Consecutive primes in short intervals and sums related to Euler's totient function

Artyom Radomskii

"SIMC welcomes postdocs–2020"
Steklov Mathematical Institute of RAS
April 22–23, 2021

1. Acknowledgements

I am grateful to the organizers of the conference for the invitation. Also I would like to thank my colleagues of the Department of Number Theory of Steklov Institute for many useful comments and suggestions. Especially, I would like to express my gratitude to my advisor Sergei Konyagin for posing problems and many useful discussions.

2. Sums related to Euler's totient function

Let us introduce some notation.
$\#A$ — the number of elements of a finite set A.
$(a_1, \ldots, a_n)$ — the greatest common divisor of integers $a_1, \ldots, a_n$.
$\varphi(n)$ — Euler's totient function:

$$\varphi(n) = \#\{1 \le m \le n : \ (m, n) = 1\}.$$

The symbol $b|a$ means that b divides a. For fixed a the sum $\sum_{b|a}$ and the product $\prod_{b|a}$ should be interpreted as being over all positive divisors of a.

We reserve the letter p for primes. In particular, the sum $\sum_{p \le K}$ should be interpreted as being over all prime numbers not exceeding K.

It is clear that $1 \leq \varphi(n) \leq n$ for any positive integer n. Therefore, if $a_1, \ldots, a_N$ are positive integers (not necessary distinct), then for any positive integer s

$$\sum_{n=1}^{N} \left(\frac{a_n}{\varphi(a_n)} \right)^s \geq N. \tag{2.1}$$

Upper bounds for such sums are of interest. A fraction $a/\varphi(a)$ and, more generally, sums (2.1) appear in sieve methods.

We mention

Theorem 2.1 (Shnirelman L., 1933)**.** *Let a be an even positive integer. Then*

$$\#\{p \le x : \ p + a \ \textit{is a prime}\} \le \frac{cx}{(\ln x)^2} \cdot \frac{a}{\varphi(a)},$$

where $c > 0$ is an absolute constant.

If we put $a = 2$ in Theorem 2.1, then we obtain an upper bound for the number of twin primes not exceeding x. We mention

Theorem 2.2 (Chen J., 1973)**.**

$$\#\{p \le x : \ p + 2 = p_1 \ \textit{or}\ p_1 p_2\} \ge \frac{cx}{(\ln x)^2},$$

where $c > 0$ is an absolute constant.

The Twin Primes conjecture, which asserts that

$$\#\{p \le x : \ p+2 \text{ is a prime}\} \to +\infty \quad \text{as } x \to +\infty,$$

is still unsolved.

We prove the following result

Theorem 2.3 (Radomskii A.)**.** *Let α be a real number with $0 < \alpha < 1$. Then there is a number $C(\alpha) > 0$, depending only on α, such that the following holds. Let M be a real number, $a_1, \dots, a_N$ be positive integers (not necessary distinct) with $a_n \le M$ for all $1 \le n \le N$, let s be a positive integer. We define*

$$\omega(d) = \omega(d, \mathcal{A}) = \#\{1 \le n \le N : \ a_n \equiv 0 \pmod{d}\}$$

for any positive integer d. Then

$$\sum_{n=1}^{N} \left(\frac{a_n}{\varphi(a_n)}\right)^s \le \big(C(\alpha)\big)^s \Big(N + \sum_{p \le (\ln M)^\alpha} \frac{\omega(p)(\ln p)^s}{p}\Big).$$

From Theorem 2.3 we obtain the following results.

Corollary 2.1 (Radomskii A.)**.** *Let ε be a real number with $0 < \varepsilon < 1$. Then there is a number $C(\varepsilon) > 0$, depending only on ε, such that the following holds. Let x and z be real numbers with $x \geq 2$ and $(\ln x)^{\varepsilon} \leq z \leq x$. Let $c_0, \ldots, c_k$ be integers with $|c_i| \leq x$ for all $0 \leq i \leq k$ and $c_k \neq 0$. By $\delta := (c_0, \ldots, c_k)$ we denote the greatest common divisor of $c_0, \ldots, c_k$. Let*

$$R(n) = c_k n^k + c_{k-1} n^{k-1} + \ldots + c_0.$$

Let s be a positive integer. Then

$$\sum_{\substack{-z \leq n \leq z \\ R(n) \neq 0}} \left(\frac{|R(n)|}{\varphi(|R(n)|)} \right)^s \leq \left(C(\varepsilon) \frac{\delta}{\varphi(\delta)} \ln(k+1) \right)^s s! z.$$

