Théorie de Ramsey

Thomas Budzinski

27 mai 2024

Introduction

La théorie de Ramsey est une branche de la combinatoire, dont les résultats sont souvent de la forme suivante : "si on colorie en r couleurs les éléments d'une certaines structure très grande, alors il existe une sous-structure assez grande dont tous les éléments sont de la même couleur". Autrement dit, le désordre complet n'existe pas. Voici les deux exemples les plus connus de résultats issus de cette théorie :

- Théorème de Ramsey : Soit $k \geq 1$. Si n est suffisamment grand, alors dans tout groupe de n personnes, on peut en trouver k qui sont deux à deux amies entre elles, ou bien k qui sont deux à deux non-amies entre elles.
- Théorème de Van der Waerden : Si on colorie les entiers naturels en r couleurs, alors il existe des progressions arithmétiques monochromes arbitrairement grandes.

De manière générale, les techniques utilisées sont très élémentaires, mais la théorie de Ramsey a des liens avec diverses branches des mathématiques : probabilités, théorie des nombres, logique, et même géométrie des espaces de Banach [1]. On verra aussi que de nombreuses questions qui peuvent sembler basiques sont toujours ouvertes. En particulier, l'écart entre les meilleures bornes inférieures et les meilleures bornes supérieures connues pour certains problèmes est parfois colossal !

Quelques références. Une référence assez complète sur le sujet est l'ouvrage de Graham, Rothschild et Spencer [12], qui couvre bien plus que le contenu de ces cours, et où les résultats sont en général présentés de la manière la plus générale possible. La partie sur la méthode probabiliste est largement issue de l'excellent livre d'Alon et Spencer [1], qui expose la méthode probabiliste ainsi que des applications à divers domaines, bien au-delà de la théorie de Ramsey. Enfin, si vous êtes curieux sur les liens entre la théorie de Ramsey, la logique et les fonctions à croissance (très, très) rapide, je recommande le livre récent de Katz et Reimann [14].

1 Théorème de Ramsey

Principe des tiroirs. Si le théorème de Ramsey est le "premier" théorème de la théorie de Ramsey, on commençons par mentionner le principe des tiroirs, qu'on peut considérer comme le "zéroième" théorème de la théorie.

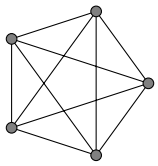
1. Ce dernier sujet a fait l'objet de travaux de Gowers, mais est trop avancé et ne sera pas abordé dans ce cours.

Théorème 0 (Principe des tiroirs)**.** Soient $k, \ell \geq 2$. Si on colorie en k couleurs les éléments d'un ensemble S de taille $k(\ell-1)+1$, alors il existe un sous-ensemble $A \subset S$ de taille ℓ dont tous les éléments sont de la même couleur.

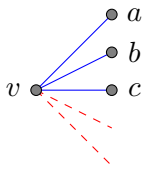
Il s'agit du résultat "de type Ramsey" le plus simple, puisque contrairement aux résultats qui suivront, l'ensemble S sous-jacent n'est muni d'aucune structure. Le principe des tiroirs sera très utilisé dans les démonstrations des prochains théorèmes.

Définition 1. Un *graphe* est un ensemble de sommets, reliés entre eux par des arêtes. Un *graphe complet* à $n \geq 2$ sommets, noté K_n, est un graphe à n sommets où toute paire de sommets distincts $x \neq y$ est reliée par un arête.

Exemple 2. Le graphe K_5 :



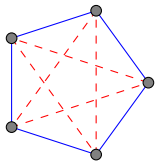
Avant d'aborder le théorème de Ramsey, commençons par un cas particulier : on colorie en bleu et rouge les arêtes de K_6. On va montrer qu'il existe soit un triangle tout bleu, soit un triangle tout rouge. Pour cela, on se fixe un sommet v.



Alors 5 arêtes sont issues de v, donc au moins 3 sont de la même couleur d'après le principe des tiroirs. Supposons par exemple qu'elles soient bleues. On appelle a, b et c les trois bouts (autres que v) de ces trois arêtes :

- si deux sommets parmi a, b et c sont reliés par une arête bleue, alors ces deux sommets forment un triangle bleu avec v ;
- sinon, alors a, b et c forment un triangle rouge.

Par ailleurs, la figure ci-dessous montre qu'il est possible de colorier K_5 sans triangle monochrome :



On dira donc que *le troisième nombre de Ramsey vaut* 6. Plus généralement, ceci motive la définition suivante.

Définition 3. Soient $\ell, m \geq 2$. On note $R(\ell, m)$ le plus petit entier n tel que tout coloriage des arêtes de K_n en bleu et rouge admet un K_ℓ bleu ou un K_m rouge (avec la convention $R(\ell, m) = +\infty$ si un tel n n'existe pas).

Commençons par quelques remarques très simples :

- On a déjà vu que $R(3,3) = 6$.
- Les nombres de Ramsey sont symétriques, i.e. $R(\ell, m) = R(m, \ell)$ pour tous ℓ, m.
- On a $R(2, \ell) = R(\ell, 2) = \ell$ pour tout ℓ. En effet, si toutes les arêtes de K_ℓ sont rouges, alors on a un K_ℓ rouge, et sinon on a une arête bleue et donc un K_2 bleu.

Passons maintenant au théorème de Ramsey.

Théorème 1 (Théorème de Ramsey)**.** Pour tous $\ell, m \geq 3$, on a

$$R(\ell, m) \leq R(\ell - 1, m) + R(\ell, m - 1). \tag{1}$$

On a donc $R(\ell, m) \leq \binom{l+m-2}{\ell-1}$ pour tous $\ell, m \geq 2$. En particulier, les nombres de Ramsey sont finis.

La finitude des nombres de Ramsey a été montrée par Ramsey en 1930 [16] avec un argument un peu compliqué. La version quantitative du théorème et la démonstration qui suit sont dues à Erdös et Szekeres [8]. Notons aussi que (1) évoque la formule du triangle de Pascal. L'apparition d'un coefficient binomial n'est donc pas surprenante.

Démonstration. On commence par l'inégalité (1). Si $R(\ell - 1, m)$ ou $R(\ell, m - 1)$ est infini, il n'y a rien à faire. Sinon, soit $n = R(\ell - 1, m) + R(\ell, m - 1)$. On considère un coloriage en bleu et rouge des arêtes de K_n. Comme plus haut, on fixe un sommet v. On note B (resp. C) l'ensemble des sommets de K_n qui sont reliés à v par une arête bleue (resp. rouge). Alors on a

$$|B| + |C| = n - 1 = R(\ell - 1, m) + R(\ell, m - 1) - 1,$$

donc $|B| \geq R(\ell - 1, m)$ ou $|C| \geq R(\ell, m - 1)$. Sans perte de généralité, supposons $|B| \geq R(\ell - 1, m)$. Par définition de $R(\ell - 1, m)$, l'ensemble B contient soit un $K_{\ell-1}$ bleu, soit un K_m rouge. Dans le premier cas, on obtient un K_ℓ bleu en ajoutant v. Dans le second, on a directement un K_m rouge, d'où $R(\ell, m) \leq n$.

On en déduit $R(\ell, m) \leq \binom{\ell+m-2}{\ell-1}$ par récurrence sur $\ell + m \geq 4$. Le résultat est immédiat pour $\ell + m = 4$. Si de plus le résultat est montré pour tous ℓ, m avec $\ell + m = k - 1$, soient $\ell, m \geq 2$ avec $\ell + m = 4$. Si ℓ ou m vaut 2, l'inégalité est immédiate. Sinon, par hypothèse de récurrence, on a

$$R(\ell, m) \leq R(\ell - 1, m) + R(\ell, m - 1) \leq \binom{m + \ell - 3}{\ell - 2} + \binom{m + \ell - 3}{\ell - 1} = \binom{\ell + m - 2}{\ell - 1}.$$

□

Remarques 4.

1. Cette borne donne $R(3, 4) \leq 10$, puis $R(4, 4) \leq 20$ et $R(3, 5) \leq 15$. On a en fait $R(3, 4) = 9$ et $R(4, 4) = 18$, ainsi que $R(3, 5) = 14$ (voir TD).
2. En revanche, la valeur $R(5, 5)$ est inconnue ! Les meilleures bornes connues sont

   $$43 \leq R(5, 5) \leq 48.$$

   La borne inférieure est issue de [9] ; et la borne supérieure de [2]. Elle ne date que de 2017. Pour $R(10, 10)$, l'écart devient important : les meilleures bornes connues sont $798 \leq R(10, 10) \leq 23556$. Pour plus d'informations sur les "petits" nombres de Ramsey, vous pouvez consulter la page Wikipédia du théorème de Ramsey.

3. Voici ce que disait Paul Erdös sur nos chances de calculer $R(5,5)$ et $R(6,6)$:

   Suppose aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find the Ramsey number for red five and blue five. We could marshal the world's best minds and fastest computers, and within a year we could probably calculate the value. If the aliens demanded the Ramsey number for red six and blue six, however, we would have no choice but to launch a preemptive attack.

4. On peut aussi s'interroger sur l'asymptotique des nombres de Ramsey quand les paramètres tendent vers l'infini. Pour les nombres diagonaux $R(\ell,\ell)$, notre borne est de la forme
$$R(\ell,\ell) \leq \binom{2\ell-2}{\ell-1} \sim (1+o(1))\frac{1}{4\sqrt{\pi}}\frac{4^\ell}{\sqrt{\ell}}.$$
   Les meilleures bornes asympotiques connues sont les suivantes :
$$(1+o(1))\frac{\sqrt{2}}{e}\ell\sqrt{2}^\ell \leq R(\ell,\ell) \leq \left(4-\frac{1}{128}\right)^\ell,$$
   où c est une constante strictement positive. La borne inférieure est dûe à Spencer [18] et date de 1970 (on donnera la preuve plus loin dans le cours), tandis que la borne supérieure n'a été prouvée u'il y a quelques mois par Campos, Griffiths, Morris et Sahasrabudhe [5]. Il s'agit d'un résultat majeur en combinatoire : la question de savoir si les nombres de Ramsey diagonaux sont d'ordre $4^{\ell+o(\ell)}$ était ouverte depuis plus de 80 ans ! La meilleure borne précédente, elle aussi très récente, était de l'ordre de $\exp\left(-c(\log\ell)^2\right)4^\ell$ et était dûe à Sah [17]. Notons que l'écart entre la borne supérieure et la borne inférieure reste considérable.

5. Enfin, on peut généraliser le théorème de Ramsey à des coloriages à plus de deux couleurs. On obtient des nombres de Ramsey de la forme $R(\ell_1,\ell_2,\ldots,\ell_k)$, qui sont eux aussi finis.

Le théorème de Ramsey infini On peut donner une version "infinie" du théorème de Ramsey.

Théorème 2. Soit S un ensemble dénombrable. On relie chaque paire d'éléments distincts de S par une arête soit bleue, soit rouge. Alors il existe un sous-ensemble A de S infini tel que toutes les arêtes reliant deux sommets de A soient de la même couleur.

Notons que le théorème de Ramsey fini nous garantit qu'on peut trouver des K_n monochromes arbitrairement grand, mais que la conclusion du théorème de Ramsey infini est bien plus forte !

Démonstration. On va construire, par récurrence sur n, une suite $(x_n)_{n\geq 1}$ d'éléments de S, ainsi qu'une suite $(S_n)_{n\geq 1}$ de sous-ensembles de S, qui vérifient les propriétés suivantes :

- pour tout n, l'ensemble S_n est infini ;
- pour tout n, on a $S_{n+1} \subset S_n$;
- pour tout n, on a $x_{n+1} \in S_n$ mais $x_n \notin S_n$;
- pour tout n, les arêtes reliant x_n aux sommets de S_n sont toutes de la même couleur, notée c_n.

L'idée derrière cette construction est que le sommets x_1 va jouer le même rôle que le v de la preuve précédente, mais comme on raisonne sur un ensemble infini, on doit itérer la construction une infinité de fois plutôt que de faire une récurrence.

Passons à la construction. On choisit x_1 quelconque, et on choisit une couleur c_1 telle qu'il y a une infinité d'arêtes de couleur c_1 issues de x_1. On note S_1 l'ensemble des sommets de S reliés à x_1 par une arête de couleur c_1. Puis on itère cette construction : si x_n et S_n ont été construits, on choisit $x_{n+1} \in S_n$ arbitrairement. Comme S_n est infini par hypothèse de récurrence, il existe une couleur c_{n+1} qui apparaît une infinité de fois dans les arêtes entre x_{n+1} et $S_n \backslash \{x_{n+1}\}$. Enfin, on note S_{n+1} l'ensemble des sommets de S_n qui sont reliés à x_{n+1} par une arête de la couleur c_{n+1}. Par construction, on vérifie facilement que x_{n+1} et S_{n+1} vérifient toutes les propriétés voulues.

Pour terminer la preuve, on note qu'il existe une couleur qui apparaît une infinité de fois dans la liste $(c_n)_{n\geq 1}$. Sans perte de généralité, supposons que c'est le bleu, et posons

$$A = \{x_n \,|\, c_n = \text{bleu}\}\,.$$

Si deux sommets sont dans A, ils sont de la forme x_m, x_n avec $m < n$. Alors $X_n \in S_{n-1} \subset S_m$, donc comme $c_m =$ bleu, ces deux sommets sont reliés par une arête bleue, ce qui conclut la preuve. □

Bien que les preuves soient similaires, le théorème de Ramsey infini est "plus fort" que la version finie. En effet, on ne peut pas déduire la version infinie de la version finie, mais l'inverse est possible. On obtient un théorème fini moins quantitatif que celui obtenu ci-dessus, mais on présente quand même l'argument car il s'agit d'un type d'argument très classique en combinatoire.

