# 素数的小间隔 (I)

龚　克

同济大学数学系　*E-mail: kkgong@gmail.com*

在这篇报告中，我们将介绍 Goldston 和 Yıldırım 关于素数间小间隔的工作，该工作的主要思想是用短的除数和逼近素数组及小区间上素数的矩.

## 1　引言

记 $p_n$ 为第 $n$ 个素数. 由素数定理

$$\pi(x) := \sum_{p\leqslant x} 1 \sim \frac{x}{\log x} \quad (\, x \to \infty \,)$$

知[1]

$$p_n \sim n \log n \quad (\, n \to \infty \,),$$

于是

$$\frac{1}{N}\sum_{n=1}^{N} \frac{p_{n+1}-p_n}{\log p_n} \to 1 \quad (\, N \to \infty \,).$$

上式表明相邻素数的间隔 $p_{n+1}-p_n$ 平均约为 $\log p_n$. 定义

$$\Delta = \liminf_{n\to\infty} \frac{p_{n+1}-p_n}{\log p_n}, \tag{1.1}$$

显见 $\Delta \leqslant 1$ 且 $\Delta$ 就是满足如下条件的最小的数：存在无穷多个相邻素数的间隔小于平均间隔的 $\Delta+\epsilon$ 倍. 经验上 $\Delta$ 明显应当等于 0，但该论断迄今未被证明. 为了给出 $\Delta$ 的非平凡估计，人们创造出了三种不同的方法. 现列举如次.

**第一种：Hardy-Littlewood 和 Bombieri-Davenport 的方法.**

[1]由素数定理，令 $y=\pi(x)$，则当 $x\to\infty$ 时有 $\dfrac{y\log x}{x}\to 1$，即 $\log y+\log\log x-\log x\to 0$. 于是有 $\dfrac{\log y}{\log x}\to 1$ 及 $\dfrac{y\log y}{x}\to 1$. 取 $x=p_n$ ( $y=\pi(p_n)=n$ )，即得 $n\to\infty$ 时 $p_n\sim n\log n$.

Hardy 和 Littlewood 在广义 Riemann 猜想下用圆法得到 $\Delta \leqslant 2/3$ [1]．1940 年，Rankin [13] 得到 $\Delta \leqslant \frac{1+4\Theta}{5}$，其中 $\Theta$ 是 Dirichlet $L$-函数所有零点实部的上确界．特别地，假定广义 Riemann 猜想 ($\Theta = 1/2$) 成立，此结果即为 $\Delta \leqslant 3/5$．

1965 年，Bombieri 和 Davenport [1] 改进了上述结果，并去掉了广义 Riemann 猜想的假定．他们得到了 $\Delta \leqslant 1/2$，证明使用圆法和大筛法．这种方法相当于 Goldston **用截断的除数和替代** von Mangoldt **函数** $\Lambda(n)$ **的二阶矩方法**．

**第二种：**Erdös **方法．** 由素数定理知 $\Delta \leqslant 1$ (证明亦见 Prachar [12]，定理 4.2)．Erdös [3] 在 1940 年首先证得无条件的结果 $\Delta < 1$．他对相差偶数 $k$ 的两个素数使用上界筛法

$$\sum_{n \leqslant N} \Lambda(n)\Lambda(n+k) \leqslant (\mathcal{B}+\epsilon)\mathfrak{S}(k)N\,, \tag{1.2}$$

其中 $\mathfrak{S}(k)$ 为奇异级数，$\mathcal{B}$ 为一常数．这种方法给出

$$\Delta \leqslant 1 - \frac{1}{2\mathcal{B}}\,. \tag{1.3}$$

**第三种：**Maier **方法．** 1988 年 Maier [10] 发现在某些区间上 ( 这些区间分布的相当稀疏 ) 能找到 $e^{\gamma}$ ( $\gamma$ 代表 Euler 常数 ) 倍于预期的素数个数那么多的素数，这就导致在这些区间上素数的平均间隔丢掉一个 $e^{-\gamma}$ 因子．因此由素数定理知 $\Delta \leqslant e^{-\gamma} \cdot 1 = 0.56145\ldots$．此方法与前两种方法结合可得到目前最好的结果 $\Delta \leqslant 0.24846\ldots$，但与前两种方法不同的是它不能说明小的素数间隔占有正比例．

本文将证明如下结果:

**定理 1** *设 $r \in \mathbb{Z}^+$，对任意固定的 $\lambda > (\sqrt{r} - \frac{1}{2})^2$ 及充分大的 $N$，我们有*

$$\sum_{\substack{p_n \leqslant N \\ p_{n+r} - p_n \leqslant \lambda \log p_n}} 1 \gg_r \pi(N)\,. \tag{1.4}$$

*特别的，对任意固定的 $\eta > 0$ 及所有充分大的 $N > N_0(\eta)$，当 $p_n \leqslant N$ 时，小于 $(\frac{1}{4}+\eta)\log N$ 的间隔 $p_{n+1} - p_n$ 占有正比例，从而*

$$\Delta \leqslant \frac{1}{4}\,. \tag{1.5}$$

## 2 逼近素数组

本文寻找素数小间隔的途径在于计算小区间上素数个数矩 ( moment ) 的逼近，在计算的过程中要用短的除数和去逼近素数组.

给定正整数 $h$，设

$$\mathcal{H}=\{h_1,h_2,\ldots,h_k\},\quad 0\leqslant h_1,h_2,\ldots,h_k\leqslant h \text{ 为互异的整数}. \tag{2.1}$$

记 $\nu_p(\mathcal{H})$ 为 $\mathcal{H}$ 中元素所占据的模 $p$ 的不同剩余类的个数. 定义奇异级数 ( singular series )

$$\mathfrak{S}(\mathcal{H})=\prod_p\left(1-\frac{1}{p}\right)^{-k}\left(1-\frac{\nu_p(\mathcal{H})}{p}\right). \tag{2.2}$$

若 $\mathfrak{S}(\mathcal{H})\neq 0$，则称 $\mathcal{H}$ 为可容许的 ( admissible ). 因此 $\mathcal{H}$ 为可容许的充要条件为对任意的 $p$ 有 $\nu_p(\mathcal{H})<p$.

记 $\Lambda(n)$ 为 von Mangoldt 函数，定义

$$\Lambda(n;\mathcal{H})=\Lambda(n+h_1)\Lambda(n+h_2)\cdots\Lambda(n+h_k). \tag{2.3}$$

Hardy-Littlewood 素数组猜想断言：若 $\mathcal{H}$ 为可容许的，则当 $N\to\infty$ 时有

$$\sum_{n\leqslant N}\Lambda(n;\mathcal{H})=N\left(\mathfrak{S}(\mathcal{H})+o(1)\right). \tag{2.4}$$

( 若 $\mathcal{H}$ 不是可容许的，则该猜想是平凡的[2]. ) 如在 Goldston 早期的工作中一样，此处先用截断的除数和[3]

$$\Lambda_R(n)=\sum_{\substack{d|n\\ d\leqslant R}}\mu(d)\log\frac{R}{d}, \tag{2.5}$$

来替代 $\Lambda(n)$，然后用

$$\Lambda_R(n;\mathcal{H})=\Lambda_R(n+h_1)\Lambda_R(n+h_2)\cdots\Lambda_R(n+h_k) \tag{2.6}$$

替代 $\Lambda(n;\mathcal{H})$. 为方便起见，在 $\mathcal{H}=\emptyset$ 时定义 $\Lambda_R(n;\mathcal{H})=1$.

[2]若 $\mathcal{H}$ 不是可容许的，即对某个素数 $p$ 有 $\nu_p(\mathcal{H})=p$，亦即对某个素数 $p$，$h_1,h_2,\ldots,h_k$ 占据了模 $p$ 的所有剩余类，则对任意的 $n$，在数 $n+h_1,n+h_2,\ldots,n+h_k$ 中必有一个为 $p$ 的倍数，于是只可能有有限多个 $k$ 生素数组 $(n+h_1,n+h_2,\ldots,n+h_k)$. 注意，这种素数组虽然很少，但确实存在！如：$p=3$，$\mathcal{H}=\{0,2,4\}$，在 $n=3$ 时 $(3,5,7)$ 即给出一个素数组.

[3]Goldston's remark [6]: The use of functions like $\Lambda_R(n)$ goes back to Selberg's work [14] on the zeros of the Riemann zeta-function $\zeta(s)$. The most beneficial feature of the truncated sums is that they can be used in place of $\Lambda(n)$ on some occasions when it is not known how to work with $\Lambda(n)$ itself.

