Bulletin AMQ

Association mathématique du Québec

Vol. LXVI, nº 2 — Mai 2026

Association mathématique du Québec

Association mathématique du Québec
Département de didactique
Pavillon Marie-Victorin
Université de Montréal
C.P. 6128, Succ. Centre-ville
Montréal (Québec) H3C 3J7

Courriel : info@amq.math.ca
Site : amq.math.ca

Comité exécutif

Présidente	Amélie Compagna	Université Laval
Vice-président	Alexandre Girouard	Université Laval
Trésorier	Yann Ricaud	Université Laval
Secrétaire	Véronique Boutet	Collège Durocher St-Lambert
Directeur	Frédéric Morneau-Guérin	Université TÉLUQ
Directice	Julie Tremblay	Cégep de Sherbrooke
Directreur	Christian Genest	Université McGill
Directeur	Nicolas Beauchemin	Collège de Bois-de-Boulogne

Conseil d'administration

Le CA est constitué des membres du Comité exécutif et des personnes suivantes :

Vacant	pour le niveau primaire
Vincent Rouleau	pour le niveau secondaire
Cédric Lamathe	pour le niveau collégial
Camille Melançon	pour le niveau collégial
Sarah Dufour	pour le niveau universitaire
Josiane Dussault	pour le niveau universitaire
Vacant	étudiant·e universitaire en mathématiques
Izabella Oliveira	pour le GDM
Frédéric Ouellet	pour le GRMS
Bruno Martin	hors maison d'enseignement

Comité de rédaction du Bulletin AMQ

Courriel : bulletin@amq.math.ca

Daniel Audet	Collège de Bois-de-Boulogne
France Caron	Université de Montréal
Yannick Delbecque	Cégep de Saint-Laurent
Paul Guertin	Collège André-Grasset
Frédéric Morneau-Guérin	Université TÉLUQ
Sophie René de Cotret	Université de Montréal

Le Bulletin AMQ (ISSN 0316-8832) est publié quatre fois l'an (mars, mai, octobre et décembre).
Montage de la couverture effectué à partir d'une affiche de Marie-Claude Asselin, conceptionmc.com.
Impression par Impressions KLM.

Dépôt légal, Bibliothèque et Archives nationales du Québec, 2026.
© Association mathématique du Québec

Association
mathématique
du Québec

Bulletin AMQ

LXVI, n° 2 — Mai 2026

Sommaire

Éditorial

Mathématiques et espoir : au-delà des nombres **4**
Amélie Compagna

AMQ en action

Concours AMQ – Ordre secondaire **7**
Concours AMQ – Ordre collégial **16**

Article

La méthode de bissection stochastique : le hasard fait-il mieux que le milieu ? **24**
Ludovick Bouthat, Philippe-André Luneau et Philippe Petitclerc,
Département de mathématiques et de statistique, Université Laval

Article

La filiation méconnue entre logique mathématique et linguistique chomskyenne **49**
Patrick Duffley, Département de langues, linguistique et traduction, Université Laval
Frédéric Morneau-Guérin, Département Éducation, Université TÉLUQ

Réflexions didactiques

Intérêt de la connaissance-obstacle pour l'enseignement **72**
Sophie René de Cotret, Université de Montréal

Chronique

Problèmes et jeux mathématiques **81**
Daniel Audet, Collège de Bois-de-Boulogne

Lu pour vous

Graph Theory in America: The First Hundred Years **88**
Frédéric Morneau-Guérin, Département Éducation, Université TÉLUQ

Éditorial

Mathématiques et espoir : au-delà des nombres

Amélie Compagna
Présidente de l'AMQ
Département de mathématiques et de statistique, Université Laval

En tant que mathématiciens et mathématiciennes, nous aimons bien les nombres. Que ce soit les nombres entiers, les nombres premiers, ou encore les nombres irrationnels. Parmi ces nombres, un nombre emblématique est connu depuis aussi longtemps que les gens font des mathématiques. Ce fameux ratio entre la circonférence d'un cercle et son diamètre, plus formellement connu par la lettre grecque π depuis le début du XVIII^{e} siècle.

$$\pi = 3{,}141592653\ldots$$

Le 14 mars est depuis maintenant presque 40 ans célébré comme la journée de pi, et cette journée coïncide depuis mars 2020, avec la Journée internationale des mathématiques. En 2026, le thème de cette journée était « Mathématiques et espoir ». Ce thème nous rappelle particulièrement que les mathématiques ne sont pas qu'un ensemble de nombres, d'outils abstraits ou de calculs formels. Non. Les mathématiques sont un véritable langage universel capable de rassembler les gens, et de traverser les frontières culturelles et géographiques.

En cherchant à sensibiliser le public, à soutenir l'éducation et à outiller la jeunesse face aux grands enjeux contemporains, la Journée internationale des mathématiques rejoint les engagements fondamentaux de l'AMQ. C'est pour cette raison que nous avons été fiers de nous associer à l'ISM et au CRM cette année encore afin d'offrir des subventions associées à la réalisation d'activités mathématiques riches, vivantes et accessibles à toutes et tous.

Nous sommes heureux de voir que tous les milieux de l'enseignement ont pu profiter des subventions offertes pour la journée des mathématiques. Voici brièvement quelques projets qui ont reçu l'aide de l'AMQ, de l'ISM et du CRM, de l'école primaire jusqu'à l'université !

Primaire
L'**école du Parc** a organisé une journée pédagogique sur le thème d'Olympiades mathématiques. L'événement s'est décliné en plusieurs espaces d'ateliers adaptés à tous les âges : la résolution d'énigmes, des jeux de réflexion et de logique (échecs, tangrams, sudokus), des activités créatives (peinture à numéros, bracelets de perles) et des zones de manipulation pour estimer ou comparer. De plus, un marché fictif et un atelier de cuisine ont permis aux enfants de manipuler de la fausse monnaie et de mesurer des ingrédients. Ils ont ainsi pu utiliser leurs compétences mathématiques, leur esprit de collaboration tout en ayant le plaisir d'apprendre par le jeu.

Secondaire
De son côté, l'**école secondaire Vanier** a fait vivre une journée d'activités variées à l'ensemble de ses élèves. Le programme comprenait une kermesse mathématique animée par le personnel enseignant, où chaque kiosque proposait des défis à saveur mathématique (échecs, sudokus, probabilités, géométrie) permettant de récolter des coupons de participation. Tous les élèves de l'école ont également pu assister au spectacle « Salto », offert par l'organisme Sciences et Mathématiques en Action (SMAC), alliant sciences, mathématiques et art du cirque. Sur l'heure du dîner, une part de tarte a été offerte à toutes et tous, accompagnée d'animations comme un concours de mémorisation des décimales de π, avant de conclure la journée par un grand tirage de prix (jeux de société, calculatrices et casse-têtes).

Collégial
Le **cégep de Sherbrooke**, quant à lui, a organisé un événement en trois volets. Tout au long de la semaine, le Centre des médias a accueilli un kiosque mettant en valeur une sélection d'ouvrages mathématiques afin de faire découvrir ce que les mathématiques ont à offrir. De plus, un rallye d'énigmes a également été organisé à travers le cégep, on pouvait y participer à son rythme et courir la chance de gagner des prix de participation. Finalement, leur Club Math a proposé une édition spéciale offrant des pizzas et des présentations sur le thème « Mathématiques et espoir » réalisées par des membres du corps professoral et de la communauté étudiante.

Universitaire
Finalement, du côté universitaire, l'**UQAM** a présenté une conférence grand public animée par la professeure Nadia Lafrenière de l'Université Concordia. Axé sur le thème de l'espérance en probabilités, cet exposé interactif a été conçu pour permettre au public de participer aux démonstrations plutôt que de seulement les observer. À travers trois expériences explorant les liens entre géométrie et hasard, la communauté universitaire de l'UQAM et le public ont ainsi pu expérimenter l'estimation du nombre π à l'aide d'observations expérimentales. Ouvert à toutes et à tous, l'événement a certainement su rendre ces concepts scientifiques profonds vivants, ludiques et accessibles.

En célébrant la Journée internationale des mathématiques sous le thème de l'espoir, nous célébrons aussi toutes celles et ceux – enseignants et enseignantes, étudiants et étudiantes, ou personnes passionnées – qui partagent jour après jour leur amour pour notre discipline.

Cet espoir réside également dans la façon dont les mathématiques nous aident à comprendre le monde qui nous entoure et nous permettent de révéler la beauté de la nature. Que ce soit la spirale logarithmique d'un coquillage, les fractales présentes dans un flocon de neige, ou encore le fragile équilibre de la dynamique des écosystèmes, les mathématiques font partie intégrante du langage de la nature. Elles offrent une fenêtre privilégiée pour émerveiller petits et grands, pour comprendre le vivant, et pour préserver la beauté qui nous entoure.

En espérant que vous avez toutes et tous passé une enrichissante Journée internationale des mathématiques !

Association
mathématique
du Québec

AMQ en action

Concours de l'Association mathématique du Québec
Ordre secondaire

Question 1. Fonction affine

Soit a et b deux nombres réels non nuls et soit f une fonction telle que $f(x) = ax + b$ pour tout x réel.

Déterminer les valeurs de a **et** b sachant que

$$f\left(\frac{a}{b}\right) = f\left(\frac{b}{a}\right)$$

et

$$f(a) = f\left(\frac{1-b}{a}\right) + f\left(\frac{-b^2}{a}\right).$$

Solution.

Avec la première équation, on trouve

$$a^2 = b^2.$$

Avec la seconde, on a

$$1 = a^2 + b^2.$$

Avec ces deux équations, on trouve que

$$a = \pm\frac{\sqrt{2}}{2}, \qquad b = \pm\frac{\sqrt{2}}{2}.$$

On a donc quatre solutions

$$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \quad \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right), \quad \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right), \quad \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right).$$

© Association mathématique du Québec

Question 2. Casse-tête

Un casse-tête standard rectangulaire a le même nombre de pièces qui se retrouvent sur les côtés que de pièces qui ne s'y retrouvent pas. **Combien ce casse-tête a-t-il de morceaux ?** Justifier clairement et donner toutes les solutions.

Note. Un casse-tête standard a le même nombre de pièces sur chaque ligne, et le même nombre sur chaque colonne.

Solution.

Soit n et m le nombre de lignes et de colonnes du casse-tête. Sans perte de généralité, on peut poser que $n \leq m$. Le nombre de pièces sur les côtés est $2n + 2m - 4$ alors que le nombre de pièces à l'intérieur est $(n-2)(m-2)$. On doit donc avoir

$$2(n + m - 2) = (n - 2)(m - 2).$$

En développant l'équation et en simplifiant, on trouve

$$nm - 4n - 4m + 8 = 0.$$

Une façon de transformer cette équation pour trouver ses solutions est de factoriser les termes en

$$(n - 4)(m - 4) = 8.$$

En posant $a = n - 4$ et $b = m - 4$, on peut trouver les solutions entières qui vérifient $ab = 8$.

a	b	n	m	nm
1	8	5	12	60
2	4	6	8	48

Il n'y a pas de solution négative, car on trouverait un nombre négatif de pièces de casse-tête.

On trouve donc deux solutions : 60 et 48.

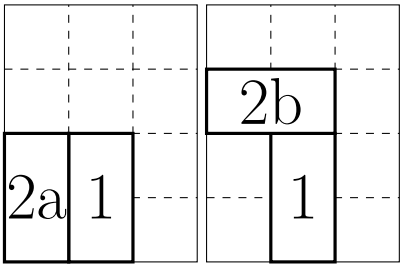
Question 3. Angie et Bastien

Angie et Bastien jouent au jeu suivant, inspiré du jeu de Tétris. À tour de rôle, en débutant par Angie, chacun place un domino 2 par 1 dans une grille rectangulaire de largeur 3 et de hauteur 4 en respectant le mouvement suivant.

Le domino est placé par le joueur en haut de la grille : horizontalement vis-à vis deux des trois colonnes, ou verticalement vis-à-vis une colonne. Puis, il est descendu vers le bas jusqu'à ce qu'il ne puisse plus descendre. Aucun autre mouvement n'est permis.

L'un des deux joueurs peut s'assurer de gagner toutes les parties en jouant bien. **Qui est ce joueur, et quelle est sa stratégie gagnante ? On doit bien expliquer la stratégie, et ce pourquoi elle mène toujours à la victoire.**

Exemples de coups possibles. Sur la figure, si Angie a inséré le domino 1 verticalement dans la deuxième colonne, alors Bastien pourrait par exemple placer le deuxième domino de sorte qu'il soit en position 2a ou 2b. (Ce ne sont pas les seuls coups possibles pour Bastien.) Le joueur qui est le dernier à pouvoir placer un domino a gagné la partie.



Déterminer quel joueur a une stratégie gagnante.

Solution.

Bastien a plus d'une stratégie gagnante. Nous en décrivons une.

Peu importe le premier coup joué par Angie, Bastien peut jouer pour compléter un carré 2 par 2. Puis, selon les coups de Angie, voici les réponses de Bastien.

Cas 1. Angie joue horizontalement directement au-dessus du carré 2 par 2, pour compléter un rectangle 2 par 3. Alors Bastien joue pour compléter un rectangle 2 par 4.

Il reste alors un seul coup possible pour Angie : verticalement dans la colonne libre. Bastien gagne en jouant le même coup pour compléter la colonne.

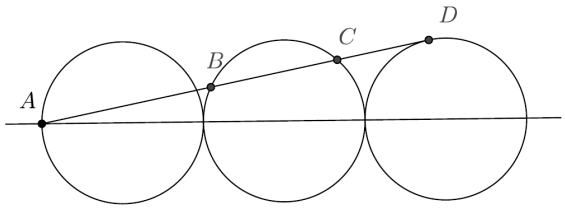
Cas 2. Angie joue horizontalement au-dessus d'une case du carré 2 par 2 et pour bloquer la colonne vide. Alors Bastien joue horizontalement au-dessus d'une case du domino de Angie et au-dessus d'une case vide (donc pas vis-à-vis le coup de Angie). Il ne reste plus de coup possible pour Angie.

Cas 3. Angie joue verticalement à côté du carré 2 par 2, donc pour compléter un rectangle 3 par 2. Bastien joue le même coup que Angie, et il reste un carré 2 par 2 de libre. Si Angie joue verticalement, Bastien pourra aussi le faire. De même si Angie joue horizontalement.

Cas 4. Angie joue verticalement au dessus du carré. Bastien joue verticalement dans la colonne libre au-dessus du carré, pour compléter un rectangle 2 par 4, comme dans le cas 1.

Question 4. Tangente et trois cercles

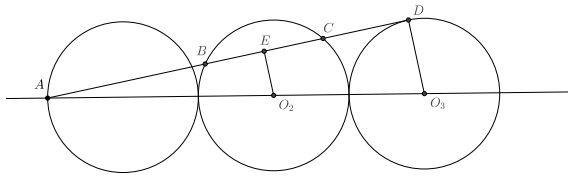
Sur la figure, les trois cercles de rayon 12 cm sont tangents entre eux. La droite AD est tangente au troisième cercle en D et coupe le deuxième cercle en B et C. **Déterminer la mesure de BC.**



Solution.

Soit O_2 et O_3 les centres des deuxième et troisième cercles.

On considère la perpendiculaire à $\overline{BC}$ issue de O_2, et on dénote par E son point d'intersection avec $\overline{BC}$. Le triangle $\triangle AEO_2$ est un triangle rectangle.



Le triangle $\triangle ADO_3$ est aussi un triangle rectangle par construction du point D. Les triangles $\triangle AEO_2$ et $\triangle ADO_3$ sont semblables, car ils ont les mêmes valeurs d'angles :

- $m\angle EAO_2 = m\angle DAO_3$,
- $m\angle AEO_2 = m\angle ADO_3 = 90°$.

On a donc

$$\frac{m\overline{AO_3}}{m\overline{AO_2}} = \frac{m\overline{DO_3}}{m\overline{EO_2}}.$$

On sait que $m\overline{AO_3} = 60\,\text{cm}$, $m\overline{AO_2} = 36\,\text{cm}$ et $m\overline{DO_3} = 12\,\text{cm}$. Après quelques calculs, on peut montrer que $m\overline{EO_2} = \frac{36}{5}$ cm. On utilise ensuite le théorème de Pythagore sur le triangle rectangle $\triangle BEO_2$ pour déterminer $m\overline{BE}$. On a

$$m\overline{BE} = \sqrt{12^2 - \left(\frac{36}{5}\right)^2}\,\text{cm} = \frac{48}{5}\,\text{cm}.$$

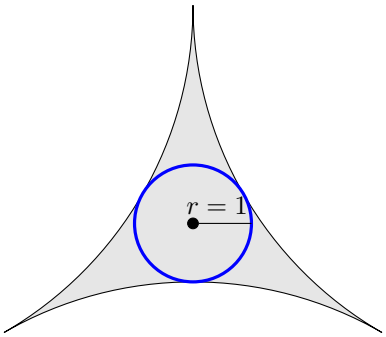
Finalement, puisque $\overline{EO_2}$ coupe perpendiculairement $\overline{BC}$ et qu'il passe sur le centre du cercle, alors il coupe la corde $\overline{BC}$ en son centre. Ainsi,

$$m\overline{BC} = 2m\overline{BE} = \frac{96}{5}\,\text{cm} = 19{,}2\,\text{cm}.$$

Question 5. Triarc

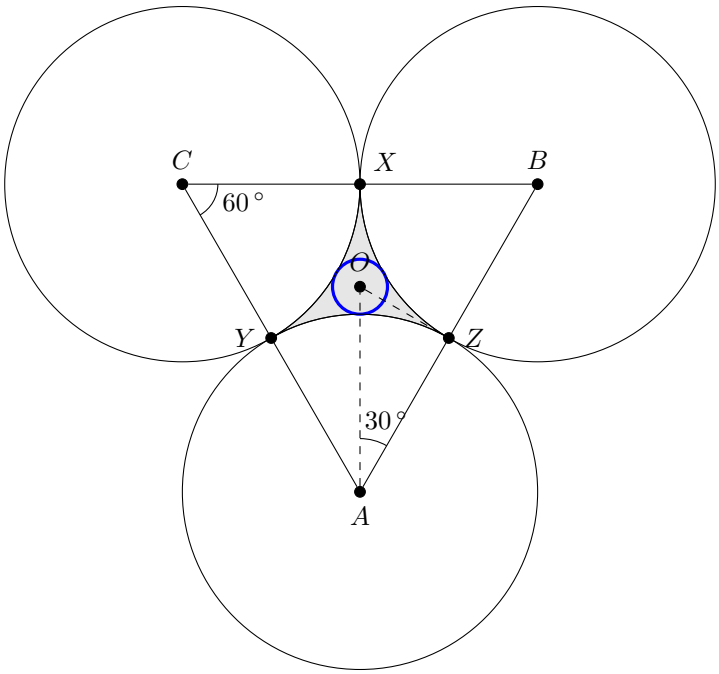
Sur la figure, un cercle de rayon 1 est tangent à trois grands arcs de cercles, tous les trois égaux et tangents entre eux. **Quel est le *périmètre* de la figure délimitée par ces trois arcs ?**

Note. Des arcs sont tangents entre eux si les cercles qui les contiennent sont tangents.



Solution.

Soit O le centre du petit cercle. On commence par dessiner les trois cercles dont on nous donne les arcs, puis on relie leurs centres respectifs, qu'on nomme A, B et C. Ces cercles sont tangents entre eux aux points X, Y et Z.



Soit $R = m\overline{AZ}$ le rayon de ces cercles. Le triangle ABC est équilatéral, de côté $2R$, et donc l'angle en chaque sommet est de 60°. Donc, $\angle OAZ = 30\,^\circ$. Puisque les tangentes d'un cercle sont perpendiculaires au rayon, alors $\angle OZA = 90\,^\circ$.

Le triangle OZA est donc rectangle avec un angle de $30\,^\circ$ et un de $60\,^\circ$. On sait par l'hypothèse du problème que son hypoténuse AO est de mesure $R + 1$. Son côté $\overline{OZ}$, opposé à l'angle de $30\,^\circ$, mesure la demie de l'hypoténuse soit $\frac{R+1}{2}$.

Le théorème de Pythagore donne alors $R^2 + (\frac{R+1}{2})^2 = (R+1)^2$, et on trouve

$$R^2 - 6R - 3 = 0.$$

On obtient finalement $R = 3 + 2\sqrt{3}$.

On peut aussi obtenir la valeur de R par trigonométrie :

$$\frac{\sqrt{3}}{2} = \cos(30\,^\circ) = \frac{m\overline{AZ}}{m\overline{AO}} = \frac{R}{R+1}.$$

On isole R pour trouver le rayon des trois cercles qui est

$$R = \frac{\sqrt{3}}{2 - \sqrt{3}} = 3 + 2\sqrt{3}.$$

Finalement, pour trouver le périmètre de la figure, il s'agit de trois fois la longueur d'un arc de $60\,^\circ$, ce qui équivaut à la longueur d'un arc de $180\,^\circ$, c'est-à-dire, la moitié de la circonférence du cercle. Le périmètre est donc :

$$p = \frac{1}{2}(2\pi R) = \frac{\pi\sqrt{3}}{2 - \sqrt{3}} = \pi\left(3 + 2\sqrt{3}\right).$$

Question 6. Les turbines

Des turbines hydrauliques fonctionnent grâce à l'eau contenue dans un grand réservoir. On suppose que toutes les turbines sont identiques et demandent le même débit constant d'eau pour fonctionner. Le réservoir est lui-même alimenté par un conduit qui fournit de l'eau à taux constant.

— Lorsque le réservoir est rempli à la moitié et que l'on continue à l'alimenter en eau par le conduit, alors on peut faire fonctionner 2 turbines pendant exactement 5 jours avant que le réservoir ne soit vide.
— Lorsque le réservoir est rempli au complet et que l'on continue à l'alimenter en eau par le conduit, alors on peut faire fonctionner 4 turbines pendant exactement 2 jours.

Quel doit être le niveau d'eau dans le réservoir pour que l'on puisse faire fonctionner 3 turbines pendant exactement 3 jours avant que le réservoir ne soit vide ?

Solution.

Soit R la quantité d'eau d'un demi-réservoir, T la quantité d'eau utilisée par une turbine en une journée, et E la quantité d'eau fournie par le conduit d'eau en une journée. On a alors les équations :

$$\begin{aligned} 10T &= R + 5E \\ 8T &= 2R + 2E \\ 9T &= xR + 3E \end{aligned}$$

En soustrayant la demie de la deuxième équation de la première, on obtient directement $6T = 4E$ et donc $T = \frac{2}{3}E$. On peut alors remplacer T par $\frac{2}{3}E$ dans la première équation pour obtenir $10 \cdot \frac{2}{3}E = R + 5E$. On isole R pour obtenir $R = \frac{5}{3}E$.

La troisième équation peut alors être exprimée uniquement en fonction de E, en remplaçant R et T. Cela donne : $6E = \frac{5x}{3}E + 3E$, d'où $x = \frac{9}{5}$. Comme R est le demi-réservoir, on obtient que le réservoir doit être rempli aux $\frac{9}{10}$.

Lauréates et lauréats du concours secondaire 2026

Voici la liste des participantes et participants ayant obtenu les 33 meilleurs résultats lors du concours 2026. Félicitations !

Nous tenons à remercier Catherine Bilodeau (Université Laval), Jérôme Fortier (Université McGill), Frédéric Gourdeau (Université Laval), Benoît Pouliot (InnovMetric), Andrée-Ann Pugin (Cégep de l'Outaouais), Kim Rioux-Paradis (Université Laval) et Vincent Rouleau (Pensionnat du Saint-Nom-de-Marie) qui ont contribué au bon déroulement du concours, que ce soit pour la gestion, la conception ou la correction du concours.

Nous remercions aussi Nicolas Berthiaume, de M2Learning, et Vincent Rouleau, du Pensionnat du Saint-Nom-de-Marie, qui ont rendu possible la participation de jeunes dont les institutions scolaires ne participaient pas au concours.

Nom	Rang	Institution
Chen Chen	1	Collège Jean de la Mennais
Travis Li	1	Collège Jean-de-Brébeuf
Zehan Li	1	Collège Jean-de-Brébeuf
Nanxuan Zhang	1	École internationale Lucille-Teasdale
Alex Xie	5	Collège Jean de la Mennais
Shanyuan Liu	6	Collège Sainte-Anne (Lachine)
Océan Cao	7	Collège Saint-Louis
Xiuqi Xu	7	École secondaire Joseph-François-Perrault
David Luo	9	Collège Sainte-Anne (Lachine)
Morgan Zhou	10	Collège Saint-Louis
Austin Haoning Tang	11	Collège Jean-Eudes
Siyuan Qiu	12	Collège Jean-Eudes
Zhixing Yuan	13	Collège Jean-Eudes
Kefeng Ding	14	Collège Jean de la Mennais
Michael Xia	15	Collège Jean-Eudes
Junteng Zhang	16	Collège Sainte-Anne (Lachine)
Aaron Chenshu Jiang	17	Collège Jean-de-Brébeuf
Aaron Liu	18	Collège Charlemagne
Marco Leonardo Bianucci	19	Collège Jean-de-Brébeuf
Alice Zhu	20	Collège Jean de la Mennais
Evan Aglat-Clayman	21	École secondaire Honoré-Mercier
Chuanshi Li	21	École internationale Lucille-Teasdale
Aaron Lou	23	École internationale de Montréal
Evan Xuan	23	Collège Jean-de-Brébeuf
Richard Runda Wang	25	Collège Jean-Eudes
Édouard Chevrette	26	Séminaire de Chicoutimi
Eric Deng	26	Collège Jean-de-Brébeuf
Ruijing Huang	26	Collège Jean de la Mennais
Bhuuvanesh Hurchand	29	École internationale Lucille-Teasdale
Huaipu Kang	30	Collège Jean-de-Brébeuf
Samuel Rometsch	30	Séminaire de Chicoutimi
Ole Sweetnam	30	Collège Sainte-Anne (Dorval)
Xu Leran	33	Collège Jean-de-Brébeuf
Yiyi Liu	33	Collège Jean-de-Brébeuf
Jiacheng Tan	33	École internationale Lucille-Teasdale

Association
mathématique
du Québec

Concours de l'Association mathématique du Québec
Ordre collégial

Question 1. La grosse division

Donner le reste de la division par 7 du nombre 999 998 997 . . . 003 002 001.

Solution.

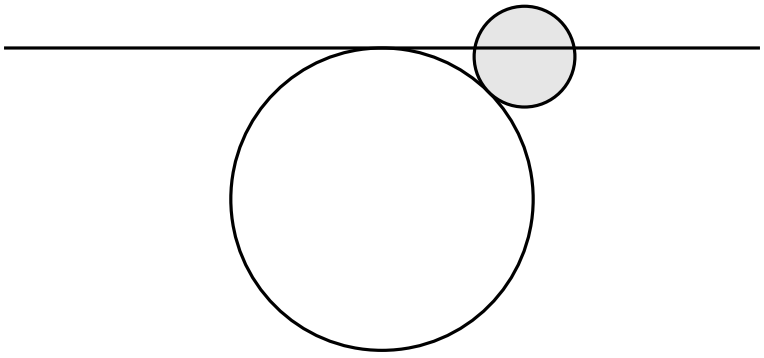
Pour faciliter la lecture de la solution, nous utiliserons la notation $a \underset{7}{\equiv} b$ qui signifie que a et b ont le même reste de division par 7. Par exemple, on a $13 \underset{7}{\equiv} 6$, $36 \underset{7}{\equiv} 1$ et $1000 \underset{7}{\equiv} 6 \underset{7}{\equiv} -1$. On a donc

$$\begin{aligned}
999\,998\,997\ldots003\,002\,001 &= 999 \cdot 1000^{998} + 998 \cdot 1000^{997} + \ldots + 3 \cdot 1000^2 + 2 \cdot 1000^1 + 1 \\
&\underset{7}{\equiv} 999 \cdot (-1)^{998} + 998 \cdot (-1)^{997} + \ldots + 3 \cdot (-1)^2 + 2 \cdot (-1)^1 + 1 \\
&= (999 - 998) + (997 - 996) + \ldots + (3 - 2) + 1 \\
&= 1 + 1 + \ldots + 1 + 1 \\
&= 500 \\
&\underset{7}{\equiv} 3.
\end{aligned}$$

Le reste de division recherché est donc 3.

Question 2. L'horizon

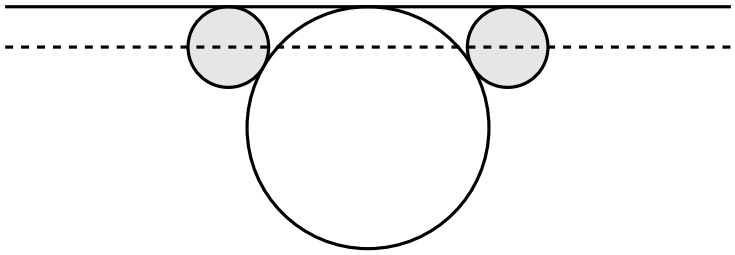
On trace deux disques tangents l'un à l'autre, comme sur le dessin. Le rayon du disque blanc est trois fois plus grand que le rayon du disque gris. On trace ensuite toutes les tangentes possibles au disque blanc. Quelle fraction de ces tangentes intersecte le disque gris ?



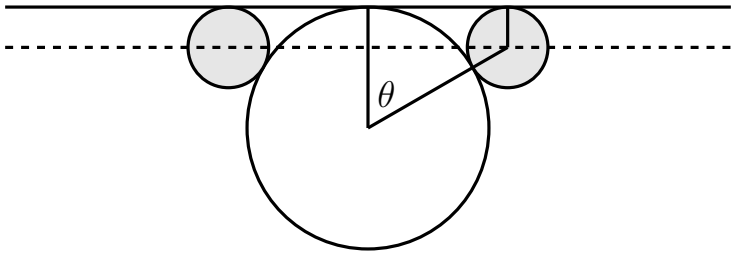
 © Association mathématique du Québec

Solution.

