Chapter 1

Expected value proof

1.1 Upper bound

Before proving part (b) of the main theorem, we will prove a lemma that shows that a cyclic group of prime order is covered by the sums of a random subset of logarithmic size almost always.

Lemma 1.1.1. *Let q be a prime number and $\mathcal{A}$ be a random subset of $\mathbb{Z}_q$ of size $4\lfloor 6\log_2 q\rfloor$. As q tends to infinity, $2\lfloor 6\log_2 q\rfloor\mathcal{A}$ covers $\mathbb{Z}_q$ almost always.*

Proof. Let $s \in \mathbb{N}$ such that $s \leq q$. Let $\mathcal{A}$ be a uniformly random subset of $\mathbb{Z}_q$ of size s, that is,

$$\Pr(\mathcal{A}) = \frac{1}{\binom{q}{s}}.$$

For a given $z \in \mathbb{Z}_q$ and $k \in \mathbb{N}$ for which $k \leq s/2$, let

$$N_z^k := \left\{ K \subseteq \mathbb{Z}_q : |K| = k, \sum_{t\in K} t = z \right\}.$$

Note that $|N_z^k| = \frac{1}{q}\binom{q}{k}$, since $K \in N_0^k$ if and only if $K + k^{-1}z \in N_z^k$ for every $z \in \mathbb{Z}_q$.

For $K \in N_z^k$, let E_K be the event that $K \subset \mathcal{A}$. Let X_K be the indicator variable of E_K. We define the random variable

$$X_z = \sum_{K\in N_z^k} X_K.$$

Note that X_z counts the number of sets of size k which add up to z. We now find $\mathrm{E}[X_z]$. Since the sum of every subset $K \subset S$ is in $\mathbb{Z}_q$,

$$\sum_{z\in Z_q} X_z = \binom{s}{k},$$

and so

$$\binom{s}{k} = \mathrm{E}\left[\sum_{z\in Z_q} X_z\right] = \sum_{z\in Z_q} \mathrm{E}[X_z].$$

As in the argument for finding $|N_z^k|$, for every $z \in \mathbb{Z}_q$,

$$\mathrm{E}[X_0] = \sum_{K \in N_0^k} \mathrm{E}[X_K] = \sum_{K \in N_0^k} \mathrm{E}[X_{K+k^{-1}z}] = \sum_{K \in N_z^k} \mathrm{E}[X_K] = \mathrm{E}[X_z].$$

Therefore, we have that

$$E[X_z] = \frac{1}{q}\binom{s}{k}. \tag{1.1}$$

Now, for $K, L \in N_z^k$, let $j \in \mathbb{N}$ such that $j \leq k$ and define

$$\Delta_j := \sum_{|K \cap L| = j} \Pr[E_K \wedge E_L].$$

If $|K \cap L| = j$,

$$\Pr[E_K \wedge E_L] = \frac{\binom{q-2k+j}{s-2k+j}}{\binom{q}{s}}.$$

We can bound the number of events for which $|K \cap L| = j$. First we choose K as any set in N_z^k and then we choose the remaining $k - j$ elements as any subset of $\mathbb{Z}_q \setminus K$ with size $k - j$. Thus,

$$\Delta_j \leq \frac{1}{q}\binom{q}{k}\binom{q-k}{k-j}\frac{\binom{q-2k+j}{s-2k+j}}{\binom{q}{s}}.$$

This implies that, using 1.1,

$$\begin{aligned}
\frac{\Delta_j}{\mathrm{E}[X_z]^2} &\leq \frac{\binom{q}{k}\binom{q-k}{k-j}\binom{q-2k+j}{s-2k+j}}{\frac{1}{q}\binom{s}{k}\frac{1}{q}\binom{s}{k}q\binom{q}{s}} \\
&= \frac{\frac{q!}{(q-k)!k!}\frac{(q-k)!}{(k-j)!(q-2k+k)!}\frac{(q-2k+j)!}{(s-2k+j)!(q-s)!}}{\frac{1}{q}\binom{s}{k}\frac{s!}{(s-k)!k!}\frac{q!}{(q-s)!s!}} \\
&= \frac{q\binom{s-k}{k-j}}{\binom{s}{k}}.
\end{aligned}$$

