



**HAL**  
open science

## Nous sommes toutes distinctes

Stéphane Flon, Karim Zayana

► **To cite this version:**

Stéphane Flon, Karim Zayana. Nous sommes toutes distinctes. *Quadrature*, 2024, 134 (134), pp.17.  
hal-04776601

**HAL Id: hal-04776601**

**<https://hal.science/hal-04776601v1>**

Submitted on 11 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Nous sommes toutes distinctes

Stéphane FLON, Karim ZAYANA

Le programme de la spécialité du cycle terminal [1, 2] comporte un volet consacré aux notations mathématiques, à la théorie des ensemble et à la logique. Cette initiation se veut progressive et congruente aux chapitres habituellement traités en classe. Il peut être utile de l'accompagner de quelques problèmes aux contextes variés. Ludique et élégante, l'étude des « ensembles aux sommes toutes distinctes » nous en offre un bon exemple. Voilà un thème qui, d'un abord *a priori* formel, jouera heureusement avec les nombres, les inégalités, la combinatoire et, par un habile détour, les probabilités. Inspiré d'un énoncé national d'Olympiades de première [3], tiré lui-même de publications scientifiques [4, 5, 6], nous en structurerons l'exposé de manière à lui rendre un caractère plus linéaire.

Dans tout ce texte, le symbole  $n$  désignera un entier naturel ; tous les ensembles étudiés seront finis et constitués de nombres réels (même quand on omettra de le préciser) ; de plus on conviendra d'écrire et d'indexer un ensemble fini  $A = \{a_1; \dots; a_n\}$  à  $n$  éléments en ordonnant toujours  $a_1, a_2, \dots, a_n$  de sorte que  $a_1 < a_2 < \dots < a_n$ . Étant donné un tel ensemble, on notera  $S(A)$  la somme de ses éléments, soit

$$S(A) = a_1 + \dots + a_n.$$

Enfin, on attribuera à la somme vide, c'est-à-dire réduite à un nombre nul d'éléments, la valeur nulle.

Définissons et illustrons donc sans plus tarder ce que nous entendrons par un « ensemble aux sommes toutes distinctes ».

## 1 Définition. Exemples et tests.

### Ensemble aux sommes toutes distinctes

**Définition.** *Un sous-ensemble  $A$  de  $\mathbb{R}$  à  $n$  éléments est à « sommes toutes distinctes » (en abrégé :  $A$  est STD) quand, pour toutes parties distinctes  $Y$  et  $Z$  de  $A$ ,  $S(Y) \neq S(Z)$ . Ceci revient à demander aux  $2^n$  sommes que l'on peut former avec des éléments de  $A$  d'être toutes distinctes.*

Pour  $n = 0$  cette définition confère à  $A = \emptyset$  le statut de STD. Plus intéressant,  $A = \{1;2;5\}$  est STD parce que les 8 nombres que sont 0 (résultant de la somme vide), 1, 2, 5 (résultant des singletons),  $1 + 2 = 3$ ,  $2 + 5 = 7$ ,  $1 + 5 = 6$  (résultant des paires),  $1 + 2 + 5 = 8$  (résultant du tout) sont tous distincts. En revanche,  $A = \{0;1;2;5\}$  n'est pas STD car la somme vide est aussi celle que renvoie le singleton  $\{0\}$ . L'ensemble  $A = \{4;6;7;8;9\}$  n'est pas non plus STD au motif que ses deux parties  $Y = \{4;8;9\}$  et  $Z = \{6;7;8\}$ , bien que distinctes, fournissent la même somme  $S = 21$ . L'élément 8, qui fait doublon, eût parfaitement pu être ôté. De façon générale, il suffirait d'envisager les parties  $Y$  et  $Z$  de  $A$  qui sont disjointes (et distinctes, afin d'éviter le cas très exceptionnel où  $Y$  et  $Z$  seraient simultanément vides). Ainsi y a-t-il souvent moyen de s'économiser pour déterminer la nature STD d'un ensemble  $A$  donné. Du reste, on établirait sans peine les propriétés suivantes, elles aussi sources de simplifications :

- Proposition 1.** (a) *Que  $A$  contienne l'élément nul et il n'est pas STD;*  
 (b) *Qu'un seul sous-ensemble  $B$  de  $A$  ne soit pas STD, et  $A$  n'est pas STD. Autrement dit, tout sur-ensemble  $A$  d'un ensemble  $B$  qui n'est pas STD ne peut l'être lui-même;*  
 (c) *Que  $A$  soit STD, et tous ses sous-ensembles le sont aussi.*