En reformulant la question, on peut dire que le problème est de trouver la proportion pour laquelle le centre du disque gris est en haut de la ligne pointillée.



Si on rajoute les rayons, on trouve un triangle rectangle.



Si R_1 et R_2 sont les rayons des disques blanc et gris respectivement, alors

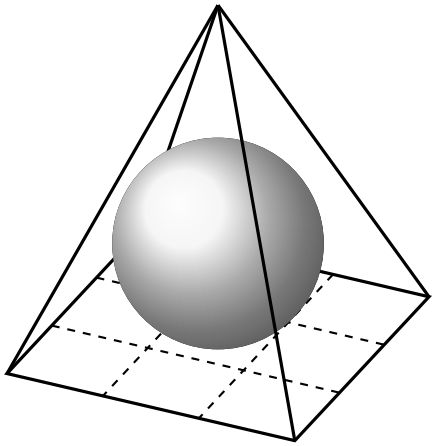
$$\cos(\theta) = \frac{R_1 - R_2}{R_1 + R_2} = \frac{1}{2}.$$

Ainsi $\theta = \pi/3$ et donc la proportion est

$$P = \frac{2\theta}{2\pi} = \frac{1}{3} = 33{,}\overline{3}\,\%.$$

Question 3. La sphère pyramidale

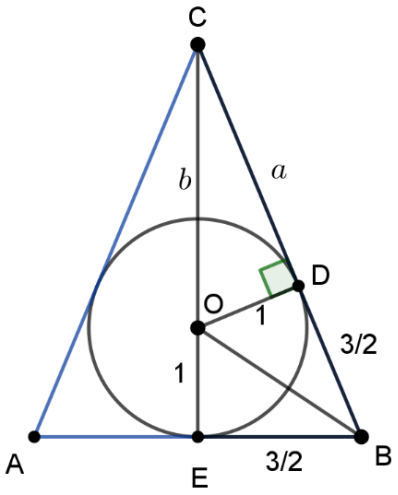
Soit une pyramide droite à base carrée dont la base est un carré 3 par 3. Une sphère de rayon 1 est tangente aux cinq faces de la pyramide. Quelle est la hauteur de la pyramide ?



Solution.

Par symétrie, la sphère est tangente au centre du carré et aux hauteurs (ou médianes) des quatre faces latérales.

Soit la coupe transversale qui coupe la base au centre. On pose C le sommet de la pyramide, A et B les points à l'intersection des faces latérales, de la base et de la coupe, O le centre de la sphère, E le point d'intersection entre la base et la sphère et D le point d'intersection de la sphère et de la face qui contient B.



Le triangle $\triangle OEB$ est congru au triangle $\triangle ODB$, car ce sont deux triangles rectangles partageant un côté, donc $m\overline{DB} = \frac{3}{2}$.

Posons $m\overline{OC} = b$ et $m\overline{DC} = a$. Par Pythagore, on a :

$$(1)\quad b^2 = a^2 + 1^2 \qquad \text{et} \qquad (2)\quad (b+1)^2 + \left(\frac{3}{2}\right)^2 = \left(a + \frac{3}{2}\right)^2.$$

La deuxième équation devient :

$$(3)\quad b^2 + 2b + 1 + \frac{9}{4} = a^2 + 3a + \frac{9}{4}.$$

En substituant b^2 dans l'équation (3) par sa valeur dans la première équation, on a :

$$(a^2 + 1) + 2b + 1 + \frac{9}{4} = a^2 + 3a + \frac{9}{4}.$$

Qui se simplifie par

$$b = \frac{3a}{2} - 1.$$

En substituant dans la première équation, on a :

$$\left(\frac{3a}{2} - 1\right)^2 = a^2 + 1$$

et si on développe, on obtient :

$$5a^2 - 12a = 0.$$

On élimine la valeur $a = 0$, donc $a = \frac{12}{5}$, on trouve $b = \frac{13}{5}$. Donc la hauteur de la pyramide est $m\overline{EC} = 1 + b = \frac{18}{5}$ unités.

Question 4. Le jeu de Kenda

Kenda a un jeu de cartes numérotées de 1 à n. Elle choisit une carte au hasard, la remet dans le paquet et brasse le paquet. Elle recommence ainsi jusqu'à avoir fait n piges. Soit P_n la probabilité que les n cartes pigées contiennent exactement $n-1$ numéros différents. Pour quelle valeur de n, la probabilité P_n est maximale ?

Solution.

Le nombre de manières d'avoir seulement 2 nombres identiques est

$$\binom{n}{2} \times n \times (n-1) \times \cdots \times 2 = \binom{n}{2} \times n!.$$

Puisqu'il y a n^n possibilités de tirages différents, alors

$$P_n = \frac{n(n-1)}{2}\frac{n!}{n^n} = \frac{(n-1)^2\,(n-2)!}{2n^{n-2}}.$$

Pour déterminer le maximum de P_n, on peut d'abord calculer les premières valeurs de P_n.

n	P_n
2	1/2
3	2/3
4	9/16
5	48/125
⋮	⋮

Le tableau nous donne l'intuition que le maximum de P_n est en $n = 3$. Pour le démontrer, on a que pour $n \geq 5$,

$$P_n = \frac{(n-1)^2\,(n-2)!}{2n^{n-2}} = \frac{(n-1)^2}{n^2}\frac{(n-2)!}{n^{n-5}}\frac{1}{2n} \leq \frac{3!}{2n} \leq \frac{3}{5} \leq \frac{2}{3} = P_3.$$

Ainsi donc, P_n est maximal en $n = 3$.

Question 5. De cinq à zéro

Soit la fonction f définie sur tous les entiers telle que pour tout couple d'entiers (m,n), on a :

$$2f(mn) + f(n-1) + 4n + 2m^2 = f(m)f(n) + f(1) + 1.$$

Trouver la valeur de $f(5)$ sachant que $f(0) = 1$.

Solution.

Remplaçons m par 0. L'équation devient

$$2f(0) + f(n-1) + 4n = f(0)f(n) + f(1) + 1.$$

Substituons $f(0)$ par 1 selon l'hypothèse

$$\begin{aligned} 2 + f(n-1) + 4n &= f(n) + f(1) + 1, \\ f(n) &= f(n-1) + 4n + 1 - f(1). \end{aligned} \tag{1}$$

On remplace n par 1, on a

$$f(1) = f(0) + 4 + 1 - f(1).$$

Donc $2f(1) = 5 + f(0)$ et $f(1) = 3$. L'équation (1) devient :

$$f(n) = f(n-1) + 4n - 2. \tag{2}$$

Avec $f(1) = 3$, on trouve récursivement à partir de l'équation (2) les valeurs $f(2) = 9$, $f(3) = 19$, $f(4) = 33$ et $f(5) = 51$.

Note : On peut montrer que la fonction $f(n) = 2n^2 + 1$ satisfait la relation.

Question 6. Réflexion Logo-Expo

Soit $k > 0$ une constante. Trouver la valeur de k pour laquelle l'équation

$$e^{kx} = \ln\left(x^{1/k}\right)$$

n'a qu'une seule solution.

Solution.

Soit

$$f(x) = e^{kx} \qquad \text{et} \qquad g(x) = \ln\left(x^{1/k}\right).$$

On peut d'abord remarquer que pour $x > 0$, ces deux fonctions sont l'inverse l'une de l'autre. (On peut ignorer le cas $x \leq 0$ car $g(x)$ n'est pas définie dans ce cas.) En effet,

$$g(f(x)) = x.$$

Ainsi donc, toutes solutions de l'équation originale se trouve, par symétrie, sur la droite d'équation $y = x$.[(1)] Nous pouvons donc seulement trouver les solutions telles que

$$e^{kx} = x.$$

On note

$$h(x) = e^{kx} - x.$$

Si x_0 est la solution unique de $h(x)$, alors

$$h(x_0) = e^{kx_0} - x_0 = 0, \tag{1}$$
$$h'(x_0) = ke^{kx_0} - 1 = 0. \tag{2}$$

En utilisant l'égalité (1) dans (2), on trouve

$$x_0 = \frac{1}{k}.$$

En réutilisant (1), on trouve que $x_0 = e$ et donc que $k = 1/e$.

[1]**Note :** Cette affirmation n'est pas vraie en général, il suffit de prendre l'involution $i(x) = -x$ pour s'en convaincre. Cette fonction a la propriété que $i(i(x)) = x$ et que toutes les valeurs dans $\mathbb{R}$ sont bien sûr solutions de $i(x) = i(x)$. Ces solutions ne sont pas toutes sur la droite $y = x$. Cela contredit ainsi ladite affirmation.

Pour être plus précis dans notre cas, il faut aussi prendre en compte que f est monotone croissante. En effet, supposons que l'équation $f(x) = g(x)$ a une solution x_0 qui n'est pas sur la droite $y = x$. Sans perte de généralité, on peut supposer que $f(x_0) > x_0$. Ainsi,

$$g(x_0) = f(x_0) > x_0.$$

Puisque f est monotone croissante, alors

$$x_0 = f(g(x_0)) > f(x_0).$$

Cela nous donne une contradiction.

Lauréates et lauréats du concours collégial 2026

Voici la liste des participants et participantes ayant obtenu les 31 meilleurs résultats lors du concours 2026. Félicitations !

Nom	Rang	Institution
Francis Audet	1	Cégep de Bois-de-Boulogne
Victor Liu	2	Cégep John Abbott
Maxime Goulet	3	Collège Jean-de-Brébeuf
Ryan Bossou Tihoun	4	Collège Marianopolis
Holly Wei	4	M2Learning
Yuqing Zhou	4	M2Learning
Arthur Delcourt	7	Cégep du Vieux Montréal
Youji Wiliam Li	8	Collège Marianopolis
Neilson Tao	8	Collège Marianopolis
Cédric Poulin	10	Cégep de Sherbrooke
Evan Lacasse	11	Cégep de Lanaudière à l'Assomption
Cédric Gaudet	12	Collège André-Grasset
Eric Tanase	12	Collège Jean-de-Brébeuf
Xavier Bélair	14	Collège Marianopolis
Maksim Alekseev	15	Cégep Édouard-Montpetit
Christophe Lewis	15	Cégep Champlain (Saint-Lambert)
Sevan Boudreau	17	Collège Jean-de-Brébeuf
Gaël Simard	18	Cégep de Granby
Charles-André Nolet	19	Cégep de Sherbrooke
Xinyi Wang	19	Collège Jean-de-Brébeuf
Feihong Qin	19	Collège Marianopolis
Oliver Rusu	22	Collège Marianopolis
Yu He Wang	22	Collège Marianopolis
Nicolas Cappiello	24	Cégep Édouard-Montpetit
Brent Suter	24	Collège Dawson
Alex Meng	24	Collège Marianopolis
Étienne Baribeault	24	Collège Mérici
Dahwi Joo	28	Collège Jean-de-Brébeuf
Léonard Jobin	29	Cégep Champlain (Saint-Lambert)
Zhiheng Liu	29	Collège Jean-de-Brébeuf
Shida Zheng	29	Collège Marianopolis

Association
mathématique
du Québec

La méthode de bissection stochastique : le hasard fait-il mieux que le milieu ? [1]

Ludovick Bouthat
Philippe-André Luneau
Philippe Petitclerc
Département de mathématiques et statistiques, Université Laval

Résumé

La méthode de bissection est l'un des algorithmes de recherche de racines les plus simples et les mieux compris. Nous en étudions une généralisation naturelle où le point de coupure, plutôt que d'être fixé au milieu de l'intervalle, est tiré aléatoirement selon une distribution arbitraire. Nous déterminons son taux de convergence moyen lorsque la racine est supposée uniformément distribuée dans l'intervalle initial, et montrons qu'il ne dépend de la loi du point de coupure c qu'à travers la quantité $\mathbb{E}[c(1-c)]$. Les prédictions théoriques, fondées sur des arguments probabilistes simples afin de rester accessibles à un large public mathématique, sont finalement confirmées par des simulations numériques.

Mots clés (MSC2020) : 65-01 Exposé introductif sur l'analyse numérique ; 65H05 Méthodes numériques en analyse non linéaire ; 60G10 Processus stochastiques stationnaires ; 37H99 Systèmes dynamiques aléatoires (divers) ; 37A50 Systèmes dynamiques ergodiques aléatoires

1. Introduction

En mathématiques, plusieurs questions reviennent sans cesse sous différentes formes : à quel moment deux quantités deviennent-elles égales ? À partir de quel point une expression passe-t-elle de valeurs positives à des valeurs négatives ? Où une fonction atteint-elle un maximum ? Ces questions, malgré leurs différentes formulations, cachent toutes la même idée : déterminer les *racines* d'une fonction, c'est-à-dire les valeurs pour lesquelles une fonction s'annule.

Bien avant l'apparition de l'algèbre moderne, on trouvait déjà des méthodes pour déterminer les racines d'une fonction. Le *papyrus Rhind*, datant de 1650 avant notre ère, présente 91 problèmes mathématiques variés dont plusieurs consistent à trouver une *quantité* à partir de relations numériques simples [8]. Le problème 26, par exemple, revient à chercher une quantité qui, ajoutée à son quart, donne 15. Cette équation linéaire correspond ainsi dans le langage moderne à trouver la racine de la fonction $f(x) = x + x/4 - 15$.

La solution proposée dans le papyrus Rhind utilise un algorithme systématique : la *méthode regula falsi* (ou *méthode de fausse position*). Celle-ci consiste à essayer une valeur candidate,

1. Ce travail a été soutenu par le CRSNG.

 © Association mathématique du Québec

observer l'écart avec le résultat attendu, puis corriger la valeur candidate proportionnellement. Cette méthode apparaît notamment dans les problèmes 24 à 29 du papyrus Rhind [2, p. 11].

Cet algorithme, couramment utilisé partout dans le monde jusqu'au XVII^{e} siècle [5], repose toutefois sur une hypothèse clé, à savoir l'*existence* même de la racine. Pendant longtemps, l'idée qu'une racine doit exister entre deux valeurs de signes opposés a été traitée comme une évidence géométrique. Si une courbe passe d'une valeur négative à une valeur positive, comment pourrait-elle ne pas traverser l'axe horizontal ? Cette intuition est fondée, mais cache une hypothèse essentielle : la *continuité*.

Cette évidence visuelle ne suffisait toutefois pas au mathématicien polymathe Bernard Bolzano. Celui-ci considérait qu'une preuve en analyse ne devait pas introduire des concepts étrangers au domaine, comme des considérations géométriques ou visuelles. Elle devait montrer, à partir de notions purement analytiques, pourquoi le passage d'une valeur négative à une valeur positive force l'existence d'une valeur intermédiaire nulle. Sa démonstration, publiée en 1817 dans son mémoire *Démonstration purement analytique du théorème : entre deux valeurs quelconques qui donnent deux résultats de signes opposés se trouve au moins une racine réelle de l'équation* [1], répond précisément à cette exigence.

La preuve de ce célèbre résultat, aujourd'hui appelé *théorème des valeurs intermédiaires*, repose sur une méthode itérative simple. Si f est une fonction continue sur un intervalle $[a,b]$, avec

$$f(a) < 0 \qquad \text{et} \qquad f(b) > 0,$$

alors on veut montrer qu'il existe un point $r \in [a,b]$ tel que $f(r) = 0$.

Pour ce faire, Bolzano commence par couper l'intervalle $[a,b]$ en deux. Si la fonction vaut zéro au milieu, la racine est trouvée. Sinon, le milieu donne une valeur ou bien positive, ou bien négative. On conserve alors la moitié de l'intervalle dont les extrémités sont encore des valeurs de signes opposés. En répétant ce procédé, une suite d'intervalles emboîtés est obtenue, de plus en plus petits, dans lesquels le changement de signe persiste. Comme leur longueur tend vers zéro, ces intervalles se resserrent autour d'un unique point r.

Bolzano argumente que ce point est bien une racine. D'un côté des intervalles, les valeurs de f restent négatives ou nulles ; de l'autre, elles sont positives ou nulles. Or, comme f est continue, les valeurs de la fonction aux deux extrémités doivent tendre vers la même valeur, soit $f(r)$. Comme cette valeur ne peut pas être simultanément strictement négative et strictement positive, elle doit être nulle : r est donc une racine de f.

La preuve de Bolzano marque ainsi un tournant : elle transforme une évidence visuelle en argument rigoureux utilisant uniquement des outils de l'analyse. En découpant progressivement l'intervalle où le changement de signe se produit, Bolzano montre que la racine peut être enfermée dans des intervalles de plus en plus petits. Le résultat est donc un *théorème d'existence*, puisqu'il affirme l'existence d'une racine sans dire où se trouve celle-ci. Toutefois, sa preuve se transpose directement en un algorithme simple pour localiser une racine d'une fonction continue f dans un intervalle donné : la *méthode de bissection*.

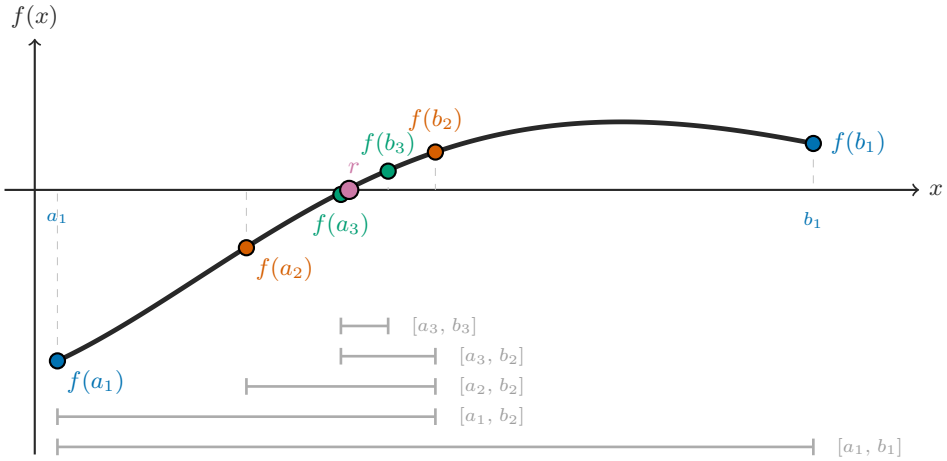


FIGURE 1 – Méthode itérative de la bissection

Dans sa forme classique, la méthode de bissection débute avec un intervalle $[a,b]$, où les valeurs de f sont de signes opposés aux extrémités. À chaque étape, on calcule le point milieu

$$m = \frac{a+b}{2},$$

puis on évalue $f(m)$. On remplace ensuite l'intervalle de départ par la moitié pour laquelle il y a toujours un changement de signe, à savoir $[a,m]$ ou $[m,b]$. Notons que si $f(m) = 0$, alors on choisi arbitrairement l'intervalle$[m,b]$. Après n étapes, l'intervalle obtenu possède une longueur égale à $2^{-n}(b-a)$. Ainsi, même si la racine exacte demeure inconnue, nous pouvons affirmer qu'elle se trouve dans un intervalle dont la taille peut être rendue aussi petite que désiré.

Si les points de coupe c_n sont considérés comme des approximations de la racine r, alors l'erreur $e_n = |c_n - r|$ est toujours bornée supérieurement par la demi-longueur de l'intervalle, c'est-à-dire

$$e_n \leq 2^{-(n+1)}(b-a).$$

Il s'ensuit ainsi que cet algorithme de recherche de racine converge linéairement avec un taux de convergence égal à $1/2$, c'est-à-dire que

$$e_{n+1} \approx \frac{1}{2} e_n, \quad \text{ou encore} \quad e_n \approx \frac{1}{2^n} e_0.$$

Malgré le fait que la méthode de bissection est l'un des algorithmes les plus anciens pour trouver une racine d'une fonction continue, elle demeure encore aujourd'hui l'un des plus simples. Plusieurs autres algorithmes découverts depuis, comme la méthode de Newton, convergent nettement plus rapidement (à savoir, quadratiquement plutôt que linéairement). Or, grâce à sa simplicité et au faible nombre d'hypothèses requises pour son application, celle-ci sert souvent de point de comparaison pour évaluer des méthodes de recherche de racines plus sophistiquées.

En particulier, la méthode de bissection est depuis devenue un incontournable de l'analyse numérique et une partie intégrante de l'enseignement de ce domaine au niveau universitaire. C'est dans ce contexte, davantage pédagogique que technique, que le problème de recherche présenté dans cet article a vu le jour.

En effet, en présentant la méthode de bissection à des étudiants, nous avons tenté de motiver l'algorithme en soulevant la question suivante : comment choisir une approximation d'une racine dont nous savons seulement qu'elle se trouve dans un certain intervalle $[a,b]$? De ce point de vue, choisir le point milieu n'est pas la seule option raisonnable : n'importe quel point de l'intervalle pourrait servir de candidat, et choisir un tel point au hasard paraît tout aussi intuitif. Cela soulève immédiatement une question naturelle : que se passe-t-il si la règle déterministe du point milieu est remplacée par un choix aléatoire ?

Cette question est vaste et possède plusieurs facettes. Toutefois, il n'est pas difficile de se convaincre que l'argument de Bolzano demeure valide sous cette variante et donc que l'algorithme associé converge encore vers une racine. Néanmoins, alors qu'il était aisément possible de déduire la longueur de l'intervalle après n étapes de la méthode de bissection, déterminer cette longueur dans le cas de la variante stochastique est un problème beaucoup plus difficile.

Il semble malgré tout raisonnable de penser que, si le point de coupure est choisi selon une distribution symétrique autour du point milieu, la méthode conserve, du moins en moyenne, une partie du comportement de la bissection classique. En effet, l'algorithme de bissection stochastique produit parfois des intervalles plus grands que la méthode classique ne le fait, et parfois des intervalles plus petits (voir la figure 2). Si l'on suppose que la racine est uniformément distribuée sur l'intervalle initial $[a,b]$, on pourrait alors s'attendre à ce que ces deux effets se compensent, et que le taux de convergence moyen coïncide avec celui du cas déterministe. Vérifier cette hypothèse correspond alors à répondre à la question : *en moyenne, sommes-nous plutôt chanceux, malchanceux, ou est-ce que les différents cas s'équilibrent* ?

L'objectif principal de cet article est de répondre à cette question en déterminant explicitement le taux de convergence moyen pour l'algorithme de bissection aléatoire lorsque la racine est initialement distribuée uniformément dans l'intervalle $[a,b]$. Pour ce faire, nous débutons dans la section 2 par introduire formellement l'algorithme que nous appellerons *méthode de bissection stochastique*. Nous rappelons également quelques concepts probabilistes élémentaires afin d'assurer l'accessibilité de nos résultats. Un lecteur familier avec ceux-ci pourra aisément ignorer cette section, quitte à y revenir au besoin.

Ensuite, à l'aide d'arguments probabilistes et géométriques élémentaires, nous obtenons dans la section 3 des formules explicites pour la fonction de répartition du facteur de réduction de l'intervalle à chaque étape de l'algorithme, et nous en déduisons des expressions pour son espérance et sa variance. De plus, nous montrons que la distribution uniforme est invariante pour le processus de position normalisée de la racine, *quelle que soit la distribution des coupures*. En d'autres termes, si l'on suppose la racine uniformément distribuée sur l'intervalle initial $[a,b]$, alors elle le sera encore sur tous les intervalles plus petits obtenus par le processus itératif. Ces résultats nous permettent ainsi de formuler des expressions explicites pour le taux de convergence moyen de la méthode de bissection stochastique, dépendant seulement de la quantité $\mathbb{E}[c(1-c)]$, et donc de répondre à la question formulée dans le titre de cet article.

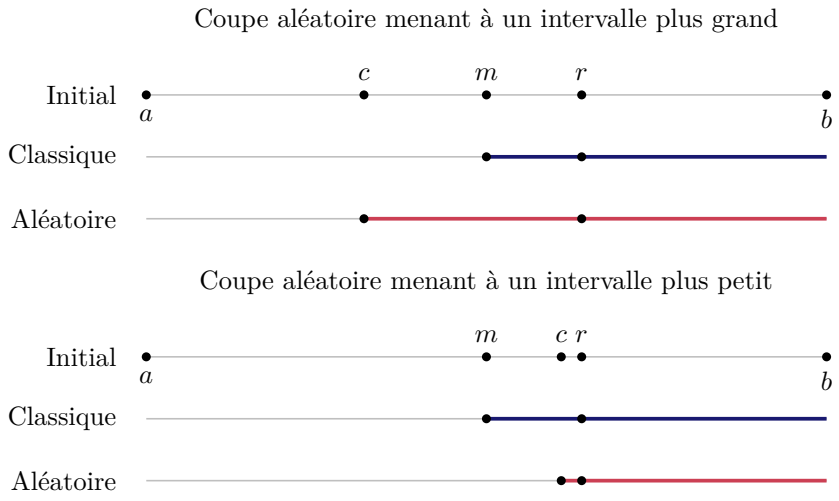


FIGURE 2 – Comparaison de l'intervalle conservé selon le point de coupe et la méthode considérée

Ces résultats sont finalement illustrés numériquement dans la section 4. En considérant différentes caractéristiques qualitatives des distributions du point de coupe, comme l'asymétrie ou la concentration de la masse, les expériences numériques présentées permettent d'évaluer la robustesse du comportement de convergence prédit et d'illustrer certaines tendances générales dans la performance de l'algorithme.

2. Rappels et définitions

2.1. La méthode de bissection stochastique

Considérons une fonction continue $f : \mathbb{R} \to \mathbb{R}$ possédant une racine r, c'est-à-dire telle que $f(r) = 0$, où la valeur de r est inconnue. La méthode de bissection permet d'approximer la valeur de r, dès lors que cette racine se trouve dans un intervalle $[a,b]$ où la fonction f prend des signes opposés aux deux extrémités. On définit alors la *fonction de coupe* comme le point milieu de l'intervalle :

$$\text{coupe}(a,b) = \frac{a+b}{2}.$$

La variante stochastique considérée dans cet article est analogue à la méthode de bissection classique, à une modification essentielle près : les points de coupe c_n, plutôt que d'être toujours le milieu de l'intervalle, sont tirés aléatoirement selon une distribution $\mathcal{D}$ sur l'intervalle $[a_n,b_n]$:

$$\text{coupe}(a_n,b_n) = c_n \sim \mathcal{D}(a_n,b_n), \qquad n \in \mathbb{N}.$$

Algorithme 1 Méthode de bissection

Entrée : a,b t.q. $a < b$, f t.q. $f(a)f(b) < 0$ ayant une racine $r \in [a,b]$, $\varepsilon > 0$, $N > 0$
Sortie : $r \in [a_n,b_n]$

```
n ← 0
a_n ← a
b_n ← b
tant que b_n − a_n ≥ ε et n < N faire
    c_n ← coupe(a_n,b_n)
    si f(a_n)f(c_n) < 0 alors
        a_{n+1} ← a_n
        b_{n+1} ← c_n
    sinon
        a_{n+1} ← c_n
        b_{n+1} ← b_n
    fin si
    n ← n + 1
fin tant que
retourne [a_n,b_n]
```

L'objectif principal de cet article est de déterminer le *taux de convergence moyen* d'une telle méthode. Sans perte de généralité, l'intervalle initial est choisi comme étant [0,1], puisqu'une fonction possédant une racine dans $[a,b]$ peut être ramenée, par une transformation affine, à une fonction dont la racine se trouve dans [0,1]. Pour simplifier l'analyse des points de coupe et des extrémités, nous étudions l'algorithme 2 plutôt que l'algorithme 1. Cet algorithme est essentiellement équivalent au précédent, à la différence près qu'après chaque itération, l'intervalle conservé est renormalisé vers [0,1]. La racine est donc elle aussi exprimée à chaque étape par sa position relative dans le nouvel intervalle.

Notons que l'algorithme 2 n'est pas destiné à être utilisé en pratique, puisque la valeur de r est généralement inconnue. De plus, on remarque que l'algorithme s'applique également aux fonctions ayant plusieurs racines dans $[a,b]$, mais nous supposerons ici par simplicité que la racine est unique. Finalement, à des fins de clarté, nous notons que la transformation de normalisation dans l'algorithme 2 peut être explicitée sous la forme $r_{n+1} = T(c_n,r_n)$, où

$$T(c,r) = \begin{cases} \frac{r}{c} & \text{si } c \geq r, \\ \frac{r-c}{1-c} & \text{si } c < r. \end{cases} \tag{1}$$

Bien que ce ne soit pas nécessaire pour notre analyse, nous notons pour le lecteur intéressé que cette transformation correspond à l'application dyadique asymétrique [10], avec paramètre c (voir la figure 3). L'itération de l'*application dyadique* dans le cas $c = 1/2$ est un système dynamique chaotique bien connu, topologiquement mélangeant [6, 9]. Cette dernière propriété découle du fait que T dilate les deux sous-intervalles $[0,c]$ et $[c,1]$ vers [0,1], mélangeant ainsi les points issus de part et d'autre de la coupure.

Algorithme 2 Méthode de bissection normalisée

Entrée : Un point $r \in [0,1]$, $\varepsilon > 0$, $N > 0$
Sortie : ℓ_n et L_n, le facteur de normalisation et la longueur de l'intervalle à l'étape n, resp.

```
n ← 0
a_n ← 0
b_n ← 1
ℓ_n ← ∞
L_n ← 1
r_n ← r
tant que L_n ≥ ε et n < N faire
    c_n ← coupe(0,1)
    si c_n > r_n alors
        a_{n+1} ← 0
        b_{n+1} ← c_n
    sinon
        a_{n+1} ← c_n
        b_{n+1} ← 1
    fin si
    r_{n+1} ← (r_n − a_{n+1}) / (b_{n+1} − a_{n+1})      ▷ Normalisation de la racine dans l'intervalle [0,1]
    ℓ_{n+1} ← (b_{n+1} − a_{n+1})
    L_{n+1} ← ℓ_{n+1} L_n
    n ← n + 1
fin tant que
retourne ℓ_n, L_n
```

2.2. Rappels de probabilité

Une variable aléatoire X formalise une quantité soumise au hasard. À chaque *événement* associé à X, par exemple « la valeur de X se trouve entre $1/2$ et $3/4$ », on attribue une valeur entre 0 et 1, appelée *probabilité*. Une probabilité de 1 signifie que l'événement est presque certain, tandis qu'une probabilité de 0 signifie qu'il est presque impossible. [2]

Une variable aléatoire est décrite par sa *fonction de répartition*, qui dicte la probabilité que X appartienne à un intervalle donné. Plus précisément, la fonction de répartition F est définie par

$$F(x) = \mathbb{P}[X \leq x],$$

à savoir la probabilité que la valeur de X soit inférieure ou égale à x. Cette probabilité sera dictée par la *loi* de X, par exemple, la loi uniforme $\mathcal{U}(0,1)$ sur $[0,1]$, pour laquelle $\mathbb{P}[X \leq x] = x$.

