Prime numbers: emergence and victories of bilinear forms decomposition

Olivier Ramaré

Prime numbers: emergence and victories of bilinear forms decomposition

Olivier Ramaré

July 2, 2013

1 Towards a proper question: before 1800

Historical papers on primes often start with a line like "*The quest for the primes has a long history that begins in ancient Greece with Euclid at least twenty-three hundred years ago*". This is fundamentally true.

Or is there a catch?

Let us change the lense to see better: there *are* primes in the ring $\mathbb{Q}[X]$ of polynomials in one variable over $\mathbb{Q}$, though no one asks whether one "*knows*" them. They are, they exist, we have a definition and sound algorithms to recognise them. When pushed further, we may answer: "*yes, there are many of them, infinitely many in fact*". And a question crops in: can you find an irreducible polynomial for any given degree?

This is a way different problem! This question mixes multiplicative properties together with some size questions! It is not about primes as such, but about their sizes. And the problem gets even more entangled in case of integers, for the size structure is closely linked with the addition –

The reader can now understand why questions about primes are often difficult: they are couched in a simple language that hides their difficulty. Take some of the "*observations*" of the early period*:

(1742) Exchanges initiated by Christian Goldbach between Leonhard Euler and himself led, on the 7th of june 1742, to the statement: every even integer ≥ 4 is a sum of two primes.†

(1752) C. Goldbach tells the same L. Euler that every odd integer can be written in the form $p+2a^2$ where p is a prime number and a an integer.‡

(1775) L. Euler wrote that every arithmetic progression starting by 1 contains infinitely many primes.

(1792) Carl Friedrich Gauss gives an argument that shows that there are approximately $x/\log x$ prime numbers below x, when x becomes large.§

(1839) Johann Dirichlet proves (in today's terminology!) that every arithmetic pro-

*Such observations had sometimes the status of theorems, sometimes the status of truth, or maybe I should say "*experimental truth*", since the very notion of proof was shaky at the time.

†Historians discovered later that René Descartes had stated this property some fifty years before. One of the modern prince of arithmetic, Pál Erdös, commented this fact with philosophy: " It is better that the conjecture be named after Goldbach because, mathematically speaking, Descartes was infinitely rich and Goldbach was very poor".

‡He was wrong: 6077 is an exception, but it seems to be the last one!

§This will become the "prime number theorem".

gressions without any constant factor has infinately many primes.

(1845) Joseph Bertand announces that, for each integer $n > 1$, there exist at least one prime p that satisfies $n < p < 2n$. *

(1849) Alphonse de Polignac announces in an equally vague manner that every even integer h is the difference of two primes. The case $h = 2$ is known since Paul Stäckel as the "prime twin conjecture", according to Heinrich Tietze en 1959 [55].

As the reader can check, all these questions mix both the additive and the multiplicative structure. We understand each structure individually perfectly well; But how do they interact? An obvious interaction is given by distributivity: $2a + 2b = 2(a + b)$, which means that if you sum two even numbers (and this latter property belongs to the multiplicative realm), you still get something that has a multiplicative property: it is... even! The question the 20th-century mathematicians endeavoured to settle is: Is this the *only* relation that exists?†

There are many other confusingly simple looking questions, as well as way too many false proofs that keep alimenting the web every year, some by genuine beginners who just missed a step, and some by well-known difficult cases‡ (some of whom sadly occupy academical positions). I hope this paper will help the beginners with mathematical background to understand where the difficulties lie, and where the field is open. The above list contains old questions, but modern times have shown deep ties between modular forms in various senses and more classical problems, in particular via the use of Kloosterman sums. Now classical problems include the evaluation of $\sum_{p\leq X} \lambda(p)$ for cusps forms in the modular case, or the Maass case, or the automorphic case, as well as that of $\sum_{p\leq X} \lambda(p)e^{2i\pi\, p\alpha}$ for any $\alpha \in \mathbb{R}/\mathbb{Z}$.

Recently, many impressive results concerning primes or the Moebius function have been proved and the second aim of this paper is to present a main tool to attack this problems. Indeed, these achievements are of course due to the work of some tenacious individuals, but also emerged after a long toiling from a large community. As an outcome, a general and flexible tool has been created, whose history I will now try to recount. If this tool is now fairly common knowledge among specialists, this does not imply, and by far!, that all the above questions have been answered. This tool is however a good weapon whose conception has reached an evolved enough stage that it should presented to a more general audience. Some of the ideas here may be useful in other contexts and other fields may also contribute; Such a crossing of borders has for instance led ergodic theorists to add their own input, among which the impressive work of Ben Green, Terence Tao, Peter Sarnak, Jean Bourgain and many others.

To be complete and before embarking in my storytelling, I should specify that several other tools have been invented: I am only concentrated on the one that is the most specific to prime numbers.

Now that we have underlined the difficulty of the diverse questions asked, let us turn towards the strategy that has developed to tackle them. We start at the very beginning of this trade: how to handle prime numbers? We make here the first turn: instead of studying the *set* of prime numbers $\mathcal{P}$, we study its *characteristic function* $\mathbb{1}_{\mathcal{P}}$. We further assume a posi-

*This not-so-easy proof is very popular; The down effect being that few know that much more is accessible, and for instance there is a prime p satisfying $2n < p < 3n$ provided $n \geq 2$.

†I am skipping here the fascinating *abc*-conjecture, which is known to hold in the case of $\mathbb{Q}[X]$ and has created some turmoil recently in the case of $\mathbb{Z}$...

‡No, I won't give you my list!

tive (large) real number X be given, and study $\mathbb{1}_{X<\mathcal{P}\leq 2X}$ with takes value 1 on prime numbers p that are such that $X < p \leq 2X$, and 0 otherwise. Studying a set of its characteristic function are of course equivalent, but we are now ready to express $\mathbb{1}_{X<\mathcal{P}\leq 2X}$ as a linear combination of functions, that do not have any special geometrical interpretation.

There has been historically two main lines of approach, that we will see will converge in 1968:

- The first branch of this story can be nicknamed *combinatorial* and will give birth to *the sieve*.
- The second branch, which I call *eulerian*, goes through what is nowadays known as *Dirichlet series*.

In the sequel of this paper, we shall present the characteristics of both approaches, try to point out how they interact in history, and see how they mingled to create the modern theory.

2 Dirichlet and Riemann: the early history

The eulerian approach has been really started by Bernhard Riemann in his 1859 memoir [52], and came to the front scene in 1896: this is the path followed by Jacques Hadamard and à Charles de la Vallée - Poussin [10] to prove the prime number theorem. The next crucial moment will be in 1968 when this method will hybridate with the combinatorial approach, but this is for later! The idea of Leonhard Euler is to consider the decomposition

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} = \prod_{p\geq 2} \left(1 - \frac{1}{p^s}\right)^{-1} \qquad (1)$$

where the variable p ranges the prime numbers. On the right hand side, one find the primes, while on the left hand side, one find only integers: we have potentially a machine to extract information on the primes from information on the integers! L. Euler used it in 1737 to prove that there are infinitely many primes, and, in 1796, C.F. Gauss refined the analysis to guess the prime number Theorem. It is only in 1837-39 that serious proofs started with G. Dirichlet, followed by B. Riemann in 1859. In fact, L. Euler restricted the variable s above to integer values (He even considered the case $s = 1$ and, in some roundabout way, the case of negative values of s as well!). G. Dirichlet applied the logarithm of both side of the equation above to handle the product, and considered $s > 1$ a real number. And shortly after this work, B. Riemann simplified that in his eight pages long epoch making memoir [52] and took the logarithmic derivative of both members:

$$\begin{aligned} -\frac{\zeta'(s)}{\zeta(s)} &= \sum_{p\geq 2} \left(\frac{\operatorname{Log} p}{p^s} + \frac{\operatorname{Log} p}{p^{2s}} + \frac{\operatorname{Log} p}{p^{3s}} + \cdots \right) \\ &= \sum_{n\geq 2} \frac{\Lambda(n)}{n^s} \end{aligned} \qquad (2)$$

where $\Lambda(n)$ will be called the van Mangoldt function... fifty years later! This formula has still the property of the Euler formula: it is potentially a machine to extract knowledge on the primes from the information on the integers. The reader may worry that the Λ-function does not only detect primes but also their powers, but these later are in negligible quantity.

