Functiones et Approximatio
XXXVII.2 (2007), 361–376

CRAMÉR VS. CRAMÉR. ON CRAMÉR'S PROBABILISTIC MODEL FOR PRIMES*

János Pintz

Dedicated to the 60th birthday
of Jean-Marc Deshouillers

Abstract: In the 1930's Cramér created a probabilistic model for primes. He applied his model to express a very deep conjecture about large differences between consecutive primes. The general belief was for a period of 50 years that the model reflects the true behaviour of primes when applied to proper problems. It was a great surprise therefore when Helmut Maier discovered in 1985 that the model gives wrong predictions for the distribution of primes in short intervals. In the paper we analyse this phenomen, and describe a simpler proof of Maier's theorem which uses only tools available at the mid thirties. We present further a completely different contradiction between the model and the reality. Additionally, we show that, unlike to the contradiction discovered by Maier, this new contradiction would be present in essentially all Cramér type models using independent random variables.
Keywords: primes, probabilistic model for primes, Cramér's model for primes.

1. Cramér's model

Cramér's probabilistic model for primes [3,4] (abbreviated further as CM), created by him in the mid 1930's, plays also today a fundamental role when formulating conjectures concerning primes.

The prime number theorem (PNT)

$$\pi(x) = \sum_{p \leqslant x} 1 \sim \operatorname{li} x = \int_0^x \frac{du}{\log u}, \tag{1.1}$$

asserts that the expected density of primes around x is $1/\log x$. Cramér's probabilistic model is a sequence of *independent* random variables $\xi(n)$, defined for $n \geqslant 3$ by

$$\mathrm{P}(\xi_n = 1) = \frac{1}{\log n}, \qquad \mathrm{P}(\xi_n = 0) = 1 - \frac{1}{\log n}. \tag{1.2}$$

2000 Mathematics Subject Classification: 11N05.
* This work was supported by OTKA grants No. 43623, 43693 and 67676.

The heuristic assumption of Cramér is that in certain problems ξ_n imitates well the behaviour of the characteristic function of primes, i.e. the function ($\mathcal{P} = \{p_n\}_{n=1}^{\infty}$ denotes the set of primes)

$$\chi_{\mathcal{P}}(n) = \begin{cases} 1 & \text{if} \quad n \in \mathcal{P}, \\ 0 & \text{otherwise.} \end{cases} \tag{1.3}$$

Cramér used this model to formulate a conjecture about large gaps between consecutive primes. For any infinite $\{0,1\}$ sequence, corresponding to the value distribution $\{\xi_n\}_{n=3}^{\infty}$ we can associate the series P_ν, where

$$P_{\nu+1} = m \quad \text{if} \quad \sum_{n=3}^{m} \xi_n = \nu \quad \text{and} \quad \sum_{n=3}^{m-1} \xi_n = \nu - 1. \tag{1.4}$$

According to Cramér's heuristic the largest possible gaps between primes behave similarly to the largest possible gaps between P_n. More precisely, he proved that with probability 1 we have

$$\limsup_{n\to\infty} \frac{P_{n+1} - P_n}{(\log P_n)^2} = 1. \tag{1.5}$$

On the basis of (1.5) he conjectured for the primes the same relation, that is,

$$\limsup_{n\to\infty} \frac{p_{n+1} - p_n}{\log^2 p_n} = 1. \tag{1.6}$$

The conjecture shows a relatively good (although not very good) agreement with empirical data.

In the present work we will examine how well this model describes the properties of primes. We will describe in detail the obvious trivial deficiencies of Cramér's model (Section 2) and the surprising result of Maier from 1985 which showed the first time a 'non-trivial contradiction' between the distribution of primes and Cramér's model (Section 3). We will present another type of contradiction in Section 4. Further we will examine whether Cramér's model may be modified in a way as to scope with the arising 'contradictory' phenomena. While such a modification is known to be possible in case of 'Maier's phenomenon', we will show (Sections 5–6) that no non-trivial Cramér-type probabilistic model (using independent random variables) exists, which would scope with the phenomenon examined in Section 5. Finally following Granville [7], we describe in Section 7 how probabilistic models can help or could have helped not only to conjecture but also prove results about primes.

2. Obvious deficiencies of Cramér's model

Since the sequence of primes p_n is a given deterministic sequence we cannot hope that the simple probabilistic model of Cramér should well reflect all properties of primes. We have, for example, with probability 1 (abbreviated in the following by w. p. 1), asymptotically as many even and odd values of 'the probabilistic primes' P_n, that is

$$\left|\{P_n \leqslant x; 2 \mid P_n\}\right| \sim \left|\{P_n \leqslant x; 2 \nmid P_n\}\right| \qquad (\text{as } x \to \infty) \tag{2.1}$$

or

$$\liminf_{n\to\infty}(P_{n+1} - P_n) = 1 \tag{2.2}$$

(that is, we can find w. p. 1, infinitely many consecutive 'probabilistic primes' P_n). CM is clearly not sensible for the property whether a given natural number n is divisible by small primes or not; a property which is crucial if we examine the primality of n. This deficiency is obviously responsible for the fact that the analogues of (2.1) and (2.2) are not true for primes, since all primes larger than two are odd.