Les variables aléatoires considérées dans cet article prennent toujours leurs valeurs dans $[0,1]$.

2. Les notions de presque certain et presque impossible font référence au concept d'événements de mesure nulle. Formellement, un événement de probabilité 0 n'est pas impossible, mais son poids est négligeable au sens de l'intégration. C'est pourquoi on dit qu'un événement de probabilité 1 se réalise « presque sûrement ».

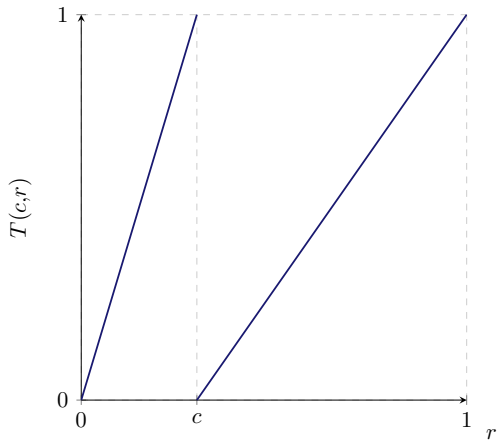


FIGURE 3 – L'application dyadique asymétrique de paramètre c

Pour de telles variables aléatoires, la fonction de répartition est :

- *non négative* : la probabilité qu'un événement se produise est positive ou nulle ;
- *non décroissante* : plus l'intervalle dans lequel X est choisi est grand, plus l'événement est probable ;

De plus, elle satisfait toujours $F(1) = 1$, puisque l'événement $X \leq 1$ est assurément réalisé.

Si F est dérivable, alors la *fonction de densité* de X est définie par $f = F'$. Les régions où f est plus élevée correspondent alors à des valeurs de X plus probables. La fonction de densité permet de donner à la probabilité une interprétation géométrique simple. Lorsqu'une variable aléatoire X prend ses valeurs dans $[0,1]$ et admet une densité f, la probabilité d'un événement ne correspond pas seulement à la taille de l'ensemble considéré, mais à son poids sous la densité. Ainsi, pour un sous-ensemble $A \subseteq [0,1]$, on a

$$\mathbb{P}[X \in A] = \int_A f(x)\,\mathrm{d}x.$$

Autrement dit, cette probabilité est l'aire sous la courbe de f au-dessus de l'ensemble A. En particulier, on a $\int_0^1 f(x)\,\mathrm{d}x = \mathbb{P}[X \in [0,1]] = 1$. Ainsi, chaque aire partielle $\int_A f(x)\,\mathrm{d}x$ représente la proportion de masse contenue dans la région correspondante. Par exemple, si X est une variable aléatoire dont la fonction de densité est f et si $A = [0,1/3] \cup [2/3,1]$, alors

$$\mathbb{P}[X \in A] = \int_0^{1/3} f(x)\,\mathrm{d}x + \int_{2/3}^1 f(x)\,\mathrm{d}x.$$

Remarque 1. *Si plusieurs ensembles sont disjoints, alors la probabilité que X soit dans l'union de ceux-ci est égale à la somme des probabilités individuelles sur chaque ensemble. Dans le cas contraire, il ne faut pas intégrer plusieurs fois sur la même région, sinon la probabilité sera surestimée ; c'est le principe d'inclusion-exclusion (voir la Figure 4).*

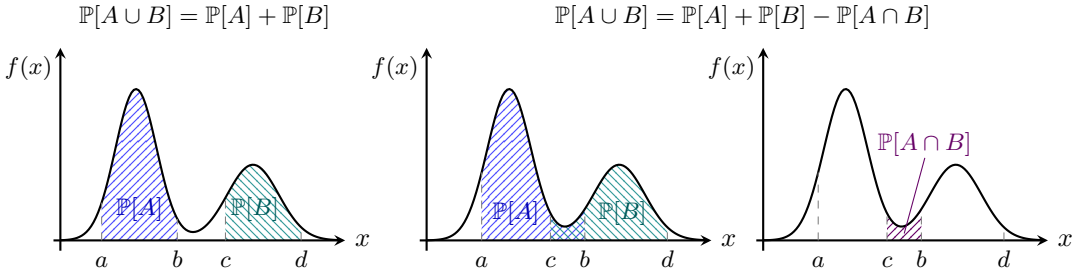


FIGURE 4 – Probabilité de l'événement « A ou B », où $A = \{X \in [a,b]\}$ et $B = \{X \in [c,d]\}$. À gauche, quand les événements sont disjoints. À droite, quand ils ne le sont pas.

L'*espérance* (ou la *moyenne*) est une mesure de la position des valeurs possibles de X ; la *variance* est une mesure de dispersion quantifiant à quel point les valeurs de X s'éloignent de la moyenne. Ces concepts sont respectivement notés et définis par

$$\mu = \mathbb{E}[X] = \int_0^1 x f(x)\,\mathrm{d}x \qquad \text{et} \qquad \sigma^2 = \mathrm{Var}[X] = \int_0^1 (x-\mu)^2 f(x)\,\mathrm{d}x.$$

La *probabilité conditionnelle* de l'événement « A sachant B » est la probabilité que l'événement A se réalise, sachant que B s'est produit. Géométriquement, cette probabilité est égale au rapport entre l'aire de l'intersection des événements A et B et l'aire totale de l'événement B, c'est-à-dire

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}. \tag{2}$$

L'*espérance conditionnelle* est une notion d'espérance qui tient compte de la connaissance d'un événement. Soit un évènement A de probabilité non-nulle, par exemple, $X \in [a,b]$. Alors

$$\mathbb{E}[X \mid A] = \frac{1}{\mathbb{P}[A]} \int_a^b x f(x)\,\mathrm{d}x, \tag{3}$$

c'est-à-dire l'espérance de X restreinte à A, normalisée par la probabilité de A. Si l'événement A se décompose en événements disjoints, à savoir $A = A_1 \cup \cdots \cup A_n$ où $A_i \cap A_j = \emptyset$ pour tout $i \neq j$, alors la *loi de l'espérance totale* affirme que l'espérance conditionnelle est égale à

$$\mathbb{E}[X \mid A] = \frac{1}{\mathbb{P}[A]} \sum_{i=1}^{n} \mathbb{P}[A_i]\,\mathbb{E}[X \mid A_i]. \tag{4}$$

Deux événements A et B sont dits *indépendants* si $\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$. Intuitivement, cela signifie que la réalisation de l'un n'influence pas la probabilité de réalisation de l'autre. En effet, par l'équation (2), cette relation est équivalente à $\mathbb{P}[A \mid B] = \mathbb{P}[A]$, c'est-à-dire que savoir que B s'est réalisé ne fournit aucune information supplémentaire sur la probabilité que A se réalise.

Il est souvent utile d'étudier les relations entre plusieurs variables aléatoires, surtout lorsqu'elles sont dépendantes. Considérons X et Y deux variables à valeurs dans [0,1]. Leur *fonction de répartition jointe* est définie par $F(x,y) = \mathbb{P}[X \leq x, Y \leq y]$. Lorsque X et Y sont dépendantes, la probabilité jointe ne se factorise généralement pas : $\mathbb{P}[X \leq x, Y \leq y] \neq \mathbb{P}[X \leq x]\mathbb{P}[Y \leq y]$. Si une densité jointe existe, cette probabilité s'interprète géométriquement comme l'aire pondérée d'une région de $[0,1]^2$. La dépendance signifie alors que cette aire ne peut généralement pas se décomposer en un produit de deux intégrales.

Afin d'énoncer nos résultats dans toute leur généralité, nous devons permettre que les variables aléatoires considérées n'admettent pas nécessairement de densité. Cette précaution est importante : certaines distributions naturelles dans notre contexte sont continues, comme la loi uniforme, mais d'autres peuvent être discrètes ou même déterministes. Par exemple, la méthode de bissection classique correspond au cas où le point de coupe vaut toujours $1/2$. Cette situation est modélisée par une variable aléatoire prenant toujours une valeur unique et ne possèdant pas de densité au sens usuel, car toute la masse probabiliste est concentrée en un seul point. Malgré tout, elle a pour fonction de répartition

$$F(x) = \begin{cases} 0 & \text{si } 0 \leq x < 1/2, \\ 1 & \text{si } 1/2 \leq x \leq 1. \end{cases}$$

Pour traiter tous ces cas simultanément, nous utilisons l'intégrale de Lebesgue–Stieltjes. Malgré la terminologie pouvant sembler intimidante, l'idée est relativement simple : remplacer l'intégration par rapport à une densité $f(x)\,\mathrm{d}x$ par une intégration directement par rapport à la fonction de répartition F. Ainsi, pour toute fonction bornée g pour laquelle l'espérance est bien définie, on écrit

$$\mathbb{E}[g(X)] = \int_0^1 g(x)\,\mathrm{d}F(x).$$

Cette notation signifie que l'on « additionne » (au sens infinitésimal) toutes les valeurs de $g(x)$ en les pondérant par la masse de probabilité attribuée à chaque région par la loi de X.

Lorsque F est dérivable et admet une densité f, on retrouve simplement la formule familière

$$\mathbb{E}[g(X)] = \int_0^1 g(x)f(x)\,\mathrm{d}x.$$

Dans le cas discret, l'intégrale de Lebesgue–Stieltjes se comporte plutôt comme une somme pondérée : si X prend les valeurs x_i avec probabilités p_i, alors

$$\int_0^1 g(x)\,\mathrm{d}F(x) = \sum_i g(x_i)p_i. \tag{5}$$

Ainsi, les parties continues de F contribuent comme des aires sous une densité, tandis que ses sauts contribuent comme des masses ponctuelles. Par exemple, si X est déterministe et vaut toujours $1/2$, alors toute la masse est concentrée en $1/2$, et l'intégrale vaut simplement

$$\int_0^1 g(x)\,\mathrm{d}F(x) = g(1/2).$$

L'intégrale de Lebesgue–Stieltjes fournit donc une notation unifiée qui couvre tous les types de distributions. Elle nous permet ainsi de formuler nos résultats sans imposer inutilement l'existence d'une densité. Toutefois, un lecteur moins familier avec cette notation peut aisément interpréter toutes les preuves dans le cas continu en remplaçant $\mathrm{d}F(x)$ par $f(x)\,\mathrm{d}x$ et dans le cas discret en utilisant l'identité (5).

3. L'algorithme aléatoire, étape par étape

3.1. La longueur de l'intervalle après une itération

En théorie, étant donné une fonction f, la racine r que l'on cherche à identifier à l'aide de l'algorithme de bissection stochastique est fixée et, donc, déterministe. Toutefois, en pratique, la racine r est inconnue et se comporte, sur plusieurs exécutions de l'algorithme, comme une variable aléatoire de distribution inconnue dans l'intervalle (0,1).

Par simplicité, et pour des raisons élégantes qui deviendront bientôt claires, nous supposons ici que la racine initiale r est uniformément distribuée sur (0,1), c'est-à-dire que $r = r_0 \sim \mathcal{U}(0,1)$. De plus, les points de coupe c_n sont choisis indépendamment des racines normalisées r_n, selon une distribution ayant la fonction de répartition F. En suivant l'algorithme 2, nous pouvons déduire que la longueur de l'intervalle après la première coupe est donnée par

$$\ell_1 = \begin{cases} c_0 & \text{si } c_0 \geq r_0, \\ 1 - c_0 & \text{si } c_0 < r_0. \end{cases} \tag{6}$$

En supposant la valeur de r_0 connue, la loi de l'espérance totale (4) nous assure que

$$\begin{aligned} \mathbb{E}[\ell_1 \mid r_0] &= \mathbb{P}[c_0 \geq r_0]\,\mathbb{E}[\ell_1 \mid c_0 \geq r_0] + \mathbb{P}[c_0 < r_0]\,\mathbb{E}[\ell_1 \mid c_0 < r_0] \\ &= \mathbb{P}[c_0 \geq r_0]\,\mathbb{E}[c_0 \mid c_0 \geq r_0] + \mathbb{P}[c_0 < r_0]\,\mathbb{E}[1 - c_0 \mid c_0 < r_0] \\ &= \int_{r_0}^{1} c\,\mathrm{d}F(c) + \int_{0}^{r_0} (1-c)\,\mathrm{d}F(c), \end{aligned}$$

où la dernière identité provient de l'équation (3). Prendre l'espérance de cette expression par rapport à la distribution de r_0 nous donne le facteur d'échelle moyen après la première coupe.

Exemple 1. *Supposons que $c \sim \mathcal{U}(0,1)$, à savoir que la coupe est choisie uniformément sur (0,1). Alors la fonction de répartition est $F(x) = x$ et on a donc $\mathrm{d}F(x) = \mathrm{d}x$. Par conséquent,*

$$\mathbb{E}[\ell_1 \mid r_0] = \frac{1 + 2r_0 - 2r_0^2}{2}.$$

La longueur moyenne de l'intervalle après la première coupe aléatoire dépend ainsi de la position de la racine (voir la figure 5). En prenant l'espérance par rapport à $r_0 \sim \mathcal{U}(0,1)$, on obtient alors un facteur de contraction moyen $\mathbb{E}[\ell_1] = 2/3$.

Exemple 2. *Supposons que* $c \sim \text{Beta}(2;\,3)$. *Alors*

$$\mathbb{E}[\ell_1 \mid r_0] = \frac{2 + 30r_0^2 - 80r_0^3 + 75r_0^4 - 24r_0^5}{5}$$

La distribution du facteur d'échelle espéré en fonction de r_0 *est représentée à la figure 6. En prenant l'espérance par rapport à* $r_0 \sim \mathcal{U}(0,1)$, *on obtient* $\mathbb{E}[\ell_1] = 3/5$.

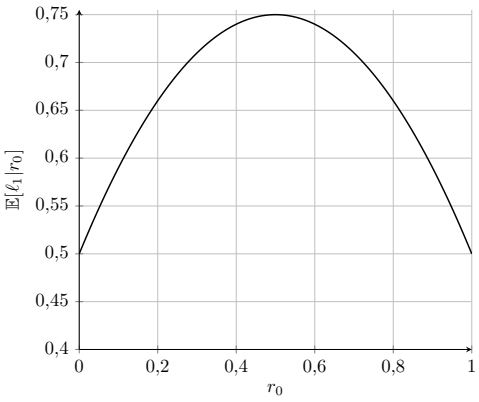


FIGURE 5 – Longueur moyenne après une itération si $c_0 \sim \mathcal{U}(0,1)$.

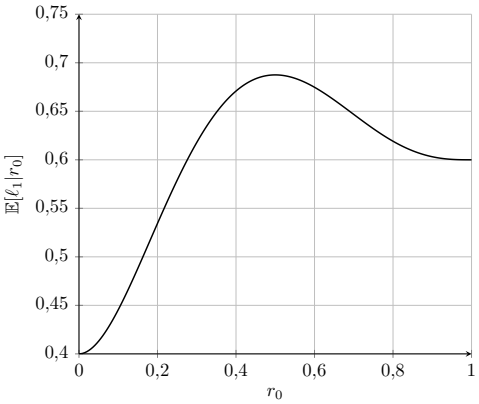


FIGURE 6 – Longueur moyenne après une itération si $c_0 \sim \text{Beta}(2;\,3)$.

Remarquons que, si r_0 est uniformément distribuée sur (0,1), alors après une première étape, l'espérance $\mathbb{E}[\ell_1]$ ne dépend que de la distribution du point de coupe. Par conséquent, si la racine normalisée r_1 demeurait elle aussi uniforme sur (0,1), les mêmes arguments s'appliqueraient aux étapes suivantes. En particulier, le facteur de contraction moyen $\mathbb{E}[\ell_n]$ resterait donc constant tout au long de l'algorithme.

3.2. L'invariance de la distribution uniforme

La proposition suivante montre que l'hypothèse formulée ci-dessus est effectivement valide. Plus précisément, nous montrons dans ce qui suit que si $r_n \sim \mathcal{U}(0,1)$, alors la distribution de r_{n+1} est elle aussi uniforme sur (0,1) après normalisation.

Proposition 1. *Si* $r_0 \sim \mathcal{U}(0,1)$, *alors* $r_n \sim \mathcal{U}(0,1)$ *pour tout* $n \geq 1$. *Autrement dit,* $\mathcal{U}(0,1)$ *est une distribution stationnaire pour le processus aléatoire décrit dans l'algorithme 2.*

Démonstration. Considérons la fonction $T(c,r)$, définie en (1), permettant de calculer r_{n+1} à partir de r_n. Par définition, la racine $r_1 = T(c_0,r_0)$ sera uniformément distribuée si $\mathbb{P}[T(c_0,r_0) \leq t] = t$ pour tout $t \in [0,1]$. Soit $L_t^-(T) = \{(c_0,r_0) : T(c_0,r_0) \leq t\}$ l'ensemble des points (c_0,r_0)

tels que $T(c_0,r_0) \leq t$. Comme mentionné à la section 2.2, le calcul de la probabilité revient à calculer l'aire pondérée d'un sous-ensemble de $[0,1]^2$. Cet ensemble, illustré à la figure 7, admet la description géométrique suivante :

$$L_t^-(T) = \big\{(c_0,r_0) : c_0 \in [0,1], r_0 \in [0,tc_0]\big\} \cup \big\{(c_0,r_0) : c_0 \in [0,1], r_0 \in (c_0, t+(1-t)c_0]\big\}.$$

Par conséquent, puisque $\int_0^1 \mathrm{d}F(c) = 1$, il découle de la définition que

$$\begin{aligned}
\mathbb{P}\big(T(c_0,r_0) < t\big) &= \int_{L_t^-(T)} \mathrm{d}r\,\mathrm{d}F(c) \\
&= \int_0^1 \left(\int_0^{tc} \mathrm{d}r + \int_c^{t+(1-t)c} \mathrm{d}r \right) \mathrm{d}F(c) \\
&= \int_0^1 \big(tc + t + (1-t)c - c\big)\,\mathrm{d}F(c) \\
&= t \int_0^1 \mathrm{d}F(c) \\
&= t.
\end{aligned}$$

Ainsi, $r_1 \sim \mathcal{U}(0,1)$. Comme l'argument ne dépend pas de n, la même méthode peut être appliquée récursivement afin de montrer que $r_n \sim \mathcal{U}(0,1)$ pour tout $n \geq 1$, démontrant donc que $\mathcal{U}(0,1)$ est une distribution stationnaire pour l'algorithme. ■

Notons que la stationnarité de $\mathcal{U}(0,1)$ pour le processus aléatoire décrit dans l'algorithme 2 tient quelle que soit la fonction de répartition F. Ainsi, à chaque étape du processus, la racine renormalisée r_n suit toujours une distribution uniforme sur $[0,1]$, indépendamment de la distribution choisie pour le point de coupe.

Cette remarquable invariance a des conséquences profondes pour notre problème. En particulier, le corollaire suivant découle alors de ce qui a été présenté dans la section 3.1.

Corollaire 1. *Supposons que $r \sim \mathcal{U}(0,1)$. Alors, il existe une constante $C > 0$ telle que les variables aléatoires ℓ_n générées par l'algorithme 2 satisfont $\mathbb{E}[\ell_n] = C$ pour tout $n \geq 1$, où la constante C dépend uniquement de la distribution $\mathcal{D}$ choisie pour les points de coupe.*

Par exemple, dans le cas où $\mathcal{D} = \mathcal{U}(0,1)$, l'**??** 1 implique que $\mathbb{E}[\ell_n] = \frac{2}{3}$ pour tout $n \geq 1$.

3.3. La distribution de ℓ_n

Le corollaire 1 révèle que l'espérance $\mathbb{E}[\ell_n]$ est constante pour tout $n \geq 1$. Nous cherchons maintenant à déterminer une formule générale pour $\mathbb{E}[\ell_n]$, en fonction d'une distribution arbitraire du point de coupe, c'est-à-dire

$$\text{coupe}(0,1) = c \sim \mathcal{D}(0,1).$$

Pour ce faire, nous commençons par établir une formule pour la fonction de répartition et pour la densité, lorsqu'elle existe, des variables aléatoires ℓ_n.

En particulier, nous montrons que les variables aléatoires ℓ_n possèdent une distribution étroitement liée à celle des points de coupe c_n.

Théorème 1. *Supposons que $r \sim \mathcal{U}(0,1)$ et que c soit une variable aléatoire de fonction de répartition F. Définissons ℓ selon la relation (6). Alors la fonction de répartition de ℓ est*

$$H(t) = \int_0^t x \,\mathrm{d}F(x) + \int_{1-t}^1 (1-x)\,\mathrm{d}F(x).$$

De plus, si c admet une densité de probabilité f, alors ℓ a une densité de probabilité donnée par

$$h(t) = t\big(f(t) + f(1-t)\big).$$

Démonstration. Nous calculons d'abord la fonction de répartition de ℓ. Soit $t \in [0,1]$. Par définition, on a $H(t) = \mathbb{P}[\ell \le t]$. Puisque

$$\ell = \begin{cases} c & \text{si } \ c \ge r, \\ 1-c & \text{si } \ c < r, \end{cases}$$

on obtient

$$\{(c,r) : \ell \le t\} = \{(c,r) : r \le c \le t\} \cup \{(c,r) : 1-t \le c < r\}.$$

Ces deux événements sont disjoints, comme l'illustre la figure 8. Par conséquent,

$$H(t) = \mathbb{P}[r \le c \le t] + \mathbb{P}[1-t \le c < r].$$

Comme $r \sim \mathcal{U}(0,1)$, on a $\mathbb{P}[r \le x \mid c = x] = x$. Ainsi,

$$\mathbb{P}[r \le c \le t] = \int_0^t \mathbb{P}[r \le x \mid c = x]\,\mathrm{d}F(x) = \int_0^t x \,\mathrm{d}F(x).$$

De même, en conditionnant à nouveau sur $c = x$, on a $\mathbb{P}[x \le r \mid c = x] = 1 - x$, d'où

$$\mathbb{P}[1-t \le c < r] = \int_{1-t}^1 \mathbb{P}[x < r \mid c = x]\,\mathrm{d}F(x) = \int_{1-t}^1 (1-x)\,\mathrm{d}F(x).$$

En combinant les deux termes, on obtient donc finalement

$$H(t) = \int_0^t x \,\mathrm{d}F(x) + \int_{1-t}^1 (1-x)\,\mathrm{d}F(x), \qquad 0 \le t \le 1.$$

De plus, si c admet une densité f, alors $\mathrm{d}F(x) = f(x)\,\mathrm{d}x$. En dérivant par rapport à t, on trouve finalement que ℓ admet la densité

$$h(t) = H'(t) = \frac{d}{dt}\left(\int_0^t x f(x)\,\mathrm{d}x + \int_{1-t}^1 (1-x) f(x)\,\mathrm{d}x \right) = t f(t) + t f(1-t),$$

concluant ainsi la preuve. ■

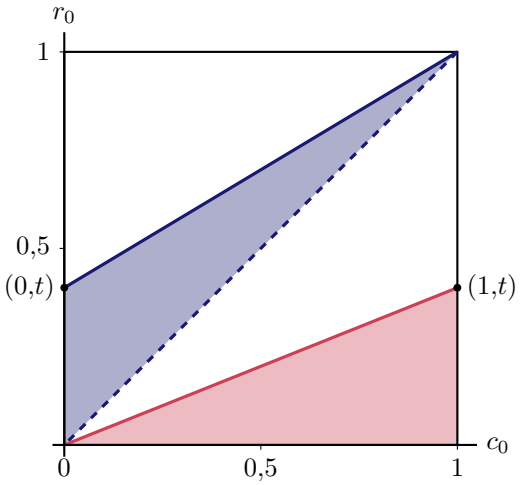


FIGURE 7 – La région $L_t^-(T)$. La section rouge correspond au cas $c_0 \geq r_0$ et la section bleue correspond au cas $c_0 < r_0$.

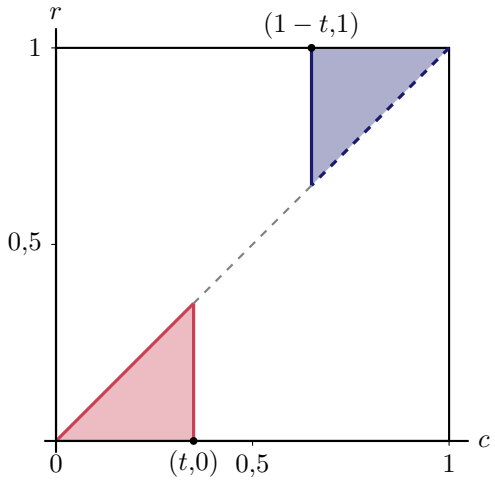


FIGURE 8 – La région $\{(c,r) : \ell \leq t\}$. La section rouge correspond au cas $c \geq r$ et la section bleue correspond au cas $c < r$.

Rappelons que nous cherchons à calculer $\mathbb{E}[\ell_n]$ pour une distribution arbitraire du point de coupe. Le théorème précédent mène à une formule explicite pour cette espérance, ainsi que pour la variance de ℓ_n. Étonnamment, la moyenne et la variance de la variable aléatoire ℓ peuvent toutes deux s'exprimer simplement comme des polynômes de la quantité $\mu - \mu^2 - \sigma^2 = \mathbb{E}[c(1-c)]$.

Corollaire 2. *Supposons que $r \sim \mathcal{U}(0,1)$ et que c soit une variable aléatoire de moyenne μ et de variance σ^2. De plus, notons $\theta := \mu - \mu^2 - \sigma^2$. Alors*

$$\mu_\ell = 1 - 2\theta \qquad \textit{et} \qquad \sigma_\ell^2 = \theta(1 - 4\theta).$$

Démonstration. Nous commençons par établir l'identité pour μ_ℓ. Puisque

$$H(t) = \int_0^t x \, \mathrm{d}F(x) + \int_{1-t}^1 (1-x) \, \mathrm{d}F(x),$$

il découle d'un simple changement de variable et du théorème fondamental du calcul que

$$\mathrm{d}H(t) = t \, \mathrm{d}F(t) + t \, \mathrm{d}F(1-t).$$

Ainsi, par le théorème 1 et un second changement de variable, on obtient

$$\begin{aligned}\mu_\ell &= \int_0^1 x \,\mathrm{d}H(x) \\ &= \int_0^1 x^2 \,\mathrm{d}F(x) + \int_0^1 x^2 \,\mathrm{d}F(1-x) \\ &= \int_0^1 x^2 \,\mathrm{d}F(x) + \int_0^1 (1-x)^2 \,\mathrm{d}F(x) \\ &= \int_0^1 \big(1 - 2x + 2x^2\big) \,\mathrm{d}F(x).\end{aligned}$$

Comme $\int_0^1 x^2 \,\mathrm{d}F(x) = \mu^2 + \sigma^2$, il s'ensuit que

$$\mu_\ell = 1 - 2\mu + 2(\mu^2 + \sigma^2) = 1 - 2(\mu - \mu^2 - \sigma^2) = 1 - 2\theta.$$

Pour la variance, un argument similaire révèle que

$$\begin{aligned}\sigma_\ell^2 + \mu_\ell^2 &= \int_0^1 x^2 \,\mathrm{d}H(x) \\ &= \int_0^1 x^3 \,\mathrm{d}F(x) + \int_0^1 (1-x)^3 \,\mathrm{d}F(x) \\ &= \int_0^1 \big(3x^2 - 3x + 1\big) \,\mathrm{d}F(x) \\ &= 3(\mu^2 + \sigma^2) - 3\mu + 1 \\ &= 1 - 3\theta.\end{aligned}$$

Par conséquent, on trouve alors finalement

$$\sigma_\ell^2 = 1 - 3\theta - \mu_\ell^2 = 1 - 3\theta - (1-2\theta)^2 = \theta(1-4\theta),$$

concluant ainsi la preuve. ■

Remarquons que, dans le cas particulièrement naturel où les points de coupe sont tirés d'une distribution symétrique, alors $\mu = \frac{1}{2}$ et on obtient l'élégant résultat

$$\mu_{\ell_n} = \frac{1}{2} + 2\sigma^2 \qquad \text{et} \qquad \sigma_{\ell_n}^2 = \sigma^2(1 - 4\sigma^2), \qquad n \geq 1.$$

Ainsi, pour ces distributions, la longueur de l'intervalle est réduite à chaque étape de l'algorithme par un facteur moyen de $\frac{1}{2} + 2\sigma^2$. Ce facteur est minimal lorsque $\sigma^2 \to 0$, ce qui correspond précisément au cas de la méthode de bissection classique. La figure 9 illustre graphiquement les relations obtenues pour μ_ℓ et σ_ℓ^2 dans ce cas.

3.4. La longueur moyenne L_n

La variable aléatoire L_n générée par l'algorithme 2 correspond à la longueur de l'intervalle après n itérations. Cette longueur est l'objet principal d'intérêt pour cet article, puisqu'elle fournit une borne supérieure sur l'erreur de l'approximation donnée par la méthode de bissection stochastique. Dans cet article, le taux de convergence moyen désigne donc le comportement de l'espérance de la longueur de l'intervalle, et non le taux presque sûr souvent associé à $\mathbb{E}[\log \ell]$.

