y)$ tels que $(x-1)^2 + (y-2)^2 = 1$.*

** soit comme l'image de $\mathbb{R}$ par l'application $t \to (1 + \cos t\,;\,2 + \sin t)$.*

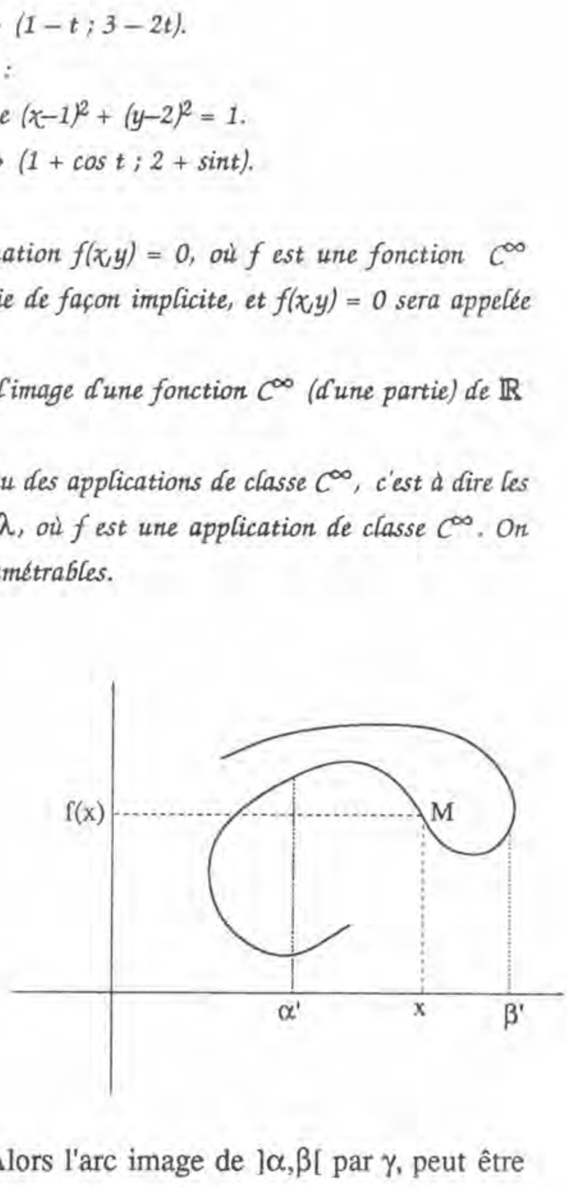
Les courbes planes peuvent ainsi apparaître

** soit comme l'ensemble des solutions d'une équation $f(x,y) = 0$, où f est une fonction C^∞ d'une partie de $\mathcal{P}$ dans $\mathbb{R}$. On dira alors que la courbe est définie de façon implicite, et $f(x,y) = 0$ sera appelée une équation cartésienne de cette courbe.*

** soit de façon paramétrique, c'est à dire comme l'image d'une fonction C^∞ (d'une partie) de $\mathbb{R}$ dans le plan $\mathcal{P}$.*

L'objet du présent chapitre est d'étudier les lignes de niveau des applications de classe C^∞, c'est à dire les sous-ensembles de $\mathcal{P}$ définis par des équations du type $f(x,y) = \lambda$, où f est une application de classe C^∞. On démontrera que 'dans les bons cas', ces lignes de niveau sont paramétrables.

Etude des paramétrages d'une courbe régulière.



Soit γ: $]a,b[\, \to P$ une courbe régulière*; on notera $x(t)$ et $y(t)$ les coordonnées de $\gamma(t)$. Considérons un point $M = \gamma(t_0)$, où le vecteur vitesse n'est pas parallèle à l'axe des ordonnées; c'est à dire tel que $x'(t_0) \neq 0$. Alors, puisque la fonction x' est continue, il existe un intervalle $]\alpha,\beta[$ contenant t_0, sur lequel x' reste non nulle, donc x strictement monotone. L'image de $]\alpha,\beta[$ par x est un intervalle $]\alpha',\beta'[$, et x définit une bijection de $]\alpha,\beta[$ sur $]\alpha',\beta'[$; nous noterons φ la fonction inverse; elle est aussi de classe C^1, puisque x' ne s'annule pas sur $]\alpha,\beta[$. Alors l'arc image de $]\alpha,\beta[$ par γ, peut être

* On dit que la courbe est régulière si la fonction γ est de clase C^1 et que le vecteur vitesse $\frac{\vec{d\gamma}}{dt}(t)$ ne s'annule pour aucune valeur de t. Pour plus de détails voir le chapitre 17.

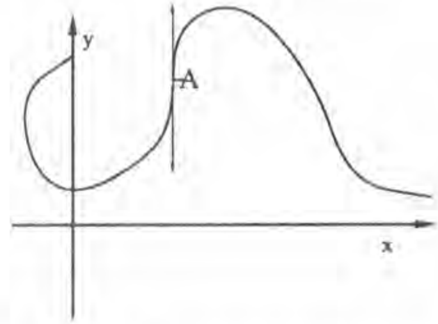
muni d'un nouveau paramétrage de classe C^1 : $\begin{cases}]\alpha',\beta'[\to P \\ \tau \to \gamma(\varphi(\tau)) \end{cases}$

Le point $\gamma(\varphi(\tau))$ a pour coordonnées $x(\varphi(\tau)) = \tau$, et $y(\varphi(\tau)) = f(\tau)$; ainsi cet arc s'écrit $y = f(x)$; il apparait comme le graphique de la fonction $f=y\circ\varphi$.

De la même façon si en un point M_1 le vecteur vitesse n'est pas parallèle à l'axe des abscisses (c'est à dire si y' n'est pas nulle en ce point), il existe un arc contenant ce point, qui peut s'écrire $x = g(y)$.

Exercice a : Marquer les arcs de la courbe ci-contre qui peuvent s'écrire sous la forme $y = f(x)$, et ceux qui peuvent s'écrire sous la forme $x = g(y)$.

Pourquoi un arc contenant le point A ne peut il être représenté sous la forme $y = f(x)$, où f est une fonction de classe C^1 ?



Sur la figure de l'exercice a, il est clair qu'en général une courbe ne peut pas être représentée globalement par une équation du type $y = f(x)$, ou $x = g(y)$. De telles représentations n'existent que localement. C'est la raison pour laquelle, dans toute la suite, nous considèrerons toujours des petits morceaux de courbes.

§ 2 : Difféomorphismes.

La notion de difféomorphisme

Soit U et V des ouverts de deux espaces vectoriels E et F de dimension finie, et f une bijection de U sur V; nous dirons que f est un difféomorphisme C^1 (resp. C^∞) de U sur V si:

**d'une part $f: U \to V$ est C^1 (resp. C^∞)*

**d'autre part $f^{-1}: V \to U$ est C^1 (resp. C^∞).*

On notera qu'il ne suffit pas que f soit C^1 (resp. C^∞), pour que f^{-1} le soit. Par exemple en dimension 1 l'application $f: x \to x^3$ est C^∞ de $U(=\mathbb{R})$ dans $V(=\mathbb{R})$, alors que $f^{-1}: x \to \sqrt[3]{x}$ n'est pas dérivable en 0.

On notera que s'il existe un difféomorphisme f d'un ouvert (non vide) U de E, sur un ouvert V de F, alors E et F ont même dimension, puisque $(\forall u\in U)$ $d_u f: E\to F$ est une application linéaire inversible (d'inverse $d_{f(u)}f^{-1}$).

Exercice b: Soit U l'ouvert de $\mathbb{R}^2$ formé des points $M(\rho,\theta)$ tels que ($0<\rho$ et $-\pi<\theta<\pi$). Soit V l'ouvert de $\mathbb{R}^2$ formé des points $P(x,y)$ tels que ($y\neq0$ ou $x>0$). Soit f l'application de U dans V qui à $M(\rho,\theta)$ associe $P(\rho \cos\theta,\rho \sin\theta)$. Montrer que f est bijective.

Montrer que $f^{-1}(x,y) = (\sqrt{x^2+y^2}, 2 \operatorname{Arccos}(\sqrt{\frac{1}{2}(1+\frac{x}{\sqrt{x^2+y^2}})}\,))$. Montrer que f est un difféomorphisme C^∞.

Exercice c: Soit U l'ouvert de $\mathbb{R}^3$ formé des points $M(\rho,\theta,z)$ tels que ($0<\rho$ et $-\pi<\theta<\pi$). Soit V l'ouvert de $\mathbb{R}^3$ formé des points $P(x,y,z)$ tels que ($y\neq0$ ou $x>0$).

Soit F l'application de U dans V qui à $M(\rho,\theta,z)$ associe $P(\rho \cos\theta,\rho \sin\theta,z)$. Montrer que F est bijective. Trouver des formules donnant F^{-1} (cf Ex.b). Montrer que F est un difféomorphisme C^∞.

Exercice d: Soit U l'ouvert de $\mathbb{R}^3$ formé des points $M(r,\theta,\varphi)$ tels que ($0<r,-\pi<\theta<\pi,-\pi/2<\varphi<\pi/2$). Soit V l'ouvert de $\mathbb{R}^3$ formé des points $P(x,y,z)$ tels que ($y\neq0$ ou $x>0$).

Soit f l'application de U dans V qui à $M(\rho,\theta,\varphi)$ associe $P(\rho \cos\varphi \cos\theta,\rho \cos\varphi \sin\theta,\rho \sin\varphi)$. Montrer que f est bijective. Trouver des formules donnant f^{-1} (cf. Ex.b). Montrer que f est un difféomorphisme C^∞.

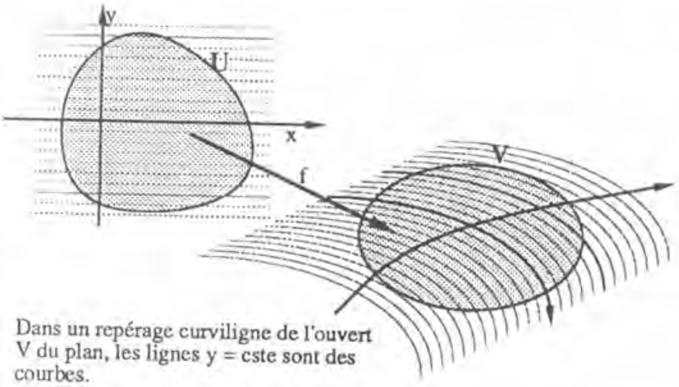
Image d'une courbe par un difféomorphisme

Soit $f: U \to V$ un difféomorphisme. Soit $\gamma:]\alpha,\beta[\to U$ une courbe de classe C^1. Alors $f\circ\gamma:]\alpha,\beta[\to V$ est aussi de classe C^1. De plus, pour tout t_0, la formule de composition des différentielles nous donne $d_{t_0} f\circ\gamma = d_{\gamma(t_0)} f \circ d_{t_0} \gamma$. C'est à dire $\frac{df\circ\gamma}{dt}(t_0) = d_{\gamma(t_0)} f\,(\frac{d\gamma}{dt}(t_0))$. Autrement dit le vecteur vitesse de $f\circ\gamma$ à l'instant t_0, est l'image par $d_{\gamma(t_0)} f$ du vecteur vitesse de γ à l'instant t_0. En particulier (puisque $d_{\gamma(t_0)} f$ est inversible) la courbe $f\circ\gamma$ est régulière dés que γ est régulière.

Le langage des coordonnées curvilignes

Si f est un difféomorphisme de $U(\subset \mathbb{R}^n)$ sur $V(\subset F)$, pour repérer un point v de V, on peut se donner le point $u = f^{-1}(v)$; c'est pourquoi f sera aussi appelé un repérage curviligne dans V. Les coordonnées de $f^{-1}(v)$ dans la base naturelle de $\mathbb{R}^n$ sont appelées les cooordonnées curvilignes de v (dans le repérage défini par f).

Expliquons le mot "curviligne": l'ensemble des points dont toutes les coordonnées sauf une - par exemple la première - sont égales à celles de $A(a_1,\ldots a_p)$, est l'image par f de l'ensemble des points de U de la forme $M(t,a_2,\ldots a_p)$; c'est la courbe $t \to f(t,a_2,\ldots a_p)$. Avec les coordonnées linéaires que l'on utilise habituellement, la même démarche aurait donné une droite.



Dans un repérage curviligne de l'ouvert V du plan, les lignes y = cste sont des courbes.

L'emploi de coordonnées curvilignes est fort répandu en physique, en mécanique, en géométrie,... Les trois systèmes de cooordonnées curvilignes les plus employés sont les coordonnées polaires du plan, les coordonnées cylindriques, et les coordonnées polaires de l'espace (voir les trois exercices ci-desssus).

Le théorème d'inversion locale (pour les applications du plan P dans $\mathbb{R}^2$).

Soit F une application de classe C^1 d'un ouvert U du plan P dans $\mathbb{R}^2$. Cherchons des conditions pour qu'il existe une application G de classe C^1 (d'une partie) de $\mathbb{R}^2$ dans P telle que $(\forall M\in U)$ $G\circ F(M) = M$. Cette relation implique que $(\forall M)$ $dG_{F(M)}\circ dF_M = Id$, donc elle n'est possible que si $(\forall M)$ dF_M est de rang 2, donc inversible. Inversement on démontre le résultat suivant connu sous le nom de "théorème d'inversion locale":

Si dF_M est inversible, il existe:

* *un ouvert ω de P contenant M et contenu dans U.*

* *un ouvert Ω de $\mathbb{R}^2$.*

Tels que $F|\omega$ soit un difféomorphisme de ω sur Ω. Si F est C^r $(r = 2,\ldots,\infty)$, l'application inverse G est aussi C^r. Nous admettrons ce résultat.

Exercice e: Soit F l'application de $\mathbb{C}$ (=P) dans $\mathbb{C}$ $(=\mathbb{R}^2)$ qui à z associe z^2. Montrer qu'elle est différentiable en tout point z_0, et que sa différentielle en z_0 est l'application qui à h $(\in \mathbb{C})$ associe $2z_0h$. Au point 0 le théorème ci-dessus ne s'applique donc pas. Montrer qu'il n'existe aucun disque de centre 0 sur lequel F soit injective. En tout point distinct de 0, le théorème s'applique. Autrement dit au voisinage de tout point de $\mathbb{C}^*$, on peut définir une fonction racine carrée qui est C^∞.

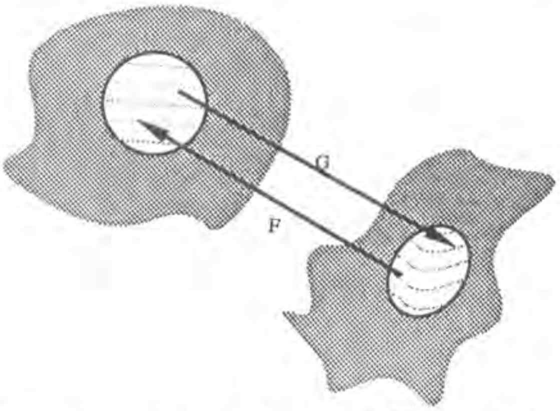
En général on ne peut pas prendre pour ω l'ouvert U tout entier. Ceci parait tout naturel si l'on remarque que, sur ω, F est injective (puisqu'elle est inversible), et que pour vérifier que la différentielle au point M soit inversible, il suffit de connaitre les valeurs de f dans un disque (aussi petit qu'on veut) centré en M; il est bien normal qu'une hypothèse qui n'utilise que les valeurs de la fonction au voisinage de M, ne puisse impliquer que F soit injective sur U tout entier.

§ 3 :Le théorème des fonctions implicites pour les fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$

Considérons maintenant un plan P muni d'un repère (non nécessairement orthonormé) $(O,\vec{i},\vec{j})$, et une application f de classe C^1 d'un ouvert U de P dans $\mathbb{R}$.

Supposons donné un point $M(x_M,y_M)$ de U où la dérivée partielle $\frac{\partial f}{\partial y}$ n'est pas nulle. Définissons $F: U \to \mathbb{R}^2$ par $F(x,y) = (x,f(x,y))$. Cette application est de classe C^1, et sa différentielle en M a pour matrice :

$$\begin{pmatrix} 1 & 0 \\ \frac{\partial f}{\partial x}f(M) & \frac{\partial f}{\partial y}f(M) \end{pmatrix}$$



Elle est donc inversible, et nous pouvons appliquer le théorème d'inversion locale au point M. Ainsi il existe un ouvert ω contenant M et un ouvert Ω contenant $F(M) = (x_M,f(M))$, tels que $F|\omega$ soit un difféomorphisme de ω sur Ω.

Soit $G: \Omega \to \omega$ le difféomorphisme inverse.

Etudions les lignes de niveau de f, au voisinage de M; plus précisément notons γ_n l'intersection de ω et de la ligne de niveau définie par $f(x,y) = n$. Dire qu'un point $P(x_P,y_P)$ est dans γ_n, équivaut à dire que $F(P) = (x_P,n)$; ou encore que F(P) est dans l'intersection δ_n de Ω et de la droite d'ordonnée n. Puisque $F|\omega$ et G sont inverses l'une de l'autre, ceci signifie encore que P est dans l'image du morceau de droite δ_n par le difféomorphisme G. Autrement dit γ_n est la courbe $G(\delta_n)$.

Choisisssons un paramétrage régulier de δ_n, par exemple $t \to (t,n)$; nous en déduisons un paramétrage $t \to G(t,n)$ de γ_n. Et puisque G est de classe C^1 (et même C^∞ si f est C^∞) le paramétrage ainsi obtenu est aussi C^1 (et même C^∞ si f est C^∞), et régulier. *Ainsi, au voisinage de M, les lignes de niveau de f sont des courbes régulières .*

Posons $G(X,Y) = (\alpha(X,Y),g(X,Y))$, on a $F(G(X,Y) = (X,Y) = (\alpha(X,Y),f(\alpha(X,Y),g(X,Y)))$. Donc $\alpha(X,Y) = X$. Il en résulte que le paramétrage $t \to G(t,n)$ est de la forme $t \to (t,g(t,n))$; autrement dit la ligne de niveau γ_n est, au voisinage de M, le graphe de la fonction $t \to g(t,n)$.

<u>Exemple</u>: Soit $f: \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^2 + 4y^2$. Plaçons nous au point M(1;1). La dérivée partielle $\frac{\partial f}{\partial y}(M)$ vaut 8; elle n'est pas nulle. Soit $F: \mathbb{R}^2 \to \mathbb{R}^2$ définie par $F(x,y) = (x,x^2+4y^2)$. L'application F est inversible au voisinage de M, son inverse est $G : (t,u) \to (t, \sqrt{\frac{1}{4}(u-t^2)}\ )$. Et la courbe de niveau de f passant par M (définie par l'équation $f(x,y)=f(M)=5$) est l'image par G des points de la forme (t,5); d'où le paramétrage $t \to (t,\sqrt{\frac{1}{4}(5-t^2)})$ de cette ligne de niveau.

Nous avons supposé que $\frac{\partial f}{\partial y}(M)$ est non nulle, et nous avons obtenu des paramétrages du type $t \to (t,y(t))$. Il est clair que si $\frac{\partial f}{\partial x}f(M)$ est non nulle, nous obtiendrons de façon analogue un paramétrage

$t \to (t,y(t))$. Il est clair que si $\frac{\partial f}{\partial x}f(M)$ est non nulle, nous obtiendrons de façon analogue un paramétrage du type $t \to (x(t),t)$. Ces paramétrages sont réguliers. Le théorème général des fonctions implicites s'écrit: *Si la différentielle $d_M f$ est non nulle (c'est à dire si l'une au moins des deux dérivées partielles de f en M est non nulle), alors au voisinage de M les lignes de niveau sont des courbes régulières.*

Que savons nous calculer explicitement ?

L'exemple ci-desssus rend très mal compte des calculs que l'on pourra faire en général. En effet nous avons pu trouver des formules pour G, ce qui n'est pratiquement jamais le cas. Généralement nous pourrons affirmer (grâce aux théorèmes énoncés ci-dessus) que G et le paramétrage existent, mais il nous sera impossible de les calculer explicitement. Ce seront (selon une vieille terminologie) des 'fonctions implicites'.

Exemple: Soit f: $\mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^3 + 4y^3+x^2+y$. Etudions la ligne de niveau passsant par $M(x_0=1;y_0=2)$. La dérivée $\frac{\partial f}{\partial y}(x,y) = 12y^2+1$, est non nulle en M. Nous pourrons donc trouver un paramétrage du type $t \to (t,\gamma(t))$, défini pour t voisin de 1. La fonction γ vérifie:

(pour tout t voisin de 1) $\quad f(t,\gamma(t)) = f(M)$

c'est à dire: $\quad t^3 + 4[\gamma(t)]^3 + t^2 + \gamma(t) = 36$

Nous savons que γ est C^∞, et que $\gamma(1) = 2$ (car le point $(1;\gamma(1))$ est le point M). Ceci nous permet de calculer les dérivées successsives de γ. En dérivant la relation ci-dessus, nous obtenons:

(pour tout t voisin de 1) $\quad 3t^2 + 12[\gamma(t)]^2\gamma'(t) + 2t + \gamma'(t) = 0.$

Pour t = 1, ceci nous donne $\quad 3 + 48\gamma'(1) + 2 + \gamma'(1) = 0$; soit $\gamma'(1) = -5/49$.

Puis nous pouvons dériver une seconde fois:

(pour tout t voisin de 1) $\quad 6t + 24\gamma(t)[\gamma'(t)]^2 + 12[\gamma(t)]^2\gamma''(t) + 2 + \gamma''(t) = 0$

Puisque l'on connait $\gamma(1)$ et $\gamma'(1)$, on en déduit $\gamma''(1)$; et ainsi de suite.

Exercice f: Continuer le calcul, et déterminer les dérivées de γ jusqu'à l'ordre 4.

Dans les exemples que nous rencontrerons, la fonction f est C^∞. Et nous pourrons ainsi calculer les dérivées successives de γ au point x_0 (abscisse du point M considéré). Mais c'est tout. Ceci nous permettra de calculer des développements limités de γ en x_0. En théorie on peut ainsi obtenir la série de Taylor de γ en x_0, mais ce calcul est rarement explicite. Encore faut il se rappeler que cette série peut ne pas converger, et que, si elle converge, sa somme n'est peut être pas γ. En fait sous des hypothèses un peu plus restrictives (que le fait que f est C^∞; et que nous ne pouvons expliciter faute d'avoir défini les 'séries entières à deux variables') on démontre que tout se passe bien, c'est à dire que la série de Taylor de γ converge vers γ (au voisinage de x_0). On remarquera l'analogie de cette situation avec les calculs de solutions d'équations différentielles du chap.12 §6. Dans les deux cas on accède à la fonction par sa série de Taylor, c'est à dire par le calcul formel de ses dérivées.

Variante du calcul: Reprenons l'exemple ci-dessus. Au lieu de rechercher les unes après les autres les dérivées de γ, nous pouvons rechercher directement un développement limité. Nous savons que γ est C^∞, donc elle a un développement d'ordre 2 (par exemple) $\gamma(x) = a + b(x-1) + c(x-1)^2 + o(x-1)^2$. Où $\gamma(1) = 2 = a$. Portons cette expression dans f. Nous aurons

$$36 = x^3 + 4(2 + b(x-1) + c(x-1)^2 + o(x-1)^2)^3 + x^2 + 2 + b(x-1) + c(x-1)^2 + o(x-1)^2$$

En regroupant tous les termes qui sont $o(x-1)^2$, nous obtenons

$$36 = 36 + (5+ 49b)(x-1) + (4 + 24b^2 + 49c)(x-1)^2 + o(x-1)^2$$

Puisque la fonction constante $x \to 36$ n'a qu'un seul développement limité à l'ordre 2 au voisinage de 1, ceci implique

$$5+49b = 0 \quad \text{et} \quad 4 + 24b^2 + 49c = 0$$

D'où l'on tire b et c (vérifier que les résultats coïncident avec ceux que l'on a trouvés ci-dessus)

Equation de la tangente

Supposons que $\frac{\partial f}{\partial y}$ soit non nul en M. Nous savons qu'au voisinage de M, la ligne de niveau de M possède un paramétrage du type $t \to (t,y(t))$. Nous avons donc $(\forall t)\ f(t,y(t)) = y_M$. D'où par dérivation $\frac{\partial f}{\partial x}(t,y(t))+y'(t)\frac{\partial f}{\partial y}(t,y(t)) = 0$. Donc $\frac{\partial f}{\partial x}(M)+y'(x_M)\frac{\partial f}{\partial y}(M) = 0$. La tangente en M s'écrit donc:

$$y = \left[-\frac{\frac{\partial f}{\partial x}(M)}{\frac{\partial f}{\partial y}(M)}\right]\ (x-x_M) + y_M$$

D'où l'expression classique $\quad (x-x_M)\,\frac{\partial f}{\partial x}(M) + (y-y_M)\,\frac{\partial f}{\partial y}(M) = 0$

Exercice d: On suppose maintenant que $\frac{\partial f}{\partial y}$ est nulle en M, mais $\frac{\partial f}{\partial x}$ non nulle. On peut alors paramétrer par $t \to (x(t),t)$. Vérifier que la formule ci-dessus est encore une équation de la tangente en M.

§ 4 : Développements à l'ordre 2.

Soit f une fonction de classe C^2 d'un ouvert U de $\mathbb{R}^2$ dans $\mathbb{R}$. Nous appellerons polynôme de Taylor à l'ordre 2 de f en $M(x_0,y_0)$, la fonction polynome de degré 2:

$$(x=x_0+h,y=y_0+k) \to P_f(h,k) = f(x_0,y_0) + h\,\frac{\partial f}{\partial x}(x_0,y_0) + k\,\frac{\partial f}{\partial y}(x_0,y_0)$$
$$+\frac{h^2}{2}\,\frac{\partial^2 f}{\partial x^2}(x_0,y_0) + hk\,\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0) + \frac{k^2}{2}\,\frac{\partial^2 f}{\partial y^2}(x_0,y_0)$$

Exercice g: Vérifier que c'est le seul polynome de degré 2 qui a même valeur et mêmes dérivées partielles d'ordre 1 et 2 que f au point M.

Le polynôme de Taylor est le développement limité à l'ordre 2, autrement dit la différence $f(x_0+h,y_0+k) - P_f(h,k)$ s'écrit $(h^2+k^2)\,\varepsilon(h,k)$, où ε est une fonction continue et nulle en $(0,0)$.

Pour nous en convaincre observons d'abord que (d'après l'exercice g) cette différence $\delta(h,k)$ est une fonction de classe C^2 nulle, ainsi que ses dérivées partielles d'ordre 1 et 2, pour $h = k = 0$. Posons

$$\varepsilon(h,k) = \frac{\delta(h,k)}{h^2+k^2} \text{ pour } h^2+k^2 \neq 0, \text{ et } \varepsilon(0,0) = 0.$$

Il s'agit de montrer que ε est continue en (0,0), c'est à dire que

$$(\forall\alpha>0)\ (\exists\eta>0) \text{ tel que } (\forall h \text{ et } k)\ \ (h^2+k^2\leq\eta) \text{ implique } |\varepsilon(h,k)| \leq \alpha.$$

Soit donc α positif. Puisque les dérivées partielles d'ordre 2 de δ sont nulles et continues en (0,0), il existe η tel que

$$h^2+k^2 \leq \eta \text{ implique } \left|\frac{\partial^2\delta}{\partial x^2}(0,0)\right| \leq \alpha/3,\ \left|\frac{\partial^2\delta}{\partial x\partial y}(0,0)\right| \leq \alpha/3 \text{ et } \left|\frac{\partial^2\delta}{\partial x^2}(0,0)\right| \leq \alpha/3)$$

Considérons alors la fonction $t \to \gamma(t) = \delta(th,tk)$. Elle est de classe C^2 (car composée de δ qui est de classe C^2 et de $t \to (th,tk)$ qui est C^∞), et ses dérivées s'écrivent :

$$\gamma'(t) = h\,\frac{\partial\delta}{\partial x}(th,tk) + k\,\frac{\partial\delta}{\partial y}(th,tk)$$
$$\gamma''(t) = h^2\frac{\partial^2\delta}{\partial x^2}(th,tk) + 2hk\,\frac{\partial^2\delta}{\partial x\partial y}(th,tk) + k^2\frac{\partial^2\delta}{\partial y^2}(th,tk)$$

Donc si $h^2+k^2 \leq \eta$ nous aurons $(\forall t\in[0,1])$:

$$|\gamma''(t)| \leq (h^2 + k^2 + 2|hk|)\,\alpha/3.$$

Soit , puisque $|hk| \leq h^2 + k^2$: $\quad |\gamma''(t)| \leq (h^2 + k^2)\,\alpha.$

Et par conséquent, par application de l'inégalité des acccroissements finis à γ' sur l'intervalle $[0,\tau]$, on obtient: $(\forall\tau\in[0,1])\ |\gamma'(0)-\gamma'(\tau)| = |\gamma'(\tau)| \leq \tau\,(h^2 + k^2)\,\alpha \leq (h^2 + k^2)\,\alpha$. Puis par application de l'inégalité des accroissements finis à γ sur $[0,1]$, on obtient:

$$|\gamma(0) - \gamma(1)| = |\gamma(1)| \leq (h^2 + k^2)\,\alpha.$$

Cette inégalité signifie que $|\varepsilon(h,k)| \leq \alpha$, ce que nous devions démontrer.

§ 5 : Recherche des maxima et minima d'une fonction.

Soit f une fonction de classe C^2 d'un ouvert U de $\mathbb{R}^2$ dans $\mathbb{R}$. Cherchons des conditions permettant de savoir si f possède un maximum ou un minimum relatif en $M(x_0,y_0)$.

Condition nécessaire: Pour que f possède un maximum ou un minimum en x_0, il est nécessaire que sa différentielle en x_0 soit nulle.

Pour nous en convaincre, nous allons montrer que si d_Mf n'est pas nulle, alors f n'a ni maximum, ni minimum en M. Choisissons un vecteur V tel que $d_Mf(V)\neq 0$. Posons $\varphi(t) = f(M+tV)$; φ est une fonction de classe C^2 définie pour t voisin de 0 (car* U est un ouvert), et $\varphi'(0) = d_Mf(V) \neq 0$ (chap 13, §2). Donc φ est strictement monotone au voisinage de 0. Ceci implique que, aussi près que l'on veut de 0, on peut trouver (**) des points $P = M+tV$ tels que $f(P) > f(M)$, et aussi des $P = M+tV$ tels que $f(P) < f(M)$. Donc f n'a ni minimum ni maximum en M.

Exercice h: Préciser les assertions marquées de (*) et de (**)

Pour que f posséde un maximum ou un minimum en M, il ne suffit pas que d_Mf soit nulle.
C'est bien connu, même pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$; la dérivée à l'origine de la fonction $x \to x^3$ est nulle, et pourtant il existe au voisinage de l'origine des x tels que $x^3>0$, et aussi des x tels que $x^3<0$. Pour les fonctions sur $\mathbb{R}^2$, citons $(x,y) \to x^2 - y^2$ qui, lorsqu'on la restreint à l'axe Ox présente un minimum en passant par l'origine, tandis que si on la restreint à Oy elle présente un maximum en 0.

Les points M tels que $d_Mf = 0$ sont appelés les points stationnaires de f. Nous allons chercher une condition suffisante permettant d'affirmer que le point stationnaire M est un maximum (ou un minimum) de f. Pour cela examinons la partie de degré 2 du développement limité en M, c'est à dire la fonction:

$$(h,k) \to \frac{h^2}{2}\frac{\partial^2 f}{\partial x^2}(x_0,y_0) + hk\,\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) + \frac{k^2}{2}\frac{\partial^2 f}{\partial y^2}(x_0,y_0)$$

C'est un polynôme "homogène de degré 2" en h et k (on dit aussi une "forme quadratique" en h et k).

Réduction de Gauss des formes quadratiques sur $\mathbb{R}^2$.

Considérons une une forme quadratique en h et k: $(h,k) \to Q(h,k) = Ah^2 + 2Bhk + Ck^2$

Si $A\neq 0$ nous pouvons écrire: $\quad Q(h,k) = A[h + \frac{B}{A}k]^2 + (C - \frac{B^2}{A})\,k^2$

Si $C\neq 0$ nous pouvons écrire: $\quad Q(h,k) = C[k + \frac{B}{C}h]^2 + (A - \frac{B^2}{C})\,h^2$

Si $A=C=0$ (et $B\neq 0$) nous pouvons écrire: $\quad Q(h,k) = \frac{B}{2}[(h+k)^2 - (h-k)^2]$

Dans les trois cas nous obtenons une expression de la forme $Q(h,k) = \lambda\,(\alpha h+\beta k)^2 + \mu(\gamma h+\delta k)^2$, dans laquelle λ et μ sont deux nombres peut être nuls, et $(h,k) \to \alpha h+\beta k$ et $(h,k) \to \gamma h+\delta k$ deux formes linéaires indépendantes (il est facile de vérifier que dans les trois cas $\alpha\delta - \beta\gamma \neq 0$).

