Let $m \in \mathbb{N}$, $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ be the set of the residues modulo m. If p is a prime, then $\mathbb{Z}_p$ is a field of order p. Let $\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{0\}$ be the set of invertible elements in $\mathbb{Z}_p$. We take an arbitrary subgroup G of the group $\mathbb{Z}_p^*$. Let $t = |G|$. For brevity, we will write $a \equiv b$ instead of $a \equiv b (\mathrm{mod} p)$.

For $u \in \mathbb{R}$ we denote $e(u) = \exp(2\pi i u)$. The function $e(\cdot)$ is 1-periodic, and this allows us to talk about $e(a/p)$ for $a \in \mathbb{Z}_p$.

The main subject of my talks is the estimation of exponential sums over G:

$$S(a, G) = \sum_{x \in G} e(ax/p), \quad a \in \mathbb{Z}_p.$$

We can estimate $S(a, G)$ trivially:

$$|S(a,G)| \le \sum_{x\in G} |e(ax/p)| = \sum_{x\in G} 1 = |G|. \tag{1.3}$$

Clearly, inequality (1.3) is equality if $a = 0$. We are interested in obtaining nontrivial estimates for $S(a, G)$:

$$S(a,G) = o(|G|) \quad (p \to \infty, a \in \mathbb{Z}_p^*) \tag{1.4}$$

or, for some $\delta > 0$.

$$S(a,G) \ll |G|p^{-\delta} \quad (a \in \mathbb{Z}_p^*). \tag{1.5}$$

We proved the simplest estimate for $|S(a, G)|$.

Theorem 1.7. *We have*

$$|S(a,G)| \le \sqrt{p} \quad (a \in \mathbb{Z}_p^*). \tag{1.15}$$

So, we have a nontrivial estimate for exponential sums over G (namely, (1.5)) provided that $|G| \ge p^{1/2+\delta}$. Our aim is to weaken this inequality for $|G|$.

To get better estimates for $S(a,G)$ we define, for $k \in \mathbb{N}$, $T_k(G)$ as the number of the solutions to the congruence

$$x_1 + \cdots + x_k \equiv x_{k+1} + \cdots + x_{2k}, \quad x_j \in G.$$

Clearly, $T_1(G) = t$, and, for any k,

$$t^k \le T_k(G) \le t^{2k-1}. \tag{1.17}$$

Also, we have

$$pT_k(G) = \sum_{a \in \mathbb{Z}_p} |S(a,G)|^{2k}. \tag{1.18}$$

It easily follows from (1.18) that

$$T_k(G) \ge |S(0,G)|^{2k}/p = t^{2k}/p. \tag{1.19}$$

We proved the following.

Proposition 1.9. *We have*

$$|S(a,G)| \le (pT_k(G)/t)^{1/(2k)} \quad (a \in \mathbb{Z}_p^*). \tag{1.21}$$

In particular, if $T_k(G)/t^{2k} \le tp^{-\varepsilon}/p$ then

$$|S(a,G)| \le |G| p^{-\varepsilon/(2k)} \quad (a \in \mathbb{Z}_p^*).$$

Observe that Theorem 1.7 is a particular case of Proposition 1.9 for $k = 1$. If we use a trivial estimate $T_k(G) \le t^{2k-1}$ we get only

$$|S(a,G)| \le \left(pt^{2k-1}/t\right)^{1/(2k)} = t(p/t^2)^{1/(2k)}.$$

This estimate is worse than the trivial one $|S(a,G)| \le t$ if $|G| < p^{1/2}$ and worse than the simplest estimate $|S(a,G)| \le p^{1/2}$ if $|G| > p^{1/2}$. However, if $|G|$ is close to $p^{1/2}$ then any improvement of the trivial inequality $T_k(G) \le t^{2k-1}$ will improve estimates for $|S(a,G)|$.

Such an improvement was made by Shparlinski who used the following result of A. Garcia and J. F. Voloch.

