Analytic number theory

Anirban Mukhopadhyay

Contents

Introduction to Sieve Methods 1

Chapter 1. Basic formulation 3
1. General set-up 3
2. Examples 3
3. Sieve weights 4
4. Composition of sieves 5
5. Sieve via Buchstab's equality 6

Chapter 2. Selberg's Sieve 9
1. Introduction 9
2. More precise estimates 15
3. Estimating the error term 18
4. Application 21

Chapter 3. Large sieve inequality 23
1. Motivation 23
2. Choosing b_n's 26
3. Stronger results 27
4. Large sive for additive characters 28
5. Discrepency of distribution in arithematic progression 29
6. Least quadratic non-residue 30
7. Arithmetic large sieve 31

Introduction to Circle Method 35

Chapter 4. Vinogradov's three primes theorem 37
1. Method of proof 37
2. Main theorem 38
3. Major and Minor arcs 39
4. Major arc 41
5. Minor arc 46

Bibliography 53

Introduction to Sieve Methods

CHAPTER 1

Basic formulation

1. General set-up

Let $\mathcal{A} = (a_n)_n$ be a sequence of non-negative reals for $n \leq x$ and $\mathcal{P}$ is a set of primes and $1 < z \leq x$. Define

$$S(\mathcal{A}, \mathcal{P}, z) = \sum_{\substack{n \leq x \\ (n, P(z))=1}} a_n, \quad \text{where} \quad \mathrm{P(z)} = \prod_{\substack{p \in \mathcal{P} \\ p < z}} \mathrm{p}.$$

Using Mobius function to remove the coprimality condition we get

$$S(\mathcal{A}, \mathcal{P}, z) = \sum_{d | P(z)} \mu(d) A_d(x), \quad \text{where} \quad \mathrm{A_d(x)} = \sum_{\mathrm{n} \leq \mathrm{xd} | \mathrm{n}} \mathrm{a_n}.$$

We assume that $A_d(x) = g(d)X + r_d$ where $g(d)$ is a multiplicative function satisfying $0 < g(p) < 1$ and X is a suitable approximation to the sum $\sum_{n \leq x} a_n$. Thus we have

$$\begin{aligned} S(\mathcal{A}, \mathcal{P}, z) &= X \sum_{d | P(z)} \mu(d) g(d) + O(\sum_{d | P(z)} |r_d(\mathcal{A})|). \\ &= XV(z) + +O(\sum_{d | P(z)} |r_d(\mathcal{A})|) \end{aligned}$$

where $V(z) = \prod_{p < z}(1 - g(p))$.

2. Examples

1. Let $\mathcal{A} = \{m \in \mathbb{Z} : x - y < m \leq x\}$ for $1 < y < x$ and $\mathcal{P}$ be the set of all primes. Then $X = y$ and $g(p) = 1/p$ and $S(\mathcal{A}, \mathcal{P}, z)$ counts the number of integers in $(x - y, x]$ which does not have prime factors $< z$. In particular it counts primes if we take $z = \sqrt{x}$.

2. Let a_n be the indicator function of integers $n \leq x$ of the form $n = m^2 + 1$ and $\mathcal{P} = \{p : p \not\equiv 3 \mod 4\}$. Then $X = \sqrt{(}x)$ and since $m^2 + 1 \equiv o(p)$ has 2 solution modulo p if $p \equiv 1 \mod 4$ and it has only

one solution modulo 2. Thus

$$g(p) = \begin{cases} \frac{2}{p} & \text{if } p \equiv 1 \mod 4, \\ \frac{1}{2} & \text{if } p = 2. \end{cases}$$

We also have $r_d(\mathcal{A}) = 2^{\omega(d)}$ where $\omega(d)$ is the number of distinct prime factors of d. This counts primes of the form $m^2 + 1$ if we take $z = \sqrt{x}$.

3. Let a_n be the indicator function of the integers n of the form $n = m(m+2)$ and $\mathcal{P}$ be the set of all primes. We get

$$g(p) = \begin{cases} \frac{2}{p} & \text{if } p \text{ is odd}, \\ \frac{1}{2} & \text{if } p = 2. \end{cases}$$

Here we count twin primes if $z = x^{1/4}$.

4. This example gives another way of counting twin primes. Let a_n be indicator function of integrs of the form $p - 2$ where $p \leq x$ is a odd prime and $\mathcal{P}$ is the set of all odd primes. Then $A_d(x) = \{p \leq x : p - 2 \equiv 0 \mod d\} = \pi(x, d, 2)$ where $\pi(x, q, a)$ denone the number pf primes $\leq x$ in the arithmetic progression $a \mod q$. Thus $X = \pi(x)$ and $g(p) = 1/\phi(p)$. The error

$$\sum_d |r_d| = \sum_d |\pi(x, d, 2) - \frac{\pi(x)}{\phi(d)}|$$

can be estimated by Bombieri - Vinogadov theorem.

3. Sieve weights

We have seen in the last section that

$$S(\mathcal{A}, \mathcal{P}, z) \sum_{d|P(z)} \mu(d) A_d(x).$$

Instead of Mobius function we can use some other sequence λ_d for defined for $d|P(z)$ and supported on $d < D$ and define

$$S^{\lambda}(\mathcal{A}, \mathcal{P}, z) = \sum_{d|(P(z)} \lambda_d A_d(x).$$

For simplicity we would use $S^\lambda(\mathcal{A}, z)$ instead of $S^\lambda(\mathcal{A}, \mathcal{P}, z)$. We call $(\lambda_d)_d$ sieve weights and D is called sieve level. Now

$$\begin{aligned} S^\lambda(\mathcal{A}, z) &= \sum_{d|P(z)} \lambda_d \sum_{\substack{d\leq x \\ d|n}} a_n \\ &= \sum_{n\leq x} a_n \sum_{d|(n,P(z))} \lambda_d \\ &= \sum_{n\leq x} a_n\theta_n \quad \text{where } \theta_n = \sum_{d|n} \lambda_d. \end{aligned}$$

If we have sequences $(\lambda_d^+)_d$ and $(\lambda_d^-)_d$ for d dividing $P(z)$ and supported on $d < D$ such that

$$\theta_n^- \leq \sum_{d|n} \mu(d) \leq \theta_n^+, \ \forall n \tag{3.1}$$

where $\theta^\pm = \lambda^\pm * 1$.

Then

$$\sum_{n\leq x} a_n\theta_n^- \leq S(\mathcal{A}, z) \leq \sum_{n\leq x} a_n\theta_n^+$$

which gives

$$\sum_{d|P(z)} \lambda_d^- A_d(x) \leq S(\mathcal{A}, z) \leq \sum_{d|P(z)} \lambda_d^+ A_d(x).$$

Thus quite naturally $(\lambda_d^-)_d$ and $(\lambda_d^+)_d$ are called the lower and upper bound sieves respectively.

Hence a sieve is given by its weights $(\lambda_d)_d$ defined for $d|P(z)$ and supported on $d < D$ and it is a lower or upper bound sieve depending on which inequality in (3.1) the corresponding $(\theta_n)_n$ satisfies.

We note that

$$S^\lambda(\mathcal{A}, z) = XV(D, z) + R(D, z)$$

where

$$V(D, z) = \sum_{\substack{d|P(z) \\ d<D}} g(d)\lambda_d \quad \text{and} \quad R(D, z) = \sum_{\substack{d|P(z) \\ d<D}} \lambda_d r_d.$$

4. Composition of sieves

Let $\Lambda' = (\lambda'_d)$ and $\Lambda'' = (\lambda''_d)$ be sieves of level D' and D'' respectively, then the composition sieve (λ_d) is given by

$$\lambda_d = \sum_{[d_1,d_2]=d} \lambda'_{d_1}\lambda''_{d_2}.$$

Then $(\lambda_d)_d$ is a sieve of level $D = D'D''$ and the corresponding $\theta_n = \theta'_n\theta''_n$ for all n. Thus composition of a lower bound and upper bound sive would be a lower bound sieve.

Upper bound sieves are relatively easy to set up. Thus one can get a lower bound sieve by composition of a strong upper bound sieve with a simple lower bound sieve. Keeping that in mind we give an example of a very simply lower bound sieve.

Let $\Lambda = (\lambda_d)_d$ be defined by $\lambda_1 = 1$ and $\lambda_p = -1$ for all primes p and $\lambda_d = 0$ in all other cases. So $\theta_n = \sum_{d|n} \lambda_d = 1 - \sum_{p|n} 1 = 1 - \omega(n) < 0$ for all $n > 1$. Hence Λ is a lower bound sieve.

Suppose $\Lambda^+ = (\lambda^+_d)_d$ is an upper bound sieve. Composing with the above lower bound sieve we get a lower bound sieve $\Lambda' = (\lambda'_d)$ given by

$$\lambda'_d = \sum_{[d_1,d_2]=d} \lambda_{d_1}\lambda^+_{d_2} = \lambda^+_d - \sum_{[p,d_2]=d} \lambda^+_{d_2} = \lambda^+_d - \sum_{p|d}(\lambda^+_d + \lambda^+_{d/p}).$$

5. Sieve via Buchstab's equality

Let $\mathcal{A}_d = \{a_n : d|n\}$ and $\mathcal{A}_1 = \mathcal{A}$ and let $|\mathcal{A}_d|$ denote $A_d(x)$ which is the sum of the elements of $\mathcal{A}_d$, not the cardinality. It is easy to see that

$$S(\mathcal{A}, z) = |\mathcal{A}| - \sum_{p<z} S(\mathcal{A}_p, p) \quad \text{where} \quad S(\mathcal{A}_p, p) = \sum_{\substack{p|n \\ (n,P(p))=1}} a_n \tag{5.1}$$

with $P(p)$ denoting the product of primes $< p$ as usual. Using (5.2) again

$$S(\mathcal{A}_p, p) = |\mathcal{A}_p| - \sum_{p_1<p} S(\mathcal{A}_{pp_1}, p_1).$$

Hence we get

$$S(\mathcal{A}, z) = |\mathcal{A}| - \sum_{p<z} |\mathcal{A}_p| + \sum \sum_{p_1<p<z} S(\mathcal{A}_{pp_1}, p_1).$$

Iterating the same process r times we get

$$S(\mathcal{A}, z) = \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)|\mathcal{A}_d| + (-1)^r \sum_{\substack{d|P(z) \\ \omega(d)=r}} S(\mathcal{A}_d, p^-(d)) \tag{5.2}$$

where $p^-(d)$ denotes the smallest prime divisor of d. Thus $\Lambda = (\lambda_d)_d$ defined by

$$\lambda_d = \begin{cases} \mu(d) & \text{if } \omega(d) < r \\ 0 & \text{if } \omega(d) \geq r \end{cases}$$

gives an upper/ lower bound sieve if r is odd/ even.

Recall that $V(z) = \prod_{p<z}(1 - g(p))$. Using an agrument similar to the Buchstab euality it is easy to see that

$$V(z) = 1 - \sum_{p<z} g(p)V(p).$$

After iterating r times

$$V(z) = \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)g(d) + (-1)^r \sum_{\substack{d|P(z) \\ \omega(d)=r}} V(p^-(d)).$$

Using $|\mathcal{A}_d| = A_d(x) = Xg(d) + r_d$ in (5.2) we get

$$S(\mathcal{A}, z) = X \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)g(d) + \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)r_d + (-1)^r \sum_{\substack{d|P(z) \\ \omega(d)=r}} S(\mathcal{A}_d, p^-(d)).$$

Using the above equation for $V(z)$ we get

$$S(\mathcal{A}, z) = XV(z) + \sum_{\substack{d|P(z) \\ \omega(d)<r}} \mu(d)r_d + (-1)^r \sum_{\substack{d|P(z) \\ \omega(d)=r}} \{S(\mathcal{A}_d, p^-(d)) - Xg(d)V(p^-(d))\}.$$

Using trivial estimates

$$S(\mathcal{A}_d, p^-(d)) \leq |\mathcal{A}_d| = g(d)X + r_d \quad \text{and} \quad g(d)V(p^-(d)) \leq g(d)X$$

we get

$$S(\mathcal{A}, z) = XV(z) + 2\theta XG_r + \theta R_r$$

where for $r > 1$

$$G_r = \sum_{\substack{d|P(z) \\ \omega(d)=r}} g(d) \quad \text{and} \quad R_r = \sum_{\substack{d|P(z) \\ \omega(d)\leq r}} |r_d|.$$

Defining $G_1 = G = \sum_{p<z} g(p)$. We note that

$$G_1 \leq \sum_{p<z} -\log(1 - g(p)) = -\log V(z) = |\log V(z)|.$$

For $r > 1$ we get

$$G_r \leq \frac{1}{r!}G^r \leq \frac{1}{e}(\frac{e}{r}|\log V(z)|)^r$$

To make this bound small we have to take $r > c|\log V(z)|$ for some $c > e$. We choose c to be the solution of

$$(\frac{c}{e})^c = e. \tag{5.3}$$

For any $b \geq c$ it is easy to see that $(b/e)^b \geq e^{b-c+1}$. For $r \geq c|\log V(z)|$ and put $b = r/|\log V(z)|$. This gives $G_r \leq e^{-r-1}V(z)^{1-c}$.