Quatre éventualités sont alors possibles:

* *Si λ et μ sont strictement positifs, Q est le carré d'une norme sur $\mathbb{R}^2$ (on dit alors que Q est définie positive).* Considérons en effet l'application de $\mathbb{R}^2$ dans $\mathbb{R}$, qui à $X(u,v)$ et $Y(u',v')$ associe

$$\langle X,Y\rangle = \lambda\,(\alpha u+\beta v)(\alpha u'+\beta v') + \mu\,(\gamma u+\delta v)(\gamma u'+\delta v')$$

Elle est clairement bilinéaire et symétrique; $(\forall X)\ \langle X,X\rangle = \lambda\,(\alpha u+\beta v)^2 + \mu\,(\gamma u+\delta v)^2 \geq 0$ et de plus

$<X,X> = \lambda\,(\alpha u+\beta v)^2 + \mu\,(\gamma u+\delta v)^2 = 0$ implique $\alpha u+\beta v = \gamma u+\delta v = 0$, ce qui entraine $u = v = 0$ (i.e. $X = 0$) puisque $\alpha\delta - \beta\gamma \neq 0$. Nous avons donc défini un produit scalaire, et Q est le carré de la norme associée à ce produit scalaire.

** Si λ et μ sont strictement négatifs, c'est $-Q$ qui est le carré d'une norme (on dit alors que Q est définie négative).*

** Si λ et μ sont non nuls mais de signes contraires (on dit alors que Q est définie mais ni positive ni négative), il existe des droites passant par l'origine sur les quelles Q est nulle , il en existe sur les quelles Q est positive , il en existe sur les quelles Q est négative.*

En effet (en supposant $\lambda > 0$ et $\mu < 0$), sur la droite d'équation

$$E(u,v) = \sqrt{\lambda}\,(\alpha u+\beta v) + K\sqrt{-\mu}\,(\gamma u+\delta v) = 0,$$

Q est nulle si $K = \pm 1$, Q est positive si $|K| > 1$, et Q est négative si $|K| < 1$.

** Si l'un (au moins) des deux nombres λ et μ est nul (on dit que Q est dégénérée),* il existe (au moins) une droite passant par l'origine sur laquelle Q est nulle. (Par exemple si $\lambda = 0$, Q est nulle sur la droite d'équation $\gamma u+\delta v = 0$).

Conditions suffisantes pour qu'un point soit un extremum.

Considérons un point stationnaire de f, c'est à dire un point $M(x_0,y_0)$ tel que d_Mf soit nulle. En ce point le développement de f à l'ordre 2 s'écrit:

$$f(x_0+h,x_0+k) = f(x_0,y_0) + \frac{h^2}{2}\frac{\partial^2 f}{\partial x^2}(x_0,y_0) + hk\,\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0) + \frac{k^2}{2}\frac{\partial^2 f}{\partial y^2}(x_0,y_0) + (h^2+k^2)\,\varepsilon(h,k)$$

Soit $$f(x_0+h,x_0+k) = f(x_0,y_0) + Q(h,k) + (h^2+k^2)\,\varepsilon(h,k)$$

Nous retrouvons alors les quatre cas précédents:

** Si Q est définie positive, la fonction f a un minimum relatif strict en M.*

En effet définisssons $\varepsilon'(h,k)$ par: $\begin{cases} \varepsilon'(0,0) = 0 \\ \varepsilon'(h,k) = \dfrac{h^2+k^2}{Q(h,k)}\,\varepsilon(h,k) \text{ si } (h,k) \neq (0,0) \end{cases}$

Cette nouvelle fonction est continue et nulle en (0,0), puisque les normes $\sqrt{Q}$ et $(h,k) \to \sqrt{h^2+k^2}$ sont équivalentes et qu'ainsi le quotient $\dfrac{h^2+k^2}{Q(h,k)}$ reste compris entre deux nombres positifs A et B(*). Donc:

$$f(x_0+h,x_0+k) = f(x_0,y_0) + Q(h,k)\,[1+\varepsilon'(h,k)]$$

Pour $(h,k) \neq (0,0)$, $Q(h,k) > 0$. Pour (h,k) assez voisin de (0,0), $\varepsilon'(h,k) > -1$(**), donc le crochet est positif. Donc pour (h,k) non nul et assez voisin de (0,0), $Q(h,k)\,[1+\varepsilon'(h,k)]$ est strictement positif, et ainsi $f(x_0+h,y_0+k) > f(x_0,y_0)$. Ce qui signifie que M est un minimum relatif strict de f.

Exercice i: Expliquer les assertions (*) et (**)

** Si Q est définie négative, Q a un maximum relatif strict en M.*

Exercice j: On suppose Q définie négative. Démontrer qu'il existe une fonction φ continue et nulle en (0,0) etlle que $f(x_0+h,x_0+k) = f(x_0,y_0) - Q(h,k)\,[-1+\varphi(h,k)]$. En déduire que f a un maximum relatif strict en M.

** Si Q est définie mais n'est ni positive ni négative, la fonction f n'a ni maximum ni minimum en M.*

En effet: Soit Δ une droites passant par l'origine et sur la quelle Q est positive (resp. négative). Soit $t \to (x_0+th_0,y_0+tk_0)$ un paramétrage de la parallèle Δ' à Δ passant par M. Alors

$$f(x_0+th_0,x_0+tk_0) = f(x_0,y_0) + Q(th_0,tk_0) + t^2\,(h_0^2+k_0^2)\,\varepsilon(th_0,tk_0)$$

Soit $$f(x_0+th_0,x_0+tk_0) = f(x_0,y_0) + t^2Q(h_0,k_0) + t^2\,(h_0^2+k_0^2)\,\varepsilon(th_0,tk_0)$$

Ou encore $\quad f(x_0+th_0,x_0+tk_0) = f(x_0,y_0) + t^2\,[Q(h_0,k_0) + (h_0^2+k_0^2)\,\varepsilon(th_0,tk_0)]$

Puisque ε est continue et nulle en (0,0), pour t assez petit, le crochet est du signe de $Q(h_0,k_0)$; donc la restriction de f à Δ' présente un minimum (resp. un maximum) en M.

Exercice k: Expliquer pourquoi pour t assez petit le crochet est du signe de $Q(h_0,k_0)$.

* *Si Q est dégénérée il n'est pas possible de conclure en ne regardant que les termes de degré 2 du développement; pour conlure il faudrait développer à l'ordre 3, 4,...*

Par exemple les fonctions $f(x,y) = x^2+ y^4$, $g(x,y) = x^2+ y^3$, $h(x,y) = x^2 - y^4$ ont toutes même développement à l'ordre 2 à l'origine, avec pour forme quadratique $Q(x,y) = x^2$, qui est dégénérée. Pourtant elles ont des comportements fort différents; f a un minimum à l'origine, tandis que g et h n'en ont pas; la restriction de h à l'axe des y, a même un maximum en 0.

§ 6 : Les lignes de niveau des polynômes du second degré.

Soit $(O,\vec{\imath},\vec{\jmath})$ un repère orthonormé du plan. Soit un polynome homogène du second degré

$$P(x,y) = ax^2 + 2bxy + cy^2$$

Pour tout nombre n, notons Γ_n l'ensemble des points M(x,y) tels que P(x,y) = n.

Le calcul qui suit repose sur le fait que la matrice à une ligne et une colonne $(x,y)\begin{pmatrix} a & b\\ b & c\end{pmatrix}\begin{pmatrix} x\\ y\end{pmatrix}$ a pour unique coefficient P(x,y). Pour étudier Γ_n nous allons changer de repère. La matrice $S = \begin{pmatrix} a & b\\ b & c\end{pmatrix}$ est symétrique donc (cf:chap.5 §5) il existe une matrice orthogonale $H = \begin{pmatrix} \alpha & \beta\\ \gamma & \delta\end{pmatrix}$ et une matrice diagonale $\Delta = \begin{pmatrix} \lambda & 0\\ 0 & \mu\end{pmatrix}$, telles que $S = H^{-1}\Delta H$. Alors la matrice $(x,y)S\begin{pmatrix} x\\ y\end{pmatrix}$ s'écrit aussi:

$$(P(x,y)) = (x,y)H^{-1}\circ\Delta\circ H\begin{pmatrix} x\\ y\end{pmatrix} = (x,y)H^t\,\Delta\,H\begin{pmatrix} x\\ y\end{pmatrix} = [H\begin{pmatrix} x\\ y\end{pmatrix}]^t\,\Delta\,[H\begin{pmatrix} x\\ y\end{pmatrix}]$$

$$= (\alpha x+\beta y,\gamma x+\delta y)\;\Delta\begin{pmatrix} \alpha x+\beta y\\ \gamma x+\delta y\end{pmatrix} = \lambda(\alpha x+\beta y)^2 + \mu(\gamma x+\delta y)^2.$$

Supposons que $\mathbb{R}^2$ soit muni de son produit scalaire naturel. Alors (puisque H est orthogonale) $X = \alpha x+\beta y$ et $Y = \gamma x+\delta y$ sont les coordonnées du point M(x,y) dans une nouvelle base orthonormée $(\vec{I},\vec{J})$. Dans cette base l'expression de la fonction $P: \mathbb{R}^2 \to \mathbb{R}$ est $(X,Y) \to \lambda X^2 + \mu Y^2$.

Pour dessiner les lignes de niveau nous distinguerons 3 cas (en oubliant le cas $\lambda=\mu=0$, c'est à dire P(x,y)=0 quel que soit (x,y)).

<u>Si λ et μ sont de même signe</u>: Alors Γ_n ($\lambda X^2+\mu Y^2 = n$) est soit vide (si n n'est pas de même signe que λ et μ), soit une ellipse (si λ,μ et n sont de même signe), soit réduite à l'origine (si n=0). C'est le cas où P est définie positive, ou définie négative.

<u>Si λ et μ sont de signes contraires</u>: Alors Γ_n est soit une hyperbole (si $n\neq 0$), soit deux droites sécantes à l'origine (si n=0). C'est le cas où P est définie, mais ni positive, ni négative.

<u>Si λ est nul et μ non nul</u>: Alors Γ_n est soit vide (si $\mu n<0$), soit une droite (si n=0), soit deux droites parallèles (si $\mu n>0$). C'est le cas où P est non définie (mais non nulle).

Exercice l: On considère la forme quadratique $Q(h,k) = h^2 + 4hk + 5k^2$.

1) Montrer qu'elle est définie positive.

2) Déterminer un repère orthonormé $(O,\vec{I},\vec{J})$, dans lequel Q s'écrit $\lambda X^2 + \mu Y^2$.

3) Dans un plan muni d'un repère orthonormé $(O,\vec{\imath},\vec{\jmath})$, soit la courbe C d'équation $Q(x,y) = 4x + 2y + 1$. Trouver une équation de C dans le repère $(O,\vec{I},\vec{J})$; et préciser le tracé de C en donnant en particulier ses axes de symétrie.

Exercice m: On considère la forme quadratique $Q(h,k) = h^2 + 4hk + k^2$.

1) Montrer qu'elle est définie mais ni positive ni négative.

2) Déterminer un repère orthonormé $(O, \vec{I}, \vec{J})$ dans lequel Q s'écrit $\lambda X^2 + \mu Y^2$.

3) Dans un plan muni d'un repère orthonormé $(O, \vec{i}, \vec{j})$, soit la courbe C d'équation $Q(x,y) = 4x + 6y + 1$. Trouver une équation de C dans le repère $(O, \vec{I}, \vec{J})$; et préciser le tracé de C en donnant ses axes de symétrie et ses asymptotes.

§ 7 : Les lignes de niveau au voisinage d'un point stationnaire.

Considérons une fonction $f : \mathbb{R}^2 \to \mathbb{R}$, et un point $M(x_0,y_0)$ tel que $d_M f = 0$. Nous allons étudier la forme des lignes de niveau de f au voisinage de M. Nous n'étudierons que le cas où la forme quadratique $Q(h,k)$ formée des termes de degré 2 du développement de Taylor est définie; car, si elle est dégénérée, l'étude des lignes de niveau de f nécessite de développer à l'ordre 3 (ou plus).

Puisque $d_M f=0$, le polynôme de Taylor à l'ordre 2 de f en $M(x_0,y_0)$ est de la forme

$$P_f(x-x_0,y-y_0) = f(x_0,y_0) + A(x-x_0)^2 + 2B(x-x_0)(y-y_0) + C(y-y_0)^2$$

Puisque nous supposons que la forme quadratique des termes du second degré est définie, les lignes de niveau de P_f sont:

Soit deux droites sécantes et les hyperboles qui les admettent pour asymptotes.

Soit un point et des ellipses centrées en ce point.

Des lignes de niveau de P_f à celles de f.

Si f est de classe C^∞ (si $d_M f = 0$, et si la forme quadratique P_f des termes du second degré du polynôme de Taylor est définie) il existe deux ouverts ω et Ω contenant M, et un difféomorphisme F de ω sur Ω, tels que $f|\omega = P_f \circ F$ (et tel que $F(M) = M$ et $d_M F = \mathrm{Id}$). On posera $F^{-1} = G$.

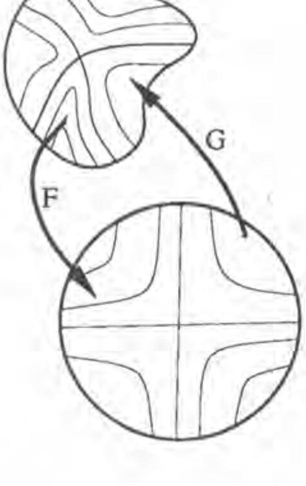
On notera que la relation $f|\omega = P_f \circ F$ implique que M est un minimum (ou un maximum) de f sur ω, si et seulement si M (=F(M)) est un minimum (ou un maximum) de P_f sur Ω. Ce qui redonne immédiatement les résultats du §4.

Les lignes de niveau de f sont donc les transformées par G de celles de P_f (on comparera avec la situation étudiée dans le théorème des fonctions implicites). Nous distinguerons alors deux cas:

Premier cas: La forme quadratique est définie mais ni positive, ni négative.

Au voisinage de M, les lignes de niveau de P_f sont:

D'une part deux morceaux de droites sécants, dont les images par G sont deux morceaux de courbes C^∞, sécants en M. Ils forment la ligne de niveau de f passant par M, qui est ainsi une courbe γ_M possédant un point double en M.



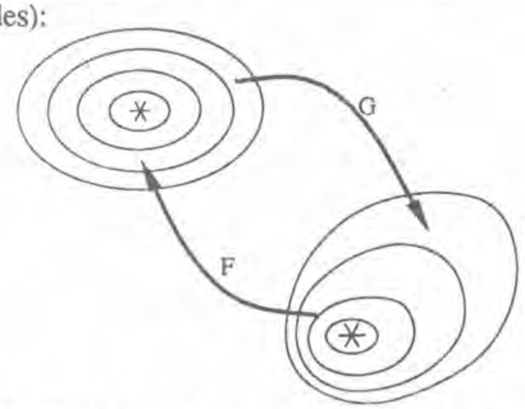
D'autre part des morceaux d'hyperboles, dont les images par G sont des courbes C^∞. Ce sont les lignes de niveau de f qui correspondent aux niveaux différents de f(M).

Nous pouvons déterminer les tangentes en M aux deux branches de γ_M. Les vecteurs tangents sont les images par $d_M G$ des vecteurs directeurs des deux droites Δ et Δ' définies par l'équation $P_f(x,y) = f(M)$; et puisque $d_M G$ $(=(d_M F)^{-1})$ est l'identité, ce sont ces vecteurs directeurs eux mêmes. Donc les deux branches de γ_M sont tangentes en M à Δ et Δ'.

Second cas: la forme quadratique est, soit définie positive, soit définie négative.

Les lignes de niveau de P_f sont (en oubliant celles qui sont vides):

D'une part le point M dont l'image par G est le point M. Ainsi l'équation f(X) = f(M) n'a qu'une solution dans ω c'est le point M.



D'autre part des ellipses concentriques à M dont les images par G sont des courbes qui entourent M. Ce sont celles qui correspondent aux niveaux supérieurs à f(M) si f a un minimum en M, et celles qui correspondent aux niveaux inférieurs à f(M) si f a un maximum en M.

Exemple: Soit $f(x,y) = (x^2+2y)(x^2+y^2) + x^2 - y^2$, et soit M(0,0). On vérifie (le lecteur fera les calculs) que les deux dérivées partielles de f sont nulles en M, et que le polynôme de Taylor à l'ordre 2 en M, s'écrit $(x,y) \to x^2 - y^2$. La forme quadratique est donc définie, mais ni positive, ni négative. L'équation $P_f(x,y) = f(M) = 0$, définit les deux bissectrices (x=y) et (x+y=0). Donc les deux branches de γ_M ont pour tangentes en M, les deux bissectrices.

Exercice n: Soit Γ la branche qui est tangente à la première bissectrice, elle possède un paramétrage du type $t \to (t,\varphi(t))$. Quelle est la dérivée de φ en 0? Utiliser le fait que $(\forall t)\ f(t,\varphi(t)) = 0$, pour calculer les dérivées secondes de φ en 0.

Comment on démontre l'existence de $F:\omega \to \Omega$.

Nous ne démontrerons pas complètement ce résultat. Nous allons seulement, sur un exemple, montrer comment on construit F. Choisissons la fonction de l'exemple ci-dessus:

$$f(x,y) = (x^2+2y)(x^2+y^2) + x^2 - y^2$$

Nous pouvons écrire:

$$f(x,y) = x^2\times(1+x^2+y^2) + 2xy\times(x) + y^2\times(-1+2y)$$

Puis, de façon assez analogue à ce que nous avons fait dans la réduction de Gauss des formes quadratiques:

$$f(x,y) = \left[x\sqrt{1+x^2+y^2} + y\frac{x}{\sqrt{1+x^2+y^2}}\right]^2 + y^2\left[-1+2y - \frac{x^2}{1+x^2+y^2}\right]$$

Nous définirons alors une application (de classe C^∞) $F: \mathbb{R}^2 \to \mathbb{R}^2$, en posant

$$X = x\sqrt{1+x^2+y^2} + y\frac{x}{\sqrt{1+x^2+y^2}} \qquad \text{et} \qquad Y = y\sqrt{1-2y + \frac{x^2}{1+x^2+y^2}}$$

Il est clair que $f(x,y) = X^2 - Y^2$; ce qui s'écrit aussi $P_f \circ F = f$. La différentielle de F en (0,0) est l'identité (vérification facile); donc (théorème d'inversion locale) F définit un difféomorphisme d'un voisinage ω de (0,0) sur un voisiange Ω de F(0,0) = (0,0). Ce qui termine la démonstration.

§ 8 : Les surfaces z = f(x,y).

Soit f une fonction de classe C^1 d'un ouvert U de $\mathbb{R}^2$ dans $\mathbb{R}$. Le graphe de cette fonction est l'ensemble S des $(x,y,z)\in U\times\mathbb{R}$ tels que z = f(x,y). Un tel sous-ensemble de $\mathbb{R}^3$ est appelé une surface.

Soit $M(x_0,y_0,f(x_0,y_0))$ un point de S, nous appellerons plan tangent à S au point M, le graphe de l'application $(x,y) \to f(x_0,y_0) + (x-x_0)\frac{\partial f}{\partial x}(x_0,y_0) + (y-y_0)\frac{\partial f}{\partial y}(x_0,y_0)$; c'est à dire le graphe du polynôme de Taylor à l'ordre 1 en $m(x_0,y_0)$. On le notera $T_M(S)$.

Exercice p: Soit $t \to \Gamma(t)=(x(t),y(t),z(t))$ une courbe régulière tracée sur S et passant par M pour $t = t_0$. Montrer que la tangente en M à cette courbe est dans $T_M(S)$.

Intersection de deux surfaces.

Considérons deux surfaces $S(z=f(x,y))$ et $\Sigma(z=g(x,y))$. Posons $\Gamma = S\cap\Sigma$, et considérons un point M de Γ. On suppose que les plans tangents $T_M(S)$ et $T_M(\Sigma)$ ne sont pas confondus (puisqu'ils ont un point commun, ils se coupent donc suivant une droite T). Alors au voisinage de M, Γ est une courbe régulière dont la tangente en M est la droite T.

Pour nous en convaincre nous allons d'abord chercher la projection γ de Γ sur le plan des (x,y). C'est l'ensemble des points $m(x,y)$ tels que $f(x,y)-g(x,y) = 0$.

La différentielle de f-g en (x_0,y_0) est non nulle, en effet c'est l'application:

$$(x-x_0,y-y_0) \to (x-x_0)\frac{\partial(f-g)}{\partial x}(x_0,y_0) + (y-y_0)\frac{\partial(f-g)}{\partial y}(x_0,y_0)$$

c'est à dire la différence des deux applications:

$$(x-x_0,y-y_0) \to f(x_0,y_0) + (x-x_0)\frac{\partial f}{\partial x}(x_0,y_0) + (y-y_0)\frac{\partial f}{\partial y}(x_0,y_0)$$

$$(x-x_0,y-y_0) \to g(x_0,y_0) + (x-x_0)\frac{\partial g}{\partial x}(x_0,y_0) + (y-y_0)\frac{\partial g}{\partial y}(x_0,y_0)$$

qui sont distinctes puisque les plans tangents sont distincts.

Nous pouvons donc appliquer le théorème des fonctions implicites à f-g et au point m. Il existe donc un ouvert ω contenant m, tel que $\omega\cap\gamma$ soit un arc régulier.

Soit $t\mapsto (x(t),y(t))$ un paramétrage régulier de cet arc, alors $t\mapsto (x(t),y(t),f(x(t),y(t)))$ est un paramétrage de $(\omega\times\mathbb{R})\cap\Gamma$. Et $\frac{d\vec{\Gamma}}{dt}(t)$ est non nul, puisque sa projection $\frac{d\vec{\gamma}}{dt}(t)$ sur le pl:an des (x,y) est non nulle. Pour montrer que la tangente à Γ en M est l'intersection des plans tangents, on se reportera à l'exercice p.

Intersection d'une surface et de son plan tangent

Cherchons l'intersection Γ de $S(z=f(x,y))$ et de $T_M(S)$. La projection γ de Γ sur le plan des (x,y) est l'ensemble des $m(x,y)$ tels que

$$G(x,y) = f(x,y) - \left[f(x_0,y_0) + (x-x_0)\frac{\partial f}{\partial x}(x_0,y_0) + (y-y_0)\frac{\partial f}{\partial y}(x_0,y_0)\right] = 0$$

La différentielle de G est nulle, donc le calcul précédent ne s'applique pas. L'étude de γ va se faire par les méthodes du § 7. Le développement de G à l'ordre 2 en m s'écrit:

$$G(x_0+h,x_0+k) = \frac{h^2}{2}\frac{\partial^2 f}{\partial x^2}(x_0,y_0) + hk\,\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0) + \frac{k^2}{2}\frac{\partial^2 f}{\partial y^2}(x_0,y_0) + (h^2+k^2)\,\varepsilon(h,k)$$

(où ε est continue et nulle en (0,0)). Notons Q la forme quadratique des termes du second degré de ce développement. Trois cas sont alors possibles:

* Si la forme Q est dégénérée, nous ne pourrons rien dire, il faudrait (supposer que f est C^3,...et) regarder les dévéloppements d'ordre 3,...

* Si la forme Q est définie positive ou définie négative, on sait qu'il existe un ouvert ω contenant m, tel que $\omega\cap\gamma$ soit réduit à m. Il en résulte que $(\omega\times\mathbb{R})\cap\Gamma$ est réduit à M. Ainsi le point M est isolé sur Γ. Au voisinage de M l'ensemble Γ n'a pas d'autre point que M.

* Si la forme Q est définie mais ni positive, ni négative, on sait qu'il existe un ouvert ω contenant m, tel que $\omega\cap\gamma$ soit formé de deux arcs réguliers se coupant en m.

Si $t \to (x(t),y(t))$ est un paramétrage régulier de l'un d'eux, la courbe de l'espace définie par $t \to (x(t),y(t),f(x(t),y(t)))$ est régulière (car sa projection est régulière), et tracée sur S et sur $T_M(S)$. Ainsi au voisinage de M, Γ est formé de deux arcs sécants en M; l'intersection de la surface et de son plan tangent est une courbe à point double.

<u>Vocabulaire</u>: Lorsque la forme Q est définie positive ou négative, on dit que la courbure de la surface est positive en M. Lorsque la forme Q est définie mais ni positive, ni négative, on dit que la courbure de la surface est négative en M. Lorsque Q est dégénérée, on dit que la courbure de la surface est nulle en M.

Exercice q: Repérer sur la surface d'un vase les points où la courbure est positive, et ceux où elle est négative.

§ 9 : Recherche des extrema d'une fonction de $\mathbb{R}^3$ dans $\mathbb{R}$.

Soit f une application de classe C^2 d'un ouvert U de $\mathbb{R}^3$ dans $\mathbb{R}$. Et soit M un point de U.

Une condition nécessaire pour que f ait un extremum en M, est que la différentielle de f en M soit nulle.

La démonstration est analogue à celle que nous avons faite en dimension 2, résumons la: si $d_M f$ est non nulle, il existe un vecteur X tel que $D_X f$ soit non nulle; et alors le long de la droite passant par M et de vecteur directeur X, la fonction f est strictement monotone au voisinage de M; ce qui exclut que f puisse présenter un extremum en M.

Les points M tels que $d_M f = 0$ sont appelés les points stationnaires de f; ce ne sont pas tous des extrema. Pour savoir si le point stationnaire M est un extremum, nous allons regarder le développement à l'ordre 2 de f en M.

Nous admettrons qu'il s'écrit (si $h = (h_x,h_y,h_z)$)

$$f(M+h) - f(M) = \frac{1}{2}\left[(h_x)^2\frac{\partial^2 f}{\partial x^2}(M) + (h_y)^2\frac{\partial^2 f}{\partial y^2}(M) + (h_z)^2\frac{\partial^2 f}{\partial z^2}(M) + 2h_xh_y\frac{\partial^2 f}{\partial x\partial y}(M) + 2h_yh_z\frac{\partial^2 f}{\partial y\partial z}(M) + 2h_zh_x\frac{\partial^2 f}{\partial z\partial x}(M)\right] + (h_x^2+h_y^2+h_z^2)\varepsilon(h)$$

(où ε est une fonction continue et nulle pour h = 0). Soit $f(M+h) = f(M) + Q(h) + N(h)^2\,\varepsilon(h)$. Où Q(h) est une forme quadratique sur $\mathbb{R}^3$, c'est à dire une expression de la forme

$$Q(h_x,h_y,h_z) = A_1\, h_x^2 + A_2\, h_y^2 + A_3\, h_z^2 + 2B_1h_xh_y + 2B_2\, h_yh_z + 2B_3\, h_zh_x$$

Réduction des formes quadratiques

Toute forme quadratique Q sur $\mathbb{R}^3$ peut s'écrire

$$Q(h) = \lambda_1\,[\varphi_1(h)]^2 + \lambda_2\,[\varphi_2(h)]^2 + \lambda_3\,[\varphi_3(h)]^2$$

Où les λ_j sont des nombres et les φ_j des formes linéaires indépendantes.

Pour nous en convaincre posons $S = \begin{pmatrix} A_1 & B_1 & B_3 \\ B_1 & A_2 & B_2 \\ B_3 & B_2 & A_3 \end{pmatrix}$; on vérifie facilement que $(h_x,h_y,h_z)\,S\begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix}$ est la matrice à une ligne et une colonne dont l'unique coefficient est Q(h). On sait que toute matrice symétrique est diagonalisable par une matrice orthogonale (chap.5 §5). Il existe donc une matrice orthogonale H et une matrice diagonale Δ telles que $S = H^{-1}\Delta H = H^t\,\Delta H$. Et alors

$$(Q(h)) = (h_x,h_y,h_z)\,S\begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = (h_x,h_y,h_z)\,H^{-1}\Delta H\begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = (h_x,h_y,h_z)\,H^t\,\Delta\,H\begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix}$$

En posant $H\begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = \begin{pmatrix} \varphi_1(h) \\ \varphi_2(h) \\ \varphi_3(h) \end{pmatrix}$ (ce qui équivaut à $(\varphi_1(h),\varphi_2(h),\varphi_3(h)) = (h_x,h_y,h_z)\,H^t$), on obtient

$$(Q(h)) = (\varphi_1,\varphi_2,\varphi_3)\,\Delta\begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} = (\lambda_1\,[\varphi_1(h)]^2 + \lambda_2\,[\varphi_2(h)]^2 + \lambda_3\,[\varphi_3(h)]^2)$$

Où les λ_j sont les coefficients diagonaux de Δ. Les φ_j sont des formes linéaires indépendantes (puisque H est une matrice inversible).

Cette démonstration nous donne une méthode un peu longue pour trouver une forme réduite de Q. Nous emploierons plutôt la méthode de Gauss, que nous allons décrire sur un exemple.

Soit: $Q(u,v,w) = u^2+v^2+8w^2+2uv+4uw-2vw$

Nous allons écrire une forme linéaire dont le carré contient tous les termes en u de Q.

$$Q(u,v,w) = (u+v+2w)^2 + 4w^2 - 6vw = (u+v+2w)^2 + Q'$$

Puis nous écrivons une forme linéaire dont le carré contient tous les termes en w de Q'.

$$Q(u,v,w) = (u+v+2w)^2+(2w-\tfrac{3}{2}v)^2 - \tfrac{9}{4}v^2$$

On vérifiera facilement que les 3 formes u+v+w, $2w - \frac{3}{2}v$, v sont indépendantes.

Exercice r: Réduire la forme quadratique $Q = 4xy + 2yz + 6zx$. On remarquera qu'ici la méthode de Gauss que nous venons de décrire ne peut s'employer directement (car il n'y a pas de termes en x^2, ni en y^2, ni en z^2). On commencera donc par employer l'identité $(x+y)^2 - (x-y)^2 = 4xy$ pour "faire apparaitre des carrés.

Il y a ainsi plusieurs méthodes pour réduire un forme quadratique, et chacune d'elles comporte un certain nombre de choix. Il y a donc de nombreuses réductions possibles; mais on démontre que:

Le nombre des coefficients λ_i positifs, le nombre des λ_i négatifs et le nombre des λ_i nuls est indépendant des choix que l'on a faits. Si les trois λ_i sont non nuls, on dit que la forme Q est définie. S'ils sont tous trois positifs, Q est dite définie positive.S'ils sont tous les trois négatifs, Q est dite définie négative. S'il y a un ou plusieurs λ_i nuls, on dit que Q est dégénérée.

Exercice s: Déterminer le nombre de carrés positifs et le nombre de carrés négatifs des formes quadratiques suivantes:

$x^2 + y^2 + z^2 - 2xy - 2yz - 2zx$ $\quad v^2+5w^2+6uv+4vw-2wu$ $\quad 2uv+4uw-2vw$

Conditions suffisantes pour qu'un point stationnaire soit un extremum.

Soit Q la forme quadratique des termes de degré 2 du développement de f en M; alors:

* *Si Q est définie positive, f a un minimum relatif strict en M.*

* *Si Q est définie négative, f a un maximum relatif strict en M.*

* *Si Q est définie, mais ni positive, ni négative, M n'est pas un extremum.*

* *Si Q est dégénérée, il est impossible de conclure en ne regardant que le développement à l'ordre 2 de f en M.*

Ces résultats sont tout à fait analogues à ceux que nous avons vus en dimension 2. Les démonstrations sont aussi fort semblables. Nous ne les ferons pas.

Complément : Le théorème des fonctions implicites pour les fonctions de $\mathbb{R}^3$ dans $\mathbb{R}$.

Le théorème d'inversion locale

Soit f une application de classe C^1 d'un ouvert U de $\mathbb{R}^3$ dans $\mathbb{R}^3$. Soit M un point de U tel que $d_M f$ soit inversible (c'est une application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$). Alors il existe un ouvert ω de $\mathbb{R}^3$ qui contient M et un ouvert Ω de $\mathbb{R}^3$, tels que $f|\omega$ soit un difféomorphisme de ω sur Ω.

Nous admettrons cet énoncé, identique à celui que nous avons vu en dimension 2. Il permet de préciser l'allure de l'ensemble des solutions de l'équation $f(P) = \alpha$, lorsque f est une fonction de classe C^1, à valeurs dans $\mathbb{R}$.

Le théorème des fonctions implicites en dimension 3

Soit $f: U(\subset\mathbb{R}^3) \to \mathbb{R}$ une fonction de classe C^1. Soit M un point où la dérivée partielle $\frac{\partial f}{\partial z}$ est non nulle. Pour étudier l'équation f(P) = f(M) au voisinage du point M, nous allons utiliser l'application $F:\mathbb{R}^3 \to \mathbb{R}^3$ définie par $F(x,y,z) = (x,y,f(x,y,z))$. Elle est de classe C^1; et sa différentielle en M a pour matrice (dans la base naturelle de $\mathbb{R}^3$)

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{\partial f}{\partial x}(M) & \frac{\partial f}{\partial y}(M) & \frac{\partial f}{\partial z}(M) \end{pmatrix}$$

Puisque $\frac{\partial f}{\partial z}(M) \neq 0$, elle est inversible. Nous pouvons donc appliquer le théorème d'inversion locale. Il existe un ouvert ω contenant M (et contenu dans U), et un ouvert Ω de $\mathbb{R}^3$ contenant F(M), tels que $F|\omega$ soit un difféomorphisme de ω sur Ω. Notons $G:\Omega \to \omega$ l'application inverse; elle est de classe C^1 (et même C^∞ si f est C^∞). Notons G_1,G_2,G_3 les trois fonctions coordonnées de G. Alors $F(G_1(X,Y,Z),G_2(X,Y,Z),G_3(X,X,Z))$ est égal à (X,Y,Z) (car F est l'inverse de G) et aussi à $(G_1(X,Y,Z),G_2(X,Y,Z), f(G_1(X,Y,Z),G_2(X,Y,Z),G_3(X,X,Z))$ (d'après la définition de F); donc $G_1(X,Y,Z) = X$ et $G_2(X,Y,Z) = Y$. Ainsi $G(X,Y,Z) = (X,Y,G_3(X,Y,Z))$.

