Séminaire BOURBAKI Mars 2005
57e année, 2004-2005, no 944, p. 229 à 246

PROGRESSIONS ARITHMÉTIQUES DANS LES NOMBRES PREMIERS [d'après B. Green et T. Tao]

par **Bernard HOST**

1. INTRODUCTION

1.1. Le résultat

Le but de cet exposé est de présenter un travail récent et spectaculaire de B. Green et T. Tao où ils montrent :

THÉORÈME 1.1 ([6]). — *L'ensemble des nombres premiers contient des progressions arithmétiques de toutes longueurs.*

En fait Green et Tao montrent un résultat plus fort : la conclusion du théorème reste valable si on remplace l'ensemble des nombres premiers par un sous-ensemble de densité relative positive. De plus, la méthode employée permet de déterminer explicitement pour tout k un entier N (très grand) tel que l'ensemble des nombres premiers plus petits que N contienne une progression arithmétique de longueur $k+1$.

Le théorème 1.1 répond à une question fort ancienne bien que difficile à dater exactement. Très peu de résultats partiels étaient connus jusqu'ici ; citons celui de van der Corput [12] qui a montré en 1939 l'existence d'une infinité de progressions de longueur 3 dans les nombres premiers.

En 1923, Hardy et Littlewood [7] ont proposé une conjecture très générale sur la répartition de certaines configurations dans les nombres premiers, qui entraînerait une version quantitative précise du théorème 1.1 si elle s'avérait exacte. Ce même théorème suivrait aussi d'une résolution positive donnée à une conjecture proposée par Erdös et Turán [1] en 1936 :

CONJECTURE. — *Tout sous-ensemble E de $\mathbb{N}^*$ vérifiant $\sum_{n\in E} 1/n = +\infty$ contient des progressions arithmétiques de toutes longueurs.*

Cette conjecture reste totalement ouverte et les méthodes de Green et Tao ne permettent pas de s'en approcher. Dans une direction voisine, Szemerédi a montré en 1975 l'existence de progressions sous l'hypothèse plus forte de la densité positive. Rappelons que la densité d'un ensemble d'entiers $E \subset \mathbb{N}$ est :

$$d^*(E) = \limsup_{N\to\infty} \frac{1}{N} \operatorname{Card}(E \cap [0, N-1]).$$

Le théorème de Szemerédi s'énonce :

THÉORÈME DE SZEMERÉDI ([9]). — *Tout ensemble d'entiers de densité positive contient des progressions arithmétiques de toutes longueurs.*

Il peut aussi s'exprimer en termes d'ensembles finis d'entiers :

VERSION FINIE DU THÉORÈME DE SZEMERÉDI. — *Pour tout entier $k \geqslant 2$ et tout réel $\delta > 0$ il existe un entier $N = N(k, \delta)$ tel que tout sous-ensemble E de $[0, N[$ ayant au moins δN éléments contienne une progression arithmétique de longueur $k+1$.*

Ce théorème ne peut évidemment pas être utilisé directement puisque les nombres premiers ont une densité nulle. Cependant il tient une place centrale dans la démonstration.

1.2. La méthode

Le travail de Green et Tao comporte deux parties très différentes.

La première partie, qui est la plus longue, contient la démonstration d'une extension de la version finie du théorème de Szemerédi (théorème 2.2).

Dans ce dernier théorème, la quantité $|E|/N \geqslant \delta$ peut être vue comme la moyenne sur $[0, N[$ de la fonction indicatrice de E. L'idée naturelle est de remplacer cette fonction indicatrice par une fonction nulle en dehors de l'ensemble des nombres premiers, mais alors cette fonction ne peut pas être choisie majorée par 1 sinon sa moyenne deviendrait arbitrairement petite pour N grand. Green et Tao montrent un théorème de Szemerédi modifié (théorème 2.2) qui s'applique à une fonction majorée par un « poids pseudo-aléatoire », c'est-à-dire par une fonction de moyenne 1 dont les corrélations sont voisines de celles qu'on obtiendrait en tirant au hasard et indépendamment les valeurs aux points $0, 1, \ldots, N-1$ (section 2.2). Cette utilisation d'une majoration fait penser à la méthode du crible.