Naturally, for the same reason it is also totally false to assume the random variables $\xi(n)$ to be independent as shown by the obvious falsity of the analogue of (2.2) for primes.

Nevertheless, we have some means to make corrections on CM and to arrive in this way at plausible conjectures about primes. Let us consider, for example, following Pólya [15], the number of twin primes below x. If the events that n and $n+2$ are both primes would be 'approximately' independent, then we would arrive at the naive conjecture:

$$N_2(x) = \left|\{n \leqslant x;\ n, n+2 \in \mathcal{P}\}\right| \sim \frac{x}{\log^2 x}, \tag{2.3}$$

which would be the result following w. p. 1 from CM. (This naive conjecture is according to general belief wrong, but we have no way to disprove it.) On the other hand, n and $n+2$ are clearly simultaneously both even or both odd, so the probability that they are both odd is $1/2$ and not $1/4 = (1/2)^2$, which would be the case if these two events would be independent.

Similarly, for any $p > 2$ the probability

$$\mathrm{P}(p \nmid n, p \nmid n+2) = \mathrm{P}(n \not\equiv 0, -2 (\operatorname{mod} p)) = \frac{p-2}{p}, \tag{2.4}$$

contrary to the wrong probability

$$\left(1 - \frac{1}{p}\right)^2, \tag{2.5}$$

suggested by the *independence* condition of CM. This suggests a correction factor $\left(1 - \frac{2}{p}\right)\left(1 - \frac{1}{p}\right)^{-2}$ for any prime $p > 2$.

Taking into account all these local correction factors (including the correction factor 2 for $p = 2$), this means that we should multiply the naive probability $(\log x)^{-2}$ of (2.3) by

$$2c_0 := 2 \prod_{p>3} \left(1 - \frac{2}{p}\right) \left(1 - \frac{1}{p}\right)^{-2} = 2 \prod_{p>3} \left(1 - \frac{1}{(p-1)^2}\right), \tag{2.6}$$

in order to arrive at the plausible conjecture

$$N_2(x) \sim 2c_0 \frac{x}{\log^2 x} \qquad (c_0 = 0.66016\ldots), \tag{2.7}$$

where $2c_0 = 1.3203\ldots$ is the well-known twin-prime constant. This formula is the same which is suggested by the singular series in Hardy–Littlewood's circle method.

The same heuristic would apply to the number of Goldbach decompositions of even integers or to the expected number of prime k-tuples

$$N_{\mathcal{H}}(x) := \bigl|\{n \leqslant x; n + h_i \in \mathcal{P} \ \text{ for } \ i = 1, 2, \ldots, k\}\bigr|, \tag{2.8}$$

when $\mathcal{H} = \{h_i\}_{i=1}^k$, $h_i \in \mathbb{Z}^+ \cup \{0\}$, $h_i \neq h_j$ for $i \neq j$, which constitutes a far-reaching generalization of the twin prime conjecture. Denoting by $\nu_p(\mathcal{H})$ the number of distinct residue classes covered by $\mathcal{H} \bmod p$, the correct probability that all $n + h_i$ are not divisible by a given prime p is

$$\mathrm{P}\bigl(p \nmid (n + h_i),\ 1 \leqslant i \leqslant k\bigr) = \frac{p - \nu_p(\mathcal{H})}{p}, \tag{2.9}$$

in contrast to the wrong naive probability (cf. (2.4)–(2.5))

$$\left(1 - \frac{1}{p}\right)^k, \tag{2.10}$$

suggested by the *independence* condition of CM. This suggests that the naive probability $(\log x)^{-k}$ arising from CM should be multiplied by the correction factor

$$\mathfrak{S}(\mathcal{H}) := \prod_p \left(1 - \frac{\nu_p(\mathcal{H})}{p}\right) \left(1 - \frac{1}{p}\right)^{-k} \tag{2.11}$$

to yield the plausible estimate

$$N_{\mathcal{H}}(x) \sim \mathfrak{S}(\mathcal{H}) \frac{x}{(\log x)^k}, \tag{2.12}$$

which is the same formula as the one conjectured by Hardy–Littlewood, on a completely different basis.

3. Maier's discovery. Serious deficiency found in CM

Despite the mentioned obvious deficiency of Cramér's model in local problems, the general belief was before 1985 that his model predicts correctly the behaviour of primes in those cases where we consider problems of *global* nature (like distribution of primes in long intervals) or of *semi-global* nature (like distribution of primes in short intervals, like, for example, the original problem (1.5)–(1.6) of Cramér about large gaps between consecutive primes). It was therefore a great surprise when H. Maier [12] showed in 1985 that the Prime Number Theorem is not true in short intervals of type

$$\big[x, x + (\log x)^{\lambda}\big], \qquad \lambda > 0 \ \text{ arbitrary}, \tag{3.1}$$

whereas CM would predict its truth for all $\lambda > 2$ with probability $= 1$.

