



COULISSES • F2. COMPLÉMENTS DE DÉRIVATION ET CONVEXITÉ

L'envers de la courbure

Le cours vous a appris à dériver la dérivée et à lire la courbure. Mais il a beaucoup tu. Un jeu de pile ou face vaut une infinité d'euros et personne n'en offre plus de trois. La règle de dérivation la plus utilisée du chapitre a attendu cent vingt et un ans qu'on daigne l'écrire. Un ingénieur des téléphones a fondé une théorie que les professeurs n'avaient pas vue. Une méthode rapportée d'Orient par Fibonacci converge au rythme du nombre d'or. Et un architecte a dessiné ses églises la tête en bas, avec des chaînes et un miroir. Numéro court, cette fois : le chapitre précédent a déjà raconté beaucoup. Chaque pièce se lit seule, aucune n'est exigible, et toutes contiennent de vraies mathématiques.

Plan de la visite

ARCHIVES • Le juste prix d'un jeu infini (9 min)	2
CONTRE-PIED • La règle que personne n'écrivait (7 min)	3
AUTRE CHEMIN • Un théorème, trois héritiers (8 min)	4
FRONTIÈRE • La théorie qui n'intéressait personne (8 min)	5
PASSERELLE • La corde au rythme doré (7 min)	6
ARCHIVES • L'architecte qui pendait ses églises (6 min)	7
ARCHIVES • Brèves de coulisses (3 min)	8

Le juste prix d'un jeu infini

L'épigraphe du cours affirme que les hommes de bon sens estiment l'argent à l'usage. Voici le jeu de hasard qui a forcé cette phrase, et la concavité qui la cache.

↔ prolonge COURS F2, GARDE ET § 4.2 · l'épigraphe de Cramer et la propriété « Milieux et moyennes »

COMBIEN paieriez-vous pour jouer au jeu suivant ? On lance une pièce équilibrée jusqu'au premier pile. S'il survient au premier lancer, vous gagnez un ducat ; au deuxième, deux ducats ; au troisième, quatre ; à chaque lancer supplémentaire, le gain double. Le jeu est posé en 1713 par Nicolas Bernoulli dans sa correspondance avec Montmort, et il embarrasse aussitôt tout le monde : le calcul dit que ce billet de loterie vaut une *infinité* de ducats, et aucune personne sensée n'en offre plus de quelques-uns. Le paradoxe prendra le nom de la ville où la famille Bernoulli le fera mûrir : Saint-Petersbourg.

Pour les curieux du fond. D'où vient l'infini ? Le pile arrive au lancer k avec la probabilité $\frac{1}{2^k}$, et rapporte alors 2^{k-1} ducats. L'espérance de gain additionne chances et gains :

$$\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{8} \times 4 + \dots = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots$$

Chaque terme vaut exactement un demi-ducats, et il y en a une infinité : la somme dépasse tout nombre. L'espérance est infinie, et pourtant, une fois sur deux, le jeu rapporte un seul ducat.

La réponse qui a traversé les siècles arrive en 1728, dans une lettre du Genevois Gabriel Cramer, vingt-quatre ans : le calcul est juste, c'est la marchandise qui est mal mesurée. « Les mathématiciens estiment l'argent à proportion de sa quantité, et les hommes de bon sens à proportion de l'usage qu'ils en peuvent faire. » Le millionième ducat rend moins service que le premier ; Cramer propose de mesurer le *plaisir* d'un gain par sa racine carrée, et de faire la moyenne des plaisirs plutôt que des sommes.

Pour les curieux du fond. Refaisons son calcul (valeurs recalculées). L'espérance d'utilité vaut

$$\sum_{k=1}^{+\infty} \frac{1}{2^k} \sqrt{2^{k-1}} = \frac{1}{2 - \sqrt{2}} \approx 1,7071,$$

et la somme, cette fois, converge : la racine carrée écrase les gains géants plus vite que leurs probabilités ne fondent. Quel gain *certain* procurerait le même plaisir ? Celui dont la racine vaut 1,7071, soit $(1,7071)^2 \approx 2,91$ ducats. Le juste prix du jeu infini : moins de trois ducats, exactement ce que le bon sens offrait. La racine carrée est **concave** : sa courbe s'essouffle, et cet essoufflement est la modération du joueur.

