Some minor arcs estimates related to the paper "Roth's theorem in the primes".

This set of notes is intended to supply details of some estimates required in the paper [2] of the title. It is normal, when writing a paper, to prove the results contained therein completely. In this case, however, we found that the estimates we needed were very close to results contained in [1, 3, 4]. The estimates are of a type which are likely to be "clear" to most analytic number theorists reading [2], and of the sort that will probably not interest any analysts reading that paper. This, together with the opportunity to practise nested subscripts in LaTeX, is why we have chosen to record them separately here.

The estimates we discuss here could be useful to anyone who has an additive problem concerning the primes restricted to an arithmetic progression $p \equiv b (\operatorname{mod} m)$, so long as no attempt is being made to obtain results with a strong dependence on m.

The treatment that follows is brief and rough. Furthermore we have made no effort to ensure that the present notes can be read independently of [2].

1. Two preliminary lemmata. We begin with two estimates of a type which were described in [2] as being "of a standard type".

Lemma 1 *Let K be a positive integer, and suppose that $q \geq 100K$, $(a, q) = 1$ and $|\theta - a/q| \leq K/q^2$. Then*

$$\sum_{1 \leq n \leq X} \min(Y, \|\theta n\|^{-1}) \ \ll\ K\left(\frac{XY}{q} + Y + (X+q)\log q\right).$$

Proof. Observe that for any j the numbers

$$(j+1)\theta, \ldots, (j + \lfloor q/2K \rfloor)\theta \tag{1}$$

are $1/2q$-separated modulo 1. Indeed if they were not we should be able to find an integer l, $0 < |l| \leq \lfloor q/2K \rfloor$, with $\|l\theta\| < 1/2q$. But then $|l\theta - al/q| \leq |l|/q^2 \leq 1/2q$, and $al \neq 0 \pmod q$, which is impossible. Divide the sum over $1 \leq n \leq X$ into at most $1 + 4XK/q$ ranges of type (1). Over each such range R, we have

$$\begin{aligned}\sum_{n \in R} \min(Y, \|\theta n\|^{-1}) \ &\leq\ \sum_{i=0}^{q} \min(Y, 2q/i) \\ &\leq\ Y + q\log q.\end{aligned}$$

Thus the whole sum over $1 \leq n \leq X$ is $\ll (1 + XK/q)(Y + q\log q)$, a bound of the form stated. □

Lemma 2 *Let K be a positive integer, and suppose that $q \geq 100K$, $(a, q) = 1$ and $|\theta - a/q| \leq K/q^2$. Then*

$$\sum_{1 \leq n \leq V} \min(U/n, \|\theta n\|^{-1}) \ \ll\ K^2 \log V \log q \left(\frac{U}{q} + V + q\right).$$

Proof. Divide the sum over n into the range $1 \leq n \leq \lfloor q/2K \rfloor$ plus dyadic ranges $2^i \leq n < 2^{i+1}$ for 2^i ranging between about q/K and V. The numbers $\|\theta n\|$, $n = 1, \dots, \lfloor q/2K \rfloor$, are $1/2q$-separated and so one has

$$\sum_{n=1}^{\lfloor q/2K \rfloor} \|\theta n\|^{-1} \ll \sum_{i=1}^{q} q/i \ll q \log q.$$

For each range $2^i \leq n < 2^{i+1}$ use the bound of Lemma 1. This gives

$$\begin{aligned}\sum_{n=\lfloor q/2K \rfloor}^{V} \min(U/n, \|\theta n\|^{-1}) &\ll K \sum_{i:2^i \in [q/4K, 2V]} \left(\frac{U}{q} + \frac{U}{2^i} + 2^i \log q + q \log q \right) \\ &\ll K^2 \left(\frac{U \log V}{q} + V \log q + q \log q \log V \right).\end{aligned}$$

The lemma follows. □

2. A minor arcs estimate. The estimate we refer to, which was used in [2], is the following.

Lemma 3 *Suppose that a, q are positive integers with $(a, q) = 1$, and let θ be a real number such that $|\theta - a/q| \leq 1/q^2$. Suppose that $q \geq 100m^2$. Then*

$$\lambda^{\wedge}_{b,m,N}(\theta) \ll m^6 (\log N)^4 \left(q^{-1/2} + N^{-1/5} + N^{-1/2} q^{1/2} \right). \tag{2}$$

Thus if $\theta \in \mathfrak{m}$ then $\lambda^{\wedge}_{b,m,N}(\theta) = O((\log N)^{-A})$.