Let $s \geq 1 - c\log V(z) = 1 + c|\log V(z)|$ and put $r = [s]$ and so $G_r \leq e^{-s}V(z)^{1-c}$.

Since $d|P(z)$ and $\omega(d) \leq r$, implies $d \leq z^r \leq z^s$, we have

$$R_r = \sum_{\substack{d|P(z)\\ \omega(d)\leq r}} |r_d| = R(\mathcal{A}, z^s)$$

where we introduce the genaral notation

$$R(\mathcal{A}, D) = \sum_{\substack{d|P(z)\\ d<D}} |r_d|.$$

Recalling

$$S(\mathcal{A}, z) = XV(z) + 2\theta XG_r + \theta R_r,$$

we get

Theorem. Let $\mathcal{A}$ be a sequence of non-negative reals and $\mathcal{P}$ be a finite set of primes. Let $z \geq 2$ and $D = z^s$ with $s \geq 1 + c|\log V(z)|$. Then

$$S(\mathcal{A}, z) = XV(z)\left(1 + 2\theta e^{-s}V(z)^{-c}\right) + \theta R(\mathcal{A}, D)$$

where $|\theta| \leq 1$ and c is as defined in (5.3).

A theorem of this form is referred as Fundamental lemma. Special importance of these resullts comes from the fact that they give assymptotic formulas, in contrast with other sieving theorems where we get upper or lower bounds.

CHAPTER 2

Selberg's Sieve

1. Introduction

A sequence of real numbers $(\lambda_d)_{d<D}$ is an upper bound sieve if $\lambda_1 = 1$ and $\sum_{d|n} \lambda_d \geq 0, \forall n > 1$. If we take an arbitrary sequence $(\rho_d)_{d<D}$ of reals such that $\rho_1 = 1$ and choose λ_d in such a way that

$$\sum_{d|n} \lambda_d = \left(\sum_{d|n} \rho_d \right)^2$$

holds, then $(\lambda_d)_d$ is obviously an upper bound sieve. Thus

$$S(\mathcal{A}, z) = \sum_{d|P(z)} \mu(d)|\mathcal{A}_d| \leq \sum_{d|P(z)} \lambda(d)|\mathcal{A}_d| = S^+(\mathcal{A}, z).$$

Expanding squares

$$S^+(\mathcal{A}, z) = \sum_{d_1, d_2|P(z)} \rho_{d_1}\rho_{d_2}|\mathcal{A}_{[d_1,d_2]}|.$$

Assuming $|\mathcal{A}_d| = g(d)X + r(d)$, we get

$$S^+(\mathcal{A}, z) = XG + R^+(\mathcal{A})$$

where

$$G = \sum_{d_1, d_2|P(z)} \rho_{d_1}\rho_{d_2} g([d_1, d_2])$$

and

$$R^+(\mathcal{A}) = \sum_{d_1, d_2|P(z)} \rho_{d_1}\rho_{d_2} r([d_1, d_2]).$$

Assuming $|\rho_d| \leq 1$ for all d, we have

$$R^+(\mathcal{A}) = \sum_{d_1, d_2|P(z)} |r([d_1, d_2])|.$$

We can think of ρ_d as variables satisfying $\rho_1 = 1$ and G as a quadratic form in these variable. To find a good upper bound sieve our aim would be to minimise the quadratic form G.

We change variable as $(d_1, d_2) = c$, $d_1 = ac$ and $d_2 = bc$, so $[d_1, d_2] = abc$ with $(a, b) = (b, c) = (a, c) = 1$. We also suppose that g is multiplicative with $0 < g(p) < 1$ for all $p < z$ and $g(p) = 0$ otherwise. Now our aim is to diagonalise G as a quadratic form in ρ_{d_1} and ρ_{d_2}.

$$\begin{aligned} G &= \sum_{abc|P(z)} g(abc)\rho_{ac}\rho_{bc} \\ &= \sum_{c} g(c) \sum_{\substack{a,b \\ (a,b)=1}} g(ab)\rho_{ac}\rho_{bc} \\ &= \sum_{c} g(c) \sum_{a,b} g(a)g(b)\rho_{ac}\rho_{bc} \sum_{d|(a,b)} \mu(d) \\ &= \sum_{c} g(c) \sum_{d} \mu(d) \sum_{d|a,d|b} g(a)g(b)\rho_{ac}\rho_{bc} \end{aligned}$$

Changing the variable to $a = dm$ and $b = dl$, the above sum equals to

$$\begin{aligned} &= \sum_{c} g(c) \sum_{d} \mu(d)g(d)^2 \sum_{m,l} g(m)g(l)\rho_{cdm}\rho_{cdl} \\ &= \sum_{c} g(c) \sum_{d} \mu(d)g(d)^2 \left(\sum_{m} g(m)\rho_{cdm} \right)^2 \\ &= \sum_{d} \mu(d) \sum_{c} \frac{1}{g(c)} \left(\sum_{m} g(cdm)\rho_{cdm} \right)^2 \\ &= \sum_{d} \mu(d) \sum_{c} \frac{1}{g(c)} \left(\sum_{cd|m} g(m)\rho_{m} \right)^2 \\ &= \sum_{c,d} \frac{\mu(d)g(d)}{g(cd)} \left(\sum_{cd|m} g(m)\rho_{m} \right)^2. \end{aligned}$$

Changing variable cd to d, we get the above is

$$\sum_{d} \frac{1}{g(d)} \left(\sum_{l|d} \mu(d)g(d) \right) \left(\sum_{cd|m} g(m)\rho_{m} \right)^2.$$

Recall that g is multiplicative, so

$$\sum_{d} \frac{1}{g(d)} \left(\sum_{l|d} \mu(d)g(d) \right) = \prod_{p|d} \frac{1 - g(p)}{g(p)}.$$

Hence we define a multiplicative function h by

$$h(p) = \frac{g(p)}{1-g(p)}, \quad \text{so } g(p) = \frac{h(p)}{1+h(p)}.$$

Therefore we get

$$G = \sum_{d|P(z)} \frac{1}{h(d)} \left(\sum_{cd|m} g(m)\rho_m \right)^2.$$

Put

$$y_d = \frac{\mu(d)}{h(d)} \sum_{cd|m} g(m)\rho_m.$$

Thus $G = \sum_{d|P(z)} h(d)y_d^2$. Now we need the following lemma which is, in a sense dual to Mobius inversion.

Lemma. For two multiplicative function f and g,

$$f(n) = \sum_{n|m} g(m) \text{ if and only if } g(n) = \sum_{n|m} \mu(m/n)f(m).$$

Applying this lemma on

$$y_d = \frac{\mu(d)}{h(d)} \sum_{cd|m} g(m)\rho_m.$$

we get

$$\rho_l = \frac{\mu(l)}{g(l)} \sum_{\substack{d|P(z) \\ l|d}} h(d)y_d.$$

Thus $\rho_1 = 1$ implies $\sum_{d|P(z)} h(d)y_d = 1$. We also restrict the support of ρ_d upto $d \leq \sqrt{D}$ i.e $\rho_d = 0$ for $d > \sqrt{D}$. Therefore our task is to Minimise the quadratic form

$$G = \sum_{d|P(z)} h(d){y_d}^2$$

under the condition $\sum_{d|P(z)} h(d)y_d = 1$. Cauchy -Schwarz inequality gives

$$1 = \left(\sum_{d|P(z)} h(d)y_d \right)^2 \leq \left(\sum_{d|P(z)} h(d){y_d}^2 \right) \left(\sum_{d|P(z)} h(d) \right) = GJ$$

where $J = \sum_{d|P(z)} h(d)$. Thus $G \geq 1/J$ and G attains that value for $y_d = 1/J$.

Remark 1.1. If we take $D > P(z)$, then $J = \sum_{d|P(z)} h(d) = \prod_{p<z}(1+h(p))$. Clearly $1 + h(p) = (1 - g(p))^{-1}$. Thus $J = \prod_{p<z}(1-g(p))^{-1} = V(z)^{-1}$. So for the optimal choice of weights $G = V(z)$, hence the main term in $S^+(\mathcal{A}, z)$ is $XV(z)$ which is the expected main term. But if we choose D to be so large then error term also becomes too large, so to control the error term we have to choose an optimal D.

Using $y_d = 1/J$ we get

$$\rho_l = \frac{\mu(l)}{Jg(l)} \sum_{l|d} h(d) = \frac{\mu(l)}{Jg(l)} h(l) \sum_{(d_1,l)=1} h(d_1).$$

Let

$$J_l = \sum_{(d,l)=1} h(d).$$

Hence $J_1 = J$ and

$$\rho_l = \frac{J_l}{J}\mu(l)j(l) \quad \text{where} \quad j(l) = \frac{h(l)}{g(l)} = \prod_{p|l}(1-g(p))^{-1}.$$

To show that $|P_l| \leq 1$ we have to show that $j(l)J_l \leq J$. We group the sum in J according to the gcd with l to get

$$J = \sum_{k|l} \sum_{\substack{d|P(z), d\leq\sqrt{D} \\ (d,l)=k}} h(d).$$

Writing $d = km$ where $k = (d, l)$ and $(m, l) = 1$ the right hand side is

$$\sum_{k|l} h(k) \sum_{\substack{m|P(z), m<\sqrt{D}/k \\ (m,l)=1}} h(m) \geq \sum_{k|l} h(k) \sum_{\substack{m<\sqrt{D}/l \\ (m,l)=1}} h(m) = J_l \sum_{k|l} h(k).$$

The above inequality holds as h is a Now by multiplicativity

$$\sum_{k|l} h(k) = \prod_{p|l}(1+h(p)) = \prod_{p|l} \frac{h(p)}{g(p)} = \frac{h(l)}{g(l)} = j(l).$$

Therefore $J \geq j(l)J_l$ for all l, this proves $|\rho_l| \leq 1$. Gathering our arguments so far we get

Theorem. Let $\mathcal{A} = (a_n)_n$ be a finite sequence of non-negative numbers and $\mathcal{A}_d, P(z), g, h$ and J as defined above we have

$$S(\mathcal{A}, z) \leq \frac{X}{J} + R^+(\mathcal{A})$$

where

$$R^+(\mathcal{A}) = \sum_{d|P(z)} \lambda_d r_d(\mathcal{A}) \text{ with } \lambda_d = \sum_{[d_1,d_2]=d} \rho_{d_1}\rho_{d_2}$$

and ρ_d is as defined above.

In order the apply this theorem we would require good lower bound for J which we pursue now. We recall that

$$J = J(D) = \sum_{\substack{d|P(z)\\ d<\sqrt{D}}} h(d).$$

Clearly $J(P(z)) = \prod_{p<z}(1-g(p))^{-1} = V(z)^{-1}$. Let $I = I(D) = V(z)^{-1} - J(D)$. Then

$$I(D) = \sum_{\substack{d|P(z)\\ d\geq\sqrt{D}}} h(d) \leq \sum_{d|P(z)} h(d)\left(\frac{d}{\sqrt{D}}\right)^{2\epsilon}$$

for any $\epsilon > 0$ since h is positive. By multiplicativity of h, we get

$$I(D) \leq D^{-\epsilon}\prod_{p<z}(1+h(p)p^{2\epsilon}).$$

Further

$$\begin{aligned} V(z)I(D) &\leq D^{-\epsilon}\prod_{p<z}(1-g(p))(1+h(p)p^{2\epsilon}) \\ &= D^{-\epsilon}\prod_{p<z}(1+g(p)(p^{2\epsilon}-1)) \\ &= D^{-\epsilon}e\left(\sum_{p<z}\log(1+g(p)(p^{2\epsilon}-1))\right) \\ &= D^{-\epsilon}e\left(\sum_{p<z}g(p)(p^{2\epsilon}-1)\right) \end{aligned}$$

where the last inequality uses $\log(1+x) \leq x$ for any $x > 0$ and $e(x) = e^x$. Now we make the following assumption. Suppose $z = D^{1/s}$ with $s \geq 1$ and $g(p)$ satisfies

$$\prod_{w\leq p<z}(1-g(p))^{-1} \leq C\left(\frac{\log z}{\log w}\right)^k. \tag{1.1}$$

for all w such that $z < w \geq 2$ and $C > 1$, $k \geq 1$ are constants. We note that the left hand side is equal to $V(w)/V(z)$.

To proceed further we need the following form of partial summation.

Lemma. Let f and g be arithmetic functions and g is smooth in (y, x). Then

$$\sum_{y\le n<x} f(n)g(n) = g(y)\sum_{y\le n<x} f(n) + \int_y^x g'(w)\left(\sum_{w\le n<x} f(n)\right)dw.$$

Now we use this lemma to estimate a sum over primes that we would require.

