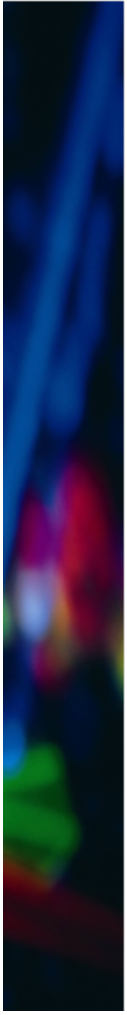
Problèmes de suites à somme nulle sur les groupes abéliens finis : une approche explicite

Hanane Zerdoum



Université Paris 8 Vincennes - Saint-Denis

École doctorale: Cognition, langage, interaction (ED 224)

Laboratoire Analyse, Géométrie et applications (UMR 7539)

Problèmes de suites à somme nulle sur les groupes abéliens finis : une approche explicite

Par Hanane Zerdoum

Thèse de doctorat en Mathématiques

Sous la direction de Wolfgang Schmid et Philippe Guillot

présentée et soutenue publiquement le 06 juillet 2021

Devant un jury composé de:

Mme Gautami Bhowmik,	Maître de conférences, Université de Lille	Pré-rapporteur
M. Philippe Guillot,	Maître de conférence, Université de Paris 8	Co-directeur
M. François Hennecart ,	Professeur, Université de Jean-Monnet	Pré-rapporteur
M. Nicolas Jouandeau,	Maître de conférence, Université de Paris 8	Examinateur
Mme Anne de Roton,	Maître de conférences, Université de Lorraine	Examinateur
M. Wolfgang Schmid,	Professeur, Université de Paris 8	Directeur de thèse
Mme Isis Truck,	Professeur, Université de Paris 8	Examinateur
M. Gilles Zémor ,	Professeur, Université de Bordeaux	Examinateur

THESE

REMERCIEMENTS

Tout d'abord, je tiens à exprimer toute ma gratitude à mon directeur de thèse, Wolfgang Schmid, qui m'a proposé ce sujet de recherche passionnant. Qu'il soit aussi remercié pour avoir été toujours présent pour guider mon cheminement malgré ses nombreuses charges. Grâce à son écoute, compétences, confiance rigueur scientifique et gentillesse, j'ai beaucoup appris à son contact. Je suis heureuse d'avoir travaillé en sa compagnie.

Je suis extrêmement reconnaissante à Philippe Guillot, pour son engagement dans le co-encadrement de ma thèse. Je le remercie pour tout le temps qu'il m'a consacré. Son talent de pédagogue, ses suggestions éclairées et critiques honnêtes, son enthousiasme et ses encouragement ont été pour moi essentiels.

Merci également aux membres du jury d'avoir accepté de s'intéresser à ces travaux.

J'adresse toute ma gratitude à toutes les personnes qui m'ont aidé dans la réalisation de ce travail. Chercheurs et enseignants du laboratoire LAGA. Ainsi que mes collègues thésards.

Je voudrais aussi remercier les responsables de l'école doctorale CLI de m'avoir permis de travailler dans des bonnes conditions et aussi toutes les personnes que j'ai rencontrée par leur biais.

Enfin, je ne pourrais finir ces remerciements sans penser à ma famille, en particulier mes chers parents et époux, dont l'affection, l'amour, le soutien et l'encouragement constants m'ont été d'un grand réconfort et ont contribué à l'aboutissement de ce travail.

TABLE DES MATIÈRES

CHAPITRE I

INTRODUCTION ET MOTIVATIONS

L'objet de cette thèse est l'étude des problèmes de type somme nulle sur des groupes abéliens finis.

Définir ces problèmes se résume à déterminer pour un groupe abélien fini G noté additivement, le plus petit entier k tel que toute suite S de k éléments de G contient une sous-suite T dont la somme des termes est nulle, c'est-à-dire la somme de tous les éléments de T vaut 0.

Il est possible de varier ce problème en rajoutant à la définition des restrictions supplémentaires sur la sous-suite T, comme par exemple sur sa longueur. En rajoutant cette restriction nous obtenons la définition de la constante de Davenport qui sera abordée dans la suite du document.

Les questions analogues sont aussi étudiées pour les ensembles, autrement dit pour les suites sans répétition d'éléments. De telles suites sont appelées suites sans facteurs carrés.

Historiquement, les motivations pour l'étude de ces types de problèmes sont nombreuses.

En effet, tout à commencé en 1961 avec un premier résultat obtenu par Erdős, Ginzburg et Ziv. Ils ont montré que sur le groupe cyclique à n éléments $\mathbb{Z}/n\mathbb{Z}$, toute suite S de longueur $2n-1$, contient une sous-suite T de longueur n et de somme nulle [EGZ61].

Leurs motivations sont issues principalement d'une question de théorie élémentaire des nombres énoncée de la façon suivante :

Quelle est la taille minimale d'une collection d'entiers qui garantit de trouver n éléments dans cette collection dont la somme est divisible par n ?

Au passage, nous remarquons que ce problème formulé quelque soit pour des suites ou des ensembles d'entiers, après réduction modulo n, induit toujours un problème

pour des suites avec répétitions dans $\mathbb{Z}/n\mathbb{Z}$.

Environ une décennie plus tard, exactement en 1973, Harborth a considéré cette question de théorie des nombres en dimension supérieure de la façon suivante :

Pour un paramètre n et une dimension r donnés, quel est le nombre minimum de points du treillis $\mathbb{Z}^r$ qu'il faut considérer pour que l'on puisse trouver n points parmi ces points dont la somme est divisible par n ? c'est-à-dire chaque coordonnée de la somme est divisible par n. [Har73]

Autrement dit, quel est le nombre minimum de points du treillis $\mathbb{Z}^r$ qu'il faut considérer pour que l'on puisse trouver n points parmi ces points dont le barycentre est aussi un point du treillis ? Cette question revient au problème de déterminer la plus petite longueur d'une suite avec répétition d'éléments de $(\mathbb{Z}/n\mathbb{Z})^r$ qui admet une sous-suite de somme nulle et de longueur n.

Plusieurs recherches ont été menées pour étudier cette question.

En effet, en 1983, Arnfried Kemnitz a obtenu des résultats pour la dimension $r = 2$ en admettant l'hypothèse que la plus petite longueur recherchée de la suite vaut $4n - 3$ [Kem83]. Ensuite, il y a eu plusieurs tentatives pour montrer cette extension du théorème de EGZ, dite « conjecture de Kemnitz ». Nous citons par exemple les travaux d'Alon et Dubiner qui l'ont démontré pour $6n - 5$ éléments [AD93], en 2000, Lajos Rónyai l'a démontré pour 4n – 2 éléments si n est premier [Rón00], et en 2001, Gao a étendu ce résultat partiel au cas où n est une puissance d'un nombre premier [GT04].

Le problème pour $r = 2$ n'a été résolu complètement qu'en 2003 par Christian Reiher [Rei07] et indépendamment par Carlos di Fiore qui ont démontré la conjecture de Kemnitz (voir l'article de Savchev et Chen [SC05] qui donne aussi une version affinée et plus générale de la conjecture de Kemnitz).

Cependant, ce problème reste ouvert en dimension supérieure à 2.

Une généralisation naturelle de ce problème à n'importe quel groupe abélien fini G consiste à déterminer le plus petit entier $k \in \mathbb{N}$ tel que toute suite de longueur supérieure ou égale à k contient une sous suite de longueur égale à l'exposant du groupe et de somme nulle.

De nos jours cette constante est souvent appelée la constante d'Erdős-Ginzburg-Ziv, notée $\mathrm{s}(G)$.

Une autre généralisation naturelle du théorème Erdős-Ginzburg-Ziv serait de considérer le problème suivant :

Quel est le plus petit entier k tel que toute suite de longueur supérieure ou égale à k, contient une sous-suite de longueur égale à l'ordre du groupe et de somme nulle ?

Cette constante est souvent notée $E(G)$, or de nos jours, cette constante est peu étudiée, puisque selon un théorème de Gao, elle peut être déduite d'une autre constante qui est plus facile à étudier, à savoir la constante de Davenport notée $\mathcal{D}(G)$ [GG06]. Concrètement, pour tout groupe abélien fini G, la constante $E(G)$ s'exprime en fonction de la fonction $\mathcal{D}(G)$ de la façon suivante :

$$E(G) = \mathcal{D}(G) + |G| - 1$$

Rappelons que trouver la constante de Davenport d'un groupe abélien fini G est de trouver :

le plus petit entier k tel que toute suite de longueur supérieure ou égale k contient une sous-suite non-vide et de somme nulle.

Notons que le problème de déterminer la valeur exacte de la constante de Davenport fait partie des problèmes dont il n'existe pas de restriction sur la longueur de la sous-suite. De plus, ce problème reste ouvert pour de nombreux groupes abéliens, et la valeur exacte de la constante de Davenport n'est connue que pour des groupes abélien finis particuliers, voir chapitre V, page 93.

Historiquement, en 1966, la constante de Davenport a été popularisé dans un contexte de théorie algébrique des nombres lors d'une intervention à une conférence donnée par Davenport. Olson en parle dans [Ols69a] et [Ols69b].

Or, effectivement, le lien entre la théorie algébrique des nombres et ce problème de suite de somme nulle a déjà été détaillé par Roger dans [Rog63]. Cet article a été oublié pendant des décennies et n'a été redécouvert qu'en 2008 dans [RST08].

Les premiers travaux approfondis sur cette constante datent de 1969 et ils sont dûs à plusieurs chercheurs, notamment Olson et un groupe de chercheurs à Amsterdam comme Baayen et van Emde Boas, par exemple : [Ols69a], [Ols69b] et [PVEB69].

Notons C_n le groupe cyclique d'ordre n. Soit G un groupe abélien fini, noté additivement. D'après le théorème de structure des groupes abéliens finis, le groupe G se décompose en somme directe de groupes cycliques.

Soit $(G, +)$ un groupe abélien fini. Il existe une séquence $(n_i)_{i=1 \text{ à } r}$ de r entiers tel que chaque entier divise le suivant et tel que G est isomorphe à la somme des r groupes cycliques C_{n_i}. L'entier r s'appelle le rang de G et l'entier n_r s'appelle l'exposant de G.

Définissons $\mathcal{D}^*(G)$ comme :

$$\mathcal{D}^*(G) = 1 + \sum_{i=1}^{r} (n_i - 1).$$

On a l'inégalité suivante [Rog63].

$$\mathcal{D}^*(G) \leq \mathcal{D}(G)$$

D'où

$$\forall n \in \mathbb{N}^*,\ \mathcal{D}(C_n) = n$$

Le minorant est obtenu en exhibant un exemple explicite d'une suite sans sous-suite de somme nulle. Ce minorant apparaît dans plusieurs travaux, par exemple dans [Rog63].

De plus Olson à démontré deux résultats [Ols69b], [Ols69a], donnant la valeur exacte de la constante de Davenport dans le cas des p-groupes abéliens finis (où l'exposant du groupe est une puissance d'un nombre premier), et des groupes abéliens finis de rang inférieur ou égal à deux. En dehors de ces résultats, peu d'autres sont connus.

D'autres motivations et applications de ce problème ont été découvertes, comme par exemple dans la factorisation des entiers algébriques. La constante de Davenport intervient dans la généralisation du théorème des nombres premiers, portant sur la répartition des éléments irréductibles dans un anneau d'entiers algébriques ([Nar13], chapitre 9.2).

La suite du document est structurée ainsi :

— Dans le chapitre 2, nous rappelons quelques notions générales.

— Dans le chapitre 3, nous donnons des résultats sur une constante appelée la constante de Harborth. Concrètement, pour un nombre premier n, nous avons déterminé cette constante pour le groupe $C_3 \oplus C_{3n}$. Nous l'avons fait en établissant d'abord un minorant reposant sur une intuition acquise à partir des résultats de calculs, ensuite nous avons prouvé le majorant correspondant en utilisant les divers théorèmes de la théorie additive des nombres, notamment les théorèmes de Cauchy-Davenport, Dias de Silva-Hamidoune et Vosper. Le contenu de ce chapitre correspond à notre article [Gui+19].

— Dans le chapitre 4, nous présentons un algorithme permettant de calculer la constante d'Erdős-Ginzburg-Ziv. Cet algorithme nous a permis de trouver la valeur de la constante pour des groupes abéliens finis assez grand. En particulier, nous l'avons déterminé pour les groupes $C_2 \oplus C_2 \oplus C_2 \oplus C_4$ et $C_2 \oplus C_2 \oplus C_2 \oplus C_6$.

— Nous avons adapté cet algorithme dans le chapitre 5 pour trouver des résultats pour une autre constante notée $\eta(G)$.

— Enfin, nous concluons en présentant un bilan des constantes étudiées. Lors des études d'algorithmes, nous avons eu besoin de manipuler les ensembles, nous discutons dans ce chapitre les différentes façons de les représenter dans un programme. L'ensemble des résultats a été mis dans une application web qui sera présentée à la fin de cette conclusion.

CHAPITRE II

PRÉLIMINAIRES

Dans ce chapitre, nous rappelons d'une manière assez détaillée les définitions de plusieurs structures algébriques (groupe, monoïdes ,monoïdes libres, etc ...) et notions liées (suite, séquences, etc ...). Certes ces notions sont déjà bien connues mais l'usage n'est pas tout à fait uniforme dans la littérature. De plus, nous travaillons avec plusieurs concepts et terminologies qui sont proches les uns des autres. Il est alors nécessaire de bien détailler ces terminologies pour éviter toute confusion. Par exemple, nous serons amenés à traiter trois différents types de collection des éléments d'une structure algébrique : les sous-ensembles où il n'y a pas de répétition des éléments et l'ordre des éléments n'est pas pertinent, les suites où la répétition des éléments est admise et l'ordre des éléments est pertinent, et enfin les séquences où il y des répétitions des éléments mais l'ordre des éléments n'est pas pertinent.

De plus, nous traitons dans la deuxième section de ce chapitre le problème de numéroter certaines collections des éléments.

Nous désignons par $\mathbb{N}$ l'ensemble des entiers positifs et par $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ l'ensemble des entiers non négatifs.

Pour tout entiers relatifs a et b tel que a est inférieure ou égale à b, nous noterons $[a,b]$, l'ensemble $\{n \in \mathbb{Z} \mid a \leq n \leq b\}$.

De même, pour tout entiers relatifs a et b tel que a est strictement inférieure à b, nous noterons $[a,b[$, l'ensemble $[a,b[= \{n \in \mathbb{Z} \mid a \leq n \leq b-1\}$.

1 Structures algébriques et collection des éléments

Nous rappelons dans les sous-sections 1 à 4, les définitions de magma, groupe, sous-groupe et la description de la structure des groupes abéliens. Nous discutons

ensuite dans les sous-sections de 5 à 7 le prolongement des applications et le morphisme de monoïde ainsi que le concept de structure libre. La sous-section 9 est particulièrement importante. Elle introduit les monoïdes commutatifs libres sur un alphabet, ses éléments sont appelés des séquences. Ces dernières sont des objets centraux dans ce type de combinatoire additive que nous traitons ici. Nous concluons ces discussions avec le lien entre les suites croissantes et les mots binaires.

1-) Un magma

Un magma est un ensemble M muni d'une loi de composition interne $*$

On dit que le magma $(M,*)$ est :

— unifère s'il possède un élément neutre e :

$$\exists e \in M,\ \forall x \in M, x * e = e * x = x$$

— un semi-groupe si la loi $*$ est associative :

$$\forall x, y,\ z \in M, (x * y) * z = x * (y * z)$$

— un monoïde s'il vérifie les deux propriétés : l'associativité et l'existence d'un élément neutre. Pour préciser l'élément neutre, nous utilisons la notation $(M,*,e)$.

— un monoïde E est dit simplifiable à gauche (respectivement à droite) si :

$$\forall (a,b,c) \in E^3,\ a * b = a * c \text{ (respectivement } b * a = c * a) \Rightarrow b = c$$

Un monoïde E simplifiable à gauche et à droite est appelé un monoïde simplifiable. Dans ces travaux et dans de nombreux travaux en théorie de la factorisation, tous les monoïdes sont supposés être simplifiables.

2-) Un groupe

Un groupe G est un ensemble muni d'une loi de composition interne notée $*$ vérifiant :

— L'associativité ;

— L'existence de l'élément neutre ;

— Tout élément admet un inverse :

$$\forall x \in G,\ \exists x^{'} \in G,\ x * x^{'} = x^{'} * x = e$$

Si de plus : $\forall x, y \in G, x * y = y * x$, alors on dit que G est commutatif ou abélien.

3-) Un sous-groupe

Soit G un groupe. On dit qu'un sous-ensemble H de G est un sous-groupe de G lorsque les trois conditions suivantes sont vérifiées :

— L'ensemble H n'est pas vide ;

— Pour tout x et y de H, le produit xy est aussi dans H ;

— Pour tout x dans H, l'inverse x^{-1} de x est aussi dans H.

Soit $(G, \times)$ un groupe (éventuellement non-commutatif) et $g \in G$.

Le sous-groupe engendré par g est noté $\langle g \rangle$ est le plus petit sous-groupe de G contenant g.

S'il existe un élément $g \in G$ tel que $G = \langle g \rangle$, on dit que G est un groupe monogène Un groupe cyclique est un groupe qui est à la fois fini et monogène. L'élément g est appelé un générateur de G ou un élément primitif de G.

Soit $(G, \times)$ un groupe et $g \in G$. Dans la suite g est un générateur du groupe G.

S'il existe un élément $k \in \mathbb{N}$, tel que $g^k = 1$ alors on dit que g est d'ordre fini et l'entier $n = \min \{k \in \mathbb{N}, g^k = 1\}$ est appelé l'ordre de g est noté $\mathrm{ord}(g)$.

Soit $(G, \times)$ un groupe et $g \in G$ un élément d'ordre n. Alors on a $\langle g \rangle = \{1, g, g^2, \ldots, g^{n-1}\}$ et $\mathrm{card}\langle g \rangle = n$.

L'usage est de noter additivement les groupes commutatifs et lorsque le groupe n'est pas nécessairement commutatif on le note multiplicativement.

4-) La somme directe

Soient $(G_1, +_1, 0_1)$ et $(G_2, +_2, 0_2)$ deux groupes abéliens.

Soit $G_1 \times G_2$ le produit cartésien des ensembles G_1 et G_2.

On appelle somme directe de G_1 et de G_2 notée $G_1 \oplus G_2$ l'ensemble $G_1 \times G_2$ muni de la loi :

$$\forall x_1, y_1 \in G_1 \text{ et } x_2, y_2 \in G_2 \ (x_1, x_2) \oplus (y_1, y_2) = (x_1 +_1 y_1, x_2 +_2 y_2)$$

Propriété 1.1. — L'opération $\oplus$ munit $G_1 \oplus G_2$ d'une structure de groupe commutatif ;

— L'élément neutre est le couple $(0_1, 0_2)$;

— Notons par $-_1 x_1$ l'inverse d'un élément x_1 par la loi $+_1$, et $-_2 x_2$ l'inverse d'un élément x_2 par la loi $+_2$.

L'opposé de (x_1, x_2) est le couple $(-_1 x_1, -_2 x_2)$.

Remarque 1.2. La somme directe de deux groupes se généralise à la somme directe de k groupes.

Remarque 1.3.
— L'ensemble des classes d'équivalences $\mathbb{Z}/n\mathbb{Z}$ est le translaté de $n\mathbb{Z}$, noté $k+n\mathbb{Z}$ avec k un entier désignant un représentant quelconque de la classe.

— La notation $(\mathbb{Z}/n\mathbb{Z},+)$, désigne les classes de congruences modulo n par rapport à l'addition. Ceci est un groupe cyclique d'ordre n.

— La notation C_n, désigne un groupe cyclique d'ordre n.

— Pour k un entier positif, la notation $k \times g$ signifie $\underbrace{g+\cdots+g}_{k \text{ fois}}$. Le groupe $C_n = \langle g \rangle$ est isomorphe au groupe $(\mathbb{Z}/n\mathbb{Z},+)$, où l'isomorphisme est donné par l'application :

$$\phi' : \begin{cases} \mathbb{Z}/n\mathbb{Z} & \to C_n \\ k+n\mathbb{Z} & \mapsto k \times g \end{cases}$$

— Notons que si $k_1 \equiv k_2 \pmod{n}$, alors $k_1 \times g = k_2 \times g$, car l'ordre de g est égale à n.

— Tout groupe cyclique d'ordre n est isomorphe au groupe $\mathbb{Z}/n\mathbb{Z}$, en particulier, tous les groupes cycliques d'ordre n sont isomorphes. Or, en général il n'y a pas d'isomorphisme canonique, on remarque que l'isomorphisme donné ci-dessus dépend du choix de l'élément générateur qui n'est en général pas unique.

Théorème 1.4 (Kronecker (1870))**.** *Soit G un groupe commutatif fini non trivial, il existe un unique ensemble d'entiers positifs $\{n_1, n_2 \ldots, n_r\}$ tel que*

— *L'entier n_1 est strictement supérieur à 1.*

— *Pour tout i de 1 à $r-1$, la quantité n_i divise n_{i+1}.*

— *G est isomorphe à la somme directe $C_{n_1} \oplus C_{n_2} \oplus \cdots \oplus C_{n_r}$.*

Remarque 1.5.
— Pour le groupe trivial, l'ensemble $\{n_1, n_2 \ldots, n_r\}$ serait l'ensemble vide. Dans ce cas, le rang du groupe est égale à 0 et son exposant est égale à 1.

— L'entier r est appelé le rang de G. De plus si le groupe G est de cardinal strictement supérieur à 1, alors n_r est l'exposant du groupe G, noté $\exp(G)$;

— Notons que l'exposant du groupe peut être défini comme le ppcm des ordres des éléments de G.

— La notation C_n^r signifie la somme directe de r groupes cycliques d'ordre n.

Soit s un entier naturel non nul et soit $(e_1, \ldots, e_s)$ une famille de s éléments de G. Tout groupe commutatif est un $\mathbb{Z}-$module. Nous avons ainsi la notion de famille libre.

— On dit que les e_i sont libres si

$$\forall (m_1,\dots,m_s) \in \mathbb{Z}^s, \sum_{i=1}^{s} m_i e_i = 0 \Rightarrow \forall i \in \{1\dots s\}, m_i = 0.$$

En d'autres termes, la seule combinaison linéaire nulle d'éléments des e_i, est celle dont les coefficients m_i sont nuls, appelée combinaison linéaire triviale.

Or dans un groupe fini il n'y a pas de familles non-triviales qui soient libres dans ce sens. Donc, dans notre contexte, la notion suivante est plus pertinente :

$$\forall (m_1,\dots,m_s) \in \mathbb{Z}^s, \sum_{i=1}^{s} m_i e_i = 0 \Rightarrow \forall i \in \{1\dots s\}, m_i e_i = 0.$$

On appelle une telle famille, une famille indépendante.

— On dit qu'une famille $(e_1,\dots,e_s)$ est génératrice si

$$\forall g \in G,\ \exists (m_1,\dots,m_s) \in \mathbb{Z}^s \text{ tel que } g = m_1 e_1 + \cdots + m_s e_s$$

— Si une famille est à la fois indépendante et génératrice du groupe G alors elle est appelée une base de G.

5-) Prolongement d'application

Soient deux applications f et $f^{'}$ telles que :

$f : E \to F$

$f^{'} : E^{'} \to F^{'}$

Et $E \subset E^{'}$.

On dit que $f^{'}$ est un prolongement de f, si : $\forall x \in E,\ f(x) = f^{'}(x)$.

6-) Morphisme de monoïdes

Soient $(E, *, e)$ et $(F, \cdot, f)$ deux monoïdes.

Soit

$$\phi : \begin{cases} E & \to F \\ x & \mapsto \phi(x) \end{cases}$$

ϕ est un morphisme de monoïde si :

— $\forall (x,y) \in E^2, \phi(x * y) = \phi(x) \cdot \phi(y)$

— $\phi(e) = f$

7-) Une structure libre

Soient F et G deux structures du même type. Soit E un sous-ensemble de F. On dit que E est libre sur F si toute application de E dans G peut se prolonger en un morphisme de F dans G.

Exemple 1.6. :

Si F et G sont des espaces vectoriels de dimension finie sur un corps $\mathbb{K}$. Si $B = (b_1, \ldots, b_s)$ est une base de F, alors F est libre sur B car toute application définie sur la base B :

$$f : \begin{cases} B & \to F \\ b_i & \mapsto f(b_i) \end{cases}$$

se prolonge de manière unique en un morphisme à tous l'espace G.

Si $x = \sum_{i=1}^{s} \lambda_i b_i$ alors $f(x) = \sum_{i=1}^{s} \lambda_i f(b_i)$.

8-) Un monoïde libre sur un alphabet

Cette sous-section introduit le monoïde libre sur un alphabet. En pratique nous n'utiliserons par ce concept, mais nous l'introduisons car il permet d'introduire celui de monoïde commutatif libre sur un alphabet qui sera l'objet de la sous-section suivante.

Considérons un alphabet de n éléments $G_0 = \{g_1 \ldots g_n\}$. Les éléments de G_0 sont appelés des lettres. Un mot sur G_0 est par définition une suite finie de symboles appartenant à G_0.

On appelle le monoïde de mots ou le monoïde libre sur un alphabet G_0 l'ensemble des suites finies d'éléments de G_0. Il est muni de la loi de concaténation notée $\cdot$ qui est associative. Il est noté $(F^*(G_0), \cdot)$.

Ce monoïde est un monoïde libre sur l'alphabet G_0.

De plus, on note par G_0^ℓ l'ensemble des $\ell-$uplets qui est par définition l'ensemble des suites de longueur ℓ sur G_0.

$F^*(G)$ est donc la réunion de toutes les suites finies de toutes longueurs d'éléments de G_0

$$F^*(G_0) = \bigcup_{\ell \in N_0} G_0^\ell$$

Pour un élément $g \in G_0$ et un entier $s \geq 0$ nous utilisons la notation exponentielle pour exprimer une séquence constituée d'un élément g répété s fois.

$$g^s = \underbrace{g \cdots g}_{s \text{ fois}}$$

Remarque 1.7. — Le monoïde libre $F^*(G_0)$ possède un **élément neutre** noté $1_{F^*(G_0)} \in G_0^0$ qui est le mot vide.

— Un mot S^* du monoïde libre sur G_0 s'écrit sous la forme suivante :

$$S^* = g_{i_1} \cdots g_{i_\ell} \in F^*(G_0) \text{ avec } i_j \in \{1 \cdots n\}$$

9-) Monoïde commutatif libre sur un alphabet

Cette sous-section est centrale dans ce chapitre, car nous traitons le concept de séquences sur un ensemble G_0 qui jouent un rôle important dans nos travaux.

Nous définissons la relation d'équivalence $\sim$ sur $F^*(G_0)$ qui identifie deux mots s'ils ne diffèrent que par l'ordre de leurs termes.

Dit autrement, on dit que deux mots $S_1 = g_{i_1} \cdots g_{i_\ell} \in F^*(G_0)$ et

$S_2 = g_{j_1} \cdots g_{j_\ell} \in F^*(G_0)$ sont équivalents et on note $S_1 \sim S_2$ s'il existe une permutation τ de l'ensemble $\{1 \ldots n\}$ telle que :

$$\begin{array}{rccc} \tau & : & \{1 \ldots n\} & \to & \{1 \ldots n\} \\ & & i & \mapsto & \tau(i) \end{array}$$

$$\forall k \in \{1 \ldots \ell\}, g_{i_k} = g_{j_{\tau(k)}}$$

Ce qui est équivalent à :

$$\forall i \in \{1 \ldots n\} \ \#\{k \mid i_k = i\} = \#\{k \mid j_k = i\}$$

Pour tout S de $F^*(G_0)$, on note $\overline{S}$ la classe d'équivalence de S défini par l'ensemble des mots équivalents à S :

$$\overline{S} = \{T \in F^*(G) : S \sim T\}$$

L'ensemble quotient de $F^*(G_0)$ par la relation $\sim$ est l'ensemble des classes d'équivalences des éléments de $F^*(G_0)$ et noté $F^*(G_0)/\sim$.

Définition 1.8 (Séquences)**.** Soit G_0 un ensemble. On appelle le monoïde des séquences, le monoïde abélien libre sur G_0 et on le note par $F(G_0, \cdot)$, l'ensemble des classes d'équivalences d'éléments de $F^*(G_0)$ par la relation $\sim$ définie ci-dessus.

$$F(G_0) = F^*(G_0)/\sim$$

On appelle une telle classe d'équivalence une séquence de G_0.

Pour tout élément g de G_0, l'ordre de multiplicité de l'élément g dans un représentant de la classe $\overline{S}$ est le nombre d'apparition de l'élément g dans un représentant de la classe $\overline{S}$ noté $\nu_g(\overline{S})$.

— Tout élément $\overline{S}$ de $F(G_0)$ a un représentant de la forme :

$$g_1^{\nu_{g_1}(\overline{S})} g_2^{\nu_{g_2}(\overline{S})} \dots g_n^{\nu_{g_n}(\overline{S})}$$

— Tout élément $\overline{S}$ de $F(G_0)$ est représenté par une application :

$$\begin{array}{rccc} \nu_{\overline{S}} & : & \{1 \dots n\} & \to & \mathbb{N} \\ & & i & \mapsto & \nu_{g_i}(\overline{S}) \end{array}$$

$F(G_0, \cdot)$ est muni de la loi de concaténation commutative.

Définition 1.9 (Concaténation commutative)**.** Soient $\overline{S}$ et $\overline{T}$ deux éléments de $F(G_0)$ et g un élément de G_0, la concaténation des deux éléments $\overline{S} \cdot \overline{T}$ est définie par une application :

$$\nu_g(\overline{S} \cdot \overline{T}) = \nu_g(\overline{S}) + \nu_g(\overline{T})$$

Exemple 1.10. Soit G_0 l'alphabet fini $G_0 = \{1, 2, 3, 4\}$ et soit la séquence T donné par $T = 2^2 \cdot 4^3 \cdot 1$. La séquence T contient deux éléments égaux à 2, trois éléments égaux à 4 et un élément égal à 1. On peut alors écrire $\nu_2(T) = 2$, $\nu_4(T) = 3$ et $\nu_1(T) = 1$.

Une séquence T est appelée une sous-séquence de S si T divise S selon la loi de $F(G_0)$. C'est-à-dire qu'il existe une séquence T' de $F(G_0)$ tel que $TT' = S$. Autrement dit, si pour tout élément g dans G_0 on a $\nu_g(T) \leq \nu_g(S)$.

De plus, si la séquence T divise la séquence S, alors on note ST' par ST^{-1} ou bien $T^{-1}S$. Elle peut s'interpréter comme la séquence obtenue en retirant T de la séquence S. On a alors :

$$\forall g \in G_0,\ \nu_g(ST^{-1}) = \nu_g(T^{-1}S) = \nu_g(S) - \nu_g(T)$$

On pourrait aussi interpréter T^{-1} comme l'inverse de T dans le groupe de fraction de $F(G_0)$, or nous avons besoin de cette notation uniquement dans le cas où T divise S avec $ST^{-1} \in F(G_0)$.

Soient S et T deux séquences de $F(G_0)$. Nous utilisons la notation du plus grand commun diviseur de S et T, noté $\mathrm{pgcd}(S, T)$. Il peut être défini comme le diviseur commun à S et T qui est maximal pour l'ordre donné par la relation « divise ». En effet, dans $F(G_0)$, il y a toujours un unique diviseur maximal commun à S et T. Concrètement :

$$\forall g \in G_0,\ \nu_g(\mathrm{pgcd}(S,T)) = \min\{\nu_g(S), \nu_g(T)\}$$

La quantité $\mathrm{pgcd}(S,T)$ correspond aussi à la plus longue séquence constituée d'éléments en communs entre les séquences S et T. Celle-ci est une sous-séquence de S et de T.

Pour une séquence

$$S = g_{i_1} \cdots g_{i_\ell} \in F(G_0)$$

— La longueur de la séquence S est noté $|S|$.

$$|S| = \ell = \sum_{g \in G_0} \nu_g(S)$$

— L'ordre de multiplicité maximal de la séquence S est noté par $\mathrm{h}(s)$.

$$\mathrm{h}(s) = \max\{\nu_g(S) : g \in G_0\} \in \{0, \ldots |S|\}$$

— Le support de la séquence S est l'ensemble des éléments de G_0 qui sont dans la séquence S.

$$\mathrm{supp}(S) = \{g \in G_0 : \nu_g(S) > 0\}$$

— La séquence S est dite sans facteurs carrés si pour tout $g \in G_0$ l'ordre de multiplicité ν_g est inférieur ou égal à 1. Autrement dit, pour tout i_j tel que $1 \leq i_j \leq n$ les g_{i_j} sont tous distincts.

Soit $(M, *)$ un monoïde, et f une application :

$$f \ : \ G_0 \ \rightarrow \ M$$

Alors, l'application f se prolonge en un unique homomorphisme de monoïde f^* :

$$f^* \ : \ F^*(G_0) \ \rightarrow \ M$$

Si $(M, *)$, est un monoïde commutatif, alors il existe un homomorphisme de monoïde commutatif donné par :

$$f^+ \ : \ F(G_0) \ \rightarrow \ M$$

Si G_0 est un sous-ensemble d'un groupe commutatif $(G, +)$ alors l'inclusion :

$$id \ : \ G_0 \ \rightarrow \ G$$

induit un homomorphisme de monoïde :

$$id^* \quad : \quad F^*(G_0) \quad \rightarrow \quad G$$

Et

$$id^+ \quad : \quad F(G_0) \quad \rightarrow \quad G$$

Concrètement, soit $S \in F(G_0)$, la séquence $S = g_1 \cdots g_\ell$. Les applications id^* et id^+ sont notées σ représentant la somme de la séquence S, où $\sigma(S)$ est donnée par :

$$\sigma(S) = \sum_{i=1}^{\ell} g_i = \sum_{g \in G_0} \nu_g(S) g \in G$$

Si G_0 et G_1 sont deux ensembles et

$$f \quad : \quad G_0 \quad \rightarrow \quad G_1$$

est une application alors, il existe un unique homomorphisme de monoïde

$$f^{**} \quad : \quad F^*(G_0) \quad \rightarrow \quad F^*(G_1)$$

Tel que pour tout g de G_0, on a $f^{**}(g) = f(g)$

Et

$$f^{++} \quad : \quad F(G_0) \quad \rightarrow \quad F(G_1)$$

Tel que pour tout g de G_0, on a $f^{++}(g) = f(g)$.

Pour simplifier la notation on note f pour f^{**} et f^{++}.

Si G_0 et G_1 sont des sous-ensembles de deux groupes abéliens, et f est un homomorphisme entre ces deux groupes, alors $f(\sigma(S)) = \sigma(f(S))$. De plus, la longueur de la séquence S est toujours égale à la longueur de la séquence $f(S)$, $|S| = |f(S)|$. Notons que ceci reste vrai même si l'application f n'est pas injective. En revanche, il est possible que le cardinal du support de la séquence S soit strictement plus grand que celui du support de la séquence $f(S)$.

Pour un entier positif k, l'ensemble des sommes des sous-séquences non vides de S et de longueur k est donné par :

$$\Sigma_k(S) = \{\sigma(T) \colon T \mid S \text{ avec } |T| = k\}$$

L'ensemble des sommes de toutes les sous-séquences T non-vides de la séquence S est donné par :

$$\Sigma(S) = \{\sigma(T) : 1 \neq T|S\}$$

Rappelons que la notation que nous utilisons pour la sous-séquence vide est 1.

Une séquence S est dite :

— libre de somme nulle si $0 \notin \Sigma(S)$;

— de somme nulle si $\sigma(S) = 0$.

Définition 1.11 (Somme d'ensembles)**.** Soient A et B deux sous-ensembles d'un groupe G, on appelle somme des deux ensembles A et B, l'ensemble $A + B$ défini par :

$$A + B = \{a + b : a \in A, b \in B\}$$

Si A contient un seul élément $A = \{g\}$, alors, nous pouvons utiliser la notation $g + B$ au lieu de $A + B$.

Nous utilisons aussi une notation semblable pour les séquences.

Définition 1.12 (Somme de séquences)**.** Soient $g \in G$ et $S = g_1 \cdots g_\ell$. La somme $g + S$ est la séquence obtenue en translatant de g chacun des termes de S, soit :

$$(g + g_1) \ldots (g + g_\ell)$$

Remarque 1.13. La séquence S contient une sous-séquence de longueur $\exp(G)$ et de somme nulle si et seulement si $g + S$ contient une sous-séquence de longueur $\exp(G)$ et de somme nulle.