Dix ans plus tard, Daniel Bernoulli publie la solution restée célèbre, avec une autre fonction d'usage, le logarithme, que le chapitre F3 construira. Peu importe la fonction choisie : ce qui règle le paradoxe, c'est sa *forme*. Une utilité concave reste sous ses cordes, et rester sous la corde, c'est précisément préférer un gain moyen sûr à un pari équilibré entre deux extrêmes.

N.B. Cette idée est devenue un pilier de l'économie : la concavité de l'utilité est l'aversion au risque. C'est elle qui explique qu'on paie une assurance plus cher que la perte moyenne qu'elle couvre : on achète la certitude, parce que la corde est sous l'arc. Le cours résume cela en une ligne, l'inégalité des milieux du § 4.2 ; l'économie en a fait une théorie.

La règle que personne n'écrivait

Une marge du cours résume l'affaire en quatre noms. Voici l'histoire complète de la formule la plus utilisée du chapitre : cent vingt et un ans de service avant d'être énoncée.

↔ prolonge COURS F2, § 1.3 · la marge Leibniz · Lagrange

ON imagine les grandes formules naissant baptisées : un théorème, un nom, une date. La dérivée d'une composée, $(v \circ u)' = (v' \circ u) \times u'$, raconte l'inverse. Sa première apparition connue date de 1676, dans un mémoire où Leibniz dérive une racine emboîtée, du type $\sqrt{a + bz + cz^2}$: il applique la chaîne exactement comme vous au paragraphe 1, et il se trompe de signe en route. La règle naît clandestine et fautive, sans énoncé, au détour d'un calcul.

Pour les curieux du fond. Refaisons proprement le calcul qui inaugura la règle. Posons, sur un intervalle I où $a + bz + cz^2 > 0$:

$$\forall z \in I, \quad g(z) = \sqrt{a + bz + cz^2}.$$

Couche intérieure $u(z) = a + bz + cz^2$, dérivable, avec $u'(z) = b + 2cz$; couche extérieure la racine, dérivable sur $]0; +\infty[$. La chaîne donne

$$\forall z \in I, \quad g'(z) = \frac{b + 2cz}{2\sqrt{a + bz + cz^2}}.$$

Deux zooms multipliés, dix secondes de calcul : le manuscrit de 1676, lui, laisse échapper un signe. Même les inventeurs se font piéger par leurs propres machines.

La suite est une histoire d'invisibilité. L'Hôpital, dont l'*Analyse des infiniment petits* (1696), nourrie des leçons de Jean Bernoulli, est le premier manuel de calcul de l'histoire, utilise la règle sans jamais l'énoncer. Euler, qui dérive des composées en virtuose dans des milliers de pages, ne l'érige jamais en règle. Il faut attendre 1797 et la *Théorie des fonctions analytiques* de Lagrange pour lire un énoncé explicite, cent vingt et un ans après le premier usage.

N.B. Pourquoi si tard ? Regardez la notation de Leibniz : $\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$. Écrivez ainsi, la règle ressemble à une simplification de fractions, une évidence qui ne mérite pas d'énoncé : les du semblent se simplifier. C'est la notation f' , introduite par Lagrange dans ce même traité de 1797, qui rend la règle visible en la privant de son déguisement : $(v \circ u)'$ ne ressemble plus à rien de connu, il faut donc dire ce que c'est. Une bonne notation peut cacher un théorème ; en changer le révèle.

Un théorème, trois héritiers

Le numéro 4 avait promis à Rolle une maison au chapitre de la dérivation. La dette est payée au § 4.3 du cours ; voici ce que l'héritage contient, radar des autoroutes compris.

↔ prolonge COURS F2, § 4.3 · le rendez-vous de Rolle, promis par la pièce 33 du n°4

UNE promesse traînait depuis le numéro précédent : le théorème de Michel Rolle, adversaire converti du calcul infinitésimal, attendait « sa vraie maison ». Le cours l'y a installé, et la maison s'est révélée être une nurserie : de ce théorème d'apparence modeste, partir et revenir à la même altitude force une tangente horizontale, descendent trois résultats qui portent tout le chapitre.