Pour tout α, notons Σ_α l'ensemble des points P de ω tels que $f(P) = \alpha$; notons P_α l'intersection de Ω et du plan d'équation $z = \alpha$. Alors $F(\Sigma_\alpha) \subset P_\alpha$ et $G(P_\alpha) \subset \Sigma_\alpha$. Donc F et G sont des bijections réciproques de Σ_α sur P_α.

A tout (x,y) tel que (x,y,α) soit dans P_α associons $G_3(x,y,\alpha)$; nous obtenons une application φ d'une partie de $\mathbb{R}^2$, à valeurs dans $\mathbb{R}$, dont le graphe est $G(P_\alpha)$ c'est à dire Σ_α.

Ainsi l'hypothèse que $\frac{\partial f}{\partial z}(M) \neq 0$ permet de démontrer que dans un ouvert ω contenant M, les solutions de l'équation $f(P) = \alpha$, forment une surface qui peut être définie comme le graphe d'une application φ de classe C^1 d'une partie de $\mathbb{R}^2$ dans $\mathbb{R}$. Et φ est C^∞ chaque fois que f est C^∞.

Si nous avions supposé $\frac{\partial f}{\partial y}(M) \neq 0$ (resp. $\frac{\partial f}{\partial x}(M) \neq 0$), nous aurions, de façon analogue, montré que les Σ_α étaient les graphes d'applications $(x,z) \to \varphi(x,z)$ (resp. $(y,z) \to \varphi(y,z)$). De façon générale si $d_M f$ n'est pas nulle, au voisinage de M les solutions de $f(P) = \alpha$ (s'il en existe !) forment une surface.

Exercice t: Soit S l'ensemble des points M(x,y,z) tels que $x^2+2xy^2+y^2z^2 = 2x+4y+2z$. Trouver trois fonctions $\varphi,\varphi',\varphi''$ telles que, au voisinage de l'origine, S puisse têre considérée soit comme l'ensemble des points de la forme $(x,y,\varphi(x,y))$, soit comme l'ensemble des points de la forme $(x,\varphi'(x,z),z)$, soit comme l'ensemble des points de la forme $(\varphi''(y,z),y,z)$.

La fonction φ est une fonction implicite, c'est à dire qu'il nous faut renoncer à la calculer explicitement (sauf dans quelques cas simples comme à l'exercice t). Nous pourrons seulement connaitre ses dérivées partielles en dérivant la formule $(\forall x \text{ et } y) f(x,y,\varphi(x,y)) = 0$.

L'équation du plan tangent

En dérivant la relation ci-dessus, d'abord par rapport à x, puis par rapport à y, on obtient:

$$\frac{\partial f}{\partial x}(x,y,\varphi(x,y)) + \frac{\partial f}{\partial z}(x,y,\varphi(x,y))\,\frac{\partial \varphi}{\partial x}(x,y) = 0$$

$$\frac{\partial f}{\partial y}(x,y,\varphi(x,y)) + \frac{\partial f}{\partial z}(x,y,\varphi(x,y))\,\frac{\partial \varphi}{\partial y}(x,y) = 0$$

Ce qui donne les dérivées partielles premières de φ en (x_M,y_M), et l'équation du plan tangent en M:

$$z = z_M + \frac{-\frac{\partial f}{\partial x}(x_M,y_M,\varphi(x_M,y_M))}{\frac{\partial f}{\partial z}(x_M,y_M,\varphi(x_M,y_M))}(x-x_M) + \frac{-\frac{\partial f}{\partial y}(x_M,y_M,\varphi(x_M,y_M))}{\frac{\partial f}{\partial z}(x_M,y_M,\varphi(x_M,y_M))}(y-y_M)$$

Ce qui donne la formule classique (valable quelle que soit la dérivée partielle non nulle en M):

$$(x-x_M)\frac{\partial f}{\partial x}(M) + (y-y_M)\frac{\partial f}{\partial y}(M) + (z-z_M)\frac{\partial f}{\partial z}(M) = 0$$

Exercice u: Soit S la surface d'équation $x^2+3y^2+2z^2 = 4x + y + 2z + 1$.

1) Ecrire l'équation de son plan tangent au point $M(1,-1,1)$.

2) Déterminer les points de S où le plan tangent est parallèle au plan $\pi(x+4y-3z = 0)$

Exercices sur le chapitre 16

* **Exercice 1**: Dans $\mathbb{R}^2$ on considère les courbes:

$$C(x^2 - 2y^3 = x - 2y) \quad \text{et} \quad \Gamma(x^2 + 2y^5 - 4y^4 = x - 2y)$$

Elles passent toutes deux par le point P(1;1). Montrer qu'elles ont même tangente en ce point. Dessiner cette tangente, préciser les positions des deux courbes par rapport à cette tangente, et les positions relatives des deux courbes.

* **Exercice 2:** Tracer la courbe d'équation $x^2 - y^3 + 3y + 2 = 0$. On pourra chercher à l'écrire sous la forme $x = \varphi(y)$. Préciser les tangentes aux points d'intersection avec les axes.

Mêmes questions pour la courbe $x^2 + y^3 - 3y - 2 = 0$.

* **Exercice 3:** La courbe Γ d'équation $x \sin \pi y + y \cos \pi x = \frac{\sqrt{2}}{4}$ passe par le point $M(\frac{1}{4} ; \frac{1}{4})$. Montrer qu'au voisinage de M c'est une courbe régulière. Quelle est sa tangente en M ?

** **Exercice 4:** Dans un plan muni d'un repère orthonormé $(O, \vec{i}, \vec{j})$, on considère les points $A(0;-1)$ et $B(0,1)$. Soit $M(x_0,y_0)$ un point non situé sur les axes. Par M passe une hyperbole H de foysers A et B, et une ellipse E de foyers A et B.

a) Donner une équation de H et une équation de E.

b) Montrer que ces deux courbes se coupent orthogonalement en M .

** **Exercice 5**: Dans $\mathbb{R}^2$ on considère l'ensemble Γ des points M(x,y) tels que:

$$f(x,y) = (x+y)(x^2+y^2) - 4xy = 0.$$

a) Montrer qu'au voisinage du point A(1;1), Γ est une courbe régulière. Ecrire une équation de la tangente à Γ en A.

b) Montrer q'au voisinage de l'origine, Γ est une courbe ayant un point double. Ecrire des équations des tangentes à l'origine aux deux arcs qui y passent.

c) Montrer que le système d'équations $\{ f(x,y) = 0 , \frac{\partial f}{\partial x}(x,y) = 0 , \frac{\partial f}{dy}(x,y) = 0 \}$ a une seule solution (qui est l'origine) Qu'en résulte-t-il pour Γ ?

d) Soit D_t la droite d'équation $y = tx$. Montrer qu'elle coupe Γ en deux points: l'origine et un autre point M_t dont on calculera les coordonnées (en fonction de t). Tracer la courbe Γ.

** **Exercice 6**: Dans un plan P, on considére un repère orthonormé $(O, \vec{i}, \vec{j})$, et l'ensemble Γ défini (dans ce repère) par l'équation:

$$f(x,y) = x(x^2 + y^2) - (x^2 - y^2) = 0$$

a) Le point A(1;0) est sur Γ. Montrer qu'au voisinage de A, Γ est une courbe régulière et donner une équation de la tangente à Γ en A.

b) Montrer que, au voisinage de l'origine, Γ est formée de deux arcs. Préciser les tangentes à l'origine de ces deux arcs. Préciser les positions de ces arcs par rapport à leurs tangentes à l'origine.

** **Exercice 7**: Soit Γ l'ensemble des points $M(x,y)$ tels que $x^{121} + x^{50}y^{31} + y^{54} + x^2 - y^2 = 0$. Montrer qu'au voisinage de l'origine Γ est une courbe à point double. Déterminer les tangentes au point double. Et montrer que le point double est d'inflexion pour chacune des deux branches.

En est il de même pour $C(x^{123} + x^{51}y^{31} + y^{18} + x^2 - y^2 = 0)$?

* **Exercice 8**: Les deux courbes $C(x^2 + xy + 3x + y = 0)$ et $\Gamma(xy + y^2 + x^2 + 3x + y = 0)$ passent par l'origine, où elles ont même tangente. Placer (au voisinage de l'origine) sur une même figure cette tangente et les deux courbes (en précisant leurs positions relatives).

* **Exercice 9**: Soit $F: \mathbb{R}^2 \to \mathbb{R}$ définie par $F(x,y) = x^3 + y^3 + x^2y - y$.

a) Tracer la "courbe" d'équation $\frac{\partial F}{\partial x}(x,y) = 0$.

b) Tracer la "courbe" d'équation $\frac{\partial F}{\partial y}(x,y) = 0$.

c) Déterminer les points stationnaires de F. Quels sont ceux qui sont des extrema ?

* **Exercice 10**: Soit $F: \mathbb{R}^2 \to \mathbb{R}$ définie par $F(x,y) = x^3 + y^3 + 3xy$.

a) Tracer la "courbe" d'équation $\frac{\partial F}{\partial x}(x,y) = 0$.

b) Tracer la "courbe" d'équation $\frac{\partial F}{\partial y}(x,y) = 0$.

c) Déterminer les points stationnaires de F. Quels sont ceux qui sont des extrema ?

** **Exercice 11**: Soit $F: \mathbb{R}^2 \to \mathbb{R}$ définie par $F(x,y) = x^3 + xy^2 + 2x^2y + 2y + x$.

a) Tracer la "courbe" d'équation $\frac{\partial F}{\partial x}(x,y) = 0$.

b) Tracer la "courbe" d'équation $\frac{\partial F}{\partial y}(x,y) = 0$.

c) Déterminer les points stationnaires de F. Quels sont ceux qui sont des extrema ?

d) Tracer la courbe γ d'équation $F(x,y) = 0$ (on pourra essayer de l'écrire sous la forme $y = \varphi(x)$)

Préciser les tangentes à γ à l'origine, et aux points d'intersection avec les axes.

* **Exercice 12**: On considère l'application F de $\mathbb{C}$ dans $\mathbb{R}$ qui à tout z associe $|\sin z|$. Soit U le complémentaire des points de la forme $k\pi$. Montrer que F est de classe C^1 sur U. Montrer que F n'a aucun extremum sur U.

* **Exercice 13**: Soit f l'application de $\mathbb{C}$ dans $\mathbb{R}$ définie par $f(z) = |z^3 + 3z|$

a) Montrer qu'elle est C^∞ en tout point où elle est non nulle. Calculer sa différentielle en $z_0 = x_0 + iy_0$.

b) Quels sont ses points stationnaires ? A-t-elle des extrema ?

*** **Exercice 14**: Dans un plan P muni d'un repère orthonormé $(O, \vec{i}, \vec{j})$, on donne les deux paraboles d'équations $y = x^2$ et $y = -x^2 - 2x - \frac{15}{8}$. Tracer ces deux courbes.

Soit $u \to M(u,u^2)$ un paramétrage de la première. Soit $v \to N(v, -v^2 - 2v - \frac{15}{8})$ un paramétrage de la seconde.

Déterminer les extréma de la fonction $(u,v) \to [\mathrm{dist}(M(u),N(v))]^2$. Montrer qu'il existe un point P_0 de la première et un point Q_0 de la seconde tels que, quels que soient P sur la première et Q sur la seconde, on ait $d(P_0,Q_0) \leq d(P,Q)$.

* **Exercice 15**: Sur le disque défini dans $\mathbb{R}^2$ par $x^2 + y^2 \leq 4$, on considère la fonction f définie par

$$f(x,y) = x^2 + 2y^2 + 6x + 3.$$

Déterminer les points stationnaires de f situés à l'intérieur de ce disque. Puis déterminer la plus petite valeur et la plus grande valeur prises par f sur ce disque.

** **Exercice 16:** On considère la fonction F: $\mathbb{R}^2 \to \mathbb{R}$ définie par $F(x,y) = x^3 + y^3 + xy^2 + x$. Montrer qu'ellle n'a aucun point stationnaire.

Au chapitre 10 §7 on a démontré que, sur le carré $C(|x|+|y| \leq 1)$, la fonction F est bornée et atteint ses bornes. Déterminer ces bornes et les points où elles sont atteintes.

** **Exercice 17** On considére dans $\mathbb{R}^3$ la surface S définie par $z = x^3 + y^3 + x^2y + 3y = 0$.

a) Quels sont les points de S où le plan tangent est parallèle au plan $\pi(5x + 7y - z = 0)$?

b) Quels sont les points où la courbure de S est positive ? négative ? nulle ?

* **Exercice 18:** Dans l'espace muni d'un repère orthonormé $(O,\vec{i},\vec{j},\vec{k})$ on considère la surface $\Sigma(z=x^4+y^3+xy+7x)$. Le point A(1;-2;-2) est sur Σ et sur la sphère S de centre l'origine et de rayon 3. Ecrire une fonction g(x,y) telle que, au voisinage de A, la sphére S s'écrive $z = g(x,y)$. Montrer que, au voisinage de A, $S\cap\Sigma$ est une courbe régulière, et déterminer sa tangente en A.

** **Exercice 19**: On considère la surface S d'équation $z = 2x^2 + 2xy + x + 1$.

a) Le point M(1,–1,2) est sur S. Quel est le plan π tangent à S en M ? Montrer que l'intersection de π et de S est formée de deux droites (on donnera des vecteurs directeurs de ces droites).

b) Dessiner les projections sur le plan des (x,y) des intersections de S et des plans parallèles à π et voisins de π.

c) Montrer que, quel que soit le point $P(x_0,y_0,z_0)$ de S, le plan tangent à S en P coupe S suivant deux droites.

** **Exercice 20**:

Dans un plan muni d'un repère orthonormé $(O,\vec{i},\vec{j})$, on considère la courbe Γ d'équation $x^2 + 2xy + 4x + 2y = 0$. Trouver un repère orthonormé $(O,\vec{I},\vec{J})$ dans le quel l'équation de Γ est de la forme $AX^2 + BY^2 + CX + DY = 0$. Trouver le centre de symétrie et les axes de symétrie de Γ. Tracer Γ.

* **Exercice 21:** Déterminer les points stationnaires des fonctions suivantes de $\mathbb{R}^3$ dans $\mathbb{R}$. Parmi ces points quels sont ceux qui sont des extréma ?

$$f(x,y,z) = x^3 + 3xy^2 + z^3 + 3xy$$
$$g(x,y,z) = x^3 + z^3 + xyz + x^2 - z^2$$
$$h(x,y,z) = 2xyz + x^2 + y^2 + z^2$$

Etude géométrique des courbes du plan et de l'espace

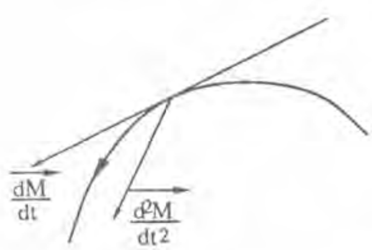
On désigne par P un plan orienté. Pour la commodité on y choisit une fois pour toutes un repère orthonormé direct $(O, \vec{i}, \vec{j})$. De même on choisit une fois pour toutes un repère orthonormé direct $(O, \vec{i}, \vec{j}, \vec{k})$ de l'espace. Les normes considérées sont les normes de la géométrie élémentaire.

§ 1 : Courbes - Points réguliers

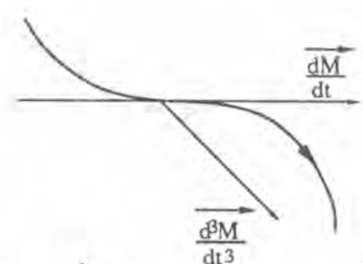
On appellera courbe paramétrée toute application de classe C^∞ d'un intervalle I dans P, ou dans l'espace. L'image Γ de ladite application est la courbe, l'application elle-même est le paramétrage. Deux telles applications peuvent avoir même image; autrement dit une courbe peut avoir plusieurs paramétrages.

Nous appellerons point régulier du paramétrage, tout point $M(t_0)$ tel que $\frac{d\vec{M}}{dt}(t_0) \neq 0$ (nous noterons $\frac{d\vec{M}}{dt}$, plutôt que $\frac{d\vec{OM}}{dt}$, le vecteur dérivée première).

En un point régulier du paramétrage, une courbe plane présente l'allure suivante:

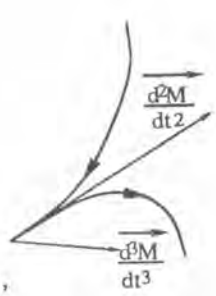


Si $\frac{d^2\vec{M}}{dt}$ n'est pas colinéaire à $\frac{d\vec{M}}{dt}$



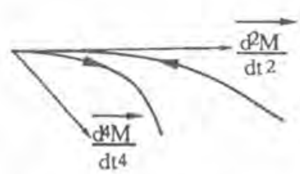
Si $\frac{d^2\vec{M}}{dt^2} = \lambda \frac{d\vec{M}}{dt}$, et si $\frac{d^3\vec{M}}{dt^3}$ n'est pas colinéaire à $\frac{d\vec{M}}{dt}$

En un point singulier, la courbe peut avoir l'une des allures suivantes



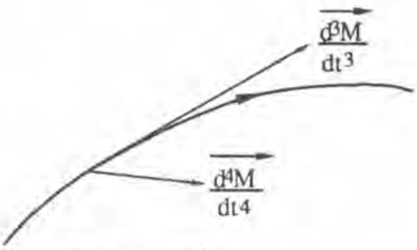
Si $\frac{d^2\vec{M}}{dt^2} \neq 0$,

et si $\frac{d^3\vec{M}}{dt^3}$ n'est pas colinéaire à $\frac{d^2\vec{M}}{dt^2}$



Si $\frac{d^2\vec{M}}{dt^2} \neq 0$, si $\frac{d^3\vec{M}}{dt^3} = \lambda \frac{d^2\vec{M}}{dt^2}$

et si $\frac{d^4\vec{M}}{dt^4}$ n'est pas colinéaire à $\frac{d^2\vec{M}}{dt^2}$



Si $\frac{d^2\vec{M}}{dt^2} = 0$, $\frac{d^3\vec{M}}{dt^3} \neq 0$,

et si $\frac{d^4\vec{M}}{dt^4}$ n'est pas colinéaire à $\frac{d^3\vec{M}}{dt^3}$

Si $\frac{d^2\vec{M}}{dt^2} = 0$, $\frac{d^3\vec{M}}{dt^3} \neq 0$, $\frac{d^4M}{dt^4} = \lambda \frac{d^3\vec{M}}{dt^3}$

et si $\frac{d^5\vec{M}}{dt^5}$ n'est pas colinéaire à $\frac{d^3\vec{M}}{dt^3}$

§ 2 : Changements de paramétrage

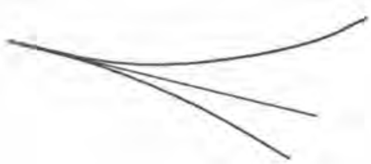
Considérons une courbe Γ définie comme l'image de l'intervalle I par l'application $t \rightarrow M(t)$. Une application bijective φ (de classe C^∞) de l'intervalle J dans I, permet de définir un nouveau paramétrage de la courbe $\theta \rightarrow M(\varphi(\theta))$.

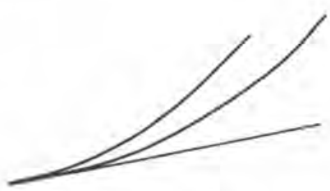
Notons qu'il est possible qu'un point soit régulier pour le paramétrage initial, et ne le soit plus pour le nouveau paramétrage. Par exemple la courbe plane ($x = \text{sh } t$, $y = \text{ch } t$) est régulière en tout point; mais ($x = \text{ch } t^5$, $y = \text{sh } t^5$) définit la même courbe, et a une singularité pour $t = 0$.

Nous dirons qu'un point est géométriquement régulier s'il existe un paramétrage pour lequel il est régulier.

En un point géométriquement régulier, une courbe plane a l'une des allures suivantes.

Des points tels que ceux ci



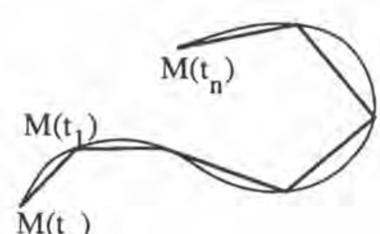


ne sont réguliers pour aucun paramétrage.

§ 3 : Longueur des courbes

Soit $t \rightarrow M(t)$ ($t \in I$) une courbe et α et β dans I, on appelle longueur de l'arc $M(\alpha)M(\beta)$ de cette courbe le nombre :

$$\ell_{(\alpha,\beta)} = \left| \int_\alpha^\beta \left\| \frac{d\vec{M}}{dt}(t) \right\| \; dt \right|$$



Ceci correspond à la notion physique de longueur, en effet : soit $t_0=\alpha<t_1<\ldots<t_n=\beta$, notons $\lambda(t_0,\ldots,t_n)$ la longueur de la ligne polygonale $M(t_0)M(t_1)\ldots M(t_n)$.

$$\lambda(t_0,\ldots,t_n) = \sum_{i=1}^{n} \|\overrightarrow{M(t_{i-1})M(t_i)}\| \; .$$

On démontre (Ex a) que $(\forall\varepsilon>0)$ $(\exists\eta>0)$ tel que $\sup|t_i-t_{i-1}|\leq\eta$ entraine $|\ell(\alpha,\beta)-\lambda(t_0,\dots,t_n)|\leq\varepsilon$.

Autrement dit $\ell(\alpha,\beta)$ est la limite des longueurs de ces lignes polygonales lorsque les t_i-t_{i-1} tendent vers 0.

Notons que $\ell(\alpha,\beta)$ ne dépend que de la courbe et des points $M(\alpha)M(\beta)$ (mais ne dépend pas du paramétrage utilisé pour faire le calcul). Ceci résulte du calcul que l'on vient d'esquisser. Pour un changement de paramétrage du type $\big(t\to M(t)\big)\to\big(\theta\to M(\varphi(\theta))\big)$ avec φ de classe C^1, ceci résulte aussi du théorème de changement de variable dans les intégrales :

$$\int_\alpha^\beta \left\|\frac{d\overrightarrow{M}}{dt}(t)\right\| dt = \int_{\varphi^{-1}(\alpha)}^{\varphi^{-1}(\beta)} \left\|\frac{d\overrightarrow{M}}{dt}(\varphi(\theta))\right\| \varphi'(\theta)\,d\theta = \int_{\varphi^{-1}(\alpha)}^{\varphi^{-1}(\beta)} \left\|\frac{d(\overrightarrow{M\circ\varphi})}{d\theta}(\theta)\right\| d\theta$$

si φ' est positif, si non les deux dernières intégrales ne sont égales qu'au signe près.

Exercice a : 1) Soit $\vec{\varphi}:[a,b]\to\vec{P}$ une application de classe C^1. On pose $f(t)=\|\vec{\varphi}(t)\|$. Montrer que, si $\vec{\varphi}$ ne s'annule pas, f est de classe C^1. Puis montrer que $(\forall t)$: $\left|\frac{df}{dt}(t)\right|\leq\left\|\frac{d\vec{\varphi}}{dt}(t)\right\|$ (On pourra dériver l'égalité $f^2=\vec{\varphi}^2$).

2)Soit une courbe $t\to M(t)$ $(t\in[a,b])$ de classe C^2.Puisque toute fonction continue sur un intervalle fermé borné est majorée , on peut affirmer qu'il existe m tel que $(\forall t\in[a,b])$ $\left\|\frac{d^2\overrightarrow{M}}{dt^2}(t)\right\|\leq m$. Justifier que , quels que soient u et v dans $[a,b]$, on a les inégalités :

$$\left\|\overrightarrow{M(u)M(v)}-(v-u)\frac{d\overrightarrow{M}}{dt}(u)\right\|\leq\frac{m(v-u)^2}{2}\quad\text{(formule de Taylor)}$$

$$\left|\ell_{u,v}-(v-u)\left\|\frac{d\overrightarrow{M}}{dt}(u)\right\|\right|\leq\frac{m(v-u)^2}{2}\quad\text{(formule de Taylor)}$$

$$\left|\ell_{u,v}-\|\overrightarrow{M(u)M(v)}\|\right|\leq m(v-u)^2\quad\text{(utiliser les deux précédentes)}$$

3) Soit $t_i=a+i\,\frac{b-a}{n}$ $(i=0,\dots,n)$. Soit λ_n la longueur de la ligne polygonale $[M(t_0)M(t_1)\dots M(t_n)]$. Démontrer que $|\lambda_n-\ell_{a,b}|\leq\frac{1}{n}m(b-a)^2$. En déduire que , quel que soit $\varepsilon>0$, il existe N tel que $(n\geq N)$ implique $|\lambda_n-\ell_{a,b}|\leq\varepsilon$.

§ 4 : Paramétrages intrinsèques - Orientations d'une courbe

Un paramétrage $s\to M(s)$ d'une courbe Γ est dit intrinsèque si $(\forall s)$ $\frac{d\overrightarrow{M}}{ds}(s)$ est de norme 1.

Construction d'un paramétrage intrinsèque d'une courbe régulière

Soit Γ une courbe régulière, et soit $t\to M(t)$ $(I\to\Gamma)$ un paramétrage partout régulier. Posons

$$\psi(t)=\int_{t_0}^{t}\left\|\frac{d\overrightarrow{M}}{dt}(u)\right\| du$$

L'application ψ est de classe C^∞ et sa dérivée est toujours strictement positive (parceque $\frac{d\overrightarrow{M}}{ds}$ ne s'annule pas). Elle définit donc une bijection de l'intervalle I sur un intervalle J. Soit φ la bijection inverse. Alors le paramétrage $s\to M(\varphi(s))$ est intrinsèque. En effet :

$$\frac{d\overrightarrow{M}}{ds}(s)=\frac{d\varphi}{ds}(s)\,\frac{d\overrightarrow{M}}{dt}(\varphi(s))$$

donc
$$\left\|\frac{d\overrightarrow{M}}{ds}(s)\right\| = \left|\frac{d\psi}{dt}(\varphi(s))\right|^{-1}\left\|\frac{d\overrightarrow{M}}{dt}(\varphi(s))\right\| = 1$$

Exercice b : On considère la courbe $(x(t) = \cos t + \cos^2 t \; ; \; y(t) = \sin t + \sin t \cos t)$ $(t \in]-\pi,\pi[)$. Construire un paramétrage intrinsèque de cette courbe. Tracer l'arc correspondant à $[\pi/2, 3\pi/2]$; pourquoi ne possède-t-il pas de paramétrage intrinsèque?

Les différents paramétrages intrinsèques

Soit $s \to M(s)$ une courbe donnée en paramétrage intrinsèque. Faisons un changement de paramétrage du type $s = \pm\sigma + \text{cste}$; nous en obtenons un nouveau. On démontre qu'il n'y a pas d'autres paramétrages intrinsèques que ceux que l'on obtient ainsi.

Il y a donc 2 familles de paramétrages intrinsèques. Chacune correspond à un sens de parcours (ou orientation) de la courbe. Dans la pratique, on travaillera sur des courbes orientées, c'est-à-dire qu'on éliminera l'une de ces deux familles de paramétrages.

Notation: Soit $M(s)$ un point d'une courbe paramétrée intrinsèquement. On note $\overrightarrow{T(s)}$ le vecteur $\frac{d\overrightarrow{M}}{ds}(s)$; on l'appelle le vecteur unitaire tangent à la courbe au point considéré. Notons que ce vecteur dépend de l'orientation choisie sur la courbe; l'orientation opposée donnerait le vecteur opposé.

§ 5 : Centre de courbure et plan osculateur.

Centre de courbure.

Le vecteur $\frac{d\vec{T}}{ds}$ $\left(= \frac{d^2M}{ds^2}\right)$ est orthogonal à $\vec{T}$ (car $2\vec{T}.\frac{d\vec{T}}{ds} = \frac{d}{ds}(\vec{T}.\vec{T}) = \frac{d}{ds}(1) = 0$); il est donc soit nul, soit non nul et perpendiculaire à $\vec{T}$. Soit M le point de paramètre s_0:

* *Si $\frac{d\vec{T}}{ds}(s_0)$ est nul, nous dirons que M est un point de courbure nulle* (*) *de l'arc considéré.*

* *Si $\frac{d\vec{T}}{ds}(s_0)$ est non nul, nous appellerons centre de courbure de la courbe en M, le point C défini par*

$$\overrightarrow{MC} = \left\|\frac{d\vec{T}}{ds}(s_0)\right\|^{-2} \frac{d\vec{T}}{ds}(s_0)$$

Ce point C ne dépend pas du sens de parcours, puisque $\vec{T}$ en dépend, mais que $\frac{d\vec{T}}{ds}$ n'en dépend pas.

Exercice b: Soit C un cercle de rayon R et de centre O. On choisit un repère orthonormé $(O, \vec{i}, \vec{j}, \vec{k})$ tel que $\vec{i}$ et $\vec{j}$ soient dans le plan du cercle. Montrer que $s \to (R \cos \frac{s}{R}, R \sin \frac{s}{R}, 0)$ est un paramétrage intrinsèque de C. Montrer que en tout point M de C, le centre de courbure est O.

Le plan osculateur

Le plan osculateur en un point M où la courbure n'est pas nulle, est le plan passant par M et parallèle aux vecteurs $\vec{T}$ et $\frac{d\vec{T}}{ds}$. C'est le plan qui contient la tangente et le centre de courbure. Si la courbe est plane, c'est le plan de la courbe.

(*) On dit aussi un point méplat.

Conventions dans le plan.

Dans le plan (orienté) P, on a coutume de noter $\vec{N}(s_0)$ le vecteur unitaire directement perpendiculaire à $\vec{T}(s_0)$. On l'appelle le vecteur normal principal en M à la courbe. Puisque le vecteur $\frac{d\vec{T}}{ds}(s_0)$ est orthogonal à $\vec{T}(s_0)$, il s'écrit

$$\frac{d\vec{T}}{ds}(s_0) = \rho\, \vec{N}(s_0)$$

Le scalaire ρ est appelé la courbure de l'arc considéré en M (il est nul si $\frac{d\vec{T}}{ds}(s_0) = 0$). Et, s'il n'est pas nul, nous avons $\overrightarrow{MC} = \frac{1}{\rho}\vec{N}(s_0)$.

Exercice c: En voiture, vous prenez un virage, dans quel sens tournez vous le volant si la courbure du virage est positive ? Négative ?

Conventions dans l'espace.

Si $\frac{d\vec{T}}{ds}(s_0) \neq 0$, nous appellerons courbure en M de l'arc considéré le scalaire $\rho = \left\| \frac{d\vec{T}}{ds}(s_0) \right\|$; et, si $\rho \neq 0$, nous noterons $\vec{N}(s_0)$ le vecteur $\frac{1}{\rho}\frac{d\vec{T}}{ds}(s_0)$; c'est un vecteur unitaire orthogonal à $\vec{T}(s_0)$, que nous appellerons vecteur normal principal en M à la courbe. Nous avons encore $\overrightarrow{MC} = \frac{1}{\rho}\vec{N}(s_0)$. Les vecteurs $\vec{T}(s_0)$ et $\vec{N}(s_0)$ forment une base orthonormée du plan osculateur.

Remarquons que dans l'espace la courbure est toujours positive ou nulle; dans le plan elle est positive si $(\vec{T}(s_0), \frac{d\vec{T}}{ds}(s_0)) = +\frac{\pi}{2}$, et négative si $(\vec{T}(s_0), \frac{d\vec{T}}{ds}(s_0)) = -\frac{\pi}{2}$. Les deux conventions sont donc très différentes, bien que les notations usuelles soient les mêmes.

Dans le plan orienté la convention que nous avons faite a l'avantage de distinguer les courbes "qui tournent à droite", de celles "qui tournent à gauche"; et ainsi de faire jouer un rôle essentiel à la caractéristique de la courbe la plus banale, celle que l'on voit immédiatement sur le dessin. Dans l'espace, en l'abscence d'une orientation du plan osculateur de la courbe au point M, il est impossible de distinguer les deux vecteurs unitaires de ce plan qui sont orthogonaux à $\vec{T}$; c'est pourquoi notre convention porte sur le signe de ρ. D'ailleurs, dans l'espace on ne peut plus distinguer les courbes qui tournent à droite de celles qui tournent à gauche, cela dépend de la position de l'observateur par rapport à la courbe, ou plutôt par rapport à son plan osculateur.

