The Sequence of Totient Values on Short Intervals

Alexia Ascott, Supervised by Joni Teräväinen

Although the behaviour of sequences with some "multiplicative structure" is well understood on long intervals (such as those of the form $[X, 2X]$), their behaviour on short intervals is significantly more difficult to understand. In 2016, Matomäki and Radziwill showed in groundbreaking work that multiplicative functions have the same behaviour on typical very short intervals as they have on long intervals [1]. Their proof method has since then been applied by various authors to show that many sequences having some "multiplicative structure" behave in the expected manner in typical very short intervals. This has been done for almost primes (Teräväinen, 2016) [2] and for several other sequences.

One very interesting sequence for which the method has not yet been applied is the sequence of totient values (numbers that are values of the Euler totient function). In particular, an open problem due to Igor Shparlinski is to show that one can find totient values on significantly shorter intervals than on which one can unconditionally find primes (where the record is due to Baker, Harman and Pintz, who showed that $[X, X+X^{0.525}]$ contains primes for all large X). The goal of this project was to make progress on Shparlinski's problem, getting close to the interval length $X^{1/2}$, below which one cannot reasonably hope to get with current methods.

In order to prove that, for any $\epsilon > 0$, there exists at least one totient value in the interval $[X, X+X^{1/2+\epsilon}]$ for sufficiently large X, it suffices to show that this interval contains at least one integer of the form $(p_1-1)(p_2-1)(p_3-1)$, where p_1, p_2, p_3 are distinct primes such that $p_1 - 1 \sim P_1$, $p_2 - 1 \sim P_2$ and $P_2 \le p_3 - 1 \le 8P_2$, with $P_1 = X^{\epsilon}$ and $P_2 = \sqrt{X/P_1}/2$ (note that $n \sim N$ denotes $N \le n \le 2N$).

We let $$S_{h_j} = \sum_{\substack{X \le n \le X+h_j \\ n=(p_1-1)(p_2-1)(p_3-1) \\ p_1-1\sim P_1,\, p_2-1\sim P_2,\, P_2 \le p_3-1 \le 8P_2}} 1 \;=\; \sum_{\substack{X \le n \le X+h_j \\ n=(p_1-1)(p_2-1)(p_3-1) \\ p_1-1\sim P_1,\, p_2-1\sim P_2}} 1$$ for $j = 1, 2$ with $h_1 = X^{\frac{1}{2}+\epsilon}$, $h_2 = \dfrac{X}{\log^b X}$

(so we want to show that $S_{h_1} \ge 1$ for all sufficiently large X). We can relatively easily prove that $S_{h_2} \gg \dfrac{h_2}{\log^3 X}$ using the Prime Number Theorem with classical error term. Hence, it suffices to show $\left|\dfrac{1}{h_1}S_{h_1} - \dfrac{1}{h_2}S_{h_2}\right| = o\left(\dfrac{1}{\log^3 X}\right)$ so that $S_{h_1} \gg \dfrac{h_1}{\log^3 X}$ (note S_{h_1} is dominated by those integers with p_1, p_2 and p_3 distinct). In fact, the idea of relating "short averages" to "long averages" is central to the Matomäki-Radziwill method, and is a key step in proving results concerning short intervals.

An effective way to relate $\dfrac{1}{h_1}S_{h_1}$ and $\dfrac{1}{h_2}S_{h_2}$ (and short and long averages in general) is to use Perron's formula (see e.g. [3, Corollary 5.3]), which gives (taking $T = X \log^a X/h_1$ for some $a > 0$);

$$S_{h_j} = \frac{1}{2\pi}\int_{-T}^{T} F(1+it)\frac{(X+h_j)^{1+it} - X^{1+it}}{1+it}\,dt + O(X/T), \text{ where } F(s) = \sum_{\substack{n=(p_1-1)(p_2-1)(p_3-1) \\ p_1-1\sim P_1,\, p_2-1\sim P_2,\, P_2 \le p_3-1 \le 8P_2}} \frac{1}{n^s}.$$

To finish the proof, we split the integral over the intervals $[-T_0, T_0]$ and $[-T, T]\backslash[-T_0, T_0]$, where $T_0 = \log^d X$ for some $d > 0$. For the former interval we can bound $\left|\dfrac{1}{h_1}S_{h_1} - \dfrac{1}{h_2}S_{h_2}\right|$ directly (using Taylor's theorem), whereas for latter it suffices to bound $\left|\dfrac{1}{h_j}S_{h_j}\right|$. By applying the Cauchy-Schwarz inequality and a mean value theorem for Dirichlet polynomials (see theorem 9.1 of [4]), this reduces to bounding $\sum_{n\sim P_1} \mathbb{1}_{\{n+1\in\mathbb{P}\}}/n^{1+it}$, which can be done using Vaughan's identity and several lemmas in [4] on bounding exponential sums.

Therefore, for any $\epsilon > 0$, there exists at least one totient value in the interval $[X, X+X^{1/2+\epsilon}]$, as desired. This gives a result as strong as what was known under the Riemann hypothesis for the primes, but for the set of totient values (thus resolving Shparlinski's problem). Moreover, the proof above easily generalises to the following result: For any integer $k \ge 3$, any constants $c_1, c_2, \ldots, c_k$, and any $\epsilon > 0$, there exists an integer in $[X, X+X^{1/2+\epsilon}]$ of the form $(p_1+c_1)(p_2+c_2)\ldots(p_k+c_k)$, where $p_1, p_2, \ldots, p_k$ are distinct primes.

The next step is to now shorten the interval to $[X, X+\sqrt{X}\log^c X]$ for some positive constant c, and then to optimise the value of c. This is very likely the best result we can hope to obtain using our method. Good progress has been made towards this result and we hope to prove it soon.

References:

[1] K. Matomäki and M. Radziwill. Multiplicative functions in short intervals. Ann. of Math. (2),183(3):1015–1056, 2016.

[2] J. Teräväinen. Almost primes in almost all short intervals. Math. Proc. Cambridge Philos. Soc., 161(2):247–281, 2016.

[3] H. L. Montgomery and R. C. Vaughan. Multiplicative number theory. I. Classical theory, volume 97 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2007.

[4] H. Iwaniec and E. Kowalski. Analytic number theory, volume 53 of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 2004.