Naturellement, puisqu'à chaque étape de l'algorithme, l'intervalle est réduit en moyenne d'un facteur $\mu_{\ell_n} = 1 - 2\theta$, il est tentant de croire que la longueur moyenne de l'intervalle après n itérations est de

$$\mathbb{E}[L_n] = \mathbb{E}[(b-a)\cdot \ell_1 \cdot \ell_2 \cdots \ell_n] = (b-a)\,\mathbb{E}[\ell_1]\,\mathbb{E}[\ell_2]\,\cdots\,\mathbb{E}[\ell_n] = (b-a)\big(1-2\theta\big)^n.$$

Or, la seconde égalité ci-dessus n'est pas automatique. Toutefois, il suffit que les facteurs de contraction ℓ_n soient mutuellement indépendants pour que celle-ci tienne toujours. Bien que cette propriété peut sembler naturelle, sa validité n'est pas triviale et demande à être démontrée.

Théorème 2. *Supposons que $r \sim \mathcal{U}(0,1)$. Alors*

(i) r_n est indépendante de ℓ_n pour tout $n > 0$;

(ii) les variables aléatoires $\ell_1, \ell_2, \ldots,$ sont mutuellement indépendantes.

Démonstration. Nous cherchons à calculer la probabilité $\mathbb{P}[r_1 \le t, \ell_1 \le u]$ et à montrer que la loi jointe de r_1 et ℓ_1 se factorise sous la forme

$$\mathbb{P}[r_1 \le t, \ell_1 \le u] = \mathbb{P}[r_1 \le t]\,\mathbb{P}[\ell_1 \le u],$$

démontrant ainsi que r_1 et ℓ_1 sont indépendantes. Puisque les événements $\{r_0 \le c_0\}$ et $\{r_0 > c_0\}$ sont disjoints, il est possible de décomposer cette probabilité en deux termes (voir remarque 1) :

$$\begin{aligned}\mathbb{P}[r_1 \le t, \ell_1 \le u] &= \mathbb{P}[r_1 \le t, \ell_1 \le u, r_0 \le c_0] + \mathbb{P}[r_1 \le t, \ell_1 \le u, r_0 > c_0] \\ &= \mathbb{P}[c_0 \le u, r_0 \le tc_0] + \mathbb{P}[1-u \le c_0, r_0 \le t(1-c_0)+c_0, r_0 > c_0] \\ &= \mathbb{P}[c_0 \le u, r_0 \le tc_0] + \mathbb{P}[1-u \le c_0 < r_0 \le t(1-c_0)+c_0].\end{aligned}$$

Le premier terme correspond à la région triangulaire $\big\{c_0 \in [0,u], r_0 \in [0,tc_0]\big\}$ illustrée en rouge à la figure 10. Puisque c_0 et r_0 sont indépendantes, et puisque $r_0 \sim \mathcal{U}(0,1)$, on trouve alors

$$\mathbb{P}[c_0 \le u, r_0 \le tc_0] = \int_0^u \mathbb{P}[r_0 \le tc_0 \mid c_0 = c]\,\mathrm{d}F(c) = t\int_0^u c\,\mathrm{d}F(c).$$

Le second terme correspond également à une région triangulaire dans le carré unité, caractérisée par $\big\{c_0 \in [1-u,1], r_0 \in (c_0, t(1-c_0)+c_0]\big\}$. Ainsi, puisque r_0 est uniforme sur $(0,1)$, on a

$$\mathbb{P}[c_0 < r_0 \le t(1-c_0)+c_0 \mid c_0] = t(1-c_0)+c_0-c_0 = t(1-c_0).$$

Il s'ensuit alors que

$$\begin{aligned}\mathbb{P}[1-u \le c_0 < r_0 \le t(1-c_0)+c_0] &= \int_{1-u}^{1} \mathbb{P}[c_0 < r_0 \le t(1-c_0)+c_0 \mid c_0 = c]\,\mathrm{d}F(c) \\ &= t\int_{1-u}^{1} (1-c)\,\mathrm{d}F(c).\end{aligned}$$

Par conséquent,

$$\begin{aligned}\mathbb{P}[r_1 \le t, \ell_1 \le u] &= t\left(\int_0^u c\,\mathrm{d}F(c) + \int_{1-u}^{1} (1-c)\,\mathrm{d}F(c)\right) \\ &= tH(u) \\ &= \mathbb{P}[r_1 < t]\,\mathbb{P}[\ell_1 < u],\end{aligned}$$

où H est la fonction de répartition de ℓ obtenue au théorème 1. Ainsi, r_1 et ℓ_1 sont indépendantes. De plus, pour tout $n \ge 1$, nous avons les propriétés suivantes :

1. $r_n \sim \mathcal{U}(0,1)$, par la proposition 1 ;
2. c_n a pour fonction de répartition F ;
3. c_n est indépendante de r_n.

Le calcul précédent peut donc être répété, *mutatis mutandis*, avec (r_{n+1},ℓ_{n+1}) à la place de (r_1,ℓ_1). On obtient ainsi que r_{n+1} et ℓ_{n+1} sont indépendantes pour tout $n > 0$.

Il reste à établir l'indépendance mutuelle des variables ℓ_n. Pour ce faire, on procède par induction forte, en montrant simultanément que $\ell_1,\ldots,\ell_n$ sont indépendantes et que r_n est indépendante du vecteur $(\ell_1,\ldots,\ell_n)$.

Le cas initial découle du calcul précédent. Supposons maintenant le résultat vrai jusqu'au rang n. Puisque c_n est choisie indépendamment de tout ce qui a été généré auparavant, et puisque r_n est indépendante de $(\ell_1,\ldots,\ell_n)$ par hypothèse d'induction, le couple (r_n,c_n) est indépendant de $(\ell_1,\ldots,\ell_n)$. Or,

$$\ell_{n+1} = \begin{cases} c_n & \text{si } c_n \ge r_n, \\ 1-c_n & \text{si } c_n < r_n, \end{cases}$$

dépend uniquement de r_n et de c_n. Il s'ensuit que ℓ_{n+1} est indépendante de $\ell_1,\ldots,\ell_n$.

De plus, le même argument montre que r_{n+1} est indépendante de ℓ_{n+1}, tandis que le couple (r_{n+1},ℓ_{n+1}), étant une fonction de (r_n,c_n), est indépendant de $(\ell_1,\ldots,\ell_n)$. Ainsi, r_{n+1} est indépendante de $(\ell_1,\ldots,\ell_n,\ell_{n+1})$. L'hypothèse d'induction est donc propagée au rang $n+1$, démontrant que les variables ℓ_n sont mutuellement indépendantes. ■

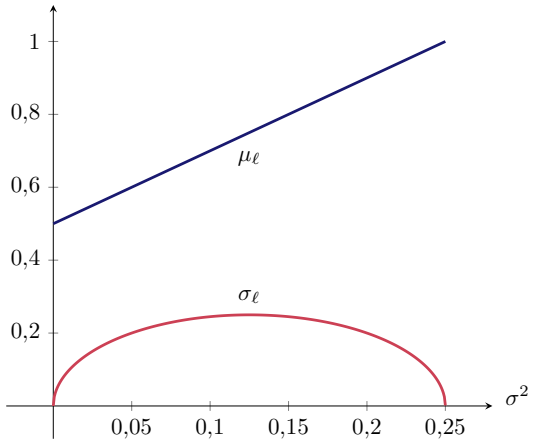


FIGURE 9 – La moyenne μ_ℓ et l'écart-type σ_ℓ de la variable aléatoire $\ell = \ell(c,r)$, exprimés en fonction de la variance σ^2 de la variable aléatoire *symétrique* c

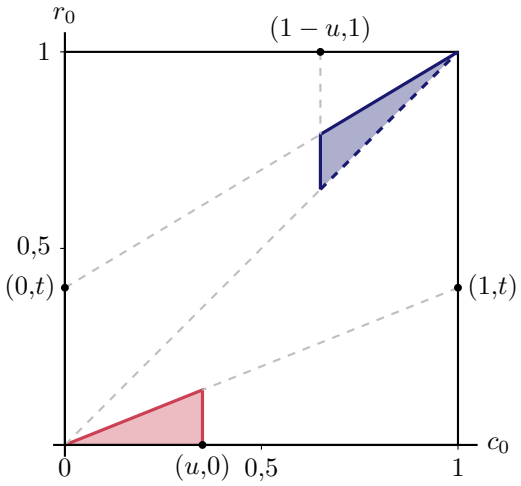


FIGURE 10 – Les régions dans la preuve du théorème 2. Notons qu'elles correspondent aux intersections des figures 7 et 8

Corollaire 3. *L'espérance de la longueur du n-ième intervalle est donnée par*

$$\mathbb{E}[L_n] = \mathbb{E}[\ell_1]^n = (1-2\theta)^n,$$

où θ est défini dans le corollaire 2.

Démonstration. Puisque les variables ℓ_i, $i > 0$, sont mutuellement indépendantes, il s'ensuit que

$$\mathbb{E}[L_n] = \mathbb{E}\left[\prod_{i=1}^{n} \ell_i\right] = \prod_{i=1}^{n} \mathbb{E}[\ell_i] = \mathbb{E}[\ell_1]^n,$$

par le corollaire 1. Le résultat s'ensuit alors du corollaire 2. ■

À la suite du corollaire 2, nous avions observé que la méthode de bissection classique minimise le facteur de contraction moyen parmi les méthodes de bissection stochastiques *symétriques*. Le corollaire 3 permet de traduire cette observation en un énoncé d'optimalité pour la longueur moyenne de l'intervalle après n itérations. En effet, lorsque $r \sim \mathcal{U}(0,1)$ et que la longueur de l'intervalle initial est normalisée à 1, on a $\mathbb{E}[L_n] = \mathbb{E}[\ell_1]^n$. Ainsi, le comportement moyen de l'algorithme est gouverné entièrement par la constante de contraction $C = \mathbb{E}[\ell_1]$. En particulier, l'algorithme 2 est en moyenne linéaire, c'est-à-dire que l'erreur e_n à la n-ième étape satisfait

$$e_{n+1} \approx C^n e_0,$$

Or, une simple analyse de la fonction $f(c) = c(1-c)$ montre que

$$0 \leq c(1-c) \leq \frac{1}{4}$$

pour tout $c \in [0,1]$. Puisque $\mathbb{E}[\ell_1] = 1 - 2\,\mathbb{E}[c(1-c)]$, il s'ensuit alors que

$$\frac{1}{2} \leq \mathbb{E}[\ell_1] \leq 1.$$

L'égalité à droite se produit si et seulement si $c \in \{0,1\}$ presque sûrement, c'est-à-dire lorsque la coupe se fait presque toujours aux extrémités de l'intervalle. À l'inverse, l'égalité à gauche se produit si et seulement si $c = \frac{1}{2}$ presque sûrement, ce qui correspond exactement à la méthode de bissection classique. Ainsi, la constante de contraction moyenne est minimisée *uniquement*, à un ensemble de mesure nulle près, par la bissection classique.

Corollaire 4. *La méthode de bissection classique est optimale en moyenne (au sens de $\mathbb{E}[L_n]$) dans la famille des algorithmes de bissection stochastique.*

4. Simulations numériques

Nous réalisons des expériences numériques afin de valider et d'illustrer les résultats théoriques des sections précédentes. Les valeurs rapportées ci-dessous sont obtenues à partir de la méthode de bissection stochastique originale (algorithme 1), plutôt qu'à partir de la version invariante par changement d'échelle, afin de confirmer que nos résultats demeurent valides dans ce cadre. Le code utilisé pour les simulations de cette section est disponible dans le répertoire GitHub [7].

4.1. Le taux de convergence moyen

Pour mesurer empiriquement le facteur de contraction moyen, nous menons l'expérience suivante :

1. On génère $r \sim \mathcal{U}(0,1)$;
2. On applique l'algorithme 1 avec des points de coupe $c_n \sim \mathcal{D}(0,1)$, pour $f(x) = x - r$, avec une petite tolérance et un grand nombre d'itérations, par exemple 10^{-15} et $N = 30$;
3. On calcule les facteurs $\frac{b_{i+1}-a_{i+1}}{b_i-a_i}$ à chaque itération et on les conserve en mémoire ;
4. On répète les étapes 1 à 3 un grand nombre de fois, par exemple 500 fois ;
5. On calcule un intervalle de confiance à 95 % pour la moyenne à l'aide du bootstrap.

Le principal avantage des intervalles de confiance bootstrap est qu'ils sont non paramétriques : ils peuvent donc être utilisés sans connaissance précise de la distribution sous-jacente, ou seulement sous des hypothèses faibles [3].

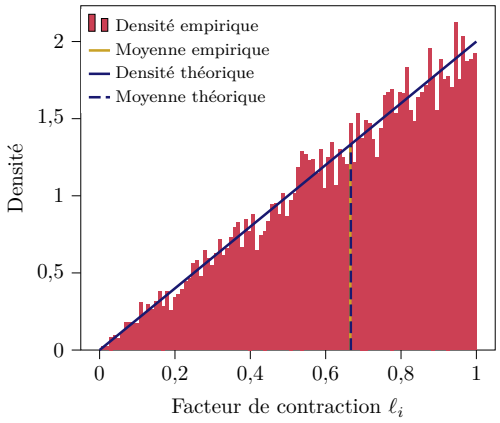


FIGURE 11 – Densités empirique et théorique des ℓ_i pour $\mathcal{U}(0,1)$.

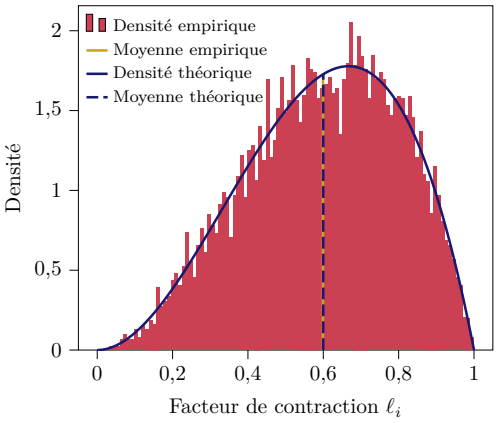


FIGURE 12 – Densités empirique et théorique des ℓ_i pour Beta(2; 2).

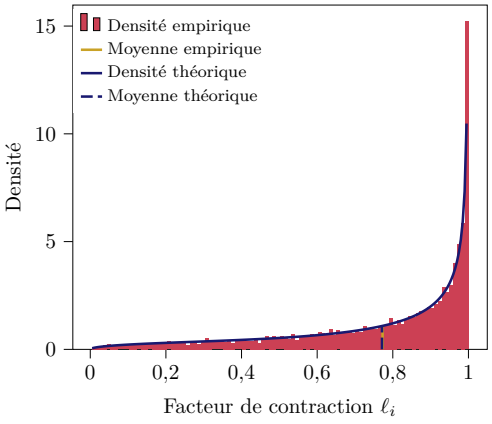


FIGURE 13 – Densités empirique et théorique des ℓ_i pour Beta(0,5; 2).

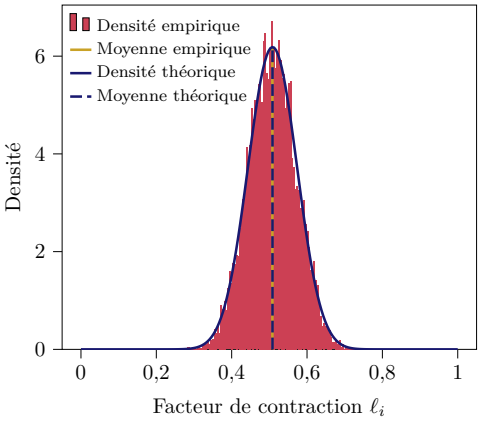


FIGURE 14 – Densités empirique et théorique des ℓ_i pour Bates(20).

Les résultats sont présentés dans le tableau 1, et les distributions empiriques sont illustrées aux figures 11 à 14, en comparaison avec les densités théoriques $h(t) = t(f(t) + f(1-t))$. Pour toutes les distributions testées, la valeur théorique se trouve à l'intérieur des intervalles de

$\mathcal{D}$	$\mathcal{U}(0{,}1)$	Beta(2; 2)	Beta(0,5; 2)	Bates(20)
$\mathbb{E}[\ell_i]$	$[0{,}654; 0{,}677]$	$[0{,}595; 0{,}602]$	$[0{,}767; 0{,}775]$	$[0{,}507; 0{,}509]$
$(\mathbb{E}[L_N])^{\frac{1}{N}}$	$[0{,}655; 0{,}676]$	$[0{,}599; 0{,}612]$	$[0{,}764; 0{,}781]$	$[0{,}508; 0{,}510]$
Théorique	$\frac{2}{3} = 0.\overline{6}$	$\frac{3}{5} = 0{,}6$	$\frac{27}{35} = 0{,}77\ldots$	$\frac{61}{120} = 0{,}508\overline{3}$

TABLEAU 1 – Le facteur de contraction moyen pour différentes distributions

r		$\mathcal{U}(0{,}1)$	Beta(2; 2)	Beta(0,5; 2)	Beta(2; 0,5)	Bates(20)
0.1	Moyenne	$[37{,}01; 37{,}75]$	$[31{,}76; 32{,}25]$	$[49{,}42; 50{,}64]$	$[53{,}99; 55{,}35]$	$[27{,}37; 27{,}49]$
	Étendue	$[19{,}56]$	$[17{,}47]$	$[17{,}89]$	$[19{,}88]$	$[25{,}31]$
	$\mathbb{P}[N \leq 27]$	$[0{,}03; 0{,}06]$	$[0{,}12; 0{,}16]$	$[0{,}00; 0{,}01]$	$[0{,}00; 0{,}01]$	$[0{,}53; 0{,}59]$
0.3	Moyenne	$[37{,}97; 38{,}71]$	$[32{,}39; 32{,}91]$	$[51{,}65; 53{,}00]$	$[53{,}68; 54{,}37]$	$[27{,}35; 27{,}49]$
	Étendue	$[22{,}56]$	$[21{,}46]$	$[17{,}95]$	$[23{,}93]$	$[24{,}31]$
	$\mathbb{P}[N \leq 27]$	$[0{,}02; 0{,}04]$	$[0{,}08; 0{,}12]$	$[0{,}00; 0{,}02]$	$[0{,}00; 0{,}01]$	$[0{,}52; 0{,}58]$
0.45	Moyenne	$[37{,}93; 38{,}65]$	$[32{,}19; 32{,}69]$	$[52{,}55; 53{,}84]$	$[53{,}89; 54{,}00]$	$[27{,}41; 27{,}54]$
	Étendue	$[20{,}62]$	$[20{,}47]$	$[21{,}95]$	$[24{,}89]$	$[25{,}31]$
	$\mathbb{P}[N \leq 27]$	$[0{,}03; 0{,}05]$	$[0{,}08; 0{,}11]$	$[0{,}00; 0{,}01]$	$[0{,}00; 0{,}01]$	$[0{,}49; 0{,}55]$

TABLEAU 2 – Statistiques du nombre d'itérations jusqu'à convergence pour différentes racines

confiance, lesquels sont très étroits. Cela fournit un appui statistique aux facteurs d'échelle théoriques obtenus précédemment. De plus, une distribution fortement concentrée autour de 1/2 et de faible variance, comme la distribution Bates(20), présente un taux de convergence très proche de celui de la méthode de bissection classique. La longueur du dernier intervalle est également mesurée pour chaque exécution afin de vérifier qu'elle satisfait elle aussi la conclusion du corollaire 3.

4.2. Performance pour des racines fixées

La seconde expérience de cette section considère des valeurs fixes de la racine. Nous examinons une racine près du bord ($r = 0{,}1$), près du centre ($r = 0{,}45$) et dans une position intermédiaire ($r = 0{,}3$). Pour atteindre une tolérance de 10^{-8}, la méthode de bissection classique nécessite 27 itérations, indépendamment de la racine. Pour chaque valeur de r, nous effectuons 1000 exécutions de l'algorithme stochastique et calculons un intervalle de confiance bootstrap à 95 % pour le nombre moyen d'itérations nécessaires afin d'atteindre la tolérance prescrite (tableau 2). Nous calculons également l'étendue des résultats, c'est-à-dire le nombre minimal et maximal d'itérations nécessaires, ainsi que le nombre d'exécutions faisant au moins aussi bien que la bissection déterministe. Nous appelons ces exécutions des *essais chanceux*. La quantité $\mathbb{P}[N \leq 27]$ est alors estimée par une proportion, avec des intervalles de confiance de Wilson [4].

Les distributions de N pour $r = 0{,}1$ sont illustrées à la figure 15. Sans surprise, Beta(2; 0,5) donne les moins bons résultats, étant initialement biaisée à l'opposé du sous-intervalle contenant la racine. À l'inverse, Beta(0,5; 2) devient problématique lorsque la racine se retrouve près du centre de l'intervalle de travail, les coupes demeurant biaisées vers la gauche et menant à des choix sous-optimaux. Des *essais chanceux* surviennent néanmoins, mais avec une probabilité très faible, inférieure à 1 %. La loi uniforme offre aussi des performances modestes, nécessitant en moyenne environ 10 itérations supplémentaires dans tous les cas, avec une faible probabilité d'*essais chanceux*. La loi Beta(2; 2) requiert environ 5 itérations de plus en moyenne, mais présente une probabilité nettement plus élevée d'*essais chanceux*, pouvant réduire le nombre d'itérations jusqu'à 10. Enfin, Bates(20) produit des *essais chanceux* dans environ 50 % à 60 % des cas ; toutefois, sa faible variance limite ces gains à 2 ou 3 itérations au maximum, et elle nécessite en moyenne tout au plus une itération de plus que la bissection déterministe.

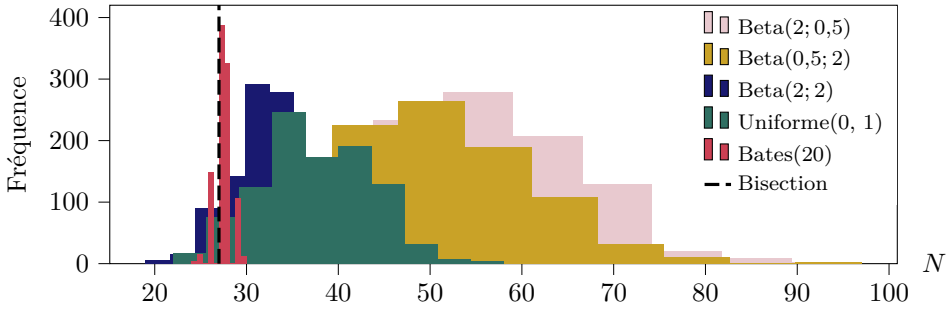


FIGURE 15 – Distribution empirique du nombre N pour $r = 0{,}1$, selon la loi de la coupe c

5. Conclusion

Dans cet article, nous avons étudié les propriétés de convergence de la méthode de bissection stochastique. Lorsque la position de la racine est supposée uniformément distribuée sur l'intervalle de recherche, nous avons obtenu des taux de convergence moyens explicites, ne dépendant que de la moyenne et de la variance de la distribution du point de coupe. En particulier, nous avons montré que la longueur relative du nouvel intervalle est indépendante de toute information précédente, ce qui nous a permis de calculer l'espérance de la longueur de l'intervalle à l'itération n. En moyenne, la convergence est linéaire, et la longueur moyenne de l'intervalle se comporte comme $C^n L_0$, où $C = 1 - 2\,\mathbb{E}[c(1-c)]$ est le facteur de contraction moyen. En particulier, nous avons pu répondre à la question posée dans le titre de cet article par une réponse simple : *Non, pas en moyenne* !

Les prédictions théoriques ont ensuite été validées numériquement à l'aide de simulations statistiques qui ont fourni des intervalles de confiance appuyant les taux de convergence prédits.

Plusieurs questions demeurent ouvertes.

(1) Nous avons supposé dans cet article que la racine se trouvant dans l'intervalle [0,1] se distribuait en moyenne uniformément sur l'intervalle. Or, la « vraie » distribution est inconnue. Que pouvons-nous dire dans le cas où la racine r suit une distribution arbitraire sur [0,1] ?
(2) Lorsque la distribution du point de coupe est concentrée autour du milieu, les résultats numériques suggèrent que l'algorithme stochastique a une probabilité supérieure à 50 % de converger au moins aussi rapidement que la méthode de bissection classique. Pouvons-nous identifier des distributions de coupe optimales, qui minimisent avec grande probabilité le nombre d'itérations nécessaires pour atteindre une tolérance prescrite ?
(3) Notre analyse a été menée dans le cas particulier où la fonction f, dont on cherche à estimer une racine, possède une racine unique dans l'intervalle $[a,b]$. Il serait alors naturel d'étudier comment la possibilité de racines multiples dans $[a,b]$ pourrait affecter l'analyse.
(4) L'algorithme de K-section, analogue à la méthode de bissection, à la différence près que l'on coupe l'intervalle en K sections à la place de deux sections, semble se prêter naturellement à l'analyse menée dans cet article. Est-il possible d'appliquer les mêmes méthodes pour obtenir un analogue avec $K-1$ coupes ?

Les observations faites dans cet article illustrent comment de simples perturbations aléatoires de procédures numériques classiques peuvent engendrer des mathématiques riches en résultats intéressants, ouvrant même potentiellement de nouvelles pistes pour l'analyse des algorithmes stochastiques et de leurs lois de convergence.

Bibliographie

[1] Bernard Bolzano : Rein analytischer Beweis des Lehrsatzes, daß zwischen je zwey Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege. 1817. Traduction anglaise par S. Russ dans *Historia Mathematica*, 7(2) :156–185, 1980.

[2] Arnold Buffum Chace et Henry Parker Manning : *The Rhind Mathematical Papyrus, British Museum 10057 and 10058. Volume I : Free Translation and Commentary.* Mathematical Association of America, Oberlin, 1927.

[3] Bradley Efron et R. J. Tibshirani : *An Introduction to the Bootstrap.* Chapman and Hall/CRC, New York, mai 1994.

[4] Carolina Franco, Roderick J.A. Little, Thomas A. Louis et Eric V. Slud : Comparative study of confidence intervals for proportions in complex sample surveys. *Journal of survey statistics and methodology*, 7(3):334–364, 2019.

[5] Jérôme Gavin et Alain Schärlig : *Longtemps avant l'algèbre, la fausse position, ou comment on a posé le faux pour connaître le vrai, des pharaons aux temps modernes.* EPFL Press, 2012.

[6] Orestis Georgiou, Carl P. Dettmann et Eduardo G. Altmann : Faster than expected escape for a class of fully chaotic maps. *Chaos*, 22(4):043115, 10, 2012.

[7] Philippe-André Luneau, Ludovick Bouthat et Philippe Petitclerc : Randombisection, 2026. DOI : 10.5281/zenodo.20778406.

[8] Gay Robins et Charles Shute : *The Rhind Mathematical Papyrus : An Ancient Egyptian Text*. British Museum Publications, London, 1987. Published for the Trustees of the British Museum.

[9] Arnaldo Rodriguez-Gonzalez : Analyzing topological mixing and chaos on continua with symbolic dynamics, 2023. Preprint (arXiv :2304.02820v2).

[10] Ummu Atiqah Mohd Roslan : Sensitive dependence on initial conditions for an example of Markov map : skewed doubling map. *Matematika (Johor)*, 34(1):13–21, 2018.

Association
mathématique
du Québec

La filiation méconnue entre logique mathématique et linguistique chomskyenne

Patrick Duffley
Département de langues, linguistique et traduction
Université Laval

Frédéric Morneau-Guérin
département Éducation
Université TÉLUQ

Résumé

Noam Chomsky est surtout connu du grand public comme essayiste et commentateur politique. Son œuvre linguistique, elle, s'inscrit dans une histoire intellectuelle beaucoup moins souvent racontée : celle de la formalisation de la logique aux XIX^{e} et XX^{e} siècles. Le présent article retrace cette filiation. De Boole, qui fait de la logique un calcul symbolique, à Frege, Russell et Hilbert, qui déplacent l'attention du contenu des énoncés vers leurs structures formelles, puis à Carnap, qui soutient qu'une langue peut être étudiée du seul point de vue de sa syntaxe, se met progressivement en place l'idée qu'un système symbolique s'analyse indépendamment de sa signification. La grammaire générative, avec ses règles de formation, ses transformations et ses procédures récursives, en hérite directement. Nous examinons enfin les tensions philosophiques que soulève cette autonomie accordée à la syntaxe lorsqu'elle est appliquée aux langues naturelles, dont l'usage effectif déborde constamment les cadres purement formels.

Mots clés : logique mathématique ; philosophie du langage ; syntaxe et sémantique ; grammaire générative ; Noam Chomsky ; Rudolf Carnap ; formalisation ; récursivité ; hiérarchie de Chomsky ; histoire des idées.

Classification mathématique par matières (MSC2020) : 03A05, 03B65, 01A60, 68Q42, 00A30

1. Introduction

Peu d'intellectuels du XX^{e} siècle auront occupé un espace médiatique comparable à celui qu'à occupé Noam Chomsky. Linguiste de formation, mais également essayiste et commentateur politique omniprésent dans l'espace public pendant plusieurs décennies, Chomsky est devenu, bien au-delà du milieu universitaire, une véritable figure intellectuelle mondiale. Son nom est associé depuis longtemps aux débats sur les médias et la liberté d'expression ainsi qu'aux grandes

© Association mathématique du Québec

controverses politiques et géopolitiques de son époque, de la guerre du Vietnam aux discussions contemporaines sur l'impérialisme américain ou le conflit israélo-palestinien. Derrière cette notoriété publique se trouve toutefois une œuvre linguistique immense, composée de centaines d'articles et de nombreux ouvrages, oeuvre qui a exercé une influence considérable sur les sciences du langage et les sciences cognitives [1]. Par son rôle de catalyseur au sein d'un vaste réseau de collaborations scientifiques, Chomsky rappelle, toutes proportions gardées, la place singulière qu'occupait Paul Erdős dans le monde des mathématiques. Dans les deux cas, il s'agit moins de comparer l'étendue ou la nature de leurs contributions que de souligner la fonction d'une figure intellectuelle reliant de nombreuses équipes et traditions de recherche, proposant des intuitions théoriques ensuite développées, raffinées et systématisées par un vaste ensemble de collaborateurs et de continuateurs.