Lemma. Let h be a continuous non-negative, non-decreasing function on $[y, z]$ and g be a multiplicative function with $0 \le g(p) < 1$ such that (1.1) holds for all $w \in [y, z]$. Then

$$\sum_{y\le p<z} g(p)h(p) \le k\int_y^z \frac{h(w)}{w\log w}dw + h(z)\log C.$$

Proof. From (1.1)

$$\sum_{w\le p<z} g(p) \le \sum_{w\le p<z} -\log(1-g(p)) \le \log C + k\log\left(\frac{\log z}{\log w}\right)$$

where we use $x \le -\log(1-x)$ for all $x > 0$. Using Lemma 1, we get

$$\begin{aligned}\sum_{y\le p<z} g(p)h(p) &= h(y)\sum_{y\le p<z} g(p) + \int_y^z h'(w)\left(\sum_{w\le p<z} g(p)\right)dw\\ &\le h(y)\left(k\log\left(\frac{\log z}{\log y}\right) + \log C\right) + \int_y^z \left(k\log\left(\frac{\log z}{\log w}\right) + \log C\right)h'(w)dw\\ &= h(z)\log C + kh(y)\log\left(\frac{\log z}{\log y}\right) + k\int_y^z \log\left(\frac{\log z}{\log w}\right)h'(w)dw.\end{aligned}$$

The lemma follows by using partial integration on the last integral.

Recall that

$$V(z)I(D) \le D^{-\epsilon}e\left(\sum_{p<z} g(p)(p^{2\epsilon}-1)\right).$$

Assuming (1.1), Lemma 3 implies

$$\sum_{p<z} g(p)(p^{2\epsilon}-1) \le k\int_1^z \frac{w^{2\epsilon}-1}{w\log w}dw + (z^{2\epsilon}-1)\log C.$$

Changing variables to $v = 2\epsilon\log w$ and putting $\alpha = 2\epsilon\log z$ we get

$$V(z)I(D) \le C^{e(\alpha)-1}e\left(kB(\alpha) - \frac{\alpha s}{2}\right)$$

where

$$B(\alpha) = \int_0^\alpha \frac{e(v)-1}{v} dv.$$

It is easy to see that $B(\alpha) \le e(\alpha) - 1$. Thus

$$C^{e(\alpha)-1} e\left(kB(\alpha) - \frac{\alpha s}{2}\right) \le e\left((e(\alpha)-1)l - \frac{\alpha s}{2}\right)$$

where $l = k + \log C$. Now let

$$Q(\alpha) = e\left((e(\alpha)-1)l - \frac{\alpha s}{2}\right).$$

The minimal value of $Q(\alpha)$ is attained at $\alpha = \log(s/2l)$ if $s > 2l$. Thus

$$Q(\alpha) \le e(-l)\left(\frac{2el}{s}\right)^{s/2}.$$

Therefore

$$V(z)I(D) \le e(-l)\left(\frac{2el}{s}\right)^{s/2}.$$

Since $J(D) + I(D) = V(z)^{-1}$ we have

$$1 \ge V(z)J(D) \ge 1 - e^{-l}\left(\frac{2el}{s}\right)^{s/2}$$

when $s > 2l$. Thus we have proved the following

Theorem. We assume all the conditions of Theorem 1 and also (1.1), then with $z = D^{1/s}$ for $s > k + \log C = l$ we get

$$S(\mathcal{A}, z) \le XV(z)\left(1 - e^{-l}\left(\frac{2el}{s}\right)^{s/2}\right)^{-1} + \sum_{\substack{d|P(z)\\ d<D}} \tau_3(d)|r_d(\mathcal{A})|.$$

2. More precise estimates

We can obtain better lower bound for $J(D)$ and better upper bound for the error term if we assume some conditions on $g(p)$ and $r_d(\mathcal{A})$. The conditions also needs to available in applications so that results obtained are useful.

Let q be a fixed positive square-free integer. We assume that $g(p) \ge k/p$ for all primes $p \nmid q$ for a fixed k. Recall that

$$J(D) = \sum_{\substack{d|P(z)\\ d<\sqrt{D}}} h(d).$$

If $z \geq \sqrt{D}$ then

$$J(D) = \sum_{d<\sqrt{D}}^{\flat} h(d)$$

where the sum is over square free d. Put $d = ab$ where $a|q$ and $(b, q) = 1$. Thus

$$J(D) = \sum_{\substack{a<\sqrt{D} \\ a|q}}^{\flat} h(a) \sum_{\substack{b<\sqrt{D}/a \\ (b,q)=1}}^{\flat} h(b).$$

Therefore we want a lower bound for

$$F(x) = \sum_{\substack{b<x \\ (b,q)=1}}^{\flat} h(b).$$

We extend g to all integers by defining

$$\sum_{b=1}^{\infty} \frac{g(b)}{b^s} = \prod_{p<z} \left(1 - \frac{g(p)}{p^s}\right)^{-1}.$$

We note that this makes the extended g completely multiplicative function supported on the integers composed of primes $p < z$. Thus

$$\begin{aligned} h(b) &= \prod_{p|b} \frac{g(p)}{1-g(p)} = \prod_{p|b} g(p)(1 + g(p) + g(p)^2 + g(p)^3 + ...) \\ &= \prod_{p|b} (g(p) + g(p^2) + g(p^3) + ...) \end{aligned}$$

Therefore

$$\sum_{b}^{\flat} h(b) \geq \sum_{b} g(b)$$

where the first sum is on square-free integers. Since $g(p) \geq k/p$, we get $g(b) \geq k^{\Omega(b)}/b \geq \tau_k(b)/b$ where $\Omega(b)$ is the number of prime factors of b counted with multiplicity and τ_k is the k-th divisor function.

$$F(x) \geq \sum_{\substack{b<x \\ (b,q)=1}} \frac{\tau_k(b)}{b} \geq \left(\frac{\varphi(q)}{q}\right)^k \sum_{b<x} \frac{\tau_k(b)}{b}.$$

For $x \geq 1$ we have

$$\sum_{b<x} \frac{\tau_k(b)}{b} = \sum_{d_1 d_2 ... d_k < x} \frac{1}{d_1 d_2 ... d_k} \geq \int \int ... \int_{\substack{x_1 x_2 ... x_k < x \\ x_1, ..., x_k \geq 1}} \frac{dx_1 .... dx_k}{x_1 ... x_k} = \frac{1}{k!} (\log x)^k.$$

So we get

$$F(x) \geq \frac{1}{k!}\left(\frac{\varphi(q)}{q}\log x\right)^k.$$

This gives

$$\begin{aligned} J(D) &\geq \frac{1}{k!}\sum_{\substack{a<\sqrt{D}\\ a|q}} h(a)\left(\frac{\varphi(q)}{q}\log\frac{\sqrt{D}}{a}\right) \\ &\geq \frac{1}{k!}\left(\frac{\varphi(q)}{q}\log\sqrt{D}\right)^k \sum_{\substack{a<\sqrt{D}\\ a|q}} h(a)\left(1-\frac{k\log a}{\log\sqrt{D}}\right) \\ &\geq \frac{1}{k!}\left(\frac{\varphi(q)}{q}\log\sqrt{D}\right)^k \sum_{a|q} h(a)\left(1-\frac{k\log a}{\log\sqrt{D}}\right) \end{aligned}$$

where in the first inequality we use $(1-y)^k \geq 1-ky$ for $y = \log a/\log\sqrt{D} < 1$ and in the second inequality we observe that the extra terms for $a \geq \sqrt{D}$ which we added are all negative. We note that $\sum_{a|q} h(a) = \prod_{p|q}(1-g(p))^{-1} = j(q)$ which we have already used earlier. also

$$\begin{aligned} \sum_{a|q} h(a)\log a &= \sum_{a|q} h(a)\sum_{p|a}\log p \\ &= \sum_{p|q} h(p)\log p\prod_{l|q/p}(1+h(l)) \\ &= \sum_{p|q} h(p)\log q(1+h(p))^{-1}\prod_{l|q}(1+h(l)) \\ &= \prod_{l|q}(1-g(l))^{-1}\sum_{p|q}\log p\frac{h(p)}{1+h(p)} \\ &= j(q)L(q) \end{aligned}$$

where we define $L(q) = \sum_{p|q} g(p)\log p$. Now we have a lower bound of $J(D)$ as follows

$$\begin{aligned} J(D) &\geq \frac{1}{k!}\left(\frac{\varphi(q)}{q}\log\sqrt{D}\right)^k j(q)\left(1-\frac{kL(q)}{\log\sqrt{D}}\right) \\ &= \frac{(\log\sqrt{D}^k}{K!H_q}\left(1-\frac{kL(q)}{\log\sqrt{D}}\right) \end{aligned}$$

where

$$H_q = \prod_{p|q}(1-g(p))\left(1-\frac{1}{p}\right)^{-k}.$$

3. Estimating the error term

We suppose that

$$|r_d(\mathcal{A})| \le g(d)d \text{ and } g(d)d \ge 1 \text{ if } d|P(z).$$

We observe that

$$g([d_1,d_2])[d_1,d_2] \le g(d_1)g(d_2)d_1d_2.$$

The error term is

$$R^+(\mathcal{A}) = \sum \sum_{d_1,d_2|P(z)} \rho_{d_1}\rho_{d_2} r_{[d_1,d_2]}(\mathcal{A}).$$

Also recall that

$$\rho_d = \frac{\mu(d)}{Jg(d)} \sum_{\substack{n<\sqrt{D}\\ d|n}} h(n).$$

Hence

$$\sum_{d<\sqrt{D}} |\rho_d| g(d)d \le \frac{1}{J} \sum_{n<\sqrt{D}} h(n)\sigma(n)$$

where $\sigma(n)$ is the sum of divisors of n. Therefore

$$\begin{aligned}
R^+(\mathcal{A}) &\le \sum_{\substack{d_1,d_2|P(z)\\ d_1,d_2<\sqrt{D}}} |\rho_{d_1}||\rho_{d_2}| g([d_1,d_2])[d_1,d_2] \\
&\le \sum_{\substack{d_1,d_2|P(z)\\ d_1,d_2<\sqrt{D}}} |\rho_{d_1}||\rho_{d_2}| g(d_1)g(d_2)d_1d_2 \\
&\le \left(\sum_{\substack{d|P(z)\\ d<\sqrt{D}}} |\rho_d| g(d)d \right)^2 \\
&\le \frac{1}{J^2} \left(\sum_{\substack{n|P(z)\\ n<\sqrt{D}}} h(n)\sigma(n) \right)^2.
\end{aligned}$$

Now we assume

$$\sum_{y\le p<x} g(p)\log p \ll \log(2x/y) \text{ for all } x,y \text{ satisfying } 2\le y<x. \tag{3.1}$$

and

$$\sum_{p} g(p)^2\log p<\infty. \tag{3.2}$$

We need the following lemma before proceeding further.

Lemma. Let f and g are multiplicative functions such that $g(n)=f(n)/n$. Let $M_f(x)=\sum_{n\le x} f(n)$. If g satisfies the condition (3.1) above then

$$M_f(x)\ll \frac{x}{\log x}M_g(x).$$

Proof. It follows from condition (3.1) that

$$\sum_{p\le x} g(p)\ll \log\log x \quad\text{and}\quad g(p)\ll 1/\log p.$$

Using

$$\sum_{n\le x} f(n)\le x\prod_{p\le x}\left(1+\frac{f(p)}{p}\right)$$

we get $\sum_{n\le x} f(n)\ll x\log x$ and this implies by partial summation

$$\sum_{p\le x} f(p)\log p\ll x.$$

Hence

$$\sum_{m\le x} f(m)\log m=\sum_{m\le x} f(m)\sum_{p|m}\log p\le x\sum_{n\le x}\frac{f(n)}{n}.$$

Then by partial summation, we get

$$\sum_{n\le x} f(n)\ll \frac{x}{\log x}\sum_{n\le x}\frac{f(n)}{n}$$

which completes the proof.