Let $s = 4\lfloor 6\log_2 q \rfloor$ and $k = 2\lfloor 6\log_2 q \rfloor$. Using that $\binom{s-k}{k-j}$ is maximized at $k - j = \lfloor (s-k)/2 \rfloor$,

$$\frac{\Delta_j}{\mathrm{E}[X_z]^2} \leq \frac{q\binom{2\lfloor 6\log_2 q \rfloor}{\lfloor 6\log_2 q \rfloor}}{\binom{4\lfloor 6\log_2 q \rfloor}{2\lfloor 6\log_2 q \rfloor}} \leq \frac{q}{\binom{2\lfloor 6\log_2 q \rfloor}{\lfloor 6\log_2 q \rfloor}} \leq \frac{q}{2^{\lfloor 6\log_2 q \rfloor}} \sim \frac{1}{q^5},$$

since $\binom{2\lfloor 6\log_2 q \rfloor}{\lfloor 6\log_2 q \rfloor}^2 \leq \binom{4\lfloor 6\log_2 q \rfloor}{2\lfloor 6\log_2 q \rfloor}$ (Proposition **??**).

Hence, by (**??**) and Theorem **??**,

$$\begin{aligned}
\Pr[X_z = 0] \leq \frac{\mathrm{E}[X_z] + \Delta}{\mathrm{E}[X_z]^2} &= \frac{1}{E[X_z]} + \sum_{j=0}^{k}\frac{\Delta_j}{\mathrm{E}[X_z]^2} \\
&\leq \frac{1}{E[X_z]} + \frac{(k+1)}{q^5} = \frac{1}{E[X_z]} + \frac{2\lfloor 6\log_2 q \rfloor + 1}{q^5}.
\end{aligned}$$

Therefore, by the union bound and since $q \to \infty$ as $p \to 0$,

$$\Pr\left[\bigvee_{z\in\mathbb{Z}_q} X_z = 0\right] \le \frac{q}{\mathrm{E}[X_z]} + \frac{2\lfloor 6\log_2 q\rfloor + 1}{q^4} \tag{1.2}$$

$$= \frac{q^2}{\binom{4\lfloor 6\log_2 q\rfloor}{2\lfloor 6\log_2 q\rfloor}} + \frac{2\lfloor 6\log_2 q\rfloor + 1}{q^4} \tag{1.3}$$

$$\le \frac{q^2}{2^{2\lfloor 6\log_2 q\rfloor}} + \frac{2\lfloor 6\log_2 q\rfloor + 1}{q^4} \tag{1.4}$$

$$\sim \frac{1}{q^{10}} + \frac{6\log q}{q^4} = o(1). \tag{1.5}$$

We conclude that, as $q \to \infty$, $X_z > 0$ for every $z \in \mathbb{Z}_q$ almost always. Thus, for every $z \in \mathbb{Z}_q$, there exists $K \in N_z^k$ such that $K \subset \mathcal{A}$ almost always. This means that $2\lfloor 6\log_2 q\rfloor\mathcal{A}$ covers $\mathbb{Z}_q$ almost always. $\square$

1.1.1 Proof of the upper bound

Lemma 1.1.2. *Let $\psi(x)$ be a function for which $x(\log x)^2 \in o(\psi(x))$ .Then*

$$\lim_{p\to 0}\Pr\left[F(\mathcal{S}) \le \psi\left(\frac{1}{p}\right)\right] = 1.$$

The proof of this theorem consists of several parts. The strategy is to prove that the Ápery set of a subsemigroup of S is completed before step $\psi\left(\frac{1}{p}\right)$ with high probability, since $F(\mathcal{S})$ is less than the maximum element of this Ápery set. The proof has the following structure:

1. First, we will find a step for which a prime q is chosen with high probability (E_1).
2. Then, in the spirit of Lemma 1.1.1 we will find a step such that a set $\mathcal{A}$ of s elements which are different modulo q are chosen with high probability (E_2).
3. Finally, we will apply Lemma 1.1.1 to $\mathrm{Ap}(\langle \mathcal{A} \cup \{q\}\rangle, q)$.

