363-364

ASTÉRISQUE

2015

SÉMINAIRE BOURBAKI
VOLUME 2013/2014
EXPOSÉ Nº 1084

Emmanuel KOWALSKI

Gaps between prime numbers and primes in arithmetic progressions

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

Publié avec le concours du CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Astérisque est un périodique de la Société Mathématique de France.

Numéro 367-368

Diffusion

Maison de la SMF	Hindustan Book Agency	AMS
Case 916 - Luminy	O-131, The Shopping Mall	P.O. Box 6248
13288 Marseille Cedex 9	Arjun Marg, DLF Phase 1	Providence RI 02940
France	Gurgaon 122002, Haryana	USA
smf@smf.univ-mrs.fr	Inde	www.ams.org

Tarifs

Vente au numéro : 90 € ($ 135)

Abonnement Europe : 650 €, hors Europe : 689 € ($ 1 033)

Des conditions spéciales sont accordées aux membres de la SMF.

Secrétariat : Nathalie Christiaën

Astérisque
Société Mathématique de France
Institut Henri Poincaré, 11, rue Pierre et Marie Curie
75231 Paris Cedex 05, France
Tél : (33) 01 44 27 67 99 • Fax : (33) 01 40 46 90 96
revues@smf.ens.fr • http://smf.emath.fr/

ISSN 0303-1179

ISBN 978-2-85629-804-6

Directeur de la publication : Marc PEIGNÉ

Séminaire BOURBAKI Mars 2014
66e année, 2013-2014, n° 1084, p. 327 à 366

GAPS BETWEEN PRIME NUMBERS AND PRIMES IN ARITHMETIC PROGRESSIONS [after Y. Zhang and J. Maynard]

by **Emmanuel KOWALSKI**

. . . utinam intelligere possim rationacinationes pulcherrimas quae e propositione concisa $|\sum x_i y_i| \leqslant \|x\|\ \|y\|$ *fluunt. . .*

(after H. Cartan)

1. INTRODUCTION

Y. Zhang proved in **[27]**, announced in May 2013, the following theorem:

Theorem 1.1 (Zhang). — *There exists an even integer* $h \geqslant 2$ *with the property that there exist infinitely many pairs of prime numbers of the form* $(p, p+h)$*. In fact, there exists such an* h *with* $h \leqslant 70{,}000{,}000$*.*

Equivalently, if p_n*,* $n \geqslant 1$*, denotes the* n*-th prime number, we have*

$$\liminf_{n\to+\infty}(p_{n+1} - p_n) < +\infty,$$

and more precisely

$$\liminf_{n\to+\infty}(p_{n+1} - p_n) \leqslant 70{,}000{,}000.$$

The equivalence of the two formulations is clear by the pigeon-hole principle. The first one is psychologically more spectacular: it emphasizes the fact that *for the first time* in history, one has proved an *unconditional* existence result for infinitely many primes p and q constrained by a binary condition $q - p = h$.

Remarkably, this already extraordinary result was improved in spectacular fashion in October 2013 by J. Maynard **[21]**:(1)

Theorem 1.2 (Maynard). — *There exists an even integer* $h \leqslant 600$ *with the property that there exist infinitely many pairs of prime numbers of the form* $(p, p+h)$*. In fact, for any fixed integer* $k \geqslant 2$*, there exist* k *distinct integers* $(h_1, \ldots, h_k)$ *such that the set of integers* n *for which* $n + h_1$*, . . . ,* $n + h_k$ *are all primes is infinite.*

(1) T. Tao derived similar results independently.

Equivalently, we have

$$\liminf_{n\to+\infty}(p_{n+m}-p_n)<+\infty \tag{1}$$

for any fixed $m \geqslant 1$.

We emphasize that Zhang's method did not apply to (1) except for $m=1$, and thus Maynard's work is a far-reaching extension, and not merely a strengthening, of Zhang's Theorem. It is all the more amazing that Maynard's proof is, in its technical aspects, much simpler than Zhang's. Indeed, it can be realistically presented from scratch in a one-year graduate course in analytic number theory. On the other hand, the main ingredient in Zhang's proof is a statement concerning the distribution of primes in arithmetic progressions to large moduli which has considerable independent interest (see Theorem 4.1), since it reflects information which lies beyond the immediate reach of the Generalized Riemann Hypothesis.

