Smoothed GPY Sieve

By Yoichi Motohashi

Abstract. We introduce a smoothing device into the now famous sieve procedure [2] of D.A. Goldston, J. Pintz, and C.Y. Yıldırım combining the arguments developed in [1] and [5]. This is to be regraded as an experimental account of the initial part of a project of ours; accordingly, details of estimation procedures are mostly suppressed.

1. Let N be a parameter increasing monotonically to infinity. There are four other basic parameters H, R, k, ℓ in our discussion. We impose the following conditions to them:

$$H \ll \log N \ll \log R \le \log N, \tag{1.1}$$

and

$$\text{integers } k, \ell > 0 \text{ are arbitrary but bounded.} \tag{1.2}$$

All implicit constants in the sequel are possibly dependent on k, ℓ at most; and besides, the symbol c stands for a positive constant with the same dependency, whose value may differ at each occurrence. It suffices to have (1.2), since our eventual aim is to look into the possibility to detect the bounded differences between primes with a certain sharpening of the GPY sieve. We surmise that such a sharpening may be obtained by introducing a smoothing device. The present article is, however, only to indicate that the GPY sieve admits indeed a smoothing; it is yet to be seen if this particular smoothing contributes to our eventual aim.

Let

$$\mathcal{H} = \{h_1, h_2, \ldots, h_k\} \subseteq [-H, H] \cap \mathbb{Z}, \tag{1.3}$$

with $h_i \ne h_j$ for $i \ne j$. Let us put, for a prime p,

$$\Omega(p) = \{\text{different residue classes among } -h (\operatorname{mod} p),\ h \in \mathcal{H}\} \tag{1.4}$$

and write $n \in \Omega(p)$ instead of $n\,(\operatorname{mod} p) \in \Omega(p)$. We call $\mathcal{H}$ admissible if

$$|\Omega(p)| < p \quad \text{for all } p, \tag{1.5}$$

and assume this unless otherwise stated. We extend Ω multiplicatively, so that $n \in \Omega(d)$ with square-free d if and only if $n \in \Omega(p)$ for all $p|d$, which is equivalent to

$$d|P(n;\mathcal{H}), \quad P(n;\mathcal{H}) = (n+h_1)(n+h_2)\cdots(n+h_k). \tag{1.6}$$

We put, with μ the Möbius function,

$$\lambda_R(d;\ell) = \begin{cases} 0 & \text{if } d > R, \\ \dfrac{\mu(d)}{(k+\ell)!}\left(\log \dfrac{R}{d}\right)^{k+\ell} & \text{if } d \le R, \end{cases} \tag{1.7}$$

and

$$\Lambda_R(n;\mathcal{H},\ell) = \sum_{n\in\Omega(d)} \lambda_R(d;\ell). \tag{1.8}$$

Also, let

$$E^*(y;a,q) = \vartheta^*(y;a,q) - \frac{y}{\varphi(q)}, \quad \vartheta^*(y;a,q) = \sum_{\substack{y<n\le 2y\\ n\equiv a \bmod q}} \varpi(n), \tag{1.9}$$

where φ is the Euler totient function; and $\varpi(n) = \log n$ if n is a prime, and $= 0$ otherwise. In all accounts [1]–[3] of the GPY sieve, it is assumed that

$$\sum_{q\le x^\theta} \max_{(a,q)=1} \max_{y\le x} |E^*(y;a,q)| \ll \frac{x}{(\log x)^A}, \tag{1.10}$$

with a certain $\theta \in (0,1)$ and an arbitrary fixed $A > 0$.

The following asymptotic formulas are the fundamental implements in the GPY sieve:

Lemma 1. *Provided* (1.1), (1.2), *and* $R \le N^{1/2}/(\log N)^C$ *hold with a sufficiently large* $C > 0$ *depending only on* k *and* ℓ, *we have*

$$\begin{aligned}\sum_{N<n\le 2N} &\Lambda_R(n;\mathcal{H},\ell)^2 \\ &= \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell)!}\binom{2\ell}{\ell} N(\log R)^{k+2\ell} + O(N(\log N)^{k+2\ell-1}(\log\log N)^c),\end{aligned} \tag{1.11}$$

where

$$\mathfrak{S}(\mathcal{H}) = \prod_p \left(1 - \frac{|\Omega(p)|}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}. \tag{1.12}$$

Lemma 2. *Provided* (1.1), (1.2), *and* (1.10), *we have, for* $R \le N^{\theta/2}/(\log N)^C$ *with a sufficiently large* $C > 0$ *depending only on* k *and* ℓ,

$$\begin{aligned}\sum_{N<n\le 2N} &\varpi(n+h)\Lambda_R(n;\mathcal{H},\ell)^2 \\ &= \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell+1)!}\binom{2(\ell+1)}{\ell+1} N(\log R)^{k+2\ell+1} + O(N(\log N)^{k+2\ell}(\log\log N)^c),\end{aligned} \tag{1.13}$$

whenever $h \in \mathcal{H}$.