Let Λ be von Mangoldt's function: $\Lambda(n) = \log p$ if $n = p^k$ is a prime power, and 0 otherwise. We will begin (and almost end) by obtaining the following estimate.

Lemma 4 *Suppose that $(a, q) = 1$, that $q \geq 100K$ and that $|\theta - a/q| \leq Kq^{-2}$. Then*

$$\sum_{\substack{x \leq N \\ x \equiv b (\operatorname{mod} m)}} \Lambda(x) e(\theta x) \ll K^2 (\log N)^4 \left(\frac{N}{q^{1/2}} + N^{4/5} m^{3/5} + N^{1/2} q^{1/2} \right).$$

Proof. Let us quote an identity from [1]. This identity, very closely related to an identity of Vaughan [4], will be instantly memorable to every reader who sets eyes on it. Let $1 \leq U \leq$

$\sqrt{N}$ be a parameter to be chosen later. Then

$$\begin{aligned}
\sum_{\substack{x \leq N \\ x \equiv b (\mathrm{mod}\, m)}} \Lambda(x) e(\theta x) &= \sum_{\substack{x \leq U \\ x \equiv b (\mathrm{mod}\, m)}} \Lambda(x) e(\theta x) \\
&+ \sum_{\substack{xy \leq N \\ x \leq U \\ xy \equiv b (\mathrm{mod}\, m)}} \mu(x)(\log y) e(\theta xy) \\
&+ \sum_{\substack{xy \leq N \\ x \leq U^2 \\ xy \equiv b (\mathrm{mod}\, m)}} f(x) e(\theta xy) \\
&+ \sum_{\substack{xy \leq N \\ x,y > U \\ xy \equiv b (\mathrm{mod}\, m)}} \mu(x) g(y) e(\theta xy), \qquad (3)
\end{aligned}$$

where

$$f(x) = \sum_{\substack{x = uv \\ uv \leq U}} \mu(u) \Lambda(v) \qquad (4)$$

and

$$g(y) = \sum_{\substack{y = uv \\ v > U}} \Lambda(v). \qquad (5)$$

We will refer to the four terms in (3) in imaginative fashion as the *first*, *second*, *third* and *fourth terms*. Observe that both $\|f\|_\infty$ and $\|g\|_\infty$ are at most $\log N$. We shall derive bounds for general bilinear forms of the fype appearing in the first, second, third and fourth terms using only this kind of L^∞ information on the coefficients.

Lemma 5 *Let $f : \mathbb{Z} \to \mathbb{C}$ be any function. Suppose that $q \geq 100K$, $(a, q) = 1$ and that $|\theta - a/q| \leq Kq^{-2}$. Suppose also that $N \geq 4mL$. Then*

$$\sum_{\substack{xy \leq N \\ L \leq x < 2L \\ y \geq T \\ xy \equiv b (\mathrm{mod}\, m)}} f(x) e(\theta xy) \ll K^2 (\log N)^2 \left(\frac{N}{q} + Lm + q \right) \|f\|_\infty.$$

Proof. Assume without loss of generality that $\|f\|_\infty = 1$. By the triangle inequality, the sum S in question is at most

$$\sum_{\substack{L \leq x < 2L \\ (x,m) = 1}} \left| \sum_{\substack{T \leq y \leq N/x \\ y \equiv bx^{-1} (\mathrm{mod}\, m)}} e(\theta xy) \right|.$$

