

Un peu d'histoire de quelques idées en  
mathématiques  
Version pré-pré-préliminaire, partielle,  
sans garantie...

R. Langevin

February 11, 2011



# Contents

<b>1</b>	<b>La notion de dimension</b>	<b>7</b>
1.1	Aristote, Didon...	7
1.1.1	Oresme, Fermat, Descartes	8
1.1.2	Monge	8
1.2	Grassmann, Riemann, Abbot, Peano	8
1.2.1	Cantor	9
1.2.2	Peano	9
1.3	Dimension topologique	10
1.3.1	Dimensions fractales	11
1.4	dimension “infinie”	13
<b>2</b>	<b>La notion de fonction</b>	<b>15</b>
2.1	L’antiquité	15
2.2	Autres	15
2.3	Le Moyen-âge et la renaissance	16
2.4	Le 17 <sup>ème</sup> siècle	17
<b>3</b>	<b>Calcul différentiel et intégral</b>	<b>21</b>
3.1	Archimède et les mathématiciens de l’antiquité	21
3.1.1	Le cercle	21
3.1.2	L’aire comprise entre une parabole et une corde	22
3.2	Stevin (1548-1620)	24
3.3	Kepler (1571-1630)	25
3.4	Galilée	25
3.4.1	Cavalieri	27
3.5	De Fermat à Barrow	28
3.6	Isaac Newton	31
3.6.1	Gottfried Wilhelm Leibniz	36
3.6.2	Après Newton et Leibniz	38
3.7	Au XIX <sup>ème</sup> siècle	39

3.7.1	L'intégrale de Lebesgue . . . . .	41
3.7.2	Une extension non commutative: le calcul de Fox . . . . .	42
<b>4</b>	<b>Probabilités et statistiques</b>	<b>45</b>
4.1	Jeux, notion de "chances" et "partis" . . . . .	45
4.2	Début des probabilités . . . . .	48
4.2.1	Le paradoxe de Saint Petersburg . . . . .	50
4.2.2	Un précurseur . . . . .	50
4.3	Naissance des statistiques . . . . .	51
4.4	Analyse et probabilités . . . . .	53
4.5	Statistiques . . . . .	56
4.6	Probabilité au 20 <sup>ème</sup> siècle . . . . .	63
4.6.1	Le mouvement Brownien et autres processus . . . . .	63
4.7	Entropie . . . . .	69
<b>5</b>	<b>La géométrie intégrale de Buffon à Poincaré</b>	<b>71</b>
5.1	Buffon . . . . .	71
5.1.1	The birth of the notion of geometric measure . . . . .	73
<b>6</b>	<b>La courbure des courbes planes et gauches</b>	<b>79</b>
<b>7</b>	<b>La courbure (extrinsèque et intrinsèque) des surfaces.</b>	<b>83</b>
7.1	L'application de Gauss et les courbures principales. . . . .	83
7.2	Euler . . . . .	88
7.2.1	Monge . . . . .	89
7.3	Le point de vue intrinsèque . . . . .	90
7.3.1	Courbure de Gauss des surfaces . . . . .	92
7.3.2	Géodésiques et courbure géodésique . . . . .	95
7.4	Modèles de surfaces de courbure négative . . . . .	95
7.4.1	Les modèles de Poincaré . . . . .	95
<b>8</b>	<b>Le cinquième postulat d'Euclide</b>	<b>99</b>
8.1	Le cinquième postulat . . . . .	99
8.2	Tentatives de démonstration . . . . .	99
8.3	Les fondateurs de la géométrie non euclidienne. . . . .	103
<b>9</b>	<b>Appendice</b>	<b>105</b>
9.1	Le mot "vecteur" (citation) . . . . .	105
9.2	Grassmanniennes . . . . .	109
9.3	Enveloppes . . . . .	109

9.4 Courbure de Gauss des surfaces polyédriques . . . . . 113



# Chapter 1

## La notion de dimension

### 1.1 Aristote, Didon...

Aristote dans “de caelo” (du ciel), discute la notion de dimension, en remarquant qu’il faut un point pour disconnecter un arc, [Ari] *On entend par continu tout ce qui peut se diviser en parties toujours divisibles; le corps est ce qui est divisible en tous sens. C’est que, parmi les grandeurs, l’une n’est divisible qu’en sens unique, c’est la ligne; l’autre l’est en deux, c’est la surface; l’autre l’est en trois, c’est le corps. Il n’y a pas de grandeurs autres que celles-là, parceque trois est tout et que trois renferme toutes les dimensions possibles. En effet, ainsi que le disent les Pythagoriciens, l’univers entier et toutes les choses dont il est composé sont déterminées par ce nombre Trois.[...]mais ce qui doit être évident dès à présent, c’est qu’il n’y a pas pour le corps de passage possible à un autre genre différent, comme par exemple, on passe de la longueur à la surface, ou de la surface au corps. Le corps, s’il était dans cette condition, ne serait plus une grandeur complète; car cette transition à un autre genre ne peut nécessairement avoir lieu que par suite d’un certain défaut; or il n’est pas possible que ce qui est complet soit défectueux, puisqu’il est tout ce qu’il doit être.*

On retrouve la séquence “point, droite, surface” dans les définitions 1, 2, 3, 5, 6 des éléments d’Euclide.

1. *Un point est ce qui ne se découpe pas*
2. *une courbe est une longueur sans largeur*
3. *Les extrémités d’une courbe sont des points*
5. *Une surface est ce qui a longueur et largeur seulement*
6. *Les extrémités d’une surface sont des courbes*

Dans l’Énéïde [Vi] livre I, vers 367, est mentionnée la légende de Didon qui explicite la différence entre aire et longueur. *Mercatique solum, facti de nomine*

*Byrsam**Taurino quantum possent circumdare tergo:*

Il achetèrent tout le terrain qu'ils pourraient entourer avec une peau de taureau, d'où lui vient le nom de Byrsa (de fait l'étymologie: byrsae: lanières de cuir très minces est controversée, on pense plutôt au mot phénicien borsa: contrefort, bastion).

Suivant la légende, en découpant une peau de taureau en de fines lanières, qui entourèrent une surface, Didon réussit à acheter suffisamment de terre pour édifier sa ville.

Qintillien reprend l'argument distinguant aire et longueur du périmètre, en donnant des exemples de carrés et rectangles de même périmètre et d'aires très différentes.

**1.1.1 Oresme, Fermat, Descartes**

On trouve chez Oresme (14<sup>ème</sup> siècle) une nomenclature des fonctions définies sur un segment: fonctions rectangle (constantes), uniformément difformes (affines) et enfin difformément difformes (le reste) voir chapitre 2. Il classe aussi les fonctions définies sur une portion du plan et mentionne enfin celles définies sur un solide. Cependant, effrayé par l'autorité d'Aristote, il se refuse à envisager l'objet de dimension 4 correspondant.

Fermat et Descartes, sans doute indépendamment, utilisent les coordonnées d'un point du plan ou de l'espace pour traiter des problèmes de géométrie.

**1.1.2 Monge**

Le livre de géométrie descriptive [Mo1] de Gaspard Monge commence par *considérations d'après lesquelles on détermine la position d'un point situé dans l'espace*. Il se donne la peine de montrer que parmi les moyens de repérer un point de l'espace à partir de distances à des objets fixes, les coordonnées cartésiennes vues comme distances (avec un signe) à un système de trois plans orthogonaux, est souvent le plus commode.

**1.2 Grassmann, Riemann, Abbot, Peano**

La première apparition explicite de la dimension 4 se trouve chez G. Grassmann [Gr]. En fait, non seulement Grassmann envisage des espaces de dimension quelconque, mais il fonde le calcul extérieur, définissant des aires, volumes, p-volumes orientés dans  $\mathbb{R}^n$ . Le mot "vecteur" est apparu en astronomie 1704: J.

Harris Lexicon Technicum I. s.v., *A Line supposed to be drawn from any Planet moving round a Center, or the Focus of an Ellipsis, to that Center or Focus, is by some Writers of the New Astronomy, called the Vector; because 'tis that Line by which the Planet seems to be carried round its Center. En Français on parle de "rayon vecteur", voir par exemple Laplace: Traité de mécanique céleste (1799-1825).*

Hamilton (1805-1865) distinguera vecteurs et scalaires lorsqu'il définit et étudie les quaternions.

(Extrait de *Earliest Known Uses of Some of the Words of Mathematics (V)*  
*Last revision: Sept. 10, 2007 (AOL)*)

Parler de Cayley (vecteurs, 1845?)

Il faudra pourtant, je crois, attendre l'article de Peano (1888) [Pea1] pour trouver une définition explicite de la structure d'espace vectoriel de dimension  $n$ , à l'occasion de l'étude de systèmes de  $n$  équations différentielles linéaires du premier ordre.

Le texte d'Abbot "Flatland" est une jolie illustration de l'existence d'espaces de dimension (entière) quelconque [Ab].

### 1.2.1 Cantor

En remarquant qu'un segment et un carré ont même cardinal (c'est à dire qu'il existe une bijection entre un segment et un carré, Cantor sème le trouble.

La démonstration de Cantor repose sur le codage de points du segment ou du carré en base 2 ou 3.

### 1.2.2 Peano

Par ailleurs, en 1890, G. Peano [Pea2], construit une courbe qui remplit un carré. L'idée de la démonstration est de construire une suite de fonctions  $f_n$  du segment dans le carré d'images une suite de courbes polygonales  $C_n$  qui sont "de plus en plus denses"; la suite converge uniformément vers une fonction  $f$  qui sera continue et surjective. La figure 1.1 montre comment passer de  $f_0$  à  $f_1$ ; il suffit de répéter le procédé sur chaque carré d'une subdivision de taille  $\frac{1}{3^n}$  pour construire la suite  $f_k$ .

L'article de Peano [Pea2] écrit directement des formules utilisant des coordonnées en base 3.

Un an plus tard, D. Hilbert propose une construction analogue pour laquelle les courbes successives  $C_n$  sont plongées. L'article de Hilbert [Hi] est illustré d'une figure.

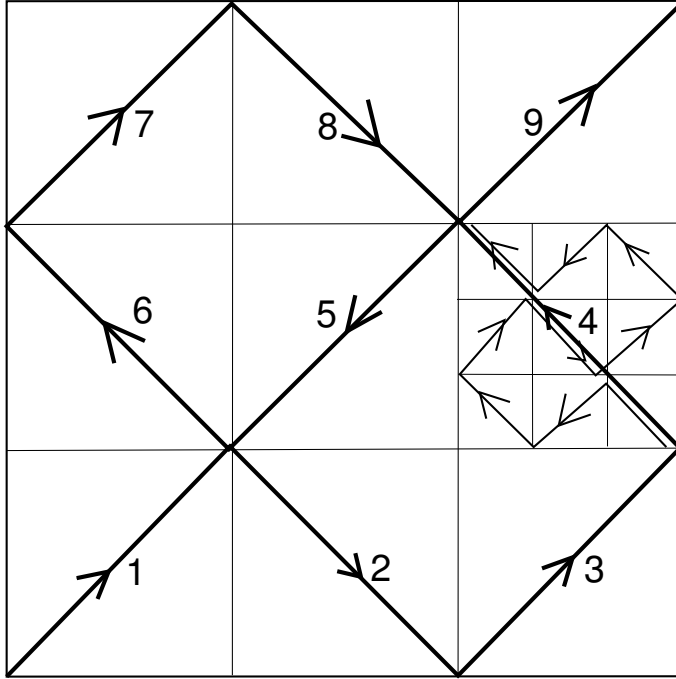


Figure 1.1: Construction de la courbe de Peano

### 1.3 Dimension topologique

Le fait que  $I^n$  et  $I^m$  ne soient pas homéomorphes quand  $m \neq n$  a été démontré en 1911 Par Brouwer [Bro] en utilisant le fait qu'il est impossible, pour  $\epsilon$  assez petit, de trouver une application continue  $f$  du cube  $I^n$  sur un polyèdre  $P$  de dimension strictement inférieure à  $n$  de sorte que pour tout  $x \in P$  les "fibres"  $f^{-1}(x)$  aient un diamètre inférieur à  $\epsilon$ . Lebesgue, la même année, arrive au même résultat en remarquant que  $I^n$  peut être recouvert,  $\forall \epsilon > 0$ , par une famille finie de fermés de diamètre plus petit que  $\epsilon$ , de sorte que toutes les intersection de  $n + 2$  membre de la famille soient vides, mais ne peut pas être recouvert par une famille finie d'ouverts de diamètre plus petit que  $\epsilon$  de sorte que toutes les intersections de  $n + 1$  membres de la famille soient vides.

La preuve de Lebesgue contient un trou qui sera comblé par Brouwer en 1913 et par Lebesgue en 1921.

Urysohn en 1922 et Menger en 1923 définissent de façon beaucoup plus générale des dimensions topologiques.

Bornons nous ici à supposer que  $X$  est un espace métrique compact.

**Définition 1.3.1.** petite dimension inductive

La petite dimension inductive  $ind$  ou dimension de Menger-Urysohn est un entier  $-1 < ind(X) < \infty$  défini par

1.  $ind(X) = -1$  si et seulement si  $X = \emptyset$ .
2.  $ind(X) \leq n$ , (pour  $n \geq 0$ ), si pour tout point  $x \in X$  et tout voisinage  $V \subset X$  du point  $x$ , il existe un ouvert  $U \subset V \subset X$  tel que  $x \in U$  et  $ind(\partial U) \leq n - 1$ .
3.  $ind(X) = n$  si  $ind(X) \leq n$  et  $ind(X) > n - 1$ , c'est à dire si l'inégalité  $ind(X) \leq n - 1$  n'est pas vraie.
4.  $ind(X) = \infty$  si  $ind(X) > n$  pour  $n = -1, 0, 1, \dots$

*Remarque.* L'ensemble triadique de Cantor est de dimension de Menger-Urysohn 0. En effet on peut choisir comme ouvert  $U$  un des intervalles ouvert-fermés de la construction, dont le bord est vide.

Maintenant on peut aussi démontrer que les cubes  $I^n$  de dimension différentes ne sont pas homéomorphes en utilisant le fait que l'homologie du complémentaire d'un point intérieur des cubes n'est pas la même.

### 1.3.1 Dimensions fractales

Le nom "fractal" a été inventé, sans doute vers 1975, par B. Mandelbrot. Cependant des exemples d'objets fractals étaient connus bien plus tôt.

Von Koch en 1906 a construit une courbe "en flocon de neige" (figure 1.2):

Cette courbe est la limite uniforme d'une suite de courbes  $C_n$  définies par récurrence: pour obtenir  $C_{n+1}$ , il suffit de remplacer chaque segment de  $C_n$  par une ligne polygonale à 4 cotés, comme indiqué à droite de la figure 1.2.

On en déduit immédiatement l'égalité:  $longueur(C_{n+1}) = \frac{4}{3} \cdot longueur(C_n)$ . La courbe limite n'est donc pas rectifiable.

Par ailleurs l'ensemble triadique de Cantor est, certes, de dimension de Menger-Urysohn 0, mais contient beaucoup de points<sup>1</sup> c'est à dire, au sens ensembliste "autant" qu'un segment.

En cherchant à définir des mesures adaptées à la géométrie d'un ensemble, Hausdorff (1868-1942) a permis la définition de dimensions à valeurs réelles, dont l'une est appelée *dimension de Hausdorff*.

#### Dimension de Hausdorff

---

<sup>1</sup>On peut même construire un ensemble de Cantor contenu dans le segment de mesure de Lebesgue positive.

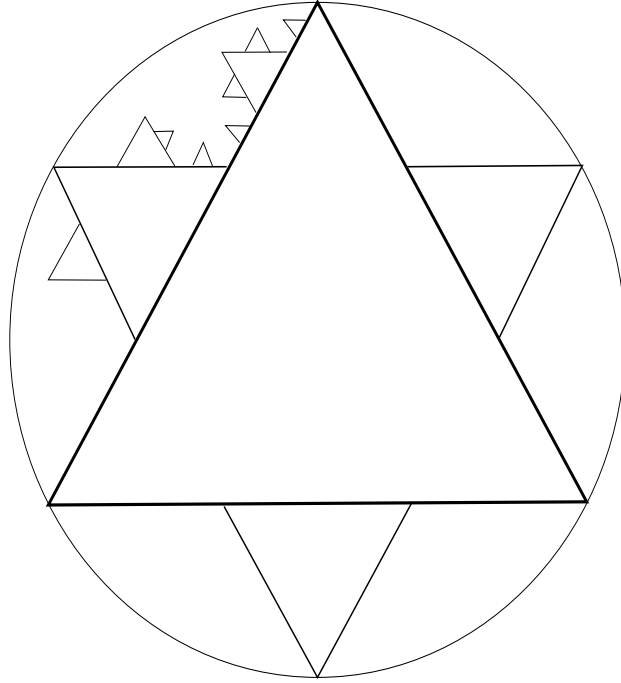


Figure 1.2: La courbe de Von Koch

Pour cela supposons que  $X$  est un espace métrique compact. Soit  $\mathcal{U} = \bigcap U_i, i \in I$  un recouvrement ouvert de  $X$ . Calculons:

$$m(\mathcal{U}, \alpha) = \sum_{i \in I} \text{diam}(U_i)^\alpha$$

$$m(X, \alpha, \epsilon) = \inf(m(\mathcal{U}, \alpha) \text{ où } \text{diam}(U_i) < \epsilon; U_i \in \mathcal{U})$$

La limite supérieure quand  $\epsilon$  tend vers 0 de  $m(X, \alpha, \epsilon)$  peut être nulle ou infinie. De fait, c'est le cas pour toute valeur de  $\alpha$  sauf au plus une.

Dans tous les cas les ensembles  $\{\alpha \text{ tels que } \overline{\text{lim}} = 0\}$  et  $\{\alpha \text{ tels que } \overline{\text{lim}} = \infty\}$  forment une coupure. Le réel  $d_H(X)$  défini par cette coupure est par définition la *dimension de Hausdorff* de  $X$ .

**Dimension capacitaire** (Dimension de Bouligand-Minkowski lorsque le compact est plongé dans  $\mathbb{R}^N$ ).

On va ici utiliser des boules toutes de même rayon  $\epsilon$ . Définissons deux nombres:

- $A(X, \epsilon, n = \inf(\#\{B_i\})$  où  $\{B_i\}$  est un recouvrement de  $X$  par des boules (centrées en des points de  $X$ ) de rayon  $\epsilon$
- $B(X, \epsilon, n = \sup(\#\{b_i\})$  où  $\{b_i\}$  est un ensemble de boules (centrées en des points de  $X$ ) de rayon  $\epsilon$  disjointes.

Considérons maintenant les deux limites:

$$\lim_{\epsilon \rightarrow 0} \frac{-1}{\log \epsilon} \log(A(X, \epsilon))$$

$$\lim_{\epsilon \rightarrow 0} \frac{-1}{\log \epsilon} \log(B(X, \epsilon))$$

**Proposition 1.3.2.** *Les deux limites précédentes existent et sont égales. Leur valeur commune  $d_{B-M}(X)$  est la dimension de Bouligand-Minkowski.*

Exemples.

La réunion des points de la suite  $u_n = \frac{1}{n}$  est de dimension de Bouligand-Minkowski  $\frac{1}{2}$ .

La réunion des points de la suite  $v_n = \frac{1}{2^n}$  est de dimension de Bouligand-Minkowski 0.

**Dimension de Tricot.** [Tri]

$X$  est comme plus haut un espace compact. Un packing de  $X$  est une collection de boules centrées sur  $X$  et deux à deux disjointes.

Considérons,  $\alpha$  et  $\epsilon$  étant deux nombres positifs, la quantité:

$$\Lambda(X, \alpha, \epsilon) = \sup \left\{ \sum_j (\text{diam } B_j)^\alpha \right\}$$

$\{B_j\}$  étant le packing de  $X$  par des boules.

La limite supérieure, quand  $\epsilon$  tend vers 0 de  $\Lambda(X, \alpha, \epsilon)$  peut être nulle ou infinie. C’est encore le cas pour toute valeur de  $\alpha$  sauf au plus une.

Dans tous les cas les ensembles  $\{\alpha \text{ tels que } \bar{\lim} = 0\}$  et  $\{\alpha \text{ tels que } \bar{\lim} = \infty\}$  forment une coupure. Le réel  $d_T(X)$  défini par cette coupure est par définition la *dimension de Tricot* de  $X$ .

On démontre que les trois dimensions que nous venons de présenter satisfont l’inégalité:

$$d_H \leq d_{B-M} \leq d_T$$

## vérifier la dernière inégalité

Lorsque l’ensemble  $X$  est assez régulier (auto-semblable comme l’ensemble de Cantor triadique ou le flocon de Von Koch par exemple, ou comme certains ensembles venant des systèmes dynamiques, on a:

$$d_H(X) = d_{B-M}(X) = d_T(X).$$

## 1.4 dimension “infinie”

L’analyse fonctionnelle (voir chapitre 2) amène à reconnaître la structure d’espace vectoriel sur des espaces de fonctions.

On peut cependant fabriquer “à la main” des espaces de dimension infinie, comme par exemple le cube de Hilbert.

Dans  $\mathbb{R}^{\mathbb{N}}$  on considère les suites  $(a_i)$  telles que  $\forall i, 0 \leq a_i \leq \frac{1}{2^i}$ . C’est un espace métrisable.

On peut aussi considérer dans l’espace de Hilbert  $\ell^2 = \{a_i, \text{ tels que } \sum (a_i)^2 < \infty\}$  le “cube”  $Q$  défini par  $\forall i, |a_i| \leq 1$ . L’application  $(a_1, a_2, \dots, a_n, \cdot) \mapsto (0, a_1, a_2, \dots, a_n, \dots)$  est une isométrie de  $Q$  sur une partie de  $Q$ .

L’application  $(a_i) \mapsto (b_i)$  où  $b_0 = 0$ ,  $b_i = a_{i-1}$  si  $i > 0$  est une isométrie de ce “cube” pour la norme  $|a| = \sqrt{\sum (a_i)^2}$ .

# Chapter 2

## La notion de fonction

### 2.1 L'antiquité

Dès l'antiquité, il existait des tables. Par exemple les babyloniens avaient des tables fournissant les carrés et les racines carrées, les cubes et les racines cubiques. Les astronomes d'Alexandrie connaissaient les longueurs des cordes d'un cercle de circonférence donnée en fonction de l'angle du secteur.

Par ailleurs l'idée de mouvement était familière pour Zénon d'Élée, et plus tard Aristote, qui considérait aussi des changements de forme.

**les trois formes principales dans le processus du monde suivant Aristote:**

- *l'altération ou changement de qualité*
- *le changement de grandeur ou de quantité, par exemple croissance ou diminution*
- *le mouvement local.*

Le paradoxe de Zénon et les controverses autour de ce paradoxe ont peut-être bloqué le développement de la notion de fonction.

### 2.2 Autres

à voir - 5ème siècle Inde Aryabhata, table des sinus

- 10ème siècle Monde Arabe Aboul-Waja nouvelle table des sinus

Une avancée d'un ordre différent, plus "algébrique" se trouve chez Diophante : utiliser des substitutions dans un calcul.

## 2.3 Le Moyen-âge et la renaissance

ORESME au milieu du 14<sup>ème</sup> siècle représentait ce que nous appellerions aujourd'hui la surface comprise entre le graphe de la fonction et l'axe des abscisses. Citons ORESME (1323-1377) :

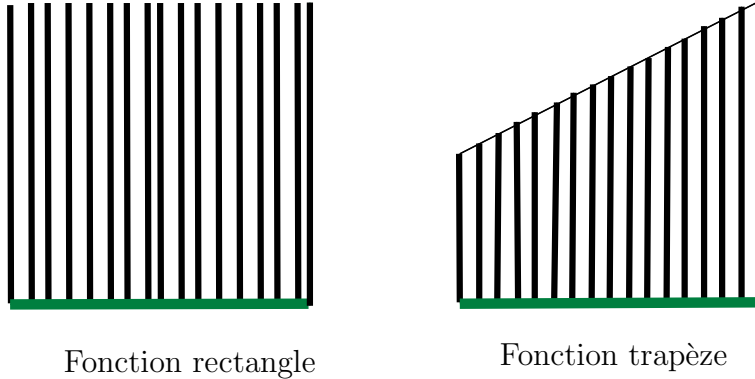


Figure 2.1: Fonctions rectangle et trapèze

- L'altitude d'une surface est estimée à l'aide des segments perpendiculaires à la base comme on le voit sur la figure (2.1).
- Une surface est dite uniformément ou également haute si toutes les segments à l'aide desquels on estime l'altitude sont égaux ; une surface est dite difformément haute si ces segments ne sont pas égaux et [si leurs extrémités] forment un segment un segment, le segment des sommets, qui n'est pas parallèle à la base.
- La hauteur est dite uniformément difforme si trois ou plus hauteurs équidistantes sont en progression arithmétique (Fig.1), c'est-à-dire si le premier est autant au dessus du second que le second au dessus du troisième, ce qui montre que la ligne des sommets est une droite qui n'est pas parallèle à la base.
- La hauteur est dite difformément difforme si les hauteurs ne sont pas disposées comme dans le cas uniformément difforme. Dans ce cas leur ligne de sommets n'est pas droite et la difformité des hauteurs varie avec la variation de cette ligne de sommets.

ORESME considère ensuite des fonctions définie sur des surfaces et même sur des volumes, bien que l'interdiction faite par Aristote de considérer des objets de dimension plus grande que trois lui complique la vie.

Visiblement les “bonnes fonctions” sont constantes ou affines.

GALILÉE (1564-1642) reprenant la “règle de Merton” (du Merton college Oxford, connue en 1535?) utilise une figure comparant l’aire d’un rectangle et d’un triangle de base double de la largeur du rectangle, pour illustrer son mémoire sur la chute des corps. Cela lui permet de reconnaître la progression arithmétique des vitesses en fonction des temps, en d’en déduire la loi satisfaite pas les espaces parcourus. Remarquons déjà ici une interaction entre mathématiques et observation (ou physique). Une manipulation importante des fonctions, pas encore au point dans sa généralité, mais indispensable pour comprendre la relation entre vitesse et distance parcourue, est l’intégration des fonctions, c’est-à-dire le calcul de l’aire de la surface comprise entre la ligne des sommets (aujourd’hui le graphe de la fonction) et l’axe horizontal (voir Fig. 1). On note maintenant l’intégrale avec une sorte de “S” étiré :  $\int$  (notation de Leibniz).

## 2.4 Le 17<sup>ème</sup> siècle

Une autre fonction va émerger de la difformité : le logarithme. La première table de NAPIER(1550-1617) (ou NEPER) a été publiée à Edinbourg en 1614. Pour faire comprendre que la correspondance entre progression géométrique et progression arithmétique peut s’étendre à des échelles continues, il imagine deux points l’un se déplaçant sur un segment, l’autre sur une demi-droite, de sorte que les rapports des distances au point fixe dans le premier cas soient égaux aux distances entre les points dans le second. Afin de bien présenter les tables, la notation décimale (chiffres après la virgule) est vite utilisée.

L’introduction de la notation à l’aide de lettres par VIETE (1540-1603) fournit un outil essentiel.

L’étape suivante est franchie par Pierre de FERMAT (1601-1665) (*Introductions aux lieux plans et solide*, écrit un peu avant 1637) et René DESCARTES (1596-1650) (“*La géométrie*” 1637) : associer l’équation reliant les coordonnées des points d’une courbe à cette courbe.

FERMAT : “*aussitôt que deux quantités inconnues apparaissent dans une ultime égalité, il y a un lieu, et le point terminal de l’une des deux quantités décrit une ligne droite ou courbe*”.

DESCARTES : “*prenant successivement infinies diverses grandeurs pour la ligne y, on en trouvera aussi infinies pour la ligne x, et ainsi on aura une infinité de divers points tels que celui qui est marqué C, par le moyen desquels on décrit la ligne courbe demandée*”. Cela lui permet de distinguer les courbes algébriques (telles que la relation entre x et y soit un polynôme) des autres.

Déjà un mémoire de James GREGORY (Ecosse, 1638-1675), (*“Vera circuli et hyperboles quadrature”*, 1667) considère des séries convergentes et leur composition, ouvrant la voie à la théorie des fonctions analytiques.

Pour lui on appelle quantité  $x$  composée d'autres quantités  $a, b, \dots$  si  $x$  résulte de  $a, b, \dots$  par les quatre opérations élémentaires, l'extraction de racines ou par tout autre opération imaginable.

5. *Quantitatem dicimus a quantitatibus esse compositum; cum a quantitatibus additione, subtractione, multiplicatione, divisione, radicum extractione, vel quacunq[ue] illa imaginabili operatione fit alia quantitas.*

6. *Quando quantitas componitur ex quantitatibus additione, subtractione divisione, radicum extractione: dicimus illam componi analyticè.*

7. *Quando quantitates a quantitatibus inter se commensurabilibus analyticè componi possint, dicimus illas esse inter se analyticas.*

La paternité du terme fonction revient sans doute à LEIBNIZ (1646-1716). Déjà il distingue les fonctions algébriques, qui peuvent être représentées par une équation d'un certain ordre, et transcendentes.

Après BARROW, à Cambridge, Isaac NEWTON (1642-1727) va enrichir encore la compréhension de la classe des “bonnes fonctions” il emploie le terme de “fluxion” à la place du terme moderne “variation de la fonction” et “fluente” à la place du terme “variable” ou “fonction”; la référence à la variable “temps” ou en tous cas à un point de vue dynamique, est implicite. Ce que nous appelons maintenant méthode de Newton est un moyen de construire une suite d'approximation d'une racine d'une équation algébrique ; parallèlement il donne un algorithme permettant d'exprimer à l'aide d'exposants fractionnaires une solution locale d'une équation algébrique à deux variables.

*Data aequatione fluentes quotcunq[ue] quantitates involvente, fluxiones invenire et vice versa. Per fluentes quantitates intelligit indeterminatas id est quae in generatione Curvarum per motum localem perpetuo augentur vel diminuuntur, et per earum fluxionem intelligit celeritatem incrementi vel decrementi.*

La concomitance de la genèse du calcul infinitésimal et de la naissance de l'analyse vue comme études des fonctions définie par une expression analytique est frappante. Jean BERNOULLI (1694-1718) note  $\bar{1}X, \bar{2}X$

*diverses quantités données d'une façon ou d'une autre par une [quantité] indéterminée  $x$  et par des constantes...vel algebraica, vel transcendentè.*

Pour L. EULER (1707-1783), les fonctions sont d'abord définies par des formules *une fonction de quantité variable est une expression analytique composée de quelque manière que ce soit de cette quantité ou de nombres ou de quantités constantes.*

EULER , considère cependant explicitement des fonctions définies comme somme d'une série entière infinie dans son "introduction", et connaissait l'existence d'autres fonctions.

Euler Opera Omnia p17 *Une quantité variable est une quantité indéterminée ou, si l'on veut, une quantité universelle, qui comprend toutes les valeurs déterminées.*

*Ainsi une quantité variable comprend tous les nombres en elle-même, tant positifs que négatifs, les nombres entiers et fractionnaires, ceux qui sont rationnels, transcendants, irrationnels. on ne doit même pas en exclure zéro ni les nombres imaginaires.*

Il appelle alors *continues* les fonctions qui sont partout définies par la même équation. Il explicite aussi la notion de changement de variable :

*"on change la forme des fonction en introduisant une nouvelle variable à la place de la première..."*

Par opposition les fonctions *discontinues* sont:

*toutes les courbes qui ne sont déterminées par aucune équation définie, de sorte que l'on a coutume de tracer par un mouvement libre de la main.*

En 1755, l'abandon de la référence à des "formules" est achevé: dans la préface des "Institutiones calculi differentialis", il écrit:

*Si certaines quantités dépendent d'autres quantités de telle manière que si les autres changent, ces quantités changent aussi, alors on a l'habitude de nommer ces quantités fonction de ces dernières; cette dénomination a la plus grande étendue et contient en elle-même toutes les manières par lesquelles un quantité peut être déterminée par d'autres. Si, par conséquent,  $x$  désigne une quantité variable, alors toutes les autres quantités qui dépendent de  $x$  de n'importe quelle manière, ou qui sont déterminées par  $x$ , sont appelées fonctions de  $x$ .*

L'intégration des fonctions non analytiques parmi les "bonnes" fonctions sera inévitable après les travaux de FOURIER (1768-1830). Citons FOURIER :

*"Il résulte de mes recherches sur cet objet que les fonction arbitraires, mêmes discontinues, peuvent toujours être représentées par des développements en sinus ou cosinus d'arcs multiples..."*. FOURIER a énoncé cette remarque dans le cadre de recherches portant sur la propagation de la chaleur. C'est encore un exemple de l'interaction féconde entre physique et mathématiques.

Cauchy donnera ensuite à l'adjectif "continu" le sens actuel. Il prouvera aussi que même une fonction infiniment différentiable en un point peut ne pas être analytique en ce point:

$$f(x) = e^{-1/x^2} \text{ si } x > 0; f(x) = 0 \text{ si } x \leq 0$$

La théorie de la mesure (l'intégration encore...) permet de définir de nom-

breux espaces de fonctions ainsi que leur position respective dans une vision maintenant géométrisée. En effet on peut par exemple mesurer une “distance” entre deux fonctions  $f$  et  $g$  définies, disons sur  $\mathbb{R}$ , en calculant :

$$\int_{\mathbb{R}} (f - g)^2$$

Ainsi un espace de fonctions, celles telles que l’intégrale  $\int_{\mathbb{R}} f^2$  est finie (de tels espaces sont maintenant appelés espaces de HILBERT (1862-1943)), a les mêmes propriétés que notre espace euclidien.

Notons aussi que l’étude des solutions d’une équation différentielle amène aussi à considérer un espace vectoriel de solutions, qui par exemple pour une équation linéaire d’ordre 2 :  $ay'' + by' + c = 0$ ,  $a \neq 0$  est de dimension 2.

C’est d’ailleurs en étudiant des espaces de fonctions solutions d’une équation différentielle que Peano [Pea1] a été amené à définir la structure d’espace vectoriel.

On peut définir bien d’autres espaces de fonctions, de dimension finie ou infinie, définir de nouveau des fonctions sur ces espaces (que l’on appelle plutôt opérateurs, pour les distinguer des fonctions “usuelles”). Maintenant la physique théorique utilise abondamment les espaces fonctionnels et les opérateurs .

Terminons ce survol partial et incomplet en citant H. WEYL : *“personne n’a jamais pu expliquer ce qu’est une fonction. Mais “une fonction  $f$  est définie si par un moyen quelconque on peut associer à un nombre  $a$  un nombre  $b$ ” ... .*

# Chapter 3

## Calcul différentiel et intégral

### 3.1 Archimède et les mathématiciens de l'antiquité

Nous commencerons avec deux exemples dus à Archimède : l'encadrement de la longueur d'un cercle par celle de la suite des polygones réguliers inscrits et circonscrits à  $2^n$  cotés et le calcul de l'aire bordée par un segment de parabole et sa corde.