Approximation d'une courbe par sa tangente, par son cercle de courbure.

Notons que le développement limité à l'ordre 2 (en s_0) de la fonction $s \to M(s)$ s'écrit

$$\overrightarrow{OM}(s_0+h) = \overrightarrow{OM}(s_0) + h\, \vec{T}(s_0) + \frac{h^2}{2}\rho(s_0)\vec{N}(s_0) + \vec{o}(h^2)$$

Le point m(h) défini par $\overrightarrow{Om}(h) = \overrightarrow{OM}(s_0) + h\, \vec{T}(s_0)$ est un point de la tangente en $M(s_0)$. La distance $m(h)M(s_0+h)$ est $\left\| \frac{h^2}{2}\rho(s_0)\vec{N}(s_0) + \vec{o}(h^2) \right\|$, elle est $o(h)$. *A fortiori la distance de $M(s_0+h)$ à sa tangente en $m(s_0)$ est $o(h)$.*

Si la courbure est nulle au point $M(s_0)$, la distance $m(h)M(s_0+h)$ est $o(h^2)$; et à fortiori la distance de $M(s_0+h)$ à la tangente est $o(h^2)$. On dit alors que la tangente est osculatrice en $M(s_0)$ à la courbe .

Autrement dit la tangente approche beaucoup mieux la courbe en un point de courbure nulle qu'en un point de courbure non nulle. Notons que si la courbe est plane, un point $M(s_0)$ où la courbure est nulle est en général (c'est-à-dire si $\frac{d^3M}{ds^3}(s_0)$ existe et est non nulle) un point d'inflexion (cf: §1).

On appelle cercle de courbure au point $M(s_0)$, le cercle Γ centré au centre de courbure $C(s_0)$, passant par $M(s_0)$ et situé dans le plan osculateur. On démontre que la distance de $M(s_0+h)$ au cercle osculateur (c'est à dire $\inf_{m\in\Gamma} mM(s_0+h)$) est $o(h^2)$.

Exercice d: Vérifier que $t \to \overrightarrow{O\mu(t)} = \overrightarrow{OM}(s_0) + \frac{1}{\rho(s_0)}\left[(1-\cos\rho(s_0)t)\ \overrightarrow{N(s_0)} + \sin\rho(s_0)t\ \overrightarrow{T(s_0)}\right]$ est un paramétrage intrinsèque du cercle de courbure en $M(s_0)$. Ecrire le développement à l'ordre 2 de la fonction $t \to \overrightarrow{O\mu(t)}$ pour $t = 0$.

En déduire que $\overrightarrow{O\mu(h)} - \overrightarrow{OM}(s_0+h) = \overrightarrow{o(h^2)}$, et que $\inf_{m\in\Gamma} mM(s_0+h) = o(h^2)$

On peut condidérer que la tangente est en quelque sorte un développement limité à l'ordre 1 géométrique de la courbe. Si la courbure est nulle c'est même un développement à l'ordre 2. En un point de courbure non nulle c'est le cercle de courbure qui joue le rôle de développement limité géométrique à l'ordre 2.

§ 6 : Calcul de la courbure

Le § 5 nous donne un moyen de déterminer $\rho(s)$ et $C(s)$ quand la courbe est définie en paramétrage intrinsèque. Malheureusement il est le plus souvent impossible de construire un paramétrage intrinsèque explicite d'une courbe donnée. En effet la méthode donnée au § 4 est toute théorique, elle utilise deux opérations (calcul d'une primitive, et calcul de l'inverse d'une bijection donnée) que l'on ne sait pas effectuer explicitement (sauf bien entendu dans quelques cas particuliers). Il est donc nécessaire de trouver des méthodes permettant de calculer ρ et C quand on ne dispose pas d'un paramétrage intrinsèque explicite.

Supposons donc donnée une courbe régulière avec un paramétrage quelconque $t \to M(t)$. Nous savons (§ 4) qu'il existe un paramétrage intrinsèque $s \to M(s)$ de cette courbe, qui définit la même orientation que le paramétrage initial. Nous avons alors :

$$\frac{\overrightarrow{dM}}{ds} = \frac{dt}{ds}\frac{\overrightarrow{dM}}{dt}$$

d'où $\quad \frac{dt}{ds} = \left\|\frac{\overrightarrow{dM}}{dt}\right\|^{-1} \quad$ et $\quad \overrightarrow{T} = \left\|\frac{\overrightarrow{dM}}{dt}\right\|^{-1} \frac{\overrightarrow{dM}}{dt} = \frac{dt}{ds}\frac{\overrightarrow{dM}}{dt}$

Puis $\quad \frac{d^2\overrightarrow{M}}{ds^2} = \frac{d}{ds}\left(\frac{dt}{ds}\right)\frac{\overrightarrow{dM}}{dt} + \left(\frac{dt}{ds}\right)^2\frac{d^2\overrightarrow{M}}{dt^2}$

soit $\quad \frac{d^2\overrightarrow{M}}{ds^2} = \frac{dt}{ds}\left[\frac{d}{dt}\left(\frac{dt}{ds}\right)\right]\frac{\overrightarrow{dM}}{dt} + \left(\frac{dt}{ds}\right)^2\frac{d^2\overrightarrow{M}}{dt^2}$

<u>Exemple 1</u>: Soit la courbe plane définie par $(x = \cos^3 t\ ;\ y = \sin t^3 t)$ $(0 < t < \frac{\pi}{2})$.

$\frac{\overrightarrow{dM}}{dt} = (-3\cos^2 t\ \sin t;\ 3\sin^2 t\ \cos t)$ $\qquad$ donc $\qquad \left\|\frac{\overrightarrow{dM}}{dt}\right\| = 3\cos t\ \sin t = \frac{ds}{dt}$

En choisissant pour orientation de la courbe, celle qui est définie par le paramètre t.

$\overrightarrow{T} = (-\cos t;\ \sin t)$ $\qquad$ donc $\qquad \overrightarrow{N} = (-\sin t;\ -\cos t)$

$$\frac{d^2\vec{M}}{ds^2} = \frac{d\vec{T}}{ds} = \frac{dt}{ds}\frac{d\vec{T}}{dt} = \frac{1}{3\cos t\sin t}(\sin t\,;\cos t) \qquad \text{donc} \quad \rho = \frac{-1}{3\cos t\sin t}$$

Le centre de courbure au point de paramètre t est donné par $\vec{MC} = -3\cos t\ \sin t\,(-\sin t, -\cos t)$.

Donc C est le point de coordonnées $(\cos^3 t + 3\cos t\sin^2 t,\ \sin^3 t + 3\sin t\cos^2 t)$.

<u>Exemple 2</u>: Soit la courbe de l'espace définie par $(x = \sin t;\ y = \cos t;\ z = \mathrm{ch}\, t)$.

$$\frac{d\vec{M}}{dt} = (\cos t; -\sin t; \mathrm{sh}\, t) \qquad \text{donc} \quad \left\|\frac{d\vec{M}}{dt}\right\| = \mathrm{ch}\, t = \frac{ds}{dt}$$

En choisissant pour orientation de la courbe, celle qui est définie par le paramètre t, on a

$$\vec{T} = \frac{1}{\mathrm{ch}\, t}(\cos t; -\sin t; \mathrm{sh}\, t)$$

$$\frac{d^2\vec{M}}{ds^2} = \frac{dt}{ds}\frac{d\vec{T}}{dt} = \frac{d\vec{T}}{ds} = \frac{1}{\mathrm{ch}^3 t}(-\mathrm{ch}\, t\ \sin t - \mathrm{sh}\, t\cos t; -\mathrm{ch}\, t\cos t + \mathrm{sh}\, t\sin t; 1)$$

Donc $\rho = \left\|\frac{d\vec{T}}{ds}\right\| = \frac{\sqrt{2}}{\mathrm{ch}^2 t}$ et $\vec{N} = \frac{1}{\sqrt{2}\ \mathrm{ch}\, t}(-\mathrm{ch}\, t\ \sin t - \mathrm{sh}\, t\cos t; -\mathrm{ch}\, t\cos t + \mathrm{sh}\, t\sin t; 1)$

D'où $\vec{MC} = \frac{\mathrm{ch}\, t}{2}(-\mathrm{ch}\, t\ \sin t - \mathrm{sh}\, t\cos t; -\mathrm{ch}\, t\cos t + \mathrm{sh}\, t\sin t; 1)$.

Quelques formules dans le plan: Supposons que, dans un repère orthonormé direct, les coordonnées de M(t) soient x(t) et y(t). Alors:

$$\frac{d\vec{M}}{dt} = (x', y') \qquad \text{donc} \quad \frac{ds}{dt} = \sqrt{x'^2+y'^2} \qquad \text{et} \qquad \vec{T} = \frac{1}{\sqrt{x'^2+y'^2}}(x', y')$$

d'où
$$\vec{N} = \frac{1}{\sqrt{x'^2+y'^2}}(-y', x')$$

Puis $\frac{d\vec{T}}{dt} = \left(\frac{y'(y'x''-x'y'')}{(x'^2+y'^2)^{3/2}}, \frac{x'(x'y''-y'x'')}{(x'^2+y'^2)^{3/2}}\right)$ donc $\frac{d\vec{T}}{ds} = \frac{y'x''-x'y''}{(x'^2+y'^2)^2}(y', -x')$

Et enfin
$$\rho = \frac{x'y''-x''y'}{(x'^2+y'^2)^{3/2}}$$

Exercice e : Soit f une fonction de classe C^2. Soit Γ le graphique de f en repère orthonormé. Soit M le point $(0,f(0))$. On suppose que le développement limité de f à l'origine s'écrit : $f(t) = a + bt + ct^2 + o(t^2)$.

1) Calculer en fonction de a,b,c les vecteurs $\vec{T}$ et $\vec{N}$ de Γ en M; calculer la courbure de Γ en M.

2) Calculer en fonction de a,b,c les coordonnées du centre de courbure de Γ en M.

Quelques formules dans l'espace:

$$\frac{d\vec{M}}{dt} = (x',y',z') \qquad \text{donc} \quad \frac{ds}{dt} = \sqrt{x'^2+y'^2+z'^2} \quad \text{et} \qquad \vec{T} = \frac{1}{\sqrt{x'^2+y'^2+z'^2}}(x',y',z')$$

$$\vec{T}\wedge\rho\,\vec{N} = \frac{d\vec{M}}{ds}\wedge\frac{d^2\vec{M}}{ds^2} = \left(\frac{dt}{ds}\right)^3\frac{d\vec{M}}{dt}\wedge\frac{d^2\vec{M}}{dt^2}$$

Donc, puisque $\rho>0$ et $\|\vec{T}\wedge\vec{N}\| = 1$, $\rho = \left(\frac{dt}{ds}\right)^3\left\|\frac{d\vec{M}}{dt}\wedge\frac{d^2\vec{M}}{dt^2}\right\|$.

On notera qu'en normant le vecteur $\frac{d\vec{M}}{ds}\wedge\frac{d^2\vec{M}}{ds^2}$ on obtient *le vecteur* $\vec{B} = \vec{T}\wedge\vec{N}$ *normal au plan osculateur (et appelé vecteur binormal à la courbe au point considéré). Les vecteurs* $\vec{T}, \vec{N}, \vec{B}$

forment une base orthonormée directe (appelée repère de Serret-Frenet de la courbe au point considéré). Ainsi la connaissance de $\vec{T}$ et $\vec{B}$ nous redonne $\vec{N}$ grâce à la formule $\vec{N} = \vec{B} \wedge \vec{T}$.

§ 7 : Torsion d'une courbe de l'espace.

Considérons un point M_0 d'un arc régulier muni d'un paramétrage intrinsèque. Supposons que M_0 ne soit pas un point de courbure nulle.

Les trois vecteurs $\vec{T}, \vec{N}, \vec{B}$ sont des fonctions dérivables de s (si $s \to M(s)$ est au moins 3 fois dérivable, ce que nous avons supposé). Le vecteur $\frac{d\vec{B}}{ds}$ est colinéaire à $\vec{N}$; car

$$\vec{B}.\frac{d\vec{B}}{ds} = \frac{1}{2}\frac{d}{ds}(\vec{B}.\vec{B}) = 0 \quad \text{et} \quad \vec{T}.\frac{d\vec{B}}{ds} = \frac{d}{ds}(\vec{T}.\vec{B}) - \frac{d\vec{T}}{ds}.\vec{B} = 0 - \rho\,\vec{N}.\vec{B} = 0.$$

Il existe donc un scalaire $\tau(s_0)$ tel que $\frac{d\vec{B}}{ds}(s_0) = \tau(s_0)\,\vec{N}(s_0)$; on l'appelle la torsion (*) de Γ en M_0.

Pour calculer $\frac{d\vec{N}}{ds}$, calculons ses trois coordonnées dans le repère de Serret-Frenet. Ce sont :

$$\frac{d\vec{N}}{ds}.\vec{T} = -\vec{N}\frac{d\vec{T}}{ds} = -\rho \qquad (\text{car } \vec{N}.\vec{T} = \text{cste})$$

$$\frac{d\vec{N}}{ds}.\vec{N} = 0 \qquad (\text{car } \vec{N}^2 = \text{cste})$$

$$\frac{d\vec{N}}{ds}.\vec{B} = -\vec{N}\frac{d\vec{B}}{ds} = -\tau \qquad (\text{car } \vec{N}.\vec{B} = \text{cste})$$

Donc
$$\frac{d\vec{N}}{ds} = -\rho\vec{T} - \tau\vec{B}$$

On retiendra les 3 formules de Serret Frenet

$$\frac{d\vec{T}}{ds} = \rho\vec{N} \qquad \frac{d\vec{B}}{ds} = \tau\vec{N} \qquad \text{et} \qquad \frac{d\vec{N}}{ds} = -\rho\vec{T} - \tau\vec{B}$$

Remarques: 1) Les vecteurs $\vec{T}$ et $\vec{B}$ sont multipliés par -1 si l'on change le sens de parcours. Tandis que $\vec{N}$ est indépendant du sens de parcours.

2) Des formules de Serret-Frenet on déduit

$$\frac{d^3M}{ds^3} = \frac{d}{ds}(\rho\vec{N}) = \frac{d\rho}{ds}\vec{N} - \rho^2\,\vec{T} - \rho\,\tau\,\vec{B}$$

Ce qui nous donne le développement de Taylor à l'ordre 3 :

$$\vec{OM}(s_0+h) = \vec{OM}(s_0) + h\vec{T} + \frac{h^2}{2}\rho\vec{N} + \frac{h^3}{6}\Big[\frac{d\rho}{ds}\vec{N} - \rho^2\vec{T} + \rho\tau\vec{B}\Big] + o(h^3)$$

§ 8 : Projections d'une courbe de l'espace sur un plan.

Considérons une courbe Γ (donnée en paramétrage intrinsèque), un point M_0 de Γ où la courbure n'est pas nulle, et la projection orthogonale π sur un plan P (on notera Δ la direction de droite perpendiculaire à P). Essayons de déterminer l'allure de la courbe $\gamma = \pi(\Gamma)$, au point $m_0 = \pi(M_0)$.

Cas général: La direction Δ n'est parallèle à aucun des vecteurs $\vec{T_0}, \vec{N_0}, \vec{B_0}$ du trièdre de Serret-Frenet du point M_0.

(*) Certains auteurs prennent une convention de signe opposée, et appellent torsion de Γ en M, le nombre $-\tau$.

Projetons le developpement limité ci-dessus:

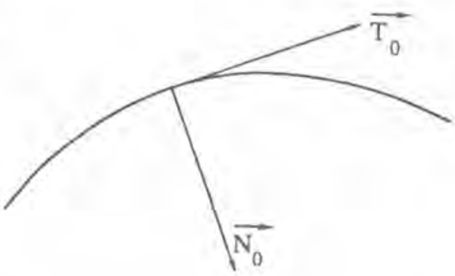
$$\overrightarrow{om}(s_0+h) = \overrightarrow{om_0} + h\pi(\overrightarrow{T_0}) + \frac{h^2}{2}\rho\pi(\overrightarrow{N_0}) + \frac{h^3}{6}\left[\frac{d\rho}{ds}\pi(\overrightarrow{N}) - \rho^2\pi(\overrightarrow{T}) + \rho\tau\pi(\overrightarrow{B})\right] + o(h^3)$$

α) $\pi(\overrightarrow{T_0}) \neq 0$ donc le point m_0 est régulier sur γ.

β) Si $\pi(\overrightarrow{N_0})$ n'est pas colinéaire à $\pi(\overrightarrow{T_0})$ (c'est à dire si $\Delta, \overrightarrow{N_0}$, et $\overrightarrow{T_0}$ ne sont pas coplanaires), la courbure de γ en m_0 est non nulle; et on sait que le sens de la concavité de la courbe est donné par le second terme de ce développement. La projection du centre de courbure de Γ n'est pas le centre de courbure de γ, mais ces deux points sont du même côté de la tangente à γ en m_0.

γ) Si $\pi(\overrightarrow{N_0})$ est colinéaire à $\pi(\overrightarrow{T_0})$ (c'est à dire si $\Delta, \overrightarrow{N_0}$, et $\overrightarrow{T_0}$ sont coplanaires), la courbure de γ en m_0 est nulle. Le point m_0 est un point d'inflexion de γ, pourvu que la torsion τ soit non nulle (car $\pi(\overrightarrow{B})$ est alors égal à $\overrightarrow{B}$).

Cas où Δ est parallèle à B_0: On obtient le développement à l'ordre 2 suivant:

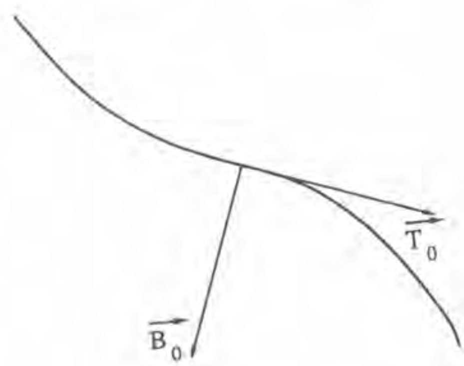


$$\overrightarrow{om}(s_0+h) = \overrightarrow{om_0} + h\overrightarrow{T_0} + \frac{h^2}{2}\rho\overrightarrow{N_0} + o(h^2).$$

Donc $\overrightarrow{T_0}$ est le vecteur unitaire tangent à γ en m_0, $\overrightarrow{N_0}$ est (au signe près) le vecteur normal de γ, et ρ est (au signe près) la courbure de γ. En particulier le centre de courbure de Γ en M_0, se projette sur le centre de courbure de γ en m_0.

Cas où Δ est parallèle à N_0: Dans ce cas on obtient le développement de l'ordre 3 suivant :

$$\overrightarrow{om}(s_0+h) = \overrightarrow{om_0} + h\overrightarrow{T_0} + \frac{h^3}{6}(-\rho^2\overrightarrow{T_0} - \rho\tau\overrightarrow{B_0}) + o(h^3)$$



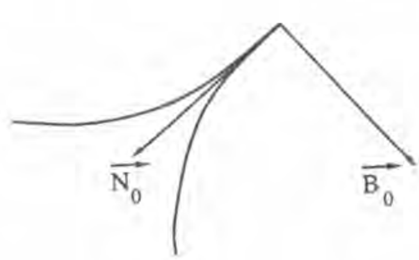
Donc le point m_0 est régulier sur γ; et c'est un point d'inflexion chaque fois que τ n'est pas nulle. La courbure de γ en m_0 est nulle.

Cas où Δ est parallèle à T_0:

Dans ce cas on obtient le développement à l'ordre 3 suivant:

$$\overrightarrow{om}(s_0+h) = \overrightarrow{om_0} + \frac{h^2}{2}\rho\overrightarrow{N_0} + \frac{h^3}{6}\left(\frac{d\rho}{ds}\overrightarrow{N_0} - \rho\tau\overrightarrow{B_0}\right) + o(h^3)$$

Puisque ρ est non nul, on obtient un point de rebroussement (la tangente de rebroussement étant parallèle à $\overrightarrow{N_0}$). Et si $\tau \neq 0$, ce point de rebroussement est de 1ère espèce.



Exercices sur le chapitre 17

* **Exercice 1:** Tracer les courbes définies paramétriquement par

(Γ_1) $x = \frac{t+2}{t\,(t-1)}$ et $y = \frac{t}{1-t^2}$ (Γ_2) $x = \cos t$ et $y = \cos\frac{t}{4} + \sin\frac{t}{4}$

(Γ_3) $x = \frac{t^2}{(t+1)\,(t-2)}$ et $y = \frac{t^2(t+2)}{t+1}$ (Γ_4) $x = \frac{t^2}{t+1}$ et $y = \frac{t^3}{t+1}$

* **Exercice 2:** Tracer les courbes suivantes, qui sont définies de façon polaire (r est le rayon polaire ; θ est l'angle polaire) On précisera leurs éventuelles asymptotes et leurs éventuelles tangentes à l'origine.

(Γ_1) $r = \sqrt{\sin 2\theta}$ (Γ_2) $r = \frac{1}{\sqrt{\cos 2\theta}}$

(Γ_3) $r = 1 + \cos 2\theta$ (Γ_4) $r = \frac{1}{1 + \cos(\theta + \frac{\pi}{4})}$

(Γ_5) $r = \frac{1}{2 + \cos(\theta - \frac{\pi}{3})}$ (Γ_6) $r = \frac{1}{1 + 2\sin(\theta + \frac{\pi}{3})}$

* **Exercice 3** : Trouver le centre de courbure, au point de paramètre t, des courbes suivantes :

(Γ_1) $x = \cos 2t + 2\cos t$ et $y = \sin 2t + 2\sin t$

(Γ_2) $x = 3t - t^3$ et $y = 3t^2$

(Γ_3) $x = \cos^2 t + \text{Log}(\sin t)$ et $y = \sin t \cos t$

** **Exercice 4** : Voici 2 courbes définies en coordonnées polaires

(Γ_1) $r = \sqrt{\cos 2\theta}$

(Γ_2) $r = 1 + \sin 2\theta$

Trouver pour chacune d'elles les coordornées cartésiennes du centre de courbure au point $M(\theta, r(\theta))$.

Tracer la trajectoire du centre de courbure.

** **Exercice 5:** On fait rouler un cercle C de rayon a sur une droite D.

α) Soit M un point du cercle ; écrire une définition paramétrique de la trajectoire Γ de M. Tracer cette trajectoire.

β) Quelle est la trajectoire du centre de courbure Γ ?

*** **Exercice 6:**

Mêmes questions qu'à l'exercice 5 en faisant rouler C à l'intérieur d'un cercle $\mathcal{C}$ de rayon na ($n = 1, 2, 3, \ldots$

* **Exercice 7:** Pour chacune des courbes suivantes, déterminer le trièdre de Serret Frenet, la courbure et la torsion au point de paraètre t.

(Γ_1) $x = t - \frac{t^3}{3}$ $y = t^2$ $z = t + \frac{t^3}{3}$ $(t \neq 0)$

(Γ_2) $x = 3t^2$ $y = 4t^3$ $z = 3t^4$ $(t \neq 0)$

(Γ_3) $x = 2abt$ $y = a^2 \text{Log}\, t$ $z = b^2 t^2$ (a,b réels quelconques)

(Γ_4) $x = e^t$ $y = e^{-t}$ $z = t\sqrt{2}$

*** **Exercice 8**: Soit Γ une courbe régulière tracée sur une sphère de rayon R. On suppose que la courbure de Γ est constante et égale à ρ.

a) Etablir les formules $\overrightarrow{OM}.\overrightarrow{T} = 0$ et $\overrightarrow{OM}.\overrightarrow{N} = \frac{-1}{\rho}$.

En déduire $\overrightarrow{OM} \wedge \overrightarrow{N} = \varepsilon \frac{1}{\rho} \sqrt{\rho^2 R^2 - 1} \ \overrightarrow{T}$ (où $\varepsilon = \pm 1$)

b) En utilisant les fomules ci-dessus, calculer α et β tels que $\overrightarrow{OM} = \alpha \overrightarrow{N} + \beta \overrightarrow{B}$.

c) Montrer que la torsion est nulle. En déduire que $\overrightarrow{B}$ est constant, puis que la courbe est plane. Quelle est la nature géométrique de Γ ?

** **Exercice 9**: Soit Γ une courbe, on suppose que sa courbure n'est jamais nulle, et que sa torsion est constamment nulle. Montrer que c'est une courbe plane.

** **Exercice 10**: Soit γ une courbe plane; on suppose que sa courbure est constante non nulle. Montrer que c'est un cercle. Conseil: Il existe une fonction θ telle que ($\vec{\imath}$ étant un vecteur unitaire fixe) $(\vec{\imath}, \overrightarrow{T}) = \theta(s)$; calculer $\theta'(s)$.

*** **Exercice 11**: Soit Γ une courbe régulière telle que la courbure ρ et la torsion τ soient constantes (et non nulles).

a) Montrer que $\frac{1}{\rho}\overrightarrow{T} - \frac{1}{\tau}\overrightarrow{B} = \vec{k}$ est un vecteur constant.

b) Soit γ la projection orthogonale de Γ sur un plan perpendiculaire à $\vec{k}$; déterminer $\overrightarrow{Om}(t)$ en fonction de $\overrightarrow{OM}(t)$ et de $\vec{k}$. Puis calculer les vecteurs $\vec{t}$ et $\vec{n}$ de γ, ainsi que la courbure r de γ.

c) Montrer que γ est un cercle, et Γ une hélice circulaire.

** **Exercuce 12**: Soit Γ une courbe régulière telle que, en tout point, la courbure et la torsion soient égales.

a) Montrer que $\vec{k} = \overrightarrow{T} - \overrightarrow{B}$ est un vecteur constant.

b) Soit γ la projection de Γ sur un plan P perpendiculaire à $\vec{k}$, déterminer $\overrightarrow{Om}(t)$ en fonction de $\overrightarrow{OM}(t)$ et de $\vec{k}$. Puis déterminer la courbure de γ en fonction de celle de Γ.

* **Exercice 13**: On considère la courbe d'équation $x^2 + 2xy + 2y^2 + x + y = 13$. Elle passe par le point M(2;1). Déterminer sa tangente en M. Déterminer son centre de courbure en M.

* **Exercice 14**: On considère la courbe d'équation $x^3 + 2xy^2 + y^3 - 2x - 2y = 0$.

a) Elle passe par le point M(1;1). Déterminer sa tangente en M. Déterminer son centre de courbure en M.

b) Elle passe par l'origine. Montrer que l'origine est un centre de symétrie. Quel est sa courbure à l'origine ?

** **Exercice 15**: On considère la courbe Γ de l'espace définie comme l'intersection du plan $P(z = 2x + y + 1)$ et de la surface $S(z = x^3 + y^3 + 2x - y)$. Le point M(1;0;3) est sur Γ. Montrer que c'est un point régulier et donner un vecteur directeur de la tangente.

Quelle est la courbure de Γ en M ? Quel est le centre de courbure ?

** **Exercice 16**: On considère la courbe Γ de l'espace définie comme l'intersection de la surface $\Sigma(z = x^2 + y + 2)$ et de la surface $S(z = x^3 + y^3 + 2x - y)$. Le point $M(1;0;3)$ est sur Γ.

a) Montrer que c'est un point régulier et donner un vecteur directeur de la tangente.

b) Quelle est la courbure de Γ en M ? Quel est le centre de courbure ?

c) Quelle est la torsion de Γ en M ?

* **Exercice 17**: La courbe d'équation $x^3 + x^2y + y^3 + x^2 - y^2 = 0$ a un point double à l'origine.

a) Déterminer les tangentes à l'origine des deux arcs qui y passent.

b) Déterminer les centres de courbure à l'origine de ces deux arcs.

* **Exercice 18**: On considère la courbe d'équation $(x^2 + y^2)^2 - 4x = 0$.

a) Elle passe par le point $M(1;1)$. Déterminer sa tangente en M. Déterminer son centre de courbure en M.

b) Elle passe par l'origine. Quel est son centre de courbure à l'origine ?

** **Exercice 19**: La courbe d'équation $(x^2 + y^2)^2 - 4xy = 0$ a un point double à l'origine.

a) Déterminer les tangentes à l'origine des deux arcs qui y passent.

b) Montrer que ces deux arcs ont une courbure nulle à l'origine.

c) La courbe passe par le point $M(1;1)$. Quel est son centre de courbure en ce point ?

*** **Exercice 20:** Dans un plan P muni d'un repère orthonormé $(O, \vec{i}, \vec{j})$, on donne une fonction f de classe C^2 qui s'annule à l'origine. On suppose que $d_O f$ n'est pas nulle; ainsi l'équation $f(x,y) = 0$ définit une courbe γ régulière au voisinage de O.

Soit $s \to \gamma(s)$ un paramétrage intrinsèque de γ au voisinage de O. Montrer que les vecteurs dérivés $\dfrac{d\vec{M}}{ds}$ et $\dfrac{d^2\vec{M}}{ds^2}$ au point O, ne dépendent que des dérivées partielles premières et secondes de f en O.

On suppose que $\dfrac{\partial f}{\partial x}(O) = 1$, $\dfrac{\partial f}{\partial y}(O) = 2$, $\dfrac{\partial^2 f}{\partial x^2}(O) = 1$, $\dfrac{\partial^2 f}{\partial x \partial y}(O) = 0$ et $\dfrac{\partial^2 f}{\partial y^2}(O) = 1$. Déterminer les centre de courbure de γ en O. On peut le faire sans aucun calcul, en commençant par écrire un polynôme de degré 2 qui a même développement à l'ordre 2 que f en O.

Les intégrales multiples

Les fonctions que nous considérons ici sont à valeurs réelles définies sur $\mathbb{R}^n$. En fait le cas où n égale 2 contient déjà toutes les difficultés. C'est pourquoi, sauf au §5, nous nous limiterons à la dimension 2. Pour simplifier, les démonstrations ont été à peu près toutes oubliées; elles sont difficiles et ne trouveraient pas place dans un tel cours.

§ 1 : La position du problème.

Nous considérons les fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ qui

* d'une part sont nulles en dehors d'un sous-ensemble borné D de $\mathbb{R}^2$ (i.e. pour toute f il existe M tel que ($|x| \geq M$ ou $|y| \geq M$) implique $f(x,y) = 0$.
* d'autre part sont bornées (i.e. pour toute f il existe M' tel que $(\forall x,y)\ |f(x,y)| \leq M'$).

Elles forment un sous-espace vectoriel $\mathcal{B}$ de $\mathcal{F}(\mathbb{R}^2,\mathbb{R})$.

Définir une théorie de l'intégrale sur $\mathbb{R}^2$ c'est:

1) Définir un sous-espace vectoriel de $\mathcal{B}$, noté $\mathcal{I}$, et appelé "espace des fonctions intégrables".

2) Définir une forme linéaire sur $\mathcal{I}$, appelée "intégrale". On notera $\iint f(x,y)\,dx\,dy$ l'image de la fonction f par cette forme.

De telle façon que:

a) Si f et g sont dans $\mathcal{I}$, et si $f \leq g$ (i.e.: si $(\forall x,y)\ f(x,y) \leq g(x,y)$), alors:

$$\iint f(x,y)\ dx\ dy \leq \iint g(x,y)\ dx\ dy.$$

b) Quelque soit le rectangle $P(a\leq x\leq b$ et $c\leq y\leq d)$, et quelle que soit la fonction f qui vaut 0 hors de P, et 1 à l'intérieur de P (et n'importe quoi sur la frontière de P), f est intégrable et

$$\iint f(x,y)\ dx\ dy = (b-a)(d-c).$$

On remarquera l'analogie avec l'intégrale des fonctions d'une variable (chap.7 §1).

§ 2 : Les sommes de Riemann.

Soit f dans $\mathcal{B}$. Soit $\alpha \in]0,\infty[$. Pour tout couple (i,j) dans $\mathbb{Z}^2$ considérons le carré

$$C_{ij} = \{(x,y) \text{ tels que } i\alpha\leq x\leq (i+1)\alpha \text{ et } j\alpha\leq y\leq (j+1)\alpha\}$$

Ces carrés forment un pavage $\mathcal{P}$ du plan. Pour tout pavé C_{ij}, nous pouvons poser :

$$M(C_{ij}) = \sup_{t\in C_{ij}} f(t) \quad \text{et} \quad m(C_{ij}) = \inf_{t\in C_{ij}} f(t)$$

Ces bornes inférieure et supérieure existent puisque f est bornée. Nous poserons alors

$$I(\alpha,+,f) = \sum_{C_{ij}} M(C_{ij})\,\alpha^2 \quad \text{et} \quad I(\alpha,-,f) = \sum_{C_{ij}} m(C_{ij})\,\alpha^2$$

Notons que ces sommes sont étendues à tous les C_{ij}; ce sont donc apparemment des sommes infinies. Mais, puisque f est nulle en dehors d'un ensemble borné, il n'existe qu'un nombre fini des C_{ij} qui contiennent des points où f est non nulle (donc pour les quels $M(C_{ij})$ et $m(C_{ij})$ peuvent être non nuls). Ce sont donc des sommes d'un nombre fini de termes non nuls, et d'une infinité de zéros; autrement dit ce sont des sommes finies. On les appelle les sommes de Riemann associées à f et au pavage.