Note that the case $h \notin \mathcal{H}$, which is not included here, seems to be ignorable for our purpose. It is in fact known that the combination of (1.11), (1.13), and (1.10) with a $\theta > \frac{1}{2}$ gives rise to bounded differences between primes. The aim of the present work is to prove a smoothed version of (1.11) and (1.13) to look into the possibility of replacing (1.10) with a $\theta > \frac{1}{2}$ by any less stringent hypothesis. With this in mind, we shall hereafter assume that

$$H = H(k,\ell) \text{ is bounded.} \tag{1.14}$$

2. We begin with a smoothing of (1.11). To this end, we shall follow [5]. Thus, let us put

$$R_0 = \exp\left(\frac{\log R}{(\log\log R)^{1/5}}\right), \quad R_1 = \exp\left(\frac{\log R}{(\log\log R)^{9/10}}\right), \quad \tau = (\log\log R)^{1/10}. \tag{2.1}$$

We divide the interval $[R_0, \infty)$ into intervals $[R_0R_1^{j-1}, R_0R_1^j)$ $(j = 1, 2, \ldots)$, denoting them by P, with or without suffix. Let D be a generic element of the commutative semi-group generated by all P's. When $D = P_1P_2\cdots P_r$, the notation $d \in D$ indicates that d has the prime decomposition $d = p_1p_2\cdots p_r$ with $p_j \in P_j$ $(1 \le j \le r)$. Note that we use the convention that $1 \in D$ if and only if D is the empty product. Further, we put $|P| = R_0R_1^j$ if $P = [R_0R_1^{j-1}, R_0R_1^j)$; and $|D| = |P_1|\cdots|P_r|$ if $D = P_1P_2\cdots P_r$. Naturally, $|D| = 1$ if D is empty. We put

$$\Delta(D) = \prod_{P|D}\left(\sum_{p\in P}\frac{|\Omega(p)|}{p}\right), \tag{2.2}$$

and

$$\Phi(D) = \prod_{P|D}\left(\sum_{p\in P}\frac{|\Omega(p)|}{p}\left(1-\frac{|\Omega(p)|}{p}\right)\right)\left(\sum_{p\in P}\frac{|\Omega(p)|}{p}\right)^{-2}. \tag{2.3}$$

Also, modifying [1, (1.21)] and [5, (6)], we put

$$\tilde\lambda_R(D;\ell) = \frac{\mathfrak{S}(\mathcal{H})}{\ell! W(R_0)}\frac{\mu(D)}{\Delta(D)}\sum_{\substack{|K|<R\\ D|K}}\frac{\mu(K)^2}{\Phi(K)}\left(\log\frac{R}{|K|}\right)^\ell, \tag{2.4}$$

where

$$W(z) = \prod_{d<z}\left(1-\frac{|\Omega(p)|}{p}\right), \tag{2.5}$$

and the empty sum is to vanish; that is, $\tilde\lambda_R(D;\ell) = 0$ for $|D| \ge R$. Note that in (2.2) and (2.3) we have $p \ge R_0$, and thus $|\Omega(p)| = k$ always. We shall, however, keep the notation $|\Omega(p)|$, because of a future purpose.

As to the interval $[1, R_0)$, which is excluded in the above, we appeal to the Fundamental Lemma in the sieve method (see e.g., [6, Sections 3.2–3.5]). Thus, there exits a set of sieve weights $\varrho(d)$ such that $|\varrho(d)| \le 1$ for any $d \ge 1$, and $\varrho(d) = 0$ either if $d \ge R_0^\tau$ with τ as above or if d has a prime factor greater than or equal to R_0, and that

$$\gamma(n;\mathcal{H}) = \sum_{n\in\Omega(d)}\varrho(d) \ge 0 \text{ for any } n \ge 1 \tag{2.6}$$

as well as

$$\sum_d \frac{\varrho(d)}{d}|\Omega(d)| = W(R_0)\left\{1 + O\left(e^{-\tau}\right)\right\}, \tag{2.7}$$

with the implied constant being absolute.