#### 3.1.1 Le cercle

Il commence, comme les “éléments” d'Euclide, par des définitions et des postulats. Entre autres il définit des courbes et surfaces concaves.

##### Postulats

(1) La droite est la plus courte des lignes ayant mêmes extrémités.

(2) Quant aux autres lignes, lorsque, situées dans un plan, elles ont les mêmes extrémités, elles sont inégales si, étant les unes et les autres concaves dans la même direction, l'une d'elles est entièrement comprise entre une autre et la droite ayant mêmes extrémités, ou est en partie comprise et en partie commune; la ligne comprise est la plus petite<sup>1</sup> Par ailleurs la démonstration est “élémentaire” si l'on se restreint à des polygones convexes fermés par une corde donnée. En encadrant un cercle entre des polygones réguliers à  $2^n$  cotés, Archimède encadre ensuite la longueur de ce cercle.

On notera que le tome I des œuvres complètes commence par l'étude de la longueur du cercle, et qu'Archimède reviendra sur le sujet, en donnant des estimations numériques, dans le chapitre “de la mesure du cercle”.

---

<sup>1</sup>Nous pouvons maintenant démontrer ce postulat en utilisant la formule de Cauchy-Crofton:  $2 \cdot \text{longueur}(C) = \int_{\mathcal{A}(2,1)} \#C \cap LdL$ ; voir la section 5.

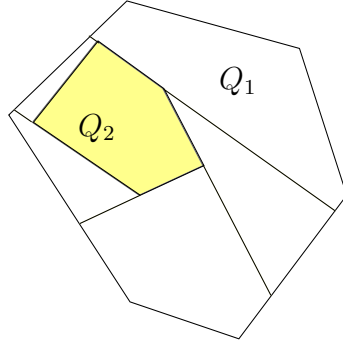


Figure 3.1: Comparaisons des longueurs de polygones convexes emboîtés

### 3.1.2 L'aire comprise entre une parabole et une corde

Archimède propose deux calculs. Le premier utilise des “pesées” convenablement construites équilibrant une aire triangulaire de côté situé sur un côté du balancier et une aire de rectangle “suspendu” de l'autre côté du point  $B$  (voir figure 3.1.2) qui sert d'axe.

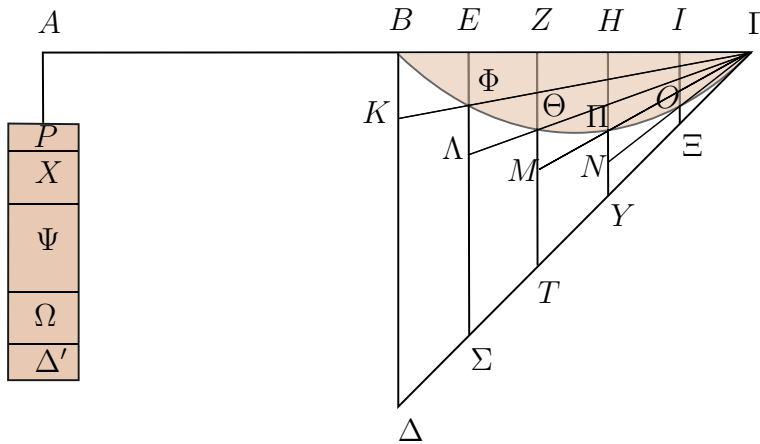


Figure 3.2: Le calcul “statique” de l'aire comprise entre une parabole et une de ses cordes

Citons Archimède: *Soit  $B\Theta\Gamma$  un segment délimité par une droite et une parabole. Soit, dans un premier cas, la droite  $B\Gamma$  parallèle au diamètre, et, par le point  $\Gamma$ , la droite  $\Gamma\Delta$  tangente à la parabole au point  $\Gamma$ ; le triangle  $B\Gamma\Delta$  sera donc rectangle. Partageons la droite  $B\Gamma$  en autant de parties égales que l'on voudra:  $BE, EZ, ZH, H\Gamma$ ; menons, par les points de division, les parallèles au diamètre  $E\Sigma, ZT, HY, I\Xi$ , et relierons au point  $\Gamma$ , par des droites que nous prolongeons, les points où ces parallèles coupent la parabole. Dès lors, je dis que le triangle  $B\Delta\Gamma$  est plus petit que le triple de l'ensemble des trapèzes  $KE, \Lambda Z, MH, NI$ , et du triangle  $\Xi I\Gamma$ , et qu'il est plus grand que le triple de l'ensemble*

des trapèzes  $Z\Phi$ ,  $H\Phi$ ,  $\Pi$  du triangle  $IO\Gamma$ .

En effet, menons la droite  $AB\Gamma$ ; découpons-en une droite  $AB$  égale à la droite  $B\Gamma$ , et imaginons que  $A\Gamma$  soit un levier, dont le milieu est le point  $B$ , et qui soit suspendu au point  $B$ . Suspendons aussi le triangle  $B\Delta\Gamma$  au levier aux points  $B$ ,  $\Gamma$ , et suspendons à l'autre partie du levier, au point  $A$ , des aires  $P$ ,  $X$ ,  $\Psi$ ,  $\Omega$ ,  $\Delta'$ . Que l'aire  $P$  fasse équilibre au trapèze  $\Delta E$  tel qu'il est placé, l'aire  $X$  au trapèze  $Z\Sigma$ , l'aire  $\Psi$  au trapèze  $TH$ , l'aire  $\Omega$  au trapèze  $YI$ , et l'aire  $\Delta'$  au triangle  $\Xi I\Gamma$ . Dès lors un ensemble fera équilibre à l'autre, en sorte que le triangle  $B\Delta\Gamma$  sera triple de l'aire  $PX\Psi\Omega\Delta'$ . De plus, puisque  $B\Gamma\Theta$  est un segment délimité par une droite et une parabole, que la droite  $B\Delta$  a été menée du point  $B$  parallèlement au diamètre, que la droite  $\Gamma\Delta$  a été mené du point  $\Gamma$  tangentielllement à la parabole en  $\Gamma$ , et qu'une autre droite  $\Sigma E$  a aussi été menée parallèlement au diamètre, le rapport de  $B\Gamma$  à  $BE$  sera le même que celui de  $\Sigma E$  à  $E\Phi$ ; en sorte que le rapport de  $BA$  à  $BE$  sera aussi le même que celui du trapèze  $\Delta E$  au trapèze  $KE$ .

De fait, Archimède montre que le moment du triangle  $B\Delta\Gamma$  est égal à l'aire entre la corde  $B\Gamma$  et la parabole, et donc à l'aire suspendue en  $A$ .

Le second calcul décompose l'aire cherchée en somme infinie d'aires de triangles de façon à se ramener à calculer la somme des termes d'une série géométrique.

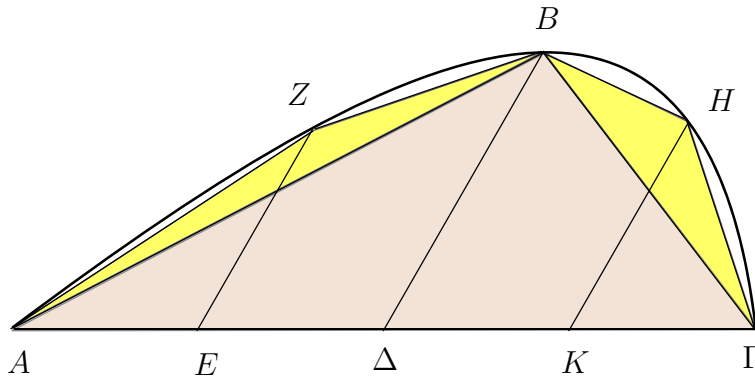


Figure 3.3: Triangles dont la réunion est la zone comprise entre un arc de parabole et sa corde.

Plus précisément, l'aire des triangles construits à chaque étape est le quart de l'aire des triangles construits à l'étape précédente.

Pour cela observons d'abord que le rapport de la longueur de  $B\Delta$  à la longueur de  $B\Theta$  est la même que celui du carré de  $A\Delta$  au carré de  $Z\Theta$  (Propriété mentionnée dans un ouvrage écrit par Aristée et Euclide qui ne nous est pas parvenu (voir la note de Ver Eecke à propos de la proposition III dans [Arch]; on retrouve cette propriété analytiquement en écrivant  $y = ax^2$  l'équation d'une parabole

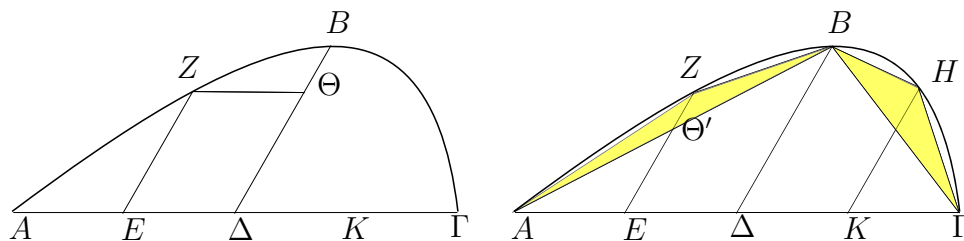


Figure 3.4: Le calcul de l'aire des triangles de la seconde étape

repérée par rapport à la tangente en un point et au diamètre d'origine ce point).

Observons ensuite qu'un diamètre d'une parabole est la demi-droite lieu des milieux des cordes parallèles à une direction donnée. Son origine est le point de la parabole où la tangente est parallèle à la direction donnée. Tous les diamètres sont parallèles entre eux [et à la direction asymptotique de la parabole]. Notons  $\Theta'$  le point d'intersection du diamètre d'origine  $Z$  [où la tangente à la parabole est parallèle à  $AB$ ] et du côté  $AB$  du triangle construit à l'étape précédente. La longueur du segment  $E\Theta'$  est donc le double de celle du segment  $\Theta'Z$ . L'aire du triangle  $AEB$  est le double de l'aire du triangle  $\Delta BA$ . Par ailleurs l'aire du triangle  $AE\Theta'$  est le double de l'aire du triangle  $A\Theta'Z$ , et l'aire du triangle  $\Theta'BE$  est le double de l'aire du triangle  $Z\Theta'B$ . En conséquence [l'aire du] triangle  $AB\Gamma$  est octuple de [l'aire du] triangle  $AZB$ , et l'on démontrerait de la même manière, qu'il [elle] est octuple [de l'aire] du triangle  $BHT$ .

Il reste maintenant à faire la somme des termes d'une série géométrique de raison  $\frac{1}{4}$ , ce qui représente un peu plus de quatre pages dans [Arch].

## 3.2 Stevin (1548-1620)

C'est un ingénieur et mathématicien associé au prince Maurice d'Orange (Pays Bas).

Pour montrer que le centre de gravité d'un triangle est le point d'intersection des médianes, il va montrer que, suspendu par un sommet, un triangle à l'équilibre a la médiane correspondante verticale.

Pour cela il remarque que de (fins) parallélogrammes de côtés parallèle au côté opposé au sommet  $A$  et à la médiane de sommet  $A$  sont en équilibre si on les suspend par un fil qui suit la médiane de sommet  $A$ . Plus les parallélogrammes sont fins, moins on laisse de côté d'aire du triangle.

La médiane de sommet  $A$  est donc une ligne centrale de gravité.

### 3.3 Kepler (1571-1630)

Montrons comment il montre que le volume d'un tore de révolution est égal à celui d'un cylindre droit dont les disques sont égaux aux méridiens du tore solide et dont la hauteur est égale à la longueur de l'âme (cercle lieu des centres des disques méridiens).

Imaginons l'axe de rotation vertical. Un secteur du tore obtenu en le découpant par un dièdre d'arête l'axe de révolution du tore est formé de disques tassés vers l'axe et au contraire écartés le plus loin de l'axe. Pour calculer le volume d'un tel secteur, regardons d'abord un secteur de l'anneau horizontal dans le plan de symétrie horizontal du tore; la somme des longueurs des arcs de cercles extérieur et intérieur est la longueur de l'arc de cercle médian, c'est à dire à un arc de l'âme correspondant au même angle. Il en va de même pour deux arcs symétriques par rapport à l'arc médian. L'aire du secteur de l'anneau est donc égale à la longueur de l'arc médian multiplié par la longueur de la corde horizontale du disque méridien (voir figure 3.5). Il suffit maintenant de répéter le raisonnement pour les anneaux horizontaux engendrés par une corde horizontale de l'un des disques méridien puis empiler des tranches fines et échanger les "sommations". Ceci permet de conclure.

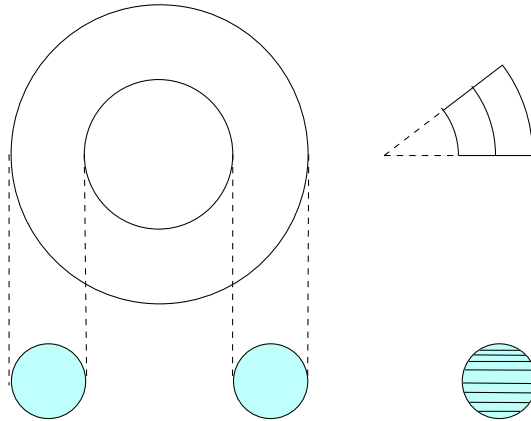


Figure 3.5: Calcul du volume du tore solide

### 3.4 Galilée

Galileo GALILEI 1564-1642.

En étudiant la chute des corps, Galilée a dû faire la différence entre immobilité et vitesse nulle à un instant donné.

*Tout corps se trouvant au repos pour une raison quelconque mais mobile par sa nature, se mettra en mouvement lorsqu'on le laissera libre pourvu qu'il ait par*

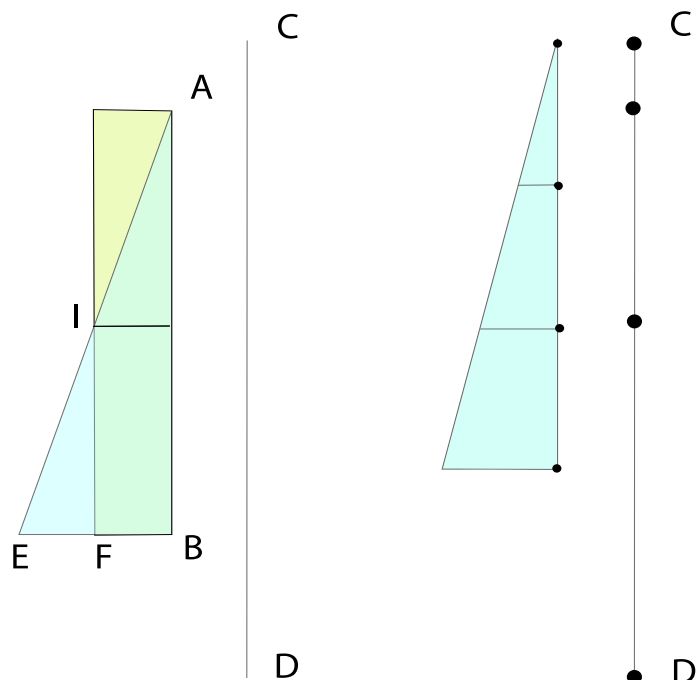


Figure 3.6: Mouvement uniformément accéléré et mouvement uniforme; temps et distances parcourues

*nature une inclination vers quelque lieu particulier. S'il est indifférent à tout lieu, il restera en repos, n'ayant pas de raison de se mouvoir plutôt vers l'un que vers l'autre. Mais s'il a une telle inclination, il s'en suivra nécessairement que dans son mouvement il ira continuellement en s'accélération. Commencant avec un mouvement extrêmement lent, il n'acquerra aucun degré de vitesse qu'il n'ait passé d'abord par tous les degrés de vitesse inférieurs, disons de lenteur plus grande. Car partant de l'état de repos (qui est le degré infini de lenteur), il n'y a aucune raison qu'il doive entrer dans un degré déterminé de vitesse avant d'entrer dans un plus faible, et dans un autre encore plus petit avant celui-là. Il semble très raisonnable de penser qu'il passera d'abord par les degrés les plus voisins de celui d'où il est parti et de ceux-là aux plus éloignés. Mais le degré d'où le mobile commence à se mouvoir est celui de la plus grande lenteur, c'est à dire du repos...*, Dialogues: première journée, voir [Zou].

Galilée étudie ensuite, en deux étapes, le mouvement uniformément accéléré.

**Théorème 3.4.1.** *Un mobile uniformément accéléré partant du repos parcourra un espace donné dans le même temps qu'il lui faudrait pour parcourir ce même espace avec une vitesse constante égale à la moitié de la vitesse maximum atteinte à la fin du mouvement uniformément accéléré.*

La démonstration du théorème 3.4.1 se ramène au théorème de Thales (voir

figure 3.6).

**Théorème 3.4.2.** *Si un mobile descend, à partir du repos, avec un mouvement uniformément accéléré, les espaces parcourus en des temps quelconques sont entre eux comme le carré des durées des parcours (voir figure 3.6).*

On se ramène au cas précédent en comparant les distances parcourues à partir du repos par deux mouvements uniformes l'un pendant le temps  $T$  avec vitesse  $v$ , un autre pendant le temps  $2T$  et la vitesse  $2v$ , et enfin un dernier pendant le temps  $aT$  avec la vitesse  $av$ . L'intégrale  $\int_0^T t \cdot dt = \frac{1}{2}T^2$  n'est plus très loin.

Un autre texte permet d'être sur que Galilée avait une notion claire de la signification du futur  $dx$ .

Pour comprendre comment un solide formé de deux cercles concentriques roule lorsque le plus grand des cercles roule sans glisser sur une droite, Galilée commence par procéder avec deux hexagones concentriques.

Figure 3.7: Cercles et hexagones qui roulent

Tandis qu'en déroulant l'hexagone extérieur, la trace de ses cotés remplit la droite, la trace des hexagones concentriques laisse des vides de plus en plus grands pendant que le sommet le plus bas du petit hexagone décrit un arc de cercle centré au sommet du grand hexagone en contact avec la droite. À l'extrême, la trajectoire du centre est une union d'arcs de cercles d'extrémités situées sur une parallèle à la droite (voir figure 3.4).

Lorsque l'hexagone est remplacé par un polygone régulier ayant plus de cotés, le phénomène se répète tandis que les intervalles se réduisent. La situation limite, lorsque les polygones sont devenus des cercles impose au cercle intérieur de glisser: *Si nous considérons un segment divisé en une infinité de morceaux infiniment petits et indivisibles, on doit pouvoir imaginer le segment étendu indéfiniment en intercalant un nombre infini de petits espaces indivisibles vides.*

### 3.4.1 Cavalieri

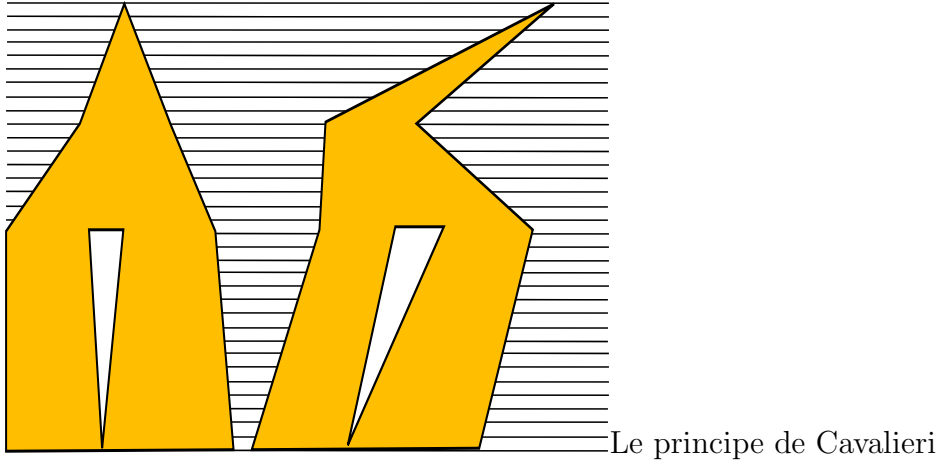
(1598-1647)

Bonaventura Cavalieri, professeur à Bologne, était un disciple de Galilée.

Le principe de Cavalieri affirme qu'un glissement horizontal des tranches horizontales d'une figure ne change par l'aire (ou le volume si la figure est un solide).

En langage moderne on dirait plutôt, en faisant glisser des segments suivant des verticales :

si  $\forall x, a \geq x \geq b$ ,



$$f_2(x) - f_1(x) = \varphi_2(x) - \varphi_1(x)$$

$$A_1 = \int_a^b [f_2(x) - f_1(x)] dx = A_2 = \int_a^b [\varphi_2(x) - \varphi_1(x)] dx.$$

Il obtient, à l'aide de démonstrations parfois périlleuses le résultat:

$$\int_0^a x^n dx = \frac{a^{n+1}}{n+1} \quad (n \text{ entier positif})$$

### 3.5 De Fermat à Barrow

**Pierre Simon de Fermat** (1601-1665). Une méthode pour l'évaluation des maxima et minima. La théorie d'évaluation des maxima et minima présuppose deux quantités inconnues et la règle suivante:

Soit  $a$  une inconnue du problème (qui est en une deux ou trois dimensions, suivant la formulation du problème). On suppose que l'expression à étudier est un polynome en  $a$ . On doit maintenant remplacer l'inconnue initiale par  $a+e$  et on doit exprimer la quantité à maximiser ou minimiser en terme de  $a$  et  $e$  (polynome en  $a$  et  $e$ ). On doit adégaler, pour utiliser le terme de Diophante, les deux expressions [avant et après ajout de  $e$ ] de la quantité maximum ou minimum et éliminer leur terme commun; Maintenant les deux termes contiennent des termes en  $e$  et ses puissances. On doit diviser tous ces termes par  $e$  ou une puissance plus élevée de  $e$  pour éliminer  $e$  d'au moins l'un de ces termes. On supprime alors tous les termes où  $e$  apparaît encore, et on adégale les autres. La solution de cette dernière équation fournit une valeur de  $a$  qui conduit au maximum ou au minimum, en utilisant de nouveau l'expression initiale.

Voici un exemple.

*Pour diviser le segment  $AC$  en  $E$  de sorte que  $AE \times EC$  soit maximum*

Figure 3.8: Recherche du maximum de  $AE \times EC$ 

Posons  $AC = b$ ; soit  $a$  [la longueur de] l'un des segments, de sorte que l'autre est [de longueur]  $b - a$ . Le produit, dont on doit trouver le maximum, est  $ba - a^2$ . Si le premier segment est de longueur  $a + e$  le second aura pour longueur  $b - a - e$ , le produit des longueurs devient:  $ba - a^2 + be - 2ae - e^2$  il faut adégaler avec la formule précédente  $ba - a^2$ . Il reste  $be \simeq 2ae$ . En supprimant  $e$ :  $b = 2a$ . Il faut donc prendre [a égal à] la moitié de  $b$ .

On peut difficilement attendre une méthode plus générale.

Remarquons qu'aujourd'hui nous dirions avoir démontré seulement que le milieu du segment est un point critique pour la fonction  $AE \times EC$ .

### Tangente en un point d'une courbe.

Considérons la parabole  $BDN$  de sommet  $D$  et de diamètre  $DC$ ; soit  $B$  un point sur la parabole où l'on cherche la droite  $BE$  tangente à la parabole et coupant le diamètre en  $E$ .

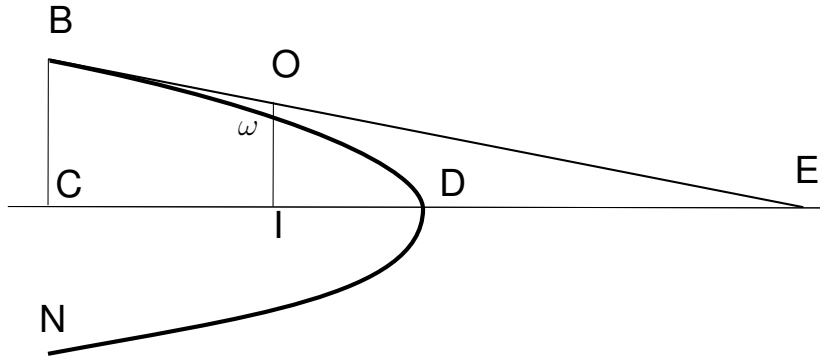


Figure 3.9: Tangente à une parabole

Nous choisissons sur  $BE$  un point  $O$  pour lequel on dessine l'ordonnée  $OI$ ; Nous construisons aussi l'ordonnée  $BC$  du point  $B$ . On a alors:  $DC/DI > BC^2/OI^2$  [l'égalité vient d'une propriété de la parabole déjà connue dans l'antiquité, et on remarque que si on choisit une autre droite passant par  $B$  que la tangente à la parabole, la quantité  $\frac{DC}{DI} - \frac{BC^2}{OI^2}$  n'admet pas un minimum (0) en  $B$ ], puisque le point  $O$  est extérieur [au domaine bordé par] la parabole [qui contient le point  $C$ , et qui est convexe]. Mais on a  $BC/OI = CE/IE$  puisque les triangles sont semblables, et donc  $BC^2/OI^2 = CE^2/IE^2$ . On a donc  $CD/DI > CE^2/IE^2$ .

Maintenant le point  $B$  est donné, et donc l'ordonnée  $BC$ , et donc le point  $C$ , donc aussi  $CD$ . Soit  $CD = d$  cette quantité donnée, posons  $CE = a$  et  $CI = e$

[ $O$  proche de  $B$  donc  $I$  proche de  $C$ ]; on obtient:

$$\frac{d}{d-e} > \frac{a^2}{a^2 + e^2 - 2ae}$$

Enlevons les fractions :

$$da^2 + de^2 - 2dae > da^2 - a^2e.$$

Adégalons, en suivant la méthode précédente; en supprimant le terme commun on trouve :

$$de^2 + a^2e \simeq -a^2e$$

ou, ce qui revient au même,

$$de^2 + a^2e \simeq 2dae$$

Divisons tous les termes par  $e$ :

$$de + a^2 \simeq 2da.$$

En enlevant  $de$ , il reste  $a^2 = da$  et donc  $a = 2d$ .

On a démontré que  $CE$  est le double de  $CD$ , ce qui est le résultat.

La méthode ne faillit jamais et peut être étendue à nombre de beaux problèmes;

...

Autre exemple : construction de la tangente à la “roulette” (cycloïde) par Roberval (Gilles Personne de Roberval 1602-1675)

### Isaac Barrow (1630-1677)

Il fut professeur à Cambridge où il eut Newton comme élève. Son point de départ est le mouvement, ce qui lui donne une approche cinématique de la notion de tangente. Il utilise avec certaines précautions la méthode des indivisibles due à Fermat.

Ses conférences contiennent, sous une forme géométrique, des théorèmes équivalents à

$$ds^2 = dx^2 + dy^2$$

et

$$(d/dx) \int_0^x y dx = y$$

souvent appelé théorème fondamental [du calcul différentiel et intégral, voir Leibniz bientôt].

## 3.6 Isaac Newton

(1643-1727)

L'introduction à la traduction en français des *Principia* par Madame du Chatelet (1756) est écrite par Voltaire et témoigne de sa profonde admiration pour Newton.

Commençons par l'introduction de la notion de limite (première et dernière raison) dans le livre 1 des *Principia*, dont la partie la plus connue concerne la cinématique. Citons Newton [New2] :

*De la méthode des premières et dernières raisons employée dans tout cet ouvrage*  
 LEMME PREMIER

*Les quantités & les raisons<sup>2</sup> des quantités qui tendent continuellement à devenir égales pendant un temps fini, & qui avant la fin de ce temps approchent tellement l'égalité, que leur différence est plus petite qu'aucune différence donnée, deviennent à la fin égales.*

Si on le nie, qu'on suppose qu'elles soient à la fin inégales, & que leur dernière différence soit D, puisqu'elles ne peuvent pas approcher plus près de l'égalité que cette différence donnée D, leur différence ne sera donc pas plus petite que toute différence donnée, ce qui est contre l'hypothèse.

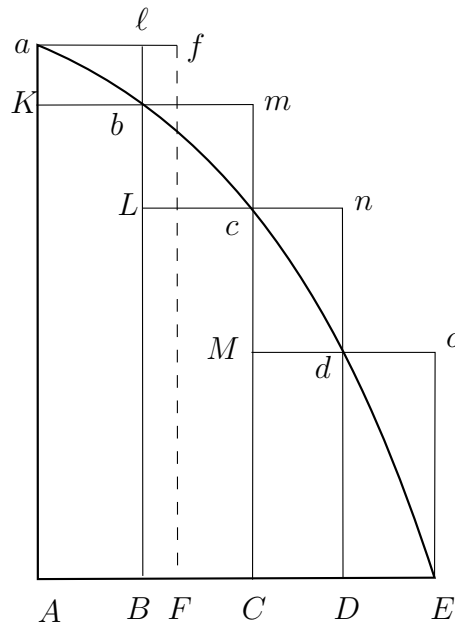


Figure 3.10: Ultima ratio : une intégrale, d'après fig 6 planche 1 des principia

LEMME II

---

<sup>2</sup>rappports

Figure 3.11: Angle entre la corde et la tangente à une courbe suivant Newton

*Si dans une figure quelconque  $AacE$ , comprise entre les droites  $Aa$ ,  $AE$ , et la courbe  $ace$ , on inscrit un nombre quelconque de parallélogrammes  $Ab$ ,  $Bc$ ,  $Cd$ , etc. compris sous les bases égales  $AB$ ,  $BC$ ,  $CD$ , etc. & sous les côtés  $Bb$ ,  $Cc$ ,  $Dd$ , etc; parallèles au côté  $Aa$  de la figure ; & qu'on acheve les parallélogrammes  $akbl$ ,  $bLcm$ ,  $cMdn$ , etc. qu'on diminue ensuite la largeur de ces parallélogrammes, & qu'on augmente leur nombre à l'infini : les dernières raisons<sup>3</sup> qu'auront entr'elles la figure inscrite  $AKbLcMD$ , la circonscrite  $AalbmcndoE$ , & la curviligne  $AabcdE$ , seront des raisons d'égalité. (voir fig 6 de la planche 1.)*

Car la différence de la figure inscrite & de la figure circonscrite, est la somme des parallélogrammes  $Kl$ ,  $Lm$ ,  $Mn$ ,  $Do$ , c'est à dire (à cause de l'égalité de toutes les bases) que cette différence est égale au rectangle  $ABla$  fait sur l'une des bases  $Kb$  & sur la somme  $Aa$ , de toutes les hauteurs ; mais ce rectangle, à cause que sa largeur diminue à l'infini, deviendra plus petite qu'aucun rectangle donné. Donc (par le Lemme premier) la figure inscrite, la figure circonscrite, & à plus forte raison la figure curviligne intermédiaire seront à la fin égales. C.Q.F.D.

Newton montre ensuite que le même résultat est vrai si les bases des rectangles sont inégales. Il se croit ensuite obligé de préciser qu'il ne s'agit pas d'une théorie physique, et que les considérations précédentes ne sont pas en contradiction avec l'existence d'atomes.

Nous voyons aujourd'hui que cette définition de l'intégrale ne permet pas, comme l'intégrale de Riemann, d'intégrer toute fonction continue sur un intervalle fermé borné. Il faut se limiter à des fonctions ayant un nombre fini d'intervalles de monotonie, ou à la rigueur de variation bornée, éventuellement en éliminant des voisinages d'un nombre fini de points.

Le lemme 4 est une manipulation d'infiniments petits proportionnels (aujourd'hui on dirait  $dx$  et  $a \cdot dx$ ), le 5 fait de même pour des aires infinitésimales.

le lemme 6 montre que si l'on considère un arc et sa corde l'angle entre la tangente et la corde tend vers 0.

La démonstration utilise de manière un peu surprenante la courbure de l'arc, supposée continue.

page 294 Struik.

Le lemme 7 énonce enfin l'équivalence de la longueur de l'arc, de la corde et de la projection de l'arc sur la tangente.

---

<sup>3</sup>limites des rapports

### La méthode des fluxions

Ce traité a été traduit par Buffon 1740

Ce traité contient des calculs de développement en série, par exemple

$$\frac{a^2}{b+x} = \frac{aa}{b} - \frac{a^2x}{b^2} + \frac{a^2x^2}{b^3} - \frac{a^2x^3}{b^4} + \dots$$

laquelle suite étant continuée à l'infini =  $\frac{aa}{b+x}$ .

Il développe ensuite en série  $\frac{1}{1+xx}$  et  $\sqrt{aa+xx}$ , ce qui est l'occasion de rappeler l'usage d'exposant entiers, positifs ou négatifs, ou rationnels.

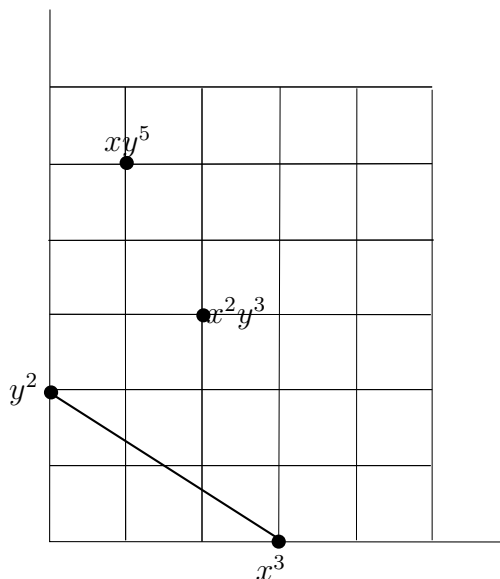
Il en vient ensuite à expliquer la construction d'une série entière solution d'une équation algébrique (le cas intéressant est lorsque l'équation admet une singularité en l'origine) en utilisant ce que l'on appelle maintenant polygone de Newton. Ce polygone le bord de l'enveloppe convexe de l'ensemble des points dont les coordonnées sont les exposants des monomes du polynôme en  $x, y$  dont on cherche une solution. La somme de la série définit un arc de courbe  $y =$  série fractionnaire en  $x$ , partant de l'origine contenu dans l'ensemble  $f(x, y) = 0$ . Newton ne vérifie pas que la série construite est convergente voir par exemple [?]

Exemple  $f(x, y) = y^2 - x^3 + 2x^2 \cdot y^3 + xy^5 = 0$ .