B. Riemann considered complex values of s. The reader may easily guess that an inversion formula, akin to the formula that expresses the Fourier coefficients of a function in terms of this function, using complex analysis links $\sum_{n\leq N} \Lambda(n)$ with $-\zeta'/\zeta$, but to make it work, one needs bounds for this function. The difficulty lies there: the $\zeta(s)$ on the denominator tells us that the *zeroes* of this function are going to give troubles. This is the beginning of

the long and yet unfinished chase for these zeroes!

To make a long story short, when this method applies, it usually gives very precise results. Moreover it has wide generalizations (to number fields, to curves, be they elliptic or not, to modular forms, ...). But the weakness of available informations on the potential zeroes reduces drastically its range. In modern times, computers have entered the course and we are now in a position to check numerically that large but finite regions do not contain any zeroes, but this is material for another paper!

3 Same problem, different tune: the Moebius function

Before introducing the combinatorial approach, let me introduce another player: the Moebius function, named after the german mathematician August Ferdinand Moebius who introduced it in 1832. This player is more discreet than the primes, because it is less geometrical, but it is equally important. Its formal definition reads as follows

$$\mu(n) = \begin{cases} (-1)^r & \text{when } n = p_1 \cdots p_r,\ (p_i \neq p_j), \\ 0 & \text{else.} \end{cases}$$

This function appears in the inclusion-exclusion formula when applied to the divisor set of an integer, and we will see below. It has been noticed, and this was put very formally in theorems by E. Landau in the early 1900's that studying this function is equivalent to studying the prime numbers. However it is often far from obvious to translate a property of the Moebius function into a property of the prime numbers: there are no direct dictionnary between these two worlds. Note that

$$1/\zeta(s) = \sum_{n\geq 1} \mu(n)/n^s. \tag{3}$$

I said before that the difficuly in (2) lies in the denominator. With the Moebius function, we study directly this denominator!

4 The combinatorial approach as seen by Legendre

No more delayong: let us embark in the combinatorial approach that I introduced so long ago!

The starting point is due to Erathosthenes and stems from the following remark: an integer from the interval $(X, 2X]$ is prime if and only if it has no divisor strictly larger than 1 and below $\sqrt{2X}$ (on assuming that $\sqrt{2X} \leq X$, i.e. that X be greater than 2). Erathosthenes deduced from this remark an efficient algorithm to build tables of all the primes below some limit. How to use theoretically this algorithmic efficiency?

Adrien Marie Legendre put this idea in a *formula* in 1808, but failed to turn it into anything efficient (though this miss will be fruitful!). Let us inspect this approach on the problem of counting the number of primes between X and $2X$. We start with the number of integers between these two bounds, that is $X + \mathcal{O}(1)$. From this number, and for each prime $p \leq \sqrt{2X}$, we remove the number of integers that are divisible by p, i.e. we consider

$$X + \mathcal{O}(1) \\ - \frac{X}{2} - \mathcal{O}(1) - \frac{X}{3} - \mathcal{O}(1) - \frac{X}{5} - \mathcal{O}(1) - \cdots$$

Now we have removed twice the integers that are divisible by a product of two primes, so we have to add

$$\frac{X}{6} + \mathcal{O}(1) + \frac{X}{10} + \mathcal{O}(1) + \frac{X}{15} + \mathcal{O}(1) + \cdots$$

This times integers divisible for instance by $2 \times 3 \times 5$ are removed three times (divisible by 2, by 3 and by 5), then added three times

(divisible by 2×3, by 2×5 and by 3×5), so we still need to remove them... The inclusion-exclusion principle is the modern way of rigorously and compactly expressing this idea but we will need later the pedestrian mecanism that I just described. Indeed, the formula of A.-M. Legendre leads to an enormous difficulty: the number of $\mathcal{O}(1)$ that appear is about $2^{\pi(\sqrt{2X})}$, where $\pi(\sqrt{2X})$ is the number of primes below $\sqrt{2X}$. The addition of all these error terms gives rise to a gigantic and final $\mathcal{O}(2^{(1+o(1))\sqrt{2X}/\log\sqrt{2X}})$, swallowing the main term with no second thoughts! One has to face a sheer wall: the combinatorial explosion. This phenomenom is often met in complexity; here it spoils the efficiency of the formula.

But Legendre formula also suffers from a congenital disease: if we ignore the error terms, the supposedly main term it gives is incorrect in view of the prime number theorem*!

5 Two earthquakes: V. Brun, 1919 and I.M. Vinogradov, 1937

The Legendre formula has been abandoned for more than a hundred years, period during which the eulerian approach was developped. In 1910, the young norwegian mathematician Viggo Brun went to Göttingen, at the time one of the main center of mathematics in europe. Edmund Landau was developping and systematizing the use of analysis in number theory, and more often than not, in prime numbers theory. There, the young Brun was introduced to classical problems in this field, among which the Goldbach's conjecture, for instance as part of E. Landau's 1912 ICM list of problems. V. Brun wanted to use combinatorial methods but E. Landau was convinced such methods would never yield anything of interest. E. Landau waited until 1921 before reading V. Brun's memoir! But I am anticipating: let me describe briefly this revolutionary work.

The first idea of V. Brun [5] is to give up hope for an equality in the Legendre formula. Since the error terms accumulate too much, let us stop the process before the end. The pedestrian approach I described gives the principle: after an even number of steps, one has a lower bound, and after an odd number of steps, an upper bound is produced! So, if we aim only at inequalities, the process can be made to work! Well, when counting the number of primes, this lower bound is... negative. But the upper bound is much stronger than what was then known!

The Brun method (later called *the Brun sieve*) is exceptionally flexible and this first work was the source of a wealth of activities. The technical side was however extremely heavy, and one had to go through pages after pages of evaluations.

For instance the Brun sieve [6] gives a sharp upper bound for the number of representations of the even integer N as a sum of two primes. This upper bound is indeed sharp: it is only a multiplicative constant larger than what is expected to be true! This is the one of the main ingredient that Lev Šnirel'man used in 1933 [53] to show the existence of a constant C such that every integer is a sum of at most C prime numbers.

The second sysmic move occured in 1937: I.M. Vinogradov [58] proved that every large enough odd integer is a sum of three primes.† This achievement relies on a magnificent discovery: Vinogradov found a way to deal with

*This "main" term is $\prod_{p\leq\sqrt{2X}}(1-\frac{1}{p})X$ which is equivalent, by one of Mertens Theorems to $2e^{-\gamma}X/\log X$. We have $2e^{-\gamma}=1.122\cdots$ while the prime number theorem will almost a century later show that it should be 1.