Fort de ces propriétés, voici un algorithme (non optimal) prenant en entrée un ensemble  $A$  quelconque et retournant en sortie un booléen de valeur « vraie » ou « fausse » selon que  $A$  est STD ou non. Il faudra parcourir  $A$ , étendre l'éventail des sommes connues formées d'éléments de  $A$  en les complétant du dernier élément  $a$  rencontré, vérifier qu'il n'y a pas de recoupement avec les anciennes, actualiser la mémoire ou bien sortir de la boucle. Tant qu'elle est réalisée, chaque itération sur  $A$  double la mise. Jusqu'au total déjà évoqué de  $2^{\text{card}(A)}$  combinaisons, toutes les sommes possibles auront donc été balayées quand  $A$  est STD. Sinon l'algorithme s'arrête avant.

Le langage Python permet d'instancier des ensembles en tant que structures de données. Un script correspondant serait alors :

```

1 def testSTD(A):
2     sommesConnues = {0}
3     ensembleVide = set()
4     for a in A:
5         sommesNouvelles = {s+a for s in sommesConnues}
6         if sommesConnues & sommesNouvelles != ensembleVide:
7             return False

```

```

8         else :
9             sommesConnues = sommesConnues | sommesNouvelles
10        return True

```

*Quelques commentaires.* En ligne 2 l'ensemble `sommesConnues` est initialisé à  $\{0\}$  : il résulte de la somme vide. En ligne 5, `sommesConnues` est translaté en bloc de la valeur de `a` ; ce nouvel ensemble reste composé de termes distincts dans l'hypothèse où l'ancien le demeurerait. Le cas échéant on les croise, et selon que l'intersection est non vide ou pas, on échappe de la boucle (ligne 10) ou on réalise une union (lignes 13, 14).

Pointons ici les limites de l'ordinateur qui, après avoir exécuté ce code sur l'ensemble  $\{1/10; 2/10; 3/10\}$ , le croit STD. C'est une illusion. La machine représente en effet l'addition  $1/10 + 2/10$  par le nombre  $0,300 \dots 04$  qu'elle distingue par conséquent de  $3/10$ . Méfiance, donc !

À la lecture de notre programme, on mesure combien enrichir  $A$  petit à petit d'éléments « radicalement différents » des précédents préserve à l'ensemble ses chances d'être STD. Évitions tout formalisme excessif ; revenons plutôt au cas de  $A = \{1; 2; 5\}$ , qui était STD. Adjoignons-lui maintenant le décimal  $1,5$  pour former  $A' = \{1; 1,5; 2; 5\}$ . Aucune somme impliquant  $1,5$  ne peut valoir une somme ne l'impliquant pas, sans quoi  $1,5$  serait la différence de deux entiers, donc entier. Il y a contradiction. Si bien que  $A' = \{1; 1,5; 2; 5\}$  est STD. Adjoignons-lui ensuite le rationnel  $1/3$  pour former  $A'' = \{1/3; 1; 1,5; 2; 5\}$ . La fraction  $1/3$  n'est pas décimale, donc par des arguments similaires,  $A''$  est STD. En filant ce raisonnement, on ajouterait encore l'irrationnel  $\sqrt{2}$  puis le nombre transcendant  $\pi$  pour former par exemple le STD  $A^{(4)} = \{1/3; 1; \sqrt{2}; 1,5; 2; \pi; 5\}$ .

Une autre façon d'introduire un nouvel élément  $a'$  dans une partie  $A$  connue pour STD, en assurant  $a'$  de s'affranchir de la portée de  $A$ , consiste à le choisir supérieur aux sommes et différences déjà considérées, et donc à la plus grande d'entre elles. Sur les calculs effectués en ligne 5, cela a pour effet de décaler loin à droite les `sommesNouvelles` des `sommesConnues`, éliminant ainsi tout risque de collision. Concrètement, on partirait de  $A = A^{(1)} = a_1$  réduit à l'élément  $a_1 \neq 0$ , on insérerait  $a_2$  avec  $|a_2| > |a_1|$ , pour former  $A^{(2)} = \{a_1; a_2\}$ , etc. Jusqu'à  $A^{(n)} = \{a_1; \dots; a_n\}$  avec  $|a_n| > |a_1| + \dots + |a_{n-1}|$ . Afin de respecter la convention d'écriture des éléments de  $A^{(n)}$ , qui doivent se présenter dans l'ordre croissant, on les prendra volontiers tous positifs. Soit  $a_1 > 0$ ,  $a_2 > a_1, \dots, a_n > a_1 + \dots + a_{n-1}$ . De proche en proche, ceci construit un ensemble STD. On s'en convainc en comparant directement les deux sommes génériques

aux supports disjoints

$$S = a_n + \sum_{i \in I} a_i \quad \text{et} \quad S' = \sum_{j \in J} a_j$$

avec  $I, J \subset \llbracket 1; n-1 \rrbracket$  et  $I \cap J = \emptyset$ . De fait,

$$a_n > \sum_{k \in \llbracket 1; n-1 \rrbracket} a_k \geq \sum_{i \in I} a_i + \sum_{j \in J} a_j \geq - \sum_{i \in I} a_i + \sum_{j \in J} a_j$$

et donc  $S > S'$ .