Définition 1.14 (Somme restreinte)**.** Soient A et B deux sous-ensembles d'un groupe G, on appelle la somme restreinte de A et B l'ensemble des sommes de termes distincts de a et de b.

$$A \widehat{+} B = \{a + b : a \in A, b \in B, a \neq b\}$$

Remarque 1.15. — Notons que

$$A \widehat{+} A = \Sigma_2(A)$$

— On définit d'une manière analogue la somme $\Sigma_2(A)$ pour les ensembles.

10-) Suites croissantes

Soient G un groupe abélien fini, G_0 un sous-ensemble de G et S une séquence du monoïde commutatif libre $F(G_0)$ définie par :

$$S = g_{i_1} \cdots g_{i_\ell} \in F(G_0) \text{ avec } i_j \in \{1 \dots n\}$$

Comme nous l'avons déjà évoqué, un concept central dans nos travaux est celui de séquences où l'ordre des éléments n'importe pas. Pour des raisons techniques, il est peut être utile d'en avoir une représentation canonique. Une possibilité de le faire est de trier ces éléments d'une manière particulière. Pour le faire, il faut un ordre sur les éléments de l'alphabet sous-jacent. Dans certains cas, un ordre naturel sur cet ensemble existe, par exemple dans le cas où nous considérons un ensemble d'entiers, cependant, dans notre cas où nous traitons des sous-ensembles d'un groupe, à priori, il n'y a pas toujours un ordre naturel sur les éléments.

On munit alors le sous-ensemble G_0 du groupe abélien fini G d'un ordre total en posant arbitrairement :

$$g_1 < \cdots < g_n$$

Soit $F^c(G_0)$ le sous-ensemble de $F^*(G_0)$ qui contient les suites croissantes sur G_0.

Pour tout élément $\overline{S}$ de $F(G_0)$, l'ensemble $F^c(G_0)$ contient exactement un élément de la classe $\overline{S}$. Cet élément est appelé le représentant canonique de $\overline{S}$.

Exemple 1.16. Soit $G_0 = \{a, b, c\}$. Définissons l'ordre total sur G_0 par $a < b < c$.

Soit $S = a \cdot b \cdot a \in F^*(G_0)$. On a $\overline{S} = \{a \cdot a \cdot b, a \cdot b \cdot a, b \cdot a \cdot a\} \in F(G_0)$.

Le représentant canonique est la suite croissante $a \cdot a \cdot b \in \overline{S}$.

Remarque 1.17. — L'ensemble des suites croissantes $F^c(G_0)$ est constitué des éléments canoniques de chaque classe d'équivalence $\overline{S} \in F(G_0)$.

11-) Les mots binaires

On peut représenter les sous-ensembles d'entiers de $\{0, \dots, \ell - 1\}$ comme des mots binaires de longueur ℓ, où chaque position du mot binaire correspond à un élément du sous-ensemble, tel que la position i du mot binaire vaut 1 si l'élément i appartient ou sous-ensemble et 0 sinon.

Soit G_0 un alphabet fini $G_0 = \{g_1, \dots, g_n\}$.

On décrit les sous-ensembles de G_0 par un mot ou suite binaire de longueur n c'est-à-dire par $x \in \{0,1\}^n$, de la façon suivante :

Pour un sous-ensemble F de G_0, le mot binaire $x_F \in \{0,1\}^n$ décrit F par la relation suivante :

$$g_i \in F \Longleftrightarrow x_i = 1.$$

Exemple 1.18. Soit $G_0 = \{a_1, a_2, a_3, a_4, a_5\}$. Le sous-ensemble $F = \{a_1, a_4\}$ de G_0 se décrit comme $x_F = 10010$.

Un tel mot binaire est appelé « bitmap » dans le langage des informaticiens.

Définition 1.19 (Le poids de Hamming)**.** Le poids de Hamming d'un bitmap noté $w_H(x)$ est égale au nombre de ces composantes non nulles, il vaut le cardinal du sous-ensemble qu'il représente.

Remarque 1.20 (Notation)**.** L'ensemble des mots binaires de longueur n et de poids k est noté $B_{n,k}$.

$$B_{n,k} = \{x \in \{0,1\}^n | w_H(x) = k\} \tag{1.1}$$

2 La numérotation spéciale des séquences, suites et sous-ensembles

Nous nous intéressons ici aux séquences, qui sont définies sur le monoïde commutatif libre $F(G_0)$, le but étant de les numéroter toutes sur un alphabet donné de taille fixe. Or, comme il existe un nombre infini de séquences sur un alphabet de taille fixe, on se limite à numéroter celles d'une longueur donnée.

Nous avons vu, qu'en imposant un ordre total sur les éléments de l'alphabet, alors toute séquence a un représentant canonique. Ce représentant est une suite croissante. Numéroter les séquences revient alors à numéroter les suites croissantes sur un alphabet donné.

Il existe une technique bien connue pour numéroter les suites croissantes sur un tel alphabet, voir [Mor17], où nous numérotons plutôt les suites strictement croissantes sur un alphabet de taille plus grande. Notons que les suites strictement croissantes correspondent aux sous-ensembles.

Rappelons qu'un ensemble de n éléments contient 2^n sous-ensembles, que nous pouvons représenter chacun par un mot binaire. Alors, une manière de numéroter tous les sous-ensembles serait d'interpréter les mots binaires comme des entiers en numération en binaire. Or, cette numérotation ne nous convient pas ; car nous voulons numéroter les sous-ensembles de sorte que le numéro d'un sous-ensemble de cardinal n est toujours plus petit que le numéro d'un sous-ensemble de cardinal $n^{'}$ si n est inférieur à $n^{'}$. Autrement dit, nous voulons imposer que tous

les sous-ensembles d'un certain cardinal forment des entiers consécutifs de l'intervalle $[0, 2^n - 1]$. Donc l'objectif est de numéroter les éléments de $\{0,1\}^n$ comme réunion de $B_{n,k}$ pour $k = 0$ à n. D'où l'utilisation d'une numérotation spéciale que nous décrivons dans cette section.

1-) Les étapes de numérotation des séquences

Soient G un groupe abélien fini, $G_0 = \{g_1, \ldots, g_n\}$ un sous-ensemble de G.

Nous voulons numéroter toutes les séquences $g_{i_1} \ldots g_{i_\ell} \in F(G_0)$ avec $i_j \in \{1 \ldots n\}$

Pour avoir cette numérotation, nous établissons une bijection entre ces séquences et des mots binaires en suivant les étapes suivantes :

1. On munit le sous-ensemble G_0 du groupe abélien fini G d'un ordre total en posant :

   $$g_1 < \cdots < g_n$$

   Avec cet ordre, nous savons qu'il y a une bijection entre l'ensemble des séquences et l'ensemble des suites croissantes $F^c(G_0)$.

2. Soit η l'application qui pour $G_0 = \{g_1, \ldots, g_n\}$ attribue respectivement les entiers $\{1, \ldots, n\} \in G_0^{'}$ tel que $\eta(g_i) = i$.

   Le prolongement η vers l'ensemble des suites $F(G_0)$ et $F(G_0^{'})$ établi une bijection entre les suites croissantes sur G_0 et les suites croissantes sur $\{1 \ldots n\}$ ce qui préserve la longueur des suites.

   Donc, nous avons besoin de parcourir l'espace de ces suites croissantes d'entiers d'une longueur donnée. De telles suites peuvent avoir des répétitions.

   Prenons l'exemple suivant :

   Soit $G_0^{'} = \{1, 2, 3, 4\}$ et S la suite croissante $S = 1 \cdot 1 \cdot 1 \cdot 2 \cdot 2 \cdot 4$. Notons que dans cet exemple la taille de l'alphabet n est égale à 4 et la longueur k de la suite est égale à 6.

3. L'étape suivante est le passage d'une suite croissante d'entiers à une suite strictement croissante d'entiers donné par l'application ϕ :

   Soit F^s l'ensemble des suites strictement croissantes d'entiers.

   $$\phi\colon \begin{cases} F^c(G_0^{'}) & \to F^s \\ S = u_1 \cdots u_\ell & \mapsto v_{u_1} \cdots v_{u_\ell} \text{ tel que } \forall i \in \{1 \ldots \ell\} \quad v_i = u_i + (i-1) \end{cases}$$

   Notons que cette application ϕ est injective. De plus, l'image de l'ensemble des suites croissantes d'une longueur donnée k est l'ensemble des suites strictement croissantes sur un alphabet de taille $n + k - 1$.

En suivant l'exemple précédent, ce passage élimine les répétions en transformant la suite S sur l'alphabet $\{1, \ldots, 4\}$ en une suite S' sur l'alphabet $\{1,\ldots,9\}$ qui est strictement croissante. Notons que $n+k-1=9$ et $S' = 1 \cdot 2 \cdot 3 \cdot 5 \cdot 6 \cdot 9$

Cette suite sans répétition est représentée par un sous-ensemble de 6 éléments choisis parmi $\{1,\ldots,9\}$.

4. Etant donné que la suite strictement croissante est sans répétition d'éléments, il existe une bijection entre l'ensemble des suites strictement croissantes de longueur k sur les entiers de 1 à n et l'ensemble des sous-ensembles de $\{1,\cdots,n+k-1\}$. Dans l'exemple précédent, le sous-ensemble est représenté par un mot binaire de poids 6 et de longueur 9. Parcourir l'espace des suites croissantes de longueur k sur alphabet de n éléments revient donc à parcourir l'ensemble des mots binaires de poids k et de longueur $n+k-1$.
5. Enfin, nous transformons la séquence binaire en un entier en parcourant une structure décrite plus loin et que nous appelons le triangle de Pascal modifié [Gui].

Pour des entiers $-1 \leq j \leq i$, on note $a_{i,j}$ le cumul des coefficients binomiaux définis précisément par :

$$a_{i,j} = \sum_{k=0}^{j} \binom{i}{k} \tag{2.1}$$

Remarques 2.1. — Les coefficients $a_{i,j}$ héritent des coefficients binomiaux de la relation du triangle de Pascal.

$$\text{pour tout } i,j \in \mathbb{N} \quad a_{i,j} = a_{i-1,j} + a_{i-1,j-1} \tag{2.2}$$

— Dans la suite, la table des coefficients $a_{i,j}$ tel que $j \leq i$ est appelée le triangle de Pascal modifié.

— Comme pour le triangle de Pascal classique, on peut illustrer les tables du triangle de Pascal modifié, comme le montre la figure II.1.

n \ k	-1	0	1	2	3	4
-1	0	0	0	0	0	0
0	0	1	1	1	1	1
1	0	1	2	2	2	2
2	0	1	3	4	4	4
3	0	1	4	7	8	8
4	0	1	5	11	15	16

FIGURE II.1 – Le triangle de Pascal modifié

Proposition 2.2 (Caractérisation des coefficients du triangle de Pascal modifié)**.**
Les relations suivantes sont tout aussi immédiates par récurrence.

- *Pour tout* $i \geq -1,\ a_{i,-1} = 0$.
- *Pour tout* $i \geq -1,\ a_{i,0} = \binom{i}{0} = 1$.
- *Pour tout* $i \geq 0,\ a_{i,i} = \sum_{k=0}^{i} \binom{i}{k} = 2^i$.
- *Pour tout* $i, j \in \mathbb{N}$ *:*

$$a_{i,j} = a_{i-1,j} + a_{i-2,j-1} + \cdots + a_{i-j-1,0} \tag{2.3}$$

n \ k	-1	0	1	2	3	4
-1	0	0	0	0	0	0
0	0	1	1	1	1	1
1	0	1	2	2	2	2
2	0	1	3	4	4	4
3	0	1	4	7	8	8
4	0	1	5	11	15	16

FIGURE II.2 – Exemple de la propriété 2.3 page 22

- *Pour tout* $i, j \in \mathbb{N}$ *:*

$$\begin{aligned} a_{i,j} = a_{i-1,j-1} + a_{i-2,j-1} + \cdots + a_{j-1,j-1} + \\ a_{j-2,j-2} + \cdots + a_{0,0} + 1 \end{aligned} \tag{2.4}$$

n \ k	-1	0	1	2	3	4
-1	0	0	0	0	0	0
0	0	1	1	1	1	1
1	0	1	2	2	2	2
2	0	1	3	4	4	4
3	0	1	4	7	8	8
4	0	1	5	11	15	16

FIGURE II.3 – Exemple de la propriété 2.4 page 22

Exemple 2.3. Soit $\{0,1\}^n$ l'ensemble des mots binaires de longueurs $n = 3$.

$\{0,1\}^n$	numéro	$w_H(x)$
000	0	0
001	1	1
010	2	1
100	3	1
011	4	2
101	5	2
110	6	2
111	7	3

— $\binom{3}{0} = 1$ mot de poids 0.

— $\binom{3}{1} = 3$ mots de poids 1 qui commencent au numéro $1 = \binom{0}{3}$.

— $\binom{3}{2} = 3$ mots de poids 2 qui commencent au numéro $4 = \binom{0}{3} + \binom{1}{3}$.

— $\binom{3}{3} = 1$ mot de poids 3 qui commencent au numéro $7 = \binom{0}{3} + \binom{1}{3} + \binom{2}{3}$.

La numérotation est donnée par la fonction suivante :

$$\begin{array}{rccc} \mathrm{Num}_n & : \{0,1\}^n & \to & \{0,\dots,2^n-1\} \\ & (x_1,\dots,x_n) & \mapsto & \sum_{i=0}^{n} a_{i,k_i} \end{array}$$

où $k_0 = -1$ et pour tout $i \in \{1 \dots n\}$ $k_i = k_{i-1} + x_i$

Cela correspond à un parcours dans le triangle de Pascal modifié où

— On part de l'indice (-1,-1).

— Le mot x est parcouru de gauche vers la droite.

— Si $x_i = 1$, on se déplace diagonalement en bas à droite.

— Si $x_i = 0$, on se déplace verticalement vers le bas.

— Le résultat est la somme de toutes les cases rencontrées.

Nous avons décrit la numérotation ci-dessus par un algorithme nommé numsubset voir Algorithme 2 page 28.

Exemple 2.4. Soit x le mot binaire de taille $n = 5$ et de poids $k = 3$ tel que $x = 10110$.

numsubset(10110, 5)=0+1+1+3+7+11=23.

n \ k	-1	0	1	2	3	4
0	0	0	0	0	0	0
1	0	1	1	1	1	1
2	0	1	2	2	2	2
3	0	1	3	4	4	4
4	0	1	4	7	8	8
5	0	1	5	11	15	16

FIGURE II.4 – Transformer le bitmap 10110 en numéro

Propriété 2.5. L'ensemble $B_{n,k}$ est l'ensemble des mots binaires de longueurs n et de poids k qui sont numérotés entre $a_{n,k-1}$ (inclus) et $a_{n,k}$ (exclu).

On peut alors écrire que $\mathrm{Num}_{n,k}$ est la restriction de Num_n à $B_{n,k}$

$$\begin{array}{lll} \mathrm{Num}_{n,k} & : B_{n,k} & \to [a_{n,k-1}, a_{n,k}[\\ & (x_1, \ldots, x_n) & \mapsto \sum_{i=0}^{n} a_{i,k_i}, \end{array}$$

où $k_0 = -1$ et pour tout $i \in \{1 \ldots n\}$ $k_i = k_{i-1} + x_i$

Remarque 2.6. Les applications Num_n ainsi que ${\mathrm{Num}_n}^{-1}$ sont calculables efficacement par un algorithme simple à déduire de la description ci-dessus, voir Algorithme 1 page 27 et Algorithme 2 page 28.

Remarque 2.7. Le cardinal de l'ensemble des mots binaires de longueurs n et de poids $B_{n,k}$ est égal au cardinal de son image $[a_{n,k-1}, a_{n,k}[$.

$\mathrm{card}(B_{n,k}) = \binom{n}{k}$ et $a_{n,k} - a_{n,k-1} = \sum_{l=0}^{k} \binom{n}{l} - \sum_{l=0}^{k-1} \binom{n}{l} = \binom{n}{k}$.

Propriété 2.8. L'application Num_n est :

— Bijective de $\{0,1\}^n$ sur $\{0 \ldots 2^n - 1\}$.

— Croissante selon le poids de son paramètre, c'est-à-dire, si $w_H(b) \leq w_H(c) \Rightarrow \mathrm{Num}_n(b) \leq \mathrm{Num}_n(c)$.

Preuve. Démontrons par récurrence que l'application Num_n est bijective.

Pour cela, montrons que $\mathrm{Num}_{n,k}$ est une bijection de $B_{n,k}$ sur $\{a_{n,k-1}, \ldots, a_{n,k} - 1\}$.

Initialisation : Pour $n = 1$ et $0 \leq k \leq 1$.

— Pour $k = 0$ on a $B_{1,0} \to [a_{1,-1}, a_{1,0}[= [0,1[= \{0\}$, est bijective.

— Pour $k = 1$ on a $B_{1,1} \to [a_{1,0}, a_{1,1}[= [1,2[= \{1\}$ est bijective.

Hypothèse de récurrence :
Pour tout k, l'application $\mathrm{Num}_{n-1,k} : B_{n-1,k} \to [a_{n-1,k-1}, a_{n-1,k}[$ est bijective

Décomposons l'ensemble $B_{n,k}$ des bitmaps de longueur n et de poids k en réunion disjointe de deux sous-ensembles selon la valeur de la dernière composante :

— $B^0_{n,k}$: ensemble des bitmaps de longueur n et de poids k dont la dernière composante est zéro. $B^0_{n,k} = \{(x_1, \ldots, x_n) \in B_{n,k} \mid x_n = 0\}$.

— $B^1_{n,k}$: ensemble des bitmaps de longueur n et de poids k dont la dernière composante est un. $B^1_{n,k} = \{(x_1, \ldots, x_n) \in B_{n,k} \mid x_n = 1\}$.

Notons que :

$$\begin{array}{rccc} \mathrm{red}^0 : & B^0_{n,k} & \to B_{n-1,k} \\ & (x_1, \ldots, x_n) & \mapsto (x_1, \ldots, x_{n-1}) \end{array}$$

et

$$\begin{array}{rccc} \mathrm{red}^1 : & B^1_{n,k} & \to B_{n-1,k-1} \\ & (x_1, \ldots, x_n) & \mapsto (x_1, \ldots, x_{n-1}) \end{array}$$

sont bijectives.

Par hypothèse de recurrence :

— $\mathrm{Num}_{n-1,k} \circ \mathrm{red}^0 : B^0_{n,k} \to [a_{n-1,k-1}, a_{n-1,k}[$ est bijective.

— $\mathrm{Num}_{n-1,k-1} \circ \mathrm{red}^1 : B^1_{n,k} \to [a_{n-1,k-2}, a_{n-1,k-1}[$ est bijective.

De plus, d'après la propriété 2.5 on a :

$$\text{Pour tout } b \in B^0_{n,k} \ \mathrm{Num}_{n,k}(b) = \mathrm{Num}_{n-1,k}(\mathrm{red}^0(b)) + a_{n-1,k-1}$$

et

$$\text{Pour tout } b \in B^1_{n,k} \ \mathrm{Num}_{n,k}(b) = \mathrm{Num}_{n-1,k-1}(\mathrm{red}^1(b)) + a_{n-1,k-1}$$

Il résulte que :

— La restriction de $\mathrm{Num}_{n,k}$ au sous-ensemble $B^0_{n,k}$ est une bijection de $B^0_{n,k}$ sur l'intervalle translaté $a_{n-1,k-1} + [a_{n-1,k-1}, a_{n-1,k}[$.

— De même, la restriction de $\mathrm{Num}_{n,k}$ au sous-ensemble $B^1_{n,k}$ est une bijection de $B^1_{n,k}$ sur l'intervalle translaté $a_{n-1,k-1} + [a_{n-1,k-2}, a_{n-1,k-1}[$.

Par conséquent comme les intervalles images sont disjoints, l'application $\mathrm{Num}_{n,k}$ est une bijection de $B_{n,k}$ sur l'intervalle translaté.

$a_{n-1,k-1} + [a_{n-1,k-1}, a_{n-1,k}[\cup a_{n-1,k-1} + [a_{n-1,k-2}, a_{n-1,k-1}[$.

Or d'après la relation du triangle de pascal 2.2 cet intervalle translaté est finalement l'intervalle $[a_{n,k-1}, a_{n,k}[$.

On rappelle que d'après le triangle de Pascal on a :

— $a_{n-1,k-1} + a_{n-1,k-2} = a_{n,k-1}$

— $a_{n-1,k-1} + a_{n-1,k} = a_{n,k}$.

Pour terminer la preuve, il reste à donner l'argument pour lequel la fonction Num_n est croissante selon le poids de son bitmap :

Si b et b' sont deux bitmaps tel que $w_H(b) < w_H(b')$. L'arrivée du chemin de b dans le triangle de Pascal modifié est à gauche de celui du chemin de b'.

□

Nous avons développé deux algorithmes qui sont inverses l'un de l'autre tels que :

— L'algorithme nommé *numsubset* rend le numéro d'un bitmap de longueur n et de poids k, ce qui correspond à la fonction $\mathrm{Num}_{n,k}$ (voir la propriété 2.5 page 24).

— L'algorithme nommé *subsetnum* rend le bitmap d'un numéro donné, ce qui correspond à la fonction Num_n^{-1} (voir Algorithme 1 page 27).

Si on mesure la complexité algorithmique par le nombre de division et multiplication, alors la complexité de cet algorithme est linéaire en n.

Algorithme 1 subsetnum

Entrée

x : un numéro

i : l'indice de la ligne courante

Sortie

w : le bitmap du numéro x

Variables

j : l'indice de la colonne courante

a : coefficient du triangle de Pascal modifié

p : coefficient du triangle de Pascal classique

Initialisation

$w \leftarrow 0$

$a \leftarrow 0$

$p \leftarrow 1$

$j \leftarrow 0$

1: **procédure**
2: Rechercher la colonne j telle que $a_{i,j} \leq x < a_{i,j+1}$
3: Initialisation de $a_{i,j}$
4: **tant que** $(i \neq (j-1)$ et $j \neq -1)$ **faire**
5: **si** $x \geq a$ **alors**
6: on peut soustraire $a_{i-1,j}$.
7: Nous mettons à jour $a_{i,j}$ avec la formule :
8: $a = \frac{a+p}{2}$ où $p = \frac{(p \times (i-j))}{i}$
9: La remontée est verticale donc on diminue le i de 1.
10: Le chiffre du bitmap est 0
11: **sinon**
12: on peut soustraire $a_{i-1,j-1}$.
13: La remontée est verticale vers la gauche
14: Nous mettons à jour $a_{i,j}$ avec la formule :
15: $a = \frac{a-(p \times (i-j)}{\frac{i}{2}}$
16: Nous mettons à jour $p_{i,j}$ avec la formule :
17: $p = \frac{(p \times j)}{i}$
18: Le chiffre du bitmap est 1
19: L'algorithme d'arrête dès qu'on est arrivé sur l'un des côtés
20: soit le côté $j = 0$, alors on complète avec des 0
21: soit le côté $i = j$, alors on complète avec des 1

Algorithme 2 numsubset

Entrée

x : le bitmap correspondant au sous-ensemble d'un ensemble de n éléments

Sortie

w : le numéro recherché

Variables

i, j : les indices du parcours du triangle de Pascal modifié

a : coefficient du triangle de Pascal modifié

p : coefficient du triangle de Pascal classique

Initialisation

$i \leftarrow -1$

$j \leftarrow -1$

$p \leftarrow 0$

$a \leftarrow 0$

$x \leftarrow 0$

1: **procédure**

2: **tant que** $(n > 0)$ **faire**

switch($w\&1$)

3: se déplacer dans le triangle selon la valeur du chiffre du bitmap

case(0)

4: se déplacer diagonalement vers le bas

5: la mise à jour du triangle de Pascal modifié par la formule :

6: $a = 2 \times a - p$

7: la mise à jour du triangle de Pascal classique par la formule :

8: $p = \frac{(p \times i)}{(i-j)}$

case(1)

9: la mise à jour du triangle de Pascal modifié par la formule :

10: $a = 2 \times a + p$

11: se déplacer diagonalement vers le bas à droite

12: la mise à jour de $p(i, j)$ en fonction de $p(i-1, j)$

endswitch

13: décaler w pour positionner le chiffre suivant

14: cumul de la valeur du sommet : $x+ = a$

15: compteur de boucle : $- - n$

CHAPITRE III

LA CONSTANTE DE HARBORTH

Sommaire

1 Introduction

Le contenu de ce chapitre correspond à notre article [Gui+19].

Soit $(G,+,0)$ un groupe abélien fini. La constante de Harborth de G, notée $\mathrm{g}(G)$, est le plus petit entier k tel que toute suite d'éléments deux-à-deux distincts de G de longueur k, de manière équivalente tout sous-ensemble de G de cardinal au moins k, admet une sous-suite de longueur $\exp(G)$ dont la somme soit nulle. Notons que $\mathrm{g}(G) \leq |G| + 1$.

Dans ce chapitre il est démontré que $\mathrm{g}(C_3 \oplus C_9) = 13$ et pour tout nombre premier $p \neq 3$, on a $\mathrm{g}(C_3 \oplus C_{3p}) = 3p + 3$.

Pour plus de détails sur l'étude de cette constante et des constantes similaires voir [GG06]. Notons qu'il est techniquement plus avantageux de travailler avec des séquences sans facteurs carrés, c'est-à-dire avec des séquences dont tous les termes sont distincts au lieu de travailler avec les ensembles. La raison est que dans la suite nous travaillerons avec des homomorphismes de groupe, et en appliquant un homomorphisme de groupe sur une séquence sans facteurs carrée, l'image de cette séquence sera une séquence de même longueur, or ceci n'est pas le cas pour les ensembles. voir la partie 9-) dans le chapitre des préliminaires.

Harborth [Har73] a considéré les constantes qui interviennent pour les séquences et pour les séquences sans facteurs carrés lorsque la longueur de la sous-séquence est égale à l'exposant du groupe.

Considérer ces constantes peut être vu comme une généralisation du problème énoncé dans le théorème d'Erdős-Ginzburg-Ziv (voir [EGZ61]) en passant des groupes cycliques aux groupes abéliens finis plus généraux.

La valeur de la constante de Harborth n'est connue que pour quelques types de groupes (voir [Baj18]).

Nous rappelons quelques résultats connus qui sont pertinents pour nos recherches actuelles.

a) Résultats connus

Voici ci-après la liste des groupes abéliens finis pour lesquels la valeur de la constante de Harborth est connue :

1. Si G est un 2-groupe, c'est-à-dire si son exposant vaut 2. Alors $\mathrm{g}(G) = |G| + 1$ signifiant qu'il n'y a pas de séquences sans facteur carré de longueur strictement supérieure au cardinal de G.

2. Pour G un 3-groupes élémentaires ${C_3}^n$, le problème de la détermination $\mathrm{g}(G)$ est particulièrement populaire car il est équivalent à plusieurs autres problèmes bien étudiés tels que les "cap-set" et les ensembles sans progression arithmétique à 3 termes.

   Néanmoins, la valeur de la constante de Harborth pour les 3-groupes élémentaires n'est connue que jusqu'au rang 6 (voir [Ede+07] pour un aperçu détaillé et voir [Pot08] pour le résultat du rang 6). Récemment Ellenberg et Gijswijt [EG17] en s'appuyant sur les travaux de Croot, Lev et Pach [CLP17], ont amélioré la borne supérieure de $\mathrm{g}(C_3^n)$ asymptotiquement quand n tend vers l'infini.

3. Lorsque G est un groupe cyclique, la seule séquence sans facteur carré dont la longueur vaut l'exposant du groupe est celle qui contient chaque élément du groupe G. Il suffit donc de vérifier si la somme de tous les éléments du groupe G est nulle ou pas.

   Plus concrètement, pour un entier strictement positif n, et pour le groupe cyclique C_n d'ordre n, on a :

$$\mathrm{g}(C_n) = \begin{cases} n & \text{si } n \text{ est impair} \\ n+1 & \text{si } n \text{ est pair} \end{cases}$$

4. Pour les groupes de rang 2, le problème de déterminer $\mathrm{g}(G)$ est encore ouvert dans certain cas.

   On sait que si p est un nombre premier, supérieur ou égal à 47 et pour $p \in \{3,5,7\}$, alors on a $\mathrm{g}(C_p \oplus C_p) = 2p-1$. Ce dernier résultat est dû à Kemnitz [Kem83], et le premier est dû à Gao et Thangadurai [GT04], avec une amélioration mineure supplémentaire en passant de $p \geq 67$ au départ à $p \geq 47$ [GGS07]. De plus, Gao et Thangadurai [GT04] ont déterminé que $\mathrm{g}(C_4 \oplus C_4) = 9$ et ont ensuite énoncé la conjecture suivante :

$$\mathrm{g}(C_n \oplus C_n) = \begin{cases} 2n-1 & \text{si } n \text{ est impair} \\ 2n+1 & \text{si } n \text{ est pair} \end{cases}$$

   De plus, Marchan et autres chercheurs [Mar+13] ont déterminé la valeur de la constante de Harborth pour les groupes de la forme $C_2 \oplus C_{2n}$:

$$\mathrm{g}(C_2 \oplus C_{2n}) = \begin{cases} 2n+3 & \text{si } n \text{ est impair} \\ 2n+2 & \text{si } n \text{ est pair} \end{cases}$$

   Enfin, Kiefer [Kie16] et [Baj18, Proposition F.104] ont montré que si n est un entier impair supérieur ou égal à 2, alors on a le minorant suivant : $\mathrm{g}(C_3 \oplus C_{3n}) \geq 3n+3$,(voir la sous-section a) page 35 pour plus de détails).

b) Théorèmes et lemmes techniques

Rappelons quelques théorèmes classiques de la théorie additive des nombres pour des groupes cycliques d'ordre premier.

1-) Le théorème de Cauchy–Davenport

Pour plus de détails concernant ce théorème, voir [Gry13, Theorem 6.2].

Théorème 1.1 (Cauchy–Davenport). *Soit p un nombre premier et soient A et B des sous-ensembles non-vides de C_p, alors :*

$$|A+B| \geq \min\{p, |A|+|B|-1\}$$

Il en découle immédiatement le résultat suivant pour les sous-ensembles non-vides $A_1, \ldots, A_h$ de C_p :

$$|A_1 + \cdots + A_h| \geq \min\Big\{p, \sum_{i=1}^{h} |A_i| - (h-1)\Big\}$$

2-) Le théorème de Vosper

Le problème qui consiste à représenter les ensembles quand la borne du théorème de Cauchy–Davenport est atteinte est résolu par le théorème de Vosper

Théorème 1.2 (Vosper). *Soient p un nombre premier et A, B des sous-ensembles non-vides de C_p. Supposons que A et B ont plus de deux éléments et que $|A+B| = |A|+|B|-1$.*

- *Si $|A+B| \leq p-2$, alors A et B sont des progressions arithmétique de même raison, c'est-à-dire, il existe $d \in C_p$ et $a, b \in C_p$ tels que : $A = \{a+id \colon i \in [0, |A|-1]\}$ et $B = \{b+id \colon i \in [0, |B|-1]\}$.*
- *Si $|A+B| = p-1$, alors il existe $c \in C_p$ tel que $A = \{c-a \colon a \in C_p \setminus B\}$.*

3-) Le théorème de Dias da Silva–Hamidoune

Il existe une généralisation du théorème de Cauchy-Davenport appelé le théorème de Dias da Silva–Hamidoune, pour plus de détails voir [Gry13, Theorem 22.5].

Théorème 1.3 (Dias da Silva–Hamidoune). *Soient p un nombre premier, $A \subset C_p$ un sous-ensemble non-vide de C_p, et $h \in \{1 \ldots |A|\}$, on a :*

$$|\Sigma_h(A)| \geq \min\Big\{p, h(|A|-h)+1\Big\}$$

Nous rappelons maintenant deux lemmes techniques. Le premier lemme ci-dessous, affirme que, sauf dans certains cas particuliers, la raison d'une progression arithmétique dans un groupe cyclique d'ordre premier est déterminée de manière unique à un signe près. Nous incluons une preuve car nous n'avons pas pu trouver de référence appropriée.

Lemme 1.4. *Soient p un nombre premier $p \geq 5$ et A un sous-ensemble de C_p de cardinal k, avec $2 \leq k \leq p-2$. Supposons que A est une progression arithmétique, c'est-à-dire qu'il existe $r \neq 0$ et a dans C_p, tel que $A = \{a + ir : i \in \{0 \dots k-1\}\}$. La raison r est déterminée de manière unique à un signe près, c'est-à-dire s'il existe s et b dans C_p tel que $A = \{b + is : i \in [0, k-1]\}$, alors $s \in \{r, -r\}$.*

Preuve. Comme A est une progression arithmétique de raison r si et seulement si le complément de A dans C_p est aussi une progression arithmétique de raison r, on peut alors supposer que $|A| \leq \frac{p-1}{2}$. Soit e un élément non nul de C_p.

On peut supposer sans perte de généralité que $A = \{0, e, 2e, \dots, (k-1)e\}$ car le problème est invariant par une transformation affine.

Supposons maintenant au vu d'une contradiction $A = \{a + ir : i \in \{0 \dots k-1\}\}$ avec $a, r \in C_p$ et $r \notin \{e, -e\}$. Sans perte de généralité, on peut supposer que $r = r'e$ avec $r' \in [2, \frac{p-1}{2}]$.

Vu que

$$k - 1 < k - 1 + r' \leq \frac{p-1}{2} + \frac{p-1}{2} - 1 = p - 2 < p,$$

il en résulte que $(k-1)e + r \notin A$ et que $(k-1)e$ est aussi le dernier élément de la progression arithmétique A lorsque la raison est représentée avec la raison r. Autrement dit, $(k-1)e = a + (k-1)r$.

Le même raisonnement montre que, lorsque l'on retire l'élément $(k-1)e$ de A alors $(k-2)e$ est le dernier élément de la progression arithmétique $A \setminus \{(k-1)e\}$ à la fois avec la raison r et la raison e. Par conséquent, $(k-2)e + r = (k-1)e$. Ainsi, $r = e$.

□

Il est possible de généraliser ce résultat aux groupes cycliques d'ordre quelconque à condition d'exclure les progressions arithmétiques dont la longueur est égale à d ou $d-1$, où d divise n.

Etant donné une séquence S, lorsqu'on cherche à prouver l'existence d'une sous-séquence T de S de somme nulle et dont la longueur est proche de celle de S, il peut être avantageux de travailler plutôt avec les quelques éléments de la séquence S qui ne sont pas contenus dans la sous-séquence présumée T.

Ce deuxième lemme nous permet de faire le lien exact.

Lemme 1.5. *Soient G un groupe abélien fini et $0 \leq r \leq k$ et S une séquence sans facteur carré de longueur k. Les assertions suivantes sont équivalentes.*

- *La séquence S contient une sous séquence R de longueur r telle que $\sigma(S) = \sigma(R)$.*
- *La séquence S contient sous-séquence T de somme nulle et de longueur $k - r$.*

Preuve. Soit S une séquence sans facteur carré et de longueur k. Soit R une sous-séquence de S et de longueur r avec $\sigma(R) = \sigma(S)$. Dans ce cas la séquence $T = R^{-1}S$ est une séquence de longueur $k - r$ et de somme $\sigma(T) = \sigma(S) - \sigma(R) = 0$. Inversement, supposons que T est une sous-séquence de longueur $k - r$. Alors la séquence $R = T^{-1}S$ est une séquence de longueur $k - (k - r) = r$ et de somme $\sigma(R) = \sigma(S) - \sigma(T) = \sigma(S)$. □

2 Résultat principal

Dans ce chapitre, nous déterminons la valeur de la constante de Harborth pour le groupe $C_3 \oplus C_{3p}$ où p est un nombre premier.

En se basant sur les notations données dans le paragraphe 1, page 7, la définition de la constante de Harborth $\mathrm{g}(G)$ peut être formulée de la façon suivante :

La constante $\mathrm{g}(G)$ est le plus petit entier k tel que toute séquence S sans facteur carré d'éléments de G et de longueur au moins k, contient une sous-séquence de longueur exposant du groupe et de somme nulle, en d'autres termes :

$0 \in \Sigma_{\exp(G)}(S)$

Il s'avère que la borne de Kiefer est généralement atteinte, à une exception près, à savoir pour $p = 3$. Plus précisément, nous allons montrer la valeur de la constante Harborth pour les groupes de la forme $C_3 \oplus C_{3p}$ où p est premier.

Théorème 2.1. *Soit p un nombre premier. On a*

$$\mathrm{g}(C_3 \oplus C_9) = 13$$

et si $p \neq 3$ alors $\mathrm{g}(C_3 \oplus C_{3p}) = 3p + 3$

La preuve utilise divers théorèmes de la combinatoire additive présentés ci-dessus, à savoir les théorèmes de Cauchy–Davenport, de Dias da Silva–Hamidoune, et de Vosper. Ceux-ci sont appliqués aux 'projections' de l'ensemble au sous-groupe C_p de $C_3 \oplus C_{3p}$. C'est une raison pour laquelle nos investigations sont limitées

aux groupes où p est premier. Nous obtenons également certains résultats par le calcul. En particulier, nous confirmons la conjecture de Gao et Thangadurai que nous avons mentionnée ci-dessus pour $C_6 \oplus C_6$.