†At the time, I.M. Vinogradov was also nearly drowned under the administrative work he had to cope with as director of the Steklov institute!

prime numbers! This method is based of the Brun sieve which was already very intricate and I propose now an anachronical but much clearer way of presenting it.

When working for my thesis, I realised in 1991 [48], [49] that one could consider that the Brun sieve produces a larger sequence $\mathcal{A}$ that contains the sequence of primes, The sequence $\mathcal{A}$ *envelops* the primes, we loose in size, but we gain in control. In functional form, this means that, for any positive function f, we have

$$\sum_{X<p\leq 2X} f(p) \leq \sum_{\substack{a\in\mathcal{A},\\ X<a\leq 2X}} f(a).$$

What is expressed in the above is that the Brun sieve does not only give an upper bound for a counting function, but also provides us with a *local* upper bound! In truth, the situation is somewhat more complicated, since the sequence $\mathcal{A}$ is maybe infinite but only serves as an upper bound for the primes when the variable is between X and $2X$. Attle Selberg in 1947 extended this setting some more: it is enough to find non-negative coefficients $\beta(n)$ such that, for any non-negative function, one has

$$\sum_{X<p\leq 2X} f(p) \leq \sum_{X<n\leq 2X} \beta(n) f(n)$$

where here and everywhere else, the letter p always denotes a prime variable. A. Selberg provides a construction of such good coefficients $\beta(n)$, which one should think of as a (weighted) sequence. It is easier to see on these coefficients β what has been gained by switching to an upper bound. There exists a parameter $D > 1$ (say something like $X^{1/4}$; anyway *strictly* less than X) and coefficients λ_d^* such that

$$\beta(n) = \sum_{\substack{d|n,\\ d\leq X^{1/4}}} \lambda_d^*. \tag{4}$$

The major feature which renders this expression tractable is that D is small enough. Furthermore, this parameter is at our disposal. I developped fully this idea of *envelopping sieve* in [51], but let us go back to I.M. Vinogradov. He writes

$$\mathbb{1}_{X<\mathcal{P}\leq 2X} = \mathbb{1}_{\mathcal{A}} - \Theta.$$

Nothing has been done so far. I.M. Vinogradov crucial observation is that Θ has the a special shape, namely

$$\Theta(n) = \text{lin. comb. of } \sum_{\ell m=n} a_\ell^{(i)} b_m^{(i)} \tag{5}$$

where the sequences $(a_\ell^{(i)})$ and $(b_m^{(i)})$ vanish as soon as m or ℓ is either too large or too small; since we have constrained n by $X < n \leq 2X$, if $a_\ell^{(i)} = 0$ as soon as $\ell \leq L$, then the m's with $m \geq 2X/L$ do not intervene, so our conditions are somewhat redundant. In practive, we will ensure that both $a_\ell^{(i)}$ and $b_m^{(i)}$ vanish when ℓ (resp. m) is small. This observation is a main turning point. I.M. Vinogradov termed *type I sums* the sums arising from (4) and *type II sums* the sums arising from (5). I prefer with many to speak of *bilinear sums* for (5) and, of course, to call (4) a *linear sum*! More will be said later on this bilinear structure.

This step being crucial, let me enunciate a simple lemma that shows the power of this bilinear structure.

Lemma 1 (Toy lemma)**.** *Given q, $L \geq q$, $M \geq q^2$ and two sequences $|a_\ell|, |b_m| \leq 1$,*

$$\left| \sum_{\substack{\ell\leq L,\\ m\leq M}} a_\ell b_m \, e^{2i\pi \, \ell m/q} \right| \leq 2LM/\sqrt{q}.$$

The condition $|b_m| \leq 1$ can be relaxed if the upper bound is replaced by $2L\sqrt{M\sum_m |b_m|^2}/\sqrt{q}$. The bound L can be replaced by $L(m) \leq L$ depending on m, provided that, given ℓ, the set of m such that $\ell \leq L(m)$ is an interval.

The last condition is typically met by conditions like $\ell \leq X/m$ for some X. An early general version of the toy lemma is to be found in [58, Lemma 4], see also [9, Lemma 8]. The proof goes simply by writing the sum to be studied in the form $\sum_{m\leq M} b_m c(m)$ and using Cauchy's inequality. In the resulting sum $\sum_{m\leq M} |c(m)|^2$, open the square and invert summations, the result will follow readily. What has been gained here? If one sets $\gamma(n) = \sum_{\ell m=n} a_\ell b_m$ (ensure that $a_\ell = 0$ when $\ell > L$ and similarly for b_m and m), we see that we are studying

$$\Big| \sum_{n\leq N=LM} \gamma_n \, e^{2i\pi \, n/q} \Big|.$$

When the only information we have on γ_n is that it is bounded above in absolute value by 1, the best possible upper bound is $\sum_{n\leq N} |\gamma_n|$ which can be as large as $N = LM$. The above lemma uses the structure and saves a factor $2/\sqrt{q}$! Note the discreet conditions $L \geq q$ and $M \geq q^2$ that are in fact essential.

Here is what A. Ingham wrote in Zentralblatt on I.M. Vinogradov paper:

> *"This is a fairly simple deduction from Cauchy's inequality, and the essential basis of the result has been available since 1910. It is hardly surprising, however, that its possibilities remainded so long unsuspected. For double sums occuring in (1) do not appear naturally in the known treatments of the above problems, and in any case a straightforward application is liable to give only crude results owing to the loss involved in the use of Cauch's inequality. It is, in fact, in devising ways of adapting the lemma to the various problems, and in elaborating techniques for bringing it to a successful conclusion, (...), that the author reveals his amazing powers."*

This is a flexible and powerful principle. Our presentation is voluntarily naive; modern versions rely heavily on the underlying bilinear structure, and on Bessel type inequalities for the relevant quasi-orthogonal system. I'm ahead of the subject, let us go back to the beginning of the previous century!

6 The eulerian approach continued

While sieves and their derivatives occupied the front scene, the eulerian approach was still under scrutiny. The main hurdle being the lack of information of the zeroes, Harald Bohr & Edmund Landau [2] decided in 1914 (somewhat before Brun's discovery) to look for regions that do not have *many* zeroes*. What they showed is a *density estimate*; for any $\theta > 1/2$, we have

$$\frac{\#\{\rho = \beta + i\gamma, \zeta(\rho) = 0, |\gamma| \leq T, \beta > \theta\}}{\#\{\rho = \beta + i\gamma, \zeta(\rho) = 0, |\gamma| \leq T, \beta > 0\}} \to 0.$$

This statement says that most of the zeroes have a real part $\leq \theta$, for any $\theta > 1/2$ (and in fact, by the functional equation, almost all zeroes with positive real part have a real part close to $1/2$). The Riemann hypothesis states that all these zeroes have indeed a real part *equal* to $1/2$; the above statement is a statistical step in this direction. And this statistical step happened to have been a crucial one, since it started a very fecund branch of investigation that delivered new results for the next eighty years (this theory is now somewhat stalled).

H. Bohr & E. Landau studied the function $(1-\zeta(s)P_D(s))^2$ where P is a finite Euler product: $P_D(s) = \prod_{p\leq D}(1-p^{-s})^{-1}$. This is one

*By the way, Harald Bohr was the rising danish star while V. Brun was the two years older norwegian rising star. By the way again, Harald Bohr was also an accomplished football player and is, together with his teammates, responsible for the sharpest defeat of the french national team (17 to 1!).

of the striking feature of analytic functions: it is possible de bound from above their number of zeroes by bounding from above some integral containing them. Here, H. Bohr & E. Landau integrated $|1-\zeta(s)P_D(s)|^2$ on a square and proved this quantity to be a multiplicative constant times larger than the number of zeroes in a smaller region, a process later improved upon by John Edensor Littlewood.