假定集合 $\mathcal{H}_1$，$\mathcal{H}_2$ 皆由 $\leqslant h$ 的互异正整数构成，且 $|\mathcal{H}_1| = k_1$，$|\mathcal{H}_2| = k_2$．命 $k = k_1 + k_2$．我们总假定 $k \geqslant 1$．下面两个命题可使我们得到关于素数的信息．

**命题 1** 设 $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$，$r = |\mathcal{H}_1 \cap \mathcal{H}_2|$．若 $R = o\left(N^{\frac{1}{k}}(\log N)^{1-\frac{|\mathcal{H}|}{k}}\right)$ 且对任意大的常数 $A > 0$ 有 $h \leqslant \log^A N$，则当 $R, N \to \infty$ 时有

$$\sum_{n \leqslant N} \Lambda_R(n; \mathcal{H}_1)\Lambda_R(n; \mathcal{H}_2) = N\left(\mathfrak{S}(\mathcal{H}) + o_k(1)\right)(\log R)^r . \tag{2.7}$$

**命题 2** 设 $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$，$r = |\mathcal{H}_1 \cap \mathcal{H}_2|$，$1 \leqslant h_0 \leqslant h$．设 $\mathcal{H}_0 = \mathcal{H} \cup \{h_0\}$，当 $h_0 \notin \mathcal{H}$ 时 $r_0 = r$；当 $h_0 \in \mathcal{H}$ 时 $r_0 = r + 1$．若 $R \ll_k N^{\frac{1}{2k}}(\log N)^{-B(k)}$，其中 $B(k)$ 为一充分大的正常数，且 $h \leqslant R^{\frac{1}{2k}}$，则当 $R, N \to \infty$ 时有

$$\sum_{n \leqslant N} \Lambda_R(n; \mathcal{H}_1)\Lambda_R(n; \mathcal{H}_2)\Lambda(n + h_0) = N\left(\mathfrak{S}(\mathcal{H}_0) + o_k(1)\right)(\log R)^{r_0} . \tag{2.8}$$

假定 Elliott-Halberstam 猜想成立，则 (2.8) 对 $R \ll_k N^{\frac{1}{k}-\epsilon}$ 成立，其中 $\epsilon$ 为任一正数．

若在命题 1 中取 $\mathcal{H}_2 = \emptyset$，则对 $R = o(N^{\frac{1}{k}})$ 及 $h \leqslant \log^A N$，$A$ 是任意给定的正常数，当 $R, N \to \infty$ 时有

$$\sum_{n \leqslant N} \Lambda_R(n; \mathcal{H}) = N\big(\mathfrak{S}(\mathcal{H}) + o_k(1)\big) , \tag{2.9}$$

这对应于 Hardy-Littlewood 素数组猜想 (2.4)．

在应用这些命题时要求某些量为正很关键．譬如，在 $\mathcal{H}_2 = \emptyset$ 的特殊情形，命题 2 有如下看起来可用来探测素数的形式：对 $R \leqslant N^{\frac{1}{2k}}(\log N)^{-B(k)}$，

$$\sum_{n \leqslant N} \Lambda_R(n; \mathcal{H})\Lambda(n + h_0) = \begin{cases} N\left(\mathfrak{S}(\mathcal{H}) + o_k(1)\right)\log R , & \text{若 } h_0 \in \mathcal{H} , \\ N\left(\mathfrak{S}(\mathcal{H}_0) + o_k(1)\right) , & \text{若 } h_0 \notin \mathcal{H} , \end{cases} \tag{2.10}$$

事实上，由于 $\Lambda_R(n; \mathcal{H})$ 不是非负的，所以我们不可能单独从 (2.10) 得到任何关于素数的结果．

另一方面，考虑 $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$ 时命题 2 的特殊情形．取 $|\mathcal{H}| = k$，对 $R \leqslant N^{\frac{1}{4k}}(\log N)^{-B(k)}$ 有[4]

$$\sum_{n \leqslant N} \Lambda_R(n; \mathcal{H})^2\Lambda(n + h_0) = \begin{cases} N\left(\mathfrak{S}(\mathcal{H}) + o_k(1)\right)(\log R)^{k+1} , & \text{若 } h_0 \in \mathcal{H} , \\ N\left(\mathfrak{S}(\mathcal{H}_0) + o_k(1)\right)(\log R)^{k} , & \text{若 } h_0 \notin \mathcal{H} . \end{cases} \tag{2.11}$$

[4]此处出现的 $k$ 的大小是命题 2 中 $k$ 的 $1/2$，故 $R \leqslant N^{\frac{1}{4k}}(\log N)^{-B(k)}$．

此处对 $R$ 大小的限制使得我们不可能由 (2.11) 得出数组 $\mathcal{H}$ 中含有多于两个的素数．但 Granville 和 Soundararajan 利用 $\Lambda_R(n;\mathcal{H})^2$ 的非负性可从 (2.11) 简短地证得

$$\Delta \leqslant \frac{3}{4}\,. \tag{2.12}$$

下面证明 (2.12)．我们需要 Gallagher [4] 的结果：当 $h \to \infty$ 时，

$$\sum_{\substack{1\leqslant h_1,h_2,\ldots,h_k\leqslant h\\ \text{互异}}} \mathfrak{S}(\mathcal{H}) = h^k + O_k(h^{k-\frac{1}{2}+\epsilon})\,. \tag{2.13}$$

固定 $k \geqslant 1$，下述讨论对任意 $k$ 皆成立．现在假定

$$R = N^{\frac{1}{4k}}(\log N)^{-B(k)}\,, \qquad h \ll \log N\,.$$

通过作差知，当方程 (2.11) 右端求和范围变为 $N < n \leqslant 2N$ 时该式仍成立，于是我们有

$$\begin{aligned}\sum_{n=N+1}^{2N}\Big(\sum_{1\leqslant h_0\leqslant h}\Lambda(n+h_0)\Big)\Lambda_R(n;\mathcal{H})^2 \\ \sim kN\mathfrak{S}(\mathcal{H})(\log R)^{k+1} + \sum_{\substack{1\leqslant h_0\leqslant h\\ h_0\neq h_i,\ 1\leqslant i\leqslant k}} N\mathfrak{S}(\mathcal{H}_0)(\log R)^k\,.\end{aligned}$$

由命题 1 作差亦可得到

$$\sum_{n=N+1}^{2N}\Lambda_R(n;\mathcal{H})^2 \sim N\mathfrak{S}(\mathcal{H})(\log R)^k$$

于是我们可对所有互异的数组 $1 \leqslant h_1, h_2, \ldots, h_k \leqslant h$ 求和，并应用 (2.13) 得：对固定的 $\rho$，当 $h, N \to \infty$ 时有

$$\begin{aligned}\sum_{n=N+1}^{2N}\Big(\sum_{1\leqslant h_0\leqslant h}\Lambda(n+h_0)-\rho\log N\Big)\Big(\sum_{\substack{1\leqslant h_1,h_2,\ldots,h_k\leqslant h\\ \text{互异}}}\Lambda_R(n;\mathcal{H})^2\Big) \\ \sim Nh^k(\log R)^k\big(k\log R + h - \rho\log N\big) \\ \sim Nh^k(\log R)^k\big(h - (\rho-\frac{1}{4})\log N\big)\,.\end{aligned}$$

由于 ( 注意 $N+1 \leqslant n \leqslant 2N$，故 $\log R \ll \log N \ll n^\epsilon$ )

$$\Lambda_R(n) = \sum_{d|n,\, d\leqslant R}\mu(d)\log\frac{R}{d} \ll \log R\sum_{d|n}1 \ll d(n)\log R \ll n^\epsilon\,,$$

显见上述和式中对应于 $n+h_0$ 为素数幂的项的和[5] $\ll N^{1/2+\epsilon}$，可以忽略，因此我们可将对 $h_0$ 的求和限制在 $n+h_0$ 为素数的项上．若 $h>(\rho-\frac{1}{4})\log N$，则上式右端为正．这表明当 $h>(\rho-\frac{1}{4})\log N$ 时存在一个 $n$ ( $N<n\leqslant 2N$ )，使得[6]

$$\sum_{\substack{1\leqslant h_0\leqslant h\\ n+h_0\ \text{为素数}}}\log(n+h_0)>\rho\log N\,.$$

若 $\rho>1$，显见当 $N$ 充分大时左端和式中至少有两项．取 $\rho\to 1^+$，即得 (2.12)．

## 3 与 Poisson 分布相关的二次型的最优化

定理 1 的证明最后依赖于一个与 Poisson 分布相关的最优化过程．

设随机变量 $X$ 服从参数为 $\lambda$ 的 Poisson 分布，它的概率分布函数为

$$p(j)=\text{Prob}\,(X=j)=e^{-\lambda}\frac{\lambda^j}{j!},\quad j=0,1,2,\ldots\,. \tag{3.1}$$

定义关于 $p(j)$ 的内积为

$$\begin{aligned}\langle f(x),g(x)\rangle&=\sum_{j=0}^{\infty}f(j)g(j)p(j)\\&=e^{-\lambda}\sum_{j=0}^{\infty}f(j)g(j)\frac{\lambda^j}{j!}\,.\end{aligned} \tag{3.2}$$

定义 Poisson 分布的 $k$ 阶矩为[7]

$$\begin{aligned}\mu_k(\lambda)=E(x^k)=\langle x^k,1\rangle&=e^{-\lambda}\sum_{j=0}^{\infty}\frac{j^k\lambda^j}{j!}\\&=e^{-\lambda}\left(\lambda\frac{d}{d\lambda}\right)^k\sum_{j=0}^{\infty}\frac{\lambda^j}{j!}\\&=e^{-\lambda}\left(\lambda\frac{d}{d\lambda}\right)^k e^{\lambda}\,.\end{aligned} \tag{3.3}$$

[5] 第二个括号中的和式 $\ll N^{\epsilon}$，第一个括号中的和式亦 $\ll N^{\epsilon}$，而对应于 $n+h_0$ 为素数幂的 $n$ 的个数 $\leqslant N^{1/2}$．

[6] $1\leqslant h_0\leqslant h\ll\log N$，$N<n\leqslant 2N$，于是当 $N$ 充分大时，$N+1\leqslant n+h_0\ll 2N+\log N$，$\log(N+1)<\log(n+h_0)\ll\log N+\log 2$．因此需取 $\rho>1$ 才能保证左端和式至少有两项．

[7]
$$\sum_{j=0}^{\infty}\frac{j^k\lambda^j}{j!}=\sum_{j=0}^{\infty}j^{k-1}j\frac{\lambda\cdot\lambda^{j-1}}{j!}=\sum_{j=0}^{\infty}j^{k-1}\left(\lambda\frac{d}{d\lambda}\right)\frac{\lambda^j}{j!}=\cdots=\left(\lambda\frac{d}{d\lambda}\right)^k\sum_{j=0}^{\infty}\frac{\lambda^j}{j!}\,.$$

于是有

$$\mu_k(\lambda) = \sum_{\nu=1}^{k} \left\{ \begin{array}{c} k \\ \nu \end{array} \right\} \lambda^{\nu}, \tag{3.4}$$

其中 $\left\{ \begin{array}{c} k \\ \nu \end{array} \right\}$ 为第二型 Stirling 数，即将一个 $k$ 元集划分成 $\nu$ 个非空子集 ( 不计子集次序 ) 的分法个数.