Limitons-nous à des éléments entiers naturels. En ajustant au plus près les inégalités dictées plus haut, nous obtenons en l'espèce l'ensemble  $A^{(n)}$  composé de  $a_1 = 1$ ,  $a_2 = 2$ ,  $a_3 = a_1 + a_2 + 1 = 4$  et, par ce procédé récurrent,  $a_n = 2^{n-1}$ .

## 2 Localisation des éléments d'un STD constitué d'entiers positifs

Concentrons-nous dans ce second paragraphe sur les seuls ensembles constitués d'entiers naturels. Dans le dernier exemple, nous avons alors constaté que ces entiers étaient assez espacés. Il y a ici matière à généralisation comme le précise la proposition qui suit.

**Proposition 2.** *Si  $A$  est un STD fait de  $n \geq 2$  entiers (strictement) positifs  $a_1 < a_2 < \dots < a_n$  alors, pour tout indice  $k$ ,  $2 \leq k \leq n$ ,*

$$a_k \geq \frac{2^k}{k}.$$

*Démonstration.* Quitte à emboîter les STD successifs  $\{a_1; a_2\}, \dots, \{a_1; a_2; \dots; a_n\}$  les uns dans les autres, il suffit de prouver le résultat voulu pour  $k = n$ . Les  $n$  éléments  $a_1, a_2, \dots, a_n$  engendrent  $2^n$  sommes toutes distinctes, toutes entières, la plus petite valant 0 et la plus grande  $S(A) = a_1 + a_2 + \dots + a_n$ . L'intervalle qui les loge doit donc être assez large, de sorte que

$$a_1 + a_2 + \dots + a_n \geq 2^n - 1.$$

Comme  $a_n$  est le plus grand (strictement) des éléments de  $A$  et que  $S(A)$  cumule  $n$  termes, l'inégalité large précédente entraîne l'inégalité stricte suivante,

$$na_n > 2^n - 1,$$

laquelle n'implique que des membres entiers et revient à une inégalité large modulo l'ajout d'un pas unité à sa droite.  $\square$

Nous allons désormais améliorer cette série d'inégalités.

**Proposition 3.** *Sous les mêmes hypothèses qu'en proposition 2, pour tout indice  $k$ ,  $2 \leq k \leq n$ ,*

$$a_k \geq \frac{2^k}{\sqrt{3k}}.$$

Cette série d'inégalités affine bien la précédente car, dès le rang 3, les mineurants nouvellement proposés sont supérieurs aux anciens. Une preuve particulièrement originale utilise la théorie des probabilités et s'avère abordable dès le lycée, ce qui en renforce l'intérêt. La voici :

*Démonstration.* Comme en proposition 2, attachons-nous au dernier indice, à savoir  $k = n$ . Considérons  $n$  variables aléatoires de Rademacher<sup>1</sup> indépendantes  $X_1, X_2, \dots, X_n$  et posons  $X$  leur somme ainsi pondérée :  $X = a_1X_1 + a_2X_2 + \dots + a_nX_n$ . Puisque  $E(X_1) = 0$  et  $V(X_1) = 1$ , espérance et variance de  $X$  valent respectivement

$$E(X) = a_1E(X_1) + \dots + a_nE(X_n) = 0$$

et

$$V(X) = a_1^2V(X_1) + \dots + a_n^2V(X_n) = a_1^2 + \dots + a_n^2.$$

La variable  $X$  prend ses valeurs parmi toutes les sommes et différences des éléments  $a_1, \dots, a_n$ , qui sont de la forme  $a_1\epsilon_1 + \dots + a_n\epsilon_n$ . où  $\epsilon_1 = \pm 1, \dots, \epsilon_n = \pm 1$ . Ces  $2^n$  sommes pondérées sont toutes distinctes. En effet, que pour deux jeux de  $n$ -uplets  $\omega = (\epsilon_k)_k$  et  $\omega' = (\epsilon'_k)_k$  de  $\{-1; 1\}^n$  on ait

$$\sum_k a_k\epsilon_k = \sum_k a_k\epsilon'_k$$

et il viendra, en sommant des deux côtés à  $a_1 + a_2 + \dots + a_n$ ,

$$2 \sum_{i, \epsilon_i=1} a_i\epsilon_i = 2 \sum_{j, \epsilon'_j=1} a_j\epsilon'_j.$$

