ISRAEL JOURNAL OF MATHEMATICS, Vol. 39, No. 3, 1981

# FINITE ABELIAN GROUP COHESION

BY
P. ERDÖS AND B. SMITH

ABSTRACT
This paper studies the evenness of set arithmetic in a finite abelian group.

Let $G$ be a finite abelian group. We use $\#$ to denote cardinality. Let $\# G = p$. For $A, B \subset G$ let $m(x, A, B) = \#\{(a,b)\colon a + b = x,\ a \in A,\ b \in B\}$. For $E \subset G$ let $E'$ denote its complement.

THEOREM. (Cohesion Equation).

$$\sum_{x \in G} |m(x, E, E) + m(x, E', E') - m(x, E, E') - m(x, E', E)|^2$$

$$= \sum_{x \in G} |m(x, E, -E) + m(x, E', -E') - m(x, E, -E') - m(x, E', -E)|^2.$$

PROOF. Let $\Gamma$ denote the dual group of $G$. Let

$$f(x) = \begin{cases} 1 & \text{if } x \in E, \\ -1 & \text{if } x \in E'. \end{cases}$$

Let $\tilde{f}(x) = f(-x)$. The Cohesion Equation states

$$\sum_{x \in G} |f * f(x)|^2 = \sum_{x \in G} |f * \tilde{f}(x)|^2.$$

Let $\hat{f}(\gamma) = \Sigma_{x \in G} f(x)\gamma(-x)$ for $\gamma \in \Gamma$. Then

$$\sum_{x \in G} |f * f(x)|^2 = \frac{1}{p} \sum_{\gamma \in \Gamma} |\hat{f}^2(\gamma)|^2 = \frac{1}{p} \sum_{\gamma \in \Gamma} |\hat{f}(\gamma)\bar{\hat{f}}(\gamma)|^2 = \sum_{x \in G} |f * \tilde{f}(x)|^2.$$

Received April 4, 1979

THEOREM 1.

$$\min_{E\subset G}\max_{x\in G}|m(x,E,E)+m(x,E',E')-2m(x,E,E')|\geqq p^{1/2}.$$

PROOF. Consider the right hand side of the Cohesion Equation.

$$\sum|m(x,E,-E)+m(x,E',-E')-m(x,E,-E')-m(x,E',-E)|^2$$

$$\geqq|m(0,E,-E)+m(0,E',-E')-m(0,E,-E')-m(0,E',-E)|^2=p^2.$$

THEOREM II. *Let $\lambda>\frac{1}{2}$. Let $G$ be a finite group with no elements of order 2. Then*

$$\min_{E\subset G}\max_{x\in G}|m(x,E,E)+m(x,E',E')-2m(x,E,E')|\leqq Kp^{\lambda}$$

*($K$ depends only on $\lambda$).*

The proof of Theorem II requires 3 Lemmas.

For the remainder of the argument let $\# G=n+1$, and let there be no elements of order 2 in $G$. We consider all ways of writing $G\backslash\{0\}=E\cup F$ with $\# E=\# F=n/2$. Let $\alpha=(n-1)/n$. For $x\in G\backslash\{0\}$ we see that $\alpha n/4$ is the expected value of $m(x,E,F)$, since $(G\backslash\{0\})\times(G\backslash\{0\})$ has cardinality $n^2$ and $(G\backslash\{0\})+(G\backslash\{0\})$ represents $x(\neq 0)$ $n-1$ times. When $E$ is understood we use $m(x)$ for $m(x,E,F)$. Let

$$A(r,s)=\sum\binom{k_1+\cdots+k_s}{k_1}\binom{k_2+\cdots+k_s}{k_2}\cdots\binom{k_s}{k_s}\frac{1}{j_1!}\frac{1}{j_2!}\cdots\frac{1}{j_t!}$$

where the summation is over $s$-tuples of integers, $k_1,\cdots,k_s$, satisfying: $k_1+\cdots+k_s=r$. $k_1\geqq k_2\geqq\cdots\geqq k_s\geqq 1$. $k_1=\cdots=k_{j_1}$; $k_{j_1+1}=\cdots=k_{j_1+j_2};\cdots;$ $k_{j_1+\cdots+j_{t-1}+1}=\cdots=k_{j_1+\cdots+j_t}$.