Démonstration de la finitude des nombres de Ramsey en utilisant le théorème infini : Soit $\ell \geq 3$. On veut montrer qu'il existe $n \geq 1$ tel que tout coloriage des arêtes de K_n en bleu et rouge admet un K_ℓ monochrome. On raisonne par l'absurde et on identifie, pour tout n, les sommets de K_n avec les entiers $1, 2, \ldots, n$. Alors pour tout n, il existe un coloriage C_n des arêtes (i, j) avec $1 \leq i < j \leq n$ pour lequel, pour tout $A \subset \{1, 2, \ldots, n\}$ de cardinal ℓ, les arêtes dans A ne sont pas toutes de la même couleur. À partir des C_n, on va construire un coloriage infini C qui contredira le théorème de Ramsey infini.

Soit donc $C(1, 2)$ une couleur qui apparaît une infinité de fois dans la liste $(C_n(1, 2))_{n\geq 1}$, et soit $S_{1,2}$ l'ensemble des n tels que $C_n(1, 2) = C(1, 2)$. Puis soit $C(1, 3)$ une couleur qui apparaît une infinité de fois dans la liste $(C_n(1, 3))_{n\in S_{1,2}}$, et soit $S_{1,3}$ l'ensemble des $n \in S_{1,2}$ tels que $C_n(1, 3) = C(1, 3)$. Puis soit $C(2, 3)$ une couleur qui apparaît une infinité de fois dans la liste $(C_n(2, 3))_{n\in S_{1,3}}$, et soit $S_{2,3}$ l'ensemble des $n \in S_{1,3}$ tels que $C_n(2, 3) = C(2, 3)$. Et on continue ainsi de suite : on définit de même $C(1, 4)$ et $S_{1,4} \subset S_{2,3}$, puis $C(2, 4)$ et $S_{2,4}$, et ainsi de suite. On définit ainsi $C(i, j)$ et $S_{i,j}$ pour toute paire $i < j$ d'entiers naturels, puisque ces paires sont dénombrables.

On note qu'alors, pour tout k, l'ensemble $S_{k,k+1}$ est infini, et pour tous $n \in S_{k,k+1}$ et $1 \leq i < j \leq k$, on a $C_n(i, j) = C(i, j)$. Soit maintenant $A \subset \mathbb{N}^*$ de taille ℓ, et soit $k = \max(A)$. Soit aussi $n \in S_{k,k+1}$. Alors pour tous $i < j$ dans A, on a $C(i, j) = C_n(i, j)$. En particulier, comme A n'est pas monochrome pour le coloriage C_n, il ne l'est pas non plus pour C. Le coloriage C n'admet donc pas de K_ℓ monochrome, ce qui contredit le théorème de Ramsey infini. □

Remarque 5. L'argument qu'on vient de voir consiste à réaliser une extraction diagonale, ce qui n'est pas sans rappeler certains arguments de topologie (en particulier, la preuve du théorème de Tychonov, dans le cas particulier d'un produit dénombrable d'espaces métriques). Ce n'est pas une coïncidence : cet argument est appelé *argument de compacité* en combinatoire. En effet, notre preuve a consisté à extraire une valeur d'adhérence de la suite de coloriages (C_n) (pour une certaine topologie [2] sur l'espace des coloriages). Le dernier paragraphe de la preuve, lui, consiste à montrer que l'ensemble des coloriages sans K_ℓ monochrome est un fermé.

2 Théorème de Van der Waerden

Le théorème de Van der Waerden est un théorème du même type que celui de Ramsey, mais où la structure sous-jacente n'est plus une structure de graphe, mais est donnée par l'addition dans les entiers naturels.

Définition 6. Soit $k \geq 2$. Une *progression arithmétique de longueur* k est un ensemble d'entiers naturels de la forme $\{a, a+d, a+2d, \ldots, a+(k-1)d\}$ avec $a, d \geq 1$.

Théorème 3 (Van der Waerden, 1927)**.** Soient $k, r \geq 2$. Il existe N tel que tout coloriage en r couleurs des entiers de 1 à N admet au moins une progression arithmétique monochrome de longueur k. On notera $W(k, r)$ le plus petit entier N vérifiant cette propriété.

Remarques 7.
- Ce théorème est équivalent à la version (plus faible en apparence) suivante : "Tout coloriage des entiers naturels en r couleurs admet une progression arithmétique monochrome de longueur k". Un sens est immédiat, et l'autre se montre avec le même argument de compacité que celui utilisé ci-dessus pour le théorème de Ramsey.
- Le cas $k = 2$ découle du principe des tiroirs, et on a $W(2, r) = r + 1$ pour tout $r \geq 2$.

Ce théorème a été montré par Van der Waerden dans [20]. La preuve qu'on va présenter est dûe à Graham et Rotschild en 1974 [13]. Signalons aussi qu'une preuve reposant sur des arguments topologiques plutôt que combinatoires a été trouvée par Furstenberg et Weiss [10].

Le cas $(k, r) = (3, 2)$. Avant de s'attaquer au cas général, on va traiter le premier cas non trivial : le cas $(k, r) = (3, 2)$. Plus précisément, on va montrer que $W(3, 2) \leq 325$. Cette borne est très mauvaise (on a en fait $W(3, 2) = 9$), mais la preuve contient une bonne partie des idées du cas général.

On sépare donc les entiers de 1 à 325 en 65 blocs $B_1, \ldots, B_{65}$ de longueur 5, i.e. $B_i = \{5(i-1)+1, 5(i-1)+2, \ldots, 5i\}$. Il y a $2^5 = 32$ manières de colorier un bloc en bleu et rouge, donc parmi les blocs $B_1, \ldots, B_{33}$, il existe deux blocs B_i et B_{i+d} coloriés de manière identique.

Parmi les trois premiers éléments du bloc B_i, deux sont de la même couleur, disons bleue. Supposons que ce sont le premier et le troisième (les autres cas se traitent de manière similaire). Alors, si il n'y a pas de progression arithmétique monochrome de longueur 3 :
- $5(i-1)+1$ et $5(i-1)+3$ sont bleus, donc $5(i-1)+5$ est rouge ;
- les blocs B_i et B_{i+d} sont identiques, donc $5(i+d-1)+1$ et $5(i+d-1)+3$ sont bleus, et $5(i+d-1)+5$ est rouge.

2. Plus précisément, cette topologie provient par exemple de la métrique $d(C, C') = 2^{-k(C,C')}$, où $k(C, C')$ est le plus petit entier k pour lequel il existe $i < k$ tel que $C(i, k) \neq C(i', k)$.

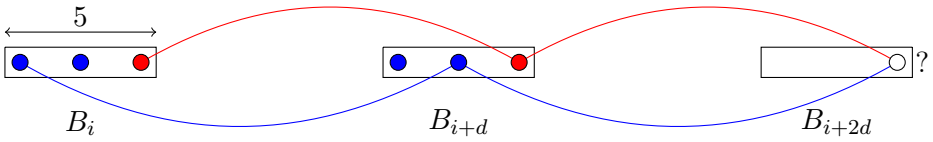


FIGURE 1 – Illustration de la preuve du théorème de Van der Waerden pour $k = 3$ et $r = 2$.

Mais alors $\{5(i-1)+5, 5(i+d-1)+5, 5(i+2d-1)+5\}$ et $\{5(i-1)+1, 5(i+d-1)+3, 5(i+2d-1)+5\}$ sont deux progressions arithmétiques de raisons respectives $5d$ et $5d+2$, donc $5(i+2d-1)+5$ ne peut être ni rouge ni bleu, ce qui est impossible.

Notons que dans cet argument, on a utilisé deux fois le principe des tiroirs, c'est-à-dire le théorème de Van der Waerden pour $r = 2$: la première fois sur un coloriage à 32 couleurs, et la deuxième fois sur un coloriage à 2 couleurs. On sent donc que pour montrer le théorème pour $(k+1, r)$, on peut avoir besoin du théorème pour (k, r') avec k' bien plus grand que r.

Le cas $(k, r) = (3, 3)$. Il semble donc que dans le théorème de Van der Waerden, il est plus facile d'augmenter le nombre de couleurs r que la couleur k. Le cas le plus naturel à traiter est donc maintenant $(k, r) = (3, 3)$. On va donc maintenant montrer la borne :

$$W(3, 3) \leq t(2 \times 3^t + 1),$$

où $t = 7(2 \times 3^7 + 1) = 30625$. À nouveau, cette borne est abominable, puisqu'on a en fait $W(3, 3) = 27$.

Comme précédemment, on sépare les entiers de 1 à $t(2 \times 3^t + 1)$ en $2 \times 3^t + 1$ blocs B_i de longueur t. Il y a 3^t manières de colorier un bloc en bleu, rouge et vert, donc il y a deux blocs identiques $B_{i_1}, B_{i_1+d_1}$ parmi les $3^t + 1$ premiers, de sorte que $i_1 + 2d_1 \leq 2 \times 3^t + 1$.

On divise maintenant chaque bloc B_i en $2 \times 3^7 + 1$ sous-blocs $B_{i,j}$ de longueur 7. Il y a 3^7 manières de colorier un sous-bloc en bleu, rouge et vert, donc il existe $1 \leq i_2 < i_2+d_2 \leq 3^7+1$ tels que les sous-blocs B_{i_1,i_2} et B_{i_1,i_2+d_2} sont coloriés de manière identique, et on a $i_2+2d_2 \leq 2\times3^7+1$.

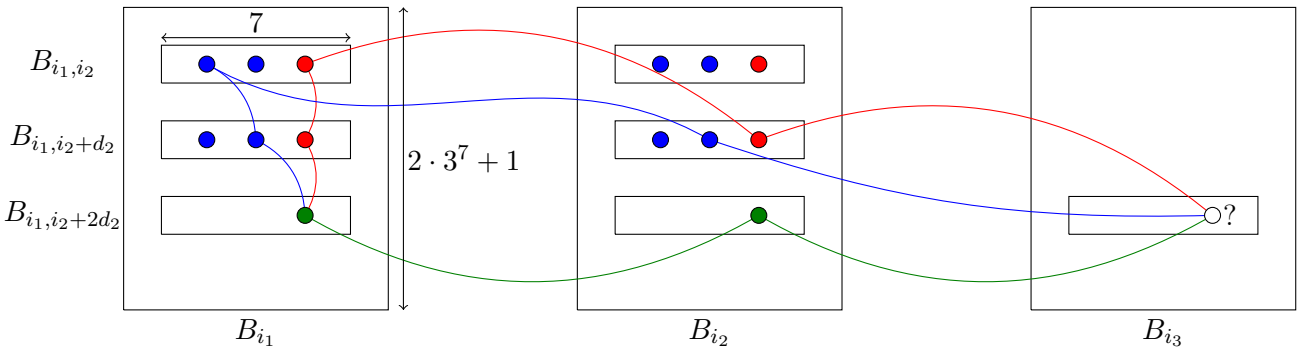


FIGURE 2 – Illustration de la preuve du théorème de Van der Waerden pour $k = r = 3$.

Parmi les 4 premiers éléments du sous-bloc B_{i_1,i_2}, il y en a 2 de la même couleur. Supposons qu'ils sont bleus, et sont en positions i_3 et $i_3 + d_3$ dans le sous-bloc. Si on a pas de progressions arithmétique monochrome de longueur 3, alors :

- l'élément en position $i_3 + 2d_3$ dans B_{i_1,i_2} (soit $(i_1 - 1)t + (i_2 - 1) \times 7 + (i_3 + 2d_3)$) n'est pas bleu, donc on suppose qu'il est rouge ;
- les éléments en position $i_3 + 2d_3$ des sous-blocs B_{i_1,i_2+d_2}, $B_{i_1+d_1,i_2}$ et $B_{i_1+d_1,i_2+d_2}$ sont aussi rouges ;
- on a les deux progressions suivantes de longueur 3 :

  $$\{(i_1-1)t+(i_2-1)\times 7+i_3, (i_1-1)t+(i_2+d_2-1)\times 7+i_3+d_3, (i_1-1)t+(i_2+2d_2-1)\times 7+i_3+2d_3\},$$

  $$\{(i_1-1)t+(i_2-1)\times 7+i_3+2d_3, (i_1-1)t+(i_2+d_2-1)\times 7+i_3+2d_3, (i_1-1)t+(i_2+2d_2-1)\times 7+i_3+2d_3\},$$

  de raisons respectives $7d_2 + d_3$ et $7d_2$. Par conséquent, le nombre $(i_1 - 1)t + (i_2 + 2d_2 - 1) \times 7 + i_3 + 2d_3$ ne peut être ni bleu ni rouge, donc il est vert. Le nombre $(i_1 + d_1 - 1)t + (i_2 + 2d_2 - 1) \times 7 + i_3 + 2d_3$ est donc aussi vert.

Mais alors, le nombre $(i_1 + 2d_1 - 1)t + (i_2 + 2d_2 - 1) \times 7 + (i_3 + 2d_3)$ n'a plus aucune couleur disponible, à cause des trois progressions arithmétiques suivantes :

$$\{(i_1-1)t+(i_2-1)7+i_3, (i_1+d_1-1)t+(i_2+d_2-1)7+i_3+d_3, (i_1+2d_1-1)t+(i_2+2d_2-1)7+i_3+2d_3\},$$

$$\{(i_1-1)t+(i_2-1)7+i_3+2d_3, (i_1+d_1-1)t+(i_2+d_2-1)7+i_3+2d_3, (i_1+2d_1-1)t+(i_2+2d_2-1)7+i_3+2d_3\},$$

$$\{(i_1-1)t+(i_2+2d_2-1)7+i_3+2d_3, (i_1+d_1-1)t+(i_2+2d_2-1)7+i_3+2d_3, (i_1+2d_1-1)t+(i_2+2d_2-1)7+i_3+2d_3\}.$$

Progressions de dimension supérieure. Une observation importante sur cet argument est que pour trouver notre progression arithmétique monochrome, on passe par un objet de "dimension 3", c'est-à-dire qu'on repère nos entiers par les trois coordonnées (i_1, i_2, i_3) (dans le cas $r = 2$, on avait seulement besoin de deux coordonnées). La première étape est donc de définir la structure de "dimension supérieure" que l'on va rechercher.