Corollary 2.2 (Radomskii A.)**.** *Let*

$$R(n) = c_k n^k + c_{k-1} n^{k-1} + \ldots + c_0$$

be a polynomial with integer coefficients, $c_k \neq 0$*. Then there is a number* $C(R) > 0$*, depending only on a polynomial* R*, such that if* s *is a positive integer and* x *is a real number with* $x \geq 1$*, then*

$$\sum_{\substack{-x \leq n \leq x \\ R(n) \neq 0}} \left(\frac{|R(n)|}{\varphi(|R(n)|)} \right)^s \leq \Big(C(R) \Big)^s s! x.$$

Let $\mathcal{L} = \{L_1, \ldots, L_k\}$ be a set of k linear functions with integer coefficients

$$L_i(n) = a_i n + b_i, \ i = 1, \ldots, k.$$

For $L(n) = an + b$, $a, b \in \mathbb{Z}$, we define

$$\Delta_L = |a| \prod_{i=1}^{k} |ab_i - ba_i|.$$

In modern sieve methods sums

$$\sum_{(a,b)\in\Omega} \frac{\Delta_L}{\varphi(\Delta_L)}$$

appear. Here (a, b) denotes a vector and Ω is a finite set in $\mathbb{Z}^2$.

Corollary 2.3 (Radomskii A.)**.** *Let ε be a real number with $0 < \varepsilon < 1$. Then there is a number $C(\varepsilon) > 0$, depending only on ε, such that the following holds. Let x and z be real numbers with $x \geq 2$ and $(\ln x)^{\varepsilon} \leq z \leq x$. Let $a, b_1, \ldots, b_k$ be integers with $a \geq 1$, $|b_i| \leq x$ for all $1 \leq i \leq k$. Let $\mathcal{L} = \{L_1, \ldots, L_k\}$ be a set of k linear functions, where*

$$L_i(n) = an + b_i, \ i = 1, \ldots, k.$$

For $L(n) = an + b$, $b \in \mathbb{Z}$, we define

$$\Delta_L = a^{k+1} \prod_{i=1}^{k} |b_i - b|.$$

Let s be a positive integer. Then

$$\sum_{\substack{-z \leq b \leq z \\ L(n)=an+b \notin \mathcal{L}}} \left(\frac{\Delta_L}{\varphi(\Delta_L)} \right)^s \leq \left(C(\varepsilon) \frac{a}{\varphi(a)} \ln(k+1) \right)^s s! z.$$

Corollary 2.3 extends a result of Maynard [2, Lemma 8.1] which showed the same result but with $s = 1$ and $x^{1/10} \leq z \leq x$.

The following result shows that Theorem 2.3 can not be improved in the following direction: a condition $p \le (\ln M)^{\alpha}$ can not be replaced by a condition $p \le (\ln M)^{o(1)}$.

Theorem 2.4 (Radomskii A.)**.** *Let $\alpha(M)$, $M = 1, 2, \dots,$ be a sequence of positive real numbers such that $\alpha(M) \to 0$ as $M \to +\infty$ and $(\ln M)^{\alpha(M)} \ge 2$ for $M \ge 3$. Then there is a number $M_0 > 0$, depending only on a sequence $\alpha(M)$ such that the following holds. For any positive integer $M \ge M_0$ there is a non-empty set $A \subset \{1, \dots, M\}$ such that*

$$\#\{n \in A : \ n \equiv 0 \pmod{p}\} = 0$$

for any prime $p \le (\ln M)^{\alpha(M)}$ and

$$\sum_{n \in A} \frac{n}{\varphi(n)} \ge \frac{c}{\alpha(M)} \# A.$$

Here $c > 0$ is an absolute constant.