It follows from condition (3.1) that $g(p)\ll 1/\log p$ hence it is easy to check that

$$\sum_{y\le p<x} h(p)\sigma(p)\frac{1}{p}\log p\ll \log(2x/y).$$

So we can use Lemma 4 with $f(m) = h(m)\sigma(m)$ to get

$$\sum_{m<\sqrt{D}} h(m)\sigma(m) \ll \frac{\sqrt{D}}{\log D} \sum_{m<\sqrt{D}} h(m)\sigma(m)\frac{1}{m}.$$

Now

$$\begin{aligned}
\sum_{m<\sqrt{D}} h(m)\sigma(m)\frac{1}{m} &= \sum_{m<\sqrt{D}} \frac{h(m)}{m} \sum_{d|m} d \\
&= \sum_{d<\sqrt{D}} d \sum_{\substack{m<\sqrt{D} \\ d|m}} \frac{h(m)}{m} \\
&= \sum_{d<\sqrt{D}} h(d) \sum_{n<\sqrt{D}/d} \frac{h(n)}{n} \\
&\leq \sum_{d<\sqrt{D}} h(d) \sum_{n<\sqrt{D}} \frac{h(n)}{n} \\
&= J \sum_{n<\sqrt{D}} \frac{h(n)}{n} \ll J
\end{aligned}$$

where we use $\sum_n h(n)/n < \infty$ which is a consequence of assumption (3.2). Thus we conclude that

$$R^+(\mathcal{A}) \ll \frac{D}{\log^2 D}$$

and he have

Theorem. Let $\mathcal{A} = (a_n)_n$ be a finite sequence of non-negative numbers and P be a finite product of disticnt primes. For every $d|P$ we write

$$|\mathcal{A}_d| := \sum_{d|n} a_n = g(d)X + r_d(\mathcal{A})$$

where g is a multiplicative function with $0 < g(p) < 1$ for all $p|P$ and h is another multiplicative function defined by

$$h(p) = \frac{g(p)}{1 - g(p)}.$$

Suppose following conditions are satisfied

$$|r_d(\mathcal{A})| \leq g(d)d \tag{3.3}$$

$$g(d)d \geq 1 \tag{3.4}$$

$$\sum_{y\leq p<x} g(p)\log p \ll \log(2x/y) \tag{3.5}$$

Then

$$S(\mathcal{A}, P) \leq \frac{X}{J} + O\left(\frac{D}{\log D}\right)$$

where

$$J = J(D) = \sum_{\substack{d|P \\ d<\sqrt{D}}} h(d).$$

4. Application

Let $a \leq q$ be integers such that $(a, q) = 1$ and a_n denote the characteristic function of the arithmetic progression $n \equiv a(q)$ satisfying $x < n \leq x + y$ where $q < y$. Let $\mathcal{P}$ be the set of primes $p \leq \sqrt{y}$ such that $p|q$. we note that

$$\pi(x + y, q; a) - \pi(x, q; a) \leq S(\mathcal{A}, \sqrt{y}) + \frac{\sqrt{y}}{q}$$

and

$$|\mathcal{A}_d| = \{x < n \leq x + y : n \equiv a(q), d|n\} = \frac{y}{dq} + O(1).$$

Hence $g(p) = 1/p$ for $p \nmid q$, $X = y/q$ and $|r_d(\mathcal{A})| \ll 1$. Thus we get

$$|r_d(\mathcal{A})| \ll dg(d), dg(d) \geq 1$$

also

$$\sum_{y<p\leq x+y} g(p) \log p = \sum_{y<p\leq x+y} \frac{\log p}{p} \ll \log(x/y)$$

and

$$\sum_{p} g(p)^2 \log p = \sum_{p} \frac{\log p}{p^2} < \infty.$$

We apply the lower bound for $J(D)$ obtained in Section 4.1 with $k = 1$ and the modulus of the arithmetic progression as q. Since $g(p) = 0$ for $p|q$, we get $L(q) = 0$ and

$$H_q = \prod_{p|q} (1 - g(p)) \left(1 - \frac{1}{p}\right)^{-1} = \frac{q}{\varphi(q)}.$$

Hence

$$J(D) \geq \frac{\log \sqrt{D}}{H_q} = \frac{\varphi(q) \log D}{2q}.$$

So we conclude

$$\pi(x + y, q; a) - \pi(x, q; a) \leq \frac{2y}{\varphi(q) \log D} + O\left(\frac{D}{\log^2 D} + \frac{\sqrt{y}}{q}\right).$$

We choose $D = y/q$ so that $z = \sqrt{y} \geq \sqrt{D} = \sqrt{y/q}$ and we get

$$\pi(x+y,q;a) - \pi(x,q;a) \leq \frac{2y}{\varphi(q)\log D} + O\left(\frac{y}{q(\log(y/q))^2}\right).$$

This result is known as Brun-Titchmarsh theorem.

CHAPTER 3

Large sieve inequality

1. Motivation

Let f is a complex valued function on the group $\mathbb{Z}/N\mathbb{Z}$. The Fourier transform of f is defined by

$$\hat{f}(\alpha) = \frac{1}{N}\sum_n f(n)e(n\alpha) \quad \text{where} \quad e(n\alpha) = e^{\frac{2\pi i n\alpha}{N}}.$$

Then the Parseval equality is

$$\sum_\alpha |\hat{f}(\alpha)|^2 = \frac{1}{N}\sum_n |f(n)|^2$$

which in can be written as

$$\sum_\alpha |\sum_n f(n)e(n\alpha)|^2 = N\sum_n |f(n)|^2.$$

The crucial fact here is orthogonality of $e(n\alpha)$ with α running over $\mathbb{R}/\mathbb{Z}$. In some important situations we do not have such orthogonality but still we need a result in the line of Parseval's equality. But it is unreasonable to expect a good result for any sequence of α's. More precisely the question is "if we have a sequence $(\alpha_r)_r$ such that $(e(n\alpha_r)_r$ is almost orthogonal, then can we prove a good inequality?" (equality being too much to expect). This question does not make much mathematical sense yet. Let us first try to guess what is reasonable to expect.

Let $\alpha_r : r = 1, 2, ..., R$ be positive real numbers and $f(n)$ be a sequence of complex numbers. We wish to prove an inequality of the form

$$\sum_{r=1}^{R}\left|\sum_{n=1}^{N} f(n)e(n\alpha_r)\right|^2 \leq C(N,R)\sum_{n=1}^{N}|f(n)|^2. \tag{1.1}$$

Before we try to guess what $C(N,R)$ to expect we need the following general lemma.

Lemma. (Duality principle) Let (c_{nr}) be a fixed $N \times R$ complex matrix and D be a positive real number. Then then following conditions are

equivalent.

$$\left|\sum_{n=1}^{N}\sum_{r=1}^{R} c_{nr}x_n y_r\right| \leq D\left(\sum_{n=1}^{N}|x_n|^2\right)^{1/2}\left(\sum_{r=1}^{R}|y_r|^2\right)^{1/2}, \tag{1.2}$$
$$\forall (x_1,...,x_N)\in\mathbb{C}^N, (y_1,...,y_R)\in\mathbb{C}^R.$$

$$\left(\sum_{r=1}^{R}\left|\sum_{n=1}^{N} c_{nr}x_n\right|^2\right)^{1/2} \leq D\left(\sum_{n=1}^{N}|x_n|^2\right)^{1/2}, \forall (x_1,...,x_N)\in\mathbb{C}^N \tag{1.3}$$

$$\left(\sum_{n=1}^{N}\left|\sum_{r=1}^{R} c_{nr}y_r\right|^2\right)^{1/2} \leq D\left(\sum_{n=1}^{N}|x_n|^2\right)^{1/2}, \forall (y_1,...,y_R)\in\mathbb{C}^R. \tag{1.4}$$

Now going back to inequality (1.1), expanding the LHS we get

$$\sum_{n_1,n_2} f(n_1)\overline{f(n_2)}\sum_{r} e(\alpha_r(n_1-n_2)).$$

So the contribution from the diagonal term $n_1 = n_2$ is $R\sum_n |f(n)|^2$. Hence we should expect $C(N,R)\geq R$. Using duality principle we get

$$\sum_{n=1}^{N}\left|\sum_{r=1}^{R} g(n)e(n\alpha_r)\right|^2 \leq C(N,R)\sum_{r=1}^{R}|g(r)|^2$$

for any sequence $g(r)$. Expanding the LHS we see that the diagonal contribution in this case is $N\sum_{r=1}^{R}|g(r)|^2$. Thus we should also expect $C(n,R)\geq R$. Now we need a concept of almost orthogonality. Examining Parseval's equality we see that for a fixed n the points $n\alpha/N$ with $\alpha\in\mathbb{R}/\mathbb{Z}$ have a spacing of $1/N$ between them. This property truns out to be the right concept of almost orthogonality. We define it now.

Definition 1.1. A sequence $\alpha_r : r = 1,...,R$ of real numbers in $(0,1]$ are said to be δ-well spaced for $0<\delta<1$ if $|\alpha_r-\alpha_s|>\delta$ whenever $r\neq s$.

Obviously the number of δ-well spaced points would be $R=[1/\delta]\geq 1/\delta-1$. Thus we should expect that

$$C(N,R)\geq N+\frac{1}{\delta}-1.$$

This is what is known as Large sieve inequality.

Theorem. (Large sieve inequality) For any δ-well spaced points $(\alpha_r)_r$ in $(0,1]$ and for any sequence $a_n : M<n<M+N$ of complex numbers,

we have

$$\sum_r \left| \sum_n a_n e(\alpha_r n) \right|^2 \leq \left(N + \frac{1}{\delta} - 1 \right) \sum_n |a_n|^2.$$

On our to prove this theorem we would first prove some easier but weaker results. We begin by slightly reformulating the problem which would give us a very good advantage.

Let $(b_n)_n$ be a sequence of real numbers such that

(1) $b_n \geq 0$ for all n.
(2) $b_n \geq 1$ for $M+1 \leq n \leq M+N$.

Thus the folowing inequality

$$\sum_r \left| \sum_{n=M+1}^{M+N} a_n e(n\alpha_r) \right|^2 \leq B(N,\delta) \sum_{n=M+1}^{M+N} |a_n|^2 \frac{1}{b_n}$$

if established for arbitrary complex numbers a_n and δ-well spaced sequence $(\alpha_r)_r$ would imply Large sieve inequality with $C(N,R) = B(N,\delta)$. Writing $a_n = c_n b_n^{1/2}$ this becomes

$$\sum_r \left| \sum_{M+1}^{M+N} c_n b_n^{1/2} e(n\alpha_r) \right|^2 \leq B(N,\delta) \sum_{n=M+1}^{M+N} |c_n|^2.$$

Now by Duality principle, using $c_{nr} = b_n^{1/2} e(n\alpha_r)$, it is enough to prove that

$$\sum_{M+1}^{M+N} \left| \sum_r y_r b_n^{1/2} e(n\alpha_r) \right|^2 \leq B \sum_r |y_r|^2$$

for all complex sequence y_r. Expanding the square on the right hand side and interchanging summation we get

$$\sum_{r,s} y_r \overline{y_s} \sum_{M+1}^{M+N} b_n e(n(\alpha_r - \alpha_s)) \leq B \sum_r |y_r|^2.$$

Defining

$$B(\alpha) = \sum_{M+1}^{M+N} b_n e(n\alpha)$$

Our aim is to prove an inequality of the form

$$\sum_{r,s} y_r \overline{y_s} B(\alpha_r - \alpha_s) \leq B \sum_r |y_r|^2.$$

The following lemma provides us with a form of B.

Lemma. Let $(c_{rs})_{rs}$ be a $R \times S$ Hermitian matrix with complex entries such that There exists $B > 0$ and positive real numbers $k_1, ..., k_R$ (suppose $S < R$) satisfying

$$\sum_s k_s |c_{rs}| \leq B k_r, \forall 1 \leq r \leq R.$$

Then for arbitrary complex numbers $y_r, ..., y_R$, we have

$$\left| \sum_{r,s} c_{rs} y_r \overline{y_s} \right| \leq B \sum_r |y_r|^2.$$

Moreover we can take

$$B = \max_r \sum_s |c_{rs}|.$$

PROOF. Using inequality $|ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$ with $a = y_r k_r^{-1}, b = \overline{y_r} k_s^{-1}$, we get

$$\left| \sum_{r,s} c_{rs} y_r \overline{y_s} \right| \leq \sum_{r,s} |c_{rs}| k_r k_s \left(\frac{1}{2} |y_r/k_r|^2 + \frac{1}{2} |y_s/k_s|^2 \right).$$

The right hand side becomes

$$= \sum_r |y_r|^2 k_r^{-1} \sum_s k_s |c_{rs}|$$

which is bouned above by $B \sum_r |y_r|^2$ by the assumed condition.

Thus we need to find upper bound for

$$\max_r \sum_s |B(\alpha_r - \alpha_s)|$$

with b_n satisfying the properties listed above. Clearly any choice of b_n's satisfying the properties will give us a large sieve type inequality with B as given by the last lemma.