Maier succeeded namely to show the existence of intervals of type (3.1) containing $\leqslant c_1(\lambda)$ times less or $\geqslant c_2(\lambda)$ times more primes than expected, where

$$c_1(\lambda) < 1 < c_2(\lambda) \quad \text{for any} \ \ \lambda. \tag{3.2}$$

Maier examined in his work the distribution of numbers, relatively prime to

$$P(z) = \prod_{p \leqslant z} p, \tag{3.3}$$

which we shall call z-quasiprimes (and denote their set by $Q(z)$) in the following.

It follows trivially from the periodicity of the set $Q(z)$ that

$$\Phi(x, z) := \big|\{n \leqslant x; (n, P(z)) = 1\}\big| \sim xW(z) := x \prod_{p \leqslant z} \left(1 - \frac{1}{p}\right), \tag{3.4}$$

for any x and $z = z(x)$ such that

$$\frac{xW(z)}{P(z)} \longrightarrow \infty. \tag{3.5}$$

Another plausible hypothesis (also true for fixed z and $x \to \infty$) is that the density of primes among z-quasiprimes is larger by a factor $B(z)$ where

$$B(z) := \frac{P(z)}{\varphi(P(z))} = \prod_{p \leqslant z} \left(1 - \frac{1}{p}\right)^{-1} = (W(z))^{-1}. \tag{3.6}$$

The above principle is true in the stronger sense also that taking any arithmetic progression of type

$$A_m = mP(z) + a, \qquad a \in Q(z) \ \big(\Longrightarrow A_m \in Q(z)\big), \tag{3.7}$$

the density of primes in the special sequence $A_m \cap [x/2, x]$ is

$$\sim \frac{B(z)}{\log x}, \tag{3.8}$$

if there are no Siegel zeros for the characters $\bmod P(z)$ and

$$\frac{\log x}{\log P(z)} \sim \frac{\log x}{z} := f(x) \longrightarrow \infty \qquad (\text{as} \quad x \to \infty). \tag{3.9}$$

The above assertion is a deep result of Gallagher [6]. This means heuristically that in order to find some intervals with more (or resp. less) primes than expected, it is sufficient to produce some intervals with more (or resp. less) z-quasiprimes than expected.

The next observation is that the number of z-quasiprimes in not too long intervals of type

$$(0, y] \quad y = z^{\lambda} \qquad (\lambda > 1 \text{ fixed}) \tag{3.10}$$

is, in contrast to the asymptotic relation (3.4), by Buchstab's theorem [1]

$$\Phi(y, z) \sim yW(z)\omega^*(\lambda), \qquad (\text{as } y \to \infty), \tag{3.11}$$

where $\omega^*(u)$ is a precisely determined continuous function with (see Iwaniec [9])

$$\min_{u \in [a, a+1]} \omega^*(u) < 1 < \max_{u \in [a, a+1]} \omega^*(u) \qquad \text{if } \ a \geqslant 1. \tag{3.12}$$

The trivial periodic nature of the set $Q(z)$ implies that similarly to the above we have $\omega^*(\lambda)$ times the average number of z-quasiprimes in intervals of type

$$(x_\nu, x_\nu + z^{\lambda}] \quad \text{if} \quad x_\nu \in \mathbb{Z}, \quad P(z) \mid x_\nu. \tag{3.13}$$

This shows that we can expect $\omega^*(\lambda)$ times the expected number of primes in intervals of type (3.13), if

$$z^{\lambda} = \left(\frac{\log x_\nu}{f(x_\nu)}\right)^{\lambda}, \quad \text{for any } \ f(x) \to \infty \ \text{ as } \ x \to \infty. \tag{3.14}$$

To make the last heuristic precise we have only to average over intervals of type (3.13) with

$$x_\nu := (M + \nu)P(z), \qquad \nu = 1, 2, \ldots M \tag{3.15}$$

with an integer M with property $\log M / \log P(z) \to \infty$, e.g.

$$M = P(z)^{[\log z]} \qquad \text{for } \ z = n \to \infty. \tag{3.16}$$

The (sketch of) proof can be finished with the observation that due to the scarcity of Siegel zeros (cf. Landau–Page theorem, Davenport [5, Chapter 14, pp. 93–95]) we can choose a subsequence $z_j = n_j \to \infty$ such that there are no Siegel zeros $\bmod q_j$, where $q_j = P(z_j)$, that is,

$$L(s,\chi) \neq 0 \quad \text{for} \quad s \in [1 - c/\log q_j, 1] \tag{3.17}$$

for all characters $\bmod q_j$ with some positive absolute constant c. The asymptotic (3.8) follows then by well-known density theorems for any modulus q satisfying (3.17) (cf. [12, Lemma 2]). We have namely in this case

$$\sum_{\substack{x<p\leqslant 2x \\ p\equiv a \pmod q}} \log p = \frac{x}{\varphi(q)}\Big(1 + O\big(e^{-\log x/\log q} + e^{-\sqrt{\log x}}\big)\Big). \tag{3.18}$$