3. Elliptic curves

An *elliptic curve* is given by an equation of the form

$$E: \ y^2 = x^3 + Ax + B, \tag{3.1}$$

with the one further requirement that the *discriminant*

$$\Delta = 4A^3 + 27B^2$$

should not vanish. The discriminant condition ensures that the cubic polynomial $P(x) = x^3 + Ax + B$ has distinct (complex) roots. For convenience, we shall generally assume that the coefficients A and B are integers.

One of the properties that make an elliptic curve E such a fascinating object is the existence of a composition law that allows us to 'add' points to one another. Let us consider the real solutions of (3.1) as points in the plane. Let P and Q be distinct points on E and let L be the line through P and Q. Then the fact that E is given by an equation (3.1) of degree 3 means that L intersects E in three points. Two of these points are P and Q. If we let R denote the third point in $L \cap E$, then the sum of P and Q is defined by

$$P + Q = (\text{the reflection of } R \text{ across the } x\text{-axis}).$$

In order to add P to itself, we let Q approach P, so L becomes the tangent line to E at P. The addition law on E is illustrated in Figure.

We define the negation of a point $P = (x, y)$ to be its reflection across the x-axis

$$-P = (x, -y).$$

The line L through P and $-P$ intersects E in only these two points, so there is no third point R to use in the addition law. To remedy these situation, we adjoin an idealized point $\mathcal{O}$ to the plane. This point $\mathcal{O}$ is called the *point at infinity*. The special rules relating to the point $\mathcal{O}$ are

$$P + (-P) = \mathcal{O} \quad \text{and} \quad P + \mathcal{O} = \mathcal{O} + P = P$$

for all points P on E.

Let $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$ be two points on E, $Q \neq P$, $Q \neq -P$. It can be shown that $P + Q = (x_{P+Q}, y_{P+Q})$, where

$$x_{P+Q} = \left(\frac{y_Q - y_P}{x_Q - x_P}\right)^2 - x_P - x_Q,$$

$$y_{P+Q} = -\frac{y_Q - y_P}{x_Q - x_P}\left(\left(\frac{y_Q - y_P}{x_Q - x_P}\right)^2 - x_P - x_Q\right) - \\ - \frac{y_P x_Q - y_Q x_P}{x_Q - x_P}.$$

If $Q = P$, then (*the duplication formula*)

$$x_{2P} = \frac{x_P^4 - 2Ax_P^2 - 8Bx_P + A^2}{4x_P^3 + 4Ax_P + 4B},$$

$$y_{2P} = -\frac{3x_P^2 + A}{2y_P}\left(\frac{x_P^4 - 2Ax_P^2 - 8Bx_P + A^2}{4x_P^3 + 4Ax_P + 4B}\right) - \\ - \frac{-x_P^3 + Ax_P - 2B}{2y_P}.$$

It can be shown that

$$P + Q = Q + P \quad \text{for all } P, Q \in E.$$

$$(P + Q) + R = P + (Q + R) \quad \text{for all } P, Q, R \in E.$$

Special cases of the duplication and composition law on elliptic curves, described algebraically, date back to Diophantus, but it appears that the first geometric description via secant lines is due to Newton.

If A and B are in a field k and if the coordinates of P and Q are in k, then the coordinates of $P \pm Q$ are also in k. We obtain

Theorem 3.1. *Let E be an elliptic curve given by an equation*

$$y^2 = x^3 + Ax + B,$$

whose coefficients A and B are in a field k and let

$$E(k) = \big\{(x, y) \in k^2 : \ y^2 = x^3 + Ax + B\big\} \cup \{\mathcal{O}\}.$$

Then the sum and difference of two points in $E(k)$ is again in $E(k)$, so $E(k)$ is a commutative group.

Theorem 3.1 was first observed by Poincaré, *Jour. Math. Pures Appl.* 7 (1901).