Né à Philadelphie en 1928 dans une famille juive d'origine ukrainienne, Chomsky est formé très tôt dans un milieu intellectuel fortement marqué par les débats politiques, le sionisme culturel et l'étude des langues. Son père, William Chomsky, spécialiste de grammaire hébraïque et professeur d'hébreu, joue notamment un rôle important dans cette formation intellectuelle précoce. Après des études à l'Université de Pennsylvanie sous la direction du linguiste Zellig Harris – figure importante de la linguistique américaine de l'époque, dont les travaux accordaient une grande importance à l'analyse formelle des structures du langage –, Chomsky entreprend dans les années 1950 une remise en question radicale de la linguistique dominante de son époque.

À l'époque où Chomsky commence ses travaux, une grande partie des sciences humaines américaines est fortement influencée par le *behaviorisme*, courant selon lequel les comportements humains – y compris le langage – doivent être expliqués principalement à partir de mécanismes d'apprentissage, d'imitation et d'habitude [2]. Dans cette perspective, parler une langue reviendrait essentiellement à reproduire progressivement des structures entendues dans son environnement. Chomsky s'oppose fortement à cette manière de voir. Selon lui, le langage humain possède une créativité beaucoup trop grande pour être expliqué uniquement par l'imitation. Un enfant peut en effet comprendre et produire des phrases entièrement nouvelles qu'il n'a pourtant jamais entendues auparavant. Pour rendre compte de cette capacité, Chomsky propose que l'esprit humain possède des structures mentales innées permettant de construire un nombre potentiellement infini de phrases à partir d'un ensemble limité de règles grammaticales. Cette approche, appelée « grammaire générative », repose sur une idée qui jouera un rôle central dans toute la suite de notre propos : une langue peut être décrite comme un système de règles capable de générer des structures linguistiques, un peu à la manière dont certains systèmes formels en logique ou en mathématiques permettent de produire des expressions à partir d'un petit nombre de principes de base [3].

1. Pour des présentations générales de l'œuvre de Chomsky et de son influence intellectuelle, voir notamment Barsky, *Noam Chomsky : A Life of Dissent* (1997), ainsi que McGilvray (dir.), *The Cambridge Companion to Chomsky* (2005).

2. Sur le contexte behavioriste et le rôle de la célèbre recension par Chomsky de l'ouvrage *Verbal Behavior* de B. F. Skinner dans l'émergence de la révolution cognitive, voir notamment Harris, *The Linguistics Wars* (1993), ainsi que Chomsky (1959).

3. Voir principalement Chomsky, *Syntactic Structures* (1957), ouvrage fondateur de la grammaire générative moderne.

Les travaux de Chomsky allaient exercer une influence considérable non seulement sur la linguistique théorique, mais également sur la psychologie cognitive, la philosophie du langage et, plus largement, sur plusieurs disciplines cherchant à comprendre les rapports entre langage, pensée et structures formelles [4].

Cette histoire constitue également un épisode révélateur de l'influence culturelle et intellectuelle qu'ont exercée les mathématiques modernes bien au-delà de leur domaine propre.

Le présent article ne vise toutefois ni à présenter en détail l'ensemble de la théorie linguistique chomskyenne, ni à trancher les nombreuses controverses qu'elle a suscitées. Son objectif est plutôt historique et explicatif. Il s'agira de montrer comment une certaine manière de concevoir le langage – non plus d'abord comme un instrument de communication chargé de sens, mais comme un système de formes réglées – s'inscrit dans une histoire intellectuelle plus ancienne, issue de la logique mathématique moderne [5].

Pour suivre cette histoire, il faudra d'abord préciser quelques notions. La philosophie du langage s'interroge sur les rapports entre les mots, les phrases, la pensée et le monde : comment une phrase signifie-t-elle quelque chose ? Comment les mots renvoient-ils à des objets, des idées ou des situations ? Comment distinguons-nous la structure grammaticale d'une phrase de ce qu'elle veut dire ? Ces questions conduisent naturellement à la distinction entre syntaxe et sémantique. La syntaxe concerne l'agencement formel des signes – l'ordre des mots, les règles grammaticales, les structures possibles d'une phrase – tandis que la sémantique concerne leur signification [6].

C'est cette distinction entre structure grammaticale et signification qui permettra de comprendre la filiation intellectuelle que nous souhaitons examiner. De Boole à Frege, de Russell à Hilbert, puis de Carnap à Chomsky, on voit progressivement apparaître l'idée qu'un système symbolique peut être étudié à partir de ses règles internes de formation et de transformation, indépendamment (au moins en partie) de son contenu sémantique immédiat. Le langage humain commence alors à être envisagé comme un objet susceptible d'une analyse formelle. C'est cette évolution intellectuelle, ainsi que les questions philosophiques qu'elle soulève, que nous replacerons dans son contexte historique.

2. Langage, pensée et philosophie du langage

Le langage accompagne la plupart des dimensions de l'existence humaine. Nous l'utilisons pour communiquer nos pensées, transmettre des connaissances, exprimer des émotions, décrire le monde, élaborer des lois, construire des théories scientifiques, ou encore rédiger des démonstrations mathématiques. Cette omniprésence explique que la réflexion sur le langage soit presque aussi ancienne que la philosophie elle-même.

4. Sur le rôle de Chomsky dans la révolution cognitive du milieu du XX^{e} siècle, voir notamment Gardner, *The Mind's New Science* (1985).

5. Parmi les travaux historiques importants sur les rapports entre logique mathématique, formalisation et linguistique générative, voir notamment Tomalin, *Linguistics and the Formal Sciences* (2006).

6. Cette distinction jouera un rôle particulièrement important dans certaines approches issues de la logique formelle du XX^{e} siècle, notamment chez Rudolf Carnap.

Déjà dans l'Antiquité grecque, Platon – notamment dans le *Cratyle* – et Aristote – particulièrement dans *De Interpretatione* – s'interrogeaient sur les rapports entre les mots, les idées et les choses. Les mots reflètent-ils la structure réelle du monde ou seulement nos manières humaines de le penser ? Les philosophes médiévaux – notamment Abélard, Thomas d'Aquin, Duns Scot et Guillaume d'Occam – reprendront ensuite ces interrogations à travers les débats sur les universaux et sur le lien entre langage, pensée et réalité. Ces questions, anciennes mais toujours actuelles, constituent l'un des points de départ de la philosophie du langage.

Plus tard, avec l'émergence des sciences modernes aux XVI^{e} et XVII^{e} siècles, la question du langage prit une importance nouvelle. Les progrès spectaculaires des mathématiques et de la physique nourrissaient alors l'espoir qu'une pensée véritablement rigoureuse devait pouvoir s'exprimer dans une langue elle aussi rigoureuse, débarrassée des ambiguïtés, des imprécisions et des équivoques du langage ordinaire.

Plusieurs penseurs rêvèrent ainsi de construire une forme de langage universel capable d'exprimer clairement les raisonnements scientifiques. René Descartes évoquait déjà la possibilité d'une langue philosophique fondée sur des idées simples et distinctes, où les erreurs de raisonnement deviendraient presque impossibles (Descartes, 1637). Gottfried Wilhelm Leibniz ira encore plus loin. Fasciné par les succès du calcul algébrique, il imagina un projet de *characteristica universalis*, sorte de langage symbolique universel dans lequel les disputes philosophiques pourraient être résolues par le calcul. Dans une formule célèbre, Leibniz envisage même qu'un jour deux savants en désaccord puissent simplement s'asseoir à une table et se dire : *Calculemus* [« calculons »] (Leibniz, 1685).

Cette ambition d'une langue parfaite ne relevait pas uniquement d'un rêve philosophique abstrait. Elle s'inscrivait dans un mouvement intellectuel beaucoup plus large : celui de la mathématisation croissante des sciences modernes. À mesure que les mathématiques semblaient capables de décrire avec précision les phénomènes physiques, plusieurs penseurs commencèrent à se demander si des méthodes comparables pourraient un jour être appliquées au raisonnement, à la pensée ou même au langage lui-même.

C'est de cette vaste tradition qu'est née ce qu'on appelle aujourd'hui la philosophie du langage (Frege, 1892 ; Russell, 1905 ; Wittgenstein, 1921). Cette discipline étudie notamment de quelle manière les mots acquièrent un sens, la façon dont les phrases expriment des pensées, les rapports entre langage et vérité, ainsi que les liens entre structures linguistiques et structures de la pensée. Certaines de ses questions sont devenues classiques : comment les mots parviennent-ils à désigner le monde ? Qu'est-ce qui fait que le mot « arbre » renvoie effectivement à des arbres réels plutôt qu'à une simple idée mentale ? Que signifie exactement comprendre une phrase ? Comprendre l'énoncé « Tous les cygnes sont blancs » consiste-t-il uniquement à saisir la signification des mots ou également à savoir dans quelles circonstances cet énoncé pourrait être vrai ou faux ? Le sens d'une phrase est-il entièrement contenu dans les mots qui la composent ou dépend-il aussi du contexte, des intentions du locuteur et des connaissances implicites des interlocuteurs ? Ainsi, lorsque quelqu'un affirme « Il y a du café sur la table », parle-t-il d'une tasse de café, de grains de café ou d'une tache laissée par une tasse renversée ? Plus profondément encore, la structure grammaticale d'une langue reflète-t-elle certaines structures fondamentales de la

pensée humaine ? Enfin, peut-on décrire le langage à l'aide de règles formelles comparables à celles utilisées en logique ou en mathématiques ? Ces interrogations, à la frontière de la linguistique, de la philosophie et de la logique, finiront par rencontrer directement certaines préoccupations de la logique mathématique moderne.

Afin de comprendre cette rencontre, il faut introduire une distinction devenue fondamentale dans la linguistique contemporaine, à savoir celle entre *syntaxe* et *sémantique*. La syntaxe désigne l'organisation formelle des expressions linguistiques : l'ordre des mots, les catégories grammaticales et les règles permettant de construire des phrases. La sémantique concerne quant à elle la signification de ces expressions. Une phrase peut ainsi être syntaxiquement correcte tout en étant difficilement interprétable sur le plan du sens (Carnap, 1934). La célèbre phrase de Chomsky « Colorless green ideas sleep furiously » – généralement traduite par « D'incolores idées vertes dorment furieusement » – respecte parfaitement les règles grammaticales de l'anglais, bien qu'elle semble dépourvue de signification claire (Chomsky, 1957). Cet exemple, sur lequel nous reviendrons plus loin, jouera un rôle important dans l'histoire que nous allons raconter.

La distinction entre syntaxe et sémantique n'est toutefois pas née avec Chomsky. Elle plonge ses racines dans une transformation beaucoup plus ancienne de la logique elle-même. Pendant des siècles, la logique héritée d'Aristote avait principalement été conçue comme une théorie du raisonnement portant sur des concepts et des catégories du réel. À partir du XIX^{e} siècle, cependant, plusieurs mathématiciens et logiciens entreprirent de transformer la logique en un véritable calcul symbolique (Boole, 1847 ; Frege, 1879). L'objectif n'était plus simplement d'étudier les raisonnements corrects, mais de construire des systèmes formels capables de manipuler des symboles selon des règles explicites, un peu à la manière des opérations algébriques.

Cette évolution fut l'un des grands bouleversements intellectuels de l'époque moderne [7]. Elle conduisit progressivement à l'apparition de la logique mathématique contemporaine, à la formalisation des fondements des mathématiques, puis à l'idée qu'un langage pouvait être étudié indépendamment de son contenu sémantique, uniquement à partir de sa structure formelle. Ce déplacement intellectuel – du sens vers la structure, des idées vers les règles de manipulation symbolique – constitue l'un des arrière-plans essentiels de la linguistique chomskyenne (Tomalin, 2006).

Pour comprendre comment une telle conception du langage a pu émerger, il faut maintenant remonter au XIX^{e} siècle et revenir à quelques figures majeures de la logique moderne, en commençant par George Boole, dont les travaux allaient profondément transformer la manière même de concevoir la logique.

7. Cette évolution intellectuelle possède également des racines philosophiques plus anciennes. Avec Descartes, la recherche d'une certitude entièrement rigoureuse conduit à une méfiance accrue envers les connaissances issues des sens et à une valorisation croissante des structures rationnelles de la pensée. Comme l'a montré Étienne Gilson, le doute cartésien vise précisément à découvrir une vérité absolument indubitable, indépendante des incertitudes de l'expérience sensible (Gilson, 1930). Sans conduire directement au formalisme moderne, cette orientation contribue néanmoins à installer l'idée qu'une connaissance véritablement certaine doit reposer sur des structures rationnelles explicites plutôt que sur les données immédiates de l'expérience.

3. Quand la logique devint un calcul

Pendant une grande partie de son histoire, la logique fut principalement conçue comme l'étude des formes correctes du raisonnement. Héritée d'Aristote, elle visait avant tout à analyser les relations entre concepts, propositions et catégories du réel. Le syllogisme classique

> *Tous les hommes sont mortels ;*
> *Socrate est un homme ;*
> *donc Socrate est mortel*

en constitue l'exemple emblématique. Dans cette tradition, la logique n'était pas encore un calcul au sens moderne du terme. Elle demeurait profondément liée à l'analyse philosophique des choses, des essences et des rapports entre les idées.

Au XIX^{e} siècle, cependant, plusieurs mathématiciens et logiciens – parmi lesquels George Boole et Augustus De Morgan [8] en Grande-Bretagne, Gottlob Frege en Allemagne et Giuseppe Peano en Italie – commencèrent à envisager la possibilité de traiter la logique elle-même comme une forme de calcul symbolique (Boole, 1847 ; Frege, 1879). L'objectif n'était plus seulement d'analyser philosophiquement les raisonnements corrects, mais de construire des systèmes capables de manipuler des symboles selon des règles explicites, un peu à la manière des opérations algébriques.

Parmi les principaux artisans de cette évolution figure le mathématicien britannique George Boole. Dans la préface de *The Mathematical Analysis of Logic : Being an Essay Towards a Calculus of Deductive Reasoning*, publié en 1847, Boole propose de concevoir la logique comme un véritable calcul formel [9]. Dès les premières pages de l'ouvrage, il écrit :

> *"The present work is not a mere collection of logical truths, but also an attempt to establish a calculus of deductive reasoning."* (Boole, 1847)
>
> *« Le présent ouvrage n'est pas un simple recueil de vérités logiques, mais aussi une tentative d'établir un calcul du raisonnement déductif. »*

Boole ne se contente pas d'affirmer que le raisonnement peut être traité comme un calcul. Il précise également ce qu'il entend par là :

> *"A true calculus, that is, a method resting upon the employment of symbols, whose laws of combination are known in general, and whose results admit of a consistent interpretation."*

8. Les travaux de De Morgan sur les relations logiques et sur l'algébrisation du raisonnement contribuèrent à plusieurs développements importants de la logique du XIX^{e} siècle. Ils influencèrent notamment le développement ultérieur de la logique relationnelle chez Peirce, Schröder, Frege et Russell.

9. Boole considérait lui-même *The Mathematical Analysis of Logic* comme une première esquisse imparfaite de son système logique, qu'il développera plus complètement dans *An Investigation of the Laws of Thought* (1854).

> *« Un véritable calcul, c'est-à-dire une méthode reposant sur l'emploi de symboles dont les lois de combinaison sont généralement connues et dont les résultats se prêtent à une interprétation cohérente. »*

La portée philosophique de cette formulation est immense. La logique cesse progressivement d'être conçue uniquement comme une théorie du raisonnement humain pour devenir un système de manipulation symbolique autonome dont les expressions peuvent être manipulées indépendamment de leur contenu immédiat.

Comme l'explique l'historien de la logique Józef Maria Bocheński, Boole ne considérait plus la logique comme une abstraction tirée des raisonnements ordinaires, mais comme une construction formelle à laquelle on cherche ensuite une interprétation (Bocheński, 1961). Le sens n'est plus nécessairement premier ; il peut devenir secondaire par rapport à la structure symbolique elle-même.

Ce renversement méthodologique allait profondément transformer la logique moderne (Frege, 1879). Il ouvrait la possibilité de construire des systèmes formels dont la cohérence dépendrait essentiellement de règles internes de manipulation symbolique. Cette idée allait bientôt trouver un terrain particulièrement fertile dans les recherches sur les fondements des mathématiques.

À la fin du XIX^{e} siècle, Gottlob Frege franchit une nouvelle étape importante avec la publication, en 1879, d'un traité de logique mathématique d'une grande technicité aujourd'hui généralement considéré comme l'un des textes fondateurs de la logique moderne, *Begriffsschrift : Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* [*Idéographie : Un langage formel de la pensée pure modelé sur celui de l'arithmétique*]. Frege y introduit un symbolisme logique très différent de la logique aristotélicienne classique. Son objectif n'est d'ailleurs pas seulement philosophique. En effet, Frege cherche avant tout à fournir aux mathématiques – et plus particulièrement à l'arithmétique – des fondements d'une rigueur entièrement nouvelle. Pour y parvenir, il développe un langage formel capable d'exprimer avec précision des relations logiques complexes que la logique traditionnelle peine à représenter adéquatement.

Ce nouveau symbolisme repose notamment sur l'usage systématique des quantificateurs [10], des variables et des relations. Grâce à ces outils, il devient possible de représenter formellement certains raisonnements mathématiques d'une grande complexité. En contrepartie toutefois, l'expression des énoncés ordinaires y devient souvent extrêmement artificielle et éloignée du langage naturel. Même des affirmations apparemment simples comme « Tous les hommes sont mortels » doivent être reconstruites dans un langage dont la structure diffère sensiblement de celle des langues ordinaires

$$\forall x \big(H(x) \to M(x)\big).$$

10. Les quantificateurs sont les symboles logiques permettant d'exprimer des énoncés généraux ou existentiels. En mathématiques, on les rencontre aujourd'hui sous les formes familières « $\forall$ » (« pour tout ») et « $\exists$ » (« il existe »). Les symboles modernes eux-mêmes seront introduits plus tard par Giuseppe Peano puis popularisés par Bertrand Russell, mais l'idée fondamentale d'une logique quantificationnelle systématique provient largement des travaux de Frege.

Or cette reformulation ne constitue pas une simple abréviation technique du langage ordinaire. Elle transforme la nature même de l'énoncé. Les mots « homme » et « mortel » cessent d'y fonctionner comme des concepts saisis dans l'expérience humaine ordinaire pour devenir des prédicats formels appliqués à des variables abstraites. Le raisonnement se trouve ainsi déplacé du contenu concret des affirmations vers la manipulation de structures symboliques régies par des règles syntaxiques explicites. Ce gain de précision logique s'accompagne donc d'un éloignement croissant des formes spontanées du langage naturel et de leur richesse sémantique. Cette tension entre puissance formelle et artificialité expressive traversera d'ailleurs une grande partie de l'histoire ultérieure de la logique mathématique.

L'influence de Frege sur la logique du XX^{e} siècle fut considérable, même si son œuvre demeura longtemps relativement peu connue en dehors des milieux spécialisés. Parmi ceux et celles qui contribuèrent le plus à faire connaître ses idées figure le philosophe, logicien et mathématicien britannique Bertrand Russell, l'une des grandes figures intellectuelles du début du XX^{e} siècle. Russell reconnaîtra lui-même l'importance fondamentale des travaux de Frege :

> *"In all questions of logical analysis, our chief debt is to Frege."* (Russell, 1944)
>
> *« En matière d'analyse logique, c'est à Frege que nous devons le plus »*

Frege ne voyait pas dans la nouvelle logique frégéenne un simple outil technique. Il croyait possible de montrer que l'arithmétique elle-même pouvait être fondée entièrement sur des principes logiques, ouvrant la voie à un programme philosophique qui prendra plus tard le nom de *logicisme*. Bertrand Russell et Alfred North Whitehead poursuivront cette ambition monumentale dans les *Principia Mathematica*, immense traité de logique mathématique et de fondements des mathématiques publié entre 1910 et 1913. Leur objectif était de reconstruire les mathématiques à partir d'un petit nombre de principes logiques fondamentaux à l'aide de règles de déduction purement formelles.

Cette période voit également se renforcer une idée appelée à jouer un rôle important dans toute l'histoire ultérieure de la logique et de la linguistique : l'attention se déplace progressivement du contenu des énoncés vers leurs structures formelles et les relations qu'ils entretiennent entre eux. Les mathématiques apparaissent alors de moins en moins comme une science portant sur des objets particuliers et de plus en plus comme l'étude de structures abstraites susceptibles d'être exprimées symboliquement.

Bertrand Russell soulignera lui-même l'importance croissante des relations dans les mathématiques modernes :

> *"A careful analysis of mathematical reasoning shows that all mathematical propositions are concerned with relations, and that the logic of relations has a more immediate bearing on mathematics than that of classes or propositions."* (Russell, 1903)
>
> *« Une analyse attentive du raisonnement mathématique montre que toutes les propositions mathématiques portent sur des relations, et que la logique des relations a une incidence plus directe sur les mathématiques que celle des classes ou des*

propositions. »

Dans cette perspective, la logique n'a plus seulement pour rôle de classifier des concepts ou d'analyser des raisonnements ordinaires ; elle devient un instrument permettant de formaliser des structures relationnelles complexes.

Ces développements transformèrent profondément la manière dont plusieurs mathématiciens et philosophes concevaient la logique. Celle-ci n'apparaissait plus seulement comme une théorie générale du raisonnement, mais de plus en plus comme l'étude de systèmes symboliques régis par des règles explicites de formation et de transformation.

Au début du XX^{e} siècle, cette orientation allait encore se renforcer avec les travaux de David Hilbert, dont les recherches sur les fondements des mathématiques accordaient une importance croissante aux systèmes symboliques, aux axiomes et aux règles de déduction formelle [11] (Hilbert, 1922 ; Hilbert & Ackermann, 1928). Les questions portant sur la cohérence, la manipulation symbolique et l'autonomie des systèmes formels prendraient alors une importance centrale. C'est dans ce climat intellectuel que commencerait progressivement à émerger l'idée qu'un langage lui-même pourrait être étudié comme une structure formelle relativement indépendante de sa signification immédiate.

Plusieurs décennies plus tard, ces idées exerceront une influence bien au-delà des fondements des mathématiques eux-mêmes. Certains linguistes et philosophes commenceront progressivement à envisager le langage naturel non plus seulement comme un véhicule de signification, mais aussi comme un système de structures symboliques gouverné par des règles formelles. C'est dans ce contexte intellectuel, largement façonné par la logique mathématique moderne, que s'inscriront plus tard les travaux de Noam Chomsky (Tomalin, 2006).

4. Hilbert, Carnap et le rêve d'une syntaxe autonome

Les recherches de David Hilbert sur les fondements des mathématiques marquent une nouvelle étape importante dans cette évolution intellectuelle. Hilbert nourrit alors une ambition considérable : montrer que les mathématiques peuvent être développées à partir de systèmes axiomatiques rigoureusement définis et manipulés selon des règles explicites de déduction (Hilbert, 1922).

Dans cette perspective, les mathématiques apparaissent de moins en moins comme une science portant directement sur des objets intuitivement accessibles — nombres, figures géométriques

11. Le terme « formalisme » est utilisé de manière très variable dans l'histoire et la philosophie des mathématiques. Il peut désigner aussi bien certaines méthodes axiomatiques que des positions philosophiques beaucoup plus radicales selon lesquelles les mathématiques ne seraient qu'un simple jeu de symboles dépourvus de signification intrinsèque. Or, les positions exactes de Hilbert sur ces questions demeurent historiographiquement complexes et ont souvent été simplifiées rétrospectivement. Hilbert lui-même ne se désignait pas comme « formaliste » au sens doctrinal fort du terme ; cette étiquette fut largement popularisée dans le contexte des controverses l'opposant notamment au mathématicien intuitionniste L. E. J. Brouwer (Grattan-Guinness, 2000). Nous éviterons donc ici d'utiliser le terme dans un sens philosophique trop rigide.

ou grandeurs — et de plus en plus comme un univers de symboles soumis à des opérations formelles rigoureusement contrôlées. Les signes ne servent plus seulement à représenter des objets mathématiques déjà compris indépendamment d'eux ; ils deviennent eux-mêmes le point de départ de l'analyse.

Cette manière de concevoir les mathématiques est souvent résumée – sans que l'on sache toutefois avec certitude si la formule provient directement de Hilbert lui-même – par l'aphorisme allemand : *« Am Anfang ist das Zeichen »* [« Au commencement est le signe »].

L'idée exprimée par cette formule est simple mais profonde : les symboles et les règles gouvernant leur manipulation tendent progressivement à acquérir une autonomie croissante par rapport aux contenus intuitifs ou aux significations immédiates qu'ils représentaient traditionnellement. Pendant une grande partie de l'histoire des mathématiques, les objets mathématiques étaient pourtant étroitement associés à certaines formes d'intuition ou d'expérience. Chez les Grecs, par exemple, la géométrie reposait largement sur les figures de l'espace visibles ou imaginables, tandis que les nombres étaient souvent conçus à travers des collections concrètes d'objets ou des rapports de grandeurs. Même les démonstrations les plus abstraites d'Euclide demeuraient profondément liées à des constructions géométriques interprétables intuitivement.

Avec le développement du formalisme moderne, cette relation commence progressivement à se transformer. Les symboles mathématiques ne sont plus seulement envisagés comme des moyens commodes de représenter des objets ou des raisonnements déjà compris intuitivement ; ils deviennent eux-mêmes des objets de manipulation soumis à des règles explicites. L'attention se porte alors de plus en plus sur la cohérence interne des systèmes symboliques et sur les opérations qu'il est permis d'y effectuer, indépendamment de toute interprétation immédiate dans le monde sensible ou dans l'intuition géométrique traditionnelle.

Cette évolution s'inscrit dans un contexte intellectuel plus large marqué par la recherche d'une rigueur absolue en mathématiques. Les paradoxes découverts dans la théorie des ensembles à la fin du XIX^{e} siècle avaient profondément ébranlé les fondements de la discipline (Russell, 1903 ; Zermelo, 1908). Le programme de Hilbert cherchait précisément à répondre à cette crise en montrant que les mathématiques pouvaient être sécurisées au moyen d'une analyse formelle des démonstrations elles-mêmes. Les célèbres théorèmes d'incomplétudes de Gödel montreront toutefois quelques années plus tard certaines limites fondamentales de ce programme (Gödel, 1931).

L'historien des mathématiques José Ferreirós, spécialiste de l'histoire de la théorie des ensembles et des fondements des mathématiques, résume bien cette évolution intellectuelle :

> *"In treating logical formulas as algebraic objects which can be manipulated independently of any semantic interpretation, Boole provided a key insight that prepared the emergence of the axiomatic orientation which later rose to prominence with David Hilbert's finitary metamathematical program."* (Ferreirós, 2001)
>
> *« En considérant les formules logiques comme des objets algébriques pouvant être manipulés indépendamment de toute interprétation sémantique, Boole a apporté une contribution décisive qui a ouvert la voie à l'émergence de l'approche axiomatique,*

laquelle a ensuite pris toute son importance avec le programme métamathematique finitaire de David Hilbert. »

Ces transformations de la logique et des fondements des mathématiques finirent par susciter une question nouvelle : une langue elle-même pourrait-elle être étudiée comme un système symbolique régi par des règles formelles ?

Cette idée jouera un rôle central dans les travaux du philosophe Rudolf Carnap, figure importante du positivisme logique et du Cercle de Vienne. Dans *The Logical Syntax of Language*, publié en 1934, Carnap défend l'idée qu'une langue peut être étudiée exclusivement du point de vue de sa syntaxe, c'est-à-dire des règles formelles gouvernant l'agencement des symboles, sans qu'il soit nécessaire de s'intéresser d'abord à leur signification. Carnap pousse même cette idée extrêmement loin, résumant sa position par une formule devenue célèbre :

"Mathematics is nothing but syntax of language." (Carnap, 1934)

« Les mathématiques ne sont rien d'autre que la syntaxe du langage. »

Cette manière de concevoir les systèmes formels exercera une influence durable sur plusieurs logiciens du XX^{e} siècle. Le logicien américain Haskell Curry – connu notamment pour ses travaux en logique combinatoire – ira même jusqu'à défendre une conception extrêmement radicale des systèmes formels, soutenant que ceux-ci ne portent en eux-mêmes aucune signification intrinsèque et ne concernent fondamentalement que *marks on paper* (Curry, 1958), c'est-à-dire de simples marques tracées sur le papier soumises à des règles de manipulation formelle.

C'est précisément cette autonomie relative de la syntaxe que Carnap cherchera à mettre au premier plan dans son analyse des langages formels. Afin d'illustrer son propos, il utilise un exemple devenu célèbre :

"Pirots karulize elatically." (Carnap, 1934)

« Les pirots karulisent élatiquement. »

La phrase ne possède aucune signification en anglais. Pourtant, soutient Carnap, il demeure possible d'en analyser correctement la structure grammaticale. On peut reconnaître qu'il s'agit d'une suite organisée selon un certain schéma syntaxique – un sujet, un verbe et un adverbe – même si les mots eux-mêmes sont dénués de sens. Autrement dit, la syntaxe peut être étudiée indépendamment du contenu sémantique des expressions.

Cette distinction entre structure syntaxique et signification exercera une influence considérable sur les développements ultérieurs de la linguistique générative et préparera plusieurs des distinctions qui deviendront centrales dans l'œuvre de Chomsky.

Dans la tradition philosophique classique, le langage était généralement conçu comme un moyen d'exprimer des pensées portant sur le monde. Chez Carnap et plusieurs représentants du positivisme logique, l'analyse se concentre plutôt sur les structures internes des systèmes symboliques et sur les règles qui gouvernent leur fonctionnement. Le langage tend alors à être

étudié comme un objet formel susceptible d'être décrit à l'aide d'axiomes, de règles de formation et de règles de transformation.

Cette idée de « règles de transformation » sera particulièrement importante pour la suite. Carnap (Carnap, 1934) distinguait notamment :

— des règles de formation permettant de construire des expressions bien formées ;

— et des règles de transformation permettant de dériver certaines expressions à partir d'autres.

Or cette distinction réapparaîtra presque directement quelques années plus tard dans la grammaire générative de Chomsky, sous la forme des règles de structure syntaxique et des transformations grammaticales.