**证明 (3.4) 式.**

*1)* 显见

$$\left\{ \begin{array}{c} k \\ \nu \end{array} \right\} = \nu \left\{ \begin{array}{c} k-1 \\ \nu \end{array} \right\} + \left\{ \begin{array}{c} k-1 \\ \nu-1 \end{array} \right\}. \tag{3.5}$$

欲证 (3.5) 只需注意到作为 $k$ 元集的最后一个元素的归属只有如下两种可能:

*i)* 该元素单独成为一个非空子集(单点集)，而前面的 $k-1$ 个元素要划分成 $\nu-1$ 个非空子集，共有 $\left\{ \begin{array}{c} k-1 \\ \nu-1 \end{array} \right\}$ 种分法.

*ii)* 该元素不作为一个非空单点集，而是可以随意放入 $\nu$ 个非空子集中的任一个，此时其余 $k-1$ 个元素已经划分成 $\nu$ 个非空子集，因此共有 $\nu \left\{ \begin{array}{c} k-1 \\ \nu \end{array} \right\}$ 种分法.

*2)* 为证 (3.4) 成立，我们还需要恒等式

$$\sum_{\nu=0}^{j} \nu! \left\{ \begin{array}{c} k \\ \nu \end{array} \right\} \left( \begin{array}{c} j \\ \nu \end{array} \right) = j^k. \tag{3.6}$$

其中 $\left( \begin{array}{c} j \\ \nu \end{array} \right)$ 为二项式符号. 上式只是是将一个 $k$ 元集划分成 $\leqslant j$ 个子集 (考虑子集次序) 分法个数的两种不同的写法. 一方面，$k$ 个元素每个都有 $j$ 种可能的归属，因此总共有 $j^k$ 中分法；另一方面，如果 $j$ 个集合中有 $\nu$ 个是非空的，则有 $\nu! \left\{ \begin{array}{c} k \\ \nu \end{array} \right\}$ 种划分办法，而又有 $\left( \begin{array}{c} j \\ \nu \end{array} \right)$ 种办法选择这 $\nu$ 个非空集合.

*3)* 将 (3.6) 写成

$$\frac{j^k}{j!} = \sum_{\nu=0}^{j} \left\{ \begin{array}{c} k \\ \nu \end{array} \right\} \frac{1}{(j-\nu)!},$$

两端分别乘以 $\lambda^j e^{-\lambda}$ 并对 $j$ 求和，由 (3.3) 得

$$\begin{aligned}
\mu_k(\lambda) &= e^{-\lambda} \sum_{j=0}^{\infty} \frac{j^k \lambda^j}{j!} \\
&= e^{-\lambda} \sum_{j=0}^{\infty} \lambda^j \sum_{\nu=0}^{j} \left\{ \begin{array}{c} k \\ \nu \end{array} \right\} \frac{1}{(j-\nu)!} \\
&= e^{-\lambda} \sum_{\nu=0}^{k} \left\{ \begin{array}{c} k \\ \nu \end{array} \right\} \sum_{j=\nu}^{\infty} \frac{\lambda^j}{(j-\nu)!} \\
&= e^{-\lambda} \sum_{\nu=0}^{k} \left\{ \begin{array}{c} k \\ \nu \end{array} \right\} \left( \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \lambda^{\nu} \right) \\
&= \sum_{\nu=0}^{k} \left\{ \begin{array}{c} k \\ \nu \end{array} \right\} \lambda^{\nu},
\end{aligned}$$

从而证得 (3.4).

寻找素数小间隔的方法引导我们定义如下的二阶双线性型 ( second bilinear form )

$$\begin{aligned}\langle f(x), g(x)\rangle_\rho :&= \langle x-\rho, f(x)g(x)\rangle \\ &= \sum_{j=0}^{\infty}(j-\rho)f(j)g(j)p(j)\end{aligned} \tag{3.7}$$

其中 $\rho \in \mathbb{R}$ . ( 注意：未必非负，故不是内积. ) 设 $\mathbf{a} = (a_0, a_1, a_2, \ldots, a_k)$，考虑

$$P_{\mathbf{a}}(x) = \sum_{i=0}^{k} a_i x^i \,, \tag{3.8}$$

及相关的二次型 ( 注意 $\mu_k(\lambda) = \langle x^k, 1\rangle$ )

$$\begin{aligned}Q = Q_{\mathbf{a}}(\lambda,\rho) :&= \langle P_{\mathbf{a}}(x), P_{\mathbf{a}}(x)\rangle_\rho = \langle x-\rho, P_{\mathbf{a}}(x)P_{\mathbf{a}}(x)\rangle \\ &= \sum_{0\leqslant i,j\leqslant k} a_i a_j \langle x-\rho, x^{i+j}\rangle \\ &= \sum_{0\leqslant i,j\leqslant k} a_i a_j \left(\mu_{i+j+1}(\lambda) - \rho\mu_{i+j}(\lambda)\right) \\ &= \sum_{0\leqslant i,j\leqslant k} a_i a_j c_{i+j} \,,\end{aligned} \tag{3.9}$$

其中定义

$$c_m = c_m(\lambda,\rho) = \mu_{m+1}(\lambda) - \rho\mu_m(\lambda) \,. \tag{3.10}$$

我们要解决的最优化问题为：固定 $\rho > 0$ ，对于由 $a_k = 1$ 正则化的向量 $\mathbf{a}$，$Q$ 何时为最大. 该问题的解与广义 Laguerre 多项式

$$L_n{}^{(\alpha)}(x) = \sum_{\nu=0}^{n}(-1)^\nu \binom{n+\alpha}{n-\nu}\frac{x^\nu}{\nu!} \,, \qquad (\alpha > -1) \tag{3.11}$$

有关. 已知 Laguerre 多项式的零点皆为实的、正的和单的 ( 见 [15]，第 6 章 )，记 $L_n{}^{(\alpha)}(x)$ 最小的零点为 $x_1(n,\alpha)$. 下面的命题即给出了上述最优化问题的解.

**命题 3** 对任意的 $k \geqslant 1$ 及固定的 $\rho > k$，当 $0 < \lambda < x_1(k+1,\, \rho-k-1)$ 时有

$$\max_{a_k=1} Q_{\mathbf{a}}(\lambda,\rho) = -(k+1)!\lambda^k \frac{L_{k+1}{}^{(\rho-k-1)}(\lambda)}{L_k{}^{(\rho-k)}(\lambda)} \,. \tag{3.12}$$

因此，对任意的 $k \geqslant 1$ 和 $\rho > k$，

$$\inf\{\lambda > 0 \,:\, Q_{\mathbf{a}}(\lambda,\rho) > 0,\, a_k = 1\} = x_1(k+1,\, \rho-k-1) \,. \tag{3.13}$$

该命题的证明最终将归结为计算行列式

$$D_k = \det \begin{vmatrix} c_0 & c_1 & c_2 & \cdots & c_k \\ c_1 & c_2 & c_3 & \cdots & c_{k+1} \\ c_2 & c_3 & c_4 & \cdots & c_{k+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_k & c_{k+1} & c_{k+2} & \cdots & c_{2k} \end{vmatrix} = \det \left[c_{i+j}\right]_{\substack{i=0,1,2,\ldots,k \\ j=0,1,2,\ldots,k}} . \tag{3.14}$$

由于最优化问题的解可通过选择能使得 $P_{\mathbf{a}}(x)$ 与所有低次多项式关于 $\langle\ ,\ \rangle_\rho$ 正交的 $\mathbf{a}$ 得到，因此我们要考虑如下 $k$ 个方程

$$\langle P_{\mathbf{a}}(x), x^i \rangle_\rho = 0, \quad i = 0, 1, 2, \ldots, k-1\,, \tag{3.15}$$

并且证明下述引理.

**引理 1** *若对一给定的 $\lambda$ 有 $D_{k-1} \neq 0$，则存在一个向量 $\mathbf{a}$，它满足 $a_k = 1$ 及 (3.15)，且使得*

$$Q_{\mathbf{a}}(\lambda, \rho) = \frac{D_k}{D_{k-1}} \tag{3.16}$$

*成立.*

**证明.** 取 $a_k = 1$. 方程 (3.15) 等价于

$$\sum_{j=0}^{k} a_j c_{i+j} = 0, \quad i = 0, 1, 2, \ldots, k-1\,. \tag{3.17}$$

若 $\mathbf{a}$ 满足方程组 (3.17)，则有

$$\begin{aligned} Q &= \sum_{i=0}^{k} a_i \left( \sum_{j=0}^{k} a_j c_{i+j} \right) \\ &= \sum_{j=0}^{k} a_j c_{j+k}\,. \end{aligned} \tag{3.18}$$

将 (3.17) 写成如下形式

$$\begin{aligned} &c_0 a_0 + c_1 a_1 + c_2 a_2 + \cdots + c_{k-1} a_{k-1} = -c_k \\ &c_1 a_0 + c_2 a_1 + c_3 a_2 + \cdots + c_k a_{k-1} = -c_{k+1} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &c_{k-1} a_0 + c_k a_1 + c_{k+1} a_2 + \cdots + c_{2k-2} a_{k-1} = -c_{2k-1} \end{aligned} \tag{3.19}$$

由 Cramer 法则知，若 $D_{k-1} \neq 0$，则这些方程的解为

$$a_j = -\frac{D_{k-1}^{(j+1)}}{D_{k-1}}, \quad j = 0, 1, \ldots, k-1\,, \tag{3.20}$$

其中 $D_{k-1}^{(i)}$ 是用列向量 $(c_k, c_{k+1}, \ldots, c_{2k-1})$ 替换 $D_{k-1}$ 的第 $i$ 列而得到的行列式．于是 (3.18) 给出

$$Q = \frac{1}{D_{k-1}} \left( -\sum_{j=0}^{k-1} D_{k-1}^{(j+1)} c_{k+j} + D_{k-1} c_{2k} \right)\,. \tag{3.21}$$

另一方面，若将 $D_k$ 按其最后一行进行余子式展开，可得

$$D_k = \sum_{j=0}^{k} (-1)^{k+j} D_{k+1,j+1} c_{k+j}\,,$$

其中 $D_{i,j}$ 是移去 $D_k$ 中第 $i$ 行第 $j$ 列而得到的行列式．从 (3.14) 知

$$D_{k+1,j+1} = (-1)^{k-j-1} D_{k-1}^{(j+1)}, \quad 1 \leqslant j \leqslant k-1, \quad D_{k+1,k+1} = D_{k-1}$$

其中因子 $(-1)^{k-j-1}$ 是由 $D_k$ 的最后一列向左平移 $k-(j+1)$ 个位置而产生的．因此我们有

$$Q = \frac{D_k}{D_{k-1}}\,. \tag{3.22}$$

下面计算 $D_k$.

