On the stochasticity parameter of quadratic residues

Mikhail Gabdullin

Colloquium of the Steklov Mathematical Institute

Moscow, 5th of March

The plan

I. The stochasticity parameter

II. Quadratic residues

III. The stochasticity parameter of quadratic residues

The main definition: circle

Let $\mathbb{T}_n = \mathbb{R}/n\mathbb{Z}$ be the circle of length n. Let $k \in \mathbb{N}$ and $U = \{0 \leqslant u_1 < u_2 < \dots < u_k < n\}$ be a k-element subset of $\mathbb{T}_n$.

Denote by $s_i = u_{i+1} - u_i \in \mathbb{R}^+$, $i = 1, \dots, k$, consecutive distances between elements of U (we set $s_k = u_1 + n - u_k$).

Following V.I.Arnold, define the stochasticity parameter $S(U)$ of the set U to be

$$S(U) = \sum_{i=1}^{k} s_i^2.$$

Since $\frac{n}{k} = \frac{1}{k}\sum_{i=1}^{k} s_i \leqslant \left(\frac{1}{k}\sum_{i=1}^{k} s_i^2\right)^{1/2}$ and $\sum_{i=1}^{k} s_i^2 < \left(\sum_{i=1}^{k} s_i\right)^2 = n^2$, it follows that

$$\inf_{|U|=k} S(U) = \frac{n^2}{k}$$

and

$$\sup_{|U|=k} S(U) = n^2.$$

Let $\mathbb{T}_n = \mathbb{R}/n\mathbb{Z}$ be the circle of length n. Let $k \in \mathbb{N}$ and $U = \{0 \leqslant u_1 < u_2 < ... < u_k < n\}$ be a k-element subset of $\mathbb{T}_n$.

Denote by $s_i = u_{i+1} - u_i \in \mathbb{R}^+$, $i = 1, ..., k$, consecutive distances between elements of U (we set $s_k = u_1 + n - u_k$).

Following V.I.Arnold, define the stochasticity parameter $S(U)$ of the set U to be

$$S(U) = \sum_{i=1}^{k} s_i^2.$$

Since $\frac{n}{k} = \frac{1}{k}\sum_{i=1}^k s_i \leqslant \left(\frac{1}{k}\sum_{i=1}^k s_i^2\right)^{1/2}$ and $\sum_{i=1}^k s_i^2 < \left(\sum_{i=1}^k s_i\right)^2 = n^2$, it follows that

$$\inf_{|U|=k} S(U) = \frac{n^2}{k}$$

and

$$\sup_{|U|=k} S(U) = n^2.$$

Let $\mathbb{T}_n = \mathbb{R}/n\mathbb{Z}$ be the circle of length n. Let $k \in \mathbb{N}$ and $U = \{0 \leqslant u_1 < u_2 < ... < u_k < n\}$ be a k-element subset of $\mathbb{T}_n$.

Denote by $s_i = u_{i+1} - u_i \in \mathbb{R}^+$, $i = 1, ..., k$, consecutive distances between elements of U (we set $s_k = u_1 + n - u_k$).

Following V.I.Arnold, define the stochasticity parameter $S(U)$ of the set U to be

$$S(U) = \sum_{i=1}^{k} s_i^2.$$

Since $\frac{n}{k} = \frac{1}{k}\sum_{i=1}^{k} s_i \leqslant \left(\frac{1}{k}\sum_{i=1}^{k} s_i^2\right)^{1/2}$ and $\sum_{i=1}^{k} s_i^2 < \left(\sum_{i=1}^{k} s_i\right)^2 = n^2$, it follows that

$$\inf_{|U|=k} S(U) = \frac{n^2}{k}$$

and

$$\sup_{|U|=k} S(U) = n^2.$$

The main definition: circle

Let $\mathbb{T}_n = \mathbb{R}/n\mathbb{Z}$ be the circle of length n. Let $k \in \mathbb{N}$ and $U = \{0 \leqslant u_1 < u_2 < ... < u_k < n\}$ be a k-element subset of $\mathbb{T}_n$.

Denote by $s_i = u_{i+1} - u_i \in \mathbb{R}^+$, $i = 1, ..., k$, consecutive distances between elements of U (we set $s_k = u_1 + n - u_k$).

Following V.I.Arnold, define the stochasticity parameter $S(U)$ of the set U to be

$$S(U) = \sum_{i=1}^{k} s_i^2.$$

Since $\frac{n}{k} = \frac{1}{k}\sum_{i=1}^{k} s_i \leqslant \left(\frac{1}{k}\sum_{i=1}^{k} s_i^2\right)^{1/2}$ and $\sum_{i=1}^{k} s_i^2 < \left(\sum_{i=1}^{k} s_i\right)^2 = n^2$, it follows that

$$\inf_{|U|=k} S(U) = \frac{n^2}{k}$$

and

$$\sup_{|U|=k} S(U) = n^2.$$

Let $\mathbb{T}_n = \mathbb{R}/n\mathbb{Z}$ be the circle of length n. Let $k \in \mathbb{N}$ and $U = \{0 \leqslant u_1 < u_2 < ... < u_k < n\}$ be a k-element subset of $\mathbb{T}_n$.

Denote by $s_i = u_{i+1} - u_i \in \mathbb{R}^+$, $i = 1, ..., k$, consecutive distances between elements of U (we set $s_k = u_1 + n - u_k$).

Following V.I.Arnold, define the stochasticity parameter $S(U)$ of the set U to be

$$S(U) = \sum_{i=1}^{k} s_i^2.$$

Since $\frac{n}{k} = \frac{1}{k}\sum_{i=1}^{k} s_i \leqslant \left(\frac{1}{k}\sum_{i=1}^{k} s_i^2\right)^{1/2}$ and $\sum_{i=1}^{k} s_i^2 < \left(\sum_{i=1}^{k} s_i\right)^2 = n^2$, it follows that

$$\inf_{|U|=k} S(U) = \frac{n^2}{k}$$

and

$$\sup_{|U|=k} S(U) = n^2.$$

The main definition: circle

Let $\mathbb{T}_n = \mathbb{R}/n\mathbb{Z}$ be the circle of length n. Let $k \in \mathbb{N}$ and $U = \{0 \leqslant u_1 < u_2 < \ldots < u_k < n\}$ be a k-element subset of $\mathbb{T}_n$.

Denote by $s_i = u_{i+1} - u_i \in \mathbb{R}^+$, $i = 1, \ldots, k$, consecutive distances between elements of U (we set $s_k = u_1 + n - u_k$).

Following V.I.Arnold, define the stochasticity parameter $S(U)$ of the set U to be

$$S(U) = \sum_{i=1}^{k} s_i^2.$$

Since $\frac{n}{k} = \frac{1}{k}\sum_{i=1}^{k} s_i \leqslant \left(\frac{1}{k}\sum_{i=1}^{k} s_i^2\right)^{1/2}$ and $\sum_{i=1}^{k} s_i^2 < \left(\sum_{i=1}^{k} s_i\right)^2 = n^2$, it follows that

$$\inf_{|U|=k} S(U) = \frac{n^2}{k}$$

and

$$\sup_{|U|=k} S(U) = n^2.$$

The stochasticity parameter of a random set

Fix k and let U be a random subset of $\mathbb{T}_n$ (we throw k points on $\mathbb{T}_n$ uniformly at random). Then for each i and $t \in (0, n)$ we have

$$\mathbb{P}(s_i > t) = \mathbb{P}(s_1 > t) = \left(\frac{n-t}{n}\right)^{k-1}$$

and, hence,

$$\mathbb{E}s_i = \int_0^n \mathbb{P}(s_i > t)\, dt = \int_0^n (1 - t/n)^{k-1}\, dt = n \int_0^1 (1-v)^{k-1} dv = n \int_0^1 v^{k-1} dv = n/k$$

(as should be, because $\mathbb{E}(\sum_{i=1}^k s_i) = n$) and

$$\mathbb{E}s_i^2 = \int_0^n 2t \cdot \mathbb{P}(s_i > t)\, dt = \int_0^n 2t(1-t/n)^{k-1} dt = 2n^2 \int_0^1 v(1-v)^{k-1} dv = \frac{2n^2}{k(k+1)},$$

and, hence,

$$\mathbb{E}S(U) = \frac{2n^2}{k+1}.$$

Fix k and let U be a random subset of $\mathbb{T}_n$ (we throw k points on $\mathbb{T}_n$ uniformly at random). Then for each i and $t \in (0, n)$ we have

$$\mathbb{P}(s_i > t) = \mathbb{P}(s_1 > t) = \left(\frac{n-t}{n}\right)^{k-1}$$

and, hence,

$$\mathbb{E}s_i = \int_0^n \mathbb{P}(s_i > t)\,dt = \int_0^n (1 - t/n)^{k-1}\,dt = n\int_0^1 (1-v)^{k-1}dv = n\int_0^1 v^{k-1}dv = n/k$$

(as should be, because $\mathbb{E}(\sum_{i=1}^k s_i) = n$) and

$$\mathbb{E}s_i^2 = \int_0^n 2t{\cdot}\mathbb{P}(s_i > t)\,dt = \int_0^n 2t(1-t/n)^{k-1}dt = 2n^2\int_0^1 v(1-v)^{k-1}dv = \frac{2n^2}{k(k+1)},$$

and, hence,

$$\mathbb{E}S(U) = \frac{2n^2}{k+1}.$$

Fix k and let U be a random subset of $\mathbb{T}_n$ (we throw k points on $\mathbb{T}_n$ uniformly at random). Then for each i and $t \in (0,n)$ we have

$$\mathbb{P}\left(s_i > t\right) = \mathbb{P}\left(s_1 > t\right) = \left(\frac{n-t}{n}\right)^{k-1}$$

and, hence,

$$\mathbb{E}s_i = \int_0^n \mathbb{P}\left(s_i > t\right)dt = \int_0^n (1-t/n)^{k-1}\,dt = n\int_0^1 (1-v)^{k-1}dv = n\int_0^1 v^{k-1}dv = n/k$$

(as should be, because $\mathbb{E}(\sum_{i=1}^k s_i) = n$) and

$$\mathbb{E}s_i^2 = \int_0^n 2t\cdot\mathbb{P}\left(s_i > t\right)dt = \int_0^n 2t(1-t/n)^{k-1}dt = 2n^2\int_0^1 v(1-v)^{k-1}dv = \frac{2n^2}{k(k+1)},$$

and, hence,

$$\mathbb{E}S(U) = \frac{2n^2}{k+1}.$$

Fix k and let U be a random subset of $\mathbb{T}_n$ (we throw k points on $\mathbb{T}_n$ uniformly at random). Then for each i and $t \in (0, n)$ we have

$$\mathbb{P}(s_i > t) = \mathbb{P}(s_1 > t) = \left(\frac{n-t}{n}\right)^{k-1}$$

and, hence,

$$\mathbb{E}s_i = \int_0^n \mathbb{P}(s_i > t)\, dt = \int_0^n (1 - t/n)^{k-1}\, dt = n \int_0^1 (1-v)^{k-1} dv = n \int_0^1 v^{k-1} dv = n/k$$