Le philosophe et historien des rapports entre logique et linguistique Marcus Tomalin, qui a étudié en détail les rapports entre logique mathématique et linguistique générative, résume cette filiation de manière très explicite :

> *"With its focus on syntax as opposed to semantics, with its use of a logic-based notational system, with its identification of the analogy between a proof and the generation of a grammatical sentence, and with its use of such procedures as recursive definitions and axiomatic deduction, Transformational Generative Grammar unambiguously reveals its associations with the formal sciences."* (Tomalin, 2006)
>
> *« En mettant l'accent sur la syntaxe plutôt que sur la sémantique, en recourant à un système de notation fondé sur la logique, en établissant une analogie entre une démonstration et la construction d'une phrase grammaticalement correcte, et en utilisant des procédures telles que les définitions récursives et la déduction axiomatique, la grammaire générative transformationnelle révèle sans ambiguïté ses liens avec les sciences formelles. »*

Cette remarque est fondamentale pour comprendre la nouveauté du projet chomskyen. La grammaire générative ne surgit pas dans un vide intellectuel. Elle hérite d'une longue tradition de formalisation issue de la logique mathématique moderne, tradition dans laquelle le langage tend progressivement à être envisagé comme un système autonome de manipulation symbolique.

Toutefois, avant d'aborder directement Chomsky lui-même, il faut encore comprendre un autre élément essentiel du contexte intellectuel de l'époque : la fascination croissante exercée par les méthodes mathématiques sur l'ensemble des sciences humaines et sociales au milieu du XX^{e} siècle.

5. La tentation formaliste dans les sciences humaines

Les idées issues de la logique mathématique et du formalisme ne demeurèrent toutefois pas confinées aux fondements des mathématiques. Au cours du XX^{e} siècle, plusieurs disciplines des sciences humaines et sociales furent progressivement marquées par une fascination croissante pour les méthodes formelles, les modèles abstraits et les structures symboliques explicites.

Cette évolution s'inscrivait dans un contexte intellectuel plus large où les mathématiques et les sciences exactes jouissaient d'un prestige considérable. Les succès spectaculaires de la physique mathématique nourrissaient l'idée qu'une discipline devenait véritablement scientifique dans la mesure où elle parvenait à formaliser ses objets à l'aide de structures abstraites et de systèmes explicites de règles.

Cette influence se manifesta dans plusieurs domaines. En économie et en science politique, la théorie des jeux cherchait à modéliser les décisions humaines à partir de structures mathématiques abstraites (von Neumann & Morgenstern, 1944 ; Nash, 1950). En psychologie cognitive et en informatique théorique, l'esprit humain commençait à être décrit comme un système de traitement computationnel de l'information (Turing, 1950 ; Fodor, 1975). La cybernétique et la théorie de l'information contribuèrent également à diffuser l'idée que des phénomènes complexes pouvaient être analysés au moyen de systèmes formels, de flux symboliques et de règles générales de transformation (Wiener, 1948 ; Shannon, 1948).

La linguistique n'échappa pas à cette tendance. Bien avant Chomsky, plusieurs linguistes américains tentaient déjà de rendre l'étude du langage plus systématique et plus « scientifique ». Sous l'influence du behaviorisme dominant dans les sciences humaines américaines, l'accent était souvent mis sur les aspects observables du langage : distribution des mots, régularités formelles, classifications syntaxiques, structures récurrentes (Skinner, 1957). Les questions relatives au sens, jugées trop subjectives ou difficilement vérifiables, tendaient fréquemment à être mises de côté.

C'est dans cet environnement intellectuel que Chomsky entreprend ses premiers travaux dans les années 1950. Le contexte est important. Les développements récents de la logique mathématique, de la théorie de la récursivité et de la métamathématique semblaient alors offrir de nouveaux outils capables de traiter scientifiquement les structures linguistiques. Cette conception du langage comme système explicite de règles apparaît clairement dans la définition que Chomsky donne de la grammaire générative :

> *"By a generative grammar I mean simply a system of rules that in some explicit and well-defined way assigns structural descriptions to sentences."* (Chomsky, 1957)
>
> *« Par grammaire générative, j'entends simplement un système de règles qui, d'une manière explicite et bien définie, attribue des descriptions structurelles aux phrases. »*

Le langage commence alors à être envisagé comme un système explicite de règles capable de générer un ensemble potentiellement infini d'énoncés bien formés.

L'influence de Rudolf Carnap et du positivisme logique sur ce climat intellectuel est manifeste. Comme nous l'avons vu, Carnap défendait l'idée qu'une langue pouvait être étudiée du point de vue purement syntaxique, indépendamment de toute signification. Cette orientation allait profondément marquer la linguistique générative naissante (Carnap, 1934).

Chez Chomsky, l'importance accordée à la syntaxe devient encore plus explicite dans *Syntactic Structures* (1957), l'ouvrage qui contribuera à imposer la linguistique générative sur la scène intellectuelle internationale. Comme il l'écrit lui-même :

> *"Syntax is the study of linguistic form."* (Chomsky, 1957)
>
> *« La syntaxe est l'étude de la forme linguistique. »*

Le langage peut alors être étudié, au moins dans un premier temps, comme un système de structures formelles analysables indépendamment de leur contenu sémantique immédiat. Quelques lignes plus loin, Chomsky précise encore plus explicitement cette orientation :

> *"The subject of investigation in the following pages will be syntactic structure, and we shall study it as an independent aspect of linguistic theory."* (Chomsky, 1957)
>
> *« Le sujet d'étude des pages qui suivent sera la structure syntaxique, et nous l'étudierons comme un aspect indépendant de la théorie linguistique. »*

La formulation est importante. La syntaxe cesse ici d'être envisagée simplement comme un instrument au service de la signification ; elle devient un objet d'étude possédant ses propres régularités et ses propres principes d'organisation. Cette idée jouera un rôle central dans toute la linguistique générative.

L'exemple le plus célèbre associé à cette conception est sans doute la phrase :

> *"Colorless green ideas sleep furiously."* (Chomsky, 1957)
>
> *« D'incolores idées vertes dorment furieusement. »*

Cette phrase est grammaticalement correcte, bien qu'elle paraisse difficilement interprétable du point de vue du sens. Chomsky utilise précisément ce type d'exemple pour montrer qu'une phrase peut être syntaxiquement bien formée même lorsqu'elle semble sémantiquement étrange ou absurde. L'objectif est de démontrer que la grammaticalité ne peut être réduite à la signification.

On retrouve ici très clairement l'héritage de Carnap et d'une tradition intellectuelle issue de la logique mathématique moderne : la structure formelle peut être étudiée indépendamment de son interprétation sémantique immédiate.

Mais l'influence de la logique mathématique sur Chomsky va beaucoup plus loin qu'une simple distinction entre syntaxe et sémantique. Plusieurs aspects fondamentaux de la grammaire générative portent la marque de cette filiation intellectuelle. Les phrases y sont produites

à partir de règles formelles, de procédures récursives, de transformations syntaxiques et de mécanismes de dérivation rappelant parfois les systèmes axiomatiques de la logique.

L'analogie entre démonstration logique et génération grammaticale devient alors centrale. De même qu'un système logique peut produire des théorèmes à partir d'axiomes et de règles de déduction, une grammaire générative doit pouvoir produire un ensemble potentiellement infini de phrases à partir d'un nombre fini de règles syntaxiques.

Cette idée de générativité fascinait profondément plusieurs intellectuels du milieu du XX^{e} siècle. Elle semblait offrir une solution élégante à un problème fondamental : comment expliquer la capacité apparemment ouverte et potentiellement illimitée des locuteurs à comprendre et produire de nouvelles phrases qu'ils n'ont jamais rencontrées auparavant ? La réponse de Chomsky consistait précisément à postuler l'existence de structures formelles innées permettant cette créativité linguistique.

Le langage humain commence alors à apparaître moins comme un répertoire d'expressions apprises que comme un système computationnel capable de générer des structures symboliques (Chomsky, 1965 ; Fodor, 1975). Le rapprochement avec certaines branches de la logique et des mathématiques devient dès lors difficile à ignorer.

Cependant, cette montée en puissance des approches formalistes dans l'étude du langage allait également soulever des questions philosophiques importantes. Peut-on réellement étudier le langage humain indépendamment du sens, du contexte et de l'expérience vécue ? Une structure syntaxique suffit-elle réellement à rendre compte de la signification ? Et jusqu'où peut-on pousser l'analogie entre langue naturelle et système formel ?

Plusieurs linguistes et philosophes du langage ont justement reproché à certaines versions de la linguistique générative de traiter les phrases comme de simples objets syntaxiques abstraits, détachés non seulement des conditions concrètes de leur usage effectif, mais également du rapport direct que les expressions linguistiques entretiennent avec le monde et l'expérience humaine. Avec cette perspective critique, la signification ne peut être réduite à une structure formelle sous-jacente ni être reléguée à un rôle secondaire par rapport à la syntaxe. Le langage humain n'est pas d'abord un système formel autonome manipulant des symboles selon des règles abstraites ; il constitue avant tout une activité signifiante enracinée dans l'expérience humaine et dans le rapport concret des locuteurs au monde. Il sert avant tout à désigner des réalités, à exprimer des expériences, à orienter l'action et à permettre à des sujets humains d'entrer en relation avec le monde qui les entoure.

Ce sont précisément ces tensions entre structure formelle et signification qu'il convient maintenant d'examiner plus attentivement.

6. Le langage peut-il être réduit à la syntaxe ?

L'autonomie de la syntaxe constitue sans doute l'une des idées les plus caractéristiques – et les plus radicales – de la linguistique chomskyenne. Dans cette perspective, le langage humain peut être étudié avant tout comme un système formel de règles génératives largement indépendant du sens, du contexte et des usages concrets des locuteurs (Chomsky, 1957 ; Carnap, 1934). L'idée rappelle fortement certaines orientations déjà présentes chez Carnap : les structures syntaxiques seraient régies par leurs propres régularités internes et pourraient être analysées indépendamment de leur interprétation sémantique immédiate.

Pour comprendre l'enjeu, il faut revenir à l'ambition scientifique qui anime une grande partie du projet génératif. Chomsky cherche à dégager les mécanismes formels sous-jacents à la production des phrases grammaticales. Dans cette conception, la syntaxe cesse alors d'être un simple instrument au service du sens ; elle devient la structure fondamentale à partir de laquelle les autres dimensions du langage doivent être reconstruites (Chomsky, 1965). La signification tend alors à apparaître comme une composante dérivée venant s'ajouter à des structures déjà produites par les mécanismes formels de la grammaire.

Chomsky ne présente pas cette séparation entre syntaxe et signification comme une simple commodité pédagogique ou méthodologique provisoire. Dans les premières formulations de la grammaire générative, la mise à distance du sens est explicitement justifiée par des considérations épistémologiques :

> *"Our desire to place no reliance on meaning in systematic development is motivated by a feeling that the theory of meaning fails to meet certain minimum requirements of objectivity and operational verifiability."* (Chomsky, 1957)
>
> *« Notre volonté de ne pas nous fonder sur la signification dans le développement systématique découle du sentiment que la théorie de la signification ne répond pas à certaines exigences minimales d'objectivité et de vérifiabilité opérationnelle »*

La phrase est révélatrice. Le problème n'est pas seulement que la signification soit difficile à étudier ; elle apparaît ici comme insuffisamment objective pour servir de fondement rigoureux à une science du langage. La syntaxe tend alors à devenir le véritable noyau scientifique de la linguistique (Tomalin, 2006).

Au fil des décennies, la théorie chomskyenne évoluera considérablement. Toutefois, malgré les nombreuses transformations, une idée demeurera relativement stable : les propriétés fondamentales du langage sont d'abord recherchées du côté des structures syntaxiques et des mécanismes formels plutôt que dans l'ancrage sémantique des expressions dans l'expérience humaine et le rapport au monde.

Le langage naturel tend alors à être envisagé comme une sorte de système computationnel interne opérant sur des symboles selon des règles formelles (Chomsky, 1965 ; Fodor, 1975).

C'est toutefois précisément ici qu'apparaissent plusieurs difficultés philosophiques importantes.

Dans l'usage réel du langage, la signification ne dépend jamais uniquement de la structure syntaxique des phrases. Elle dépend constamment : du contexte ; de la situation concrète d'énonciation ; des connaissances partagées entre les interlocuteurs ; des intentions du locuteur ; et, plus largement, du rapport des sujets parlants au monde et à leur expérience vécue (Duffley, 2020).

Une difficulté classique apparaît immédiatement dès qu'on tente de réduire le langage à sa seule structure formelle. Dans l'usage réel des langues naturelles, la compréhension d'un énoncé dépend constamment d'éléments qui débordent largement la syntaxe elle-même. Prenons une phrase aussi simple que :

« Il fait froid ici. »

Selon le contexte, cette phrase peut : (1) constituer une simple observation météorologique ; (2) exprimer une plainte ; (3) servir de demande implicite pour fermer une fenêtre ; (4) ou encore fonctionner comme une invitation indirecte à modifier la température de la pièce.

La structure grammaticale demeure pourtant identique dans chacun de ces cas. Ce qui varie, c'est : (1) la situation concrète ; (2) les intentions du locuteur ; (3) les attentes des interlocuteurs ; (4) et l'ensemble des connaissances implicites partagées.

Plus révélateur encore, certaines expressions syntaxiquement atypiques demeurent néanmoins parfaitement intelligibles pour les locuteurs. L'énoncé :

« Étonnamment de gens ne le savent pas. »

paraît difficilement analysable du point de vue d'une syntaxe strictement régulière, puisqu'un adverbe comme « étonnamment » ne peut normalement remplir ici une fonction déterminative. Pourtant, l'interprétation ne pose pratiquement aucune difficulté : le locuteur comprend spontanément qu'il est question d'« une quantité étonnamment grande de gens ». Le sens mobilise ainsi des mécanismes interprétatifs et pragmatiques qui débordent les seules règles syntaxiques explicites (Duffley, 2018).

Le sens effectif d'un énoncé semble ainsi difficilement réductible à sa seule organisation syntaxique. Autrement dit, comprendre une phrase exige souvent beaucoup plus que l'application mécanique de règles syntaxiques. Le sens mobilise des connaissances implicites, des attentes culturelles, des habitudes d'interprétation et une familiarité générale avec le monde vécu.

Il ne s'agit pas de nier l'utilité des modèles formels en linguistique. Les approches génératives ont produit des outils analytiques extrêmement puissants, profondément transformé l'étude scientifique du langage et permis de mettre au jour d'importantes régularités structurelles. Les difficultés évoquées ci-dessus invitent cependant à distinguer la fécondité d'un formalisme de sa portée explicative. Comme plusieurs penseurs du XX^{e} siècle l'ont souligné, un système symbolique cohérent ne garantit pas, à lui seul, une compréhension adéquate des phénomènes humains qu'il cherche à décrire. Le langage naturel possède peut-être une dimension irréductiblement contextuelle et sémantique qui résiste, au moins en partie, à une reconstruction purement syntaxique.

Le philosophe Henry Veatch insistait précisément sur cette distinction entre cohérence symbolique et rapport au réel (Veatch, 1974). Un système formel peut posséder une grande rigueur interne tout en demeurant relativement éloigné de l'expérience concrète du monde. Cette remarque touche directement certaines versions de la linguistique générative : une théorie syntaxique très élaborée ne suffit pas nécessairement à rendre compte de la manière dont les êtres humains utilisent effectivement le langage pour signifier, comprendre, agir et entrer en relation avec le monde qui les entoure.

Le cas de Chomsky apparaît ainsi particulièrement révélateur d'une tension intellectuelle plus large ayant marqué une partie importante du XX^{e} siècle : jusqu'où peut-on étendre aux phénomènes humains les méthodes initialement développées pour les systèmes formels et les structures symboliques abstraites ? L'histoire de la linguistique générative montre ainsi à la fois la puissance explicative de cette orientation et les limites philosophiques qui apparaissent lorsque la syntaxe tend à être séparée du sens, du contexte et de l'expérience vécue.

7. De la logique mathématique à la linguistique générative

Vue rétrospectivement, la parenté intellectuelle entre la logique mathématique moderne et la linguistique chomskyenne est difficile à ignorer. Bien sûr, Chomsky ne cherchait pas à transformer la linguistique en branche des mathématiques au sens strict. Son objectif demeurait l'étude du langage humain. Mais plusieurs éléments centraux de la grammaire générative portent clairement la marque du grand mouvement de formalisation issu de la logique des XIX^{e} et XX^{e} siècles (Tomalin, 2006).

Cette filiation repose notamment sur une idée fondamentale : celle selon laquelle un système complexe peut être engendré à partir d'un petit nombre de règles explicites. Chez Frege et Russell, cette idée concernait principalement le raisonnement logique et les mathématiques ; chez Hilbert, les systèmes axiomatiques ; chez Carnap, les langages formels eux-mêmes. Chez Chomsky enfin, elle devient le fondement d'une théorie générale de la faculté de langage (Frege, 1879 ; Carnap, 1934 ; Chomsky, 1957).

La proximité structurelle est frappante. Comme les systèmes logiques formels, la grammaire générative repose sur des symboles primitifs, des règles de formation, des mécanismes de transformation et des procédures récursives permettant de produire un ensemble potentiellement infini d'expressions. Plusieurs notions centrales de la linguistique chomskyenne – dérivation, transformation, générativité, computation – s'inscrivent directement dans cet univers conceptuel hérité des sciences formelles modernes (Tomalin, 2006).

Un exemple particulièrement suggestif est fourni par une observation de Chomsky lui-même, reprise notamment par Hinzen et Uriagereka (2006). Dans le cadre du Programme minimaliste, l'opération fondamentale est appelée *Merge*. Elle consiste à prendre deux objets syntaxiques et à les combiner en une nouvelle unité syntaxique, laquelle peut ensuite être combinée à son tour avec d'autres objets par la même opération. C'est cette possibilité d'application répétée qui confère à Merge son caractère génératif.

Supposons maintenant qu'un lexique ne contienne qu'un seul élément, noté *one*. Si l'on applique successivement l'opération Merge, on obtient

$$\begin{aligned} &\mathit{one}, \\ \mathrm{Merge}(\mathit{one}) &= \{\mathit{one}\}, \\ \mathrm{Merge}(\{\mathit{one}\}) &= \{\{\mathit{one}\}\}, \\ \mathrm{Merge}(\{\{\mathit{one}\}\}) &= \{\{\{\mathit{one}\}\}\}, \\ &\vdots \end{aligned}$$

Chomsky avance que cette construction fournit un analogue de la fonction successeur utilisée pour construire les entiers naturels. À partir d'un unique élément lexical et d'une seule opération syntaxique, on obtient ainsi une suite potentiellement infinie de structures hiérarchiques. Comme l'observent Hinzen et Uriagereka, cette analogie rapproche directement l'opération Merge de certaines constructions fondamentales de la théorie des ensembles et de l'arithmétique. Les mathématiciens y reconnaîtront en effet immédiatement une construction directement inspirée de celle des entiers naturels de Zermelo en théorie des ensembles, où chaque nouvel entier est obtenu en prenant le singleton du précédent.

L'intérêt de cet exemple ne réside évidemment pas dans la construction elle-même, mais dans l'idée générale qu'elle illustre. À partir d'un ensemble extrêmement réduit de primitives et d'une unique opération combinatoire, il devient possible d'engendrer une famille potentiellement illimitée de structures de complexité croissante. C'est précisément ce que Chomsky entend par une grammaire générative : un système fini de principes capables de produire un ensemble potentiellement infini d'expressions linguistiques.

Cette idée devient encore plus remarquable lorsque l'on considère le rôle de la récursivité. La récursivité intervient lorsqu'une règle ou une opération peut être réappliquée à ses propres résultats, engendrant ainsi une suite potentiellement infinie de nouvelles structures. Dans l'exemple précédent, l'opération Merge peut être réappliquée indéfiniment au résultat de son application précédente, sur un principe analogue à celui de la construction itérative des entiers naturels de Zermelo. Cette possibilité d'obtenir une infinité potentielle à partir d'un nombre fini de principes constitue l'un des liens les plus profonds entre la logique mathématique, la théorie de la calculabilité et la linguistique générative.

Cette influence ne concerne toutefois pas seulement certaines méthodes particulières. Elle touche plus profondément une certaine manière de concevoir la scientificité elle-même. Une théorie rigoureuse tend alors à être envisagée comme un système explicite, formalisable et génératif, capable de produire systématiquement les objets qu'elle étudie à partir d'un nombre limité de principes abstraits (Hilbert, 1922).

L'influence de ces idées dépasse aujourd'hui largement la linguistique. Les différentes classes de grammaires distinguées par Chomsky – grammaires régulières, hors contexte, sensibles au contexte et récursivement énumérables – constituent la célèbre hiérarchie de Chomsky, qui joue désormais un rôle fondamental en informatique théorique. Elle permet notamment de

classifier les langages formels selon leur puissance expressive et les automates nécessaires à leur reconnaissance. Cette classification établit un lien profond entre la linguistique, la logique mathématique et la théorie du calcul, illustrant la fécondité des échanges entre ces disciplines.

C'est précisément cette orientation qui contribuera à faire de Chomsky l'une des figures majeures de la révolution cognitive du XX^{e} siècle. Son opposition au behaviorisme ne consistait pas uniquement à réhabiliter l'existence de structures mentales internes ; elle participait également à l'émergence d'une nouvelle conception de la cognition comme traitement formel de structures symboliques (Gardner, 1985 ; Fodor, 1975).

8. Conclusion

Il serait évidemment excessif de prétendre que la linguistique chomskyenne constitue simplement une application directe de la logique mathématique au langage humain. Les différences entre ces domaines demeurent considérables, et Chomsky lui-même ne s'est jamais présenté comme un logicien au sens strict du terme. Pourtant, le survol historique que nous venons de proposer révèle une continuité intellectuelle profonde souvent négligée : celle reliant le grand mouvement de formalisation de la logique moderne à l'ambition de construire une théorie scientifique du langage fondée sur des structures syntaxiques explicites et génératives (Tomalin, 2006).

De Boole à Frege, de Russell à Hilbert, puis de Carnap à Chomsky, on voit progressivement se développer une manière nouvelle de concevoir les systèmes symboliques. Chez les premiers logiciens du XIX^{e} siècle, il s'agit surtout de transformer la logique en un calcul formel capable de manipuler des symboles selon des règles explicites. Avec les recherches sur les fondements des mathématiques au début du XX^{e} siècle, l'attention accordée aux structures symboliques et aux systèmes axiomatiques devient encore plus centrale.

Chez Carnap et certains représentants du positivisme logique, cette orientation atteint toutefois un degré particulièrement radical : le langage lui-même commence désormais à être étudié principalement du point de vue de sa syntaxe formelle, relativement indépendamment de sa signification. Carnap ira même jusqu'à soutenir que certains problèmes philosophiques traditionnels proviennent essentiellement d'une mauvaise compréhension de la syntaxe logique du langage (Carnap, 1934) [12].

C'est néanmoins cette phase fortement syntaxique et formaliste de son œuvre qui exercera une influence importante sur le climat intellectuel dans lequel émergera plus tard la linguistique générative chomskyenne.

L'histoire de la linguistique chomskyenne apparaît ainsi comme un chapitre particulier d'une évolution intellectuelle beaucoup plus vaste, soit celle de l'influence croissante exercée par les sciences formelles dans la culture moderne (Tomalin, 2006 ; Gardner, 1985). À partir du moment

12. Fait intéressant cependant, Carnap lui-même prendra plus tard une certaine distance par rapport aux versions les plus strictement syntaxiques de cette approche, accordant progressivement une importance croissante aux questions sémantiques (Carnap, 1942).

où les mathématiques semblaient capables de décrire avec une précision remarquable des pans entiers de la réalité physique, il devenait naturel que d'autres disciplines cherchent à importer, au moins partiellement, certaines de leurs méthodes, de leurs structures et de leurs idéaux de rigueur.

Mais cette histoire rappelle également que le langage humain possède des caractéristiques singulières qui compliquent toute réduction purement formelle. Contrairement aux systèmes symboliques artificiels construits par les logiciens, les langues naturelles demeurent profondément enracinées dans l'expérience vécue, le contexte, les pratiques culturelles, les intentions des locuteurs et le rapport concret au monde (Duffley, 2020 ; Putnam, 1975).

La signification déborde constamment les structures syntaxiques qui cherchent à l'organiser (Quine, 1951). Or cette résistance du sens aux reconstructions purement formelles constitue précisément l'une des difficultés philosophiques majeures auxquelles se heurtent les approches inspirées des sciences formelles lorsqu'elles sont appliquées au langage humain.

C'est peut-être là que réside l'intérêt philosophique le plus durable de la filiation intellectuelle que nous avons esquissée. Elle montre à quel point le prestige considérable acquis par les mathématiques et la logique modernes a contribué à imposer progressivement l'idée qu'un phénomène devient véritablement scientifique lorsqu'il peut être reconstruit sous forme de système formel explicite. Dans ce contexte, la cohérence symbolique et la puissance des structures abstraites ont pu tendre à être confondues avec une compréhension adéquate des phénomènes eux-mêmes (Veatch, 1974).

Or, l'histoire du langage humain rappelle précisément les limites possibles d'une telle assimilation. Les formalismes peuvent sans contredit révéler certaines régularités profondes et produire des outils analytiques puissants. Mais lorsqu'ils sont étendus à des phénomènes aussi riches, contextuels et sémantiquement flexibles que les langues naturelles, ils risquent également de laisser dans l'ombre des dimensions essentielles de l'expérience linguistique concrète. La cohérence interne d'un système symbolique ne garantit pas automatiquement qu'il saisisse adéquatement la réalité humaine qu'il prétend modéliser.

Le cas de Chomsky est particulièrement révélateur à cet égard. Sa linguistique demeure profondément marquée par l'héritage intellectuel de la logique mathématique moderne et par l'ambition d'étudier le langage comme un système génératif régi par des règles explicites de combinaison symbolique. Mais l'histoire de cette filiation met également en lumière les tensions qui apparaissent lorsque des méthodes initialement développées pour les mathématiques et les langages formels sont étendues aux langues naturelles, dont les usages effectifs débordent constamment les cadres purement syntaxiques (Chomsky, 1957 ; Tomalin, 2006).

Qu'on adhère ou non aux ambitions les plus fortes du projet chomskyen, la filiation reliant logique mathématique et linguistique générative demeure un épisode particulièrement révélateur de l'histoire intellectuelle moderne. Elle illustre la puissance d'attraction exercée par les sciences formelles au XX^{e} siècle, mais aussi les difficultés qui apparaissent lorsque les exigences de rigueur syntaxique en viennent à occuper une place disproportionnée dans l'étude de phénomènes profondément enracinés dans le sens, le contexte et les usages humains concrets.

Bibliographie

[1] Aristote. (1984). *De Interpretatione* (J. Tricot, trad.). Paris : Vrin. (L'œuvre originale date du IV^{e} siècle av. J.-C.)

[2] Barsky, R. F. (1997). *Noam Chomsky : A life of dissent*. Cambridge, MA : MIT Press.

[3] Bocheński, J. M. (1961). *A History of Formal Logic*. South Bend, IN : University of Notre Dame Press.

[4] Boole, G. (1847). *The mathematical analysis of logic : Being an essay towards a calculus of deductive reasoning*. Cambridge, UK : Macmillan.

[5] Carnap, R. (1934). *Logische Syntax der Sprache*. Vienne, Autriche : Springer. (Traduction anglaise : *The logical syntax of language*, A. Smeaton, trad., London, UK : Kegan Paul, 1937.)

[6] Carnap, R. (1942). *Introduction to semantics*. Cambridge, MA : Harvard University Press.

[7] Chomsky, N. (1957). *Syntactic structures*. The Hague, Netherlands : Mouton.

[8] Chomsky, N. (1959). Review of *Verbal behavior*, by B. F. Skinner. *Language*, *35*(1), 26–58.

[9] Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA : MIT Press.

[10] Curry, H. B. (1958). *Outlines of a formalist philosophy of mathematics*. Amsterdam, Netherlands : North-Holland.

[11] Descartes, R. (1637). *Discours de la méthode*. Paris : Flammarion.

[12] Duffley, P. J. (2018). *How can an adverb define an extension and have subject function in a sentence ? – Pragmatiquement, bien sûr !*. Communication présentée au 4th Usage-Based Linguistics Conference, University of Tel Aviv.

[13] Duffley, P. J. (2020). *Linguistic meaning meets linguistic form*. Oxford, UK : Oxford University Press.

[14] Ferreirós, J. (2001). The road to modern logic — An interpretation. *The Bulletin of Symbolic Logic*, *7*(4), 441–484.

[15] Ferreirós, J. (2007). *Labyrinth of thought : A history of set theory and its role in modern mathematics* (2^{e} éd.). Basel, Switzerland : Birkhäuser.

[16] Fodor, J. A. (1975). *The language of thought*. New York, NY : Crowell.

[17] Frege, G. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle, Allemagne : Louis Nebert.

[18] Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, *100*, 25–50.

[19] Gardner, H. (1985). *The mind's new science : A history of the cognitive revolution*. New York, NY : Basic Books.

[20] Gilson, É. (1930). *Études sur Descartes*. Paris : Vrin.

[21] Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, *38*, 173–198.

[22] Grattan-Guinness, I. (2000). *The Search for Mathematical Roots, 1870–1940 : Logics, Set Theories and the Foundations of Mathematics from Cantor through Russell to Gödel*. Princeton : Princeton University Press.

[23] Harris, R. A. (1993). *The linguistics wars*. New York, NY : Oxford University Press.