Soit (x,y) un point, il appartient à un des carrés C_{ij} ou à un nombre fini d'entre eux (s'il est sur leur bord). Notons $\lambda(x,y) = \sup_{(x,y)\in C_{ij}} M_{ij}$ et $\mu(x,y) = \inf_{(x,y)\in C_{ij}} m_{ij}$. Nous obtenons ainsi deux fonctions. Pour tout (x,y), on a $\lambda(x,y) \geq f(x,y) \geq \mu(x,y)$.

Exercice a: Montrer que si $(\forall(x,y))\ |f(x,y)| \leq A$, alors $(\forall(x,y))\ |\lambda(x,y)| \leq A$ et $|\mu(x,y)| \leq A$. Montrer que λ et μ sont nulles en dehors d'un ensemble borné.

Les fonctions λ et μ sont combinaisons linéaires de fonctions du type de celles qui interviennent dans la condition b (cf §1). Elles sont donc intégrables. Et leurs intégrales sont respectivement $I(\alpha,+,f)$ et $I(\alpha,-,f)$. Comme f est comprise entre λ et μ, si f est intégrable, son intégrale est comprise entre $I(\alpha,+,f)$ et $I(\alpha,-,f)$.

Dés lors, comme pour les fonctions d'une variable, nous poserons

$$I(+,f) = \inf_{\alpha} I(\alpha,+,f) \text{ et } I(-,f) = \sup_{\alpha} I(\alpha,-,f) .$$

Et nous dirons que f est intégrable au sens de Riemann, si $I(+,f) = I(-,f)$; ce nombre est alors appelé l'intégrale de f.

C'est, généralisée au cas de la dimension 2, la définition que nous avons donnée au chapitre 7 pour les fonctions d'une variable. On démontre (et nous admettrons) que nous avons ainsi défini une théorie de l'intégrale, c'est à dire que les fonctions intégrables au sens de Riemann forment un sous-espace $\mathcal{I}$ de $\mathcal{B}$, et que l'intégrale est une forme linéaire sur $\mathcal{I}$ possédant les propriétés a et b du § 1.

§ 3 : Sous-ensembles quarrables du plan, et fonctions intégrables.

Ensembles quarrables.

Soit E un sous-ensemble borné du plan, nous dirons que E est quarrable si sa fonction caractéristique χ_E est intégrable au sens de Riemann. L'intégrale de cette fonction est alors appelée l'aire de E.

Nous dirons que E est négligeable, s'il est quarrable et d'aire nulle.

Exercice b: Soit E le segment défini par $\{x=1 \text{ et } 0\leq y\leq 1\}$. Pour tout entier n, calculer $I(1/n,-,\chi_E)$ et $I(1/n,+,\chi_E)$. En déduire que E est négligeable.

Même question en prenant pour E le segment $\{x=y \text{ et } 0\leq x\leq 1\}$.

Exercice c: On suppose que E est négligeable. Montrer que pour tout partie F de E, $I(\alpha,+,\chi_F) \leq I(\alpha,+,\chi_E)$. En déduire que F est négligeable.

De façon générale, on démontre que tout arc simple (au sens du chap. 17) qui est borné, est un ensemble négligeable. Nous l'admettrons.

Frontière d'un fermé de $\mathbb{R}^2$.

Nous appellerons frontière du fermé D l'ensemble des points (de $\mathbb{R}^2$) qui sont à la fois limite d'une suite de points de U, et limite d'une suite de points du complémentaire de U.

Exercice d: Soit D un carré, montrer que la frontière de D est la réunion des quatre cotés de ce carré.

Soit D le disque limité par le cercle de centre O et de rayon r, quelle est la frontière de D ?

Soit P un demi-plan, quelle est la frontière de P ?

Soit D un segment, quelle est la frontière de D ?

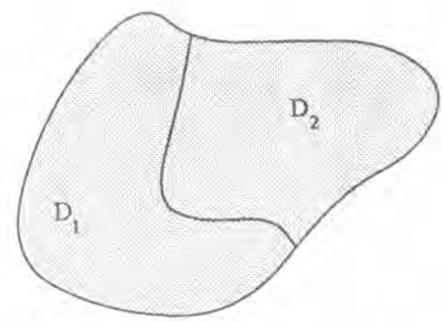
Domaines simples du plan

Nous appellerons "domaine simple" du plan tout fermé borné dont la frontière est négligeable. Les domaines simples que nous rencontrerons auront tous pour frontière une réunion (finie) d'arcs simples. Un carré est un domaine simple, un disque également. Un demi-plan n'est pas un domaine simple car il n'est pas borné.

Nous admettrons que toute fonction bornée $f: \mathbb{R}^2 \to \mathbb{R}$ qui est continue sur un domaine simple D, et nulle hors de D, est intégrable. Etant donnés un sous-ensemble A du plan qui contient D, et une fonction f continue sur A, nous noterons $\iint_D f(x,y)\,dx\,dy$ l'intégrale de la fonction qui est nulle hors de D, et qui est égale à f sur D

Dans le calcul des intégrales doubles nous utiliserons souvent le résultat suivant (que nous admettons) Soit D un domaine simple et f une fonction continue sur A ($\supset$D). Supposons que D soit la réunion de deux domaines simples D_1 et D_2 ayant en commun une partie de leur frontière. Alors

$$\iint_D f(x,y)\,dx\,dy = \iint_{D_1} f(x,y)\,dx\,dy + \iint_{D_2} f(x,y)\,dx\,dy$$



§ 4 : Le calcul des intégrales doubles.

Pour calculer une intégrale double nous ne disposons pour l'instant que d'une méthode: en déterminer un encadrement en calculant $I(\alpha,+,f)$ et $I(\alpha,-,f)$ pour une valeur de α très petite (de façon que la différence $I(\alpha,+,f) - I(\alpha,-,f)$ soit elle même petite). Procédé analogue au calcul d'une intégrale simple par la méthode des trapèzes. En voici une autre, connue sous le nom de théorème de Fubini, qui nous donnera la valeur exacte de l'intégrale pour peu qu'on puissse effectuer certaines primitivations.

Soit D un domaine simple de E, et soit f une fonction continue sur D, et nulle hors de D.

Pour tout x dans $\mathbb{R}$, considérons la fonction partielle f_x: $\mathbb{R} \to \mathbb{R}$ (qui à y associe f(x,y)), et définissons une application φ de $\mathbb{R}$ dans $\mathbb{R}$, en posant:

$$\begin{cases} \varphi(x) = \int f_x(y)\,dy \quad \text{si } f_x \text{ est intégrable} \\ \varphi(x) = 0 \quad \text{si } f_x \text{ n'est pas intégrable} \end{cases}$$

Alors, si φ est intégrable, son intégrale est égale à $\iint_D f(x,y)\,dx\,dy$

Exemple 1: Dans $\mathbb{R}^2$ muni de sa structure euclidienne naturelle considérons le rectangle D défini par $(0 \leq x \leq 2$ et $0 \leq y \leq 1)$, et calculons l'intégrale $I = \iint_D x^2y\,dx\,dy$.

Pour $x \notin [0,2]$ on a $f_x = 0$. Donc $\int f_x(y)\,dy = 0$.

Pour $x \in [0,2]$ on a $\begin{cases} f_x(y) = x^2y & \text{si } y \in [0,1] \\ f_x(y) = 0 & \text{si } y \notin [0,1] \end{cases}$; donc $\int f_x(y)\,dy = \int_0^1 x^2y\,dy = \frac{x^2}{2}$.

Le théorème de Fubini nous dit que $I = \iint_D x^2y \, dx \, dy = \int_0^2 \frac{x^2}{2} dx = \frac{4}{3}$.

Exemple 2: Dans $\mathbb{R}^2$ le domaine D défini par $(x^2+y^2 \leq 1$ et $x+y \geq 0)$, et calculons $I = \iint_D xy \, dx \, dy$.

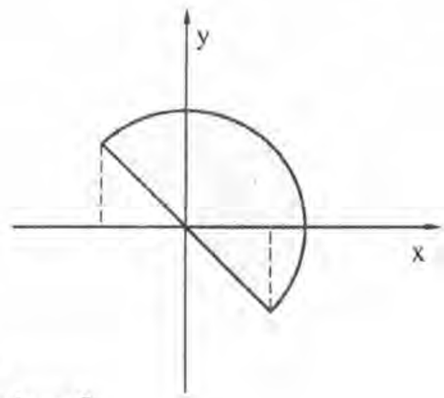
Calculons d'abord l'intégrale $\int f_x(y) \, dy$ .Elle est nulle si $x>1$ et si $x < -1/\sqrt{2}$.

Pour $-1/\sqrt{2} \leq x \leq 1/\sqrt{2}$ elle est égale à

$$\int_{-x}^{\sqrt{1-x^2}} xy \, dy = x \left[\frac{1-x^2}{2} - \frac{x^2}{2} \right] = \frac{x}{2} - x^3.$$

Pour $1/\sqrt{2} \leq x \leq 1$, elle est égale à

$$\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} xy \, dy = x\left[\frac{1-x^2}{2} - \frac{1-x^2}{2} \right] = 0.$$



D'après le théorème de Fubini $I = \int_{-1/\sqrt{2}}^{1/\sqrt{2}} (\frac{x}{2} - x^3) \, dx + \int_{1/\sqrt{2}}^{1} 0 \, dx = 0.$

Remarque: Dans tout ce qui précède nous avons calculé d'abord l'intégrale en y, puis l'intégrale en x. Nous pouvons évidemment échanger les rôles de y et de x, et calculer d'abord les intégrales des fonctions f_y: $x \to f(x,y)$.

Exercice f: Reprendre le calcul des intégrales des exemples ci dessus en intégrant d'abord en y. Vérifier que l'on trouve bien les mêmes résultats.

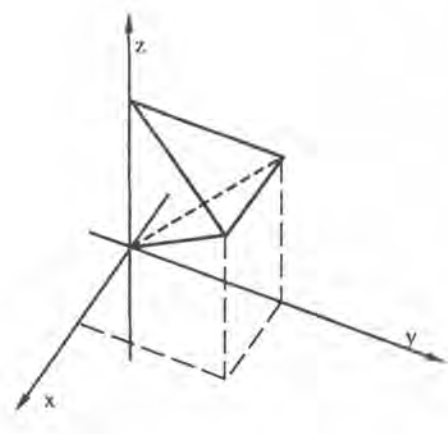
§ 5 : Le calcul des intégrales triples

La théorie des intégrales triples, quadruples,... est tout à fait analogue à celle des intégrales doubles. Le théorème de Fubini s'étend au calcul des intégrales triples, quadruples,... Nous ne donnerons pas d'énoncé général, mais simplement un exemple d'utilisation en dimension 3.

Dans $\mathbb{R}^3$ muni de son produit scalaire naturel, on note D le domaine défini par $(0 \leq x \leq y, 0 \leq y \leq z, 0 \leq z \leq 1)$. Nous allons calculer $I = \iiint_D (xy + z) \, dx \, dy \, dz$.

Pour tout point $m(y,z)$, notons f_{yz} la fonction qui à x associe $xy+z$ si $M(x,y,z)$ est dans D, et 0 sinon. En particulier f_{yz} est nulle si m n'est pas dans le domaine plan V défini par $(0 \leq y \leq z, 0 \leq z \leq 1)$. Pour m appartenant à V, nous avons:

$$\int f_{yz}(x) \, dx = \int_0^y (xy+z) \, dx = \frac{y^3}{2} + zy$$



Alors $I = \iint_V (\frac{y^3}{2} + zy) \, dy \, dz$, et nous calculons cette intégrale double comme précédemment. Soit g_z la fonction qui à y associe $\frac{y^3}{2} + zy$ si m(y,z) est dans V, et 0 sinon. Cette fonction est nulle si z n'appartient pas à [0,1]. Pour z appartenant à [0,1], nous avons:

$$\int g_z(y)\,dy = \int_0^z \left(\frac{y^3}{2} + zy\right) dy = \frac{z^4}{8} + \frac{z^3}{2}$$

D'où $I = \int_0^1 \left(\frac{z^4}{8} + \frac{z^3}{2}\right) dz = \frac{1}{40} + \frac{1}{8} = \frac{3}{20}$.

Exercice g : Démontrer que D peut aussi être défini par les inégalités ($y \le z \le 1, x \le y \le 1, 0 \le x \le 1$). Recalculer cette intégrale en intégrant d'abord en z, puis en y, enfin en x.

§ 6 : Changements de variables

Le théorème de changement de variable

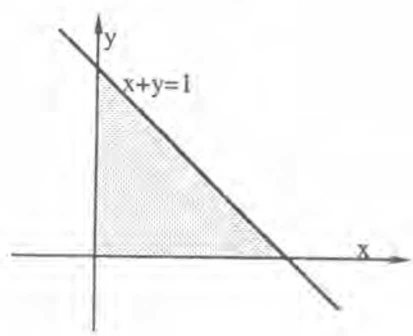
Supposons que D et Δ soient des "domaines simples", que f soit une fonction de classe C^1, telle que $f(\Delta) = D$, et qui induit un difféomorphisme de l'intérieur de Δ sur l'intérieur de D. Nous appellerons fonction de Jacobi de f la fonction de Δ dans $\mathbb{R}$ définie par $(\forall u \in \Delta)\ J_f(u) = \text{dét}\,(d_u f : \mathbb{R}^2 \to \mathbb{R}^2)$. Alors pour toute fonction φ continue sur D

$$\iint_D \varphi(X,Y)\,dX\,dY = \iint_\Delta \varphi(f(x,y))\ |J_f(x,y)|\ dx\,dy.$$

Nous admettrons ce résultat.

Exemple: Soit D le domaine de $\mathbb{R}^2$ défini par ($0 \le y < \le 1-x$ et $0 \le x$), et soit $I = \iint_D \frac{xy}{x^2+y^2}\,dx\,dy$.

Pour calculer I nous pouvons travailler en coordonnées polaires. Soit Δ le domaine de $\mathbb{R}^2$ défini par ($0 \le \theta \le \pi/2$ et $0 \le \rho \le \frac{1}{\cos\theta + \sin\theta}$). On définit $f : \Delta \to D$ par les formules $(\rho,\theta) \to (\rho\cos\theta, \rho\sin\theta)$. La fonction de Jacobi est le déterminant de $\begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$ Donc $J(\rho,\theta) = \rho > 0$. Le théorème de changement de variable nous dit alors que:



$$I = \iint_\Delta \frac{\rho\cos\theta\ \rho\sin\theta}{\rho^2\cos^2\theta + \rho^2\sin^2\theta}\,\rho\,d\rho\,d\theta = \iint_\Delta \rho\cos\theta\sin\theta\,d\rho\,d\theta$$

Exercice h: Terminer le calcul de cette intégrale, en intégrant d'abord par rapport à ρ. Puis calculer I directement sans changement de variable.

Les trois changements de variable classiques sont les passages en coordonnées polaires. Il peut être utile de connaître leurs fonctions de Jacobi.

* *Coordonnées polaires planes: $(\rho,\theta) \to (\rho\cos\theta, \rho\sin\theta)$ nous donne la fonction de Jacobi $J(\rho,\theta) = \rho$.*

* *Coordonnées cylindriques dans l'espace: $(\rho,\theta,z) \to (\rho\cos\theta, \rho\sin\theta, z)$ nous donne la fonction de Jacobi $J(\rho,\theta,z) = \rho$.*

* *Coordonnées polaires dans l'espace: $(\rho,\theta,\varphi) \to (\rho\cos\varphi\cos\theta, \rho\cos\varphi\sin\theta, \rho\sin\varphi)$ nous donne la fonction de Jacobi $J(\rho,\theta,\varphi) = \rho^2\cos\varphi$.*

<u>Le cas particulier de la dimension 1</u>: Soit à calculer $\int_a^b \varphi(t)\,dt$, où a<b et où φ est continue; et soit f une application bijective, croissante et continument dérivable de $[\alpha,\beta]$ sur [a,b]. Nous avons l'habitude d'écrire $\int_a^b \varphi(t)\,dt = \int_\alpha^\beta f'(\tau)\varphi(f(\tau))\,d\tau$. C'est, en dimension 1, le théorème de changement de variable exposé ci-dessus (la dérivée f' étant la fonction de Jacobi) avec des hypothèses un peu plus faibles, puisque nous n'avons pas supposé que f^{-1} est ce classe C^1. Mais si f est décroissante, nous écrirons encore la formule $\int_a^b \varphi(t)\,dt = \int_\alpha^\beta f'(\tau)\varphi(f(\tau))\,d\tau$ (où $f(\alpha) = a$ et $f(\beta) = b$, donc $\alpha>\beta$); alors que la fonction de Jacobi est $t \to |f'(t)|$. C'est qu'en dimension 1, nous avons fait des conventions particulières. Si nous calquons en dimension 1 la théorie des intégrales multiples, nous ne définirons pas $\int_a^b \varphi(t)\,dt$ mais $\int_{[a,b]} \varphi(t)\,dt$ (cf: chap.7 §1). Et si a<b nous aurons $\int_a^b \varphi(t)\,dt = \int_{[a,b]} \varphi(t)\,dt$.

Tandis que le symbole $\int_a^b \varphi(t)\,dt$, où a>b, n'a pas d'équivalent en termes d'intégrales multiples*. C'est, par convention spécifique à la dimension 1, $-\int_{[a,b]} \varphi(t)\,dt$). Et ceci nous donne deux écritures de la formule de changement de variable (valables que f' soit positif ou négatif):

* Celle qui utilise les conventions d'écriture spécifiques à la dimension 1 qui s'écrit:

$$\int_a^b \varphi(t)\,dt = \int_\alpha^\beta f'(\tau)\varphi(f(\tau))\,d\tau.$$

* Celle qui utilise les conventions d'écritures des dimensions supérieures qui s'écrit:

$$\int_{[a,b]} \varphi(t)\,dt = \int_{[\alpha,\beta]} |f'(\tau)|\,\varphi(f(\tau))\,d\tau.$$

Complément : Convergence à l'infini des intégrales doubles

Etant donnée f: $\mathbb{R}^2 \to \mathbb{R}$ (continue) nous allons définir $\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy$; intuitivement, nous allons prendre la limite de $\iint_K f(x,y)\,dx\,dy$ lorsque le domaine borné K "croit indéfiniment". Mais il faut donner un sens précis à cette expression.

<u>Une première tentative</u>: Posons $I(r) = \iint_{D(r)} f(x,y)\,dx\,dy$, où D(r) est le disque de centre l'origine et de rayon r. Nous pouvons considérer la limite λ (si elle existe!) de I(r) pour r tendant vers l'infini.

* Il n'est introduit que pour avoir une notation reliant agréablement la théorie de l'intégrale aux primitives.

Une seconde tentative: Posons $J(a) = \iint_{C(a)} f(x,y)\, dx\, dy$, où C(a) est le carré centré à l'origine de cotés parallèles aux axes et de coté 2a. Nous pouvons considérer la limite μ (si elle existe!) de J(a) pour a tendant vers l'infini.

D'autres possibilités: Mais il n'y a aucune raison de choisir des carrés ou des disques. Nous pouvons choisir un domaine simple quelconque U, contenant l'origine; considérer, pour tout t, l'image U(t) de U par l'homothétie de centre l'origine et de rapport t, et l'intégrale F(t) de f sur U(t); nous chercherons alors la limite ν de F(t) pour t tendant vers l'infini.

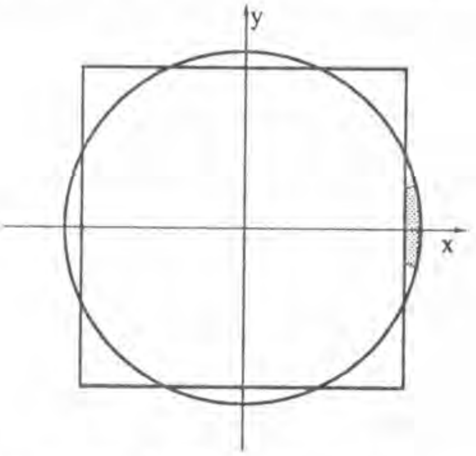
Malheureusement tous ces procédés ne donnent pas (pas toujours!) la même limite; l'exercice qui suit donne un exemple où λ existe, mais n'est pas la limite des J(a).

Exercice k: On donne f: $\mathbb{R}^2 \to \mathbb{R}$ définie en coordonnées polaires par $f(\rho,\theta) = \dfrac{\rho \cos 4\theta}{1+\rho^2}$

1) Calculer I(r). Montrer que λ existe.

2) Dessiner sur une même figure D(r) et C(r cos $\pi/8$). Vérifier que sur les parties de D(r) qui ne sont pas dans C(r cos $\pi/8$) la fonction f est positive. Vérifier qu'elle est négative sur les parties de C(r cos $\pi/8$) qui ne sont pas dans D(r). Qu'en déduire pour J(a) ?

3) Soit X(r) la partie du disque D(r) qui est extérieure au carré C(r cos $\pi/8$), et comprise entre les demi-droites d'angles polaires $\pi/16$ et $-\pi/16$ (hachurée sur la figure).



Montrer qu'il existe A tel que Aire(X(r)) = A r^2. Montrer que, pour r assez grand, en tout point de X la fonction f est supérieure à $\dfrac{r \cos \pi/4}{1+r^2}$. En déduire que, pour r grand, $J(r \cos \pi/8) \geq \dfrac{A\, r^3 \cos \pi/4}{1+r^2}$, et que J(a) tend vers l'infini lorsque a tend vers l'infini.

De tels phénomènes ne peuvent se produire si f est positive; en effet dans ce cas, quel que soit r : $I(r) \leq J(r) \leq I(r \cos \pi/4)$ (car plus le domaine est grand, plus l'intégrale est grande); donc si I a une limite, J a même limite. Ainsi dans le plan il n'est pas possible de définir des intégrales convergentes, il faudra se contenter de définir les intrégrales absolument convergentes(*).

Nous dirons que l'intégrale de f: $\mathbb{R}^2 \to \mathbb{R}$ est (absolument) convergente à l'infini, si:

$$(\forall\varepsilon>0)\ (\exists\eta>0)\ (\forall r)\ (\forall r')\ (r\geq\eta \text{ et } r'\geq r) \text{ implique } \left(0\leq\right) \iint_{D(r')} |f(x,y)|\, dx\, dy - \iint_{D(r)} |f(x,y)|\, dx\, dy \leq \varepsilon$$

Ceci signifie que la fonction $r \to \iint_{D(r)} |f(x,y)|\, dx\, dy$ vérifie le critère de Cauchy. Elle a donc une limite (que nous noterons $\iint_{\mathbb{R}^2} |f(x,y)|\, dx\, dy$).

Mais ceci entraine aussi que la fonction $r \to I(r) = \iint_{D(r)} f(x,y)\, dx\, dy$ *vérifie la condition de Cauchy, donc elle a une limite à l'infini que nous noterons* $\iint_{\mathbb{R}^2} f(x,y)\, dx\, dy$.

(*) Si nous n'avons pas rencontré de tels problèmes dans $\mathbb{R}$, c'est parceque nous avons une fois pour toutes choisi des domaines simples particuliers: les intervalles. Tandis qu'ici nous hésitons entre les carrés, les cercles,...

En effet (en notant A(r,r') la couronne comprise entre les deux disques):

$$|I(r')-I(r)| = \left|\iint_{A(r,r')} f(x,y)\,dx\,dy\right| \leq \iint_{A(r,r')} |f(x,y)|\,dx\,dy = \iint_{D(r')} |f(x,y)|\,dx\,dy - \iint_{D(r)} |f(x,y)|\,dx\,dy$$

Donc, quel que soit ε, si nous choisissons η comme ci-dessus, et r et r' supérieurs ou égaux à η, nous pourrons affirmer que $|I(r')-I(r)| \leq \varepsilon$. C'est le critère de Cauchy annoncé.

Et pour obtenir une valeur approchée de $\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy$ *, il suffit de calculer l'intégrale de f sur un domaine assez grand (quelle que soit sa forme!) puisque:*

$(\forall\varepsilon>0)$ $(\exists\eta>0)$ $(\forall\Delta$ domaine simple$)$ $(\Delta\supset D(\eta))$ implique $\left|\iint_{\Delta} f(x,y)\,dx\,dy - \iint_{\mathbb{R}^2} f(x,y)\,dx\,dy\right| \leq \varepsilon$

En effet choisissons η comme ci-dessus. Supposons que Δ contienne $D(\eta)$, et consirérons r' tel que $D(r')\supset\Delta$. Alors

$$\left|\iint_{D(r')} f(x,y)\,dx\,dy - \iint_{\Delta} f(x,y)\,dx\,dy\right| = \left|\iint_{D(r')-\Delta} f(x,y)\,dx\,dy\right|$$

$$\leq \iint_{D(r')-\Delta} |f(x,y)|\,dx\,dy \leq \iint_{A(r,r')} |f(x,y)|\,dx\,dy \leq \varepsilon$$

Et puisque cette inégalité est vraie quel que soit r' assez grand (i.e.: pour $D(r')\supset\Delta$), nous pouvons prendre la limite pour r' tendant vers l'infini, ce qui nous donnera l'inégalité annoncée.

Un critère de convergence:

Il est clair que si $|f| \leq |g|$, alors $\iint_{A(r,r')} |f(x,y)|\,dx\,dy \leq \iint_{A(r,r')} |g(x,y)|\,dx\,dy$. Il en résulte que si l'intégrale de g converge, celle de f converge également. Dés lors pour démontrer la convergence des intégrales doubles à l'infini, il nous faut connaitre un certain nombre de cas simples. Nous retiendrons que l'intégrale de la fonction $\frac{1}{(x^2+y^2)^\alpha}$ est convergente à l'infini si et seulement si $\alpha > 1$.

Exercice 1: Dés que α est positif, la fonction $\frac{1}{(x^2+y^2)^\alpha}$ n'est pas continue en 0, nous ne pourrons donc pas l'intégrer sur le disque de centre l'origine et de rayon r (sauf à évoquer des problèmes de convergence à l'origine, ce que nous ne ferons pas). Calculer son intégrale sur A(r,r'). En déduire que cette intégrale converge à l'infini pourvu que $\alpha>1$, et calculer l'intégrale sur le complémentaire du disque de rayon r

Exercices sur le chapitre 18

* **Exercice 1** : Calculer $I = \iint_U e^{x+y}\,dx\,dy$, où U est défini par ($x \geq 0, y \geq 0$ et $x+y \leq 1$).

* **Exercice 2**: Calculer $I = \iint_U (x^2+y^2)\,dx\,dy$, où U est le triangle A(1;2)B(-2,3)C(2;5).

* **Exercice 3** : Calculer $I = \iint_U (x^2+y^2)\,dx\,dy$, où U est défini par ($x^2+y^2 \leq 1$ et $x+y\sqrt{3} \geq 1$).

* **Exercice 4** : Calculer $I = \iint_U \frac{xy}{x^2+y^2}\,dx\,dy$, où U est le triangle A(1;0)B(0,0)C(1;1). On fera le calcul en coordonnées polaires.

** **Exercice 5** : Calculer $I = \iint_U \frac{xy}{x^2+y^2}\,dx\,dy$, où U est le cercle de rayon $\sqrt{2}$, centré au point W(1,1).

** **Exercice 6** : Calculer $\iiint_U xyz\,dx\,dy\,dz$, où U est le domaine défini par ($0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1$).

*** **Exercice 7** : Calculer $I = \iiint_U (x^2+y^2)\,dx\,dy\,dz$, où U est le domaine défini par ($0 \leq z \leq x^2+y^2 \leq 1$). On fera le calcul en coordonnées cylindriques, en justifiant le changement de variable.

*** **Exercice 8** : Calculer $I = \iiint_U (x^2+y^2)\,dx\,dy\,dz$, où U est le domaine défini par ($x \geq 0, y \geq 0, z \geq 0$ et $x^2+y^2+z^2 \leq 1$). On fera le calcul en coordonnées cylindriques, puis en coordonnées polaires.

* **Exercice 9** : Calculer $I = \iint_U x^2\,dx\,dy$, où U est l'ellipse $4x^2+y^2 \leq 1$. On fera d'abord un changement de variable tranformant l'ellipse en un cercle, puis on fera le calcul en coordonnées polaires.

* **Exercice 10** : Calculer $I = \iint_U (x^2+y^2)\,dx\,dy$, où U est le domaine défini par ($x^2+y^2+y \leq 1$).

*** **Exercice 11** : Calculer $I = \iiint_U (x^2+y^2)\,dx\,dy\,dz$, où U est la partie de la sphère de centre l'origine et de rayon 1 qui vérifie $x+y+z \geq 1$.

** **Exercice 12:** Calculer $I = \iiint_U x^2 \, dx \, dy \, dz$, où U est le domaine compris entre la sphère de centre l'origine et de rayon 3, et la sphère de centre $\Omega(1,0,0)$ et de rayon 1.

* **Exercice 14**: Calculer $I = \iint_U (x^2+y^2) \, dx \, dy$ où U est le domaine défini par les inégalités $x^2+y^2 \leq 1$, $x^2+ y^2 -x-y \leq 0$ et $x \geq 1/2$.

* **Exercice 15**: Calculer $I = \iint_U x^2 \, dx \, dy$, où U est le carré de sommets A(1;3), B(2;5), C(4;4), D(3;2). On pourra faire un changement de variable.

* **Exercice 16**: Soit U un domaine simple, et soit U_λ son image dans l'homothétie de centre l'origine et de rapport λ. Posons $I_\lambda = \iint_{U_\lambda} (x^2+y^2) \, dx \, dy$. Calculer U_λ/U.

Réduction des endomorphismes

Dans ce qui suit E est un espace vectoriel (sur $\mathbb{R}$ ou $\mathbb{C}$) de dimension n finie ; E est muni d'une base $\{e_1,\ldots,e_n\}$ dite base naturelle; A est un endomorphisme de E et M est sa matrice dans cette base.

Nous supposerons connus les définitions et résultats du chapitre 5.

§ 1: Triangulation des matrices

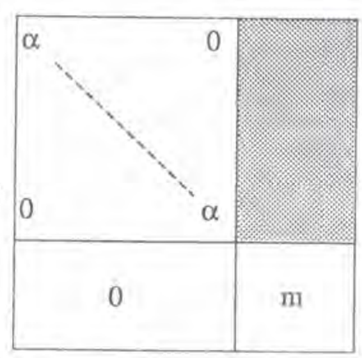
Dans tout ce paragraphe on suppose que dét(M-λI) est un produit de facteurs du premier degré. Si le corps de base est $\mathbb{C}$, ce n'est pas une restriction ; si ce corps est $\mathbb{R}$, ceci veut dire que le polynôme caractéristique n'a pas de racine non réelle.

Sous l'hypothèse ci-dessus, il existe une base de E dans laquelle la matrice T de A est triangulaire (supérieure) (autrement dit les coefficients situés sous la diagonale principale sont nuls).

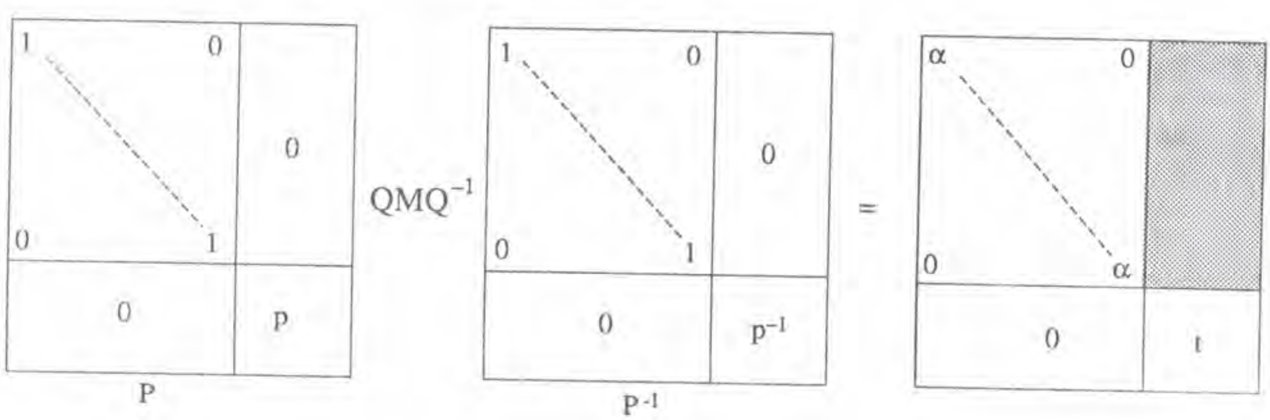
Il revient au même de dire qu'il existe P inversible telle que $PMP^{-1} = T$ soit triangulaire.

Démonstration : Nous ferons une récurrence sur la dimension. Supposons le résultat vrai en dimension inférieure à n, et soit A en dimension n. Soit α une racine du polynôme caractéristique, et soit B_α une base du noyau de Ker (A–αId). C'est une famille libre de E; complètons la en une base B de E; la matrice de A dans cette base a la forme ci-contre.