Now the inner sum here is a geometric progression with common ratio $e(\theta xm)$ and length $\leq 2N/mx$ (this fact relies on the assumption about m, L, N in the statement of the lemma). Thus we have, using Lemma 2,

$$\begin{aligned} S &\ll \sum_{\substack{L \leq x < 2L \\ (x,m)=1}} \min\left(\frac{N}{mx}, \|\theta xm\|^{-1}\right) \\ &\leq \sum_{x' \leq 2Lm} \min\left(\frac{N}{x'}, \|\theta x'\|^{-1}\right) \\ &\ll K^2 (\log N)^2 \left(\frac{N}{q} + Lm + q\right). \end{aligned}$$

This concludes the proof of the Lemma. □

Lemma 6 *Let $f, g : \mathbb{Z} \to \mathbb{C}$ be any two functions. Suppose that $q \geq 100K$, $(a, q) = 1$, that $|\theta - a/q| \leq Kq^{-2}$, that $m \geq 4L$ and that $N \geq 4Lm$. Then*

$$\sum_{\substack{xy \leq N \\ L \leq x < 2L \\ xy \equiv b \, (\mathrm{mod}\, m)}} f(x)g(y)e(\theta xy)$$

$$\ll K^2 \|f\|_\infty \|g\|_\infty \left(\frac{N}{q^{1/2}} + \frac{N^{1/2}L^{1/2}}{m^{1/2}} + \frac{m^{1/2} N \log q}{L^{1/2}} + N^{1/2} q^{1/2} \log q\right).$$

Proof. Again, suppose without loss of generality that $\|f\|_\infty, \|g\|_\infty \leq 1$. Write the sum as

$$S = \sum_{\substack{b_1, b_2 \, (\mathrm{mod}\, m) \\ b_1 b_2 \equiv b \, (\mathrm{mod}\, m)}} S_{b_1, b_2}, \tag{6}$$

where

$$S_{b_1, b_2} = \sum_{\substack{xy \leq N \\ L \leq x < 2L \\ x \equiv b_1 \, (\mathrm{mod}\, m) \\ y \equiv b_2 \, (\mathrm{mod}\, m)}} f(x)g(y)e(\theta xy).$$

Using Cauchy-Schwarz in the variable x, one gets

$$
\begin{aligned}
|S_{b_1,b_2}|^2 &\ll \frac{L}{m} \sum_{\substack{L \le x < 2L \\ x \equiv b_1 (\bmod m)}} \left| \sum_{\substack{y \le N/x \\ y \equiv b_2 (\bmod m)}} g(y) e(\theta x y) \right|^2 \\
&\ll \frac{L}{m} \sum_{\substack{L \le x < 2L \\ x \equiv b_1 (\bmod m)}} \sum_{\substack{y \le N/x \\ y \equiv b_2 (\bmod m)}} \sum_{\substack{y' \le N/x \\ y' \equiv b_2 (\bmod m)}} g(y)\overline{g(y')} e(\theta x (y - y')) \\
&\le \frac{L}{m} \sum_{\substack{y \le N/L \\ y \equiv b_2 (\bmod m)}} \sum_{\substack{y' \le N/L \\ y' \equiv b_2 (\bmod m)}} \left| \sum_{\substack{L \le x < 2L \\ x \equiv b_1 (\bmod m) \\ x \le \min(N/y, N/y')}} e(\theta x (y - y')) \right| .
\end{aligned}
$$

Now the inner sum here is a geometric progression with common difference $e(\theta m(y-y'))$ and length at most $2L/m$. Thus

$$
\begin{aligned}
|S_{b_1,b_2}|^2 &\ll \frac{L}{m} \sum_{\substack{y \le N/L \\ y \equiv b_2 (\bmod m)}} \sum_{\substack{y' \le N/L \\ y' \equiv b_2 (\bmod m)}} \min\left(\frac{L}{m}, \|\theta m (y - y')\|^{-1} \right) \\
&\ll \frac{N}{m^2} \sum_{0 \le j \le N/L} \min\left(\frac{L}{m}, \|\theta m j\|^{-1} \right) \\
&\ll \frac{N}{m^2} \sum_{i \le mN/L} \min\left(\frac{L}{m}, \|\theta i\|^{-1} \right) \\
&\le \frac{NK}{m^2} \left(\frac{N}{q} + \frac{L}{m} + \left(\frac{mN}{L} + q \right) \log q \right)
\end{aligned}
$$