[24] Hilbert, D. (1922). Neubegründung der Mathematik. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, *1*, 157–177.
[25] Hilbert, D., & Ackermann, W. (1928). *Grundzüge der theoretischen Logik*. Berlin, Allemagne : Springer.
[26] Hinzen, W., & Uriagereka, J. (2006). On the metaphysics of linguistics. *Erkenntnis*, *65*(1), 71–96.
[27] Leibniz, G. W. (1685). *Philosophical essays* (R. Ariew & D. Garber, Eds.). Indianapolis, IN : Hackett.
[28] McGilvray, J. (Ed.). (2005). *The Cambridge companion to Chomsky*. Cambridge, UK : Cambridge University Press.
[29] Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences of the United States of America*, *36*(1), 48–49.
[30] Occam, G. d'. (1974). *Summa logicae* (P. Boehner, trad.). St. Bonaventure, NY : Franciscan Institute. (Oeuvre originale publiée vers 1323.)
[31] Platon. (1967). *Cratyle* (L. Méridier, trad.). Paris : Les Belles Lettres.
[32] Putnam, H. (1975). The meaning of "meaning". In K. Gunderson (Ed.), *Language, mind and knowledge* (pp. 131–193). Minneapolis, MN : University of Minnesota Press.
[33] Quine, W. V. O. (1951). Two dogmas of empiricism. *The Philosophical Review*, *60*(1), 20–43.
[34] Russell, B. (1903). *The principles of mathematics*. Cambridge, UK : Cambridge University Press.
[35] Russell, B. (1905). On denoting. *Mind*, *14*(56), 479–493.
[36] Russell, B. (1944). My mental development. In P. A. Schilpp (Ed.), *The philosophy of Bertrand Russell* (pp. 1–20). Evanston, IL : Library of Living Philosophers.
[37] Russell, B., & Whitehead, A. N. (1910–1913). *Principia mathematica* (Vols. 1–3). Cambridge, UK : Cambridge University Press.
[38] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423.
[39] Skinner, B. F. (1957). *Verbal behavior*. New York, NY : Appleton-Century-Crofts.
[40] Tomalin, M. (2006). *Linguistics and the formal sciences : The origins of generative grammar*. Cambridge, UK : Cambridge University Press.
[41] Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433–460.
[42] Veatch, H. B. (1974). *Aristotle : A Contemporary Appreciation*. Bloomington, IN : Indiana University Press.
[43] von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ : Princeton University Press.
[44] Wiener, N. (1948). *Cybernetics : Or control and communication in the animal and the machine*. Cambridge, MA : MIT Press.
[45] Wittgenstein, L. (1921). *Tractatus logico-philosophicus* (D. Pears & B. McGuinness, trad.). London, UK : Routledge.
[46] Zermelo, E. (1908). Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen*, *65*, 261–281.

Association
mathématique
du Québec

Chronique didactique

Intérêt de la connaissance-obstacle pour l'enseignement [1]

Sophie René de Cotret
Université de Montréal

Mots clés : Connaissance-obstacle, domaine de validité, erreurs, rencontre de l'ignorance, sens de la division.

Dans une précédente chronique didactique (*Pourquoi résiste-t-on à la preuve ?* mai 2025), l'idée de connaissance-obstacle avait été évoquée en suggérant que nous y reviendrions. C'est ce que nous faisons aujourd'hui.

Voici, à peu de choses près, ce que nous en disions :

> Une connaissance-obstacle est une connaissance qui a un domaine de validité, c'est-à-dire un ensemble de problèmes pour lesquels son usage est efficace, et qui, notamment du fait de son succès éprouvé, résiste à l'adaptation à une nouvelle situation. Plus la connaissance a bien fonctionné, plus il devient coûteux de la modifier. On l'utilisera donc possiblement à tort, faute de pouvoir reconnaître son domaine de validité et d'en contrôler l'utilisation, ou faute de pouvoir trouver mieux.

La présente chronique revient sur ce concept de connaissance-obstacle en didactique afin d'en montrer l'intérêt pour l'enseignement et, tout particulièrement, pour le traitement des erreurs.

1. Pour plonger dans le vif du sujet

Pour plonger dans le vif du sujet, je vous propose la tâche suivante :

> Composez un problème « à contexte » qui demande comme solution la division $\frac{2}{3} \div \frac{3}{5}$ et résolvez-le.

Je vous laisse quelques minutes...

1. Je tiens à remercier France Caron pour sa lecture attentive et ses précieux commentaires.

 © Association mathématique du Québec

Il m'est évidemment difficile de présumer de votre réponse. Toutefois, si je me fie à mes 20 dernières années d'expérience comme professeure de didactique, je parierais que vous avez d'abord cherché un contexte de partage. C'est ce que j'ai pu constater en soumettant au fil des ans cette tâche à près de 500 étudiants et étudiantes de 2e année au baccalauréat en enseignement des mathématiques au secondaire. En effet, très peu réussissaient à composer un problème « à contexte » dont la solution demandait effectivement de faire l'opération $\frac{2}{3} \div \frac{3}{5}$. Les textes obtenus, souvent à propos de partage de pizzas, étaient plutôt alambiqués et pouvaient renvoyer à une multiplication par $\frac{3}{5}$ plutôt qu'à une division, sans toujours que le sujet s'en aperçoive. Et, dans plusieurs cas, la feuille restait blanche.

À la retraite depuis quelques années, j'ai voulu vérifier si les réponses étaient toujours les mêmes aujourd'hui. J'ai donc soumis cette tâche à 25 étudiants et étudiantes de 1re année au programme de baccalauréat ou de maîtrise qualifiante en enseignement des mathématiques au secondaire. Les réponses recueillies sont semblables à celles que j'ai toujours obtenues.

Voici quelques exemples, en commençant avec deux des 20 productions inadéquates recueillies.

- Tentative d'adaptation d'un contexte de partage (8 productions)

  Exemple 1 – Dans un groupe de 5 amis, trois aiment les barres au chocolat. Les deux autres n'aiment pas le chocolat. Le jour de l'Halloween, ce groupe a reçu trois paquets surprises de bonbons. Un paquet contient des dragées. Deux autres paquets contiennent chacun une barre de chocolat.

  Les trois amis de ce groupe de cinq veulent partager les deux paquets de chocolat des trois barres de chocolat.

  Solution proposée : $\frac{2}{3} \div \frac{3}{5} = \frac{2}{3} \times \frac{5}{3} = \frac{10}{9}$ *Ça ne marche pas...*

- Problème dont la solution est la multiplication $\frac{2}{3} \times \frac{3}{5}$ (4 productions)

  Exemple 2 – Léa veut faire une recette de biscuits à partager entre ses 2 sœurs et elle-même. La recette qu'elle a trouvée est pour 5 personnes, mais elle veut une quantité de biscuits raisonnable pour ses sœurs et elle, puisqu'elle déteste le gaspillage.

  Si la recette indique qu'il faut ajouter $\frac{2}{3}$ de tasse de sucre, combien de sucre doit-elle ajouter pour que la recette convienne à ses deux sœurs et elle ?

  Solution proposée : $\frac{2}{3} \div \frac{3}{5} = \frac{2}{3} \times \frac{5}{3} = \frac{10}{9}$ *de tasse de sucre*

Parmi les huit autres productions inadéquates, on retrouve notamment deux feuilles restées blanches ainsi que deux cas où la solution au problème demande de faire $\frac{2}{3} \times \frac{5}{3}$, évitant ainsi la division demandée, tout en conservant le résultat attendu.

Cinq productions adéquates, ou presque, ont aussi été recueillies.

- Contexte de groupement (2 productions)

  Exemple 3 – Justin fait un gâteau. Sa recette nécessite $\frac{3}{5}$ de tasse de farine. Justin possède $\frac{2}{3}$ de tasse de farine. Combien de gâteau peut-il faire avec la quantité de farine qu'il possède ?

  Solution proposée : $\frac{2}{3} \div \frac{3}{5} = \frac{2}{3} \times \frac{5}{3} = \frac{10}{9} = 1,\overline{1}$ *Il peut faire un gâteau complet*

- Contexte de comparaison ou de rapport (3 productions)

  Exemple 4 – Sacha a rempli les $\frac{2}{3}$ de sa piscine. Carlos a quant à lui rempli les $\frac{3}{5}$ de sa piscine. Combien de fois de plus Sacha a progressé dans le remplissage de sa piscine ?

  Solution proposée : $\frac{2}{3} \div \frac{3}{5} = \frac{2}{3} \times \frac{5}{3} = \frac{10}{9} \approx 1,11$ *Réponse*[2] *: environ 11 %*

Ces différents résultats conduisent à se demander comment il se fait que des étudiants et des étudiantes se destinant à l'enseignement des maths au secondaire, et pouvant déjà avoir une formation universitaire en mathématiques, ont tant de mal à composer un problème dont la solution est une simple division de fractions ?

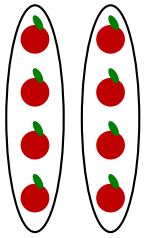
Le concept de connaissance-obstacle pourrait offrir une hypothèse de réponse. En effet, l'explication pourrait se trouver dans le fait que leur connaissance de la fraction renvoie d'emblée à un partage ; diviser c'est partager. C'est le sens attribué spontanément à la division. Cette conception de la division, du fait de son efficacité et de sa pertinence dans de nombreux cas, notamment lorsqu'on divise par un nombre naturel, est si prégnante qu'elle fait en quelque sorte obstacle à une autre conception qui conviendrait mieux à la division par une fraction, par exemple une conception de groupement.

Partage et groupement

Voici un exemple qui illustre une même division, ici $8 \div 2$, conçue comme un partage et conçue comme un groupement.

Contexte de partage

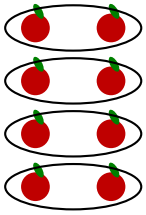
J'ai 8 pommes que je veux partager entre 2 personnes, combien de pommes aura chaque personne ?



Réponse : 4 pommes par personne

Contexte de groupement

J'ai 8 pommes et je veux faire des paquets de 2 pommes. Combien de paquets puis-je faire ?



Réponse : 4 paquets de 2 pommes

Voyons maintenant plus précisément comment se définit, en didactique, la connaissance-obstacle.

2. Remarquons que, bien que le résultat du calcul corresponde effectivement à combien de fois plus Sacha a progressé dans le remplissage, soit 1,11 fois, la réponse donne plutôt le pourcentage de plus.

2. Connaissance-obstacle, quelques caractéristiques

Le mot obstacle évoque d'emblée quelque chose à éviter ou à contourner, une entrave sur le chemin. Nous allons voir que l'obstacle, tel qu'on le conçoit en didactique, n'est pas extérieur à la connaissance et qu'il s'agit plutôt d'une connaissance qui cherche à s'adapter [3], sans y parvenir adéquatement.

Comme le disait Bachelard, à l'origine de cette notion d'obstacle :

> « Et il ne s'agit pas de considérer des obstacles externes, comme la complexité et la fugacité des phénomènes, ni d'incriminer la faiblesse des sens et de l'esprit humain : c'est dans l'acte même de connaître, intimement, qu'apparaissent, par une sorte de nécessité fonctionnelle, des lenteurs et des troubles. [...] En fait, on connaît *contre* une connaissance antérieure [...]. » (Bachelard, 1969, p. 13, c'est nous qui soulignons)

Afin de définir plus précisément la connaissance-obstacle, nous retenons les quatre critères, proposés par Duroux (1983), auxquels une telle connaissance devrait répondre.

> « 1. Il s'agit d'une **connaissance** qui fonctionne comme telle sur un ensemble de situations et pour certaines valeurs des variables de ces situations. [...]
>
> 2. L'obstacle est une connaissance qui, en tentant de s'adapter à d'autres situations ou à d'autres valeurs des variables, va provoquer des erreurs spécifiques, repérables, analysables.
>
> 3. L'obstacle est une connaissance stable. Dans les situations qui sortent de son domaine de validité, son rejet coûtera plus à l'élève qu'une tentative d'adaptation à tout prix, même si cela alourdit notablement les processus de résolution employés. C'est le symptôme classique de l'élève qui "fait compliqué" alors qu'il pourrait "faire simple".
>
> 4. L'obstacle ne pourra donc être franchi que dans des situations spécifiques de rejet et ce rejet sera constitutif du savoir. [...] Le retour même sur la conception obstacle sera partie intégrante du nouveau savoir. » (Duroux, 1983, p. 54, c'est l'auteur qui souligne)

La connaissance de la division partage, mise en évidence dans les productions précédentes, semble répondre à ces critères. En effet, (1) la conception de la division comme un partage est une connaissance qui fonctionne sur un grand ensemble de situations ; (2) des tentatives d'adaptation de cette conception à la division par une fraction apparaissent problématiques et conduisent à des erreurs, (3) erreurs que l'on retrouve d'année en année. (4) La tâche de composition d'un problème « à contexte » pour une division par une fraction s'avère porteuse

3. Précisons que ce concept d'obstacle s'inscrit dans une perspective d'apprentissage par adaptation où l'on construit des connaissances en s'adaptant à son environnement ou à des situations problématiques sources de déséquilibre.

pour le rejet de la connaissance-obstacle, comme en témoigne le commentaire *« Ça ne marche pas »* relevé dans l'exemple 1 ci-dessus et dans quelques autres. Nous verrons plus loin comment ce rejet pourrait être constitutif du savoir.

Mentionnons, enfin, que la notion de connaissance-obstacle se décline en obstacles épistémologiques (relatifs au développement historique de la notion), ontogéniques (relatifs au développement de la personne) et didactiques (relatifs aux choix didactiques faits). Nous ne les développerons pas ici. Pour en savoir davantage, voir Brousseau 1983.

3. Intérêt du concept d'obstacle pour l'enseignement des mathématiques

En quoi ce concept d'obstacle pourrait-il être utile pour l'enseignement ? Nous allons voir qu'un des intérêts de la prise en compte des connaissances-obstacles dans l'enseignement des mathématiques est de contribuer à **comprendre la source des erreurs** des élèves (3.1) et **d'offrir des pistes d'intervention en conséquence** (3.2 & 3.3).

3.1. Identifier la connaissance à la source de l'erreur

Il n'est pas rare que l'on attribue les réponses inadéquates produites par ses élèves à un manque de connaissance : l'élève n'a pas fait ce qu'il fallait, il aurait dû faire ceci, il ne sait pas cela. De sorte que l'on cherche à enseigner ce qui manque, à combler ce manque. L'idée d'obstacle invite à modifier le regard en cherchant à dégager ce qui a été utilisé et qui a conduit à un résultat inadéquat, plutôt que ce qui manque.

S'attarder à ce que l'élève sait sous-entend toutefois que l'on présume qu'il est de bonne foi, c'est-à-dire que le résultat produit constitue la meilleure réponse étant donné ses connaissances ; comme une forme de *présomption de connaissance*. On considère que l'erreur peut être le symptôme d'une connaissance qui cherche à s'adapter, peut-être maladroitement, mais sincèrement. Ainsi, comme le souligne Brousseau, devant un résultat inadéquat, on cherchera à trouver quel raisonnement a présidé à cette solution, et non ce qui manque ou qui aurait dû être fait :

> « Très souvent, c'est parmi les "difficultés" qu'il faut chercher les indices d'un obstacle mais pour satisfaire la première condition qui dit qu'un **obstacle est une connaissance**, le chercheur devra faire un effort pour reformuler la "difficulté" qu'il étudie en termes, non pas de manque de connaissance, mais de connaissance (fausse, voire incomplète…). » (Brousseau, 1983, p. 190, c'est l'auteur qui souligne).

Pour s'aider à reformuler la difficulté en termes de connaissance, il peut être utile de s'astreindre à compléter la phrase suivante : « L'élève pense que … ce qui est inadéquat ici » [4] et, ainsi, se prémunir contre le désir de s'attarder directement à ce qui manque.

4. Dans l'exemple à propos de la division $\frac{2}{3} \div \frac{3}{5}$, on pourrait dire : « L'élève pense que … la division est un partage… ce qui est inadéquat ici. »

Pourquoi est-ce si tentant de donner directement ce qui manque ?
Qui, face à une erreur de son élève, n'a jamais dit : « Non, voici, je vais te montrer comment faire ». Donner ce qui manque est tellement plus facile que de chercher et, surtout, de trouver ce qui a amené l'élève à cette production. De plus, proposer une bonne façon de faire permet à l'élève de réussir et, du même coup, l'enseignement apparaîtra réussi lui aussi. C'est donc très tentant ! Mais la réussite de l'élève témoigne-t-elle alors de sa compréhension ?
Le fait de « donner ce qui manque » est décrit par Brousseau comme une forme de court-circuit entre deux rôles de l'enseignant, soit celui de faire produire la connaissance par l'élève et celui de reconnaître l'existence de cette connaissance dans l'édifice du savoir mathématique ou, en d'autres termes, celui d'institutionnaliser.

> « On voit bien les deux parties du rôle du maître qui sont assez contradictoires : faire vivre la connaissance, la faire produire par les élèves comme réponse raisonnable à une situation familière, et de plus, transformer cette « réponse raisonnable » en « événement » cognitif extraordinaire identifié reconnu à l'extérieur.
>
> La tentation est grande pour le professeur de court-circuiter ces deux phases et d'enseigner directement le savoir en tant qu'objet culturel en faisant l'économie de cette double manœuvre. On présente le savoir et l'élève se l'approprie comme il peut. » (Brousseau 1988, p. 14)

Donner ce qui manque en enseignant directement le savoir peut ainsi apparaître économique au regard du temps et de l'énergie investis, d'où la tentation. Mais cette économie peut parfois se faire au détriment de la compréhension des élèves.

Après avoir identifié la connaissance-obstacle qui semble à l'origine de la difficulté, il s'agit d'en définir le domaine de validité. En effet, si cette connaissance existe au sein du répertoire de connaissances des élèves, c'est qu'il doit y avoir un ensemble de problèmes pour lesquels elle est efficace, sinon on peut penser qu'elle n'aurait pas été retenue. C'est d'ailleurs pourquoi je privilégie souvent le terme *inadéquate* au terme *fausse* pour qualifier la connaissance. Les connaissances dont on parle ne sont pas fondamentalement fausses, elles sont simplement utilisées à mauvais escient ou de manière inadéquate, en d'autres termes, hors de leur domaine de validité.

3.2. Déstabiliser la connaissance-obstacle

Une fois l'obstacle identifié et son domaine de validité défini, on cherchera des problèmes qui sortent de ce domaine en vue de déstabiliser la connaissance de l'élève et de lui faire rencontrer son ignorance.

En proposant un tel problème à l'élève on cherche à provoquer l'erreur dans l'espoir qu'il constate l'inadéquation de sa procédure pour ce problème. Il importe que l'élève puisse *par lui-même* constater l'invalidité de sa réponse, de sorte que le besoin de modifier sa connaissance, modification souvent très coûteuse, vienne de lui, lui devienne nécessaire pour résoudre un possible conflit cognitif.

Si l'invalidation vient du professeur, l'élève ne « rencontre » pas son ignorance, elle lui est simplement « montrée ». Tant que l'élève n'est pas convaincu qu'il ne sait pas, il n'éprouve pas le besoin d'apprendre. Il y a donc lieu d'organiser la rencontre avec son ignorance. Ainsi que le dit Mercier :

> « Comment peut-on avoir besoin de chercher ce que l'on croit déjà connaître ? Par le fait qu'une institution parvient à nous persuader qu'on ne le connaît que très mal en nous faisant rencontrer notre ignorance et en nous offrant de la combler. » (Mercier 1996, p. 360)

Pour que l'élève puisse invalider son résultat, il importe qu'il ait les moyens de cette invalidation, c'est-à-dire que les conditions entourant le problème proposé lui permettent de juger par lui-même de la validité de la solution ou de la réponse. Il n'est pas toujours facile de concevoir de tels problèmes. Souvent, en raison d'implicites que les élèves ignorent et qui sont nécessaires à la reconnaissance de la résolution attendue, les problèmes proposés aux élèves ne leur permettent pas de conclure à la validité de la réponse obtenue [5], ce qui les contraint à s'en remettre au professeur pour savoir si la réponse est valide ou non. Mais quand on arrive à créer (ou à trouver) des problèmes propices à la remise en question par les élèves de ce qu'ils croyaient pouvoir utiliser, ils en viennent à réclamer la nouvelle connaissance.

3.3. Rejeter la connaissance-obstacle, levier de l'apprentissage

Après avoir identifié une connaissance-obstacle, après l'avoir sollicitée en dehors de son domaine de validité de manière à faire rencontrer l'ignorance et à susciter le désir d'apprendre de ses élèves, reste à faire développer une connaissance adéquate en s'appuyant sur l'ancienne dont le rejet sera constitutif du savoir. Pour illustrer cela, revenons à la tâche de départ que j'utilisais dans mes cours de didactique.

Cette tâche de composition de problème a été conçue pour mettre en jeu la connaissance-obstacle *la division est un partage*, inadéquate dans le cas d'une division par une fraction, afin de déstabiliser les étudiants et étudiantes en enseignement et de leur faire rencontrer l'ignorance. Et ils la rencontrent !

En effet, comme les productions retenues l'ont montré, les étudiants et les étudiantes tentent de trouver un contexte de partage pour l'opération demandée, mais n'y arrivent pas. Leur connaissance de la division dans son sens de partage semble faire obstacle au déploiement d'une connaissance plus adéquate. Ils réalisent que ça ne fonctionne pas et le disent parfois en autant de mots, comme on l'a vu plus tôt : *Ça ne marche pas…* Leur difficulté, voire leur incapacité, à rédiger un tel problème crée un choc. Ils rencontrent leur ignorance. Ce choc de l'ignorance

5. Par exemple, le problème suivant, tiré de AlloProf, où le rythme constant nécessaire à la reconnaissance d'un modèle proportionnel reste implicite : *Simone plante des fleurs sur son terrain. Aujourd'hui, elle a travaillé pendant 4 heures et a réussi à planter les fleurs de 24 boîtes. Combien de boites de fleurs réussira-t-elle à planter si, demain, elle n'a que 3 heures à consacrer à cette tâche ?*

sert alors de levier à l'apprentissage. La limite du domaine de validité se dessine, le désir de savoir se manifeste.

Il est important de mentionner que c'est la tâche de **composer** un problème dont la solution demande une division par une fraction qui déstabilise les étudiants et les étudiantes en enseignement. S'il s'était agi de résoudre plutôt que de composer un tel problème, ils auraient réussi [6]. En effet, bien qu'ils aient sûrement déjà résolu des problèmes où la division se concevait comme un partage, un groupement, un rapport ou un autre sens, ils n'ont généralement pas eu l'occasion d'expliciter ces différents sens ni, par conséquent, senti le besoin de faire les distinctions entre ces diverses conceptions de la division et encore moins d'en définir les domaines de validité. Ce n'est que confrontés à l'impossibilité de trouver un contexte qui appelle une division par une fraction que leur besoin de savoir est né. Ce qu'ils expérimentaient de manière implicite (un partage, un groupement, un rapport ou autre) devra maintenant s'expliciter pour pouvoir être utilisé dans leur travail d'enseignement. Bien que ces différents sens ne soient habituellement pas explicités dans les cours de maths, il est essentiel pour les futurs enseignants et enseignantes de connaître ces distinctions, non pas pour les enseigner à leurs élèves, mais pour s'assurer de leur présenter une richesse de contextes qui toucheront à tant au partage qu'au groupement ainsi qu'aux autres sens, et ce, dès le travail dans les nombres naturels.

Reste à organiser une situation de classe qui amènera les futurs enseignants et enseignantes à trouver des contextes mettant en jeu une division par une fraction, puis à caractériser ces contextes en fonction du sens de la division qui est alors sollicité. Cela peut se faire par des discussions entre les étudiants et étudiantes afin de déterminer lesquelles de leurs productions conviennent ou non et pourquoi. Il s'agit alors, en quelque sorte, de faire faire par la classe le travail de validation et d'invalidation des diverses productions. Les échanges, les arguments et les conclusions servant alors à distinguer les différents sens (partage, groupement, rapport, etc.), à les caractériser en fonction de leur domaine de validité et à en définir ainsi les espaces et les limites d'usage.

Les futurs enseignants et enseignantes en identifiant explicitement d'autres sens à la division que le partage, en distinguant les domaines de validité de chacun de ces sens et en en traçant les limites, enrichissent leur connaissance de la division. Le rejet de la connaissance-obstacle *la division est un partage*, pour les problèmes de division par une fraction, devient alors constitutif de la connaissance de la division que les futurs enseignants et enseignantes développent. Ce rejet du sens de partage, pour les problèmes qui sortent de son domaine de validité, informe sur la division et peut conduire à se questionner sur d'autres sens possibles de la division selon les contextes.

En conclusion, c'est autant par la reconnaissance de ce qui fonctionne que par celle de ce qui ne fonctionne pas que la connaissance s'enrichit.

6. Un peu comme on comprend une langue sans pouvoir la parler, ou on comprend une preuve mais on n'aurait pas pu la produire.

Bibliographie

AlloProf, https://www.alloprof.qc.ca/fr/eleves/bv/mathematiques/la-resolution-de-situations-directement-ou-invers-m1429, page consultée le 28 septembre 2026.

Bachelard, G. (1938). *La formation de l'esprit scientifique*. Paris : Vrin.

Brousseau, G. (1983). Les obstacles épistémologiques et les problèmes en mathématiques. *Recherches en Didactique des Mathématiques*, 4(2), 165–198.

Brousseau, G. (1988). Les différents rôles du maître. *Bulletin de l'Association Mathématique du Québec*, 28(2), 14–24.

Duroux, A. (1983). La valeur absolue. Difficultés majeures pour une notion mineure. *petit x*, no 3, 43–67.

Mercier, A. (1996). La création d'ignorance, condition de l'apprentissage. *Revue des sciences de l'éducation*, 22(2), 345–363.

Association
mathématique
du Québec

Chronique

Problèmes et jeux mathématiques

Daniel Audet
Collège de Bois-de-Boulogne

Nous vous encourageons à envoyer vos solutions à bulletin@amq.math.ca. Les noms des personnes ayant résolu les problèmes seront publiés avec les solutions dans la prochaine édition.

Les problèmes

Problème 1 — Carrés en somme

Déterminer tous les triplets d'entiers (a,b,c) qui satisfont simultanément

$$a^2 = b + c + 5, \qquad b^2 = a + c - 1, \qquad c^2 = a + b - 1.$$

Problème 2 — Un triangle équilatéral dans une grille carrée

Considérons une grille formée de carrés identiques. Est-il possible de choisir trois points d'intersection de la grille qui soient les sommets d'un triangle équilatéral ? Si oui, donnez un exemple et si non démontrez cette impossibilité.

Problème 3 — Paires sans collision On considère des points également espacés sur la circonférence d'un cercle et on les regroupe au hasard en paires. Les deux points de chaque paire sont reliés par une corde.

1. Avec 4 points, quelle est la probabilité qu'aucune paire de cordes ne se croise ?
2. Même question avec 6 points.
3. Plus généralement, avec $2n$ points, quelle est la probabilité qu'aucune paire de cordes ne se croise ?

© Association mathématique du Québec

Solutions aux problèmes de la chronique précédente

Problème 1 — La tour d'intégrales

Résoudre

$$\int_1^{\int_1^{\int_1^{\cdots}(5-2x)\,dx}(5-2x)\,dx}(5-2x)\,dx$$

Solution.

Notons I la valeur de la tour d'intégrales, si cette valeur existe. Puisque la tour qui apparaît comme borne supérieure de la première intégrale est la tour elle-même, on doit avoir

$$I = \int_1^I (5-2x)\,dx.$$

Ainsi,

$$I = \left[5x - x^2\right]_1^I = 5I - I^2 - 4.$$

Par conséquent,

$$I^2 - 4I + 4 = 0,$$

c'est-à-dire

$$(I-2)^2 = 0.$$

La seule valeur possible est donc

$$\boxed{I = 2}.$$

On vérifie d'ailleurs immédiatement que

$$\int_1^2 (5-2x)\,dx = 2,$$

de sorte que 2 est bien une valeur de la tour.

Problème 2 — L'impossible carré dans une grille triangulaire

Considérer une grille triangulaire formée de triangles équilatéraux identiques. Autrement dit, une grille isométrique. Montrer qu'on ne peut pas former de carré dont les sommets sont aux intersections de cette grille.

Solution.

Soit $\vec{u}$ et $\vec{v}$ deux vecteurs unitaires formant un angle de 60 °. Les sommets de la grille triangulaire sont précisément les points dont les coordonnées dans la base $(\vec{u},\vec{v})$ sont entières.

Considérons la rotation R d'angle 90 °. Dans une base orthonormée, on peut prendre

$$\vec{u} = (1,0) \qquad \text{et} \qquad \vec{v} = \left(\frac{1}{2},\frac{\sqrt{3}}{2}\right).$$

On vérifie alors que

$$R(\vec{u}) = -\frac{1}{\sqrt{3}}\,\vec{u} + \frac{2}{\sqrt{3}}\,\vec{v}$$

et

$$R(\vec{v}) = -\frac{2}{\sqrt{3}}\,\vec{u} + \frac{1}{\sqrt{3}}\,\vec{v}.$$

Ainsi, dans la base $(\vec{u},\vec{v})$, la matrice de la rotation est

$$[R] = \frac{1}{\sqrt{3}}\begin{pmatrix} -1 & -2 \\ 2 & 1 \end{pmatrix}.$$

Supposons maintenant qu'un carré ait ses quatre sommets sur la grille. Le vecteur correspondant à l'un de ses côtés s'écrit alors

$$\vec{w} = p\,\vec{u} + q\,\vec{v}, \qquad p,q \in \mathbb{Z},$$

avec $(p,q) \neq (0,0)$.

Le côté suivant du carré est obtenu en faisant subir à $\vec{w}$ une rotation de 90°. Ses coordonnées dans la base $(\vec{u},\vec{v})$ sont donc

$$[R]\begin{pmatrix} p \\ q \end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix} -p-2q \\ 2p+q \end{pmatrix}.$$

Puisque les deux extrémités de ce second côté sont également des sommets de la grille, ces deux coordonnées doivent être entières.

Or p et q sont entiers et $\sqrt{3}$ est irrationnel. Pour que

$$\frac{-p-2q}{\sqrt{3}} \qquad \text{et} \qquad \frac{2p+q}{\sqrt{3}}$$

soient tous deux entiers, il faut donc que

$$-p-2q = 0 \qquad \text{et} \qquad 2p+q = 0.$$

Ce système n'a que la solution

$$p = q = 0,$$

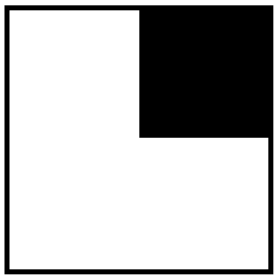
ce qui contredit le fait que $\vec{w}$ est le vecteur d'un côté du carré.

