AN ALTERNATIVE TO VAUGHAN'S IDENTITY

ANDREW GRANVILLE

ABSTRACT. We exhibit an identity that plays the same role as Vaughan's identity but is arguably simpler.

1. Introduction

Let 1_y denote the characteristic function of the integers free of prime factors $\leqslant y$, The idea is to work with the identity

$$\Lambda(n) = \log n - \sum_{\substack{\ell m = n \\ \ell, m > 1}} \Lambda(\ell), \tag{1.1}$$

summing it up over integers $n \leqslant x$ for which $1_y(n) = 1$, where we might select $y := \exp(\sqrt{\log x})$ or larger. In this case the second sum can be written a sum of terms $1_y(\ell)\Lambda(\ell) \cdot 1_y(m)$ which have the bilinear structure that is used in "Type II sums". We will see the identity in action in two key results in analytic number theory:

2. The Bombieri-Vinogradov Theorem

Theorem 1 (The Bombieri-Vinogradov Theorem). *If $x^{1/2}/(\log x)^B \leqslant Q \leqslant x^{1/2}$ then*

$$\sum_{q \leqslant Q} \max_{(a,q)=1} \left| \pi(x;q,a) - \frac{\pi(x)}{\phi(q)} \right| \ll Q x^{1/2} (\log\log x)^{1/2}. \tag{2.1}$$

This is a little stronger than the results in the literature (for example Davenport [2] has the $(\log\log x)^{1/2}$ replaced by $(\log x)^5$). The reason for this improvement is the simplicity of our identity, and some slight strengthening of the auxiliary results used in the proof.

Proof. Let $y = x^{1/\log\log x}$. We will instead prove the following result, in which the ψ function replaces π, and deduce (2.1) by partial summation:

$$\sum_{q \leqslant Q} \max_{(a,q)=1} \left| \psi(x;q,a) - \frac{\psi(x)}{\phi(q)} \right| \ll Q x^{1/2} (\log x)(\log\log x)^{1/2}. \tag{2.2}$$

Using (1.1) for integers n with $1_y(n) = 1$, the quantity on the left-hand side of (2.2) is $\leqslant S_I + S_{II} + E$ where

$$S_I = \sum_{q \leqslant Q} \max_{(a,q)=1} \left| \sum_{\substack{n \leqslant x \\ n \equiv a \pmod q}} 1_y(n) \log n - \frac{1}{\phi(q)} \sum_{\substack{n \leqslant x \\ (n,q)=1}} 1_y(n) \log n \right|$$

1991 *Mathematics Subject Classification.* 11N13.

Key words and phrases. Large sieve. Bombieri-Vinogradov Theorem. Type II sums. Sieve identity.

The author has received funding from the European Research Council grant agreement n° 670239, and from NSERC Canada under the CRC program.

which is $\ll xu^{-u+o(u)} \ll_A \frac{x}{(\log x)^A}$ by the small sieve, where $x/Q = y^u$; and E is the contribution of the powers of primes $\leqslant y$, which contribute $\leqslant \pi(y)\log x$ to each sum and therefore $\leqslant Q\pi(y)\log x \ll_A \frac{x}{(\log x)^A}$ in total. Most interesting is

$$S_{II} = \sum_{q\leqslant Q} \max_{a:\ (a,q)=1} \left| \sum_{n\equiv a \pmod q} f(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} f(n) \right|$$

where $f(n) = \sum_{\ell m=n, \ell,m>y} \Lambda(\ell) 1_y(\ell)\cdot 1_y(m)$. Its bilinearity means that this is a Type II sum, and we can employ the following general result.[1]

Theorem 2. *For each integer $n \leqslant x$ we define*

$$f(n) := \sum_{\ell m = n} \alpha_\ell \beta_m$$

where $\{\alpha_\ell\}$ and $\{\beta_m\}$ are sequences of complex numbers, for which

- *The $\{\alpha_\ell\}$ satisfy the Siegel-Walfisz criterion;*
- *The $\{\alpha_\ell\}$ are only supported in the range $L_0 \leqslant \ell \leqslant x/y$;*
- *$\sum_{\ell\leqslant L} |\alpha_\ell|^2 \leqslant aL$ and $\sum_{m\leqslant M} |\beta_m|^2 \leqslant bM$ for all $L, M \leqslant x$.*