(as should be, because $\mathbb{E}(\sum_{i=1}^k s_i) = n$) and

$$\mathbb{E}s_i^2 = \int_0^n 2t \cdot \mathbb{P}(s_i > t)\, dt = \int_0^n 2t(1-t/n)^{k-1} dt = 2n^2 \int_0^1 v(1-v)^{k-1} dv = \frac{2n^2}{k(k+1)},$$

and, hence,

$$\mathbb{E}S(U) = \frac{2n^2}{k+1}.$$

Fix k and let U be a random subset of $\mathbb{T}_n$ (we throw k points on $\mathbb{T}_n$ uniformly at random). Then for each i and $t \in (0, n)$ we have

$$\mathbb{P}(s_i > t) = \mathbb{P}(s_1 > t) = \left(\frac{n-t}{n}\right)^{k-1}$$

and, hence,

$$\mathbb{E}s_i = \int_0^n \mathbb{P}(s_i > t)\, dt = \int_0^n (1 - t/n)^{k-1}\, dt = n \int_0^1 (1-v)^{k-1} dv = n \int_0^1 v^{k-1} dv = n/k$$

(as should be, because $\mathbb{E}(\sum_{i=1}^k s_i) = n$) and

$$\mathbb{E}s_i^2 = \int_0^n 2t \cdot \mathbb{P}(s_i > t)\, dt = \int_0^n 2t(1-t/n)^{k-1} dt = 2n^2 \int_0^1 v(1-v)^{k-1} dv = \frac{2n^2}{k(k+1)},$$

and, hence,

$$\mathbb{E}S(U) = \frac{2n^2}{k+1}.$$

We see that

$$\mathbb{P}(s_i > t\mathbb{E}s_i) = \mathbb{P}(s_i > tn/k) = (1 - t/k)^{k-1} = e^{-t}(1 + o(1)), \quad k \to \infty,$$

uniformly for $0 \leqslant t \leqslant t_0$ for any fixed t_0, and so for large k the normalized gaps $\tilde{s}_i = s_i/\mathbb{E}s_i$ have the exponential distribution with parameter 1.

Hence, the function $N(t) = \#\{i : u_i \leqslant t\}$ behaves like the Poisson point process with constant rate 1 and the number $N(a, b)$ of points u_i in an interval $(a, b]$ has Poisson distribution with mean $b - a$.

We see that

$$\mathbb{P}(s_i > t\mathbb{E}s_i) = \mathbb{P}(s_i > tn/k) = (1 - t/k)^{k-1} = e^{-t}(1 + o(1)), \quad k \to \infty,$$

uniformly for $0 \leqslant t \leqslant t_0$ for any fixed t_0, and so for large k the normalized gaps $\tilde{s_i} = s_i/\mathbb{E}s_i$ have the exponential distribution with parameter 1.

Hence, the function $N(t) = \#\{i : u_i \leqslant t\}$ behaves like the Poisson point process with constant rate 1 and the number $N(a, b)$ of points u_i in an interval $(a, b]$ has Poisson distribution with mean $b - a$.

Let $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$. Let $k \in \mathbb{N}$ and $U = \{0 \leqslant u_1 < u_2 < ... < u_k < n\}$ be a k-element subset of $\mathbb{T}_n$.

Denote by $s_i = u_{i+1} - u_i \in \mathbb{N}$, $i = 1, ..., k$, consecutive distances between elements of U (we set $s_k = u_1 + n - u_k$).

As before, define the stochaticity parameter of the set U to be the quantity

$$S(U) = \sum_{i=1}^{k} s_i^2,$$

and again

$$\frac{n^2}{k} \leqslant S(U) \leqslant n^2.$$

The main definition: $\mathbb{Z}_n$

Let $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$. Let $k \in \mathbb{N}$ and $U = \{0 \leqslant u_1 < u_2 < \ldots < u_k < n\}$ be a k-element subset of $\mathbb{T}_n$.

Denote by $s_i = u_{i+1} - u_i \in \mathbb{N}$, $i = 1, \ldots, k$, consecutive distances between elements of U (we set $s_k = u_1 + n - u_k$).

As before, define the stochaticity parameter of the set U to be the quantity

$$S(U) = \sum_{i=1}^{k} s_i^2,$$

and again

$$\frac{n^2}{k} \leqslant S(U) \leqslant n^2.$$

The main definition: $\mathbb{Z}_n$

Let $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$. Let $k \in \mathbb{N}$ and $U = \{0 \leqslant u_1 < u_2 < \dots < u_k < n\}$ be a k-element subset of $\mathbb{T}_n$.

Denote by $s_i = u_{i+1} - u_i \in \mathbb{N}$, $i = 1, \dots, k$, consecutive distances between elements of U (we set $s_k = u_1 + n - u_k$).

As before, define the stochaticity parameter of the set U to be the quantity

$$S(U) = \sum_{i=1}^{k} s_i^2,$$

and again

$$\frac{n^2}{k} \leqslant S(U) \leqslant n^2.$$

Let $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$. Let $k \in \mathbb{N}$ and $U = \{0 \leqslant u_1 < u_2 < \dots < u_k < n\}$ be a k-element subset of $\mathbb{T}_n$.

Denote by $s_i = u_{i+1} - u_i \in \mathbb{N}$, $i = 1, \dots, k$, consecutive distances between elements of U (we set $s_k = u_1 + n - u_k$).

As before, define the stochaticity parameter of the set U to be the quantity

$$S(U) = \sum_{i=1}^{k} s_i^2,$$

and again

$$\frac{n^2}{k} \leqslant S(U) \leqslant n^2.$$

It is easy to see that $S(U)$ is minimal when s_i are equal (or close) to n/k, and is maximal when U is an interval of length k (so $\max_{|U|=k} S(U) = (n-k+1)^2 + k - 1$).

So too small or too large values of $S(U)$ indicate that U is far from a random set.

One can find the mean value $s(k) = s(M, k)$ of $S(U)$ over all k-element subsets of $\mathbb{Z}_M$.

Proposition 1. *We have*

$$s(k) = M\frac{2M-k+1}{k+1} .$$

Note that $s(k) \sim \frac{2M^2}{k+1}$ whenever $k = o(M)$.

(Recall that in the case $\mathbb{R}/n\mathbb{Z}$ we have $\mathbb{E}S(U) = \frac{2n^2}{k+1}$.)

It is easy to see that $S(U)$ is minimal when s_i are equal (or close) to n/k, and is maximal when U is an interval of length k (so $\max_{|U|=k} S(U) = (n-k+1)^2 + k - 1$).

So too small or too large values of $S(U)$ indicate that U is far from a random set.

One can find the mean value $s(k) = s(M,k)$ of $S(U)$ over all k-element subsets of $\mathbb{Z}_M$.

Proposition 1. *We have*

$$s(k) = M\frac{2M-k+1}{k+1} .$$

Note that $s(k) \sim \frac{2M^2}{k+1}$ whenever $k = o(M)$.

(Recall that in the case $\mathbb{R}/n\mathbb{Z}$ we have $\mathbb{E}S(U) = \frac{2n^2}{k+1}$.)

It is easy to see that $S(U)$ is minimal when s_i are equal (or close) to n/k, and is maximal when U is an interval of length k (so $\max_{|U|=k} S(U) = (n-k+1)^2 + k - 1$).

So too small or too large values of $S(U)$ indicate that U is far from a random set.

One can find the mean value $s(k) = s(M,k)$ of $S(U)$ over all k-element subsets of $\mathbb{Z}_M$.

Proposition 1. *We have*

$$s(k) = M\frac{2M-k+1}{k+1} \ .$$

Note that $s(k) \sim \frac{2M^2}{k+1}$ whenever $k = o(M)$.

(Recall that in the case $\mathbb{R}/n\mathbb{Z}$ we have $\mathbb{E}S(U) = \frac{2n^2}{k+1}$.)

It is easy to see that $S(U)$ is minimal when s_i are equal (or close) to n/k, and is maximal when U is an interval of length k (so $\max_{|U|=k} S(U) = (n-k+1)^2 + k - 1$).

So too small or too large values of $S(U)$ indicate that U is far from a random set.

One can find the mean value $s(k) = s(M, k)$ of $S(U)$ over all k-element subsets of $\mathbb{Z}_M$.

Proposition 1. *We have*

$$s(k) = M\frac{2M - k + 1}{k + 1} .$$

Note that $s(k) \sim \frac{2M^2}{k+1}$ whenever $k = o(M)$.

(Recall that in the case $\mathbb{R}/n\mathbb{Z}$ we have $\mathbb{E}S(U) = \frac{2n^2}{k+1}$.)

It is easy to see that $S(U)$ is minimal when s_i are equal (or close) to n/k, and is maximal when U is an interval of length k (so $\max_{|U|=k} S(U) = (n-k+1)^2 + k - 1$).

So too small or too large values of $S(U)$ indicate that U is far from a random set.

One can find the mean value $s(k) = s(M, k)$ of $S(U)$ over all k-element subsets of $\mathbb{Z}_M$.

Proposition 1. *We have*

$$s(k) = M\frac{2M-k+1}{k+1} \ .$$

Note that $s(k) \sim \frac{2M^2}{k+1}$ whenever $k = o(M)$.

(Recall that in the case $\mathbb{R}/n\mathbb{Z}$ we have $\mathbb{E}S(U) = \frac{2n^2}{k+1}$.)

It is easy to see that $S(U)$ is minimal when s_i are equal (or close) to n/k, and is maximal when U is an interval of length k (so $\max_{|U|=k} S(U) = (n-k+1)^2 + k - 1$).

So too small or too large values of $S(U)$ indicate that U is far from a random set.

One can find the mean value $s(k) = s(M, k)$ of $S(U)$ over all k-element subsets of $\mathbb{Z}_M$.

Proposition 1. *We have*

$$s(k) = M\frac{2M-k+1}{k+1} \ .$$

Note that $s(k) \sim \frac{2M^2}{k+1}$ whenever $k = o(M)$.

(Recall that in the case $\mathbb{R}/n\mathbb{Z}$ we have $\mathbb{E}S(U) = \frac{2n^2}{k+1}$.)

Quadratic residues

Recall that an element $x \in \mathbb{Z}_n$ is called a quadratic residue if there is $y \in \mathbb{Z}_n$ with $x = y^2$. Let R_n be the set of quadratic residues modulo n.

If $p \geqslant 2$ is a prime, then there are $(p+1)/2$ quadratic residues modulo p: they are exactly

$$0^2, 1^2, 2^2, ..., \left(\frac{p-1}{2}\right)^2$$

(since $k^2 = (p-k)^2$ and if $0 \leqslant a < b \leqslant (p-1)/2$, then $(a-b)(a+b) \neq 0$ in $\mathbb{Z}_p$).

