PROCEEDINGS OF THE
ROMAN NUMBER THEORY ASSOCIATION
Volume 3, *Number* 1, *March* 2018, *pages* 37-44

Pieter Moree

Counting constrained almost primes

written by Carlo Pagano

1 General problem and main results

As the title suggests, this talk was concerned with asymptotic problems on *constrained k-almost primes*, where k is a fixed positive integer.

A positive integer n is said to be k-almost prime if $\Omega(n) = k$, i.e. if n has precisely k prime factors counted with multiplicities (for example 6 and 9 are both 2-almost prime). The *constrained* refers to the fact that additional constraints will be considered on the prime factors constituting the number n. As we shall see in the next section, the motivation of constraining the prime factors of n comes from the two main applications of this subject, namely cryptography and a problem about the size of the coefficients of the cyclotomic polynomials.

For a positive real number x, define

$$\pi(x,k) := \#\{n \leqslant x : \omega(n) = k\},\ N(x,k) := \#\{n \leqslant x : \Omega(n) = k\},$$

where $\omega(n)$ denotes the number of prime factors of n without multiplicity. By definition $N(x,k)$ is precisely the number of k-almost prime integers up to x. The two functions have the same order of magnitude, which is established in a classical theorem of Landau [6].

Theorem 1 (Landau, 1909)**.** *Let k be a fixed positive integer. Then, asymptotically in x,*

$$\pi(x,k) \sim N(x,k) \sim \frac{x}{\log(x)} \frac{(\log\log(x))^{k-1}}{(k-1)!}.$$

We recall that given two functions $A_1(x)$ and $A_2(x)$ the notation $A_1(x) \sim A_2(x)$ is a shorthand for

$$\lim_{x\to\infty} \frac{A_1(x)}{A_2(x)} = 1.$$

The speaker stressed that this is only the easiest version of a broad spectrum of results of Erdős, Sathe, Selberg, Hensley and Hildebrand-Tenenbaum, valid on larger regions of the plane (x,k).

Clearly for $k = 1$ one recovers (a simple version of) the prime number theorem, which in its cruder form states that the counting of the number of primes up to x, i.e. $\pi(x,1)$, satisfies

$$\pi(x,1) \sim \frac{x}{\log(x)}.$$

With an eye towards applications, the rest of the talk has been focused on the cases $k = 2$ and $k = 3$ with the constraints that we next explain.

1.1 Binary integer with prime factors within a given factor

Fix $r > 1$ a real number. Denote by

$$C_r(x) := \#\{pq \leqslant x : p < q < rp\}.$$

In words, this is the set of square-free 2-almost prime numbers, whose 2 distinct prime factor are within each other of a multiplicative factor at most r. Decker and the author [1] dubbed these integers `RSA-integers`, as integers with two prime factors of roughly the same size play an important role in the RSA cryptosystem. One of the two goals of the

talk was to show recent results that the author obtained regarding the asymptotic evaluation of $C_r(x)$. In increasing precision of the asymptotic formula, one has:

Theorem 2 (Decker and Moree, [1]). *As x tends to infinity we have*

$$C_r(x) = \frac{2x\log\log(r)}{(\log(x))^2} + O\Big(\frac{rx\log(er)}{(\log(x))^3}\Big).$$

Theorem 3 (Moree and Saad Eddin, [5]). *For $x \geqslant 2r$ and x tending to infinity we have*

$$C_r(x) = \int_{2r}^{x} \frac{\log\log(\sqrt{rt}) - \log\log(\sqrt{\frac{t}{r}})}{\log(t)} + O(rxe^{-c_1\sqrt{\log(x)}})$$

for some constant $c_1 > 0$.