**引理 2**

$$D_{k-1} = (-1)^k 1!\,2!\,3!\,\cdots\,k!\,\lambda^{\frac{k(k-1)}{2}} L_k{}^{(\rho-k)}(\lambda)\,. \tag{3.23}$$

**证明．** 我们断言

$$D_{k-1} = (-1)^k E_k\,, \tag{3.24}$$

其中

$$E_k = \det \begin{vmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_k \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{k+1} \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{k+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{k-1} & \mu_k & \mu_{k+1} & \cdots & \mu_{2k-1} \\ 1 & \rho & \rho^2 & \cdots & \rho^k \end{vmatrix} = \det \begin{vmatrix} \mu_{i+j} \\ \rho^j \end{vmatrix}_{\substack{i=0,1,2,\ldots,k-1 \\ j=0,1,2,\ldots,k}}\,. \tag{3.25}$$

在 $E_k$ 中，依次将第 $(\ell+1)$ 列减去第 $\ell$ 列的 $\rho$ 倍 ( $1 \leqslant \ell \leqslant k$ )，得

$$E_k = \det \begin{vmatrix} \mu_0 & c_0 & c_1 & \dots & c_{k-1} \\ \mu_1 & c_1 & c_2 & \dots & c_k \\ \mu_2 & c_2 & c_3 & \dots & c_{k+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{k-1} & c_{k-1} & c_k & \dots & c_{2k-2} \\ 1 & 0 & 0 & \dots & 0 \end{vmatrix},$$

再最后一行进行余子式展开即得 $E_k = (-1)^k D_{k-1}$.

引进微分算子

$$D = \frac{d}{d\lambda}, \quad \delta = \lambda D = \lambda \frac{d}{d\lambda}, \quad \Delta = \delta + \lambda = \lambda \frac{d}{d\lambda} + \lambda\,. \tag{3.26}$$

显然有关系式

$$\delta^k = \lambda^k D^k + \sum_{j=1}^{k-1} a_j(\lambda) D^j, \quad \Delta^k = \delta^k + \sum_{j=0}^{k-1} b_j(\lambda) \delta^j \tag{3.27}$$

成立，其中 $a_j(\lambda)$ 和 $b_j(\lambda)$ 是 $\lambda$ 的 $j$ 次多项式．由 (3.4) 及 (3.5) 知 $\mu_k = \Delta \mu_{k-1}$．一般地，

$$\mu_k = \Delta^i \mu_{k-i}\ ,\ 0 \leqslant i \leqslant k; \qquad \mu_k = \Delta^k 1\,. \tag{3.28}$$

于是

$$E_k = \det \begin{vmatrix} \mu_0 & \mu_1 & \mu_2 & \dots & \mu_k \\ \Delta\mu_0 & \Delta\mu_1 & \Delta\mu_2 & \dots & \Delta\mu_k \\ \Delta^2\mu_0 & \Delta^2\mu_1 & \Delta^2\mu_3 & \dots & \Delta^2\mu_k \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Delta^{k-1}\mu_0 & \Delta^{k-1}\mu_1 & \Delta^{k-1}\mu_2 & \dots & \Delta^{k-1}\mu_k \\ 1 & \rho & \rho^2 & \dots & \rho^k \end{vmatrix} = \det \begin{vmatrix} \Delta^i \mu_j \\ \rho^j \end{vmatrix}_{\substack{i=0,1,2,\dots,k-1 \\ j=0,1,2,\dots,k}}. \tag{3.29}$$

由 (3.27) 的第二个关系式知，我们可用 $\delta^i$ 和 $\delta$ 的较低次幂的一个线性组合来替代 $\Delta^i$，而这个线性组合可以用行运算消去．因此我们可以在上述行列式中用 $\delta$ 替代 $\Delta$ 而不改变行列式的值．由 (3.27) 的第一个关系式和行运算知，我们可用 $\lambda^i D^i$ 替代 $\delta^i$，在每一行中移去含有 $\lambda$ 的因式，得

$$\begin{aligned} E_k &= \lambda^{\frac{k(k-1)}{2}} \det \begin{vmatrix} \mu_0 & \mu_1 & \mu_2 & \dots & \mu_k \\ D\mu_0 & D\mu_1 & D\mu_2 & \dots & D\mu_k \\ D^2\mu_0 & D^2\mu_1 & D^2\mu_3 & \dots & D^2\mu_k \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ D^{k-1}\mu_0 & D^{k-1}\mu_1 & D^{k-1}\mu_2 & \dots & D^{k-1}\mu_k \\ 1 & \rho & \rho^2 & \dots & \rho^k \end{vmatrix} \\ &= \lambda^{\frac{k(k-1)}{2}} \det \begin{vmatrix} D^i \mu_j \\ \rho^j \end{vmatrix}_{\substack{i=0,1,2,\dots,k-1 \\ j=0,1,2,\dots,k}}. \end{aligned} \tag{3.30}$$

我们需要关系式

$$q(q-1)\cdots(q-h+1)=h!\begin{pmatrix} q \\ h \end{pmatrix}=\sum_{j=0}^{h}(-1)^{h-j}\begin{bmatrix} h \\ j \end{bmatrix}q^j\,, \tag{3.31}$$

其中 $\begin{bmatrix} h \\ j \end{bmatrix}$ 是第一型的 Stirling 数．由 (3.3) 得

$$\begin{aligned}\lambda^h=e^{\lambda}\lambda^h e^{-\lambda}&=\sum_{q=0}^{\infty}\frac{\lambda^{q+h}}{q!}e^{-\lambda}=\sum_{q=h}^{\infty}\frac{\lambda^q}{(q-h)!}e^{-\lambda}=\sum_{q=0}^{\infty}q(q-1)\cdots(q-h+1)\frac{\lambda^q}{q!}e^{-\lambda}\\&=\sum_{q=0}^{\infty}\left(\sum_{j=0}^{h}(-1)^{h-j}\begin{bmatrix} h \\ j \end{bmatrix}q^j\right)\frac{\lambda^q}{q!}e^{-\lambda}=\sum_{j=0}^{h}(-1)^{h-j}\begin{bmatrix} h \\ j \end{bmatrix}\mu_j(\lambda)\,.\end{aligned} \tag{3.32}$$

由行运算得

$$\begin{aligned}E_k&=\lambda^{\frac{k(k-1)}{2}}\det\begin{vmatrix} 1 & \lambda & \lambda^2 & \dots & \lambda^k \\ D1 & D\lambda & D\lambda^2 & \dots & D\lambda^k \\ D^21 & D^2\lambda & D^2\lambda^2 & \dots & D^2\lambda^k \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ D^{k-1}1 & D^{k-1}\lambda & D^{k-1}\lambda^2 & \dots & D^{k-1}\lambda^k \\ 1 & 1!\begin{pmatrix} \rho \\ 1 \end{pmatrix} & 2!\begin{pmatrix} \rho \\ 2 \end{pmatrix} & \dots & k!\begin{pmatrix} \rho \\ k \end{pmatrix}\end{vmatrix}\\&=\lambda^{\frac{k(k-1)}{2}}\det\left|\begin{matrix} D^i\lambda^j \\ j!\begin{pmatrix} \rho \\ j \end{pmatrix}\end{matrix}\right|_{\substack{i=0,1,2,\dots,k-1\\ j=0,1,2,\dots,k}}\,.\end{aligned} \tag{3.33}$$

按最后一行展开可得

$$\det\left|\begin{matrix} D^i\lambda^j \\ j!\begin{pmatrix} \rho \\ j \end{pmatrix}\end{matrix}\right|_{\substack{i=0,1,2,\dots,k-1\\ j=0,1,2,\dots,k}}=\sum_{h=0}^{k}(-1)^{k-h}h!\begin{pmatrix} \rho \\ h \end{pmatrix}\det[D^i\lambda^j]_{\substack{i=0,1,\dots,k-1\\ j=0,1,\dots,k;\,j\neq h}}\,. \tag{3.34}$$

下面先证

$$F_h=\det[D^i\lambda^j]_{\substack{i=0,1,\dots,k-1\\ j=0,1,\dots,k;\,j\neq h}}=1!\,2!\,\cdots\,(k-1)!\begin{pmatrix} k \\ h \end{pmatrix}\lambda^{k-h}\,, \tag{3.35}$$

然后由 (3.24)，(3.33)，(3.34)，(3.35) 倒推过去

$$\begin{aligned}D_{k-1}&=(-1)^k\lambda^{\frac{k(k-1)}{2}}1!\,2!\,\cdots\,(k-1)!\sum_{h=0}^{k}(-1)^{k-h}h!\begin{pmatrix} \rho \\ h \end{pmatrix}\begin{pmatrix} k \\ h \end{pmatrix}\lambda^{k-h}\\&=(-1)^k\lambda^{\frac{k(k-1)}{2}}1!\,2!\,\cdots\,k!L_k{}^{(\rho-k)}(\lambda)\,,\end{aligned}$$