Also it is easy to show that

$$\frac{p^{r-1}(p-1)}{2} \leqslant |R_{p^k}| \leqslant \frac{p^{r-1}(p+1)}{2}$$

The function $|R_n|$ is multiplicative, that is, $|R_{nm}| = |R_n||R_m|$ whenever $(n,m) = 1$ (because of the Chinese Remainder Theorem).

Quadratic residues

Recall that an element $x \in \mathbb{Z}_n$ is called a quadratic residue if there is $y \in \mathbb{Z}_n$ with $x = y^2$. Let R_n be the set of quadratic residues modulo n.

If $p \geqslant 2$ is a prime, then there are $(p+1)/2$ quadratic residues modulo p: they are exactly

$$0^2, 1^2, 2^2, ..., \left(\frac{p-1}{2}\right)^2$$

(since $k^2 = (p-k)^2$ and if $0 \leqslant a < b \leqslant (p-1)/2$, then $(a-b)(a+b) \neq 0$ in $\mathbb{Z}_p$).

Also it is easy to show that

$$\frac{p^{r-1}(p-1)}{2} \leqslant |R_{p^k}| \leqslant \frac{p^{r-1}(p+1)}{2}$$

The function $|R_n|$ is multiplicative, that is, $|R_{nm}| = |R_n||R_m|$ whenever $(n,m) = 1$ (because of the Chinese Remainder Theorem).

Recall that an element $x \in \mathbb{Z}_n$ is called a quadratic residue if there is $y \in \mathbb{Z}_n$ with $x = y^2$. Let R_n be the set of quadratic residues modulo n.

If $p \geqslant 2$ is a prime, then there are $(p+1)/2$ quadratic residues modulo p: they are exactly

$$0^2, 1^2, 2^2, ..., \left(\frac{p-1}{2}\right)^2$$

(since $k^2 = (p-k)^2$ and if $0 \leqslant a < b \leqslant (p-1)/2$, then $(a-b)(a+b) \neq 0$ in $\mathbb{Z}_p$).

Also it is easy to show that

$$\frac{p^{r-1}(p-1)}{2} \leqslant |R_{p^k}| \leqslant \frac{p^{r-1}(p+1)}{2}$$

The function $|R_n|$ is multiplicative, that is, $|R_{nm}| = |R_n||R_m|$ whenever $(n,m) = 1$ (because of the Chinese Remainder Theorem).

Quadratic residues

Recall that an element $x \in \mathbb{Z}_n$ is called a quadratic residue if there is $y \in \mathbb{Z}_n$ with $x = y^2$. Let R_n be the set of quadratic residues modulo n.

If $p \geqslant 2$ is a prime, then there are $(p+1)/2$ quadratic residues modulo p: they are exactly

$$0^2, 1^2, 2^2, ..., \left(\frac{p-1}{2}\right)^2$$

(since $k^2 = (p-k)^2$ and if $0 \leqslant a < b \leqslant (p-1)/2$, then $(a-b)(a+b) \neq 0$ in $\mathbb{Z}_p$).

Also it is easy to show that

$$\frac{p^{r-1}(p-1)}{2} \leqslant |R_{p^k}| \leqslant \frac{p^{r-1}(p+1)}{2}$$

The function $|R_n|$ is multiplicative, that is, $|R_{nm}| = |R_n||R_m|$ whenever $(n,m) = 1$ (because of the Chinese Remainder Theorem).

Quadratic residues

Denote by $\omega(n) = \sum_{p|n} 1$ the number of prime divisors of n.

We see that

$$\frac{n}{|R_n|} = \prod_{p|n} (2 + O(1/p))$$

and

$$\frac{n}{|R_n|} \to \infty \text{ if and only if } \omega(n) \to \infty.$$

Note that if n is taken from $[1, x] \in \mathbb{Z}$ uniformly at random, then

$$\mathbb{E}\omega(n) = \log\log x + O(1)$$

and

$$\operatorname{Var} \omega(n) \leqslant \log\log x + O(1).$$

Then by Chebyshev's inequality we obtain a theorem of Hardy and Ramanujan: if $f(x) \to \infty$ as $x \to \infty$, then

$$\omega(n) = \log\log x + O(f(x)\sqrt{\log\log x})$$

for all but $o(x)$ numbers $n \leqslant x$.

Denote by $\omega(n) = \sum_{p|n} 1$ the number of prime divisors of n.
We see that

$$\frac{n}{|R_n|} = \prod_{p|n} (2 + O(1/p))$$

and

$$\frac{n}{|R_n|} \to \infty \text{ if and only if } \omega(n) \to \infty.$$

Note that if n is taken from $[1, x] \in \mathbb{Z}$ uniformly at random, then

$$\mathbb{E}\omega(n) = \log\log x + O(1)$$

and

$$\operatorname{Var}\omega(n) \leqslant \log\log x + O(1).$$

Then by Chebyshev's inequality we obtain a theorem of Hardy and Ramanujan: if $f(x) \to \infty$ as $x \to \infty$, then

$$\omega(n) = \log\log x + O(f(x)\sqrt{\log\log x})$$

for all but $o(x)$ numbers $n \leqslant x$.

Quadratic residues

Denote by $\omega(n) = \sum_{p|n} 1$ the number of prime divisors of n.
We see that

$$\frac{n}{|R_n|} = \prod_{p|n} (2 + O(1/p))$$

and

$$\frac{n}{|R_n|} \to \infty \text{ if and only if } \omega(n) \to \infty.$$

Note that if n is taken from $[1, x] \in \mathbb{Z}$ uniformly at random, then

$$\mathbb{E}\omega(n) = \log\log x + O(1)$$

and

$$\operatorname{Var} \omega(n) \leqslant \log\log x + O(1).$$

Then by Chebyshev's inequality we obtain a theorem of Hardy and Ramanujan: if $f(x) \to \infty$ as $x \to \infty$, then

$$\omega(n) = \log\log x + O(f(x)\sqrt{\log\log x})$$

for all but $o(x)$ numbers $n \leqslant x$.

Denote by $\omega(n) = \sum_{p|n} 1$ the number of prime divisors of n.
We see that

$$\frac{n}{|R_n|} = \prod_{p|n} (2 + O(1/p))$$

and

$$\frac{n}{|R_n|} \to \infty \text{ if and only if } \omega(n) \to \infty.$$

Note that if n is taken from $[1, x] \in \mathbb{Z}$ uniformly at random, then

$$\mathbb{E}\omega(n) = \log\log x + O(1)$$

and

$$\operatorname{Var}\omega(n) \leqslant \log\log x + O(1).$$

Then by Chebyshev's inequality we obtain a theorem of Hardy and Ramanujan: if $f(x) \to \infty$ as $x \to \infty$, then

$$\omega(n) = \log\log x + O(f(x)\sqrt{\log\log x})$$

for all but $o(x)$ numbers $n \leqslant x$.

Denote by $\omega(n) = \sum_{p|n} 1$ the number of prime divisors of n.
We see that

$$\frac{n}{|R_n|} = \prod_{p|n} (2 + O(1/p))$$

and

$$\frac{n}{|R_n|} \to \infty \text{ if and only if } \omega(n) \to \infty.$$

Note that if n is taken from $[1, x] \in \mathbb{Z}$ uniformly at random, then

$$\mathbb{E}\omega(n) = \log\log x + O(1)$$

and

$$\operatorname{Var} \omega(n) \leqslant \log\log x + O(1).$$

Then by Chebyshev's inequality we obtain a theorem of Hardy and Ramanujan: if $f(x) \to \infty$ as $x \to \infty$, then

$$\omega(n) = \log\log x + O(f(x)\sqrt{\log\log x})$$

for all but $o(x)$ numbers $n \leqslant x$.

Denote by $\omega(n) = \sum_{p|n} 1$ the number of prime divisors of n.
We see that

$$\frac{n}{|R_n|} = \prod_{p|n} (2 + O(1/p))$$

and

$$\frac{n}{|R_n|} \to \infty \text{ if and only if } \omega(n) \to \infty.$$

Note that if n is taken from $[1, x] \in \mathbb{Z}$ uniformly at random, then

$$\mathbb{E}\omega(n) = \log\log x + O(1)$$

and

$$\operatorname{Var} \omega(n) \leqslant \log\log x + O(1).$$

Then by Chebyshev's inequality we obtain a theorem of Hardy and Ramanujan: if $f(x) \to \infty$ as $x \to \infty$, then

$$\omega(n) = \log\log x + O(f(x)\sqrt{\log\log x})$$

for all but $o(x)$ numbers $n \leqslant x$.

Denote by $\omega(n) = \sum_{p|n} 1$ the number of prime divisors of n.
We see that

$$\frac{n}{|R_n|} = \prod_{p|n} (2 + O(1/p))$$

and

$$\frac{n}{|R_n|} \to \infty \text{ if and only if } \omega(n) \to \infty.$$

Note that if n is taken from $[1, x] \in \mathbb{Z}$ uniformly at random, then

$$\mathbb{E}\omega(n) = \log\log x + O(1)$$

and

$$\operatorname{Var}\omega(n) \leqslant \log\log x + O(1).$$

Then by Chebyshev's inequality we obtain a theorem of Hardy and Ramanujan: if $f(x) \to \infty$ as $x \to \infty$, then

$$\omega(n) = \log\log x + O(f(x)\sqrt{\log\log x})$$

for all but $o(x)$ numbers $n \leqslant x$.

The stochasticity parameter of quadratic residues: $M = p$ is a prime

Let R_p be the set of quadratic residues modulo a prime p. A special case of result of M.Z.Garaev, S.V.Konyagin and Yu.V.Malykhin is the following.

Theorem (G.-K.-M., 2012). *Let $M = p$ be a prime. Then*

$$S(R_p) = s(|R_p|)(1 + o(1)), \quad p \to \infty.$$

So we can say that the set of quadratic residues behaves like a random set (of the same size) with respect to the stochasticity parameter.

We turn to this problem an arbitrary modulo M.

The stochasticity parameter of quadratic residues: $M = p$ is a prime

Let R_p be the set of quadratic residues modulo a prime p. A special case of result of M.Z.Garaev, S.V.Konyagin and Yu.V.Malykhin is the following.

Theorem (G.-K.-M., 2012). *Let $M = p$ be a prime. Then*

$$S(R_p) = s(|R_p|)(1 + o(1)), \quad p \to \infty.$$

So we can say that the set of quadratic residues behaves like a random set (of the same size) with respect to the stochasticity parameter.

We turn to this problem an arbitrary modulo M.

The stochasticity parameter of quadratic residues: $M = p$ is a prime

Let R_p be the set of quadratic residues modulo a prime p. A special case of result of M.Z.Garaev, S.V.Konyagin and Yu.V.Malykhin is the following.