By partial integration one obtains from the latter result:

Theorem 4 (Moree and Saad Eddin [5]). *Let $r > 1$ be an arbitrary fixed real number and $n \geqslant 2$ be an arbitrary integer. As x tends to infinity, we have*

$$C_r(x) = \sum_{j=1}^{n-1} a_j(r)\frac{x}{(\log(x))^{k+1}} + O_n\Big(\frac{x\log(2r)^{2\lfloor n/2\rfloor+1}\log(r)}{(\log(x))^{n+1}}\Big) + O(rxe^{-c(\varepsilon)\sqrt{\log(x)}}).$$

Where $c(\varepsilon) = (1-\varepsilon)c/\sqrt{2}$, for a constant $c > 0$ and $0 < \varepsilon < 1$ is arbitrary, and where

$$a_k(r) := \sum_{j=1}^{[\frac{k+1}{2}]} \frac{k!}{(2j-1)!}\frac{2\log(r)^{2j-1}}{2j-1}.$$

Here $[x]$ denotes the integral part of x and $\lfloor x\rfloor$ its floor.

1.2 Binary integers within a given factor and in a given congruence

The results of this section are motivated by a recently observed *bias* in the congruence of the two factors of a binary integer modulo 4. Namely consider the counting function

$$s(x) := \frac{\#\{pq \leqslant x : p \equiv q \equiv 3 \bmod 4\}}{\frac{1}{4}\#\{pq \leqslant x\}}.$$

One would obviously expect that numerical simulation should show rather rapidly that $s(x)$ approaches to 1. So it might be quite surprising that one finds numerically that $s(10^6)$ is roughly 1.183 and $s(10^7)$ is roughly 1.162. This suggested the presence of a large *secondary term*, as indeed can be proved:

Theorem 5 (Dummit, Granville and Kisilevsky [2])**.**

$$s(x) = 1 + \frac{\beta + o(1)}{\log\log(x)},$$

with β *of order* 0.334 *up to the third decimal digit.*

Clearly the theorem explains the above empirical bias, as $\beta/\log\log(x)$ decreases immensely slowly to zero!

Given the cryptography-application of binary integers of the restricted form discussed in the previous subsection called RSA-integers, it is interesting to see if they display a similar bias. If so one might perhaps be able to speed up cracking the RSA-system by assuming that both prime factors are congruent to 3 modulo 4 first.

The result obtained is as follows:

Theorem 6 (Moree and Saad Eddin [5])**.** *Let* a_1, d_1, a_2, d_2 *be natural numbers with* $(a_1, d_1) = (a_2, d_2) = 1$*. Set*

$$\mathcal{S}(x) = \{pq \leqslant x : p \equiv a_1 \bmod d_1, q \equiv a_2 \bmod d_2\}.$$

We then have:

$$\frac{\#\mathcal{S}(x)}{\frac{1}{\phi(d_1)\phi(d_2)}\#\{pq \leqslant x\}} = 1 + O_r\big((\log(x))^2 e^{-c(\varepsilon)\sqrt{\log(x)}}\big).$$

Note that the error decreases very quickly to zero and thus there is at most a very weak bias. In particular, there is no usable bias in the RSA-integer case (with $d_1 = d_2 = 4$).

1.3 Ternary integers

Consider the constrained ternary problem of asymptotically estimating the cardinality of

$$\mathcal{T}(x)=\{pqr \leqslant x : 3 \leqslant p < q < r < \frac{p-1}{p-2}(q-1), r \equiv q \equiv \pm 1 \bmod p\}.$$

The speaker, together with a number of coauthors, obtained the following result [4]:

Theorem 7 (Luca, Moree, Osburn, Saad Eddin and Sedunova, 2017)**.** *We have that*

$$\#\mathcal{T}(x) = C_1 \frac{x}{(\log(x))^2} + O\Big(\frac{x \log\log(x)}{(\log(x))^3}\Big),$$

where

$$C_1 = \frac{1}{4}\sum_{l \geqslant 3} \frac{1}{l(l-1)^2} \log\Big(\frac{l-1}{l-2}\Big),$$

with l running over all odd prime numbers.