For any $B > 0$ we have

$$\sum_{q\leqslant Q} \max_{a:\ (a,q)=1} \left| \sum_{n\equiv a \pmod q} f(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} f(n) \right| \ll (ab)^{1/2} Q x^{1/2} \log x, \quad (2.3)$$

where $Q = x^{1/2}/(\log x)^B$, with $x/y \leqslant \frac{Q^2}{(\log x)^2}$ and $L_0 \geqslant y, \exp((\log x)^\epsilon)$.

We deduce that $S_{II} \ll Qx^{1/2}(\log x)(\log\log x)^{1/2}$ by Theorem 2 since $a \ll \log x$ and $b \ll \frac{1}{\log y} = \frac{\log\log x}{\log x}$. Then (2.2) follows from which we deduce the result. □

3. A couple of remarks

One can further deduce that

$$\sum_{q\leqslant Q} \max_{y\leqslant x} \max_{(a,q)=1} \left| \pi(y;q,a) - \frac{\pi(y)}{\phi(q)} \right| \ll Q^{1/2} x^{3/4} (\log\log x)^{1/2}. \quad (3.1)$$

To prove this we select N to be the nearest integer to $Q^{-1/2}x^{1/4} \ll (\log x)^{B/2}$, and then let $y_k = \frac{k}{N}x$ for $k = 0, 1, \ldots, N$. For each q select y for which the error term $|\pi(y;q,a) - \frac{\pi(y)}{\phi(q)}|$ is maximal, and $k = [Ny/x] + 1$, in which case this error term is

$$\leqslant \left| \pi(y_k;q,a) - \frac{\pi(y_k)}{\phi(q)} \right| + (\pi(y_k;q,a) - \pi(y;q,a)) + \frac{\pi(y_k) - \pi(y)}{\phi(q)}$$

Since $y_k - y \leqslant \frac{x}{N}$ the second two terms are $\ll \frac{1}{\phi(q)} \cdot \frac{x/N}{\log x}$ by the Brun-Titchmarsh Theorem. Summing over q, and then summing(2.1) for $x = y_k$ for each $k \leqslant N$, we deduce (3.1).

These methods do not allow one to take $Q > x^{1/2}$. This has been achieved by Bombieri, Friedlander and Iwaniec [1] for a fixed with a weaker error term, and for larger Q but with restricted moduli q in Yitang Zhang's famous work [5]. Maynard has recently announced some further improvements.

[1]This can be obtained by taking the ideas in proving (5) of chapter 28 of [2], along with the method of proof of Theorem 9.16 of [3]; in any case it is only a minor improvement on either of these results. For full details see chapter 59 of [4].

4. A general bound for a sum over primes

Proposition 1. *For any given function $F(.)$ and $y \leqslant x$ we have*

$$\left|\sum_{\substack{n\leqslant x\\ p(n)>y}} \Lambda(n)F(n)\right| \ll S_I \log x + (S_{II}\, x(\log x)^5)^{1/2}$$

where S_I is the Type I sum given by

$$S_I := \max_{t\leqslant x}\left|\sum_{\substack{n\leqslant t\\ p(n)>y}} F(n)\right| \leqslant \sum_{\substack{d\geqslant 1\\ P(d)\leqslant y}}\left|\sum_{m\leqslant t/d} F(dm)\right|,$$

and S_{II} is the Type II sum given by

$$S_{II} := \max_{\substack{y<L\leqslant x/y\\ y<m\leqslant 2x/L}} \sum_{m/2<n\leqslant 2m}\left|\sum_{\substack{L<\ell\leqslant 2L\\ \ell\leqslant \frac{x}{m},\frac{x}{n}}} F(\ell m)\overline{F(\ell n)}\right|$$

This simplifies, and slightly improves chapter 24 of [2], which is what is used there to bound exponential sums over primes.

