A BOUND FOR PROGRESSIONS OF LENGTH k IN THE PRIMES

BEN GREEN AND TERENCE TAO

ABSTRACT. From our paper in [1], we extract a quantitative upper bound for the size of the first arithmetic progression of primes of length k; the argument shows that those primes are bounded by an exponential tower in k of height seven.

1. NOTATION

We assume complete familiarity with our paper [1], and all equations, Lemmas, etc. refer to that paper.

We use $x \uparrow y$ as shorthand for x^y. Similarly $x \uparrow y \uparrow z$ for x^{y^z}, etc.

We will take $w = w(k)$ to be a fixed large number depending only on k (as opposed to depending on N as is done in the main argument). The quantity W is basically e^w. The quantity $W/\phi(W)$ is basically $\log w$.

2. QUANTITATIVE BOUNDS FOR THE GOLDSTON-YILDIRIM ESTIMATES

The purpose of this section is to quantify the dependence of constants in Propositions 8.5 and 8.6. For us the parameters m and t are exponential in k, thus at most $2 \uparrow O(k)$; we can thus set R to be $N \uparrow 2 \uparrow -O(k)$, and σ to be $2 \uparrow -O(k)$. In Lemma 10.1 we can then take $\beta = 2 \uparrow -O(k)$.

In Lemma 10.2, the quantity B is always $O(1)$. we can take δ to be $\sqrt{\beta}/4$, and thus $\delta = 2 \uparrow -O(k)$. The factor of $(\beta \log R)^{B/2}$ can then be absorbed as long as N is sufficiently large, e.g. $N \gg 2 \uparrow 2 \uparrow 2 \uparrow O(k)$ will certainly work here; in that case we can take the $O_B(1)$ bound in (10.2) to just be $O(1)$. The bound in (10.3) is also $O(1)$.

In Lemma 10.3, the bounds $O_m(1)$ can in fact just be $O(1)$ by taking N large again, for instance $N \gg 2 \uparrow 2 \uparrow 2 \uparrow O(k)$.

In Lemma 9.4, the factors of $O_{m,\delta}(1)$ that we lose are at most $\delta^{-O(m)}$, which is $2 \uparrow 2 \uparrow O(k)$. The second factor of $O_{m,\delta}(1)$ can thus be absorbed by shrinking δ and assuming $N \gg 2 \uparrow 2 \uparrow 2 \uparrow O(k)$.

Now we begin the inspection of the proof of Proposition 8.5. The first large constants appear in Lemma 9.3. Here the C^m bounds on G_1 are of the form $m^{O(m)}$, which is $2 \uparrow 2 \uparrow O(k)$. The $o_m(1)$ error is of the form $m^{O(m)}O(w^{-1/O(m)})$, which is at most $w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$ if we assume $w \gg 2 \uparrow 2 \uparrow 2 \uparrow O(k)$. The bound on G_2 is like $O(mw) \uparrow O(mw)$, which is $2 \uparrow w \uparrow O(1)$ again taking $w \gg 2 \uparrow 2 \uparrow 2 \uparrow O(k)$. Putting this together we see that the C^m bounds for G_1 are also of the form $2 \uparrow w \uparrow O(1)$; this allows one to make the hypothesis (9.10) obeyed with bounded constants if $N \gg 2 \uparrow w \uparrow O(1)$.

Of the two error terms in Lemma 9.4, the first is the larger one, contributing $O(\|G\|_{C^m}/\log R)$ to the $o_{m,t}(1)$ term in (9.2). By the preceding estimates, this error is basically $(2 \uparrow$

$w \uparrow O(1))/\log N$, which will be certainly smaller than $w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$ if $N \gg 2 \uparrow 2 \uparrow w \uparrow O(1)$. The other error term to (9.1) is much smaller for this range of N and w and can thus be ignored. We have thus obtained a quantitative error term of $w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$ assuming that $w \gg 2 \uparrow 2 \uparrow 2 \uparrow O(k)$ and $N \gg 2 \uparrow 2 \uparrow w \uparrow O(1)$.

Now we turn to Proposition 8.6. In (9.12), the constants $O_m(1)$ are of the form $m \uparrow O(m) = 2 \uparrow 2 \uparrow O(k)$. Similarly in (9.13), but we can eliminate those constants by taking N large (we are already needing $N \gg 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow O(k)$, and this is more than sufficient). Now consider the three terms used to estimate (9.5) at the end of the argument. The third term is very small compared to the second and will be ignored. The second term contributes an error of the form $(2 \uparrow w \uparrow O(1))/\log\log N$, which will be $O(1)$ if $N \gg 2 \uparrow 2 \uparrow 2 \uparrow w \uparrow O(1)$. We can then replace the right-hand side of (8.1) by

$$O(1)(\frac{W \log R}{\phi(W)})^m \prod_{p|\Delta}(1 + (2 \uparrow 2 \uparrow O(k))p^{-1/2})$$

(estimating the $o_m(1)$ term crudely by $O(1)$).