Proof.

Part 1

Consider the event D_1 that a prime q is selected, such that

$$\frac{200}{p}\log\frac{4}{p} \le q \le \left(\frac{4}{p}\log\frac{1}{p}\right)\log\left(\frac{4}{p}\log\frac{1}{p}\right).$$

Then

$$\Pr[\neg D_1] \le (1-p)^{\frac{1}{p}\left(4\log\frac{1}{p}-200\right)} \le e^{-\left(4\log\frac{1}{p}-200\right)} \in O(p^4).$$

Part 2

Given D_1, let D_2 be the event that more than $24\log q$ generators are selected. Let $X \sim \mathrm{Bin}(q,p)$. Since

$$q \le \left(\frac{4}{p}\log\frac{1}{p}\right)\log\left(\frac{4}{p}\log\frac{1}{p}\right),$$

then

$$24 \log q \leq 24 \log \left(\frac{4}{p}\right)^4 \leq 100 \log \frac{4}{p}.$$

Also, since

$$q \geq \frac{200}{p} \log \frac{4}{p},$$

then

$$\mathrm{E}[X] = qp \geq 200 \log \frac{4}{p}.$$

Remember Chernoff's bound:

$$\Pr[X \leq \mathrm{E}[X] - \lambda] \leq e^{-\frac{\lambda^2}{2\mathrm{E}[X]}} \tag{1.6}$$

Thus, using $\lambda = \frac{\mathrm{E}[X]}{2}$,

$$\Pr[\neg D_2] \leq \Pr\left[X \leq \mathrm{E}[X] - \frac{\mathrm{E}[X]}{2}\right] = e^{-\frac{\mathrm{E}[X]}{8}} \leq e^{-25 \log \frac{4}{p}} \in O(p^{25}). \tag{1.7}$$

Part 3

Finally, assume D_1 and D_2. Let $\mathcal{A}$ be the set of generators of chosen before q. Since the generators are chosen randomly and $|\mathcal{A}| \geq 24 \log q$, we can apply Lemma 1.1.1 to $\mathbb{Z}_q \cong \mathrm{Ap}(\langle \mathcal{A} \cup \{q\}\rangle, q)$. Consider the event D_3 that $\mathrm{Ap}(\langle \mathcal{A} \cup \{q\}\rangle, q)$ will be completed before step

$$12q \log q \in O\left(\frac{1}{p}\left(\log \frac{1}{p}\right)^3\right).$$

Applying Lemma 1.1.1 (Equation 1.5), we have that

$$\begin{aligned}
\Pr[\neg D_3] &\leq \frac{1}{q^{10}} + \frac{6 \log q}{q^4} \\
&\leq \frac{1}{\left(\frac{96}{p} \log \frac{4}{p}\right)^{10}} + \frac{6 \log \left(\frac{4}{p}\right)^4}{\left(\frac{96}{p} \log \frac{4}{p}\right)^4} \in O(p^4).
\end{aligned}$$

Thus, there exists $K > 0$ such that

$$\lim_{p \to 0} \Pr\left[F(\langle \mathcal{A} \cup \{q\}\rangle) \leq \frac{K}{p}\left(\log \frac{1}{p}\right)^3\right] = 1.$$

Since $F(\mathcal{S}) \leq F(\langle \mathcal{A} \cup \{q\}\rangle)$, we conclude that

$$\lim_{p \to 0} \Pr\left[F(\mathcal{S}) \leq \frac{K}{p}\left(\log \frac{1}{p}\right)^3\right] = 1. \quad \square$$