The suedish mathematician Fritz Carlson in 1920 [8] simply replaced the product P_D by a sum $M_D(s) = \sum_{d\leq D} \mu(d)/d^s$ and considered $(1-\zeta M_D)^2$. He obtained in this manner much better bounds for the number of zeroes with real parts $> \theta$. Many authors continued this line of work, J.E. Littlewood, Alan Titchmarsh, Albert Ingham, Pál Turán, Atle Selberg, Askold Vinogradov, Enrico Bombieri, to name but a few! One uses since A.I. Vinogradov a *zero detection method* instead of the method described above but this is not my subject here! It is still relevant to note that hermitian methods became more and more important and the large inequality (see (13) below) proved to be an essential tool. Yu Linnik is the pioneer of line of investigation [39], that became understood as the use of Bessel type inequalities for quasi-orthogonal systems... More will be said on this subject later!

There are two main highpoints of the theory of density estimates: Y. Linnik theorem concerning primes in arithmetic progressions in 1944 and the Bombieri-Vinogradov theorem in 1965. In essence, these authors prove statements concerning density of zeroes, and this is the major part of their work. These statements are then converted in results concerning the prime numbers, via some adhoc explicit formula.

The Linnik theorem [40], [41] says that there exists two constants C_0 and q_0 so that, when q is larger than q_0, there exists for each residue class a modulo q, a prime number congruent to a modulo q and below q^{C_0}.

The Bombieri-Vinogradov theorem [3], [57]* says that for each positive constant A, we have

$$\sum_{q\leq \frac{\sqrt{x}}{(\log x)^{A+4}}} \max_{\substack{1\leq a\leq q,\\ \gcd(a,q)=1}} \Big|\Delta(X;q,a)\Big| \ll \frac{x}{(\log x)^A},$$

$$\text{where } \Delta(X;q,a) = \sum_{\substack{p\leq X,\\ p\equiv a[q]}} \log p - \frac{x}{\varphi(q)}. \quad (6)$$

This theorem can be seen as a statistical Generalized Rieman Hypothesis and serves in many situation as a replacement. And we should also recall the Guido Hoheisel theorem from 1930 [33]†: there exists two constants $X_0 \geq 1$ and $\delta_2 \in (0,1)$ such that every interval $[X, x+X^{\delta_2}]$ contains at least a prime number when $X \geq X_0$. This proof created some turmoil (check) when it was published, as the existence of such a $\delta_2 < 1$ was only known under the hypothesis than no zero with real part $\geq \delta_2$ existed (to be precise, a slightly stronger hypothesis is needed) and seemed close to be equivalent to it. Nowadays, we term a *Hoheisel theorem* any theorem that proves a similar statement with some definite value of δ_2. The initial value provided by Hoheisel was very close to 1‡.

7 The '68 generation

Patrick Gallagher [19] remarked in 1968 that the process is abnormaly convoluted: the proof starts from the series $\sum 1/n^s$, retrieves in some fashion informations on its zeroes and deduces

*The russian mathematician Askold Ivanovich Vinogradov is not to be confused with the other russian mathematician, mathematical great-grandchild of Pafnuty Lvovich Chebyshev, and whose work is at the heart of this paper: Ivan Matveyevich Vinogradov

†G. Hoheisel was a mathematical grandchild of David Hilbert.

‡Any $\delta_2 > 1 - (1/33\,000)$ would do. This is real small!

from that informations on the primes. Why not use a shortcut and skip the zeroes? This is easier said than done, but P.X. Gallagher found such a shortcut. He simply multiplied $-\zeta'/\zeta$ by the kernel used for density estimates! This amounts to writing

$$\frac{\zeta'}{\zeta} = 2M_D\zeta' - M_D^2\zeta'\zeta + \frac{\zeta'}{\zeta}(1-\zeta M_D)^2 \quad (7)$$

where M_D is the Dirichlet polynomial we have defined above. The difficult term is the last one: it is the only one that still has a denominator. But this term is essentially of the shape identified by I.M. Vinogradov! A short dictionnary is called for here: one has to convert operations on Dirichlet series on operations on their coefficients. The main rule is that the (arithmetic) convolution is trivialised when expressed in terms of Dirichlet series, i.e.:

$$\gamma_n = \sum_{\ell m=n} a_\ell b_m \Leftrightarrow \sum_{n\geq 1}\frac{\gamma_n}{n^s} = \sum_{\ell\geq 1}\frac{a_\ell}{\ell^s}\sum_{m\geq 1}\frac{b_m}{m^s}.$$

This equality is either formal, either valid in the domain of absolute convergence of the three series (lighter hypotheses are possible!). I have to point out specifically that this implies that

$$\zeta(s)M_D(s) - 1 = \sum_{n>D}\frac{v_n}{n^s}.$$

The coefficients v_n are not important, as we noted earlier; they are bounded by a divisor function*. What is really important is that the variable d *cannot* be small.

The first two terms in (7) give rise to linear sums while the last one gives rise to a bilinear one. This transformation has been completed by Robert Vaughan in 1975 [56] by removing a finite polynomial to ζ'/ζ; he introduced, for some parameter y typically like $X^{1/4}$, the finite Dirichlet polynomial

$$F_y(s) = \sum_{n\leq y}\Lambda(n)/n^s$$

so that

$$\frac{\zeta'(s)}{\zeta(s)} + F_y(s) = \sum_{n>y}\frac{-\Lambda(n)}{n^s}.$$

Now, multiply together $(\zeta'/\zeta) + F_y$ and $(1-\zeta M_D)$ and expand as above. This product is a Dirichlet series whose coefficients are a convolution product of two sequences that both vanish when the variable is small: it is indeed of the special shape highlighted by I.M. Vinogradov!

What has been gained in the process?

- The sieve part in I.M. Vinogradov process was *not* in most cases the main term, while the linear part is *expected* to carry the main term. We will say more on this point later.

- The method is simple and flexible: one can change the kernel; it applies to other functions instead of the Rieman zeta functions, like the Dedekind zeta functions or Hecke L-series. Note that the multiplicativity is essential, but not the functional equation. The method applies also the Moebius function, but so did I.M. Vinogradov method as already noticed by Harald Davenport [9]. However H. Davenport reduced the problem to the case of primes while the present proof is direct.

On using this approach, P.X. Gallagher obtained in 1970 [20] a major theorem that unifies the Linnik and the Bombieri-Vinogradov theorem: the Gallagher prime number theorem which is still unsurpassed in strength.