Note that we already need

$$N \gg 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow O(k)$$

and

$$w \gg 2 \uparrow 2 \uparrow 2 \uparrow O(k).$$

Finally, we observe that the density $\delta := k^{-1}2^{-k-5}\epsilon_k$ we need in the proof of Theorem 1.1 is $2 \uparrow -k \uparrow O(1)$, so by Gowers' estimate, the quantity $c(k, \delta)$ is

$$c(k, \delta) = 2 \uparrow -2 \uparrow 2 \uparrow 2 \uparrow O(k).$$

Thus the ε we need in the proof of Theorem 3.5 is also

$$\varepsilon = 2 \uparrow -2 \uparrow 2 \uparrow 2 \uparrow O(k),$$

and the maximal number of steps K_0 in the Furstenberg tower is

$$K_0 = 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow O(k).$$

3. Pseudorandomness bounds

Now we can quantify the pseudorandomness of ν. We begin with the linear forms condition. We are certainly taking w large enough so that $\sqrt{w}/2 > (k + 1)!$. All the $O_{m,t}(1)$ bounds are at most $2 \uparrow 2 \uparrow O(k)$, and if we take $Q := w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$, one can then replace the $o(1)$ error in the linear forms condition by $w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$.

Now we look at Lemma 8.9. As computed earlier, the $O_m(1)$ constants are $2 \uparrow 2 \uparrow O(k)$, which effectively replaces q by $\tilde{q} := (2 \uparrow 2 \uparrow O(k))q$ (this is the $O_m(q)$ expression in the text). The final $O_{m,q}(1)$ bound is polynomial in $2 \uparrow \tilde{q} \uparrow O(1)$, and thus

$$\mathbb{E}(\tau^q) \leqslant 2 \uparrow 2 \uparrow 2 \uparrow O(k) + 2 \uparrow q \uparrow O(1).$$

4. Uniformity and anti-uniformity

Now we quantify Sections 5,6. The $o(1)$ bound in Lemma 5.2 comes from the linear forms condition and is thus $w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$. Similarly for the Von Neumann theorem, Proposition 5.3, as well as the uniform bound in (6.5) (at this point w is so huge that the exponential factors of k appearing are negligible).

Now look at Lemma 6.3. Assume the coefficients of P are bounded by some M; then we lose a factor of Md^K in order to reduce to the dK-monomial. Applying the correlation condition with $q = dK$ we thus see that

$$O_{K,d,P}(1) \leqslant Md^K(2 \uparrow 2 \uparrow 2 \uparrow O(k) + 2 \uparrow dK \uparrow O(1)).$$

Now turn to Lemma 6.5. The bound for (6.9) is of the form $O(\varepsilon^{-2K}\sigma^{1/100})$. Next, we observe that the function Φ is essentially the Kth tensor power of a σ-smoothed out indicator of the unit cube, rescaled by ε. One can approximate this fairly well (up to errors of $O(K)\sigma$) by a polynomial of degree $d = O(K\sigma^{-O(1)})$ and coefficients $M = 2 \uparrow O(K\sigma^{-O(1)})$. Thus

$$\mathbb{E}((\nu - 1)\Phi) \leqslant O(K)\sigma + Md^K(2 \uparrow 2 \uparrow 2 \uparrow O(k) + 2 \uparrow dK \uparrow O(1))w \uparrow -1/(2 \uparrow 2 \uparrow O(k)).$$

Write $K' := \max(K, 2 \uparrow 2 \uparrow O(k))$, then this simplifies a bit to

$$\mathbb{E}((\nu - 1)\Phi) \leqslant O(K')\sigma + 2 \uparrow ((K')^{O(1)}\sigma^{-O(1)}) \times w \uparrow -1/(2 \uparrow 2 \uparrow O(k)).$$

If one then sets $\sigma := (K')^{O(1)}(\log \log N)^{-c}$, we thus obtain

$$\mathbb{E}((\nu - 1)\Phi) \leqslant (K')^{O(1)}w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$$

and thus the bound in (6.10) is also of the form $(K')^{O(1)}w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$, while the bound in (6.9) is $O(\varepsilon^{-2K'}(K')^{O(1)}w \uparrow -1/(2 \uparrow 2 \uparrow O(k)))$. Inserting our bounds for ε and K_0 from the Furstenberg argument, we thus obtain a bound of

$$(2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow O(k))w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$$

for (6.10) and

$$(2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow O(k))w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$$

for (6.9). Both of these are at most $w \uparrow -1/(2 \uparrow 2 \uparrow O(k))$ if we assume

$$w \gg 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow O(k)$$

and hence

$$N \gg 2 \uparrow 2 \uparrow w \uparrow O(1) \gg 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow\uparrow O(k).$$

This is also ensures that the errors $o_{j,\varepsilon}(1)$ in the Furstenberg tower construction are manageable (i.e. smaller than ε). Also with N this large it is clear that $c(k, \delta)N^2$ is much larger than N, and so we have shown that a non-trivial progression of primes of length k exists with all primes less than

$$O(NW) \leqslant 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow 2 \uparrow O(k).$$

It is likely that the implicit constant in $O(k)$ is quite manageable, e.g. $100k$ will certainly suffice, basically because almost all of the $O(1)$ bounds encountered in the above analysis will eventually be hit with several logarithms.

REFERENCES

[1] B.J. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions,* preprint.

PACIFIC INSTITUTE FOR THE MATHEMATICAL SCIENCES, ROOM 205, 1933 WEST MALL, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER BC, CANADA,
  E-mail address: bjg23@hermes.cam.ac.uk

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT LOS ANGELES, LOS ANGELES CA 90095
  E-mail address: tao@math.ucla.edu