L'ensemble  $A$  de travail étant STD, les supports  $I = \{i, \epsilon_i = 1\}$  et  $J = \{j, \epsilon'_j = 1\}$  se confondent. Les jeux  $(\epsilon_k)_k$  et  $(\epsilon'_k)_k$  aussi. Par conséquent, la variable  $X$

1. Probabilité 1/2 de valoir 1, et 1/2 de valoir -1.

suit une loi uniforme sur l'ensemble des  $2^n$  entiers relatifs  $\pm a_1 \pm \dots \pm a_n$ . Ces entiers, qui forment l'image de  $X$ , se répartissent symétriquement par rapport à l'origine (on retrouve que la loi est centrée), sont tous de la même parité (la différence de deux valeurs images étant toujours paire), et sont non nuls (par l'absurde, via le caractère STD de  $A$ ). Ceci va nous permettre d'exprimer la variance de  $X$  d'une deuxième manière. En transférant à  $X^2$  la distribution des  $n$ -uplets  $\omega$ ,  $V(X) = E((X - E(X))^2) = E(X^2)$  s'écrit aussi

$$V(X) = \sum_{\omega \in \{-1;1\}^n} (a_1 \epsilon_1 + \dots + a_n \epsilon_n)^2 \frac{1}{2^n}.$$

Une fois ordonnés, des nombres  $|a_1 \epsilon_1 + \dots + a_n \epsilon_n|$  (dont une moitié d'entre eux, sans les valeurs absolues, seraient strictement positifs) occupent une première position au moins égale à 1, une deuxième au moins égale à 3, ..., une dernière au moins égale à  $2 \times 2^{n-1} - 1 = 2^n - 1$ . Par suite,

$$V(X) \geq 2 \times \frac{1}{2^n} \sum_{\ell=1}^{2^{n-1}} (2\ell - 1)^2.$$

Après avoir reconnu une somme classique,

$$V(X) \geq \frac{4^n - 1}{3}.$$

Or la variance de  $X$ , déjà calculée plus tôt, est entière et strictement majorée par l'entier  $na_n^2$ . Enchaîner nos inégalités conduit à

$$na_n^2 \geq \frac{4^n - 1}{3} + 1,$$

et de  $a_n \geq 0$  la conclusion tombe. Notons qu'elle reste valable pour  $n = 1$ .  $\square$

Stéphane FLON est professeur de spéciales au lycée Henri IV (Paris) et Karim ZAYANA est inspecteur général de l'éducation, du sport et de la recherche et chercheur associé au laboratoire ÉRCAÉ (université d'Orléans).

## Remerciements

Les auteurs remercient la rédaction du magazine Quadrature ainsi qu'Anne-Laure BIOLLEY, Baptiste COLLIN-BAUX et Mathieu RICHARD, tous trois membres du jury national des Olympiades de première (session 2024), pour leurs échanges autour de ce texte.

## Références

- [1] « Programme d'enseignement de spécialité de mathématiques de la classe de première de la voie générale ». In : *Bulletin officiel spécial n°1 du 22 janvier 2019* (2019). URL : <https://www.education.gouv.fr/bo/19/Special1/MENE1901632A.htm>.
- [2] « Programme de l'enseignement de spécialité de mathématiques de la classe de terminale de la voie générale ». In : *Bulletin officiel spécial n°8 du 25 juillet 2019* (2019). URL : <https://www.education.gouv.fr/bo/19/Special8/MENE1921246A.htm>.
- [3] « Olympiades nationales de mathématiques ». In : *sujet national Métropole – Europe – Afrique – Orient* (2024). URL : [https://euler.ac-versailles.fr/IMG/pdf/national\\_spe\\_vf.pdf](https://euler.ac-versailles.fr/IMG/pdf/national_spe_vf.pdf).
- [4] Paul ERDÖS. « Problems and results in additive number theory ». In : *Actes du Colloque sur la Théorie des Nombres, Bruxelles* (1955).
- [5] John Horton CONWAY et Richard Kenneth GUY. « Sets of natural numbers with distinct sums ». In : *Notices Of The American Mathematical Society* 15 (1968).
- [6] Jaegug BAE. « On subset-sum-distinct sequences ». In : *Progress in Mathematics, Analytic number theory, Proceedings of a Conference In Honor of Heini Halberstam* 138 (1996).