Theorem 2.1. *For $b \in \mathbb{Z}_p$ denote by $N_2(b)$ the number of solutions to the congruence $x_1 + x_2 \equiv b$, $x_1, x_2 \in G$. If*

$$|G| < \frac{p-1}{(p-1)^{1/4}+1}, \tag{2.1}$$

then for any $b \in \mathbb{Z}_p^*$ we have

$$N_2(b) \le 4|G|^{2/3}. \tag{2.2}$$

Using (2.2), one can nontrivially estimate $T_2(G)$ provided that (2.1) holds. Recall that $T_2(G)$ is the number of solutions to

$$x_1 + x_2 \equiv x_3 + x_4, \quad x_j \in G. \tag{2.3}$$

The number of solutions to (2.3) with $x_3 + x_4 \equiv 0$ is at most $|G|^2$. Next, if $x_3 + x_4 \not\equiv 0$, then, by (2.2), there are at most $4|G|^{2/3}$ pairs (x_1, x_2) satisfying (2.3) Therefore,

$$T_2(G) \leq p^2 + 4p^{8/3} < 5p^{8/3}. \tag{2.4}$$

Now we can estimate exponential sums using Proposition 1.9

$$|S(a, G)| \leq (pT_k(G)/t)^{1/(2k)} \quad (a \in \mathbb{Z}_p^*). \tag{1.21}$$

for $k = 2$:

$$|S(a, G)| \leq (5p)^{1/4}|G|^{5/12} \quad (a \in \mathbb{Z}_p^*).$$

This is better than the estimate $p^{1/2}$ for $|G| \leq p^{3/5-\delta}$, $p \geq p(\delta)$, and better than the trivial $|G|$ for $|G| \geq p^{3/7+\delta}$, $p \geq p(\delta)$. Observing that (2.1) holds for $|G| \leq p^{3/4-\delta}$, $p \geq p(\delta)$. Thus, the improvement was made for $p^{3/7+\delta} \leq |G| \leq p^{3/5-\delta}$, $p \geq p(\delta)$.

D. R. Heath-Brown succeeded in applying Stepanov's method to the proof of the theorem of Garcia and Voloch. Moreover, in our joint paper we used his technique to improve estimate (2.4) for $T_2(G)$ if $|G| \le p^{2/3}$.

Theorem 2.2. *If* $|G| \le p^{2/3}$, *then*

$$T_2(G) \ll |G|^{5/2}. \tag{2.5}$$

We are not able to improve the estimate of Garcia and Voloch

$$N_2(b) \ll |G|^{2/3}$$

for all $b \in \mathbb{Z}_p^*$, but it can be improved in average, and this implies (2.5). I shall present the proof of (2.5), but first let us discuss its applications. To estimate exponential sums $S(a, G)$, one can use Proposition 1.9; however, the following more general fact sometimes gives better estimates.

Theorem 2.3. *If $k, l \in \mathbb{N}$, $a \in \mathbb{Z}_p^*$, then*

$$|S(a,G)| \le \left(pT_k(G)T_l(G)\right)^{1/(2kl)} t^{1-1/k-1/l}. \tag{2.6}$$

Clearly, for $l = 1$ Theorem 2.3 is just Proposition 1.9. For $k = l$ (2.6) can be written as

$$|S(a,G)| \le \left(\frac{T_k(G)p^{1/2}}{t^{2k}}\right)^{1/(k^2)} t. \tag{2.7}$$

Clearly, (2.7) supersedes the trivial estimate $|S(a,G)| \le t$ if and only if

$$T_k(G) < t^{2k}p^{-1/2}. \tag{2.8}$$

In the most interesting case $|G| < p^{1/2}$ (2.8) is weaker than the condition $T_k(G) < t^{2k}t/p$ required to have any benefit from Proposition 1.9.

Theorem 2.3 probably has to be attributed to A. A. Karatsuba who in fact proved the following.