The ingenious idea of Maier was later used and further developed in a series of works by Granville, Friedlander (partly in collaboration with Hildebrand and Maier) to yield a number of *unexpected* irregularities in the distribution of primes in arithmetic progressions, in the sense that they all *contradict* the predicting of Cramér's model. These results also showed limitations to the extension of Bombieri–Vinogradov's theorem if the modulus q approaches x (that is, limitations to Elliott–Halberstam type conjectures). To formulate the simplest of their results: the relation

$$\pi(x,q,a) \sim \frac{\pi(x)}{\varphi(q)} \tag{3.19}$$

is false for some values of a_ν, q_ν and x_ν satisfying $(a_\nu, q_\nu) = 1$,

$$q_\nu > \frac{x_\nu}{(\log x_\nu)^B} \qquad (x_\nu \to \infty), \tag{3.20}$$

for any fixed B. On the contrary, CM predicts (3.19) to be uniformly true for any fixed $B > 2$.

These results and many interesting aspects of Cramér's model are described in more details in the excellent survey papers [7, 8] of Granville.

4. A global deficiency in CM

Few years after Maier's discovery [12], the present author also observed (but did not publish) another type of deficiency in CM. We might list different features of the two type of deficiencies as follows:

(i) the contradiction to CM is quantitatively much less significant than in Maier's case;

on the other hand,

(ii) the contradiction is completely global as it refers to the distribution of primes in intervals $[0, x]$, whereas in Maier's case the contradiction refers to the distribution of primes in short intervals, so it is of a semi-global nature;

(iii) the contradiction is true for all values of x, whereas Maier's result refers only for some very special short intervals of a given length;

(iv) the proof of the contradiction is quite simple, it does not require any deep results about the distribution of primes, as density theorems in Maier's proof;

(v) contrary to Maier's case, it seems to be impossible to correct CM in order to avoid the deficiency.

Some of the mentioned features of the contradiction discovered by Maier are obvious. As of (iii) we remark that the intervals of type (3.13) really represent a very thin subset of all intervals of length $(\log x)^N$, even if we would assume the absence of Siegel-zeros or the Generalized Riemann Hypothesis, for example. In the original proof of Maier even $z_j = n_j$ has to run over a scarce sequence.

Concerning the possibility of the correction of Cramér's model we explained in Section 3 that the irregularities of the distribution of primes in short intervals were 'caused' in Maier's case by the irregularities of the distribution of z-quasiprimes in short intervals, while the density of primes reflected the density of z-quasiprimes in the relevant short intervals (cf. 3.7–3.9). This shows the following possibility of correcting CM for $n \in (x/2, x]$, as described by Granville [7, 8]. Let us choose the 'possible primes' only among the z-quasiprimes and let us take composite all numbers having a prime divisor $\leqslant z$. Putting it into an exact form let us choose for all $n \in (x/2, x]$ the independent random variables ζ_n as

$$\zeta_n = 0 \qquad \text{if} \quad n \notin Q(z); \tag{4.1}$$

and for $n \in Q(z)$ let

$$\mathrm{P}(\zeta_n = 1) = \frac{x}{\Phi(x,z)\log n}, \quad \mathrm{P}(\zeta_n = 0) = 1 - \frac{x}{\Phi(x,z)\log n}, \tag{4.2}$$

where the parameter $z = z(x)$ could be chosen in different ways, but it should satisfy for $x \to \infty$ with any constant $A > 0$

$$\frac{z}{(\log x)^A} \to \infty, \quad z \leqslant x^c \quad (c < 1/2). \tag{4.3}$$

It is clear that in this way we can exclude also the 'obvious' deficiencies discussed in Section 2. We mention that if we allow z to reach or exceed $\sqrt{x}$, then the z-quasiprimes below x will be exactly the primes above z, so it is no surprise that then all the contradictions would disappear, since in this case every prime in $(z, x]$ is chosen as a probabilistic prime with probability 1, and all composite numbers in $(0, x]$ with probability 0. On the other hand, for $z = 1$ this is exactly the original model of Cramér.