LEMMA 1.

$$\text{Expectation}\left(\sum_{\substack{x\in G\\ x\neq 0}}(m(x)-\alpha n/4)^p\right)=$$

$$=E\left(\sum_{\substack{x\in G\\ x\neq 0}}\left\{\binom{p}{p}(m(x))^p+\binom{p}{p-1}(-1)\left(\frac{\alpha n}{4}\right)(m(x))^{p-1}\right.\right.$$

$$\left.\left.+\cdots+\binom{p}{0}(-1)^p\left(\frac{\alpha n}{4}\right)^p\right\}\right)$$

$$
\begin{aligned}
&= \binom{p}{p}\left\{\left(\frac{\alpha}{4}\right)^{p} n^{2}(n-2(1))(n-2(2))\cdots(n-2(p-1))\right. \\
&\qquad + \left(\frac{\alpha}{4}\right)^{p-1} A(p,p-1)n^{2}(n-2(1))\cdots(n-2(p-2)) \\
&\qquad + \left(\frac{\alpha}{4}\right)^{p-2} A(p,p-2)n^{2}(n-2(1))\cdots(n-2(p-3)) \\
&\qquad \left. + \cdots + \left(\frac{\alpha}{4}\right) A(p,1)n^{2}\right\} \\
&+(-1)^{1}\binom{p}{p-1}\left\{\left(\frac{\alpha}{4}\right)^{p} nn^{2}(n-2(1))\cdots(n-2(p-2))\right. \\
&\qquad\qquad + \left(\frac{\alpha}{4}\right)^{p-1} A(p-1,p-2)nn^{2}(n-2(1))\cdots(n-2(p-3)) \\
&\qquad\qquad \left. + \cdots + \left(\frac{\alpha}{4}\right)^{2} A(p-1,1)nn^{2}\right\} \\
&+(-1)^{2}\binom{p}{p-2}\left\{\left(\frac{\alpha}{4}\right)^{p} n^{2}n^{2}(n-2(1))\cdots(n-2(p-3))\right. \\
&\qquad\qquad + \left(\frac{\alpha}{4}\right)^{p-1} A(p-2,p-3)n^{2}n^{2}(n-2(1))\cdots(n-2(p-4)) \\
&\qquad\qquad \left. + \cdots + \left(\frac{\alpha}{4}\right)^{3} A(p-2,1)n^{2}n^{2}\right\} \\
&+\cdots+ \\
&\qquad +(-1)^{p-1}\binom{p}{1}\left\{\left(\frac{\alpha}{4}\right)^{p} n^{p-1}n^{2}\right\} + (-1)^{p}\binom{p}{0}\left\{\left(\frac{\alpha}{4}\right)^{p} n^{p}n\right\}.
\end{aligned}
$$

**Proof.** The proof of this lemma is done in analogy to the proof of the lemma on page 130 of [1]. The $p=2$ argument is completely general.

$$
E\left(\sum_{\substack{x\in G\\ x\neq 0}}\left(m(x)-\frac{\alpha n}{4}\right)^{2}\right)=
$$

$$
E\left(\sum_{\substack{x\in G\\ x\neq 0}}\left\{\binom{2}{2}(m(x))^{2}-\binom{2}{1}m(x)\left(\frac{\alpha n}{4}\right)+\binom{2}{0}\left(\frac{\alpha n}{4}\right)^{2}\right\}\right),
$$

$$
m(x)=\sum_{\varepsilon_1 a_1+\varepsilon_2 a_2+\cdots+\varepsilon_{n/2}a_{n/2}+\delta_1 b_1+\cdots+\delta_{n/2}b_{n/2}=x} 1
$$