Nous commençons par établir que les valeurs données dans le théorème sont des minorants. Ensuite nous montrons que ces valeurs sont aussi des majorants de la constante de Harborth. Pour le faire, il faut montrer l'existence des sous-séquences de somme nulle. L'argument est séparé dans plusieurs cas où des hypothèses supplémentaires sur les séquences sont imposées. Enfin, nous combinons tous ces résultats pour achever la démonstration.

a) Un minorant

Dans cette partie, nous déterminons un minorant de la constante de Harborth. Nous commençons par énoncer un lemme général. Un aspect intéressant de ce lemme est qu'il combine des constantes pour les séquences et pour des séquences sans facteur carré. De plus, il améliore le résultat du [Mar+13, Lemma 3.2], où la constante d'Olson a été utilisée au lieu de la constante de Davenport.

Lemme 2.2. *Soient G_1, G_2 deux groupes abéliens finis avec $\exp(G_2) \mid \exp(G_1)$. Alors*

$$\mathrm{g}(G_1 \oplus G_2) \geq \mathrm{g}(G_1) + \mathsf{D}(G_2) - 1.$$

Preuve. Soit S_1 une séquence sans facteur carré de G_1 de longueur $\mathrm{g}(G_1) - 1$ qui ne contient aucune sous-séquence de longueur $\exp(G_1)$ et de somme nulle. Soit S_2' une séquence de G_2 de longueur $\mathsf{D}(G_2) - 1$ qui ne contient aucune sous-séquence de somme nulle. Posons $S_2' = \prod_{g \in G_2} g^{v_g}$. Puisque S_2' ne contient aucune sous-séquence de somme nulle alors pour tout $g \in G_2$ l'ordre de multiplicité v_g vérifie : $v_g < \exp(G_2) \leq \exp(G_1)$. Soient $\{h_1, \ldots, h_{\exp(G_1)-1}\}$ des éléments distincts de G_1, et

$$S_2 = \prod_{g \in G_2} \left(\prod_{i=1}^{v_g} (g + h_i) \right).$$

Alors, S_2 est une séquence sans facteur carré de somme nulle de $G_1 \oplus G_2$. Notons que $S_1 S_2$ est aussi une séquence sans facteur carré de $G_1 \oplus G_2$. Afin de démontrer notre résultat, il suffit de montrer que la séquence $S_1 S_2$ ne contient aucune sous-séquence de longueur $\exp(G_1 \oplus G_2)$ et de somme nulle. Supposons au vu d'une contradiction que T est une sous-séquence de $S_1 S_2$ de longueur $\exp(G_1 \oplus G_2)$ et de somme nulle. Posons $T = T_1 T_2$ avec $T_i \mid S_i$. Comme $\exp(G_1 \oplus G_2) = \exp(G_1)$, il s'ensuit que T n'est pas une sous-séquence de S_1, c'est-à-dire que, T_2 n'est pas la séquence vide. Soit $\pi : G \to G_2$ l'application de projection de G sur G_2 en suivant la direction de G_1 par rapport à la décomposition $G = G_1 \oplus G_2$. Puisque

$\sigma(\pi(T_1)) = 0$, il s'ensuit que $\sigma(\pi(T_2)) = 0$. Ce qui est impossible car, comme $\pi(T_2)$ est une sous-séquence non vide de somme nulle de S_2', alors par hypothèse S_2' ne contient aucune sous-séquence non-vide de somme nulle. □

En utilisant ce lemme et en le combinant avec les résultats sur les constantes pour les groupes cycliques, on obtient la minoration suivante, qui est donnée dans [Baj18], [Proposition F.102].

Lemme 2.3. *Soient n_1 et n_2 des entiers strictement positifs avec $n_1 \mid n_2$. Alors*

$$\mathrm{g}(C_{n_1} \oplus C_{n_2}) \geq \begin{cases} n_1 + n_2 - 1 & \textit{si } n_2 \textit{ est impair} \\ n_1 + n_2 & \textit{si } n_2 \textit{ est pair} \end{cases}.$$

En particulier,

$$\mathrm{g}(C_3 \oplus C_{3n}) \geq \begin{cases} 3n+2 & \textit{si } n \textit{ est impair} \\ 3n+3 & \textit{si } n \textit{ est pair} \end{cases}.$$

Preuve. Le lemme 2.2 donne $\mathrm{g}(C_{n_1} \oplus C_{n_2}) \geq \mathrm{g}(C_{n_2}) + \mathsf{D}(C_{n_1}) - 1$. Cette affirmation s'appuie sur le fait que

$$\mathrm{g}(C_{n_2}) = \begin{cases} n_2 & \text{si } n_2 \text{ est impair} \\ n_2 + 1 & \text{si } n_2 \text{ est pair} \end{cases}$$

et $\mathsf{D}(C_{n_1}) = n_1$ (voir, e.g., [Gry13, Theorem 10.2]). l'affirmation pour $C_3 \oplus C_{3n}$ est une conséquence directe. □

Le minorant pour $\mathrm{g}(C_3 \oplus C_{3n})$ peut être amélioré pour les n nombres impairs. Cela a été fait initialement par Kiefer (voir [Kie16], voir aussi [Baj18, Proposition F.104]).

Nous donnons la preuve, car notre construction est légèrement différente.

Lemme 2.4. *Pour un entier $n \geq 2$, soit G le groupe de la forme $G = C_3 \oplus C_{3n}$. Alors $\mathrm{g}(G) \geq 3n + 3$.*

Preuve. Pour un entier n pair, l'inégalité est connue par le lemme 2.3 ci-dessus. Ainsi, il suffit de la prouver pour n impair.

Pour prouver ce lemme, il suffit d'exhiber une séquence sans facteur carré de longueur $3n + 2$ et qui ne contient pas de sous-séquence de longueur $\exp(G) = 3n$ et de somme nulle.

Soit $G = \langle e_1 \rangle \oplus \langle e_2 \rangle$ avec $\mathrm{ord}(e_1) = 3$ and $\mathrm{ord}(e_2) = 3n$.

Soit π_1 la projection sur $\langle e_1 \rangle$ en suivant la direction de $\langle e_2 \rangle$:

$$\pi_1 : G = \langle e_1 \rangle \oplus \langle e_2 \rangle \to \langle e_1 \rangle$$

et soit π_2 la projection sur $\langle e_2\rangle$ en suivant la direction de $\langle e_1\rangle$:

$$\pi_2 : G = \langle e_1\rangle \oplus \langle e_2\rangle \to \langle e_2\rangle$$

De plus, soit

$$T_1 = \prod_{g\in\langle e_2\rangle\setminus\{0,-e_2,e_2\}} (e_1 + g)$$

et $T_2 = 0(e_2)(2e_2)(3e_2)(-6e_2)$. Alors $T = T_1T_2$ est une séquence sans facteur carré de longueur $|T| = 3n - 3 + 5 = 3n + 2$.

Pour obtenir l'inégalité souhaitée, il suffit d'affirmer que la séquence T ne contient pas de sous-séquence de longueur $3n$ et de somme nulle.

Supposons au vu d'une contradiction que T contient une sous-séquence R de longueur $3n$ et de somme nulle. Il est clair que l'on a $\sigma(\pi_1(R)) = \sigma(\pi_2(R)) = 0$.

Soit $R = R_1R_2$ avec $R_1|T_1$ et $R_2|T_2$. On a $|R| = 3n$ et $|T_1| = 3n - 3$ et $|T_2| = 5$ d'où $|R_2| \geq 3$ et $3n - 5 \leq |R_1| \leq 3n - 3$.

Notons que $\sigma(\pi_1(R)) = \sigma(\pi_1(R_1)) = |R_1|e_1$.

Par conséquent, comme $\sigma(\pi_1(R)) = 0$ il est nécessaire que 3 divise $|R_1|$. Il s'ensuit que $|R_1| = 3n - 3$, c'est-à-dire $R_1 = T_1$. Il s'ensuit que $|R_2| = 3$.

Maintenant, $\sigma(\pi_1(R_1)) = |R_1|e_1 = 0$. En outre,

$$\sigma(\pi_2(R_1)) = \sum_{h\in\langle e_2\rangle\setminus\{-e_2,e_2,0\}} h = \Big(\sum_{h\in\langle e_2\rangle} h\Big) - (-e_2 + e_2 + 0),$$

qui est également égal à 0, puisque la somme de tous les éléments du groupe cyclique $\langle e_2\rangle$ est nulle du fait que $3n$ est impair.

Ainsi, $\sigma(R_1) = 0$, et il s'ensuit que : $\sigma(R) = 0$ si et seulement si $\sigma(R_2) = 0$. Cependant, T_2 ne contient pas de sous-séquence de longueur 3 et de somme nulle.

Ainsi, T ne contient pas de sous-séquence de longueur $3n$ et de somme nulle.

□

Il s'avère que pour $p = 3$, il y a une meilleure construction.

Lemme 2.5. *Pour le groupe $G = C_3 \oplus C_9$ on a $\mathrm{g}(G) \geq 13$*

Preuve. Pour prouver ce lemme, comme $\exp(G) = 9$, il suffit de donner l'exemple d'une séquence T sans facteur carré sur G de longueur 12 et qui n'admet aucune sous-séquence T_1 de longueur $\exp(G) = 9$ et de somme nulle.

Soit $G = \langle e_1\rangle \oplus \langle e_2\rangle$ avec $\mathrm{ord}(e_1) = 3$, et $\mathrm{ord}(e_2) = 9$.

Considérons la séquence suivante :

$$T = R(e_1 + R)(e_2 + R)(e_1 + e_2 + R), \text{ avec } R = 0(3e_2)(6e_2),$$

C'est une séquence sans facteur carré de longueur 12 qui satisfait $\sigma(T) = 0 + 0 + 3e_2 + 3e_2 = 6e_2$. Par le lemme 1.5 avec $k = 12$ et $r = 9$, la séquence T contient une sous-séquence de longueur 9 et de somme nulle si et seulement si T contient une sous-séquence T_2 avec $|T_2| = 3 = 12 - 9$ et $\sigma(T) = \sigma(T_2) = 6e_2$.

En vue d'une contradiction, supposons qu'une telle sous-séquence T_2 existe.

Soit $H = \{0, 3e_2, 6e_2\}$ et soit l'épimorphisme $\pi : G \to G/H$.

On a $G/H \cong C_3 \oplus C_3$, et ce groupe est engendré par $f_1 = \pi(e_1)$ et $f_2 = \pi(e_2)$.

Puisque $\sigma(T_2) = 6e_2$, on a que $\pi(T_2)$ est une sous-séquence de somme nulle de $\pi(T)$ et $\pi(\sigma(T)) = \pi(6e_2) = 0$.

Mais, notons que les seules sous-séquences de $\pi(T) = 0^3 {f_1}^3 {f_2}^3 (f_1 + f_2)^3$ de longueur 3 qui ont la somme nulle sont 0^3, ${f_1}^3$, ${f_2}^3$ et $(f_1 + f_2)^3$. Il reste à vérifier si l'une des sous-séquences correspondantes de T contient une somme égale à $6e_2$. Ce qui n'est pas le cas. Concrètement, nous avons $\sigma(R) = 0$, $\sigma(e_1 + R) = 0$, $\sigma(e_2 + R) = 3e_2$, et $\sigma((e_1 + e_2) + R) = 3e_2$. Ainsi, la séquence T ne contient pas de sous-séquence de longueur 3 avec une somme $6e_2$. Ceci établit l'inégalité.

□

b) L'existence d'une sous-séquence avec les conditions souhaitées

Dans cette partie, nous allons établir l'existence d'une sous-séquence de longueur $\exp(G)$ et de somme nulle sous différentes hypothèses.

Mettons au point quelques notations qui seront utilisées tout au long de la sous-section. Pour un nombre premier $p \neq 3$ soit $G = C_3 \oplus C_{3p}$. Notons que $G = H_1 \oplus H_2$ où $H_1 \cong C_3^2$ est le sous-groupe des éléments dont l'ordre divise 3 et $H_2 \cong C_p$ est le sous-groupe des éléments dont l'ordre divise p. Pour $i \in \{1, 2\}$, soit $\pi_i : G \to H_i$ la projection de G sur H_i par rapport à la décomposition $G = H_1 \oplus H_2$.

Pour une séquence S de G, il existe une unique décomposition $S = \prod_{h \in H_1} S_h$ où S_h est la sous-séquence des éléments de S avec $\pi_1(g) = h$.

Si S est une séquence sans facteur carré, alors pour chaque élément $h \in H_1$ la séquence $\pi_2(S_h)$ est une séquence sans facteur carré sur H_2.

Pour établir le majorant $\mathrm{g}(G) \leq 3p + 3$, nous devons montrer que chaque séquence sans facteur carré S de longueur $3p + 3$ sur G contient une sous-séquence de longueur $3p$ de somme nulle. Par le lemme 1.5 cela équivaut à montrer que chaque

séquence sans facteurs carrés S de longueur $3p+3$ sur G contient une sous-séquence R de longueur 3 de somme égale à celle de la séquence S.

Pour obtenir une telle séquence de longueur 3, nous nous limitons d'abord à la recherche d'une sous-séquence pour laquelle $\pi_1(\sigma(S)) = \pi_1(\sigma(R))$; cette condition peut être établie par des arguments explicites, car le groupe H_1 est fixé et petit.

Ensuite, en utilisant les outils de la combinatoire additive rappelés dans la sous-section b), nous montrons que parmi les sous-séquences R de S telles que $\pi_1(\sigma(S)) = \pi_1(\sigma(R))$ il y en a une pour laquelle nous avons aussi $\pi_2(\sigma(S)) = \pi_2(\sigma(R))$ et donc qui satisfait $\sigma(S) = \sigma(R)$ selon les besoins.

Nous formulons un lemme technique qui est un outil clé dans notre argumentation. Notez que pour la preuve de ce lemme, il est crucial que p soit premier. Plus tard, par exemple dans la proposition 2.8, le cas des petits nombres premiers crée une difficulté supplémentaire dans la preuve. Comme nous pouvons traiter le cas $p=2$ par le calcul, nous excluons celui-ci immédiatement pour éviter de considérer les cas particuliers.

Lemme 2.6. *Soit $p \geq 5$ un nombre premier et soit S une séquence sans facteur carré de longueur $3p+3$ sur $G = C_3 \oplus C_{3p}$. Soit $S = \prod_{h \in H_1} S_h$ où S_h est la sous-séquence des éléments G de S tel que $\pi_1(g) = h$.*

1. *S'il existe des éléments distincts x, y, z dans H_1 avec $x+y+z = \pi_1(\sigma(S))$ tels que S_x, S_y, S_z sont tous non vides et $|S_x| + |S_y| + |S_z| - 2 \geq p$, alors S contient une sous-séquence de longueur $3p$ et de somme nulle. .*
2. *S'il existe des éléments distinct x, y dans H_1 avec $2x + y = \pi_1(\sigma(S))$ tels que $|S_x| \geq 2$ et $|S_y| \geq 1$ et $2|S_x| + |S_y| - 4 \geq p$, alors S contient une sous-séquence de longueur $3p$ et de somme nulle .*
3. *S'il existe $x \in H_1$ avec $3x = \pi_1(\sigma(S))$ tel que $|S_x| \geq 3$ et $3|S_x| - 8 \geq p$, alors S contient une sous-séquence de longueur de $3p$ et de somme nulle.*

Nous utilisons et combinons les théorèmes de Cauchy–Davenport et de Dias da Silva–Hamidoune.

Preuve. Pour chaque cas, nous montrons que sous les hypothèses du lemme, S contient une sous-séquence R de longueur 3 avec une somme égale à celle de la séquence S. Par le lemme 1.5 avec $k = 3p+3$ et $r = 3$ cela établit notre résultat.

(1). Soient x, y, z des éléments distincts de H_1 avec $x+y+z = \pi_1(\sigma(S))$. Si $g_x, g_y, g_z \in G$ sont tels que g_x divise S_x, g_y divise S_y et g_z divise S_z, alors $g_x g_y g_z$ est une sous-séquence de S et $\pi_1(\sigma(g_x g_y g_z)) = x+y+z = \pi_1(\sigma(S))$.

Ainsi, pour montrer que S contient une sous-séquence R de longueur 3 il suffit de montrer qu'il existe des éléments g_x, g_y et g_z dans G tels que g_x divise S_x, g_y divise S_y et g_z divise S_z avec $\pi_2(\sigma(g_xg_yg_z)) = \pi_2(\sigma(S))$.

Soit Ω l'ensemble de toutes les séquences $g_xg_yg_z$ de longueur 3 avec g_x divise S_x, g_y divise S_y et g_z divise S_z. Nous observons que

$$\Big\{\pi_2(\sigma(R))\colon R\in\Omega\Big\} = \operatorname{supp}(\pi_2(S_x)) + \operatorname{supp}(\pi_2(S_y)) + \operatorname{supp}(\pi_2(S_z)).$$

D'après le théorème de Cauchy-Davenport on a :

$$|\operatorname{supp}(\pi_2(S_x)) + \operatorname{supp}(\pi_2(S_y)) + \operatorname{supp}(\pi_2(S_z))| \geq$$

$$\min\Big\{p, |\operatorname{supp}(\pi_2(S_x))| + |\operatorname{supp}(\pi_2(S_y))| + |\operatorname{supp}(\pi_2(S_z))| - 2\Big\}$$

Comme la séquence S est sans facteur carré, pour chaque $h\in H_1$, la séquence $\pi_2(S_h)$ est également sans facteur carré. Par conséquent, $|\operatorname{supp}(\pi_2(S_h))| = |S_h|$. Ainsi, si $|S_x|+|S_y|+|S_z|-2\geq p$, alors $\operatorname{supp}(\pi_2(S_x)) + \operatorname{supp}(\pi_2(S_y)) + \operatorname{supp}(\pi_2(S_z))$ doit être égale au groupe entier H_2. En particulier, il existe une séquence $R\in\Omega$ avec $\pi_2(\sigma(R)) = \pi_2(\sigma(S))$, ce qui achève la preuve.

(2). L'argument est similaire à celui de la première partie, mais de plus, nous devons utiliser le théorème de Dias da Silva–Hamidoune. Soient x, y des éléments distincts de H_1 avec $2x+y = \pi_1(\sigma(S))$. Si $g_xg'_x \mid S_x$ et $g_y \mid S_y$, alors $g_xg'_xg_y$ est une sous-séquence de S et $\pi_1(\sigma(g_xg'_xg_y)) = 2x+y = \pi_1(\sigma(S))$.

Soit Ω l'ensemble de toute la séquence $g_xg'_xg_y$ avec $g_xg'_x$ divise S_x et g_y divise S_y. On note que $\{\pi_2(\sigma(R))\colon R\in\Omega\} = \Sigma_2(\operatorname{supp}(\pi_2(S_x)) + \operatorname{supp}(\pi_2(S_y))$.

Selon les théorèmes de Dias da Silva–Hamidoune et Cauchy–Davenport (voir les théorèmes 1.3 et 1.1), on obtient, comme p est supposé premier :

$$\Big|\Sigma_2\Big(\operatorname{supp}(\pi_2(S_x))\Big) + \operatorname{supp}\Big(\pi_2(S_y)\Big)\Big| \geq$$
$$\min\Big\{p, 2|\operatorname{supp}(\pi_2(S_x))| + |\operatorname{supp}(\pi_2(S_y))| - 4\Big\}.$$

Comme dans (1), si $2|S_x|+|S_y|-4\geq p$, alors il existe des $R\in\Omega$ avec $\pi_2(\sigma(R)) = \pi_2(\sigma(S))$, ce qui achève la preuve.

(3). Soit $x\in H_1$ avec $3x = \pi_1(\sigma(S))$. Soit $g_xg'_xg''_x$ une sous-séquence de S_x de longueur 3. On a $g_xg'_xg''_x$ est une sous-séquence de S avec $\pi_1(\sigma(g_xg'_xg''_x)) = 3x = \pi_1(\sigma(S))$.

Soit Ω l'ensemble de toutes les séquences $g_xg'_xg''_x$ de S_x. On note que

$$\{\pi_2(\sigma(R))\colon R\in\Omega\} = \Sigma_3\Big(\operatorname{supp}(\pi_2(S_x))\Big)$$

.

De même, par le théorème de Dias da Silva–Hamidoune (voir le théorème 1.3), on obtient :

$$\Big|\Sigma_3\Big(\mathrm{supp}(\pi_2(S_x))\Big)\Big| \geq \min\Big\{p, 3|\mathrm{supp}(\pi_2(S_x))| - 8\Big\}.$$

Comme dans (1), si $3|S_x| - 8 \geq p$, alors il existe une sous-séquence $R \in \Omega$ telle que $\pi_2(\sigma(R)) = \pi_2(\sigma(S))$. Ce qui achève la preuve. □

Dans le présent contexte, il existe essentiellement deux types de séquences S sur G de longueur $3p+3$: celles pour lesquelles $\pi_1(\sigma(S))$ est égal à zéro et celles pour lesquelles $\pi_1(\sigma(S))$ est non nul. Il est clair que cette propriété est préservée en appliquant un automorphisme du groupe, et lorsque la longueur de la séquence est un multiple de 3, elle est également préservée par translation. Nous traitons ces deux types de séquences séparément. Dans les deux cas, il sera pertinent d'inclure les sous-séquences de longueurs trois de $\pi_1(S)$ qui ont une somme égale à $\pi_1(\sigma(S))$. Dans le premier cas, la séquence est formée par les trois éléments d'une classe modulo H_1' qui ne contient qu'un seul élément de multiplicité trois.

Ce dernier type est traité dans la proposition 2.10. Pour le premier type, on distingue deux cas : le cas où le support de $\pi_1(S)$ est le groupe H_1 tout entier (voir la proposition 2.7) et les cas où il ne l'est pas (voir la proposition 2.8).

Proposition 2.7. *Soit $p \geq 5$ un nombre premier et soit S une séquence sans facteur carré de longueur $3p+3$ sur $G = C_3 \oplus C_{3p}$. Si $\sigma(\pi_1(S)) = 0$ et $\mathrm{supp}(\pi_1(S)) = H_1$, alors S contient une sous-séquence de longueur $3p$ et de somme nulle.*

Preuve. Pour clarifier les explications qui suivent, remarquons que nous pouvons supposer que $\sigma(\pi_2(S)) = 0$ (et donc $\sigma(S) = 0$).

En effet, il suffit de noter que si pour un $h \in G$, la séquence translatée $h + S$ contient une sous-séquence de longueur $3p$ et de somme nulle, alors la séquence S contient une sous-séquence de longueur $3p$ et de somme nulle.

Il existe $h' \in H_2$ tel que $(3p+3)h' = -\sigma(\pi_2(S))$; notons que comme p et $3p+3$ sont premiers entre eux, la multiplication $h \mapsto (3p+3)h$ est un automorphisme de H_2. Maintenant, on peut considérer $h' + S$ au lieu de S à condition que la condition supplémentaire $\sigma(\pi_1(S)) = 0$ ne soit pas modifiée. Puisque $\sigma(\pi_1(h' + S)) = |S|h' + \sigma(\pi_1(S))$ et $|S|h' = (3p+3)h' = 0$, c'est effectivement vrai et l'assertion $\mathrm{supp}(\pi_1(h' + S)) = H_1$ reste satisfaite.

Soit H_1' un sous-groupe cyclique non trivial de H_1 et soit $g \in H_1$, et $\{x, y, z\} = g + H_1'$ une classe modulo H_1'. Puisque $x + y + z = 0 = \sigma(\pi_1(S))$, il s'ensuit que

si $|S_xS_yS_z| - 2 \geq p$, alors à partir de la partie (1) du lemme 2.6, le résultat reste satisfait.

Il reste à considérer le cas où pour chaque classe modulo H_1 qui est de cardinal trois, noté $\{x, y, z\}$, on a $|S_xS_yS_z| \leq p + 1$.

Notons que cela n'est possible que si pour chaque classe $\{x, y, z\}$ d'équivalence modulo H_1, on a $|S_xS_yS_z| = p + 1$. En effet, H_1 peut être décomposer comme union disjointe de trois de ces classes, disons, $H_1 = \{x_1, y_1, z_1\} \cup \{x_2, y_2, z_2\} \cup \{x_3, y_3, z_3\}$. Ensuite, d'une part

$$|S_{x_i}S_{y_i}S_{z_i}| \leq p + 1, \text{ pour tout } i \in \{1, 2, 3\}$$

et d'autre part $|S_{x_1}S_{y_1}S_{z_1}| + |S_{x_2}S_{y_2}S_{z_2}| + |S_{x_3}S_{y_3}S_{z_3}| = |S| = 3p + 3$. Il est donc nécessaire que pour chaque $i \in \{1, 2, 3\}$ on ait $|S_{x_i}S_{y_i}S_{z_i}| = p + 1$.

Ensuite, nous affirmons que cela n'est possible que si chacune des 9 séquences est de même longueur.

Soit $H_1 = \{q_1, q_2, \cdots, q_9\}$ tel que $|S_{q_1}| \geq |S_{q_2}| \geq \cdots \geq |S_{q_9}|$. Soit $v_i = |S_{q_i}|$. Il existe $j \in \{3 \ldots 9\}$ tel que $\{q_1, q_2, q_j\}$ est une classe d'équivalence modulo H_1', notamment c'est le cas pour $q_j = 2q_2 - q_1$, et il existe un entier $i \in \{1 \ldots 7\}$ tel que $\{q_8, q_9, q_i\}$ est une classe d'équivalence modulo H_1.

On a donc $v_1 + v_2 + v_j = p + 1$ et $v_8 + v_9 + v_i = p + 1$.

Il s'ensuit que $(v_1 - v_9) + (v_2 - v_8) + (v_j - v_i) = 0$, et donc $(v_1 - v_9) + (v_2 - v_8) = v_i - v_j$. Pourtant, comme $(v_1 - v_9) \geq (v_i - v_j)$ nous obtenons que $v_2 - v_8 = 0$. Par conséquent, on a $v_2 = v_8$ et de plus $v_2 = v_3 = \cdots = v_8 = v$. Puisque il y a une classe d'équivalence $\{q_i, q_j, q_k\}$ avec i, j, k dans $[2, 8]$ on a $3v = p + 1$. Donc, cette valeur commune doit être $\frac{p+1}{3}$. Il reste à montrer que $v_1 = v$ et $v_9 = v$. Il existe une classe d'équivalence modulo H_1 de cardinal 3 qui contient q_9 et qui ne contient pas q_1, donc $v_9 + 2v = p + 1$ et donc $v_9 = \frac{p+1}{3}$. De la même manière, on obtient que $v_1 = v$.

Nous reconsidérons maintenant, pour une classe $\{x, y, z\}$ d'équivalence modulo H_1, le cardinal de l'ensemble

$$\text{supp}(\pi_2(S_x)) + \text{supp}(\pi_2(S_y)) + \text{supp}(\pi_2(S_z)).$$

Par le théorème de Cauchy-Davenport, on a

$$\begin{aligned}
&|\text{supp}(\pi_2(S_x)) + \text{supp}(\pi_2(S_y)) + \text{supp}(\pi_2(S_z))| \geq \\
&\min\{p, |\text{supp}(\pi_2(S_x)) + \text{supp}(\pi_2(S_y))| + |\text{supp}(\pi_2(S_z))| - 1\} \geq \\
&\min\{p, \min\{p, |\text{supp}(\pi_2(S_x))| + |\text{supp}(\pi_2(S_y))| - 1\} + |\text{supp}(\pi_2(S_z))| - 1\}.
\end{aligned}$$

Cela permet de simplifier :

$$\min\{p, |\mathrm{supp}(\pi_2(S_x))| + |\mathrm{supp}(\pi_2(S_y))| + |\mathrm{supp}(\pi_2(S_z))| - 2\} =$$
$$\min\{p, |S_x| + |S_y| + |S_z| - 2\} = p - 1.$$

Si l'on sait que $|\mathrm{supp}(\pi_2(S_x)) + \mathrm{supp}(\pi_2(S_y)) + \mathrm{supp}(\pi_2(S_z))| \geq p$, alors

$$\mathrm{supp}(\pi_2(S_x)) + \mathrm{supp}(\pi_2(S_y)) + \mathrm{supp}(\pi_2(S_z)) = H_2,$$

et nous pouvons conclure comme dans le lemme 2.6.

Ainsi, il reste à considérer le cas où $|\mathrm{supp}(\pi_2(S_x)) + \mathrm{supp}(\pi_2(S_y)) + \mathrm{supp}(\pi_2(S_z))| = p - 1$. Cela n'est possible que lorsqu'on a :

$$|\mathrm{supp}(\pi_2(S_x)) + \mathrm{supp}(\pi_2(S_y))| = |\mathrm{supp}(\pi_2(S_x))| + |\mathrm{supp}(\pi_2(S_y))| - 1,$$

car sinon la deuxième inégalité de l'équation ci-dessus serait stricte. De la même manière, on obtient que $|\mathrm{supp}(\pi_2(S_x)) + \mathrm{supp}(\pi_2(S_z))| = |\mathrm{supp}(\pi_2(S_x))| + |\mathrm{supp}(\pi_2(S_z))| - 1$.

Le théorème de Vosper (voir le théorème 1.2) permet d'annoncer que $\mathrm{supp}(\pi_2(S_x))$ et $\mathrm{supp}(\pi_2(S_y))$ sont des progressions arithmétiques de même raison, et les ensembles $\mathrm{supp}(\pi_2(S_x))$ et $\mathrm{supp}(\pi_2(S_z))$ sont des progressions arithmétiques de même raison. D'après le lemme 1.4, il s'ensuit qu'il existe une raison commune aux trois progressions $\mathrm{supp}(\pi_2(S_x))$, $\mathrm{supp}(\pi_2(S_y))$, $\mathrm{supp}(\pi_2(S_z))$. En effet, tous les 9 ensembles $\mathrm{supp}(\pi_2(S_h))$ où $h \in H_1$ sont des progressions arithmétiques de même raison. Puisque, on peut appliquer l'argument pour n'importe quelle paire de ces 9 ensembles. Ainsi, Appelons cette raison e ; bien sûr, il s'agit d'un élément générateur de H_2.

S'il existe $h \in H_1$ tel que $\pi_2(S_h)$ contient une sous-séquence de longueur 3 et de somme nulle, alors S_h contient une séquence de longueur 3 et de somme nulle. Puisque nous avons supposé au départ que $\sigma(S) = 0$, en invoquant le lemme 1.5 notre argument permet d'aboutir au résultat.

Ainsi, nous supposons que pour aucun des éléments h de H_1 la séquence $\pi_2(S_h)$ contient une sous-séquence de longueur 3 et de somme nulle. En particulier, $\pi_2(S_h)$ n'a pas $(-e)e0$, comme sous-séquence. Ainsi, pour tout h de H_1 on a $\pi_2(S_h)$ a la forme suivante $\pi_2(S_h) = \prod_{j=s_h}^{s_h+\frac{p-2}{3}} (je)$ où s_h est un entier tel que $0 \leq s_h \leq s_h + \frac{p-2}{3} < p - 1$. Il est facile d'observer que :

$$\Sigma_3(\pi_2(S_h)) = \Big\{je \colon j \in [3s_h + 3, 3s_h - 3 + (p-2)]\Big\}.$$

Pour que cet ensemble ne contienne pas 0, nous avons besoin que l'entier s_h satisfasse $3s_h - 3 + (p-2) < p$. Donc, que $3s_h < 5$, c'est-à-dire $s_h \in \{0, 1\}$.

S'il existe une classe $\{x,y,z\}$ tel que $s_x = s_y = s_z = 0$, alors clairement $\pi_2(S_x) + \pi_2(S_y) + \pi_2(S_z)$ contient 0. Mais s'il n'y a pas de classe $\{x,y,z\}$ tel que $s_x = s_y = s_z = 0$, alors il y a une classe $\{x',y',z'\}$ tel que $s_{x'} + s_{y'} + s_{z'} \geq 2$; en effet, il suffit de noter que par la première condition il doit y avoir au moins deux éléments $h, h' \in H_1$ avec $s_h \geq 1$ et $s_{h'} \geq 1$. Cependant, cela donne que l'ensemble $\pi_2(S_{x'}) + \pi_2(S_{y'}) + \pi_2(S_{z'})$ contient l'élément $(2 \cdot \frac{p+1}{3} + \frac{p-2}{3})e = pe = 0$. Ainsi, la preuve est établie. □

Pour le résultat suivant, nous gardons la condition que $\sigma(\pi_1(S)) = 0$, mais considérons plutôt le cas où $\mathrm{supp}(\pi_1(S)) \neq H_1$.

Proposition 2.8. *Soit $p \geq 5$ un nombre premier et soit S une séquence sans facteur carré de longueur $3p+3$ sur le groupe $G = C_3 \oplus C_{3p}$. Si $\sigma(\pi_1(S)) = 0$ et $\mathrm{supp}(\pi_1(S)) \neq H_1$, alors S contient une sous-séquence de longueur $3p$ et de somme nulle.*

Preuve. Soit $h \in H_1$ tel que $|S_h| = 0$; un tel élément h existe par hypothèse. Maintenant, comme rappelé dans la section b), la séquence S contient une sous-séquence T de longueur $3p$ et de somme nulle ; donc la séquence $S - h$ contient $T - h$ comme une sous-séquence de longueur $3p$ et de somme nulle. Puisque $\mathrm{supp}(\pi_1(-h+S)) = -h + \mathrm{supp}(\pi_1(S))$ il s'ensuit de $h \notin \mathrm{supp}(\pi_1(S))$ que $0 \notin \mathrm{supp}(\pi_1(-h+S))$.

Puisque $\sigma(\pi_1(-h+S)) = |S|(-h) + \sigma(\pi_1(S))$ et puisque $|S|h = (3p+3)h = 0$, on peut considérer $-h+S$ au lieu de S, car la condition supplémentaire $\sigma(\pi_1(S)) = 0$ reste satisfaite. Ainsi, par translation, on peut supposer sans perte de généralité que $|S_0| = 0$.

Maintenant, nous distinguons les cas selon le cardinal de $\mathrm{supp}(\pi_1(S))$. Par hypothèse, elle est strictement inférieure à $|H_1| = 9$.

Supposons que $|\mathrm{supp}(\pi_1(S))| = 8$. Nous constatons que H_1 a exactement 8 classes de cardinal 3 qui ne contiennent pas 0. Chaque élément non nul est contenu dans exactement 3 d'entre elles. Il existe donc une classe $\{x,y,z\}$ telle que $|S_x S_y S_z| \geq \frac{3}{8}|S| = \frac{9p+9}{8}$. L'existence de la sous-séquence requise découle alors de la partie (1) du lemme 2.6 en notant que $(9p+9)/8 > p+1$.

Supposons que $|\mathrm{supp}(\pi_1(S))| = 7$. Soit $-x \in H_1$ l'élément non nul tel que $|S_{-x}| = 0$. On note qu'il y a 4 classes de cardinal 3 qui contiennent x, et que 3 d'entre elles ne contiennent ni $-x$ ni 0. Il s'ensuit donc qu'il existe une classe $\{x,y,z\}$ tel que $|S_x S_y S_z| \geq |S_x| + \frac{1}{3}(|S_x^{-1} S|) = (p+1) + \frac{2|S_x|}{3} > p+1$. En utilisant à nouveau la partie (1) du lemme 2.6, l'existence de la sous-séquence apparaît.

Supposons que $|\mathrm{supp}(\pi_1(S))| = 6$. Soient deux éléments non nuls $g, h \in H_1$ tels que $|S_g| = |S_h| = 0$. Si $g = -h$, alors il y a une classe $\{x, y, z\}$ modulo le sous-groupe $\{0, g, -g\}$ telle que $|S_x S_y S_z| \geq \frac{1}{2}|S| = \frac{3p+3}{2} > p + 1$ (notons que les deux classes autres que $\{0, g, -g\}$ recouvrent elles-mêmes les 6 éléments restants de H_1). Là encore, la partie (1) du lemme 2.6 conduit au résultat.

Si $g \neq -h$, alors pour l'une des 2 classes $\{x, y, z\} = \{-g, -g + h, -g - h\}$ ou $\{x, y, z\} = \{-h, -h + g, -h - g\}$, nous avons :

$$|S_x S_y S_z| \geq |S_{-g-h}| + \tfrac{1}{2}|(S_{g+h} S_{-g-h})^{-1} S| = (3p + 3 - |S_{g+h}| + |S_{-g-h}|)/2$$

notons que ces deux classes contiennent $-g - h$, et que l'union des deux classes contient tous les éléments de $H_1 \setminus \{0, g, h\}$ à l'exception de l'élément $g + h$. Comme $|S_{g+h}| \leq p$, il s'ensuit que $|S_x S_y S_z| > p + 1$. De nouveau la partie (1) du lemme 2.6 permet de conclure.