L'idée est de chercher une première approximation de la solution de l'équation de la forme  $x = t^q y = ct^p$ . Quels sont les monomes de l'équation qui fournisse(nt) le ou les termes de plus bas degré en  $t$  peuvent être éliminés, à condition de bien choisir le coefficient  $c$ , lorsque l'on a choisi  $p$  et  $q$ ? Un monome  $x^\alpha \cdot y^\beta$  devient  $t^{p\alpha+q\beta}$ . Il faut regarder dans le plan  $\alpha, \beta$  les droites d'équation  $p\alpha + q\beta = d$ . En partant "d'en bas à gauche" les couples  $(\alpha, \beta)$  que contient la première droite qui touche le polygone de Newton est/sont le ou les lexposants du ou des monomes cherché. Seules certaines pentes permettent de trouver ainsi au moins deux monomes. Cette ou ces pentes fournissent les approximations cherchées. Il reste encore à trouver une suite d'approximations successives qui va déterminer un arc analytique contenu dans l'ensemble d'équation  $f(x, y) = 0$ .

### Fluents et fluxions

J'appellerai **quantité fluentes**, ou simplement *fluentes* ces quantités que je connais comme augmentées graduellement et indéfiniment, je les représenterai par les dernières lettres de l'alphabet  $v, x, y$ , et  $z$  pour les distinguer des autres quantités qui dans les équations sont considérées comme connues et déterminées qu'on représente par des lettres initiales  $a, b$ , etc, et je représenterai par les mêmes dernières lettres surmontées d'un point  $\dot{v}, \dot{x}, \dot{y}$  et  $\dot{z}$  les vitesses dont les

Figure 3.12: Polygone de Newton ( $y^2 - x^3 + 2x^2y^3 + xy^5$ )

fluentes sont augmentées par le mouvement qui les produit, et que par conséquent on peut appeler **fluxions**. Ainsi pour la vitesse ou la fluxion de  $v$  je mettrai  $\dot{v}$ .

Problème 1 Étant donnée la relation des quantités fluentes, trouver la relation de leurs fluxions.

Disposez l'équation par laquelle la relation donnée est exprimée suivant les dimensions de l'une de ses quantités fluentes et multipliez les termes par une progression arithmétique quelconque et ensuite par  $\frac{\dot{x}}{x}$ , après quoi égalez à zéro la somme de tous les produits et vous aurez l'équation cherchée.

$$\begin{array}{cccc|ccc}
 x^3 & -ax^2 & +axy & -y^3 & -y^3 & +axy & -ax^2 + x^3 \\
 3\frac{\dot{x}}{x} & 2\frac{\dot{x}}{x} & \frac{\dot{x}}{x} & 0 & 3\frac{\dot{y}}{y} & \frac{\dot{y}}{y} & 0 \\
 & & & & & & \\
 3\dot{x}x^2 & -2a\dot{x}x & +\dot{a}xy & * & -3\dot{y}y^2 & +\dot{a}yx & *
 \end{array}$$

La somme des produits est :

$$3\dot{x}x^2 - 2a\dot{x}x + \dot{a}xy - 3\dot{y}y^2 + \dot{a}yx$$

qui, étant égale à zéro, donne la relation des fluxions  $\dot{x}$  et  $\dot{y}$ , car si vous donnez à volonté une valeur à  $x$ , l'équation  $x^3 - ax^2 + axy - y^3 = 0$  donnera une valeur de  $y$  ; ce qui étant déterminé, l'on aura  $\frac{\dot{x}}{\dot{y}} = \frac{3y^2 - ax}{3x^2 - 2ax + ay}$

Réciproquement :

Problème II Étant donnée la relation des fluxions, trouver celle des quantités fluentes [équation différentielle à deux ou trois variables].

Comme le problème est l'inverse du précédent, on peut le résoudre en procédant d'une façon contraire, c'est à dire il faudra disposer suivant les dimensions de  $x$  les termes multipliés par  $\dot{x}$ , ensuite les diviser par  $\frac{\dot{x}}{x}$  et enfin par le nombre de leur dimension ou peut-être par quelqu'autre progression arithmétique. On répètera la même opération pour les termes multipliés par  $\dot{y}$ ,  $\dot{z}$  et on égalera toute la somme à zéro en rejetant les termes superflus.

**Exemple.**

	$3\dot{x}x^2$	$-2a\dot{x}x$	$+a\dot{x}y$		$-3\dot{y}y^2$	*	$+a\dot{y}x$
divisé par $\frac{\dot{x}}{x}$	$3x^3$	$-2ax^2$	$+axy$		divisés par $\frac{\dot{y}}{y}$	$-3y^3$	* $+axy$
divisés par	3	2	1		divisés par	3	2 1
quotient	$x^3$	$-ax^2$	$+axy$		quotient	$-y^3$	* $axy$

La somme est

$$x^3 - ax^2 + axy - y^3$$

qui égalée à zéro donne la relation des quantités fluentes  $x$  et  $y$ . On peut observer que, quoique le terme  $axy$  se trouve deux fois, je ne le met qu'une fois. [...] [Le traitement des termes  $a\dot{x}y$  et  $a\dot{y}x$  correspond bien à:  $\frac{d}{dt}(axy) = a\dot{x}y + a\dot{y}x$ ].

Il y a d'autres circonstances à observer[...]

J'ajouterai seulement que quand cette méthode nous donne la relation des fluentes, il faut par le pb I en chercher les fluxions, et quand elles se retrouvent les mêmes, l'opération est bonne, mais sans cela, non.

Puis, ...ça se complique comme tout traité de résolution d'équations différentielles.

Newton traite ensuite les problèmes suivants :

- Problèmes d'extrema
- Tangentes
- Calculs d'aires
- Longueur d'arcs de courbe : la fluxion d'une ligne courbe est égale à la racine carrée de la somme des carrés de l'abscisse et de l'ordonnée; [On écrirait aujourd'hui  $ds^2 = dx^2 + dy^2$ ]. Newton traite ensuite le cas des graphes en prenant " $t = x$ ".

### 3.6.1 Gottfried Wilhelm Leibniz

(1646-1716) Né à Leipzig, il devient conseiller diplomatique au service de l'électeur de Mayence. Cela lui permet de séjourner à Paris entre 1672 et 1676 où il a développé ses idées sous l'influence de Huygens.

Il aura aussi des contacts avec l'Angleterre, et en particulier en mains deux lettres de Newton qui contenaient beaucoup des résultats de celui-ci sur les séries mais peu de choses sur le calcul différentiel et intégral.

**Une nouvelle méthode pour les maxima et minima comme pour les tangentes...** (1684) (traduction du latin)

C'est dans cet article que Leibniz introduit la notation  $dx$  et donne des règles formelles de calcul pour utiliser ces *segments de droites choisis arbitrairement [petits]*

*Soit l'axe AX et les différentes courbes VV, WW, YY, ZZ; soient VX, WX, YX, ZX leurs ordonnées perpendiculaires à l'axe. que je nommerai respectivement v, w, y, z, et AX, que je nommerai x, leur abscisse sur l'axe. Soient VB, WC, YD, ZE, les tangentes, coupant celui-ci respectivement aux ponts B, C, D, E. Appelons alors dx un segment de droite choisi arbitrairement, et dv (dw, dy ou dz), c'est à dire différence de v (de w, y, ou z) un segment qui soit avec dx comme à l'ordonnée correspondante de la courbe VV), dy sera égal à dv.*

**Addition et soustraction:** *si  $z - y + w + x = v$ ,  $d(z - y + w + x)$ , autrement dit dv [on écrirait aujourd'hui  $d(z - y + w + x)$  ], sera égal à  $dz - dy + dw + dx$ .*

**Multiplication:**  *$dxv$  égal  $x dv + v dx$ . Car on a tout loisir d'employer soit l'expression  $xv$ , soit à sa place, une lettre, par exemple y, pour abréger. Remarquons que dans ce calcul x et dx sont traités sur le même pied, de même que y et dy, ou toute autre lettre indéterminée et sa différentielle. Remarquons également que la démarche inverse, à partir de l'équation différentielle, n'existe pas toujours, ou tout au moins suppose une certaine précaution, dont je parlerai ailleurs.*

*Enfin la Division:*

$$d\frac{v}{y} \text{ ou (en posant } z \text{ égal à } \frac{v}{y}\text{), } dz \text{ égal } \frac{+-vdy - +ydv}{yy}.$$

*En ce qui concerne les Signes, remarquons bien que lorsque dans le calcul on passe simplement d'une lettre à sa différentielle, on conserve naturellement les mêmes signes, on remplace  $+z$  par  $+dz$ ,  $-z$  par  $-dz$ , comme le montrent l'addition et la soustraction établies plus haut ;*

Leibniz explique ensuite que lorsque  $v$  admet un maximum ou un minimum,  $dv$  devient égal à 0, tandis que si  $dv$  est infini en regard de  $dx$  la tangente à la courbe  $\{(x, v)\}$  est perpendiculaire à l'axe des abscisse (vertical sur la figure). En prenant les différences de différences, il étudie ensuite la concavité d'une courbe.

Donnons un exemple d'application : la trajectoire d'un rayon lumineux qui traverse deux milieux différents (réfraction).

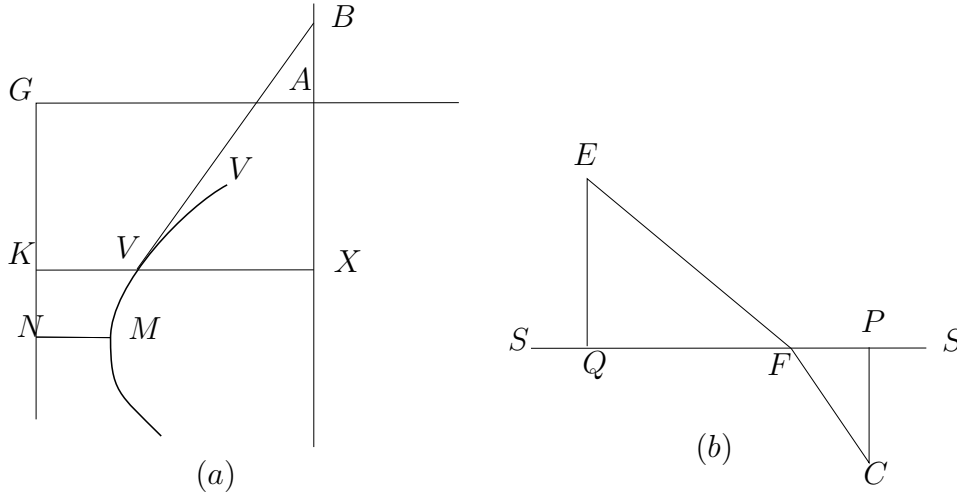


Figure 3.13: Réfraction

Soit deux points  $C$  et  $E$  et une droite  $SS$  dans le plan contenant  $C$  et  $E$ . On cherche le point  $F$  sur  $SS$  tel que, lorsque l'on joint  $E$  et  $C$  à  $F$  la somme du rectangle de  $CF$  et d'une droite donnée  $h$  et du rectangle de  $FE$  et d'une droite  $r$  soit aussi petite que possible. En d'autres termes, si  $SS$  est la ligne séparant les deux milieux, et si  $h$  représente la densité du milieu du côté de  $C$  (disons de l'eau), et  $r$  celle du milieu du côté de  $EE$  (disons, de l'air), on cherche le point  $F$  tel que le chemin de  $C$  à  $E$  via  $F$  soit le plus court [parcouru en temps minimum] possible. Supposons que toutes les sommes possibles des rectangles, pour tous les chemins possibles, soient représentés par les ordonnées  $KV$  de la courbe  $VV$  perpendiculairement à la ligne  $GK$  [figure 3.13, (a)]. Nous appellerons ces ordonnées  $w$ . On veut maintenant trouver leur minimum  $NM$ . Puisque  $C$  et  $E$  sont donnés [figure 3.13 (b)], leurs perpendiculaires à  $SS$  sont aussi données, soit  $CP$  (que nous noterons  $c$ ) et  $EQ$  (que nous noterons  $e$ ); de plus  $PQ$  (que nous appelons  $p$ ) est donné. On note  $x$  [la longueur]  $QF = GN$  (ou  $AX$ ),  $f$  [la longueur]  $CF$  et  $g$  [la longueur]  $EF$ . on a :  $FP = p - x$ ,  $f = \sqrt{cc + pp - 2px + xx}$  ou  $= \sqrt{\ell}$  pour abrégier;  $g = \sqrt{ee + xx}$  ou  $= \sqrt{m}$  pour abrégier. on a :

$$w = h\sqrt{\ell} + r\sqrt{m}.$$

L'équation différentielle (puisque  $dw = 0$  pour un minimum) est, en suivant notre mode de calcul,

$$0 = +hdl : 2\sqrt{\ell} + rdm : 2\sqrt{m}.$$

Cependant  $d\ell = -2(p-x)dx$ ,  $dm = 2xdx$ ; donc

$$h(p-x) : f = rx : g.$$

Lorsque l'on applique cela au dioptre, et que l'on prend  $f$  et  $g$ , c'est à dire  $CF$  et  $EF$  égaux (puisque la réfraction au point  $F$  est la même quelque soit la longueur de la droite  $CF$ ), on a  $h(p-x) = rx$  ou  $h : r = x : (p-x)$ , ou  $h : r = QF : FP$ , donc les sinus des angles d'incidence et de réfraction,  $FP$  et  $QF$ , sont dans le rapport inverse de celui de  $r$  et  $h$ , les densités des milieux dans lesquels la réfraction se place. Cependant cette densité ne doit pas se comprendre par rapport à nous, mais par rapport à la résistance que rencontrent les rayons lumineux.

C'est aussi Leibniz qui introduira le signe  $\int$  dans *De geometria recondita et analysi indivisibilium atque infinitorum* (juin 1686) et énonce ce que les anglo-saxons appellent "théorème fondamental du calcul": la relation inverse entre intégration et différentiation. voir *Supplementum geometriae dimensoriae...* Acta Eruditorum (1693).

### 3.6.2 Après Newton et Leibniz

- l'Hôpital, Taylor, Jakob Bernoulli : le développement de l'analyse.
- D'Alembert puis Daniel Bernoulli : les cordes vibrantes
- Johann et Jakob Bernoulli : la courbe brachystochrone (un mobile glissant le long de la courbe soumis à une accélération constante à partir du repos à l'origine de la courbe atteint l'extrémité en un temps minimum.
- Euler et Lagrange : le calcul des variations

Notons cependant que l'acceptation des idées de Newton n'allait pas de soi : Citons un petit extrait de "the analyst, or a discourse addressed to an infidel mathematician" de George BERKELEY (1685-1753) *Whatever therefore is got by such exponents and proportions is to be ascribed to fluxions: which must therefore be previously understood. And what are these fluxions? The velocity of evanescent increments? And what are the same evanescent increments? They are neither finite quantities, nor quantities infinitely small, nor yet nothing. May we not call them the ghosts of departed quantities?*

Rappelons aussi la critique mordante que Jonathan SWIFT (1667-1745) fait des mathématiciens dans les "voyages de Gulliver" (1726).

### 3.7 Au XIX<sup>ème</sup> siècle

**Augustin-Louis Cauchy** (1789-1857)

Extrait du cours d'analyse donné à l'École Royale Polytechnique (paru en 1821).

#### PREMIÈRE LEÇON

#### *Des Variables, de leurs Limites et des Quantités infiniment petites*

*On nomme quantité variable celle que l'on considère comme devoir recevoir successivement plusieurs valeurs différentes les unes des autres. On appelle au contraire quantité constante toute quantité qui reçoit une valeur fixe et déterminée. Lorsque les valeurs successivement attribuées à une même variable s'approchent indéfiniment d'une valeur fixe, de manière à finir par en différer aussi peu que l'on voudra, cette dernière est appelée la limite de toutes les autres. Ainsi, par exemple, la surface du cercle est la limite vers laquelle convergent les surfaces des polygones réguliers inscrits, tandis que le nombre de leurs côtés croît de plus en plus ; et le rayon vecteur, mené du centre d'une hyperbole à un point de la courbe qui s'éloigne de plus en plus de ce centre, forme avec l'axe des  $x$  un angle qui a pour limite l'angle formé par l'asymptote avec le même axe ; &c..... Nous indiquerons la limite vers laquelle converge une variable donnée par l'abréviation lim. placée devant cette variable [...]*

*Lorsque les valeurs numériques successives d'une même variable décroissent indéfiniment de manière à s'abaisser au-dessous de tout nombre donné, cette variable devient ce qu'on nomme un infiniment petit ou une quantité infiniment petite. Une variable de cette espèce a zéro pour limite. Telle est la variable  $\alpha$  dans les calculs qui précèdent.*

*Lorsque les valeurs numériques successives d'une même variable croissent de plus en plus, de manière à s'élever au-dessus de tout nombre donné, on dit que cette variable a pour limite l'infini positif indiqué par le signe  $\infty$ , s'il s'agit d'une variable positive ; et l'infini négatif indiqué par la notation  $-\infty$ , s'il s'agit d'une variable négative. Tel est le nombre variable  $m$  que nous avons employé ci-dessus.*

La deuxième leçon est consacrée à la continuité.

#### TROISIÈME LEÇON

#### *Dérivées des Fonctions d'une seule variable*

*lorsque la fonction  $y = f(x)$  reste continue entre deux limites donnée de la variable  $x$ , et que l'on assigne à cette variable une valeur comprise entre les*

deux limites dont il s'agit, un accroissement infiniment petit, attribué à la variable, produit un accroissement infiniment petit de la fonction elle-même. Par conséquent, si l'on pose alors  $\Delta x = i$  les deux termes du rapport aux différences

$$\frac{\Delta y}{\Delta x} = \frac{f(x+i) - f(x)}{i} \quad (3.1)$$

seront des quantités infiniment petites. Mais, tandis que ces deux termes s'approchent indéfiniment et simultanément de la limite zéro, le rapport lui-même pour converger vers une autre limite, soit positive, soit négative. Cette limite, lorsqu'elle existe, a une valeur déterminée, pour chaque valeur particulière de  $x$  ; mais elle varie avec  $x$ . Ainsi, par exemple, si l'on prend  $f(x) = x^m$ ,  $m$  désignant un nombre entier, le rapport entre les différences infiniment petites sera

$$\frac{(x+i)^m - x^m}{i} = mx^{m-1} + \frac{m(m-1)}{1 \cdot 2} x^{m-2} i + \dots i^{m-1}$$

et aura pour limite la quantité  $mx^{m-1}$ , c'est à dire, une nouvelle fonction de la variable  $x$ . Il en sera de même en général ; seulement, la forme de la fonction nouvelle qui servira de limite au rapport  $\frac{f(x+i)-f(x)}{i}$  dépendra de sa forme de la fonction proposée  $y = f(x)$ . Pour indiquer cette dépendance, on donne à la nouvelle fonction de nom de fonction dérivée, et on la désigne, à l'aide d'un accent, par la notation

$$y' \text{ ou } f'(x)$$

La quatrième leçon est consacrée au différentielles des fonctions d'une seule variable ; Cauchy y utilise les notations  $dx$  et  $dy$ .

La sixième leçon s'intitule *Usage des Différentielles et des Fonctions dérivées dans la solution de plusieurs problèmes. Maxima et minima des Fonctions d'une seule variable. Valeur des Fractions qui se présentent sous la forme  $\frac{0}{0}$ .*

Donnons en un extrait: *Déterminer l'inclinaison d'une courbe en un point donné.*

Solution. *Considérons la courbe qui a pour équation en coordonnées rectangulaires  $y = f(x)$ . Dans cette courbe, la corde menée du point  $(x, y)$ <sup>4</sup> au point  $(x + \Delta x, y + \Delta y)$  forme, avec l'axe des  $x$  prolongé dans le sens des  $x$  positives, deux angles, l'un aigu, l'autre obtus, dont le premier mesure l'inclinaison de la*

---

<sup>4</sup>Nous indiquons ici les points à l'aide de leurs coordonnées renfermées entre deux parenthèses, ce que nous ferons toujours par la suite. souvent aussi, nous indiquerons les courbes ou surfaces courbes par leurs équations.

corde, par rapport à l'axe des  $x$ . Si le second point vient à se rapprocher à une distance infiniment petite du premier, la corde se confondra sensiblement avec la tangente menée à la courbe par ce premier point ; et l'inclinaison de la corde, par rapport à l'axe des  $x$ , deviendra l'inclinaison de la tangente, ou ce qu'on nomme inclinaison de la courbe par rapport au même axe. Cela posé, comme l'inclinaison de la corde aura pour tangente trigonométrique la valeur numérique de la limite vers laquelle ce rapport converge, c'est à dire, de la fonction dérivée  $y' = \frac{dy}{dx}$ .

Si la valeur de  $y'$  est nulle ou infinie, la tangente à la courbe sera parallèle ou perpendiculaire à l'axe des  $x$ . C'est ordinairement ce qui arrive, quand l'ordonnée  $y$  devient un maximum ou un minimum.

Exemples.  $y = x^2$ ,  $y = x^3$ ,  $y = x^m$ ,  $y = x^{\frac{2}{3}}$ ,  $y = x^a$ ,  $y = A^x$ ,  $y = \sin x$ ; &c.

### 3.7.1 L'intégrale de Lebesgue

C'est par une note aux Comptes rendus de l'académie des sciences *Sur une généralisation de l'intégrale définie* parue en 1901 que Lebesgue expose cette nouvelle théorie.

Une référence agréable est l'article [Kah] que J-P. Kahane a publié dans la gazette des mathématiciens.

Voici comment Lebesgue explique son point de vue dans une conférence à Copenhague en 1926 *Avec le procédé de Riemann, on sommait les indivisibles dans l'ordre où ils étaient fournis par la variation de  $x$ . On opérerait donc comme le ferait un commerçant sans méthode qui compterait pièces et billets au hasard de l'ordre où ils lui tomberaient sous la main; tandis que nous opérons comme le commerçant méthodique qui dit :*

*j'ai  $m(E_1)$  pièces de une couronne, valant  $1.m(E_1)$   
j'ai  $m(E_2)$  pièces de deux couronnes, valant  $2.m(E_2)$   
j'ai  $m(E_5)$  pièces de cinq couronnes, valant  $5.m(E_5)$   
etc; donc j'ai en tout*

$$S = 1m(E_1) + 2m(E_2) + 5m(E_5) + \dots$$

Pour une fonction  $y = f(x)$  bornée définie sur un intervalle borné, Lebesgue découpe l'intervalle où varie  $y$  et associe à chaque intervalle  $(y_i, y_{i+1})$  la mesure de l'ensemble des  $x$  tels que

$$y_i \geq f(x) \geq y_{i+1}$$

Si cette mesure est  $m_i$ , une valeur approchée de l'intégrale sera

$$\Sigma m_i y_i$$

Pour cela il a besoin de la mesure des ensembles  $y_i \geq f(x) \geq y_{i+1}$ , et de la propriété d'additivité de la mesure. C'est fait (ou plus exactement en projet) par Borel pour les ensembles que l'on appelle aujourd'hui boréliens; il suffit à Lebesgue de compléter cette théorie de la mesure avec les ensembles de mesure nulle.

Quelques applications considérées par Lebesgue :

- Si une fonction  $f$  est dérivable sur un intervalle et si sa dérivée  $f'$  est bornée, la différence

$$f(x) - \int_0^x f'(t) dt$$

est une constante. C'est clair si  $f'$  est intégrable au sens de Riemann ; cependant il existe (Volterra) des exemples où  $f'$  est bornée et n'est pas intégrable au sens de Riemann.

- Série de Fourier. Si une fonction somme d'une série trigonométrique est bornée, elle est intégrable en son sens, et les coefficients de la serie  $f(x) = \Sigma c_n e^{int}$  s'obtiennent par les formules de Fourier :

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{int} dt$$

### Autres prolongements

- analyse fonctionnelle, espaces  $L^2$ ,  $L^p$ . Operateurs et théorie spectrale
- Mesure de Wiener sur un espace de chemins et mouvement brownien.
- etc.

### 3.7.2 Une extension non commutative: le calcul de Fox

(Mort dans les années 1970)

(Une construction algébrique plutôt "à la Leibniz").

Afin de calculer une présentation du groupe fondamental du complémentaire d'un nœud Fox a développé un calcul différentiel dans l'anneau d'un groupe, en particulier dans l'anneau d'un groupe libre à plusieurs générateurs.

On notera  $G$  le groupe et  $JG$  l'anneau (à coefficients dans  $\mathbb{Z}$ ) de ce groupe.

Définissons d'abord le *trivialisateur* de L'anneau  $JG$  comme le morphisme dans  $\mathbb{Z}$  satisfaisant  $\forall g \in G, \mathbf{t}(g) = 1$

**Définition 3.7.1.** Nous appellerons dérivation dans un anneau de groupe  $JG$  une application  $D : JG \rightarrow JG$  qui satisfait

$$D(\nu_1 + \nu_2) = D\nu_1 + D\nu_2 \quad (3.2)$$

$$D(\nu_1\nu_2) = (D\nu_1)\mathbf{t}(\nu_2) + \nu_1 D\nu_2, \quad (3.3)$$

où  $\mathbf{t}$  est le trivialisateur de l'anneau de groupe.

Pour des éléments de  $G$ , la propriété 3.3 prend la forme plus simple

$$D(g_1g_2) = Dg_1 + g_1Dg_2 \quad g_1 \in G \quad g_2 \in G.$$

Quelques conséquences des axiomes satisfaits par une dérivation :

$$D(\Sigma n_i g_i) = \Sigma n_i Dg_i \quad (3.4)$$

$$Dn = 0, \quad \forall n \in \mathbb{Z} \quad (3.5)$$

$$Dg^{-1} = -g^{-1}Dg, \quad \forall g \in G. \quad (3.6)$$

Il est commode de définir

$$\begin{aligned} \frac{g^n - 1}{g - 1} &= 0 \quad \text{si } n = 0 \\ \frac{g^n - 1}{g - 1} &= \Sigma_{i=0}^{n-1} g^i, \quad \text{si } n > 0 \\ \frac{g^n - 1}{g - 1} &= -\Sigma_{i=n}^{-1} g^i, \quad \text{si } n < 0 \end{aligned}$$



# Chapter 4

## Probabilités et statistiques

### 4.1 Jeux, notion de “chances” et “partis”

Les jeux de hasard, d’abord avec un osselet (astragale) puis aussi avec un dé cubique, existent depuis l’antiquité. Les jeux ont servi à effectuer des prédictions dès 3500 av. J.C.

Cependant les premiers témoignages de dénombrement servant à calculer les “chances” apparaissent au début 15<sup>ème</sup> en occident.

Un peu plus tard, Fra Luca des Borgo (Paccioli) publie en 1494 la “Summa de Arithmetica...”. On y trouve un problème de “partis”, c’est à dire de répartition anticipée équitable de l’argent en jeu, lorsque celui-ci ne peut être mené à son terme.

Pour une partie en 6 manches, A en a gagné 5 et B: 3; il propose (solution fausse) de répartir l’enjeu dans la proportion 5 pour A et 3 pour B.

Mentionnons aussi la “loi de la mer” à Rhodes (dès environ 1500). Les premières assurances maritimes supposent en effet, fut-ce de manière intuitive, une estimation de probabilités.

CARDAN (Gerolamo CARDANO, Pavie 1501-Rome 1576) voir *Liber de Ludo Aleae* trouvé après sa mort et publié en 1663 à Lyon.

Il donne des énoncés de chances égales au jeu de dé, avec comme point de départ des chances égales pour chaque point (en un jet) :

- pour qu’un point choisi parmi 3 apparaisse en un jet.

- pour qu’un point apparaisse en trois jets (comparer avec  $1 - (3/5)^3$ , et, à

suivre, avec une espérance) **vérifier**

On lui doit aussi un premier énoncé rudimentaire annonçant la loi des grands nombres : sur une longue période le nombre d’occurrences d’un événement de probabilité  $p$  après  $n$  tentatives est approximativement  $np$ .

**GALILÉE** (1564-1642)

Il remarque que certains nombres (somme des chiffres indiqués par les dés) obtenus en jetant 3 dés sont plus avantageux que d'autres.

- 3 et 18 sont obtenus d'une façon: respectivement 1,1,1 et 6,6,6;
- 4 et 17 sont obtenus de 3 façons (une somme 1+ 1 +2)
- 5 et 16 sont obtenus de 6 façons (deux sommes 1+1+3 et 1+2+2)
- 6 et 15 sont obtenus de 10 façons (trois sommes: 1+2+3; 2+2+2; 1+1+4)
- 7 et 14 sont obtenus de 15 façons (quatre sommes: 1+1+5; 1+2+4; 1+3+3; 2+2+3).
- 8 et 13 sont obtenus de 21 façons (cinq sommes: 1+1+6; 1+2+5; 1+3+4; 2+2+4; 2+3+3)
- 9 et 12 sont obtenus de 25 façons (six sommes: 1+2+6; 1+3+5; 1+4+4; 2+2+5; 2+3+4; 3+3+3)
- 10 et 11 sont obtenus de 27 façons (six sommes: 1+3+6; 1+4+5; 2+2+6; 2+3+5; 2+4+4; 3+3+4)

On vérifie que le total est bien  $216 = 6 \times 6 \times 6$ .

Remarque: les sommes de trois chiffres différents correspondent à 6 façons, celles à deux chiffres différents correspondent à 3 façons, celles où les trois chiffres sont les mêmes à une façon.

**Blaise PASCAL** (1623-1662) Citons Pascal : *J'appelle Triangle arithmétique, une figure dont la construction est telle.*

*Je mène d'un point quelconque,  $G$ , deux lignes perpendiculaires l'une à l'autre,  $GV$ ,  $G\zeta$ , dans chacune desquelles je prends tant que je veux de parties égales et continues, à commencer par  $G$ , que je nomme 1, 2, 3, 4, etc.; et ces nombres sont les exposants des divisions des lignes.*

*Ensuite je joins les points de la première division qui sont dans chacune des deux lignes par une autre ligne qui forme un triangle dont elle est la base.*

*Je joins ainsi les deux points de la seconde division par une autre ligne, qui forme un second triangle dont elle est la base. [...]* Après un repérage un peu compliqué des cellules, Pascal arrive à la règle que nous connaissons sous la forme

$$\binom{n}{p} = \binom{n-1}{p-1} + \binom{n-1}{p}.$$

Une autre manière de construire le triangle est de l'extraire du carré dont la première ligne est formée de 1, la seconde des entiers naturels, La troisième des nombres "triangulaires", somme des  $k$  premiers entiers, la troisième des nombres "pyramidaux", somme des  $k$  premiers nombres triangulaires etc. en choisissant une base diagonale (voir la figure 4.2).

*Remarque.* Le "triangle de Pascal" (que Pascal appelle triangle arithmétique) se

Figure 4.1: Triangle arithmétique

Figure 4.2: Triangle arithmétique

trouve déjà chez Jia Xiang (1<sup>ère</sup> moitié du 11<sup>ème</sup> siècle (Chine, Song du Nord).

Pascal cherche à déterminer comment répartir entre deux joueurs de façon équitable un enjeu si la partie, jouée en plusieurs manches est interrompue avant la fin. *Étant proposés deux joueurs, à chacun desquels il manque un certain nombre de parties pour achever, trouver par le triangle arithmétique le parti qu'il faut faire (s'ils veulent se séparer sans jouer), eu égard aux parties qui manquent à chacun.*

*Soit prise dans le triangle la base dans laquelle il y a autant de cellules qu'il manque de parties aux deux ensemble. Ensuite soient prises dans cette base autant de cellules continues à commencer par la première qu'il manque de parties au premier joueur, et qu'on prenne la somme de leur nombres. Donc il reste autant de cellules qu'il manque de parties à l'autre. Qu'on prenne encore la somme la somme de leurs nombres. Ces sommes sont l'une à l'autre comme les avantages des joueurs réciproquement ; de sorte que si la somme qu'ils jouent est égale à la somme des nombres de toutes les cellules de la base, il en appartiendra à chacun ce qui est contenu en autant de cellules qu'il manque de parties à l'autre; et s'il jouent une autre somme, il leur en appartiendra en proportion. [...]*

La démonstration se fait par récurrence sur les nombres de manches (parties chez Pascal) qui manquent à chaque joueur. En effet imaginons qu'il puissent jouer une manche supplémentaire; deux cas se présentent :

- le premier,  $A$  gagne, il reste donc à  $A$  une manche de moins à gagner, tandis que le deuxième,  $B$  a toujours le même nombre de manche à gagner.
- le deuxième,  $B$  gagne, il reste donc à  $B$  une manche de moins à gagner, tandis que le deuxième,  $A$  a toujours le même nombre de manche à gagner.

Si l'on suppose connue les partis dans les deux situations ainsi obtenues il ne reste plus qu'à faire une moyenne si l'on pense déjà en termes de probabilités de gain pour déterminer les partis dans la situation considérée.

La présentation que fait Pascal du raisonnement est plus progressive. Pascal considère d'abord le cas où il manque à chaque joueur une manche, ou un nombre éga de manches. Puis le cas où il manque une manche à l'un et deux à l'autre. Il considère ensuite les cas où il manque une manche à l'un et 3, 4, 5 manches

à l'autre. Enfin il analyse le cas où il manque deux manches à l'un et trois à l'autre, avant de prouver la proposition citée plus haut.

**Christian HUYGENS** (La Haye 1629-La Haye 1695)

de rationiis in Aleae ludo publié en 1657

exemple de calcul utilisant des probabilités conditionnelles :

(en jouant avec 2 dés **VÉRIFIER**) je gagne si j'obtiens 7 (avant que l'autre n'obtienne la somme 6)

l'autre gagne si il obtient 6 (avant que je fasse 7)

si de plus mon adversaire commence, quelles sont les probabilités pour chacun de gagner.

Il introduit la notion d'espérance mathématique, seulement implicite dans le calcul des "partis" de PASCAL. Rappelons que l'espérance mathématique d'une variable aléatoire discrète  $X$  qui prend les valeurs  $x_i \in \mathbb{R}$  est  $E(X) = \sum_i x_i \mathbb{P}(X = x_i)$ .

## 4.2 Début des probabilités

(titre à revoir)

**Jacques (Jakob) BERNOULLI**<sup>1</sup>

Il vient d'une famille anversoise qui dû t s'exiler lorsque le duc d'Albe fit régner la terreur dans les Flandres qu'il gouverne au nom du roi d'Espagne Philippe II de 1567 à 1573.

Il démontre (Ars conjectandi, publication en 1713): étant données  $n$  variables aléatoires  $X_1, X_2, \dots, X_n$  telles que  $P(X = 1) = p$ ;  $P(X = 0) = 1 - p$  (maintenant appelées *épreuves de Bernoulli*; aujourd'hui nous ajouterions explicitement que ces variables sont indépendantes),

$$\lim P\left(\left|\frac{Y_n}{n} - p\right| < \epsilon\right) = 1$$

où  $Y_n = X_1 + X_2 + \dots + X_n$ .