Theorem (G.-K.-M., 2012). *Let $M = p$ be a prime. Then*

$$S(R_p) = s(|R_p|)(1 + o(1)), \quad p \to \infty.$$

So we can say that the set of quadratic residues behaves like a random set (of the same size) with respect to the stochasticity parameter.

We turn to this problem an arbitrary modulo M.

Let R_p be the set of quadratic residues modulo a prime p. A special case of result of M.Z.Garaev, S.V.Konyagin and Yu.V.Malykhin is the following.

Theorem (G.-K.-M., 2012). *Let $M = p$ be a prime. Then*

$$S(R_p) = s(|R_p|)(1 + o(1)), \quad p \to \infty.$$

So we can say that the set of quadratic residues behaves like a random set (of the same size) with respect to the stochasticity parameter.

We turn to this problem an arbitrary modulo M.

Let $R_M = \{0 = r_1 < r_2 < ... < r_{|R_M|}\}$ be the set of quadratic residues modulo M. As earlier, set $r_{|R_M|+1} := r_1 + M = M$.

Take an index j randomly and uniformly in $1, ..., |R_M|$. On average we of course have

$$\mathbb{E}(r_{j+1} - r_j) = \frac{M}{|R_M|}.$$

In 1999/2000 P.Kurlberg and Z.Rudnick found the limit distribution of spaces between quadratic residues.

Proposition 2. *We have*

$$\mathbb{P}\left(r_{j+1} - r_j > u\frac{M}{|R_M|}\right) = e^{-u}(1 + o(1)), \qquad \omega(M) \to \infty.$$

uniformly in the range $0 \leqslant u \leqslant u_0$ *for any fixed* u_0.

Let $R_M = \{0 = r_1 < r_2 < ... < r_{|R_M|}\}$ be the set of quadratic residues modulo M. As earlier, set $r_{|R_M|+1} := r_1 + M = M$.

Take an index j randomly and uniformly in $1, ..., |R_M|$. On average we of course have

$$\mathbb{E}(r_{j+1} - r_j) = \frac{M}{|R_M|}.$$

In 1999/2000 P.Kurlberg and Z.Rudnick found the limit distribution of spaces between quadratic residues.

Proposition 2. *We have*

$$\mathbb{P}\left(r_{j+1} - r_j > u\frac{M}{|R_M|}\right) = e^{-u}(1 + o(1)), \qquad \omega(M) \to \infty.$$

uniformly in the range $0 \leqslant u \leqslant u_0$ *for any fixed* u_0.

Let $R_M = \{0 = r_1 < r_2 < ... < r_{|R_M|}\}$ be the set of quadratic residues modulo M. As earlier, set $r_{|R_M|+1} := r_1 + M = M$.

Take an index j randomly and uniformly in $1, ..., |R_M|$. On average we of course have

$$\mathbb{E}(r_{j+1} - r_j) = \frac{M}{|R_M|}.$$

In 1999/2000 P.Kurlberg and Z.Rudnick found the limit distribution of spaces between quadratic residues.

Proposition 2. *We have*

$$\mathbb{P}\left(r_{j+1} - r_j > u\frac{M}{|R_M|}\right) = e^{-u}(1 + o(1)), \qquad \omega(M) \to \infty.$$

uniformly in the range $0 \leqslant u \leqslant u_0$ *for any fixed* u_0.

Now let $\omega(M) \to \infty$. The result of P.Kurlberg and Z.Rudnick supports the conjecture that

$$S(R_M) = \sum_{i=1}^{|R_M|} (r_{i+1} - r_i)^2 = \frac{M^2}{|R_M|} \mathbb{E}\left(\frac{r_{i+1} - r_i}{M/|R_M|}\right)^2 \sim \frac{M^2}{|R_M|} \int_0^\infty x^2 e^{-x} dx = \frac{2M^2}{|R_M|}$$

(however, we need good upper bounds for the contribution of large gaps between residues).

Also, if $\omega(M) \to \infty$, then $|R_M| \to \infty$ and $M/|R_M| \to \infty$, and hence, as we mentioned before,

$$s(|R_M|) = M \frac{2M - |R_M| + 1}{|R_M| + 1} \sim \frac{2M^2}{|R_M|}.$$

Recall Garaev, Konyagin, Malykhin proved $S(R_p) \sim s(|R_p|) = p\frac{2p-(p+1)/2+1}{(p+1)/2} \sim 3p$ (not $\frac{2p^2}{(p+1)/2)} \sim 4p$!)

So in general case $s(|R_M|)$ seems to be a better approximation for $S(R_M)$ than $\frac{2M^2}{|R_M|}$, because the latter quantity does not agree with the case where $\omega(M)$ is bounded.

Now let $\omega(M) \to \infty$. The result of P.Kurlberg and Z.Rudnick supports the conjecture that

$$S(R_M) = \sum_{i=1}^{|R_M|} (r_{i+1} - r_i)^2 = \frac{M^2}{|R_M|} \mathbb{E} \left(\frac{r_{i+1} - r_i}{M/|R_M|} \right)^2 \sim \frac{M^2}{|R_M|} \int_0^\infty x^2 e^{-x} dx = \frac{2M^2}{|R_M|}$$

(however, we need good upper bounds for the contribution of large gaps between residues).

Also, if $\omega(M) \to \infty$, then $|R_M| \to \infty$ and $M/|R_M| \to \infty$, and hence, as we mentioned before,

$$s(|R_M|) = M \frac{2M - |R_M| + 1}{|R_M| + 1} \sim \frac{2M^2}{|R_M|}.$$

Recall Garaev, Konyagin, Malykhin proved $S(R_p) \sim s(|R_p|) = p\frac{2p-(p+1)/2+1}{(p+1)/2} \sim 3p$ (not $\frac{2p^2}{(p+1)/2)} \sim 4p$!)

So in general case $s(|R_M|)$ seems to be a better approximation for $S(R_M)$ than $\frac{2M^2}{|R_M|}$, because the latter quantity does not agree with the case where $\omega(M)$ is bounded.

Now let $\omega(M) \to \infty$. The result of P.Kurlberg and Z.Rudnick supports the conjecture that

$$S(R_M) = \sum_{i=1}^{|R_M|} (r_{i+1} - r_i)^2 = \frac{M^2}{|R_M|} \mathbb{E} \left(\frac{r_{i+1} - r_i}{M/|R_M|} \right)^2 \sim \frac{M^2}{|R_M|} \int_0^\infty x^2 e^{-x} dx = \frac{2M^2}{|R_M|}$$

(however, we need good upper bounds for the contribution of large gaps between residues).

Also, if $\omega(M) \to \infty$, then $|R_M| \to \infty$ and $M/|R_M| \to \infty$, and hence, as we mentioned before,

$$s(|R_M|) = M \frac{2M - |R_M| + 1}{|R_M| + 1} \sim \frac{2M^2}{|R_M|}.$$

Recall Garaev, Konyagin, Malykhin proved $S(R_p) \sim s(|R_p|) = p\frac{2p-(p+1)/2+1}{(p+1)/2} \sim 3p$ (not $\frac{2p^2}{(p+1)/2)} \sim 4p$!)

So in general case $s(|R_M|)$ seems to be a better approximation for $S(R_M)$ than $\frac{2M^2}{|R_M|}$, because the latter quantity does not agree with the case where $\omega(M)$ is bounded.

Now let $\omega(M) \to \infty$. The result of P.Kurlberg and Z.Rudnick supports the conjecture that

$$S(R_M) = \sum_{i=1}^{|R_M|} (r_{i+1} - r_i)^2 = \frac{M^2}{|R_M|} \mathbb{E}\left(\frac{r_{i+1} - r_i}{M/|R_M|}\right)^2 \sim \frac{M^2}{|R_M|} \int_0^\infty x^2 e^{-x} dx = \frac{2M^2}{|R_M|}$$

(however, we need good upper bounds for the contribution of large gaps between residues).

Also, if $\omega(M) \to \infty$, then $|R_M| \to \infty$ and $M/|R_M| \to \infty$, and hence, as we mentioned before,

$$s(|R_M|) = M \frac{2M - |R_M| + 1}{|R_M| + 1} \sim \frac{2M^2}{|R_M|}.$$

Recall Garaev, Konyagin, Malykhin proved $S(R_p) \sim s(|R_p|) = p\frac{2p-(p+1)/2+1}{(p+1)/2} \sim 3p$ (not $\frac{2p^2}{(p+1)/2)} \sim 4p$!)

So in general case $s(|R_M|)$ seems to be a better approximation for $S(R_M)$ than $\frac{2M^2}{|R_M|}$, because the latter quantity does not agree with the case where $\omega(M)$ is bounded.

Now let $\omega(M) \to \infty$. The result of P.Kurlberg and Z.Rudnick supports the conjecture that

$$S(R_M) = \sum_{i=1}^{|R_M|} (r_{i+1} - r_i)^2 = \frac{M^2}{|R_M|} \mathbb{E} \left(\frac{r_{i+1} - r_i}{M/|R_M|} \right)^2 \sim \frac{M^2}{|R_M|} \int_0^\infty x^2 e^{-x} dx = \frac{2M^2}{|R_M|}$$

(however, we need good upper bounds for the contribution of large gaps between residues).

Also, if $\omega(M) \to \infty$, then $|R_M| \to \infty$ and $M/|R_M| \to \infty$, and hence, as we mentioned before,

$$s(|R_M|) = M \frac{2M - |R_M| + 1}{|R_M| + 1} \sim \frac{2M^2}{|R_M|}.$$

Recall Garaev, Konyagin, Malykhin proved $S(R_p) \sim s(|R_p|) = p \frac{2p-(p+1)/2+1}{(p+1)/2} \sim 3p$ (not $\frac{2p^2}{(p+1)/2)} \sim 4p$!)

So in general case $s(|R_M|)$ seems to be a better approximation for $S(R_M)$ than $\frac{2M^2}{|R_M|}$, because the latter quantity does not agree with the case where $\omega(M)$ is bounded.

Two heuristics

Now let $\omega(M) \to \infty$. The result of P.Kurlberg and Z.Rudnick supports the conjecture that

$$S(R_M) = \sum_{i=1}^{|R_M|} (r_{i+1} - r_i)^2 = \frac{M^2}{|R_M|} \mathbb{E} \left(\frac{r_{i+1} - r_i}{M/|R_M|} \right)^2 \sim \frac{M^2}{|R_M|} \int\limits_0^\infty x^2 e^{-x} dx = \frac{2M^2}{|R_M|}$$

(however, we need good upper bounds for the contribution of large gaps between residues).