where $(a_1, a_2, \cdots, a_{n/2})$, $(b_1, b_2, \cdots, b_{n/2})$ represent choices of $E, F$ and where $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{n/2})$, $(\delta_1, \cdots, \delta_{n/2})$ represent $n/2$-tuples consisting of one 1 and $(n/2 - 1)$ zeroes. So,

$$\sum_{\substack{x \in G \\ x \neq 0}} E((m(x))^2) = \sum P(0 \neq \varepsilon_1 a_1 + \cdots + \varepsilon_{n/2} a_{n/2} + \delta_1 b_1 + \cdots + \delta_{n/2} b_{n/2}$$

$$= \varepsilon'_1 a_1 + \cdots + \varepsilon'_{n/2} a_{n/2} + \delta'_1 b_1 + \cdots + \delta'_{n/2} b_{n/2}$$

where $(\varepsilon_1, \cdots, \varepsilon_{n/2})$, $(\delta_1, \cdots, \delta_{n/2})$, $(\varepsilon'_1, \cdots, \varepsilon'_{n/2})$, $(\delta'_1, \cdots, \delta'_{n/2})$ are allowed to run independently.

We abbreviate $(\varepsilon_1, \cdots, \varepsilon_{n/2}) = \varepsilon$, $(\delta_1, \cdots, \delta_{n/2}) = \delta$, $(a_1, \cdots, a_{n/2}) = a$, $(b_1, \cdots, b_{n/2}) = b$, $\varepsilon_1 a_1 + \cdots + \varepsilon_{n/2} a_{n/2} = (\varepsilon, a)$ and $\delta_1 b_1 + \cdots + \delta_{n/2} b_{n/2} = (\delta, b)$. We use $P(A \mid B)$ to denote the conditional probability of $A$ given $B$.

If $\varepsilon = \varepsilon'$, $\delta = \delta'$, then $P((\varepsilon, a) + (\delta, b) = (\varepsilon', a) + (\delta', b)) = 1$ and $P(0 \neq (\varepsilon, a) + (\delta, b)) = \alpha$. So

$$\sum_{\substack{\varepsilon = \varepsilon' \\ \delta = \delta'}} P(0 \neq (\varepsilon, a) + (\delta, b) = (\varepsilon', a) + (\delta', b)) = \left(\frac{n}{2}\right)^2 \alpha.$$

If $\varepsilon = \varepsilon'$ and $\delta \neq \delta'$ or $\varepsilon \neq \varepsilon'$ and $\delta = \delta'$, then $P((\varepsilon, a) + (\delta, b) = (\varepsilon', a) + (\delta', b)) = 0$. So

$$\sum_{\substack{(\varepsilon = \varepsilon' \text{ and } \delta \neq \delta') \\ \text{or} \\ (\varepsilon \neq \varepsilon' \text{ and } \delta = \delta')}} P(0 \neq (\varepsilon, a) + (\delta, b) = (\varepsilon', a) + (\delta', b)) = 0.$$

If $\varepsilon \neq \varepsilon'$ and $\delta \neq \delta'$, then

$$P(0 \neq (\varepsilon, a) + (\delta, b) = (\varepsilon', a) + (\delta', b)) =$$

$$= P(0 \neq (\varepsilon, a) + (\delta, b)) \cdot P((\varepsilon, a) + (\delta, b) \neq 2(\varepsilon', a) \mid 0 \neq (\varepsilon, a) + (\delta, b))$$

$$\cdot P(0 \neq (\varepsilon, a) + (\delta, b) - (\varepsilon', a) \mid 0 \neq (\varepsilon, a) + (\delta, b), (\varepsilon, a) + (\delta, b) \neq 2(\varepsilon', a))$$

$$\cdot P((\delta', b) = (\varepsilon, a) + (\delta, b) - (\varepsilon', a) \mid 0 \neq (\varepsilon, a) + (\delta, b),$$

$$(\varepsilon, a) + (\delta, b) \neq 2(\varepsilon', a), 0 \neq (\varepsilon, a) + (\delta, b) - (\varepsilon', a)) = \alpha \cdot \frac{n-3}{n-2} \cdot \alpha \cdot \frac{1}{n-3}$$

$$\sum_{\substack{\varepsilon \neq \varepsilon' \\ \delta \neq \delta'}} P(0 \neq (\varepsilon, a) + (\delta, b) = (\varepsilon', a) + (\delta', b)) = \left(\frac{n}{2}\right)^2 \left(\frac{n}{2} - 1\right)^2 \alpha^2 \frac{1}{n-2}.$$