Since there has been recently a flourish of works on the Moebius function, and since many people asked how this theory handles this case,

*We have $v_n = \sum_{\substack{d|n,\\ d\leq D}}\mu(d)$ when $n\geq 2$ and 0 otherwise. And thus $|v_n|\leq d(n)$.

let me be more precise here. The identity I propose to use is a simplification of one I devised recently, as explained later in this survey. It relies on the simpler kernel $(1-\zeta M_D)$. We consider the identity

$$\frac{1}{\zeta} = \Big(\frac{1}{\zeta} - M_D\Big)(1-\zeta M_D) + 2M_D - \zeta M_D^2.$$

There only remains to identify the coefficients! It is best for applications to express the result in functional form. For any function f and provided $D \le X$, we have:

$$\sum_{X<n\le 2X} f(n)\mu(n) = \sum_{m\le D^2} u_m \sum_{\substack{\ell\ge 1,\\ X<\ell m\le 2X}} f(\ell m) + \sum_{\substack{\ell>D, m>D,\\ X<\ell m\le 2X}} \mu(\ell) v_m f(\ell m) \quad (8)$$

where v_m has been defined above and where

$$u_m = -\sum_{\substack{hk=m,\\ h,k\le D}} \mu(h)\mu(k).$$

In the first summation on the right-hand side of (8), we hope to be able to evaluate the summation over ℓ. For instance, in the toy lemma case, one selects $f(x) = \exp(2i\pi x/q)$ and the sum over ℓ is bounded by q, giving rise to a total contribution bounded by $\mathcal{O}(qD^2)$. Concerning the second sum, we first note that the variable m ranges $(D, 2X/D)$. We cover this interval by at most $(\log(2X/D^2)/\log 2$ disjoint intervals of the shape $(M, M']$ for some $M' \le 2M$. Our toy lemma applies provided that $D \ge q^2$. We note that $\sum_{m\in[M,M']} |v_m|^2 \ll M(\log M)^3$. Collecting our estimates, we have proved that

$$\sum_{X<n\le 2X} \mu(n) e^{2i\pi\, n/q} \ll qD^2 + (\log X)^{5/2}\frac{X}{\sqrt{q}}$$

provided $D \ge q^2$ and $D < X$. On selecting $D = q^2$ and assuming that $q \le X^{2/7}$, this gives our case study result:

$$\sum_{X<n\le 2X} \mu(n) e^{2i\pi\, n/q} \ll (\log X)^{5/2}\frac{X}{\sqrt{q}}. \quad (9)$$

This simple result is way beyond the power of the classical eulerian approach! But the proof we gave requires not more than half a page!

There has been recently a renewed activity around the Moebius function, as in [24] and [25] and around a conjecture due to Peter Sarnak*. This subject is somewhat off our main road, though we have to specify that getting to the Moebius function is done as above. Recently Jean Bourgain, Peter Sarnak & Tamar Ziegler have given [4] another way to handle the Moebius function that follows a very combinatorial path closer to the I.M. Vinogradov one.

In short, we have reached a point where the eulerian approach derived sufficiently to ressemble the combinatorial one! In both cases, the idea is to represent the characteristic function of the primes as a linear combination of linear forms and of bilinear forms. This idea is one of the main ingrédient of Christian Mauduit & Joël Rivat [43] in their proof of the forty years old conjecture of Gelfand: there exists up to an error term as many prime numbers whose sum of digits in base 2 is odd or even. It is at the heart of the proof of Terence Tao [54] that every odd integer $\neq 1$ is a sum of at most five primes, and also at the heart of Harald Helfgott's proof that every odd integer $\neq 1$ [31], [32] is a sum of at most three primes.† The second paper has a final result better than the first one, of course, but T. Tao's paper develops ideas around small intervals containing sums of two primes that are of independant interest.

*See also [34, (13.7)]. P. Sarnak's conjecture somehow quantifies this statement. See also [23].

†Both papers have been submitted, there are good reasons to believe in their solidity, but rules are rules and the checking should be complete before the result be fully accepted!

H.A. Helfgott closes after about seventy-five years the proof of I.M. Vinogradov: we knew that the statement was true for large enough integers, and large enough meant real large, and bringing this bound down was no small achievement.

8 Sad news: there are limitations! The main term problem

Let us resume our general analysis. The problem addressed in this section: in the initial Vinogradov method, the sieve part did not yield the main term. To understand properly why, here is a simplified presentation of the Vinogradov method I developped some years back. In the Brun sieve, the sequence $\mathcal{A}$ is the sequence of integers that do not have any prime factors less than some given bound z. This parameter is typically between a high power of $\log X$ and X to a power that tends very slowly to 0. To reach the primes from the interval $(X, 2X]$, we still need to remove all the integers that have a prime factor between z and $\sqrt{2X}$. Say that p is such a prime. The bad candidates have thus the form $pm...$ and this is bilinear! Well almost, but not quite: m has to be required to have no prime factors below p if we want the representation pm to be unique, and this ties p and m together... Before continuing, let me precise that the process used is known as the Buchstab iteration [7]. I learned recently while reading the notes of [47] that, in the late seventies, Hans-Egon Richert had performed a similar analysis from the Selberg sieve.

The problem encountered is well identified: to get a proper bilinear form, one needs to *separate* both variables. This problem is serious but often not deadly. One can introduce here the number $\omega_{z,\sqrt{2X}}(m)$ of prime factors of m that lies within $(z, \sqrt{2X}]$ and the representation pm has multiplicity $\omega^*(pm)$ say, so it is enough to divide by this number. Though now p and m are tied in $\omega^*(pm)$! Well, yes, but less so. For most m's, i.e. for the ones that are not divisible by p, we have $\omega^*(pm) = 1 + \omega^*(m)$; the other ones correspond to integers of the shape p^2k, and since p is large enough, they are in a (usually) negligible quantity.

Since the sequence $\mathcal{A}$ that comes from the Brun sieve is larger than the primes, it leads to a larger main term! Hence what we treat like an error term contains in fact part of the main term. In the linear/bilinear approach, the linear part can in usual problems be shown to have the proper size, at least if believed conjectures do hold. But *the way* the bilinear form is treated induces a loss of precision that can be deadly! On our toy problem for instance, (9) is way less than what is expected, namely at least:

$$\sum_{X<n\leq 2X} \mu(n) e^{2i\pi\, n/q} \ll (\log X)^{100} \frac{X}{q} \qquad (10)$$

for $q \leq X^{2/7}$. But, if we were to prove such a statement, we would prove that there are no Siegel zero, or equivalently that the class number of the imaginary quadratic field $\mathbb{Q}(\sqrt{-q})$ is at least $\gg q^{1/2}/(\log q)^{200}$. In fact "simply" improving the power of q from $1/2$ to $(1/2)+\delta$ for any $\delta > 0$ would be a major achievement. We will see below more as to where this limitation comes from. The best result to date [50, Corollary 5]* reads

$$\sum_{X<n\leq 2X} \mu(n) e^{2i\pi\, n/q} \ll \frac{X}{\sqrt{q}} \prod_{p|q} \Big(1 + \frac{1}{\sqrt{p}}\Big) \qquad (11)$$

for $q \leq X^{1/9}$. The last product is just an annoying blemish. In the case of a prime modulus q, proving that the implied constant is < 1 would prove that there are no Siegel zeroes!

*As I said, rules are rules and the result I mention now has only been submitted!

John Friedlander & Henryk Iwaniec managed in their awesome work [17], [18] to overcome in some delicate cases this enormous difficulty.

There is another major limitation. I have presented the toy lemma with an additive character $e^{2i\pi\, n/q}$, but what happens with a multiplicative character? The bilinear forms becomes trivial and nothing can be gained anymore. A path could be to express these multiplicative characters, modulo q say, in terms of additive ones modulo q. Such a process loses $\sqrt{q}$ due to the size of the Gauss sum, reducing to nil the saving acquired!

9 Other identities and divisors: the philosophy extends

I said several times that this method offers flexibility, but the reader has seen only two identities so far. In his state thesis in 1980, Étienne Fouvry used recursively Vaughan's identity*. At the time of writing up the corresponding paper [14], Roger Heath-Brown had published in [29] and [28] a systematized version which E. Fouvry prefered to use. This systematized version consists in selecting the kernel $(1-\zeta M_D)^k$ for an integer parameter k to be chosen (É. Fouvry's took $k = 12$ for instance, later reduced to $k = 7$).