Theorem 2.4. *Let $X \subset \mathbb{Z}_p^*$. For $k \in \mathbb{N}$ by $T_k(X)$ denote the number of the solutions to the congruence*

$$x_1 + \cdots + x_k \equiv x_{k+1} + \cdots + x_{2k}, \quad x_j \in X.$$

Then for $k, l \in \mathbb{N}$, $a \in \mathbb{Z}_p^*$, we have

$$\left| \sum_{x,y \in X} e(axy/p) \right| \le (pT_k(X)T_l(X))^{1/(2kl)} |X|^{2-1/k-1/l}.$$

Theorem 2.4 is similar to the results proven for estimates of H. Weil's sums by I. M. Vinogradov's method. Theorem 2.3 is contained in Theorem 2.4 since

$$\sum_{x,y \in G} e(axy/p) = |G| \sum_{z \in G} e(az/p) = |G| S(a, G).$$

Combining Theorem 2.2 with Theorem 2.3 for $k = 1, l = 2$ if $p^{1/2} < |G| \le p^{2/3}$ and for $k = l = 2$ if $|G| \le p^{1/2}$ we get for $a \in \mathbb{Z}_p^*$

$$|S(a, G)| \ll p^{1/4} |G|^{3/8} \quad (p^{1/2} < |G| \le p^{2/3}), \tag{2.9}$$

$$|S(a, G)| \ll p^{1/8} |G|^{5/8} \quad (|G| \le p^{1/2}). \tag{2.10}$$

Observe that (2.9) supersedes the simplest estimate $|S(a,G)| \le p^{1/2}$ for $|G| \le p^{2/3-\delta}$, $p \ge p(\delta)$, and (2.10) supersedes the trivial estimate $|S(a,G)| \le |G|$ for $|G| \ge p^{1/3+\delta}$, $p \ge p(\delta)$. For $|G| \ge p^{2/3}$ we cannot prove anything better than $|S(a,G)| \ll p^{1/2}$.

Let me recall the definition of $1/p$-pseudo-random generators of Blum, Blum, and Shub. Take an integer $g \ge 2$. We consider the g-ary expansion of $1/p$. If g is fixed then we can expect (and this is true indeed) that for many primes p there is no large correlation among close digits in this expansion, and we can talk about a pseudo-random generator. Let G be the subgroup of $\mathbb{Z}_p^*$ generated by g, $t = |G|$. It is easy to see that t is the (least) period of the g-ary expansion of $1/p$. We are interested in appearances of a sequence $(d_1, \dots, d_k)$ of g-ary digits in the expansion. We have proved that if

$$\max_{a \in \mathbb{Z}_p^*} |S(a,G)|/t \le \eta \tag{1.7}$$

and

$$\frac{p}{2g^k} - 1 \ge \eta p/(1+\eta) \tag{2.11}$$

then the g-ary expansion of $1/p$ contains any string of length k. It is easy to see that (2.11) holds if $k \le (\log(1/\eta) - C)/\log g$ for some absolute constant C.

Let me stress that we do not expect that the digits of the g-ary expansion of $1/p$ are well-distributed for ALL large p. For example, take $g = 2$. If p is a Mersenne prime (that is, $p = 2^q - 1$), then the expansion has the string $(0, \dots, 0, 1)$ of size q as its period; thus, the sequence is very far from being pseudo-random. However, we can say that for ALMOST ALL primes the sequence of digits is in a sense well-distributed.

Fix g and take a large $L \in \mathbb{N}$. Also,let $T \in \mathbb{N}$. Let us estimate the number N of primes $p \le g^L$ such that the order of g in $\mathbb{Z}_p$ is at most T. We have

$$N \le \sum_{t \le T} |\{p : \, g^t \equiv 1 (\mathrm{mod} p)\}| = \sum_{t \le T} w(g^t - 1)$$
$$\ll \sum_{t \le T} t \le T^2.$$

On the other hand, the number of primes $p \le g^L$ is $\gg g^L/L$. Therefore, for every fixed $\varepsilon > 0$, specifying $T = g^{(1/2-\varepsilon)L}$, we see that for almost all primes $p \le g^L$ the order of g in $\mathbb{Z}_p$ is $> T \ge p^{1/2-\varepsilon}$. This means that the proportion of exceptional primes amongst all the primes $\le g^L$ tends to 0 as $L \to \infty$.