Il montre aussi que la somme des  $2n$  termes du milieu dans le développement de  $(r + s)^{(r+s) \cdot n}$ ,  $r = 30$ ,  $s = 20$  (même en excluant le terme maximal du milieu)

---

<sup>1</sup>La famille BERNOULLI:

Jacques (Jakob) 1654-1705

Jean (frère de Jacques) 1667-1748

Daniel (fils de Jean) (1700-1782)

Nicolas (fils de Jean) (1695-1726)

Jean (petit-fils du Jean précédent) (1744-1807) (Bâle 1654- Bâle 1705)

Nicolas (neveu de Jacques) (1687-1759)

est plus grande que  $c > 0$  fois la somme des autres termes du développement si le nombre de tirages satisfait:  $nt := n(r + s) \geq 25.500 + 5758 \log(\frac{c}{1000}) = 8226 + 5758 \log(c)$ .

La démonstration de Bernoulli est combinatoire. Maintenant une preuve plus courte utilise l'inégalité de Bienaymé-Tchebytchev.

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq \frac{\text{Var}(X)}{t^2}$$

où  $\text{Var}(X) = \mathbb{E}(|X|^2)$ . En effet comme  $Y_n$  est la moyenne des variables  $X_i$ ,  $(Y_n/n) - p$  est la moyenne des  $n$  variables aléatoires centrées  $X_i - p$  considérées plus haut), la variance de  $Y_n$  est égale à la variance de  $Y_n - p$  qui est  $\frac{1}{n}$ , ce qui démontre le résultat, aussi petit que soit  $\epsilon$ .

### Nicolas BERNOULLI

Après la mort de Jacques, mais avant la publication de “Ars conjectandi”

Il démontre (presque **vérifier**) la loi faible des grands nombres

plus précisément Nicolas BERNOULLI **lequel?** donne une estimation approximative du rapport entre la partie du milieu d’une série binomiale et le reste en utilisant des données dues à Arbuthnot (1712) sur la régularité constante observée concernant les naissances des deux sexes.

*convergence en loi vers la loi de Laplace-Gauss.*

Soit  $X_n$  la variable “gain ou perte du joueur A” après  $n$  jeu de pile ou face, disons pile implique A gagne 1, et face implique A perd 1. L’espérance de  $X_n$  est nulle. La variance de  $X_n$  est  $\sqrt{n}$ ; si l’on pose  $\tilde{X}_n = \frac{1}{\sqrt{n}}X_n$ , la variable  $\tilde{X}_n$  converge en loi vers une variable de loi maintenant appelée de Laplace-Gauss ou  $\mathcal{N}(0, 1)$ . Cela veut dire que si la loi  $G_n = \{\mathbb{P}(\tilde{X}_n = x)\}$  est la distribution de probabilité de la variable  $\tilde{X}_n$ , on a

$$\lim_{n \rightarrow \infty} [\sum_{a < x < b} \mathbb{P}(\tilde{X}_n = x)] = \int_a^b f(t) dt$$

où  $f(t)$  est la fonction “de Gauss” :  $f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ .

Une conséquence (presque une reformulation) est énoncée ainsi par Bachelier dans son traité “le jeu, la chance et le hasard” (publié en 1914) : *quel que soit le jeu considéré, pourvu qu’il soit composé de parties équitables, indépendantes et nombreuses (pas nécessairement identiques), le rapport des écarts qui a une chance sur trois d’être dépassé à l’écart qui a une chance sur deux d’être dépassé est toujours 30/21.*

Bachelier étudie précédemment une suite de parties, chacune d’enjeu 1, et calcule ( $\pm 21$ ) ce qu’il appelle “écart probable” : valeur telle qu’il y ait une chance

sur quatre que le joueur perde plus de 21 ou gagne plus de 21, puis calcule ( $\pm 30$ ) l'écart tel que le joueur a une chance sur six de gagner plus que 30 et une chance sur six de perdre plus que 30.

Un énoncé équivalent utilisant une variable aléatoire de loi binomiale  $\mathcal{B}(n, p)$  est:  $(X_n - np) \left( \frac{1}{\sqrt{npq}} \right)$  converge en loi vers  $\mathcal{N}(0, 1)$ .

### 4.2.1 Le paradoxe de Saint Petersburg

**Nicolas BERNOULLI** (Bâle 1695-Saint-Petersbourg 1726)

Un joueur mise une somme d'argent pour avoir le droit de participer au jeu suivant : On lance une pièce de monnaie, si la pièce tombe sur face, le jeu s'arrête, sinon le joueur gagne un et rejoue. À la deuxième partie le gain éventuel est de 2, et de  $2^n$  à la  $n^{\text{ème}}$  partie, si le jeu se poursuit jusque là. Quelle mise initiale le joueur peut-il accepter ( c'est à dire est telle que l'espérance de gain du joueur, tenant compte de cette mise initiale)? Réponse : il peut accepter n'importe quelle somme, puisque, dans tous les cas l'espérance sera  $+\infty$ . Ceci ne tient pas compte de l'argument de Buffon [Bu] qui tient compte du résultat moral désastreux d'une ruine complète conséquence éventuelle du jeu.

### 4.2.2 Un précurseur

**Abraham de MOIVRE** (Vitry-le-François, 1667-Londres, 1754) (A quitté la France pour l'Angleterre à la révocation de l'édit de Nantes)

Dans "The doctrine of chances" publié en 1718 puis 1738 puis après sa mort en 1756 Il se convainc qu'une loi binomiale tend vers une loi de la forme  $k \cdot e^{-ax^2}$  et remarque (montre, **vérifier** que  $\int_0^\infty e^{-x^2} dx = \sqrt{\pi}$ ).

La "courbe des erreurs" apparait déjà dans son œuvre (publiée en 1738)

*...if after taking a great number of experiments, it should be perceived that the happening and failings at any one time assigned will be very near in that proportion, and that the greater the number of experiments has been, so much nearer the truth will the conjectures be that are derived from them*

*But suppose it should be said, that notwithstanding the reasonableness of building of conjectures upon observations, still considering the great power of chance, events might at long run fall out in a different proportion from the real bent (according to) which they have to happen one way or the other; and that supposing for instance that an event might as easily happen as not to happen, wether after thousand experiments it may not be possible it should have*

Il donne, la même année (1730) que **James STIRLING** ((1692-1770) un

équivalent à l'infini de la factorielle

$$n! \simeq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

**Thomas BAYES** (Londres,1702-Turnbridge Wells (Kent),1761) Étudie en particulier avec de MOIVRE.

Formule de Bayes

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$

ou encore si les  $B_i$  forment un système complet d'événements

$$P(A) = \sum P(A|B_i) \cdot P(B_i)$$

ce qui donne aussi

$$P(B_i|A) = \frac{p(A|B_i) \cdot P(B_i)}{\sum_k P(A|B_k) \cdot P(B_k)}$$

Marie Jean Antoine Nicolas de Caritat, marquis de CONDORCET (1743-1794)

Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix (1785) *Il est nécessaire d'ailleurs de faire attention dans toutes les circonstances à ce minimum de pluralité. En effet il ne suffit point, pour la sûreté, d'avoir une très grande probabilité que l'on ne sera pas jugé d'après un jugement dont la probabilité soit très petite, il faut faire en sorte que cette probabilité soit toujours très grande dans chaque [pour l'ensemble de tous les] jugement particulier.*

## 4.3 Naissance des statistiques

(en partie tiré de Wikipedia)

Bien que le nom de statistique soit relativement récent - on attribue en général l'origine du nom au XVIIIe siècle de l'allemand Staatskunde - cette activité semble exister dès la naissance des premières structures sociales. D'ailleurs, les premiers textes écrits retrouvés étaient des recensements du bétail, des informations sur son cours et des contrats divers. On a ainsi trace de recensements en Chine au XXIIIe siècle av. J.-C. ou en Égypte au XVIIIe siècle av. J.-C. Parallèlement l'amplitude des crues du Nil a donné lieu à des relevés systématiques dès l'Égypte des pharaons. Il faudrait mentionner aussi des recueils de données organisés dès l'antiquité (bible, empire Romain), à l'époque de Charlemagne, dans les cités-états de l'Italie du 13<sup>ème</sup> siècle et en particulier à Venise); le nom

“les offices” du grand musée de Florence vient de ce que ces bâtiments hébergeait une administration devenue importante pendant la renaissance.

Ce système de recueil de données se poursuit jusqu’au XVIIe siècle. En Europe, le rôle de collecteur est souvent tenu par des guildes marchandes, puis par les intendants de l’État. On appelle statistiques descriptives cette collecte et présentation organisée de données.

Ce n’est qu’au XVIIIe siècle que l’on voit apparaître le rôle prévisionnel des statistiques avec la construction des premières tables de mortalité.

On trouve chez Ludwig HUYGENS (frère de Christian voir plus haut “De rationis in Aleae ludo” et calcul différentiel), pionnier des statistiques médicales et démographiques, des tables donnant des espérances de vie (combien de temps peut “espérer” vivre une personne à un age donné.

**John GRAUNT** (né le 24 Avril 1620, à Londres, mort le 18 Avril 1674, à Londres)

Statisticien anglais, généralement considéré comme le fondateur de la démographie (étude statistique des populations humaines). Son analyse des tables de mortalité de la population londonniene ont influencé le travail pionnier de son ami Sir William Petty, et sans doute encore plus important de l’astronome Edmond Halley... (encyclopedia britanica)

**William PETTY** (1623-1687)

William Petty est né à Romsey, Hampshire le 26 May 1623 et est mort à Londres le 16 Décembre 1687. Contacts with Mersenne, John Pell, Sir Charles Cavendish. Études médicales, partisan de Cromwell (travaillera pour lui comme “physician-general”). Importantes contributions aux statistiques, à la démographie et à l’économie. On lui doit par exemple les premières analyses rigoureuses du produit national. En démographie il montre l’absurdité de spéculation supposant un taux d’accroissement constant. Par exemple il projeta la croissance de la ville de Londres : *Si la ville double sa population en 40 ans alors que le nombre de ses habitants est de 6700 et si tout le pays contient 7400 000 habitants et double en 360 ans, il apparaît que’en 1840 la population de la ville sera de 10 718 880 et celle de tout le pays de 10 917 389. D’où il est nécessaire que la croissance de la ville s’interrompe avant pour atteindre son maximum dans la période immédiatement précédente....*

Gregory KING (1648-1712) Estimations de populations à partir du nombre de maisons taxées.

Peut-être peut-on faire commencer l’essor des statistiques à la maison de café Lloyds fondée en 1688 par Edward Lloyd. En effet au 17<sup>ème</sup> siècle Londres était devenu un important centre commercial ce qui à conduit à une demande

importante d'assurance de navires et de leur cargaison. Le café Lloyd's est devenu un endroit connu pour obtenir des assurances maritimes.

Autres précurseurs: PLAYFAIR (1759-1863) ingénieur, spécialiste d'économie politique: représentation graphique de données statistiques diagrammes en bâton, "camemberts"..., (CROME?), BRAVAIS (1811-1863)

\*\*\*\*

On observe à la fois le développement du recueil des données et des tentatives de constructions théoriques permettant des extrapolations.

## 4.4 Analyse et probabilités

On voit apparaître l'utilisation de différentes fonctions associées à la loi d'une variable aléatoire. Le théorème "central limit" (de la limite centrale?) se met en place avec Gauss et Laplace et donne des conditions suffisantes qui assurent qu'une somme finie de variables aléatoires indépendantes, une fois convenablement normalisée, suit approximativement une loi normale.

**Définition 4.4.1.** La fonction génératrice associée à la mesure de probabilité  $P = \sum_{i=0}^{\infty} \alpha_i \cdot \delta_i$  à support dans  $\mathbb{N}$  est  $G_P(s) = \sum_{i=0}^{\infty} \alpha_i \cdot s^i$ .

Remarque En particulier la fonction caractéristique d'une variable aléatoire de loi  $\mathcal{N}(0, 1)$  est  $e^{-t^2/2}$ .

**Pierre Simon de LAPLACE** (1749-1827).

En langage moderne il démontre le théorème: (Théorème de la limite centrale "central limit theorem"): Soit  $X_n$  une suite de variables aléatoires dont le terme général suit la loi binomiale  $\mathcal{B}(n, p)$  ( $n \geq 1, 0 < p < 1, q = 1 - p$ ). On pose  $Y_n = (X_n - np) / \sqrt{npq}$ ; alors la suite  $(Y_n) (N \geq 1)$  converge en loi vers une variable aléatoire de loi  $\mathcal{N}(0, 1)$ . Le cas  $p = 1/2$  avait déjà été étudié par de Moivre.

(Théorie analytique des probabilités (préface 3<sup>ème</sup> édition, 1820))

Il y définit aussi

**Définition 4.4.2.** La fonction caractéristique de la variable aléatoire réelle  $X$  est

$$\varphi_X(t) = E(e^{itX})$$

On démontre deux résultats

1) si une suite de variables aléatoires  $X_n$  converge en loi vers une variable  $X$ , la suite des fonctions caractéristiques  $\varphi_{X_n}$  converge uniformément vers la fonction caractéristique  $\varphi$  de  $X$ .

Il met ensuite la loi normale en rapport avec la théorie des erreurs.

**Théorème 4.4.3.** Soit  $X_n$  une suite de variables aléatoires indépendantes, centrées, du second ordre ( $\text{var}(X) < +\infty?$ ); on pose  $\sigma_k^2 = \text{Var}(X_k)$  et l'on suppose  $0 < \sigma_k < +\infty$ . Introduisons pour tout entier  $n \geq 1$  les variables aléatoires

$$S_n = X_1 + \dots + X_n, \quad Y_n = \frac{S_n}{\sigma(S_n)}, \quad \text{où } \sigma^2(S_n) = \sigma_1^2 + \dots + \sigma_n^2.$$

Pour que  $Y_n$  converge en loi vers une variable aléatoire de loi  $\mathcal{N}(0, 1)$  il suffit que

- 1)  $\sup_{k=1 \dots n} \frac{\sigma_k}{\sigma(S_n)} \rightarrow 0 \quad (n \rightarrow +\infty)$
- 2) Une condition technique portant sur la seconde fonction caractéristique  $\psi_k = \ln(\varphi_k)$ : si l'on pose:  $\psi_k = -\frac{\sigma_k^2 t^2}{2} [1 + \epsilon_k(\sigma_k t)]$ , les  $\epsilon_k(t)$  convergent vers 0 uniformément en  $k$ .

1) Théorème (Paul Levy et Daniel Dugué) Si la suite des fonctions caractéristiques  $\varphi_{X_n}(t)$  converge simplement vers une fonction  $\varphi$  dont la partie réelle est continue à l'origine, alors  $\varphi$  est la fonction caractéristique d'une variable aléatoire  $X$ ...

#### Karl Fredrich GAUSS (1777-1855)

Motivé en particulier par l'utilisation en astronomie de plusieurs mesures pour obtenir (probablement) une meilleure approximation il retrouve la distribution maintenant appelée *distribution de Gauss-Laplace*.

Plus précisément, soit  $\theta$  une quantité inconnue dont  $n$  observations indépendantes ont fourni  $n$  approximations  $x_1, \dots, x_n$ . Gauss propose d'estimer  $\theta$  à partir des valeurs observées.

1) Une première estimation est fournie par la méthode des moindres carrés : on prend comme estimation de  $\theta$  la valeur  $\tilde{\theta}$  qui réalise le minimum de la fonction  $\theta \mapsto \sum_{k=1}^n (x_k - \theta)^2$ ; cette fonction quadratique admet un unique point critique  $\tilde{\theta} = \frac{x_1 + \dots + x_n}{n} = \bar{x}$ .

1) Gauss propose aussi une autre méthode, que l'on appelle aussi *méthode du maximum de vraisemblance*. Introduisons une fonction  $f(x)$  qui représente la densité de probabilité de l'erreur  $x$  commise lors d'une observation; il fait sur  $f$  les hypothèses suivantes (H):

$$f > 0, \quad \int_{-\infty}^{+\infty} f(x) dx = 1;$$

$f$  est paire (les erreurs négatives sont aussi probables que les erreurs positives)

$f(x)$  est décroissante lorsque  $|x| \rightarrow +\infty$  (les grandes erreurs sont moins probables que les petites)

(je l'ajoute; peut-être pas nécessaire, mais simplifie la preuve)  $f$  de classe  $\mathcal{C}^1$ .

Les erreurs correspondant aux valeurs observées  $x_1, \dots, x_n$  de  $\theta$  sont  $x_1 - \theta, \dots, x_n - \theta$ . La densité de probabilité pour que ces erreurs aient été commise simultanément est:

$$L(x_1, \dots, x_n; \theta) = f(x_1 - \theta)f(x_2 - \theta) \cdots f(x_n - \theta)$$

puisque les observations sont indépendantes. On appelle aujourd'hui  $L$  la fonction de vraisemblance.

**Théorème 4.4.4.** *Les seules densités  $f$  vérifiant (H) telles que l'estimation de  $\theta$  par la méthode des moindres carrés coïncide avec son estimation par le maximum de vraisemblance sont les densités normales centrées (lois  $\mathcal{N}(0, \sigma)$ ).*

**Démonstration :** La fonction  $l_n(L)$  s'écrit:

$$l_n(L(\theta)) = l_n f(x_1 - \theta) + \cdots + l_n f(x_n - \theta).$$

On cherche  $f$  vérifiant les hypothèses (H) telle que pour tout  $n$  et toute suite  $(x_1, \dots, x_n)$ , on ait

$$\left[ \frac{d}{d\theta} l_n(L(\theta)) \right]_{\theta=\bar{x}} = 0,$$

soit, en posant  $g = l_n f$ :

$$g'(x_1 - \theta) + \cdots + g'(x_n - \theta) = 0$$

Prenons en particulier  $x_2 = \cdots = x_n = x_1 - ny$  avec  $y$  réel. La moyenne des observations est dans ce cas  $\bar{x} = x_1 - (n-1)y$ ; on a alors  $x_1 - \bar{x} = (n-1)y$  et  $x_2 - \bar{x} = \cdots = x_n - \bar{x} = -y$ . On a donc pour tout  $n$  et tout  $y$

$$g'[(n-1)y] + (n-1)g'(-y) = 0.$$

La fonction  $f$  est paire, la fonction  $g'$  est donc impaire,  $g'(-y) = -g'(y)$  ce qui implique  $g'[(n-1)y] = (n-1)g'(y)$ , où encore, en:

$$\frac{g'[(n-1)y]}{(n-1)y} = \frac{g'(y)}{y}$$

On en déduit que  $\frac{g'(py)}{py} = \frac{g'(qy)}{qy}$ ,  $\forall p \in \mathbb{N}, \forall q \in \mathbb{N}$ , et comme la fonction  $\frac{g'(x)}{x}$  est continue,

$$\frac{g'(x)}{x} = k.$$

On a donc  $g'(x) = kx$ ,  $g(x) = k\frac{x^2}{2} + C$  Comme  $g(x) = l_n(f(x))$  on obtient

$$f(x) = \lambda \cdot e^{k\frac{x^2}{2}}, \lambda > 0$$

Après avoir constaté que  $k$  est nécessairement négatif pour que l'intégrale converge pour une valeur  $C > 0$ , il ne reste plus qu'à choisir la constante  $\lambda$  pour que l'intégrale de  $f$  soit égale à 1.

□

Remarquons qu'un autre choix de  $f$  implique en général que l'estimation  $\theta = \bar{x}$  donnée par la méthode des moindres carrés ne coïncide pas avec celle par le maximum de vraisemblance.

**(Siméon) Denis POISSON** (1781-1840)

Distribution de POISSON : En statistique, la loi de Poisson de paramètre  $\lambda$ , ou loi des événements rares, correspond au modèle suivant :

Sur une période  $T$ , un événement arrive en moyenne  $\lambda$  fois. On appelle  $X$  la variable aléatoire déterminant le nombre de fois où l'événement se produit dans la période  $T$ . La variable  $X$  prend des valeurs entières : 0, 1, 2, ...

Cette variable aléatoire suit une loi de probabilité définie pour tout entier naturel  $k$  par :

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Ce calcul peut se faire de manière déductive en travaillant sur une loi binomiale de paramètres  $(T; \lambda/T)$ . Pour  $T$  grand, on démontre que la loi binomiale converge vers la loi de Poisson.

**Irénée Jules BIENAYMÉ** (1796-1876)

**Pafnouti Lvovitch TCHEBYTCHEV** (1821-1894)

**Inégalité de Tchebytchev** Elle fournit une majoration de  $\mathbb{P}(|X| > t), t > 0$  en fonction du moment d'ordre  $r$ :  $m_r = (\mathbb{E}|X|^r)$ :

$$\mathbb{P}(|X| \geq t) \leq \frac{\mathbb{E}|X|^r}{t^r}$$

Bienaymé a déjà considéré le cas  $r = 2$ .

**Démonstration :** Donnons la preuve pour une variable aléatoire à valeurs discrètes.

$$\begin{aligned} \mathbb{E}|X|^r &= \sum \mathbb{P}(X = x_i) |x_i|^r \\ &= \sum_{|x_i| < t} \mathbb{P}(X = x_i) |x_i|^r + \sum_{|x_i| \geq t} \mathbb{P}(X = x_i) |x_i|^r \\ &\leq \sum_{|x_i| \geq t} \mathbb{P}(X = x_i) |x_i|^r \\ &\leq t^r \cdot \sum_{|x_i| \geq t} \mathbb{P}(X = x_i) = t^r \cdot \mathbb{P}(|X| \geq t) \end{aligned}$$

□

## 4.5 Statistiques

De nouveau, **LAPLACE** (1749-1827)

**Statistiques sur les naissances de garçons et de filles** ([La] Livre deuxième, p 414 et suivantes)

*C'est principalement aux naissances que l'analyse précédente est applicable, et l'on peut en déduire non-seulement pour l'espèce humaine, mais pour toute espèces d'êtres organisés, des résultats intéressants. Jusqu'ici les observations de ce genre n'ont été faites en grand nombre que sur l'espèce humaine ; nous allons soumettre au calcul les principales.*

*Considérons d'abord les naissances observées à Paris, à Londres et dans le royaume de Naples. Dans l'espace des quarante années écoulées depuis le commencement de 1745, époque où l'on a commencé à distinguer à Paris, sur les registres, les naissances des deux sexes, jusqu'à la fin de 1784, on a baptisé, dans cette capitale, 393 386 garçons et 377 555 filles, les enfants trouvés étant compris dans ce nombre ; cela donne à peu près  $\frac{25}{24}$  pour le rapport des baptêmes des garçons à ceux des filles.*

*Dans l'espace des quatre-vingt-quinze années écoulées depuis le commencement de 1664 jusqu'à la fin de 1758, il est né, à Londres, 737 629 garçons et 698 958 filles ; ce qui donne  $\frac{19}{18}$  à peu près, pour le rapport des naissances des garçons à celles des filles.*

*Enfin, dans l'espace de neuf années écoulées depuis le commencement de 1774 jusqu'à la fin de 1782, il est né dans le royaume de Naples, la Sicile non comprise, 782 352 garçons et 746 821 filles ; ce qui donne  $\frac{22}{21}$  pour le rapport des naissances des garçons à celles des filles.*

*[...] et si dans quelques petites villes, où l'on a observé qu'un nombre peu considérable de naissances, la nature semble s'en [en=la supériorité du nombre de naissances des garçons sur les naissances de filles] écarter, il y a tout lieu de croire que cet écart n'est qu'apparent, et qu'à la longue, les naissances observées dans ces villes offriraient, en se multipliant, un résultat semblable à celui des grandes villes. [...] [après un calcul] Il y a donc 328 268 à parier contre un, qu'à Londres la possibilité des baptêmes des garçons est plus grande qu'à Paris. Cette probabilité est tellement proche de la certitude, qu'il y a lieu de rechercher la cause de cette supériorité.*

*Parmi les causes qui peuvent la produire, il m'a paru que les baptêmes des enfants trouvés, qui font partie de la liste annuelle des baptêmes à Paris, devaient avoir une influence sensible sur le rapport des baptêmes des garçons à ceux des filles ; et qu'il devaient diminuer ce rapport, si, comme il est naturel de le croire, les parents des campagnes environnantes, trouvant de l'avantage à retenir près d'eux les enfants mâles, en avaient envoyé à l'hospice des enfants trouvés de Paris, dans un rapport moindre que celui des naissances des deux sexes. C'est ce que le relevé des registres de cet hospice m'a fait voir avec une très-grande probabilité. Depuis le commencement de 1745 jusqu'à la fin de 1809, on y a*

baptisé 163 499 garçons et 159 405 filles, nombre dont le rapport est  $\frac{39}{38}$  : il diffère trop du rapport  $\frac{24}{25}$  des baptêmes des garçons et des filles à Paris pour être attribué au simple hasard.

**William FARR** (1807-1883) Inoculation (vaccin) contre la variole. Un des Bernoulli calcule que la durée moyenne de vie gagne 3 ans avec l'inoculation.

Citer aussi Emmanuel Etienne DUVILLARD de Durand (1745-1832) mathématicien-démographe table de mortalité, "Analyse et tableaux de l'influence de la petite vérole sur la mortalité à chaque âge, et de celle d'un préservatif tel que la vaccine peut avoir sur la population et la longévité" (Paris 1806).

**QUETELET** (1796-1874)

(Wikipedia) Adolphe Quételet est né le 7 février 1796 à Gand et mort le 17 février 1874 à Bruxelles.

Dans "Sur l'homme et le développement de ses facultés, essai d'une physique sociale" (1835), Quételet présenta sa conception de "l'homme moyen" comme valeur centrale autour de laquelle les mesures d'une caractéristique humaine étaient groupés suivant une courbe normale. L'homme moyen est une réalité sui generis, différente des individus. Quetelet s'oppose ainsi au nominalisme de Guillaume d'Ockham. Quetelet distingue trois types de moyennes, reprises par Adolphe Bertillon : la moyenne objective, qui correspond à un objet réel, soumis à un certain nombre de mesures ; la moyenne subjective, résultat du calcul d'une tendance centrale, dans le cas où la distribution obéit à la loi binomiale (courbe de Gauss) ; le troisième cas, qui ne mérite pas vraiment le nom de moyenne, s'appelle moyenne arithmétique, afin de marquer le fait que c'est une pure fiction, n'obéissant pas à la loi binomiale (la hauteur des toits d'une rue), ne justifiant pas la création d'un type moyen.

Ses études sur la consistance numérique des crimes suscitèrent une large discussion entre liberté et déterminisme social. Pour son gouvernement, il rassemblait et analysait les statistiques sur le crime, la mortalité, etc. améliorant le système de recensement. Il apporta des améliorations dans les prises de sanctions. Son travail suscita d'ailleurs une grande controverse parmi les sociologues du XIXe siècle.

On lui doit aussi le système de mesure internationale de l'obésité, connu sous le nom d'indice de Quételet, ou encore Indice de masse corporelle.

**Sir Francis GALTON** (près de Birmingham, 1822-Londres, 1911) coefficient de corrélation, fondateur avec K. Pearson de la revue "biometrika". Motivé par des travaux sur l'origine des espèces. De fait déjà considéré environ 50 ans auparavant par Auguste BRAVAIS<sup>2</sup> (1811-1863) On considère deux variables

---

<sup>2</sup>Auguste Bravais a mené des recherches en astronomie, météorologie, physique, botanique

statistiques  $X$  et  $Y$ . On peut les représenter par un nuage de points dans le plan.

On cherche à résumer par un nombre la forme du nuage de points, plus précisément si la forme du nuage se rapproche d'un segment.

Exemple: hérédité de la taille des graines de petit pois (sweet pea seeds). Galton 1877 Autour de 1875 il conduisait des expériences avec des graines de petit pois (sweet-pea). Il utilisait 100 graines provenant de chacune de 7 classes de graines de différents diamètres. Ensuite il construisait un nuage plan en appariant les diamètres des graines initiales et ceux des graines de la génération suivante. Il a remarqué que le diamètre médian des descendant des graines de grand diamètre était inférieur à celui de leurs parents, tandis que le diamètre médian des descendants des graines de petit diamètre était plus grands que celui de leurs parents. Galton s'est rendu compte que les diamètres des descendants tendaient à se rapprocher du diamètre moyen. Il ne se rendait sans doute pas compte que ses découvertes pouvaient s'appliquer à tout nuage à deux dimensions, pensant plutôt qu'il s'agissait de particularité du phénomène qu'il étudiait. Il appela d'abord ce phénomène "réversion" puis changea le nom pour "régression".

$X$	15	16	17	18	19	20	21
$Y$	15,2	16	15,6	16,25	16	17,2	17,5

*Remarque.* Si l'on veut comparer les points du nuage à des droites, il faut calculer la covariance  $cov(X, Y) = \sigma x_i y_i$ ; on peut d'abord la comparer au produit des écart-types respectifs de  $X$  et  $Y$

$$r(X, Y) = \frac{cov(X, Y)}{s(X)s(Y)}$$

On peut ensuite calculer les droites qui approchent "le mieux" les points du nuage, maintenant encore appelées *droites de régression*, ce que fera Pearson.

**Karl PEARSON** (Londres, 1857- Londres 1936) titulaire de la chaire d'eugénisme! travaille sur la loi du  $\chi_2$ , donne à la méthode de corrélation sa forme moderne.

Les *droites de régressions* minimisent un écart (somme des carrés d'écarts verticaux ou horizontaux) entre une droite et le nuage.

Citer **A.J. LOTKA** (1880-1949) mathématicien-démographe(?) (Wikipedia) Alfred James Lotka (March 2, 1880 - December 5, 1949) était un mathématicien américain (États Unis), aussi physicien-chimiste et statisticien, connu pour son travail sur les dynamiques de population.

Les equations de Lotka-Volterra, aussi appelées équations proie-prédateur, sont deux équations différentielles non-linéaires du premier ordre, souvent utilisées  
et ... cristallographie.

pour construire un modèle dans lequel deux espèces, un prédateur et une proie, interagissent. Ce modèle a été proposé indépendamment par Alfred J. Lotka en 1925 et par Vito Volterra en 1926 [?], voir aussi [?] et [?].

$$\begin{aligned}\frac{dx}{dt} &= x(\alpha - \beta y) \\ \frac{dy}{dt} &= -y(\gamma - \delta x)\end{aligned}\tag{4.1}$$

où

$y$  est l'effectif d'un prédateur (par exemple, des loups);

$x$  est l'effectif des proies (par exemple, des lapins);

$dy/dt$  et  $dx/dt$  représentent la croissance des deux populations en fonction du temps  $t$ ;  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  sont des paramètres associés à l'interaction des deux espèces.

Ces équations ont des solutions périodiques, qui ne sont cependant pas de simples fonctions  $\sin at + b$ .

### Comportement dynamique du système

Avec ce le modèle, les prédateurs se développent quand il y a un grand nombre de proies, mais leur besoins finissent par dépasser la quantité de nourriture disponible ce qui amène la population de prédateur à décliner. Pendant que la population de prédateurs est basse, la population de proie augmente à nouveau. On obtain ainsi un cycle de croissance et décroissance pour chacune des populations.

Équilibre Il faut pour cela que les effectifs des deux populaion restent constants. Pour cela il faut que :

$$\begin{aligned}x(\alpha - \beta y) &= 0 \\ -y(\gamma - \delta x) &= 0\end{aligned}\tag{4.2}$$

En résolvant le système (4.5)on obtient les deux équilibres :

$$x = 0, y = 0$$

et

$$x = \frac{\gamma}{\delta}, y = \frac{\alpha}{\beta}$$

La première solution correspond à l'xtinction des deux espèces. Si à un moment les deux population sont nulles, elles le restent. La seconde solution est un point fixe où les deux populations restent constantes. Les valeurs des population à l'équilibre dépendent des paramètres.

### Stabilité des points fixes

La stabilité de l'origine peut être déterminée en linéarisant le système. Il faudra des méthodes un peu plus fines dans les autres cas.

La matrice jacobienne du modèle proie-prédateur est :

$$J(x, y) = \begin{pmatrix} \alpha - \beta y & -\beta x \\ \delta y & \delta x - \gamma \end{pmatrix} \quad (4.3)$$

**Premier point fixe** Évaluée au premier point fixe  $(0, 0)$ , la matrice jacobienne devient :

$$J(0, 0) = \begin{pmatrix} \alpha & 0 \\ 0 & -\gamma \end{pmatrix} \quad (4.4)$$

Les valeurs propres de la matrice sont :

$$\lambda_1 = \alpha, \lambda_2 = -\gamma$$

Pour ce modèle,  $\alpha$  et  $\gamma$  sont toujours strictement positifs, ce qui force les signes des valeurs propres à être différents. L'origine est donc un point-selle.

Il reste à décider de la stabilité éventuelle de ces points critiques. Un point critique stable va attirer les valeurs des populations. Par exemple, si  $(0, 0)$  est stable des valeurs (en particulier assez petites) des niveau de population vont conduire à l'extinction des deux espèces. Si l'origine est un point de selle, donc est instable, cela rendra l'exDe fait ceci ne peut arriver que si la proie est éradiquée de manière artificielle, ce qui conduit le prédateur à mourir de faim. L'éradication des prédateurs conduit à une croissance sans limite de la population de proies.

### Second point fixe

En évaluant  $J$  au second point fixe, on obtient :

$$J\left(\frac{\gamma}{\delta}, \frac{\alpha\delta}{\beta}, \right) = \begin{pmatrix} 0 & -\frac{\beta\gamma}{\delta} \\ \frac{\alpha\delta}{\beta} & 0 \end{pmatrix} \quad (4.5)$$

Les valeurs propres de cette matrice sont :

$$\lambda_1 = i\sqrt{\alpha\gamma}, \lambda_2 = -i\sqrt{\alpha\gamma}$$

Comme les deux valeurs sont imaginaires pures, ce point fixe n'est pas hyperbolique. Pour conclure, remarquons que le système admet une intégrale première :

$$K = y^\alpha e^{-\beta y} x^\gamma e^{-\delta x}$$

Les courbes de niveau  $K = \text{constante}$  sont des trajectoires fermées qui entourent le point fixe. Par conséquent, les niveaux des populations de proies et de prédateurs décrivent un cycle, c'est à dire chacune oscille autour de la valeur donnée par la coordonnée du point fixe correspondant à l'espèce .

**Francesco Paolo CANTELLI** (Palerme 1875- Rome 1966) Actuaire (mathématiques financières) Chaire à Naples et finalement Rome (jusqu'en 1951)

**Ronald Aylmer FISHER** (Londres 1890-Adélaïde 1962)

Spécialiste de la théorie de l'évolution, il développe des méthodes statistiques pour comparer deux échantillons.