2. Choosing b_n's

The simplest choice for b_n is just the value of the characteristic function of interval $[M+1, M+N]$ which is

$$b_n = \begin{cases} 1 & \text{if } M+1 \leq n \leq M+N, \\ 0 & \text{otherwise} \end{cases}$$

It is easy to see that $B(0) = N$ and for $\alpha \neq 0$

$$|B(\alpha)| = \left| \frac{\sin(N\pi\alpha)}{\sin(\pi\alpha)} \right|.$$

Using the fact that $|\sin(\pi\alpha)| \geq 2\|\alpha\|$ where $\|\alpha\|$ denotes the diatance of α from its nearest integer, we have

$$B \leq N + \max_r \sum_{s \neq r} \frac{1}{\|\alpha_r - \alpha_s\|} \leq N + \sum_{k=1}^{[1/\delta]} \frac{1}{k\alpha} \leq N + \frac{1}{\delta}\log(\frac{1}{\delta}).$$

Next we try to get a better result by choosing b_n a continuous function of n. Let $A > 0$ be a fixed positive number and let b_n be defined as follows

$$b_n = \begin{cases} 1 & \text{if } n \leq M+1-A, \\ 1 - \left(\frac{M+1-n}{A}\right) & \text{if } M+1-A < n \leq M+1, \\ 1 & \text{if } M+1 < n \leq M+N, \\ 1 - \left(\frac{n-M-N}{A}\right) & \text{if } M+N < n \leq M+N+A, \\ 0 & \text{if } \ M+N+A < n \end{cases}$$

It is east to prove that for $\alpha \neq 0$

$$|B(\alpha)| = \left|\frac{\sin(\pi A\alpha)\sin(\pi(N+A)\alpha)}{A(\sin(\pi\alpha))^2}\right| \leq \frac{1}{4A\|\alpha\|^2}$$

also $B(0) = N + A$. As before we get

$$\max_r \sum_{s \neq r} \frac{1}{\|\alpha_r - \alpha_s\|} \leq \frac{1}{2A}\sum_m \frac{1}{m^2\delta^2} \leq \frac{1}{2A\delta^2}\frac{\pi^2}{6} = \frac{\pi^2}{12A\delta^2}.$$

Thus $B \leq N + A + \frac{\pi^2}{12A\delta^2}$. Choosing $A = \frac{\pi}{2\sqrt{3}\delta}$ we get $B \leq N + \frac{\pi}{\sqrt{3}\delta}$. This gives us for any sequence $a_n : M < n < M + N$ of complex numbers, we have

$$\sum_r \left|\sum_n a_n e(\alpha_r n)\right|^2 \leq \left(N + \frac{\pi}{\sqrt{3}\delta} - 1\right)\sum_n |a_n|^2.$$

3. Stronger results

Recall that

$$B = \max_r \sum_s |B(\alpha_r - \alpha_s)|$$

where α_r's are δ-well spaced. Suppose we choose b_n satisfying an additional condition that

$$B(\alpha) = 0, \ \forall \|\alpha\| \geq \delta.$$

Then because of the well spacing property of α's, we get $B = B(0)$. By Poisson summation,

$$B(\alpha) = \sum_n \hat{b}(n - \alpha).$$

Now we construct a function of real variable $b(x)$ such that

(1) $b(x) \geq 0$ for all $x \in \mathbb{R}$.
(2) $b(x) \geq 1$ for $M+1 \leq x \leq M+N$
(3) $b \in L^1(\mathbb{R})$ and $\hat{b}(t) = 0$ for $|t| \geq \delta$.

Then put $b_n = b(n)$ and hence by above argument $B(\alpha) = 0$ for all α with $\|\alpha\| \geq \delta$ and the Large sieve inequality holds with $C(N,R) = B(0) = \hat{b}(0)\int_{\mathbb{R}} b(x)dx$. It follows from works of Selberg and Beurling, that the following provides a good choice for $b(x)$.

Lemma. Let

$$F(z) = \left(\frac{\sin \pi z}{\pi}\right)^2 \left(\sum_{n=0}^{\infty} (z-n)^{-2} + (z+n)^{-2} + 2z^{-1}\right)$$

be a fucntion of complex variable z. Then F is entire, $F(z) = O(e^{2\pi|Imz|})$, $F(x) \geq sgn(x)$ for all real x and

$$\int_{-\infty}^{\infty} (F(x) - sgn(x))dx = 1$$

where $sgn(x)$ is defined by $sgn(x) = 1$ for $x \geq 0$ and $sgn(x) = -1$ for $x < 0$.

For $d > 0$, define $G_d(x) = \frac{1}{2}F(x) + \frac{1}{2}F(d-x)$. Then $G_d(x) \geq 0$ and $G_d(x) = 1_{[0,d]}(x)$, the indicator function of $[0,d]$. Considering the properties of F given by the above lemma it follows from Paley-Weiner theory, that $\hat{G}(t) = 0$ for $|t| \geq 1$. Now putting $d = \delta(N-1)$ we see that $b(x) = G_d(\delta x)$ satisfies all the properties we wanted and $\hat{b}(0) = N - 1 + 1/\delta$. Therefore this proves the following best possible large sieve inequality

$$\sum_r \left|\sum_n a_n e(\alpha_r n)\right|^2 \leq \left(N + \frac{1}{\delta} - 1\right) \sum_n |a_n|^2.$$

4. Large sive for additive characters

Let $Q > 1$ be a integer. Consider the rational numbers a/q for $1 \leq a \leq q$, $(a,q) = 1$ with $q \leq Q$. Then for $a/q \neq a'/q'$

$$\left|\frac{a}{q} - \frac{a'}{q'}\right| = \left|\frac{a'q - aq'}{qq'}\right| \geq \frac{1}{qq'} \geq \frac{1}{Q^2}.$$

So these points are $1/Q^2$-well spaced. For $M+1 \leq n \leq M+N$ let a_n be arbitrary complex numbers, $0 < \alpha < 1$ be a real number and we

define

$$S(\alpha) = \sum_{M+1}^{M+N} a_n e(n\alpha).$$

The large sieve inequality proved in the last section gives

$$\sum_{q\leq Q}\sideset{}{^*}\sum_{a(q)} \left|S\left(\frac{a}{q}\right)\right|^2 \leq (Q^2+N-1)\sum_n |a_n|^2.$$

5. Discrepency of distribution in arithematic progression

In this and in the following two sections we are going to discuss applications of Large sieve inequality. Let a_n be complex numbers for $M+1 \leq n \leq M+N$ and $q>1$ be an integer. Let

$$X = \sum_n a_n,\ X(q,v) = \sum_{n\equiv v(q)} a_n, \text{and } \Delta(q,v) = X(q,v) - \frac{X}{q}$$

where $1 \leq v \leq q$. Hence $\Delta(q,v)$ measures the disprepency which we expect to be small for a well distributed sequence. But without any additional assumption on a_n's we can show that it is small on an average. By orthogonality of additive characters, we see that

$$\Delta(q,v) = \frac{1}{q}\sum_{a\neq 0(q)} e\left(-\frac{av}{q}\right) S\left(\frac{a}{q}\right).$$

Thus $\Delta(q,v)$ can be thought of as the Fourier transform $\hat{S}(v)$. It is easy to check the corresponding Parsevals identity

$$\sum_{v(q)} |\Delta(q,v)|^2 = \frac{1}{q}\sum_{a\neq 0(q)} \left|S\left(\frac{a}{q}\right)\right|^2.$$

Obviously we wish to sum over $q \leq Q$ and apply large sieve inequality on the RHS. But we are faced with a problem. The sum on RHS is over all non-zero residues whereas we can apply large sieve when the sum is over all reduced (coprime to q) residue classes. Although the genral case can be done with a little more effort, since our aim here is to illustrate an application of large sieve inequality, we resort to the easiest case, namely when $q=p$, a prime as in that case the sums coincide, saving us the extra work. So we have

$$\sum_{p\leq Q} p\sum_{v(q)} |\Delta(p,v)|^2 \leq \sum_{p\leq Q}\sideset{}{^*}\sum_{a(p)} \left|S\left(\frac{a}{p}\right)\right|^2 \leq \sum_{q\leq Q}\sideset{}{^*}\sum_{a(q)} \left|S\left(\frac{a}{q}\right)\right|^2$$

where the sums with $p \leq Q$ are sums over primes. For the second inequality we have added some extra terms which are positive. Now large sieve inequality applies and we get

$$\sum_{p \leq Q} p \sum_{v(q)} |\Delta(p,v)|^2 \leq (Q^2 + N - 1) \sum_n |a_n|^2.$$

Specialising to the case when a_n is characteristic function of a set $A \in [M+1, M+N]$ and also choosing $Q = \sqrt{N}$ we have

$$\sum_{p \leq \sqrt{N}} p \sum_{v(p)} \left| |\{m \in A : m \equiv u(p)\}| - \frac{|A|}{p} \right|^2 \leq 2N|A|.$$

We observe that, obviously the inequality hold if we consider a subset of primes in $[1, \sqrt{N}]$ and a subset of residue classes.

6. Least quadratic non-residue

In this section we apply large sieve inequality to prove an average version of a famous theorem of Linnik.

Definition 6.1. For a prime p let $q(p)$ denote the least positive integer $< p$ which is not a square modulo p. Such $q(p)$ is called the least quadratic non-residue modulo p.

It is clear from the above definition that $q(p)$ is always a prime. It has been conjectured that $q(p) \ll_\epsilon p^\epsilon$ for any $\epsilon > 0$ and sufficiently large p. The best known result gives $q(p) \ll_\epsilon p^{\frac{1}{4\sqrt{e}} + \epsilon}$. In this context we prove the following theorem.

Theorem. For any $0 < \epsilon < 1$ and sufficiently large positive integer N we have

$$|\{p \leq N : q(p) > N^\epsilon\}| = O_\epsilon(1).$$

Proof. We define the following sets

$$\begin{aligned} A &= \{1, 2, ..., N\}, \\ \mathcal{P} &= \{p \leq \sqrt{N} : \left(\frac{n}{p}\right) = 1, \ \forall n \leq N^\epsilon\}, \\ \Omega_p &= \{v \mod p : \left(\frac{v}{p}\right) = -1\}. \end{aligned}$$

We see that $\omega(p) := |\Omega_p| = (p-1)/2$. Let

$$(A, \mathcal{P}, \Omega) := \{m \in A : m(mod p) \notin \Omega_p, \ \forall p \in \mathcal{P}\}.$$

Now we use the discrepency result obtained in the previous section with $X = (A, \mathcal{P}, \Omega)$ to obtain

$$\sum_{p\in\mathcal{P}}\sum_{v\in\Omega_p}\frac{|X|^2}{p^2} \leq 2N|X|.$$

Since cardinality of Ω_p is $(p-1)/2$ we finally get

$$\sum_{p\in\mathcal{P}}\left(1-\frac{1}{p}\right) \leq \frac{4N}{|X|}.$$

We observe that to prove the result it is enough to show that cardinality of $\mathcal{P}$ is bounded and it follows if we show that

$$\sum_{p\in\mathcal{P}}\left(1-\frac{1}{p}\right) \ll_\epsilon 1.$$

Therefore it is enough to prove that $|X| \gg_\epsilon N$.

Recall that

$$X = \{m \leq N : \left(\frac{m}{p}\right) = 1, \forall p \in \mathcal{P}\}.$$

If an integer m has all prime factors less than N^ϵ, then $m \in X$. Consider an integer m of the form $m = np_1...p_k \leq N$ such that $N^{\epsilon-\epsilon^2} < p_j < N^\epsilon$ and $k = [1/\epsilon]$. Then $p_1...p_k > N^{1-\epsilon}$ so $n \leq N^\epsilon$, therefore $(\frac{n}{p}) = 1$ for all $p \in \mathcal{P}$. Since for all $j = 1, ..., k$, $p_j < N^\epsilon$, we have $(\frac{p_j}{p}) = 1$ for all $p \in \mathcal{P}$. Thus $(\frac{m}{p}) = 1$ for all $p \in \mathcal{P}$ implies that $m \in X$. Counting these integers we get the following lower bound for $|X|$.

$$X \geq \sum_{\substack{p_1,...,p_k \\ N^{\epsilon-\epsilon^2}<p_j<N^\epsilon}}\left[\frac{N}{p_1...p_k}\right] \geq N\left(\sum_{N^{\epsilon-\epsilon^2}<p<N^\epsilon}\frac{1}{p}\right)^k \gg_\epsilon N.$$

7. Arithmetic large sieve

In this section we are going to explore sieving properties of Large sieve inequality. We continue using notations from the last section. Recall that

$$(A, \mathcal{P}, \Omega) := \{m \in A : m(mod p) \notin \Omega_p,\ \forall p \in \mathcal{P}\}.$$

Let a_n's be arbitrary complex numbers for $n \in (A, \mathcal{P}, \Omega)$ and define $a_n = 0$ for $n \notin (A, \mathcal{P}, \Omega)$. We are going to find an upper bound for

$$Z = \sum_{n\in(A,\mathcal{P},\Omega)} a_n.$$

Here we are sieving by a subset Ω_p of residues modulo $p \in \mathcal{P}$ and the number of such classes can be θp with $0 < \theta < 1$. This is much "large scale" sieving compared to the sieves we studied earlier where we sieved only by primes less than z, i.e choosing $\Omega_p = 0$. Hence the name Large sieve in contrast with "small" sieves. Coming back to the proof, we define a multiplicative function h supported on square-free numbers by

$$h(p) = \frac{\omega(p)}{p - \omega(p)}.$$

(the $g(p)$ as defined in the Selberg sieve would be, in the present case given by $g(p) = \omega(p)/p$. Thus h defined above is exactly same as in Selberg Sieve)

Let

$$S(\alpha) = \sum_n a_n e(n\alpha).$$

Lemma. For any $q > 1$,

$$h(q)|S(0)|^2 \leq \sum_{a(q)}^{*} \left| S\left(\frac{a}{q}\right)\right|^2.$$

Proof. We first prove it for $q = p$ a prime. We see that $X(p, v) = \sum_{n \equiv v(p)} a_n = 0$ for all $v \in \Omega_p$. For any $p \in \mathcal{P}$, we have

$$|S(0)|^2 = \left|\sum_n a_n\right|^2 = \left|\sum_{v(p)} \sum_{n \equiv v(p)} a_n\right|^2 = \left|\sum_{v(p)} 1(p, v) X(p, v)\right|^2$$

where

$$1(p, v) = \begin{cases} 1 & \text{if } X(p, v) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Using Cauchy-Schwarz inequality

$$|S(0)|^2 \leq \sum_{v(p)} 1(p, v)^2 \sum_{v(p)} |X(p, v)|^2.$$

The first sum clearly is $\leq p - \omega(p)$ whereas the second sum is equal to $1/p \sum_{a(p)} |S(a/p)|^2$. Hence

$$|S(o)|^2 \leq (p - \omega(p)) \frac{1}{p} \sum_{a(p)} \left| S\left(\frac{a}{p}\right)\right|^2.$$

Seperating the summand corresponding to $0(modp)$ in the sum on the right, we get

$$h(p)|S(0)|^2 \sideset{}{^*}\sum_{a(p)} \left|S\left(\frac{a}{p}\right)\right|^2$$

proving the result in this case.