$N = QMQ^{-1} =$



Puisque Dét (M–λId) = Dét (N–λId) = $(\alpha-\lambda)^k$ dét (m–λId), le polynôme caractéristique de m est produit de facteurs du 1[er] degré. Et puisque m est de dimension inférieure à n, d'après l'hypothèse de récurrence, il existe p telle que $pmp^{-1} = t$ soit triangulaire. Alors



Nota : Dans la matrice triangulaire T ainsi construite, les coefficients diagonaux sont les valeurs propres de A, comptées avec leur ordre de multiplicité. Pour s'en persuader il suffit de calculer dét(T−λId).

Exercice a : Trouver A telle que AMA^{-1} soit triangulaire.

$$M_1 = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & 2 & 2 \end{pmatrix} \text{ (val. propres: 2,1,1)} \qquad M_2 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 2 \\ -1 & 0 & 2 & 1 \end{pmatrix} \text{ (val. propres: 2,2,0,0)}$$

§ 2 : P.G.C.D. de deux polynômes

Soit deux polynômes A et B, on considère l'ensemble I(A,B) des polynômes de la forme UA+VB (où U et V sont des polynômes). Il est clair que

α) si P et Q sont dans I(A,B), P+Q est aussi dans I(A,B).

β) si P est dans I(A,B), (∀Q) PQ est aussi dans I(A,B).

On dit que I(A,B) est un idéal de l'anneau des polynômes.

Soit P un élément de I(A,B) − {0} dont le degré soit le plus petit possible. Pour tout Q dans I(A,B), on peut écrire Q = PZ+R où d° R < d° P (et R = 0 si d° P = 0) ; c'est la division suivant les puissances décroissantes. Le polynôme R = Q − PZ est dans I(A,B) (à cause de α et β) et comme son degré est plus petit que celui de P, il est nécessairement nul. Donc P divise Q. Ainsi tout polynôme de I est un multiple de P. Inversement tout multiple de P est dans I. Donc I(A,B) est l'ensemble des multiples de P.

Le polynôme P est appelé le PGCD des polynômes A et B. Il divise A et B (puisque A et B sont dans I(A,B)). Tout polynôme S qui divise A et B, divise aussi P (puisque P est de la forme UA+VB).*

<u>1ère méthode de calcul du PGCD</u> : Divisons A par B, on a $A = BQ_1+R_1$, et l'ensemble des diviseurs de A et B, coïncide avec l'ensemble des diviseurs de B et R_1 ; donc

$$\mathrm{PGCD}(A,B) = \mathrm{PGCD}(B,R_1).$$

Exercice b: Donner le détail de la preuve de cette égalité.

Divisons B par R_1, on a $B = R_1Q_2+R_2$; et, de même:

$$\mathrm{PGCD}(B,R_1) = \mathrm{PGCD}(R_1,R_2).$$

Etc.

A chaque étape on a un reste R_i de degré strictement inférieur au degré du reste précédent. Au bout d'un nombre fini d'étapes, le reste est nul. Le dernier reste non nul est le PGCD cherché.

<u>2de méthode de calcul du PGCD</u> : Décomposons A et B en un produit de facteurs du 1er degré.

$$A(X) = \prod (X-a_i)^{\alpha_i} \quad (a_i \text{ tous distincts})$$
$$B(X) = \prod (X-b_j)^{\beta_j} \quad (b_j \text{ tous distincts})$$

On cherche les racines communes, et chaque fois que $a_i = b_j$ on prend le facteur $(X-a_i)^{\inf(\alpha_i,\beta_j)}$. Le produit de tous ces facteurs est le PGCD.

Notons que si A et B sont réels, leur PGCD sur $\mathbb{C}$ est égal à leur PGCD sur $\mathbb{R}$; par conséquent, si A et B sont réels, on peut faire le calcul ci-dessus avec toutes les racines, réelles ou complexes de A et B. Le résultat trouvé est réel.

* Notons que le PGCD n'est défini qu'au produit par un scalaire près.

Vocabulaire : Si le PGCD de A et B est de degré 0, on dit que A et B sont premiers entre eux. Dans ce cas 1 est dans I(A,B), il existe donc U et V tels que $UA+VB = 1$ (Formule de Bezout).

Exercice c: Trouver (par les deux méthodes ci-dessus) les PGCD

1) de X^6+1 et X^4-1
2) de X^9-1 et X^6-1
3) de X^7-1 et X^5+1.

§ 3 : Le théorème général de réduction

A tout polynôme $P(X) = \sum_{i=0}^{k} p_i X^i$, on associe** l'endomorphisme $P(A) = \sum_{i=0}^{k} p_i A^i$, et la matrice $P(M) = \sum_{i=0}^{k} p_i M^i$. On définit ainsi des homomorphismes de l'anneau des polynômes dans l'anneau $\mathcal{L}(E,E)$ des endomorphismes d'une part, et dans l'anneau des matrices $n \times n$ d'autre part. Evidemment P(M) est la matrice de P(A) dans la base canonique. On retiendra (c'est en fait la signification du mot "homomorphisme") que toute identité polynomiale, reste vraie si on remplace X par A, ou par M. Par exemple $(X^2+1)(X^2-1) = X^4-1$ implique ($\forall$A et $\forall$M) $(A^2+Id)(A^2-Id) = A^4-Id$ et $(M^2+I_n)(M^2-I_n) = (M^4-I_n)$.

En particulier quels que soient P et Q, les endomorphismes P(A) et Q(A) commutent.

Il existe (au moins) un polynôme P ($\neq 0$) tel que $P(A) = 0$, et $P(M) = 0$.

En effet considérons l'espace E_{n^2} des polynômes de degré au plus n^2. Il est de dimension n^2+1. L'application $A \to P(A)$ (de E_{n^2} dans $\mathcal{L}(E,E)$) est linéaire. Comme $\mathcal{L}(E,E)$ est de dimension n^2, elle a un noyau non réduit à {0}; autrement dit il existe P ($\neq 0$ et de degré au plus n^2) tel que $P(A) = 0$.

L'ensemble des polynômes P tels que $P(A) = 0$, est l'ensemble des multiples d'un polynôme P_0 appelé polynôme minimal de A.

Démonstration : Cet ensemble est un idéal (il possède les propriétés α et β); on reprend le raisonnement du § 2.

Soit P et Q deux polynômes de PGCD π, alors : $\mathrm{Ker}\,\pi(A) = \mathrm{Ker}\,P(A) \cap \mathrm{Ker}\,Q(A)$.

Démonstration: Il existe Z et T tels que $P = Z\pi$ et $Q = T\pi$ donc $P(A) = Z(A) \circ \pi(A)$ et $Q(A) = T(A) \circ \pi(A)$ donc $\mathrm{Ker}\,P(A) \supset \mathrm{Ker}\,\pi(A)$ et $\mathrm{Ker}\,Q(A) \supset \mathrm{Ker}\,\pi(A)$; ce qui prouve l'inclusion $\mathrm{Ker}\,\pi(A) \subset \mathrm{Ker}\,P(A) \cap \mathrm{Ker}\,Q(A)$.

Réciproquement il existe U et V tels que $\pi = UP+VQ$ donc $\pi(A) = U(A) \circ P(A) + V(A) \circ Q(A)$, donc tout vecteur qui est annulé à la fois par P(A) et par Q(A), est aussi annulé par $\pi(A)$; ce qui démontre l'inclusion $\mathrm{Ker}\,\pi(A) \supset \mathrm{Ker}\,P(A) \cap \mathrm{Ker}\,Q(A)$.

Si P et Q sont des polynômes premiers entre eux, Ker(P(A)Q(A)) est somme directe de Ker P(A) et de Ker Q(A).

** Par convention $A^0 = Id$ et $M^0 = I_n$.

Démonstration: Nous savons déjà que Ker P(A) et Ker Q(A) ont pour intersection le noyau de l'indentité (i.e. de $\Pi(A)$ lorsque Π est le polynôme 1), c'est à dire $\{0\}$. Il reste à montrer qu'ils engendrent Ker (P(A)Q(A)). Soit donc x tel que $P(A)Q(A)(x) = 0$. Puisque 1 s'écrit PU+QV, nous avons $Id = P(A)U(A) + Q(A)V(A)$, d'où:

$$x = P(A)U(A)(x) + Q(A)V(A)(x) = x_1+x_2$$

Et $\quad Q(A)(x_1) = [Q(A)P(A)]U(A)(x) = U(A)[Q(A)P(A)](x) = 0$; donc $x_1 \in \text{Ker}\, Q(A)$

de même $\quad P(A)(x_2) = [P(A)Q(A)]V(A)(x) = V(A)[P(A)Q(A)](x) = 0$; donc $x_2 \in \text{Ker}\, P(A)$.

Plus généralement si P, Q, R sont premiers entre eux 2 à 2, alors Ker P(A), Ker Q(A) et Ker R(A) sont des sous-espaces supplémentaires de Ker (PQR(A)).

Exercice d: 1) Démontrer que si P,Q,R sont premiers entre eux 2 à 2, alors P est premier avec QR (utiliser la formule de Bezout).

2) En déduire que Ker P(A), Ker Q(A) et Ker R(A) sont supplémentaires dans Ker (PQR(A)) (cf. chap.1 §6).

Conséquence:*Soit P un polynôme tel que P(A) = 0. Supposons que P soit produit de facteurs de d° 1 (sur le corps de base). Alors en regroupant les facteurs égaux, on a :*

$$P(X) = \prod_{i=1,\dots,k} (X-r_i)^{\alpha_i} \quad (r_i \text{ tous distincts}).$$

Les polynômes $(X-r_i)^{\alpha_i}$ sont premiers deux à deux. Posons $E_i = Ker(A-r_iId_E)^{\alpha_i}$. Alors

1) Les E_i forment une décomposition de E en somme directe.

2) $A(E_i) \subset E_i$.

Si le corps de base est $\mathbb{R}$, il est possible que le polynôme P ne soit pas produit de facteurs réels de degré 1. Alors on peut écrire (en regroupant les racines complexes conjuguées) :

$$P = \prod_{i=1,\dots,k} (X-r_i)^{\alpha_i} \times \prod_{j=1,\dots,\ell} [X^2 - (s_j+\bar{s}_j)X+s_j\bar{s}_j]^{\beta_j}$$

Les polynômes $(X-r_i)^{\alpha_i}$ et $[X^2 - (s_j+\bar{s}_j)X+s_j\bar{s}_j]^{\beta_j}$ sont premiers deux à deux.

Posons $E_i = Ker(A-r_iId_E)^{\alpha_i}$ et $F_j = Ker(A^2-(s_j+\bar{s}_j)A+s_j\bar{s}_jId_E))^{\beta_j}$. Alors

1) Les E_i et les F_j forment une décomposition de E en somme directe.

2) ($\forall$ i et j) $A(E_i) \subset E_i$; $A(F_j) \subset F_j$.

Exemple: Si $A^2 - Id = 0$, E (= Ker $(A^2 - Id)$) = Ker (A − Id)$\oplus$Ker (A + Id) ; ce que nous avons démontré en exercice au chapitre 1. Rechercher les exercices du chapitre 1 qui deviennent ainsi des évidences lorsque l'on connait ce théorème de réduction.

§ 4 : Etude des endomorphismes $A|E_i: E_i \to E_i$

Posons $\varphi_i = A|E_i - r_iId_{E_i} : E_i \to E_i$. Par définition de E_i, *il existe n tel que $(\varphi_i)^n = 0$. On dit que φ_i est nilpotent.* Soit $X_k = Ker(\varphi_i)^k$. Les X_k forment une suite croissante:

$$\{0\} = X_0 \subset X_1 \subset X_2 \dots \subset X_n = E_i.$$

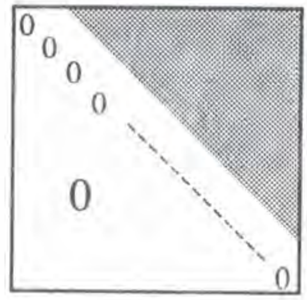
Si $X_p = X_{p+1}$, alors $X_{p+1} = X_{p+2}$. Il en résulte que la suite (X_p) est strictement croissante jusqu'à un certain rang p_0, et constante à partir du rang p_0. Donc $X_{p_0} = E_i$. Et $p_0 \le \dim E_i$.

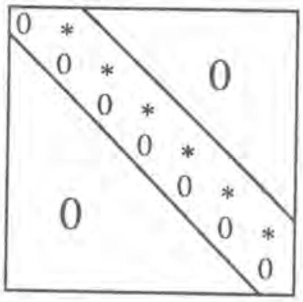
Exercice e: Démontrer que $Ker(\varphi_i)^p = Ker(\varphi_i)^{p+1}$ implique $Ker(\varphi_i)^{p+1} = Ker(\varphi_i)^{p+2}$.

Prenons alors une base de E_i adaptée à la suite

$$X_1 \subset X_2 \subset \ldots \subset X_{p_0}$$

C'est-à-dire formée d'une base de X_1, complètée en une base de X_2, complètée en une base de X_3, etc... Dans cette base la matrice de φ_i a la forme ci contre.

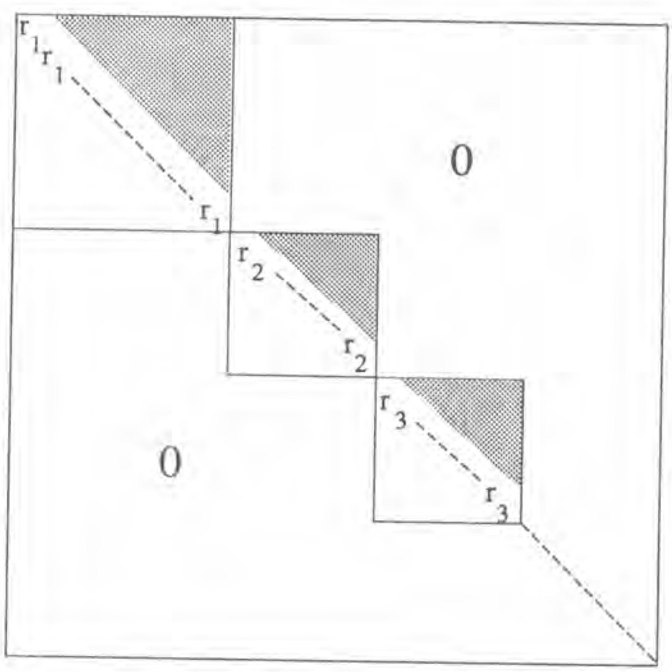




Elle est triangulaire supérieure, et les termes diagonaux sont nuls. En particulier le polynôme caractéristique de φ_i s'écrit $\lambda^{\dim E_i}$; et celui de $A \mid E_i$ est $(\lambda - r_i)^{\dim E_i}$.

Nous ne démontrerons pas qu'en choisisant bien cette base cette matrice a la forme ci-contre (où chaque * désigne un 0 ou un 1).

§ 5 : Forme réduite d'un endomorphisme



Supposons qu'on ait trouvé un polynôme P, tel que P(A) = 0, et qui soit produit de facteurs de degré 1 (sur le corps de base). Alors E est somme directe des E_i. Prenons, dans chaque E_i, une base comme au § 4 . La réunion de ces bases, est une base de E. Dans cette base la matrice A a la forme ci-contre.

Conséquence: *Sur cette dernière matrice nous voyons que les r_i sont les valeurs propres, et que ($\forall i$) la dimension du sous-espace E_i est égale à l'ordre de multiplicité de r_i comme racine du polynôme caractéristique.*

Le théorème de Hamilton Cayley: *Pour toute matrice M (réelle ou complexe), notons P_M le polynôme caractéristique de M. Alors $P_M(M) = 0$.* Nous savions que pour toute M, il existe des polynômes P tels que $P(M) = 0$; mais nous n'avions aucun procédé pratique pour en calculer un; ceci nous en donne un.

Démonstration : On peut toujours considérer M comme la matrice d'un endomorphisme A d'un espace complexe E. Il suffit alors de démontrer que $P_A(A) = 0$.

Considérons un polynôme P tel que $P(A) = 0$ et construisons des espaces E_i comme ci-dessus. Pour montrer que $P_A(A)$ est nul, il suffit de montrer que ($\forall i$) $P_A(A) \mid E_i = P_A(A \mid E_i)$ est nul. Or

$P_A(\lambda) = \prod (\lambda - r_j)^{\dim E_j}$ (pour s'en convaincre, on regarde la forme réduite ci-dessus). Donc $P_A(A \,|\, E_i)$ est la composée des $(A \,|\, E_i - r_j \, Id_{E_i})^{\dim E_j}$. Et puisque $(A \,|\, E_i - r_i \, Id_{E_i})^{\dim E_i}$ est nul, $P_A(A \,|\, E_i)$ est nul.

Récapitulons:

Soit un endomorphisme A, dont le polynôme caractéristique est produit de facteurs de degré 1 dans le corps de base. Ce qui n'est pas une restriction si l'on travaille sur $\mathbb{C}$. Dans le cas réel ceci signifie que le polynôme caractéristique n'a pas de racine non réelle.

Alors on peut reprendre tout ce qui a été fait ci-dessus, en choisissant pour A le polynôme P_A. On a $P_A(X) = \prod_{i=1,\dots,k} (X-r_i)^{\alpha_i}$ *(où les r_i deux à deux distincts). Posons $E_i = Ker(A - r_i \, Id_E)^{\alpha_i}$, alors $\dim E_i = \alpha_i$, et E est somme directe des E_i.*

Notons Δ l'endomorphisme de E qui $(\forall i)$ vaut $r_i \, Id_{E_i}$ sur E_i. Alors l'endomorphisme $N = A - \Delta$, est nilpotent. (En effet, si on l'exprime dans la base construite ci-dessus, sa matrice est triangulaire et a des 0 sur la diagonale). *Donc A s'écrit $\Delta + N$, où Δ est un endomorphisme diagonalisable, et N un endomorphisme nilpotent. De plus $\Delta N = N \Delta$.*

Vocabulaire :

- $Ker(A - r_i \, Id_E)$ est le sous-espace propre associé à la vakeur propre r_i.
- $Ker(A - r_i \, Id_E)^{\alpha_i}$ est appelé le sous-espace spectral associé à r_i; il est de dimension α_i.
- "Le" plus petit polynôme P tel que $P(A) = 0$ (ie: "le" générateur de l'idéal des polynômes qui s'annulent sur A) est appelé le polynôme minimal de A. Il divise P_A. Il a les mêmes racines que P_A, mais avec des ordres de multiplicité qui peuvent être plus petits.

Exemple: Soit f: $\mathbb{R}^4 \to \mathbb{R}^4$ dont la matrice, dans la base naturelle, s'écrit

$$M = \begin{pmatrix} 0 & -1 & 1 & 1 \\ -3 & -4 & 4 & 3 \\ -2 & -3 & 3 & 2 \\ -3 & -5 & 5 & 4 \end{pmatrix}$$

Ses valeurs propres sont 1 (triple) et 0 (simple). Les sous-espaces spectraux de f sont donc Ker f (de dimension 1, engendré par (0;1;1;0)) et $Ker\,(f-Id)^3$ (de dimension 3). Pour déterminer ce dernier, on peut calculer $(M-I_4)^3$

$$(M-I_4)^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Ce qui montre que (x,y,z,t) est dans $Ker\,(f-Id)^3$ si et seulement si $x+y = t$; ainsi les vecteurs (1;0;0;1), (0;1;0;1) et (0;0;1;0) engendrent $Ker\,(f-Id)^3$.

Les vecteurs $A_1(1;0;0;1)$, $A_2(0;1;0;1)$, $A_3(0;0;1;0)$ et $A_4(0;1;1;0)$ forment une base de $\mathbb{R}^4$. La matrice de f dans cette base est

$$M' = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 4 & 0 \\ 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Elle est formée d'un bloc 3×3 qui correspond à $Ker\,(f-Id)^3$, et d'un bloc 1×1 (qui est nul et) qui correspond à la restriction à Ker f.

Ici le polynôme minimal est égal au polynôme caractéristique (car Ker $(f-Id)^2 \neq$ Ker $(f-Id)^3$).
Posons

$$M' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & -2 & 4 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Nous avons ainsi écrit M' comme la somme d'une matrice diagonale et d'une matrice nilpotente, qui commutent entre elles.

Variante: Nous savons qu'il existe une base de Ker $(f-Id)^3$ dans la quelle le bloc 3×3 est triangulaire. Nous pouvons modifier notre procédé de recherche de Ker $(f-Id)^3$ pour obtenir une telle base. Nous chercherons d'abord Ker (f–Id); il est engendré par $V_1(1,0,0,1)$. Nous chercherons ensuite Ker $(f\text{-}Id)^2$, qui est l'ensemble des vecteurs X tels que (f–Id)(X) soit dans Ker (f–Id). Nous résoudrons donc l'équation $(f-Id)(X) = \alpha V_1$ (α arbiraire). On trouve

$$X = \alpha(0;2;1;2) + xV_1 \quad (x \text{ arbiraire})$$

Nous prendrons V_1 et $V_2(0;2;1;2)$ comme base de Ker $(f-Id)^2$. Enfin nous chercherons Ker $(f-Id)^3$; c'est à dire les vecteurs Y tels que (f–Id)(Y) soit dans Ker $(f-Id)^2$. Pour cela nous résoudrons l'équation $(f-Id)(Y) = \beta V_2 + \gamma V_1$ (β et γ arbitraires). On trouve

$$Y = \beta(0:-1;0;-1) + \gamma V_2 + yV_1 \quad (y \text{ arbitraire})$$

On prendra pour base de Ker $(f-Id)^3$ les vecteurs V_1, V_2 et $V_3(0;-1;0;-1)$, le bloc 3×3 s'écrit alors

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Exercices sur le chapitre 19

* **Exercice 1**: Trouver de PGCD

a) de $X^8 + X^4 + 1$ et de $X^6 - 1$.

b) de $X^6 + X^3 + 1$ et de $X^4 - 1$.

c) de $X^6 + X^4 + X^2 + 1$ et de $X^6 + X^3 + 1$.

* **Exercice 2**: Un endomorphisme A de $\mathbb{C}^n$ vérifie $A^3 + A^2 + A + Id = 0$ et $A^4 + 2A^2 + 1 = 0$. Montrer que $A^2 + Id = 0$. Montrer qu'il est diagonalisable.

* **Exercice 3**: Pour chacun des endomorphismes de $\mathbb{R}^n$ (ou de $\mathbb{C}^n$) dont les matrices sont écrites ci-dessous, trouver

- Les valeurs propres
- Les sous espaces propres
- Les sous espaces spectraux
- Une base dans laquelle leur matrice a une forme réduite.

$$\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix} \qquad \begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}$$

$$\begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & -1 & 2 & -1 \\ 2 & 0 & 1 & -4 & -1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & -3 & 3 & -1 \end{pmatrix} \text{(val propres 1 triple et -1 double)}$$

$$\begin{pmatrix} 3 & -1 & 1 & -1 \\ 1 & 3 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & -3 & 5 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 \\ 1 & 1 & 4 & 2 \end{pmatrix} \qquad \begin{pmatrix} 0 & \sin\alpha & \sin 2\alpha \\ \sin\alpha & 0 & \sin 2\alpha \\ \sin 2\alpha & \sin\alpha & 0 \end{pmatrix}$$

** **Exercice 4**: α) On se donne un endomorphisme A de $\mathbb{R}^n$ qui vérifie $A^4 = I$ et $A^6 = I$.

Trouver un générateur de l'idéal engendré par $X^4 - 1$ et $X^6 - 1$. Peux-t-on affirmer que A est diagonalisable ? (Si la réponse est non, donner un exemple d'endomorphisme qui vérifie ces relations, et n'est pas diagonalisable).

β) Mêmes questions avec les relations $A^4 - 2A^2 + I = 0$ et $A^6 = I$.

γ) Mêmes questions avec les relations $A^3 - 3A^2 + 3A = I$ et $A^4 - 2A^2 + I = 0$.

δ) Mêmes questions avec les relations $A^8 - 2A^4 + I = 0$ et $A^6 + 2A^3 + I = 0$.

* **Exercice 5**: On considère l'endomorphisme $f : \mathbb{R}^3 \to \mathbb{R}^3$ dont la matrice (dans la base naturelle) est $M = \begin{pmatrix} -2 & 2 & 0 \\ -2 & 2 & 0 \\ -2 & 1 & 1 \end{pmatrix}$

a) Déterminer ses valeurs propres puis ses sous-espaces spectraux.

b) Ecrire M comme la somme d'une matrice diagonalisable Δ et d'une matrice nilpotente N.

c) Montrer que $(\forall n)\ \Delta^n = \Delta$. Calculer N^2 . En déduire M^n .

d) On munit l'espace E des matrices 3x3 d'une norme arbitraire. Pour quelles valeurs de λ, la série $((\lambda^n \Delta^n))_{n\in\mathbb{N}}$ est elle sommable ? Pour quelles valeurs de λ la série $((\lambda^n M^n))_{n\in\mathbb{N}}$ est elle sommable ?

** **Exercice 6:** Soit M une matrice inversible, montrer qu'il existe un polynôme P tel que $M^{-1} = P(M)$

Calculer un tel polynôme lorsque $M = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 2 & 1 \\ 4 & 1 & 2 \end{pmatrix}$.

* **Exercice 7:** Déterminer le polynôme minimal des matrices suivantes:

$$M = \begin{pmatrix} 1 & 2 & 1 & 3 & 1 \\ 0 & 1 & 3 & 1 & 0 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad N = \begin{pmatrix} 1 & 2 & 1 & 2 & 3 \\ 0 & 1 & 2 & 1 & 4 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

** **Exercice 8:** Ecrire chacune des matrices suivantes comme somme d'une matrice nilpotente et d'une matrice diagonalisable qui commutent entre elles.

$$M = \begin{pmatrix} 1 & 2 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \qquad N = \begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

* **Exercice 9:** Ecrire chacune des matrices suivantes comme somme d'une matrice nilpotente et d'une matrice diagonalisable qui commutent entre elles.

$$M = \begin{pmatrix} -1 & -2 & -2 \\ 1 & 1 & -1 \\ 1 & 2 & 4 \end{pmatrix} \qquad N = \begin{pmatrix} 2 & 2 & 4 \\ -1 & -1 & -4 \\ 1 & 2 & 5 \end{pmatrix}$$

** **Exercice 10:** On dit que deux matrices M et N sont semblables s'il existe une matrice inversible P telle que $P^{-1}MP = N$. On sait que si deux matrices sont semblables elles ont le même polynôme caractéristique.

a) Montrer qu'il existe deux matrices 2×2, qui ont même polynôme caractéristique et qui ne sont pas semblables.

b) Montrer que si M et N (matrices $n\times n$) ont même polynôme caractéristique, et si ce polynôme n'a que des racines simples, alors elles sont semblables.

** **Exercice 11:** Soit M une matrice et P_M son polynôme caractéristique. Soit Q le polynôme caractéristique de M^2. Montrer que si α est racine de P_M, alors α^2 est racine de Q.

Montrer que si $P_M(X) = (X+\alpha)^r (X-\alpha)^s P_1(X)$ où P_1 ne s'annule ni en α ni en $-\alpha$, alors α^2 est une racine d'ordre r+s de Q. On pourra commencer par regarder le cas où M est triangulaire.

* **Exercice 12:** Calculer M^{1990}, où $M = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$

** **Exercice 13:** Soit M une matrice $n\times n$. Montrer qu'elle est diagonalisable si et seulement si son polynôme minimal n'a que des racines simples.

* **Exercice 14:** Montrer que M est diagonalisable si et seulement si M^t est diagonalisable.

* **Exercice 15:** Calculer M^{1990}, où $M = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 1 & 1 & 5 & 9 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$

** **Exercice 16**: Soit $M = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

a) Ecrire le polynôme caractéristique de M, et son polynôme minimal Q.

b) Soit E l'espace vectoriel des matrices de la forme P(M) (où P est un polynôme). Montrer que E a pour dimension le degré de Q.

* **Exercice 17:** Ecrire chacune des matrices suivantes comme somme d'une matrice nilpotente et d'une matrice diagonalisable qui commutent entre elles.

$$M = \begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \qquad N = \begin{pmatrix} 3 & 0 & 0 \\ -1 & -1 & -4 \\ 1 & 2 & 5 \end{pmatrix}$$

Calculer M^p et N^p $(p \in \mathbb{Z})$

* **Exercice 18**: Soit M et N deux matrices $n \times n$. Montrer que, si N est inversible, MN et NM ont même polynôme caractéristique. On comparera les polynômes caractéristiques de MN et de NM à $\det(M - \lambda N^{-1})$.

* **Exercice 19:** Calculer M^n et N^n $(n \in \mathbb{Z})$ où

$$M = \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix} \qquad N = \begin{pmatrix} -1 & 2 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Equations différentielles

Soit $F :]\alpha,\beta[\times U \to \mathbb{R}^p$ (où U est un ouvert de $\mathbb{R}^p$), nous lui associerons le problème suivant: Trouver tous les couples $(]a,b[,f)$ où $]a,b[\subset]\alpha,\beta[$ et où f est une application dérivable de $]a,b[$ dans U, telle que $(\forall t \in]a,b[)\ f'(t) = F(t,f(t))$.

Ce problème est appelé une équation différentielle d'ordre 1. Les couples $(]a,b[,f)$ sont appelés les solutions de cette équation.

Plus précisément, étant donné $(t_0,x_0) \in]\alpha,\beta[\times U$, nous pouvons chercher les solutions $(]a,b[,f)$ telles que $t_0 \in]a,b[$ et $f(t_0) = x_0$. Nous dirons alors que nous cherchons les solutions "correspondant à la donnée initiale (t_0,x_0)".

§ 1 : Le théorème d'existence et d'unicité des solutions.

Le théorème d'existence.

Soit $F :]\alpha,\beta[\times U \to \mathbb{R}^p$ comme ci-dessus; faisons l'hypothèse que F est de classe C^1. Alors pour toute donnée initiale $(t_0,x_0) \in]\alpha,\beta[\times U$, il existe une solution $(]a,b[,f)$ correspondant à (t_0,x_0).

Nous ne démontrerons pas ce résultat général; nous allons seulement, sur un exemple, décrire le principe de la construction de cette solution.

Considérons l'équation $x' = \cos(\sin x)$. Choisissons un nombre η dans $]0,1[$, nous allons chercher une solution définie sur l'intervalle $I = [-\eta,\eta]$ et correspondant à la donnée initiale $(0,x_0)$. La fonction continue $\varphi : I \to \mathbb{R}$ est solution si et seulement si $(\forall t)\ \varphi(t) = x_0 + \int_0^t \cos(\sin(\varphi(\tau)))\,d\tau$. En effet si cette relation est vérifiée, φ est dérivable (car primitive d'une fonction continue), et sa dérivée en t est $\cos(\sin(\varphi(t)))$. Réciproquement si $(\forall t)\ \varphi'(t) = \cos(\sin(\varphi(t)))$, par primitivation on obtient la relation annoncée, puisque $\varphi(0) = x_0$.

Une telle solution, s'il en existe, est un élément de l'espace E des fonctions continues de I dans $\mathbb{R}$, que nous munirons de la norme de la convergence uniforme. Nous définirons une application T de E dans lui même en posant:

$$(\forall \mu \in E) \qquad T(\mu)(t) = x_0 + \int_0^t \cos(\sin(\mu(\tau)))\,d\tau$$

On vérifie facilement que la dérivée de la fonction $t \to \cos(\sin t)$ est bornée par 1. Le théorème des accroissements finis nous permet donc d'affirmer que $(\forall t \text{ et } u)$

$$|\cos(\sin t) - \cos(\sin u)| \le |t - u|$$

Il en résulte que $(\forall \mu \text{ et } \nu)$ et $(\forall t)$

$$|T(\mu)(t)-T(\nu)(t)| = |\int_0^t \cos(\sin(\mu(\tau)))\,d\tau - \int_0^t \cos(\sin(\nu(\tau)))\,d\tau|$$

$$|T(\mu)(t)-T(\nu)(t)| = |\int_0^t \cos(\sin(\mu(\tau))) - \cos(\sin(\nu(\tau)))\,d\tau|$$

$$|T(\mu)(t)-T(\nu)(t)| \le |\int_0^t |\cos(\sin(\mu(\tau))) - \cos(\sin(\nu(\tau)))|\,d\tau|$$

$$|T(\mu)(t)-T(\nu)(t)| \le |\int_0^t |\mu(\tau)-\nu(\tau)|\,d\tau| \le |\int_0^t N_\infty(\mu-\nu)\,d\tau| \le \eta\, N_\infty(\mu-\nu)$$

Puisque ceci est vrai pour tout t, nous avons $N_\infty(T(\mu)-T(\nu)) \le \eta\, N_\infty(\mu-\nu)$. Puisque $\eta<1$, ceci signifie que $T: E \to E$ est contractante pour la norme uniforme.