Test de Fisher-Snedecor<sup>3</sup>. Étant donnés deux échantillons de taille  $n_1$  et  $n_2$ , peut-on les admettre dans une même population relativement à la variable étudiée, ces deux échantillons ayant été prélevés indépendamment l'un de l'autre?

**William Sealy GOSSET** (Canterbury 1876-Beaconsfield 1937, connu sous le pseudonyme de STUDENT)

modestie ou contrat avec Guinness?

Lois de distributions des petits échantillons (qui s'appellent maintenant lois de Student).

**Jerzy NEYMAN** (Moldavie 1894-États-Unis 1981)

**Egon Shape PEARSON** (Londres 1895-1980, fils du précédent) Test d'hypothèses, contrôle de qualité.

exemple:

Un test d'ajustement, le test du  $\chi^2$ . Le but est de décider s'il est raisonnable de supposer qu'un échantillon provient d'une variable aléatoire de distribution connue  $F_0$ .

Soit  $X$  une variable aléatoire discrète ou discrétisée, c'est à dire divisée en  $k$  classes de probabilités  $p_1, p_2, \dots, p_k$ , Soit un échantillon de taille  $n$  fournissant les effectifs aléatoires:  $N_1, N_2, \dots, N_k$ . On a  $E(N_i) = np_i$  il est donc naturel de calculer

$$D = \sum_{i=1}^k \frac{(N_i - np_i)^2}{np_i}$$

La déviation  $D$  mesure un écart aléatoire entre les effectifs réalisés et les effectifs espérés. Elle dépend du nombre de termes de la somme, cependant comme  $\sum N_i = n$  il reste  $k - 1$  *degrés de liberté* (termes indépendants).

**Théorème 4.5.1.** *Quand  $n \rightarrow +\infty$ ,  $D^2$  est asymptotiquement distribué comme une variable de  $\chi_{k-1}^2$ , et ceci quelque soit la loi de  $X$ .*

Le test cherche à proposer un choix entre deux hypothèses à l'aide d'une statistique portant sur une répartition en classes de  $n$  valeurs d'une variable aléatoire  $X$ . On compare les valeurs  $p_i$  des proportions à celle  $\mu_i$  attendues à partie de la loi supposée de  $X$ .

---

<sup>3</sup>George Waddell SNEDECOR ((1882-1974); pioneering work in agriculture research, correlation and machine calculation, established a mathematical statistical service in 1927.

$$\begin{aligned} H_0 & F(x) = F_0(x) \\ H_1 & F(x) \neq F_0(x) \end{aligned}$$

On rejettera l'hypothèse nulle  $H_0$  si l'écart  $d^2 = \sum(p_i - \mu_i)^2$  constaté est trop grand, c'est à dire supérieur à une valeur qui n'a qu'une probabilité  $1 - \alpha$  d'être dépassée par une variable de loi une des loi "du  $\chi^2$ ". Le nombre de degré de liberté est  $I - 1$  lorsque le nombre de proportions considérées es  $I$ . Il faut cependant pratiquement supposer que les nombre  $n\mu_i$  ne sont pas trop petits, disons plus grands que 5.

On peut aussi utiliser une loi du  $\chi^2$  pour tester l'indépendance de deux variables. La comparaison se fait avec les probabilités calculées à partir des distribution marginales, en supposant l'indépendance. Si l'on a  $I$  probabilités pour la variable  $X$  et  $J$  probabilités pour la variable  $Y$ , le nombre de degrés de liberté est  $(I - 1)(J - 1)$ .

Plus récemment : **BENZECRI à faire**

## 4.6 Probabilité au 20<sup>ème</sup> siècle

**Émile BOREL**(1871-1956) (loi forte des grands nombres: Convergence presque sûre de  $Y_n = (\sum_{k=1}^n X_k)/n$  vers zéro (proba =1 pour la valeur 0; preuve fausse?), si les variables sont centrées et indépendantes et si de plus elles sont dans  $L^2$  et si la série  $\sum_{n \geq 1} \sigma_n^2/n^2$  est convergente.

Ou encore si les  $X_n$  sont des variables aléatoires indépendantes de moyenne  $\mu$  de lois convenables cf supra

$$\mathbb{P}(\lim_{n \rightarrow +\infty} (\frac{X_1 + X_2 + \dots + X_n}{n} = moy(X_i)) = 1$$

Lemme de Borel-Cantelli (Cantelli 1875-1966) soit  $A_n$  une suite d'événement d'un espace probabilisé

1) si la série  $\sum(\mathbb{P}(A_n))$  converge, la probabilité de voir se réaliser simultanément une infinité d'événements est nulle.

2) si la série  $\sum(\mathbb{P}(A_n))$  diverge, alors la probabilité de voir se réaliser une infinité d'événements est égale à 1.

### 4.6.1 Le mouvement Brownien et autres processus

**Andreï Andreïevitch MARKOV**(Riazan 1856-Petrograd 1922)

**Définition 4.6.1.** On appelle *processus de Markov* une famille de variables aléatoires  $X_t, t \in \mathbb{R}, t \in \mathbb{R}^+$  ou  $t \in \mathbb{N}$  qui vérifie, pour tout  $t < u$ ,  $\mathbb{P}(X_u|X_t) = X_t$ . Lorsque l'indexation se fait sur  $\mathbb{N}$  on parle de chaîne de Markov.

**BROWN** Observation du mouvement Brownien (mouvement erratique d'une particule en suspension dans un liquide du aux chocs moléculaires) (1827).

**Extrait de Wikipedia.**

### Notion de processus stochastique

La difficulté de modélisation du mouvement brownien réside dans le fait que ce mouvement est aléatoire et que statistiquement, le déplacement est nul : il n'y a pas de mouvement d'ensemble, contrairement à un vent ou un courant. Plus précisément :

à un instant donné, la somme vectorielle des vitesses de toutes les particules s'annule (il n'y a pas de mouvement d'ensemble) ; si l'on suit une particule donnée au cours du temps, le barycentre de sa trajectoire est son point de départ, elle « virevolte » autour du même point. Difficile dans ces conditions de caractériser le mouvement... La solution fut trouvée par Louis Bachelier le 29 mars 1900. Il démontra que ce qui caractérise le mouvement, ce n'est pas la moyenne arithmétique des positions  $\langle X \rangle$  mais la moyenne quadratique  $\sqrt{\langle X^2 \rangle}$  : si  $x(t)$  est la distance de la particule à sa position de départ à l'instant  $t$ , alors :

$$\langle X^2(t) \rangle = \frac{1}{t} \int_0^t x^2(\tau) d\tau$$

On démontre que le déplacement quadratique moyen est proportionnel au temps[3] :

$$\langle X^2(t) \rangle = 2dDt$$

où  $d$  est la dimension du mouvement (linéaire, plan, spatial),  $D$  le coefficient de diffusion, et  $t$  le temps écoulé.

Définition mathématique

On peut définir de façon formelle un mouvement brownien : c'est un processus stochastique  $(B_t)_{t>0}$  dont les accroissements disjoints sont indépendants et tels que  $B_{t+s} - B_t$  suit une loi normale de moyenne nulle et de variance  $s$ .

Cette définition permet de démontrer des propriétés du mouvement brownien, comme par exemple sa continuité (presque sûre), le fait que presque sûrement, la trajectoire est différentiable nulle part, et de nombreuses autres propriétés.

On pourrait également définir le mouvement brownien par rapport à sa variation quadratique moyenne. Cette définition, classiquement appelée théorème de Levy, donne la caractérisation suivante: un processus stochastique à trajectoires continues dont la variation quadratique est  $t$  est un mouvement brownien. Ceci

se traduit mathématiquement par le fait que pour une filtration donnée,  $(B_t)_{t>0}$  et  $(B_t^2 - t)_{t>0}$  sont des martingales.

On peut aussi utiliser un modèle de marche aléatoire (ou au hasard), où le mouvement se fait par sauts discrets entre positions définies (on a alors des mouvements en ligne droite entre deux positions), comme par exemple dans le cas de la diffusion dans les solides. Si les  $x_i$  sont les positions successives d'une particule, alors on a après  $n$  sauts :

$$\langle X_n^2 \rangle = \frac{1}{n} \sum_{i=1}^n x_i^2$$

### Marches aléatoires sur $\mathbb{Z}$ , $\mathbb{Z}^2$ , $\mathbb{Z}^3$ et d'autres réseaux.

Marche aléatoire à une dimension d'espace (Exemple)

Considérons la marche aléatoire d'une particule sur l'axe  $Ox$ . On suppose que cette particule effectue des sauts de longueur  $a$  entre deux positions contiguës situées sur le réseau  $\{na, n \in \mathbb{Z}\}$ : de maille  $a$  sur l'axe, chaque saut ayant une durée  $t$ .

Il faut encore se donner un nombre  $p$  tel que :  $0 < p < 1$ . L'interprétation physique de ce paramètre est la suivante :

$p$  représente la probabilité que la particule fasse un saut vers la droite à chaque instant ;  $q = 1 - p$  représente la probabilité que la particule fasse un saut vers la gauche à chaque instant. Le cas du mouvement brownien correspond à faire l'hypothèse d'isotropie spatiale. Toutes les directions de l'espace physique étant a priori équivalentes, on pose l'équiprobabilité :

$$p = q = \frac{1}{2}$$

\*\*\*\*\*

Dans une grille euclidienne  $\mathbb{Z}^2$ , les marches aléatoires reviennent encore presque sûrement. Ce n'est plus le cas à partir de la dimension 3 ( $\mathbb{Z}^3$ ). Ce n'est plus le cas non plus si le graphe qui porte la marche est un analogue du plan hyperbolique, comme par exemple un arbre où trois arêtes partent de chaque sommet.

### **BACHELIER** (1870-1946)

\*\*\*\*(Wikipedia) était un mathématicien français. Il est aujourd'hui considéré comme un précurseur de la théorie moderne des probabilités et comme le fondateur des mathématiques financières. Dans sa thèse intitulée Théorie de la spéculation soutenue le 29 mars 1900, il a introduit l'utilisation en finance du mouvement brownien (découvert par Brown, biologiste), qui est à la base de la plupart des modèles de prix en finance, notamment la formule de Black-Scholes (1973) \*\*\*\*. Il étudie des marches aléatoires dans  $a\mathbb{Z} \subset \mathbb{R}$ ) et montre que si un objet mathématique, "mouvement brownien" existe, sa distribution de

probabilité serait la limite de celles de marches aléatoires de pas de plus en plus petit.

**Albert EINSTEIN**(1879-1955) articles sur le mouvement Brownien 1905 et 1906, qui posèrent les bases théoriques de son étude.

**Jean PERRIN**(1870-1942) Article sur le mouvement brownien 1909, “Les atomes” 1913, rééditions 1924,1935,1939.

\*\*\*(Wikipedia) Après que Albert Einstein a publié (1905) son explication théorique du mouvement Brownien en fonction du mouvement aléatoire des molécules, c’est Perrin qui a effectué les expériences pour vérifier les prédictions d’Einstein. Il a démontré (1908) un accord complet entre théorie et expérience, ce qui a confirmé une fois pour tout l’existence réelle des atomes, proposés un siècle avant par John Dalton, et détermine en même temps une valeur précise du nombre d’Avogadro. \*\*\*\*\*

Quelques citation tirées de “Les atomes”. *L’agitation moléculaire échappe à notre perception directe comme le mouvement des vagues de la mer à un observateur trop éloigné. Cependant, si quelque bateau se trouve alors en vue, le même observateur pourra voir un balancement qui lui révélera l’agitation qu’il ne soupçonnait pas. Ne peut-on de même espérer, si des particules microscopiques se trouvent dans un fluide, que ces particules, encore assez grosses pour être suivies sous le microscope, soient déjà assez petites pour être notablement agitées par les chocs moléculaires?[...]*

*Si séduisante que soit l’hypothèse qui place dans l’agitation moléculaire l’origine du mouvement brownien, c’est cependant encore une hypothèse.[...]*

*L’ÉQUILIBRE STATISTIQUE DES ÉMULSIONS* Un calcul justifiant la répartition d’équilibre dans une colonne gazeuse verticale.

### **Les lois du mouvement brownien**

*Mais de telles évaluations[de la vitesse moyenne d’agitation, de l’ordre de quelques microns par seconde pour des grains de l’ordre du micron] sont grossièrement fausses. Les enchevêtrements de la trajectoire sont si nombreux et si rapides, qu’il est impossible de les suivre et que la trajectoire notée est infiniment plus simple et plus courte que la trajectoire réelle. De même, la vitesse moyenne apparente d’un grain pendant un temps donné varie follement en grandeur et en direction sans tendre vers une limite quand le temps de l’observation décroît, comme on le voit de façon simple, en notant les positions d’un grain à la chambre claire de minute en minute, puis, par exemple, de 5 en 5 secondes, et mieux encore en les photographiant de vingtième en vingtième de seconde, comme ont fait MM. Victor Henri, Commandon ou de Broglie, pour cinématographier le mouvement. On ne peut non plus fixer une tangente, même de façon approchée,*

à aucun point de la trajectoire, et c'est un cas où il est vraiment naturel de penser à ces fonctions continues sans dérivées que les mathématiciens ont imaginées, et que l'on regardait à tort comme de simples curiosités mathématiques, puisque la nature les suggère aussi bien que les fonctions à dérivée.[...] [Afin d'éviter l'effet du champ de gravité, Perrin considère la projection des trajectoire de mouvement brownien sur un plan horizontal] Ceci admis, et absolument sans autre hypothèse, on peut prouver que le déplacement moyen d'un grain devient seulement double quand la durée du déplacement devient quadruple, seulement décuple quand cette durée devient centuple et ainsi de suite. De façon plus précise, on prouve que le carré moyen  $e^2$  du déplacement horizontal pendant une durée  $t$  grandit seulement de façon proportionnelle à cette durée. Il en est de même par suite pour la moitié de ce carré, c'est à dire pour le carré moyen  $X^2$  de la projection du déplacement horizontal sur un axe horizontal arbitraire.

**Norbert WIENER** (Columbia 1894-Stockholm 1964)

Mesure sur l'espace des chemins, définition correcte de  $X(t)$  comprenant l'existence. (Levy parle de fonction aléatoire, pour chaque  $t \in \mathbb{R}$ ,  $X(t)$  est une variable aléatoire). Un point important est de montrer que presque toute trajectoire du mouvement brownien est continue.

Reprenons P. Levy (fonction aléatoire, on dirait aujourd'hui processus stochastique)

**Définition 4.6.2.** Une fonction aléatoire à valeurs réelles de la variable réelle  $t$  est dite *additive* si ses accroissements successifs sont des variables aléatoires indépendantes.

Elle est définie à une constante près si l'on connaît la loi de ses accroissements.

**Définition 4.6.3.** La fonction  $X(t)$  du mouvement brownien linéaire, est le processus stochastique (la fonction aléatoire additive) défini(e) à une constante près par la formule:

$$X(t_1) - X(t_0) = \xi \cdot \sqrt{t_1 - t_0} \quad (t_1 \geq t_0)$$

où  $\xi$  est une variable aléatoire gaussienne réduite, c'est à dire telle que

$$\mathbb{P}(\xi < x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$$

Les propriétés géométriques du mouvement brownien apparaissent dans  $\mathbb{R}^2$ ,  $\mathbb{R}^n$ ,  $n \geq 3$ , et les espaces hyperbolique  $\mathbb{H}^n$ ,  $n \geq 2$ .

**rem** dans  $\mathbb{R}^2$  la distribution de probabilité au temps  $t$  est

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{u_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{u_2^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{\rho^2}{2}}$$

où  $\rho$  est la distance à l'origine.

Dans le plan euclidien, étant donné un ouvert  $v$  quelconque, en particulier une petite boule autour de n'importe quel point, une trajectoire du mouvement brownien passe par  $v$ .

Ceci implique que presque toute trajectoire du mouvement brownien  $X(t)$  telle que  $X(0) = 0$  revient une infinité de fois dans une boule centrée en 0, aussi petite soit-elle.

dimension capacitaire 2 de presque toute trajectoire, mais aire nulle.

Invariance conforme du mouvement brownien plan.

Par contre la propriété de récurrence n'est plus vraie dans l'espace euclidien  $\mathbb{R}^3$  et encore moins dans  $\mathbb{R}^n, n > 3$ .

**Francis PERRIN**(Paris, 1901 - id., 1992) Mouvement brownien sur la sphère Une manière intuitive de voir les trajectoires est de "rouler sans glisser" sur la sphère un plan euclidien sur lequel est tracée à l'encre fraîche une trajectoire du mouvement brownien. A partir des trajectoires du brownien dans le plan, on "imprime" ainsi toutes les trajectoires du brownien sur la sphère. On pourrait procéder de même sur n'importe quelle surface, et en travaillant un peu plus sur n'importe quelle variété riemannienne.

Le même procédé peut être utilisé (en utilisant le modèle de Klein, hyperboloïde dans un espace de Lorentz) pour "voir" les trajectoires du brownien dans le plan hyperbolique. La propriété de retour presque sure dans un voisinage n'est plus vraie dans le plan hyperbolique où presque toute trajectoire du mouvement brownien va en un point à l'infini (un seul point, cependant on ne sait pas lequel) (voir Douady? Sullivan?).

Paul LEVY (points doubles d'une trajectoire...sujet de recherches actuel (nombre de rotation moyen...))

L'utilisation du langage de la théorie de la mesure pour obtenir des résultats classiques de probabilité commence avec STEINHAUS en 1930.

**KOLMOGOROV** (1903-1987)

La monographie de KOLMOGOROV (1933) est la référence pour ce point de vue.

- 1)  $(\Omega, \mathcal{B}(\sigma\text{algèbre}), P \text{ ou } \mu)$
  - 2) les variables aléatoires sont les analogues des fonctions: l'image est dans  $\Omega', P'$ .
  - 3) espérance: intégrale par rapport à la mesure
  - 4) Probabilité, probabilité conditionnelle:  $P(A \cap B)/P(B)$  Espérance conditionnelle d'une variable aléatoire.
- autres processus: processus de contact, (modèle du votant) Spins ...

## 4.7 Entropie

**KOLMOGOROV** (1903-1987)

Soit  $E$  un espace muni d'une mesure  $\mu$  de masse totale 1.

**Définition 4.7.1.** L'entropie d'une partition mesurable  $\mathcal{P} = P_1, P_2, \dots, P_k$  est

$$h(\mathcal{P}) = -\sum p_i \cdot \ln(p_i)$$

où  $p_i$  est la mesure  $\mu(P_i)$ .

C'est un nombre positif ou nul puisque  $\ln(p_i) \leq 0$  pour tout  $i$ . Si l'on fixe le nombre  $k$  de parties, l'entropie de  $\mathcal{P}$  est maximale et vaut  $\ln(k)$  lorsque toutes les parties ont même mesure  $\frac{1}{k}$ .

Soit  $E$  un espace mesuré. Étant donnée une application mesurable  $f : E \rightarrow E$  qui préserve la mesure  $\mu$  définie sur  $E$ , on peut obtenir une suite de partitions  $\mathcal{P}_n$  de  $E$  en considérant les intersections d'images inverses par  $f^\ell$ ,  $\ell = 0, 1, \dots, n$  des ensembles de la partition initiale. Plus précisément soit  $\omega = (i_0, i_1, \dots, i_n)$ , posons  $P_\omega = P_{i_0} \cap f^{-1}(P_{i_1}) \cap f^{-2}(P_{i_2}) \cap \dots \cap f^{-n}(P_{i_n})$ . La famille  $\mathcal{P}_n = \{P_\omega\}$  est une partition mesurable de  $E$ .

**Proposition 4.7.2.** *La limite*

$$h_P(f) = \lim_{n \rightarrow +\infty} \frac{1}{n} h(\mathcal{P}_n)$$

*existe*

On appelle entropie de  $f$  la limite supérieure (prise sur les partitions  $\mathcal{P}$ ) des nombres  $h_P(f)$  est appelée entropie de l'application mesurée de  $f$ .

Exemples d'applications d'entropie nulle: L'identité, une rotation restreinte à la sphère unité munie de la mesure d'aire usuelle divisée par  $\text{volume}(S^n)$ , en particulier une rotation sur le cercle unité, ou ce qui revient au même l'application l'application  $x \mapsto x + \alpha$  dans  $S^1 = \mathbb{R}/\mathbb{Z}$  où le cercle est muni de la mesure image de la mesure de Lebesgue usuelle, le temps 1 d'un champ de vecteurs défini sur une surface muni d'une métrique riemannienne et de la mesure d'aire correspondante (normalisée).

Exemples d'applications d'entropie strictement positive **SINAI, RUELLE, BOWEN**



# Chapter 5

## La géométrie intégrale de Buffon à Poincaré

### 5.1 Buffon

En 1777 Buffon a publié son *Essai d'arithmétique morale* [Bu], où il décrit l'expérience de l'aiguille. Le premier paragraphe de l'essai est :

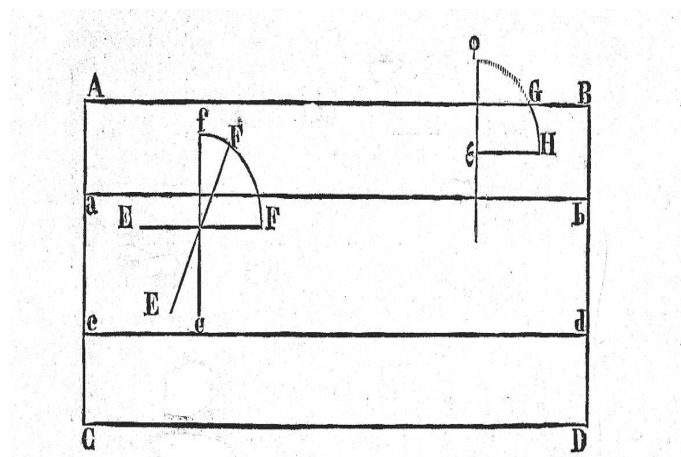


Figure 5.1: Le calcul de Buffon

*La mesure des choses incertaines est ici mon objet : je vais tâcher de donner quelques règles pour estimer les rapports de vraisemblance, les degrés de probabilité, le poids des témoignages, l'influence des hasards, l'inconvénient des risques, et juger en même temps de la valeur réelle de nos craintes et de nos espérances.*

Après quelques considérations à propos d'un jeu appelé "franc-carreau", où les joueurs parient sur la position d'une pièce jetée sur un carrelage (entièrement

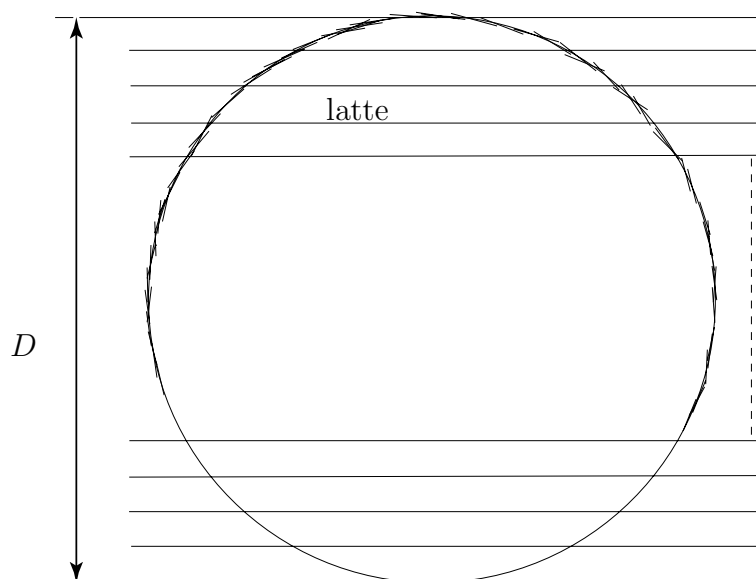


Figure 5.2: Aiguilles et aiguilles replacées

sur un carreau ou à cheval sur un joint), Buffon montre que, quand une aiguille est jetée “au hasard” sur les lattes d’un parquet, quand la longueur de l’aiguille est égale à la largeur des lattes, la probabilité qu’elle tombe à cheval est  $2/\pi$ . Il admet sans l’ombre d’un doute que la bonne mesure de probabilité sur l’espace des positions de l’aiguille est :  $\frac{1}{2\pi}|dx \wedge d\theta|$  que nous rencontrerons de nouveau plus loin.

La présence du nombre  $\pi$  cache un cercle. Le physicien Paul Langevin a décrit en 1908 une manière de visualiser la démonstration du résultat de Buffon. Jetons des milliers d’aiguilles et déplaçons-les en utilisant seulement des translations parallèles aux lattes ou des translations perpendiculaires aux lattes de longueur un multiple entier de la largeur des lattes. Comme toutes les positions relatives (angle, distance des aiguilles aux rainures du parquet) sont également probables, on peut réorganiser les aiguilles autour d’un très grand cercle comme dans la figure 5.2 de façon à ce que la “densité” d’aiguilles soit à peu près constante le long du cercle. Le nombre total d’aiguilles est proche de  $N \cdot L$ , où  $L$  est la longueur du cercle; le nombre d’aiguilles qui coupent les rainures est proche de :

$$N \cdot (\text{nombre de points d'intersection des rainures avec le cercle}),$$

c’est à dire  $2N \cdot D$ , où  $D$  est le diamètre du cercle. La probabilité que nous cherchons est donc  $\frac{2N \cdot D}{N \cdot L} = \frac{2}{\pi}$ .

Il faudra une centaine d’années pour clarifier la notion de probabilité en jeu.

Avant d'en arriver là, donnons une preuve sans doute très proche de celle de Buffon.

repérons l'aiguille sur le parquet à l'aide de la position de sa pointe (il s'agit d'une "vraie" aiguille dont l'autre extrémité est un chas) et de l'angle que fait l'aiguille avec les rainures. En utilisant encore des translations parallèles aux rainures, ou perpendiculaire à celles-ci d'amplitude un multiple entier de la largeur des lattes, nous pouvons supposer que la pointe de l'aiguille se trouve sur le segment vertical  $AB$  de la figure (5.3). supposons que ce segment  $AB$  est de longueur 1. Notons  $x$  La distance entre l'extrémité de la'aiguille et  $A$ . Therefore

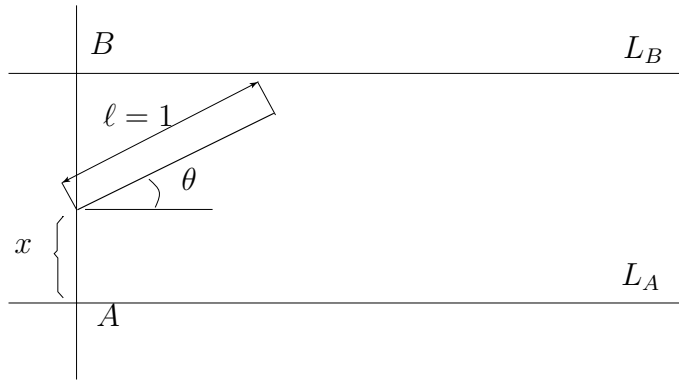


Figure 5.3: Localisation de l'aiguille

the set of all possible positions of the needle is  $[AB[.S^1$ , (or rather  $\mathbb{R}/\mathbb{Z} \times S^1$ ). The needle meets the line  $L_B$  if  $x + \sin\theta \geq 1$  and  $L_A$  if  $x + \sin\theta \leq 0$ .

The ratio between the dashed area and the area of the rectangle  $]0, 1[.]0, \pi[$  is  $2/\pi$ .

### 5.1.1 The birth of the notion of geometric measure

#### Cauchy and Crofton

In 1832, in a communication to the French Academy of Sciences, Cauchy noticed that the length of a convex curve is proportional to the average of the lengths of the orthogonal projection of the convex curve on all lines through the origin. More generally, for any rectifiable planar curve  $C$ , denote by  $m(C,L)$  the "absolute length" of the orthogonal projection of  $C$  on the line  $L$ , the length of the projection counted with multiplicity. In modern language:

$$m(C, L) = \int \text{card}(p^{-1}(y)dy) ; y \in L$$

Then:

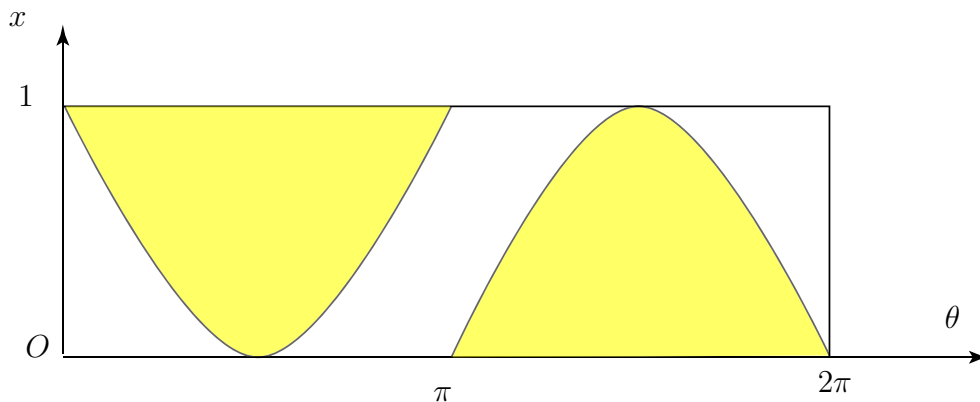


Figure 5.4: aire dans  $[0, 1] \times [0, 2\pi]$  de l'ensemble des aiguilles qui traversent une rainure

**Théorème 5.1.1.** *Cauchy formula [Cau]*

$$\int_{-\pi/2}^{\pi/2} m(C, L_\theta) d\theta = 2(\text{length of } C)$$

Cauchy's proof amounted to prove the formula for a segment, and then approximate any curve by inscribed polygons.

From Cauchy's communication to the french academy in 1832 [Cau], to Crofton's mémoire (1868) [Cro], 36 years were needed to clarify the notion of a measure on the set of affine lines. Let us quote Crofton: *The expression "at random" has in common language a very clear and definite meaning; one which cannot be better conveyed than by Mr Wilson's expression "according to no law"... There is always a direct reference to the assemblage of things to which it belongs and from which we take, and not till then, we can proceed to sum up the favorable cases,... But there are several classes or questions in which the totality of cases is not merely infinite, but of an inconceivable nature... We can thus continually suppose variations of the experiment, each variation giving a new infinity of cases (then Crofton justifies the choice of the measure on the plane). What means: an infinity of lines drawn at random on the plane, what is the nature of this aggregate? First, since any direction is as likely as the others, as many of the lines are parallel to any direction as to any other. As this infinite system of parallels is drawn at random, they are as thickly disposed along any part of the perpendicular as along any other...*

Crofton did find the right answer as we will see in next section. Nevertheless, at the turn of the century, the choice of a measure on a continuum was not obvious, because there were too many possibilities.

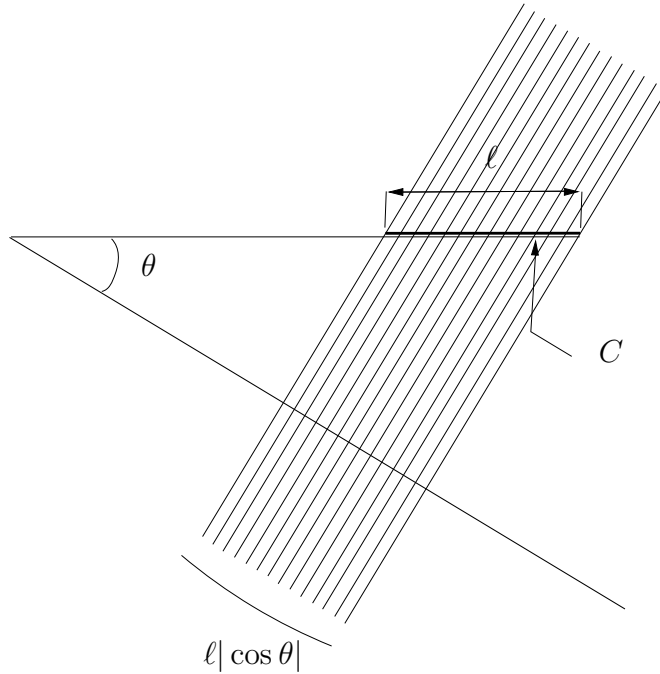


Figure 5.5: mesure des droites qui coupent un segment et sont orthogonales à une direction donnée

### Bertrand's paradoxes

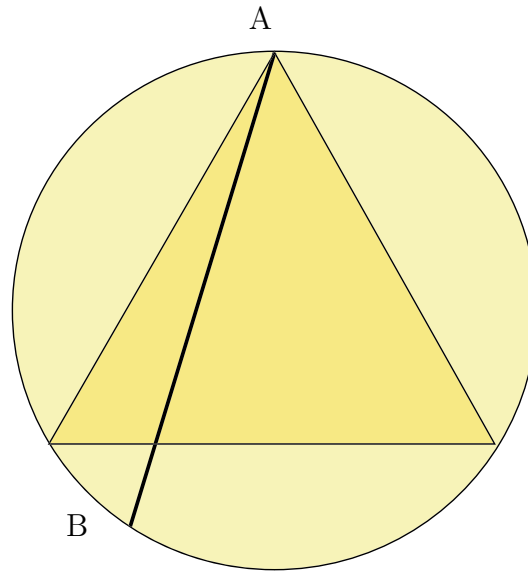
Let us give three different answers proposed by the probabilist Bertrand to the same problem of elementary geometry. At that point, integral geometry was close to disappear. The question is (see pictures below): what is the probability for a chord of a circle taken at random to be longer than the side of an equilateral triangle inscribed in the circle? The three different answers Bertrand proposes will come from three different ways to choose the chord.

1) Chose an arbitrary point  $A$  on the circle. Using the rotational symmetry of the picture we can forget about  $A$  and choose now another point  $B$  on the circle, endowed with arc length measure. The chord is then longer than the side of the inscribed equilateral triangle with probability  $1/3$ .