Le premier héritier s'obtient en penchant le tableau. Si les altitudes de départ et d'arrivée diffèrent, la tangente forcée n'est plus horizontale mais *parallèle à la corde* : c'est le théorème des accroissements finis. Quelque part sur le trajet, la vitesse instantanée égale exactement la vitesse moyenne.

Pour les curieux du fond. Ce théorème roule sur nos autoroutes. Un radar-tronçon ne mesure aucune vitesse instantanée : il note l'heure d'entrée et l'heure de sortie, et divise la distance par la durée. Si votre moyenne sur le tronçon vaut 151 km/h, de quel droit vous reproche-t-on d'avoir *roulé* à 151 km/h ? Du droit des accroissements finis : la position $t \mapsto x(t)$ est continue et dérivable, donc il a existé un instant c où

$$x'(c) = \frac{x(b) - x(a)}{b - a} = 151 \text{ km/h.}$$

L'infraction instantanée est un théorème d'analyse : à un moment précis, votre compteur affichait très exactement la moyenne. Le radar ne sait pas quand ; le théorème garantit que l'instant a existé.

Le deuxième héritier est le grand théorème de la classe de première : si $f' \geq 0$ sur un intervalle, alors f y est croissante. Vous l'utilisiez depuis un an sur parole ; le cours l'a ramené en trois lignes aux accroissements finis, donc à Rolle. Et le troisième héritier est le chapitre entier : les caractérisations de la convexité reposent sur le lien entre signe de f' et variations, donc sur le deuxième héritier, donc sur Rolle. La courbure elle-même est de la descendance.

N.B. La généalogie ne tourne pas en rond, et cela mérite d'être vérifié une fois dans sa vie. Rolle se démontre sans jamais parler de croissance : une fonction continue sur un segment atteint son maximum et son minimum ; si les deux sont aux bornes, f est constante et tout c convient ; sinon un extremum vit à l'intérieur, et en un extremum intérieur d'une fonction dérivable, la tangente est horizontale. Aucun maillon n'utilise ce que la chaîne va produire : l'édifice de première tient debout, avec cinq ans de retard sur son utilisation.

La théorie qui n'intéressait personne

Le cours date la convexité de 1906, trois siècles après la dérivée. Voici pourquoi la théorie est si jeune, et pourquoi son fondateur n'était pas professeur.

↔ prolonge COURS F2, § 3.3 · les caractérisations de la convexité et leur marge

R IEN ne prédisposait Johan Jensen à fonder quoi que ce soit. Ingénieur à la compagnie des téléphones de Copenhague, où il fera toute sa carrière jusqu'à en diriger le service technique, il fait des mathématiques le soir, en autodidacte, sans poste ni étudiants. En 1906, il publie dans une grande revue l'article qui donne à la convexité son statut de théorie : définition par les cordes, inégalité générale, et cette phrase de programme restée célèbre en substance : les fonctions convexes engendrent la majorité des inégalités classiques.

Le plus étrange n'est pas l'amateur, c'est le retard. La dérivée avait trois siècles, l'inégalité elle-même traînait dans les tiroirs : Otto Hölder l'avait démontrée en 1889 pour les fonctions deux fois dérivables, exactement le point de vue $f'' \geq 0$ de votre cours, et un polytechnicien français oublié, Jules Grolous, avait établi le cas discret dès les années 1870. Mais chacun voyait là un lemme de passage, un outil local. Jensen a vu une *classe de fonctions* : ranger ensemble toutes celles qui restent sous leurs cordes, et étudier ce que ce seul trait géométrique impose. C'est ce déplacement du regard, de l'inégalité vers la famille, qui fonde une théorie.

Pour les curieux du fond. Le cours a démontré l'inégalité de Jensen à poids égaux : pour f convexe sur I et $x_1, \dots, x_n$ dans I ,

$$f\left(\frac{x_1 + \dots + x_n}{n}\right) \leq \frac{f(x_1) + \dots + f(x_n)}{n}.$$

Tournons la manivelle deux fois. Avec f la fonction carré (convexe) : *la moyenne des carrés dépasse le carré de la moyenne*, l'inégalité qui gouverne l'écart type en statistiques. Avec la racine carrée (concave, inégalité renversée) : *la moyenne des racines est sous la racine de la moyenne*. Deux inégalités de concours pour un seul geste ; en choisissant mieux la fonction et les poids, la même manivelle produit presque toutes les autres.