Now that the expected value of the Frobenius number can be bounded by

$$\mathrm{E}[F(\mathcal{S})] = \mathrm{E}[F(\mathcal{S}) | D_1 \wedge D_2 \wedge D_3] \Pr[D_1 \wedge D_2 \wedge D_3]$$

$$
\begin{aligned}
&+ \mathrm{E}[F(\mathcal{S})|\neg(D_1 \wedge D_2 \wedge D_3)]\Pr[\neg(D_1 \wedge D_2 \wedge D_3)] \\
&\leq \mathrm{E}[F(\mathcal{S})|D_1 \wedge D_2 \wedge D_3] \\
&+ \mathrm{E}[F(\mathcal{S})|\neg(D_1 \wedge D_2 \wedge D_3)]\Pr[\neg(D_1 \wedge D_2 \wedge D_3)] \\
&= \frac{K}{p}\left(\log \frac{1}{p}\right)^3 + \mathrm{E}[F(\mathcal{S})|\neg(D_1 \wedge D_2 \wedge D_3)]O(p^4).
\end{aligned}
$$

Now,

$$
\begin{aligned}
\mathrm{E}[F(\mathcal{S})|\neg(D_1 \wedge D_2 \wedge D_3)] &\leq \mathrm{E}[(\min\{(2n)^2|2n \text{ and } 2n+1 \text{ are selected}\})] \\
&= \sum_{n=0}^{\infty}(2n)^2(1-p^2)^{n-1}p^2 \\
&= p^2\sum_{n=0}^{\infty}(2n)^2 e^{-p^2(n-1)} \\
&\leq p^2\int_0^{\infty}(2x)^2 e^{-p^2(x-1)}\,\mathrm{d}x \\
&= \frac{8e^{p^2}}{p^4}.
\end{aligned}
$$

Therefore,

$$
\mathrm{E}[F(\mathcal{S})] \leq \frac{K}{p}\left(\log \frac{1}{p}\right)^3 + C,
$$

for some constants K and C that do not depend on p.

Corollary 1.1.1. *Let $\psi(x)$ be a function for which $x(\log x)^2 \in o(\psi(x))$ .Then*

$$
\lim_{p\to 0}\Pr\left[g(\mathcal{S}) \leq \psi\left(\frac{1}{p}\right)\right] = 1.
$$

Proof. Use Proposition **??**. □

Corollary 1.1.2. *Let $\varphi(x)$ be a function for which $(\log x)^2 \in o(\varphi(x))$ .Then*

$$
\lim_{p\to 0}\Pr\left[e(\mathcal{S}) \leq \varphi\left(\frac{1}{p}\right)\right] = 1.
$$

Proof. Since

$$
\lim_{p\to 0}\Pr\left[F(\mathcal{S}) \leq \psi\left(\frac{1}{p}\right)\right] = 1,
$$

and the maximal element of the minimal generating set is at most $2F(\mathcal{S})$, the elements of the minimal generating set are chosen before step $2\psi\left(\frac{1}{p}\right)$ with high probability. Since

$$
\left|\mathcal{A} \cap \left\{1, \ldots, \left\lfloor 2\psi\left(\frac{1}{p}\right)\right\rfloor\right\}\right| \sim \mathrm{Bin}\left(\left\lfloor 2\psi\left(\frac{1}{p}\right)\right\rfloor, p\right),
$$

by the bound on the right tail of the binomial distribution (Proposition **??**), we have that

$$
\lim_{p\to 0}\Pr\left[e(\mathcal{S}) \leq (3p)\psi\left(\frac{1}{p}\right)\right] = 1.
$$

Thus, if $\varphi(x) = \frac{3}{x}\psi\,(x)$, then $(\log x)^2 \in \varphi(x)$ and

$$\lim_{p \to 0} \Pr \left[e(\mathcal{S}) \leq \varphi\left(\frac{1}{p}\right) \right] = 1. \quad \square$$