Supposons que $|\mathrm{supp}(\pi_1(S))| \leq 5$. Dans ce cas, il existe $h \in H_1$ tel que $|S_h| \geq \frac{1}{5}|S| = \frac{3p+3}{5}$. Si $p > 5$, en appliquant la partie (3) du lemme 2.6 à S_h, alors on peut achever la preuve ; pour $p \geq 11$ c'est immédiat et pour $p = 7$ on observe qu'on a $|S_h| \geq 5$.

Il reste à considérer le cas particulier $p = 5$. La partie (3) du lemme 2.6 peut être appliquée s'il existe $h \in H_1$ avec $|S_h| = 5$. Supposons donc que $|S_h| \leq 4$ pour tous les $h \in H_1$. Cela implique $|\mathrm{supp}(\pi_1(S))| = 5$ car sinon il existerait quelques $h \in H_1$ avec $|S_h| \geq 18/4 > 4$.

Soit $\{h_1, \ldots, h_5\} \subset H_1$ tel que $|S_{h_i}| \neq 0$ pour chaque entier $i \in \{1 \ldots 5\}$. Puisque $\mathsf{g}(C_3^2) = 5$, comme rappelé dans l'introduction, il existe des entiers $i, j, k \in \{1 \ldots 5\}$ distincts tels que $h_i + h_j + h_k = 0$. Or, $|S_{h_i} S_{h_j} S_{h_k}| = |S| - 2\max\{|S_h| \colon h \in H_1\} \geq 18 - 2 \cdot 4 = 10$. Là encore, nous pouvons appliquer la partie (1) du lemme 2.6 pour établir la preuve. □

Remarque 2.9. Lorsque l'entier p est assez grand, une preuve plus courte est possible. Il existe quelques $h \in H_1$ tels que $|S_h| \geq \frac{1}{8}|S| = \frac{3p+3}{8}$. Nous pouvons appliquer la partie (3) du lemme 2.6 si 3 $|S_h| - 8 \geq p$. Cela est vrai à condition qu'on ait : $3 \cdot \frac{3p+3}{8} - 8 \geq p$, ce qui équivaut à $\frac{p}{8} - \frac{55}{8} \geq 0$. Donc pour $p \geq 55$ nous pouvons établir la preuve de cette façon.

Nous passons maintenant au cas où $\sigma(\pi_1(S)) \neq 0$.

Proposition 2.10. *Soit p un nombre premier supérieur ou égal à 5 et soit S une séquence sans facteur carré de longueur $3p + 3$ sur le groupe $G = C_3 \oplus C_{3p}$. Si $\sigma(\pi_1(S)) \neq 0$, alors S contient une sous-séquence de longueur de $3p$ et de somme nulle.*

Preuve. Soit $c = \sigma(\pi_1(S))$. Sans perte de généralité, on peut supposer que pour h de H_1, l'entier $|S_c|$ est maximal parmi tous les $|S_h|$. L'argument est le même que dans la preuve de la proposition 2.8.

Notons que $\frac{|S|}{9} = \frac{3p+3}{9} = \frac{p+1}{3}$ et donc $|S_c| \geq \frac{p+1}{3}$. Montrons que $|S_c| > \frac{p+1}{3}$. Si pour chaque $h \in H_1$ on a $|S_h| = \frac{p+1}{3}$, alors $\sigma(\pi_1(S)) = \frac{p+1}{3} \sum_{h \in H_1} h$. Or, $\sum_{h \in H_1} h = 0$, contredit $\sigma(\pi_1(S)) = c \neq 0$. Donc $|S_h| \neq \frac{p+1}{3}$ pour $h \in H_1$, et donc $|S_c| > \frac{p+1}{3}$.

Le principe de la preuve consiste de nouveau à appliquer le lemme 2.6. Pour cela, il faut trouver une sous-séquence de $\pi_1(S)$ de longueur 3 de somme c.

Une possibilité est de considérer une telle sous-séquence formée seulement par des éléments du sous-groupe cyclique de $C = \{-c, 0, c\}$. Ainsi, les sous-séquences de longueur 3 et de somme c de ce sous-groupe sont : 0^2c, $(-c)^20$ et $c^2(-c)$. Cette approche s'applique si ce sous-groupe contient un nombre suffisant d'éléments de la séquence S. Ce point est détaillé dans le cas 1 ci-dessous.

Une autre possibilité consiste à considérer des sous-séquences de la forme $ch(-h)$ avec $h \notin C$. Bien qu'elle ne soit pas formulée explicitement dans ce document, la distinction des sous-cas dans le cas 2 correspond au nombre de sous-séquences distinctes de la forme $\pi_1(S)$ compté sans multiplicité.

Soit $v_C = |S_0 S_c S_{-c}|$.

Cas 1 : $v_C \geq p + 3$. Si $|S_{-c}| = 0$, alors $|S_0| + |S_c| = v_C \geq p + 3$. Ainsi, comme $|S_c| \leq p$, nous avons $|S_0| \geq 3$. Donc $|S_c| + 2|S_0| = |S_c| + |S_0| + |S_0| \geq p + 3 + 2 = p + 5$. Puisque $|S_c| > \frac{p+1}{3} \geq 2$, la partie (2) du lemme 2.6 appliqué avec $x = 0$ et $y = c$, permet de conclure.

Si $|S_0| \leq 1$, alors $|S_{-c}| \geq 2$ et donc $2 \cdot |S_c| + |S_{-c}| \geq v_C - 1 + |S_c| \geq p + 4$. Le résultat découle de la partie (2) du lemme 2.6, appliquée avec $x = c$ et $y = -c$.

Si $|S_{-c}| \geq 1$ et $|S_0| \geq 2$, et que l'une des affirmations $|S_c| + 2|S_0| \geq p + 4$ ou $2|S_c| + |S_{-c}| \geq p + 4$ est vraie, alors le résultat découle de la partie (2) du lemme 2.6 ; notons que $c + 2 \cdot 0 = c$ et $2c + (-c) = c$. Ainsi, supposons $|S_c| + 2|S_0| \leq p + 3$ et $2|S_c| + |S_{-c}| \leq p + 3$.

En additionnant ces deux inégalités, il s'ensuit que $3|S_c| + 2|S_0| + |S_{-c}| \leq 2p + 6$. Puisque $v_C = |S_c| + |S_0| + |S_{-c}| \geq p + 3$, il s'ensuit que $v_C = p + 3$ et $|S_c| = |S_{-c}|$. Puisque $|S_c| \geq \frac{v_C}{3} = \frac{p+3}{3}$, qui n'est pas un nombre entier, il s'ensuit qu'en effet $|S_c| \geq \frac{p+4}{3}$. Donc $2|S_c| + |S_{-c}| = 3|S_c| \geq p + 4$, et l'affirmation est donc vérifiée à nouveau.

Cas 2 : $v_C \leq p + 2$. L'ensemble $H_1 \setminus C$, peut être divisé en trois sous-ensembles de taille deux chacun, contenant un élément et son opposé, disons $H_1 \setminus C =$

$\{g_1, -g_1, g_2, -g_2, g_3, -g_3\}$. En échangeant éventuellement le sens de g_i et de $-g_i$, on peut supposer que pour chaque $i \in \{1,2,3\}$ on a $|S_{g_i}| \geq |S_{-g_i}|$. De plus, en renumérotant si nécessaire, on peut supposer que $|S_{-g_1}| \geq |S_{-g_2}| \geq |S_{-g_3}|$. En adoptant cette convention, on obtient $|S_{-g_3}| > 0$, ce qui implique qu'en fait les six séquences S_h pour $h \in H_1 \setminus C$ sont toutes non vides. Cependant, notons que nous ne savons pas si $|S_{g_1}| \geq |S_{g_2}|$; nous ne connaissons que $|S_{g_1}| \geq |S_{-g_1}| \geq |S_{-g_2}|$ et $|S_{g_2}| \geq |S_{-g_2}|$.

Cas 2.1 : $|S_{-g_3}| > 0$. Soit $i \in \{1,2,3\}$ tel que $|S_{g_i} S_{-g_i}|$ soit maximal parmi $|S_{g_1} S_{-g_1}|$, $|S_{g_2} S_{-g_2}|$, et $|S_{g_3} S_{-g_3}|$. Ainsi $|S_{g_i} S_{-g_i}| \geq \frac{3p+3-v_C}{3}$.

D'où

$$|S_c| + |S_{g_i}| + |S_{-g_i}| \geq \frac{3p+3-v_C}{3} + |S_c| = (p+1) + \left(|S_c| - \frac{v_C}{3}\right).$$

Ainsi, $|S_c| + |S_{g_i}| + |S_{-g_i}| \geq p+1$ avec égalité si et seulement si $|S_c| = \frac{v_C}{3}$ et $|S_{g_i}| + |S_{-g_i}| = \frac{3p+3-v_C}{3}$. Si l'égalité n'est pas satisfaite, alors comme on a $|S_c| + |S_{g_i}| + |S_{-g_i}| > p+1$, la déduction découle de la partie (1) du lemme 2.6

Supposons donc que l'on ait une égalité, c'est-à-dire $|S_c| = \frac{v_C}{3}$ et $|S_{g_i}| + |S_{-g_i}| = \frac{3p+3-v_C}{3}$. La dernière égalité implique qu'en effet $|S_{g_j}| + |S_{-g_j}| = \frac{3p+3-v_C}{3}$ pour chaque $j \in [1,3]$, tandis que la première implique que $|S_c| = |S_{-c}| = |S_0|$.

Puisque $|S_c| \geq \frac{p+2}{3}$ (rappelons l'argument au tout début de la preuve) tandis que $v_C \leq p+2$ (c'est l'hypothèse du cas 2) nous obtenons qu'en effet $v_C = p+2$, et donc $|S_c| = \frac{p+2}{3}$.

De plus, nous pouvons maintenant déduire que $|S_{g_j}| + |S_{-g_j}| = \frac{2p+1}{3}$ pour chaque $j \in \{1,2,3\}$. Or, puisque $|S_{g_j}| \leq |S_c|$, cela n'est possible que si $|S_{g_j}| = \frac{p+2}{3}$ et $|S_{-g_j}| = \frac{p-1}{3}$. Par conséquent, on a

$$\begin{aligned} c = \sigma(\pi_1(S)) = \\ \frac{p+2}{3}\Big(c + (-c) + 0 + g_1 + g_2 + g_3\Big) + \frac{p-1}{3}(-g_1 - g_2 - g_3) = \\ g_1 + g_2 + g_3. \end{aligned}$$

Maintenant, nous pouvons appliquer la partie (1) du lemme 2.6 avec g_1, g_2, g_3 ; notons que $|S_{g_1}| + |S_{g_2}| + |S_{g_3}| = 3 \cdot \frac{p+2}{3} = p+2$. En fait, on peut voir que $g_1 + g_2 + g_3 = c$ est impossible. Affirmer cela serait une autre façon de conclure.

Cas 2.2 : $|S_{-g_2}| > 0$ et $|S_{-g_3}| = 0$. On a $|S_{g_1} S_{-g_1} S_{g_2} S_{-g_2}| = 3p+3-v_C-|S_{g_3}| \geq 3p+3-v_C-|S_c|$. Soit $i \in \{1,2\}$ tel que $|S_{g_i} S_{-g_i}|$ soit maximal parmi $|S_{g_1} S_{-g_1}|$ et $|S_{g_2} S_{-g_2}|$. alors,

$$|S_c| + |S_{g_i}| + |S_{-g_i}| \geq \frac{3p+3-v_C-|S_c|}{2} + |S_c| = p+1+\frac{p+1-v_C+|S_c|}{2}.$$

Or, puisque $v_C \leq p+2$ et $|S_c| \geq \frac{p+2}{3} \geq 2$, il s'ensuit que $|S_c| + |S_{g_i}| + |S_{-g_i}| \geq p+2$ et on peut appliquer la partie (1) du lemme 2.6 avec $c, g_i, -g_i$.

Cas 2.3 : $|S_{-g_1}| > 0$ et $|S_{-g_2}| = |S_{-g_3}| = 0$. On a $|S_{g_1}S_{-g_1}| = 3p+3-v_C - |S_{g_2}S_{g_3}| \geq 3p+3-v_C-2\cdot|S_c|$. Ainsi, $|S_c|+|S_{g_1}|+|S_{-g_1}| \geq 3p+3-v_C-|S_c|$. Puisque $v_C \leq p+2$ et $|S_c| \leq p$, il s'ensuit que $|S_c|+|S_{g_1}|+|S_{-g_1}| \geq p$ avec égalité si et seulement si $v_C = p+2$ et $|S_c| = p$. Si l'égalité n'est pas atteinte, la preuve découle de la partie (1) du lemme 2.6 avec $c, g_1, -g_1$. Ainsi, nous supposons que $v_C = p+2$ et $|S_c| = p$. Comme $2|S_c|+|S_{-c}| \geq 2p \geq p+4$, nous avons :

— Si $|S_{-c}| \neq 0$, alors l'affirmation découle de la partie (2) du lemme 2.6 avec $x = c$ et $y = -c$.

— Si $|S_{-c}| = 0$, alors $|S_0| = v_C - |S_c| = 2$ et la preuve découle de la partie (2) du Lemme 2.6 avec $x = 0$ et comme $2|S_0|+|S_c| = p+4$, alors $y = c$.

Cas 2.4 : $|S_{-g_1}| = |S_{-g_2}| = |S_{-g_3}| = 0$. On a $|S_{g_1}S_{g_2}S_{g_3}| = 3p+3-v_C \geq 2p+1$ Si $|S_c| = p$, alors nous pouvons supposer que $v_C \leq p+1$ (voir l'argument à la fin du cas précédent). Ainsi, dans le cas présent $|S_{g_1}S_{g_2}S_{g_3}| \geq 2p+2$. Il s'ensuit que pour chaque $i \in \{1,2,3\}$, on a $|S_{g_i}| \geq 2$, et donc $2|S_{g_i}|+|S_{g_j}| \geq 2+(2p+2-p) = p+4$, pour chaque choix de i et j dans $\{1,2,3\}$.

Si $|S_c| \leq p-1$, alors il s'ensuit que pour chaque $i \in [1,3]$, on a$|S_{g_i}| \geq 2p+1-2(p-1) = 3$, et donc $2|S_{g_i}|+|S_{g_j}| \geq 3+(2p+1-(p-1)) = p+5$, pour chaque choix de i et j dans $\{1,2,3\}$.

Ainsi, s'il y a un choix de i et j pour lequel $2g_i+g_j = c$, en appliquant la partie (2) du lemme 2.6, le résultat cherché est $2|S_{g_i}|+|S_{g_j}| \geq p+4$.

On peut constater que ce choix existe toujours. En effet, pour un élément d de H_1 tel que $H_1 = \langle c \rangle \oplus \langle d \rangle$ nous remarquons que :

$$\{\{g_1,-g_1\},\{g_2,-g_2\},\{g_3,-g_3\}\} = \{\{d,-d\},\{c+d,-c-d\},\{c-d,-c+d\}\}.$$

Il existe huit possibilités pour l'ensemble $\{g_1,g_2,g_3\}$ (notez que l'ordre des éléments est indifférent), et pour chacun de ces huit choix, il existe une relation de la forme $2\cdot g_i+1\cdot g_j+0\cdot g_k = c$ avec $\{i,j,k\} = \{1,2,3\}$. Plus précisément :

— $2\cdot d+1\cdot(c+d)+0\cdot(c-d) = c$

— $1\cdot d+0\cdot(c+d)+2\cdot(-c+d) = c$

— $0\cdot d+1\cdot(-c-d)+2\cdot(c-d) = c$

— $1\cdot d+0\cdot(-c-d)+2\cdot(-c+d) = c$

— $2\cdot(-d)+0\cdot(c+d)+1\cdot(c-d) = c$

— $0\cdot(-d)+2\cdot(c+d)+1\cdot(-c+d) = c$

— $1\cdot(-d)+2\cdot(-c-d)+0\cdot(c-d) = c$

— $1\cdot(-d)+2\cdot(-c-d)+0\cdot(-c+d) = c$

Ceci achève la preuve. □

c) Preuve du résultat principal

Pour établir notre résultat principal, nous combinons les résultats partiels obtenus jusqu'à présent.

Par les lemmes 2.3 et 2.4 nous savons que pour chaque $n \geq 2$ on a $\mathrm{g}(C_3 \oplus C_{3n}) \geq 3n+3$.

Maintenant, supposons que p est un nombre premier supérieur ou égal à 5. Nous voulons montrer que $\mathrm{g}(C_3 \oplus C_{3p}) \leq 3p+3$. Soit S une séquence sans facteur carré de longueur $3p+3$ sur $C_3 \oplus C_{3p}$. Nous devons montrer que S contient une sous-séquence de longueur $3p$ et de somme nulle. Nous utilisons les applications π_1 et π_2 introduites dans la sous-section précédente.

Si $\sigma(\pi_1(S)) \neq 0$, alors d'après la proposition 2.10, la séquence S admet une sous-séquence de somme nulle.

Si $\sigma(\pi_1(S)) = 0$, alors, soit par la proposition 2.7, soit par la proposition 2.8, S admet une sous-séquence de somme nulle.

Ainsi, dans tous les cas, S admet une sous-séquence de longueur $3p$ et de somme nulle, et donc $\mathrm{g}(C_3 \oplus C_{3p}) \leq 3p+3$. En combinant avec le minorant, cela implique que pour chaque p premier supérieur ou égal à 5 on a $\mathrm{g}(C_3 \oplus C_{3p}) = 3p+3$.

Il reste à déterminer la valeur de $\mathrm{g}(C_3 \oplus C_6)$ et de $\mathrm{g}(C_3 \oplus C_9)$. Nous savons par les lemmes 2.3 et 2.5 que les valeurs respectives sont des minorants. Pour montrer que ces minorants sont les valeurs exactes de la constante de Harborth, nous avons utilisé un algorithme.

La suite de ce chapitre concerne la description de notre algorithme.

3 Algorithme et résultats calculatoires

Pour la description de l'algorithme, nous utilisons le vocabulaire des ensembles plutôt que celui des séquences, car la description semble plus naturelle. Déterminer la constante de Harborth de G revient à trouver le plus petit entier k tel que chaque sous-ensemble de G de cardinal k ait un sous-ensemble de cardinal $\exp(G)$ et de somme nulle. Nous présentons ci-dessous l'algorithme que nous avons utilisé.

Dans un premier temps, tous les sous-ensembles de G de cardinal $\exp(G)$ et de somme nulle sont construits (voir la discussion à la fin de cette section pour les détails à ce sujet). Si les sous-ensembles de cardinal $\exp(G)$ et de somme nulle sont tous les sous-ensembles de G de cardinal $\exp(G)$, alors cela signifie que la constante de Harborth est $\exp(G)$.

Sinon, nous considérons tous les sous-ensembles de G qui sont les successeurs directs pour la relation d'inclusion d'un ensemble de cardinal égal à $\exp(G)$ et de somme nulle ; en d'autres termes, nous étendons chaque sous-ensemble de cardinal $\exp(G)$ et de somme nulle de toutes les manières possibles de l'inclure dans un sous-ensemble de cardinal $\exp(G)+1$.

Ainsi, nous obtenons tous les sous-ensembles de G de cardinal $\exp(G)+1$ qui contiennent un sous-ensemble de cardinal $\exp(G)$ et de somme nulle. Si les sous-ensembles de G ainsi obtenus sont tous des sous-ensembles de G de cardinal $\exp(G)+1$, alors nous avons établi que la constante de Harborth de G est $\exp(G)+1$. Sinon, nous continuons comme ci-dessus jusqu'à ce que, pour un entier k, l'ensemble des sous-ensembles de cardinal k obtenus de cette manière coïncide avec l'ensemble de tous les sous-ensembles de cardinal k de G.

Nous détaillons ci-dessous un peu plus les étapes de l'algorithme. Cependant, une étude plus complète du problème algorithmique sera présentée dans le paragraphe a), et nous passons ici sur des aspects plus techniques. Pour plus de détails, le code source est disponible sur https://github.com/Zerdoum/Harborth_constant.

a) Les étapes de l'algorithme

Entrée : Un groupe abélien fini G d'ordre n et un exposant e.

- `[Initialisation]` Soit $Z(e)$ la collection de tous les sous-ensembles de G de cardinal e qui ont la somme 0. Initialisons $k \leftarrow e$.
- `[vérifications]` Si $|Z(k)| = \binom{n}{k}$, alors retourner $\mathrm{g}(G) = k$ et terminer. Sinon, incrémenter k.
- `[itération]` Soit $Z(k)$ la collection de tous les sous-ensembles de cardinal k de G, obtenu en insérant un élément dans les ensembles de $Z(k+1)$.
- Allez en `[vérification]`.

Sortie : $\mathrm{g}(G)$, la constante Harborth du groupe G.

Nous ajoutons quelques explications et remarques supplémentaires.

- Le groupe n'intervient que dans l'étape [Initialisation], il est représenté par les procédures d'addition et d'inverse. Le reste de l'algorithme ne fonctionne qu'avec des sous-ensembles d'un ensemble courant donné. Pour trouver tous les sous-ensembles de cardinal e avec somme 0, nous parcourons tous les sous-ensembles E de cardinal $e-1$. Pour chacun de ces ensembles, nous calculons la somme des $e-1$ éléments de cet ensemble et nous vérifions si cette somme appartient à E. Si ce n'est pas le cas, la somme est ajoutée

à E pour obtenir un ensemble de cardinal e et de somme 0. Pour cette étape, les sous-ensembles de G sont représentés par un bitmap. Procéder ainsi permet de n'explorer que $\binom{n}{e-1}$ ensembles au lieu de $\binom{n}{e}$.

— Pour les dernières parties de l'algorithme, en plus des représentations sous forme de bitmaps, la numérotation judicieusement choisie des sous-ensembles de G présentée dans le paragraphe 2 est utilisée. En particulier, la numérotation est choisie de manière à ce que, pour chaque sous-ensemble, son cardinal soit au moins aussi grand que le cardinal de tous ses prédécesseurs. Cela est utile car de cette manière, à chaque étape, notre recherche peut être efficacement limitée aux sous-ensembles $\binom{n}{k}$ d'un cardinal donné k au lieu de devoir considérer tous les 2^n sous-ensembles à chaque étape.

— Les sous-ensembles de cardinal k qui ne sont pas dans $Z(k)$ sont tous les sous-ensembles de G de cardinal k qui n'ont pas de sous-ensemble à somme nulle d'éléments e. Ainsi, dans la dernière étape avant que l'algorithme ne se termine, nous avons effectivement tous les sous-ensembles de G de cardinal $\mathrm{g}(G)-1$ qui n'ont pas de sous-ensemble à somme nulle d'éléments e.

— L'algorithme est valable pour tout groupe abélien fini. Avec le matériel dont nous disposons, il est possible de calculer la constante de Harborth pour les groupes abéliens finis d'ordre jusqu'à environ 45. Le principal facteur limitant est la mémoire. Afin d'augmenter la taille des groupes accessibles, la perspective existe de travailler avec une représentation plus efficace des sous-ensembles basée sur la compression des données.

— Le fait que e soit égal à l'exposant du groupe n'est pas critique pour l'algorithme. Des modifications minimes peuvent permettre de calculer d'autres constantes.

Nous terminons en mentionnant deux autres résultats de calcul.

Proposition 3.1. *1.* $\mathrm{g}(C_6 \oplus C_6) = 13$.

2. $\mathrm{g}(C_3 \oplus C_{12}) = 15$.

Le premier confirme la conjecture de Gao et Thangadurai, mentionnée dans l'introduction, $\mathrm{g}(C_n \oplus C_n) = 2n+1$ pour un montant pair de n dans le cas où $n = 6$. Cette dernière montre que $\mathrm{g}(C_3 \oplus C_{3n}) = 3n+3$ vaut également pour $n = 4$, ce qui conforte l'idée que $\mathrm{g}(C_3 \oplus C_{3n})$ pourrait valoir $3n+3$ pour des entiers n qui ne sont pas premiers.

CHAPITRE IV

LA CONSTANTE D'ERDŐS-GINZBURG-ZIV

Sommaire

1 Introduction

Soit $(G,+,0)$ un groupe abélien fini. La constante d'Erdős-Ginzburg-Ziv de G notée $\mathrm{s}(G)$, est le plus petit entier $\ell \in \mathbb{N}$ tel que toute séquence S de G de longueur ℓ, admet une sous-séquence T de longueur $\exp(G)$ dont la somme est nulle. Rappelons que dans une séquence, la répétition des éléments de G est admise.

D'après le théorème d'Erdős-Ginzburg-Ziv prouvé en 1960, voir [EGZ61], nous connaissons la valeur de cette constante dans le cas des groupes cycliques.

$$\mathrm{s}(C_n) = 2n - 1$$

Pour une démonstration plus moderne, voir [GHK06]. La valeur de la constante d'Erdős-Ginzburg-Ziv est également connue pour des groupes de rang 2, voir [FGZ11].

$$\text{Soit } G = C_{n_1} \oplus C_{n_2} \text{ avec } n_1 | n_2 \text{ alors } \mathrm{s}(G) = 2n_1 + 2n_2 - 3$$

Le cas crucial où $n_1 = n_2 = p$ avec p un nombre premier, a été prouvé dans [Rei07], et [SC05].

Nous connaissons aussi d'après [FGZ11] les résultats suivants : Soient G un groupe abélien fini, $r,\ a,\ b$ des entiers positifs.

- Si $G = C_{2^a} \oplus C_{2^b}^{r-1}$ où $r \geq 2,\ b \geq 1$ et $a \in [1,b]$, alors $\mathrm{s}(G) = 2^{r-1}(2^a + 2^b - 2) + 1$. [Ede+07]
- $\mathrm{s}(C_{3^a 5^b}^3) = 9(3^a 5^b - 1) + 1$, où $a + b \geq 1$. [Gao+07]
- $\mathrm{s}(C_{3^a}^4) = 20(3^a - 1) + 1$, où $a \geq 1$. [Ede+07]
- $\mathrm{s}(C_3^5) = 91$ [Ede+02] et $\mathrm{s}(C_3^6) = 255$.[Pot08]. Ces résultats ne sont pas formulés pour la constante s mais pour les ovoïdes. Or pour les groupes de la forme C_3^r, il est connu que déterminer la constante s et déterminer le plus grand cardinal d'un ovoïde sont des problèmes équivalents (mais les valeurs numériques ne sont pas identiques). Voir [Ede+07]
- $\mathrm{s}(C_{3\times 2^a}^3) = 8(3 \times 2^a - 1) + 1$, où $a \geq 1$. [Gao+07]
- $\mathrm{s}(C_2^2 \oplus C_{2n}) = 4n + 3$ avec $n \geq 2$. [FZ16]
- $\mathrm{s}(C_2^3 \oplus C_{2n}) = 4n + 5$ avec $n \geq 36$. [FZ16].

  Notons que pour les petites valeurs de n, y compris pour $n = 2$ et $n = 3$ la valeur de la constante d'Erdős-Ginzburg-Ziv pour le groupe $C_2^3 \oplus C_{2n}$ restait jusqu'à présent inconnus.

— Rappelons l'inégalité suivante du théorème 1.2 de l'article [SZ10] :

Soient p un nombre premier impaire, et G un p-groupe abélien fini avec $\exp(G) = n$ et $D(G) \leq 2n-1$. Alors :

$$2D(G) - 1 \leq \mathrm{s}_{\leq n}(G) + n - 1 \leq s(G) \leq D(G) + 2n - 2$$

La définition de la constante $D(G)$ a déjà été évoquée dans l'introduction. La constante $\mathrm{s}_{\leq n}(G)$ est définie comme étant le plus petit entier ℓ tel que toute séquence de longueur ℓ contient une sous-séquence de longueur au plus égale à n et de somme nulle. Cette constante est notée aussi $\eta(G)$. Voir le prochain chapitre pour une définition plus formelle et des informations complémentaires sur ces constantes.

Dans ce chapitre il est établi que :

— $\mathrm{s}(C_2 \oplus C_2 \oplus C_2 \oplus C_4) = 13$.

— $\mathrm{s}(C_2 \oplus C_2 \oplus C_2 \oplus C_6) = 17$.

D'après le théorème 1.2 de l'article [SZ10], on a $\mathrm{s}(C_p^{p+1} \oplus C_{p^2}) = 4p^2 - 3$ pour p un nombre premier impair.

Notre premier résultat, montre que cette égalité est aussi vraie pour $p = 2$.

Nous pouvons alors annoncer le théorème suivant :

Théorème 1.1. *Soit p un nombre premier. Alors*

$$\mathrm{s}(C_p^{p+1} \oplus C_{p^2}) = 4p^2 - 3$$

Nous avons trouvé que $\mathrm{s}(C_2^3 \oplus C_{2^2}) = 13$ d'une manière calculatoire. Notons que nous pouvons aussi trouver ce résultat d'une manière théorique.

En effet, d'après le théorème $6.6.1$ de article [GG06] on a :

$$\mathrm{s}_{\leq 4}(C_2^3 \oplus C_{2^2}) = \mathrm{s}(C_2^3 \oplus C_{2^2}) - 3$$

d'après le résultat récent de Luo, théorème 4.1 de l'article [Luo17] on a :

Soit G un $p-$groupe abélien fini avec $D(G) \leq \exp(G) - 1$. Alors pour tout entier a non divisible par p, le groupe $G^{'} = C_a \oplus G$ satisfait l'égalité suivante :

$$2D(G^{'}) - 1 = \mathrm{s}_{\leq \exp(G^{'})}(G^{'}) + \exp(G^{'}) - 1$$

En combinaison ces deux résultats et en prenant $a = 1$ on a :

$s(C_2^3 \oplus C_{2^2}) = 2D(C_2^3 \oplus C_{2^2}) - 1 = 13.$

Ces résultats ont été trouvés grace à un algorithme que nous avons développé.

Avec nos machines, nous avons pu déterminer la valeur de la constante d'Erdős-Ginzburg-Ziv pour des groupes abéliens finis d'ordre inférieur ou égal à 48.

En utilisant des machines plus puissantes, il serait possible de calculer la constante pour des groupes abéliens finis d'ordre plus grand.

La description de cet algorithme est donnée dans la section suivante.

2 Description de l'algorithme

Le principe général de l'algorithme est le suivant :

Tout d'abord, nous considérons l'ensemble de toutes les séquences de longueur exposant du groupe et de somme nulle.

Puis, nous calculons les successeurs immédiats de cet ensemble pour la relation de précédence. La relation de précédence est une relation d'ordre définie sur l'ensemble des séquences. Elle est définie comme suit :

Définition 2.1 (Relation de précédence sur les séquences)**.** On dit qu'une séquence S_1 précède une séquence S_2, si la séquence S_1 est une sous-séquence propre de S_2. Dans ce cas, on dit aussi que S_1 est un prédecesseur de la séquence S_2 ou encore que, la séquence S_2 est un successeur de la séquence S_1.

Définition 2.2 (Successeur immédiat d'une séquence)**.** On dit que S_1 précède immédiatement S_2 ou encore S_2 est un successeur immédiat de S_1 s'il n'y pas de séquence $S_1^{'}$ tel que S_1 précède $S_1^{'}$ et $S_1^{'}$ précède S_2.

Remarque 2.3. Si S_2 est un successeur immédiat de la séquence S_1 alors

$|S_2| = |S_1| + 1$.

Nous obtenons alors l'ensemble des séquences de longueur égale à l'exposant du groupe plus un qui sont des successeurs de séquences de longueur égale à l'exposant du groupe et de somme nulle.

Nous répétons l'opération précédente, jusqu'à ce que nous obtenions comme successeurs l'ensemble de toutes les séquences possibles d'une longueur spécifique.

Ainsi, toutes les séquences possibles de cette longueur, sont forcément successeurs pour la relation de précédence, d'une séquence de longueur exposant du groupe et de somme nulle.

Notre algorithme trouve alors cette longueur qui représente par définition la valeur de la constante d'Erdős-Ginzburg-Ziv.

Nous avons su implémenter cet algorithme d'une manière efficace grâce aux points suivants :

— Chaque séquence est représentée par un numéro. L'ensemble des séquences est alors représenté par un ensemble d'entiers. Nous utilisons la numérotation spéciale des séquences d'une longueur donnée introduite dans la section 2, page 19. Cette numérotation est particulièrement bien adaptée pour trouver les successeurs d'une séquence.

— Les ensembles d'entiers sont représentés comme une réunion d'intervalles d'entiers. Ce qui permet d'une part de réduire les données nécessaires pour représenter un sous-ensemble, et d'autre part de traiter un sous-ensemble en ne considérant si possible que les bornes des intervalles. Donc on économise d'une part de la mémoire et d'autre part du temps de calculs.

— En effet, au lieu de parcourir tout l'intervalle d'entiers pour calculer les successeurs de chaque entier individuellement, nous avons développé une méthode permettant de trouver les successeurs d'un intervalle d'entiers $[a,b]$ en considérant uniquement ses bornes a et b.

La suite du chapitre, est consacrée à une description détaillée de cet algorithme.

Tout d'abord, les suites croissantes d'entiers de longueur k et de somme nulle sont numérotées en suivant les étapes décrites dans la section 2, page 19.

Rappelons que cette numérotation spéciale permet de numéroter les mots binaires en des entiers de manière consécutive lorsque celles-ci possèdent le même poids de Hamming.

L'étape de l'initialisation est détaillée à travers l'exemple ci-dessous.

Exemple 2.4. Soit $G = C_2 \oplus C_2$.

$$G = \{(0,0),(0,1),(1,0),(1,1)\} = \{g_0, g_1, g_2, g_3\}$$

Les éléments de G sont représentés par l'indice i de g.

Concrètement, pour {(0,0),(0,1),(1,0),(1,1)} nous attribuons respectivement les numéros {0,1,2,3}.

L'ordre du groupe G est $n = 4$.

Et l'exposant du groupe G est $k = 2$.

Le tableau suivant explore toutes les suites croissantes d'entiers S de $F(G_0')$ de deux termes choisis parmi $\{0,1,2,3\}$.

La suite S' est la suite rendue strictement croissante par l'application ϕ. Voir la section 2, page 19.

Le mot binaire w représente l'ensemble des termes de S'.

Le numéro du sous-ensemble représenté par w est Num_4.

La somme des termes de la suite S est $\sigma(S)$.

S	S'	w	Num_4	$\sigma(S)$
00	01	11000	0	(0,0)+(0,0)=(0,0)
01	02	10100	1	(0,0)+(0,1)=(0,1)
02	03	10010	2	(0,0)+(1,0)=(1,0)
03	04	10001	3	(0,0)+(1,1)=(1,1)
11	12	01100	4	(0,1)+(0,1)=(0,0)
12	13	01010	5	(0,1)+(1,0)=(1,1)
13	14	01001	6	(0,1)+(1,1)=(1,0)
22	23	00110	7	(1,0)+(1,0)=(0,0)
23	24	00101	8	(1,0)+(1,1)=(0,1)
33	34	00011	9	(1,1)+(1,1)=(0,0)

TABLEAU IV.1 – Les suites croissantes d'entiers de deux termes choisis parmi $\{0,1,2,3\}$

— Par abus de notation, nous ne séparons pas les éléments d'une séquence par un point. Attention à ne pas confondre une séquence $S = 12$ avec l'entier 12.

— Les séquences sont identifiées par leurs numéros. Par exemple on parlera indistinctement de la séquence $S = 12$ comme étant la séquence numéro 5. Puisque $\mathrm{Num}_4(12) = 5$.

— Nous pouvons alors à partir de maintenant considérer des entiers à la place des séquences.

— L'ensemble des séquences de somme nulle est représenté comme la réunion d'intervalle trié d'entiers suivant : $[0,0] \cup [4,4] \cup [7,7] \cup [9,9]$. Lorsqu'un intervalle $[a,b]$ est réduit au singleton {a} tel que $a = b$, il est simplement noté $[a]$. Notons qu'à ce stade, la représentation des séquences comme une réunion d'intervalles ne permet pas d'économiser de la mémoire.

Nous cherchons maintenant les successeurs immédiats de l'ensemble des séquences de somme nulle et de longueur k.

a) Le calcul des successeurs d'une séquence

Comme nous l'avons vu dans la définition 2.1, on appelle successeurs d'une séquence S par la relation de précédence, l'ensemble de toutes les séquences S_i telles que la séquence S est une sous-séquence de la séquence S_i.

Trouver tous les successeurs immédiats d'une séquence S peut se faire de la manière suivante :

— Commençons par un représentant canonique de la séquence S.

— On rajoute successivement l'un des éléments du groupe G "derrière" cette suite. on aura en tout n successeurs immédiats pour chaque séquence, où n est l'ordre du groupe.