Dés lors pour résoudre l'équation $\varphi = T(\varphi)$ qui nous est proposée, nous allons employer la méthode des approximations successives (cf: chap.11 §7). Nous choisissons φ_0 arbitraire dans E, et nous définissons une suite $(\varphi_n)_{n\in\mathbb{N}}$ par la relation de récurrence $(\forall n)$ $\varphi_{n+1} = T(\varphi_n)$. Soit $((u_n))_{n\in\mathbb{N}}$ la série associée (définie par $u_n = \varphi_{n+1} - \varphi_n$); on a $(\forall n\ge 1)$ $N_\infty(u_n) \le \eta\, N_\infty(u_{n-1})$, donc elle est normalement sommable (critère de D'Alembert*). Autrement dit la suite $(\varphi_n)_{n\in\mathbb{N}}$ converge pour N_∞ vers une limite φ (ce qui signifie que la suite $(N_\infty(\varphi-\varphi_n))_{n\in\mathbb{N}}$ converge vers 0). Puisque T est contractante on a

$$N_\infty(\varphi-\varphi_n) = N_\infty(T(\varphi)-T(\varphi_n)) \le \eta\, N_\infty(\varphi-\varphi_n)$$

Donc la suite $(N_\infty(T(\varphi)-T(\varphi_n)))_{n\in\mathbb{N}} = (N_\infty(T(\varphi)-\varphi_{n+1}))_{n\in\mathbb{N}}$ converge aussi vers 0. Puisque les suites $(\varphi_n)_{n\in\mathbb{N}}$ et $(\varphi_{n+1})_{n\in\mathbb{N}}$ ont la même limite ceci entraine que $\varphi = T(\varphi)$, autrement dit que φ est solution du problème posé.

Le théorème d'unicité

Nous supposons toujours F de classe C^1; si deux solutions $(]u,v[,h)$ et $(]c,d[,g)$ prennent la même valeur en un point $t_1 \in]u,v[\cap]c,d[$, alors f et g coïncident sur $]u,v[\cap]c,d[$.

Nous utiliserons le lemme suivant:

Lemme : Si h et g coïncident en un point t_0, il existe un intervalle centré en t_0 sur lequel elles coïncident.

Nous ne démontrerons pas ce lemme. Nous allons cependant observer sur l'exemple ci-dessus le principe de sa démonstration.

Supposons que deux solutions h et g de l'équation $x' = \cos(\sin x)$ soient définies au voisinage de 0, et prennent la même valeur x_0 en 0. Reprenons les notations de la démonstration précédente. Nous pouvons choisir η assez petit pour que h et g soient définies sur $I = [-\eta,\eta]$. Posons $\gamma = g|I$ et $\delta = h|I$. Puisque (I,γ) et (I,δ) sont deux solutions qui prennent la même valeur x_0 en 0, nous aurons $\gamma = T(\gamma)$ et $\delta = T(\delta)$. Puisque T est contractante, on en déduira que

$$N_\infty(\gamma-\delta) = N_\infty(T(\gamma)-T(\delta)) \le \eta\, N_\infty(\gamma-\delta)$$

Puisque $\eta<1$, ceci implique $N_\infty(\gamma-\delta) = 0$; c'est à dire $\gamma = \delta$.

Démontrons (à partir du lemme) que h et g coïncident sur $]u',v'[=]c,d[\cap]u,v[$. Nous allons montrer que h et g coïncident sur $[t_1,v'[$ (on démontrerait de façon analogue que h et g coïncident sur $]u',t_1]$). Notons X l'ensemble des points t tels que f et g coïncident sur $[t_1,t]$; il s'agit de montrer que $X =]t_1,v'[$. Pour cela on va montrer par l'absurde que $\sup X = v'$.

* On utilise ici le fait que E, muni de la norme uniforme, est complet.

Si $\sup X = \theta < v'$ *, de deux choses l'une, ou bien $\theta \in X$, ou bien $\theta \notin X$.

- Si $\theta \in X$, d'après le lemme, il existe un intervalle $]\theta,\theta+\varepsilon[$ dans lequel h et g coïncident. Donc les points de $]\theta,\theta+\varepsilon[$ sont aussi dans X, ce qui est contradictoire avec le fait que $\theta = \sup X$.

- Si $\theta \notin X$, considérons une suite $(x_n)_{n\in \mathbb{N}}$ de points de X qui converge vers θ; pour tout n, on a donc $h(x_n) = g(x_n)$; et (puisque h et g sont continues) à la limite $h(\theta) = g(\theta)$. Donc θ est dans X, ce qui est contredit l'hypothèse que $\theta \notin X$.

C'est donc qu'on ne peut pas avoir $\sup X < v'$.

§ 2 : Solutions maximales

L'ensemble des solutions est ordonné: on dit que $(]a,b[,f)$ est plus grande que $(]c,d[,g)$ si $]a,b[\supset]c,d[$ et si g est la restriction de f à $]c,d[$. *On dira qu'une solution $(]a,b[,f)$ est maximale si elle n'est pas prolongeable en une solution définie sur un intervalle plus grand; autrement dit s'il n'existe pas de solution strictement plus grande au sens de la relation d'ordre que l'on vient de définir* (cf: chap.10b §2)

Quels que soient (t_0,x_0) il existe une solution maximale et une seule qui correspond à la donnée initiale (t_0,x_0).

L'unicité résulte du théorème d'unicité du § 1 ci-dessus. Pour démontrer l'existence considérons toutes les solutions $(]a_i,b_i],f_i)$ correspondant à la donnée initiale (t_0,x_0). Posons $]a,b[= \bigcup_i]a_i,b_i[$. Si t appartient à $]a,b[$, il appartient à un certain nombre d'intervalles $]a_i,b_i[$, et par conséquent il existe plusieurs des fonctions f_i qui sont définies en t. Mais d'après le théorème d'unicité du § 1 toutes ces fonctions prennent la même valeur en t; cette valeur commune sera notée f(t). Alors $(]a,b[,f)$ est solution de l'équation différentielle. En effet si $t\in]a,b[$, il existe i tel que $t\in]a_i,b_i[$, et nous avons $f_i'(t) = F(t,f_i(t))$; ce qui s'écrit aussi $f'(t) = F(t,f(t)$ puisque, au voisinage de t, les fonctions f et f_i coïncident. Cette solution est maximale puisque toute solution correspondant à la valeur initiale (t_0,x_0) est l'une des $(]a_i,b_i[,f_i)$, donc est la restriction de f à l'intervalle $]a_i,b_i[$.

Nous admettrons le résultat suivant: *Si $F:]a,b[\times\mathbb{R}^p \to \mathbb{R}^p$, une solution $f:]\alpha,\beta[\to U$ de l'équation $x' = F(t,x)$ est maximale, si et seulement si :*

* *d'une part ou bien $\alpha = a$, ou bien f tend vers l'infini en α.*

* *d'autre part ou bien $\beta = b$, ou bien f tend vers l'infini en β.*

Exercice a : Considérons l'équation $y' = \sqrt{1-y^2}$. On a donc $F(t,y) = \sqrt{1-y^2}$ définie pour $-1 < y < 1$ (nous sommes obligés d'éliminer les cas $y = 1$ et $y = -1$, puisque F doit être de classe C^1).

a) Démontrer que, pour tout φ, le couple $(]\varphi - \frac{\pi}{2}, \varphi + \frac{\pi}{2}[, t \to \sin(t-\varphi))$ est une solution de l'équation.

b) Montrer qu'elle est maximale , c'est à dire qu'elle n'est prolongeable ni pour $t < \varphi - \frac{\pi}{2}$, ni pour $t > \varphi+ \frac{\pi}{2}$.

c) Montrer qu'on a ainsi toutes les solutions maximales.

§ 3: Exponentielles de matrices.

Soit $M = (m_{ij})$ une matrice p×p, posons $v(M) = \sup_{i,j}|m_{ij}|$. Nous obtenons une norme sur l'espace des matrices p×p (cf: chap.10). Si M et N sont deux matrices, les coefficients de MN sont, en valeur absolue inférieurs à $p\,v(M)v(N)$. Donc $v(MN) \le p\,v(M)v(N)$.

Considérons la série $u = \left(\left(\frac{M^n}{n!}\right)\right)_{n\in\mathbb{N}}$. De l'inégalité ci-dessus, on déduit

* v' et sup X sont finis ou $+\infty$

$$\nu\left(\frac{M^n}{n!}\right) \leq \frac{p^{n-1}(\nu(M))^n}{n!}$$

Donc la série u est normalement sommable. Sa somme est appelée l'exponentielle de la matrice M, et notée exp(M).

Calcul de Exp(M)

Si deux matrices M et N commutent alors exp(M+N) = exp(M)exp(N). Nous ne ferons pas la démonstration, qui est une adaptation de celle du chap.12 §4. Notons seulement que celle-ci repose sur l'égalité $\sum_{i+j=n} \frac{M^i}{i!}\frac{N^j}{j!} = \frac{(M+N)^n}{n!}$, qui n'est autre que la formule du binôme. C'est pourquoi nous avons besoin de l'hypothèse que MN = NM.

En particulier $\exp(-M) = [\exp(M)]^{-1}$.

Cas des matrices diagonalisables

Si $M = A\Delta A^{-1}$, où Δ est diagonale, pour tout n nous avons $\frac{M^n}{n!} = A\ \frac{\Delta^n}{n!}\ A^{-1}$; d'où l'on déduit que $(\forall N)\ \sum_{n=0}^{N} \frac{M^n}{n!} = A\left[\sum_{n=0}^{N} \frac{\Delta^n}{n!}\right]A^{-1}$. Et en passant à la limite $\exp(M) = A\exp(\Delta)A^{-1}$. Ceci nous permet de calculer l'exponentielle d'une matrice diagonalisable, puisque l'exponentielle de la matrice diagonale de coefficients δ_i est - c'est à peu près évident - la matrice diagonale de coefficients $\exp(\delta_i)$.

Cas des matrices nilpotentes

Si $M^n = 0$, la série est nulle à partir du rang n, donc $\exp(M) = Id+M+\frac{M^2}{2}+\ldots+\frac{M^{n-1}}{(n-1)!}$.

Cas général

Au chapitre 19, on a appris à écrire toute matrice M sous la forme M_0+N, où M_0 est diagonalisable et N nilpotente, et où $M_0N = NM_0$. Il en résulte que $\exp(M) = \exp(M_0)\exp(N)$. Et nous savons calculer $\exp(M_0)$ et $\exp(N)$.

Exercice b : Reprendre les matrices M que l'on a réduites au chapitre 19, et calculer pour chacune d'elles exp(tM).

§ 4 : Equations différentielles linéaires à coefficients constants.

Equations linéaires homogènes à coefficients constants.

Ce sont les équations de la forme y' = Ay, où (y est une fonction à valeurs dans $\mathbb{R}^p$, et) A est une matrice (qui ne dépend pas de t). Nous avons déjà étudié ces équations lorsque p = 1 (cf: Chap.3 §3). Leurs solutions peuvent être calculées explicitement.

Soit $y_0 \in \mathbb{R}^p$ et $t_0 \in \mathbb{R}$, alors la fonction $t \to \varphi(t) = \exp((t-t_0)A).y_0$ est solution de l'équation homogène y' = Ay. C'est la solution maximale (unique) qui prend la valeur y_0 en t_0.

En effet: Il est clair que $\varphi(t_0) = y_0$, et que φ étant définie sur $\mathbb{R}$ tout entier, si c'est une solution elle est maximale. Il nous faut donc montrer que φ est dérivable et que $(\forall t)\ \varphi'(t) = A\varphi(t)$.

Lemme: Soit ν la norme sur l'espace des matrices déjà utilisée au § 3, alors

$$\nu(\exp(hA) - Id - hA) \leq \frac{1}{p}\left[e^{p|h|\nu(A)} - p|h|\nu(A) - 1\right] = o(h)$$

Démonstration: Posons $\Sigma_n = \sum_{i=2}^{n} \frac{h^iA^i}{i!}$. Alors:

$$\nu(\Sigma_n) \leq \sum_{i=2}^{n} \nu\left(\frac{h^iA^i}{i!}\right) = \sum_{i=2}^{n} \frac{|h|^i}{i!}\nu(A^i)$$

$$\leq \sum_{i=2}^{n} \frac{|h|^i}{i!} p^{i-1} \nu(A)^i \qquad\qquad (\text{car } \nu(MN) \leq p\, \nu(M)\, \nu(N))$$

Et $\displaystyle\sum_{i=2}^{n} \frac{|h|^i}{i!} p^{i-1} \nu(A)^i \leq \sum_{i=2}^{\infty} \frac{|h|^i}{i!} p^{i-1} \nu(A)^i = \frac{1}{p}\left[e^{p|h|\nu(A)} - p|h|\nu(A) - 1\right]$

Donc $(\forall n)$ $\quad \nu(\Sigma_n) \leq \frac{1}{p}\left[e^{p|h|\nu(A)} - p|h|\nu(A) - 1\right]$.

Et puisque la suite $(\Sigma_n)_{n\in\mathbb{N}}$ converge vers $\exp(hA) - hA - Id$, en passant à la limite on obtient

$$\nu(\exp(hA) - hA - Id) \leq \frac{1}{p}\left[e^{p|h|\nu(A)} - p|h|\nu(A) - 1\right].$$

Ce qui termine la démonstration du lemme, et nous permet de montrer que φ est dérivable. En effet pour tout t et pour tout h, nous avons:

$$\begin{aligned}\varphi(t+h) - \varphi(t) - hA\varphi(t) &= [\exp((t+h-t_0)A) - \exp((t-t_0)A) - hA\exp((t-t_0)A)]y_0 \\ &= [\exp(hA) - Id - hA]\,[\exp((t-t_0)A)y_0].\end{aligned}$$

Notons N_∞ la norme sur $\mathbb{R}^p$ définie par $N_\infty(y_1,\ldots,y_p) = \sup_i(|y_i|)$. Pour toute matrice B et tout vecteur y, nous avons $N_\infty(By) \leq \nu(B)\, N_\infty(y)$ (vérification facile). Donc

$$\begin{aligned}N_\infty(\varphi(t+h) - \varphi(t) - hA\varphi(t)) &\leq \nu(\exp(hA) - Id - A)\; N_\infty(\exp((t-t_0)A)y_0) \\ &\leq \frac{1}{p}\left[e^{p|h|\nu(A)} - p|h|\nu(A) - 1\right]\; N_\infty(\exp((t-t_0)A)y_0)\end{aligned}$$

Donc $\qquad N_\infty(\varphi(t+h) - \varphi(t) - hA\varphi(t)) = o(h)\, N_\infty(\exp((t-t_0)A)y_0) = o(h)$

Formule qui prouve qu'au point t, φ est dérivable de dérivée $A\varphi(t)$; ce qui termine la démonstration.

Les équations linéaires non homogènes à coefficients constants. Ce sont les équations du type $y' = Ay + B(t)$; où A est une matrice (indépendante de t) et B une fonction continue de t.

La méthode dite "de variation de la constante" consiste à chercher les fonctions $z: \mathbb{R} \to \mathbb{R}^p$ telles que la fonction $t \to \exp(tA)\, z(t)$ soit une solution de $y' = Ay+B(t)$.

Nous admettrons que la dérivée de $t \to \exp(tA)z(t)$ est $t \to A\exp(tA)\, z(t)+\exp(tA)\, z'(t)$; et l'équation devient donc $(\forall t)\ \exp(tA)\, z'(t) = B(t)$. Ce qui s'écrit aussi $z'(t) = \exp(-tA)B(t)$. Et nous obtiendrons z par une primitivation.

Calculs pratiques

Ce qui précède nous donne un calcul des solutions d'une équation linéaire à coefficients constants, à partir du calcul de l'exponentielle des matrices. Cette méthode est souvent assez longue à mettre en œuvre. Dans la pratique il est quelquefois plus simple de ne pas chercher à réduire complètement la matrice A, mais de se contenter de faire un changement de base qui la rend triangulaire. On a alors un système qui se résout étapes par étapes. Voici deux exemples.

Exercice d: Soit à résoudre $\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} t \\ 0 \end{pmatrix}$

a) On pose $M = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Vérifier que $M^{-1}\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}M$ est une matrice triangulaire T.

On pose $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = M^{-1}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Montrer que l'équation proposée est équivalente à $\begin{pmatrix} z_1' \\ z_2' \end{pmatrix} = T\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + B_1$ où B_1 est une fonction de t que l'on précisera.

b) Calculer alors z_2 **puis** z_1. En déduire y_1 et y_2.

Exercice e : On considère dans $\mathbb{R}^3$ l'équation $y' = Ay$ où $A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ -1 & -1 & 3 \end{pmatrix}$.

a) Trouver une matrice M telle que $M^{-1}AM$ soit triangulaire.

b) En s'inspirant de l'exercice d déterminer les fonctions $z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} : \mathbb{R} \to \mathbb{R}^3$ telles que $t \to M\,z(t)$ soit solution de l'équation proposée. Puis en déduire les solutions de cette dernière.

§ 5 : Les avatars d'une équation différentielle.

On prendra garde qu'une même équation différentielle peut se présenter sous des formes très différentes. Nous ne ferons pas de théorie générale sur les diverses présentations que l'on peut donner d'une même équation. Voici, sur un exemple, quelques unes d'entre elles.

Du second ordre au premier ordre

Considérons l'équation linéaire du second ordre

(1) $$y'' = -2y + 2y'$$

(où la fonction inconnue y est à valeurs dans $\mathbb{R}$). Nous pouvons la transformer en une équation du premier ordre dont la fonction inconnue Y est à valeurs dans $\mathbb{R}^2$. Considérons en effet l'équation linéaire à coefficients constants

(2) $$\begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix} \quad \text{soit} \quad \begin{cases} y' = z \\ z' = -2y + 2z \end{cases}$$

Si la fonction y est solution de l'équation du second ordre, alors $Y = \begin{pmatrix} y \\ y' \end{pmatrix}$ est solution de l'équation vectorielle. Réciproquement si $Y = \begin{pmatrix} y \\ z \end{pmatrix}$ est solution de l'équation vectorielle, alors y est solution de l'équation du second ordre.

De façon générale toute équation du second ordre (dont l'inconnue est à valeurs dans $\mathbb{R}^p$) peut ainsi être transformée en une équation du premier ordre (dont l'inconnue est à valeurs dans $\mathbb{R}^{2p}$). Pour cela il suffit de prendre pour nouvelle inconnue le couple formé de l'inconnue à valeurs dans $\mathbb{R}^p$ et de sa dérivée. En particulier les équations linéaires du second ordre à coefficients constants à valeurs dans $\mathbb{R}^p$ nous donneront des équations linéaires à coefficients constants du premier ordre à valeurs dans $\mathbb{R}^{2p}$.

C'est pourquoi dans toute étude théorique sur les équations différentielles, on oublie systématiquement les équations d'ordre 2.

Notons qu'une donnée initiale pour une fonction à valeurs dans $\mathbb{R}^{2p}$, est la donnée pour un certain t_0 d'un vecteur de $\mathbb{R}^{2p}$, c'est à dire de 2p scalaires. Si celle ci provient d'une équation du second ordre (à valeurs dans $\mathbb{R}^p$), ces 2p scalaires sont les valeurs en t_0 de la fonction inconnue à valeurs dans $\mathbb{R}^p$ et de sa dérivée.

Par exemple dans les équations de la mécanique, qui sont du second ordre, une donnée initiale est le couple formé de la position du mobile, et de sa vitesse à l'instant t_0. Dans ce cas le théorème d'unicité des solutions signifie que, dans un champ de forces donné (c'est à dire pour une équation donnée), la position initiale et la vitesse initiale déterminent complètement le mouvement. C'est là l'un des principes de base de la mécanique, admis depuis Newton.

Les courbes solutions d'une équation à valeurs dans $\mathbb{R}^2$.

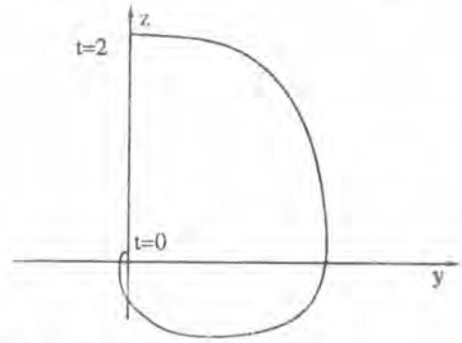
Les solutions de $\begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix}$ sont des fonctions (d'une variable t) à valeurs dans $\mathbb{R}^2$. Toute solution $t \to \begin{pmatrix} y(t) \\ z(t) \end{pmatrix}$ définit une courbe paramétrée Γ du plan. Soit Γ_1 et Γ_2 deux de ces courbes passant au même point M pour t égal, respectivement, à t_1 et à t_2. Puisque l'équation est à coefficients constants la fonction $t \to \Gamma_2(t_2-t_1+t) = \begin{pmatrix} y_2(t_2-t_1+t) \\ z_2(t_2-t_1+t) \end{pmatrix}$ est aussi une solution. Cette solution prend la valeur M pour $t = t_1$, donc (théorème d'unicité des solutions) c'est la solution $t \to \Gamma_1(t)$. Ainsi les deux courbes Γ_1 et Γ_2 sont une seule et même courbe qui a reçu deux paramétrages différents.

Nous avons ainsi démontré que deux courbes solutions ne peuvent avoir un point commun que si elles sont confondues; par ailleurs (puisqu'à toute donnée initiale correspond une solution) par tout point du plan passe une courbe solution. Dans le cas présent une des solutions est la courbe Γ_1 définie par :

$$\begin{cases} y(t) = e^t \cos t \\ z(t) = e^t \cos t - e^t \sin t \end{cases}$$

Pour tracer Γ_1 on trace d'abord l'arc correspondant à $0 \le t \le 2\pi$, puis on trace le reste de la courbe par les homothéties de rapports $e^{2n\pi}$ (où $n \in \mathbb{Z}$).

Les courbes Γ_K homothétiques de Γ_1 dans les homothéties de centre l'origine et de rapport K sont aussi des solutions. Et puisque par tout point du plan, sauf l'origine, passe une unique courbe Γ_K, nous avons ainsi obtenu toutes les courbes solutions, sauf une: celle qui correspond à la solution constante égale à 0, et qui se réduit à l'origine.



C'est une famille de spirales s'enroulant les unes autour des autres.

Variations sur le paramétrage des courbes solutions

Considérons maintenant l'équation $\begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -4 & 4 \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix}$. Il est clair que $t \to \begin{pmatrix} y(t) \\ z(t) \end{pmatrix}$ est une solution du système (1), si et seulement si $t \to \begin{pmatrix} y(2t) \\ z(2t) \end{pmatrix}$ est une solution de ce nouveau système. Autrement dit les courbes solutions des deux systèmes sont identiques.

Plus généralement si φ est une fonction (de classe C^1) de $\mathbb{R}$ dans $\mathbb{R}$, nous pouvons considérer l'équation

(3) $$\begin{pmatrix} y' \\ z' \end{pmatrix} = \varphi \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix}$$

(qui n'est plus à coefficients constants si φ n'est pas constante). Alorts si Φ est une primitive de φ, et si $t \to \begin{pmatrix} y(t) \\ z(t) \end{pmatrix}$ est une solution de l'équation (1), $t \to \begin{pmatrix} y(\Phi(t)) \\ z(\Phi(t)) \end{pmatrix}$ est une solution de (3). Les équations (1) et (3) ont mêmes courbes solutions, mais elles donnent des paramétrages différents de ces courbes.

Si F est une fonction de classe C^1 de $\mathbb{R}^2$ dans $\mathbb{R}$, nous pouvons considérer l'équation:

(3') $$\begin{pmatrix} y' \\ z' \end{pmatrix} = F(y,z) \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} y \\ z \end{pmatrix}$$

Si $t \to \begin{pmatrix} y(t) \\ z(t) \end{pmatrix}$ est solution de l'équation (2), et G une primitive de la fonction $t \to F(y(t),z(t))$, alors $t \to \begin{pmatrix} y(G(t)) \\ z(G(t)) \end{pmatrix}$ est solution de (3'). Ici encore nous retrouvons les mêmes courbes solutions que pour l'équation (2), mais avec un autre paramétrage.

Retour à une équation de degré 1 à valeurs dans $\mathbb{R}$.

Un cas particulier de (3') particulièrement intéressant est celui où $F(y,z) = \frac{1}{z}$ (notons qu'ici F n'est pas définie pour $z = 0$, nous nous restreignons donc au complémentaire de la droite $z = 0$ dans le plan de (y,z)). Dans ce cas l'équation (3') devient:

$$(4) \qquad \begin{cases} y' = 1 \\ z' = -2\,\dfrac{y}{z} + 2 \end{cases}$$

Puisque $y' = 1$, c'est que $y = t+h$, donc $z' = \dfrac{dz}{dy}$, par conséquent notre nouvelle équation s'écrit aussi :

$$(4') \qquad \frac{dz}{dy} = -2\frac{y}{z} + 2$$

C'est donc une équation d'ordre 1, de la fonction inconnue à valeurs réelles $y \to z(y)$. Le graphe d'une solution de (4') est encore une courbe solution de (2) (ou plutôt un morceau de courbe solution de (2), puisque nous travaillons dans le complémentaire de la droite $z = 0$). Comme dans (3) ou (3') nous avons fait un changement de paramètre sur les courbes solutions; ici nous avons choisi y comme nouveau paramètre (ce qui n'est possible que là où la tangente à ces courbes n'est pas parallèle à Oz, c'est pourquoi il a fallu se placer dans le complémentaire de la droite $z = 0$).

Equations différentielles et champs de vecteurs

Nous pouvons encore modifier l'aspect de l'équation 2 de la façon suivante. Considérons l'application $X: \mathbb{R}^2 \to \mathbb{R}^2$ qui à (y,z) associe $(z,-2y+2z)$; une telle application peut être considérée comme un champ de vecteurs sur $\mathbb{R}^2$. Les solutions de l'équation (2) sont les fonctions $\gamma: t \to \begin{pmatrix} y(t) \\ z(t) \end{pmatrix}$ dont le vecteur dérivé en t, est le vecteur du champ au point $\gamma(t)$. Ainsi l'équation différentielle s'écrit aussi:

$$(5) \qquad \frac{d\gamma}{dt}(t) = X(\gamma(t))$$

Les courbes solutions sont aussi appelées les courbes intégrales du champ de vecteurs X.

Remplaçons le champ $M \to X(M)$, par le champ $M \to F(M)X(M)$ (où F est une quelconque application de classe C^1 de $\mathbb{R}^2$ dans $\mathbb{R}$). C'est une modification analogue à celle qui nous a donné l'équation (3') ci-dessus; en termes de champs de vecteurs, elle nous donne l'équation:

$$(5') \qquad \frac{d\gamma}{dt}(t) = F(\gamma(t))\, X(\gamma(t))$$

Les courbes intégrales du champ $M \to F(M)X(M)$ sont les mêmes que celles du champ $M \to X(M)$ (puisque (3') et (2) ont mêmes courbes solutions).

§ 6 : Quelques méthodes de calcul pour les équations non linéaires.

Nous allons donner ici quelques recettes permettant de résoudre certaines équations (à valeurs dans $\mathbb{R}$) de types bien particuliers.

Equations à variables séparées.

Elles sont de la forme $f(x) = g(y)\dfrac{dy}{dx}$ (ou elles peuvent être ramenées à cette forme). Si F et G sont des primitives de f et g, cette équation s'écrit aussi $\dfrac{dF}{dx}(x) = \dfrac{dG\circ y}{dx}(x)$. Donc pour toute solution $x \to y(x)$ il existe une constante k telle que $F(x) + k = G(y(x))$.

Si G est une fonction inversible, ceci nous donne $y(x) = G^{-1}(F(x)+k)$. La solution correspondant à la donnée initiale (x_0,y_0) est donnée par $k = G(y_0) - F(x_0)$. Et puisque, à chaque donnée initiale, nous savons ainsi faire correspondre une solution, c'est que nous avons toutes les solutions de l'équation. Si G est seulement localement inversible au voisinage d'un point y_0 (et pour cela il suffit que $G'(y_0) \neq 0$, c'est à dire $g(y_0) \neq 0$), nous obtiendrons ainsi les solutions dont les valeurs sont voisines de y_0.

Equations d'ordre 2 où le temps ne figure pas

Elles sont de la forme $y'' = F(y,y')$ (ou se ramènent à ce type). Avant de chercher les solutions $t \to y(t)$, nous chercherons une fonction z telle que $(\forall t)\ \ z(y(t)) = y'(t)$.

En dérivant cette relation nous obtenons: $(\forall t)\ \ \frac{dz}{dy}(y(t))\, y'(t) = y''(t)$. Alors l'équation s'écrit

$$(\forall t) \qquad (y''(t) =)\ \frac{dz}{dy}(y(t))\, z(y(t)) = F(y(t),z(y(t)))$$

Pour calculer z nous sommes ainsi amenés à résoudre (si nous savons le faire !!) l'équation d'ordre 1 (où $\xi = y$ est la variable) $\qquad z'(\xi)\, z(\xi) = F(\xi,z(\xi))$

Pour toute solution $\xi \to z(\xi)$ de cette nouvelle équation, nous résoudrons ensuite l'équation (à variables séparées):

$$y' = z(y)$$

Nous avons décomposé le problème posé en deux équations d'ordre 1 successives.

Remarquons que si $t \to y(t)$ est une solution de l'équation $y'' = F(y,y')$, alors (quelle que soit la constante h) $t \to y(t+h)$ en est une autre.

Equations du second ordre où la fonction inconnue ne figure que par ses dérivées

Elles sont du type $F(t,y',y'') = 0$. Nous pouvons considérer une telle équation comme une équation d'ordre 1 dont l'inconnue est la fonction y'. Lorsque nous aurons résolu cette équation, il nous restera à calculer les primitives de ses solutions.

Equations homogènes

Ce sont les équations du premier ordre $\frac{dy}{dx} = F(y,x)$ telles que $(\forall \lambda \neq 0)\ F(\lambda y,\lambda x) = F(y,x)$. Les graphes des solutions d'une telle équation sont les courbes solutions de l'équation à valeurs dans $\mathbb{R}^2$:

$$\begin{cases} y'(t) = F(y(t),x(t)) \\ x'(t) = 1 \end{cases}$$

Soit $\Gamma(t \to x(t),y(t))$ une telle courbe, l'homothétie de centre l'origine et de rapport λ, la transforme en $\Gamma_\lambda\ (t \to \lambda x(t),\lambda y(t))$ (ou, par changement du paramétrage $t \to \lambda x(\frac{t}{\lambda}),\lambda y(\frac{t}{\lambda})$), qui est clairement une autre courbe solution. Ainsi dans $\mathbb{R}^2$ privé de l'origine, les courbes solutions sont

* Soit des droites passant par l'origine (chacune d'elles est sa propre transformée dans les homothéties de centre l'origine).

* Soit des familles de courbes Γ_λ, liées par la relation: $(\forall \lambda,\mu)\ \Gamma_\mu$ est la transformée de Γ_λ dans l'homothétie de centre l'origine et de rapport μ/λ.

Pour résoudre effectivement nous chercherons à calculer les fonctions $x \to p(x)$ telles que $x \to y(x) = xp(x)$ soit solution de l'équation. Ce qui nous amène à résoudre:

$$xp'(x) + p(x) = F(p(x),x)$$

Mais on peut aussi passer en polaires dans d'équation à valeurs dans $\mathbb{R}^2$, c'est à dire chercher deux fonctions $\rho(t)$ et $\omega(t)$ telles que la courbe $(x(t) = \rho(t) \cos \omega(t)\ ,\ y(t) = \rho(t) \sin \omega(t))$ en soit une courbe solution. Nous obtenons alors:

$$\begin{cases} \rho' \cos \omega - \rho\, \omega' \sin \omega = 1 \\ \rho' \sin \omega + \rho\, \omega' \cos \omega = F(\rho \cos \omega,\rho \sin \omega) \end{cases}$$

Ce qui nous donne:

$$\begin{cases} \rho' = \cos\omega + \sin\omega\, F(\rho\cos\omega, \rho\sin\omega) = \cos\omega + \sin\omega\, F(\cos\omega,\sin\omega) \\ \rho\,\omega' = -\sin\omega + \cos\omega\, F(\rho\cos\omega, \rho\sin\omega) = -\sin\omega + \cos\omega\, F(\cos\omega,\sin\omega) \end{cases}$$

Soit encore:

$$\frac{d\rho}{\rho} = \frac{\cos\omega + \sin\omega\, F(\cos\omega,\sin\omega)}{-\sin\omega + \cos\omega\, F(\cos\omega,\sin\omega)}\, d\omega$$

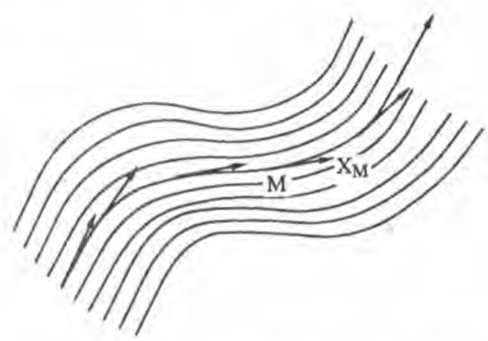
Ce qui est une équation à variables séparées.

Complément : Linéarisation d'une équation différentielle.