In 1961, Y. Linnik produced [42] another kind of identity by considering $\log\zeta = \log(1-(1-\zeta))$ together with the Taylor expansion of $\log(1-z)$ around $z = 1$. We present the modification due to Heath-Brown in [29, Lemma 3] in which the zeta function in multiplied by the finite Euler product P_D introduced above in F. Carlson's proof. The function $\log(\zeta P_D)$ is also the Dirichlet series $\sum^*_{n\geq 2} \frac{\Lambda(n)}{\log n}\frac{1}{n^s}$ where the star means that n has only prime factors $> D$; it is expedient here to introduce the product Π_D of all the primes not more than D. The condition on n is then simply that n and D are coprime. On the other side the function $(1-\zeta P_D)^k$ has for any positive integer k the Dirichlet series representation $(-1)^k \sum^*_{n\geq 1} d_k^\sharp(n)/n^s$ where the summation is again restricted to integers n prime to Π_D and where $d_k^\sharp(n)$ is the number of ways of writing the integer n as a product of k integers, all *strictly larger than 1.* We get, when n is prime to Π_D,

$$\frac{\Lambda(n)}{\log n} = \sum_{k\geq 1} \frac{(-1)^{k+1}}{k} d_k^\sharp(n). \qquad (12)$$

On restricting n to the range $(X, 2X]$ and assuming that $D^{K+1} > 2X$, the summation above can be truncated at $k \leq K$. Moreover, if we abort the summation at an odd (resp. even) number of steps, we get an upper (resp. lower) bound, as in the inclusion-exclusion principle!

The reader will also find a family of identities in [17, section 3].

But a closer look at Linnik's identity is called for: *it transposes problems for primes in problems for divisor functions.* This is what transpires from É. Fouvry's work [14]: if one knows well enough the divisor functions, up to products of six divisors, then this implies an improved Bombieri-Vinogradov theorem of the primes: the inequality $q \leq \sqrt{x}/(\log x)^{A+4}$ could be replaced by $q \leq x^{\frac{1}{2}+\delta}$ for some positive δ. Such a theorem would be *stronger* than the Generalized Rieman Hypothesis! The reader can see rapidly how modular forms come into play here: the distribution of the divisor function is linked with the distribution of Kloosterman sums, which are in turn coefficients of modular forms.

This feature of Y. Linnik's identity can be found again in R. Heath-Brown identity if one forces k to be so large that $D^k > X$. In this manner, the bilinear part does not come into

*He iterated it twelve times! The formulas obtained were so long that he printed them in landscape format...

play for integers below X! In functional form, this reads:

$$\sum_{n\leq X}\Lambda(n)f(n)=\sum_{1\leq r\leq k}\binom{k}{r}(-1)^{r+1}\times \sum_{\substack{n_1,\cdots,n_r\leq D,\\ n_{r+1},\cdots,n_{2r},\\ n_1\cdots n_{2r}\leq X}}\mu(n_1)\cdots\mu(n_r)\log n_{2r}\, f(n_1\cdots n_{2r}).$$

Well and good, but we have already difficulties to treat products of three divisors (see the groundbreaking [35]) not to speak of products of four of them, so it may be more efficient to consider these divisor functions simply as convolution products and resort to I.M. Vinogradov bilinear form approach. In the above, one may tie some variables together, say d_1 and d_2, in a single $m=d_1d_2$ affected with the coefficient u_m defined above!

The divisor angle can however be made to bear with more efficiency if one aims at a result weaker than a Bombieri-Vinogradov Theorem. É. Fouvry put in [15] this philosophy in practice: the quantity considered is, for some fixed a and some positive δ:

$$\sum_{q\leq x^{\frac{1}{2}+\delta}}c(q)\Delta(X;q,a)$$

where the weights $c(q)$ are fairly general and, yet again, convolution products of special kinds (Δ is defined in (6)). This is at the heart of the recent breakthrough of Ytang Zhang [59]: there exists infinitely many pairs of primes p and p' such that $|p-p'|\leq 7\cdot 10^7$. The argument follows the pathway opened by in 2006 by Daniel Goldston, Janós Pintz et Cem Yıldırım [22], [21], but the main novelty comes from the treatment of the error term. Or more precisely in curbing the proof so that it produces an error term of a special form, as already noted by A. Ingham in his assessment of I.M. Vinogradov's work. Studying this error term is also no small task! Let us note that this entails controling bilinear terms of the form we have already seen but also some convolution of three divisors; Or a three-linear form; or, as Y. Zhang puts it: *a type III sum.*

10 The combinatorial approach, revival time

While the work on identities has been going strong, a different line continued from the I.M. Vinogradov approach. We have seen that the correcting term from the sieve part contained part of the main term and that a coarse treatment via Cauchy's inequality was not enough. Some authors however developped an gentler treatment is some cases; such a line started in [30] where the authors obtained a strong improvement on the Hoheisel theorem (any $\delta_2 > 11/20$ is accessible; compare with Hoheisel initial value!). Combinatorial ideas are put to effect and show their teeth! This has been amplified, developped and refined by Glyn Harman in several papers [26], [27], [1] in what this author calls his *adaptative sieve.* This is surely a very accomplished work in this direction and it leads to the best results in many problems (like $\delta_2 = 0.525$).

In this section I should mention the development in [11, Section 6], and in particular Theorem S thereof. This subtle theorem ensures that the sequence of primes will be properly distributed in some sequences, provided one knows how to bound some linear sums as well as some bilinear ones. This step is extremely difficult in their case of application, as the accessible information is not enough for a usual approach! A special combinatorial treatment is required which is contained in the Theorem S I mentioned, in particular for handling products of three divisors of about the same size.* This falls within the general philosophy we have de-

*For the reader who would want to follow this proof in a gentler manner, and who can read french, I recommend the book [37].

velopped so far. The novelty here is that the bilinear form arising from I.M. Vinogradov approach is treated with more care, and the main term extracted from it.

11 Treating the bilinear forms

I have told you at length that a bilinear structure was involved in all these representations, whether directly via eulerian identities, or after more work via combinatorial means (and now, both methods mix happily!), but the way to treat this bilinear part has up to now remained extremely coarse, essentially via the toy lemma above. Even at this level of coarseness, the method yields impressive results, but a better understanding is called for. And it will show again how the sieves ideas and the eulerian approach mingle together. In most of the problems on primes, the treatment of the bilinear sum is the most difficult part, and in return, what we are able to prove at this level conditions the kind of identity one has to prove or choose.

One way to start telling this part of the story starts from the Bombieri-Vinogradov theorem. In the initial proof, the one that dealt with density estimates of zeroes, a major role was played by an inequality that finds its origin in the work of Y. Linnik: the large sieve inequality. It was later discovered by Hugh Montgomery [44] that this inequality could be used in sieve context and... led to results as strong to the Selberg sieve! In some sense, this inequality is *dual* to the Selberg sieve [36] and this notion of duality has to be understood in the usual sense, i.e. when a bilinear coupling is at stake. Let me state a special case of this inequality in the strong form given by H.L. Montgomery & R.C. Vaughan [45], and at the same time by A. Selberg with a different proof:

$$\sum_{q\leq Q}\sum_{\substack{1\leq a\leq q,\\ \gcd(a,q)=1}}\Big|\sum_{n\leq X} b_n e^{2i\pi\, na/q}\Big|^2 \leq \sum_{n\leq X}|b_n|^2(X+Q^2), \quad (13)$$

valid for any sequence of complex numbers (b_n). Such an inequality is of course reminiscent of our toy lemma above. It should be looked upon as a Bessel type inequality for a quasi-orthogonal system. From a practical viewpoint, if we were given any single sum above, say $\sum_{n\leq X} b_n e^{2i\pi\, na_0/q_0}$, Cauchy's inequality would give us the bound $\sqrt{\sum_{n\leq X}|b_n|^2 X}$ and it is best possible at this level of generality. The above inequality tells us that, for the same price, we can bound many more sums! Provided Q^2 be less than X, which will be our case of use. So the idea is to put as much as we can in the left hand side and use this bound. The reader will not be surprised to find an inequality of this type in H.A. Helfgott's work.