Now we consider composite moduli. Let $q = q_1q_2$ with $(q_1, a_2) = 1$ and suppose that we have proved the result for both q_1 and q_2. Each a mod q can be represented as $a = a_1q_2 + a_2q_1$ where $a_1 \in (\mathbb{Z}/q_1\mathbb{Z})^*$ and $a_2 \in (\mathbb{Z}/q_2\mathbb{Z})^*$. This gives us

$$\begin{aligned}
\sideset{}{^*}\sum_{a(q)} \left|S\left(\frac{a}{q}\right)\right|^2 &= \sideset{}{^*}\sum_{a_1(q_1)} \sideset{}{^*}\sum_{a_2(q_2)} \left|S\left(\frac{a_1}{q_1} + \frac{a_2}{q_2}\right)\right|^2 \\
&= \sideset{}{^*}\sum_{a_1(q_1)} \sideset{}{^*}\sum_{a_2(q_2)} \left|\sum_n a_n e\left(\frac{a_1 n}{q_1}\right) e\left(\frac{a_2 n}{q_2}\right)\right|^2 \\
&= \sideset{}{^*}\sum_{a_1(q_1)} \sideset{}{^*}\sum_{a_2(q_2)} \left|\sum_n b_n e\left(\frac{a_2 n}{q_2}\right)\right|^2
\end{aligned}$$

where

$$b_n = a_n e\left(\frac{a_1 n}{q_1}\right).$$

Let

$$S_1(\alpha) = \sum_n b_n e(n\alpha).$$

Using the result for q_2, with $S_1(\alpha)$ in place of $S(\alpha)$.

$$\sideset{}{^*}\sum_{a_2(q_2)} \left|S_1\left(\frac{a_2}{q_2}\right)\right|^2 \geq |S_1(0)|^2 h(q_2).$$

Continuing from above

$$\begin{aligned}
\sideset{}{^*}\sum_{a(q)} \left|S\left(\frac{a}{q}\right)\right|^2 &= \sideset{}{^*}\sum_{a_1(q_1)} \sideset{}{^*}\sum_{a_2(q_2)} \left|S_1\left(\frac{a_2}{q_2}\right)\right|^2 \\
&\geq h(q_2) \sideset{}{^*}\sum_{a_1(q_1)} |S_1(0)|^2 \\
&= h(q_2) \sideset{}{^*}\sum_{a_1(q_1)} \left|\sum_n a_n e\left(\frac{a_1 n}{q_1}\right)\right|^2 \\
&\geq h(q_2)h(q_1)|S(0)|^2.
\end{aligned}$$

The last step uses the result of the lemma for q_1. This finishes the proof as h is a multiplicative function.

Summing over $q \leq Q$, we get

$$|S(0)|^2 \sum_{q \leq Q} h(q) \leq \sum_{q \leq Q} \sum_{a(q)}^{*} \left| S\left(\frac{a}{q}\right) \right|^2 \leq (Q^2 + N) \sum_{n} |a_n|^2.$$

This gives

$$|Z|^2 \leq \frac{Q^2 + N}{J} \sum_{n} |a_n|^2$$

where $J = \sum_{q \leq Q} h(q)$.

In particular if a_n is the indicator function of Z, then

$$|\{n \leq N : n(mod p) \notin \Omega_p, \forall p \in \mathcal{P}\}| \leq \frac{Q^2 + N}{J}.$$

Introduction to Circle Method

CHAPTER 4

Vinogradov's three primes theorem

We are going to discuss the proof of the following theorem in this section.

Theorem. Every sufficiently large odd integer can be written as a sum of three primes.

1. Method of proof

Let 1_P be the indicator function of primes. For a large odd integer N, let $R(N)$ be the number of ways N can be wrtitten as a sum of three primes. Then

$$R(N) = \sum_{n_1+n_2+n_3=N} 1_P(n_1)1_P(n_2)1_P(n_3) = 1_P * 1_P * 1_P(N)$$

where for a functions $f, g : \mathbb{Z} \longrightarrow \mathbb{C}$ convolution $f * g$ is defined by

$$f * g(n) = \sum_m f(m)g(n-m).$$

Since the Pontrygin dual of $\mathbb{Z}$ is the torus $\mathbb{R}/\mathbb{Z}$, Fourier transform $\hat{f}$ of f is defined by

$$\hat{f}(r) = \sum_n f(n)e(-nr), \ \forall r \in \mathbb{R}/\mathbb{Z} \ \text{ where } e(z) = e^{2\pi iz}.$$

Fourier inversion formula gives

$$f(n) = \int_{\mathbb{R}/\mathbb{Z}} \hat{f}(\alpha)e(n\alpha)d\alpha.$$

Recalling that $\widehat{f * g} = \hat{f}\hat{g}$ and using Fourier inversion we get from above

$$R(N) = \int_{\mathbb{R}/\mathbb{Z}} \hat{1_P}(\alpha)^3 e(N\alpha)d\alpha.$$

This integral would be evaluated by Circle method discovered by Hardy, Littlewood and Ramanujan. We can think of $\mathbb{R}/\mathbb{Z}$ as the interval $[0, 1]$ with the end points indentified. Roughly speaking the main contribution to the above integral comes from the reals $\alpha \in [0, 1]$ which can be approximated by a rational with small denominators, collection

of all such reals is called Major arc and rest is called Minor arc. We would make these notions precise in the coming sections.

1.1. Expected size of $R(N)$. Suppose primes upto N are uniformly distributed with probability $1/\log N$. Consider prime triplets $p_1, p_2, p_3 \leq N/3$. There are $\sim N^3/(\log N)^3$ such triplets whose sum $p_1 + p_2 + p_3 \leq N$. If every odd number upto N are equaly represented by such sums then we would expect that every odd integer upto N is a sum of three primes in $\sim N^2/(\log N)^3$ ways. Instead of the indicator function 1_P of primes which is difficult ot handle we would be using the Van-Mangoldt function defined by

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^r \text{ power of a prime} \\ 0 & \text{otherwise} \end{cases}$$

Thus instead of $R(N)$ we would be estimating

$$\Lambda^*(N) = \sum_{n_1+n_2+n_3=N} \Lambda(n_1)\Lambda(n_2)\Lambda(n_3).$$

We can think of $\Lambda^*(N)$ as $R(N)$ with each prime weighted by a $\log N$. Hence the expected order is $\sim N^2$.

2. Main theorem

To establish Theorem1 it is enough to prove that $R(N) > 0$ for all sufficientky large odd integer N. We would see a proof of the following much more precise result by Vinogradov.

Theorem. For any fixed $A > 0$,

$$\Lambda^*(N) = \frac{1}{2}\mathcal{C}(N)N^2 + O\left(\frac{N^2}{\log^A N}\right)$$

where

$$\mathcal{C}(N) = \prod_{p|N}\left(1 - \frac{1}{(p-1)^2}\right)\prod_{p\not|N}\left(1 + \frac{1}{(p-1)^3}\right).$$

As observed in case of $R(N)$ using Fourier inversion formula we get

$$\Lambda^*(N) = \int_0^1 (\hat{\Lambda}(\alpha))^3 e(N\alpha)d\alpha$$

where we assume that Λ is supported on integers in $[1, N]$. Observe that the integral over $[0, 1]$ is the same as integral over $\mathbb{R}/\mathbb{Z}$ as the integrad is a periodic function with period 1. In the next section we set up the method to estimate such integrals assymptotically.

3. Major and Minor arcs

Dirichlets box principle (also known as Pigeon hole principle) gives this following lemma.

Lemma. Let $\alpha \in (0,1)$ and $Q > 1$. Then there exists $q \leq Q$ such that

$$|\alpha - \frac{a}{q}| \leq \frac{1}{qQ} \leq \frac{1}{q^2}.$$

For some $a \leq q$ with $(a,q) = 1$.

The following is one of the most crucial result in this method, proof of which we defer till the end.

Proposition 1. *Let $\alpha \in (0,1]$ be such that $|\alpha - a/q| \leq 1/q^2$ for some $a \leq q$ and $(a,q) = 1$. Suppose that Λ is supported on $[1,N]$. Then*

$$|\hat{\Lambda}(\alpha)| \ll \left(\frac{N}{\sqrt{q}} + N^{4/5} + \sqrt{N}\sqrt{q}\right) \log^4 N.$$

Following is an easy observation from Proposition 1.

Corollary 3.1. *If q satisfies $\log^A N \leq q \leq N/\log^A N$ for some $A \geq 1$. Then*

$$|\hat{\Lambda}(\alpha)| \ll \frac{N}{\log^{A/2-4} N}.$$

Using Dirichlet's principle with $Q = N/\log^A N$ we get $q \leq N/\log^A N$ such that

$$|\alpha - \frac{a}{q}| \leq \frac{\log^A N}{qN} \leq \frac{1}{q^2}.$$

If $q > \log^A N$ then we can use the above Corollary to estimate $\hat{\Lambda}(\alpha)$. If $q \leq \log^A N$ then trivially

$$|\alpha - \frac{a}{q}| \leq \frac{\log^A N}{qN} \leq \frac{\log^A N}{N}.$$

For $a \leq q$ with $(a,q) = 1$ define

$$M_{a/q} = \{\alpha \in (0,1] || \alpha - \frac{a}{q}| \leq \frac{\log^A N}{N}.$$

This is an interval of length $2\log^A N/N$ around a/q. We take the union of all such intervals with $q \leq \log^A N$

$$\mathfrak{M} = \cup_{q=1}^{\log^A N} \cup_{\substack{a \leq q \\ (a,q)=1}} M_{a/q}.$$

This is called the **Major arc** and the complement $\mathfrak{m} = (0,1] - \mathfrak{M}$ is called the **Minor arc**.

The following lemma shows that the intervals $M_{a/q}$ in the Major arc are disjoint.

Lemma. For distinct rationals a/q and a'/q' in reduced form we have $M_{a/q} \cap M_{a'/q'} = \phi$

PROOF. If possible let $\alpha \in M_{a/q} \cap M_{a'/q'}$. Then

$$\frac{1}{qq'} \leq |\frac{a}{q} - \frac{a'}{q'}| \leq |\alpha - a/q| + |\alpha - a'/q'| \leq \frac{2\log^A N}{N}.$$

Thus

$$\max(q.q') \geq \left(\frac{N}{\log^A N}\right)^{1/2}$$

which contradics that $M_{a/q}$ and $M_{a'/q'}$ belongs to Major arc and so $q, q' \leq \log^A N$.

Lemma. If $\alpha \in \mathfrak{m}$ then $|\hat{\Lambda}(\alpha)| \ll N/\log^B N$ with $B = A/2 - 4$.

PROOF. By Dirichlet's principle with $Q = N/\log^A N$ we have

$$|\alpha - \frac{a}{q}| \leq \frac{\log^A N}{qN}, \quad \text{for some } q \leq \frac{N}{\log^A N}.$$

Trivially the right hand side is less than $\frac{\log^A N}{N}$. Since $\alpha \in \mathfrak{m}$ this implies $q > \log^A N$. We have

$$\log^A N < q \leq \frac{N}{\log^A N}.$$

Thus Corollary 1 applies and gives the result.

Now we are ready to estimate the integral over Minor arc.

Lemma. For any $B > 0$ we have

$$|\int_{\mathfrak{m}} \hat{\Lambda}(\alpha)^3 e(N\alpha) d\alpha| \ll \frac{N^2}{\log^B N}.$$

PROOF. Using Lemma 3, we get

$$|\int_{\mathfrak{m}} \hat{\Lambda}(\alpha)^3 e(N\alpha) d\alpha| \leq \frac{N}{\log^{A/2-4} N} \int_0^1 |\hat{\Lambda}(\alpha)|^2 d\alpha.$$

It is easy to see that

$$\int_0^1 |\hat{\Lambda}(\alpha)|^2 d\alpha = \sum_{k=1}^{N} \Lambda(k)^2 \ll \frac{N^2 \log N}{\log^{A/2-4} N}.$$

Now the result follows by taking $A = 2(B + 5)$.