即可证得引理 2．

考虑如下的完全上三角矩阵

$$
\begin{aligned}
M = [D^i\lambda^j]_{\substack{i=0,1,\dots,k\\ j=0,1,\dots,k}} &= \begin{pmatrix} 1 & \lambda & \lambda^2 & \dots & \lambda^k \\ D1 & D\lambda & D\lambda^2 & \dots & D\lambda^k \\ D^21 & D^2\lambda & D^2\lambda^2 & \dots & D^2\lambda^k \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ D^k1 & D^k\lambda & D^k\lambda^2 & \dots & D^k\lambda^k \end{pmatrix} \\
&= \left[ i!\binom{j}{i}\lambda^{j-i} \right]_{\substack{i=0,1,\dots,k\\ j=0,1,\dots,k}} .
\end{aligned}
$$

显见 $\det M = 1!\,2!\,\cdots k!$ 及

$$
M = TP, \quad \text{其中 } T = [\delta_{ij} i!]_{\substack{i=0,1,\dots,k,\\ j=0,1,\dots k}}, \quad P = \left[ \binom{j}{i}\lambda^{j-i} \right]_{\substack{i=0,1,\dots,k\\ j=0,1,\dots,k}} .
$$

由矩阵的的求逆公式

$$
M^{-1} = \frac{1}{\det M}[(-1)^{i+j} D_{j,i}]_{\substack{i=0,1,\dots,k\\ j=0,1,\dots,k}} ,
$$

$F_h$ 在这个矩阵中作为子式 $D_{k,h}$ 出现．进一步，

$$
M^{-1} = P^{-1}T^{-1} ,
$$

其中

$$
T^{-1} = [\delta_{ij}\frac{1}{i!}]_{\substack{i=0,1,\dots,k,\\ j=0,1,\dots,k}}, \quad P^{-1} = \left[ (-1)^{j-i}\binom{j}{i}\lambda^{j-i} \right]_{\substack{i=0,1,\dots,k\\ j=0,1,\dots,k}} ,
$$

我们用到恒等式

$$
\sum_{s=0}^{k}(-1)^s\binom{j}{s}\binom{s}{i} = (-1)^i\delta_{ij} .
$$

令 $M^{-1} = [\bar{m}_{ij}]$，从最后这个关系式可知

$$
F_h = (-1)^{k+h}(\det M)\bar{m}_{hk} = (-1)^{k+h}\frac{(-1)^{k-h}}{k!}1!\,2!\,\cdots\,k!\binom{k}{h}\lambda^{k-h} ,
$$

即为所证．

**命题 3 的证明．** 设 $\mathbf{a}$ 是引理 2 中给出的 (3.15) 的解，此解当 $D_{k-1} \neq 0$ 时对任意的 $\lambda$ 都存在．设 $\mathbf{b}$ 是满足 $b_k = 1$ 的任一与 $\mathbf{a}$ 不相同的 $k$-向量．于是 $P_{\mathbf{b}-\mathbf{a}}(x)$ 是一个次数 $\leqslant k-1$ 的多项式．由正交性 (3.15) 得

$$
\begin{aligned}
Q_{\mathbf{b}}(\lambda,\rho) &= \langle P_{\mathbf{a}}(x) + P_{\mathbf{b}-\mathbf{a}}(x), P_{\mathbf{a}}(x) + P_{\mathbf{b}-\mathbf{a}}(x)\rangle_\rho \\
&= \langle P_{\mathbf{a}}(x), P_{\mathbf{a}}(x)\rangle_\rho + \langle P_{\mathbf{b}-\mathbf{a}}(x), P_{\mathbf{b}-\mathbf{a}}(x)\rangle_\rho \\
&= Q_{\mathbf{a}}(\lambda,\rho) + Q_{\mathbf{b}-\mathbf{a}}(\lambda,\rho) .
\end{aligned}
$$

一般地，假定 $\rho > 0$ 固定，对任意的 $\mathbf{c} \neq \mathbf{0}$，由 (3.7) 和 (3.9) 知，对于 $0 < \lambda \leqslant \lambda_0(\mathbf{c}, \rho)$，有

$$\begin{aligned} Q_{\mathbf{c}}(\lambda, \rho) &= \sum_{j=0}^{\infty} (j - \rho)(P_{\mathbf{c}}(j))^2 p(j) \\ &= -\rho \langle (P_{\mathbf{c}}(x))^2, 1 \rangle + O_{\mathbf{c}}(\lambda) \\ &< 0\,, \end{aligned}$$

其中 $\lambda_0(\mathbf{c}, \rho)$ 是依赖于 $\mathbf{c}$ 的小的正常数．因此对 $0 < \lambda < \lambda_0(\mathbf{c})$ 有

$$Q_{\mathbf{b}}(\lambda, \rho) \leqslant Q_{\mathbf{a}}(\lambda, \rho)\,,$$

从而证得当 $\lambda$ 充分小时 $Q_{\mathbf{a}}$ 为最大．

显然只要对任意的 $(k-1)$-向量 $\mathbf{c}$ 有 $Q_{\mathbf{c}} < 0$ 成立，或者只要对任意的 $(k-1)$-向量 $Q_{\mathbf{c}}$ 的最大值 $Q$ 是负的，则上述结论对更大的 $\lambda$ 依然成立．由 Szegö [15] 中的 (5.1.14) 知

$$\frac{d}{dx} L_n{}^{(\alpha)}(x) = -L_{n-1}{}^{(\alpha+1)}(x)\,,$$

即 Laguerre 多项式序列 $\{L_k{}^{(\rho-k)}\}$ 各项的负导数等于其前一项．( 故此 (3.12) 分子中 Laguerre 多项式的负导数就是分母中的 Laguerre 多项式．) 进一步， Laguerre 多项式序列在第一个 (正) 零点之前皆为递减函数，因此最小正零点 $x_1(k, \rho - k)$ 的序列是一个递减序列．从 $k = 1$ 的平凡情形开始我们看到当 $\mathbf{a}$ 满足 (3.15) 时，$Q_k$ 可在 $0 < \lambda < x_1(k+1, \rho - k - 1)$ 上达到最大．命题 3 证毕！

下面的引理渐近地计算当 $n \to \infty$ 时的最小正零点 $x_1(n, \alpha)$．

**引理 3** 设 $L_n{}^{(\alpha)}(x)$，$\alpha > -1$ 为 Laguerre 多项式．$L_n{}^{(\alpha)}(x)$ 的零点是实的、正的和单的．记 $x_1(n, \alpha)$ 为 $L_n{}^{(\alpha)}(x)$ 的最小零点．若 $\alpha = \beta(n) - n$ 且 $\lim_{n\to\infty} \frac{\beta(n)}{n} = A > 0$，则

$$\lim_{n\to\infty} \frac{x_1(n, \alpha)}{n} = (\sqrt{A} - 1)^2\,. \tag{3.36}$$

**证明．** $L_n{}^{(\alpha)}(x)$ 的性质可参见 Szegö [15]．方程 (3.36) 是 [2] 中定理 4.4 的特殊情形．由 Szegö [15] 的 (5.1.2)，微分方程

$$u'' + \left( \frac{n + (\alpha+1)/2}{x} + \frac{1 - \alpha^2}{4x^2} - \frac{1}{4} \right) u = 0 \tag{3.37}$$

有解

$$u = e^{-x/2} x^{(\alpha+1)/2} {L_n}^{(\alpha)}(x)\,.$$

设

$$\begin{aligned} H_n(x) :&= \frac{n+(\alpha+1)/2}{x} + \frac{1-\alpha^2}{4x^2} - \frac{1}{4} \\ &= -\frac{x^2-(4n+2(\alpha+1))x+(\alpha^2-1)}{4x^2}\,, \end{aligned}$$

且记分子上二次多项式较小的根为 ${x_n}^-$．由 [11] 中的 Sturm 比较定理并注意到 $\lim\limits_{n\to\infty}\frac{\alpha}{n} = A-1$，我们有

$$\begin{aligned} \lim_{n\to\infty}\frac{x_1(n,\alpha)}{n} &= \lim_{n\to\infty}\frac{{x_n}^-}{n} \\ &= \lim_{n\to\infty}\frac{2n+\alpha+1-\sqrt{4n^2+2\alpha+2+4n\alpha+4n}}{n} \\ &= 2+\lim_{n\to\infty}\frac{\alpha}{n} - \sqrt{4+4\lim_{n\to\infty}\frac{\alpha}{n}} \\ &= (\sqrt{A}-1)^2\,. \end{aligned}$$

## 4 素数的小间隔

本节证明定理 1．我们将从统计的角度观察当 $N < n \leqslant 2N$，$N\to\infty$ 时区间 $(n, n+h]$ 中素数的个数．在这个范围之内相邻素数的平均间隔为 $\log N$，因此可将 $h$ 取为 $\log N$ 的某个倍数．记

$$\psi(x) = \sum_{n\leqslant x}\Lambda(n)\,, \tag{4.1}$$

$$\psi(n,h) = \psi(n+h)-\psi(n)\,, \tag{4.2}$$

$$h = \lambda\log N\,. \tag{4.3}$$

假定

$$\lambda \ll 1\,. \tag{4.4}$$

Gallagher 在 [4] 中证明若 Hardy-Littlewood 猜想 (2.4) 对 $h \ll \log N$ 一致成立，则我们可渐近地估计出长度为 $h$ 的区间中素数个数的各阶矩．即，假定 (2.4) 成立，则当 $N\to\infty$ 时有

$$M_k(N,h,\psi) := \frac{1}{N(\log N)^k}\sum_{n=N+1}^{2N}(\psi(n,h))^k \sim \mu_k(\lambda)\,, \tag{4.5}$$

其中 $\mu_k(\lambda)$ 是 (3.3) 和 (3.4) 中的 Poisson 矩．下面的论证仿照 [4]中的方法.