Also, if $\omega(M) \to \infty$, then $|R_M| \to \infty$ and $M/|R_M| \to \infty$, and hence, as we mentioned before,

$$s(|R_M|) = M \frac{2M - |R_M| + 1}{|R_M| + 1} \sim \frac{2M^2}{|R_M|}.$$

Recall Garaev, Konyagin, Malykhin proved $S(R_p) \sim s(|R_p|) = p \frac{2p-(p+1)/2+1}{(p+1)/2} \sim 3p$ (not $\frac{2p^2}{(p+1)/2)} \sim 4p$!)

So in general case $s(|R_M|)$ seems to be a better approximation for $S(R_M)$ than $\frac{2M^2}{|R_M|}$, because the latter quantity does not agree with the case where $\omega(M)$ is bounded.

Conjecture: the first attempt

The first idea is to ask whether we have $S(R_M) \sim s(|R_M|)$ as $M \to \infty$.

It turns out to be false.

Theorem 1. *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

On the other hand, for these modulus M Proposition 1 gives us

$$s(|R_M|) = \left(\frac{4A^2}{|R_A|} - A\right)p + O_A(1).$$

Direct computations show that $2f_A(0.5) < \frac{4A^2}{|R_A|} - A$ for all $3 \leqslant A \leqslant 200$ except values $A = 89, 109, 178, 197$, for which $2f_A(0.5) > \frac{4A^2}{|R_A|} - A$.

The first idea is to ask whether we have $S(R_M) \sim s(|R_M|)$ as $M \to \infty$.

It turns out to be false.

Theorem 1. *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

On the other hand, for these modulus M Proposition 1 gives us

$$s(|R_M|) = \left(\frac{4A^2}{|R_A|} - A\right)p + O_A(1).$$

Direct computations show that $2f_A(0.5) < \frac{4A^2}{|R_A|} - A$ for all $3 \leqslant A \leqslant 200$ except values $A = 89, 109, 178, 197$, for which $2f_A(0.5) > \frac{4A^2}{|R_A|} - A$.

Conjecture: the first attempt

The first idea is to ask whether we have $S(R_M) \sim s(|R_M|)$ as $M \to \infty$.

It turns out to be false.

Theorem 1. *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

On the other hand, for these modulus M Proposition 1 gives us

$$s(|R_M|) = \left(\frac{4A^2}{|R_A|} - A\right)p + O_A(1).$$

Direct computations show that $2f_A(0.5) < \frac{4A^2}{|R_A|} - A$ for all $3 \leqslant A \leqslant 200$ except values $A = 89, 109, 178, 197$, for which $2f_A(0.5) > \frac{4A^2}{|R_A|} - A$.

The first idea is to ask whether we have $S(R_M) \sim s(|R_M|)$ as $M \to \infty$. It turns out to be false.

Theorem 1. *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

On the other hand, for these modulus M Proposition 1 gives us

$$s(|R_M|) = \left(\frac{4A^2}{|R_A|} - A\right) p + O_A(1).$$

Direct computations show that $2f_A(0.5) < \frac{4A^2}{|R_A|} - A$ for all $3 \leqslant A \leqslant 200$ except values $A = 89, 109, 178, 197$, for which $2f_A(0.5) > \frac{4A^2}{|R_A|} - A$.

The first idea is to ask whether we have $S(R_M) \sim s(|R_M|)$ as $M \to \infty$. It turns out to be false.

Theorem 1. *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

On the other hand, for these modulus M Proposition 1 gives us

$$s(|R_M|) = \left(\frac{4A^2}{|R_A|} - A\right)p + O_A(1).$$

Direct computations show that $2f_A(0.5) < \frac{4A^2}{|R_A|} - A$ for all $3 \leqslant A \leqslant 200$ except values $A = 89, 109, 178, 197$, for which $2f_A(0.5) > \frac{4A^2}{|R_A|} - A$.

Corollary. *We have*

$$\underline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} < 1 < \overline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)}$$

and the conjecture does not hold in general.

Since $S(R_M) \geqslant \frac{M^2}{|R_M|}$ and $s(|R_M|) = M\frac{2M-|R_M|+1}{|R_M|+1} \leqslant \frac{2M^2}{|R_M|}$, we have

$$\underline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} \geqslant 0.5\,.$$

But is it true that

$$\overline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} < \infty \quad ?$$

We do not know, but

Theorem (Aryan, 2015) *For a square-free M we have*

$$S(R_M) \ll M2^{\omega(M)} \log M \prod_{p|M} \left(1+\frac{1}{\sqrt{p}}\right)\left(1-\frac{1}{p}\right).$$

Corollary. *We have*

$$\underline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} < 1 < \overline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)}$$

and the conjecture does not hold in general.

Since $S(R_M) \geqslant \frac{M^2}{|R_M|}$ and $s(|R_M|) = M\frac{2M-|R_M|+1}{|R_M|+1} \leqslant \frac{2M^2}{|R_M|}$, we have

$$\lim_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} \geqslant 0.5\,.$$

But is it true that

$$\overline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} < \infty \quad ?$$

We do not know, but

Theorem (Aryan, 2015) *For a square-free M we have*

$$S(R_M) \ll M2^{\omega(M)} \log M \prod_{p|M} \left(1+\frac{1}{\sqrt{p}}\right)\left(1-\frac{1}{p}\right).$$

The bottom line

Corollary. *We have*

$$\varliminf_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} < 1 < \varlimsup_{M\to\infty} \frac{S(R_M)}{s(|R_M|)}$$

and the conjecture does not hold in general.

Since $S(R_M) \geqslant \frac{M^2}{|R_M|}$ and $s(|R_M|) = M\frac{2M-|R_M|+1}{|R_M|+1} \leqslant \frac{2M^2}{|R_M|}$, we have

$$\varliminf_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} \geqslant 0.5\,.$$

But is it true that

$$\varlimsup_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} < \infty \quad ?$$

We do not know, but

Theorem (Aryan, 2015) *For a square-free M we have*

$$S(R_M) \ll M2^{\omega(M)} \log M \prod_{p|M} \left(1+\frac{1}{\sqrt{p}}\right)\left(1-\frac{1}{p}\right).$$

Corollary. *We have*

$$\underline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} < 1 < \overline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)}$$

and the conjecture does not hold in general.

Since $S(R_M) \geqslant \frac{M^2}{|R_M|}$ and $s(|R_M|) = M\,\frac{2M-|R_M|+1}{|R_M|+1} \leqslant \frac{2M^2}{|R_M|}$, we have

$$\underline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} \geqslant 0.5\,.$$

But is it true that

$$\overline{\lim}_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} < \infty \quad ?$$

We do not know, but

Theorem (Aryan, 2015) *For a square-free M we have*

$$S(R_M) \ll M2^{\omega(M)} \log M \prod_{p|M} \left(1 + \frac{1}{\sqrt{p}}\right)\left(1 - \frac{1}{p}\right).$$

Corollary. *We have*

$$\lim_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} < 1 < \overline{\lim_{M\to\infty}} \frac{S(R_M)}{s(|R_M|)}$$

and the conjecture does not hold in general.

Since $S(R_M) \geqslant \frac{M^2}{|R_M|}$ and $s(|R_M|) = M\frac{2M-|R_M|+1}{|R_M|+1} \leqslant \frac{2M^2}{|R_M|}$, we have

$$\lim_{M\to\infty} \frac{S(R_M)}{s(|R_M|)} \geqslant 0.5\,.$$

But is it true that

$$\overline{\lim_{M\to\infty}} \frac{S(R_M)}{s(|R_M|)} < \infty \quad ?$$

We do not know, but

Theorem (Aryan, 2015) *For a square-free M we have*

$$S(R_M) \ll M2^{\omega(M)} \log M \prod_{p|M} \left(1+\frac{1}{\sqrt{p}}\right)\left(1-\frac{1}{p}\right).$$

$$f_1(y) = 1 + y$$

$$f_3(y) = \frac{5y^2 + 8y + 5}{1 + y}$$

$$f_4(y) = \frac{10y^2 + 12y + 10}{1 + y}$$

$$f_5(y) = \frac{11y^3 + 14y^2 + 14y + 11}{1 + y + y^2}$$

$$f_7(y) = \frac{15y^4 + 24y^3 + 20y^2 + 24y + 15}{1 + y + y^2 + y^3}$$

$$f_8(y) = \frac{26y^3 + 38y^2 + 38y + 26}{1 + y + y^2}$$

Examples of f_A

$$f_1(y) = 1 + y$$

$$f_3(y) = \frac{5y^2 + 8y + 5}{1 + y}$$

$$f_4(y) = \frac{10y^2 + 12y + 10}{1 + y}$$

$$f_5(y) = \frac{11y^3 + 14y^2 + 14y + 11}{1 + y + y^2}$$

$$f_7(y) = \frac{15y^4 + 24y^3 + 20y^2 + 24y + 15}{1 + y + y^2 + y^3}$$

$$f_8(y) = \frac{26y^3 + 38y^2 + 38y + 26}{1 + y + y^2}$$

Examples of f_A

$$f_1(y) = 1 + y$$

$$f_3(y) = \frac{5y^2 + 8y + 5}{1 + y}$$

$$f_4(y) = \frac{10y^2 + 12y + 10}{1 + y}$$

$$f_5(y) = \frac{11y^3 + 14y^2 + 14y + 11}{1 + y + y^2}$$

$$f_7(y) = \frac{15y^4 + 24y^3 + 20y^2 + 24y + 15}{1 + y + y^2 + y^3}$$

$$f_8(y) = \frac{26y^3 + 38y^2 + 38y + 26}{1 + y + y^2}$$

Examples of f_A

$$f_1(y) = 1 + y$$

$$f_3(y) = \frac{5y^2 + 8y + 5}{1 + y}$$

$$f_4(y) = \frac{10y^2 + 12y + 10}{1 + y}$$

$$f_5(y) = \frac{11y^3 + 14y^2 + 14y + 11}{1 + y + y^2}$$

$$f_7(y) = \frac{15y^4 + 24y^3 + 20y^2 + 24y + 15}{1 + y + y^2 + y^3}$$

$$f_8(y) = \frac{26y^3 + 38y^2 + 38y + 26}{1 + y + y^2}$$

Examples of f_A

$$f_1(y) = 1 + y$$

$$f_3(y) = \frac{5y^2 + 8y + 5}{1 + y}$$

$$f_4(y) = \frac{10y^2 + 12y + 10}{1 + y}$$

$$f_5(y) = \frac{11y^3 + 14y^2 + 14y + 11}{1 + y + y^2}$$

$$f_7(y) = \frac{15y^4 + 24y^3 + 20y^2 + 24y + 15}{1 + y + y^2 + y^3}$$

$$f_8(y) = \frac{26y^3 + 38y^2 + 38y + 26}{1 + y + y^2}$$

Examples of f_A

$$f_1(y) = 1 + y$$

$$f_3(y) = \frac{5y^2 + 8y + 5}{1 + y}$$

$$f_4(y) = \frac{10y^2 + 12y + 10}{1 + y}$$

$$f_5(y) = \frac{11y^3 + 14y^2 + 14y + 11}{1 + y + y^2}$$

$$f_7(y) = \frac{15y^4 + 24y^3 + 20y^2 + 24y + 15}{1 + y + y^2 + y^3}$$

$$f_8(y) = \frac{26y^3 + 38y^2 + 38y + 26}{1 + y + y^2}$$

$$f_{11}(y) = \frac{27y^6 + 38y^5 + 34y^4 + 44y^3 + 34y^2 + 38y + 27}{1 + y + y^2 + y^3 + y^4 + y^5}$$

$$f_{13}(y) = \frac{37y^7 + 38y^6 + 54y^5 + 40y^4 + 40y^3 + 54y^2 + 38y + 37}{1 + y + y^2 + y^3 + y^4 + y^5 + y^6}$$

Let $\{s_1, ..., s_{|R_A|}\}$ be consecutive distances between quadratic residues modulo A. In fact,

$$f_A(y) = \frac{F(y)}{Q(y)}$$

where $Q(y) = Q_A(y) = 1 + y + \ldots + y^{|R_A|-1}$ and

$$F(y) = F_A(y) = \sum_{k=0}^{|R_A|} \beta_k y^k$$

is the reciprocal polynomial with the coefficients $\beta_0 = \beta_{|R_A|} = \sum_i s_i^2 = S(R_A)$ and $\beta_k = 2\sum_{i=1}^{|R_A|} s_i s_{i+k}$ for $0 < k < |R_A|$ (we think of indices i as elements of $\mathbb{Z}_{|R_A|}$).