4. Major arc

In this section we would find an assymptotic formula for

$$\hat{\Lambda}(\alpha) = \sum_{n \leq N} \Lambda(n) e(n\alpha), \quad \text{when } |\alpha - \frac{a}{q}| \leq \frac{\log^A N}{N}$$

with $a \leq q \leq \log^A$ and $(a.q) = 1$.

The first lemma reduces the problem to finding the sum of $\Lambda(n)$ over arithmetic progressions.

Lemma. Let $a \leq q \leq \log^A N$ and $(a, q) = 1$ and $|\alpha - a/q| \leq \log^A N/qN$. Suppose $G : [1, N] \longrightarrow \mathbb{R}$ be such that $|G(x)| \leq \log N, \forall x \in [1, N]$ and

$$|\sum_{x \in X} G(x)| = O\left(\frac{N}{\log^B N}\right), \quad \text{with } B \geq 4A + 2$$

for any arithmetic progression $X = \{b, b+q, \cdots, b+(m-1)q\}$ modulo q in $[1, N]$. Then

$$|\sum_{x \in X} G(x) e(\alpha x)| = O\left(\frac{N}{\log^A N}\right).$$

Proof. Let $\beta = \alpha - a/q$ and $X = \{b, b+q, ..., (m-1)q + b\}$. Let $x, y \in X$ be given by $x = b + l_1 q$ and $y = b + l_2 q$. Then

$$\begin{aligned} |e(\beta x) - e(\beta y)| &= |e(\beta(b + l_1 q)) - e(\beta(b + l_2 q))| \\ &= |1 - e(\beta(l_1 - l_2)q)| \\ &\leq 2\pi |x - y||\beta| \leq 2\pi m q \beta. \end{aligned}$$

Fix $x_0 \in X$ and note that $e(ax/q) = e(ab/q)$ for all $x \in X$. Now

$$\begin{aligned} \left|\sum_{x \in X} G(x) e(\alpha x)\right| &= \left|\sum_{x \in X} G(x) e(\beta x)\right| \\ &\leq \left|\sum_{x \in X} G(x) e(\beta(x - x_0))\right| \\ &+ \left|\sum_{x \in X} G(x) e(\beta x_0)\right| \\ &\leq 2\pi m q \beta \sum_{x \in X} |G(x)| + \left|\sum_{x \in X} G(x)\right| \\ &\leq 2\pi m q \beta m \log N + c \frac{N}{\log^B N}. \end{aligned}$$

The last step utilises the assumptions of the lemma. Divide the $[1, N]$ into N/m_0 arithmetic progression modulo q of length $\leq m_0$. This gives

$$\left|\sum_{x\leq N} G(x)e(\alpha x)\right| \leq \frac{N}{m_0}\left(m_0^2 q\beta \log N + \frac{N}{\log^B N}\right).$$

We finish the proof by choosing $m_0 = N/(\log N)^{B-A}$.

Let $X = \{b, b+q, \cdots, b+(m-1)q\}$ be an arithmetic progression modulo q in $[1, N]$. Then by Siegel-Walfisz theorem

$$\sum_{n\in X} \Lambda(n) = \frac{mq}{\phi(q)} + O(\frac{N}{\log^C N})$$

for any $C > 0$. Define $H_q : [1, N] \longrightarrow \mathbb{R}$ by

$$H_q(x) = \begin{cases} \frac{q}{\phi(q)} & \text{if } (x.q) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.1}$$

Since $(a, q) = 1$ all terms of X are co-prime to q. Thus $\sum_{x\in X} H_q(x) = mq/\phi(q)$. Thus Siegel-Walfisz theorem acn be restated as

$$|\sum_{x\in X} (\Lambda(x) - H_q(x))| \ll \frac{N}{\log^C N}.$$

This holds for all arithmetic progression modulo q in $[1, N]$, so the above Lemma 5 can be applied to conclude that

$$|\sum_{x\leq N} (\Lambda(x) - H_q(x))\, e(x\alpha)| \ll \frac{N}{\log^B N}$$

for all $\alpha \in (0, 1]$ with $|\alpha - a/q| \leq \log^A N/Nq$. This can be written as

$$\hat{\Lambda}(\alpha) = \sum_{x\leq N} H_q(x)e(x\alpha) + O\left(\frac{N}{\log^B N}\right).$$

The following lemma simplifies the right hand side

Lemma. Suppose $q \leq \log^A N$ and $b \leq q,\ (b, q) = 1$. Let $\alpha \in (0, 1]$ be such that $|\alpha - b/q| \leq \log^A NqN$. Then

$$\sum_{x\leq N} H_q(x)e(x\alpha) = \frac{\mu(q)}{\phi(q)} \sum_{x\leq N} e(\beta x) + O(\log^{2A} N)$$

where $\beta = \alpha - a/q$.

PROOF. Since $H_q(x) = 0$ if $(x,q) > 1$, defining $X_a = \{n \leq N : n \equiv a(q)\}$ we have

$$\begin{aligned}\sum_{x\leq N} H_q(x)e(x\alpha) &= \sum_{a(q)}^{*}\sum_{x\in X_a} H_q(x)e(x\alpha)\\ &= \frac{q}{\varphi(q)}\sum_{a(q)}^{*}\sum_{x\in X_a} e\left(\frac{bx}{q}\right)e(\beta x)\\ &= \frac{q}{\varphi(q)}\sum_{a(q)}^{*} e\left(\frac{ba}{q}\right)\sum_{x\in X_a} e(\beta x).\end{aligned}$$

Fix a_0 with $(a_0,q)=1$ and we observe

$$\sum_{k=0}^{m-1}(e(\beta(a+kq)) - e(\beta(a_0+kq))) = (e(\beta(a-a_0))-1)\sum_{k=0}^{m-1} e(\beta(a_0+kq)).$$

Since X_a has $[(N-a)/q]$ terms in $[1,N]$, the number of terms in X_{a_1} and X_{a_2} may differ by atmost 1. Thus we get from above

$$\begin{aligned}\sum_{x\in X_a} e(\beta x) - \sum_{x\in X_{a_0}} e(\beta x) &\leq |e(\beta(a-a_0))-1|\left|\sum_{x\in X_{a_0}} e(\beta x)\right| + 1\\ &\leq 2\pi|a-a_0||\beta|\left|\sum_{x\in X_{a_0}} e(\beta x)\right| + 1\\ &\leq 2\pi\frac{\log^A N}{q} + 1.\end{aligned}$$

Now we have

$$\begin{aligned}\left|\sum_{x\leq N} e(\beta x) - q\sum_{x\in X_{a_0}} e(\beta x)\right| &= \left|\sum_{a(q)}\sum_{x\leq X_a} e(\beta x) - q\sum_{x\in X_{a_0}} e(\beta x)\right|\\ &= \left|\sum_{a(q)}\left(\sum_{x\leq X_a} e(\beta x) - \sum_{x\in X_{a_0}} e(\beta x)\right)\right|\\ &\leq \sum_{a(q)}\left|\sum_{x\leq X_a} e(\beta x) - \sum_{x\in X_{a_0}} e(\beta x)\right|\\ &\leq \sum_{a(q)}\left(2\pi\frac{\log^A N}{q} + 1\right) \ll \log^A N.\end{aligned}$$

Therefore we get

$$\begin{aligned}\sum_{x\leq N} H_q(x)e(x\alpha) &= \frac{q}{\varphi(q)}\sideset{}{^*}\sum_{a(q)} e\left(\frac{ba}{q}\right)\sum_{x\in X_a} e(\beta x)\\ &= \frac{q}{\varphi(q)}\sideset{}{^*}\sum_{a(q)} e\left(\frac{ba}{q}\right)\left(\frac{1}{q}\sum_{x\leq N} e(\beta x)+O\left(\frac{\log^A N}{q}\right)\right)\\ &= \frac{1}{\varphi(q)}\sideset{}{^*}\sum_{a(q)} e\left(\frac{ba}{q}\right)\sum_{x\leq N} e(\beta x)+O(\log^A N).\end{aligned}$$

It is a standard fact and not difficult to check that the Ramanujan sum

$$\sideset{}{^*}\sum_{a(q)} e\left(\frac{ba}{q}\right)=\mu(q)$$

for all b coprime to q. This establishes the lemma.

Recalling the relation between $\hat{\Lambda}$ and H_q, we have

$$\hat{\Lambda}(\alpha)=\frac{\mu(q)}{\varphi(q)}\sum_{x\leq N} e(\beta x)+O\left(\frac{N}{\log^B N}\right).$$

Now we are ready to compute the Major arc contribution. Using Lemma 6, we write

$$\begin{aligned}&\int_{\mathfrak{M}} \hat{\Lambda}(\alpha)^3 e(-N\alpha)d\alpha\\ =\ &\sum_{q=1}^{\log^A N}\sideset{}{^*}\sum_{a(q)}\int_{|\beta|\leq\frac{\log^A N}{N}}\left(\frac{\mu(q)}{\phi(q)}\sum_{x\leq N} e(\beta x)+O\left(\frac{N}{\log^B N}\right)\right)^3 e(-N\alpha)d\alpha.\end{aligned}$$

where the sum over $a \mod q$ runs over reduced residue classes.

By choosing a sufficiently large C in Siegel-Walfisz theorem we get a large B. With such a B it is easy to see that the O-terms are bounded above by $N^2/\log^A N$. Now we have to estimate the main term

$$\sum_{q=1}^{\log^A N}\frac{\mu(q)}{\phi(q)}\sideset{}{^*}\sum_{a(q)} e\left(-\frac{aN}{q}\right)\int_{|\beta|\leq\frac{\log^A N}{N}}\left(\sum_{x\leq N} e(\beta x)\right)^3 e(-\beta N)d\beta.$$

The next lemma extends the integral to $(0,1]$

Lemma. With assupmtions on a, q and β as above we have

$$\begin{aligned}&\int_{|\beta|\leq\frac{\log^A N}{N}}\left(\sum_{x\leq N}e(\beta x)\right)^3 e(-\beta N)d\beta\\ =\ &\int_0^1\left(\sum_{x\leq N}e(\beta x)\right)^3 e(-\beta N)d\beta+O\left(\frac{N^2}{\log^A N}\right).\end{aligned}$$

Proof. It is easy to see that

$$|\sum_{x\leq N}e(\beta x)|\leq\min\{N,\frac{1}{\|\beta\|}\}\leq\frac{CN}{1+N\|\beta\|}$$

for some constant $C>4$.
Thus

$$\begin{aligned}|\int_{|\beta|\leq\frac{\log^A N}{N}}\left(\sum_{x\leq N}e(\beta x)\right)^3 e(-\beta N)d\beta\ &-\ \int_0^1\left(\sum_{x\leq N}e(\beta x)\right)^3 e(-\beta N)d\beta|\\ &\leq\ \int_{\frac{\log^A N}{N}}^{1-\frac{\log^A N}{N}}|\sum_{x\leq N}e(\beta x)|^3 d\beta\\ &\leq\ \int_{\frac{\log^A N}{N}}^{1-\frac{\log^A N}{N}}\left(\frac{CN}{1+N\|\beta\|}\right)^3 d\beta\\ &\leq\ \int_{\frac{\log^A N}{N}}^{1/2}\left(\frac{CN}{1+N\beta}\right)^3 d\beta.\end{aligned}$$

Now it is easy to see that this last integral is bounded by $O(N^2/\log^A N)$ which finishes the lemma.

Now let

$$I_N=\int_0^1\left(\sum_{x\leq N}e(\beta x)\right)^3 e(-\beta N)d\beta.$$

It is easy to see that I_N is exactly the number of ways N can be written as $N=n_1+n_2+n_3$ where $1\leq n_i\leq N$ for $i=1,2,3$. Thus

$$I_N=\frac{(N-1)(N-2)}{2}=\frac{N^2}{2}+O(N).$$

Hence the main term is

$$\begin{aligned}
&\sum_{q=1}^{\log^A N} \frac{\mu(q)}{\phi(q)^3} \sideset{}{^*}\sum_{a(q)} e(\frac{aN}{q}) \left(\frac{N^2}{2} + O(N)\right) \\
= \quad & \frac{N^2}{2} \sum_{q=1}^{\log^A N} \frac{\mu(q)}{\phi(q)^3} \sideset{}{^*}\sum_{a(q)} e(\frac{aN}{q}) + O(N) \\
= \quad & \frac{N^2}{2} \sum_{q=1}^{\infty} \frac{\mu(q)}{\phi(q)^3} \sideset{}{^*}\sum_{a(q)} e(\frac{aN}{q}) + O\left(N^2 \sum_{q>\log^A N} \frac{1}{\phi(q)^2}\right) \\
= \quad & \frac{N^2}{2} \sum_{q=1}^{\infty} \frac{\mu(q)}{\phi(q)^3} \sideset{}{^*}\sum_{a(q)} e(\frac{aN}{q}) + O\left(\frac{N^2}{\log^A N}\right) \\
= \quad & \frac{N^2}{2} \sum_{q=1}^{\infty} \frac{\mu(q)}{\phi(q)^3} C_q(N) + O\left(\frac{N^2}{\log^A N}\right)
\end{aligned}$$

where

$$C_q(N) = \sideset{}{^*}\sum_{a(q)} e(\frac{aN}{q})$$

is called Ramanujan sum which is multiplicative and

$$C_p(N) = \begin{cases} p-1 \text{ if } \ p|N \\ -1 \text{ if } \ p \nmid N \end{cases}$$

Thus we get

$$\begin{aligned}
&\int_{\mathfrak{M}} \hat{\Lambda}(\alpha)^3 e(-N\alpha) d\alpha \\
= \quad & \frac{N^2}{2} \prod_{p|N} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{p \nmid N} \left(1 + \frac{1}{(p-1)^3}\right) + O\left(\frac{N^2}{\log^A N}\right)
\end{aligned}$$

finishing the proof of the main theorem. The next section in entirely devoted to the proof of Proposition 1.