Hence,

$$E\left(\sum\left(m(x)-\frac{\alpha n}{4}\right)^2\right)=\binom{2}{2}\left\{\left(\frac{\alpha}{4}\right)^2 n^2(n-2(1))+\frac{\alpha}{4}n^2\right\}-\binom{2}{1}\left\{\left(\frac{\alpha}{4}\right)^2 n\cdot n^2\right\}$$
$$+\binom{2}{0}\left\{\left(\frac{\alpha}{4}\right)^2 n^2\cdot n\right\}.$$

Let $f_r(s)=\sum_{1\le j_1<j_2<\cdots<j_r\le s}(j_1 j_2\cdots j_r)$ where $j_1,\cdots,j_r$ are integers.

LEMMA 2. (Cohesion Identities). *The* $1/4^{2k}$ *identities are true.*

$$\sum_{j=0}^{2k}(-1)^j\binom{2k}{j}A(j,j)=0,$$

$$\sum_{j=2}^{2k}(-1)^j\binom{2k}{j}A(j,j)f_1(j-1)=0,$$

$$\vdots$$

$$\sum_{j=k}^{2k}(-1)^j\binom{2k}{j}A(j,j)f_{k-1}(j-1)=0.$$

*The* $1/4^{2k-1}$ *identities are true.*

$$\sum_{j=2}^{2k}(-1)^j\binom{2k}{j}A(j,j-1)=0,$$

$$\sum_{j=3}^{2k}(-1)^j\binom{2k}{j}A(j,j-1)f_1(j-2)=0,$$

$$\vdots$$

$$\sum_{j=k}^{2k}(-1)^j\binom{2k}{j}A(j,j-1)f_{k-2}(j-2)=0.$$

*The* $1/4^{2k-2}$ *identities are true.*

$$\sum_{j=3}^{2k}(-1)^j\binom{2k}{j}A(j,j-2)=0,$$

$$\sum_{j=4}^{2k}(-1)^j\binom{2k}{j}A(j,j-2)f_1(j-3)=0,$$

$$\vdots$$

$$\sum_{j=k}^{2k}(-1)^j\binom{2k}{j}A(j,j-2)f_{k-3}(j-3)=0.$$

$$\cdots$$

*The* $1/4^{k+1}$ *identity is true.*

$$\sum_{j=k}^{2k}(-1)^j\binom{2k}{j}A(j,j-(k-1))=0.$$

PROOF. The proof depends on the following Prelemma.

PRELEMMA. *Let* $k\in\{1,2,3,\cdots\}$. *In the following it is understood that the* $x$ *in* $f_k(x+l)$ *satisfies* $x$ *is an integer and* $x\geqq k-l$. $a_0,a_1,a_2,\cdots$ *are constants but are allowed to change as we go from one identity to the next. For each* $l\in\{-1,0,1,2,\cdots\}$ *we have the following list of identities. (We are only interested in these identities for* $x\in\{0,1,2,\cdots\}$.)

$$f_1(x+l)=a_0\cdot 1+a_1x+a_2x(x-1),$$

$$\vdots$$

$$f_l(x+l)=a_0\cdot 1+a_1x+a_2x(x-1)+\cdots+a_{2l}x(x-1)\cdots(x-(2l-1)),$$

$$f_{l+1}(x+l)=a_1x+a_2x(x-1)+\cdots+a_{2(l+1)}x(x-1)\cdots(x-(2(l+1)-1)),$$

$$\vdots$$

$$f_{l+j}(x+l)=a_jx(x-1)\cdots(x-(j-1))+\cdots+a_{2(l+j)}x(x-1)\cdots(x-(2(l+j)-1)).$$

$$\vdots$$

PROOF OF THE PRELEMMA. The $l=-1$ identities are proved by induction using the equation $f_k(x)=xf_{k-1}(x-1)+f_k(x-1)$. The others follow from the $-1$ identities and substitution.