Une inégalité classique est presque toujours une convexité qui s'ignore.

N.B. La théorie n'a plus jamais cessé de grandir : la convexité est aujourd'hui le partage des eaux de l'optimisation. Sur une fonction convexe, tout minimum local est global, aucun creux ne piège la descente : c'est la brève en fin de numéro. Et en classe préparatoire, l'inégalité de Jensen reviendra avec l'espérance en curseur, sous la forme $f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)]$: la version que l'économie de la pièce 41 utilisait déjà sans le dire.

La corde au rythme doré

Le laboratoire du cours a chronométré la sécante et sa marge a promis un joyau : l'ordre de convergence est le nombre d'or. Voici d'où vient la corde, et d'où vient l'or.

↔ prolonge COURS F2, § 6 · le laboratoire de la sécante et la marge du nombre d'or

B IEN avant les dérivées, on cherchait déjà des racines avec des cordes. La *fausse position* consiste à essayer deux valeurs fausses, à mesurer leurs erreurs, et à corriger au prorata : c'est exactement couper à la corde. La méthode remonte aux scribes égyptiens pour les problèmes linéaires, transite par les mathématiciens arabes sous le nom de *hisâb al-khatâ'ayn*, le calcul par les deux erreurs, et entre en Europe en 1202 dans le *Liber Abaci* de Fibonacci, qui l'appelle *elchataym* en latinisant l'arabe. La sécante de votre cours est la petite-fille itérative de ce calcul de marchands.

Et voici la coïncidence qui n'en est pas une. Le cours annonce en marge que l'ordre de convergence de la sécante est le nombre d'or $\varphi = \frac{1+\sqrt{5}}{2} \approx 1,618$. Or le nombre d'or est précisément la limite des rapports de la suite de Fibonacci, 1, 1, 2, 3, 5, 8, 13, ..., où chaque terme est la somme des deux précédents. La méthode que Fibonacci a rapportée d'Orient converge au rythme de la suite qui porte son nom.

Pour les curieux du fond. D'où sort l'or ? Près de la racine, l'analyse montre que l'erreur de la sécante obéit approximativement à

$$e_{n+1} \approx C e_n e_{n-1} :$$

la nouvelle erreur est proportionnelle au *produit* des deux dernières, car la corde s'appuie sur les deux derniers points. Comptons les décimales exactes, c'est-à-dire l'ordre de grandeur de $-\log e_n$: un produit d'erreurs devient une *somme* de décimales, et le compte obéit alors, à une constante près, à

$$d_{n+1} \approx d_n + d_{n-1},$$

la récurrence de Fibonacci elle-même. Sur la table du cours, les décimales exactes valent environ 0,9 ; 1,0 ; 1,4 ; 2,1 ; 3,3 ; 5,1 ; 8,1 : chacune est presque la somme des deux précédentes, et leurs rapports grimpent vers 1,62. L'ordre exact est le nombre p qui vérifie $p^2 = p + 1$, l'équation même du nombre d'or.

N.B. Le tableau de famille est désormais complet : la dichotomie, ordre 1, ne demande que la continuité et ne rate jamais ; la sécante, ordre $\varphi \approx 1,618$, ne demande aucune dérivée ; les tangentes, ordre 2, exigent f' et un bon départ. Entre la tortue et le lièvre, la corde est le compromis doré, et c'est souvent elle que choisissent les bibliothèques de calcul quand la dérivée coûte cher. La course chronométrée des deux premières méthodes vit déjà sur le compagnon numérique du chapitre F1 : la troisième coureuse vient de s'inscrire.

L'architecte qui pendait ses églises

La marge d'ouverture du cours montrait Gaudí devant un miroir. Voici la maquette tête en bas, la courbe qui pend dans toutes les chaînes, et une légende à tirer au clair.