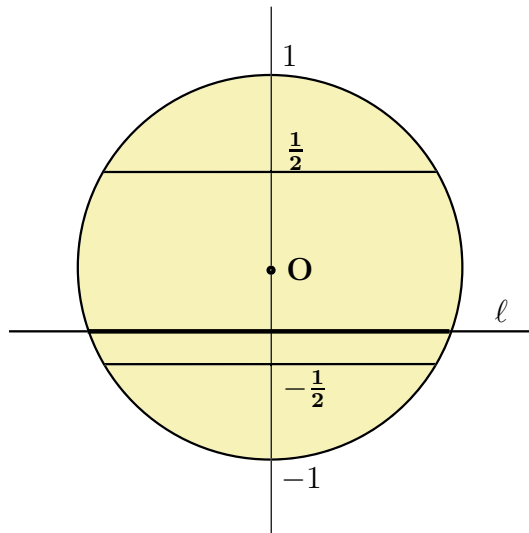
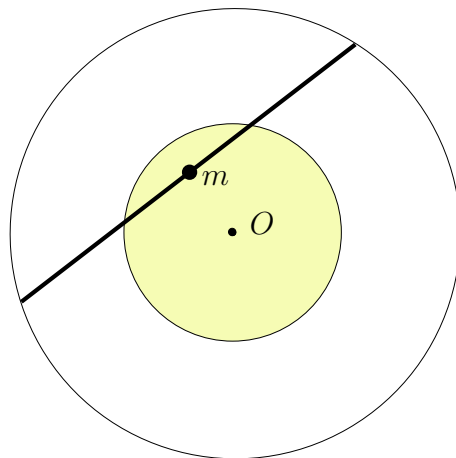
2) Chose at random the affine line supporting the chord. The rotational symmetry of the picture allows us to forget about the direction of the line. As  $\cos(\pi/3) = 1/2$ , the probability is now  $1/2$ .

3) Chose at random the middle of the chord in the disc (the measure is the Lebesgue measure on the disc). We ignore chords through the origin, as they form a set of measure zero. Then the probability is  $1/4$ .

Poincaré will take integral geometry out of this dead end. For him, (see

Figure 5.6: Probabilité  $1/3$ 

for example his book published in 1912 [Poin]), the most interesting measure is the one which is invariant under the group of affine isometries of the plane. Only isometries preserving the origin are allowed by presentations 1) and 3). In presentation 2) translations also act on the set of affine lines and preserve our measure. It was also Crofton's answer.

Figure 5.7: Probabilité  $1/2$ Figure 5.8: Probabilité  $1/4$



# Chapter 6

## La courbure des courbes planes et gauches

D'après l'histoire de la géométrie de Proclus (412-486) Aristote (384-322 av. J. C.) voyait trois sortes de courbes : droites, circulaires et mixtes. Un siècle après environ Apollonius de Perge (v. 262-v. 180 av. J.-C.) montre qu'en chaque point d'une conique il y a exactement une normale.

Il faut attendre ensuite Nicole Oresme, au 14<sup>ème</sup> siècle pour proposer que la courbure d'un cercle soit proportionnelle à l'inverse du rayon. Il compare ensuite les courbures de courbes convexes tangente à l'aide de l'inclusion des domaines qu'elles bordent localement

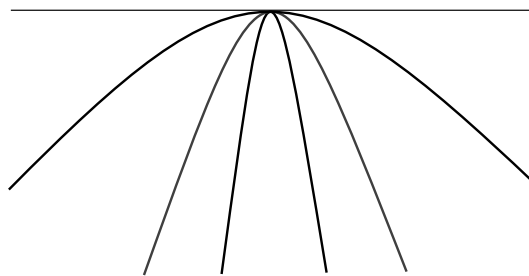


Figure 6.1: Ordre entre les courbures de courbes planes

La notion de cercle osculateur apparait chez Kepler(1571-1630) lorsque, généralisant le problème d'Alhazin ( $\simeq 1000$ , réflexion sur un cercle) il considère la réflexion sur une courbe plane quelconque.

Il reviendra à Cristiaan HUYGENS d'étudier les développées (enveloppe des normales) d'une courbe plane, et l'opération inverse : la construction d'une développante en imaginant la trajectoire d'un point extrémité d'une ficelle tendue progressivement décollée de la courbe. Il le fait dans "Horologium oscillato-

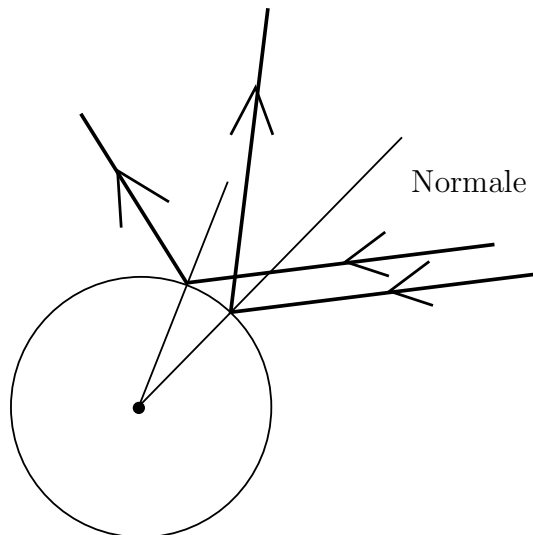


Figure 6.2: Réflexion sur un cercle

rium sive de motu pendulorum ad horologia aptato demonstrationes geometrica” (1673). En effet une courbe: la cycloïde permet de construire un pendule dont la période ne depend pas de l’amplitude du mouvement, contrairement au pendule ordinaire. Le fait qu’un pendule amarré au point singulier de la cycloïde, où l’on suppose la tangente verticale ai une longueur non constante compense exactement, pour cette courbe l’effet de l’amplitude sur la période. Le mouvement de l’extrémité du balancier est une developpante de la cycloïde.

Avant de présenter les démonstrationde HUYGENS, donnons une preuve du fait que l’enveloppe des normales à une courbe (pour la quelle on supposera que la courbure et la dérivée de la courbure sont non nulles) est le lieu des centres des cercles osculateurs de la courbe.

Notons  $T(m) = \frac{dm}{ds}$  le vecteur unitaire tangent à la courbe  $C$ , où  $s$  est un paramétrage de  $C$  par la longueur de l’arc  $m_0m \subset C$ ,  $N(m)$  le vecteur normal unitaire  $N(m) = \mathcal{R}_{\pi/2}(T(m))$ , et  $k(m)$  la courbure de la courbe  $C$  en  $m$  définie par  $\frac{dT}{ds} = k(m) \cdot N(m)$ .

Les droite  $D(m)$  ont une enveloppe dont les points sont obtenus en considérant l’intersection de  $D(m)$  et de la droite  $D'(m)$  obtenue en dérivant l’équation de  $D(m)$  par rapport au paramètre  $s$ . L’équation de la normale  $D(m)$  est  $D(m) \cdot \langle \overrightarrow{mM} | T(m) \rangle = 0$ . L’équation de la droite dérivée est :  $D'(m) \cdot \langle T(m) | T(m) \rangle + \langle \overrightarrow{mM} | k(m) \cdot N(m) \rangle = 0$  Remarquons que la droite  $D'(m)$  est orthogonale à  $D(m)$  et coupe  $D(m)$  en un point  $F(m)$  à distance  $R(m) = \frac{1}{k(m)}$  de  $m$ , qui est le centre du cercle osculateur à la courbe en  $m$ .

Il reste à vérifier que la famille  $D(m)$  vérifie aussi la condition garantissant

que l'ensemble des points  $F(m)$  forme une courbe ayant en chaque point comme tangente la droite  $D(m)$ .

Pour cela il faut dériver une fois de plus et vérifier que la droite obtenue  $D''(m)$  coupe les deux droites  $D(m)$  et  $D'(m)$  en un point différent de  $F(m)$ .

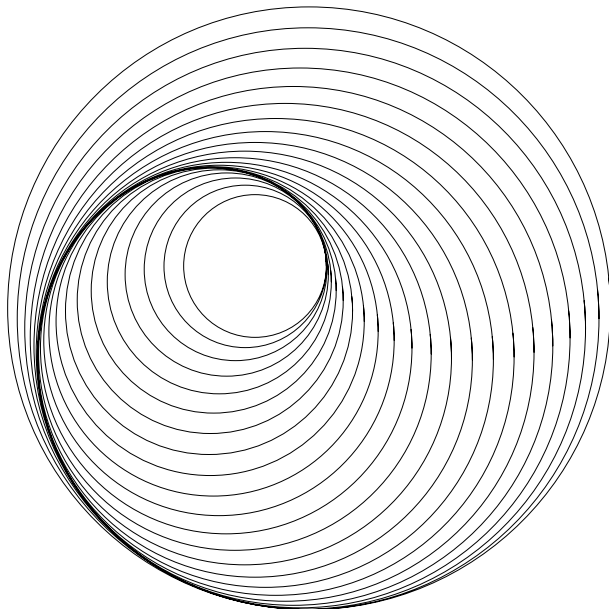


Figure 6.3: Cercles osculateurs à un arc de courbe

Suivons maintenant HUYGENS (Horologium oscillatorum, partie III).

*Proposition 1*

*Les tangentes à la développée coupent la développante à angle droit.*

démonstration : voir [Str] p. 264

*Proposition 2*

*Tout arc de courbe concave tel que  $ABD$  peut être divisé en assez de partie de sorte que si l'on trace les cordes qui sous-tendent chacun des arcs,  $AB, BC$ , et  $CD$  par exemple, et si l'on trace les tangentes  $AN, BO, CP$*



# Chapter 7

## La courbure (extrinsèque et intrinsèque) des surfaces.

Les courbures principales d'une surface ont sans doute été d'abord définies par Euler [Eu1]. Monge les retrouve par une méthode du type "enveloppe". Nous allons cependant d'abord présenter le point de vue de Gauss, qui est repris dans les livres actuels de géométrie.

### 7.1 L'application de Gauss et les courbures principales.

Commençons par considérer des courbes planes, que nous supposerons de classe  $\mathcal{C}^\infty$ . L'application de Gauss  $\gamma$  associe à chaque point  $m$  d'une courbe orientée  $C \subset \mathbb{R}^2$  sa normale orientée,  $N(m) = R_{\pi/2}(T(m))$  où  $T(m)$  est le vecteur tangent unitaire orienté à la courbe en  $m$ .

$$\gamma : C \rightarrow S^1$$

Le jacobien  $k(m)$  de  $\gamma$  au  $m \in C$  est appelé courbure de  $C$  en ce point.

Remarquons aussi que l'application tangente  $T : C \mapsto S^1$  qui envoie un point  $m \in C$  sur le vecteur unitaire tangent en  $m$  déterminé par l'orientation de  $C$ , a le même jacobien  $k(m)$ .

Considérons maintenant une surface plongée  $M \subset \mathbb{R}^3$ . Au moins si elle est sans bord, elle est orientée, car on peut choisir le vecteur normal  $N(m)$  en  $m \in M$  qui pointe vers l'extérieur de la composante bornée de  $\mathbb{R}^3 \setminus M$ . L'application de Gauss est alors:

$$\gamma : M \rightarrow S^2, m \mapsto N(m)$$

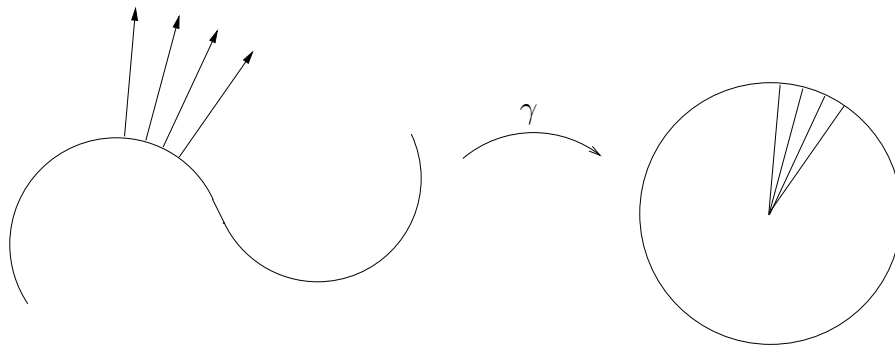


Figure 7.1: L'application de Gauss en dimension 1

La *courbure de Gauss*  $K(m)$  en  $m \in M$  est le jacobien en  $m$  de l'application de Gauss.

Les valeurs propres de  $d\gamma(m)$ :  $k_1, k_2$  (elles peuvent coïncider) sont appelées les *courbures principales* de  $M$  en  $m$ . À chacune correspond une direction propre. Il nous restera à démontrer que ces directions propres sont orthogonales.

Sur un voisinage d'un point où les directions principales sont différentes, on peut intégrer les deux champs de droites et obtenir ainsi deux familles de courbes: les *lignes de courbure*. Géométriquement, lorsque l'on suit une ligne de courbure, le plan tangent ne fait que rouler. Lorsque l'on suit une courbe quelconque son mouvement est un mélange de roulis et de tangage. Sur un domaine où la courbure de Gauss  $K(m)$  est strictement négative, il existe deux directions: les *directions asymptotiques* le long desquelles le plan tangent ne fait que rouler. Ce sont les directions engendrées par les vecteurs  $v$  qui satisfont:  $\langle d\gamma(v)|v \rangle = 0$ . ce sont aussi les directions tangentes à l'intersection de la surface avec son plan tangent.

La seconde forme fondamentale  $II(m)$  est définie par :

$$II(m)(v) = \langle d\gamma(m)(v)|v \rangle$$

La matrice, symétrique, de la seconde forme fondamentale peut, elle, être diagonalisée à l'aide d'une base orthonormée. Il reste cependant à vérifier que les vecteurs d'une telle base déterminent bien les directions propres de  $d\gamma$ .

Étudions maintenant la surface  $M \subset \mathbb{R}^3$  au voisinage du point  $m$  à l'aide d'une équation:

$$z = a \cdot x^2 + 2b \cdot xy + cy^2 + o(x^2 + y^2)$$

où les coordonnées  $(x, y, z)$  sont associées à une base orthonormée, dont les deux premiers vecteurs sont contenus dans le plan tangent  $T_{m_0}M$ , tandis que le troisième est normal au plan tangent.

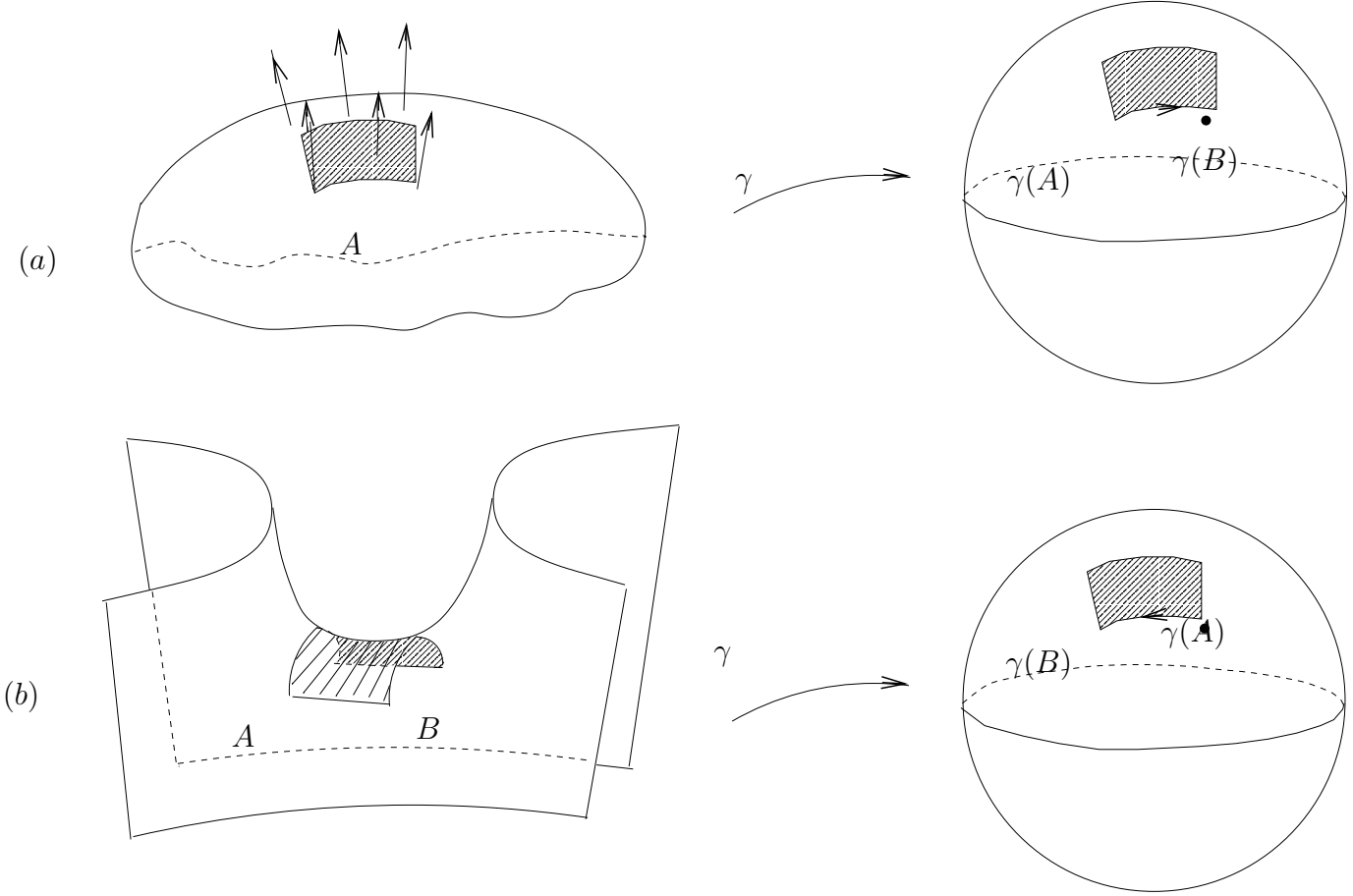


Figure 7.2: L'application de Gauss

Si, de plus on choisit comme deux premiers vecteurs de notre base deux vecteurs propres de la forme quadratique:  $a \cdot x^2 + 2b \cdot xy + cy^2$ , l'équation précédente devient (nous démontrerons plus loin que les coefficients  $k_i$  sont bien les valeurs propres de  $d\gamma$ ):

$$z = \frac{1}{2}[k_1 \cdot x^2 + k_2 \cdot y^2] + o(x^2 + y^2). \quad (7.1)$$

Justifions maintenant le nom "courbures principales" donné aux coefficients  $k_i$ . Remarquons d'abord que l'équation d'une sphère  $S_R$  de rayon  $R$  tangente à l'origine au plan horizontal est:  $(z - R)^2 + y^2 + x^2 = R^2$ . Remarquons que, lorsque  $x$ ,  $y$ , et  $z$  sont petits,  $z$  est de l'ordre de  $x^2 + y^2$ . On peut donc écrire:

$$z = \frac{1}{2}\left[\frac{1}{R}x^2 + \frac{1}{R}y^2\right] + o(x^2 + y^2) \quad (7.2)$$

Ceci nous permet de calculer la différence entre la coordonnée  $z$  d'un point sur la surface  $M$  et la coordonnée  $z_R$  d'un point près de l'origine  $m$  sur la sphère  $S_R$  d'équation  $(z - R)^2 + y^2 + x^2 = R^2$ , situés au dessus du même point de

coordonnées  $(x, y)$  du plan horizontal:

$$z - z_R = \frac{1}{2}[(k_1 - \frac{1}{R}) \cdot x^2 + (k_2 - \frac{1}{R}) \cdot y^2] + o(x^2 + y^2).$$

La partie quadratique de cette différence est dégénérée quand le rayon  $R$  prend les valeurs  $R = R_1 = \frac{1}{k_1}$  où  $R = R_2 = \frac{1}{k_2}$ . Les deux sphères correspondantes  $S_1$  et  $S_2$  sont appelées sphères osculatrices en  $m \in M$ .

On peut aussi reconnaître géométriquement les sphères osculatrices parmi les sphères du pinceau tangent à l'origine au plan horizontal (tangent à  $M$  en  $m$ ). Les équations de l'intersection de la sphère  $S_R$  avec  $M$  sont:

$$\frac{1}{2}[(k_1 - \frac{1}{R}) \cdot x^2 + (k_2 - \frac{1}{R}) \cdot y^2] + o(x^2 + y^2) = z - z_R = 0. \quad (7.3)$$

et

$$z = \frac{1}{2}[\frac{1}{R}x^2 + \frac{1}{R}y^2] + o(x^2 + y^2) \quad (7.4)$$

Lorsque la courbure  $k = \frac{1}{R}$  de la sphère  $S_R$  est comprise entre  $k_1$  et  $k_2$ , cette intersection est, au voisinage de l'origine, l'union de deux arcs qui se coupent transversalement en l'origine. L'équation des tangentes à l'origine à ces arcs est donnée par la partie quadratique de l'équation 7.3:

$$\frac{1}{2}[(k_1 - \frac{1}{R}) \cdot x^2 + (k_2 - \frac{1}{R}) \cdot y^2] = 0$$

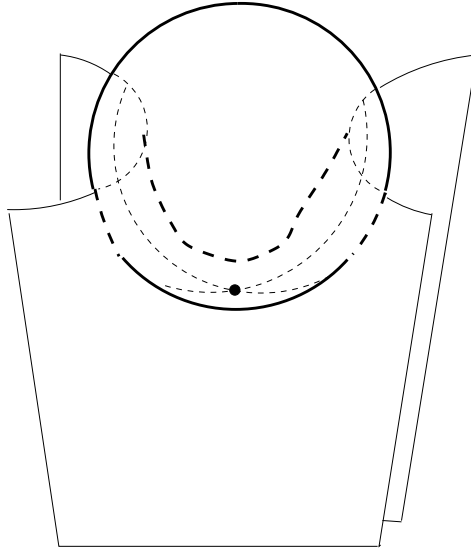


Figure 7.3: Contact de type selle entre une sphère et une surface

Lorsque  $k \notin [k_1, k_2]$  l'intersection de  $S_R$  et  $M$  au voisinage de l'origine se réduit à un point, l'origine.

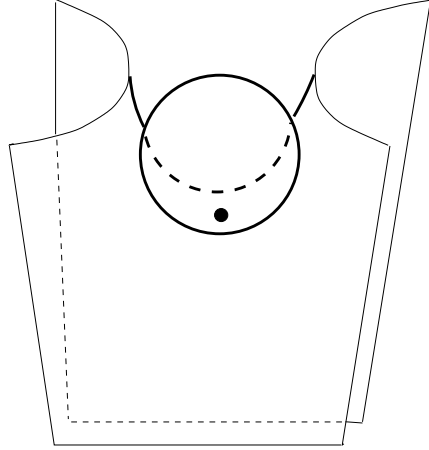


Figure 7.4: Contact de type centre entre une sphère et la surface

Les deux sphères du pinceau restantes (quand  $k_1 \neq k_2$ ) sont les sphères osculatrices.

Quand  $k = k_1, k_1 \neq k_2$ , génériquement, c'est à dire quand le terme  $o(x^2 + y^2)$  de l'équation:

$$\frac{1}{2}[(k_1 - \frac{1}{R}) \cdot x^2 + (k_2 - \frac{1}{R}) \cdot y^2] + o(x^2 + y^2) = z - z_R = 0. \quad (7.5)$$

contient un terme en  $x^3$ , cette intersection a localement une arête de rebroussement de première espèce, tangente à l'axe des  $x$ , puisque les termes principaux qui restent dans l'équation 7.5 quand  $\frac{1}{R} = \frac{1}{R_1} = k_1$  sont:

$$(k_2 - \frac{1}{R_1}) \cdot y^2 + ax^3; a \neq 0$$

Implicitement, en choisissant l'orientation de l'axe des  $z$ , nous avons défini une local orientation locale de la surface, celle telle que le repère de  $\mathbb{R}^3$  obtenu en ajoutant l'axe des  $z$  à un repère positif du plan tangent à la surface donne l'orientation de l'espace ambiant  $\mathbb{R}^3$ .

Nous avons justifié le nom "courbure", nous n'avons pas encore démontré que ces courbures sont les valeurs propres de  $d\gamma$ . Pour cela il faut écrire le vecteur normal en fonction de l'équation 7.1 de la surface au voisinage de  $m$ . En notant:

$$F(x, y, z) = z - \frac{1}{2}[k_1 \cdot x^2 + k_2 \cdot y^2] + o(x^2 + y^2)$$

on a:

$$N(m) = \frac{\text{grad}(F)}{|\text{grad}(F)|} = [1/(1 + (\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2)^{1/2}] \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ 1 \end{pmatrix}$$

En supposant  $f$  de classe suffisante, on voit qu'à un  $o(\sqrt{x^2 + y^2})$  près,  $N(m) = \begin{pmatrix} 1 \\ k_1 x \\ k_2 y \end{pmatrix} + o(\sqrt{x^2 + y^2})$ .

En supposant que les dérivées partielles secondes de  $F$  sont continues, la matrice de  $d\gamma$  en  $m = (0, 0, 0)$  est alors:

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Cette dernière matrice est appelée *hessienne* de  $f$ . Au point  $m$ , il reste:

$$d\gamma(m) = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} \quad (7.6)$$

En résumé:

**Proposition 7.1.1.** • *La courbure de Gauss de  $M$  en  $m$  est le produit des deux courbures principales  $K(m) = k_1(m) \cdot k_2(m)$ .*

- *La courbure de Gauss est aussi le déterminant de la seconde forme fondamentale.*
- *En choisissant une équation locale de la surface  $M$  où  $(x, y)$  sont des coordonnées orthonormées du plan tangent  $T_m$  et  $z$  une coordonnée portée par le vecteur normal  $N(m)$  qui complète  $(x, y)$  en un système de coordonnées orthonormées de l'espace:*

$$z = f(x, y) = \frac{1}{2}[k_1 \cdot x^2 + k_2 \cdot y^2] + o(x^2 + y^2)$$

*Les directions propres de la différentielle  $d\gamma(m)$  sont les directions propres de la partie quadratique de  $z = f(x, y)$ , et sont donc orthogonales.*

## 7.2 Euler

Cependant Euler commence par chercher à déterminer les rayons des cercles osculateurs à toutes les sections planes d'une surface  $M$  passant par un point  $m \in S$ . Plutôt que d'effectuer, comme Euler, un calcul en prenant des coordonnées telle que simplement  $M$  soit un graphe au voisinage du point  $m$ , prenons, comme le fera Gauss, des coordonnées  $(x, y, z)$  telles que  $m$  soit l'origine et le plan engendré par  $(y, z)$  le plan tangent à la surface en  $m$ . Soit, dans ce système,  $z = F(x, y)$  l'équation locale de la surface.

Une section  $C_\varphi$  de la surface par un plan vertical  $H_\varphi$  contenant l'origine, repéré par les coordonnées  $(\rho, z)$  est:

$$z = f_\varphi(\rho) = F(\rho \cos \varphi, \rho \sin \varphi),$$

où  $\varphi$  est l'angle du plan de section  $H_\varphi$  avec le plan vertical contenant l'axe des  $x$ .

Le rayon de courbure de cette section est  $R_\varphi = \frac{1}{d^2z/d^2\rho}$ . En effet un cercle tangent en  $m$  à la section verticale  $C_\varphi$  est de la forme  $(z - R)^2 + \rho^2 = R^2$ , où  $R \neq 0$ ; l'équation de ce cercle peut se mettre sous la forme  $z = f_R(\rho) = \frac{1}{2R}\rho^2 + o(\rho^2)$ . Parmi les cercles tangent en  $m$  à  $C_\varphi$ , un seul est tel que la différence  $f - f_R$  soit un  $o(\rho^2)$ : c'est le cercle osculateur. Le rayon de ce cercle est  $R_\varphi$ .

Ce rayon se calcule en fonction des dérivées partielles secondes de  $z$  par rapport à  $x$  et  $y$ . Il est plus commode de travailler avec la courbure  $k_\varphi = \frac{1}{R_\varphi}$ .

$$k_\varphi = \cos^2 \varphi \frac{\partial^2 z}{\partial x \partial x} + 2 \cos \varphi \sin \varphi \frac{\partial^2 z}{\partial x \partial y} + \sin^2 \varphi \frac{\partial^2 z}{\partial y \partial y}$$

Les extrema de  $k_\varphi$  sont obtenus lorsque  $-\sin 2\varphi \frac{\partial^2 z}{\partial x \partial x} + 2 \cos 2\varphi \frac{\partial^2 z}{\partial x \partial y} + \sin 2\varphi \frac{\partial^2 z}{\partial y \partial y} = 0$ . Cela donne deux plans verticaux orthogonaux pour lesquels la courbure  $k_\varphi$  (et donc aussi le rayon de courbure  $R_\varphi$ ) sont extrema (en fait on obtient un maximum et un minimum dès que ces deux valeurs sont différentes).

### 7.2.1 Monge

Le point de vue de Monge est un peu différent: il va étudier les normales à une surface, et montrer que le long de deux familles de courbes tracées sur la surface, ces normales forment des familles de surfaces développables.

Monge suppose la surface  $M$  donnée par une équation  $z = f(x, y)$ . La famille de sphères obtenue en faisant varier  $x$  (ou  $y$ ) admet comme surface caractéristique en  $m$  un plan normal en  $m$  à  $M$ . On le voit, puisqu'en dérivant par rapport à  $x$  l'équation  $(X - x)^2 + (Y - y)^2 + (Z - z)^2 = R^2$  des sphères de la famille, on obtient l'équation d'un plan:

$$(X - x) + (Z - z)p = 0, \text{ où } p = \partial z / \partial x \quad (7.7)$$

Comme le rayon ne dépend pas de  $x$ , le plan est normal à la courbe  $z = f(x, y)$  ( $x$  varie). En faisant varier  $y$  on obtient de même un autre plan normal d'équation:

$$(Y - y) + (Z - z)q = 0, \text{ où } q = \partial z / \partial y. \quad (7.8)$$

Ces deux plans se coupent suivant la normale à la surface.

Notons  $r = \partial q / \partial x$ ;  $s = \partial q / \partial x = \partial p / \partial y$ ;  $t = \partial q / \partial y$ . En différenciant les équations 7.7 et 7.8, on obtient:

$$\begin{aligned} dx + p^2 dx + pq dy + (z - Z)(r dx + s dy) &= 0 \\ dy + pq dx + q^2 dy + (z - Z)(s dx + t dy) &= 0 \end{aligned} \quad (7.9)$$

Si dans chacune des équations 7.9, on fixe le rapport  $dy/dx = a$  on considère en fait les courbes  $\gamma$  tracées sur  $M$  tangente à la droite de pente  $a$  dans le plan  $T_m M$  tangent en  $m$  à  $M$ , et le caractéristique de la famille de plans 7.7 ou 7.8 variant le long d'une de ces courbes. Lorsque les quatre plans d'équations 7.7, 7.8 et 7.9 ont un point commun, les normales à  $M$  en deux points infinitésimalement voisins le long de  $\gamma$  se coupent en ce point (en général ces quatre plans ne sont pas concourants).

Supposons maintenant que le plan horizontal  $(x, y)$  soit le plan tangent à  $M$  en  $m$ ; on a  $p = q = 0$ ;  $z = 0$ . En éliminant  $dy/dx$  puis  $Z$  entre les équations 7.9, et en reprenant 7.7 et 7.8, on obtient:

$$\begin{aligned} 1 + (r + t)Z + (rt - s^2).Z^2 &= 0 \\ (dy/dx)^2 + (dy/dx)\frac{r-t}{s} - 1 &= 0 \end{aligned} \quad (7.10)$$

La seconde a pour solutions deux pentes  $dy/dx = a$ ;  $dy/dx = b$ , qui comme le produit  $a \cdot b$  des racines de la deuxième équation est  $-1$ , correspondent à deux directions orthogonales.

La seconde a pour racines les coordonnées  $Z$  des deux centres de courbure que nous avons interprétés à l'aide des sphères osculatrices plus haut, puisque, dans le système de coordonnées que nous avons choisi plus haut  $r = k_1$ ,  $s = 0$ ,  $t = k_2$  et donc  $Z_1 = \frac{1}{k_1}$ ;  $Z_2 = \frac{1}{k_2}$ .

### 7.3 Le point de vue intrinsèque

Dans un espace métrique  $E$ , on se donne une fonction sur  $E \times E$ : la distance entre deux points.

Lorsque cet espace métrique est un espace euclidien  $\mathbb{R}^n$ , la distance entre deux points  $m$  et  $m'$  est aussi la longueur du plus court chemin joignant  $m$  à  $m'$ . Soit maintenant  $M$  une sous-variété de  $\mathbb{R}^N$ . On peut construire à partir de la distance euclidienne  $d_{\mathcal{E}}$  de  $\mathbb{R}^N$  deux fonctions distance sur  $M$ :  $d_R(m, m') = d_{\mathcal{E}}(m, m')$ , et  $d_I(m, m')$  la borne inférieure de la longueur des chemins tracés sur  $M$  joignant  $m$  à  $m'$ . Pour savoir calculer la longueur de tous les chemins différentiables tracés sur  $M$ , il suffit de savoir calculer la norme de tous les vecteurs tangents à ce chemin, ce qui est possible, puisque chaque espace tangent  $T_z M$  est muni

Figure 7.5: The distances  $d_I$  and  $d_E$  on the sphere  $S^2$ 

d'une forme quadratique : la restriction de la métrique euclidienne de  $\mathbb{R}^N$ . Nous pouvons maintenant "oublier" que  $M$  est une sous-variété de  $\mathbb{R}^N$ , et tout de même étudier ses propriétés pour la métrique  $d_I$ .

Remarquons que ces deux définitions ne coïncident pas, mais que l'on a :

$$\lim_{m \rightarrow m'} \frac{d_I(m, m')}{d_E(m, m')} = 1$$

Il est maintenant naturel, lorsque  $M$  est une variété abstraite, de définir les formes quadratiques  $g_z$  sur chaque espace tangent  $T_z M$ . On supposera ici de plus que la forme quadratique  $g_z$  dépend continument de  $z \in M$ . On note indifféremment  $g$  la forme quadratique ou la forme bilinéaire associée.

Pour comparer globalement ou localement deux variétés riemanniennes, définissons.

**Définition 7.3.1.** Une isométrie entre deux variétés riemanniennes  $M_1, g_1$  et  $M_2, g_2$  est un difféomorphisme  $\Phi$ , tel que  $d\Phi$  est une isométrie euclidienne entre chaque espace tangent  $T_m M$  et  $T_{\Phi(m)} M_2$ , soit tel que :

$$\forall (m, v) \in TM_1, g_2(d\Phi(m)(v)) = g_1(v).$$

Il suffit de remplacer difféomorphisme par difféomorphisme local pour définir une isométrie locale.

Lorsque l'on a une paramétrisation locale de la surface  $M$  on note :

$$\begin{aligned} - E &= g\left(\frac{\partial}{\partial u}, \frac{\partial}{\partial u}\right) \\ - F &= g\left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right) \\ - G &= g\left(\frac{\partial}{\partial v}, \frac{\partial}{\partial v}\right) \end{aligned}$$

Se donner  $E, F, G$  permet de calculer avec une métrique riemannienne donnée dans une carte.

### Exemples

- Le tore plat  $\mathbf{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ , où l'espace tangent en  $z \in \mathbf{T}^2$  est muni de la métrique euclidienne de  $T_{z\zeta} \mathbb{R}^2$  où  $\zeta \in \mathbb{R}^2$  est l'un des point qui se projette sur  $z$ .