Next, if G is the subgroup of $\mathbb{Z}_p^*$ generated by g, $t = |G| > p^{1/2-\varepsilon}$, than, by (2.9) and (2.10),

$$|S(a,G)| \ll p^{1/4}|G|^{3/8} \quad (p^{1/2} < |G| \le p^{2/3}), \tag{2.9}$$

$$|S(a,G)| \ll p^{1/8}|G|^{5/8} \quad (|G| \le p^{1/2}). \tag{2.10}$$

we have

$$\max_{a\in\mathbb{Z}_p^*} |S(a,G)|/t \le \eta$$

with $\eta \ll p^{-\frac{1}{16}+\frac{3}{8}\varepsilon}$. This implies, that the g-ary expansion of $1/p$ contains any string of length
$\le (\frac{1}{16}-\frac{3}{8}\varepsilon)L - C$. Moreover, for large L all the strings of length $\le (\frac{1}{16}-\varepsilon)L$ will appear with approximately the same frequency. Observe that we cannot prove any results of this type using the simplest estimate $|S(a,G)| \le p^{1/2}$.

We (SK, I. Shparlinski) can prove more: for almost all primes $p \le g^L$ the g-ary expansion of $1/p$ contains any string of length $\le \frac{3}{37}L$.

Now we shall make some preparations to prove the estimate for $T_2(G)$. Take some cosets $G_1, \dots, G_s$ of the group G in $\mathbb{Z}_p^*$. For any coset G_j denote

$$N_j = |\{x \in G : \, x - 1 \in G_j\}|.$$

Lemma 2.5. *Let $|G| = t$ and suppose that a positive integer L satisfies the conditions*

$$L < t, \quad tL \le p, \quad s < L^3/(2t). \tag{2.12}$$

Then

$$\sum_{j=1}^{s} N_j \le \frac{2tL}{[t/L]}.$$

Proof. Let $K = [t/L]$. We shall begin by taking a polynomial $\Phi(X, Y, Z)$, for which

$$\deg_X \Phi < K, \quad \deg_Y \Phi < L, \quad \deg_Z \Phi < L.$$

For $j = 1, \dots, s$ we define the sets

$$R_j = \{x \in G : \, x - 1 \in G_j\}, \quad R = \bigcup_{j=1}^{s} R_j.$$

Clearly,

$$\sum_{j=1}^{s} N_j = |R|.$$

The underlying idea is then to arrange that the polynomial

$$\Psi(X) = \Phi(X, X^t, (X-1)^t)$$

has a zero of order at least K at each point $x \in R$. We will therefore be able to conclude that

$$K \sum_{j=1}^{s} N_j \leq \deg \Psi,$$

provided that Ψ does not vanish identically. We note that

$$\deg \Psi \leq \deg_X \Phi + t \deg_Y \Phi + t \deg_Z \Phi \leq K - 1 + 2t(L-1),$$

whence

$$\sum_{j=1}^{s} N_j \leq \frac{K - 1 + 2t(L-1)}{K} < \frac{2tL}{[t/L]},$$

provided that Ψ does not vanish identically.

In order for Ψ to have a zero of multiplicity at least K at a point x, we need

$$\left(\frac{d}{dx}\right)^n \Psi(X)\Big|_{X=x} = 0 \quad (n < K).$$

Since $x \neq 0, 1$ for $x \in R$, this will be equivalent to

$$(X(X-1))^n \left(\frac{d}{dx}\right)^n \Psi(X)\Big|_{X=x} = 0. \tag{2.13}$$

We now observe that

$$X^m \left(\frac{d}{dx}\right)^m X^u = \frac{u!}{(u-m)!} X^u,$$

$$X^m \left(\frac{d}{dx}\right)^m X^{tv} = \frac{(tv)!}{(tv-m)!} X^{tv},$$

$$(X-1)^m \left(\frac{d}{dx}\right)^m (X-1)^{tw} = \frac{(tw)!}{(tw-m)!} (X-1)^{tw}.$$