Définition 8.
- Soient $d_1, \dots, d_m \geq 1$. Une *progression arithmétique de dimension m, de longueur ℓ et de raison* $(d_1, \dots, d_m)$ est un ensemble de la forme

  $$\{a + d_1x_1 + \cdots + d_mx_m \,|\, 0 \leq x_1, \dots, x_m \leq \ell - 1\},$$

  où $a \geq 1$.
- De plus, pour $0 \leq i \leq m$, on définit le *i-ème bord* de cette progression comme l'ensemble des $a + d_1x_1 + \cdots + d_mx_m$ où $x_1 = \cdots = x_i = \ell$, et $0 \leq x_{i+1}, \dots, x_m \ \leq \ell - 1$.

Dans ce cadre, notre preuve dans le cas $(k, r) = (3, 2)$ peut s'interpréter de la manière suivante : on commence par trouver une progression bleue de longueur 2 et dimension 2. On montre ensuite que tous les points de son premier bord sont rouges, puis on obtient une contradiction en regardant le second bord. De même, pour $(k, r) = (3, 3)$, on trouve une progression de dimension 3 et longueur 2 telle que la progression est monochrome (bleue), de même que son premier bord (rouge) et son second (vert). Le troisième bord fournit alors une contradiction.

Dans le cas général, on va montrer l'énoncé suivant, qu'on notera $S(\ell, m)$:

"Pour tout $r \geq 2$, il existe $N(\ell, m, r)$ tel que, pour tout coloriage en r couleurs de $\{1, 2, \dots, N(\ell, m, r)\}$, il existe une progression arithmétique P de longueur ℓ et de dimension m telle que pour tout $0 \leq i \leq m$, le i-ème bord de P est monochrome."

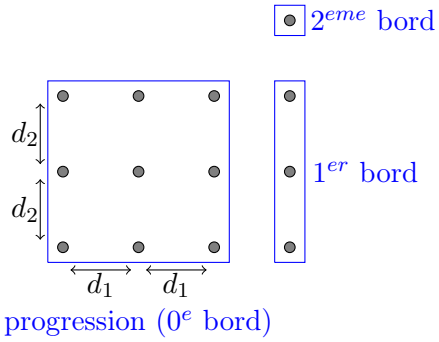


FIGURE 3 – Illustration d'une progression arithmétique de dimension 2 et de longueur 3. Le point de coordonnées (x_1, x_2) représente l'entier $a + d_1x_1 + d_2x_2$.

Exemple 9. Pour $\ell = 3$ et $m = 2$, cela revient à trouver $a, d_1, d_2 \geq 1$ tels que $a+3d_1+3d_2 \leq N$ et tels que :

- a, $a + d_1$, $a + 2d_1$, $a + d_2$, $a + d_1 + d_2$, $a + 2d_1 + d_2$, $a + 2d_2$, $a + d_1 + 2d_2$, $a + 2d_1 + 2d_2$ sont de la même couleur ;
- $a + 3d_1$, $a + 3d_1 + d_2$, $a + 3d_1 + 2d_2$ sont de la même couleur ;
- $\{a + 3d_1 + 3d_2\}$ est monochrome (ce dernier point est évident).

Le théorème de Van der Waerden étant le cas $m = 1$, il est suffisant de montrer que l'énoncé $S(\ell, m)$ est vrai pour tous $\ell \geq 2$ et $m \geq 1$. On va montrer cet énoncé par récurrence : comme les petits cas traités plus haut suggèrent de trouver une progression de longueur 2 et de grande dimension avant de trouver une progression de dimension 3, on va chercher à augmenter d'abord la dimension, puis la longueur.

Preuve de $S(\ell, m)$. On va montrer l'énoncé $S(\ell, m)$ par récurrence double. Plus précisément :

- $S(2, 1)$ est immédiat par le principe des tiroirs, avec $N(2, 1, r) = r + 1$;
- on va montrer que l'énoncé ($S(\ell, 1)$ et $S(\ell, m)$) implique $S(\ell, m + 1)$ (étape 1) ;
- on va montrer que l'énoncé ($\forall m \geq 1, S(\ell, m)$) implique $S(\ell + 1, 1)$ (étape 2).

Commençons par l'étape 1. On suppose que $S(\ell, 1)$ et $S(\ell, m)$ sont vrais, et on montre $S(\ell, m+1)$. Pour cela, on fixe $r \geq 2$. On pose $M = N(\ell, m, r)$ et $M' = N(\ell, 1, r^M)$, et on va montrer que $N(\ell, m + 1, r) \leq MM'$.

Pour cela, soit C un coloriage en r couleurs des entiers de 1 à MM'. On coupe $\{1, 2, \ldots, MM'\}$ en M' blocs de taille M, et on définit un coloriage C' de $\{1, 2, \ldots, M'\}$ en r^M couleurs comme suit :

$$C'(i_1) = C'(i_2) \quad \text{si et seulement si} \quad \forall j \in \{1, \ldots, M\}, C\left((i_1 - 1)M + j\right) = C\left((i_2 - 1)M + j\right).$$

Autrement dit, on voit chaque manière possible de colorier en r couleurs un bloc de longueur M comme le choix d'une "super-couleur" parmi r^M possibles. Par choix de M', il existe alors

$a', d' \geq 1$ tels que

$$C'(a') = C'(a'+d') = \cdots = C'(a'+(\ell-1)d').$$

Cela signifie que pour tout $1 \leq j \leq M$, on a

$$C((a'-1)M+j) = C((a'+d'-1)M+j) = \cdots = C((a'+(\ell-1)d'-1)M+j). \qquad (2)$$

De plus, par définition de M, le bloc $\{(a'-1)M+1, \ldots, a'M\}$ contient une progression arithmétique de longueur ℓ et de dimension m dont tous les bords sont monochromes. Soient $a \in \{(a'-1)M+1, \ldots, a'M\}$ son premier terme et $(d_1, \ldots, d_m)$ sa raison. On vérifie maintenant la conclusion de $S(\ell, m+1)$ avec la progression de premier terme a, et de raison $(d_1, \ldots, d_m, d'M)$. Pour cela, soit $0 \leq i \leq m+1$, et considérons le i-ème bord de cette progression :

- si $i = m+1$, alors le i-ème bord est un singleton, donc il est monochrome ;
- si $i \leq m$, alors les nombres de la forme

$$a + x_1 d_1 + \cdots + x_m d_m + x_{m+1} d'M$$

avec $x_1 = \cdots = x_i = \ell$ et $0 \leq x_{i+1}, \ldots, x_m \leq \ell - 1$ et $x_{m+1} = 0$ sont tous de la même couleur (il s'agit du i-ème bord de la progression de dimension m qu'on a trouvée dans le bloc). De plus, ces nombres restent de la même couleur si on fait varier x_{m+1} entre 0 et $\ell - 1$ d'après (2).

Passons maintenant à l'étape 2. On suppose que $S(\ell, m)$ et vraie pour tout m, et on fixe $r \geq 2$. On va montrer que $N(\ell+1, 1, r) \leq N(\ell, r, r)$, ce qui prouvera $S(\ell+1, 1)$. Cela revient à exiger une progression de longueur ℓ et de dimension le nombre de couleurs afin de passer à $\ell + 1$, comme on l'a déjà fait sur deux exemples.

Soit donc C un coloriage en r couleurs des entiers de 1 à $N(\ell, r, r)$. On sait qu'il existe une progression P de dimension r et de longueur ℓ dont tous les bords sont monochromes. Soient a son premier terme et $(d_1, \ldots, d_r)$ sa raison. Alors d'après le principe des tiroirs, il existe $0 \leq u < v \leq r$ tels que le u-ème et le v-ème bords de P sont de la même couleur. Mais alors, le nombre

$$a + \ell d_1 + \cdots + \ell d_u + x\,(d_{u+1} + \cdots + d_v)$$

est sur le u-ème bord de P pour $0 \leq x \leq \ell - 1$, et sur le v-ème bord de P pour $x = \ell$. Or, ces $\ell + 1$ nombres forment une progression arithmétique, donc on a bien trouvé une progression monochrome de longueur $\ell + 1$ et de dimension 1. Son premier (et seul) bord est un singleton, donc il est monochrome aussi. □

Étude quantitative de la borne obtenue. On rappelle que la notation $W(k, r)$ désigne les nombres de Van der Waerden, et que $N(\ell, m, r)$ désigne la borne fournie par l'énoncé $S(\ell, m)$ (i.e. jusqu'où aller pour trouver une progression de longueur ℓ et dimension m en partant d'un coloriage à r couleurs). Alors on a $W(k, r) \leq N(k, 1, r)$. De plus, en reprenant les étapes de notre preuve, on a :

- $N(2, 1, r) = r + 1$,
- $N(\ell, m+1, r) = N(\ell, m, r) \times N(\ell, 1, r^{N(\ell,m,r)})$,
- $N(\ell+1, 1, r) = N(\ell, r, r)$.

Regardons ce que cela donne pour de petites valeurs de ℓ et m. On a :

- $N(2, 2, r) = (r+1)N(2, 1, r^{r+1}) > r^r$,

- $N(2,3,r) \geq N(2,1,r^{N(2,2,r)}) > r^{r^r}$,
- $N(2,4,r) \geq N(2,1,r^{N(2,3,r)}) > r^{r^{r^r}}$

et ainsi de suite, soit

$$N(2,m,r) > \underbrace{r^{r^{\cdots^r}}}_{m \text{ fois}}.$$

On a donc

$$N(3,1,r) = N(2,r,r) > \underbrace{r^{r^{\cdots^r}}}_{r \text{ fois}},$$

ce qui donne en particulier l'ordre de grandeur de notre borne pour $W(3,r)$. On a ensuite

$$N(3,2,r) = N(3,1,r)N\left(3,1,r^{N(3,1,r)}\right) \geq N\left(3,1,\underbrace{r^{r^{\cdots^r}}}_{r \text{ fois}}\right) > \underbrace{r^{\cdots^r}}_{\underbrace{r^{\cdots^r}}_{r \text{ fois}} \text{ fois}},$$

puis

$$N(3,3,r) > \underbrace{r^{\cdots^r}}_{\substack{\underbrace{r^{\cdots^r}}_{\substack{\underbrace{r^{\cdots^r}}_{r \text{ fois}} \text{ fois}}} \text{ fois}}},$$

et plus généralement

$$N(3,m,r) > \underbrace{r^{\cdots^r}}_{\substack{\underbrace{r^{\cdots^r}}_{\substack{\underbrace{r^{\cdots^r}}_{\substack{\vdots \\ r \text{ fois}}} \text{ fois}}} \text{ fois}}}$$

avec m étage. En particulier, $N(4,1,r)$ est donné par la même expression mais avec r étages, ce qui donne la borne sur $W(4,r)$. Le nombre $N(4,2,r)$ s'écrit ensuite de la même manière mais nécessite deux colonnes, et $N(5,1,r)$ nécessite r colonnes. Ensuite, il devient trop compliqué d'essayer d'écrire les bornes explicitement.

Pour se donner un point de comparaison, on va définir une hiérarchie de croissance, c'est-à-dire une famille de fonctions croissant de plus en plus vite. On définit donc par récurrence la suite de fonctions (f_n) sur les entiers naturels de la manière suivante :

- $f_1(x) = 2x$,
- $f_{n+1}(x) = \underbrace{f_n \circ f_n \circ \cdots \circ f_n}_{x \text{ fois}}(1)$,
- $f_\infty(x) = f_x(x)$.

Alors on a $f_2(x) = 2^x$, puis

$$f_3(x) = \underbrace{2^{2^{\cdots^2}}}_{x \text{ fois}},$$

et

$$f_4(x) = \underbrace{2^{\cdot^{\cdot^{\cdot^{2}}}}}_{\underbrace{2^{\cdot^{\cdot^{\cdot^{2}}}}}_{\underbrace{2^{\cdot^{\cdot^{\cdot^{2}}}}}_{\vdots}\text{ fois}}\text{ fois}}\ 2 \text{ fois}}$$

avec x étages. On peut alors vérifier que la borne sur $W(k,2)$ fournie par la preuve ci-dessus est comprise entre $f_\infty(k-2)$ et $f_\infty(k)$. C'est le même ordre de grandeur que la fonction d'Ackermann. À titre de comparaison, on verra plus tard dans le cours une borne en $f_4(k)$, et la meilleure borne supérieure connue sur les nombres de Van der Waerden est la suivante, dûe à Gowers [11] :

$$W(k,r) \leq 2^{2^{r^{2^{2^{k+9}}}}}. \tag{3}$$

D'un autre côté, les meilleures bornes inférieures sont de type exponentiel (on verra une borne d'ordre r^k).

3 Théorème de Hales–Jewett

La preuve du théorème de Van der Waerden qu'on a vue suggère qu'il peut être intéressant de montrer des théorèmes de type Ramsey pour des structures de dimension plus grande. C'est le but du théorème de Hales–Jewett, qui montre en quelque sorte qu'un morpion en dimension suffisamment grande ne peut pas se terminer par un match nul.

Pour tous $n \geq 1$ et $t \geq 2$, on définit donc l'hypercube C_t^n comme l'ensemble des n-uplets $\mathbf{x} = (x_1, \ldots, x_n)$ avec $x_i \in \{1, \ldots, t\}$ pour tout $1 \leq i \leq n$.

Définition 10. Une *ligne* de C_t^n est une suite de t points $\mathbf{x}^1, \ldots, \mathbf{x}^t$ de C_t^n tels que :
- pour tout $1 \leq i \leq n$, on a soit $x_i^1 = x_i^2 = \cdots = x_i^t$, soit $x_i^j = j$ pour tout $1 \leq j \leq t$;
- il existe au moins un i pour lequel on est dans le second cas.

Remarque 11. Cette notion ne coïncide pas tout à fait avec la notion de droite en algèbre linéaire, car on autorise pas une coordonnée à décroître tandis qu'une autre croît. Par ailleurs, la seconde condition garantit que nos lignes ne sont pas réduites à un point.