Il est donc impossible de former un carré dont les quatre sommets sont des sommets d'une grille triangulaire.

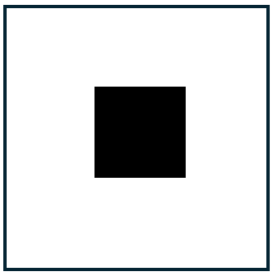
Problème 3 — Le tournant du coin

Deux points sont choisis au hasard uniformément et indépendamment à l'intérieur de la partie blanche d'une forme. Quelle est la probabilité qu'ils se voient ? C'est-à-dire, quelle est la probabilité qu'on puisse tracer un segment de droite, entièrement contenu dans la partie blanche, reliant ces deux points ?

1. Première forme



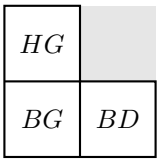
2. Deuxième forme



Solution.

1. Première forme

Divisons la région blanche en trois carrés de même aire : le carré en haut à gauche, que nous noterons HG, le carré en bas à gauche, BG, et le carré en bas à droite, BD.



Les deux points peuvent ne pas se voir seulement lorsque l'un appartient à HG et l'autre à BD. Comme les trois carrés ont la même aire et que les deux points sont choisis uniformément et indépendamment, la probabilité que cela se produise est

$$2 \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{2}{9}.$$

Supposons donc qu'un point soit choisi dans HG et l'autre dans BD. Le segment qui les relie coupe nécessairement la droite verticale passant par le coin rentrant de la région. Les deux points se voient si cette intersection se trouve sous le coin rentrant ; si elle se trouve au-dessus, le segment traverse le carré manquant.

Or, la symétrie centrale de centre le coin rentrant échange HG et BD. Elle transforme un couple de points dont le segment coupe cette droite au-dessus du coin en un couple dont le segment la coupe au-dessous du coin, et réciproquement. Ces deux possibilités ont donc la même probabilité. Le cas où le segment passe exactement par le coin étant de probabilité nulle, on en déduit que, conditionnellement au fait que les deux points appartiennent à HG et BD, la probabilité qu'ils ne se voient pas est $\frac{1}{2}$.

Ainsi,

$$P(\text{les deux points ne se voient pas}) = \frac{2}{9} \cdot \frac{1}{2} = \frac{1}{9}.$$

Finalement,

$$P(\text{les deux points se voient}) = 1 - \frac{1}{9} = \boxed{\frac{8}{9}}.$$

2. Deuxième forme

Divisons maintenant la région blanche en huit carrés de même aire. Nous noterons les quatre carrés situés dans les coins NO, NE, SO et SE, selon leur position : nord-ouest, nord-est, sud-ouest et sud-est. Les quatre autres carrés, situés sur les côtés, seront notés CN, CS, CE et CO, pour côté nord, côté sud, côté est et côté ouest.

NO	CN	NE
CO		CE
SO	CS	SE

Pour chacune des paires de régions, calculons la probabilité que les deux points se voient, conditionnellement au fait qu'ils appartiennent à ces deux régions. Par symétrie et par des

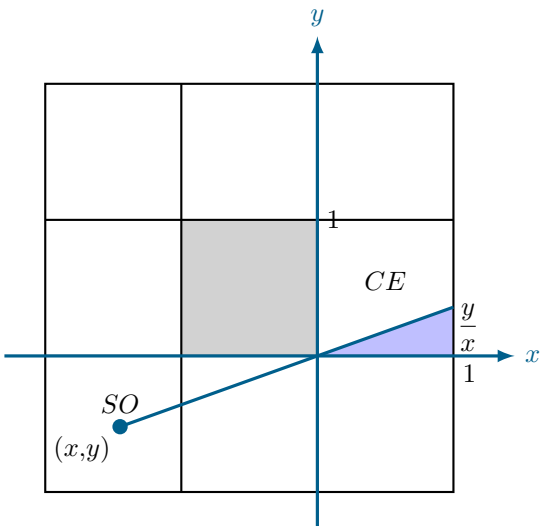
considérations géométriques simples, on obtient le tableau suivant :

	NO	NE	SO	SE	CN	CS	CE	CO
NO	1	1	1	0	1	p	p	1
NE	1	1	0	1	1	p	1	p
SO	1	0	1	1	p	1	p	1
SE	0	1	1	1	p	1	1	p
CN	1	1	p	p	1	0	$\frac{1}{2}$	$\frac{1}{2}$
CS	p	p	1	1	0	1	$\frac{1}{2}$	$\frac{1}{2}$
CE	p	1	p	1	$\frac{1}{2}$	$\frac{1}{2}$	1	0
CO	1	p	1	p	$\frac{1}{2}$	$\frac{1}{2}$	0	1

Dans ce tableau, il ne reste donc qu'à déterminer la valeur de p.

Pour calculer p, considérons par exemple les régions SO et CE. Plaçons l'origine au coin inférieur gauche de CE, l'axe des x horizontal et orienté vers la droite, et l'axe des y vertical et orienté vers le haut. Ainsi,

$$SO = [-2,-1] \times [-1,0] \qquad \text{et} \qquad CE = [0,1] \times [0,1].$$



Fixons un point $(x,y) \in SO$. La droite passant par (x,y) et par l'origine a pour équation

$$Y = \frac{y}{x} X.$$

Elle délimite, dans CE, la partie visible depuis (x,y). Comme

$$0 \leq \frac{y}{x} \leq 1,$$

cette partie est le triangle de sommets

$$(0,0), \qquad (1,0), \qquad \left(1, \frac{y}{x}\right).$$

Son aire vaut donc

$$\frac{1}{2} \cdot 1 \cdot \frac{y}{x} = \frac{y}{2x}.$$

Puisque les carrés SO et CE sont tous deux d'aire 1, la probabilité p s'obtient en faisant la moyenne de cette aire lorsque (x,y) parcourt SO. Ainsi,

$$p = \int_{-2}^{-1} \int_{-1}^{0} \frac{y}{2x}\, dy\, dx.$$

L'intégrale se sépare :

$$\begin{aligned} p &= \frac{1}{2} \left(\int_{-2}^{-1} \frac{1}{x}\, dx \right) \left(\int_{-1}^{0} y\, dy \right) \\ &= \frac{1}{2}(-\ln 2) \left(-\frac{1}{2} \right) \\ &= \boxed{\frac{\ln 2}{4}}. \end{aligned}$$

Il reste à combiner les valeurs du tableau. Celui-ci contient 32 fois la valeur 1, 8 fois la valeur 0, 8 fois la valeur 1/2 et 16 fois la valeur p. Comme les huit carrés ont la même aire, chacun des 64 couples ordonnés de régions a probabilité 1/64. La probabilité cherchée est donc

$$\frac{32 + 8\left(\frac{1}{2}\right) + 16p}{64} = \frac{9}{16} + \frac{p}{4}.$$

En remplaçant p par sa valeur, on obtient finalement

$$\boxed{\frac{9 + \ln 2}{16}} \approx 0{,}6058217.$$

Association
mathématique
du Québec

Lu pour vous

Graph Theory in America: The First Hundred Years

Frédéric Morneau-Guérin
Département Éducation, Université TÉLUQ
Frederic.Morneau-Guerin@teluq.ca

Il y a quelque chose d'un peu paradoxal dans notre rapport à l'histoire des mathématiques. Nous sommes entourés de noms propres. Nous parlons, par exemple, du théorème de Cauchy, du théorème de Pythagore, du dernier théorème de Fermat, du déterminant de Vandermonde, des graphes hamiltoniens ou eulériens. À première vue, peu de disciplines semblent conserver aussi soigneusement la mémoire de ceux et celles qui les ont façonnées. Et pourtant, comme le souligne Gary Chartrand dans la préface de *Graph Theory in America : The First Hundred Years*, nous n'avons guère l'habitude de penser les mathématiques historiquement. Le nom survit, mais le contexte disparaît.

Cela s'explique assez bien. Pour le mathématicien, ce qui importe finalement dans un résultat est sa portée, sa précision, sa véracité, ses conséquences et ses relations avec les autres résultats. Un théorème vieux de deux siècles est volontiers reformulé dans la notation d'aujourd'hui, avec le vocabulaire d'aujourd'hui et dans un cadre axiomatique qui aurait parfois été parfaitement étranger à son auteur. On lui donne une nouvelle démonstration—plus courte, plus conceptuelle, plus générale—et personne ne songe pour autant à changer son nom. Pythagore n'aurait pas reconnu toutes les dizaines de formulations contemporaines du théorème qui porte son nom (en géométrie analytique, dans les espaces euclidiens de dimension quelconque, dans les espaces préhilbertiens ou de Hilbert, ou encore sous les formes apparentées que prend l'identité de Pythagore en analyse de Fourier) ; cela ne nous trouble guère.

La chose est encore accentuée par l'ancienneté d'une bonne partie des mathématiques. Remonter à l'origine exacte d'une idée secondaire en analyse, en géométrie ou en théorie des nombres peut exiger de traverser plusieurs siècles de changements de notation, de terminologie et même de conception des objets étudiés. À force d'être réorganisées, simplifiées et systématisées, les mathématiques finissent par donner l'impression d'avoir toujours existé.

Et puis les mathématiques se veulent universelles. Un théorème n'est pas canadien, français ou allemand par nature. Sa validité ne dépend pas du lieu où il a été développé. Cela n'empêche évidemment pas que les intérêts, les traditions et les styles de recherche puissent présenter une forte coloration locale ou nationale ; pensons, par exemple, à l'école polonaise de mathématiques et à l'importance qu'y ont prise la topologie et la théorie descriptive des ensembles, dont les

 © Association mathématique du Québec

« espaces polonais » gardent jusque dans leur nom la trace historique. Les idées ont d'ailleurs toujours traversé les frontières, les langues et les cultures, parfois avec retard, certes, mais suffisamment pour que l'idée même d'écrire l'histoire d'une théorie mathématique en fonction d'un territoire ne nous vienne pas spontanément à l'esprit.

C'est précisément ce qui rend le titre *Graph Theory in America* intrigant.

Or, s'il est une branche des mathématiques dont on peut assez naturellement raconter l'histoire en l'ancrant dans un lieu, des institutions et un réseau de chercheurs, c'est bien la théorie des graphes. Elle est jeune, du moins à l'échelle des mathématiques. On peut bien sûr faire commencer l'histoire avec Euler et les ponts de Königsberg au XVIII^{e} siècle, et rechercher encore plus loin quelques préfigurations, mais l'émergence véritable d'une théorie structurée des graphes se produit essentiellement aux XIX^{e} et XX^{e} siècles. Les acteurs sont donc assez proches de nous pour que leurs lettres, leurs institutions, leurs revues, leurs collaborations et parfois leurs querelles soient encore relativement accessibles.

Elle possède en outre une physionomie singulière. Pour le dire sans suggérer le moins du monde qu'elle serait simple ou superficielle, la théorie des graphes a un faible coût d'entrée conceptuel. Il n'est pas nécessaire d'avoir traversé plusieurs couches d'abstraction avant de comprendre ce qu'est un sommet, une arête, un chemin, un cycle ou une coloration. Beaucoup de problèmes et même certains résultats peuvent être expliqués à un étudiant débutant, parfois même à un enfant : le lemme des poignées de main, le lemme des mariages, le problème des ponts de Königsberg ou encore l'affirmation selon laquelle, parmi six personnes, on en trouve toujours trois qui se connaissent mutuellement ou trois qui sont mutuellement étrangères. On peut en comprendre l'énoncé bien avant d'être en mesure d'apprécier les difficultés mathématiques qu'il recèle. La simplicité d'un énoncé peut être trompeuse : il tient parfois sur une serviette de table, alors que sa démonstration peut résister pendant un siècle.

Le problème des quatre couleurs en est évidemment l'exemple par excellence. Il en existe bien d'autres : la théorie des couplages, les problèmes de Ramsey sur les graphes, le nombre chromatique du plan. Quantité de questions extrémales ou de problèmes d'existence ont cette propriété remarquable d'être immédiatement intelligibles tout en étant mathématiquement redoutables.

Et si la théorie des graphes présente ce faible seuil d'entrée, elle déploie ses ramifications dans un nombre impressionnant de directions. Elle touche à la combinatoire, naturellement, mais aussi à la topologie, à l'algèbre linéaire et à la théorie spectrale, aux probabilités, à l'optimisation, à la recherche opérationnelle, à l'informatique théorique, à la théorie de l'information, aux réseaux électriques, à la théorie des matroïdes, à certains aspects de la théorie des représentations et à bien d'autres domaines encore. Le graphe est tantôt l'objet mathématique lui-même, tantôt un modèle, tantôt une représentation, tantôt un langage permettant de voir une structure qui restait invisible autrement.

Cette extraordinaire faculté de ramification est justement ce que les auteurs—Robin Wilson, John J. Watkins et David J. Parks—font apparaître tout au long du livre.

Les « cent premières années » du titre sont les années 1876 à 1976. Le choix des bornes est fort bien trouvé. En 1876, James Joseph Sylvester arrive à la nouvelle Université Johns Hopkins, à Baltimore. En 1976, Kenneth Appel et Wolfgang Haken annoncent leur démonstration du théorème des quatre couleurs, dont une partie substantielle repose sur des calculs informatiques. Entre ces deux dates, les mathématiques américaines ont changé de monde ; et la théorie des graphes avec elles.

Mais avant même d'arriver à Sylvester, Wilson, Watkins et Parks remontent aux débuts de l'enseignement supérieur dans les colonies britanniques. Les mathématiques y occupent d'abord une place bien modeste. Les collèges ont essentiellement une mission d'enseignement (souvent même d'enseignement mathématique relativement élémentaire) ; l'idée de l'université comme lieu de production systématique de connaissances scientifiques reste encore à construire.

L'ouvrage évoque au passage les effets de la guerre de 1812 sur les relations scientifiques américaines. Le refroidissement avec la Grande-Bretagne favorise un rapprochement avec la France, dont la tradition mathématique et scientifique accorde alors aux mathématiques fondamentales une place beaucoup plus importante que celle qu'elles occupent dans une culture britannique davantage tournée vers certaines applications pratiques. Ce sont ces petites bifurcations historiques, que l'on ne rencontre pratiquement jamais dans un cours de mathématiques, qui donnent au livre une grande partie de son intérêt.

L'idée, évoquée plus haut, de l'université comme lieu de production du savoir commence véritablement à prendre corps aux États-Unis avec la fondation de l'Université Johns Hopkins. Son premier président, Daniel Coit Gilman, souhaite bâtir aux États-Unis une université qui ne se contente pas de transmettre le savoir, mais qui en produise. Pour lancer les mathématiques, il recrute James Joseph Sylvester.

Quel personnage !

Sylvester est alors un mathématicien anglais éminent, aujourd'hui bien connu notamment des algébristes. Sa carrière n'a pourtant rien eu d'un long fleuve tranquille. Juif dans une Angleterre où les restrictions confessionnelles lui avaient fermé plusieurs portes universitaires, il avait dû emprunter des voies professionnelles inattendues, allant jusqu'à pratiquer le droit pendant une période de sa vie. Lorsqu'on vient chercher cet homme de plus de soixante ans, alors retraité, pour lui proposer de participer à l'édification d'un nouveau centre de recherche mathématique en Amérique, la perspective est manifestement assez séduisante pour lui faire traverser l'Atlantique.

Son influence sera considérable. Il forme des étudiants, crée autour de lui un milieu de recherche et contribue à la fondation de l'*American Journal of Mathematics*, qui demeure le plus ancien périodique américain de recherche mathématique encore publié.

C'est ici que le livre commence véritablement à montrer ce qu'est l'histoire concrète d'une discipline. Une théorie ne se développe pas seulement parce que quelqu'un démontre un bon théorème. Il faut des étudiants, des professeurs, des revues, des postes, des correspondants, des lieux où présenter les résultats, des réseaux permettant aux idées de circuler. Il faut aussi, parfois, quelques fortes personnalités décidées à faire exister ces structures alors qu'elles n'existent pas encore.

Les auteurs s'intéressent beaucoup à ces personnalités, et c'est un des traits distinctifs de leur ouvrage. Ils n'hésitent pas à s'éloigner pendant quelques pages d'un problème mathématique pour raconter une carrière, une amitié, une brouille ou une correspondance lorsque celle-ci aide à comprendre comment les idées ont circulé.

Ainsi en est-il de la longue relation entre Sylvester et Arthur Cayley. Les deux hommes s'étaient connus pendant une période creuse de la carrière de Sylvester, alors que celui-ci s'était temporairement tourné vers le droit. Leur correspondance devient ensuite un remarquable vecteur d'échanges entre l'Angleterre et les États-Unis.

Ainsi, encore, de la dispute entre Sylvester et William Story, peu avant le retour définitif de Sylvester en Angleterre. Les tensions autour de l'*American Journal of Mathematics* contribuent à isoler Story, mais participent aussi indirectement à déplacer des personnes et à disséminer des idées. L'histoire des mathématiques a parfois ses effets de ricochet.

Les auteurs abordent également des éléments humains moins reluisants, lorsque ceux-ci contribuent au portrait historique. L'antisémitisme de George D. Birkhoff est notamment évoqué, de même que les conséquences que certains comportements ont pu avoir sur le milieu universitaire américain. On retrouve ici une distinction importante : si les résultats mathématiques aspirent à l'universalité, les milieux où ils prennent forme, circulent et se développent restent profondément tributaires de leur époque, des institutions et des individus qui les portent.

Sylvester est aussi à l'origine d'un épisode particulièrement révélateur de la nature de la théorie des graphes. Il s'intéresse aux diagrammes utilisés par les chimistes pour représenter les molécules et croit y reconnaître des liens avec certaines structures algébriques qu'il étudie. La parenté s'avérera en partie plus notationnelle que conceptuelle, mais l'intuition est intéressante : un même dessin, un même réseau de points et de relations, peut servir de langage à des objets d'origines complètement différentes. Cette capacité des graphes à devenir un pont entre des disciplines sera d'ailleurs appelée à se répéter maintes et maintes fois.

Mais le véritable personnage récurrent du livre n'est pas un mathématicien. C'est un problème : celui des quatre couleurs. Le choix n'a rien d'étonnant puisque Robin Wilson lui-même est l'auteur de l'excellent *Four Colours Suffice*, consacré à l'histoire de ce problème. Dans le présent ouvrage, celui-ci revient continuellement, disparaît quelques pages, puis ressurgit sous une autre forme, avec une autre génération de chercheurs. Il fournit ainsi au livre une sorte de colonne vertébrale.

L'un des épisodes les plus connus est celui d'Alfred Kempe. En 1879, celui-ci publie dans l'*American Journal of Mathematics* ce qu'il croit être une démonstration du théorème des quatre couleurs. Pendant plus d'une décennie, la question semble réglée. Puis Percy Heawood découvre une faille.

Mais une mauvaise démonstration peut faire de très bonnes mathématiques. Certaines des idées introduites par Kempe survivent parfaitement à l'effondrement de sa preuve. Les chaînes qui portent aujourd'hui son nom continuent d'occuper une place importante dans l'étude des colorations. L'épisode illustre admirablement la différence entre un raisonnement faux et un raisonnement stérile : les deux notions sont loin d'être synonymes.

Le choix de Kempe de publier sa démonstration en Amérique eut lui aussi des conséquences. Le problème attire l'attention de chercheurs américains, et l'on voit successivement intervenir Paul Wernicke, Oswald Veblen, George Birkhoff, Philip Franklin et plusieurs autres.

Il est assez amusant de constater, avec le recul, combien la théorie des graphes a longtemps souffert d'un problème de respectabilité. Pour certains mathématiciens européens, ces histoires de cartes à colorier et de réseaux d'arêtes avaient quelque chose de récréatif, sinon de suspect. On trouvait les problèmes trop élémentaires dans leur formulation, trop artificiels ou insuffisamment rattachés aux grandes structures alors jugées nobles. On ira même jusqu'à qualifier les théorie des graphes de « *slum of topology* ».

Il est difficile de ne pas sourire en lisant cela aujourd'hui. Ce prétendu taudis allait devenir l'un des quartiers les plus animés des mathématiques modernes.

Et le livre montre précisément comment cela s'est produit. Le problème des quatre couleurs demeure constamment présent, mais autour de lui apparaissent les réseaux électriques de Kirchhoff, les arbres, les flots, l'optimisation combinatoire, les travaux de Ford et Fulkerson, les idées de Claude Shannon en théorie de l'information, les problèmes de transport et de plus courts chemins. Peu à peu, les graphes cessent d'être seulement associés à quelques problèmes divertissants et deviennent le langage naturel de phénomènes scientifiques, industriels et technologiques de première importance.

Le décor institutionnel change lui aussi. Les universités ne sont plus les seuls lieux où se développent des mathématiques sophistiquées. Bell Labs, RAND, General Electric et d'autres centres de recherche apparaissent dans le récit. Les besoins en télécommunications, en planification, en logistique, en optimisation et, plus tard, en informatique font des mathématiques discrètes un outil de plus en plus central.

Les deux guerres mondiales jouent évidemment leur rôle. L'interaction entre les mathématiciens universitaires et la société s'intensifie dans le cadre des efforts de guerre, puis ne revient jamais complètement à son état antérieur. Avec l'apparition de l'ordinateur, un nouveau changement de régime se produit. Les graphes sont particulièrement bien adaptés à ce nouvel univers fait d'algorithmes, de réseaux et d'objets discrets.

Il n'est donc pas anodin que le livre s'arrête en 1976. La démonstration d'Appel et Haken du théorème des quatre couleurs boucle magnifiquement le récit. Le problème qui avait accompagné l'essor de la théorie des graphes pendant près d'un siècle trouve enfin sa solution au moment même où la notion de preuve mathématique commence à être transformée par l'intervention de l'ordinateur. Une histoire s'achève et une autre commence.ç

Le mot *America* du titre appelle par ailleurs une petite précision géographique. Il faut officiellement l'entendre comme *North America*, et le Canada n'est pas entièrement absent. William Tutte, Anglais d'origine installé au Canada, y occupe naturellement une place importante, et l'Université de Waterloo apparaît dans la partie plus récente du récit. Mais l'Amérique de ce livre est, dans une très large mesure, celle des États-Unis. Ce n'est pas tant une thèse qu'un reflet du centre de gravité choisi par l'ouvrage : Baltimore, Harvard, Princeton, Chicago et les réseaux qui se constituent autour de ces institutions occupent l'essentiel de la scène.

Cette scène n'est toutefois jamais isolée de l'Europe. Deux intermèdes substantiels interrompent volontairement le récit américain pour aller voir ce qui se passe de l'autre côté de l'Atlantique. Guthrie, Tait, Heawood, Petersen, Heffter, Tietze et Minkowski apparaissent dans le premier ; König, Errera, Sainte-Laguë, Menger et Kuratowski, entre autres, dans le second. Cette structure est particulièrement efficace. Elle évite de donner l'illusion d'une théorie des graphes américaine ayant grandi en vase clos : les idées voyagent continuellement, sont importées, transformées, réexportées.

Au fil de ces quelque cent années, le lecteur fait ainsi connaissance avec plusieurs dizaines de mathématiciens aujourd'hui moins connus et retrouve sous un jour nouveau une vingtaine de noms que l'on rencontre régulièrement dans les cours et les livres. Benjamin Peirce et son fils Charles Sanders Peirce, Sylvester, Veblen, Birkhoff, Whitney, Shannon, Tutte et plusieurs autres cessent momentanément d'être des noms attachés à un théorème ou à une définition pour redevenir des personnes placées dans un réseau de collègues, d'étudiants et d'institutions.

On prend surtout conscience du caractère encore très petit de ce réseau à la fin du XIX^{e} siècle. Quelques personnes pouvaient avoir une influence immense simplement parce que les chercheurs étaient peu nombreux et que les institutions étaient encore en train de se constituer. La force de caractère, la vision, les amitiés et les inimitiés comptaient alors énormément.

Les femmes sont naturellement beaucoup moins nombreuses dans cette histoire, à l'image du monde universitaire de la période considérée. La présence de Sister Mary Petronia Van Straten, religieuse et mathématicienne qui travailla notamment sur le problème des quatre couleurs, n'en est que plus remarquable. Elle rappelle, à sa manière, que l'histoire d'une théorie se confond aussi avec celle des trajectoires individuelles rendues possibles, infléchies ou contrariées par les institutions de leur époque ; comme on l'avait déjà vu, sous une autre forme, avec Sylvester et les obstacles rencontrés en Angleterre en raison de sa judaïté.

La présentation du livre mérite enfin qu'on s'y arrête. Plusieurs extraits d'articles ayant joué un rôle important sont reproduits dans des encadrés gris, accompagnés d'explications et replacés dans leur contexte. C'est une excellente idée. Le lecteur ne se contente pas d'apprendre que tel article fut important : il peut en voir un fragment et mesurer la distance entre le langage de l'époque et celui que nous utiliserions aujourd'hui.

L'ouvrage réussit également à fonctionner à deux niveaux. Il peut être lu par quelqu'un dont l'intérêt est avant tout historique et qui ne connaît que peu la théorie des graphes. Les explications sont suffisamment accessibles pour que le fil général ne soit jamais perdu. Mais le mathématicien professionnel n'est pas pour autant condamné à rester à la surface : les auteurs fournissent assez de détails sur les problèmes, les méthodes et les filiations intellectuelles pour que le contenu mathématique conserve toute sa saveur.

L'une des grandes réussites de *Graph Theory in America* tient précisément à cela : le livre ne raconte pas seulement quels théorèmes ont été démontrés ; il montre comment une communauté mathématique se constitue.

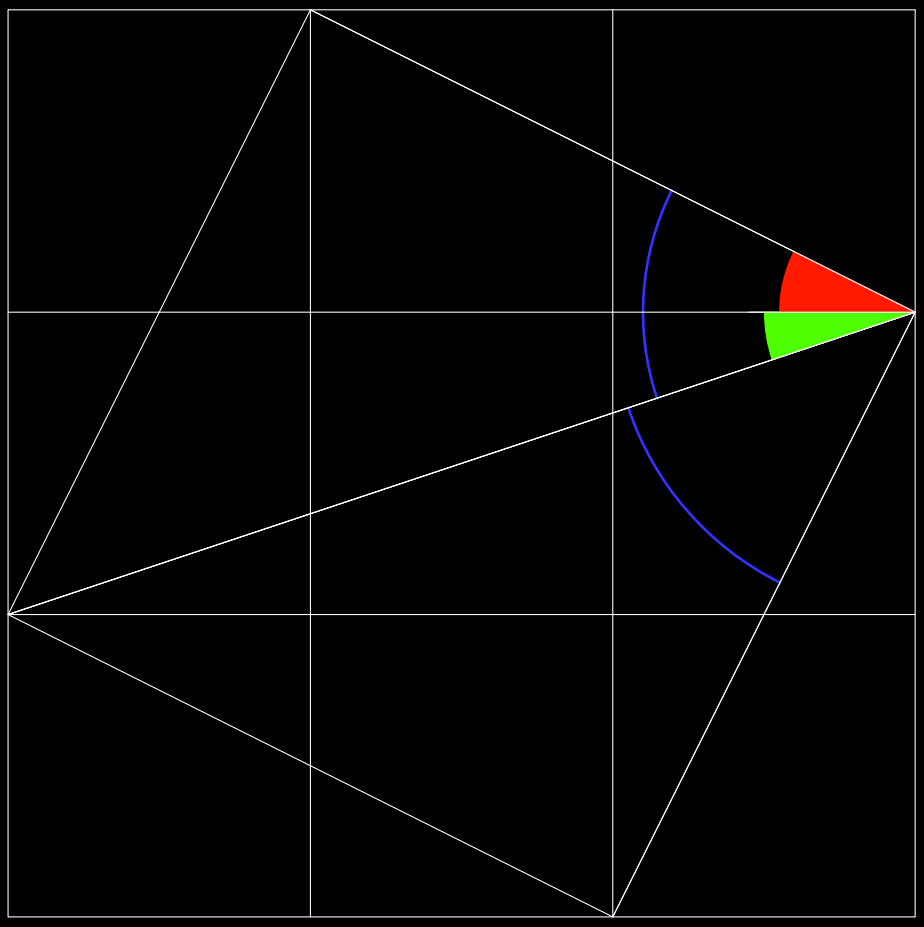
En l'espace d'un siècle, on voit ainsi se constituer tout un écosystème : les universités se donnent une véritable mission de recherche, les revues et les réseaux de chercheurs favorisent la

circulation des idées, les applications industrielles et militaires ouvrent de nouvelles directions, puis l'informatique transforme à son tour les méthodes et les problèmes. Au milieu de ces mutations, certaines questions traversent les générations, se transforment, engendrent de nouvelles techniques et finissent parfois par trouver leur solution dans un contexte mathématique devenu méconnaissable par rapport à celui qui les avait vues naître. Le problème des quatre couleurs en fournit ici l'exemple emblématique.

Nous pouvons parfaitement enseigner le théorème des quatre couleurs sans savoir tout cela.

Mais après avoir lu Wilson, Watkins et Parks, il est difficile de prétendre que savoir tout cela n'ajoute rien. La vérité mathématique est intemporelle, du moins dans l'idéal que nous nous en faisons. La vie des mathématiques ne l'est pas. Elle dépend de lieux, de revues, de voyages, de lettres, de rencontres, de préjugés, de querelles, de guerres, de technologies et, surtout, de personnes. Il fallait peut-être trois spécialistes des graphes pour nous rappeler que l'histoire des mathématiques est elle-même une histoire de réseaux.

$$\arctan\left(\frac{1}{2}\right) + \arctan\left(\frac{1}{3}\right) = \arctan(1)$$



Reproduction d'une preuve sans mots de Christiane Rousseau, sous licence *Creative Commons Attribution 4.0 International* 2020.