La démonstration de ce « théorème de Green-Tao Szemerédi » est écrite dans le langage des probabilités. Comme tous les espaces de probabilité sont finis et munis de la mesure uniforme, on pourrait dire qu'elle utilise seulement des arguments de dénombrement. Cette façon de voir serait formellement correcte mais trop réductrice. En fait la démarche de Green et Tao s'inspire directement de la théorie ergodique, et plus précisément de la démonstration ergodique du théorème de Szemerédi donnée par Furstenberg ([2], voir aussi [3]). Dans les deux cas, le cœur de la preuve est un résultat de décomposition (proposition 3.4) consistant à écrire une fonction comme la

somme de son espérance conditionnelle sur une σ-algèbre bien choisie et d'un reste. L'espérance conditionnelle est « lissée » et dans le cas considéré par Green et Tao elle est même uniformément bornée, ce qui permet d'utiliser le théorème de Szemerédi classique. Le reste se comporte comme une oscillation aléatoire et sa contribution dans les calculs est négligeable. Les ergodiciens reconnaîtront la façon dont les « facteurs » interviennent dans de nombreux problèmes. Pour les autres, nous ajoutons que l'article n'utilise aucun résultat provenant de la théorie ergodique et que sa lecture ne demande aucune connaissance dans ce domaine.

Cette inspiration ergodique dans une démonstration combinatoire est encore plus apparente dans la nouvelle démonstration que T. Tao vient de donner du théorème de Szemerédi [10]. Nous ne pensons pas que cette démarche soit artificielle. Jusqu'à présent les relations entre ces domaines se résumaient pratiquement au principe de correspondance de Furstenberg qui permet de montrer, à partir de théorèmes ergodiques, des résultats combinatoires dont beaucoup n'ont aujourd'hui pas d'autre preuve. Il apparaît depuis peu des ressemblances de plus en plus prononcées quoiqu'encore mal comprises entre les objets et les méthodes des deux théories. Nous reviendrons dans ces notes sur ce point qui mérite sans doute d'être approfondi.

Une fois démontré le théorème de Szemerédi modifié, il reste à construire un poids pseudo-aléatoire adapté au problème posé. Il s'agit donc ici de théorie des nombres. Dans cette partie de l'article [6] les auteurs utilisent une fonction de von Mangoldt tronquée et font appel à des outils sophistiqués provenant des travaux de Goldston et Yıldırım [4] mais, dans une note non publiée [11], T. Tao explique comment l'argument peut être modifié pour n'utiliser que les propriétés les plus élémentaires des nombres premiers et de la fonction ζ. C'est cette approche que nous suivons ici en nous inspirant de notes manuscrites de J.-C. Yoccoz.

Dans cet exposé, qui ne contient aucune démonstration complète, on se propose de présenter de façon assez détaillée l'organisation de la preuve et de donner une idée des méthodes employées à chaque étape. Le lecteur pressé pourra se limiter à la section 2 qui contient la formulation précise des définitions et résultats correspondant aux deux grandes parties auxquelles on vient de faire allusion, encore que la définition des normes de Gowers (sous-sections 3.1 et 3.2) ait son intérêt propre. Le résultat de décomposition (proposition 3.4) est énoncé dans la sous-section 3.4 et montré dans la section 4. La deuxième partie de la preuve, c'est-à-dire la construction du poids pseudo-aléatoire, est contenue dans la section 5.

1.3. Conventions et notations

Quand f est une fonction définie sur un ensemble fini A, l'*espérance* de f sur A, notée $\mathbb{E}(f(x) \mid x \in A)$ ou $\mathbb{E}(f \mid A)$, est la moyenne arithmétique de f sur A ; la même est utilisée pour les fonctions de plusieurs variables.

Dans toute la suite, $k \geqslant 2$ est un entier que nous considérons comme une constante. L'objectif est de montrer l'existence d'une progression arithmétique de longueur $k+1$

dans les nombres premiers. La progression est cherchée dans l'intervalle $[0, N[$, où N est un (grand) entier qu'il est souvent nécessaire de supposer premier. On identifie $[0, N[$ au groupe $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$.

Il est crucial dans la preuve de contrôler la manière dont toutes les estimations dépendent de N et nous adoptons les conventions suivantes. Dans chaque énoncé, N est supposé fixé mais toutes les constantes sont indépendantes de N. Nous notons $o(1)$ une quantité tendant vers 0 quand N tend vers l'infini, uniformément par rapport à tous les paramètres sauf éventuellement ceux notés en indice. La notation $O(1)$ est employée avec une signification similaire.

2. POIDS PSEUDO-ALÉATOIRES ET THÉORÈME DE GREEN-TAO SZEMERÉDI

Green et Tao généralisent une formulation classique du théorème de Szemerédi, qui est celle sous laquelle Gowers [5] l'a redémontré récemment.