取 $N < n \leqslant 2N$，显见

$$\begin{aligned}\psi(n,h)^k &= \sum_{1\leqslant h_1,h_2,\ldots,h_k\leqslant h} \Lambda(n+h_1)\Lambda(n+h_2)\cdots\Lambda(n+h_k) \\ &= (1+o(1)) \sum_{1\leqslant h_1,h_2,\ldots,h_k\leqslant h} (\log N)^{k-|\mathcal{H}|}\Lambda(n;\mathcal{H})\,.\end{aligned} \tag{4.6}$$

设向量 $\mathbf{H}$ 的分量是集合 $\mathcal{H}$ 中的元素．为定义此处的替代，我们需将 $\Lambda_R(n;\mathcal{H})$ 的定义扩充到向量 $\mathbf{H} = (h_1, h_2, \ldots, h_k)$．定义

$$\Lambda_R(n;\mathbf{H}) := (\log R)^{k-|\mathcal{H}|}\Lambda_R(n;\mathcal{H})\,. \tag{4.7}$$

于是我们可用

$$\psi_R{}^{(k)}(n,h) := \sum_{1\leqslant h_1,h_2,\ldots,h_k\leqslant h} \Lambda_R(n;\mathbf{H})\,. \tag{4.8}$$

来替代 $\psi(n,h)^k$．为方便起见，定义 $\psi_R{}^{(0)}(n,h) = 1$.

设 $k = i + j$，定义 approximate moments

$$M_{ij}(R) = \frac{1}{N(\log R)^k} \sum_{n=N+1}^{2N} \psi_R{}^{(i)}(n,h)\psi_R{}^{(j)}(n,h)\,. \tag{4.9}$$

( 注意 $\psi_R{}^{(0)}(n,h) = 1$，$k = i + j = 0 \Rightarrow M_{00}(R) = 1$ ).

及混合矩 ( mixed moments )

$$\tilde{M}_{ij}(R) = \frac{1}{N(\log R)^{k+1}} \sum_{n=N+1}^{2N} \psi_R{}^{(i)}(n,h)\psi_R{}^{(j)}(n,h)\psi(n,h)\,. \tag{4.10}$$

由素数定理知[8]

$$\tilde{M}_{00}(R) = \frac{1}{N\log R} \sum_{n=N+1}^{2N} \psi(n,h) \sim \frac{h}{\log R} = \frac{\lambda\log N}{\log R} \sim \frac{\lambda}{\theta} = \mu_1\left(\frac{\lambda}{\theta}\right)$$

[8]

$$\begin{aligned}\sum_{n=N+1}^{2N}\psi(n,h) &= \sum_{n=N+1}^{2N}(\psi(n+h)-\psi(n)) = \sum_{n=N+h+1}^{2N+h}\psi(n) - \sum_{n=N+1}^{2N}\psi(n) \\ &= \sum_{n=2N+1}^{2N+h}\psi(n) - \sum_{n=N+1}^{N+h}\psi(n) = \sum_{x=1}^{h}\psi(2N+x) - \sum_{x=1}^{h}\psi(N+x) \\ &= \sum_{x=1}^{h}(\psi(2N+x)-\psi(N+x)) \sim Nh\,.\end{aligned}$$

其中 $\theta$ 由 (4.11) 定义．此结果与 (4.13) 一致．

下面用命题 1 和 2 来证明这些 approximate moments 渐近的是关于截断阶 (truncation level) $R$ 的具有递增期望值的 Poisson 矩．定义 $\theta$ 如下

$$R = N^{\theta}. \tag{4.11}$$

**命题 4** 当 $N \to \infty$ 时，对于 $k = i + j \geqslant 1$ 及任意固定的 $0 < \theta < \frac{1}{k}$，有

$$M_{ij}(R) = (1 + o_k(1))\mu_k\left(\frac{\lambda}{\theta}\right); \tag{4.12}$$

对任意的 $0 < \theta < \frac{1}{2k}$，有

$$\tilde{M}_{ij}(R) = (1 + o_k(1))\mu_{k+1}\left(\frac{\lambda}{\theta}\right). \tag{4.13}$$

**证明．** 对命题 1，2 中的式子两边进行作差可知这两个命题当求和范围变为 $N < n \leqslant 2N$ 时仍成立．我们先将命题 1 扩展到对向量 $\mathbf{H}_1$ 和 $\mathbf{H}_2$ 成立．$|\mathbf{H}|$ 代表向量 $\mathbf{H}$ 中分量的个数．记 $k = |\mathbf{H}_1| + |\mathbf{H}_2|$，$\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$，其中 $\mathcal{H}_i$ 是 $\mathbf{H}_i$ 中不同分量所构成的集合．由 (4.7) 式和命题 1 ( 注意命题 1 中的 $k$ 在此时等于 $|\mathcal{H}_1| + |\mathcal{H}_2|$ )，对 $R = o(N^{\frac{1}{k}})$ 有

$$\begin{aligned}
&\sum_{n=N+1}^{2N} \Lambda_R(n;\mathbf{H}_1)\Lambda_R(n;\mathbf{H}_2) \\
&\qquad = \sum_{n=N+1}^{2N} (\log R)^{|\mathbf{H}_1|-|\mathcal{H}_1|}\Lambda_R(n;\mathcal{H}_1)(\log R)^{|\mathbf{H}_2|-|\mathcal{H}_2|}\Lambda_R(n;\mathcal{H}_2) \\
&\qquad = (\log R)^{k-|\mathcal{H}_1|-|\mathcal{H}_2|} \sum_{n=N+1}^{2N} \Lambda_R(n;\mathcal{H}_1)\Lambda_R(n;\mathcal{H}_2) \\
&\qquad = (\log R)^{k-|\mathcal{H}_1|-|\mathcal{H}_2|} N\left(\mathfrak{S}(\mathcal{H}) + o(1)\right)(\log R)^{|\mathcal{H}_1\cap\mathcal{H}_2|} \\
&\qquad = N\left(\mathfrak{S}(\mathcal{H}) + o(1)\right)(\log R)^{k-|\mathcal{H}|}.
\end{aligned} \tag{4.14}$$

由此可见该结果只依赖于 $k$，而不依赖于单独的 $|\mathbf{H}_1|$ 和 $|\mathbf{H}_2|$．用 $h_1, h_2, \ldots, h_k$ 表示 $\mathbf{H}_1$ 和 $\mathbf{H}_2$ 的所有分量 ( 任意次序 )，则有 ( 假定 $R = o(N^{\frac{1}{k}})$ )

$$\begin{aligned}
M_{ij}(R) &= \frac{1}{N(\log R)^k} \sum_{1\leqslant h_1,h_2,\ldots,h_k\leqslant h} \sum_{n=N+1}^{2N} \Lambda_R(n;\mathbf{H}_1)\Lambda_R(n;\mathbf{H}_2) \\
&= \frac{1}{N(\log R)^k} \sum_{1\leqslant h_1,h_2,\ldots,h_k\leqslant h} N\left(\mathfrak{S}(\mathcal{H}) + o_k(1)\right)(\log R)^{k-|\mathcal{H}|} \\
&= \sum_{1\leqslant h_1,h_2,\ldots,h_k\leqslant h} \left(\mathfrak{S}(\mathcal{H}) + o_k(1)\right)(\log R)^{-|\mathcal{H}|}.
\end{aligned}$$

为利用 Gallagher 的结果，将上式按照 $h_1, h_2, \ldots, h_k$ 中互异元素的个数 $\nu$ ( $1 \leqslant \nu \leqslant k$ ) 分成 $k$ 个和式．由于将 $k$ 个 $h_i$ 划分进 $\nu$ 个不相交子集的方法共有 $\left\{{k \atop \nu}\right\}$ 种，而这 $\nu$ 个不相交子集则对应上述一个和式，因此 ( 注意 $h/\log R = \lambda/\theta$ )

$$\begin{aligned} M_{ij}(R) &= \sum_{\nu=1}^{k} \left\{{k \atop \nu}\right\} h^{\nu}(1+o_k(1))(\log R)^{-\nu} \\ &= (1+o_k(1))\mu_k\left(\frac{\lambda}{\theta}\right), \end{aligned}$$

这就证得命题 1．命题 2 按同样方法亦可证得．

现在考虑

$$\begin{aligned} \mathcal{S}_k &= \mathcal{S}_k(N, R, \lambda, \rho) \\ &= \frac{1}{N(\log R)^{2k+1}} \sum_{n=N+1}^{2N} \left(\psi(n,h) - \rho \log N\right) \left(P_k(\psi_R(n,h))\right)^2, \end{aligned} \tag{4.15}$$

其中

$$P_k(\psi_R(n,h)) = \sum_{\ell=0}^{k} a_\ell {\psi_R}^{(\ell)}(n,h)(\log R)^{k-\ell}, \tag{4.16}$$

$a_\ell$ 是 $N$, $R$, $k$, $\lambda$ 和 $\rho$ 的任意函数，是最优化问题中待定的量．将上式乘开得

$$\begin{aligned} \mathcal{S}_k &= \frac{1}{N(\log R)^{2k+1}} \sum_{0 \leqslant i,j \leqslant k} a_i a_j (\log R)^{2k-i-j} \\ &\qquad \times \sum_{n=N+1}^{2N} \left(\psi(n,h) - \frac{\rho}{\theta}\log R\right) {\psi_R}^{(i)}(n,h){\psi_R}^{(j)}(n,h) \\ &= \sum_{0 \leqslant i,j \leqslant k} a_i a_j \mathcal{M}_{ij}\,. \end{aligned} \tag{4.17}$$

设

$$\tilde{\lambda} = \frac{\lambda}{\theta}, \qquad \tilde{\rho} = \frac{\rho}{\theta}\,. \tag{4.18}$$