Nous avons appris à résoudre les équations linéaires à coefficients constants; ce sont en fait à peu près les seules que l'on sache résoudre explicitement (le § 6 donne quelques autres cas où la résolution complète est possible). Que faire d'une équation du type $y' = F(t,y)$, où F est de classe C^1 ?

Pour simplifier nous supposerons que F ne dépend pas de t, c'est donc une fonction de la seule variable y. Ceci nous permettra de la considérer comme un champ de vecteurs sur $\mathbb{R}^p$. Abandonnons la recherche des solutions maximales, pour nous intéresser à l'allure des solutions "au voisinage de (t_0,y_0)". Nous allons remplacer F par une fonction "voisine", par exemple par un de ses développements de Taylor.

Au voisinage d'un point où F ne s'annule pas



Essayons d'abord le développement de Taylor d'ordre 0. Nous résoudrons donc l'équation $y' = F(y_0)$. Interprétée comme un champ de vecteurs, cette nouvelle équation nous donne un champ constant. Elle a donc pour courbes intégrales des droites paralléles. Mais il faut démontrer que, au voisinage de y_0, les courbes intégrales de l'équation initiale "ressemblent" à ces droites parallèles.

On démontre (nous l'admettrons) que, au voisinage de z_0, les courbes solutions sont les images par un difféomorphisme d'une famille de droites parallèles.

Au voisinage d'un point où F s'annule.

Si $F(y_0) = 0$, il n'est évidemment plus question d'utiliser le développement d'ordre 0; nous obtiendrions l'équation $y' = 0$, dont les solutions constantes n'ont rien de commun avec celles de l'équation donnée. *Nous allons remplacer f par son développement d'ordre 1, c'est à dire par sa différentielle. Nous obtenons l'équation:*

$$y' = d_{y_0}F.(y-y_0)$$

*Elle est linéaire à coefficients constants: on dit que c'est la linéarisation de l'équation donnée au voisinage de y_0.** .

Bien entendu les solutions de notre équation linéaire ne sont pas les solutions de l'équation initiale; tout ce qu'on peut espérer c'est que, au voisinage de (t_0,y_0), elles leur ressemblent. L'un des résultats les plus importants, et les plus faciles à démontrer est le suivant.

---

* Par exemple, dans tous les cours de mécanique on remplace l'équation du pendule simple $y'' + k \sin y = 0$, par l'équation $y'' + k\, y = 0$. C'est un cas typique de linéarisation, faite directement sur l'équation de degré 2.

<u>Le problème de la stabilité</u> : Puisque $f(y_0) = 0$, la fonction constante $t \to y_0$ est une solution de l'équation initiale. On dit que y_0 est un point d'équilibre de l'équation. *Nous dirons que cette position d'équilibre est stable s'il existe un ouvert Ω contenant y_0 tel que si une solution $t \to y(t)$ passe en un point de Ω pour $t = t_0$ (i.e. si $y(t_0) \in \Omega$), alors:*

a) Pour $t > t_0$ elle reste dans Ω

b) $y(t)$ tend vers y_0 lorsque t tend vers $+\infty$.

Autrement dit si une solution est, à l'instatnt t_0 suffisamment proche de la position d'équilibre y_0, elle ne peut ensuite que s'en rapprocher encore plus. Au contraire si y_0 est un point d'équilibre instable, aussi près qu'on veut de y_0, on pourra trouver un point y_1 tel que la solution correspondant à la donnée initiale (t_0,y_1) s'éloigne de y_0 pour $t>t_0$ (et ne tende pas vers y_0 pour t tendant vers l'infini).

Le point y_0 est un point d'équilibre de l'équation linéarisée, cherchons d'abord à quelle condition il est stable. Deux raisonnements sont possibles.

Premier raisonnement: Nous savons que la solution qui correspond à la valeur initiale (t_0,y_1) s'écrit $t \to \exp((t-t_0)d_{y_0}F)(y_1-y_0)$. Une étude attentive du calcul de $\exp((t-t_0)d_{y_0}F)$, montre que les coefficients de cette matrice sont de la forme $\sum e^{r_i(t-t_0)} P_i(t-t_0)$ (où les r_i sont les valeurs propres de $d_{y_0}F$, et les P_i des polynômes). Si les parties réelles des r_i sont toutes négatives, tous ces coefficients convergent vers 0, pour t tendant vers l'infini, et l'équilibre est stable. Si certaines des r_i ont des parties réelles positives ou nulles, certains de ces coefficients ne tendent pas vers 0, pour t tendant ves l'infini, et l'équilibre est instable. Il resterait à écrire en détail ce raisonnement basé uniquement sur le calcul des exponentielles de matrices; nous allons en donner un autre.

Deuxième raisonnement : Supposons que toutes les valeurs propres de $d_{y_0}F$ (qui sont à priori de nombres complexes) aient des parties réelles négatives. Dans ce cas on peut démontrer (lemme en fin de paragraphe) qu'il existe un produit scalaire sur $\mathbb{R}^p$ (noté $<\ ,\ >$) et un nombre λ ($\lambda<0$) tels que

$$(\forall u) \qquad <d_{y_0}F(u),u> \leq \lambda <u,u>$$

Soit y une solution de l'équation $y' = d_{y_0}F.(y-y_0)$ différente de la solution constante y_0 (c'est à dire telle que $(\forall t)\ y(t) \neq y_0$); nous aurons:

$$\frac{d}{dt}(<y(t)-y_0,y(t)-y_0>) = 2<y'(t),y(t)-y_0> = 2<d_{y_0}F.(y(t)-y_0),y(t)-y_0> \leq 2\lambda<y(t)-y_0,y(t)-y_0>$$

Donc la dérivée du logarithme de $<y(t)-y_0,y(t)-y_0>$ est au plus égale à 2λ, il en résulte que

$$(\forall t>t_0) \qquad \mathrm{Log}(<y(t)-y_0,y(t)-y_0>) - \mathrm{Log}(<y(t_0)-y_0,y(t_0)-y_0>) \leq 2\lambda\,(t-t_0)$$

Soit, en notant N la norme associée à ce produit scalaire:

$$N(y(t)-y_0) \leq N(y(t_0)-y_0)\, e^{\lambda(t-t_0)}$$

Et, puisque $\lambda < 0$, ceci implique que $(\forall t>t_0)\ N(y(t)-y_0) < N(y(t_0)-y_0)$, et que $N(y(t)-y_0)$ tend vers 0 si t tend vers $+\infty$.

<u>Le théorème de Liapounov</u>.

Nous allons démontrer que l'équation $y' = F(y)$ se comporte comme sa linéarisée, c'est à dire que: *si toutes les valeurs propres de $d_{y_0}F$ (qui sont à priori des nombres complexes) ont une partie réelle négative* *, *alors la position d'équilibre y_0 de l'équation $y' = F(y)$ est stable*. Ce résultat est connu sous le nom de théorème de Liapounov.

Démonstration : Il existe un produit scalaire sur $\mathbb{R}^p$ (noté $<\ ,\ >$) et un nombre λ (<0) tel que $(\forall u)$ $<d_{y_0}F(u),u> \leq \lambda <u,u>$ (lemme en fin de paragraphe). Notant toujours N la norme associée à ce produit scalaire, nous pouvons trouver $\alpha(>0)$ tel que

* On remarquera que cette condition implique qu'aucune valeur propre n'est nulle, donc que $d_{y_0}F$ est inversible.

$$N(v-y_0) \leq \alpha \text{ implique } N(F(v) - d_{y_0}F(v-y_0)) \leq \frac{|\lambda|}{2}\; N(v-y_0)$$

(puisque $F(y_0) = 0$, ceci signifie seulement que F est différentiable en y_0). Donc $N(v-y_0) \leq \alpha$ implique

$$<F(v) - d_{y_0}F(v-y_0)),v-y_0> \;\leq\; N(F(v) - d_{y_0}F(v-y_0))\, N(v-y_0) \;\leq\; \frac{|\lambda|}{2}\; (N(v-y_0))^2$$

D'où
$$<F(v),v-y_0> \;\leq\; <d_{y_0}F(v-y_0),v-y_0> + \frac{|\lambda|}{2}\; (N(v-y_0))^2$$

Et, puisque $(\forall u)\; <d_{y_0}F(u),u> \leq \lambda <u,u>$, on en déduit:

$$N(v-y_0) \leq \alpha \text{ implique } <F(v),v-y_0> \;\leq\; \lambda\, (N(v-y_0))^2 + \frac{|\lambda|}{2}\; (N(v-y_0))^2$$

D'où
$$<F(v),v-y_0> \;\leq\; \frac{\lambda}{2}(N(v-y_0))^2$$

Nous pouvons alors reprendre la démonstration que nous avons faite dans le cas de l'équation linéarisée. Soit $t \to y(t)$ une solution telle que $N(y(t_0)-y_0) < \alpha$, alors pour $t>t_0$, et tant que $N(y(t)-y_0)$ reste inférieur à α, nous pouvons écrire:

$$\frac{d}{dt}(<y(t)-y_0,y(t)-y_0>) = 2<y'(t),y(t)-y_0> = 2<F(y(t)),y(t)-y_0> \leq 2\,\frac{\lambda}{2} <y(t)-y_0,y(t)-y_0>$$

Donc la dérivée du logarithme de $<y(t)-y_0,y(t)-y_0>$ est au plus égale à λ, il en résulte que

$$(\forall t>t_0) \qquad \mathrm{Log}(<y(t)-y_0,y(t)-y_0>) - \mathrm{Log}(<y(t_0)-y_0,y(t_0)-y_0>) \leq \lambda\,(t-t_0)$$

Soit , en notant N la norme associée à ce produit scalaire:

$$N(y(t)-y_0) \leq N(y(t_0)-y_0)\; e^{\lambda(t-t_0)/2}$$

Et, puisque $\lambda < 0$, ceci implique que $N(y(t)-y_0) < N(y(t_0)-y_0)$. Donc $N(y(t)-y_0)$ reste inférieur à α pour toute valeur de t supérieure à t_0; et $N(y(t)-y_0)$ tend vers 0 si t tend vers $+\infty$.

A notre époque il existe des programmes d'ordinateurs permettant de calculer avec une précision raisonnable la solution correspondant à la donnée initiale (t_1,y_1) de l'équation $y' = \mathcal{F}(y)$. Les plus simples sont basés sur le principe suivant .

Nous choisissons $\alpha>0$. Sur l'intervalle $[t_1,t_1+\alpha]$ nous poserons $\varphi(t) = y_1 + \mathcal{F}(y_1)\,(t-t_1)$. C'est une fonction affine, et si α est petit elle est 'voisine d'une solution' (par exemple au sens de la convergence uniforme) puisque $\mathcal{F}(\varphi(t))$ reste voisin de $\mathcal{F}(y_1)$. Mais il ne faut pas s'écarter trop de t_1; c'est pourquoi lorsque t est entre $t_1+\alpha$ et $t_1+2\alpha$, nous poserons $\varphi(t) = \varphi(t_1) + \mathcal{F}(\varphi(t_1+\alpha))\,(t-t_1-\alpha)$. C'est aussi une fonction affine, mais nous avons réactualisé la dérivée en passant au point $t_1+\alpha$. Nous changerons aussi la dérivée en $t_1+2\alpha$, puis en $t_1+3\alpha$,... Et nous obtiendrons une fonction affine par morceaux (qui est donc de classe C^1 sauf en un nombre fini de points). On démontre qu'elle est proche de la solution correspondant à (t_1,y_1), pourvu que α soit choisi petit, mais surtout que l'on reste dans un intervalle $[t_1,t_2]$ pas trop grand.

Plus cet intervalle sera grand, plus il faudra choisir α petit, donc plus le nombre de pas du calcul deviendra grand (et il devient vite inacceptable !). En particulier ce calcul approché ne permet pas de connaitre le comportement de la solution lorsque t tend vers l'infini. On comprend mieux l'intrérêt du théorème de Liapounov, qui nous dit avec très peu de précision, mais avec certitude, ce qui se passe lorsque t devient très grand.

Pour terminer la démonstration du théorème de Liapounov, il nous reste à démontrer le:

<u>Lemme</u>: Soit A un endomorphsime de $\mathbb{R}^p$ dont toutes les valeurs propres λ_i ont une partie réelle négative (on supposera que $(\forall i)\; \mathcal{R}(\lambda_i) \leq 2k < 0$). Alors il existe un produit scalaire sur $\mathbb{R}^p$ tel que, quel que soit u, $<A(u),u> \leq k <u,u>$.

Cas où $A = \lambda Id + N$ (où $\lambda=2k$ est réel négatif, et N nilpotent). Choisissons une base $\{e_1,\ldots,e_n\}$ dans la quelle la matrice ν de N est triangulaire (avec sur la diagonale les valeurs propres, c'est à dire des zéros).

Si nous remplaçons le vecteur e_i par αe_i cette matrice v est modifiée; sa i-ème colonne est multipliée par α, et sa i-ème ligne est divisée par α (ce qui ne change pas les coefficients diagonaux). En multipliant successivement les vecteurs $e_2,\ldots,e_n$ par des scalaires convenables nous pouvons donc fabriquer une nouvelle base $\{e'_1,e'_2\ldots,e'_n\}$ (où $e'_1 = e_1$) qui donnera à N une matrice v' dont tous les coefficients (non diagonaux) sont au plus égaux à $\varepsilon = \frac{|\lambda|}{p(p-1)}$. Nous choisirons le produit scalaire qui rend cette base orthormée. Alors pour tout $u = \sum_i x_i e_i$ nous aurons (en notant v'_{ij} les coefficients de v', et $\|u\|$ la norme associée du vecteur u):

$$<Au,u> = <\lambda\sum_i x_i e_i ,\sum_i x_i e_i> + <\sum_{i\neq j} v'_{ij} x_i e_j,\sum_i x_i e_i>$$

$$<Au,u> = \lambda \|u\|^2 + \sum_{i\neq j} v'_{ij} x_i x_j$$

D'où, puisque $(\forall i)\ |x_i| \leq \|u\|$, et puisque v' a seulement $\frac{p(p-1)}{2}$ coefficients (non diagonaux) éventuellement non nuls.

$$|<Au,u> - \lambda \|u\|^2| \leq \sum_{i\neq j} |v'_{ij}|\ \|u\|^2 \leq \frac{p(p-1)}{2}\ \varepsilon \|u\|^2 = \frac{|\lambda|\ \|u\|^2}{2}$$

Et puisque λ est négatif, nous en déduisons $<Au,u> \geq \frac{\lambda}{2} \|u\|^2$.

Cas où A a toutes ses valeurs propres réelles: Soit E_i les sous-espaces spectraux de A. Nous construirons $(\forall i)$ un produit scalaire sur E_i (noté $< , >_i$), et un scalaire k_i tels que $(\forall x\in E_i)$ $<A(u),u>_i \leq k_i <u,u>_i$ (comme ci-dessus). Il suffira alors de mettre sur E le seul produit scalaire qui rend les E_i deux à deux orthogonaux, et induit sur chacun d'eux le produit $<.,.>_i$.

Exercice f: Montrer qu'il existe un unique produit scalaire $<.,.>$ sur E, pour lequel les E_i sont deux à deux orthogonaux, et qui induit sur chacun d'eux le produit $<.,.>_i$.

Montrer que, pour ce produit scalaire, on a $(\forall u)$ $<Au,u> \leq \sup_i k_i <u,u>$.

Cas où certaines valeurs propres ne sont pas réelles: Considérons $\mathbb{R}^p$ comme un sous-ensemble de $\mathbb{C}^p$. Soit M la matrice de A dans la base naturelle de $\mathbb{R}^p$. Soit $A_{\mathbb{C}}$ l'endomorphisme de $\mathbb{C}^p$ dont la matrice dans la base naturelle (de $\mathbb{C}^p$) est M; alors, pour tout u dans $\mathbb{R}^p$, on a $A(u) = A_{\mathbb{C}}(u)$. Nous pouvons construire (comme ci-dessus) un produit hermitien dans $\mathbb{C}^p$ tel que $(\forall u \in \mathbb{C}^p)$ $\mathcal{R}(A_{\mathbb{C}}(u)|u) \leq k\ (u|u)$. La restriction à $\mathbb{R}^p$ de ce produit hermitien est un produit scalaire; et $(\forall u \in \mathbb{R}^p)$ $<A(u),u> \leq k <u,u>$.

Exercices sur le chapitre 20

* **Exercice 1:** Trouver les solutions des équations différentielles suivantes :

a) $y''+3y'+2y = e^{-t}$ b) $y''-3y'-2y = \text{ch}t \cos t$ c) $y''+2y'+1 = \text{ch}t$

* **Exercice 2:** Trouver les solutions des équations différentielles suivantes :

a) $y' = \dfrac{xy}{x^2+y^2}$ b) $y' = \dfrac{1+y^2}{1+x^2}$ c) $y' = \dfrac{1+x^2}{1+y^2}$

d) $y'\sqrt{x^2+y^2} = y$ e) $y'(x^2+y^2) = x^2-3y^2$ f) $4y^2y'+(x-y)^2 = 0$

* **Exercice 3:** Trouver les solutions des équations différentielles suivantes :

a) $(x+yy')^2 = x^2+y^2$ b) $y'-y = x\sqrt{y}$ c) $y'' = yy'$

d) $y'' = (y')^2y$ (Le calcul se termine par une quadrature que l'on ne peut effectuer)

e) $y-xy' = x^2+y^2$ f) $y'-y \text{ tg } x + \cos^2 x = 0$ g) $y'\text{sh } 2x - 2y = 2 \text{ sh}^3 x$

* **Exercice 4:** Calculer $\exp tM$, où $M = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ (diagonaliser M).

* **Exercice 5:** Calculer $\exp tM$, où $M = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ (écrire M comme la somme d'une matrice diagonalisable, et d'une matrice nilpotente qui commutent entre elles).

** **Exercice 6:** On considère les fonctions α, β et γ définies par les séries entières suivantes

$$\alpha(t) = \sum_{p\geq 0} \frac{t^{3p}}{(3p)!} \qquad \beta(t) = \sum_{p\geq 0} \frac{t^{3p+1}}{(3p+1)!} \qquad \gamma(t) = \sum_{p\geq 0} \frac{t^{3p+2}}{(3p+2)!}$$

a) Montrer que les rayons de convergence de ces séries sont infinis. Calculer les sommes $\alpha+\beta+\gamma$, $\alpha+j\beta+j^2\gamma$ et $\alpha+j^2\beta+j\gamma$. En déduire une expressions des fonctions α,β et γ en fonction des fonctions exponentielle, cosinus et sinus.

b) Calculer $\exp tM$, où $M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ (calculer d'abord M^3).

* **Exercice 7:**

a) Soit A une symétrie (i.e: un endomorphisme tel que $A^2 = \text{Id}$). Montrer que $\exp tA = (\text{ch } t)\,\text{Id} + (\text{sh } t)\,A$.

b) Calculer $\exp tM$ où $M = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}$. On cherchera d'abord λ et μ tels que $(\lambda M + \mu \text{Id})^2 = \text{Id}$.

* **Exercice 8:**

Trouver la solution du système différentiel $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} e^{2t} \\ 0 \\ 0 \end{pmatrix}$ qui prend la valeur $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ pour $t = 0$.

* **Exercice 9:** Trouver les solutions du système différentiel $\begin{cases} x' = z+e^t \\ y' = y+z+\cos t \\ z' = x+\sin t \end{cases}$

** **Exercice 10:** Résoudre le système différentiel $\begin{cases} x' = x+z+u+e^t \\ y' = y+z+u \\ z' = x+y+u+e^t \\ u' = x+y+z \end{cases}$

** **Exercice 11:** Trouver la solution du système différentiel $\begin{pmatrix} x' \\ y' \\ z' \\ u' \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ u \end{pmatrix} + \begin{pmatrix} e^t \\ t \\ e^t \\ t \end{pmatrix}$

qui prend la valeur $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ pour $t = 0$.

* **Exercice 12:** On considère l'équation différentielle $y' = 1+y^2$.

a) Vérifier que $t \to \operatorname{tg} t$ en est une solution maximale.

b) Utiliser le fait que la variable t n'apparait pas dans l'équation pour en déduire les autres solutions maximales.

* **Exercice 13:** Soit l'équation différentielle $y'' - 2y' + y = e^t$.

a) Montrer qu'elle est équivalente à l'équation vectorielle $\begin{cases} u' = 2u - v + e^t \\ v' = u \end{cases}$

b) Résoudre la première équation directement.

c) Résoudre la seconde équation au moyen des exponentielles de matrices. Comparer les résultats obtenus.

* **Exercice 14:** On considère l'équation différentielle $\begin{cases} y' = 2y + z \\ z' = y - 2z \end{cases}$

a) Par un changement de coordonnées dans le plan, l'écrire sous la forme $\begin{cases} Y' = \lambda Y \\ Z' = \mu Z \end{cases}$

b) Dessiner les courbes solutions. Comment faut-il choisir $M_0(y_0, z_0)$ pour que la solution qui passe par M_0 pour $t = 0$ tende vers l'infini en $+\infty$.

* **Exercice 15:** On considère l'équation différentielle $\begin{cases} y' = -2y + z \\ z' = 3y - 2z \end{cases}$

a) Par un changement de coordonnées dans le plan, l'écrire sous la forme $\begin{cases} Y' = \lambda Y \\ Z' = \mu Z \end{cases}$

b) Dessiner les courbes solutions. Comment faut-il choisir $M_0(y_0, z_0)$ pour que la solution qui passe par M_0 pour $t = 0$ tende vers l'origine en $+\infty$.

** **Exercice 16:** Etude de l'équation différentielle $y'^2 = 1 - y^2$.

a) On considère l'équation $y' = \sqrt{1-y^2}$ $(|y| < 1)$. Quelle est sa solution maximale correspondant à la donnée initiale $(t_0 = 0, y_0 = 0)$. Utiliser le fait que la variable t n'apparait pas dans l'équation, pour déterminer toutes les solutions maximales de cette équation.

b) Quelles sont les solutions maximales de l'équation $y'^2 = 1 - y^2$ qui prennent la valeur $\frac{1}{2}$ en $\frac{\sqrt{3}}{2}$?

** **Exercice 17:** On considère l'équation différentielle $4y + y'^2 = 4xy'$ (on essayera de la mettre sous la forme $y' = F(x,y)$). Montrer que ses solutions sont d'une part la fonction $x \to x^2$, d'autre part les fonctions affines dont les graphiques sont les tangentes à la parabole $y = x^2$.

** **Exercice 18:** Soit M une matrice telle que $M^3 + 3M^2 + M = \mathrm{Id}$.

a) Montrer que $(M+\mathrm{Id})^3 = 2(M+\mathrm{Id})$.

b) Calculer $\exp t(M+I)$ puis $\exp tM$.

* **Exercice 19:** Résoudre l'équation différentielle $y''' + 3y'' + 3y' + y = \cos t$ (où y est une fonction inconnue à valeurs réelles). On pourra la ramener à une équation d'ordre 1 à valeurs vectorielles.

Outils fondamentaux du raisonnement
Cinquième partie: Relations d'équivalence et ensembles quotient.

§ 1 : Partitions et relations d'équivalence.

Une partition $(A_i)_{i\in I}$ d'un ensemble E est une famille de sous-ensembles, tous non vides, de E deux à deux disjoints et dont la réunion est E. Chaque élément de E appartient donc à un A_i et à un seul.

Une partition peut être finie (par exemple l'ensemble des nombres pairs, et l'ensemble des nombres impairs forment une partition finie de $\mathbb{Z}$). Elle peut être infinie (par exemple les intervalles de la forme $[n,n+1[$ (où $n\in\mathbb{Z}$) forment une partition infinie de $\mathbb{R}$).

La partition de E associée à une application φ:E $\to$ F.

Soit φ:E $\to$ F, considérons les sous-ensembles de la forme $\varphi^{-1}(f)$, où f est un élément de F. Si φ n'est pas surjective, certains de ces sous-ensembles sont vides; nous les éliminons. Il reste les sous-ensembles $\varphi^{-1}(f)$ pour tous les f dans Im φ; ils forment une partition de E. En effet

a) Ils sont deux à deux disjoints (pour $f\neq f'$, $\varphi^{-1}(f)$ et $\varphi^{-1}(f')$ sont disjoints puisqu'un élément de E ne peut avoir pour image à la fois f et f').

b) Leur réunion est E (l'élément e de E appartient à $\varphi^{-1}(\varphi(e))$).

Exemples:

a) Soit $\varphi:\mathbb{Z}\to\{0,1\}$ l'application qui à tout entier associe le reste de sa division par 2. La partition associée est formée de $\varphi^{-1}(0)$ (c'est à dire de l'ensemble des nombres pairs), et de $\varphi^{-1}(1)$ (c'est à dire de l'ensemble des nombres impairs).

b) Soit π la projection du plan P sur sa droite D parallèlement à la direction Δ. La partition associée à π est formée des droites de direction Δ.

c) Si φ:E $\to$ F est injective, la partition associée à φ est formée de tous les sous-ensembles à un élément de E.

§ 2 : Partition associée à une relation d'équivalence

Une relation binaire(*) (notée par exemple $\approx$) sur l'ensemble E est appelée une relation d'équivalence, si elle vérifie les trois propriétés suivantes:

α) $(\forall x\in E)$ $x\approx x$.

β) $(\forall x\in E)$ $(\forall y\in E)$ $x\approx y$ implique $y\approx x$.

γ) $(\forall x\in E)$ $(\forall y\in E)$ $(\forall z\in E)$ $(x\approx y$ et $y\approx z)$ implique $x\approx z$.

Soit $(A_i)_{i\in I}$ une partition de E. La relation "$x\approx y \Longleftrightarrow$ x et y appartiennent au même A_i", est une relation d'équivalence (c'est clair). Inversement pour toute relation d'équivalence $\approx$ sur E, nous pouvons faire la construction suivante:

A tout élément x de E nous associons sa classe $\mathscr{C}_x$, c'est le sous-ensemble de E formé des éléments de E qui sont équivalents à x. Si $x\approx y$, la classe de x et la classe de y coïncident. En effet d'après

(*) c'est à dire portant sur les couples d'éléments de E.

γ, tout z équivalent à y, est aussi équivalent à x (donc $\mathscr{C}_y \subset \mathscr{C}_x$); et réciproquement, d'après γ, tout u équivalent à x est aussi équivalent à y (donc $\mathscr{C}_x \subset \mathscr{C}_y$). Si la classe de x et celle de y ont un point commun z, on a $z \approx x$ et $z \approx y$ (c'est à dire $y \approx z$), donc $x \approx y$; et ainsi elles sont confondues. *Ainsi les classes sont des sous-ensembles non vides; deux classes distinctes sont disjointes; et la réunion de toutes les classes est E (tout élément x de E appartient à une classe). On a bien défini une partition de E.*

Remarque: Soit X l'ensemble des classes d'une relation d'équivalence sur E (X est donc un sous-ensemble de $\mathscr{P}(E)$). En associant à tout e dans E sa classe, nous obtenons une application $\chi: E \to X$ qui est surjective.

Exemples: Sur $\mathbb{Z}$ définissons la relation

$$x \approx y \iff x - y \text{ est pair}$$

C'est une relation d'équivalence. Elle a deux classes, l'une est formée des nombres pairs, l'autre des nombres impairs.

Soit P un plan; choisissons un vecteur $\vec{V}$ parallèle à P. Nous pouvons définir une relation par

$$A \approx B \iff (\exists\lambda) \text{ tel que } \overrightarrow{AB} = \lambda \vec{V}$$

C'est une relation d'équivalence; ses classes sont les droites parallèles à $\vec{V}$.

Exercices sur le chapitre 20b.

* **Exercice 1**: On considère l'ensemble E des bases orthonormées du plan P. Sur E on définit une relation par:

$$(\vec{i}, \vec{j}) \approx (\vec{u}, \vec{v}) \iff \text{Dét}_{(\vec{i}, \vec{j})}(\vec{u}, \vec{v}) \text{ est positif}$$

Montrer que l'on a défini une relation d'équivalence, et que cette relation a deux classes. Comment caractériser chacune de ces deux classes.

* **Exercice 2**: Soit E un espace vectoriel, et soit $\mathscr{N}$ l'ensemble de toutes les normes que l'on peut définir sur E. On définit une relation sur $\mathscr{N}$ par

$$N \approx N' \iff (\exists \text{ a et b } (\neq 0)) \text{ tels que } (\forall e \in E) \;\; a\,N(e) \le N'(e) \le b\,N(e).$$

Montrer que l'on a ainsi défini une relation d'équivalence sur $\mathscr{N}$. A quelle condition cette relation a-t-elle une seule classe ? Connaissez vous un E pour lequel elle a plusieurs classes ?

** **Exercice 3**: Sur l'ensemble E de toutes les matrices n×n, on définit la relation

$$M \approx N \iff (\exists A \text{ et } B \text{ inversibles}) \text{ telle que } AM = NB.$$

Montrer que l'on a ainsi défini une relation d'équivalence sur E. Montrer que deux matrices sont dans la même classe si et seulement si elles ont même rang.

** **Exercice 4**: Sur $\mathscr{C}^0([0,1],\mathbb{R})$ on définit la relation

$$f \approx g \iff f \underset{1/2}{\sim} g$$

a) Montrer que c'est une relation d'équivalence. Montrer que cette relation a une infinité de classes.

b) Montrer que $f \approx g$ et $f' \approx g'$ implique $f \times f' \approx g \times g'$. Soit α et β deux classes, choisissons f dans α et g dans β, on pose $\alpha\times\beta$ = "la classe de $f \times g$". Montrer que la notation $\alpha\times\beta$ est cohérente, c'est à dire que $\alpha\times\beta$ ne dépend que de α et β.

c) Montrer que $(\forall \alpha, \beta \text{ et } \gamma)\ (\alpha\times\beta)\times\gamma = \alpha\times(\beta\times\gamma)$.

** **Exercice 5:** Sur l'ensemble E des suites numériques on définit la relation

$$(u_n)_{n\in\mathbb{N}} \approx (v_n)_{n\in\mathbb{N}} \iff \begin{cases} (u_n)_{n\in\mathbb{N}} = O(v_n)_{n\in\mathbb{N}} \\ \text{et } (v_n)_{n\in\mathbb{N}} = O(u_n)_{n\in\mathbb{N}} \end{cases}$$

a) Montrer que c'est une relation d'équivalence.

b) Soit α et β deux classes, choisissons $(u_n)_{n\in\mathbb{N}}$ dans α et $(v_n)_{n\in\mathbb{N}}$ dans β, on pose $\alpha\times\beta$ = "la classe de $(u_n\times v_n)_{n\in\mathbb{N}}$". Montrer que la notation $\alpha\times\beta$ est cohérente, c'est à dire que $\alpha\times\beta$ ne dépend que de α et β.

c) Montrer que $(\forall\alpha,\beta \text{ et } \gamma)\ (\alpha\times\beta)\times\gamma = \alpha\times(\beta\times\gamma)$.

* **Exercice 6:** Sur l'ensemble E des suites numériques on définit la relation

$$(u_n)_{n\in\mathbb{N}} \approx (v_n)_{n\in\mathbb{N}} \iff \begin{cases} (u_n)_{n\in\mathbb{N}} = o(v_n)_{n\in\mathbb{N}} \\ \text{et } (v_n)_{n\in\mathbb{N}} = o(u_n)_{n\in\mathbb{N}} \end{cases}$$

Montrer que ce n'est pas une relation d'équivalence.

** **Exercice 7:**

a) Sur l'ensemble $\mathcal{N}_2$ des matrices nilpotentes 2×2, on définit la relation

$$M \approx N \iff (\exists A \text{ inversible}) \text{ telle que } AM = NA.$$

Montrer que c'est une relation d'équivalence. Montrer qu'elle a deux classes dont l'une a un seul élément.

b) Sur l'ensemble $\mathcal{N}_3$ des matrices nilpotentes 3×3, on définit la relation

$$M \approx N \iff (\exists A \text{ inversible}) \text{ telle que } AM = NA.$$

Montrer que c'est une relation d'équivalence. Combien a-t-elle de classes ?

* **Exercice 8:** Soit un ensemble A muni de deux relations d'équivalences notées $\approx_1$ et $\approx_2$. On suppose que

$$(\forall x \text{ et } y)\ x \approx_1 y \ \text{ implique } x \approx_2 y.$$

On note U l'ensemble des classes de la realtion $\approx_1$. On note V l'ensemble des classes de la relation $\approx_2$. Notons χ_1: A → U (resp. χ_2: A → V) l'application qui a x associe sa classe pour $\approx_1$ (resp. sa classe pour $\approx_2$) Montrer qu'il existe une unique application φ: U → V telle que $(\forall x\in A)\ \chi_2(x) = \varphi\circ\chi_1(x)$.

* **Exercice 9:** Soit {a,b,c,d} un ensemble à quatre éléments. Faire la liste de toutes les relations d'équivalence que l'on peut y définir.

** **Exercice 10:** Soit A un ensemble qui a 30 éléments. On considère toutes les relations d'équivalence sur A qui ont 3 classes. Combien en existe-t-il ?

** **Exercice 11:** Soit A un ensemble qui a 30 éléments. On considère toutes les relations d'équivalence sur A dont toutes les classes ont 10 éléments. Combien en existe-t-il ?

Table des matières

Second fascicule.