We return to the proof of the Cohesion Identities

The 1st identity in each $1/4^j$ list can be gotten directly from multinomial expansions.

We are left the problem of showing

$$\sum_{j=r}^{2k}(-1)^j\binom{2k}{j}A(j,j-t)f_s(j-t-1)=0 \tag{1}$$

where $0\leqq t$; $t+2\leqq r\leqq k$; $s=r-(t+1)$. We must use multinomial expansions, differentiation and the Prelemma to see this.

Given $t\geqq 0$ we call an expression of the form $u=h_1+\cdots+h_{u-t}$; $h_1\geqq\cdots\geqq h_{u-t}\geqq 1$; $h_1,\cdots,h_{u-t}\in Z$ a *t-partition*. We say that the $t$ partitions

$$u=h_1+\cdots+h_{u-t},\qquad v=i_1+\cdots+i_{v-t}$$

are *similar* if (let us suppose $u-t<v-t$) $h_1=i_1$, $h_2=i_2,\cdots,h_{u-t}=i_{u-t}$, $i_{u-t+1}=\cdots=i_{v-t}=1$.

Fix a $t$-decomposition and let $u$ be the smallest number $\geqq r$ with a similar decomposition.

$$u=h_1+\cdots+h_{u-t}.$$

If $h_{u-t}>1$, let $q=0$. Otherwise let $q$ be the largest number satisfying $h_{u-t-q-1}=\cdots=h_{u-t}=1$. Let $u-t-q-2=p$. Our objective is to prove

$$\text{(2)}\quad \sum_{j=u}^{2k}(-1)^j\binom{2k}{j}\binom{j}{h_1}\binom{j-h_1}{h_2}\cdots\binom{j-h_1-\cdots-h_{p-1}}{h_p}f_s(j-t-1)=0.$$

We know:

$$(y+x_1+\cdots+x_p-1)^{2k}=\cdots+\left\{(-1)^{u-q}\binom{2k}{u-q}\binom{u-q}{h_1}\cdots\binom{h_p}{h_p}\right.$$
$$+(-1)^{u-q+1}\binom{2k}{u-q}\cdots\binom{h_p+1}{h_p}y+\cdots+(-1)^u\binom{2k}{u}\binom{u}{h_1}\cdots\binom{h_p+q}{h_p}y^q$$
$$\left.+(-1)^{u+1}\binom{2k}{u+1}\cdots\binom{h_p+q+1}{h_p}y^{q+1}+\cdots\right\}x_1^{h_1}x_2^{h_2}\cdots x_p^{h_p}+\cdots.$$

We apply $D_y^j$ ($D_y^j$ denotes the operator that takes the $j$-th derivative with respect to $y$) to this equation for $j=q,\cdots,2s$. Let $h_1+\cdots+h_p=H$. We arrive at the equations

$$(2k)(2k-1)\cdots(2k-(q-1))(y+x_1+\cdots+x_p-1)^{2k-q}$$
$$=\cdots+\left\{(-1)^u\binom{2k}{u}\cdots\binom{h_p+q}{h_p}q\cdot(q-1)\cdots1\right.$$
$$\left.+(-1)^{u+1}\binom{2k}{u+1}\cdots\binom{h_p+q+1}{h_p}(q+1)q\cdots2y+\cdots\right\}x_1^{h_1}\cdots x_p^{h_p}+\cdots,$$

$$(2k)(2k-1)\cdots(2k-q)(y+x_1+\cdots+x_p-1)^{2k-q-1}$$
$$=\cdots+\left\{(-1)^{u+1}\binom{2k}{u+1}\cdots\binom{h_p+q+1}{h_p}(q+1)q\cdots1\right.$$
$$\left.+(-1)^{u+2}\binom{2k}{u+2}\cdots\binom{h_p+q+2}{h_p}(q+2)(q+1)\cdots2\cdot y+\cdots\right\}x_1^{h_1}\cdots x_p^{h_p}$$
$$+\cdots,$$