— L'ensemble des successeurs immédiats de la séquence S est l'ensemble des séquences qui correspondent à ces suites. Or les suites ainsi obtenues ne sont pas nécessairement les représentants canoniques de ces séquences. Il convient alors de réordonner les termes pour retrouver le représentant canonique de chaque successeur.

— Notons qu'après deux étapes on a $\frac{n\times(n+1)}{2}$ successeurs d'une séquence et non pas n^2 successeurs.

Dans l'exemple suivant, on commence par la séquence $S = 11$ et nous déterminons tous les successeurs de cette séquence et leurs numéros correspondants.

Exemple 2.5. Les successeurs de la séquence $S = 11$ sont :

— $S_1 = 11\mathbf{0} = \mathbf{0}11$,

— $S_2 = 11\mathbf{1} = \mathbf{1}11$,

— $S_3 = 11\mathbf{2} = 11\mathbf{2}$,

— $S_4 = 11\mathbf{3} = 11\mathbf{3}$

Chaque séquence S peut être identifiée par son numéro. Nous donnons dans le tableau ci-dessous la liste complète des séquences de longueur 3 sur un groupe d'ordre 4.

La colonne 1 contient les représentants canoniques de la séquence S. La colonne 2 contient les suites croissantes associés. La colonne 3 contient les bitmaps. La colonne 4 contient les numéros des séquences.

Les lignes en rouge sont celles qui correspondent aux successeurs de la séquence $S = 11$

L'ensemble des successeurs de la séquence $S = 11$ est donc l'ensemble des séquences de longueur 3 représentées par les numéros : $4, 10, 11$ et 12. On peut alors représenter l'ensemble des successeurs de la séquence S comme une réunion d'intervalles triés d'entiers : $[4] \cup [10, 12]$.

Cherchons maintenant les successeurs immédiats des séquences de longueur 2 et de somme nulle. Comme indiqué dans le tableau de l'exemple 2.4, il s'agit des séquences représentées par les numéros : 0, 4, 7 et 9.

Le tableau suivant, nous montre ces successeurs.

— La colonne 1 représente les séquences comme représentants canoniques.

— La colonne 2 donne les numéros de ces séquences.

— la colonne 3 contient l'ensemble des successeurs des séquences de la colonne 1 comme représentants canoniques.

S	S'	w	Num_4
000	012	111000	0
001	013	110100	1
002	014	110010	2
003	015	110001	3
011	023	101100	4
012	024	101010	5
013	025	101001	6
022	034	100110	7
023	035	100101	8
033	045	100011	9
111	123	011100	10
112	124	011010	11
113	125	011001	12
122	134	010110	13
123	135	010101	14
133	145	010011	15
222	234	001110	16
223	235	001101	17
233	245	001011	18
333	345	000111	19

TABLEAU IV.2 – La liste complète des séquences de longueur 3 sur un groupe d'ordre 4

— La colonne 4 représente l'ensemble des numéros des successeurs des séquences de la colonne 1.

1	2	3	4
00	0	{ 000, 001, 002, 003}	{0,1,2,3}
11	4	{ 011, 111, 112, 113}	{4,10,11,12}
22	7	{ 022, 122, 222, 223}	{7,13,16,17}
33	9	{ 033, 133, 233, 333}	{9,15,18,19}

Il résulte que l'ensemble des successeurs immédiats de l'ensemble des séquences de deux termes dont les numéros sont dans $[0] \cup [4] \cup [7] \cup [9]$ est l'ensemble des séquences de 3 termes représentés par les numéros :

$0, 1, 2, 3, 4, 7, 9, 10, 11, 12, 13, 15, 16, 17, 18$ et 19, qu'on peut représenter sous forme d'une réunion d'intervalles triée : $[0,4] \cup [7] \cup [9,13] \cup [15,19]$.

L'objectif de l'algorithme est de calculer directement la colonne 4 à partir de la colonne 2.

Remarque 2.6. A partir de maintenant on parlera de :

— successeurs d'un entier comme étant un ensemble d'entiers.

— successeurs d'un intervalle d'entiers comme étant la réunion des successeurs d'entiers pour tous les entiers qui appartiennent à l'ensemble.

Dans la suite de la sous-section, nous allons voir, comment trouver les successeurs d'une séquence représentée par son entier.

1-) Définition : successeurs d'une séquence représentée par son numéro

Pour calculer les successeurs d'une séquence représentée par son numéro, nous présentons une méthode directe qui suit la définition. Cette méthode est loin d'être intéressante en raison de son inefficacité, mais est nécessaire pour disposer d'une version de référence et valider les versions plus efficaces présentées plus loin.

1. Traduire l'entier en un mot binaire en parcourant le triangle de Pascal modifié décrit dans la section 1-), page 21.
2. Calculer les successeurs du mot binaire. Cette étape se décompose en plusieurs étapes :
   - Traduire le mot binaire en une suite strictement croissante et sans répétition d'éléments de G.
   - Convertir par soustraction la suite strictement croissante en une séquence croissante au sens large.
   - Trouver les successeurs en ajoutant un élément de G.
   - Procéder inversement en convertissant le résultat en une suite strictement croissante puis en un mot binaire.
3. Traduire les mots binaires en entiers.

Dans la suite du paragraphe, nous montrons comment les successeurs peuvent être déterminés directement sur le mot binaire et obtenir un algorithme considérablement plus efficace.

Voici les détails du calcul :

Rappelons que d'après la formule 1.1 page 19 l'ensemble des mots binaires de longueur $n+k-1$ et de poids k est noté $B_{n+k-1,k}$.

$$B_{n+k-1,k} = \{x \in \{0,1\}^{n+k-1} \mid w_H(x) = k\}$$

Pour x appartenant à l'ensemble $B_{n+k-1,k}$, l'ensemble des successeurs de x contient n éléments qui sont notés $x_i^{'}$ pour i allant de 1 jusqu'à n, où $x_i^{'}$ appartient à l'ensemble $B_{n+k,k+1}$.

- Soit x, le mot binaire qui correspond à la séquence croissante S.
- D'après l'application 2.5 page 24, $\mathrm{Num}_{n+k-1,k}(x)$ est le numéro qui identifie la séquence x.

Les séquences $x_i^{'}$ sont obtenus en plaçant un "1" consécutivement dans des positions bien précises du mot binaire x. Deux règles complémentaires permettent de déterminer ces positions.

1. Ajouter un "1" au début du mot binaire x revient à ajouter l'élément 0 à la séquence S.
2. En comptant les "0" dans le mot binaire à partir de 1, ajouter un "1" dans le mot binaire après le 0 numéro i revient à rajouter l'élément i à la séquence.

Montrons maintenant comment on peut directement manipuler les numéros des mots binaires pour déterminer ses successeurs.

D'après la première règle, rajouter un "1" au début du mot binaire, revient à faire une bifurcation à la fin du parcours du triangle de Pascal modifié vers la droite. On aura donc :

Pour i allant de 1 jusqu'à n les $x_i^{'}$ sont les successeurs de x et son calculer par les formules suivantes :

$$\mathrm{Num}_{n+k,k+1}(x_1^{'}) = \mathrm{Num}_{n+k-1,k}(x) + a_{n+k-1,k} \tag{2.1}$$

D'après la règle 2, soit j la position où un 1 est rajouté dans le mot binaire. La position j correspond à l'endroit de la bifurcation dans le triangle de Pascal modifié en allant vers la droite. Soit t la colonne qui correspond à l'endroit de la bifurcation. Nous pourrons alors pour i allant de 2 jusqu'à n calculer les successeurs $x_i^{'}$ de la manière suivante :

$\forall i \in \{2,\ldots,n\}$, et $(j,t) \in \mathbb{N}^2$,

$$\mathrm{Num}_{n+k,k+1}(x_i^{'}) = \mathrm{Num}_{n+k,k+1}(x_{i-1}) + \binom{j}{t} \tag{2.2}$$

Exemple 2.7. Soient $n = 5$, $k = 3$ et $x = 0011010$, $\mathrm{Num}_{7,3}(0011010) = 48$.

$\mathrm{Num}_{8,4}(x_1') = \mathrm{Num}_{8,4}(00110101) = 48 + a_{7,3} = 48 + 64 = 112$

$\mathrm{Num}_{8,4}(x_2') = \mathrm{Num}_{8,4}(00110110) = 112 + \binom{6}{3} = 132$

$\mathrm{Num}_{8,4}(x_3') = \mathrm{Num}_{8,4}(00111010) = 132 + \binom{4}{2} = 138$

$\mathrm{Num}_{8,4}(x_4') = \mathrm{Num}_{8,4}(01011010) = 138 + \binom{1}{0} = 139$

$\mathrm{Num}_{8,4}(x_5') = \mathrm{Num}_{8,4}(10011010) = 139 + \binom{0}{0} = 140$

n \ k	-1	0	1	2	3	4	4	5	6
-1	0	0	0	0	0	0	0	0	0
0	0	1	1	1	1	1	1	1	1
1	0	1	2	2	2	2	2	2	2
2	0	1	3	4	4	4	4	4	4
3	0	1	4	7	8	8	8	8	8
4	0	1	5	11	15	16	16	16	16
5	0	1	6	16	26	31	32	32	32
6	0	1	7	22	42	57	63	64	64
7	0	1	8	29	64	99	120	127	128

FIGURE IV.1 – Les différentes bifurcations pour le calcul des successeurs d'un numéro

Remarque 2.8. Les zéros consécutifs au début du mot binaire x correspondent à des successeurs de x dont les numéros sont consécutifs.

Dans l'exemple 2.7 le mot binaire $x = \mathbf{00}11010$ commence par deux zéros consécutifs. Par conséquent les numéros des successeurs $\mathrm{Num}_{8,4}(x_3')$, $\mathrm{Num}_{8,4}(x_4')$ et $\mathrm{Num}_{8,4}(x_5')$ sont consécutifs.

En effet, d'après l'équation 2.2, les zéros au début du mot binaire correspondent à rajouter au $\mathrm{Num}_{n+k,k+1}(x_{i-1})$ un coefficient binomial $\binom{j}{0} = 1$.

Le procédé du calcul des successeurs d'une séquence que nous avons décrit ci-dessus est issue de la définition des successeurs d'une séquence.

En termes algorithmiques, en appliquant ce procédé, il est nécessaire de parcourir toutes les séquences ce qui conduit à un temps de calcul considérable. Afin de gagner en efficacité, nous allons décrire une deuxième méthode. Cette dernière, permettra de calculer les successeurs d'une séquence à partir de son numéro

seulement sans avoir besoin de passer par le mot binaire. Enfin, en étendant cette méthode, nous arriverons à calculer les successeurs d'un intervalle d'entiers en considérant uniquement ses bornes.

2-) Deuxième méthode de calcul des successeur d'une séquence

Nous avons vu que les successeurs d'une séquence représentée par son numéro peuvent être calculés en faisant un parcours dans le triangle de Pascal modifié. Cette première méthode est une implémentation assez directe de la définition.

Comme nous l'avions vu dans la formule 2.2 le calcul des successeurs consiste à rajouter consécutivement un certain coefficient binomial au successeur précédent.

Avec la deuxième méthode que nous décrivons ici, nous pouvons trouver pour chaque successeur, le coefficient binomial à rajouter d'une manière directe. Ce qui nous évite de passer par les étapes décrites dans la section 1-), page 62.

En particulier, cette amélioration nous permet de générer les successeurs d'un intervalle de numéros de séquences en explorant uniquement ses bornes.

Voici la description de cette méthode :

Calculer les successeurs d'une séquence représentée par son numéro revient à effectuer une recherche dans une structure arborescente de coefficients binomiaux.

Nous rappelons qu'une structure arborescente est un arbre, donc un graphe sans cycle avec un sommet particulier appelé la racine de l'arborescence à partir duquel il existe un chemin unique vers tous les autres sommets, voir [FLP14].

Plus spécifiquement l'arborescence que nous allons considérer est l'arbre syntaxique de la formule itérée de Pascal. Elle exprime le coefficient binomial $\binom{n+k}{k+1}$ comme une somme de coefficients binomiaux de $\binom{i}{k}$. La formule est la suivante :

$$\binom{n+k}{k+1} = \sum_{i=k}^{n+k-1} \binom{i}{k} \tag{2.3}$$

Cette formule se démontre facilement par récurrence en utilisant la formule du triangle de Pascal.

Ci-dessous, nous rappelons la forme de cette arborescence en détails.

Description de l'arborescence

— Nous notons l'arborescence par $\mathcal{P}$ et sa racine par $\mathcal{R}$.

— Dans l'arbre syntaxique de cette forme, chaque nœud de l'arbre est étiqueté par un coefficient binomial.

— La racine $\mathcal{R}$ est étiquetée par le coefficient binomial $\binom{n+k}{k+1}$.

— La racine $\mathcal{R}$ a n fils qui sont étiquetés par les coefficients binomiaux $\binom{i}{k}$ qui apparaissent dans la formule 2.3.

Nous adoptons la convention de dessiner l'arborescence $\mathcal{P}$ de sorte que les nœuds fils apparaissent dans un ordre croissant allant de gauche vers la droite, comme le montre l'exemple ci-dessous.

Exemple 2.9. Soient $n = 5$ et $k = 2$. La formule itérée de Pascal est :

$$\binom{7}{3} = \sum_{i=2}^{6} \binom{i}{2} = \binom{2}{2} + \binom{3}{2} + \binom{4}{2} + \binom{5}{2} + \binom{6}{2}$$

Son arbre syntaxique est :

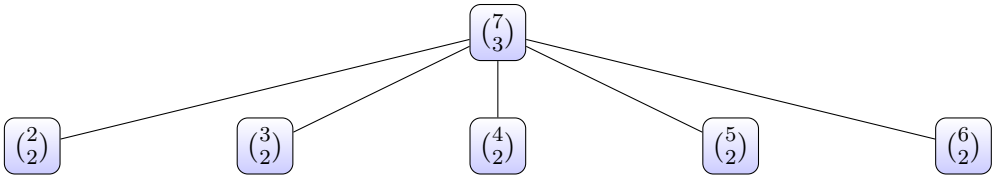


FIGURE IV.2 – Exemple d'arborescence sur un niveau

Nous décrivons maintenant comment on continue la construction de cet arbre.

— Pour un sommet de l'arbre $\mathcal{P}$ étiqueté par le coefficient binomial $\binom{i'}{k'}$. La descente en profondeur dans l'arbre implique la ré-application de la formule itérée de Pascal 2.3 sur les $\binom{i'}{k'}$. Soient $k' = k'' + 1$ et $i' = n'' + k''$.

$$\binom{i'}{k'} = \binom{n'' + k''}{k'' + 1} = \sum_{i''=k''}^{n''+k''-1} \binom{i''}{k''}$$

— On remarque que les sommets du même niveau de l'arbre $\mathcal{P}$ peuvent engendrer des coefficients binomiaux identiques avec ceux du niveau de dessous. Afin d'éviter d'avoir des sous-arbres identiques, nous ne recopions pas les mêmes coefficients binomiaux.

— Cette construction s'arrête lorsqu'on arrive aux coefficients binomiaux $\binom{i}{0}$ avec i allant de 0 jusqu'à $n - 1$. Ces coefficients représentent les feuilles de cette arborescence. Voir figure IV.3 ci-dessous.

Exemple 2.10. Soient $n = 5$ et $k = 2$

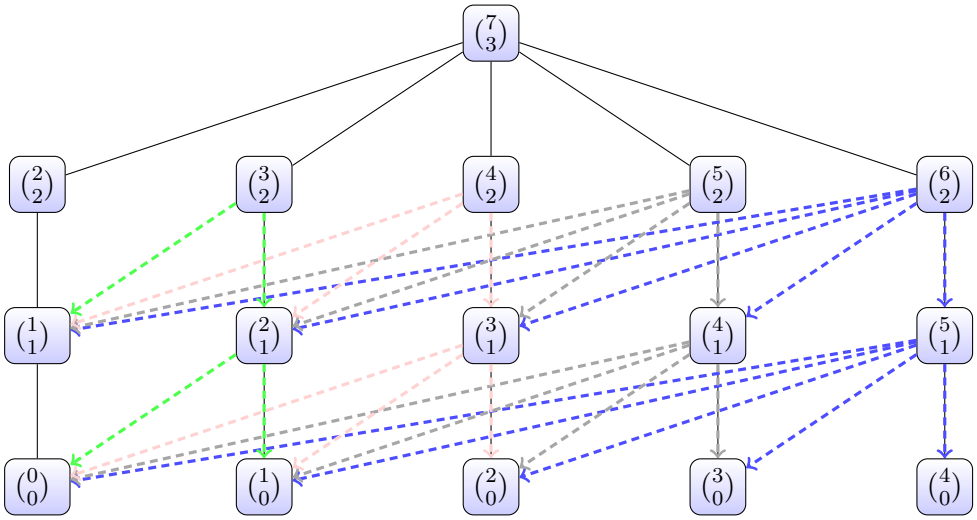


FIGURE IV.3 – Exemple d'arborescence entière

— À l'exception de la racine, on remarque qu'on a $k+1$ niveaux. On commence la numérotation des niveaux par 0 et la numérotation se fait de sorte que le $i^{\text{ème}}$ niveau correspond aux noeuds étiquetés par les coefficients $\binom{n'}{k-i}$.

Chaque niveau de l'arborescence $\mathcal{P}$ a exactement n nœuds. Ces nœuds sont triés d'une manière naturelle, on peut alors les ranger dans une forme rectangulaire. Ainsi, nous utilisons le vocabulaire de lignes et de colonnes.

En prenant $n = 5$ et $k = 2$, la table IV.4 ci-dessous range chaque nœud de l'arborescence IV.3 sauf la racine dans une case. Le i désigne l'indice des lignes et le j désigne l'indice des colonnes.

i \ j	0	1	2	3	4
0	$\binom{2}{2}$	$\binom{3}{2}$	$\binom{4}{2}$	$\binom{5}{2}$	$\binom{6}{2}$
1	$\binom{1}{1}$	$\binom{2}{1}$	$\binom{3}{1}$	$\binom{4}{1}$	$\binom{5}{1}$
2	$\binom{0}{0}$	$\binom{1}{0}$	$\binom{2}{0}$	$\binom{3}{0}$	$\binom{4}{0}$

FIGURE IV.4 – Rangement des noeuds dans une table

— Pour deux entiers n et k donnés, en utilisant cette arborescence,

nous obtenons tous les coefficients binomiaux $\binom{j}{t}$ possibles permettant de calculer les successeurs comme indiqué dans la formule 2.2. La question qui se pose maintenant est :

Comment déterminer les entiers j et t permettant de calculer chaque successeur d'une séquence représentée par son numéro ?

Afin de répondre à cette question, nous parcourons notre arborescence d'une manière spécifique. À chaque étape, nous calculons un certain intervalle d'entiers $[a,b[=\{a,\ldots,b-1\}$. Nous retenons aussi certains paramètres qu'on notera min_i et coefficients_i pour i allant de 1 jusqu'à r tel que, r est le nombre de fois où on doit retenir les paramètres. Cet entier r dépend du parcours dans l'arbre $\mathcal{P}$. Les détails de la description des intervalles et la détermination de ces paramètres sont décrits par la suite.

Les successeurs S_i d'un numéro de séquence sont calculés à partir de ces paramètres comme suit :

Soit $x = S_0$ le numéro de séquence dont on cherche ses successeurs S_1 jusqu'à S_n où n est l'ordre du groupe.

Pour i allant de 1 jusqu'à r, on a :

$$S_i = \text{coefficients}_i - \text{min}_i + S_{i-1} \tag{2.4}$$

Si l'entier r est strictement inférieur à l'entier n, les successeurs pour i allant de $r+1$ jusqu'à n seront consécutifs :

$$S_i = S_{i-1} + 1 \tag{2.5}$$

En particulier, on a :

$$S_n = S_r + (n-r) \tag{2.6}$$

Notons que pour i à partir de r, nous tombons sur le cas de la remarque 2.8.

Nous allons décrire comment on doit parcourir cet arbre et comment les paramètres coefficients_i et min_i sont définis. Mais avant cela, nous expliquons comment calculer les intervalles d'entiers.

Le calcul des intervalles :

Le calcul d'un intervalle associé à un sommet, qui n'est pas une feuille (nous n'associons pas un intervalle à une feuille), change en fonction du parcours dans l'arbre $\mathcal{P}$. Cependant, il se fait d'une manière plutôt locale.

Pour décrire son calcul, nous introduisons d'abord certains termes :

Nous nous retrouvons dans un certain nœud $\mathcal{N}$ de l'arbre $\mathcal{P}$. Ce nœud est appelé nœud courant. L'intervalle associé à ce nœud est noté $[a_{\mathcal{N}}, b_{\mathcal{N}}[$. Il est étiqueté par le coefficient $\binom{i_{\mathcal{N}}}{j_{\mathcal{N}}}$.

En fonction du numéro de la séquence, et de l'intervalle associé au nœud $\mathcal{N}$, deux types de mouvements interviennent dans le parcours de l'arbre $\mathcal{P}$.

— Aller vers le nœud qui se trouve à gauche du nœud courant $\mathcal{N}$, appelé le nœud voisin de gauche. L'intervalle associé à ce nœud est $[a_{\mathcal{N}_{\text{voisin}_\text{g}}}, b_{\mathcal{N}_{\text{voisin}_\text{g}}}[$. Il est étiqueté par le coefficient $\binom{i_{\mathcal{N}}-1}{j_{\mathcal{N}}}$. Notons que certains nœuds n'ont pas de voisin de gauche. Or, pour de tels nœuds, on ne sera jamais amené à aller vers la gauche.

— Descendre dans l'arbre $\mathcal{P}$ à partir du nœud courant $\mathcal{N}$ vers le nœud qui se trouve à l'extrémité droite. Ce dernier est appelé nœud fils. Notons qu'en rangeant les nœuds de $\mathcal{P}$ dans une table, le nœud fils est le nœud qui se trouve directement en bas du nœud $\mathcal{N}$.

Le nœud fils est noté $\mathcal{N}_{\text{fils}}$. Il a pour intervalle $[a_{\mathcal{N}_{\text{fils}}}, b_{\mathcal{N}_{\text{fils}}}[$. Et a comme coefficient $\binom{i_{f\mathcal{N}}}{j_{f\mathcal{N}}} = \binom{i_{\mathcal{N}}-1}{j_{\mathcal{N}}-1}$.

Nous utilisons aussi la terminologie de nœud voisin de droite. Il est noté $\mathcal{N}_{\text{voisin}_\text{d}}$. Et Il a pour intervalle $[a_{\mathcal{N}_{\text{voisin}_\text{d}}}, b_{\mathcal{N}_{\text{voisin}_\text{d}}}[$.

Si le nœud $\mathcal{N}$ est le nœud fils du nœud $\mathcal{N}_1$, alors on appelle aussi le nœud $\mathcal{N}_1$, le nœud parent du nœud $\mathcal{N}$.

Voici le détail du calcul des intervalles :

En descendant en profondeur dans l'arbre, l'entier $a_{\mathcal{N}}$ du nœud courant $\mathcal{N}$ obtenu après la descente, possède la même valeur que l'entier $a_{\mathcal{N}_{\text{parent}}}$ du nœud parent. Tandis que l'entier $b_{\mathcal{N}}$ a pour valeur la somme de l'entier $a_{\mathcal{N}}$ et du coefficient binomial du nœud fils.

$$\begin{cases} a_{\mathcal{N}} = a_{\mathcal{N}_{\text{parent}}} \\ b_{\mathcal{N}} = a_{\mathcal{N}} + \binom{i_{f\mathcal{N}}}{j_{f\mathcal{N}}} \end{cases}$$

En allant vers la gauche dans l'arbre. L'entier $a_{\mathcal{N}}$ du nœud courant $\mathcal{N}$ obtenu après le mouvement vers la gauche, possède la même valeur que l'entier $b_{\mathcal{N}_{\text{voisin}_\text{d}}}$.Tandis que l'entier $b_{\mathcal{N}}$ possède la somme de l'entier $b_{\mathcal{N}}$ et le coefficient binomial du nœud fils $\mathcal{N}_{\text{fils}}$.

$$\begin{cases} a_{\mathcal{N}} = b_{\mathcal{N}_{\text{voisin}_\text{d}}} \\ b_{\mathcal{N}} = b_{\mathcal{N}} + \binom{i_{f\mathcal{N}}}{j_{f\mathcal{N}}} \end{cases}$$

Pour un visuel de l'arbre $\mathcal{P}$ après le calcul des intervalles, voir l'exemple 2.11.

Maintenant que nous avons expliqué comment le calcul des intervalles se fait, nous décrivons le parcours de l'arbre $\mathcal{P}$.

Description du parcours :

Le parcours dans l'arbre $\mathcal{P}$ se fait de la manière suivante :

— On commence toujours par la racine $\mathcal{R}$.

— La racine $\mathcal{R}$ a pour intervalle associé $[a_{\mathcal{R}}, b_{\mathcal{R}}[$. Nous effectuons les initialisations suivantes : $a_{\mathcal{R}} = 0$, $b_{\mathcal{R}} = \binom{n+k-1}{k}$. De plus nous introduisons deux tables appelées min et coefficients tel que $\text{min}_1 = 0$ et $\text{coefficients}_1 = 0$.

— La continuité du parcours dépend de l'appartenance du numéro x de la séquence à l'intervalle associé au nœud courant $\mathcal{N}$.

Si le numéro de la séquence appartient à l'intervalle du nœud courant $\mathcal{N}$, alors, il faut descendre en profondeur dans l'arbre, sinon, il faut aller vers le nœud voisin de gauche $\mathcal{N}_{\text{voisin}_g}$.

— Soit $\mathcal{N}$ le dernier nœud visité au niveau t de l'arbre $\mathcal{P}$ ayant pour intervalle associé $[a_{\mathcal{N}}, b_{\mathcal{N}}[$. Si le parcours de l'arbre continue jusqu'au nœud qui se trouve à l'extrémité gauche au niveau $t+1$ de l'arbre $\mathcal{P}$, alors certainement x est un élément de l'intervalle associé à ce nœud. En effet, nous savons que x appartient à l'intervalle $[a_{\mathcal{N}}, b_{\mathcal{N}}[$ et la réunion de tous les intervalles visités au niveaux $t+1$ couvre l'intervalle $[a_{\mathcal{N}}, b_{\mathcal{N}}[$, donc si aucun des intervalles précédents ne contient x, alors l'intervalle associé au dernier nœud (qui se trouve à l'extrémité gauche au niveau $t+1$) le contient.

— En parcourant l'arbre, si la valeur de l'entier $a_{\mathcal{N}}$ du nœud courant change, c'est-à-dire, si le mouvement effectué est celui d'aller vers la gauche, nous effectuons les opérations suivantes :

  — Ajouter la valeur de l'entier $b_{\mathcal{N}} - a_{\mathcal{N}}$ dans une table appelée min.

  — Ajouter le coefficient binomial du nœud voisin de droite N_{voisin_d} dans une table appelée coefficients.

— Ces informations serviront ultérieurement pour calculer les successeurs d'une séquence.

— On note r la taille des tables min et coefficients obtenues à la fin du parcours. On remarque que r est égale à un plus le nombre de fois où la valeur de l'entier $a_{\mathcal{N}}$ change (l'initialisation $a_{\mathcal{R}} = 0$ ne compte pas comme un changement).

— Le parcours s'arrête lorsque le numéro de la séquence dont nous cherchons les successeurs est l'unique élément de l'intervalle du nœud courant.

Nous rappelons à partir de la propriété 2.5 de la section 2, page 19 que les séquences de paramètres N et k, où N est la taille du bitamp et k son poids, ont des numéros dans l'intervalle $[a_{N,k-1}, a_{N,k}[$. Leurs successeurs ont des numéros dans l'intervalle $[a_{N,k}, a_{N,k+1}[$.

Pour éviter d'opérer avec des nombres trop grands pour les calculs, on utilise une numérotation modifiée de sorte que l'intervalle commence par 0. On translate alors les numéros de séquence par $a_{N,k-1}$ et les successeurs par $a_{N,k}$.

Exemple 2.11. Soient $n = 5$ et $k = 2$.

Cherchons les successeurs de la séquence numéro 11 (par rapport à la numérotation translatée).

La figure IV.5 ci-dessous, présente une partie de l'arborescence $\mathcal{P}$ pour le parcours associé à la séquence numéro 11. Les pas du parcours sont indiqués par des flèches. Pour ne pas encombrer la présentation, nous ne dessinons pas toutes les branches de l'arborescence. Pour la présentation de l'arborescence complète voir la figure IV.3 ci-dessus. Pour les sommets visités lors du parcours, nous indiquons les intervalles associés en couleur bleue, et en rouge l'information si le numéro 11 appartient à l'intervalle ou pas.

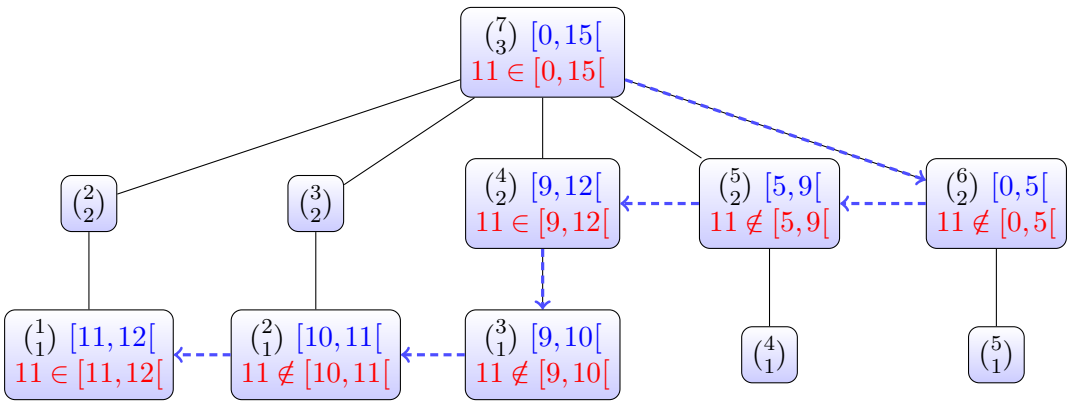


FIGURE IV.5 – Exemple de recherche dans l'arborescence

Suivant la description donnée précédemment, nous avons retenu les valeurs des tables coefficients et min. nous présentons ci-dessous leur contenu pour cet exemple.

i	min_i	coefficients_i
1	0	0
2	$5-0=5=\binom{5}{1}$	$\binom{6}{2}=15$
3	$12-9=4=\binom{4}{1}$	$\binom{5}{2}=10$
4	$11-10=1=\binom{2}{0}$	$\binom{3}{1}=3$
5	$12-11=1=\binom{1}{0}$	$\binom{2}{1}=2$

Dans l'exemple, les numéros des 5 successeurs de la séquence numéro 11 calculés avec la formule 2.4 sont :

$S_1 = 0 - 0 + 11 = 11$

$S_2 = \binom{6}{2} - \binom{5}{1} + 11 = \binom{5}{2} + 11 = 21$

$S_3 = \binom{5}{2} - \binom{4}{1} + 21 = \binom{4}{2} + 21 = 27$

$S_4 = 3 - 1 + 27 = \binom{3}{1} - \binom{2}{0} + 27 = \binom{2}{1} + 27 = 29$

$S_5 = 2 - 1 + 29 = \binom{2}{1} - \binom{1}{0} + 29 = \binom{1}{1} + 29 = 30$

L'ensemble des successeurs est donc : $\{11, 21, 27, 29, 30\}$. On peut aussi les présenter comme une réunion d'intervalles $[11] \cup [21] \cup [27] \cup [29, 30]$.

Remarque 2.12. Par abus, on appellera parfois successeurs du numéro x, l'ensemble des numéros des séquences qui sont des successeurs de la séquence numéro x.

Remarque 2.13.
- Notons que la quantité *coefficients*$_i$ − *min*$_i$ est simple à calculer. En effet par construction de l'arbre $\mathcal{P}$, si coefficients$_i$ est égale à $\binom{n'}{p'}$ alors min$_i$ est égale à $\binom{n'-1}{p'-1}$
  Donc :
  coefficients$_i$ − min$_i$ $= \binom{n'}{p'} - \binom{n'-1}{p'-1} = \binom{n'-1}{p'}$
  Les entier t et j de la formules 2.2 recherchés sont donc $j = n' - 1$ et $t = p'$.
- Soit $\mathcal{N}$ un nœud courant étiqueté par le coefficient $\binom{i_\mathcal{N}}{j_\mathcal{N}}$ et qui a pour intervalle $[a_\mathcal{N}, b_\mathcal{N}[$. Si *coefficients*$_i$ $= \binom{i_\mathcal{N}}{j_\mathcal{N}}$ et *min*$_i$ $= b_N - a_N$, alors coefficients$_i$ − min$_i$=$\binom{i_\mathcal{N}-1}{j_\mathcal{N}}$. Autement dit la quantité *coefficients*$_i$ − *min*$_i$ est égale au coefficient binomial du nœud voisin de gauche du nœud $\mathcal{N}$.

Cette méthode conduit au pseudo code décrit dans le paragraphe suivant.

Remarque 2.14.
- Afin de réduire la complexité du calcul, les coefficients binomiaux ne sont pas recalculés entièrement à chaque parcours mais sont simplement mis à jour par la formule : $\binom{n}{k} = \binom{n-1}{k-1} \times \frac{n}{k}$.
- Dans le programme, les valeurs des entiers coefficients$_i$ et min$_i$ sont accumulées.

b) Pseudo code : successeur d'une séquence

Soit n l'ordre du groupe G et k son exposant.

Cherchons les numéros des successeurs d'une séquence S représentée par son numéro x.

Algorithme 3 successeur d'un élément

Entrée

x : le numéro d'une séquence de k termes dans un alphabet de taille n

Sortie

dim : un entier représentant le nombre des successeurs consécutifs de l'entier x

S : L'ensemble des numéros des successeurs de l'entier x

Variables

min et max : les bornes d'intervalle

b : les valeurs des coefficients binomiaux

coefficients : les valeurs des coefficients binomiaux quand la valeur de min change

Initialisation

$m \leftarrow n + k$

$K \leftarrow k + 1$

min $\leftarrow 0$

coefficients $\leftarrow 0$

$i \leftarrow 1$

dim $\leftarrow n$

$S[0] \leftarrow x$

Mise à jour des coefficients binomiaux : $b \leftarrow \binom{m}{K}$

max $\leftarrow b$

1: **procédure**
2: **tant que** (l'entier x n'est pas atteint : $x >$ min) **faire**
3: **si** (l'entier x est dans l'intervalle $[\text{min}, \text{max}[$: $x <$ max) **alors**
4: Descendre en profondeur dans l'arbre
5: La mise à jour du coefficient binomial : $b \leftarrow b \times \frac{K}{m}$
6: $K \leftarrow K - 1$
7: $m \leftarrow m - 1$
8: max $\leftarrow$ min $+ b$
9: Continuer
10: **sinon** (tant que l'entier x est hors de l'intervalle)
11: aller vers le nœud voisin de gauche
12: la mise à jour de coefficients :
13: coefficients $\leftarrow$ coefficients $+ b \times \frac{(m+1)}{(k+1)}$
14: La mise à jour du coefficient binomial : $b \leftarrow b \times \frac{(m-k)}{m}$
15: $m \leftarrow m - 1$
16: min $\leftarrow$ max
17: max $\leftarrow$ min $+ b$
18: $i \leftarrow i + 1$
19: dim $\leftarrow$ dim $- 1$
20: Le calcul des successeurs :
21: $S[i] \leftarrow$ coefficients $-$ min $+ x$
22: Continuer
Le calcul des successeurs consécutifs :
23: **pour** j allant de $i + 1$ jusqu'à n **faire**
24: $S[j] \leftarrow S[j - 1] + 1$

c) Le calcul des successeurs d'un intervalle de séquences

La méthode immédiate pour déterminer l'ensemble des successeurs d'un intervalle est de calculer la réunion des successeurs de tous les éléments de cet intervalle. Cette méthode demande de parcourir tout l'intervalle, ce qui a une complexité prohibitive.

Nous présentons maintenant une méthode permettant de calculer les successeurs d'un intervalle $[a,b]$ uniquement à partir des successeurs de a et de b.

Cette méthode repose sur le théorème suivant.

Théorème 2.15. *Posons $b = a + j$, pour i allant de 1 jusqu'à n, soient $S_i(a)$ les successeurs de l'entier a, et $S_i(b)$ les successeurs de l'entier b.*

Si $S_i(b) - S_i(a) \neq j$, alors il existe un entier t qu'on peut déterminer à partir des entiers a et b tel que les successeurs de l'intervalle $[a,b]$ est donné par la réunion suivante des intervalles :

$$[S_1(a), S_1(b)] \cup \cdots \cup [S_{r_t-1}(a), S_{r_t-1}(b)] \cup [S_{r_t}(a), S_{r_t}(t-1)] \cup [S_{r_a}(a), S_n(b)]$$

Si $S_i(b) - S_i(a) = j$, alors les successeurs de l'intervalle $[a,b]$ est donné par la réunion suivante des intervalles :

$$[S_1(a), S_1(b)] \cup [S_2(a), S_2(b)] \cup \cdots \cup [S_{r_a-1}(a), S_{r_a-1}(b)] \cup [S_{r_a}(a), S_n(b)]$$

Notre but dans cette section est de démontrer ce théorème.