It follows that

$$(X(X-1))^k \left(\frac{d}{dX}\right)^k X^u X^{tv} (X-1)^{tw}$$
$$= P_{k,u,v,w}(X) X^{tv} (X-1)^{tw}$$

where $P_{k,u,v,w}$ either vanishes or is a polynomial of degree at most $k+u$. We therefore deduce that for any $j = 1, \ldots, s$ and for any $x \in R_j$, we have

$$(X(X-1))^k \left(\frac{d}{dx}\right)^k X^u X^{tv} (X-1)^{tw} \Big|_{X=x}$$
$$= a_j^w P_{k,u,v,w}(x)$$

where $a_j = y^t$ for $y \in G_j$; the crucial argument here is that y^t does not depend on the choice of $y \in G$ or $y \in G_j$.

We now write

$$\Phi(X, Y, Z) = \sum_{u,v,w} \lambda_{u,v,w} X^u Y^v Z^w$$

and

$$P_{k,j}(X) = \sum_{u,v,w} \lambda_{u,v,w} a_j^w P_{k,u,v,w}(X)$$

so that $\deg P_{k,j} < A + k$ and

$$(X(X-1))^k \left(\frac{d}{dX}\right)^k \Phi(X, X^t, (X-1)^t)\Big|_{X=x} = P_{k,j}(x)$$

for any $x \in R_j$. We shall arrange, by appropriate choice of the coefficients $\lambda_{u,v,w}$, that $P_{k,j}(X)$ vanishes identically for $k < K$. This will ensure that

$$(X(X-1))^n \left(\frac{d}{dx}\right)^n \Psi(X)\Big|_{X=x} = 0 \tag{2.13}$$

holds at every point $x \in R$. Each polynomial $P_{k,j}(X)$ has at most $K + k < 2K$ coefficients which are linear forms in the original $\lambda_{u,v,w}$. Thus if

$$sK(2K) < KL^2, \tag{2.14}$$

there will be a set of coefficients $\lambda_{u,v,w}$, not all zero, for which the polynomials $P_{k,j}(X)$ vanish for all $k < K$. But, since $K = [t/L] \le t/L$ and $s < L^3/(2t)$,

$$sK(2K) = 2sK^2 \le 2sKt/L < KL^2,$$

and (2.14) holds.

We must now consider whether the polynomial $\Phi(X, X^t, (X-1)^t)$ can vanish if $\Phi(X, Y, Z)$ does not. We shall write

$$\Phi(X, Y, Z) = \sum_w \Phi_w(X, Y) Z^w,$$

and take w_0 to be the smallest value w for which $\Phi_w(X, Y)$ is not identically zero. It follows that

$$\begin{aligned} &\Phi(X, X^t, (X-1)^t) \\ &= (X-1)^{tw_0} \sum_{w_0 \le w \le B} \Phi_w(X, X^t)(X-1)^{t(w-w_0)}, \end{aligned}$$

so that if $\Phi(X, X^t, (X-1)^t)$ is identically zero, we must have

$$\Phi_{w_0}(X, X^t) \equiv 0 (\mathrm{mod} (X-1)^t). \tag{2.15}$$

We show, by induction on N, that if a polynomial $f(X) \in \mathbb{Z}_p[X]$ of degree $\deg f < p$ is a sum of $N \geq 1$ distinct monomials, then $(X-1)^N$ cannot divide $f(X)$. The case $N = 1$ is trivial. Now suppose that $N > 1$ and let

$$f(X) = \sum_w c_w x^W$$

where w runs over N distinct values. Then the polynomial

$$g(X) = Xf'(X) - Wf(X) = \sum_w c_w(w - W)X^w,$$

where $W = \deg w$, contains exactly $N - 1$ terms. (Notice that $c_w(w - W) \in \mathbb{Z}_p$ is nonzero for $w < W$ since $W < p$.) We then see that if $(X-1)^N$ divides $f(X)$, then $(X-1)^{N-1}$ divides $g(X)$ contrary to our induction hypothesis.