Exemple 12. Dans C_3^2 (morpion classique), la diagonale $\{11, 22, 33\}$ est une ligne, mais pas $\{13, 22, 31\}$. Dans C_4^4, l'ensemble $\{2111, 2122, 2133, 2144\}$ est une ligne.

Théorème 4 (Hales–Jewett, 1963)**.** Soient $r \geq 2$ et $t \geq 1$. Il existe N tel que tout coloriage de C_t^N en r couleurs admet une ligne monochrome. On note $HJ(t,r)$ le plus petit N vérifiant cette propriété.

Remarque 13. Si on avait choisi de limiter la définition d'une ligne aux lignes parallèles aux axes (i.e. où $x_i^1 = \cdots = x_i^n$ pour toutes les coordonnées i sauf une), le théorème ne pourrait pas être vrai même pour $r = t = 2$, comme le montre le coloriage où $\mathbf{x}$ est colorié en bleu si $\sum_{i=1}^n x_i$ est pair et en rouge si cette somme est impaire.

Une des motivations du théorème de Hales–Jewett est qu'il implique celui de Van der Waerden, avec des bornes "proches".

Proposition 14. Le théorème de Hales–Jewett implique celui de Van der Waerden, avec

$$W(k,r) \leq k^{HJ(k,r)}.$$

Démonstration. Soit $N = HJ(k,r)$. À tout $0 \leq a \leq k^N - 1$, on associe ses chiffres $(a)_0, (a)_1, \ldots, (a)_{N-1}$ en base k, c'est-à-dire que

$$a = \sum_{i=0}^{N-1} (a)_i k^i,$$

avec $0 \leq (a)_i \leq k-1$ pour tout t. On note alors $\tilde{a} = ((a)_0 + 1, \ldots, (a)_{N-1} + 1) \in C_k^N$. Alors tout coloriage de $\{0, \ldots, k^N - 1\}$ induit un coloriage de C_k^N via l'application $a \to \tilde{a}$. D'après le théorème de Hales–Jewett, ce coloriage contient une ligne monochrome. Il existe donc $a^1, \ldots, a^k$ de la même couleur tels que pour tout $0 \leq i \leq N-1$, on ait

$$(a^1)_i = \cdots = (a^k)_i$$

ou

$$(a^1)_i = 0, \quad , (a^2)_i = 1, \quad \ldots, \quad (a^k)_i = k-1.$$

On note I l'ensemble des indices pour lesquels on est dans le second cas, de sorte que $I \neq \emptyset$. Alors pour tout j, on a

$$a^{j+1} - a^j = \left(\sum_{i=0}^{N-1} (a^{j+1})_i k^i \right) - \left(\sum_{i=0}^{N-1} (a^j)_i k^i \right) = \sum_{i \in I} k^i,$$

qui est non nul et ne dépend pas de j. Les nombres $a^1, \ldots, a^k$ forment donc une progression arithmétique monochrome de longueur k et de raison $\sum_{i \in I} k^i$. □

On va donner deux preuves du théorème de Hales–Jewett (la première ne sera qu'une esquisse et la seconde sera complète).

Première démonstration du théorème de Hales–Jewett. On adapte la preuve du théorème de Van der Waerden : dans le cube C_{t+1}^n, on définit une notion de "sous-espace de dimension k et de longueur t", et de bords de ce sous-espace. L'adaptation est très naturelle, puisque la définition des progressions arithmétiques de dimension supérieure et de leurs bords faisait déjà appel à des coordonnées. On montre ensuite par récurrence double (sur k, puis sur t) que si n est assez grand, alors tout coloriage de C_{t+1}^n admet des "sous-espaces" de dimension k dont tous les bords sont monochromes. La borne obtenue sur les nombres $HJ(t,r)$ par cette preuve est alors de type Ackermann, tout comme celle qu'on a déjà obtenu sur les nombres de Van der Waerden. □

Seconde démonstration du théorème de Hales–Jewett. L'idée grossière est la suivante : on raisonne à nouveau par récurrence sur t. On va d'abord chercher à construire un sous-espace de dimension assez grande à l'intérieur duquel les couleurs ne changent pas si une coordonnée égale à $t-1$ est remplacée par t. On utilise ensuite l'hypothèse de récurrence pour trouver une ligne de longueur $t-1$ dans ce sous-espace, puis l'hypothèse garantira qu'on peut prolonger cette ligne de 1. Un avantage majeur de cette preuve est que la récurrence se fera à r fixé, ce qui explique en partie qu'on obtienne une borne meilleure que dans la preuve précédente (en f_4 plutôt que f_∞).

On a donc besoin d'une notion de sous-espace de dimension plus grande que 1.

Définition 15. Soient $n_1, \ldots, n_k \geq 1$, et soit $n = n_1 + \cdots + n_k$. On identifie C_t^n avec $C_t^{n_1} \times C_t^{n_2} \times \ldots C_t^{n_k}$ en regroupant les coordonnées par blocs de longueurs $n_1, \ldots, n_k$. Un *sous-espace de dimension* k de C_t^n est un ensemble de sommets de la forme $L_1 \times L_2 \times \cdots \times L_k$, où L_i est une ligne de $C_t^{n_i}$ pour tout i.

.

Exemple 16. Pour $n = 10$, $t = 9$, $k = 2$, $n_1 = 4$ et $n_2 = 6$, l'ensemble

$$\{1a5abb9b42 \,|\, 1 \leq a, b \leq 9\}$$

est un sous-espace de dimension 2. En revanche, l'ensemble

$$\{1a5bab9a42 \,|\, 1 \leq a, b \leq 9\}$$

n'en est pas un, car on ne peut pas couper le mot $1a5bab9a42$ en deux avec les a d'un côté et les b de l'autre. Notons que la notion de sous-espace ne dépend pas seulement de n, t et k, mais aussi de la décomposition de n en les n_i.

Une remarque importante est que si $L_1 \times \cdots \times L_k$ est un sous-espace de dimension k, alors il existe une correspondance naturelle entre $L_1 \times \cdots \times L_k$ et le cube C_t^k. Sur notre exemple, cette correspondance s'écrit :

$$1a5abb9b42 \in L_1 \times L_2 \quad \leftrightarrow \quad (a, b) \in C_9^2.$$

On va rechercher des sous-espace particuliers, qui permettront de prolonger facilement des lignes de longueur $t-1$ en lignes de longueur t. On définit d'abord la notion qui nous intéresse sur un cube "complet" C_t^k.

Définition 17. Un coloriage de C_t^k est dit *invariant* si il a la propriété suivante : soient $1 \leq j \leq k$ et $\mathbf{x}, \mathbf{y}$ deux points de C_t^k tels que $x_j = t-1$, $y_j = t$ et $x_i = y_i$ pour tout $i \neq j$. Alors $\mathbf{x}$ et $\mathbf{y}$ ont la même couleur.

Exemple 18. Pour $k = 5$ et $t = 9$, un coloriage invariant doit donner la même couleur à 12838, 12839, 12938 et 12939. il n'y a pas d'autre contraintes sur ces points.

On étend maintenant cette notion à un sous-espace d'un gros cube, en utilisant la correspondance naturelle décrite plus tôt.

Définition 19. Soit $L_1 \times \ldots, \times L_k$ un sous-espace de dimension k de C_t^n, et $\varphi : L_1 \times \cdots \times L_k \to C_t^k$ l'isomorphisme naturel décrit plus haut. Un coloriage C de $L_1 \times \cdots \times L_k$ est dit *invariant* si le coloriage $C \circ \varphi$ de C_t^k est invariant.

Exemple 20. En reprenant l'exemple du sous-espace $\{1a5abb9b42\} \subset C_9^{10}$, un coloriage invariant de ce sous-espace doit attribuer la même couleur à 1858119142 et à 1959119142, mais aussi à 1858889842, 1858999942, 1959889842 et 1959999942.

Passons maintenant à la construction de sous-espaces invariants de dimension 1. Le résultat qui suit servira d'initialisation dans la récurrence qui nous permettra plus tard de construire des sous-espaces de grande dimension sur lesquels un coloriage est invariant.

Lemme 21. Si $n \geq r$, alors tout coloriage de C_t^n en r couleurs admet une ligne sur laquelle le coloriage est invariant.

Démonstration. On remarque qu'un coloriage est invariant sur une ligne si et seulement si il attribue la même couleur aux deux derniers points de cette ligne. Considérons donc les points de la forme

$$z_i = (t-1, \ldots, t-1, t, \ldots, t),$$

où les i premières coordonnées valent $t-1$ et les $n-i$ suivantes valent t. Alors on a $0 \leq i \leq n$, donc il y a $n+1 \geq r+1$ tels points. D'après le principe des tiroirs, il existe $i_1 < i_2$ tels que z_{i_1} et z_{i_2} sont de la même couleur. Ces points sont les deux derniers points de la ligne

$$\{t-1, \ldots, t-1, a, \ldots, a, t, \ldots, t \,|\, 1 \leq a \leq t\},$$

où les i_1 premières coordonnées sont des $t-1$, les $i_2 - i_1$ suivantes sont des a et les $n - i_2$ dernières sont des t. On a donc trouvé une ligne sur laquelle le coloriage est invariant. □

La proposition-clé est la suivante : elle permet de trouver des sous-espaces invariants de grande dimension, et permettra de faire tourner notre récurrence sur t.

Proposition 22 (Proposition-clé)**.** Soient $k \geq 1$ et $r, t \geq 2$. On définit les nombres $n_1, \ldots, n_k$ et $A_1, \ldots, A_k$ par

$$\begin{cases} n_1 = r^{t^{k-1}}, \\ A_i = t^{k-1-i} \times \prod_{j=1}^{i} t^{n_i}, \\ n_{i+1} = r^{A_i}. \end{cases}$$

Enfin, soit $n = n_1 + \cdots + n_k$. Alors pour tout coloriage χ de C_t^n en r couleurs, il existe un sous-espace de dimension k sur lequel χ est invariant.

Démonstration. On identifie C_t^n à $C_t^{n_1} \times \cdots \times C_t^{n_k}$. Si $\mathbf{y} \in C_t^n$, alors on l'écrit $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_k)$ avec $\mathbf{y}_i \in C_t^{n_i}$ pour tout $1 \leq i \leq k$. Étant donné un coloriage χ de C_t^n en r couleurs, on définit un coloriage $\chi^{(k)}$ de $C_t^{n_k}$ par $\chi^{(k)}(\mathbf{y}_k) = \chi^{(k)}(\mathbf{y}'_k)$ si et seulement si

$$\forall \mathbf{y}_1 \in C_t^{n_1}, \ldots, \forall \mathbf{y}_{k-1} \in C_t^{n_{k-1}}, \quad \chi(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}, \mathbf{y}_k) = \chi(\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}, \mathbf{y}'_k).$$

Estimons maintenant le nombre de couleurs du coloriage $\chi^{(k)}$. Il y a A_{k-1} manières de choisir des points $\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}$, donc il y a $r^{A_{k-1}}$ manières de choisir la couleur de $\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}, \mathbf{y}_k$ pour tous $\mathbf{y}_1, \ldots, \mathbf{y}_{k-1}$, donc $\chi^{(k)}$ a $r^{A_{k-1}} = n_k$ couleurs. D'après le Lemme 21, le coloriage $\chi^{(k)}$ de $C_t^{n_k}$ admet donc une ligne L_k sur laquelle il est invariant. L'idée va maintenant être de construire par récurrence une ligne L_{k-1} dans $C_t^{n_{k-1}}$, puis une ligne L_{k-2}, et ainsi de suite, et de prendre le produit de ces lignes à la fin.

Plus précisément, soit $1 \leq i \leq k-1$, et supposons qu'on ait construit des lignes $L_k \subset C_t^{n_k}, \ldots, L_{i+1} \subset C_t^{n_{i+1}}$. On définit alors un coloriage $\chi^{(i)}$ de $C_t^{n_i}$ par $\chi^{(i)}(\mathbf{y}_i) = \chi^{(i)}(\mathbf{y}'_i)$ si et seulement si

$$\forall \mathbf{y}_1 \in C_t^{n_1}, \ldots, \forall \mathbf{y}_{i-1} \in C_t^{n_{i-1}}, \quad \forall \mathbf{z}_{i+1} \in L_{i+1}, \ldots, \forall \mathbf{z}_k \in L_k,$$
$$\chi(\mathbf{y}_1, \ldots, \mathbf{y}_{i-1}, \mathbf{y}_i, \mathbf{z}_{i+1}, \ldots, \mathbf{z}_k) = \chi(\mathbf{y}_1, \ldots, \mathbf{y}_{i-1}, \mathbf{y}'_i, \mathbf{z}_{i+1}, \ldots, \mathbf{z}_k).$$

Alors pour tout $j < i$, il y a t^{n_j} manières de choisir $\mathbf{y}_j$, et pour tout $j > i$, il y a t manières de choisir $\mathbf{z}_j$. Il y a donc A_{i-1} manières de choisir tous les $\mathbf{y}_j$ et les $\mathbf{z}_j$, donc le coloriage $\chi^{(i)}$ de $C_t^{n_i}$ a $r^{A_{i-1}} = n_i$ couleurs. D'après le Lemme 21, il existe donc une ligne $L_i \subset C_t^{n_i}$ sur laquelle $\chi^{(i)}$ est invariant. Notons qu'il est crucial ici de s'être restreint à $\mathbf{z}_j \in L_j$ pour $j > i$ plutôt que d'écrire $\mathbf{z}_j \in C_t^{n_j}$: on évite ainsi une dépendance circulaire entre les différents n_i.