Let $\{s_1, ..., s_{|R_A|}\}$ be consecutive distances between quadratic residues modulo A. In fact,

$$f_A(y) = \frac{F(y)}{Q(y)}$$

where $Q(y) = Q_A(y) = 1 + y + \ldots + y^{|R_A|-1}$ and

$$F(y) = F_A(y) = \sum_{k=0}^{|R_A|} \beta_k y^k$$

is the reciprocal polynomial with the coefficients $\beta_0 = \beta_{|R_A|} = \sum_i s_i^2 = S(R_A)$ and $\beta_k = 2\sum_{i=1}^{|R_A|} s_i s_{i+k}$ for $0 < k < |R_A|$ (we think of indices i as elements of $\mathbb{Z}_{|R_A|}$).

Refined conjecture: believed for almost all moduli

So, a more accurate question is the following. Do we have

$$S(R_M) \sim s(|R_M|), \quad M \to \infty,$$

for almost M ?

It was proved in an unpublished work of Konyagin.

So, a more accurate question is the following. Do we have

$$S(R_M) \sim s(|R_M|), \quad M \to \infty,$$

for almost M ?

It was proved in an unpublished work of Konyagin.

Let us turn to the second term in the asymptotics.

Theorem 1 (once again). *There exists absolute constant* $c > 0$ *such that for any fixed* A *and* $M = Ap$ *we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4p^{1-c})$$

where f_A *is a function determined by the number* A.

Why only one large prime factor? Why only a fixed A?

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of all positive integers M such that $M = Am$, where
(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;
(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \ldots < p_t$ are primes greater than $2^{C_0 t}$,

Let us turn to the second term in the asymptotics.

Theorem 1 (once again). *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

Why only one large prime factor? Why only a fixed A?

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of all positive integers M such that $M = Am$, where
(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;
(ii) $m = p_1 \ldots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \ldots < p_t$ are primes greater than $2^{C_0 t}$,

Let us turn to the second term in the asymptotics.

Theorem 1 (once again). *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

Why only one large prime factor? Why only a fixed A?

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of all positive integers M such that $M = Am$, where

(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;

(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \dots < p_t$ are primes greater than $2^{C_0 t}$,

Let us turn to the second term in the asymptotics.

Theorem 1 (once again). *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

Why only one large prime factor? Why only a fixed A?

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of all positive integers M such that $M = Am$, where

(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;

(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \dots < p_t$ are primes greater than $2^{C_0 t}$,

Let us turn to the second term in the asymptotics.

Theorem 1 (once again). *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

Why only one large prime factor? Why only a fixed A?

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of all positive integers M such that $M = Am$, where

(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;

(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \dots < p_t$ are primes greater than $2^{C_0 t}$,

Let us turn to the second term in the asymptotics.

Theorem 1 (once again). *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

Why only one large prime factor? Why only a fixed A?

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of all positive integers M such that $M = Am$, where

(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;
(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \dots < p_t$ are primes greater than $2^{C_0 t}$,

Let us turn to the second term in the asymptotics.

Theorem 1 (once again). *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

Why only one large prime factor? Why only a fixed A?

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of all positive integers M such that $M = Am$, where

(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;

(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \dots < p_t$ are primes greater than $2^{C_0 t}$,

Let us turn to the second term in the asymptotics.

Theorem 1 (once again). *There exists absolute constant $c > 0$ such that for any fixed A and $M = Ap$ we have*

$$S(R_M) = 2f_A(0.5)p + O(A^4 p^{1-c})$$

where f_A is a function determined by the number A.

Why only one large prime factor? Why only a fixed A?

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of all positive integers M such that $M = Am$, where

(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;
(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \dots < p_t$ are primes greater than $2^{C_0 t}$,

Our main result is the following.

Theorem 2. *There exists absolute constant $c > 0$ such that for $M \in \Omega$ we have*

$$S(R_M) = m2^{t+1}A^2|R_A|^{-1} - A^2|R_A|^{-1}m + E,$$

where

$$E \ll m2^{3t}A^4p_1^{-c} + mA^2|R_A|2^{-t} = o(m), \quad M \to \infty, M \in \Omega.$$

Moreover, the set Ω has positive lower density.

(Let us note that $m2^{t+1}A^2|R_A|^{-1} = 2M \cdot \frac{A}{|R_A|}2^t \sim \frac{2M^2}{|R_M|}$).

On the other hand, for $M \in \Omega$ Proposition 1 gives us

$$s(|R_M|) = m2^{t+1}A^2|R_A|^{-1} - Am + O(A^2|R_A|^{-1}m2^{2t}p_1^{-1}).$$

Our main result is the following.

Theorem 2. *There exists absolute constant* $c > 0$ *such that for* $M \in \Omega$ *we have*

$$S(R_M) = m2^{t+1}A^2|R_A|^{-1} - A^2|R_A|^{-1}m + E,$$

where

$$E \ll m2^{3t}A^4p_1^{-c} + mA^2|R_A|2^{-t} = o(m), \quad M \to \infty, M \in \Omega.$$

Moreover, the set Ω *has positive lower density.*

(Let us note that $m2^{t+1}A^2|R_A|^{-1} = 2M \cdot \frac{A}{|R_A|}2^t \sim \frac{2M^2}{|R_M|}$).

On the other hand, for $M \in \Omega$ Proposition 1 gives us

$$s(|R_M|) = m2^{t+1}A^2|R_A|^{-1} - Am + O(A^2|R_A|^{-1}m2^{2t}p_1^{-1}).$$

Our main result is the following.

Theorem 2. *There exists absolute constant* $c > 0$ *such that for* $M \in \Omega$ *we have*

$$S(R_M) = m2^{t+1}A^2|R_A|^{-1} - A^2|R_A|^{-1}m + E,$$

where

$$E \ll m2^{3t}A^4p_1^{-c} + mA^2|R_A|2^{-t} = o(m), \quad M \to \infty, M \in \Omega.$$

Moreover, the set Ω *has positive lower density.*

(Let us note that $m2^{t+1}A^2|R_A|^{-1} = 2M \cdot \frac{A}{|R_A|}2^t \sim \frac{2M^2}{|R_M|}$).

On the other hand, for $M \in \Omega$ Proposition 1 gives us

$$s(|R_M|) = m2^{t+1}A^2|R_A|^{-1} - Am + O(A^2|R_A|^{-1}m2^{2t}p_1^{-1}).$$

Our main result is the following.

Theorem 2. *There exists absolute constant $c > 0$ such that for $M \in \Omega$ we have*

$$S(R_M) = m2^{t+1}A^2|R_A|^{-1} - A^2|R_A|^{-1}m + E,$$

where

$$E \ll m2^{3t}A^4p_1^{-c} + mA^2|R_A|2^{-t} = o(m), \quad M \to \infty, M \in \Omega.$$

Moreover, the set Ω has positive lower density.

(Let us note that $m2^{t+1}A^2|R_A|^{-1} = 2M \cdot \frac{A}{|R_A|}2^t \sim \frac{2M^2}{|R_M|}$).

On the other hand, for $M \in \Omega$ Proposition 1 gives us

$$s(|R_M|) = m2^{t+1}A^2|R_A|^{-1} - Am + O(A^2|R_A|^{-1}m2^{2t}p_1^{-1}).$$

Our main result is the following.

Theorem 2. *There exists absolute constant* $c > 0$ *such that for* $M \in \Omega$ *we have*

$$S(R_M) = m2^{t+1}A^2|R_A|^{-1} - A^2|R_A|^{-1}m + E,$$

where

$$E \ll m2^{3t}A^4p_1^{-c} + mA^2|R_A|2^{-t} = o(m), \quad M \to \infty, M \in \Omega.$$

Moreover, the set Ω *has positive lower density.*

(Let us note that $m2^{t+1}A^2|R_A|^{-1} = 2M \cdot \frac{A}{|R_A|}2^t \sim \frac{2M^2}{|R_M|}$).

On the other hand, for $M \in \Omega$ Proposition 1 gives us

$$s(|R_M|) = m2^{t+1}A^2|R_A|^{-1} - Am + O(A^2|R_A|^{-1}m2^{2t}p_1^{-1}).$$

Corollary. *We have*

$$S(R_M) = s(|R_M|)(1 + o(1)), \quad M \to \infty, M \in \Omega.$$

It is a generalization of the mentioned result of Garaev, Konyagin, Malykhin.

Corollary (weak repulsion) *For all sufficiently large $M \in \Omega$ with $A \geqslant 3$ we have*

$$S(R_M) < s(|R_M|).$$

Corollary. *We have*

$$S(R_M) = s(|R_M|)(1 + o(1)), \quad M \to \infty, \, M \in \Omega.$$

It is a generalization of the mentioned result of Garaev, Konyagin, Malykhin.

Corollary (weak repulsion) *For all sufficiently large $M \in \Omega$ with $A \geqslant 3$ we have*

$$S(R_M) < s(|R_M|).$$

Corollary. *We have*

$$S(R_M) = s(|R_M|)(1 + o(1)), \quad M \to \infty, \, M \in \Omega.$$

It is a generalization of the mentioned result of Garaev, Konyagin, Malykhin.

Corollary (weak repulsion) *For all sufficiently large* $M \in \Omega$ *with* $A \geqslant 3$ *we have*

$$S(R_M) < s(|R_M|).$$

Gravitation Conjecture. *The set*

$$\{M \in \mathbb{N} : S(R_M) > s(|R_M|)\}$$

also has positive lower density.