We have

$$\deg \Phi_{w_0}(X, X^t) \leq K - 1 + t(L-1) < tL.$$

Therefore, the congruence

$$\Phi_{w_0}(X, X^t) \equiv 0 (\mathrm{mod} (X-1)^t) \tag{2.15}$$

is impossible provided that $KL \leq t$, $tL \leq p$. But these inequalities hold, and Lemma 2.5 is proven.

Now take all the cosets $G_1, \dots, G_n$ of the group G in $\mathbb{Z}_p^*$; thus, $n = (p-1)/t$. Again, for any coset G_j we denote

$$N_j = |\{x \in G : \, x - 1 \in G_j\}|.$$

Hence,

$$N_j = |\{x \in G, y \in G_j : \, x - 1 \equiv y\}|,$$

$$tN_j = |\{x_1, x_2 \in G, y \in G_j : \, x_1 - x_2 \equiv y\}|,$$

and for any $y \in G_j$ we have

$$N_j = |\{(x_1, x_2) \in G : \, x_1 - x_2 \equiv y\}|.$$

Therefore,

$$\begin{aligned} T_2(G) &= |\{(x_1, x_2, x_3, x_4) : \, x_j \in G, x_1 - x_2 \equiv x_3 - x_4\}| \\ &= \sum_{y \in \mathbb{Z}_p} |\{(x_1, x_2) : \, x_1, x_2 \in G, x_1 - x_2 \equiv y\}|^2 \\ &\le t^2 + \sum_{j=1}^{n} \sum_{y \in G_j} |\{(x_1, x_2) : \, x_1, x_2 \in G, x_1 - x_2 \equiv y\}|^2 \\ &= t^2 + \sum_{j=1}^{n} \sum_{y \in G_j} N_j^2 = t^2 + t \sum_{j=1}^{n} N_j^2. \end{aligned} \tag{2.16}$$

Also, observe that

$$t^2 = \sum_{y\in\mathbb{Z}_p} |\{(x_1,x_2):\, x_1,x_2\in G, x_1-x_2\equiv y\}|$$
$$\geq \sum_{j=1}^{n}\sum_{y\in G_j} |\{(x_1,x_2):\, x_1,x_2\in G, x_1-x_2\equiv y\}|$$
$$= \sum_{j=1}^{n}\sum_{y\in G_j} N_j = t\sum_{j=1}^{n} N_j.$$

Hence,

$$\sum_{j=1}^{n} N_j \leq t. \tag{2.17}$$

Now we are in position to prove Theorem 2.2.

Theorem 2.2. *If $|G|\leq p^{2/3}$, then*

$$T_2(G) \ll |G|^{5/2}. \tag{2.5}$$

We assume that $t=|G|$ is large enough and the cosets $G_1,\dots,G_n$ are ordered in such a way that

$$N_1 \geq N_2 \cdots \geq N_n.$$

Then for $1 \le s \le t^{1/2}/3$ and $L = [(2st)^{1/3}] + 1$ the conditions

$$L < t, \quad tL \le p, \quad s < L^3/(2t). \tag{2.12}$$

of Lemma 2.5 are satisfied, and it can be applied giving

$$\sum_{j=1}^{s} N_j \ll s^{2/3} t^{2/3}.$$

Hence,

$$N_s \ll s^{-1/3} t^{2/3} \quad (s \le t^{1/2}/3). \tag{2.18}$$

For $s > t^{1/2}/3$ the following estimate holds:

$$N_s \le N_{[t^{1/2}/3]} \ll t^{1/2}. \tag{2.19}$$

Using (2.16) and combining the bounds (2.18) and (2.19) with (2.17) we get

$$T_2(G) \leq t^2 + t\sum_{s=1}^{n} N_s^2$$

$$\leq t^2 + t \sum_{s \leq t^{1/2}/3} N_s^2 + t \sum_{s > t^{1/2}/3} N_s^2$$

$$\ll t^2 + t \sum_{s \leq t^{1/2}/3} \left(s^{-1/3}t^{2/3}\right)^2 + t \sum_{s > t^{1/2}/3} t^{1/2} N_s$$

$$\ll t^2 + t \sum_{s \leq t^{1/2}/3} \left(s^{-1/3}t^{2/3}\right)^2 + t(t^{1/2})t \ll t^{5/2},$$

and we have the desired result.