With this, our smoothed counterpart of (1.8) is defined to be

$$\tilde{\Lambda}_R(n;\mathcal{H},\ell)=\sum_{D}\tilde{\lambda}_R(D;\ell)\sum_{\substack{d\in D\\ n\in\Omega(d)}}1. \tag{2.8}$$

Then our first task is to evaluate asymptotically the sum

$$\sum_{N<n\le 2N}\gamma(n;\mathcal{H})\tilde{\Lambda}_R(n;\mathcal{H},\ell)^2. \tag{2.9}$$

Expanding out the square and incorporating (2.6)–(2.7), we have

$$\begin{aligned} & N\mathcal{T}\sum_{d}\frac{\varrho(d)}{d}|\Omega(d)| \\ & \quad +O\left(\sum_{D_1,D_2}|\tilde{\lambda}_R(D_1;\ell)||\tilde{\lambda}_R(D_2;\ell)|\sum_{d_1\in D_1,d_2\in D_2}\sum_{d\le R_0^\tau}|\Omega(d[d_1,d_2])|\right) \\ & =NW(R_0)\mathcal{T}\left(1+O(e^{-\tau})\right)+O\left(R_0^\tau R^2(\log N)^c\right)\end{aligned} \tag{2.10}$$

with

$$\mathcal{T}=\sum_{D_1,D_2}\tilde{\lambda}_R(D_1;\ell)\tilde{\lambda}_R(D_2;\ell)\sum_{d_1\in D_1,\,d_2\in D_2}\frac{|\Omega([d_1,d_2])}{[d_1,d_2]}. \tag{2.11}$$

The reasoning in [5, pp. 1056 – 1057] can be applied to $\mathcal{T}$, and we get

$$\begin{aligned}\mathcal{T}&=\sum_{|D|<R}\Phi(D)\left(\sum_{D|K}\Delta(K)\tilde{\lambda}_R(K;\ell)\right)^2 \\ &=\left(\frac{\mathfrak{S}(\mathcal{H})}{\ell!W(R_0)}\right)^2\sum_{|D|<R}\frac{\mu(D)^2}{\Phi(D)}\left(\log\frac{R}{|D|}\right)^{2\ell};\end{aligned} \tag{2.12}$$

the second line is due to the definition (2.4). By [5, pp. 1057–1061], we have, for $y\le R$,

$$\sum_{|D|<y}\frac{\mu(D)^2}{\Phi(D)}=\frac{W(R_0)}{k!\mathfrak{S}(\mathcal{H})}(\log y)^k\left(1+O((\log\log R)^{-1/5})\right). \tag{2.13}$$

Summing by parts in the second line of (2.12), we get readily

Lemma 3. *Under* (1.14) *and the same assumption as in Lemma 1, we have*

$$\begin{aligned}&\sum_{N<n\le 2N}\gamma(n;\mathcal{H})\tilde{\Lambda}_R(n;\mathcal{H},\ell)^2 \\ &=\frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell)!}\binom{2\ell}{\ell}N(\log R)^{k+2\ell}\left(1+O((\log\log N)^{-1/5})\right).\end{aligned} \tag{2.14}$$

3. Next, we shall consider a twist of (2.14) with primes:

$$\sum_{N<n\le 2N}\varpi(n+h)\gamma(n;\mathcal{H})\tilde{\Lambda}_R(n;\mathcal{H},\ell)^2 \tag{3.1}$$
$$=\sum_{N<n\le 2N}\varpi(n+h)\gamma(n;\mathcal{H}\backslash\{h\})\tilde{\Lambda}_R(n;\mathcal{H}\backslash\{h\},\ell)^2,$$

provided it holds that $\mathcal{H}\ni h$, $R<N$. We may suppose that

$$\mathcal{H}\ni 0,\quad h=0; \tag{3.2}$$

otherwise one may sift the sum appropriately with a negligible error. Thus we shall consider instead

$$\sum_{N<n\le 2N}\varpi(n)\gamma(n;\mathcal{H}^-)\tilde{\Lambda}_R(n;\mathcal{H}^-,\ell)^2, \tag{3.3}$$

with $\mathcal{H}^-=\mathcal{H}\backslash\{0\}$.