Let us consider

$$E_1:\ y^2 = x^3 + 7, \qquad E_2:\ y^2 = x^3 - 43x + 166,$$
$$E_3:\ y^2 = x^3 - 2, \qquad E_4:\ y^2 = x^3 + 17.$$

The curve E_1 has no rational points, so $E_1(\mathbb{Q}) = \{\mathcal{O}\}$. $E_2(\mathbb{Q})$ is a finite group with 7 elements

$$E_2(\mathbb{Q}) = \{(3, \pm 8), (-5, \pm 16), (11, \pm 32), \mathcal{O}\}.$$

The group $E_3(\mathbb{Q})$ is freely generated by the single point $P = (3, 5)$, in the sense that every point in $E_3(\mathbb{Q})$ has the form nP for a unique $n \in \mathbb{Z}$. Similarly, the points $P = (-2, 3)$ and $Q = (2, 5)$ freely generate $E_4(\mathbb{Q})$ in the sense that every point in $E_4(\mathbb{Q})$ has the form $mP + nQ$ for a unique pair of integers $m, n \in \mathbb{Z}$. We note that none of these assertions concerning E_1, E_2, E_3, E_4 is obvious.

Repeated addition and negation allows us to 'multiply' points of E by an arbitrary integer m. This function from E to itself is called the *multiplication-by-m* map,

$$\phi_m : E \to E, \quad \phi_m(P) = mP = \text{sign}(m)(P + \cdots + P)$$

(the sum contains $|m|$ terms). By convention, we also define $\phi_0(P) = \mathcal{O}$.

The multiplication-by-m map is defined by rational functions. Maps $E \to E$ defined by rational functions and sending $\mathcal{O}$ to $\mathcal{O}$ are called *endomorphisms* of E. Endomorphisms can be added and multiplied according to the rules

$$(\phi + \psi)(P) = \phi(P) + \psi(P) \text{ and } (\phi\psi)(P) = \phi(\psi(P)),$$

and one can show that with these operations, the set of endomorphisms $\text{End}(E)$ becomes a ring.

For most elliptic curves (over fields of characteristic 0), the only endomorphisms are the multiplication-by-m maps, so for these curves $\text{End}(E) = \mathbb{Z}$. Curves that admit additional endomorphisms are said to have *complex multiplication*. Examples of such curves include

$$E_5 : y^2 = x^3 + Ax,$$

which has the endomorphism $\phi_i(x, y) = (-x, iy)$, and

$$E_6 : y^2 = x^3 + B,$$

which has the endomorphism $\phi_\rho(x, y) = (\rho x, y)$ (here $i = \sqrt{-1}$ and $\rho = e^{(2/3)\pi i}$). These endomorphisms satisfy

$$\phi_i^2(P) = -P \quad \text{and} \quad \phi_\rho^2(P) + \phi_\rho(P) + P = \mathcal{O}.$$

One can show that $\mathrm{End}(E_5)$ is isomorphic to the ring of Gaussian integers (i.e. $a+bi$, $a,b\in\mathbb{Z}$) and that $\mathrm{End}(E_6)$ is the ring of integers in $\mathbb{Q}(\rho)$ (i.e. $a+b\rho$, $a,b\in\mathbb{Z}$).

Given a prime p, by $\mathbb{F}_p$ we denote the field of classes of residues modulo p. Let $\pi(x)$ denote a number of primes not exceeding x. From Theorem 2.3 we obtain

Corollary 3.1 (Radomskii A.)**.** *Let E be an elliptic curve given by an equation*

$$y^2=x^3+Ax+B,$$

where A and B are integers with $\Delta=4A^3+27B^2\neq 0$. Suppose that E does not have complex multiplication. Let s be a positive integer and x be a real number with $x\geq 2$. Then

$$\pi(x)\leq\sum_{p\leq x}\left(\frac{\#E(\mathbb{F}_p)}{\varphi(\#E(\mathbb{F}_p))}\right)^s\leq C(E,s)\pi(x),$$

where $C(E,s)>0$ is a number, depending only on an elliptic curve E and a number s.

To apply Theorem 2.3 we need an upper bound for $\omega(p)$. We used a result of David and Wu [1], which is the following. Suppose that E does not have complex multiplication. Then for integers a and $t\geq 1$ we have

$$\#\{p\leq x:\#E(\mathbb{F}_p)\equiv a\pmod t\}\leq$$
$$\leq C(E)\Big(\frac{\pi(x)}{\varphi(t)}+x\cdot\exp(-ct^{-2}\sqrt{\ln x})\Big),$$

if $\ln x\geq t^{12}\ln t$. Here $c>0$ is an absolute constant, $C(E)>0$ is a number depending only on an elliptic curve E.