It seems that we are unable to prove this using our method.

Gravitation Conjecture. *The set*

$$\{M \in \mathbb{N} : S(R_M) > s(|R_M|)\}$$

also has positive lower density.

It seems that we are unable to prove this using our method.

We can write

$$S(R_M) = \sum_{l \geqslant 1} N_l l^2,$$

where

$$N_l = \#\{x \in \mathbb{Z}_M : x, x+l \in R_M, x+1, \ldots, x+l-1 \notin R_M\}.$$

Let $M = p$ be a prime for simplicity. Consider the Legendre symbol $\left(\frac{\cdot}{p}\right) : \mathbb{Z}_p \to \mathbb{C}$, defined by

$$\left(\frac{n}{p}\right) = \begin{cases} 0, & n = 0; \\ 1, & n \text{ is a quadratic residue;} \\ -1, & n \text{ is a quadratic nonresidue.} \end{cases}$$

It is a character $\bmod p$, that is, a homomorphism between $\mathbb{Z}_p^*$ and $\mathbb{C}^*$.

Denote the Legendre symbol $\bmod p$ by χ_p for the brevity. Our benefit is that

$$R_p(x) = \frac{1}{2}(1 + \chi_p(x) + 1(x = 0))$$

and

$$1 - R_p(x) = \frac{1}{2}(1 - \chi_p(x) - 1(x = 0)).$$

We can write

$$S(R_M) = \sum_{l \geqslant 1} N_l l^2,$$

where

$$N_l = \#\{x \in \mathbb{Z}_M : x, x+l \in R_M, x+1, \dots, x+l-1 \notin R_M\}.$$

Let $M = p$ be a prime for simplicity. Consider the Legendre symbol $\left(\frac{\cdot}{p}\right) : \mathbb{Z}_p \to \mathbb{C}$, defined by

$$\left(\frac{n}{p}\right) = \begin{cases} 0, & n = 0; \\ 1, & n \text{ is a quadratic residue;} \\ -1, & n \text{ is a quadratic nonresidue.} \end{cases}$$

It is a character $\bmod p$, that is, a homomorphism between $\mathbb{Z}_p^*$ and $\mathbb{C}^*$.

Denote the Legendre symbol $\bmod p$ by χ_p for the brevity. Our benefit is that

$$R_p(x) = \frac{1}{2}(1 + \chi_p(x) + 1(x = 0))$$

and

$$1 - R_p(x) = \frac{1}{2}(1 - \chi_p(x) - 1(x = 0)).$$

We can write

$$S(R_M) = \sum_{l \geqslant 1} N_l l^2,$$

where

$$N_l = \#\{x \in \mathbb{Z}_M : x, x+l \in R_M, x+1, \ldots, x+l-1 \notin R_M\}.$$

Let $M = p$ be a prime for simplicity. Consider the Legendre symbol $\left(\frac{\cdot}{p}\right) : \mathbb{Z}_p \to \mathbb{C}$, defined by

$$\left(\frac{n}{p}\right) = \begin{cases} 0, & n = 0; \\ 1, & n \text{ is a quadratic residue;} \\ -1, & n \text{ is a quadratic nonresidue.} \end{cases}$$

It is a character $\bmod p$, that is, a homomorphism between $\mathbb{Z}_p^*$ and $\mathbb{C}^*$.

Denote the Legendre symbol $\bmod p$ by χ_p for the brevity. Our benefit is that

$$R_p(x) = \frac{1}{2}(1 + \chi_p(x) + 1(x = 0))$$

and

$$1 - R_p(x) = \frac{1}{2}(1 - \chi_p(x) - 1(x = 0)).$$

We can write

$$S(R_M) = \sum_{l \geqslant 1} N_l l^2,$$

where

$$N_l = \#\{x \in \mathbb{Z}_M : x, x+l \in R_M, x+1, \dots, x+l-1 \notin R_M\}.$$

Let $M = p$ be a prime for simplicity. Consider the Legendre symbol $\left(\frac{\cdot}{p}\right) : \mathbb{Z}_p \to \mathbb{C}$, defined by

$$\left(\frac{n}{p}\right) = \begin{cases} 0, & n = 0; \\ 1, & n \text{ is a quadratic residue;} \\ -1, & n \text{ is a quadratic nonresidue.} \end{cases}$$

It is a character $\bmod p$, that is, a homomorphism between $\mathbb{Z}_p^*$ and $\mathbb{C}^*$.

Denote the Legendre symbol $\bmod p$ by χ_p for the brevity. Our benefit is that

$$R_p(x) = \frac{1}{2}(1 + \chi_p(x) + 1(x = 0))$$

and

$$1 - R_p(x) = \frac{1}{2}(1 - \chi_p(x) - 1(x = 0)).$$

We can write

$$S(R_M) = \sum_{l \geqslant 1} N_l l^2,$$

where

$$N_l = \#\{x \in \mathbb{Z}_M : x, x+l \in R_M, x+1, \dots, x+l-1 \notin R_M\}.$$

Let $M = p$ be a prime for simplicity. Consider the Legendre symbol $\left(\frac{\cdot}{p}\right) : \mathbb{Z}_p \to \mathbb{C}$, defined by

$$\left(\frac{n}{p}\right) = \begin{cases} 0, & n = 0; \\ 1, & n \text{ is a quadratic residue}; \\ -1, & n \text{ is a quadratic nonresidue}. \end{cases}$$

It is a character $\bmod p$, that is, a homomorphism between $\mathbb{Z}_p^*$ and $\mathbb{C}^*$.

Denote the Legendre symbol $\bmod p$ by χ_p for the brevity. Our benefit is that

$$R_p(x) = \frac{1}{2}(1 + \chi_p(x) + 1(x = 0))$$

and

$$1 - R_p(x) = \frac{1}{2}(1 - \chi_p(x) - 1(x = 0)).$$

We can write

$$S(R_M) = \sum_{l \geqslant 1} N_l l^2,$$

where

$$N_l = \#\{x \in \mathbb{Z}_M : x, x+l \in R_M, x+1, \ldots, x+l-1 \notin R_M\}.$$

Let $M = p$ be a prime for simplicity. Consider the Legendre symbol $\left(\frac{\cdot}{p}\right) : \mathbb{Z}_p \to \mathbb{C}$, defined by

$$\left(\frac{n}{p}\right) = \begin{cases} 0, & n = 0; \\ 1, & n \text{ is a quadratic residue;} \\ -1, & n \text{ is a quadratic nonresidue.} \end{cases}$$

It is a character $\bmod p$, that is, a homomorphism between $\mathbb{Z}_p^*$ and $\mathbb{C}^*$.

Denote the Legendre symbol $\bmod p$ by χ_p for the brevity. Our benefit is that

$$R_p(x) = \frac{1}{2}(1 + \chi_p(x) + 1(x = 0))$$

and

$$1 - R_p(x) = \frac{1}{2}(1 - \chi_p(x) - 1(x = 0)).$$

$$N_l = \sum_{x\in\mathbb{Z}_p} R_p(x)R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)+1(x=0)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) + O(1) =$$

$$2^{-l}\sum_{x\in\mathbb{Z}_p}(1+\chi_p(x))(1+\chi_p(x+l))\prod_{i=1}^{l-1}(1-\chi_p(x+i)) + O(l).$$

Here we have the main term $p2^{-l}$, the error term $O(l)$ and 2^l-1 character sums of the type $\sum_{x\in\mathbb{Z}_p}\chi_p(x+a_1)...\chi_p(x+a_r)$ with distinct $a_1,...,a_r$. Such a sum is estimated by $rp^{1/2}$ in magnitude (famous Weil's theorem).

Hence,

$$N_l = p2^{-l} + O(l2^l p^{1/2})$$

and it is the asymptotics for N_l if $l \ll \log p$.

$$N_l = \sum_{x\in\mathbb{Z}_p} R_p(x)R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)+1(x=0)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) + O(1) =$$

$$2^{-l}\sum_{x\in\mathbb{Z}_p}(1+\chi_p(x))(1+\chi_p(x+l))\prod_{i=1}^{l-1}(1-\chi_p(x+i)) + O(l).$$

Here we have the main term $p2^{-l}$, the error term $O(l)$ and 2^l-1 character sums of the type $\sum_{x\in\mathbb{Z}_p}\chi_p(x+a_1)...\chi_p(x+a_r)$ with distinct $a_1,...,a_r$. Such a sum is estimated by $rp^{1/2}$ in magnitude (famous Weil's theorem).

Hence,

$$N_l = p2^{-l} + O(l2^l p^{1/2})$$

and it is the asymptotics for N_l if $l \ll \log p$.

$$N_l = \sum_{x\in\mathbb{Z}_p} R_p(x)R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)+1(x=0)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) + O(1) =$$

$$2^{-l}\sum_{x\in\mathbb{Z}_p}(1+\chi_p(x))(1+\chi_p(x+l))\prod_{i=1}^{l-1}(1-\chi_p(x+i)) + O(l).$$

Here we have the main term $p2^{-l}$, the error term $O(l)$ and 2^l-1 character sums of the type $\sum_{x\in\mathbb{Z}_p}\chi_p(x+a_1)...\chi_p(x+a_r)$ with distinct $a_1,...,a_r$. Such a sum is estimated by $rp^{1/2}$ in magnitude (famous Weil's theorem).

Hence,

$$N_l = p2^{-l} + O(l2^lp^{1/2})$$

and it is the asymptotics for N_l if $l \ll \log p$.

Theorems 1 and 2: method of the proof

$$N_l = \sum_{x\in\mathbb{Z}_p} R_p(x)R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)+1(x=0)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) + O(1) =$$

$$2^{-l}\sum_{x\in\mathbb{Z}_p}(1+\chi_p(x))(1+\chi_p(x+l))\prod_{i=1}^{l-1}(1-\chi_p(x+i)) + O(l).$$

Here we have the main term $p2^{-l}$, the error term $O(l)$ and 2^l-1 character sums of the type $\sum_{x\in\mathbb{Z}_p}\chi_p(x+a_1)...\chi_p(x+a_r)$ with distinct $a_1,...,a_r$. Such a sum is estimated by $rp^{1/2}$ in magnitude (famous Weil's theorem).

Hence,

$$N_l = p2^{-l} + O(l2^l p^{1/2})$$

and it is the asymptotics for N_l if $l \ll \log p$.

$$N_l = \sum_{x\in\mathbb{Z}_p} R_p(x)R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)+1(x=0)}{2}R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)}{2}R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) + O(1) =$$

$$2^{-l}\sum_{x\in\mathbb{Z}_p}(1+\chi_p(x))(1+\chi_p(x+l))\prod_{i=1}^{l-1}(1-\chi_p(x+i)) + O(l).$$

Here we have the main term $p2^{-l}$, the error term $O(l)$ and 2^l-1 character sums of the type $\sum_{x\in\mathbb{Z}_p}\chi_p(x+a_1)...\chi_p(x+a_r)$ with distinct $a_1, ..., a_r$. Such a sum is estimated by $rp^{1/2}$ in magnitude (famous Weil's theorem).