Expanding out the square, we have

$$N\mathcal{T}^*\sum_d \frac{\varrho(d)}{\varphi(d)}|\Omega^*(d)|+\mathcal{E}, \tag{3.4}$$

where

$$\mathcal{T}^*=\sum_{D_1,D_2}\tilde{\lambda}_R(D_1;\ell)\tilde{\lambda}_R(D_2;\ell)\sum_{d_1\in D_1,\,d_2\in D_2}\frac{|\Omega^*([d_1,d_2])|}{\varphi([d_1,d_2])}, \tag{3.5}$$

$$\mathcal{E}=\sum_{D_1,D_2}\tilde{\lambda}_R(D_1;\ell)\tilde{\lambda}_R(D_2;\ell)\sum_{d_1\in D_1,\,d_2\in D_2}\sum_d\varrho(d)\sum_{a\in\Omega^*(d[d_1,d_2])}E^*(N;a,d[d_1,d_2]). \tag{3.6}$$

Here Ω^* is defined by the relation $\Omega^*(p)=\Omega(p)\backslash\{0\}$ for all p; thus $|\Omega^*(p)|=|\Omega(p)|-1$.

First we have, corresponding to (2.7),

$$\sum_d\frac{\varrho(d)}{\varphi(d)}|\Omega^*(d)|=\frac{W(R_0)}{V(R_0)}\left(1+O(e^{-\tau})\right),\quad V(z)=\prod_{p<z}\left(1-\frac{1}{p}\right), \tag{3.7}$$

again by [6, Sections 3.2–3.5]. The diagonalization procedure in [5, pp. 1056–1057] can be employed again, and we find that

$$\mathcal{T}^*=\sum_D\Phi^*(D)\left(\sum_{D|K}\Delta^*(K)\tilde{\lambda}_R(K;\ell)\right)^2, \tag{3.8}$$

where

$$\Delta^*(D)=\prod_{P|D}\left(\sum_{p\in P}\frac{|\Omega^*(p)|}{p-1}\right), \tag{3.9}$$

and

$$\Phi^*(D) = \prod_{P|D} \left(\sum_{p\in P} \frac{|\Omega^*(p)|}{p-1} \left(1 - \frac{|\Omega^*(p)|}{p-1} \right) \right) \left(\sum_{p\in P} \frac{|\Omega^*(p)|}{p-1} \right)^{-2}. \tag{3.10}$$

Note that here $|\Omega^*(p)|$ does vanish, provided $k \geq 2$, which we may of course assume.

We are about to compute $\mathcal{T}^*$ asymptotically. To this purpose, we remark that

$$\Phi^*(D) = \Delta^*(D)^{-1} \left(1 + O\left(R_0^{-1/2} \right) \right), \tag{3.11}$$

where the implied constant is absolute; (1.1) and (2.1) are relevant. Hence, we may consider instead of (3.8) the expression

$$\sum_{D} \frac{1}{\Delta^*(D)} \left(\sum_{D|K} \Delta^*(K) \tilde{\lambda}_R(K;\ell) \right)^2. \tag{3.12}$$

Further, we may approximate (2.4), in the same sense, by

$$\frac{\mathfrak{S}(\mathcal{H})}{\ell! W(R_0)} \frac{\mu(D)}{\Delta(D)} \sum_{\substack{|K|<R \\ D|K}} \mu(K)^2 \Delta(K) \left(\log \frac{R}{|K|} \right)^{\ell}. \tag{3.13}$$

Then the sum over K in (3.12) is replaced by

$$\frac{\mathfrak{S}(\mathcal{H})}{\ell! W(R_0)} \mu(D)\Delta^*(D) \sum_{\substack{|L|<R/|D| \\ (D,L)=1}} \mu(L)^2 \prod_{P|L} \left(\sum_{p\in P} \frac{1}{p-1} \right) \left(\log \frac{R/|D|}{|L|} \right)^{\ell}, \tag{3.14}$$

with an admissible error. This is to be compared with [1, (4.15)].