5. Minor arc

5.1. Vaughan's identity.

Lemma (Vaughan's identity). For any $y, z \geq 1$ and $n > z$, we have

$$\Lambda(n) = \sum_{\substack{b|n \\ b \leq y}} \mu(b) \log \frac{n}{b} - \sum \sum_{\substack{bc|n \\ b \leq y, c \leq z}} \mu(b)\Lambda(c) + \sum \sum_{\substack{bc|n \\ b > y, c > z}} \mu(b)\Lambda(c).$$

PROOF. It is easy to show that $\sum_{d|n} \Lambda(d) = \log n$ (i.e $1 * \Lambda = \log$) and by Mobius inversion we get

$$\Lambda(n) = \sum_{b|n} \mu(b) \log(\frac{n}{b}).$$

We partition the sum according to b *ley* and $b > y$ to obtain

$$\Lambda(n) = \sum_{\substack{b|n \\ b \leq y}} \mu(b) \log(\frac{n}{b}) + \sum_{\substack{b|n \\ b > y}} \mu(b) \log(\frac{n}{b}).$$

Using $1 * \Lambda = \log$ in the second sum and dividing the sum again we get

$$\begin{aligned} & \sum_{\substack{b|n \\ b>y}} \mu(b) \log(\frac{n}{b}) \\ = & \sum_{\substack{bc|n \\ b>y}} \mu(b)\Lambda(c) \\ = & \sum_{\substack{bc|n \\ b>y, c>z}} \mu(b)\Lambda(c) + \sum_{\substack{bc|n \\ b>y, c\leq z}} \mu(b)\Lambda(c) \end{aligned}$$

Now

$$\sum_{\substack{bc|n \\ b>y, c\leq z}} \mu(b)\Lambda(c) + \sum_{\substack{bc|n \\ b\leq y, c\leq z}} \mu(b)\Lambda(c) = \sum_{\substack{bc|n \\ c\leq z}} \mu(b)\Lambda(c) = \sum_{c<z} \Lambda(c) \sum_{b|\frac{n}{c}} \mu(b) = 0$$

where the inner sum of the last term vanishes as $c < z < n$ implies $n/c > 1$ The Lemma follows by putting these together.

5.2. Proof of Proposition 1. Using Vaughan's identity we divide

$$\sum_{n\leq X} \Lambda(n) e(n\alpha)$$

into three subsums. The first sum is

$$\begin{aligned}
&\sum_{n\le X} e(n\alpha)\sum_{\substack{b|n\\ b\le y}}\mu(b)\log(n/b)\\
=\;&\sum_{b\le y}\mu(b)\sum_{\substack{n\le X\\ b|n}}\log(n/b)e(n\alpha)\\
=\;&\sum_{b\le y}\mu(b)\sum_{a\le X/b}\log a e(ab\alpha)\\
=\;&\sum_{b\le y}\mu(b)\sum_{a\le X/b}e(ab\alpha)\int_1^a\frac{dt}{t}\\
=\;&\sum_{b\le y}\mu(b)\sum_{a\le X/b}e(ab\alpha)\int_1^{X/b}\chi_a(t)\frac{dt}{t}\\
=\;&\sum_{b\le y}\mu(b)\int_1^{X/b}\left(\sum_{a\le X/b}\chi_a(t)e(ab\alpha)\right)\frac{dt}{t}\\
=\;&\sum_{b\le y}\mu(b)\int_1^{X/b}\sum_{t\le a\le X/b}e(ab\alpha)\frac{dt}{t}
\end{aligned}$$

Using

$$\sum_{n\le M}e(n\alpha)\le\min\{\frac{1}{\|\alpha\|},M\}$$

we get the first subsum is

$$\le\log x\sum_{b\le y}\min\{\frac{1}{\|b\alpha\|},\frac{X}{b}\}.$$

Now we need the following general Lemma

Lemma. Let Q and R are positive integers and α be a real number such that $|\alpha-a/q|\le 1/q^2$ holds for some $q\le Q$. Then

$$\sum_{x=1}^{R}\min\{\frac{1}{\|x\alpha\|},\frac{Q}{x}\}\le C\log Q\log R\left(q+R+\frac{Q}{q}\right)$$

where $C>0$ is an absolute constant.

Using Lemma 5.2 we see that the first term is bounded by

$$(\log X)^3\left(q+y+\frac{X}{q}\right).$$

Now we treat the second term

$$\begin{aligned}\sum_{n\leq X} e(n\alpha) \sum_{\substack{bc|n\\ b\leq y, c\leq z}} \mu(b)\Lambda(c) &= \sum_{b\leq y}\mu(b)\sum_{c\leq z}\Lambda(c)\sum_{a\leq X/bc} e(abc\alpha)\\ &\leq \sum_{b\leq y}\left|\sum_{c\leq z}\Lambda(c)\sum_{a\leq X/bc} e(abc\alpha)\right|\\ &\leq \sum_{l\leq yz}\sum_{\substack{c\leq z\\ c|l}}\Lambda(c)\left|\sum_{a\leq X/l} e(al\alpha)\right|\end{aligned}$$

where the second sum

$$\sum_{\substack{c\leq z\\ c|l}}\Lambda(c)\leq\sum_{c|l}\Lambda(c)=\log l.$$

Hence the RHS is bounded above by

$$(\log y+\log z)\sum_{l\leq yz}\left|\sum_{a\leq X/l} e(al\alpha)\right|.$$

As in the case of first subsum, using Lemma 5.2 this is bounded above by

$$(\log X)^3\left(q+yz+\frac{X}{q}\right).$$

Now we are left with the last term

$$\begin{aligned}U &= \sum_{n\leq x}\left(\sum_{\substack{bc|n\\ b<y, c>z}}\mu(b)\Lambda(c)\right)e(n\alpha)\\ &= \sum_{n\leq x}\sum_{\substack{b|n\\ y<b<n/z}}\mu(b)\sum_{\substack{c|n/b\\ c>z}}\Lambda(c)e(n\alpha)\\ &= \sum_{y<k<x/z}\sum_{\substack{b|k\\ y<b<x/z}}\mu(b)\sum_{z<c<x/k}\Lambda(c)e(ck\alpha)\end{aligned}$$

where the last sum was obtained by putting $n = ck$. We dyadically partition the interval $(y, x/z)$ and write $U = \sum_M U_M$ where

$$U_M = \sum_{M<k\le 2M} \left(\sum_{\substack{b|k \\ y<b<x/z}} \mu(b) \right) \left(\sum_{z<c<x/k} \Lambda(c) e(ck\alpha) \right).$$

Applying Cauchy-Schwarz's inequality

$$U_M^2 \le \left(\sum_{M<k\le 2M} \left(\sum_{\substack{b|k \\ y<b<x/z}} \mu(b) \right)^2 \right) \left(\sum_{M<k\le 2M} \left(\sum_{z<c<x/k} \Lambda(c) e(ck\alpha) \right)^2 \right).$$

Trivially

$$\left(\sum_{\substack{b|k \\ y<b<x/z}} \mu(b) \right)^2 \le \left(\sum_{b|k} 1 \right)^2 \le \tau(k)^2.$$

Hence

$$\sum_{M<k\le 2M} \left(\sum_{\substack{b|k \\ y<b<x/z}} \mu(b) \right)^2 \le \sum_{M<k\le 2M} \tau(k)^2 \ll M(\log M)^3.$$

The other term

$$\begin{aligned} \sum_{M<k\le 2M} & \left(\sum_{z<c<x/k} \Lambda(c) e(ck\alpha) \right)^2 \\ &= \sum_{M<k\le 2M} \sum_{z<c_1,c_2<x/k} \Lambda(c_1)\Lambda(c_2) e(k(c_1-c_2)\alpha) \\ &= \sum_{z<c_1,c_2<x/k} \Lambda(c_1)\Lambda(c_2) \sum_{M<k\le 2M} e(k(c_1-c_2)\alpha) \end{aligned}$$

which in absolute value is

$$\begin{aligned}
&\leq \sum_{z<c_1,c_2<x/M} \Lambda(c_1)\Lambda(c_2) \left| \sum_{\substack{M<k\leq 2M \\ k\leq \min\{x/c_1, x/c_2\}}} e(k(c_1-c_2)\alpha) \right| \\
&\leq \sum_{z<c_1,c_2<x/M} \Lambda(c_1)\Lambda(c_2) \min\{\|\alpha(c_1-c_2)\|^{-1}, M\} \\
&\leq (\log x)^2 \frac{x}{M} \sum_{l<x/M} \min\{\|\alpha l\|\|^{-1}, M\}
\end{aligned}$$

where in the last inequality we used $l = c_1 - c_2$, so each l occurs $\leq x/M$ times and we also majorise $\Lambda(c)$ by $\log x$.

Now we need the following lemma

Lemma. Let $q, Q, R \in \mathbb{N}$ and $\alpha \in \mathbb{R}$ be such that $Q \geq 2$ and $|\alpha - a/q| \leq 1/q^2$. Then

$$\sum_{x=0}^{R} \min\{\frac{1}{\|\alpha x + \beta\|}, Q\} \ll \log Q \left(Q + q + R + \frac{QR}{q} \right).$$

Using this lemma, we get

$$\sum_{l\leq x/M} \min\{\frac{1}{\|\alpha l\|}, M\} \ll \log x \left(M + q + \frac{x}{M} + \frac{x}{q} \right).$$

Therefore we conclude that

$$U_M^2 \ll x \log^6 x \left(M + q + \frac{x}{M} + \frac{x}{q} \right).$$

Recall that $U = \sum_M U_M$ where M runs over dyadic partition of $(y, x/z)$, thus $M < x/z$ and $x/M < x/y$ and we get

$$U_M^2 \ll x \log^6 x \left(\frac{x}{z} + \frac{x}{y} + q + \frac{x}{q} \right).$$

Choosing $z = y$ we conclude that

$$U_M \ll \sqrt{x} \log^3 x \left(\sqrt{\frac{x}{y}} + \sqrt{q} + \sqrt{\frac{x}{q}} \right) = \log^3 x \left(\frac{x}{\sqrt{y}} + \sqrt{qx} + \frac{x}{\sqrt{q}} \right).$$

We get

$$U \ll \log^4 x \left(\frac{x}{\sqrt{y}} + \sqrt{qx} + \frac{x}{\sqrt{q}} \right).$$

Adding upper bounds for three subsums together we get

$$\begin{aligned}&\sum_{n\le x}\Lambda(n)e(n\alpha)\\ \ll\ &\log^3 x\left(q+y+\frac{x}{q}\right)+\log^3 x\left(q+y^2+\frac{x}{q}\right)+\log^4 x\left(\frac{x}{\sqrt{y}}+\sqrt{qx}+\frac{x}{\sqrt{q}}\right).\end{aligned}$$

We choose $y=x^{2/5}$ (to make $x/\sqrt{y}=y^2$) and conclude

$$\sum_{n\le x}\Lambda(n)e(n\alpha)\ll\log^4 x\left(x^{4/5}+\sqrt{xq}+\frac{x}{\sqrt{q}}\right)$$

proving the Proposition 1.

Remark 5.1. The following references are extensively used in this note. Since the list is short it is not necessary to pin-point which particular result came from which source. Interested readers are strongly encouraged to look into all of these valuable books and articles for the material covered in this notes and beyond.

Bibliography

[1] A. C. Cojocaru and M. Ram Murty, *An Introduction to Sieve methods and their applications,* London Math. Soc. Students Texts, Vol. **66**,(2001), Cambridge Uni. Press.

[2] H. davenport, *Multiplicative Number Theory*, GTM, 3rd Edition, Springer-Verlag, 2000.

[3] J. Friedlander and H. Iwaniec, *Opera de Cribro*, AMS Colloquium Publications, Vol. **53**. 2010.

[4] H. Halberstam and H. E. Richert, *Sieve Methods*,Courier Dover Publications, 2013.

[5] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, AMS Colloquium Publication, Vol. **53**, 2004.

[6] H. L. Montgomery, *The analytic principle of large sieve,* Bull. AMS, Vol. **84**, 4(1978), 547- 567.

[7] O. Ramare, *Arithmetical aspects of the large sieve inequality* Hindustan Book Agency, 2009.