Le tore plat est localement isométrique à l'espace euclidien  $\mathbb{R}^2$ . C'est d'ailleurs ce que veut dire "plat".

### Un cas important: les surfaces développables

Montrer que la surface formée par les tangentes à une courbe de  $\mathbb{R}^3$  est localement isométrique au plan euclidien.

Soit  $c : [a, b] \rightarrow \mathbb{R}^3$  une courbe différentiable (notée  $C$ ) paramétrée par la longueur de l'arc d'origine  $c(a) :s$ . La surface balayée par les tangentes est paramétrée par :

$$(s, t) \mapsto c(s) + t \cdot \dot{c}(s)$$

La métrique induite sur la surface (de la métrique euclidienne de  $\mathbb{R}^3$  est donnée, dans la carte  $(s, t)$ , par :

$$\begin{aligned} - E &= |\dot{c}(s) + t \cdot \ddot{c}(s)|^2 \\ - F &= \langle \dot{c}(s) + t \cdot \ddot{c}(s) | \dot{c}(s) \rangle \\ - G &= |\dot{c}(s)|^2 \end{aligned}$$

Considérons maintenant la courbe plane  $\Gamma$  paramétrée par la longueur de l'arc d'origine  $\gamma(a) : \gamma : [a, b] \rightarrow \mathbb{R}^2$ , dont la courbure en fonction de la longueur de l'arc est la même fonction que la courbure de  $C$  en fonction de la longueur de l'arc sur  $C$ . Nous savons que si la courbure  $k(s)$  est nulle part nulle, cette condition détermine la courbe plane.

Les tangentes à  $\Gamma$  balayent une région du plan (bordée par l'arc  $\Gamma$  lorsque la courbure est partout non nulle sur cet arc). Cette région admet donc un paramétrage pour le quel il suffit de remplacer  $c$  par  $\gamma$  dans le paramétrage précédent. Par construction (coïncidence des valeurs de la courbure en fonction de  $s$ ), le domaine plan et la surface sont localement isométriques, au moins hors des points singuliers du paramétrage : la courbe  $C$  dans l'espace et la courbe  $\Gamma$  dans le plan.

### 7.3.1 Courbure de Gauss des surfaces

La courbure de Gauss a une propriété importante: elle est intrinsèque. Cela veut dire que pour calculer  $K(m)$ , on n'a pas besoin de connaître le germe du plongement au point  $m$ . La courbure de Gauss dépend seulement de la métrique induite sur un voisinage de  $m$ , comme un être infinitésimal (une fourmi est une bonne approximation) pourrait la percevoir.

Nous allons établir ce résultat par un calcul direct, comme le fait Gauss en 1827 (*Disquisitiones generales circa superficies curvas*) vite traduit en Français (voir [Ga0] et [Ga]). Ici la présentation suit [dCa].

La construction est plus intuitive dans le cas polyédrique (voir appendice).

Supposons donc la surface paramétrée  $M = \{m = m(u, v)\}$ . Chaque espace tangent  $T_m M$  est engendré par les vecteurs  $m_u = \frac{\partial m}{\partial u}$  et  $m_v = \frac{\partial m}{\partial v}$ . On note  $N = N(u, v)$  le vecteur normal en  $m(u, v)$  à  $M$  correspondant à l'orientation de  $M$  que nous choisissons (un choix local suffit). La première forme fondamentale de  $M$  en  $m$  est:  $I = E(du)^2 + Fdu \cdot dv + G(dv)^2$ , ce qui veut dire qu'en chaque espace tangent  $T_m M$  on se donne une forme quadratique, maintenant appelée

métrique riemannienne:

$$\begin{aligned}g(u, u) &= \langle m_u | m_u \rangle = E(u, v) \\g(u, v) &= \langle m_u | m_v \rangle = F(u, v) \\g(v, v) &= \langle m_v | m_v \rangle = G(u, v)\end{aligned}$$

(ici la métrique  $g$  est induite par la métrique euclidienne de  $\mathbb{R}^3$ , puisque définie à partir de produits scalaires de vecteurs de  $\mathbb{R}^3$ ).

Les dérivées des vecteurs  $m_u$  et  $m_v$  s'écrivent dans la base  $m_u, m_v, N$ :

$$\begin{aligned}x_{uu} &= \Gamma_{11}^1 x_u + \Gamma_{11}^2 x_v + L_1 N \\x_{uv} &= \Gamma_{12}^1 x_u + \Gamma_{12}^2 x_v + L_2 N \\x_{vv} &= \Gamma_{22}^1 x_u + \Gamma_{22}^2 x_v + L_3 N \\(\text{on n'écrit pas : } x_{vu} &= \Gamma_{21}^1 x_u + \Gamma_{21}^2 x_v + \bar{L}_2 N, \text{ puisque } x_{uv} = x_{vu})\end{aligned}$$

Posons encore (ce qui revient à écrire la matrice de  $d\gamma$  dans la base  $x_u, x_v$ ):

$$\begin{aligned}N_u &= a_{11} x_u + a_{12} x_v \\N_v &= a_{12} x_u + a_{22} x_v\end{aligned}$$

**Définition 7.3.2.** La *seconde forme fondamentale* de  $M$  est la forme quadratique définie sur  $T_m M$  par :

$$II(w) = - \langle d\gamma(w) | w \rangle \text{ (il faut bien mettre ce signe quelque part...)}$$

Posons:

$$\begin{aligned}e &= II(x_u) = \langle N_u | x_u \rangle = \langle N | x_{uu} \rangle \\f &= - \langle N_v | x_u \rangle = \langle N | x_{uv} \rangle = \langle N | x_{vu} \rangle = - \langle N_u | x_v \rangle \\g &= II(x_v) = \langle N_v | x_v \rangle = - \langle N | x_{vv} \rangle\end{aligned}$$

(on a, par exemple,  $e = \langle N | x_{uu} \rangle$  puisque  $0 = \frac{\partial}{\partial u} \langle N | x_u \rangle = \langle N_u | x_u \rangle + \langle N | x_{uu} \rangle$ ).

Remarquons encore que:  $L_1 = e$ ,  $L_2 = f$ ,  $L_3 = g$ .

Les  $\Gamma_{jk}^i$  sont appelés *symboles de Christoffel*. On les voit apparaître en particulier dans l'équation des géodésiques de  $M$ .

**Lemme 7.3.3.** *Les symboles de Christoffel sont solutions d'un système linéaire dont les coefficients sont intrinsèques.*

**Démonstration :** Trois systèmes de deux équations sont obtenus en considérant le produit scalaire de chaque équation de 7.3.1 par respectivement  $x_u$  et  $x_v$ :

$$\begin{aligned}\Gamma_{11}^1 E + \Gamma_{11}^2 F &= \langle x_{uu} | x_u \rangle = \frac{1}{2} E_u \\ \Gamma_{11}^1 F + \Gamma_{11}^2 G &= \langle x_{uv} | x_v \rangle = F_u - \frac{1}{2} E_v\end{aligned}$$

En effet:

$$\begin{aligned} F_u &= \frac{\partial}{\partial u} \langle x_u | x_v \rangle = \langle x_{uu} | x_v \rangle + \langle x_u | x_{uv} \rangle \\ E_v &= 2 \langle x_{uv} | x_u \rangle \end{aligned}$$

Comme le déterminant de la matrice  $\begin{pmatrix} E & F \\ F & G \end{pmatrix}$  est non nul (effet, par définition, sur chaque espace tangent la première forme fondamentale est définie positive).

On obtient de même:

$$\begin{aligned} \Gamma_{12}^1 E + \Gamma_{12}^2 F &= \langle x_{uv} | x_u \rangle = \frac{1}{2} E_v \\ \Gamma_{12}^1 F + \Gamma_{12}^2 G &= \langle x_{uv} | x_v \rangle = \frac{1}{2} G_u \end{aligned}$$

$$\begin{aligned} \Gamma_{22}^1 E + \Gamma_{22}^2 F &= \langle x_{vv} | x_u \rangle = F_v - \frac{1}{2} G_u \\ \Gamma_{22}^1 F + \Gamma_{22}^2 G &= \langle x_{vv} | x_v \rangle = \frac{1}{2} G_v \end{aligned}$$

systèmes de même déterminant non nul qui permettent donc aussi de calculer des coefficients de Christoffel en fonction de  $E, F, G$  et de leurs dérivées par rapport à  $u$  et  $v$ .  $\square$

Déterminons maintenant des relations entre les trois familles  $\Gamma_{ij}^k; a_{ij}, e, f, g$ . Pour cela nous allons utiliser les trois relations:

$$\begin{aligned} (x_{uu})_v - (x_{uv})_u &= 0 \\ (x_{vv})_u - (x_{vu})_v &= 0 \\ N_{uv} - N_{vu} &= 0 \end{aligned}$$

Chacune de ces 3 équations, en décomposant les vecteurs dans la base  $x_u, x_v, N$  se réécrit sous la forme:

$$A_i x_u + B_i x_v + C_i N = 0, \quad i = 1, 2, 3$$

où les coefficients  $A_i, B_i, C_i$  dépendent de  $E, F, G, e, f, g$  et de leurs dérivées jusqu'à l'ordre 2, car les coefficients  $a_{ij}$ , se calculent en fonction de  $E, F, G, e, f, g$  (équations de Weingarten). Cependant pour démontrer le theorema egregium, nous allons choisir la paramétrisation pour qu'au point  $m_0$  on ait:  $E_0 = G_0 = 1, F_0 = 0$ . On peut de plus supposer que  $x_u$  et  $x_u$  sont des vecteurs propres de  $d\gamma$  (nous ne démontrons pas les équations de Weingarten, mais, ce qui revient presque au même, nous avons démontré que les directions propres sont orthogonales et qu'avec ce choix de base  $e(m_0) = a_{11} = k_1(m_0), a_{12} = 0 = f(m_0), g(m_0) = a_{22} = k_2(m_0)$ ).

Remarquons aussi qu'avec ce choix de base,  $K(m_0) = a_{11} \cdot a_{22} = ef$ , et rappelons que  $L_1 = e, L_2 = f$ .

Écrivons enfin:

$$(\Gamma_{11}^1 x_u + \Gamma_{11}^2 x_v + L_1 N)_v - (\Gamma_{12}^1 x_u + \Gamma_{12}^2 x_v + L_2 N)_u = 0$$

soit:

$$\begin{aligned} & (\Gamma_{11}^1 x_{uv} + \Gamma_{11}^2 x_{vv} + e N_v) + ((\Gamma_{11}^1)_v x_u + (\Gamma_{11}^2)_v x_v + e_v N) = \\ & = (\Gamma_{12}^1 x_{uu} + \Gamma_{12}^2 x_{vu} + f N_u + ((\Gamma_{12}^1)_u x_u + (\Gamma_{12}^2)_u x_v + f_u N) = 0 \end{aligned}$$

Remplaçons maintenant les vecteurs  $x_{uu}, x_{uv}, x_{vv}$  par leurs valeurs à l'aide des équations 7.3.1, on obtient:

$$\Gamma_{11}^1 \Gamma_{12}^2 + \Gamma_{11}^2 \Gamma_{22}^2 + e \cdot a_{22} + (\Gamma_{11}^2)_v = \Gamma_{12}^1 \Gamma_{11}^2 + \Gamma_{12}^2 \Gamma_{12}^2 + f \cdot a_{21} + (\Gamma_{12}^2)_u$$

en  $m_0$ ; comme de plus  $f \cdot a_{21} = 0$  en ce point, on obtient bien:

$$K(m_0) = e \cdot f = e \cdot a_{22} = \text{une combinaison des } \Gamma_{ij}^k \text{ et de leurs dérivées}$$

ce qui démontre le theorema egregium de Gauss.

### 7.3.2 Géodésiques et courbure géodésique

## 7.4 Modèles de surfaces de courbure négative

### 7.4.1 Les modèles de Poincaré

Poincaré considère le disque unité et le demi plan supérieur où les cercles orthogonaux au bord vont jouer le rôle de géodésiques.

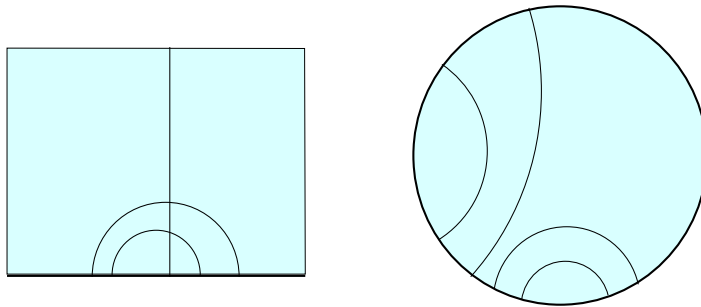


Figure 7.6: Le demi-plan de Poincaré et le disque de Poincaré

Les métriques riemanniennes à considérer sur ces deux modèles pour obtenir les cercles orthogonaux au bord comme géodésiques sont :

$$D = \{|z| < 1\} \quad g_D = \frac{4}{(1 - |z|^2)^2} \cdot g_{eucl} \quad 1/2 \text{ plan} = \{(x, y) \mid y > 0\} \quad g_{1/2 \text{ plan}} = \frac{1}{y^2} \cdot g_{eucl}$$

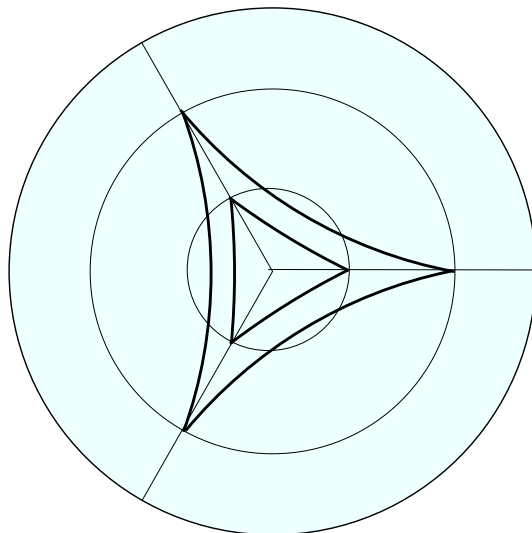


Figure 7.7: les triangles équilatéraux en géométrie hyperbolique

C'est sans doute à l'aide du disque de Poincaré que l'on voit le mieux la relation entre aire et angles d'un triangle équilatéral :

C'est aussi le modèle du disque de Poincaré qui permet relativement facilement de montrer l'existence de surfaces de genre  $g \geq 2$  munie d'une métrique de courbure constante  $K = -1$ .

Pour obtenir une surface de genre 2 il suffit de partir d'un octogone régulier centré au centre du disque. Les homothéties euclidiennes transforment les sommets de ce polygone; en les joignant par des géodésiques on obtient une famille continue de polygones réguliers entre un polygone régulier asymptotique, dont les sommets sont situés sur le cercle à l'infini, et un polygone infinitésimal, de coté aussi petit que l'on voudra, qui est lui presque euclidien. Le quotient par un groupe d'isométries convenable d'un polygone d'angles  $\frac{\pi}{4}$ .

### Les modèles de Klein et de Lorenz

L'espace  $\mathbb{R}^3$  est muni de la forme quadratique  $\mathcal{L}$ , définie par :

$$\mathcal{L}(x) = x_1^2 + x_2^2 - x_3^2$$

On note  $L^3$  l'espace  $\mathbb{R}^3$  muni de la forme quadratique  $\mathcal{L}$ .

Le cône de lumière est l'ensemble  $\mathcal{L}(x) = 0$ .

Nous dirons qu'une droite est de type espace si tout vecteur non nul  $v$  de cette droite satisfait  $\mathcal{L}(v) > 0$ , qu'elle est de type temps si tout vecteur non nul de cette droite satisfait  $\mathcal{L}(v) < 0$ . Un sous-espace de  $L^3$  sera dit de type espace si tout ses vecteurs non nuls sont de type espace. Si un sous espace contient à la

fois des vecteurs de type espace et des vecteurs de type temps, nous dirons qu'il est de type mixte.

Cette terminologie est inspirée par la théorie de la relativité.

La forme quadratique  $\mathcal{L}$  admet un groupe d'isométries  $\mathcal{G} \simeq O(2, 1)$ .

La quadrique  $\mathcal{L} = 1$  est un hyperboloïde à deux nappes. La nappe  $\mathcal{L} = 1; x_3 > 0$  est un modèle de l'espace hyperbolique  $\mathbb{H}^2$ . La restriction de  $\mathcal{L}$  à l'espace tangent en un point de  $\mathbb{H}^2$  est définie positive, et définit donc une métrique riemannienne sur  $\mathbb{H}^2$ .

**Affirmation** Cette métrique est de courbure négative  $K = -1$ .

Les droites de ce plan sont les intersection de ce plan avec les plans de type mixte de  $\mathbb{R}^3$ .

Klein propose un modèle où les points de  $\mathbb{H}^2$  sont encore ceux du disque unité tandis que les "droites" sont cette fois les segments d'extrémités contenues dans le bord.

La métrique hyperbolique  $g_H$  s'écrit alors, en fonction de la métrique euclidienne du disque unité :

$$g_H = \left( \frac{1}{(1-r^2)^2} dr + \frac{r^2}{1-r^2} d\theta^2 \right) \cdot g_{eucl}$$

Si l'on voit le disque unité dans le plan  $x_3 = 1$ , bordé par le cercle intersection du cône de lumière avec le plan  $x_3 = 1$ , les segments sont les traces de plans de type mixte. Les demi-droites de type temps donnent la correspondance, une isométrie, entre la nappe supérieure de l'hyperboloïde  $\mathcal{L} = -1$ ; on voit que les plans de type mixte coupent et le disque unité et cette nappe suivant les géodésiques.



# Chapter 8

## Le cinquième postulat d'Euclide

### 8.1 Le cinquième postulat

Le premier livre des éléments d'Euclide commence par 23 définitions, 5 postulats et 5 notions de sens commun.

Euclide (environ 330-275 avant J.C.) définit d'abord point, courbe et surface, à partir des notions de mesure de longueur d'aire et de volume, puis le plan et la droite, les angles et les figures (domaines à bord); la 23<sup>ème</sup> définition est:

**Définition 8.1.1.** 23 Des droites parallèles sont des droites qui, étant dans le même plan, et prolongées infiniment dans les deux directions, ne se rencontrent dans aucune des deux directions.

Le cinquième postulat s'énonce:

#### Postulat

*Si une droite coupant deux autres droites fait avec celle-ci des angles intérieurs du même côté de somme inférieure à deux angles droits, les deux droites, si on les prolonge indéfiniment, se rencontrent du côté où les angles sont de somme inférieure à deux angles droits.*

### 8.2 Tentatives de démonstration

La nécessité de ce postulat a été vite mise en cause comme en témoigne Proclus (410-485) Celui ci mentionne une "preuve" proposée par Ptolémée (2<sup>ème</sup> siècle après J.C.) :

*Étant données deux droites  $AB$ ,  $CD$ , et une troisième  $EFGH$  qui les rencontre, en formant les angles  $BFG$ ,  $FGD$  de somme deux angles droits. Je dis que  $AB$ ,  $CD$  sont parallèles, c'est à dire ne se rencontrent pas.*

En effet, supposons que  $FB$  et  $GD$  se rencontrent en un point  $K$ . Puisque la

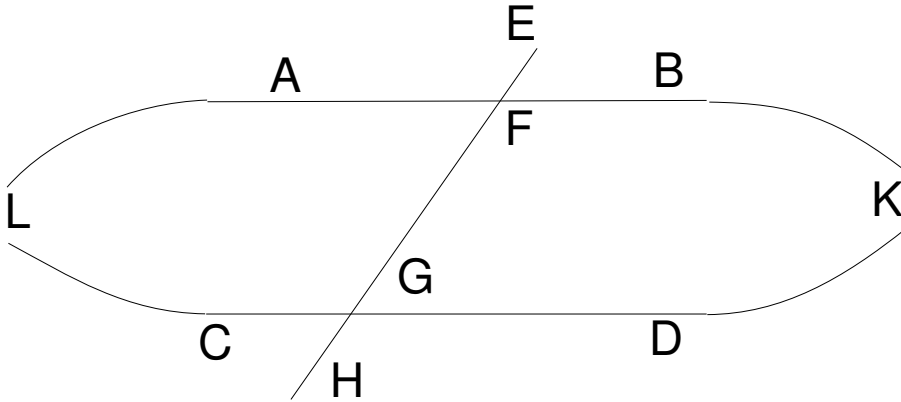


Figure 8.1: “preuve” de l’axiome des parallèles suivant Ptolémée

somme des angles  $BFG$  et  $FGD$  est deux angles droits, tandis que la somme des 4 angles  $AFG, BFG, FGD$ , et  $FGC$  est 4 angles droits, la somme des angles  $AFG$  et  $FGC$  est deux angles droits. Si donc  $FB$  et  $GD$ , alors que la somme des angles intérieurs est deux angles droits, se rencontrent en  $K$ , les droites  $FA$  et  $GC$  se rencontreront aussi, puisque la somme des angles  $AFG$  et  $CGF$  est aussi égale à deux angles droits.

Donc les droites, soit se rencontreront des deux côtés, soit elles ne se rencontreront ni d’un côté ni de l’autre, quand la somme des deux angles intérieurs est égale à deux angles droits.

Dans le premier cas, soit  $L$  le point d’intersection de  $FA$  et  $GC$ .

Les droites  $LABK$  et  $LCDK$  entourent alors un espace, ce qui est impossible. [C’est ce qui se passe en géométrie sphérique].

Il n’est donc pas possible que deux droites se rencontrent lorsque la somme des angles intérieurs est égale à deux angles droits ; elles sont donc parallèles.

Proclus critique cette “preuve”, et en propose une autre qui repose sur un autre argument, postulat équivalent au 5<sup>ème</sup> postulat d’Euclide : la distance entre deux parallèles reste finie.

Le sujet est repris pas des mathématiciens arabes, en particulier **Nasîr-Eddîn** (1201-1274). Son point de départ est le suivant: *Si deux droites  $r$  et  $s$  sont la première perpendiculaire à un segment  $AB$ , la seconde oblique, les segments de perpendiculaire menés de  $s$  sur  $r$  sont plus courts que  $AB$  du côté où  $s$  fait un angle aigu avec  $AB$  et plus longs du côté où  $s$  fait un angle obtus avec  $AB$ .*

Le sujet est repris au 16<sup>ème</sup> siècle. Au 17<sup>ème</sup> **Wallis** (1616-1703) abandonne l’idée d’équidistance des parallèles pour la remplacer par l’axiome suivant :

Chaque figure admet une figure semblable dans un rapport arbitraire.

C'est cependant **Gerolamo Saccheri** (1667-1733), qui, le premier, va tenter, certes toujours pour en déduire une preuve par l'absurde du cinquième postulat, de construire une géométrie en supposant le cinquième postulat faux. La figure fondamentale de sa construction est le quadrilatère isocèle ayant deux angles droits, plus précisément, un quadrilatère dont deux coté opposés sont égaux et perpendiculaires à la base.

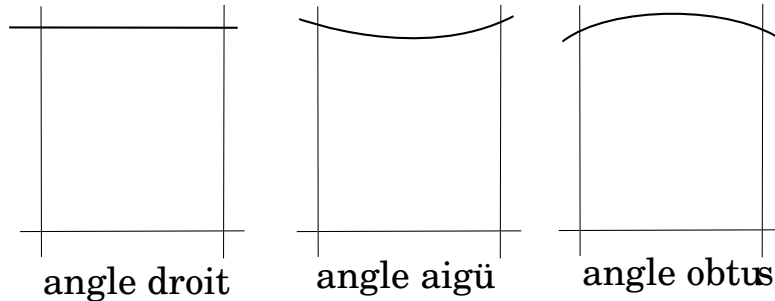


Figure 8.2: quadrilatères de Saccheri

On peut donc maintenant choisir entre trois hypothèses:

- l'hypothèse de l'angle droit : l'angle en  $C$  est égal à l'angle en  $D$  est égal à un angle droit
- l'hypothèse de l'angle obtus : l'angle en  $C$  est égal à l'angle en  $D$  est supérieur à un angle droit.
- l'hypothèse de l'angle aigu : l'angle en  $C$  est égal à l'angle en  $D$  est inférieur à un angle droit.

Un des résultats importants qu'il obtient est :

**Proposition 8.2.1.** *Suivant que l'on fasse l'hypothèse de l'angle droit, de l'angle obtus ou de l'angle aigu, on doit avoir dans le quadrilatère isocèle à deux angles droits  $AB = CD$ ,  $AB > CD$  ou  $AB < CD$  respectivement.*

Il en déduit ensuite que suivant l'hypothèse choisie, la somme des angles d'un triangle doit être égale à 2 angles droits, plus grande, ou plus petite.

Il élimine ensuite l'hypothèse de l'angle obtus. Cela lui permet de constater que par un point  $A$  donné situé hors d'une droite, il doit dans l'hypothèse de l'angle aigu, exister dans le pinceau des droites passant par le point extérieur  $A$ , une famille de droites ne rencontrant pas la droite donnée, bordée par deux droites particulières  $p$  et  $q$ .

Une autre conséquence de l'hypothèse de l'angle aigu est, étant donné une droite  $L$  et un point  $A$  hors de la droite, de deux droites passant par  $A$  asymptotes

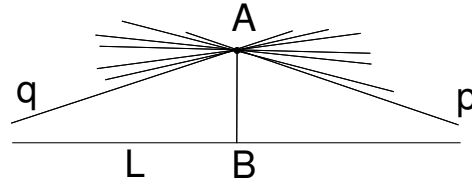


Figure 8.3: Droites non sécantes

à  $L$  et telles que les droites passant par  $A$  ne rencontrant pas  $L$  sont celle du pinceau passant par  $A$  “entre” les deux asymptotes.

**Johann Heinrich Lambert** (1728-1777) Il distingue les données relatives, obtenues par comparaison de figures du même type, et absolues. Ce lui permet de montrer qu’éliminer l’existence d’une longueur absolue lui permet d’éliminer l’hypothèse de l’angle aigü. [Par exemple en géométrie hyperbolique, on aura la distance d’où l’on voit une droite sous un angle droit; en géométrie hyperbolique il n’y a pas un groupe de similitudes contenant le groupe des isométries.]

Il montre aussi que, dans l’hypothèse de l’angle aigü, l’aire d’un triangle sera proportionnelle au “défaut d’angle”, différence entre deux droits et la somme des angles d’un triangle.

**Adrien Marie Legendre** (1752-1833) Montrons comment, une centaine d’années après Saccheri, il rejette l’hypothèse de l’angle obtus.

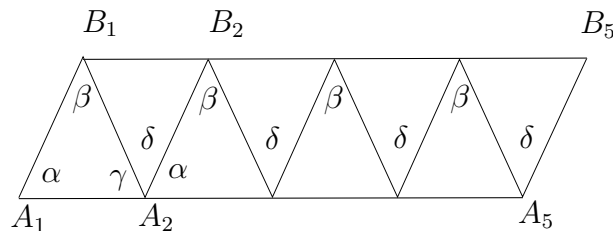


Figure 8.4: Conséquence pour Legendre de l’hypothèse de l’angle obtus

Il aligne sur une droite des triangles égaux, en commençant par un triangle  $A_1B_1A_2$  dont le côté  $A_1A_2$  repose sur la droite. Notons respectivement  $\alpha$ ,  $\beta$  et  $\gamma$  les angles du premier triangle aux points  $A_1$ ,  $B_1$  et  $A_2$ . Notons  $\delta$  l’angle  $B_1A_2B_2$ . Si l’on suppose que la somme  $\alpha + \beta + \gamma$  des angles du premier triangle est supérieure à deux droits ( $\pi$ ), on doit avoir  $\delta \leq \beta$ . En effet  $\gamma + \delta + \alpha = \pi$  puisque l’on obtient un angle plat en observant les trois angles en  $A_2$  sommet commun de trois triangles (voir figure 8.4). Euclide a démontré que dans ce cas, comme les deux triangles  $A_1B_1A_2$  et  $B_1A_2B_2$  ont, de part et d’autre des angles respectivement  $\beta$  et  $\delta$ , des côtés respectivement égaux, les côtés  $A_1A_2$  et  $B_1B_2$  satisfont l’inégalité :

$$|B_1B_2| < |A_1A_2|$$

Enfin, en juxtaposant assez de triangle on obtient une ligne brisée  $A_1B_1B_2\dots B_{n+1}A_{n+1}$  qui sera plus courte que le segment  $A_1A_{n+1}$ , le terme  $n(|A_1A_2| - |B_1B_2|)$  finissant par être plus grand que la somme  $|A_1B_1| + |B_{n+1}A_{n+1}|$  (on utilise ici le postulat d'Archimède : on peut toujours, à pas aussi petits que l'on voudra, finir par dépasser un point sur la droite).

On a construit une ligne brisée plus courte qu'un segment de droite, ce qui est la contradiction cherchée.

### 8.3 Les fondateurs de la géométrie non euclidienne.

**Carl Friedrich Gauss**(1777-1855) Que l'auteur du "theorema egregium" soit le premier à concevoir une géométrie où par un point passent plusieurs droites qui ne coupent pas une droite donnée n'est pas une surprise.

Gauss définit ainsi deux droites parallèles (géodésiques asymptotes dans le plan hyperbolique) : *Si deux droites coplanaires  $AM$ ,  $BN$  ne se coupent pas, tandis que par ailleurs toute droite passant par  $A$  contenue entre  $AM$  et  $AB$  rencontre  $BN$ , on dira que  $AM$  est parallèle à  $BN$ .*

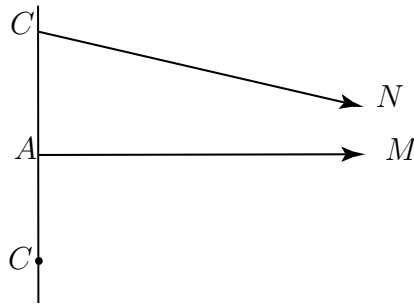


Figure 8.5: Parallèles suivant Gauss

En étudiant les familles de droites parallèles, il trouve une correspondance naturelle entre leur points; Lobatchewsky appellera ces courbes des *horocycles*. Par ailleurs il calcule la longueur de cercles de rayon  $r$  :

$$\ell = \pi(e^{\frac{r}{k}} - e^{-\frac{r}{k}})$$

ce qui lui permet de retrouver la géométrie euclidienne lorsque  $k$  tend vers  $+\infty$ .

**Franz Adolf Taurinus**(1794-1874) Partant de la relation fondamentale de la trigonométrie sphérique dans une sphère de rayon  $k$  :

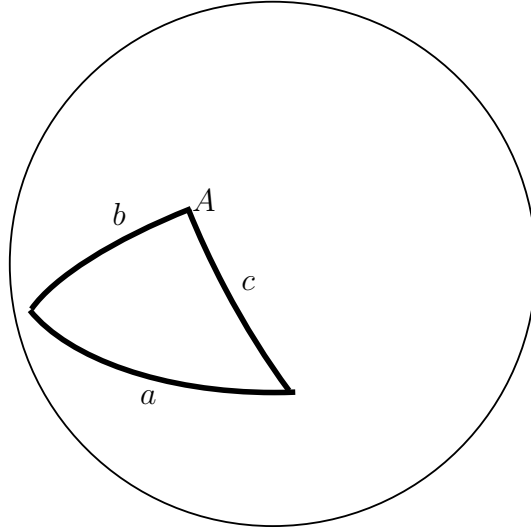


Figure 8.6: Triangle sphérique

$$\cos \frac{a}{k} = \cos \frac{b}{k} \cos \frac{c}{k} + \sin \frac{b}{k} \sin \frac{c}{k} \cos A$$

il transforme le rayon réel  $k$  en un rayon imaginaire, et obtient une formule de trigonométrie hyperbolique (logarithmique-sphérique chez Taurinus) :

$$\text{Cosh} \frac{a}{k} = \text{Cosh} \frac{b}{k} \text{Cosh} \frac{c}{k} + \text{Sinh} \frac{b}{k} \text{Sinh} \frac{c}{k} \cos A$$

**Nicolai Ivanovitch Lobatchewsky**(1793-1856) Dans la géométrie de Lobatchewsky il y a dans le pinceau des droites passant par un point  $A \notin BC$  une famille de droites qui ne coupent pas la droite  $BC$  (voir figure 8.3 (Saccheri)).

L'angle entre la perpendiculaire  $AD$  à la droite  $BC$  et une de ces non sécantes  $a$  est noté  $\pi(a)$ . en géométrie euclidienne on a toujours  $\pi(a) = \text{angle droit}$ .

Il obtient aussi une formule de trigonométrie hyperbolique, définit horocycles et horosphères, et constate que sur une horosphère la géométrie est euclidienne.

**Johann Bolyai**(1802-1860)

Son travail contient aussi la définition des parallèles et leurs propriétés indépendantes du cinquième postulat, l'existence de cercles et sphères de rayon infinis, "horocycle, horosphères", trigonométrie sphérique, euclidienne et hyperbolique. Tandis que le point de vue de Lobatchewsky est plus analytique, celui de Bolyai est plutôt celui d'une construction sans supposer a priori que le cinquième postulat soit vrai ou faux. Il dira même : *j'ai créé un nouvel univers à partir de rien*. Cependant la réaction de Gauss, reconnaissant ses propres idées non publiées chez Bolyai, le laissera très amer.

# Chapter 9

## Appendice

### 9.1 Le mot “vecteur” (citation)

Extrait de *Earliest Known Uses of Some of the Words of Mathematics (V) Last revision: Sept. 10, 2007 (AOL)*

VECTOR, VECTOR ANALYSIS and VECTOR SPACE. The entries below follow changes in the use of the term vector over the past 150 years or so. The story begins with a technical term in theoretical astronomy, “radius vector,” in which “vector” signified “carrier.” In the first and biggest change “radius” was dropped and “vector” was given a place in the algebra of quaternions (c. 1840). The recognition that vectors are more interesting than quaternions, especially in physics, led to vector analysis (c. 1880). In the 20th century vector was extended from triples of numbers to n-tuples and then to elements of abstract linear spaces.

The word VECTOR (which, like the word vehicle, derives ultimately from the Latin *vehere* to carry) was first a technical term in astronomical geometry. The OED’s earliest entry is from a technical dictionary of 1704: J. Harris *Lexicon Technicum* I. s.v., “A Line supposed to be drawn from any Planet moving round a Center, or the Focus of an Ellipsis, to that Center or Focus, is by some Writers of the New Astronomy, called the Vector; because ’tis that Line by which the Planet seems to be carried round its Center.”

Vector usually appeared in the phrase radius vector. The French term was *rayon vecteur* and can be found in e.g. Laplace’s *Traité de mécanique céleste* (1799-1825).