取 $i+j=\kappa$，假定 $0 < \theta < \frac{1}{2\kappa}$．由命题 4 得[9]

$$\begin{aligned} \mathcal{M}_{ij} &= \tilde{M}_{ij}(R) - \frac{\rho}{\theta} M_{ij}(R) \\ &= (1+o_\kappa(1))\mu_{\kappa+1}(\tilde{\lambda}) - \frac{\rho}{\theta}(1+o_\kappa(1))\mu_\kappa(\tilde{\lambda}) \\ &= \mu_{\kappa+1}(\tilde{\lambda}) - \tilde{\rho}\mu_\kappa(\tilde{\lambda}) + o_\kappa(\mu_{\kappa+1}(\tilde{\lambda})) + o_\kappa(\tilde{\rho}\mu_\kappa(\tilde{\lambda})) \\ &= \mu_{\kappa+1}(\tilde{\lambda}) - \tilde{\rho}\mu_\kappa(\tilde{\lambda}) + o_\kappa\left(\frac{1}{\theta^{\kappa+1}}\right) \\ &= c_\kappa(\tilde{\lambda}, \tilde{\rho}) + o_\kappa\left(\frac{1}{\theta^{\kappa+1}}\right), \end{aligned} \tag{4.19}$$

[9]由 (3.4) 知 $\mu_k(\lambda)$ 是一个 $k$ 次多项式；第四式是由于 $\lambda \ll 1$，$\rho$ 介于 $r$ 和 $r+1$ 之间，$r$ 是一固定的正整数．

其中 $c_\kappa(\tilde\lambda,\tilde\rho)$ 的记号来自 (3.10) 式．为了估计 $\mathcal{S}_k$，我们要对 $0\leqslant\kappa\leqslant 2k$ 应用这些结果 ( 因为 $0\leqslant i,j\leqslant k$，$\kappa=i+j$ )，还须加上条件[10]

$$\frac{1}{4k+1}\leqslant\theta<\frac{1}{4k}\,. \tag{4.20}$$

以保证这些结果成立．

于是

$$\begin{aligned}\mathcal{S}_k&=\sum_{0\leqslant i,j\leqslant k}a_ia_jc_{i+j}(\tilde\lambda,\tilde\rho)+o_k(\max_{1\leqslant\ell\leqslant k}|a_\ell|^2)\\&=Q_{\mathbf{a}}(\tilde\lambda,\tilde\rho)+o_k(1)\,,\end{aligned} \tag{4.21}$$

最后一式是由于固定 $\lambda$ 和 $\rho$ 时，$\max_{1\leqslant\ell\leqslant k}|a_\ell|^2$ 只依赖于 $k$ 而得．由命题 3 知 $Q_{\mathbf{a}}(\tilde\lambda,\tilde\rho)$ 在 Laugerre 多项式 $L_{k+1}{}^{(\tilde\rho-k-1)}(\tilde\lambda)$ 的最小零点 $x_1(k+1,\tilde\rho-k-1)$ 处改变符号：当 $0<\tilde\lambda<x_1(k+1,\tilde\rho-k-1)$ 时 $Q_{\mathbf{a}}(\tilde\lambda,\tilde\rho)$ 为负，当 $x_1(k+1,\tilde\rho-k-1)<\tilde\lambda<x_1(k,\tilde\rho-k)$ 时 $Q_{\mathbf{a}}(\tilde\lambda,\tilde\rho)$ 为正．结合 (4.21) 知 $\mathcal{S}_k$ 在 $x_1(k+1,\tilde\rho-k-1)+o_k(1)<\tilde\lambda<x_1(k,\tilde\rho-k)-o_k(1)$，$N\to\infty$ 时为正．应用引理 3，取 $\beta(k)=\tilde\rho$，及序列 $\theta=\theta_k\to\frac{1}{4k}{}^-$，$\rho=\rho_k\to r^+$ ( $k\to\infty$ )，则 $A=4r$，且必存在常数 $0<c_k<c_k'$，$c_k,c_k'\to 0$ 使得当 $(\sqrt{r}-\frac12)^2+c_k\leqslant\lambda\leqslant(\sqrt{r}-\frac12)^2+c_k'$ 时，有

$$\mathcal{S}_k\gg_k 1,\quad \mathcal{S}_k>0\,. \tag{4.22}$$

**注意：**

(1) 由条件 $x_1(k+1,\tilde\rho-k-1)+o_k(1)<\tilde\lambda<x_1(k,\tilde\rho-k)-o_k(1)$，$N\to\infty$ 及 $\tilde\lambda=\frac{\lambda}{\theta}$，$\theta=\theta_k\to\frac{1}{4k}{}^-$ $(k\to\infty)$ 及引理 3 结论 $\lim\limits_{k\to\infty}\frac{x_1(k+1,\tilde\rho-k-1)}{k}=\lim\limits_{k\to\infty}\frac{x_1(k,\tilde\rho-k)}{k}=(\sqrt{a}-1)=4(\sqrt{r}-\frac12)^2$ 知当 $(\sqrt{r}-\frac12)^2+c_k\leqslant\lambda\leqslant(\sqrt{r}-\frac12)^2+c_k'$，$0<c_k<c_k'$ 且 $c_k,c_k'\to 0$ 时，$\mathcal{S}_k$ 不只为正，而且还可以被某个正常数从下方界住，即 $\mathcal{S}_k\gg_k 1$．

(2) 由 (4.20) 知上述讨论中出现的 Laugerre 多项式是良定义的．因为由 (4.20)，$\frac{1}{4k+1}\leqslant\theta<\frac{1}{4k}\Rightarrow k\theta\leqslant\frac14$；而 $\rho$ 为 $>k\geqslant 1$ 的固定的实数，故 $\rho-k\theta>0$，$\tilde\rho-k=\frac{\rho}{\theta}-k=\frac{1}{\theta}(\rho-k\theta)>0$．从而可得证 $\tilde\rho-k-1>-1$，使其满足 “$\alpha>-1$” 的要求．

现在由 (4.22) 可给出定理 1 的证明．定义

$$Q_r^+(N,h)=\sum_{\substack{n=N+1\\ \pi(n+h)-\pi(n)>r}}^{2N}1\,. \tag{4.23}$$

[10] $\theta$ 不能太小，否则 $\frac{1}{\theta^{k+1}}$ 会很大，另外 (4.30) 的证明亦要求 $\theta$ 下方有界．

若 $n \in \mathbb{Z}$ 使得 $\pi(n+h)-\pi(n) > r$，则必存在 $j$ 使得 $n \leqslant p_j$ 且 $p_{j+r} \leqslant n+h$. 由此知 $p_{j+r}-p_j \leqslant h$ 且 $p_{j+r}-h \leqslant n \leqslant p_j < p_{j+r}$，故至多存在 $h$ 个这样的 $n$ 对应于上述长度为 $h$ 而又含有多于 $r$ 个素数的间隔. 因此

$$Q_r^+(N,h) \ll_r h \sum_{\substack{N<p_n\leqslant 2N \\ p_{n+r}-p_n\leqslant h}} 1+O\left(Ne^{-c\sqrt{\log N}}\right), \tag{4.24}$$

其中余项的出现是由于使用素数定理移除了那些跨过端点的区间.

**注：**（参见 [7]，第 62 页）.

设 $p_{j_0}=p_{j_0}(N)$ 及 $p_{j_1}=p_{j_1}(N)$ 分别代表 $[N+1,2N]$ 中最小和最大的素数. 对 $r \geqslant 1$，有

$$Q_r^+(N,h)=\sum_{\substack{N+1\leqslant n\leqslant 2N \\ \pi(n+h)-\pi(n)>r}} 1=\sum_{\substack{N+1\leqslant n<p_{j_0} \\ \pi(n+h)-\pi(n)>r}} 1+\sum_{j=j_0}^{j_1}\sum_{\substack{p_j\leqslant n<p_{j+1} \\ p_{j+1+r}\leqslant n+h}} 1-\sum_{\substack{2N<n<p_{j_1+1} \\ \pi(n+h)-\pi(n)>r}} 1.$$

第一、三项的估计可由素数定理的余项给出，即 $O\left(Ne^{-c\sqrt{\log N}}\right)$. 因此

$$\begin{aligned}
Q_r^+(N,h) &= \sum_{j=j_0}^{j_1}\sum_{\substack{p_j\leqslant n<p_{j+1} \\ p_{j+1+r}\leqslant n+h}} 1+O\left(Ne^{-c\sqrt{\log N}}\right) \\
&= \sum_{N+1\leqslant p_j\leqslant 2N}\sum_{\substack{p_j\leqslant n<p_{j+1} \\ p_{j+1+r}\leqslant n+h}} 1+O\left(Ne^{-c\sqrt{\log N}}\right) \\
&= \sum_{N+1\leqslant p_j\leqslant 2N}(p_{j+1}-\max(p_j,p_{j+1+r}-[h]))+O\left(Ne^{-c\sqrt{\log N}}\right) \\
&\leqslant h\sum_{\substack{N+1\leqslant p_j\leqslant 2N \\ p_{j+1+r}-p_{j+1}<h}} 1+O\left(Ne^{-c\sqrt{\log N}}\right).
\end{aligned}$$

另外，当 $N$ 充分大时，我们还有

$$Q_r^+(N,h)=\sum_{\substack{n=N+1 \\ \psi(n+h)-\psi(n)\geqslant\rho\log N}}^{2N} 1+O(N^{\frac{1}{2}}), \tag{4.25}$$

其中 $\rho \in (r,r+1)$. 此处的余项是因为我们扔掉了素数方幂对应的项. 由 (4.25) 式及 Cauchy 不等式得[11]

[11] 第一个不等式扔掉了 (4.15) 中的负项. 第二个不等式使用 Cauchy 不等式，其中后一个括号里扔掉了限制条件，变大. 第三个不等式再次使用 Cauchy 不等式.