Il reste maintenant à vérifier que le coloriage de départ χ est bien invariant sur le sous-espace $L_1 \times \cdots \times L_k$ de dimension k. Soient donc $1 \leq i \leq k$, et $\mathbf{y}_i, \mathbf{y}_i'$ les deux derniers points de L_i. Il suffit de montrer que la couleur χ ne change pas quand on remplace $\mathbf{y}_i$ par $\mathbf{y}_i'$ dans un point de $L_1 \times \cdots \times L_k$. Soient donc $\mathbf{y}_j \in L_j$ pour tout $j \neq i$. On veut montrer que

$$\chi\left(\mathbf{y}_1, \ldots, \mathbf{y}_{i-1}, \mathbf{y}_i, \mathbf{y}_{i+1}, \ldots, \mathbf{y}_k\right) = \chi\left(\mathbf{y}_1, \ldots, \mathbf{y}_{i-1}, \mathbf{y}_i', \mathbf{y}_{i+1}, \ldots, \mathbf{y}_k\right).$$

Or, ceci est bien vrai par choix de L_i, puisqu'on a bien $\mathbf{y}_j \in C_t^{n_j}$ pour tout $j < i$, et $\mathbf{y}_j \in L_j$ pour tout $j > i$. □

Pour conclure la preuve, il suffit maintenant du lemme suivant, qui permettra de prolonger des lignes dans un sous-espace invariant.

Lemme 23. Soient $t \geq 3$ et $k = HJ(r, t-1)$. Alors tout coloriage invariant χ de C_t^k admet une ligne monochrome.

Démonstration. On a une inclusion évident $C_{t-1}^k \subset C_t^k$. Par définition de k, il existe une ligne monochrome $\left(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(t-1)}\right)$ dans C_{t-1}^k. Soit $\mathbf{x}^{(t)} \in C_t^k$ le point qui prolonge cette ligne (i.e. si $x_j^{(i)} = i$ pour tout i, alors $x_j^{(t)} = t$, et si $x_j^{(i)} = a$ pour tout i, alors $x_j^{(t)} = a$). Alors toutes les coordonnées qui diffèrent entre $\mathbf{x}^{(t-1)}$ et $\mathbf{x}^{(t)}$ valent $t-1$ dans $\mathbf{x}^{(t-1)}$ et t dans $\mathbf{x}^{(t)}$, donc comme χ est invariant, on a $\chi\left(\mathbf{x}^{(t-1)}\right) = \chi\left(\mathbf{x}^{(t)}\right)$, donc la ligne étendue $\left(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(t)}\right)$ est bien monochrome. □

Fin de la démonstration du théorème de Hales–Jewett. Comme annoncé, on raisonne par récurrence sur t, à r fixé. Notons que pour $t = 2$, une ligne monochrome est une ligne sur laquelle le coloriage est invariant, donc le Lemme 21 montre que $HJ(2, r) \leq r < +\infty$. De plus, soit $t \geq 3$ et supposons $HJ(t-1, r) < +\infty$. Soit n donné par la Proposition 22 pour $k = HJ(t-1, r)$. Soit χ un coloriage de C_t^n, et soit $S = L_1 \times \cdots \times L_k$ un sous-espace de C_t^n de dimension k sur lequel χ est invariant. Soit $\varphi : L_1 \times \cdots \times L_k \to C_t^k$ l'isomorphisme canonique. Par définition de k et par le Lemme 23, le coloriage $\chi \circ \varphi^{-1}$ de C_t^k admet une ligne monochrome. En appliquant φ^{-1}, on obtient une ligne monochrome dans S, et donc dans C_t^n. □

Aspect quantitatif. Dans la Proposition 22, on a

$$n_{i+1} = r^{A_i} \approx r^{t^{n_1 + \cdots + n_i}} \approx r^{t^{n_i}},$$

donc n_k est décrit par une tour d'exposants de hauteur $2k$. Ainsi, quand on fait notre récurrence sur t, on a

$$HJ(t, r) \lesssim n_{HJ(t-1,r)},$$

qui est une tour d'exposants de hauteur $HJ(t-1,r)$. On a donc une borne de la forme

$$HJ(t,r) \lesssim \underbrace{r^{\cdot^{\cdot^{\cdot^{r}}}}}_{\underbrace{r^{\cdot^{\cdot^{\cdot^{r}}}}}_{\underbrace{r^{\cdot^{\cdot^{\cdot^{r}}}}}_{\vdots}\ r \text{ fois}}\ r \text{ fois}}\ \text{fois}$$

où le membre de droite a t étages, ce qui donne une croissance en $f_4(t)$ dans la hiérarchie d'Ackermann.

4 La méthode probabiliste

On va maintenant passer à la recherche de bornes inférieures dans les théorèmes de type Ramsey. Puisqu'on veut construire de grands objets (par exemple de grands coloriages) avec "le moins de structure possible", une idée naturelle est de construire ces objets au hasard. Cette idée est le point de départ de la *méthode probabiliste* : pour montrer l'existence d'un objet vérifiant une certaine propriété P, on construit notre objet au hasard, et on montre que la propriété P est vérifiée avec probabilité strictement positive.

4.1 La borne de l'union

Borne inférieure "brutale" pour les nombres de Ramsey. Une des premières utilisations de la méthode probabiliste en combinatoire est le résultat suivant [6].

Théorème 5 (Erdös, 1947)**.** Soient $n, \ell \geq 2$. Si

$$\binom{n}{\ell} 2^{1-\binom{\ell}{2}} < 1,$$

alors $R(\ell,\ell) > n$. En particulier, on a

$$R(\ell,\ell) \geq (1+o(1))\frac{1}{e\sqrt{2}}\ell 2^{\ell/2}.$$

Démonstration. On colorie chaque arête du graphe complet K_n en bleu avec probabilité $\frac{1}{2}$ et en rouge avec probabilité $\frac{1}{2}$, indépendamment les unes des autres[3]. On suppose $\binom{n}{\ell}2^{1-\binom{\ell}{2}} < 1$, et on va montrer que la probabilité qu'il n'y ait aucun K_ℓ monochrome est strictement positive.

Pour tout ensemble A de ℓ sommets, soit E_A l'événement "les sommets de A forment un K_ℓ monochrome". Alors

$$\begin{aligned}\mathbb{P}(E_A) &= \mathbb{P}\left(\text{toutes les arêtes entre sommets de } A \text{ sont bleues}\right)\\ &\quad + \mathbb{P}\left(\text{toutes les arêtes entre sommets de } A \text{ sont rouges}\right)\\ &= \left(\frac{1}{2}\right)^{\binom{\ell}{2}} + \left(\frac{1}{2}\right)^{\binom{\ell}{2}}\\ &= 2^{1-\binom{\ell}{2}}.\end{aligned}$$

3. Cette précision est importante : quand on construit un modèle probabiliste, il est important que le modèle soit décrit complètement, par exemple en précisant comme ici que les couleurs sont indépendantes.

On a donc

$$\begin{aligned}\mathbb{P}\left(\text{il y a un } K_\ell \text{ monochrome}\right) &= \mathbb{P}\left(\bigcup_{\substack{A\subset K_n\\|A|=\ell}} E_A\right)\\ &\leq \sum_{\substack{A\subset K_n\\|A|=\ell}} \mathbb{P}\left(E_A\right)\\ &= \sum_{\substack{A\subset K_n\\|A|=\ell}} 2^{1-\binom{\ell}{2}} \qquad\qquad = \binom{n}{\ell} 2^{1-\binom{\ell}{2}} < 1,\end{aligned}$$

donc $\mathbb{P}\left(\text{il n'y a pas de } K_\ell \text{ monochrome}\right) > 0$. En particulier, il existe bien un coloriage sans K_ℓ monochrome.

Le deuxième point consiste simplement à faire l'analyse asymptotique de l'inégalité qu'on a trouvée. Notons que $\binom{n}{\ell} = \frac{n(n-1)\ldots(n-\ell+1)}{\ell!} \leq \frac{n^\ell}{\ell!}$. Par conséquent, si $\frac{n^\ell}{\ell!} 2^{1-\binom{\ell}{2}} < 1$, alors $R(\ell,\ell) > n$. Cette condition équivaut à

$$n < \left(\ell!\, 2^{\binom{\ell}{2}-1}\right)^{1/\ell}.$$

On note $n_0(\ell)$ le membre de droite de cette dernière inégalité. En utilisant la formule de Stirling, on a

$$\begin{aligned}n_0(\ell) &= \left((1+o(1))\ell^\ell e^{-\ell}\sqrt{2\pi\ell} 2^{\frac{\ell(\ell-1)}{2}} \times \frac{1}{2}\right)^{1/\ell}\\ &= (1+o(1))\frac{\ell}{e} 2^{\frac{\ell-1}{2}} \left(\sqrt{\frac{\pi\ell}{2}}\right)^{1/\ell}\\ &= (1+o(1))\frac{\ell}{e\sqrt{2}} 2^{\ell/2}.\end{aligned}$$

□

Notons que quitte à diminuer de 1 la valeur de $n_0(\ell)$, notre calcul montre en fait un résultat bien plus fort : non seulement il existe des coloriages de K_n sans K_ℓ monochrome, mais en fait la grande majorité des coloriages de K_n n'admet pas de K_ℓ monochrome ! De manière un peu étonnante, cela ne signifie pas pour autant que de tels coloriages sont faciles à construire explicitement. Dans certaines applications de la méthode probabiliste, il est cependant possible de "dérandomiser" la preuve pour en tirer des algorithmes de construction déterministes, voir par exemple le chapitre 15 de [1].

Borne inférieure pour les nombres de Van der Waerden. Essayons maintenant d'appliquer cette même idée aux nombres de Van der Waerden. On obtient alors le résultat suivant (qui sera significativement amélioré un peu plus loin dans le cours).

Théorème 6. Soient $k, r \geq 2$. Alors on a

$$W(k,r) \geq \sqrt{k-1} \times r^{\frac{k-1}{2}}.$$

Démonstration. Soit $n \geq 1$ un entier. Comme précédemment, on choisit notre coloriage uniformément au hasard, i.e. pour tout i dans $\{1, 2, \ldots, n\}$, on choisit la couleur de i uniformément parmi les r possibles, et ce de manière indépendante. Pour toute progression arithmétique P de longueur k, notons E_P l'événement $\{P \text{ est monochrome}\}$. Alors on a

$$\mathbb{P}\ (E_P) = \sum_{i=1}^{r} \mathbb{P}\ (P \text{ est monochrome de couleur } i) = r \times \frac{1}{r^k} = \frac{1}{r^{k-1}}.$$

D'autre part, bornons le nombre de progressions arithmétiques de longueur k dont les termes sont dans $\{1, \ldots, n\}$. Une telle progression a sa raison bornée comprise entre 1 et $\frac{n}{k-1}$, et son premier terme entre 1 et n. Il y a donc au plus $\frac{n^2}{k-1}$ telles progression [4]. Comme précédemment, on peut maintenant effectuer une borne de l'union :

$$\begin{aligned}
\mathbb{P}\left(\text{il existe une progression monochrome de longueur } k\right) &= \mathbb{P}\left(\bigcup_{\substack{P \subset \{1,\ldots,n\} \\ P \text{ progression}}} E_p\right) \\
&\leq \sum_{\substack{P \subset \{1,\ldots,n\} \\ P \text{ progression}}} \mathbb{P}\left(E_P\right) \\
&\leq \frac{n^2}{k-1} \frac{1}{r^{k-1}}.
\end{aligned}$$

En particulier, pour $n < \sqrt{k-1} \times r^{\frac{k-1}{2}}$, cette probabilité est strictement plus petite que 1, donc il existe un coloriage de $\{1, \ldots, n\}$ à r couleurs sans progression monochrome de longueur k, ce qui montre le résultat. □

4.2 Altération

Les deux arguments probabilistes qu'on vient de voir sont assez brutaux pour deux raisons : d'une part, la construction aléatoire choisie est la plus évidente possible, et d'autre part, elle est analysée en utilisant la borne de l'union, qui peut parfois être brutale. Ceci suggère que des améliorations peuvent être apportées soit en trouvant des constructions plus astucieuses, soit en estimant les probabilités de réussite de manière plus fine.

Dans un premier temps, on se concentre sur la première idée. L'idée de la méthode d'altération est la suivante : plutôt que de chercher une construction aléatoire qui fonctionne directement, on va chercher à montrer qu'une construction aléatoire marche "presque" avec grande probabilité, puis régler "à la main" les derniers détails, cette fois sans probabilités. Autrement dit, puisque l'inconvénient de la méthode probabiliste est que "tout ce qui peut arriver finira par arriver quelque part", on ne fait pas confiance au hasard jusqu'au bout et on passe derrière pour corriger ses erreurs. Cette idée mène à l'amélioration suivante de la borne inférieure sur les nombres de Ramsey.

Théorème 7. Soient $\ell, n \geq 2$. Alors

$$R(\ell, \ell) \geq n - \binom{n}{\ell} 2^{1-\binom{\ell}{2}}. \tag{4}$$

4. C'est une borne brutale : le vrai nombre est plus proche de $\frac{n^2}{2(k-1)}$. Un comptage plus précis améliorerait le résultat final d'un facteur $\sqrt{2}$

En particulier, on a

$$R(\ell,\ell) \geq (1+o(1))\frac{1}{e}\ell 2^{\ell/2}. \tag{5}$$

Démonstration. Dans un premier temps, on colorie chaque arête de K_n en bleu (resp. rouge) avec probabilité $\frac{1}{2}$, de manière indépendante, et on note C le coloriage aléatoire obtenu. On note X le nombre de copies monochromes de K_ℓ pour le coloriage C. On va estimer l'espérance de X en la décomposant comme une somme d'indicatrices. Pour cela, pour tout ensemble A de ℓ sommets, on pose

$$Y_A = \begin{cases} 1 & \text{si les arêtes entre sommets de } A \text{ sont toutes de la même couleur,} \\ 0 & \text{sinon.} \end{cases}$$

Alors on a

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}\left[\sum_{|A|=\ell} Y_A\right] \\ &= \sum_{|A|=\ell} \mathbb{E}[Y_A] \\ &= \sum_{|A|=\ell} 2^{1-\binom{\ell}{2}} \\ &= \binom{n}{\ell} 2^{1-\binom{\ell}{2}}. \end{aligned}$$

En particulier, il existe un coloriage C qui admet $m \leq \binom{n}{\ell}2^{1-\binom{\ell}{2}}$ copies monochromes de K_ℓ, qu'on va noter $A_1,\dots,A_m$. On choisit des sommets $x_1 \in A_1,\dots,x_m \in A_m$. Alors le coloriage C restreint à $K_n\backslash\{x_1,\dots,x_m\}$ est un coloriage de K_{n-m} (ou d'un graphe complet encore plus gros, si certains x_i sont confondus) sans copie monochrome de K_ℓ, d'où

$$R(\ell,\ell) > n-m \geq n - \binom{n}{\ell}2^{1-\binom{\ell}{2}}.$$

Il reste à montrer (5) à partir de (4). Pour cela, on prend $n = \lfloor \frac{1}{e}\ell 2^{\ell/2} \rfloor$. On a alors

$$\binom{n}{\ell}2^{1-\binom{\ell}{2}} \leq \frac{n^\ell}{\ell!}\frac{2}{2^{\frac{\ell(\ell-1)}{2}}} \leq \frac{\ell^\ell 2^{\ell^2/2}}{e^\ell \ell!}\frac{2}{2^{\frac{\ell(\ell-1)}{2}}} \underset{n\to+\infty}{\sim} \frac{2^{\ell/2+1}}{\sqrt{2\pi\ell}} = o(n)$$

d'après la formule de Stirling. On peut donc écrire

$$R(\ell,\ell) \geq n - \binom{n}{\ell}2^{1-\binom{\ell}{2}} \underset{n\to+\infty}{\sim} n = (1+o(1))\frac{1}{e}\ell 2^{\ell/2}.$$

□

Remarque 24.