Now we will prove a corollary from Lemma 2.5. If $*$ is a binary operation on $\mathbb{Z}_p$, $A, B \subset \mathbb{Z}_p$, then we denote

$$A * B = \{a * b : \ a \in A, b \in B\}.$$

Corollary 2.7. *(A. Glibichuk.) Let $B \subset G$ and $0 < |B| \le p^{1/2}$. Then*

$$|G(B-B)| \gg |B|^{3/2}. \tag{2.20}$$

Proof. Let $G_1, \ldots, G_s$ be all the cosets of G in $\mathbb{Z}_p^*$ containing elements from $B-B$. Then $G_j \subset G(B-B)$ for $j = 1, \ldots, s$, and hence

$$|G(B-B)| = s|G| + 1. \tag{2.21}$$

Inequality (2.20) follows immediately from (2.21) for $s > |B|^{3/2}/(17|G|)$ (and, in particular, for $|G| > |B|^{3/2}/17$). Thus, we can assume that

$$|G| \le |B|^{3/2}/17, \quad s \le |B|^{3/2}/(17|G|). \tag{2.22}$$

Also, assume that $|B|$ is large enough. Fixed $x_0 \in B$. Recall that

$$N_j = |\{x \in G : \, x - 1 \in G_j\}|.$$

Equivalently,

$$N_j = |\{x \in G : \, x - x_0 \in G_j\}|.$$

Since for every $x \in B \setminus \{x_0\}$ we have $x - x_0 \in G_j$ for some $j = 1, \ldots, s$,

$$|B| - 1 = \sum_{j=1}^{s} |\{x \in B : x - x_0 \in G_j\}| \le \sum_{j=1}^{s} N_j. \tag{2.23}$$

Take $L = [(2st)^{1/3}] + 1$. Now we can use Lemma 2.5.

Lemma 2.5. *Let $|G| = t$ and suppose that a positive integer L satisfies the conditions*

$$L < t, \quad tL \le p, \quad s < L^3/(2t). \tag{2.12}$$

Then

$$\sum_{j=1}^{s} N_j \le \frac{2tL}{[t/L]}.$$

We have

$$L \le [(2|B|^{3/2}/17)^{1/3}] + 1 < (|B| - 1)^{1/2}/2. \tag{2.24}$$

Therefore,

$$L < |B|^{1/2} \le |B| \le t,$$

$$tL < (|B|^{3/2}/17)(|B|^{1/2}) < |B|^2 < p.$$

So, (2.12) are fulfilled. By Lemma 2.5 and (2.24),

$$\sum_{j=1}^{s} N_j \le 4L^2 < |B| - 1,$$

but his does not agree with (2.23), and Corollary 2.7 follows.

Using Stepanov— Heath-Brown's method, Theorem 2.2 can be extended to $k > 2$ provided that $|G| \le p^{1/2}$.

Theorem 2.8. *If* $|G| \le p^{1/2}$, $k \in \mathbb{N}$, *then*

$$T_k(G) \ll_k |G|^{2k-2+2^{1-k}}. \tag{2.25}$$

It follows from Theorem 2.3 that we can get nontrivial estimates for exponential sums if for some k and $\varepsilon > 0$ we have

$$T_k(G) \ll_{k,\varepsilon} |G|^{2k} p^{-1/2-\varepsilon}. \tag{2.26}$$

Namely, (2.26) implies $|S(a, G)| \ll_{k,\varepsilon} p^{-\varepsilon/k^2}|G|$ for $a \in \mathbb{Z}_p^*$. By Theorem 2.8, (2.26) holds for

$$|G| \ge p^{1/4+\varepsilon} \tag{2.27}$$

and $k \ge k(\varepsilon)$. Thus, we have nontrivial estimates for exponential sums under supposition (2.27).