Proof. We again use (1.1) so that

$$\sum_{\substack{n\leqslant x\\ p(n)>y}} \Lambda(n)F(n) = \sum_{\substack{n\leqslant x\\ p(n)>y}} F(n)\log n - \sum_{\substack{\ell\\ p(\ell)>y}} \Lambda(\ell)\sum_{\substack{m\leqslant x/\ell\\ p(m)>y}} F(\ell m).$$

where $p(n)$ denotes the smallest prime factor of n. Now

$$\sum_{\substack{n\leqslant x\\ p(n)>y}} F(n)\log n = \sum_{\substack{n\leqslant x\\ p(n)>y}} F(n)\int_1^n \frac{dt}{t} = \int_1^x \sum_{\substack{t\leqslant n\leqslant x\\ p(n)>y}} F(n)\frac{dt}{t} \leqslant 2\log x\cdot\max_{t\leqslant x}\left|\sum_{\substack{n\leqslant t\\ p(n)>y}} F(n)\right|.$$

Moreover for $P = \prod_{p\leqslant y} p$,

$$\sum_{\substack{n\leqslant t\\ p(n)>y}} F(n) = \sum_{n\leqslant t} F(n)\sum_{d|P,d|n}\mu(d) = \sum_{d|P}\mu(d)\sum_{m\leqslant t/d} F(dm).$$

For the second sum we first split the sums into dyadic intervals ($L<\ell\leqslant 2L$, $M<m\leqslant 2M$) and then Cauchy, so that the square of each subsum is

$$\leqslant \sum_{\ell:p(\ell)>y}\Lambda(\ell)^2\cdot\sum_{\ell}\left|\sum_{\substack{m\leqslant x/\ell\\ p(m)>y}} F(\ell m)\right|^2 \ll L\log L\sum_{\substack{M<m,n\leqslant 2M\\ p(m),p(n)>y}}\ \sum_{\ell\leqslant\frac{x}{\max\{m,n\}}} F(\ell m)\overline{F(\ell n)}$$

$$\ll x\log x\cdot\max_{M<m\leqslant 2M}\sum_{\substack{m/2<n\leqslant 2m\\ p(n)>y}}\left|\sum_{\substack{L<\ell\leqslant 2L\\ \ell\leqslant\frac{x}{m},\frac{x}{n}}} F(\ell m)\overline{F(\ell n)}\right|$$

since $m,n\in(M,2M]$, and the result follows. □

5. Genesis

The idea for using (1.1) germinated from reading the proof of the Bombieri-Vinogradov Theorem (Theorem 9.18) in [3], in which Friedlander and Iwaniec used Ramaré's identity, that if $\sqrt{x} < n \leqslant x$ and n is squarefree then

$$1_{\mathbb{P}}(n) = 1 - \sum_{\substack{pm=n \\ p \text{ prime} \leqslant \sqrt{x}}} \frac{1}{\omega_{\sqrt{x}}(m)}$$

where $1_{\mathbb{P}}$ is the characteristic function for the primes, and $\omega_z(m) = 1 + \sum_{p|m,\ p\leqslant z} 1$. They also had to sum this over all integers free of prime factors $> y$.

6. Uniformity

Although this method to employ type II sums on questions about prime number asymptotics is somewhat easier than those already in the literature, it seems harder to incorporate strong estimates for various exponential sums to get better and more uniform results in wide ranges.

References

[1] E. Bombieri, J.B. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli. III*, J. Amer. Math. Soc. **2** (1989), 215–224.

[2] H.M. Davenport, *Multiplicative number theory* (3rd ed), Graduate Texts in Mathematics, vol. 74, Springer-Verlag, New York, vol. 74, 2000.

[3] J.B. Friedlander and H. Iwaniec, *Opera de Cribro*, AMS Colloquium Publications vol. 57, 2010.

[4] Andrew Granville, *Analytic number theory revealed: The distribution of prime numbers*, American Mathematical Society (to appear)

[5] Yitang Zhang, *Bounded gaps between primes*, Ann. of Math. **179** (2014), 1121–1174.

Départment de Mathématiques et Statistique, Université de Montréal, CP 6128 succ Centre-Ville, Montréal, QC H3C 3J7, Canada; and Department of Mathematics, University College London, Gower Street, London WC1E 6BT, England.

Email address: andrew@dms.umontreal.ca