In the following we show the mentioned global contradiction to CM. For technical reasons we weight primes and 'probabilistic primes' by $\log n$, so we will work with the independent random variables

$$\xi_n' = \xi_n \log n \qquad (n \geqslant 3), \tag{4.4}$$

where ξ_n is defined in (1.2). In this way we have $E(\xi_n') = 1$. Therefore we might expect that the square $(\Delta'(x))^2$ of the error term in the PNT considered in the form

$$\Delta'(x) := \sum_{2<p\leqslant x} \log p - \sum_{2<n\leqslant x} 1, \quad x \in \mathbb{Z} \tag{4.5}$$

might be approximated well with the variance of the 'probabilistic primes', arising from CM,

$$\begin{aligned}\mathrm{D}^2\Biggl(\sum_{n=3}^{x} \xi_n'\Biggr) &= \sum_{n=3}^{x} \mathrm{D}^2(\xi_n') = \sum_{n=3}^{x} \bigl(\mathrm{E}((\xi_n')^2) - (\mathrm{E}(\xi_n'))^2\bigr)\\ &= \sum_{n=3}^{x} \biggl(\frac{1}{\log n}\log^2 n - 1\biggr) \sim x\log x.\end{aligned} \tag{4.6}$$

Our knowledge about the oscillation of the error term was by the theorems of von Koch [10], Phragmén [13] and Littlewood [11] already in 1914

$$\Delta'(x) = O\bigl(\sqrt{x}\log^2 x\bigr) \qquad \text{(on RH)}, \tag{4.7}$$

$$\Delta'(x) = \Omega(x^{\theta-\varepsilon}) \qquad \text{if } \theta = \sup_{\zeta(\varrho)=0} \operatorname{Re}\varrho, \tag{4.8}$$

$$\Delta'(x) = \Omega\bigl(\sqrt{x}\log\log\log x\bigr). \tag{4.9}$$

It is interesting to note that the results (4.7) and (4.9) are still today, after about one hundred years, the best known ones. Concerning (4.8), the best oscillation result, supposing the existence of a ζ-zero $\varrho_0 = \beta_0 + i\gamma_0$, the relation

$$\limsup_{x\to\infty} \frac{|\Delta'(x)|}{(x^{\beta_0}/|\varrho_0|)} \geqslant \frac{\pi}{2} \tag{4.10}$$

is due to S. G. Révész [16], improving earlier results of the author [14] and Paul Turán [18].

Here (4.8) and (4.10) imply that the correctness of CM is inconsistent with $\theta > 1/2$, the falsity of RH. On the other hand, the expected oscillation $\sqrt{x\log x}$ of (4.6) is still consistent with both (4.7) and (4.9), if RH is supposed to be true.

However, there is a result from 1920, preceding Cramér's model by 15 years, stating that on the RH one has with $\Delta(x) = \sum_{n\leqslant x} \Lambda(n) - x = \Delta'(x) + O(x^{1/4+\varepsilon})$

$$\frac{1}{Y}\int_2^Y \Delta^2(x)dx \ll Y \Longrightarrow \frac{1}{Y}\int_2^Y \bigl(\Delta'(x)\bigr)^2 dx \ll Y, \tag{4.11}$$

contradicting CM, since (4.11) shows that $|\Delta'(x)|$ is in average of size $O(\sqrt{x})$ compared to the expectation $\sqrt{x\log x}$ arising from CM, as shown in (4.6). So we see that the error term of the PNT which already reflects properties of all primes

$\leqslant x$, is not predicted correctly by CM. Further this prediction is not only for some scarce sequence of values of x wrong, but in average for all $x \leqslant X$ for any large X. It might be worth to mention that if a conjecture of Montgomery,

$$\Delta(x) = O\big(\sqrt{x}(\log\log\log x)^2\big) \tag{4.12}$$

is correct, then the prediction of CM would be false even for all sufficiently large single values of x. These remarks show that this contradiction is really of completely global nature.

What makes the story more interesting is that the person who proved the crucial relation (4.11) in 1920, is Harald Cramér himself [2]!

Remark. The usual weight $\log n$, resp. $\log p$ used in this context is not essential, since the analogue of (4.11) is also true for the error term

$$\Delta_1(x) := \sum_{2<p\leqslant x} 1 - \sum_{2<n\leqslant x} \frac{1}{\log n}. \tag{4.13}$$

We obtain namely from (4.11) easily by partial summation

$$\int\limits_2^Y \Delta_1^2(x)dx \ll \frac{1}{\log^2 Y} \int\limits_2^Y \big(\Delta'(x)\big)^2 dx \ll \frac{Y^2}{\log^2 Y}. \tag{4.14}$$

5. Is a correction of CM possible?

Whereas the 'unexpected' irregularities of Maier turn to be consistent with the corrected CM (abbreviated further only CCM(z)) described by Granville [7, 8], that is, with the model (4.1)–(4.3), there seems to be no a priori reason why CCM(z) should explain the global contradiction (4.6) vs. (4.11). On the other hand, choosing $z = \sqrt{x}$, that is $c = 1/2$, the contradiction necessarily disappears (whereas it is clearly present for $z = 1$, when CCM(1) = CM).