It is likely that Theorem 2.8 and restriction (2.27) correspond to natural thresholds of Stepanov— Heath-Brown's method.

Let me mention a corollary from Theorem 2.8. For $b \in \mathbb{Z}_p$, $k \in \mathbb{N}$ we denote by $N_k(b)$ the number of the solutions to the congruence

$$x_1 + \cdots + x_k \equiv b, \quad x_1, \ldots, x_k \in G.$$

It is not difficult to prove that

$$\sum_{b \in kG} N_k(b) = |G|^k,$$

$$\sum_{b \in kG} N_k(b)^2 = T_k(G)$$

(we have checked this for $k = 2$). Hence, by Cauchy— Schwartz inequality

$$|kG| \geq |G|^{2k}/T_k(G),$$

and from Theorem 2.8 we get the following.

Corollary 2.9. *If* $|G| \leq p^{1/2}$, $k \in \mathbb{N}$, *then*

$$|kG| \gg_k |G|^{2-2^{1-k}}. \tag{2.28}$$

To weaken restriction

$$|G| \geq p^{1/4+\varepsilon} \tag{2.27}$$

we had to show that for $|G| \leq p^{1/4}$ and for some k and ε

$$T_k(G) \ll |G|^{2k-2-\varepsilon}.$$

This would imply

$$|kG| \gg |G|^{2+\varepsilon}.$$

But before 2003 it was not clear how to exclude the situation

$$\forall k \, \exists p, G : \, |G| \leq p^{1/4}, |kG| < |G|^2. \tag{2.29}$$

Now it is time to have an excursion to a very exciting number theoretical and combinatorial problem.

P. Erdős and E. Szemerédi asked the following question.

Problem 2.9. *Is it true that for every nonempty finite* $A \subset \mathbb{Z}$ *and for every* $\varepsilon > 0$

$$\max(|A+A|, |AA|) \gg_\varepsilon |A|^{2-\varepsilon}?$$

They proved that for some $\alpha > 0$

$$\max(|A+A|, |AA|) \gg |A|^{1+\alpha}. \tag{2.30}$$

M. Nathanson established (2.30) for $\alpha = 1/31$. This value was being improved by K. Ford, G. Elekes. J. Solymosi proved (2.30) for $\alpha = 3/11 - \varepsilon$ with an arbitrary $\varepsilon > 0$; moreover, (2.30) is true for any nonempty finite $A \subset \mathbb{C}$.

It was naturally to ask if (2.30) holds for $\mathbb{Z}_p$, but it was clear that it could not hold in full generality: indeed, for $A = \mathbb{Z}_p$ we have $A + A = AA = A$. But it was reasonable to conjecture the validity of (2.30) for small A, say, $|A| \le p^{1/2}$. This would exclude

$$\forall k \, \exists p, G : \, |G| \le p^{1/4}, N_k(G) < |G|^2. \tag{2.29}$$

Indeed, take a large k and use (2.29) with k replaced by k^2. Then we have $|G| \le p^{1/4}$,

$$|kG| \gg_k |G|^{2-2^{1-k}}, \tag{2.28}$$

but, by (2.29),

$$|k^2G| < |G|^2. \tag{2.31}$$

This inequality implies

$$|kG| \le |k^2G| < p^{1/2}.$$

Since

$$kG + kG = 2kG, \quad (kG)(kG) \subset k^2G,$$

we deduce from conjectural (2.30)

$$|k^2G| \ge \max(kG + kG, (kG)(kG)) \gg_k |G|^{\left(2-2^{1-k}\right)(1+\alpha)},$$

but this does not agree with (2.30) for $k = k(\alpha)$ and sufficiently large p.

Unfortunately no existing proofs of (2.30) for integer, real or complex numbers could be used for $\mathbb{Z}_p$.