4. Consecutive primes in short intervals: an upper bound

Let p_n denote the n^{th} prime.

We shall use the notation of I. M. Vinogradov: the symbol $A \ll_{m,\alpha} B$ means that $|A| \leq c(m,\alpha)B$, where $c(m,\alpha) > 0$ is a number depending only on m and α. We prove

Theorem 4.1 (Radomskii A.)**.** *Let m be a positive integer, α and x be real numbers with $\alpha > 0$ and $x \geq 3$. Then*

$$\sum_{x/2<p_n\leq x} \frac{1}{(p_{n+m}-p_n)^\alpha} \ll_{m,\alpha} \begin{cases} x/(\ln x)^{\alpha+1}, & \text{if } \alpha < m; \\ (x \ln\ln x)/(\ln x)^{m+1}, & \text{if } \alpha = m; \\ x/(\ln x)^{m+1}, & \text{if } \alpha > m. \end{cases}$$

Theorem 4.2 (Radomskii A.)**.** *Let m be a positive integer, α and x be real numbers with $\alpha > 0$ and $x \geq 3$. Then the series*

$$\sum_{p_n > x} \frac{1}{p_n (p_{n+m} - p_n)^{\alpha}}$$

is convergent and

$$\sum_{p_n > x} \frac{1}{p_n (p_{n+m} - p_n)^{\alpha}} \ll_{m,\alpha} \begin{cases} 1/(\ln x)^{\alpha}, & \text{if } \alpha < m; \\ \ln\ln x/(\ln x)^{m}, & \text{if } \alpha = m; \\ 1/(\ln x)^{m}, & \text{if } \alpha > m. \end{cases}$$

Theorem 4.2 with $\alpha = 1$, $m = 2$, and $\alpha = m = 1$ was proved by S. Konyagin, H. Queffelec, E. Saksman, and K. Seip.

Theorem 4.3 (Radomskii A.)**.** *Let m be a positive integer. Then there is a number $c(m) > 0$, depending only on m, such that if x and y are real numbers with $x \geq 3$ and $y > 0$, then*

$$\#\{x/2 < p_n \leq x : p_{n+m} - p_n \leq y\} \leq$$
$$\leq c(m)\pi(x)\Big(\frac{y}{\ln x}\Big)^m \ln\ln x.$$

5. Consecutive primes in short intervals: a lower bound

We prove the following result.

Theorem 5.1 (Radomskii A.)**.** *There are positive absolute constants c and C such that the following holds. Let ε be a real number with $0 < \varepsilon < 1$. Then there is a number $c_0(\varepsilon) > 0$, depending only on ε, such that if $x \in \mathbb{R}$, $y \in \mathbb{R}$, $m \in \mathbb{Z}$, $q \in \mathbb{Z}$, $a \in \mathbb{Z}$ are such that*

$$c_0(\varepsilon) \le y \le \ln x,$$

$$1 \le m \le c \cdot \varepsilon \ln y, \quad 1 \le q \le y^{1-\varepsilon}, \quad (a, q) = 1,$$

then

$$\#\{x/2 < p_n \le x : \ p_n \equiv \cdots \equiv p_{n+m} \equiv a \pmod q,$$

$$p_{n+m} - p_n \le y\} \ge \pi(x) \left(\frac{y}{2q \ln x} \right)^{\exp(Cm)}.$$

Theorem 5.1 extends a result of Maynard [2, Theorem 3.3] which showed the same result but with $y = \varepsilon \ln x$.

From Theorem 5.1 we obtain

Corollary 5.1 (Radomskii A.)**.** *There is an absolute constant $C > 0$ such that if m is a positive integer, x and y are real numbers satisfying $\exp(Cm) \leq y \leq \ln x$, then*

$$\#\{x/2 < p_n \leq x : p_{n+m} - p_n \leq y\} \geq \\ \geq \pi(x) \left(\frac{y}{2 \ln x}\right)^{\exp(Cm)}.$$

References

[1] C. David and J. Wu, Pseudoprime reductions of elliptic curves, *Canad. J. Math.* **64** (2012), no. 1, 81–101.

[2] J. Maynard, Dense clusters of primes in subsets, *Compositio Mathematica*, **152** (2016), 1517–1554.

Thank you for your attention!