THÉORÈME 2.1. — *Pour tout réel $\delta > 0$ il existe une constante $c(\delta) > 0$ telle que, pour toute fonction $f\colon \mathbb{Z}_N \to \mathbb{R}$ avec*

$$0 \leqslant f(x) \leqslant 1 \text{ pour tout } x \text{ et } \mathbb{E}\big(f \mid \mathbb{Z}_N\big) \geqslant \delta$$

on ait

$$\mathbb{E}\big(f(x)f(x+t)\cdots f(x+kt) \mid x, t \in \mathbb{Z}_N\big) \geqslant c(\delta). \tag{1}$$

La version finie du théorème de Szemerédi se déduit de ce théorème en prenant pour f la fonction indicatrice d'un sous-ensemble de $[0, N[$. Green et Tao s'affranchissent de la condition $f \leqslant 1$ en la remplaçant par l'hypothèse que f est majorée par un *poids pseudo-aléatoire* ; cette notion sera définie plus loin.

2.1. Les deux ingrédients de la preuve du théorème 1.1

Nous appelons « théorème de Green-Tao Szemerédi » l'extension suivante du théorème de Szemerédi :

THÉORÈME 2.2. — *Soit $\nu\colon \mathbb{Z}_N \to \mathbb{R}^+$ un poids pseudo-aléatoire (voir la sous-section 2.2). Pour tout réel $\delta > 0$ il existe une constante $c'(\delta) > 0$ satisfaisant la propriété suivante. Pour toute fonction $f\colon \mathbb{Z}_N \to \mathbb{R}$ telle que*

$$0 \leqslant f(x) \leqslant \nu(x) \text{ pour tout } x \text{ et } \mathbb{E}\big(f \mid \mathbb{Z}_N\big) \geqslant \delta$$

on a

$$\mathbb{E}\big(f(x)f(x+t)\cdots f(x+kt) \mid x, t \in \mathbb{Z}_N\big) \geqslant c'(\delta) - o(1). \tag{2}$$

La démonstration de ce théorème, qui occupe une part importante de l'article de Green et Tao, est résumée dans les sections 3 et 4. Pour l'appliquer aux nombres premiers, il faut une fonction f et un poids ν convenables dont l'existence est donnée par le théorème suivant.

THÉORÈME 2.3. — *Il existe une constante positive δ, un poids pseudo-aléatoire $\nu\colon \mathbb{Z}_N \to \mathbb{R}^+$ et une fonction $f\colon \mathbb{Z}_N \to \mathbb{R}$ avec*

$$f \text{ est nulle en dehors de l'ensemble des nombres premiers ;}$$
$$0 \leqslant f(x) \leqslant \nu(x) \text{ pour tout } x \text{ ;}$$
$$\mathbb{E}(f \mid \mathbb{Z}_N) \geqslant \delta;$$
$$\|f\|_{L^\infty} = O(1)\ \log N.$$

La construction de f et ν est faite dans la section 5. Nous montrons maintenant comment le théorème 1.1 découle des théorèmes 2.2 et 2.3.

Démonstration. — Soient δ, f et ν comme dans le théorème 2.3. Il existe un intervalle $J \subset [0, N[$, de longueur plus petite que $N/2$ et tel que $\mathbb{E}(\mathbf{1}_J f \mid \mathbb{Z}_N) \geqslant \delta/3$. Nous utilisons le théorème 2.2 avec la fonction f remplacée par $\mathbf{1}_J f$ et le réel δ remplacé par $\delta/3$.

La contribution dans l'espérance (2) des termes où $t = 0$ est majorée par $N^{-1}\|f\|_{L^\infty}^{k+1} = o(1)$, et est donc inférieure à $c'(\delta)$ si N est assez grand. Il existe donc dans ce cas $x, t \in \mathbb{Z}_N$ avec $t \neq 0$ tels que $f(x)f(x+t)\dots f(x+kt) \neq 0$. Rappelons que dans cette expression $x, x+t, \dots, x+kt$ sont considérés comme des éléments de $\mathbb{Z}_N$ et que donc l'addition est modulo N. Si nous considérons x et t comme des entiers appartenant à l'intervalle $[0, N[$ nous obtenons que f est non nulle aux points $x, x+t \bmod N, \dots, x+kt \bmod N$. Comme elle est nulle en dehors de l'intervalle J de longueur $< N/2$, tous ces entiers appartiennent à cet intervalle et on en déduit facilement qu'ils forment une progression arithmétique non triviale de longueur $k+1$. Enfin, f est nulle en dehors de l'ensemble des nombres premiers et on a bien une progression formée de nombres premiers. □

2.2. Définition des poids pseudo-aléatoires

Dans les théorèmes précédents nous avons considéré un *poids pseudo-aléatoire* comme une fonction définie sur $\mathbb{Z}_N$. Il s'agit plus précisément de la donnée, pour chaque nombre premier N, d'une fonction $\nu = \nu_N\colon \mathbb{Z}_N \to \mathbb{R}^+$, de sorte que soient satisfaites deux conditions asymptotiques appelées *condition sur les formes linéaires* et *condition sur les corrélations*.