The lemma follows quickly from this and (6). □

Now we use Lemmas 5 and 6 to estimate the terms in (3). The first term does not need these lemmas.

$$
\text{first term} \ll U. \tag{7}
$$

To estimate the second term we use "partial summation", getting

$$\begin{aligned}\text{second term} &= \sum_{\substack{xy\le N\\ x\le U\\ xy\equiv b (\bmod m)}} \mu(x)e(\theta xy)\int_1^y \frac{dt}{t}\\ &\ll \int_1^N \frac{dt}{t}\Bigg|\sum_{\substack{xy\le N\\ x\le U\\ y\ge t\\ xy\equiv b (\bmod m)}} \mu(y)e(\theta xy)\Bigg|.\end{aligned}$$

To estimate the sum, use Lemma 5 and a dyadic decomposition over the range $x \le U$, confirming that it is at $\ll K^2(\log N)^3\left(\frac{N}{q}+Um+q\right)$. Integrating over t adds an extra logarithm, and so

$$\text{second term} \ll K^2(\log N)^4\left(\frac{N}{q}+Um+q\right).$$

The third term may be estimated by dividing the range $x \le U^2$ into dyadic ranges $2^i \le x < 2^{i+1}$ for $i = 0,\dots,\log U$ and using Lemma 5. One gets

$$\text{third term} \ll K^2(\log N)^4\left(\frac{N}{q}+U^2m+q\right). \tag{8}$$

The fourth term may be estimated by dividing the range $U \le x \le N/U$ into dyadic ranges and using Lemma 6. One gets

$$\text{fourth term} \ll K^2(\log N)^4\left(\frac{N}{q^{1/2}}+\frac{m^{1/2}N}{U^{1/2}}+N^{1/2}q\right). \tag{9}$$

Putting all this together, and absorbing terms which are obviously smaller than other terms into those terms, gives

$$\sum_{\substack{x\le N\\ x\equiv b (\bmod m)}} \Lambda(x)e(\theta x) \ll K^2(\log N)^4\left(\frac{N}{q^{1/2}}+U^2m+\frac{m^{1/2}N}{U^{1/2}}+N^{1/2}q^{1/2}\right).$$

Set $U = N^{2/5}m^{-1/5}$, and we get Lemma 4. □

It remains to derive Lemma 3 from Lemma 4. Suppose then that $|\theta - a/q| \le q^{-2}$. Observe that

$$\lambda^{\wedge}_{b,m,N}(\theta) = \frac{\phi(m)}{mN}\sum_{\substack{n\le N\\ mn+b \text{ is prime}}} \log(mn+b)e(n\theta).$$

It is easy to see from this that

$$\lambda^{\wedge}_{b,m,N}(\theta) \;=\; \frac{\phi(m)}{mN} e(-\theta b/m) \sum_{\substack{x \leq Nm \\ x \equiv b (\mathrm{mod}\, m)}} \Lambda(x) e(\theta x/m) \;+\; O(N^{-1/2}).$$

Now we can apply Lemma 4 with $\theta' = \theta/m$, noting that this satisfies the estimate $|\theta' - a/qm| \leq q^{-2}$, and hence certainly *some* estimate of the form $|\theta' - a'/q'| \leq Kq'^{-2}$ where $(a', q') = 1$ and $K \leq m^2$. The requirement, in Lemma 3, that $q \geq 100m^2$ means that certainly $q' \geq q \geq 100K$, a condition required in Lemma 4. It is easy to check that one gets an estimate of the required form. □

References

[1] Balog, A. and Perelli, A. *Exponential sums over primes in an arithmetic progression*, Proc. Amer. Math. Soc. **93** (1985), no. 4, 578–582.

[2] Green, B.J. *Roth's theorem in the primes*, preprint.

[3] Gowers, W.T. *Vinogradov's three-primes theorem*, notes available at `http://www.dpmms.cam.ac.uk/~wtg10/`

[4] Vaughan, R. C. *Sommes trigonométriques sur les nombres premiers*, C. R. Acad. Sci. Paris Sér. A-B **285** (1977), no. 16, A981–A983.