Plus précisément, nous calculons d'abord les successeurs de a et de b avec la méthode décrite dans la section précédente (voir formules 2.4 et 2.5 page 68).

Trois situations se présentent pour calculer les successeurs d'un intervalle $[a,b]$. Chaque situation permet de donner un sous-ensemble de l'ensemble des successeurs de $[a,b]$. La réunion des trois situations donne d'une manière exhaustive l'ensemble de tous les successeurs de $[a,b]$.

Les deux première situations représentent les cas pour lesquels les sous-ensembles des successeurs de $[a,b]$ forment toujours un intervalle. La troisième situation représente le cas ou le sous-ensemble des successeurs de $[a,b]$ ne forme pas toujours un intervalle.

Avant de présenter chacune des situations ainsi que leur preuves, nous introduisons d'abord certaines notations et remarques.

— Nous notons par $S_i(a)$ le numéro du $i^{\text{ème}}$ successeur de a avec i allant de 1 jusqu'à n, où n est l'ordre du groupe.

— Soit a un entier désignant le numéro dont nous cherchons les successeurs. Nous notons par r_a l'entier à partir duquel les successeurs de a sont consécutifs. Rappelons qu'à partir de r_a et jusqu'à n, les successeurs de a sont consécutifs voir formule 2.5. Par conséquent, il y a $n - r_a + 1$ numéros consécutifs.

Pour illustrer chacune des situations, nous donnons à la fin de cette section l'exemple 2.24, qui donne les successeurs de l'intervalle $[7, 11]$ pour le cas $n = 5$ et $k = 3$.

Proposition 2.16. *Le $r_{a+1}{}^{ème}$ successeur de $a+1$ est le numéro suivant du $n^{ème}$ successeur de a, c'est-à-dire,*

$$S_{r_{a+1}}(a+1) = S_n(a) + 1 \tag{2.7}$$

Preuve. Soit $\mathcal{N}^{'}$ un nœud étiqueté par le coefficient $\binom{i_{\mathcal{N}'}}{j_{\mathcal{N}'}}$ et ayant pour intervalle $[a_{\mathcal{N}'}, b_{\mathcal{N}'}[$.

Supposons qu'à partir de la racine de l'arborescence $\mathcal{P}$ et jusqu'au nœud $\mathcal{N}^{'}$, le chemin du parcours dans $\mathcal{P}$ est identique pour les entiers a et $a+1$. Ensuite, le chemin du parcours dans l'arborescence $\mathcal{P}$ pour les entiers a et $a+1$ se branche à partir de ce nœud $\mathcal{N}^{'}$ de sorte que :

— Pour l'entier a nous effectuons un mouvement vers le bas car l'entier a est dans l'intervalle $[a_{\mathcal{N}'}, b_{\mathcal{N}'}[$.

— Pour l'entier $a+1$ nous effectuons un mouvement vers la gauche car $a+1 \geq b_{\mathcal{N}'}$.

Ceci implique qu'au niveau du nœud $\mathcal{N}^{'}$ on a :

$$a_{\mathcal{N}'} \leq a < b_{\mathcal{N}'} \text{ et } a+1 = b_{\mathcal{N}'}$$

À partir de la racine de $\mathcal{P}$ et jusqu'au nœud $\mathcal{N}^{'}$, nous désignons par $r^{'}$ le nombre de fois que nous effectuons un mouvement vers la gauche. Rappelons que l'entier $r^{'}$ est aussi la taille des tables min et coefficients.

D'après la méthode de calcul des successeurs d'un numéro vue précédemment, voir la formule 2.4 page 68 on sait que pour i allant de 1 jusqu'à $r^{'}$, on a :

$$S_i(a+1) = S_i(a) + 1 \tag{2.8}$$

Pour l'entier $a+1$, on peut calculer $S_{r'+1}(a+1)$, car le mouvement qu'on effectue à partir du nœud $\mathcal{N}^{'}$ est vers la gauche.

Soit $\mathcal{N}^{''}$, le nœud voisin de gauche du nœud $\mathcal{N}^{'}$. Le nœud $\mathcal{N}^{''}$ est étiqueté par le coefficient $\binom{i_{\mathcal{N}'}-1}{j_{\mathcal{N}'}}$ et a pour intervalle $[a_{\mathcal{N}''}, b_{\mathcal{N}''}[$.

D'après la remarque 2.13 et la formule 2.4 on a :

$$S_{r'+1}(a+1) = \binom{i_{\mathcal{N}'}-1}{j_{\mathcal{N}'}} + S_{r'}(a+1)$$

Or d'après la formule 2.8, $S_{r'}(a+1) = S_{r'}(a) + 1$. On a :

$$S_{r'+1}(a+1) = \binom{i_{\mathcal{N}'}-1}{j_{\mathcal{N}'}} + S_{r'}(a) + 1 \tag{2.9}$$

De plus, on a $r_{a+1} = r^{'} + 1$ car $a_{\mathcal{N}''} = a+1$, donc les seuls mouvements qui peuvent être effectués à partir du nœud $\mathcal{N}^{''}$ est la descente. On a alors :

$$S_{r_{a+1}}(a+1) = \binom{i_{\mathcal{N}'}-1}{j_{\mathcal{N}'}} + S_{r'}(a) + 1 \tag{2.10}$$

Pour l'entier a, On sait que $a = b_{\mathcal{N}'} - 1$. Pour atteindre le nœud contenant l'intervalle $[a, a+1[$, nous distinguons deux situations :

- Première situation : si la taille de l'intervalle du nœud $\mathcal{N}^{'}$ est différente de 1. Autrement dit, si $j_{\mathcal{N}'} > 1$. Rappelons que le nœud $\mathcal{N}^{'}$ est étiqueté par le coefficient $\binom{i_{\mathcal{N}'}}{j_{\mathcal{N}'}}$.
- Deuxième situation : si la taille de l'intervalle du nœud $\mathcal{N}^{'}$ est égale à 1. Autrement dit, si $j_{\mathcal{N}'} = 1$.

Etudions d'abord la première situation.

Si $j_{\mathcal{N}'} > 1$, en descendant dans l'arborescence $\mathcal{P}$ on tombe sur le nœud fils du nœud $\mathcal{N}^{'}$. On sait que ce nœud fils est par définition étiqueté par le coefficient $\binom{i_{\mathcal{N}'}-1}{j_{\mathcal{N}'}-1}$. Comme nous l'avons vu précédemment, on peut ranger les nœuds de l'arborescence $\mathcal{P}$ dans une table (voir table IV.4 page 67). Soient i le nombre de lignes, et $j = n - r^{'}$ le nombre de colonnes, où le nœud fils $\binom{i_{\mathcal{N}'}-1}{j_{\mathcal{N}'}-1}$ est rangé dans la table.

Comme $a = b_{\mathcal{N}'} - 1$ alors le nœud contenant l'intervalle $[a, a+1[$ se trouve à la colonne 0 et à la ligne i de la table. En particulier, l'entier a n'appartient pas à tous les intervalles des nœuds précédents, c'est-à-dire, les nœuds se trouvant à la $i^{\text{ème}}$ ligne et la $t^{\text{ème}}$ colonne avec t allant de j jusqu'à 1.

Par conséquent, pour atteindre le nœud contenant l'intervalle $[a, a+1[$, on doit encore effectuer $n - r^{'}$ mouvement vers la gauche.

On a alors :

$S_n(a) = S_{r'}(a) + \sum_{s=1}^{n-r'} \binom{i_{\mathcal{N}'}-1-s}{j_{\mathcal{N}'}-1} = S_{r'}(a) + \binom{i_{\mathcal{N}'}-1}{j_{\mathcal{N}'}}$.

D'où $S_{r_{a+1}}(a+1) = S_n(a) + 1$.

Etudions maintenant la deuxième situation.

Si $j_{\mathcal{N}'} = 1$, alors le nœud $\mathcal{N}'$ a un intervalle contenant le seul élément a, tel que $a_{\mathcal{N}'} = a$ et $b_{\mathcal{N}'} = a+1$.

Dans cette situation, l'entier $r' = r_a$, et la procédure du parcours dans l'arborescence $\mathcal{P}$ s'arrête pour l'entier a.

Pour l'entier $a+1$, nous devons effectuer un mouvement vers la gauche. Nous tombons alors sur le nœud voisin de gauche du nœud $\mathcal{N}$ qu'on notera $\mathcal{N}''$ ayant pour intervalle $[a_{\mathcal{N}''}, b_{\mathcal{N}''}[$ tel que $a_{\mathcal{N}''} = a+1$.

Dans ce cas, $r_{a+1} = r' + 1 = r_a + 1$. Et le parcours pour l'entier $a+1$ s'arrête.

On sait que $S_{r_{a+1}}(a+1) = \binom{i_{\mathcal{N}'}-1}{j_{\mathcal{N}'}} + S_r(a) + 1$.

Or, on est dans le cas où $j_{\mathcal{N}'} = 1$ et on sait que nous avons $n - r_a + 1$ successeurs consécutifs pour l'entier a. De plus, on se retrouve à la $n - r_a$ ème colonne donc, $n - r_a = \binom{i_{\mathcal{N}'}-1}{j_{\mathcal{N}'}}$.

On a alors :

$S_n(a) = S_{r_a}(a) + \binom{i_{\mathcal{N}'}-1}{j_{\mathcal{N}'}}$. D'où $S_{r_{a+1}}(a+1) = S_n(a) + 1$

□

Proposition 2.17. *Le numéro du $i^{\text{ème}}$ successeur de la séquence numéro $a+1$ est supérieur ou égal au numéro du $i^{\text{ème}}$ successeur de la séquence numéro a plus un, c'est-à-dire,*

$$\forall\ 0 \leq i \leq n\ ,\ S_i(a+1) \geq S_i(a) + 1$$

Preuve. Soit $\mathcal{N}$ un nœud étiqueté par le coefficient $\binom{i_{\mathcal{N}}}{j_{\mathcal{N}}}$ et ayant pour intervalle $[a_{\mathcal{N}}, b_{\mathcal{N}}[$.

Supposons qu'à partir de la racine de l'arborescence $\mathcal{P}$ et jusqu'au nœud $\mathcal{N}$, le chemin du parcours dans $\mathcal{P}$ est identique pour les entiers a et $a+1$. Ensuite, le chemin du parcours dans l'arborescence $\mathcal{P}$ pour les entiers a et $a+1$ se branche à partir de ce nœud $\mathcal{N}$, tel qu'on a :

$$a_{\mathcal{N}} \leq a < b_{\mathcal{N}} \text{ et } a+1 = b_{\mathcal{N}}$$

Nous désignons par l'entier r, le nombre de fois que effectuons un mouvement vers la gauche (à partir de la racine jusqu'au nœud $\mathcal{N}$).

Nous distinguons deux cas :

Si $a_{\mathcal{N}} = a$:

Dans ce cas, le nœud $\mathcal{N}$ a un intervalle contenant un seul élément : $[a, a+1[$.

Il est clair que $r = r_a$ et vu que les valeurs des entiers coefficients$_i$ et min$_i$ sont identiques, pour i allant de 0 jusqu'à r_a, on a $S_i(a+1) - S_i(a) = 1$. De plus $r_{a+1} = r_a + 1$, or on sait que pour i allant de 0 à $n-1$ on a $S_n(a) > S_i(a)$ et d'après la proposition 2.7 on a $S_{r_{a+1}}(a+1) = S_n(a) + 1$, donc $S_{r_{a+1}}(a+1)$ est plus grand que tous les successeurs de a. D'où pour tout i allant de 1 jusqu'à n on a

$S_i(a+1) \geq S_i(a) + 1$.

Si $a_{\mathcal{N}} < a$:

Dans ce cas, on a $a = b_{\mathcal{N}} - 1$, pour l'entier a nous effectuons un mouvement vers le bas puis des mouvement vers la gauche jusqu'à atteindre le nœud se trouvant à l'extrémité gauche. On a alors $r_a = n$. On sait que pour i allant de 0 jusqu'à r, on a $S_i(a+1) - S_i(a) = 1$ et d'après la proposition 2.7 on a $S_{r_{a+1}}(a+1) = S_n(a) + 1$. On en déduit que $S_i(a+1) \geq S_i(a) + 1$.

□

Proposition 2.18. *Le numéro du $i^{\text{ème}}$ successeur de la séquence numéro $a+j$ est supérieur ou égal au numéro du $i^{\text{ème}}$ successeur de la séquence numéro a plus j, c'est-à-dire,*

$$\forall\ 0 \leq i \leq n\ ,\ S_i(a+j) \geq S_i(a) + j$$

Preuve. Soient a et b deux entiers tel que $b = a + j$

Nous montrons par récurrence que

$$\forall\ 0 \leq i \leq n\ ,\ S_i(b) \geq S_i(a) + j. \tag{2.11}$$

Pour $j = 1$ on a $\forall\ 0 \leq i \leq n\ ,\ S_i(a+1) \geq S_i(a) + 1$. Cette inégalité est vraie par 2.17.

Montrons que l'inégalité 2.11 est vraie au rang $j+1$.

Soit $c = b + 1$, on sait d'après la proposition 2.17 que $\forall\ 0 \leq i \leq n\ ,\ S_i(c) \geq S_i(b) + 1$ on a alors $S_i(a+j+1) \geq S_i(a+j) + 1$. D'après l'hypothèse de récurrence on en déduit que $S_i(a+j+1) \geq S_i(a+j) + 1 \geq S_i(a) + j + 1$. □

Nous présentons maintenant les deux situations pour lesquelles les sous-ensembles des successeurs de $[a, b]$ forment toujours un intervalle.

Proposition 2.19 (Première situation)**.** *Soient a et b deux entiers.*

Posons $b = a + j$.

Pour i allant de 1 jusqu'à $r_a - 1$. Supposons que la différence entre le numéro du $i^{\text{ème}}$ successeur de b et le $i^{\text{ème}}$ successeur de a est égale à $b - a = j$, c'est-à-dire, $S_i(a + j) - S_i(a) = j$, alors l'intervalle $[S_i(a), S_i(b)]$ est inclus dans l'ensemble des successeurs de l'intervalle $[a, b]$.

Preuve. Soient a et b deux entiers tel que $b = a + j$.

Soit $\mathcal{N}$ un nœud de l'arborescence $\mathcal{P}$.

Supposons qu'à partir de la racine de $\mathcal{P}$ et jusqu'au nœud $\mathcal{N}$, le parcours dans $\mathcal{P}$ est identique pour les entiers a et $a + s$, tel que l'entier s varie entre 1 jusqu'à j.

D'après la méthode du calcul des successeurs vue dans la section 2-), page 65, les valeurs des entiers *coefficients*$_i$ et *min*$_i$ pour pour $S_i(a)$ et $S_i(a + s)$ sont identiques. Par conséquent, $S_i(a + s) = S_i(a) + s$, d'où l'intervalle $[S_i(a), S_i(b)]$ est inclus dans l'ensemble des successeurs de l'intervalle $[a, b]$. □

Proposition 2.20 (Deuxième situation)**.** *Soient a et b deux entiers représentant des numéros de séquences.*

Le $r_a^{\text{ème}}$ successeur de a et le $n^{\text{ème}}$ successeur de b forment les extrémités d'un intervalle noté $[S_{r_a}(a), S_n(b)]$. Cet intervalle est inclus dans l'ensemble des successeurs de l'intervalle $[a, b]$.

Preuve. Pour cela montrons par récurrence que pour $b = a + j$, le $n^{\text{ème}}$ successeur de b est la somme du $r^{\text{ème}}$ successeur de a et l'accumulation des nombres de successeurs consécutifs de tous les numéros entre a et b.

$$S_n(a + j) = S_{r_a}(a) + \sum_{k=0}^{j} (n - r_{a+k} + 1) - 1 \tag{2.12}$$

Pour $j = 0$, on a : $S_n(a) = S_{r_a}(a) + n - r_a$. Cette égalité est vraie par 2.6

montrons que l'égalité 2.12 est vraie au rang $j + 1$.

On sait que d'après l'égalité 2.7 on a $S_{r_{a+1}}(a + 1) = S_n(a) + 1$, de plus on sait que $S_n(a + 1) = S_{r_{a+1}}(a + 1) + (n - r_{a+1})$. On a alors

$S_n(a + 1) = S_n(a) + (n - r_{a+1}) + 1$

D'où

$S_n(a + j + 1) = S_n(a + j) + (n - r_{a+j+1} + 1)$

Donc

$S_n(a+j+1) = S_{r_a}(a) + \sum_{k=0}^{j}(n - r_{a+k} + 1) - 1 + (n - r_{a+j+1}) = S_{r_a}(a) + \sum_{k=0}^{i+1}(n - r_{a+k} + 1) - 1$.

□

Avant de passer à la troisième situation, nous énonçons la remarque suivante.

Remarque 2.21. Les premiers successeurs $S_1(a)$ jusqu'à $S_1(a+j)$ sont toujours consécutifs. En effet, d'après la formule 2.4 du calcul des successeurs. Les valeurs de min_1 et $coefficient_1$ sont toujours nulles.

Passons maintenant à la situation pour laquelle le sous-ensemble des successeurs de $[a, b]$ ne forme pas toujours un intervalle.

Proposition 2.22 (Troisième situation)**.** *Pour i allant de 2 jusqu'à $r_a - 1$*

Si $S_i(a+j) - S_i(a) \neq j$, alors, il existe un entier j' strictement inférieur que j tels que :

- *l'intervalle $[S_i(a), S_i(a+j'-1)]$ est inclus dans l'ensemble des successeurs des séquences dont les numéros sont dans l'intervalle $[a, a+j]$.*
- *et l'ensemble $\{S_i(a), \ldots, S_i(a+j)\}$ est inclus dans la réunion des intervalles $[S_i(a), S_i(a+j'-1)] \cup [S_r(a), S_n(a+j)]$.*

Remarque 2.23.
- Rappelons que l'intervalle $[S_{r_a}(a), S_n(a+j)]$ est inclus dans l'ensemble des successeurs des séquences dont les numéros sont dans l'intervalle $[a, a+j]$ par la proposition 2.20.
- Le i varie à partir de 2 et non pas à partir de 1, car d'après la remarque 2.21, $S_1(a+j) - S_1(a)$ est toujours égale à j.

Preuve. Si $S_i(a+j) - S_i(a) \neq j$, cela signifie qu'il existe un plus petit entier t tel que $S_i(t) > S_i(t-1) + 1$. Dans ce cas, $S_i(t) = S_{r_t}(t)$ et l'entier t est déterminer d'une manière précise lors du parcours de l'arbre $\mathcal{P}$

En effet, soit $\mathcal{N}$ un nœud étiqueté par le coefficient $\binom{i_{\mathcal{N}}}{j_{\mathcal{N}}}$ et ayant pour intervalle $[a_{\mathcal{N}}, b_{\mathcal{N}}[$.

Supposons qu'à partir de la racine de $\mathcal{P}$ et jusqu'au nœud $\mathcal{N}$, le parcours dans $\mathcal{P}$ est identique pour les entiers a et $a+j$. Ensuite, le chemin du parcours dans l'arborescence $\mathcal{P}$ pour les entiers a et $a+j$ se branche à partir de ce nœud $\mathcal{N}$.

Ceci implique qu'au niveau du nœud $\mathcal{N}$ on a : $a \in [a_{\mathcal{N}}, b_{\mathcal{N}}[$ et $a + j \geq b_{\mathcal{N}}$. L'entier t est alors égale à $b_{\mathcal{N}}$.

De plus on sait que $S_{r_a}(a) < S_{r_t}(t)$ alors :

$[S_{r_t}(t), S_n(a+j)] \subset [S_{r_a}(a), S_n(a+j)]$.

Le cas de l'intervalle $[S_{r_a}(a), S_n(a+j)]$ est déjà inclus dans la proposition 2.19.

Il ne reste plus qu'inclure $[S_i(a), S_i(t-1)]$ aux successeurs des séquences dont les numéros sont dans l'intervalle $[a,b]$.

La valeur de l'entier $S_i(t-1)$ est donnée par :

$S_i(t-1) = S_i(a) + t - 1 - a.$

□

Nous concluons en donnant une démonstration du théorème 2.15

Preuve. Les trois situations 2.19, 2.20 et 2.22 permettent de couvrir tous les successeurs des séquences dont les numéros sont dans l'intervalle $[a,b]$.

En effet, d'après la troisième situation 2.22 on sait que si $S_i(b) - S_i(a) \neq b - a$ alors il existe un entier t tel que $a < t < b$ où $S_i(t) = S_{r_t}(t)$.

Vu que $[a,b] = [a,t-1] \cup [t,b]$, considérons d'abord les successeurs de l'intervalle $[t,b]$.

D'après la deuxième situation 2.20, l'intervalle $[S_{r_t}(t), S_n(b)]$ est inclus dans l'ensemble des successeurs des séquences dont les numéros sont dans l'intervalle $[t,b]$. De plus, on sait que pour j allant de $t+1$ jusqu'à b, l'entier r_t est supérieurs ou égale au r_j, et que le $r_t^{\text{ème}}$ successeur de la séquence numéro t est plus petit que tous les $r_j^{\text{ème}}$ successeurs de la séquence numéro j, et le $n^{\text{ème}}$ successeur de la séquence numéro b est supérieur à tous les successeurs entre les séquences numéros t et b. De plus, pour i allant de 1 jusqu'à $r_t - 1$, la différence des successeurs $S_i(b) - S_i(t) = b - t$. Donc la réunion d'intervalles :

$[S_1(t), S_1(b)] \cup \ldots\ldots \cup [S_{r_t-1}(t), S_{r_t-1}(b)] \cup [S_{r_t}(t), S_n(b)]$

couvre tous les successeurs entre t et b.

De même entre les entiers a et $t-1$, la réunion d'intervalles :

$[S_1(a), S_1(t-1)] \cup \cdots \cup [S_{r_t}(a), S_{r_t}(t-1)] \cup [S_{r_a}(a), S_n(t-1)]$

couvre tous les successeurs entre a et $t-1$.

D'où, la réunion d'intervalles :

$$[S_1(a), S_1(b)] \cup \cdots \cup [S_{r_t-1}(a), S_{r_t-1}(b)] \cup [S_{r_t}(a), S_{r_t}(t-1)] \cup [S_{r_a}(a), S_n(b)]$$

couvre tous les successeurs entre a et b.

□

Ci-dessous nous donnons un exemple du calcul des successeurs d'un intervalle $[a,b]$ à partir des successeurs de a et de b.

Exemple 2.24. Le tableau ci-dessous représente tous les numéros de séquences de paramètre $n = 5$ et $k = 3$ ainsi que leur successeurs.

Les successeurs consécutifs pour chaque numéro de séquence sont coloriés.

numéro de séquence	successeurs				
0	0	1	2	3	4
1	1	5	6	7	8
2	2	6	9	10	11
3	3	7	10	12	13
4	4	8	11	13	14
5	5	15	16	17	18
6	6	16	19	20	21
7	7	17	20	22	23
8	8	18	21	23	24
9	9	19	25	26	27
10	10	20	26	28	29
11	11	21	27	29	30
12	12	22	28	31	32
13	13	23	29	32	33
14	14	24	30	33	34
15	15	35	36	37	38
16	16	36	39	40	41
17	17	37	40	42	43
18	18	38	41	43	44
19	19	39	45	46	47
20	20	40	46	48	49
21	21	41	47	49	50
22	22	42	48	51	52
23	23	43	49	52	53
24	24	44	50	53	54
25	25	45	55	56	57
26	26	46	56	58	59
27	27	47	57	59	60
28	28	48	58	61	62
29	29	49	59	62	63
30	30	50	60	63	64
31	31	51	61	65	66
32	32	52	62	66	67
33	33	53	63	67	68
34	34	54	64	68	69

TABLEAU IV.3 – Numéros des séquences de paramètre $n = 3$ et $k = 3$

Cherchons les successeurs de l'intervalle $[7,11]$.

Nous calculons d'abord les successeurs de 7 puis les successeurs de 11.

Les successeurs de l'entier 7 :

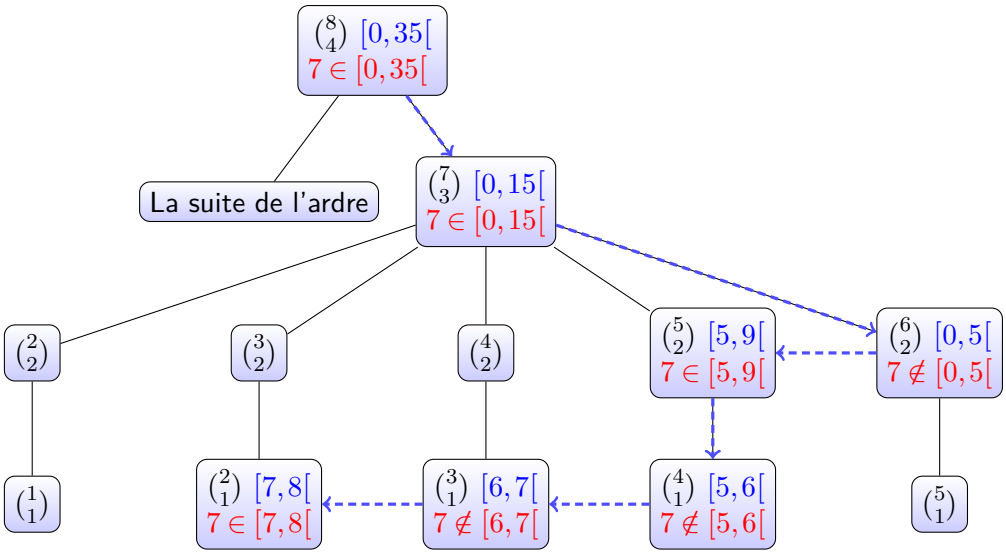


FIGURE IV.6 – Exemple du parcours pour déterminer les successeurs d'un entier a

i	$\min_i$	coefficients_i
1	0	0
2	$5-0=5$	$\binom{6}{2}=15$
3	$6-5=1$	$\binom{4}{1}=4$
4	$7-6=1$	$\binom{3}{1}=3$

Les successeurs de 7 sont alors :

$S_1(7) = 0 - 0 + 7 = 7$

$S_2(7) = 15 - 5 + 7 = 17$

$S_3(7) = 4 - 1 + 17 = 20$

$S_4(7) = 3 - 1 + 20 = 22$

$S_5(7) = 22 + 1 = 23$

Notons, que dans ce cas, l'entier t vaut 9.

Les successeurs de l'entier 11 :

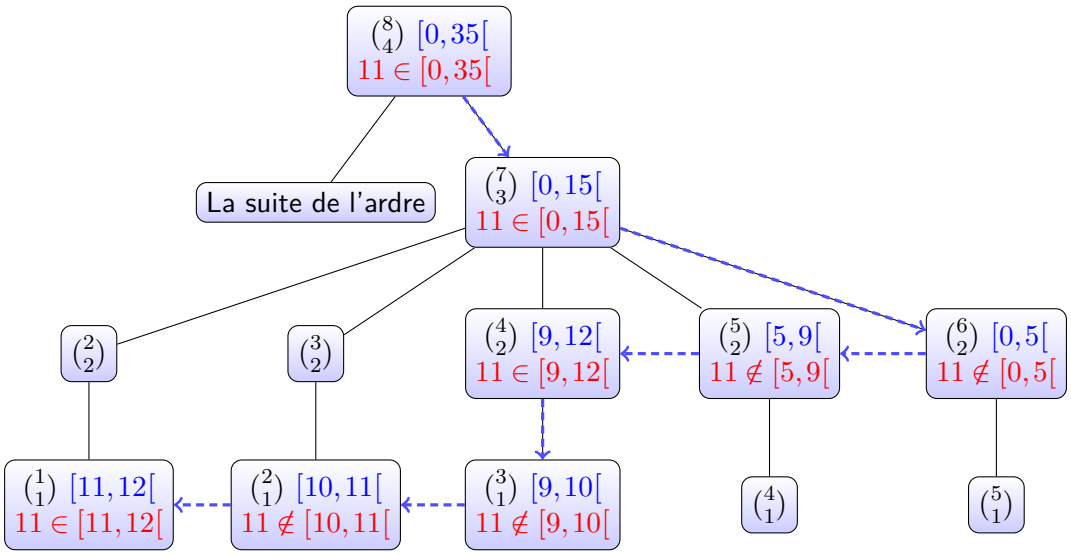


FIGURE IV.7 – Exemple du parcours pour déterminer les successeur d'un entier b

i	$\min_i$	coefficients_i
1	0	0
2	$5-0=5$	$\binom{6}{2}=15$
3	$9-5=4$	$\binom{5}{2}=10$
4	$10-9=1$	$\binom{3}{1}=3$
5	$11-10=1$	$\binom{2}{1}=2$

Les successeurs de 11 sont alors :

$S_1(11) = 0 - 0 + 11 = 11$

$S_2(11) = 15 - 5 + 11 = 21$

$S_3(11) = 10 - 4 + 21 = 27$

$S_4(11) = 3 - 1 + 27 = 29$

$S_5(7) = 2 - 1 + 29 = 30$

Notons que, dans cet exemple $S_1(11) - S_1(7) = 11 - 7 = 4$ et $S_2(11) - S_2(7) = 4$. D'après la première situation, voir 2.19, les intervalles $[7,11] \cup [17,21]$ sont dans l'ensemble des successeurs de $[7,11]$.

D'après la deuxième situation, voir 2.20, l'intervalle $[22,30]$ est dans l'ensemble des successeurs de $[7,11]$.

Et enfin, d'après la troisième situation, voir 2.22, $S_3(8) = 20 + 9 - 1 - 7 = 21$. Donc, l'intervalle $[20,21]$ est dans l'ensemble des successeurs de $[7,11]$.

L'ensemble de tous les successeurs de l'intervalle $[7,11]$ est alors $[7,11] \cup [17,21] \cup [22,30] = [7,11] \cup [17,30]$

Pour conclure, les intervalles donnés par les trois situations recouvrent tous les successeurs de l'intervalle $[a,b]$.

d) Pseudo code : successeur d'un intervalle de séquences

Soit n l'ordre du groupe G et k son exposant.

Algorithme 4 successeurs d'un intervalle [a,b]

Entrée

a, b : les bornes d'intervalles

Sortie

p : L'ensemble des successeurs de l'intervalle [a,b]

Variables

m : un entier

i : l'indice de parcours de la table des successeurs

$j^{'}$: Un entier représentant la borne supérieur de l'intervalle correspondant au dernier nœud visité au niveau $k-1$ de l'arbre

r_a : un entier représentant le nombre des successeurs consécutifs de l'entier a

$S(a)$: la table des successeurs de l'entier a

$S(b)$: la table des successeurs de l'entier b

Initialisation

$m \leftarrow n + k - 1$

initialiser le nœud p au nœud vide

1: **procédure**
2: **si** (l'intervalle contient un seul élément $(a = b)$) **alors**
3: Retourner les successeurs de l'entier a
4: **sinon**
5: Calculer les successeurs $S(b)$ de l'entier b
6: Inclure dans p l'intervalle $[S_{r_a}, S_n(b)]$
7: **pour** $(i = 0; i < r_a; i++)$ **faire**
8: **si** $(S_i(b) - S_i(a) = b - a)$ **alors**
9: Inclure l'intervalle $[S_i(a), S_i(b)]$ dans p
10: **sinon** Inclure l'intervalle $[S_i(a), S_i(a) + j^{'} - 1 - a]$ dans p

Les opérations ensemblistes sont rassemblées dans une API (Application Programming Interface) afin de rendre le programme modulaire. Cette API comprend les fonctions suivantes :

— *set_min_max* : définit l'entier minimum et l'entier maximum manipulés de telle sorte que $[\text{min}, \text{max}]$ est l'ensemble complet.

— *empty_out*(s) : vide l'ensemble s et libère sa mémoire.

— *is_full*(s) : renvoie 1 si l'ensemble s est complet $[\text{min}, \text{max}]$.

— *print_set*(s) : affiche l'ensemble s comme une réunion d'intervalles triés.

— *include_elt*(x,s) : ajoute un élément x dans un ensemble s.

— *include_intvl*(x,y,s) : ajoute l'intervalle $[x,y]$ dans un ensemble s.

— *include_set*(s,t) : ajoute un ensemble s dans un ensemble t.

— *exclude_elt*(x,s) : exclut un élément x d'un ensemble s.

— *exclude_intvl*(x,y,s) : exclut un intervalle $[x,y]$ d'un ensemble s.

— *exclude_set* (s,t) : exclure un ensemble s d'un ensemble t.

Voir https://github.com/Zerdoum/EGZ/blob/master/set.h pour plus de détails.

De plus le code source est disponible sur https://github.com/Zerdoum/EGZ/tree/master

3 Résultats et perspectives

En appliquant l'algorithme décrit dans ce chapitre, nous avons pu déterminer les valeurs suivantes de la constante d'Erdős-Ginzburg-Ziv qui étaient auparavant inconnues. :

— $\mathrm{s}(C_2 \oplus C_2 \oplus C_2 \oplus C_4) = 13$.

— $\mathrm{s}(C_2 \oplus C_2 \oplus C_2 \oplus C_6) = 17$.

a) Observations et comparaisons

Il est difficile d'évaluer la complexité algorithmique de notre algorithme. En effet, celle-ci est proportionnelle au nombre d'intervalles disjoints nécessaires pour décrire l'ensemble des successeurs. Cependant, nous avons observé l'évolution du nombre d'intervalles en fonction de la taille k des séquences.

— Au départ, l'entier k vaut l'exposant du groupe. L'ensemble des toutes les séquences de longueur k et de somme nulle forme pratiquement des singletons.

— Ensuite, en calculant les successeurs, la taille de la séquence augmente et on remarque qu'il y a des intervalles non réduit à des singletons qui apparaissent.

— Lorsqu'on avance dans l'algorithme, le nombre d'intervalles croit jusqu'à atteindre un maximum.

— Ensuite, le nombre d'intervalles décroit grâce aux réunions qui rassemblent des intervalles disjoints.

— A la fin, on atteindra un seul intervalle regroupant tous les entiers.

Exemple 3.1. Posant $G = C_3 \oplus C_6$.

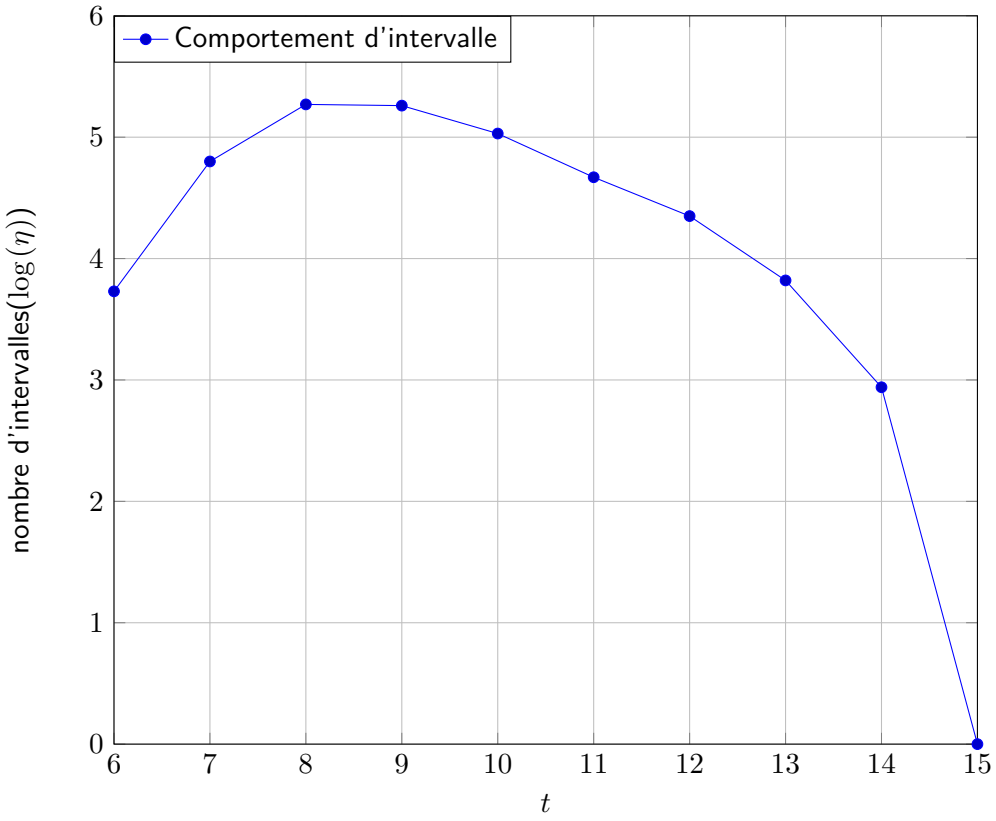
L'exposant du groupe G est $k = 6$.