2 Motivation and applications

2.1 Motivation

The security of the RSA system is based on the fact that (with classical computers) it is considered difficult to factor RSA-integers in reasonable

time. This motivated the speaker to estimate the number of such integers up to x and to see how much of a bias there is when congruence conditions are specified for the two prime factors.

2.2 Applications

Let n be a positive integer. Let $\zeta_n := e^{\frac{2\pi i}{n}}$. Recall that the n-th cyclotomic polynomial is defined as

$$\Phi_n(T) = \prod_{1 \leqslant j \leqslant n, (j,n)=1} (x - \zeta_n^j).$$

We will look at its coefficients:

$$\Phi_n(T) := \sum_{i=0}^{\phi(n)} a_n(i) T^i,$$

where $\phi(n)$ denotes Euler's totient function. One defines the height of $\Phi_n(T)$ to be the largest value of $|a_n(i)|$ as i ranges through $\{0, \ldots, \phi(n)\}$.

It is a known fact that $A(n) = 1$ whenever n is at most binary. Then one has to look at at least ternary integers to see some non-trivial behaviour.

It was proven by Bang, in 1895, that $A(pqr) \leqslant p - 1$, for $p < q < r$. Therefore one has that $\max_{q,r}\{A(pqr : p < q < r\}$ exists; denote it by $M(p)$. It is the main open problem in the subject, to determine a formula or an efficient algorithm to compute $M(p)$. It has been conjectured by a Sister, named Sister Beiter, that $M(p) \leqslant \frac{p+1}{2}$. This conjecture has been disproved by the speaker and Gallot in 2008 [3] for all primes p. They proposed a corrected version of this conjecture, by conjecturing that $M(p) \leqslant \frac{2}{3}p$. They called this the *corrected Beiter conjecture*. They were able to show that, given an $\varepsilon > 0$, the inequality $M(p) > (2/3 - \varepsilon)p$ holds for all large enough p.

Using Theorem 7, the authors of [4] established the following result.

Theorem 8 (Luca, Moree, Osburn, Saad Eddin and Sedunova (2017))**.** *The number $T(x)$ of ternary $n = pqr \leqslant x$ such that $A(pqr) \leqslant \frac{2p}{3}$ satisfies*

$$T(x) \geqslant \Big(\frac{25}{27} + o(1)\Big)\pi(x,3) \geqslant \Big(\frac{25}{27} + o(1)\Big)\frac{x(\log\log(x))^2}{2\log(x)}.$$

This implies that a very big proportion of the ternary integers $\leq x$, namely at least 0.925 for all x large enough, respects the corrected Beiter conjecture, that is

$$|a_{pqr}(i)| \leqslant \frac{2}{3}p \text{ for each } i \in \{0,\dots,\phi(pqr)\}.$$

References

[1] A. Decker and P. Moree, *C*ounting RSA-integers, Result. Math. **5**2 (2008), 35–39.

[2] D. Dummit, A. Granville, and H. Kisilevsky, Big biases amongst product of two primes, *M*athematika **6**2, (2016), 502–507.

[3] Y. Gallot and P. Moree, Ternary cyclotomic polynomials having a large coefficient, *J*. Reine Angew. Math. **6**32 (2009), 105–125.

[4] F. Luca, P. Moree, R. Osburn, S. Saad Eddin, and A. Sedunova, Constrained ternary integers, `arXiv:1710.08403`, (2017).

[5] P. Moree and S. Saad Eddin, Products of two proportional primes, *I*nt. J. Number Theory **1**3 (2017), 2583–2596.

[6] G. Tenenbaum, Introduction to analytic and probabilistic number theory, *C*ambridge University Press, (1995).

Carlo Pagano
Mathematisch Instituut
Universiteit Leiden
Niels Bohrweg 1
2333 CA Leiden, The Netherlands.
email: c.pagano@math.leidenuniv.nl