The term *rayon vecteur* is used in a non-astronomical context by Monge “*Application de l’Analyse à la Géométrie*” (1807). On p. 24 he writes, “on nomme la droite *r* le RAYON VECTEUR du point, et l’origine des coordonnées

devient un pôle, d'où partent les rayons vecteurs des différens points de l'espace." Monge does not seem to use the terminology later in the text, however. Cauchy in his *Leçons sur les Applications du Calcul Infinitésimal à la Géométrie* of 1826 uses the term freely after introducing it in the initial "Preliminaries" chapter: "Une droite AB, menée d'un point A supposé fixe à un point B suppose mobile, sera généralement désignée sous le nom de rayon vecteur." (p.14) (Information from François Ziegler.) Radius vector appears in English in 1831 in *Elements of the Integral Calculus* (1839) by J. R. Young: "...when the angle O between the radius vector and fixed axis is taken for the independent variable, the formula is...."

Hamilton would create a new meaning for vector (see VECTOR & SCALAR) but he used radius vector in the conventional way in *On a General Method in Dynamics* *Philosophical Transactions Royal Society* (part II for 1834, pp. 247-308); see article 14.

A list of matrix and linear algebra terms having entries on this web site is here.

VECTOR and SCALAR. Both the terms vector and scalar were introduced by William Rowan Hamilton (1805-1865).

terms appear in "On quaternions" a paper presented by Hamilton at a meeting of the Royal Irish Academy on November 11, 1844. This paper adopts the convention of denoting a vector by a single (Greek) letter, and concludes with a discussion of formulae for applying rotations to vectors by conjugating with unit quaternions. It is on pages 1-16 in volume 3 of the *Proceedings of the Royal Irish Academy*, covering the years 1844-1847, and the volume is dated 1847. The following is from page 3:

On account of the facility with which this so called imaginary expression, or square root of a negative quantity, is constructed by a right line having direction in space, and having  $x, y, z$  for its three rectangular axes, he has been induced to call the trinomial expression itself, as well as the line which it represents, a VECTOR. A quaternion may thus be said to consist generally of a real part and a vector. The fixing a special attention on this last part, or element, of a quaternion, by giving it a special name, and denoting it in many calculations by a single and special sign, appears to the author to have been an improvement in his method of dealing with the subject: although the general notion of treating the constituents of the imaginary part as coordinates had occurred to him in his first researches. The following is from page 8: It is, however, a peculiarity of the calculus of quaternions, at least as lately modified by the author, and one which seems to him important, that it selects no one direction in space as eminent above

another, but treats them as all equally related to that extra-spacial, or simply SCALAR direction, which has been recently called "Forward." In Hamilton's time radius-vector was an established term in astronomy (see previous entry). Hamilton explains that he is giving the term vector a new sense in Lecture I of his Lectures on Quaternions. In article 17 (p. 16) he described the difference between vector and radius-vector: 17. To illustrate more fully the distinction which was just now briefly mentioned, between the meanings of the "Vector" and the "Radius Vector" of a point, we may remark that the RADIUS-VECTOR, in astronomy, and indeed in geometry also, is usually understood to have only length; and therefore to be adequately expressed by a SINGLE NUMBER, denoting the magnitude (or length) of the straight line which is referred to by this usual name (radius-vector) as compared with the magnitude of some standard line, which has been assumed as the unit of length. Thus, in astronomy, the Geocentric Radius-Vector of the Sun is, in its mean value, nearly equal to ninety-five millions of miles: if, then, a million of miles be assumed as the standard or unit of length, the sun's geocentric radius-vector is equal (nearly) to, or is (approximately) expressible by, the number ninety-five: in such a manner that this single number, 95, with the unit here supposed, is (at certain seasons of the year) a full, complete and adequate representation or expression for that known radius vector of the sun. For it is usually the sun itself (or more fully the position of the sun's centre) and NOT the Sun's radius-vector, which is regarded as possessing also certain other (polar) coordinates of its own, namely, in general, some two angles, such as those which are called the Sun's geocentric right-ascension and declination; and which are merely associated with the radius-vector, but not inherent therein, nor belonging thereto... But in the new mode of speaking which it is here proposed to introduce, and which is guarded from confusion with the older mode by the omission of the word "RADIUS," the VECTOR of the sun HAS (itself) DIRECTION, as well as length. It is, therefore NOT sufficiently characterized by ANY SINGLE NUMBER, such as 95 (were this even otherwise rigorous); but REQUIRES, for its COMPLETE NUMERICAL EXPRESSION, a SYSTEM OF THREE NUMBERS; such as the usual and well-known rectangular or polar co-ordinates of the Sun or other body or point whose place is to be examined...

A VECTOR is thus (as you will afterwards more clearly see) a sort of NATURAL TRIPLET (suggested by Geometry): and accordingly we shall find that QUATERNIONS offer an easy mode of symbolically representing every vector by a TRINOMIAL FORM ( $ix + jy + kz$ ); which form brings the conception and expression of such a vector into the closest possible connexions with Cartesian

and rectangular co-coordinates.

Hamilton, in his “Lectures on Quaternions”, is not satisfied with having introduced vector. Within a few pages we find vectum, vehend, revector, provector, provectum, transvehend, transvectum, etc., and identities such as

Provectum = Provector + Vector + Vehend. Vector and scalar also appear in 1846 in a paper “On Symbolical Geometry,” in the The Cambridge and Dublin Mathematical Journal vol. I:

If then we give the name of scalars to all numbers of the kind called usually real, because they are all contained on the one scale of progression of number from negative to positive infinity [...] Next Hamilton goes on to tell us about another “chief class” of the “geometrical quotients,” namely the class in which the dividend is a line perpendicular to the divisor. A quotient of this latter class we shall call a vector, to mark its connection (which is closer than that of a scalar) with the conception of space [...] David Wilkins believes that the paper “On quaternions” in the Proceedings of the Royal Irish Academy probably appeared earlier than the CDMJ, probably some time in the first half of 1845.

The first occurrence of vector and scalar in the London, Edinburgh, and Dublin Philosophical Magazine is in volume XXIX (1846) in the article “On Quaternions; or on a new System of Imaginaries in Algebra”:

The separation of the real and imaginary parts of a quaternion is an operation of such frequent occurrence, and may be regarded as being so fundamental in this theory, that it is convenient to introduce symbols which shall denote concisely the two separate results of this operation. The algebraically real part may receive, according to the question in which it occurs, all values contained on the one scale of progression from number from negative to positive infinity; we shall call it therefore the scalar part, or simply the scalar of the quaternion, and shall form its symbol by prefixing, to the symbol of the quaternion, the characteristic Scal., or simply S., where no confusion seems likely to arise from using this last abbreviation. On the other hand, the algebraically imaginary part, being geometrically constructed by a straight line, or radius vector, which has, in general, for each determined quaternion, a determined length and determined direction in space, may be called the vector part, or simply the vector of the quaternion; and may be denoted by prefixing the characteristic Vect. or V... Information for this article was provided by David Wilkins and Julio González Cabillón. See HAMILTON and QUATERNION.

*Fin de la citation.*

## 9.2 Grassmanniennes

Les variétés grassmanniennes sont les variétés dont les points sont les sous-espaces d'une dimension donnée d'un espace  $\mathbb{R}^n$  (ou  $\mathbb{C}^n$  on parle alors de grassmannienne complexes).

**Définition 9.2.1.** - La variété grassmannienne  $G(n, p)$  est l'ensemble des  $p$ -plans (vectoriels) de  $\mathbb{R}^n$ .

- La variété grassmannienne affine  $\mathcal{A}(n, p)$  est l'ensemble des  $p$ -plans affines de  $\mathbb{R}^n$ .

**Proposition 9.2.2.** *Les ensembles que nous venons de définir ont une structure naturelle de variété;  $G(n, p)$  est de dimension  $n \cdot p$  et  $\mathcal{A}(n, p)$  est de dimension  $np + q$  où  $q = n - p$ .*

## 9.3 Enveloppes

À une famille à un paramètre de droites affines du plan, correspond en général une courbe, qui est l'enveloppe de la famille de droites. Soit  $D_t = \{a(t)x + b(t)y + c(t) = 0\}$  une famille lisse de droites pour lesquelles les fonctions  $a(t), b(t), c(t)$  sont des fonctions lisses de  $t$ . Si elles ne sont pas parallèles, les droites  $D_t$  et  $D_{t+h}$  ont un point d'intersection à distance finie. Quand  $h$  tend vers zero, ce point d'intersection peut avoir une limite  $m(t)$ . Donnons une condition suffisante pour que cette limite  $m(t)$  existe, et pour que l'ensemble des points  $m(t)$  définissent une courbe  $C$  qui a en chacun des points  $m(t)$  une tangente  $D_t$ .

**Théorème 9.3.1.** *Let  $D_t$  be a smooth family of lines of equations:  $a(t)x + b(t)y + c(t) = 0$ ;  $(x, y) \in \mathbb{R}^2$ . If for all  $t \in [\alpha, \beta]$ , the determinant  $\det \begin{pmatrix} a(t) & b(t) \\ a'(t) & b'(t) \end{pmatrix}$  is different from zero, the family envelopes a curve  $C$ , that is, the curve is the union of the points:  $m(t) = D_t \cap D'_t$ , where  $D'_t$  is the affine line of equation  $a'(t)x + b'(t)y + c'(t) = 0$ .*

Moreover if the determinant  $\det \begin{pmatrix} a(t) & b(t) & c(t) \\ a'(t) & b'(t) & c'(t) \\ a''(t) & b''(t) & c''(t) \end{pmatrix}$  is also different from zero, the curve is smooth at  $m(t)$  and the tangent to  $C$  at  $m(t)$  is  $D_t$ .

We will note  $D''_t$  the line of equation  $a''(t)x + b''(t)y + c''(t) = 0$ .

**Démonstration :** Let us find the intersection point of  $D_t$  and  $D_{t+h}$ . We need to solve the linear system:

$$\begin{aligned} a(t)x + b(t)y + c(t) &= 0 \\ a(t+h)x + b(t+h)y + c(t+h) &= 0 \end{aligned}$$

A first order Taylor expansion of the second equation gives:

$$\begin{aligned} a(t)x + b(t)y + c(t) &= 0 \\ (a(t) + a'(t)h + o(h))x + (b(t) + b'(t)h + o(h))y + (c(t) + c'(t)h + o(h)) &= 0 \end{aligned}$$

This is equivalent to the system:

$$\begin{aligned} a(t)x + b(t)y + c(t) &= 0 \\ (a'(t)h + o(h))x + (b'(t)h + o(h))y + (c'(t)h + o(h)) &= 0 \end{aligned}$$

If the determinant

$$\det \begin{pmatrix} a(t) & b(t) \\ a'(t) & b'(t) \end{pmatrix} \neq 0, \quad (9.1)$$

the limit of the solution, when  $h$  goes to zero, is the solution  $m(t)$  of the system:

$$\begin{aligned} a(t)x + b(t)y + c(t) &= 0 \\ a'(t)x + b'(t)y + c'(t) &= 0 \end{aligned}$$

To avoid giving a special role to the coefficients  $a, b$  we can consider the same problem in the projective space  $\mathbb{P}_2$ . The lines are then projective lines of homogeneous equation  $ax + by + cz = 0$ ; the condition 9.1 is now:

$$\text{rank} \begin{pmatrix} a(t) & b(t) & c(t) \\ a'(t) & b'(t) & c'(t) \end{pmatrix} = 2$$

The exceptional solution “at infinity” of the system 9.3 is treated as any other intersection point of two non-parallel lines.

The condition  $\det \begin{pmatrix} a(t) & b(t) & c(t) \\ a'(t) & b'(t) & c'(t) \\ a''(t) & b''(t) & c''(t) \end{pmatrix} \neq 0$  guarantees that the three lines  $D, D'$  and  $D''$  do not belong to the same linear pencil. Up to terms negligible compared with  $h$  the point  $m(t+h)$  is the point  $D_t \cap D'_{t+h}$ , which show that the limit of the line containing the chord  $m(t), m(t+h)$  is  $D_t$ ; see next picture  $\square$

Linear pencils are in that sense “degenerate” envelopes.

De la même manière on peut étudier des enveloppes de familles à un ou deux paramètres de plans.

Si la matrice

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$$

est de rang 2 l'intersection des plans  $P_t$  d'équation:

$$a(t)x + b(t)y + c(t)z + d(t) = 0$$

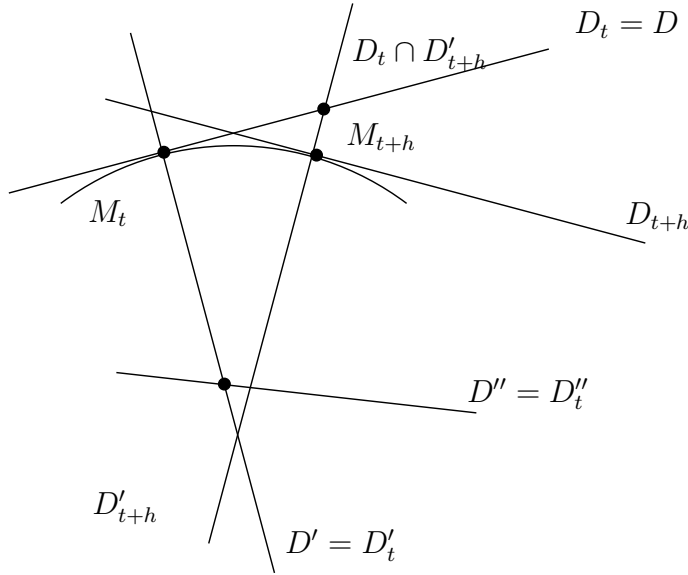


Figure 9.1: A non degenerate piece of envelope

et

$$P_{t+h}$$

d'équation:

$$a(t+h)x + b(t+h)y + c(t+h)z + d(t+h) = 0$$

admet une limite: la droite  $G_t$  d'équations:

$$\begin{aligned} a(t)x + b(t)y + c(t)z + d(t) &= 0 \\ a'(t)x + b'(t)y + c'(t)z + d'(t) &= 0 \end{aligned}$$

Si la matrice

$$\begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix}$$

est de rang 3, on obtient une surface  $E$  qui admet en tous les points de chacune des droites  $G_t$  un plan tangent qui est le plan  $P_t$ .

Si de plus la matrice

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \tag{9.2}$$

est de rang 3 on voit que, infinitésimalement, les droites  $G_t$  et  $G_{t+h}$  se coupent en un point  $F_t$ . Remarquons que ce n'est en général pas le cas pour des droites infinitésimalement proches dans une famille à un paramètre.

en supposant toujours que la matrice 9.2 est de rang 3, la surface  $M$  engendrée par les droites  $G_t$  est une surface réglée très particulière: elle est développable, ce qui veut dire que l'on peut la faire rouler sans glisser sur un plan. Déjà, puisque la normale à la surface le long d'une des droites  $G_t$ , que l'on appelle *génératrices*, définie sauf au point  $F_t$ , est constante, la courbure de Gauss de la surface  $M$  est nulle en ces points.

Les points  $F_t$  forment une courbe, en général gauche,  $C$ . Nous pouvons la paramétrer par la longueur  $s$  de l'arc d'origine un point arbitraire  $m_0 \in C$ . Paramétrons aussi (par  $t$ ) chaque tangente isométriquement à partir du point de tangence  $c(s)$  en respectant l'orientation définie par  $T(s)$ . Lorsque  $C$  est une courbe de classe  $\mathcal{C}^2$ , de courbure  $k(s)$  partout non nulle, il existe une courbe plane  $C_P$ , dont la courbure  $k(s)$  exprimée en fonction de la longueur de l'arc est la même que celle de  $C$ . Cette courbe est unique aux isométriques du plan près. Paramétrons comme celle de  $C$  les tangentes à  $C_P$  par  $t$ .

Notons  $T(s) = \partial c / \partial s$  le vecteur tangent à  $C$  défini par le paramétrage  $s$ , et par  $T_P(s) = \partial c_P / \partial s$  le vecteur tangent à  $C_P$  correspondant.

**Proposition 9.3.2.** *L'application  $dev : c(s) + tT(s) \mapsto c_P(s) + tT_P(s)$  qui "développe"  $M$  sur le plan est une isométrie locale de  $M \setminus C$  dans son image, un "coté" de la courbe plane  $C_P$ .*

**Démonstration :** Pour démontrer la proposition il faut calculer la différentielle de l'application  $dev$  et vérifier que c'est une isométrie.

Pour cela utilisons les repères orthonormés  $T(s), N(s) = \frac{1}{k(s)} dT/ds$  du plan tangent au point de paramètres  $(s, t)$  de  $M$  et  $T_P(s), N_P(s) = \frac{1}{k(s)} dT_P/ds$  au point de paramètres  $(s, t)$  du plan  $\mathbb{R}^2$ .

La matrice de chacune des paramétrisations de  $M$  et du plan  $\mathbb{R}^2$  par  $(s, t)$  s'écrit:

$$\begin{pmatrix} 1 & 0 \\ t \cdot k(s) & 1 \end{pmatrix}$$

ce qui, en regardant le diagramme:

$$\begin{array}{ccc} M & \rightarrow & \mathbb{R}^2 \\ \uparrow (s, t) & & (s, t) \uparrow \\ & Id & \\ \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \end{array}$$

démontre la proposition. □

## 9.4 Courbure de Gauss des surfaces polyédriques

Ce paragraphe se base sur l'article [Ban1] de T. Banchoff.

Le Theorema Egregium de Gauss est relativement facile à démontrer pour les surfaces polyédriques de  $\mathbb{R}^3$  en utilisant la géométrie intégrale.

For us a *polyedral surface* will be a union of planar triangles, such that if two triangles meet, they have in common a full edge or a vertex. The union of these triangles should be topologically a surface.

To prove the theorema egregium, we need to define the Gauss image of a vertex.

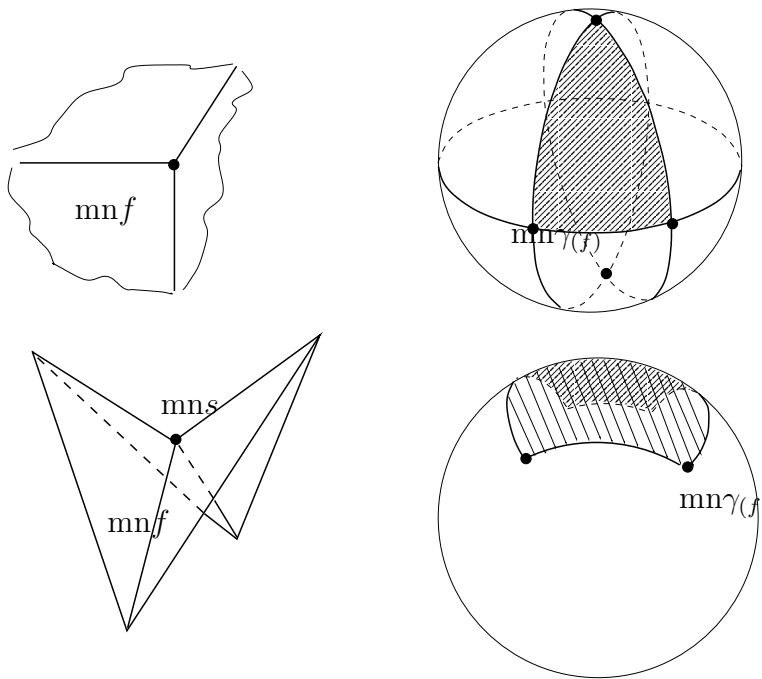


Figure 9.2: The Gauss image of a vertex

Let us first define the *link* of a vertex. It is the union of the sides opposite to the vertex in the triangles containing the vertex.

Observe that, if the vertex  $x_o$  is a useless one, that is a point in a plane  $P_o$ , with the triangles with vertex that point contained in the plane  $P_o$ , then the link  $\mathcal{L}(x_o)$  of the vertex is contained in the plane  $P_o$ . Therefore its intersection with any plane  $P \neq P_o$  containing  $x_o$  is exactly two points.

**Définition 9.4.1.** Gauss image of a vertex It is a function on  $\mathbb{P}^2$  defined by

$$\Gamma(P) = 1 - \frac{1}{2}[\#(P \cap \mathcal{L}(x_o))]$$

Notice that the function  $\Gamma(P)$  is identically zero when the vertex is useless.

We can see this function as a subset of  $\mathbb{P}^2$  counted with multiplicity (and draw it as a subset of  $\mathbf{S}^2$ ).

**Définition 9.4.2.** • The total curvature of the vertex  $x_o$  is:

$$|k|(x_o) = \int_{\mathbb{P}^2} |\Gamma(P)|$$

• The curvature of the vertex  $x_o$  is:

$$k(x_o) = \int_{\mathbb{P}^2} \Gamma(P)$$

Notice that these two curvatures are intrinsic, that is depend only on information we can get staying inside the polyedral surface.

**Théorème 9.4.3.** *Theorema egregium.* The curvature at a vertex  $x_o$  satisfies:

$$k(x_o) = \int_{\mathbb{P}^2} \Gamma(P) = 2\pi - \sum \alpha_i$$

Gauss considered the analogous smooth statement as important as its relates two definitions, one extrinsic, one intrinsic. In the polyedral case, the angles can be measured “inside” the surface, and are therefore intrinsic. **Démonstration :** Let us compute the contribution of each side of the link  $\mathcal{L}(x_o)$  to the integral  $\int_{\mathbb{P}^2} \Gamma(P)$ . We see that it is  $2\alpha$ , where  $\alpha$  is the angle at  $x_o$  of the triangle containing

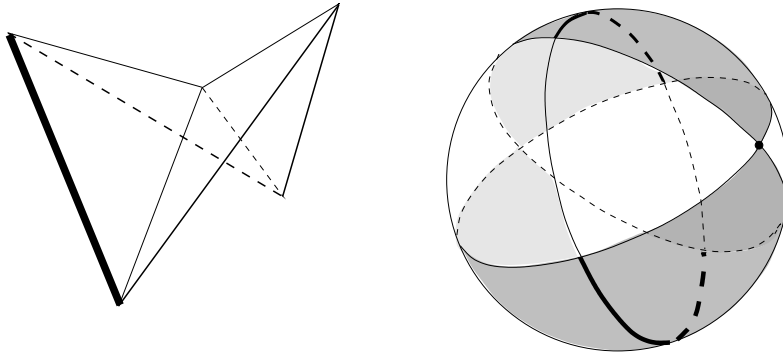


Figure 9.3: The normals to planes intersecting a side of the link of a vertex the edge  $E$ . Summing up, we get

$$k(x_o) = 2\pi - \sum \alpha_i$$

where the  $\alpha_i$  are all the angles with vertex  $x_o$ . □

The proof of Gauss-Bonnet formula:

**Théorème 9.4.4.** *Gauss-Bonnet theorem* When the surface  $M$  is compact without boundary, the sum of the curvatures of all the vertices is the Euler characteristic of the surface.

$$\sum_{x \text{ vertex of } M} k(x) = 2\pi \cdot \chi(M)$$

where the Euler characteristic of  $M$  is the sum: number of faces - number of edges + number of vertices  $\chi(M) = f - e + v$ .

**Démonstration :** Let us consider the set  $\mathcal{D}$  of pairs  $\mathcal{D} = \{edge \in face\}$ . As an edge is the boundary of two faces, and a face has three boundary edges, we have  $\#\mathcal{D} = 3f = 2e$ .

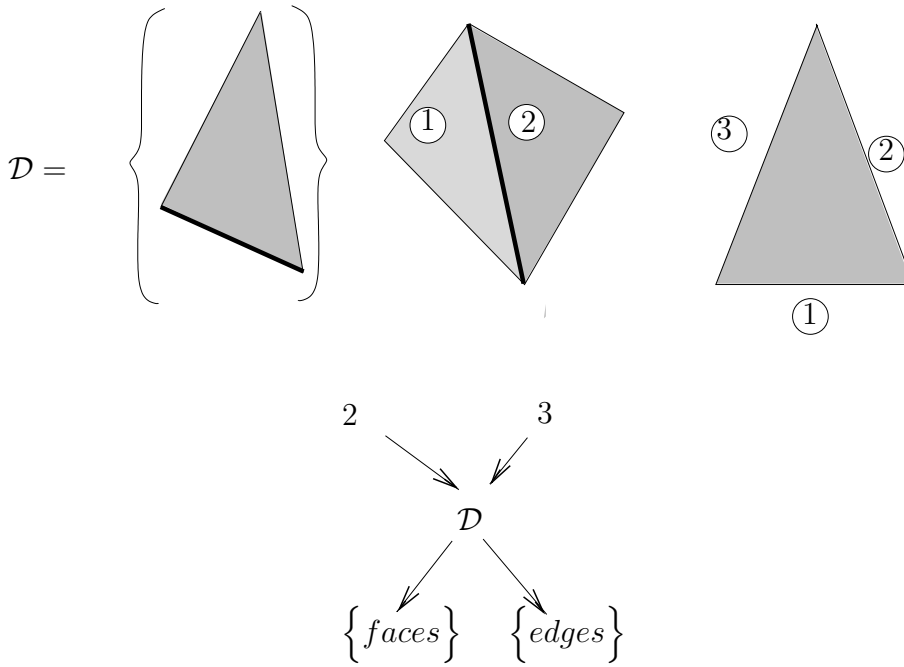


Figure 9.4: The pairs  $edge \in face$

The sum on all vertices of their curvatures is  $\sum k(x)$ . It is also, as the sum of angles of a triangle is  $\pi$ :

$$\sum_{vertices} k(x) = \sum_{vertices} [2\pi - \sum \alpha_i] = 2\pi v - \pi f$$

Then:

$$2\pi v - \pi f + O = 2\pi v - \pi f + 3\pi f - 2\pi e = \chi(M)$$

□

**Exposés****2002**

Dimension 4; relativité.

L'inégalité isopérimétrique.

Le nombre d'enlacement (Gauss...).

Gauss: construction de polygones réguliers.

Polygone de Newton.

Le théorème fondamental de l'algèbre.

Essai d'arithmétique morale et géométrie intégrale.

Polyèdres réguliers.

Euler : logarithmes.

Théorème de Riemann sur les nombres premiers.

Histoire de l'infini.

Polémique entre Berkeley et les disciples de Newton.

Polygones réguliers constructibles, théorème de Gauss.

Constructions graphiques, mécaniques et approchées.

Lindeman : transcendance de  $\pi$ .

Le 10<sup>ème</sup> problème de Hilbert.

Construction de Dedekind des nombres réels.

**2003**

Archimède : l'aire entre un arc de parabole et sa corde.

La géométrie descriptive chez Monge.

Pascal: l'intégration des puissances du sinus.

Les paradoxes de Bertrand et les formules de Poincaré et Blaschke.

L'inégalité isopérimétrique.

Le programme d'Erlangen.

Les fractions continues.

Complexes, quaternions, octaves de Cayley.

Le théorème de Cantor Bernstein.

Monge et les courbures principales d'une surface.

## 2005

Euler, nombres complexes.

Archimède, aire entre un arc de parabole et une corde.

L'inégalité isopérimétrique.

Condorcet, probabilités: constitution de Jurys.

Fréquence des nombres premiers.

Polyèdres réguliers.

Pascal, intégration de  $\sin x$ ,  $\sin^n x$ .

Coniques.

## Propositions supplémentaires pour 2008

La (les?) perspective(s).

Les systèmes de représentation de la terre et du ciel et leurs propriétés géométriques.

L'astrolabe.



# Bibliography

- [Ab] Edwin A. Abbot. *Flatland, a romance of many dimensions*. Dover 1992.
- [Arch] Archimède. *Les oeuvres complètes*, traduction Paul Ver Eecke Desclée, de Brouwer et compagnie Paris, Bruxelles (1921).
- [Ari] Aristote. *de caelo*, ??
- [Ban1] Thomas F. Banchoff. *Critical points and curvature for embedded polyedral surfaces*, American Mathematical Monthly vol 77 (1970) p.475-486.
- [Bar] Barbier. *Note sur le problème de l'aiguille et le jeu du joint ouvert*. Journal de mathématiques pures et appliquées (2) 5. (1860) p.534-541.
- [Bo-Fe] T. Bonnesen und W. Fenchel. *Theorie de konvexen Körper*, Springer (1934).
- [Bo] Roberto Bonola. *Non-Euclidean geometry* The open court publishing company, Chicago 1912.
- [Bro] Brouwer. *Beweis der Invarianz de Dimensionzahl*, Math. Ann. **70**(1911) p.161-165.
- [Bu] Georges-Louis Leclerc, comte de Buffon. *Essai d'arithmétique morale*, (1777) 4<sup>o</sup> volume des suppléments de l'édition in quarto de l'imprimerie royale (France).
- [Bu-Za] Yu.D. Burago and V.A. Zalgaler. *Geometric inequalities*, Grundlehren der mathematischen Wissenschaften 285, Springer-Verlag (1988).
- [Cau] A.L. Cauchy. *Mémoire sur la rectification des courbes et la quadrature des surfaces courbes*, mémoires de l'académie des sciences, Paris **22** (1850) p.3-15.
- [Cau2] A.L. Cauchy *Résumé des leçons données à l'École Royale Polytechnique sur le calcul infinitésimal* Debure frères, Paris (1823).

- [Cha] Jean-Luc Chabert. *La vraie fausse démonstration du Cinquième Postulat*. Histoire de problèmes, histoire des mathématiques; commission inter I.R.E.M. Ellipse Paris (1993).
- [Cro] W. Crofton. *On the theory of local probability*, Phil. Trans. of the Royal soc. London **158** (1868) p.181-199.
- [Da] G. Darboux. *Leçons sur la théorie générale des surfaces*. Gauthier-Villars (1888; 2<sup>ème</sup> édition 1915).
- [dCa] M. do Carmo. *Differential geometry of curves and surfaces*, Prentice Hall (1976).
- [D-K-P-P] P. Deleham, G. Kientz, J-C. Penin, P. Perrin. *Histoires de tangentes*, Publication de l'I.R.E.M. de Reims (1994).
- [Eu] Euclide. *The thirteen books of the elements*. Traduction et commentaires par Sir Thomas L. Heath, Dover publications, New York, 1956.
- [Eu1] L. Euler. *Recherches sur la courbure des surfaces*, Mémoires de l'académie des sciences de Berlin [16] (1760), 1767, p.119-143; in Opera omnia I 28, Commentationes geometricae Lausanne 1955 p.1-22.
- [Ga0] Carl Friedrich Gauss. *Recherches générales sur les surfaces courbes*, (traduction Capitaine Tiburce Abadie) Nouvelles annales de mathématiques vol 11 (1852).
- [Ga] Carl Friedrich Gauss. , traduction française *recherche générales sur les surfaces courbes* par E. Roger Albert Blanchard, Paris (1967).
- [Gr] Hermann Günther Grassmann. *La "lineale Audehnungslehre"*, (1840), traduction française : Dominique Flamend et Bernd Bekeneier, revue par Eberart Knobloch, Blanchard (1994).
- [Gam] Gamov. *One, two, three...infinity, ?* Penguin ? traduction française *Un, deux, trois... l'infini*, Dunod?
- [Hi] D. Hilbert. *Ueber die stetige Abbildung einer Linie auf ein Flächenstück*, Math. Annalen **38** (1891) p.459-450.
- [Hi-Co] D. Hilbert and S. Cohn-Vossen. *Anschauliche Geometrie*, Springer (1932); traduction anglaise *Geometry and imagination*, Chelsea (1952).

- [Hu] Hurewicz and Wallman. *Dimension theory*, Princeton university press (1941).
- [Kah] J-P. Kahane. *Naissance et postérité de l'intégrale de Lebesgue*, Gazette des mathématiciens (société mathématique de France) N° **89** (2001) p.5-20.
- [La] Pierre-Simon Laplace. *Théorie analytique des probabilités* Oeuvres tome VII Paris imprimerie Royale, (1847) et réédition: éditions J. Gabai (1995).
- [Mil1] J. Milnor. *Topology from the differential viewpoint*, the university press of Virginia, second edition (1969).
- [Mil2] J. Milnor. *Morse theory*, Princeton university press **51** (1963).
- [Min] H. Minkowski. *Sur les surfaces convexes fermées, comptes rendus de l'académie des sciences*, Paris, (1901) t **132**, p.21-24.
- [Mona] A.F. Monna. *Functional analysis in historical perspective*, Academische paperbacks, Oosthoek publishing company, Utrecht, Pays bas (1973).
- [Mo1] G. Monge. *Géométrie descriptive*. Paris, Beaudoin an VII (1798 ou 1799); réédition Jacques Gabay, Sceaux, France (1989).
- [New2] I. Newton. *Principia*, ( , 1726); traduction par Madame la Marquise du Chastellet (1756); réédition Blanchard, Paris (1966).
- [New2] I. Newton. *La méthode des fluxions et des suites infinies*, (traduction anglaise et première édition 1736); traduction M. de Buffon (1740), réédition Blanchard, Paris (1966)
- [d'O] d'Ocagne. *Sur la courbure du contour apparent d'une surface projetée orthogonalement*, Nouvelles annales de mathématiques (école polytechnique) p. 262-264 (1895).
- [Pea1] G. Peano. *Intégration par séries des équations différentielles linéaires*, Math. annalen **32** (1888) p 450-456.
- [Pea2] G. Peano. *Sur une courbe, qui remplit toute une aire plane*, Math. annalen **36** p.157-160.
- [Poin] H. Poincaré. *Calcul des probabilités*, Gauthier-Villars 2ème édition (1912).
- [Ri] B. Riemann. *Über die Hypothesen, welche der Geometrie zu Grunde liegen (sur les hypothèses à la base de la géométrie)*, Göttingen (1834). Traduit en anglais dans [Spi].

- [Sa1] L.A. Santalò. *Introduction à la géométrie intégrale*, Paris (1951).
- [Sa2] L.A. Santalò. *Integral geometry and geometric probability, Encyclopedia of mathematics and its applications*, Addison Wesley (1976).
- [Spi] M. Spivak. *A comprehensive introduction to differential geometry*, Publish or perish, Berkeley 1979.
- [St] S. Stevin. *Beghinselen der Weegconst (éléments de l'art de peser)*, Leiden 1585.
- [Str] D.J. Struik. *A source book in mathematics 1200-1800*, Harvard (1969).
- [Tri] C. Tricot. *Thèse*, Orsay (1983).
- [Vi] Virgile. *L'Énéïde*, (19 avant Jésus Christ, publication posthume) Texte avec traduction française de M. Rat, Garnier Paris (1947).
- [You] A.P. Youschkevitch. *Le concept de fonction jusqu'au milieu du XIXème siècle*, Brochure APMEP N° 41 (1981).
- [Zou] R. Zouckermann. *Galilée penseur libre*, édition de l'union rationaliste (1968).