- On peut voir cette méthode comme une généralisation de la méthode probabiliste "basique" : on a pu optimiser l'équation (4) pour prendre la meilleure valeur de n possible, tandis que la méthode basique consistait à se limiter aux n pour lesquels $\mathbb{E}[X] < 1$.

- Cependant, on ne gagne qu'un facteur $\sqrt{2}$, ce qui n'est pas très spectaculaire. Par ailleurs, la même approche pour les nombres de Van der Waerden semble plus compliquée : supprimer m éléments de $\{1, \dots, n\}$ ne donne pas un ensemble isomorphe à $\{1, \dots, n-m\}$, et une altération qui changerait certaines couleurs risquerait de créer de nouvelles progression monochromes.
- Enfin, nous renvoyons à [1, Chapitre 3] pour quelques exemples plus fructueux d'utilisations de l'altération.

5 Lemme local de Lovász

Comme évoqué ci-dessus, une deuxième piste d'amélioration de la méthode probabiliste "basique" est d'estimer de manière plus fine la probabilité d'obtenir une configuration satisfaisante. Plus précisément, la borne de l'union $\mathbb{P}\left(\bigcup_i E_i\right) \leq \sum_i \mathbb{P}(E_i)$ semble trop brutale quand les événements E_i sont indépendants ou presque. On voudrait alors pouvoir écrire

$$\mathbb{P}\left(\text{aucun des } E_i \text{ ne se produit}\right) \geq \prod_i (1 - \mathbb{P}(E_i)) > 0,$$

mais c'est bien sûr trop demander, car les événements E_i ne sont jamais complètement indépendants. Le lemme de Lovász est en quelque sorte un intermédiaire entre la borne de l'union (universelle, mais brutale) et l'indépendance (condition trop forte pour les exemples qui nous intéresse, mais qui donne toujours une probabilité strictement positive). Il indique que si des événements sont "presque indépendants" et ont chacun une probabilité suffisamment faible, alors la probabilité qu'aucun ne se produise n'est pas trop petite.

En particulier, sur les exemples qu'on a vus, la borne de l'union ne fonctionnait que quand les configurations souhaitées étaient très majoritaires. À l'inverse, le lemme local de Lovász est plus sensible et permet de "détecter" des événements de faible probabilité. Commençons par préciser ce qu'on entend par "presque indépendants".

Définition 25. Soient $A_1, \dots, A_k, B$ des événements. On dit que B est *indépendant* de la famille $(A_i)_{1 \leq i \leq k}$ si pour tous $S_1, S_2 \subset \{1, \dots, k\}$, l'événement B est indépendant de

$$\left(\bigcap_{i \in S_1} A_i\right) \cap \left(\bigcap_{i \in S_2} A_i^c\right),$$

où A_i^c désigne le complémentaire de A_i.

Remarque 26. Cette définition est (strictement !) plus forte que de demander que B soit indépendant de chacun des A_i. En revanche, elle est (strictement) plus faible que de demander que les événements $A_1, \dots, A_k, B$ soient indépendants.

Exemple 27. Considérons un coloriage de $\{1, \dots, n\}$ où les couleurs des éléments sont choisies de manière indépendante. Soient (E_i) et F des sous-ensembles de $\{1, \dots, n\}$, et considérons les événements

$$A_i = \{\text{l'ensemble } E_i \text{ est monochrome}\},$$

$$B = \{\text{l'ensemble } F \text{ est monochrome}\}.$$

Si $F \cap (\bigcup_i E_i) = \emptyset$, alors B est indépendant de la famille (A_i). En effet, B ne dépend que des couleurs des éléments de F, et n'importe quel événement construit à partir des A_i ne dépend que des couleurs des éléments de $\bigcup_i E_i$, donc est indépendant de B.

Définition 28. Soient $A_1, \dots, A_n$ des événements, et soit G un graphe orienté (i.e. on distingue l'arête (i,j) de l'arête (j,i)), dont les sommets sont étiquetés de 1 à n. On dit que G est un *graphe de dépendance* pour $A_1, \dots, A_n$ si pour tout $i \in \{1, \dots, n\}$, l'événement A_i est indépendant de la famille $(A_j)_{(i,j) \notin G}$.

Autrement dit, on demande que A_i soit indépendant de tout événement construit à partir d'événements qui ne sont pas reliés à lui. Insistons sur le fait que pour vérifier que G est un graphe de dépendance, il ne suffit pas de vérifier deux à deux l'indépendance de A_i et A_j quand i n'est pas relié à j !

On peut maintenant énoncer le lemme local de Lovász, montré par Erdös et Lovász en 1975. On en donnera deux formulations : une très générale en premier, puis une plus pratique à utiliser.

Théorème 8 (Lemme local de Lovász général, [7])**.** Soient $A_1, \dots, A_n$ des événements et G un graphe de dépendance de $(A_i)_{1 \leq i \leq n}$. On suppose qu'il existe des nombres $x_1, \dots, x_n$ dans $[0,1[$ tels que pour tout i, on ait :

$$\mathbb{P}(A_i) \leq x_i \prod_{(i,j) \in G} (1 - x_j). \tag{6}$$

Alors $\mathbb{P}\left(\bigcap_{i=1}^n A_i^c\right) \geq \prod_{i=1}^n (1 - x_i) > 0$.

Pour interpréter ce résultat, on peut voir x_i comme une version "augmentée" de la probabilité $\mathbb{P}(A_i)$. Plus il y a de dépendances dans le graphe, plus le produit à droite de (6) est petit, et donc plus x_i doit être grand comparé à $\mathbb{P}(A_i)$. Notons aussi que le cas où les événements A_i sont indépendants correspond au cas où G n'a aucune arête. Le lemme est alors immédiat en prenant $x_i = \mathbb{P}(A_i)$.

Démonstration. On va d'abord montrer le résultat intermédiaire suivant : sous les hypothèses du lemme local de Lovász, pour tous $S \subset \{1, \dots, n\}$ et $i \notin S$, on a

$$\mathbb{P}\left(A_i \middle| \bigcap_{j \in S} A_j^c\right) \leq x_i. \tag{7}$$

À nouveau, cette inégalité semble raisonnable si on voit x_i comme une "version augmentée" de $\mathbb{P}(A_i)$. On va la montrer par récurrence sur $s = |S|$, avec $0 \leq s \leq n-1$. Pour $s = 0$, il n'y a pas de conditionnement, et on a bien

$$\mathbb{P}(A_i) \leq x_i \prod_{(i,j) \in G} (1 - x_j) \leq x_i.$$

.

Soit maintenant $1 \leq s \leq n-1$, et supposons que (7) est vérifiée pour tout $s' < s$. Afin d'utiliser au mieux l'hypothèse de l'énoncé, on sépare S en deux en définissant

$$S_1 = \{j \in S | (i,j) \in G\} \quad \text{et} \quad S_2 = \{j \in S | (i,j) \notin G\}.$$

On peut alors écrire

$$\mathbb{P}\left(A_i \Big| \bigcap_{j\in S} A_j^c\right) = \frac{\mathbb{P}\left(A_i \cap \left(\bigcap_{j\in S_1} A_j^c\right) \cap \left(\bigcap_{j\in S_2} A_j^c\right)\right)}{\mathbb{P}\left(\left(\bigcap_{j\in S_1} A_j^c\right) \cap \left(\bigcap_{j\in S_2} A_j^c\right)\right)}$$
$$= \frac{\mathbb{P}\left(A_i \cap \left(\bigcap_{j\in S_1} A_j^c\right) \Big| \bigcap_{j\in S_2} A_j^c\right)}{\mathbb{P}\left(\bigcap_{j\in S_1} A_j^c \Big| \bigcap_{j\in S_2} A_j^c\right)}, \tag{8}$$

en divisant le numérateur et le dénominateur chacun par $\mathbb{P}\left(\bigcap_{j\in S_2} A_j^c\right)$. Comme A_i est indépendant de $(A_j)_{j\in S_2}$, on peut borner le numérateur par

$$\mathbb{P}\left(A_i \Big| \bigcap_{j\in S_2} A_j^c\right) = \mathbb{P}(A_i) \leq x_i \prod_{(i,j)\in G} (1-x_j).$$

Il reste à minorer le dénominateur de (8). Pour cela, on écrit $S_1 = \{j_1, \ldots, j_r\}$. Si $r = 0$, alors le dénominateur vaut 1 et (7) est immédiat. Sinon, pour intégrer $j_1, \ldots, j_r$ "un par un", on écrit

$$\mathbb{P}\left(A_{j_1}^c \cap A_{j_2}^c \cap \cdots \cap A_{j_r}^c \Big| \bigcap_{j\in S_2} A_j^c\right) = \prod_{k=1}^{r} \left(1 - \mathbb{P}\left(A_{j_k} \Big| A_{j_1}^c \cap \cdots \cap A_{j_{k-1}}^c \cap \bigcap_{j\in S_2} A_j^c\right)\right).$$

Dans chaque facteur du membre de droite, le nombre de A_i qui apparaissent dans le conditionnement est strictement plus petit que s (car $j_k \in S$ n'y apparaît pas). On peut donc appliquer l'hypothèse de récurrence dans chaque facteur. On trouve

$$\mathbb{P}\left(A_{j_1}^c \cap A_{j_2}^c \cap \cdots \cap A_{j_r}^c \Big| \bigcap_{j\in S_2} A_j^c\right) \geq \prod_{k=1}^{r} (1-x_{j_k}) \geq \prod_{(i,j)\in G} (1-x_j),$$

par définition de S_1. En combinant nos estimées sur le numérateur et le dénominateur de (8), on obtient

$$\mathbb{P}\left(A_i \Big| \bigcap_{j\in S} A_j^c\right) \leq \frac{x_i \prod_{(i,j)\in G}(1-x_j)}{\prod_{(i,j)\in G}(1-x_j)} = x_i,$$

ce qui montre (7) par récurrence.

On peut maintenant finir la preuve du lemme de Lovász. À nouveau, l'idée est d'ajouter les événements A_i un par un, en utilisant (7) à chaque étape :

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i^c\right) = \prod_{i=1}^{n} \left(1 - \mathbb{P}\left(A_i | A_1^c \cap \cdots \cap A_{i-1}^c\right)\right) \geq \prod_{i=1}^{n} (1-x_i) > 0.$$

□

Le lemme local de Lovász symétrique. Cette version du lemme de Lovász est certes très générale, mais elle est difficilement maniable : elle nécessite dans chaque situation de choisir les x_i de manière adéquate, ce qui peut s'avérer technique. Heureusement, dans le cas où les degrés du graphe de dépendance ne sont pas trop élevés, il existe un choix naturel des x_i, qui donne le résultat suivant.

Théorème 9 (Lemme local de Lovász symétrique)**.** Soient $p \in [0,1]$ et $d \geq 0$, et soient $A_1, \ldots, A_n$ des événements. On suppose que chaque A_i est indépendant d'une famille d'événements qui contient tous les autres A_j, sauf au plus d. On suppose aussi que $\mathbb{P}(A_i) \leq p$ pour tout $1 \leq i \leq n$. Si $ep(d+1) \leq 1$, alors

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i^c\right) > 0.$$

Une caractéristique remarquable de cet énoncé est que la condition énoncée ne dépend pas de n, mais seulement de p et de d. À titre de comparaison, la borne de l'union ne nous permettrait de conclure qu'à condition que $pn < 1$. Le lemme de Lovász est donc beaucoup plus puissant dès que d est petit devant n, c'est-à-dire que tout événement est indépendant de la plupart des autres.

Démonstration. L'hypothèse d'indépendance signifie qu'il existe un graphe de dépendance G des événements A_i tel que de tout i sont issues au plus d arêtes. Si $d = 0$, les A_i sont indépendants et la conclusion est immédiate. Sinon, on choisit $x_i = \frac{1}{d+1} < 1$ dans la version générale du lemme de Lovász. Vérifions que les hypothèses sont bien satisfaites. Pour tout $1 \leq i \leq n$, on a

$$x_i \prod_{(i,j)\in G} (1 - x_j) = \frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^{|\{j|(i,j)\in G\}|} \geq \frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^{d}.$$

L'inégalité $1 - \frac{1}{d+1} \leq \exp\left(-\frac{1}{d+1}\right)$ est alors dans le mauvais sens, mais on peut écrire

$$\left(1 - \frac{1}{d+1}\right)^{d} = \left(1 + \frac{1}{d}\right)^{-d} \geq \left(e^{1/d}\right)^{-d} \geq e^{-1},$$

d'où

$$x_i \prod_{(i,j)\in G} (1 - x_j) \geq \frac{e^{-1}}{d+1} \geq p \geq \mathbb{P}(A_i)$$

par hypothèse. Le lemme général s'applique donc bien, d'où

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i^c\right) \geq (1-p)^n > 0.$$

□

Application aux nombres de Ramsey. On va maintenant appliquer le lemme de Lovász à l'obtention d'une borne inférieure sur les nombres de Ramsey.