In fact, a repetition of the calculation (4.6) and some well-known facts about the distribution of almost primes would lead with $z = x^\alpha$ for any fixed $\alpha < 1/2$ to the relation (with $\zeta_n' = \zeta_n \log n$)

$$\mathrm{D}^2\left(\sum_{n=3}^{x} \zeta_n'\right) \sim g(\alpha) x \log x \tag{5.1}$$

where $g(\alpha)$ is a monotonically decreasing positive continuous function with

$$g(\alpha) \to 0 \quad \text{as} \quad \alpha \to 1/2,\ \alpha < 1/2. \tag{5.2}$$

This proves that the most obvious correction does not help.

We will consider therefore more generally arbitrary models preserving the idea of the original and corrected Cramér Model as

(i) to designate a set S_x of 'possible primes' among the integers $1, 2, \ldots, x$ containing all primes below x, that is

$$\mathcal{P}_x := \{p \in \mathcal{P}; p \leqslant x\} \subset S_x; \tag{5.3}$$

and afterwards

(ii) consider all elements outside S_x composite, while choosing independently all elements n of S_x to be prime with a probability

$$\frac{x}{|S_x|} \cdot \frac{1}{\log n}. \tag{5.4}$$

We will call a model satisfying the above conditions a modified Cramér model (abbreviated by MCM). We will show that an MCM also contradicts the true distribution laws of primes *unless* the set S_x of 'possible primes' *essentially* coincides with $\mathcal{P}_x$, the set of actual primes.

We will show that a good modified model is already impossible for any interval $(x/2, x]$. We can formulate our assertion as

Theorem 1. *Let x be a large even number, $I = (x/2, x] \cap \mathbb{Z}$. Let S_x^* be arbitrary with*

$$\mathcal{P}_x^* := \mathcal{P} \cap I \subseteq S_x^* \subseteq I, \quad A = \frac{|I|}{|S_x^*|}. \tag{5.5}$$

Let us define independent random variables η_n for all $n \in I$ as

$$\eta_n = 0 \qquad \text{if } \ n \notin S_x^*; \tag{5.6}$$

while for $n \in S_x^*$ let

$$\mathrm{P}(\eta_n = 1) = \frac{A}{\log n}, \quad \mathrm{P}(\eta_n = 0) = 1 - \frac{A}{\log n}. \tag{5.7}$$

Then the truth of the relation

$$\mathrm{D}^2\left(\sum_{n \in I} \eta_n\right) \ll \frac{x}{\log^2 x} \tag{5.8}$$

implies

$$|S_x^* \setminus \mathcal{P}_x^*| \ll \frac{x}{\log^2 x}. \tag{5.9}$$

This shows that in view of (4.14) any possibly correct probabilistic model which is non-trivial in the sense that (5.9) is not satisfied must operate with dependent random variables.

6. Proof of Theorem 1

For $n \in S_x^*$ we have $\mathrm{E}(\eta_n) = \frac{A}{\log n}$ and

$$\begin{aligned}
\mathrm{D}^2\left(\sum_{n\in I}\eta_n\right) &= \sum_{n\in I}\mathrm{D}^2(\eta_n) = \sum_{n\in S_x^*}\big(\mathrm{E}(\eta_n^2) - (\mathrm{E}(\eta_n))^2\big)\\
&= \sum_{n\in S_x^*}\left(\frac{A}{\log n} - \frac{A^2}{\log^2 n}\right) \qquad (6.1)\\
&= \frac{|I|}{\log^2 x}\left(\log x + O(1) - A\left(1 + O\left(\frac{1}{\log x}\right)\right)\right) \ll \frac{x}{\log^2 x}
\end{aligned}$$

if and only if $A = \log x + O(1)$, that is

$$|S_x^*| = \frac{|I|}{\log x}\left(1 + O\left(\frac{1}{\log x}\right)\right) \Leftrightarrow |S_x^* \setminus \mathcal{P}_x^*| \ll \frac{x}{\log^2 x}. \qquad (6.2)$$

Remark. This calculation verifies also (5.1)–(5.2) since in case of $z = x^\alpha$ we have $\zeta_n' = \eta_n \cdot \log n$ and

$$A = \frac{|I|}{|S_x^*|} = \frac{|I|}{\big|\{n \in I; n \in Q(x^\alpha)\}\big|} \sim f(\alpha)\log x \qquad (6.3)$$

with $0 < f(\alpha) < 1$, $f(\alpha) \to 1$ as $\alpha \to 1/2$, $\alpha < 1/2$.

7. Concluding remarks

In this last section we will indicate

A) how probabilistic models can help to prove results about primes;

B) how the deep theorem of Gallagher in Maier's proof can be substituted by classical results available in 1918 already;

C) how the very weak implicit localization of 'irregular' short intervals in Maier's proof can be improved.

A) It is obvious that any probabilistic model yields only conjectures about primes and no proofs. It is therefore interesting to remark that a good probabilistic model might supply ideas for proofs of important results about primes.