A simple modification of [5, pp. 1057–1061], i.e., a specialization of Φ there, yields

$$\begin{aligned} &\sum_{\substack{|L|<y \\ (D,L)=1}} \mu(L)^2 \prod_{P|L} \left(\sum_{p\in P} \frac{1}{p-1} \right) \\ &\qquad = \frac{V(R_0) \log y}{\prod_{P|D} \left(1 + \sum_{p\in P} \frac{1}{p-1} \right)} \left(1 + O((\log\log R)^{-1/5}) \right), \end{aligned} \tag{3.15}$$

where the implied constant is absolute. This with summation by parts implies that (3.14) is equal to

$$\frac{\mathfrak{S}(\mathcal{H}) V(R_0)}{(\ell+1)! W(R_0)} \frac{\mu(D)\Delta^*(D)}{\prod_{P|D} \left(1 + \sum_{p\in P} \frac{1}{p-1} \right)} \left(\log \frac{R}{|D|} \right)^{\ell+1} \left(1 + O((\log\log R)^{-1/5}) \right), \tag{3.16}$$

and (3.12) to

$$\left(\frac{\mathfrak{S}(\mathcal{H})V(R_0)}{(\ell+1)!W(R_0)}\right)^2 \times \sum_{|D|<R} \frac{\mu(D)^2\Delta^*(D)}{\prod_{P|D}\left(1+\sum_{p\in P}\frac{1}{p-1}\right)^2}\left(\log\frac{R}{|D|}\right)^{2(\ell+1)}\left(1+O((\log\log R)^{-1/5}))\right). \tag{3.17}$$

To evaluate the last sum we employ [5, pp. 1057–1061] again with an appropriate changes, and are led to

$$\mathcal{T}^* = \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell+1)!}\binom{2(\ell+1)}{\ell+1}\frac{V(R_0)}{W(R_0)}(\log R)^{k+2\ell+1}\left(1+O((\log\log R)^{-1/5}))\right) \tag{3.18}$$

Collecting the above assertions, we obtain

Lemma 4. *Under* (1.14) *and the same assumption as in Lemma 1, we have, for any* $h \in \mathcal{H}$*,*

$$\sum_{N<n\le 2N} \varpi(n+h)\gamma(n;\mathcal{H})\tilde\Lambda_R(n;\mathcal{H},\ell)^2 = \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell+1)!}\binom{2(\ell+1)}{\ell+1}N(\log R)^{k+2\ell+1}\left(1+O((\log\log N)^{-1/5})\right)+E, \tag{3.19}$$

with

$$E \ll (\log N)^c \sup_{\alpha,\beta}\left|\sum_{\substack{d_1,d_2<R\\(d_1,d_2)=1}}\alpha_{d_1}\beta_{d_2}\sum_{a\in\Omega^*(d_1d_2)}E^*(N;a,d_1d_2)\right|. \tag{3.20}$$

Here α, β *are arbitrary complex vectors such that* $|\alpha_{d_1}|\le 1$, $|\beta_{d_2}|\le 1$.

The assertion (3.20) is proved as in [5, pp. 1063–1064]. The bound (1.10) implies obviously that $E \ll N/(\log N)^{A/2}$ for any large A if $R \le N^{\theta/2}$. We surmise that the same estimation of E should hold with a larger R by virtue of the bilinear structure in (3.20).

Another possible way to introduce a smoothing into the GPY sieve problem is to apply the device [4, Section 2.3] upon the Rosser sieve of dimension $k+\ell$. It could lead us to a far more flexible error term than (3.20). However, our experiment indicates that the main term thus obtained has a dependency on the basic parameter ℓ that is not so effective as what the Selberg sieve has yielded in the above.

References

[1] G.A. Goldston, S.W. Graham, J. Pintz, and C.Y. Yıldırım. Small gaps between primes or products of two primes. Preprint. arXiv: math.NT/506067.

[2] D.A. Goldston, J. Pintz, and C.Y. Yıldırım. Small gaps between primes II (Preliminary). Preprint, February 8, 2005.

[3] G.A. Goldston, Y. Motohashi, J. Pintz, and C.Y. Yıldırım. Small gaps between primes exist. Preprint. arXiv: math.NT/0505300.

[4] Y. Motohashi. Sieve Methods and Prime Number Theory. Tata IFR LN, vol. **72**, Springer Verlag, Berlin etc 1983.

[5] Y. Motohashi. On the remainder term in the Selberg sieve. Number Theory in Progress - A. Schinzel Festschrift, Walter de Gruyter, Berlin · New York 1999, pp. 1053–1064.

[6] Y. Motohashi. An overview of the sieve method and its history. Preprint (to appear in Sugaku Expositions, AMS). arXiv: math.NT/0505521.