Théorème 10 (Spencer, [18])**.** Soient $\ell, n \geq 2$. Si

$$e\binom{\ell}{2}\binom{n-2}{\ell-2}2^{1-\binom{\ell}{2}} < 1,$$

alors $R(\ell,\ell) > n$. En particulier, on a

$$R(\ell,\ell) \geq (1+o(1))\frac{\sqrt{2}}{e}\ell 2^{\ell/2}.$$

Cette borne peut sembler un peu décevante : on ne gagne qu'un facteur $\sqrt{2}$ par rapport à l'altération, et qu'un facteur 2 par rapport à la méthode probabiliste "basique". Cependant, bien qu'elle date d'environ 50 ans, il s'agit de la meilleure borne inférieure connue sur les nombres de Ramsey symétriques !

Démonstration. Comme dans la méthode probabiliste "habituelle", on colorie les arêtes de K_n en bleu avec probabilité $\frac{1}{2}$ et en rouge avec probabilité $\frac{1}{2}$, et ce de manière indépendante. Pour tout ensemble S de ℓ sommets de K_n, on notera A_S l'événement

$$\{\text{les sommets de } S \text{ forment un } K_\ell \text{ monochrome}\}.$$

Alors pour tout S, on peut écrire $\mathbb{P}(A_S) = 2^{1-\binom{\ell}{2}} = p$. Il reste à trouver un d qui marche. Pour cela, on note que :

- A_S ne dépend que des couleurs des arêtes $\{x, y\}$ avec $x, y \in S$;
- si T est un autre ensemble de ℓ sommets vérifiant $|S \cap T| \leq 1$, alors l'événement A_T ne dépend que des couleurs des arêtes $\{x, y\}$, où $x \notin S$ ou $y \notin S$.

Par conséquent, pour tout S, l'événement A_S est indépendant de la famille $(A_T)_{|S\cap T|\leq 1}$. Cette famille contient tous les A_T, sauf ceux pour lesquels $|S \cap T| \geq 2$. Or, il y en a moins de

$$\binom{\ell}{2}\binom{n-2}{\ell-2}.$$

En effet, pour choisir un ensemble T de ℓ éléments tel que $|S \cap T| \geq 2$, il faut d'abord choisir deux éléments de S qui seront dans l'intersection, puis $\ell - 2$ autres éléments parmi les $n - 2$ restants. De plus, l'inégalité est stricte, car on ne compte pas l'ensemble S lui-même. On peut donc appliquer le lemme local de Lovász symétrique avec $d = \binom{\ell}{2}\binom{n-2}{\ell-2} - 1$. On a alors

$$ep(d+1) = e\binom{\ell}{2}\binom{n-2}{\ell-2}2^{1-\binom{\ell}{2}} \leq 1$$

par hypothèse, donc le LLL symétrique donne

$$\mathbb{P}\left(\bigcap_{\substack{S\subset K_n\\|S|=\ell}} A_S^c\right) > 0.$$

Il existe donc un K_ℓ monochrome, donc $R(\ell, \ell) > n$.

Il reste maintenant à traiter cette condition de manière asymptotique. Comme $\binom{n}{\ell} \leq \frac{n^\ell}{\ell!}$, il est suffisant d'avoir

$$e\binom{\ell}{2}\frac{n^{\ell-2}}{(\ell-2)!}2^{1-\binom{\ell}{2}} \leq 1.$$

Ceci équivaut à $n \leq \left(\frac{(\ell-2)!2^{\binom{\ell}{2}}}{2e\binom{\ell}{2}}\right)^{\frac{1}{\ell-2}}$. En utilisant la formule de Stirling, le membre de droite équivaut à

$$\left(\frac{(\ell-2)^{\ell-2}e^{-(\ell-2)}\sqrt{2\pi\ell}2^{\frac{\ell(\ell-1)}{2}}}{e\ell(\ell-1)}\right)^{\frac{1}{\ell-2}} \underset{\ell\to+\infty}{\sim} \frac{\ell}{e}2^{\frac{\ell(\ell-1)}{2(\ell-2)}}$$
$$= \frac{\ell}{e}2^{\frac{(\ell-2)(\ell+1)+2}{2(\ell-2)}}$$
$$\underset{\ell\to+\infty}{\sim} \frac{\sqrt{2}}{e}\ell 2^{\ell/2}.$$

□

Remarque 29. Une raison qui peut expliquer la faible amélioration par rapport à la méthode "basique" est que le degré d qui mesure le "niveau de dépendance" des A_i reste assez élevé : il est plus grand que $\binom{n-2}{\ell-2} \approx n^{\ell-2}$. L'amélioration n'est pas spectaculaire par rapport au maximum possible $\binom{n}{\ell} \approx n^{\ell}$, qui correspond à la borne de l'union. Une autre manière de voir que les A_i sont "assez peu indépendants" est que la simple présence d'un $K_{2\ell}$ monochrome va déclencher la présence d'un grand nombre de K_ℓ monochromes. Il existe donc des événements qui peuvent faire se produire simultanément beaucoup des événements A_S, chose qui serait bien plus dure avec des événements indépendants.

Application aux nombres de Van der Waerden. On rappelle que $W(k,r)$ est le plus petit n tel que tout coloriage des entiers de 1 à n en r couleurs admet une progression arithmétique monochrome de longueur k.

Théorème 11. Pour tous $k, r \geq 2$, on a

$$W(k,r) \geq \frac{1}{e}\frac{k-1}{k^2} r^{k-1}.$$

Démonstration. Soit $n \geq 1$. On attribue aux entiers de 1 à n des couleurs indépendantes, uniformes dans $\{1,\dots,r\}$. Pour toute progression arithmétique $P \subset \{1,\dots,n\}$ de longueur k, pn note A_P l'événement

$$\{P \text{ est monochrome}\}.$$

Alors $\mathbb{P}(A_P) = \frac{1}{r^{k-1}}$ pour toute P, donc on peut prendre $p = \frac{1}{r^{k-1}}$ dans le LLL. De plus, pour tout P, l'événement A_P est indépendant de la famille $(A_Q)_{P\cap Q=\emptyset}$. On doit donc majorer le nombre de progressions Q telles que $P \cap Q \neq \emptyset$. Pour cela, si $P \cap Q \neq \emptyset$, alors il existe des indices $1 \leq i,j \leq k$ tels que le i-ème terme de P coïncide avec le j-ème terme de Q. Si i et j sont fixés, la progression Q est alors déterminée par sa raison b qui vérifie $1 \leq b \leq \frac{n}{k-1}$. On en conclut que $(A_Q)_{P\cap Q=\emptyset}$ contient tous les A_Q, sauf au plus $k^2\frac{n}{k-1}$ (l'inégalité est à nouveau stricte, car on ne compte pas A_P). On peut donc appliquer le LLL avec $d = k^2\frac{n}{k-1} - 1$: on a $W(k,r) > n$ si

$$e \times \frac{1}{r^{k-1}} \times \frac{k^2 n}{k-1} < 1,$$

soit $n < \frac{k-1}{ek^2} r^{k-1}$. □

Remarque 30.

- Dans la preuve ci-dessus, les événements A_P et A_Q restent indépendants si $|P \cap Q| = 1$. Cependant, l'événement A_P n'est pas indépendant de la famille $(A_Q)_{|P\cap Q|\leq 1}$. On peut par exemple le vérifier dans le cas $k = 2$, $n = 3$ et $r = 2$, avec les progressions $P = \{1,2\}$, $Q_1 = \{1,3\}$ et $Q_2 = \{2,3\}$: l'événement A_P n'est pas indépendant de la famille (A_{Q_1}, A_{Q_2}), puisque si ni A_{Q_1} ni A_{Q_2} ne se produit, alors A_P se produit forcément.
- On rappelle que la borne obtenue par la méthode probabiliste basique était plutôt d'ordre $r^{\frac{k-1}{2}}$. L'amélioration ici est donc très nette. La différence avec les nombres de Ramsey est que les dépendances sont ici plus faibles, car il est très difficile pour deux progressions arithmétiques d'avoir une grande intersection.

- La borne inférieure donnée n'est pas la meilleure connue, mais elle s'en approche. Dans le cas $r = 2$, la meilleure borne connue à ce jour est dûe à Szabó [19] et est de la forme $W(k, 2) \geq \frac{2^k}{k^{o(1)}}$. Elle n'améliore donc notre résultat que d'un facteur polynomial, et l'approche utilise également le lemme local de Lovász. Notons que l'écart entre les meilleures bornes supérieures et inférieures connues reste gigantesque (comparer avec 3.)

6 Constructions arithmétiques

On va conclure ce cours par un autre type de constructions permettant d'obtenir des bornes inférieures en théorie de Ramsey : les constructions arithmétiques. Un exemple déjà vu en TD est le graphe de Paley, où les sommets sont les éléments de $\mathbb{Z}/17\mathbb{Z}$, et où i est relié à j si $i - j$ est un carré modulo 17, et qui permet de montrer que $R(4, 4) > 17$. Une manière de lire cette construction est que du point de vue de la structure additive de $\mathbb{Z}/p\mathbb{Z}$, l'ensemble des carrés de $\mathbb{Z}/p\mathbb{Z}$ se comporte comme un ensemble aléatoire [5], mais sans les "accidents" qui peuvent arriver dans des constructions aléatoires.

On va maintenant voir un autre exemple un peu plus élaboré de cette idée, qui améliore partiellement la borne inférieure sur les nombres de van der Waerden donnée dans la section précédente, dans le cas où il n'y a que deux couleurs. Ce résultat est dû à Berlekamp en 1968.

Théorème 12 ([3])**.** Soit q un nombre premier. Alors $W(q + 1, 2) > q(2^q - 1)$.

Remarque 31.
- Cette borne est meilleure que celle de [19] évoquée précédemment, mais à la condition que q soit premier.
- Ce résultat se limite à 2 couleurs, mais une généralisation a été donnée récemment dans [4]. La borne inférieure obtenue est $W(q+1, r) > q^{r-1}2^q$ pour r quelconque et q premier. C'est bien mieux que nos bornes probabilistes quand $r \to +\infty$ à q fixé, mais moins bon quand $q \to +\infty$ à r fixé.

L'outil principal dans la preuve du résultat de Berlekamp sont les *corps finis*, et plus particulièrement le corps fini de cardinal 2^q. Avant de présenter cette construction, commençons par les bases de la théorie des corps finis.

Théorème 13. Soit $n \geq 2$ un entier. Alors il existe un corps commutatif [6] $\mathbb{F}_n$ à n éléments si et seulement si n est une puissance d'un nombre premier. De plus, dans ce cas $\mathbb{F}_n$ est unique à isomorphisme près.

Remarque 32. Pour p premier, le corps fini $\mathbb{F}_p$ est juste $\mathbb{Z}/p\mathbb{Z}$.

Démonstration. Ce résultat sera admis, voir par exemple [15, Chapitre III]. Expliquons tout de même la partie la plus facile du résultat, à savoir que si $\mathbb{F}_n$ existe, alors n est une puissance d'un nombre premier. Si $\mathbb{F}_n$ existe, soit p le plus petit entier tel que

$$\underbrace{1 + \cdots + 1}_{p \text{ fois}} = 0.$$

5. C'est complètement faux du point de vue de la structure multiplicative, puisque le produit de deux carrés est toujours un carré.

6. Tout corps fini est en fait commutatif, d'après le théorème de Wedderburn.

Notons qu'un tel p existe car d'après le principe des tiroirs il existe k, ℓ tels que

$$\underbrace{1+\cdots+1}_{k \text{ fois}} = \underbrace{1+\cdots+1}_{\ell \text{ fois}}.$$

Si p peut s'écrire sous la forme $p = qr$ avec $q, r \geq 1$. Alors on a

$$\left(\underbrace{1+\cdots+1}_{q \text{ fois}}\right)\left(\underbrace{1+\cdots+1}_{r \text{ fois}}\right) = 0.$$

Un des deux facteurs du produit est donc nul, donc $q \geq p$ ou $r \geq p$, donc p est premier. De plus, l'ensemble

$$\left\{0, 1, 1+1, \ldots, \underbrace{1+\cdots+1}_{p-1 \text{ fois}}\right\}$$

est un sous-corps de $\mathbb{F}_n$ isomorphe à $\mathbb{F}_p$. On a donc sur $\mathbb{F}_n$ une structure d'espace vectoriel sur $\mathbb{F}_p$, qui est de dimension finie k car $\mathbb{F}_n$ est fini. Si on note $(e_1, \ldots, e_k)$ une base de $\mathbb{F}_n$ en tant que $\mathbb{F}_p$-espace vectoriel, on peut donc décrire $\mathbb{F}_n$ comme l'ensemble des combinaisons linéaires

$$\sum_{i=1}^{k} a_i e_i$$

avec $a_i \in \mathbb{F}_p$, où ces p^k combinaisons linéaires sont deux à deux distinctes car la famille (e_i) est libre. On a donc $|\mathbb{F}_n| = p^k$. □

Pour notre construction, on aura besoin de deux propriétés des corps finis. La première concerne le degré du polynôme minimal des éléments de $\mathbb{F}_{p^k}$.

Proposition 33. Soit $\alpha \in \mathbb{F}_{p^k} \backslash \{0\}$. Alors l'ensemble des polynômes $Q \in \mathbb{F}_p[X]$ tels que $Q(\alpha) = 0$ est de la forme

$$\{P_\alpha R | R \in \mathbb{F}_p[X]\},$$

où P_α est un polynôme unitaire. De plus, le degré de P_α divise k.