Le tableau suivant donne l'évolution du nombre d'intervalles lorsque k varie.

Rappelons que $\mathrm{s}(G) = 15$

k	nombres d'intervalles (η)	temps de calculs (.sec)
6	5412	0.066417
7	63129	0.130332
8	187156	0.835978
9	181894	3.078315
10	106910	5.102083
11	46126	6.188090
12	22329	6.623897
13	6678	6.817039
14	865	6.864774
15	1	6.867947

TABLEAU IV.4 – L'évolution du nombre d'intervalles lorsque k varie



Exemple 3.2. Le tableau suivant donne le nombre maximum d'intervalles observés pour le calcul sur différents groupes G.

Soient t le temps de calculs en secondes et max le nombre maximum d'intervalles.

G	$C_2 \oplus C_4$	$C_3 \oplus C_3$	$C_2 \oplus C_8$	$C_3 \oplus C_6$	$C_2 \oplus C_2 \oplus C_6$
max	220	283	379896	187156	1924181
$s(G)$	9	9	17	15	15
t	0.001916	0.003210	9.019284	6.282718	116.735440

Remarque 3.3. Ces résultats ont été obtenu en compilant le programme sur une machine ayant les caractéristiques suivantes :

— Le système utilisé est macOs Catalina avec le noyau Darwin Kernel Version 19.0.0

— Le SHELL utilisé est bash, version 3.2.57

— Le compilateur gcc est utilisé avec la version 11.0.0

— Processeur 1,7 GHz Intel Core i7 double cœur

— Mémoire 8 Go 1600 MHz DDR3

4 Implémentation

Les entiers sont codés sur 64 chiffres binaires, ce choix impose une limite au nombre maximal de séquences représentables $\binom{n+k-1}{k} < 2^{64}$, où n est l'ordre du groupe et k est la taille des séquences. Passer à une taille plus importante permet de reculer cette limite au prix d'une plus grande mémoire nécessaire et d'un ralentissement des calculs.

La structure fondamentale est le nœud d'un arbre binaire qui contient l'information d'un intervalle. Cette structure contient les champs suivants :

— Les bornes a et b de l'intervalle $[a, b]$ (64 symboles binaires chacune).

— Information d'équilibre nécessaire à l'arbre AVL, pour plus de détails sur les arbres AVL, voir section 3, page 110.

— Les fils gauche et droit du nœud, codés comme des entiers de 32 symboles binaire qui représente l'index dans un tableau de nœuds au lieu des 64 symboles binaire d'un pointeur. Cette taille limite le nombre de nœuds à 2^{32} avec pour avantage de nécessiter une taille moindre pour chaque nœud.

Passer à des pointeurs ou des entiers de 64 symboles binaires permet de dépasser cette limite au prix d'un doublement de mémoire occupée par un nœud.

Avec la présente implémentation, un nœud occupe une taille de 25 octets.

Traiter un maximum de 2^{32} nœuds nécessite une mémoire de 100 Go, ce qui est plus grand que les machines personnelles sur lesquelles l'algorithme a été expérimenté.

Remarques 4.1. — Afin de compacter au mieux les données et d'outrepasser les effets d'alignement des compilateurs pour traiter les données structurées, chaque champs d'un nœud fait l'objet d'un tableau séparé, un nœud étant désigné par son index dans chaque tableau.

Afin de faciliter le passage à d'autres modèles, les accès sont définis avec des macros.

— Une machine disposant de 8 Go de mémoire ne peut théoriquement représenter que 320 million de nœuds. Cette limite peut être dépassée grace au mécanisme de « mémoire virtuelle »proposée de façon transparente par le compilateur GCC. Ce mécanisme consiste à effectuer des échanges entre la mémoire vive et le disque sur sa partition « SWAP ».

Nous avons expérimenté l'usage d'un disque SSD (Solid State Disk) qui propose des échanges bien plus rapide que les traditionnels disques à base de support magnétique.

a) Performance de l'algorithme

Cet algorithme nous a permis de trouver la valeur de la constante d'Erdős-Ginzburg-Ziv pour des groupes abéliens finis assez grands.

La limite des calculs dépend de la quantité du coefficient binomial $\binom{n+k}{k+1}$ de la formule itérée de Pascal (voir la formule 2.3 page 65). Dans notre implémentation, ce coefficient ne doit pas dépasser 64 bits. On pourrait envisager des calculs en multiprécision pour aller au-delà de cette limite.

Nous avons essayé d'améliorer l'efficacité de notre algorithme en changeant la numérotation des éléments du groupe. En effet, la numérotation que nous avons prise est triviale, mais en essayant de changer l'ordre de numérotation nous n'avons pas obtenus des résultats significatifs.

A l'avant dernière itération de l'algorithme, nous obtenons toutes les séquences de G de longueur $s(G)-1$ qui n'ont pas de sous-séquences de longueur exposant du groupe et de somme nulle. Autrement dit, l'algorithme peut être immédiatement modifié afin de résoudre le problème inverse associé à la constante $s(G)$.

CHAPITRE V

AUTRES CONSTANTES

Sommaire

1 La constante de Davenport

Le problème de déterminer la valeur exacte de la constante de Davenport a été introduit dans les année 1960. Pour plus de détails sur le développement historique de cette constante, voir le chapitre I, page 1. Dans cette section, nous rappelons la définition de cette constante ainsi que quelques résultats connus.

Soit G un groupe abélien fini noté additivement. La constante de Davenport notée $\mathcal{D}(G)$ est définie comme le plus petit entier t tel que toutes suites S de G de longueur supérieur ou égale à t contienne une sous-suite de somme nulle.

Si $G = \oplus_{i=1}^{r} C_{n_i}$ est la somme directe de groupes cycliques C_{n_i}, où n_i représente l'ordre du groupe cyclique vérifiant $n_i | n_{i+1}$ et r est le rang du groupe G.

Alors on peut définir $\mathcal{D}^*(G)$ comme

$$\mathcal{D}^*(G) = 1 + \sum_{i=1}^{r} (n_i - 1).$$

on a les deux inégalités suivantes

$$\mathcal{D}^*(G) = \sum_{i=1}^{r} (n_i - 1) + 1 \leq \mathcal{D}(G) \leq n_r \left(1 + log \frac{|G|}{n_r}\right).$$

D'où

$$\forall n \in \mathbb{N}^*,\ \mathcal{D}(C_n) = n$$

Jusqu'à la fin des années 1960, il était connu que $\boxed{\mathcal{D}(G) = \mathcal{D}^*(G)}$ dans certains cas. Ci dessous un récapitulatif des cas possibles pour avoir cette égalité.

- Les $p-$groupes et les groupes de rang au plus 2. [Ols69b], [Ols69a], [Boa69]
- Pour le groupe $G^{'} \oplus C_n$ avec $G^{'}$ est un $p-$groupe et $\mathcal{D}(G^{'}) \leq 2\exp(G^{'}) - 1$ et $pgcd(n, \exp(G^{'})) = 1$. voir [Boa69] et [Cha+02] pour la généralisation.
- Les groupes de rang 3 (certains cas). (dans la suite $n \in \mathbb{N}$)
  - Pour le groupe $C_2^2 \oplus C_{2n}$. voir [Boa69].
  - Pour le groupe $C_3^2 \oplus C_{3n}$. voir [BSP07]
  - Pour le groupe $C_3 \oplus C_{3n}^2$ tel que n et 6 sont premiers entre eux. voir [BHSP09]
  - Pour le groupe $C_{n_1} \oplus C_{n_2} \oplus C_{n_3 m}$ avec $n_1 | n_2 | n_3$ et $m \in \mathbb{N}$ (sous les conditions suivantes) : voir [Cha+02]
    - Si $\mathcal{D}(C_{n_1} \oplus C_{n_2} \oplus C_{n_3}) = \mathcal{D}^*(C_{n_1} \oplus C_{n_2} \oplus C_{n_3})$
    - Et si $(n_1 n_2^2 - 2n_2 - n_1 - 2) \leq n_3$

- Pour le groupe $C_4^2 \oplus C_{4n}$. [Sch11]
- Pour le groupe $C_6^2 \oplus C_{6n}$. [Sch11]
- Les groupes de la forme C_n^3 (sous conditions) voir [Boa69] et [PVEB69]
  - Si $n = 2p^k$ avec p un nombre premier.
  - Si $n = 32^k$.

- Le groupe de rang 4 de la forme $C_2^3 \oplus C_{2n}$. (voir [Baa69])
- Pour les groupe de rang 5 sous la forme $C_2^4 \oplus C_{2k}$ avec k un nombre pair.

Il est aussi connu que $\boxed{\mathcal{D}(G) = \mathcal{D}^*(G) + 1}$ dans les cas suivants :

- Pour les groupes sous la forme : $C_2^{r-1} \oplus C_6$ où $r \in \{5, 6, 7\}$.
  (dû à V. Ponomarenko, voir [Sch11])
- Pour le groupe $C_3^3 \oplus C_6$. (dû à V. Ponomarenko, voir [Sch11])

Pour le cas où $\boxed{\mathcal{D}(G) = \mathcal{D}^*(G) + 2}$ on connait la valeur de la constante de Davenport pour le groupe :

- $C_2^7 \oplus C_6$. (dû à V. Ponomarenko, voir [Sch11])

On sait aussi que la valeur de la constante de Davenport pour les groupes de la forme $C_2^4 \oplus C_{2k}$ avec $k \geq 70$ est (voir [CS14]) :

$$\mathcal{D}(C_2^4 \oplus C_{2k}) = \begin{cases} 2k + 4 = \mathcal{D}^*(C_2^4 \oplus C_{2k}) & \text{si k est pair} \\ 2k + 5 = \mathcal{D}^*(C_2^4 \oplus C_{2k}) + 1 & \text{si k est impair} \end{cases}$$

2 Généralisation de la constante Erdős-Ginzburg-Ziv

a) Introduction

Définition 2.1 (La constante $\mathrm{s}_{\leq k}(G)$)**.** Rappelons que la constante $\mathrm{s}_{\leq k}(G)$ est définie comme étant le plus petit entier positif ℓ tel que toute séquence de longueur ℓ contient une sous-séquence de longueur au plus égale à k et de somme nulle.

La constante $\mathrm{s}_{\leq k}(G)$ a été introduite en 2001 par Delorme, Ordaz et Quiroz [DOQ01]. En général, le problème de déterminer la valeur de la constante $s_{\leq k}(G)$ est difficile et reste ouvert pour beaucoup de groupes.

Rappelons quelques résultats connus.

b) Résultats connus

Voici la liste des groupes abéliens finis pour lesquels la valeur de la constante $\mathrm{s}_{\leq k}(G)$ est connue.

— Dans le cas particulier où l'entier k est égal à l'exposant du groupe. On utilise cette notation $\mathrm{s}_{\leq \exp(G)}(G) = \eta(G)$. En d'autres termes $\eta(G)$ est le plus entier positif ℓ tel que chaque séquence S de G de longueur $|S| \geq \ell$ contient une sous-séquence T non vide de longueur $|T| \leq \exp(G)$ et de somme nulle.

Remarque 2.2. La constante $\eta(G)$ est notamment connue dans tous les cas où la constante d'Erdős-Ginzburg-Ziv est connue. Nous renvoyons vers le chapitre IV page 54 pour plus de détails sur la constante d'Erdős-Ginzburg-Ziv. Nous avons la conjecture suivante (voir conjecture 6.5 de l'article [GG06]).

Conjecture 2.3. Soit G un groupe abélien fini,

$$\eta(G) = \mathrm{s}(G) - \exp(G) + 1$$

— Pour les valeurs de k strictement inférieures à l'exposant du groupe, la constante $\mathrm{s}_{\leq k}(G)$ est infinie puisque pour g un élément d'ordre $\exp(G)$ la séquence g^n ne contient que des sous-séquences de somme nulles dont le longueur est une multiple de $\exp(G)$.

$$\text{Si } 1 \leq k < \exp(G) \text{ alors } \mathrm{s}_{\leq k}(G) = \infty$$

— Par contre pour les valeurs de k suffisamment grandes, en particulier $k \geq \mathcal{D}(G)$, la constante $\mathrm{s}_{\leq k}(G)$ est égale à $\mathcal{D}(G)$, puisque toute séquence de

longueur $\mathcal{D}(G)$ contient une sous-séquence de somme nulle dont la longueur est forcement au plus $\mathcal{D}(G)$ et il y a séquence de longueur $\mathcal{D}(G)-1$ qui ne contient aucune sous-séquence de somme nulle.

$$\text{Si } k \geq \mathcal{D}(G) \text{ alors } \mathrm{s}_{\leq k}(G) = \mathcal{D}(G)$$

— Pour les groupes de rangs 2, le problème de déterminer la constante $\mathrm{s}_{\leq k}(G)$ est résolu grâce aux résultats suivants (voir théorèmes 2 et 3 de l'article [WZ17]) :

1. Pour $G = C_m \oplus C_n$, où m et n sont des entiers tels que $1 \leq m \mid n$.

$$\forall k \in [0, m-1],\ s_{\leq \mathcal{D}(G)-k}(G) = \mathcal{D}(G) + k$$

2. Pour $G = C_2^r$ avec $r \in \mathbb{N}$.

$$\forall r-k \in \left[\left\lceil \frac{2r+2}{3}, r \right\rceil\right],\ s_{\leq r-k}(G) = r+2$$

— Pour les groupes abéliens finis de rang supérieur à 2, $r(G) \geq 2$ le problème de déterminer la constante $\mathrm{s}_{\leq k}(G)$ est connu pour le cas suivant :
(voir le lemme 8 de l'article [WZ17])

1.

$$s_{\leq \mathcal{D}(G)-1}(G) = \mathcal{D}(G) + 1$$

2. Pour le groupe C_3^3 on connait la valeur de la constante $\mathrm{s}_{\leq k}(G)$ pour tout k supérieur ou égale à 3 (voir [BSP07]).
$\mathrm{s}_{\leq 3}(C_3^3) = 17,\ \mathrm{s}_{\leq 4}(C_3^3) = 10,\ \mathrm{s}_{\leq 5}(C_3^3) = 9,\ \mathrm{s}_{\leq 6}(C_3^3) = 8.$

— Pour le groupe $G = C_2 \oplus C_2 \oplus C_2 \oplus C_4$, on sait que $\mathrm{s}(G) = 13$, voir le chapitre 3, page 88. D'où $\mathrm{s}_{\leq 4}(G) = 10$.

— Pour les groupes de la forme $G = C_2 \oplus C_{n_2} \oplus C_{n_3}$ avec $2 \mid n_2 \mid n_3$ la valeur de la constante $\eta(G)$ est connue (voir théorème 3.1 de l'article [GS19]). D'où $\mathrm{s}_{\leq 4}(C_2 \oplus C_4 \oplus C_4) = 14$.

— Pour le groupe $G = C_4 \oplus C_4 \oplus C_4$, la constante $\mathrm{s}(G) = 25$ (voir [Har73]). D'où $\mathrm{s}_{\leq 4}(G) = 22$.

Ci-dessous un tableau regroupant les groupes pour lesquels la valeur des constantes $\mathcal{D}(G)$ et $s_{\leq k}(G)$ sont connues. Nous ne donnons pas systématiquement les groupes de rang inférieur ou égal à 2, car la valeur des constantes est connue pour ces groupes. Les nouveaux résultats sont marqués en rouge et ceux pour lesquels le problème reste ouvert sont marqués par un point d'interrogation.

c) Les différentes valeurs des constantes $\mathcal{D}(G)$ et $\mathrm{s}_{\leq k}(G)$

$\|G\|$	$D(G)$	k	$s_{\leq k}(G)$	Le groupe
2	2	2	2	C_2
3	3	3	3	C_3
4	4	4	4	C_4
4	3	$D(G)-1=2$	$D(G)+1=4$	$C_2 \oplus C_2$
4	3	$D(G)+t\ \forall t \geq 0$	$D(G)=3$	$C_2 \oplus C_2$
5	5	5	5	C_5
6	6	6	6	$C_2 \oplus C_3 \simeq C_6$
8	5	$D(G)-1=4$	$D(G)+1=6$	$C_2 \oplus C_4$
8	5	$D(G)+t\ \forall t \geq 0$	$D(G)=5$	$C_2 \oplus C_4$
8	4	2	8	$C_2 \oplus C_2 \oplus C_2$
8	4	3	5	$C_2 \oplus C_2 \oplus C_2$
8	4	$\forall k \geq 4$	4	$C_2 \oplus C_2 \oplus C_2$
9	5	3	7	$C_3 \oplus C_3$
9	5	4	6	$C_3 \oplus C_3$
9	5	$\forall k \geq 5$	5	$C_3 \oplus C_3$
12	7	6	8	$C_2 \oplus C_6$
12	7	$\forall k \geq 7$	7	$C_2 \oplus C_6$
16	16	2	16	$C_2 \oplus C_2 \oplus C_2 \oplus C_2$
16	16	3	9	$C_2 \oplus C_2 \oplus C_2 \oplus C_2$
16	16	4	6	$C_2 \oplus C_2 \oplus C_2 \oplus C_2$
16	16	$\forall k \geq 5$	5	$C_2 \oplus C_2 \oplus C_2 \oplus C_2$
16	9	8	10	$C_2 \oplus C_8$
16	9	$\forall k \geq 9$	9	$C_2 \oplus C_8$
16	7	$\forall k \geq 7$	7	$C_4 \oplus C_4$
16	7	6	8	$C_4 \oplus C_4$
16	7	5	9	$C_4 \oplus C_4$
16	7	4	10	$C_4 \oplus C_4$
16	6	4	8	$C_2 \oplus C_2 \oplus C_4$
16	6	5	7	$C_2 \oplus C_2 \oplus C_4$
16	6	$\forall k \geq 6$	6	$C_2 \oplus C_2 \oplus C_4$
18	8	6	10	$C_3 \oplus C_6$
18	8	7	9	$C_3 \oplus C_6$
18	8	$\forall k \geq 8$	8	$C_3 \oplus C_6$
20	11	10	12	$C_2 \oplus C_{10}$
20	11	$\forall k \geq 11$	11	$C_2 \oplus C_{10}$
24	8	6	10	$C_2 \oplus C_2 \oplus C_6$
24	8	7	9	$C_2 \oplus C_2 \oplus C_6$
24	8	$\forall k \geq 8$	8	$C_2 \oplus C_2 \oplus C_6$
25	9	5	13	$C_5 \oplus C_5$
25	9	6	12	$C_5 \oplus C_5$
25	9	7	11	$C_5 \oplus C_5$
25	9	8	10	$C_5 \oplus C_5$
25	9	$\forall k \geq 9$	9	$C_5 \oplus C_5$

$\|G\|$	$D(G)$	k	$s_{\leq k}(G)$	Le groupe
27	7	3	17	$C_3 \oplus C_3 \oplus C_3$
27	7	4	10	$C_3 \oplus C_3 \oplus C_3$
27	7	5	9	$C_3 \oplus C_3 \oplus C_3$
27	7	6	8	$C_3 \oplus C_3 \oplus C_3$
27	7	$\forall k \geq 7$	7	$C_3 \oplus C_3 \oplus C_3$
27	11	$\forall k \geq 11$	11	$C_3 \oplus C_9$
27	11	10	12	$C_3 \oplus C_9$
27	11	9	13	$C_3 \oplus C_9$
32	7	4	10	$C_2 \oplus C_2 \oplus C_2 \oplus C_4$
32	7	5	9	$C_2 \oplus C_2 \oplus C_2 \oplus C_4$
32	7	6	8	$C_2 \oplus C_2 \oplus C_2 \oplus C_4$
32	7	$\forall k \geq 7$	7	$C_2 \oplus C_2 \oplus C_2 \oplus C_4$
32	10	8	12	$C_2 \oplus C_2 \oplus C_8$
32	10	9	11	$C_2 \oplus C_2 \oplus C_8$
32	10	10	10	$C_2 \oplus C_2 \oplus C_8$
32	8	4	14	$C_2 \oplus C_4 \oplus C_4$
32	8	5	11	$C_2 \oplus C_4 \oplus C_4$
32	8	6	10	$C_2 \oplus C_4 \oplus C_4$
32	8	7	9	$C_2 \oplus C_4 \oplus C_4$
32	8	$\forall k \geq 8$	8	$C_2 \oplus C_4 \oplus C_4$
32	6	2	32	$C_2 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_2$
32	6	3	17	$C_2 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_2$
32	6	4	7	$C_2 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_2$
32	6	5	7	$C_2 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_2$
32	6	$\forall k \geq 6$	6	$C_2 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_2$
48	9	6	12	$C_2 \oplus C_2 \oplus C_2 \oplus C_6$
48	9	7	11	$C_2 \oplus C_2 \oplus C_2 \oplus C_6$
48	9	8	10	$C_2 \oplus C_2 \oplus C_2 \oplus C_6$
48	9	$\forall k \geq 9$	9	$C_2 \oplus C_2 \oplus C_2 \oplus C_6$
40	12	10	14	$C_2 \oplus C_2 \oplus C_{10}$
40	12	11	13	$C_2 \oplus C_2 \oplus C_{10}$
40	12	$\forall k \geq 12$	12	$C_2 \oplus C_2 \oplus C_{10}$
48	14	12	16	$C_2 \oplus C_2 \oplus C_{12}$
48	14	13	15	$C_2 \oplus C_2 \oplus C_{12}$
48	14	$\forall k \geq 14$	14	$C_2 \oplus C_2 \oplus C_{12}$
54	10	6	14	$C_3 \oplus C_3 \oplus C_6$
54	10	7	13	$C_3 \oplus C_3 \oplus C_6$
54	10	8	12	$C_3 \oplus C_3 \oplus C_6$
54	10	9	11	$C_3 \oplus C_3 \oplus C_6$
54	10	$\forall k \geq 10$	10	$C_3 \oplus C_3 \oplus C_6$
64	8	4	?	$C_2 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_4$
64	8	5	?	$C_2 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_4$
64	8	6	?	$C_2 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_4$
64	8	7	9	$C_2 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_4$
64	8	$\forall k \geq 8$	8	$C_2 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_4$
64	11	8	?	$C_2 \oplus C_2 \oplus C_2 \oplus C_8$
64	11	9	?	$C_2 \oplus C_2 \oplus C_2 \oplus C_8$
64	11	10	12	$C_2 \oplus C_2 \oplus C_2 \oplus C_8$
64	11	$\forall k \geq 11$	11	$C_2 \oplus C_2 \oplus C_2 \oplus C_8$

$\|G\|$	$D(G)$	k	$s_{\leq k}(G)$	Le groupe
64	9	4	?	$C_2 \oplus C_2 \oplus C_4 \oplus C_4$
64	9	5	?	$C_2 \oplus C_2 \oplus C_4 \oplus C_4$
64	9	6	?	$C_2 \oplus C_2 \oplus C_4 \oplus C_4$
64	9	7	?	$C_2 \oplus C_2 \oplus C_4 \oplus C_4$
64	9	8	10	$C_2 \oplus C_2 \oplus C_4 \oplus C_4$
64	9	$\forall k \geq 9$	9	$C_2 \oplus C_2 \oplus C_4 \oplus C_4$
64	10	4	22	$C_4 \oplus C_4 \oplus C_4$
64	10	5	?	$C_4 \oplus C_4 \oplus C_4$
64	10	6	?	$C_4 \oplus C_4 \oplus C_4$
64	10	7	?	$C_4 \oplus C_4 \oplus C_4$
64	10	8	?	$C_4 \oplus C_4 \oplus C_4$
64	10	9	11	$C_4 \oplus C_4 \oplus C_4$
64	10	$\forall k \geq 10$	10	$C_4 \oplus C_4 \oplus C_4$
64	12	8	?	$C_2 \oplus C_4 \oplus C_8$
64	12	9	?	$C_2 \oplus C_4 \oplus C_8$
64	12	10	?	$C_2 \oplus C_4 \oplus C_8$
64	12	11	13	$C_2 \oplus C_4 \oplus C_8$
64	12	$\forall k \geq 12$	12	$C_2 \oplus C_4 \oplus C_8$
81	13	9	?	$C_3 \oplus C_3 \oplus C_9$
81	13	10	?	$C_3 \oplus C_3 \oplus C_9$
81	13	11	?	$C_3 \oplus C_3 \oplus C_9$
81	13	12	14	$C_3 \oplus C_3 \oplus C_9$
81	13	$\forall k \geq 13$	13	$C_3 \oplus C_3 \oplus C_9$
108	16	12	?	$C_3 \oplus C_3 \oplus C_{12}$
108	16	13	?	$C_3 \oplus C_3 \oplus C_{12}$
108	16	14	?	$C_3 \oplus C_3 \oplus C_{12}$
108	16	15	17	$C_3 \oplus C_3 \oplus C_{12}$
108	16	$\forall k \geq 16$	16	$C_3 \oplus C_3 \oplus C_{12}$

TABLEAU V.1 – Tableau récapitulatif des valeurs des constantes $D(G)$ et $\mathrm{s}_{\leq k}(G)$

d) Description de l'algorithme

Dans cette sous-section nous décrivons une adaptation de l'algorithme présenté au chapitre précédent pour calculer ces constantes.

Soit ℓ un entier supérieur ou égal à l'exposant du groupe.

Rappelons que nous cherchons à déterminer le plus petit entier t tel que toute séquence de longueur au moins t contienne une sous-séquence de somme nulle et de longueur au plus égale à ℓ.

Notons que l'entier ℓ doit être supérieur ou égal à l'exposant du groupe. Sinon comme nous l'avons vu, la valeur de la constante sera infinie.

En effet, si l'entier ℓ est inférieur que l'exposant du groupe, il existera toujours des séquences pour lesquelles la somme n'est pas nulle.

Le principe général de l'algorithme est le suivant :

Soit $E_k(i)$: L'ensemble des séquences de longueur k et de somme i.

Et soit $S(E_k(i))$ l'ensemble des successeurs de l'ensemble $E_k(i)$

Dans un premier temps, nous considérons l'ensemble des séquences de somme nulle et de longueur au plus ℓ.

Pour cela, nous suivons les étapes suivantes :

- On commence par l'élément neutre $\{0\}$ représentant $E_1(0)$, l'ensemble des séquences de longueur 1 et de somme nulle.
- Puis, nous calculons $S(E_1(0))$ les successeurs de 0 avec la méthode décrite dans la section 2-), page 65 et on rajoute $E_2(0)$, l'ensemble des séquences de longueur 2 et de somme nulle. Appelons la réunion d'ensembles $S(E_1(0)) \cup E_2(0)$ l'ensemble R_2.
- L'ensemble R_2 est alors l'ensemble de toutes les séquences de longueur 2 qui ont une sous-séquence de somme nulle.
- Nous calculons ensuite les successeurs de l'ensemble R_2 auquel on ajoute l'ensemble $E_3(0)$, l'ensemble des séquences de longueur 3 et de somme nulle. Appelons la réunion de l'ensemble $S(R_2) \cup E_3(0)$ l'ensemble R_3.
- L'ensemble R_3 est alors l'ensemble de toutes les séquences qui ont une sous-séquence de longueur au plus 3 et de somme nulle.
- Nous répétons successivement l'opération précédente, jusqu'à ce que nous obtenions R_ℓ, l'ensemble de toutes les séquences de longueur ℓ qui ont une sous-séquence de somme nulle. Notons que forcément ces sous-séquences sont de longueur plus petite ou égale à ℓ.

Nous calculons, ensuite les successeurs des séquences jusqu'à ce que nous obtenions comme successeurs l'ensemble des toutes les séquences possibles d'une longueur donnée.

Si nous prenons n'importe quelle séquence de cette longueur-là, alors par la relation d'inclusion, cette séquence est forcément le successeur d'une séquence de somme nulle et de longueur au plus égale à ℓ.

Pour calculer l'ensemble des suites croissantes de longueurs k et de somme 0, nous avons appliqué une formule de récurrence permettant d'éviter la recherche exhaustive.

Nous décrivons dans la sous-section suivante cette formule et le détail des calculs.

e) Le calcul des séquences de somme i et de longueur k

Nous avons développé un algorithme permettant de calculer pour tout $k \leq \ell$ les séquences de somme nulle et de longueur k à chaque étape.

L'idée de cet algorithme est de créer une table qui contient les séquences de somme nulle et de longueur k, pour tout $k \leq \ell$.

Puis nous appliquons la formule récursive qui dit qu'une séquence de longueur $k+1$ et de somme i s'obtient comme une séquence de longueur k et de somme u dont on lui rajoute le terme $i-u$.

$$E_{k+1}(i) = \cup_{u \in G} E_k(u) \cdot (i-u) \tag{2.1}$$

Exemple 2.4. Soit $G = C_3$.

En appliquant la formule de récurrence 2.1, nous obtenons :

$E_3(0) = (E_2(0) \cdot 0) \cup (E_2(1) \cdot 2) \cup (E_2(2) \cdot 1)$

$= [(E_1(0) \cdot 0 \cup E_1(1) \cdot 2 \cup E_1(2) \cdot 1)] \cdot 0 \cup [(E_1(0) \cdot 1 \cup E_1(1) \cdot 0 \cup E_1(2) \cdot 2)] \cdot 2$

$\cup [(E_1(0) \cdot 2 \cup E_1(1) \cdot \cup E_1(2) \cdot 0)] \cdot 1$

Si on veut appliquer la formule de récurrence 2.1, on s'aperçoit qu'on calcule plusieurs fois les mêmes données, ce qui engendre une perte de temps de calcul, dans l'exemple 2.4, les données $E_1(0)$, $E_1(1)$ et $E_1(2)$, se répètent plusieurs fois. Pour éviter de les recalculer, nous mémorisons chacune de ces valeurs déjà calculées.

Ce qui donne l'algorithme 5

Soient $\rho_{i,u,k}$ l'ensemble des numéros de suites croissantes de somme $i-u$ et de longueur k et $\nu_{i,u,k}$ l'ensemble des numéros de suites croissantes de somme i

i	u	k	$E_k(i-u)$	$\rho_{i,u,k}$	$E_{k+1}(i)$	$\nu_{i,u,k}$
0	0	1	0	$[0]$	$\{00\}$	$[0]$
0	1	1	2	$[2]$	$\{12\}$	$[4]$
0	2	1	1	$[1]$	$\emptyset$	$[]$
1	0	1	1	$[1]$	$\{01\}$	$[1]$
1	1	1	0	$[0]$	$\emptyset$	$[]$
1	2	1	2	$[2]$	$\{22\}$	$[5]$
2	0	1	2	$[2]$	$\{02\}$	$[2]$
2	1	1	1	$[1]$	$\{11\}$	$[3]$
2	2	1	0	$[0]$	$\emptyset$	$[]$
0	0	2	$\{00,12\}$	$[0] \cup [4]$	$\{000,012\}$	$[0] \cup [4]$
0	1	2	$\{02,11\}$	$[2,3]$	$\{012,111\}$	$[6]$
0	2	2	$\{01,22\}$	$[1] \cup [5]$	$\{012,222\}$	$[9]$

TABLEAU V.2 – Détails des calculs de la formule de récurrence pour le groupe C_3

et de longueur $k+1$. Notons que l'ensemble $E_{k+1}(i)$ est obtenu en rajoutant à l'ensemble $E_k(i-u)$ le terme u.

Le tableau ci-dessous donne le détail du calcul de la formule de récurrence 2.1 appliquée pour le groupe $G = C_3$.

Nous avons observé que l'ensemble translaté $\nu_{i,u,k}$ se calcule à partir de l'ensemble $\rho_{i,u,k}$, ce qui permet d'accélérer les calculs.

En traduisant le numéro de la suite croissante de longueur $k+1$ et de somme i en un bitmap de longueur $n+k$ et de poids $k+1$ et en parcourant le triangle de Pascal modifié comme expliqué dans 1-), page 23, il est clair que ce parcours n'est qu'un décalage du parcours du bitmap de la suite croissante de longueur $k+1$ et de somme $i-u$ comme l'illustre la table V.1, en prenant comme exemple un élément $x = 5$ de l'ensemble $\rho_{0,2,2}$ de la dernière linge du tableau V.2

Pour cet exemple : le numéro 5 correspond au bitamp 11001, son bitmap translaté est 11100 qui correspond au numéro $y = 9$ qui est un élément de l'ensemble $\nu_{0,2,2}$.

Ces bitmaps sont schématisés dans la table V.1 ci-dessous.

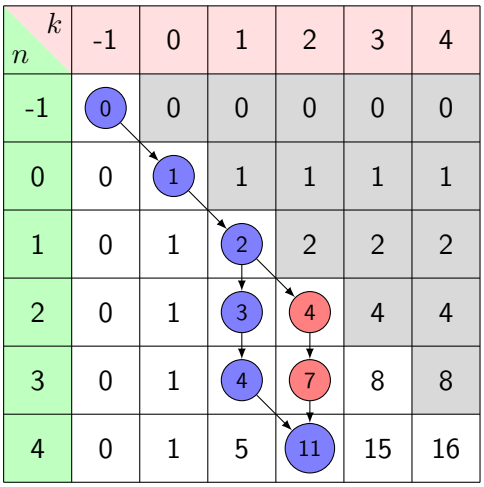


FIGURE V.1 – Le décalage du parcours dans le triangle de Pascal modifié

Dans cet exemple, l'élément y est obtenu en rajoutant à l'élément x la quantité suivante :

$(a_{2,2} - a_{2,1}) + (a_{3,2} - a_{3,1}) = \binom{2}{2} + \binom{3}{2}$

De manière générale, cette translation du parcours dans le triangle de Pascal d'un élément x de l'ensemble $\rho_{i,u,k}$ correspond à rajouter à l'élément x une certaine quantité de coefficients binomiaux donnés par la formule suivante :

$$y = x + \sum_{j=1}^{u} \binom{n+k-j-1}{k} = \sum_{s=k-2}^{k-u-1} \binom{n+s}{k} \tag{2.2}$$

Pour l'exemple 2.4, en utilisant la propriété 2.4, page 22 du triangle de Pascal modifié on a :

$(a_{2,2} + a_{3,2}) - (a_{2,1} + a_{3,1}) = (a_{4,3} - a_{2,3}) - (a_{4,2} - a_{2,2}) = \binom{4}{3} - \binom{2}{3}$

La formule 2.2 peut alors être simplifié par la formule suivante :

$$\sum_{s=k-2}^{k-u-1} \binom{n+s}{k} = \binom{n+k-1}{k+1} - \binom{n+k-1-u}{k+1} \tag{2.3}$$

f) Pseudo code : le calcul de la constante $\mathrm{s}_{\leq k}(G)$

Le code source est disponible sur https://github.com/Zerdoum/EGZ-infval/tree/master

Algorithme 5 Le calcul de la constante $s_{\leq k}(G)$

Entrée

n : l'ordre du groupe G

k : un entier tel que k est supérieur ou égal à l'exposant du groupe

Sortie

$e-1$

Variables

p, q, r, S : des ensembles

M, N : des tableaux

i, u, v, e : des entiers

Initialisation

$N[0] \leftarrow 0$

procédure

Calculer R_ℓ :

initialiser le tableau M

pour $(i = 0; i < n; i++)$ **faire**

  $M[u] = u$

initialiser l'ensemble p

Au départ p contient les suites croissantes de longueur 1 et de somme $u \in G$:

pour $(u = 0; u < n; u++)$ **faire**

  inclure l'élément i dans $p[i]$

pour $(e = 0; e \leq k; e++)$ **faire**

  pour $(u = 1; u \leq n; u++)$ **faire**

    $N[u] = N[u-1] + \binom{n+k-u-1}{k-1}$

calculer $E_e(u)$ en appliquant la formule de récurrence :

$$E_e(u) = \cup_{u \in G} E_{e-1}(u) + (i - u)$$

l'ensemble p contient $E_{e-1}(u)$ et l'ensemble q contiendra $E_e(u)$

  pour $(i = 0; i < n; i++)$ **faire**

    pour $(u = 0; u < n; u++)$ **faire**

      $v = i - u$

      $r = p[v]$

      si $u > 0$ **alors**

        Exclure l'intervalle $[0, M[u] - 1]$ de l'ensemble r

        ainsi nous ne prenons en comptes que les nouvelles suites croissantes

      On translate r par $N[u] - M[u]$ pour obtenir le numéro de la suite de

       longueur e et de somme u

      Inclure l'ensemble r dans l'ensemble $q[i]$

    si $e = k$ **alors**

      break

   inclure dans l'ensemble $q[0]$ l'ensemble S_1

   calculer les successeurs de S_1

   l'ensemble p deviendra l'ensemble q

   l'ensemble q deviendra l'ensemble p

   N deviendra M

   M deviendra N

tant que tant que l'ensemble S_1 n'est pas rempli **faire**

  nous calculons seulement successeurs S_1

  $e \leftarrow e + 1$

g) Résultats et perspectives

Nous avons pu déterminer les valeurs suivantes de la constante $s_{\leq k}(G)$:

- $s_{\leq 5}(C_2 \oplus C_2 \oplus C_2 \oplus C_4) = 9$.
- $s_{\leq 5}(C_2 \oplus C_4 \oplus C_4) = 11$.
- $s_{\leq 6}(C_2 \oplus C_4 \oplus C_4) = 10$.
- $s_{\leq 6}(C_2 \oplus C_2 \oplus C_2 \oplus C_6) = 12$.
- $s_{\leq 7}(C_2 \oplus C_2 \oplus C_2 \oplus C_6) = 11$.
- $s_{\leq 6}(C_3 \oplus C_3 \oplus C_6) = 14$.
- $s_{\leq 7}(C_3 \oplus C_3 \oplus C_6) = 13$.
- $s_{\leq 8}(C_3 \oplus C_3 \oplus C_6) = 12$.