$$
\begin{aligned}
\mathcal{S}_k &\leqslant \frac{1}{N(\log R)^{2k+1}} \sum_{\substack{n=N+1 \\ \psi(n+h)-\psi(n)\geqslant \rho\log N}}^{2N} \psi(n,h)\left(P_k(\psi_R(n,h))\right)^2 \\
&\leqslant \frac{1}{N(\log R)^{2k+1}} \left( \sum_{\substack{n=N+1 \\ \psi(n+h)-\psi(n)\geqslant \rho\log N}}^{2N} \left(P_k(\psi_R(n,h))\right)^2 \right)^{\frac{1}{2}} \\
&\qquad\qquad \times \left( \sum_{n=N+1}^{2N} \psi(n,h)^2 \left(P_k(\psi_R(n,h))\right)^2 \right)^{\frac{1}{2}} \\
&\leqslant \frac{1}{N(\log R)^{2k+1}} \left( \sum_{\substack{n=N+1 \\ \psi(n+h)-\psi(n)\geqslant \rho\log N}}^{2N} 1 \right)^{\frac{1}{4}} \left( \sum_{n=N+1}^{2N} \left(P_k(\psi_R(n,h))\right)^4 \right)^{\frac{1}{4}+\frac{1}{4}} \\
&\qquad\qquad \times \left( \sum_{n=N+1}^{2N} (\psi(n,h))^4 \right)^{\frac{1}{4}} \\
&\leqslant \frac{\sqrt[4]{Q_r^+(N,h)+O(N^{1/2})}}{N(\log R)^{2k+1}} \left( \sum_{n=N+1}^{2N} \psi(n,h)^4 \right)^{\frac{1}{4}} \left( \sum_{n=N+1}^{2N} \left(P_k(\psi_R(n,h))\right)^4 \right)^{\frac{1}{2}} .
\end{aligned}
\tag{4.26}
$$

由 (4.22) 式成立，得

$$
Q_r^+(N,h)+O(N^{1/2}) \gg_k \frac{\left(N(\log R)^{2k+1}\right)^4}{\left( \sum_{n=N+1}^{2N} \psi(n,h)^4 \right) \left( \sum_{n=N+1}^{2N} \left(P_k(\psi_R(n,h))\right)^4 \right)^2} . \tag{4.27}
$$

后面我们将证明当 $h \ll \log N$ 时有

$$
\sum_{n=N+1}^{2N} \psi(n,h)^4 \ll N(\log N)^4 \tag{4.28}
$$

及

$$
\sum_{n=N+1}^{2N} \left(P_k(\psi_R(n,h))\right)^4 \ll N(\log N)^{4k} . \tag{4.29}
$$

由 (4.24) 及 (4.27) 知，对 $\lambda = (\sqrt{r}-\frac{1}{2})^2 + c_k$，有

$$
\sum_{\substack{N<p_n\leqslant 2N \\ p_{n+r}-p_n\leqslant h}} 1 \gg \frac{1}{h} Q_r^+(N,h) \gg_k \frac{N}{h} \gg_k \pi(N) , \tag{4.30}
$$

其中 $c_k \to 0^+$ ( $k \to \infty$ )．定理 1 证毕．

先证 (4.28) 式．利用上界筛法 ( [8]，定理 5.7 )

$$\sum_{n\leqslant N}\Lambda(n;\mathcal{H}_k)\leqslant(2^k k!+\epsilon)\mathfrak{S}(\mathcal{H}_k)N\,, \tag{4.31}$$

由 (4.6) 及 (2.13) 式知 ( 注意 $h=\lambda\log N$ )

$$\begin{aligned}\sum_{n=N+1}^{2N}\psi(n;h)^4&=\sum_{n=N+1}^{2N}(1+o(1))\sum_{1\leqslant h_1,h_2,h_3,h_4\leqslant h}(\log N)^{4-|\mathcal{H}|}\Lambda(n;\mathcal{H})\\&=(1+o(1))\sum_{\nu=1}^{4}\left\{{4\atop\nu}\right\}(\log N)^{4-\nu}\sum_{\substack{1\leqslant h_1,\ldots,h_\nu\leqslant h\\ \text{互异}}}\left(\sum_{n=N+1}^{2N}\Lambda(n;\mathcal{H}_\nu)\right)\\&\leqslant(N+o(N))\sum_{\nu=1}^{4}\left\{{4\atop\nu}\right\}2^\nu\nu!(\log N)^{4-\nu}\left(\sum_{\substack{1\leqslant h_1,\ldots,h_\nu\leqslant h\\ \text{互异}}}\mathfrak{S}(\mathcal{H}_\nu)\right)\\&\leqslant\left(\sum_{\nu=1}^{4}\left\{{4\atop\nu}\right\}\nu!(2\lambda)^\nu+\epsilon\right)N\log^4 N\,,\end{aligned} \tag{4.32}$$

(4.28) 式得证．

下证 (4.29) 式．$k\geqslant 1$，$\mathcal{H}=\{h_1,h_2,\ldots,h_r\}$，$h_i$ 为互异的整数，$\mathbf{a}=(a_1,a_2,\ldots,a_r)$，$a_i\geqslant 1$ 且 $\sum\limits_{i=1}^{r}a_i=k$．设

$$\mathcal{S}_k(N,\mathcal{H},\mathbf{a})=\sum_{n=1}^{N}\Lambda_R(n+h_1)^{a_1}\Lambda_R(n+h_2)^{a_2}\cdots\Lambda_R(n+h_r)^{a_r}\,. \tag{4.33}$$

对于 $\max\limits_i|h_i|\leqslant R$，$R\geqslant 2$，我们有

$$\mathcal{S}_k(N,\mathcal{H},\mathbf{a})=\left(\mathcal{C}_k(\mathbf{a})\mathfrak{S}(\mathcal{H})+o_k(1)\right)N(\log R)^{k-r}+O(R^k)\,, \tag{4.34}$$

其中 $\mathcal{C}_k(\mathbf{a})$ 是可计算的有理常数．将 (4.29) 式的左端展开得到 $(k+1)^4$ 个如下形式项的线性组合：

$$\mathcal{T}(\ell_1,\ell_2,\ell_3,\ell_4)=(\log R)^{4k-\ell_1-\ell_2-\ell_3-\ell_4}\sum_{n=N+1}^{2N}\prod_{i=1}^{4}\psi_R{}^{(\ell_i)}(n,h)\,,$$

其中 $0\leqslant\ell_1,\ell_2,\ell_3,\ell_4\leqslant k$．设 $\ell=\ell_1+\ell_2+\ell_3+\ell_4$，于是 $0\leqslant\ell\leqslant 4k$．我们有

$$\mathcal{T}(\ell_1,\ell_2,\ell_3,\ell_4)=(\log R)^{4k-\ell}\sum_{1\leqslant m_1,m_2,\ldots,m_\ell\leqslant h}\sum_{n=N+1}^{2N}\prod_{i=1}^{4}\Lambda_R(n,\mathbf{H}_i),$$

其中 $m_1, m_2, \ldots, m_\ell$ 跑遍 $\mathbf{H}_i$，$1 \leqslant i \leqslant 4$ 的所有分量．设 $\mathcal{H}$ 是不同分量所构成的集合，由 (4.7) 及 (4.34) 得

$$\sum_{n=N+1}^{2N} \prod_{i=1}^{4} \Lambda_R(n, \mathbf{H}_i) \ll_k N\mathfrak{S}(\mathcal{H})(\log R)^{\ell-|\mathcal{H}|},$$

再由 (2.13) 得

$$\begin{aligned}
\mathcal{T}(\ell_1, \ell_2, \ell_3, \ell_4) &\ll_k (\log R)^{4k-\ell} \sum_{j=1}^{\ell} \sum_{\substack{1 \leqslant h_1, \ldots, h_j \leqslant h \\ \text{互异}}} N\mathfrak{S}(\mathcal{H}_j)(\log R)^{\ell-j} \\
&\ll_k N(\log R)^{4k} \sum_{j=1}^{\ell} \frac{h^j}{(\log R)^j} \\
&\ll_k N(\log N)^{4k},
\end{aligned}$$

于是 (4.29) 式得证．

## 参 考 文 献


[1] E. Bombieri and H. Davenport, *Small differences between prime numbers*, Proc. Roy. Soc. Ser. **A**, **293** (1966), 1–18.

[2] H. Dette and W. J. Studden, *Some new asymptotic properties for the zeros of Jacobi, Laguerre and Hermite polynomials*, Constr. Approx. **11** (1995), no. 2, 227–238.

[3] P. Erdös, *The difference of consecutive primes*, Duke Math. J. **6** (1940), 438–441.

[4] P. X. Gallagher, *On the distribution of primes in short intervals*, Mathematika **23** (1976), 4–9.

[5] D. A. Goldston, J. Pintz and C. Y. Yıldırım, *Primes in tuples III: On the difference* $p_{n+\nu} - p_n$, Functiones et Approximatio, **XXXV** (2006), 79–89.

[6] D. A. Goldston, J. Pintz and C. Y. Yıldırım, *The pathway to recent progress on small gaps between primes*, Clay Math. Proc. **7** (2007), 129--139.

[7] D. A. Goldston and C. Y. Yıldırım, *Higher correlations of divisor sums related to primes. I: Triple correlations*, Integers **3** (2003), A5, 66 pp.

[8] H. Halberstam and H. -E. Richert, *Sieve methods*, Academic Press, London-New York, 1974.

[9] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio Numerorum': III On the expression of a number as a sum of primes*, Acta Math. **44** (1923), 1–70.

[10] H. Maier, *Small differences between prime numbers*, Michigan Math. J. **35** (1988), 323–344.

[11] D. S. Moak, E. B. Saff, R. S. Varga, *On the zeros of Jacobi polynomials* $P_n^{(\alpha_n,\beta_n)}(x)$, Trans. Amer. Math. Soc., **249** (1979), no. 1, 159–162.

[12] K. Prachar, *Primzahlenverteilung*, Springer-Verlag, 1958.

[13] R. A. Rankin, *The difference between consecutive prime numbers. II*, Proc. Cambridge Philos. Soc. **36** (1940), 255–266.

[14] A. Selberg, *On the zeros of Riemann's zeta-function*, Skr. Norske Vid. Akad. Oslo I. **1942** (1942). no. 10, 59 pp.

[15] G. Szegö, *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ. vol. 23, Amer. Math. Soc., Providence, R. I., 1939.