Hence,

$$N_l = p2^{-l} + O(l2^lp^{1/2})$$

and it is the asymptotics for N_l if $l \ll \log p$.

$$N_l = \sum_{x\in\mathbb{Z}_p} R_p(x)R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)+1(x=0)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) + O(1) =$$

$$2^{-l}\sum_{x\in\mathbb{Z}_p}(1+\chi_p(x))(1+\chi_p(x+l))\prod_{i=1}^{l-1}(1-\chi_p(x+i)) + O(l).$$

Here we have the main term $p2^{-l}$, the error term $O(l)$ and 2^l-1 character sums of the type $\sum_{x\in\mathbb{Z}_p}\chi_p(x+a_1)...\chi_p(x+a_r)$ with distinct $a_1,...,a_r$. Such a sum is estimated by $rp^{1/2}$ in magnitude (famous Weil's theorem).

Hence,

$$N_l = p2^{-l} + O(l2^lp^{1/2})$$

and it is the asymptotics for N_l if $l \ll \log p$.

$$N_l = \sum_{x\in\mathbb{Z}_p} R_p(x)R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)+1(x=0)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) =$$

$$\sum_{x\in\mathbb{Z}_p} \frac{1+\chi_p(x)}{2} R_p(x+l)\prod_{i=1}^{l-1}(1-R_p(x+i)) + O(1) =$$

$$2^{-l}\sum_{x\in\mathbb{Z}_p}(1+\chi_p(x))(1+\chi_p(x+l))\prod_{i=1}^{l-1}(1-\chi_p(x+i)) + O(l).$$

Here we have the main term $p2^{-l}$, the error term $O(l)$ and 2^l-1 character sums of the type $\sum_{x\in\mathbb{Z}_p}\chi_p(x+a_1)...\chi_p(x+a_r)$ with distinct $a_1, ..., a_r$. Such a sum is estimated by $rp^{1/2}$ in magnitude (famous Weil's theorem).

Hence,

$$N_l = p2^{-l} + O(l2^l p^{1/2})$$

and it is the asymptotics for N_l if $l \ll \log p$.

For composite moduli of the type $p_1 \dots p_t$ with distinct large p_j (larger than 2^t) and small l (roughly less than $M/|R_M| \sim 2^t$) we can find the asymptotics for N_l using a simple version of sieve method and estimates of character sums. Large values of l give negligible contribution to the sum $\sum_l N_l l^2$. Small arbitrary factor A gives additional technical difficulties.

In fact, we prove that

$$S(R_M) = m2^t f_A(y_t) + O(m2^{3t} A^4 p_1^{-c}),$$

where $y = 1 - 2^{-t}$ and f_A is the function defined earlier.

We see that to have a reasonable error term we should require $p_1^{-c} \gg A^4$

Theorems 1 and 2: method of the proof

For composite moduli of the type $p_1 ... p_t$ with distinct large p_j (larger than 2^t) and small l (roughly less than $M/|R_M| \sim 2^t$) we can find the asymptotics for N_l using a simple version of sieve method and estimates of character sums. Large values of l give negligible contribution to the sum $\sum_l N_l l^2$. Small arbitrary factor A gives additional technical difficulties.

In fact, we prove that

$$S(R_M) = m2^t f_A(y_t) + O(m2^{3t} A^4 p_1^{-c}),$$

where $y = 1 - 2^{-t}$ and f_A is the function defined earlier.

We see that to have a reasonable error term we should require $p_1^{-c} \gg A^4$

For composite moduli of the type $p_1...p_t$ with distinct large p_j (larger than 2^t) and small l (roughly less than $M/|R_M| \sim 2^t$) we can find the asymptotics for N_l using a simple version of sieve method and estimates of character sums. Large values of l give negligible contribution to the sum $\sum_l N_l l^2$. Small arbitrary factor A gives additional technical difficulties.

In fact, we prove that

$$S(R_M) = m2^t f_A(y_t) + O(m2^{3t}A^4 p_1^{-c}),$$

where $y = 1 - 2^{-t}$ and f_A is the function defined earlier.

We see that to have a reasonable error term we should require $p_1^{-c} \gg A^4$

We are able to prove that

$$f_A(1) = \frac{2A^2}{|R_A|} \text{ and } \quad f'_A(1) = \frac{A^2}{|R_A|}.$$

Hence by Taylor's expansion

$$m2^t f_A(y_t) = m2^{t+1}A^2|R_A|^{-1} - mA^2|R_A|^{-1} + \frac{1}{2}f''_A(\theta_t)m2^{-t}.$$

Also we prove the bound $f''_A(y) \ll A^{O(1)}$ for $y \in (1/2, 1]$.

Getting all of this together we get Theorem 2.

We are able to prove that

$$f_A(1) = \frac{2A^2}{|R_A|} \text{ and } \quad f'_A(1) = \frac{A^2}{|R_A|}.$$

Hence by Taylor's expansion

$$m2^t f_A(y_t) = m2^{t+1}A^2|R_A|^{-1} - mA^2|R_A|^{-1} + \frac{1}{2}f''_A(\theta_t)m2^{-t}.$$

Also we prove the bound $f''_A(y) \ll A^{O(1)}$ for $y \in (1/2, 1]$.

Getting all of this together we get Theorem 2.

We are able to prove that

$$f_A(1) = \frac{2A^2}{|R_A|} \text{ and } \quad f'_A(1) = \frac{A^2}{|R_A|}.$$

Hence by Taylor's expansion

$$m2^t f_A(y_t) = m2^{t+1}A^2|R_A|^{-1} - mA^2|R_A|^{-1} + \frac{1}{2}f''_A(\theta_t)m2^{-t}.$$

Also we prove the bound $f''_A(y) \ll A^{O(1)}$ for $y \in (1/2, 1]$.

Getting all of this together we get Theorem 2.

Recall the definition of Ω.

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of positive integers M such that $M = Am$ where

(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;
(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \dots < p_t$ are primes greater than $2^{C_0 t}$.

It is well-known that $\omega(M)$ is close to $\log\log X$ most of the time $(M \leqslant X)$, so $A \leqslant 2^{c_0 t} \leqslant (\log X)^{\alpha}$; whereas it can be shown by sieve methods that

$$\Omega_0(X) := \#\{m \leqslant X : \mu(m) \neq 0, (m, P((\log X)^{\alpha})) = 1\} \sim \frac{X}{\alpha \log\log X}.$$

So

$$\#(\Omega \cap [1, X]) \gg \sum_{A \leqslant (\log X)^{\alpha}} \Omega_0(X/A) \gg \sum_{A \leqslant (\log X)^{\alpha}} \mu^2(A) \frac{X}{A \log\log(X/A)} \gg X.$$

Recall the definition of Ω.

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of positive integers M such that $M = Am$ where

(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;
(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \dots < p_t$ are primes greater than $2^{C_0 t}$.

It is well-known that $\omega(M)$ is close to $\log\log X$ most of the time $(M \leqslant X)$, so $A \leqslant 2^{c_0 t} \leqslant (\log X)^{\alpha}$; whereas it can be shown by sieve methods that

$$\Omega_0(X) := \#\{m \leqslant X : \mu(m) \neq 0, (m, P((\log X)^{\alpha})) = 1\} \sim \frac{X}{\alpha \log\log X}.$$

So

$$\#(\Omega \cap [1, X]) \gg \sum_{A \leqslant (\log X)^{\alpha}} \Omega_0(X/A) \gg \sum_{A \leqslant (\log X)^{\alpha}} \mu^2(A) \frac{X}{A \log\log(X/A)} \gg X.$$

Recall the definition of Ω.

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of positive integers M such that $M = Am$ where

(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;
(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \dots < p_t$ are primes greater than $2^{C_0 t}$.

It is well-known that $\omega(M)$ is close to $\log\log X$ most of the time $(M \leqslant X)$, so $A \leqslant 2^{c_0 t} \leqslant (\log X)^{\alpha}$; whereas it can be shown by sieve methods that

$$\Omega_0(X) := \#\{m \leqslant X : \mu(m) \neq 0, (m, P((\log X)^{\alpha})) = 1\} \sim \frac{X}{\alpha \log\log X}.$$

So

$$\#(\Omega \cap [1, X]) \gg \sum_{A \leqslant (\log X)^{\alpha}} \Omega_0(X/A) \gg \sum_{A \leqslant (\log X)^{\alpha}} \mu^2(A) \frac{X}{A \log\log(X/A)} \gg X.$$

Recall the definition of Ω.

Let c_0 and C_0 be positive absolute constants, c_0 is small, C_0 is large. Let Ω be the set of positive integers M such that $M = Am$ where

(i) A is square-free, $(A, m) = 1$ and $A \leqslant 2^{c_0 t}$;
(ii) $m = p_1 \dots p_t$, $t \geqslant 0.4 \log\log M$ and $p_1 < p_2 < \dots < p_t$ are primes greater than $2^{C_0 t}$.

It is well-known that $\omega(M)$ is close to $\log\log X$ most of the time ($M \leqslant X$), so $A \leqslant 2^{c_0 t} \leqslant (\log X)^{\alpha}$; whereas it can be shown by sieve methods that

$$\Omega_0(X) := \#\{m \leqslant X : \mu(m) \neq 0, (m, P((\log X)^{\alpha})) = 1\} \sim \frac{X}{\alpha \log\log X}.$$

So

$$\#(\Omega \cap [1, X]) \gg \sum_{A \leqslant (\log X)^{\alpha}} \Omega_0(X/A) \gg \sum_{A \leqslant (\log X)^{\alpha}} \mu^2(A) \frac{X}{A \log\log(X/A)} \gg X.$$

Theorem 3. *We have*

$$S(R_M) \geqslant mA^2|R_A|^{-1}(2^{t+1}-1) + O\left(\frac{M^{2-c}}{|R_M|}\right)$$

for almost all M.

Let $M \in [1, X]$ be "a standard" number and $p_1 \leqslant p_2 \leqslant p_3 \leqslant ... \leqslant p_k$ are all prime divisors of M written with multiplicity. Let $u > 0$ be fixed. How often do we have

$$\prod_{i<j} p_i < p_j^u \quad ?$$

Erdös-Bovey result (~ 1970):

There exists a continious increasing function $\tau(u)\colon [0, \infty) \to [0, 1]$ *with* $\tau(0) = 0$, $\lim_{u \to \infty} \tau(u) = 1$ *such that*

$$\#\{j : \prod_{i<j} p_i < p_j^u\} \sim \tau(u) \log \log X$$

for almost all $M \in [1, X]$.

THANK YOU FOR YOUR ATTENTION !