When compared with our toy lemma result, the reader may worry about the missing $1/\sqrt{q}$... And rightly so! But we above have a summation of length about q over a modulo q: in the toy lemma, simply split the variable ℓ according to its residue class modulo q. There remains a slight difficulty, as ℓ is not guaranted to be coprime with q, but this hurdle is easily overcomed.

This principle can be pushed very far and many more sums incorporated in the left-hand side! To prove (11), and elaborating on unpublished material by A. Selberg in 1972-73, and of Yoishi Motohashi [46], I developped in [50] an quasi-orthogonal family of identities for the primes, where the polynomial M_D is replaced by a family $M_D^{(r)}$. One of the first lemma of the proof is the following (version of an) inequality due to Y. Motohashi [46, Lemma 3], with

$R = \sqrt{N/T}$:

$$\sum_{\substack{r\le R/q,\\ \gcd(r,q)=1}} \frac{1}{\varphi(r)} \sum_{\substack{1\le a\le q,\\ \gcd(a,q)=1}} \int_{-T}^{T} \Big| \sum_{n\le N} \frac{b_n c_r(n)}{n^{it}} e^{2i\pi \frac{na}{q}} \Big|^2 dt \ll X \sum_{n\le X} |b_n|^2$$

provided that b_n vanishes as soon as n has a factor in common with q. Here $c_r(n)$ is the Ramanujan sum. The reader should not be scared by such an inequality, for it is a gentle monster! If we take $t = 0$, $a = 1$ and $r = 1$, the inner sum is simply $\sum_{n\le N} b_n e^{2i\pi\, n/q}$ as in the toy lemma case. But for the same price, we have added an integration over t in a large range (this part is classical since P.X. Gallagher [20, Theorem 3]), as well as a summation over r which comes from A. Selberg.

The inequality above says that the three families of "characters" $(c_r(n))_r$, $(n^{it})_t$ and $(e^{2i\pi na/q})_a$ are quasi orthogonal in themselves but also when mixed one with the other.

How can one put that in practice? That's more easily said than done, but here are some hints: when p is a prime number prime to r, the Ramanujan function of order r takes value -1 at p, i.e. $c_r(p) = -1$. As a consequence, when $r \le X$, and for any function f, we have

$$\sum_{X<p\le 2X} f(p) = -\sum_{X<p\le 2X} f(p) c_r(p).$$

We can use this fact to introduce an average over r, and for instance, for any non-negative function g, we find that

$$\Big| \sum_{X<p\le 2X} f(p) \Big|^2 = \frac{\sum_{r\le X} g(r) \Big| \sum_{X<p\le 2X} f(p) c_r(p) \Big|^2}{\sum_{r\le X} g(r)}.$$

On using bilinear form representation, this $c_r(p)$ will become a $c_r(\ell m)$ and if we can prove a proper Bessel inequality, only one term on the right hand side will contribute: we will save the denominator! A similar process is used in [13]. See also [12] and [38] for more comments on *amplification* techniques.

The summation over r can be regained by the process above, but a similar process does not apply to the other "characters", and this is where the limitation comes from: we do not know that the sequence (b_n) does not conspire with some $e^{i\pi na_0/q}/n^{it_0}$ for instance to give rise to a large contribution. We would be surprised if this were to happen, of course, but at this level, we do not know how to eliminate this possibility. We say in short that the *diagonal* contribution matters most. There has been a good amount of work to try to dispense with it. The general theme is to go back to the proof of the large sieve type inequalities we use: in these proofs, the Fourier transform has a important role, very often in conjunction with Poisson summation formula or spectral theory (as for instance in [11]). So the idea is to introduce such a Fourier transform where a smooth variable occurs and use the Poisson summation formula. This is for instance what is used [16]. The Linnik dispersion method [42] is another similar L^2-mecanism that (see for instance [14, Section 3]) eliminates the diagonal contribution.

12 This is *not* the end!

I hope the reader has now a proper idea of the flexible tool I praised so much in the introduction! There remains large parts of unexplored territory, as well as some peaks in the distance... A.-M. Legendre asked long ago whether every interval $(N^2, (N+1)^2)$ contains a prime when N is a positive integer. This is roughly equivalent to showing that an interval

of length $\sqrt{x}$ around x contains a prime: the methods we have give $x^{0.525}$ but, even after all the plausible refinements, reaching $x^{0.5}$ will require a novel input. The sum $\sum_{p\leq X}\mu(p-1)$ is still a mysterious entity, and there are many other problems on primes that are unsolved, at present time just out of our grasp, but who knows what will happen tomorrow!

References

[1] R. Baker, G. Harman, and J. Pintz. The difference between consecutive primes, III. *Proc. London Math. Soc.*, 83(3):532–562, 2001.

[2] H. Bohr and E. Landau. Sur les zéros de la fonction $\zeta(s)$ de *Riemann. C. R.*, 158:106–110, 1914.

[3] E. Bombieri. On the large sieve method. *Mathematika*, 12:201–225, 1965.

[4] J. Bourgain, P. Sarnak, and T. Ziegler. Distjointness of Moebius from horocycle flows. page 17pp, 2012. arXiv:1110.0992.

[5] V. Brun. La série $\frac{1}{5}+\frac{1}{7}+\frac{1}{11}+\frac{1}{13}+\frac{1}{17}+\frac{1}{19}+\frac{1}{29}+\frac{1}{31}+\frac{1}{41}+\frac{1}{43}+\frac{1}{59}+\frac{1}{61}+\cdots$ où les dénominateurs sont "nombres premiers jumeaux" est convergente ou finie. *Darboux Bull.*, 43(2):100–104, 124–128, 1919.

[6] V. Brun. Le crible d'Erathosthène et le théorème de Goldbach. *C.R.*, 168:544–546, 1919.

[7] A.A. Buchstab. On a relation for the function $\pi(x)$ expressing the number of primes that do not exceed x. *Rec. Math. [Mat. Sbornik] N.S.*, 12 (54):152–160, 1943.

[8] F. Carlson. Sur les zéros des séries de Dirichlet. *C. R.*, 171:339–341, 1920. `http://gallica.bnf.fr/ark:/12148/bpt6k3124m.f339`.

[9] H. Davenport. On some infinite series involving arithmetical functions. II. *Quart. J. Math., Oxf. Ser.*, 8:313–320, 1937.

[10] Ch. de la Vallée-Poussin. Sur la fonction $\zeta(s)$ de Riemann et le nombre des nombres premiers inférieurs à une limite donnée. *Belg. Mém. cour. in* 8°, LIX:74pp, 1899.

[11] W. Duke, J.B. Friedlander, and H. Iwaniec. Equidistribution of roots of a quadratic congruence to prime moduli. *Ann. of Math. (2)*, 141(2):423–441, 1995.