Démonstration. On vérifie que l'ensemble des $Q \in \mathbb{F}_p[X]$ qui annulent α est un idéal de $\mathbb{F}_p[X]$:

- si $Q_1(\alpha) = Q_2(\alpha) = 0$, alors $(Q_1 + Q_2)(\alpha) = 0$;
- si $Q_1(\alpha) = 0$, alors $(RQ_1)(\alpha) = R(\alpha)Q_1(\alpha) = 0$ pour tout $R \in \mathbb{F}_p[X]$.

De plus, cet idéal est non vide car la famille $\{1, \alpha, \alpha^2, \ldots, \alpha^k\}$ est de cardinal $k+1$, donc elle est liée dans le $\mathbb{F}_p$-espace vectoriel $\mathbb{F}_{p^k}$ de dimension k, donc il existe Q de degré au plus k qui annule α. L'existence de P_α vient donc du fait que tous les idéaux de $\mathbb{F}_p[X]$ sont principaux.

De plus, soit d le degré de P_α. On note $\mathbb{K}$ le sous-$\mathbb{F}_p$-espace vectoriel de $\mathbb{F}_{p^k}$ engendré par $\{1, \alpha, \ldots, \alpha^{d-1}\}$.Par définition de P_α la famille $\{1, \alpha, \ldots, \alpha^{d-1}\}$ est libre, donc $\mathbb{K}$ est de dimension d sur $\mathbb{F}_p$. On vérifie maintenant que $\mathbb{K}$ est un sous-corps de $\mathbb{F}_{p^k}$. Pour cela, on écrit

$$P_\alpha = a_0 + a_1 X + \cdots + a_{d-1} X^{d-1} + X^d$$

avec $a_0, a_1, \ldots, a_{d-1} \in \mathbb{F}_p$. Le fait que $P_\alpha(\alpha) = 0$ se réécrit

$$\alpha^d = -a_{d-1} X^{d-1} - \cdots - a_1 \alpha - a_0, \tag{9}$$

qui est dans $\mathbb{K}$. En multipliant (9) par des puissances de α, on en déduit par récurrence sur i que $\alpha^i \in \mathbb{K}$ pour tout $i \geq 0$, donc que $\mathbb{K}$ est stable par multiplication. Enfin, soit $x \in \mathbb{K}\backslash\{0\}$. L'application de $\mathbb{K}$ dans $\mathbb{K}$ qui à y associe xy est injective (car un corps est intègre), donc bijective (car $\mathbb{K}$ est fini), donc 1 admet un antécédent, ce qui signifie que x^{-1} est dans $\mathbb{K}$, donc $\mathbb{K}$ est bien un sous-corps, et on a $\mathbb{F}_p \subset \mathbb{K} \subset \mathbb{F}_{p^k}$.

Pour conclure, soit d' la dimension de $\mathbb{F}_{p^k}$ en tant que $\mathbb{K}$-espace vectoriel. On va montrer que $k = dd'$, ce qui impliquera en particulier que d divise k. Comme on l'a vu dans la preuve d'une partie du Théorème 13, le cardinal d'un espace vectoriel de dimension k sur un corps fini de cardinal m et m^k. En utilisant cela deux fois, on obtient

$$p^k = \left|\mathbb{F}_{p^k}\right| = |\mathbb{K}|^{d'} = \left(|\mathbb{F}_p|^d\right)^{d'} = p^{dd'},$$

d'où $k = dd'$, ce qui conclut. □

Avant de passer au résultat de Berlekamp, il nous manque une autre propriété, qui concerne la structure multiplicative des corps finis.

Proposition 34. Soient p premier et $k \geq 1$. Alors le groupe multiplicatif $\left(\mathbb{F}_{p^k}\backslash\{0\}, \times\right)$ est cyclique, i.e. il existe $\alpha \in \mathbb{F}_{p^k}$ tel que $\mathbb{F}_{p^k}\backslash\{0\} = \{\alpha^i | i \geq 0\}$.

Ce dernier résultat est en fait plus un résultat de théorie des groupes que de théorie des corps. Notons qu'il n'est pas évident, puisqu'il existe des groupes commutatifs non cycliques, comme $(\mathbb{Z}/2\mathbb{Z})^2$.

Définition 35.
- Soient G un groupe fini et $\alpha \in G$. L'*ordre* de x est le plus petit $n \geq 1$ tel que $\alpha^n = 1$.
- L'*exposant* du groupe G est le plus petit n tel que $\alpha^n = 1$ pour tout α dans G. C'est aussi le PPCM des ordres des éléments de G.

Lemme 36. Soit G un groupe fini commutatif, et soit e son exposant. Alors il existe $x \in G$ d'ordre exactement e.

Démonstration. On commence par écrire $e = q_1^{\beta_1} \dots q_\ell^{\beta_\ell}$ la décomposition de e en facteurs premiers. Dans un premier temps, on va trouver des éléments d'ordre exactement $q_i^{\beta_i}$ pour tout i. Pour cela, par définition de e, on sait qu'il existe $\alpha \in G$ dont l'ordre est divisible par $q_i^{\beta_i}$. L'ordre de α s'écrit alors $q_i^{\beta_i} \times r$, et on a $(\alpha^r)^j = 1$ ssi rj est divisible par $q_i^{\beta_i} \times r$, ce qui équivaut à $j | q_i^{\beta_i}$, donc α^r est bien d'ordre $q_i^{\beta_i}$.

Pour conclure, on va montrer que si l'ordre de a_1 vaut n_1 et l'ordre de a_2 vaut n_2 avec n_1, n_2 premiers entre eux, alors l'ordre de $a_1 a_2$ vaut $n_1 n_2$. En effet, soit d l'ordre de $a_1 a_2$. Alors on a

$$(a_1 a_2)^{n_1 n_2} = (a_1^{n_1})^{n_2} (a_2^{n_2})^{n_1} = 1^{n_2} 1^{n_1} = 1,$$

donc $d | n_1 n_2$. D'autre part, on a $a_1^d a_2^d = 1$ donc $a_2^d = a_1^{-d}$, donc

$$a_2^{n_1 d} = a_1^{-n_1 d} = 1,$$

donc $n_2 | n_1 d$. Comme n_1 et n_2 sont premiers entre eux, on en déduit $n_2 | d$, et on montre de même $n_1 | d$, donc $n_1 n_2 | d$ et finalement $d = n_1 n_2$. Finalement, pour tout i, soit γ_i un élément de G d'ordre $q_i^{\beta_i}$. On en déduit par récurrence sur i que $\gamma_1 \dots \gamma_i$ est d'ordre $q_1^{\beta_1} \dots q_i^{\beta_i}$, donc $\gamma_1 \dots \gamma_\ell$ est d'ordre exactement e. □

Preuve de la Proposition 34. Soit e l'exposant du groupe $\big(\mathbb{F}_{p^k}\backslash\{0\}, \times\big)$. Étant donné le lemme précédent, il suffit maintenant de montrer que $e = p^k - 1$. D'après le théorème de Lagrange appliqué au sous-groupe multiplicatif engendré par un élément, on sait que e divise $p^k - 1$. De plus, on a $x^e = 2$ pour tout $x \in \mathbb{F}_{p^k}\backslash\{0\}$, donc le polynôme $X^e - 1$ a $p^k - 1$ racines dans $\mathbb{F}_{p^k}$, donc $e \geq p^k - 1$, d'où finalement $e = p^k - 1$. □

Munis de tous ces outils, nous pouvons enfin démontrer le résultat de Berlekamp. La construction n'utilise les résultats qui précèdent que pour $p = 2$.

Démonstration du Théorème 12. Soit $\alpha \in \mathbb{F}_{2^q}$ un élément qui engendre le groupe multiplicatif $(\mathbb{F}_{2^q}\backslash\{0\}, \times)$ (qui existe d'après la Proposition 34), et soit $\{v_1, \ldots, v_q\}$ une base de $\mathbb{F}_{2^q}$ sur $\mathbb{F}_2$. Pour tout $j \geq 1$, on décompose α^j dans cette base :

$$\alpha^j = \sum_{i=1}^{q} x_{ij} v_i$$

avec $x_{ij} \in \mathbb{F}_2$. Pour tout $1 \leq j \leq q(2^q - 1)$, on colorie j en bleu si $x_{1j} = 0$, et en rouge si $x_{1j} = 1$. On va maintenant vérifier qu'il n'y a pas de progression monochrome de longueur $q + 1$. S'il y en a une, il existe $b, c \geq 1$ tels que les α^{b+ic} sont tous de la même couleur, avec $1 \leq c < 2^q - 1$. En posant $\beta = \alpha^b$ et $\gamma = \alpha^c$, on a $\gamma \neq 1$ car $c < 2^q - 1$, et les éléments $\beta, \beta\gamma, \ldots, \beta\gamma^q$ de $\mathbb{F}_{2^k}$ sont tous de la même couleur.

- si c'est bleu, alors $\beta, \beta\gamma, \ldots, \beta\gamma^{q-1}$ sont tous dans un hyperplan de $\mathbb{F}_{2^q}$ (celui engendré par $\{v_2, \ldots, v_q\}$), donc ils forment une famille liée. Il y a donc des coefficients $a_i \in \mathbb{F}_2$ non tous nuls tels que
$$\sum_{i=0}^{q-1} a_i \beta\gamma^i = 0.$$
Comme $\beta \neq 0$, cela signifie qu'il existe un polynôme $P \in \mathbb{F}_2[X]$ non nul, de degré au plus $q - 1$, et qui annule γ. En particulier, le polynôme minimal P_γ de γ est de degré au plus $q - 1$. D'un autre côté, d'après la Proposition 33, le degré de P_γ divise q qui est premier, donc $\deg(P_\gamma) = 1$ donc $\gamma \in \mathbb{F}_2$. Mais on sait que $\gamma \neq 0$ et que $\gamma \neq 1$, d'où la contradiction. Notons que c'est ici qu'on a utilisé de manière cruciale le fait que q est premier.
- Si la couleur est le rouge, alors $\beta, \beta\gamma, \ldots, \beta\gamma^q$ ont tous comme première coordonnée 1 dans la base $\{v_1, \ldots, v_q\}$. Par conséquent, les éléments
$$\beta\gamma - \beta, \beta\gamma^2 - \beta, \ldots, \beta\gamma^q - \beta$$
ont tous pour première coordonnée 0, donc ils sont dans un hyperplan de $\mathbb{F}_{2^q}$ en tant que $\mathbb{F}_2$-espace vectoriel, donc ils forment une famille liée, donc il existe des $a_i \in \mathbb{F}_2$ non tous nuls tels que
$$\sum_{i=1}^{q} a_i \beta(\gamma^i - 1) = 0.$$
Comme $\beta \neq 0$ et $\gamma \neq 1$, on peut factoriser cette dernière formule par $\beta(\gamma - 1)$ pour obtenir
$$\sum_{i=1}^{q} a_i (1 + \gamma + \cdots + \gamma^{i-1}) = 0,$$

donc il existe $P \in \mathbb{F}_2[X]$ non nul, de degré au plus $q-1$ et qui annule γ. On en déduit une contradiction comme dans le premier cas.

□

Remarque 37. On pourrait être tenté d'utiliser le corps fini $\mathbb{F}_{p^q}$ pour obtenir une borne inférieure sur les nombres de Van der Waerden $W(q+1,p)$ pour tous p et q premiers. Malheureusement, notre argument ne fonctionne pas. En effet, on a pu conclure car on savait que $\gamma \notin \{0,1\}$, mais pour $p \geq 3$ on pourrait avoir $\gamma \in \mathbb{F}_p \backslash \{0,1\}$.

Références

[1] N. Alon, J. Spencer, and P. Erdos. *The Probabilistic Method*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 1992.

[2] V. Angeltveit and B. D. McKay. R(5,5) and 48. *Journal of Graph Theory*, 89(1) :5–13, 2018.

[3] E. Berlekamp. A construction for partitions which avoid long arithmetic progressions. *Canadian Mathematical Bulletin*, 11(3) :409–414, 1968.

[4] T. Blankenship, J. Cummings, and V. Taranchuk. A new lower bound for van der waerden numbers. *European Journal of Combinatorics*, 69 :163–168, 2018.

[5] M. Campos, S. Griffiths, R. Morris, and J. Sahasrabudhe. An exponential improvement for diagonal Ramsey. *arXiv :2303.09521*, 2023.

[6] P. Erdös. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4) :292 – 294, 1947.

[7] P. Erdös and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. *Infinite and finite sets*, 10(2) :609–627, 1975.

[8] P. Erdös and G. Szekeres. A combinatorial problem in geometry. *Compositio Mathematica*, 2 :463–470, 1935.

[9] G. Exoo. A lower bound for r(5, 5). *Journal of Graph Theory*, 13(1) :97–98, 1989.

[10] H. Furstenberg and B. Weiss. Topological dynamics and combinatorial number theory. *Journal d'Analyse Mathématique*, 34 :61–85.

[11] T. Gowers. A new proof of Szemerédi's theorem. *Geometric and Functional Analysis*, 11 :465–588, 2001.

[12] R. Graham, B. Rothschild, and J. Spencer. *Ramsey Theory*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 1991.

[13] R. L. Graham and B. Rothschild. A short proof of van der Waerden's theorem on arithmetic progressions. *Proceedings of the AMS*, 42 :385–386.

[14] M. Katz and J. Reimann. *An Introduction to Ramsey Theory*. Student Mathematical Library. American Mathematical Society, 2018.

[15] D. Perrin. *Cours d'Algèbre*. Collection CAPES/Agrégation. 1996.

[16] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, s2-30(1) :264–286, 1930.

[17] A. Sah. Diagonal Ramsey via effective quasirandomness, 2020.

[18] J. Spencer. Ramsey's theorem—a new lower bound. *Journal of Combinatorial Theory, Series A*, 18(1) :108–115, 1975.

[19] Z. Szabó. An application of Lovász' local lemma-a new lower bound for the van der waerden number. *Random Structures & Algorithms*, 1(3) :343–360, 1990.

[20] van der Waerden B. L. Beweis einer Baudetschen Vermutung. *Nieuw Arch. Wiskunde*, 15 :212–216, 1927.