↔ prolonge COURS F2, OUVERTURE · la marge Gaudí · Colònia Güell

EN 1898, Antoni Gaudí reçoit commande d'une église pour la colonie ouvrière Güell, près de Barcelone. Plutôt que de dessiner, il suspend : au plafond de son atelier, un réseau de ficelles lestées de petits sacs de plomb, patiemment ajusté pendant des années. Chaque ficelle pend selon la courbe d'équilibre d'un fil, en **traction pure**. Sous la maquette, un miroir : l'image renversée montre l'église à l'endroit, et le renversement change les tractions en **compressions** pures. Les arcs ainsi lus tiennent par leur forme seule, sans contreforts : l'architecte calcule avec la gravité, et lit ses plans dans un miroir.

L'idée du renversement était chiffrée depuis longtemps. En 1675, Robert Hooke la publie sous forme d'anagramme, à déchiffrer plus tard : telle pend la ligne flexible, telle, inversée, tient la voûte rigide. Restait à savoir *quelle* est cette courbe. Galilée la croyait parabolique ; Huygens, à dix-sept ans, démontre que non ; et en 1691, Leibniz, Huygens et Jean Bernoulli publient ensemble la solution du défi lancé par Jacques Bernoulli. La courbe de la chaîne pendue, la *chaînette*, s'écrit avec l'exponentielle du cours.

Pour les curieux du fond. La chaînette la plus simple est la fonction c définie sur $\mathbb{R}$ par :

$$\forall x \in \mathbb{R}, \quad c(x) = \frac{e^x + e^{-x}}{2}.$$

Sa convexité se lit en une ligne de F2 : en dérivant deux fois, $\forall x \in \mathbb{R}, c''(x) = \frac{e^x + e^{-x}}{2} = c(x) \geq 1 > 0$. La fonction est sa propre dérivée seconde, comme l'exponentielle est sa propre dérivée : convexe partout, et courbée en chaque point à proportion de sa hauteur. C'est cette proportion exacte qui en fait la courbe d'équilibre d'un fil pesant, là où la parabole, dont f'' est constante, échoue de peu.

Vrai ou légende ? Les tours de la Sagrada Família ont été conçues grâce à la maquette de chaînettes suspendues de Gaudí.

Faites votre pari. La réponse vous attend sur le compagnon numérique.

Brèves de coulisses

Trois éclats du chapitre qui n'ont pas trouvé leur place, trop beaux pour être coupés.

La courbe des ingénieurs. La clothoïde du cours a son héros discret : Arthur Talbot, professeur dans l'Illinois, qui décrit en 1891 la courbe de raccordement des voies ferrées, la même que la spirale d'Euler, et lui consacre en 1899 un traité entier. Le calcul restant pénible à la main, la courbe attendra les tables spécialisées puis l'ordinateur pour conquérir les autoroutes. Aujourd'hui elle est partout : chaque bretelle que vous prenez en car déroule sa courbure linéairement, de zéro au virage, pour que le chauffeur tourne son volant à vitesse constante.

Les épaules de la cloche. Le cours a calculé les points d'inflexion de la courbe en cloche $x \mapsto e^{-x^2}$: ils marquent ses « épaules », là où le dôme du sommet devient jupe évasée. Sur la cloche de Gauss des statistiques, ces épaules

tombent exactement à un écart type du centre, en $\mu - \sigma$ et $\mu + \sigma$. L'écart type n'est donc pas qu'un nombre dans une formule : il se voit, c'est l'abscisse où la cloche change de courbure. Savoir lire une inflexion, c'est savoir lire σ sur un dessin.

Le minimum sans piège. Pourquoi l'optimisation moderne aime-t-elle tant les fonctions convexes ? Parce qu'on ne s'y perd pas : sur une fonction convexe, tout minimum local est un minimum global, aucun creux intermédiaire ne peut piéger une descente. Itinéraires, logistique, apprentissage automatique : quand un problème peut se reformuler en problème convexe, sa résolution devient une promenade en bol, où descendre suffit. Quand il ne peut pas, on paie le prix des paysages à mille vallées.

LA SUITE DANS LES COULISSES NUMÉRIQUES

Trois tiroirs restent ouverts sur le compagnon numérique : la réponse au Vrai ou légende, le script qui détecte le pic d'une épidémie sur la courbe des cas cumulés, et le tableau de la course à trois entre la dichotomie, la corde et la tangente.