For example, the discussed celebrated result of Maier could have been achieved along the following lines. (The argument – Steps 1–3 – presented below is very close to that of Granville [7, 8].)

Step 1. The obvious local deficiencies, like the existence of infinitely many even 'probabilistic primes' indicate easily a change of CM to CCM(z) (see Section 5) as described in Granville [7] and here in (4.1)–(4.3). Hence a corrected CM could have been created right after CM.

Step 2. Examine the distribution of probabilistic primes in short intervals of length $(\log x)^A$ using the new CCM(z) in place of CM (for example, with the aim to check the original conjecture (1.6)).

Since in this model the possible primes are exactly the z-quasiprimes and the probabilities of being primes are essentially (apart from the insignificant difference in the value $(\log n)^{-1}$) equal for $n, m \in Q(z)$ if $n < m < 2n$ (so even more if they are in the same short intervals), the CCM(z) predicts the same distribution laws for primes and z-quasiprimes. As the irregularities in the distribution of z-quasiprimes were proved by Buchstab [1] in 1937 one could have expected the same irregularities as those discovered by Maier much earlier, based on Buchstab's result and on the corrected CM.

Step 3. Try to show that the density of primes really follow that of z-quasiprimes, that is, that primes in short intervals are in general by a factor

$$B(z) = \prod_{p<z} \left(1 - \frac{1}{p}\right)^{-1} = \frac{P(z)}{\varphi(P(z))} \tag{7.1}$$

more dense among z-quasiprimes than among all integers.

This seems to be very hard (and is in fact hopeless) for any given single short interval. On the other hand, the factor (7.1) exactly coincides with that describing the higher concentration of primes in an arithmetic progression of type

$$mP(z) + a \quad (a, P(z)) = 1. \tag{7.2}$$

In this way one might be led to the idea of Maier: to average over many short intervals, and establish the resulting irregularities in average.

B) Further, we note that the most deep analytic part, the application of a result of Gallagher, based on log-free zero density theorems can be substituted by a much simpler one. Namely, differently from (3.16) we may choose, for example for any $N > 2$

$$M = \left\lceil \frac{x}{P(z)} \right\rceil = P(z)^{z^{N-1}} = e^{z^N(1+o(1))} \Leftrightarrow z = (1 + o(1))(\log x)^{1/N} \tag{7.3}$$

in which case the classical formula (cf. (9) of Chapter 20 of Davenport [5]) yields immediately

$$\sum_{\substack{p \leqslant x \\ p \equiv a (\mathrm{mod}\ P(z))}} \log p = \frac{x}{\varphi(P(z))} \left(1 - \overline{\chi}_1(a) \frac{x^{\beta_1 - 1}}{\beta_1}\right) + O\big(xe^{-c\sqrt{\log x}}\big), \tag{7.4}$$

where the second term in the bracket appears only in case of the existence of an exceptional zero β_1 (called today also Siegel-zero) with

$$\beta_1 > 1 - \frac{c_0}{\log P(z)}. \tag{7.5}$$

We remark that (7.4) was known already in 1918 (p. 93–95 of Davenport [5]) when Landau proved his result about scarcity of exceptional zeros which made possible for Maier to choose a sequence $z_0 \to \infty$ for which zeros with (7.5) do not exist. If we have no zeros with (7.5) then (7.4) yields the required asymptotic for $z < c_1\sqrt{\log x}$, which is true by (7.3). Choosing $y = z^{\lambda N} = (1 + o(1))\log^\lambda x$ we arrive at Maier's result (with other values of $c_i(\lambda)$, satisfying still (3.2)).

C) We also remark that it is possible to avoid exceptional zeros in another way. In that way we can obtain much more (but still a thin subset of all) short intervals with the same irregular distribution of primes, and they can be localized, for example, between

$$[N, 3N] \tag{7.6}$$

for any $N > N_0$, in place of the localization

$$\left[N, e^{c\sqrt{N}}\right] \tag{7.7}$$

which would follow after some considerations from Maier's original work [12].

The needed modification is the following. If there is an exceptional real zero of $L(s, \chi_1, P(z))$ satisfying (7.5) for a z, then consider the real primitive character χ_1^* with conductor

$$P^*(z) \mid P(z) \tag{7.8}$$

inducing $\chi_1 \bmod P(z)$, and let p_0 denote the greatest prime factor of $P^*(z)$.