[12] W. Duke, J.B. Friedlander, and H. Iwaniec. Representations by the determinant and mean values of L-functions. In *Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995)*, volume 237 of *London Math. Soc. Lecture Note Ser.*, pages 109–115. Cambridge Univ. Press, Cambridge, 1997.

[13] P.D.T.A. Elliott. On maximal variants of the Large Sieve. II. *J. Fac. Sci. Univ. Tokyo, Sect. IA*, 39(2):379–383, 1992.

[14] É. Fouvry. Autour du théorème de Bombieri-Vinogradov. *Acta Math.*, 152:219–244, 1984.

[15] É. Fouvry. Autour du théorème de Bombieri-Vinogradov. II. *Ann. Sci. École Norm. Sup. (4)*, 20(4):617–640, 1987.

[16] J. Friedlander and H. Iwaniec. A mean-value theorem for character sums. *Mich. Math. J.*, 39(1):153–159, 1992.

[17] J. Friedlander and H. Iwaniec. Asymptotic sieve for primes. *Ann. of Math. (2)*, 148(3):1041–1065, 1998.

[18] J. Friedlander and H. Iwaniec. The polynomial X^2+Y^4 captures its primes. *Ann. of Math. (2)*, 148(3):945–1040, 1998.

[19] P.X. Gallagher. Bombieri's mean value theorem. *Mathematika*, 15:1–6, 1968.

[20] P.X. Gallagher. A large sieve density estimate near $\sigma=1$. *Invent. Math.*, 11:329–339, 1970.

[21] D.A. Goldston, János J. Pintz, and C.Y. Yıldırım. Primes in tuples. III. On the difference $p_{n+\nu}-p_n$. *Funct. Approx. Comment. Math.*, 35:79–89, 2006.

[22] D.A. Goldston, J. Pintz, and C.Y. Yıldırım. Primes in Tuples I. *Ann. of Math.*, (to appear):36p, 2005. available at arxiv under reference math.NT/0508185.

[23] B. Green. On (not) computing the Möbius function using bounded depth circuits. *Combin. Probab. Comput.*, 21(6):942–951, 2012.

[24] B. Green and T. Tao. Quadratic uniformity of the Möbius function. *Ann. Inst. Fourier (Grenoble)*, 58(6):1863–1935, 2008.

[25] B. Green and T. Tao. The Möbius function is strongly orthogonal to nilsequences. *Ann. of Math. (2)*, 175(2):541–566, 2012.

[26] G. Harman. On the distribution of αp modulo one. *J. London Math. Soc. (2)*, 27(1):9–18, 1983.

[27] G. Harman. On the distribution of αp modulo one. II. *Proc. London Math. Soc. (3)*, 72(2):241–260, 1996.

[28] D.R. Heath-Brown. Prime numbers in short intervals and a generalized Vaughan identity. *Canad. J. Math.*, 34(6):1365–1377, 1982.

[29] D.R. Heath-Brown. Sieve identities and gaps between primes. In *Journées Arithmétiques (Metz, 1981)*, volume 94 of *Astérisque*, pages 61–65. Soc. Math. France, Paris, 1982.

[30] D.R. Heath-Brown and H. Iwaniec. On the difference between consecutive primes. *Invent. Math.*, 55:49–69, 1979.

[31] H.A. Helfgott. Minor arcs for goldbach's problem. *Submitted*, 2012. arXiv:1205.5252.

[32] H.A. Helfgott. Major arcs for goldbach's theorem. *Submitted*, 2013. arXiv:1305.2897.

[33] G. Hoheisel. Primzahlprobleme in der Analysis. *Sitzungsberichte Akad. Berlin*, pages 580–588, 1930.

[34] H. Iwaniec and E. Kowalski. *Analytic number theory*. American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 2004. xii+615 pp.

[35] H. J.B. Friedlander and Iwaniec. Incomplete Kloosterman sums and a divisor problem. *Ann. of Math. (2)*, 121(2):319–350, 1985. With an appendix by Bryan J. Birch and Enrico Bombieri.

[36] I. Kobayashi. A note on the Selberg sieve and the large sieve. *Proc. Japan Acad.*, 49(1):1–5, 1973.

[37] E. Kowalski. *Un cours de théorie analytique des nombres*, volume 13 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, 2004.

[38] E. Kowalski. Amplification arguments for large sieve inequalities. *Arch. Math. (Basel)*, 94(5):443–457, 2010.

[39] Yu.V. Linnik. The large sieve. *Doklady Akad. Nauk SSSR*, 30:292–294, 1941.

[40] Yu.V. Linnik. On the least prime in an arithmetic progression. I: the basic theorem. *Mat. Sb., N. Ser.*, 15(57):139–178, 1944.

[41] Yu.V. Linnik. On the least prime in an arithmetic progression. II: the Deuring- Heilbronn theorem. *Mat. Sb., N. Ser.*, 15(57):139–178, 1944.

[42] Yu.V. Linnik. The dispersion method in binary additive problems. *Leningrad*, page 208pp, 1961.

[43] C. Mauduit and J. Rivat. Sur un problème de Gelfond: la somme des chiffres des nombres premiers. *Ann. of Math. (2)*, 171(3):1591–1646, 2010.

[44] H.L. Montgomery. A note on the large sieve. *J. London Math. Soc.*, 43:93–98, 1968.

[45] H.L. Montgomery and R.C. Vaughan. The large sieve. *Mathematika*, 20(2):119–133, 1973.

[46] Y. Motohashi. Primes in arithmetic progressions. *Invent. Math.*, 44(2):163–178, 1978.

[47] Y. Motohashi. Sieve Methods and Prime Number Theory. *Tata Lectures Notes*, page 205, 1983.

[48] O. Ramaré. *Contribution au problème de Goldbach : tout entier > 1 est d'au plus 13 nombres premiers.* 1–70pp, Université Bordeaux I, 1991.

[49] O. Ramaré. On Snirel'man's constant. *Ann. Scu. Norm. Pisa*, 21:645–706, 1995. `http://math.univ-lille1.fr/~ramare/Maths/Article.pdf`.

[50] O. Ramaré. A sharp bilinear form decomposition for primes and moebius function. *Submitted to Acta Mathematica Sinica*, page 45pp, 2013.

[51] O. Ramaré and I.M. Ruzsa. Additive properties of dense subsets of sifted sequences. *J. Théorie N. Bordeaux*, 13:559–581, 2001.

[52] B. Riemann. Üeber die Anzahl der Primzahlen unter einer gegebenen Grösse. *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, 1859.

[53] L.G. Šnirel'man. Über additive Eigenschaften von Zahlen. *Math. Ann.*, 107:649–690, 1933.

[54] T. Tao. Every odd number greater than 1 is the sum of at most five primes. *Math. Comp.*, 2013.

[55] H. Tietze. *Gelöste und ungelöste mathematische Probleme aus alter und neuer Zeit*, volume Bd. 1, xx+256 pp. (10 plates); Bd. 2, iv+298 pp. (8 plates). Verlag C. H. Beck, München, 1959.

[56] R.C. Vaughan. Mean value theorems in prime number theory. *J. London Math Soc. (2)*, 10:153–162, 1975.

[57] A.I. Vinogradov. The density hypothesis for Dirichet l-series. *Izv. Akad. Nauk SSSR Ser. Mat.*, 29, 1965.

[58] I.M. Vinogradov. A new estimate of a certain sum containing primes. 1937.

[59] Ytang Zhang. Bounded gaps between primes. *Ann. of Math. (2)*, page 60pp.