Nous avons vu dans le chapitre précédent qu'à l'étape d'initialisation de la détermination de la constante Erdős-GinzBurg-Ziv, nous parcourons de manière exhaustive, l'ensemble de toutes les suites croissantes de longueur exposant du groupe et nous ne retenons que celles qui ont une somme nulle, c'est-à-dire nous ne retenons que l'ensemble $E_e(0)$.

Nous avons alors essayé d'appliquer la formule de récurrence 2.1 en disant :

$$E_e(0) = \cup_{u \in G} E_{e-1}(u) \cdot (-u) \tag{2.4}$$

pour tenter de gagner en temps de calcul lors de l'initialisation, or on s'est rendu compte que le temps de calcul était plus long.

En effet, nous avons deux types de complexités, une complexité absolue qui se mesure avec le temps de calcul et dépend de toutes les opérations utilisées, et une complexité relative qui dépend du nombre d'opérations ensemblistes.

La complexité de la méthode de recherche exhaustive est exponentielle, c'est la recherche exhaustive de toutes les suites croissantes de longueur e.

Notons que le nombre d'opérations ensembliste est le nombre d'éléments de G, car d'après la formule de récurrence 2.4 nous faisons une réunion sur tous les éléments de G.

On s'attendait à constater une accélération des calculs en appliquant la formule 2.4 mais ce n'est pas le cas ; car la complexité de chaque réunion dépend du nombre d'intervalles, or le nombre d'intervalles croit.

Pour conclure, en pratique la recherche exhaustive est plus rapide que la méthode de la formule de récurrence.

CHAPITRE VI

CONCLUSION ET PERSPECTIVES

Sommaire

Les problème des suites de somme nulle sur les groupes abéliens fini sont des problèmes de la théorie additive des nombres. Ces problèmes consistent à étudier le nombre minimal d'éléments avec ou sans répétitions d'éléments du groupe qu'il faut considérer pour être certain d'en extraire une suite de somme nulle, potentiellement avec des restrictions sur la longueur de cette suite. Ceci définit des constantes associées à chaque groupe abélien fini.

Dans cette conclusion, nous allons présenter un bilan des constantes étudiées en donnant la liste des nouveaux résultats obtenus. Nous renvoyons aux chapitres précédents pour la définition précise de ces constantes. Nous focalisons ensuite sur la partie algorithmique de notre travail, en particulier, nous présentons les différentes façons de représenter les ensembles dans un programme, notre choix de représentation ainsi que quelques perspectives envisageables. Et enfin, nous décrivons notre application web.

1 Le bilan des constantes étudiées

Le tableau suivant donne un récapitulatif des résultats trouvés durant cette thèse.

La constante	La valeur trouvé
La constante de Harborth	$\mathrm{g}(C_3 \oplus C_{3p}) = 3p+3$ où p est un nombre premier différent de 3. $\mathrm{g}(C_3 \oplus C_9) = 13$
La constante Erdős-Ginzburg-Ziv	$\mathrm{s}(C_2^3 \oplus C_4) = 13$ $\mathrm{s}(C_2^3 \oplus C_6) = 17$
La constante $\mathrm{s}_{\leq k}(G)$	$\mathrm{s}_{\leq 5}(C_2^3 \oplus C_4) = 9$ $\mathrm{s}_{\leq 5}(C_2 \oplus C_4^2) = 11$ $\mathrm{s}_{\leq 6}(C_2 \oplus C_4^2) = 10$ $\mathrm{s}_{\leq 6}(C_2^3 \oplus C_6) = 12$ $\mathrm{s}_{\leq 7}(C_2^3 \oplus C_6) = 11$ $\mathrm{s}_{\leq 6}(C_2^3 \oplus C_6) = 14$ $\mathrm{s}_{\leq 7}(C_2^3 \oplus C_6) = 13$ $\mathrm{s}_{\leq 8}(C_2^3 \oplus C_6) = 12$

TABLEAU VI.1 – Nos résultats trouvés

Deux approches ont étés menées au cours de cette étude. Une approche calculatoire et une approche théorique.

L'approche calculatoire consiste à développer des algorithmes pour trouver les constantes. Ils ont permis d'en trouver les valeurs pour certains groupes pour lesquels elles n'étaient pas connues. Et d'autre part, ils nous ont donné des idées pour

déterminer certaines bornes inférieures de constantes étudiées. En effet, examiner les sous-ensembles les plus grands qui n'ont pas de somme nulle, nous a conduits à des idées sur un minorant. C'est ce qui a été fait au troisième chapitre pour la constante de Harborth.

Dans le paragraphe qui suit nous présentons certains détails de notre travail pour la partie algorithmique ainsi que des problèmes d'implémentation informatique.

Un des problèmes à résoudre est la représentation des ensembles. Nous avons considéré plusieurs types de représentations. Les ensembles interviennent à plusieurs niveaux dans notre travail. Pour le premier niveau, nous avons l'ensemble sous-jacent des groupes dont les éléments sont numérotés arbitrairement.

Par exemple $C_2 \oplus C_2 = \{e_0, e_1, e_2, e_3\}$.

Pour le deuxième niveau, nous devons aussi considérer des suites d'éléments de ces groupes avec ou sans répétition. Une suite sans répétition d'éléments n'est rien d'autre qu'un sous-ensemble du groupe. Lorsque les répétitions sont admises, on se ramène comme décrit à la section 2, page 19 à un sous-ensemble d'un ensemble plus grand.

Notons que dans les deux cas, on se ramène toujours à des sous-ensembles.

Pour le deuxième niveau, il nous a fallu travailler avec une numérotation des séquences compatible avec la taille de la séquence. Plus précisément, le numéro d'une séquence de k termes sera toujours inférieur à celui d'une séquence de $k+1$ termes. Voir la section 2, page 19.

Examinons maintenant les façons de représenter un ensemble de numéros de séquence dans un programme.

2 La représentation informatique des ensembles

Dans le paragraphe suivant, nous présentons quelques façons classiques de représenter un ensemble dans un programme avec ses avantages et ses inconvénients. Nous justifierons le choix de la représentation que nous avons adoptée.

1. **Le bitmap** : un sous-ensemble F d'un ensemble E de n éléments est représenté par un mot binaire de n termes. Chaque position désigne un élément de E. La valeur 1 signifie que cet élément appartient à F et la valeur 0 signifie que qu'il n'y appartient pas. L'avantage de cette représentation est la simplicité. En contrepartie, la complexité des opérations est constante et nécessite de parcourir tout le bitmap, ce qui peut s'avérer très complexe.

   En effet, explorer les suites croissantes de k termes d'un ensemble de n éléments, nécessite des bitmaps de taille $\binom{n+k-1}{k}$. Il n'est pas raisonnable

d'utiliser cette représentation quand cette taille dépasse, par exemple 2^{40}. Notons que 2^{40} est égale à mille milliards, avec un milliards d'opération par seconde, le temps est milles secondes soit environs vingt minutes.

2. **La liste des éléments** : Un sous-ensemble est représenté par la liste chainée de ses éléments. Cette représentation peut être avantageuse lorsque les sous-ensembles contiennent peu d'éléments, mais ne l'est absolument pas pour les sous-ensembles de taille importante. Il faut en effet, ajouter les données nécessaires à la gestion des listes. Par exemple, sur l'hypothèse d'un entier codé sur 64 chiffres binaires et des pointeurs de la même taille, la taille peut aller jusqu'à 128 fois celle du bitmap. La complexité des opérations n'est pas constante mais proportionnelle au nombre d'éléments des sous-ensembles. Comme nous ne disposons d'aucune information sur la taille à priori des sous-ensembles, cette représentation n'a pas été retenue.
3. **La réunion en intervalles disjoints** : Cette représentation a été retenue car nous avons réussi à concilier d'une part, la compacité de la représentation et d'autres part, l'efficacité. Nous avons trouvé une manière efficace de calculer les successeurs d'un intervalle de séquences en n'explorant que les bornes des intervalles. La complexité des opérations s'en trouve proportionnelle au nombre d'intervalles et non pas au nombre d'éléments.

   Exemple 2.1. Considérons l'ensemble $E = \{1, \ldots, 29\}$ de sous-ensemble
   $F = \{1,2,3,4,5,6,7,8,16,17,18,19,20,21,22,23,24,25,26,27,28,29\}$
   F est la liste des éléments représentée par le bitmap :
   $F = 11111111000000011111111111111$
   Cette représentation correspond à la réunion d'intervalle suivante :
   $F = [1,8] \cup [16,29]$.

   Pour manipuler ces réunions d'intervalles, la structure qui s'impose est celle d'arbre de recherche équilibré. Elle permet de :

   — tester simplement l'appartenance d'un élément à un sous-ensemble en temps logarithmique.

   — d'insérer et d'ôter des éléments ou des intervalles également en temps logarithmique.

   Pour rester efficace, l'arbre doit demeurer équilibré.

   Il existe plusieurs façon d'équilibrer les arbres. Nous avons tout d'abord testé une structure appelée B-tree. Les œuds parents peuvent posséder plus de deux nœuds enfants. Cette structure s'est avérée moins efficace pour notre besoin car dans la majorité des cas, deux nœuds ne comportaient que deux

successeurs. En conséquence, la mémoire n'était pas utilisée d'une manière optimale.

Nous nous sommes repliés sur la structure d'arbre AVL. [AVL62] qui forme une famille particulière des arbres binaires de recherche équilibrés. La dénomination « arbre AVL » provient des noms respectifs de ses deux inventeurs, respectivement Georgii Adelson-Velsky et Evguenii Landis.

La recherche, l'insertion et la suppression sont toujours logarithmique en le nombre d'éléments $\mathcal{O}(\log_2(N))$ où N désigne le nombre de nœuds de l'arbre, c'est-à-dire ici le nombre d'intervalles du sous-ensemble.

Notons qu'il aurait été possible d'envisager d'utiliser d'autres structures de données pour représenter les réunions d'intervalles, comme par exemple les tables de hachage.

La représentation des ensembles comme une réunion d'intervalles consécutifs, peut s'interpréter comme une compression des données binaires avec l'algorithme RLE (Run-Lenght-Encoding) [RC67].

Le principe du codage RLE est d'indiquer le nombre de répétions d'un élément puis le compléter par l'élément à répéter. Sur l'exemple ci-dessus, l'ensemble F est représenté par l'unique intervalle $[\mathsf{min}, \mathsf{max}]$, ce codage est donc très efficace lorsqu'il existe plusieurs symboles consécutifs dans un ensemble. Mais il est inefficace lorsqu'il ya beaucoup de variation. Le pire des cas est celui de la réunion de singleton non consécutive comme par exemple l'ensemble des nombres pairs :

Exemple 2.2. Soit F le bitmap donné par :

$F = 1010101010101010$

F se décrit par la réunion d'intervalle suivante :

$F = [0] \cup [2] \cup [4] \cup [6] \cup [8] \cup [10] \cup [12] \cup [14]$, ce qui est très inefficace.

À titre de perspectives, il est possible d'explorer d'autres types d'algorithmes de compression des données binaires pour avoir une représentation compacte et efficace des ensembles.

Une perspective envisageable serait de developper ce type de représentation pour dépasser les limites en complexité et en mémoire de la présente représentation.

Citons par exemple :

— l'algorithme de Ziv-Lempel [Wel84].

Cet algorithme est plus efficace en terme de compacité car on arrive à réduire d'avantage la taille par rapport à l'algorithme RLE. En revanche, il reste moins efficace lors de la compression car il faut un dictionnaire. De

plus, l'effectivité n'est pas résolue car il reste à implémenter les opérations ensemblistes sur ce type de modèle.

Rappelons que les opérations ensemblistes sont :

- test d'appartenance d'un élément à un sous-ensemble.
- réunion, comme ajouter un élément à un sous-ensemble.
- intersection, comme retirer un élément d'un sous-ensemble.
- complémentaire.

— le codage de Huffman [Huf52].

Le principe de ce codage est de coder avec des séquences binaires très courtes les mots les plus fréquents et avec des séquences binaires plus longues les mots les moins fréquents.

— le codage arithmétique [MMK03].

Ce codage représente une autre façon de coder les mots les plus fréquents avec des séquences binaires très courtes, et les mots les moins fréquents avec des séquences binaires plus longues.

Mais le problème à résoudre pour chacun de ces codages est de pouvoir répondre à ces questions :

— Est-ce que cela permet de réaliser efficacement les opérations ensemblistes ?

— Est ce que cela permet de réaliser le calculs des successeurs, de façon que la complexité dépende uniquement de la taille des données sans avoir à explorer exhaustivement tout le sous-ensemble ?

3 Description de l'application web

Comme nous l'avons vu, il existe de nombreux résultats pour les différentes constantes étudiées. Afin de faciliter la recherche pour la communauté scientifique, nous avons développé une application web. Concrètement, notre site web permet de récupérer des valeurs connues pour les constantes suivantes :

- La constante de Harborth
- La constante de Davenport
- La constante d'Erdős-Ginzburg-Ziv
- La constante $s_{\leq k}(G)$

L'utilisateur demande la valeur de la constante pour un groupe abélien fini donné. L'application vérifie si la constante de ce groupe est connue d'après les résultats

existant dans la littérature ou bien selon nos nouveaux résultats trouvés dans le cadre de cette thèse. Si c'est le cas, elle retourne le résultat de cette constante sinon elle envoie un message nous informant que la constante pour ce groupe n'est pas encore connue ou implémentée.

Notons qu'il est tout à fait envisageable d'enrichir cette application pour l'étude d'autres constantes. Nous prévoyons aussi de la mettre à jour lorsque de nouveaux résultas seront obtenus.

Cette application implémente plusieurs pages web :

— La page d'accueil présente des liens vers les descriptions de chaque constante.

FIGURE VI.1 – Capture d'écran de la page d'accueil

— Les pages de description définit chaque constante et présente les valeurs connues.

— Une page de recherche contient un formulaire pour calculer une constante spécifique en introduisant le groupe souhaité.

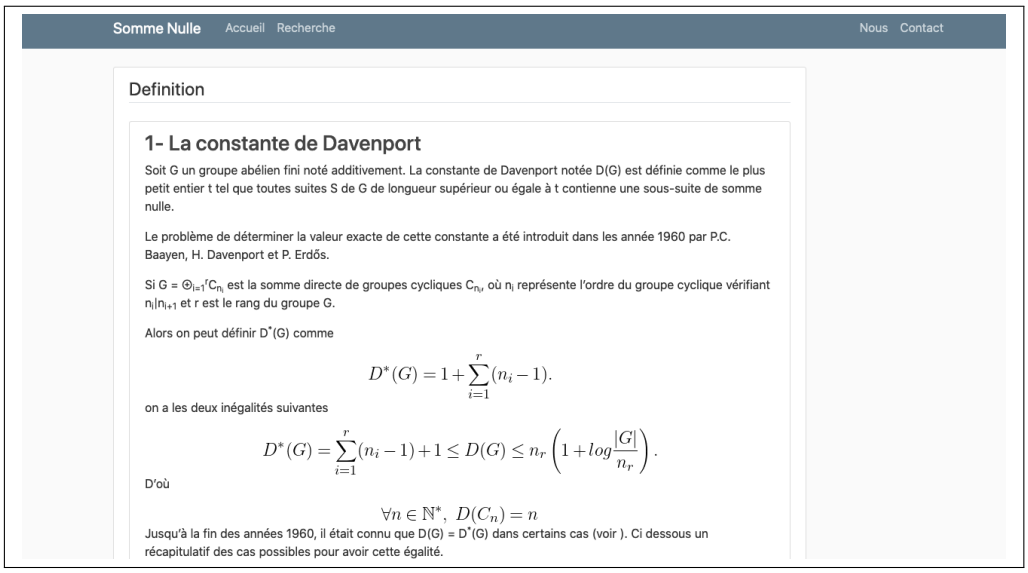


FIGURE VI.2 – Capture d'écran de la page de description de la constante de Davenport

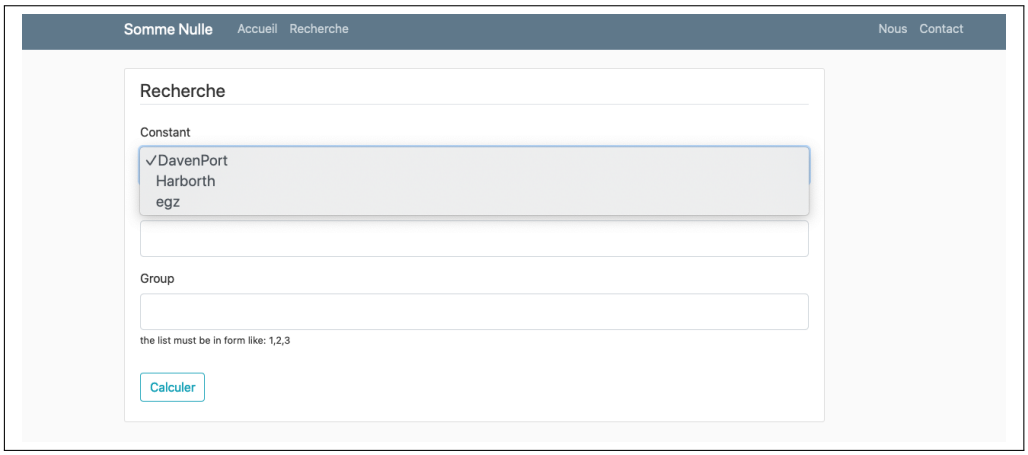


FIGURE VI.3 – Capture d'écran de la page du calcul

— Une fois la recherche lancée, une page affiche le résultat demandé.

L'application est implémentée en utilisant le framework flask. Flask est un framework python qui permet de développer des applications web. Il se base sur le modèle d'implémentation RESTfull. Le framework offre un ensemble de packages qui implémentent les fonctionnalités de base pour faire fonctionner l'application.

Les packages python utilisés sont :

— Flask, render_template, url_for, flash, redirect,request

— forms ,RegistrationForm, LoginForm, SearchForm, SearchOneForm

— wtforms.validators, DataRequired, Length, Email, EqualTo

Pour plus de détails de ce framework (voir [Fla]).

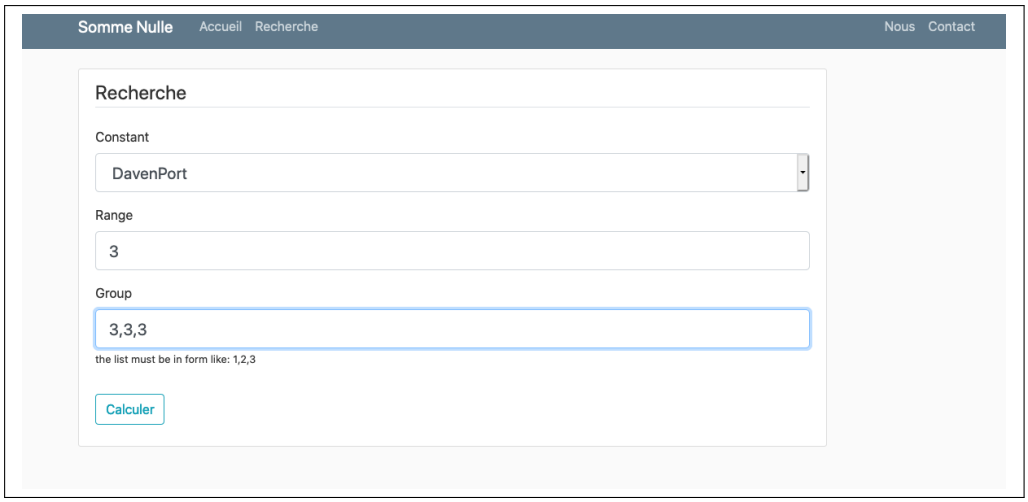


FIGURE VI.4 – Capture d'écran des cases de recherche remplies

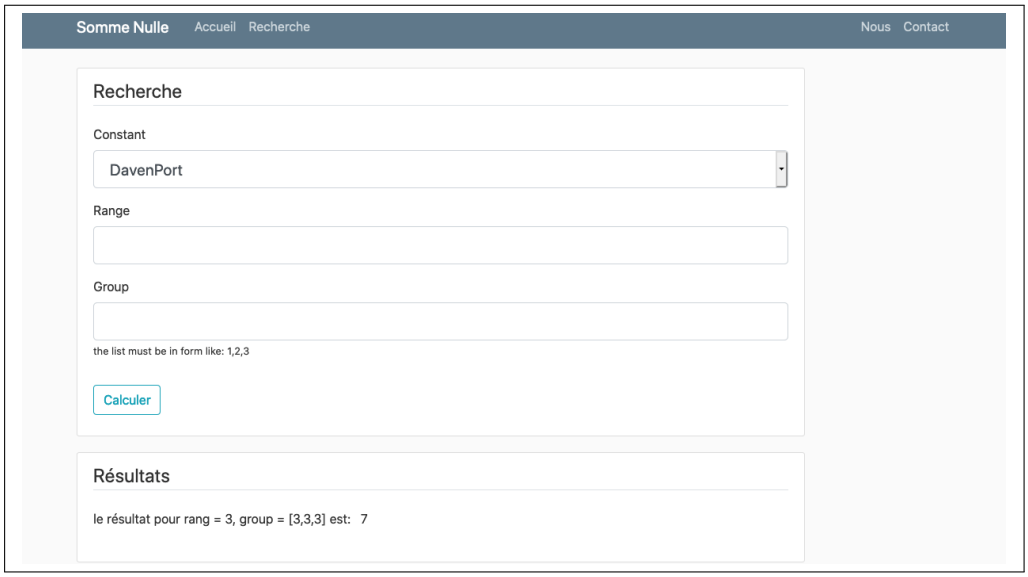


FIGURE VI.5 – Capture d'écran du résultat final

L'application est en cours de développement et n'est pas encore accessible au public. Cependant, il est possible de la lancer depuis son poste local (de préférence depuis MacOS ou Linux) en téléchargeant le code source disponible sur https://github.com/Zerdoum/sommeNulle et en suivant les étapes d'installation comme décrit dans le fichier README, voir https://github.com/Zerdoum/sommeNulle/blob/master/README.md

Le dossier racine contient les fichiers python suivants :

— main.py : c'est le premier fichier qui sera appelé par l'application. Il crée une instance d'application et définit toutes les routes (les URL) qui appellent les différentes fonctions et pages de l'application.

— core.py : est appelé par le fichier main.py. Ce fichier vérifie quelle constante

est appelée par le formulaire de recherche. Il fait appel aux différents fichiers python qui implémentent le calcul des constantes. : davenPort.py, egz.py, harborth.py, others.py.

— forms.py : définit la structure des classes d'objets utilisés pour l'affichage des différentes pages, comme : forme texte, liste, buttons de validations etc.

— davenPort.py, egz.py, harborth.py, others.py : ces fichiers implémentent les différentes fonctions qui font le calcul des constantes selon un paramétrage saisi par l'utilisateur. Ces paramètres sont : le nom de la constante et le groupe abélien fini. Ce dernier doit être saisi sous forme d'une liste numérique séparée par des virgules. De plus, on impose la contrainte qu'un nombre doit être divisible par le suivant. Prenant l'exemple de la liste numérique, $2, 2, 4$, cette liste représente le groupe $G = C_2 \oplus C_2 \oplus C_4$. Il serait sans doute aisé de compléter ce travail par une saisie moins contraignante.

— Les dossiers /static, /templates contiennent les différents fichiers .html, .css etc, qui implémentent la partie visuelle de l'application web, c'est-à-dire, les pages web, les formulaires dont celui de la recherche, les boutons, les couleurs, etc.

— run.sh est le fichier à lancer pour démarrer l'application. Il définit quelques paramètres qui permettent le lancement de l'application.

Le dessin suivant visualise le plan du site.

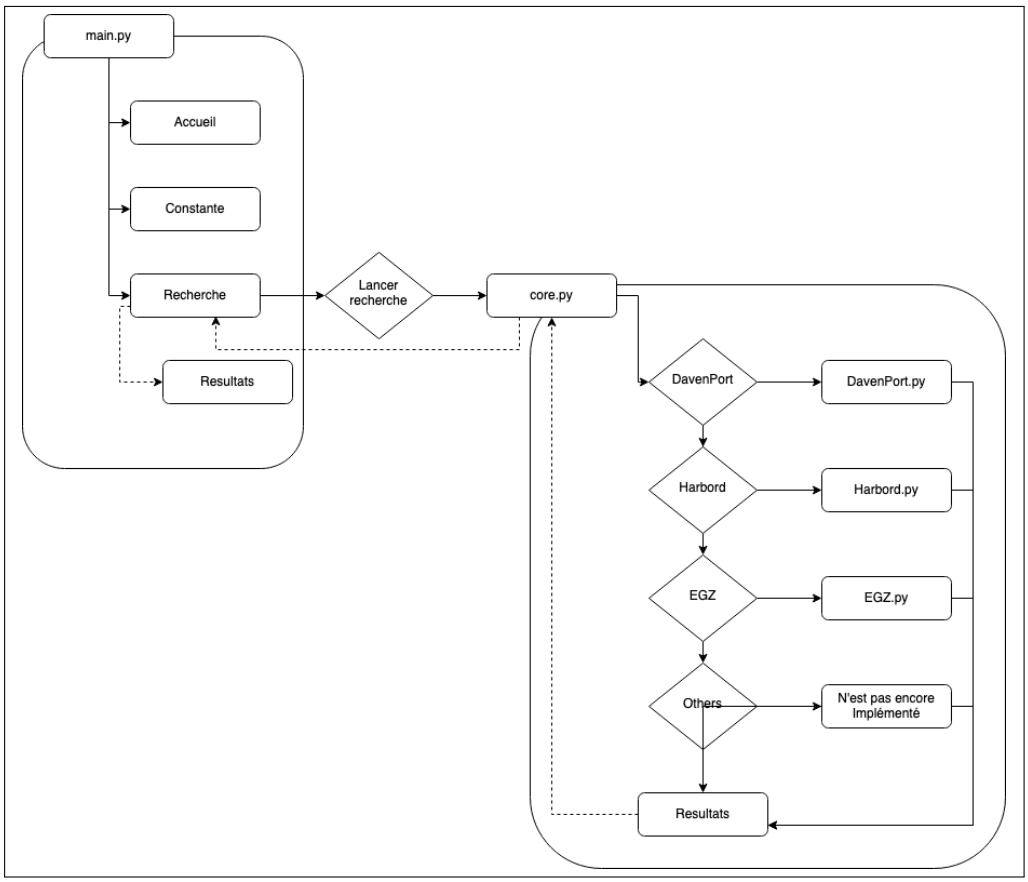


FIGURE VI.6 – Le plan du site

LISTE DES FIGURES

LISTE DES TABLEAUX

TOUTE LA BIBLIOGRAPHIE

[AVL62] George M Adel'son-Vel'skii et Evgenii Mikhailovich Landis. « An algorithm for organization of information ». In : *Doklady Akademii Nauk*. T. 146. 2. Russian Academy of Sciences. 1962, p. 263-266.

[AD93] Noga Alon et Moshe Dubiner. « Zero-sum sets of prescribed size ». In : *Combinatorics, Paul Erdos is Eighty* 1 (1993), p. 33-50.

[Baa69] P.C. Baayen. « $(C_2 \oplus C_2 \oplus C_2 \oplus C_{2n})$! » In : ZW 6/69 (1969).

[Baj18] Bela Bajnok. *Additive Combinatorics : A Menu of Research Problems*. Chapman et Hall/CRC, 2018.

[BHSP09] Gautami Bhowmik, Immanuel Halupczok et Jan-Christoph Schlage-Puchta. « Inductive methods and zero-sum free sequences ». In : *Integers* 9.5 (2009), p. 515-536.

[BSP07] Gautami Bhowmik et Jan-Christoph Schlage-Puchta. « Davenport's constant for groups of the form $\mathbb{Z}_3 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_{3d}$ ». In : *Additive combinatorics* 43 (2007), p. 307-326.

[Boa69] Peter vanEmde Boas. « A combinatorial problem on finite abelian groups, 2 ». In : *Stichting Mathematisch Centrum. Zuivere Wiskunde* ZW 7/69 (1969).

[Cha+02] ST Chapman et al. « On Davenport's constant of finite abelian groups ». In : *FAR EAST JOURNAL OF MATHEMATICAL SCIENCES* 5.1 (2002), p. 47-54.

[CS14] Fang Chen et Svetoslav Savchev. « Long minimal zero-sum sequences in the groups $C_{r-1}^2 \oplus C_{2k}$ ». In : *Integers* 14 (2014), A23.

[CLP17] Ernie Croot, Vsevolod F Lev et Péter Pál Pach. « Progression-free sets in are exponentially small ». In : *Annals of Mathematics* (2017), p. 331-337.

[DOQ01] Charles Delorme, Oscar Ordaz et Domingo Quiroz. « Some remarks on Davenport constant ». In : *Discrete Mathematics* 237.1-3 (2001), p. 119-128.

[Ede+02] Yves Edel et al. « The classification of the largest caps in AG (5, 3) ». In : *Journal of Combinatorial Theory, Series A* 99.1 (2002), p. 95-110.

[Ede+07] Yves Edel et al. « Zero-sum problems in finite abelian groups and affine caps ». In : *Quarterly journal of mathematics* 58.2 (2007), p. 159-186.

[EG17] Jordan S Ellenberg et Dion Gijswijt. « On large subsets of with no three-term arithmetic progression ». In : *Annals of Mathematics* (2017), p. 339-343.

[EGZ61] Paul Erdos, Abraham Ginzburg et Abraham Ziv. « Theorem in the additive number theory ». In : *Bull. Res. Council Israel F* 10 (1961), p. 41-43.

[FGZ11] Yushuang Fan, Weidong Gao et Qinghai Zhong. « On the Erdős–Ginzburg–Ziv constant of finite abelian groups of high rank ». In : *Journal of Number Theory* 131.10 (2011), p. 1864-1874.

[FZ16] Yushuang Fan et Qinghai Zhong. « On the Erdős–Ginzburg–Ziv constant of groups of the form $C_2^r \oplus C_n$ ». In : *International Journal of Number Theory* 12.04 (2016), p. 913-943.

[FLP14] Robert Faure, Bernard Lemaire et Christophe Picouleau. *Précis de recherche opérationnelle-7e éd. : Méthodes et exercices d'application*. Dunod, 2014.

[Fla] Flask. https://flask.palletsprojects.com/en/1.1.x/.

[GGS07] WD Gao, Alfred Geroldinger et Wolfgang A Schmid. « Inverse zero-sum problems ». In : *ACTA ARITHMETICA-WARSZAWA-* 128.3 (2007), p. 245.

[GT04] WD Gao et R Thangadurai. « A variant of Kemnitz conjecture ». In : *Journal of Combinatorial Theory, Series A* 107.1 (2004), p. 69-86.

[Gao+07] WD Gao et al. « On short zero-sum subsequences II ». In : *Integers* 7.1 (2007), Paper-A21.

[GG06] Weidong Gao et Alfred Geroldinger. « Zero-sum problems in finite abelian groups : a survey ». In : *Expositiones Mathematicae* 24.4 (2006), p. 337-369.

[GHK06] Alfred Geroldinger et Franz Halter-Koch. *Non-unique factorizations : Algebraic, combinatorial and analytic theory*. Chapman et Hall/CRC, 2006.

[GS19] Benjamin Girard et Wolfgang A Schmid. « Direct zero-sum problems for certain groups of rank three ». In : *Journal of Number Theory* 197 (2019), p. 297-316.

[Gry13] David J Grynkiewicz. *Structural additive theory*. T. 30. Springer Science & Business Media, 2013.

[Gui] Philippe Guillot. « Communication privée ». In : ().

[Gui+19] Philippe Guillot et al. « On the Harborth constant of $C_3 \oplus C_{3p}$ ». In : *Journal de Théorie des Nombres de Bordeaux* 31.3 (2019), p. 613-633.

[Har73] Heiko Harborth. « Ein Extremalproblem für Gitterpunkte. » In : *Journal für die reine und angewandte Mathematik* 262 (1973), p. 356-360.

[Huf52] David A Huffman. « A method for the construction of minimum-redundancy codes ». In : *Proceedings of the IRE* 40.9 (1952), p. 1098-1101.

[Kem83] Arnfried KEMNITZ. « On a lattice point problem ». In : *Ars Combin* 16 (1983), p. 151-160.

[Kie16] C KIEFER. « Examining the maximum size of zero-h-sum-free subsets ». In : *Research Papers in Mathematics (B. Bajnok, ed.), Gettysburg College* 19 (2016).

[Luo17] Sammy LUO. « Short zero-sum sequences over abelian p-groups of large exponent ». In : *Journal of Number Theory* 177 (2017), p. 28-36.

[MMK03] David JC MACKAY et David JC MAC KAY. *Information theory, inference and learning algorithms*. Cambridge university press, 2003.

[Mar+13] Luz E MARCHAN et al. « Some exact values of the Harborth constant and its plus-minus weighted analogue ». In : *Archiv der Mathematik* 101.6 (2013), p. 501-512.

[Mor17] Joy MORRIS. *Combinatorics : An Upper-level Introductory Course in Enumeration, Graph Theory, and Design Theory*. Joy Morris, 2017.

[Nar13] Wladyslaw NARKIEWICZ. *Elementary and analytic theory of algebraic numbers*. Springer Science & Business Media, 2013.

[Ols69a] John E OLSON. « A combinatorial problem on finite abelian groups, I ». In : *Journal of number theory* 1.1 (1969), p. 8-10.

[Ols69b] John E OLSON. « A combinatorial problem on finite abelian groups, II ». In : *Journal of Number Theory* 1.2 (1969), p. 195-199.

[PVEB69] D. KRUYSWIJK P. VAN EMDE BOAS. « A combinatorial problem on finite abelian groups, 3 ». In : *Stichting Mathematisch Centrum. Zuivere Wiskunde* (1969).

[Pot08] Aaron POTECHIN. « Maximal caps in AG (6, 3) ». In : *Designs, Codes and Cryptography* 46.3 (2008), p. 243-259.

[RST08] P RATH, K SRILAKSHMI et R THANGADURAI. « On Davenport's constant ». In : *International Journal of Number Theory* 4.01 (2008), p. 107-115.

[Rei07] Christian REIHER. « On Kemnitz'conjecture concerning lattice-points in the plane ». In : *The Ramanujan Journal* 13.1-3 (2007), p. 333-337.

[RC67] AH ROBINSON et Colin CHERRY. « Results of a prototype television bandwidth compression scheme ». In : *Proceedings of the IEEE* 55.3 (1967), p. 356-364.

[Rog63] Kenneth ROGERS. « A combinatorial problem in Abelian groups ». In : *Mathematical Proceedings of the Cambridge Philosophical Society*. T. 59. 3. Cambridge University Press. 1963, p. 559-562.

[Rón00] Lajos RÓNYAI. « On a conjecture of Kemnitz ». In : *Combinatorica* 20.4 (2000), p. 569-573.

[SC05] Svetoslav SAVCHEV et Fang CHEN. « Kemnitz'conjecture revisited ». In : *Discrete mathematics* 297.1-3 (2005), p. 196-201.

[Sch11] Wolfgang A. SCHMID. « The inverse problem associated to the Davenport constant for $C_2 \oplus C_2 \oplus C_{2n}$, and applications to the arithmetical characterization of class groups ». In : *Electron. J. Combin.* 18.1 (2011), Paper 33, 42.

[SZ10] Wolfgang A SCHMID et JJ ZHUANG. « On short zero-sum subsequences over p-groups ». In : *Ars Combin., to appear* (2010).

[WZ17] Chunlin WANG et Kevin ZHAO. « On zero-sum subsequences of length not exceeding a given number ». In : *Journal of Number Theory* 176 (2017), p. 365-374.

[Wel84] Terry A. WELCH. « A technique for high-performance data compression ». In : *Computer* 6 (1984), p. 8-19.