In this case one can work instead of z-quasiprimes with a slightly modified set of numbers: we can consider instead numbers defined by

$$n \in \widetilde{Q}(z) \Leftrightarrow \big(n, \widetilde{P}(z)\big) = 1 \quad \text{where} \quad \widetilde{P}(z) = \frac{P(z)}{p_0}. \tag{7.9}$$

Any character $\chi_2 \bmod \widetilde{P}(z)$ induces a character, $\bmod P(z)$, different from χ_1, so we have by the Landau–Page theorem (cf. pp. 93–95 of Chapter 14, Davenport [5]) for their possible real zeros β_1 and β_2 the inequality

$$\min(\beta_1, \beta_2) < 1 - \frac{2c_0}{\log P(z)} \tag{7.10}$$

(if c_0 in (7.5) was chosen appropriately) and so

$$\beta_2 < 1 - \frac{2c_0}{\log P(z)} < 1 - \frac{c_0}{\log \widetilde{P}(z)} \tag{7.11}$$

is already satisfied for all real zeros of all $L(s, \chi)$ functions with $\chi \bmod \widetilde{P}(z)$. It is easy to see that the whole proof runs mutatis mutandis when elements of $Q(z)$ (z-quasiprimes) are substituted by elements of $\widetilde{Q}(z)$. We finally remark that although the analogue of (3.11) for $\widetilde{Q}(z)$ in place of $Q(z)$ might be proved analogously, the relation (3.11) itself implies the same relation for $\widetilde{Q}(z)$, since

$$S := \left|n \leqslant y;\ n \in \widetilde{Q}(z) \setminus Q(z)\right| \leqslant \frac{y}{p_0} = o(yW(z)) \tag{7.12}$$

unless

$$p_0 = O(\log z). \tag{7.13}$$

On the other hand, if (7.13) holds, then

$$P^*(z) \leqslant \prod_{p \leqslant p_0} p \leqslant e^{C \log z} = z^c. \tag{7.14}$$

Hence, by Siegel's theorem [17], any real zero of a character $\bmod P^*(z)$ would satisfy

$$\beta_1 < 1 - \frac{c(\varepsilon)}{z^{c\varepsilon}} < 1 - \frac{c_0}{\log P(z)} \qquad \text{if} \quad \varepsilon < 1/c, \tag{7.15}$$

in contradiction with (7.5).

Another alternative is to use sieve methods to show in place of (7.12) the stronger relation

$$S \ll \frac{y}{p_0} W(z) = o(yW(z)) \tag{7.16}$$

unless

$$p_0 = 0(1) \Longrightarrow P^*(z) = O(1). \tag{7.17}$$

References

[1] A.A. Buchstab, Asymptotic estimates of a general number-theoretic function (Russian), *Mat. Sbornik* (N.S.) **2** (44) (1937), 1239–1246.

[2] H. Cramér, Some theorems concerning prime numbers, *Arkiv f. Math. Astr. Fys.* **15**, No. 5 (1920), 1–33.

[3] H. Cramér, Prime numbers and probability, *Skand. Math. Kongr.* **8** (1935), 107–115.

[4] H. Cramér, On the order of magnitude of the difference between consecutive prime numbers, *Acta Arith.* **2** (1936), 23–46.

[5] H. Davenport, *Multiplicative Number Theory*, Revised by Hugh L. Montgomery, 2nd edition, Springer, Berlin, Heidelberg, New York, 1980.

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

[7] A. Granville, Harald Cramér and the Distribution of Prime Numbers, *Scand. Actuarial J.* 1995, No. 1, 12–28.

[8] A. Granville, Unexpected irregularities in the distribution of prime numbers, in: *Proceedings of the International Congress of Mathematicians*, Vol. 1, 2 (Zürich, 1994), 388–399, Birkhäuser, Basel, 1995.

[9] H. Iwaniec, The sieve of Eratosthenes–Legendre, *Ann Scuola Norm. Sup. Pisa Cl. Sci.* (4) **4** (1977), 257–268.

[10] H. von Koch, Sur la distribution des nombres premiers, *Acta Math.* **24** (1901), 159–182.

[11] J.E. Littlewood, Sur la distribution des nombres premiers, *Comptes Rendus Acad. Sci. Paris* **158** (1914), 1869–1872.

[12] H. Maier, Primes in short intervals, *Michigan Math. J.* **32** (1985), 221–225.

[13] E. Phragmén, Sur le logarithme intégral et la fonction $f(x)$ de Riemann, *Ofversight Kongl. Vet.-Akad. Förh. Stockholm* **48** (1891), 599–616.

[14] J. Pintz, On the remainder term of the prime number formula I. On a problem of Littlewood, *Acta Arith.* **36** (1980), 341–365.

[15] G. Pólya, Heuristic reasoning in the theory of numbers, *Amer. Math. Monthly* **66** (1959), 375–384.

[16] Sz.Gy. Révész, Effective oscillation theorems for a general class of real-valued remainder terms, *Acta Arith.* **49** (1988), 481–505.

[17] C.L. Siegel, Über die Classenzahl quadratischer Zahlkörper, *Acta Arith.* **1** (1935), 83–86.

[18] P. Turán, On the remainder term of the prime number formula I, *Acta Math. Hungar.* **1** (1950), 48–63.

Address: Rényi Mathematical Institute of the Hungarian Academy of Sciences, Budapest, Reáltanoda u. 13–15, H-1053 Hungary
E-mail: pintz@renyi.hu
Received: 12 March 2007