The goal of this text is to present the context of these spectacular results, and to sketch some of the main steps of the proofs, with an emphasis on the key ideas (and some bias related to the author's taste). There already exist at least three independent full (or almost full) expositions of the results of Zhang (by the Polymath 8 group **[24]**, Granville **[14]** and Friedlander–Iwaniec **[11]**) and two of Maynard's which, as we indicated, are in any case technically simpler (in a blog post of Tao **[26]** and in **[14]**). There is therefore no doubt concerning the correctness of the results, and we will emphasize conceptual aspects instead of trying to write another complete proof.

We will also not attempt to describe the techniques that lead to the best possible bounds of h currently known, and to the best upper-bound for the liminf in (1) (as a function of m). These are for the essential part found in the versions of Theorems 1.1 and 1.2 proved in the course of the Polymath8 project (see **[24]** and the ongoing Polymath8b work).

Notation. — p will always refer to a prime number and n to positive integers; it will be convenient to denote by $\mathbf{M}$ (like "moduli") the set of squarefree numbers (i.e., those positive integers q such that no square of a prime divides q); given $y \geqslant 2$, an integer $n \geqslant 1$ is called y-friable if and only if it has no prime factor $p \mid n$ such that $p > y$.[2]

$\Lambda(n)$ denotes the von Mangoldt function, equal to $\log p$ for all prime powers p^k, $k \geqslant 1$, and to zero for other integers, and we also denote by $\theta(n)$ the function equal to $\log p$ for a prime p, and equal to 0 otherwise.

(2) A standard terminology is "smooth" instead of friable.

For $X \geqslant 2$, $q \geqslant 1$ and $a \geqslant 1$ (or $a \in \mathbf{Z}/q\mathbf{Z}$), we let

$$\psi(X; q, a) = \sum_{\substack{X < n \leqslant 2X \\ n \equiv a \,(\mathrm{mod}\ q)}} \Lambda(n).$$

We extend this slightly as follows: an *idelette* a is an element of

$$\mathbb{I} = \prod_p (\mathbf{Z}/p\mathbf{Z})^\times,$$

which therefore defines, for every modulus $q \in \mathbf{M}$, a unique invertible residue class a_q modulo q by the Chinese Remainder Theorem; we then put

$$\psi(X; q, a) = \psi(X; q, a_q).$$

$\mu(n)$ denotes the Möbius function, $\varphi(n)$ the Euler function, $\tau(n)$ the "number of divisors" function and $\tau_k(n)$ its generalization to the number of representations $n = d_1 \cdots d_k$, i.e.,

$$\tau_k(n) = \sum_{d_1 \cdots d_k = n} 1, \qquad \sum_{n \geqslant 1} \tau_k(n) n^{-s} = \zeta(s)^k$$

(for instance $\tau_k(p) = k + 1$ for p prime), so that $\tau = \tau_2$.

$f \star g$ denotes the Dirichlet convolution of arithmetic functions f and g, defined by

$$f \star g(n) = \sum_{d \mid n} f(d) g(n/d) = \sum_{ab = n} f(a) g(b)$$

for all $n \geqslant 1$.

The notation $f = O(g)$ and $f \ll g$ are synonymous: $f(x) = O(g(x))$ for all $x \in D$ means that there exists an "implied" constant $C \geqslant 0$ (which often depends on other parameters, which are clearly mentioned) such that $|f(x)| \leqslant Cg(x)$ for all $x \in D$. This definition (for $O(\cdots)$) differs from that of Bourbaki [**2**, Chap. V], which is of topological nature.

We denote by $\mathbf{1}_X$ the characteristic function of a set X. If $\boldsymbol{x} = (x_1, \ldots, x_k)$ is a tuple of real or complex numbers, we write $|\boldsymbol{x}| = x_1 \cdots x_k$.

Acknowledgments. — Thanks to É. Fouvry and H. Iwaniec as well as to R. de la Bretèche, Ph. Michel and P. Nelson, and to all the Polymath8 participants, especially T. Tao. Thanks also to V. Le Dret for reading the first version and correcting a number of typographical mistakes, and to R. Heath-Brown, Y. de Cornulier, B. Green and Y. Motohashi for comments and corrections.