$$\cdots$$

$$(2k)\cdots(2k-(2s-1))(y+x_1+\cdots+x_p-1)^{2k-2s}$$
$$=\cdots+\left\{(-1)^{H+2s}\binom{2k}{H+2s}\cdots\binom{h_p+2s}{h_p}(2s)\cdots(1)\right.$$
$$\left.+(-1)^{H+2s+1}\binom{2k}{H+2s+1}\cdots\binom{h_p+2s+1}{h_p}(2s+1)\cdots 2\cdot y+\cdots\right\}x_1^{h_1}\cdots x_p^{h_p}$$
$$+\cdots.$$

We need to note that $h_1+\cdots+h_p \leqq 2k-2s-2$. Among all $t$ partitions the largest value $h_1+\cdots+h_p$ obtains is when $p=t$ and $h_1=\cdots=h_t=2$. Hence $h_1+\cdots+h_p\leqq 2t$. For a fixed $t$ the largest value $s$ can have is $k-(t+1)$. So $2(k-s)\geqq 2t+2$.

So upon setting $y=1$ in the above list of equations and equating coefficients we can conclude that the following expressions are 0:

$$\mathrm{e}_q)\qquad \begin{aligned}&(-1)^u\binom{2k}{u}\cdots\binom{h_p+q}{u}q(q-1)\cdots 1\\ &+(-1)^{u+1}\binom{2k}{u+1}\cdots\binom{h_p+q+1}{h_p}(q+1)q\cdots 2+\cdots\end{aligned}$$

$$\mathrm{e}_{q+1})\qquad \begin{aligned}&(-1)^{u+1}\binom{2k}{u+1}\cdots\binom{h_p+q+1}{h_p}(q+1)q\cdots 1\\ &+(-1)^{u+2}\binom{2k}{u+2}\cdots\binom{h_p+q+2}{h_p}(q+2)\cdots 2+\cdots\end{aligned}$$

$$\cdots$$

$$\mathrm{e}_{2s})\qquad \begin{aligned}&(-1)^{H+2s}\binom{2k}{H+2s}\cdots\binom{h_p+2s}{h_p}(2s)(2s-1)\cdots 1\\ &+(-1)^{H+2s+1}\binom{2k}{H+2s+1}\cdots\binom{h_p+2s+1}{h_p}(2s+1)\cdots 2+\cdots\end{aligned}$$

We need to see that there are constants $a_q,\cdots,a_{2s}$ such that:

$$f_s(x+(s-q)+(u-r))=a_qx(x-1)\cdots(x-(q-1))+\cdots$$
$$+a_{2s}x(x-1)\cdots(x-(2s-1)).$$

If $q=0$, there is no problem since we are allowed $1,x,\cdots,x(x-1),\cdots,(x-(2s-1))$ in our expansion of $f_s$. If $q>0$, then $u=r$. We will be able to find

$a_q, \cdots, a_{2s}$ from the Prelemma if we establish $s-q \geqq -1$. The largest value of $q$ will occur when our partition is

$$r = \underbrace{r_0 + 1 + \cdots + 1}_{r-t}$$

where $r_0 = 1$ if $t = 0$ and $r_0 > 1$ if $t > 0$. So $q = r$ for $t = 0$ and $q = r - t - 1$ for $t > 0$. Now $s - q = r - (t+1) - q$. For $t = 0$, $s - q \geqq r - 1 - r = -1$. For $t > 0$, $s - q \geqq r - (t+1) - (r - t - 1) = 0$.

This finishes the proof of Lemma 2 since the left hand side of (2) is

$$a_q e_q + \cdots + a_{2s} e_{2s} = 0 + 0 + \cdots + 0$$

and (1) is a linear combination of expressions of the form (2).

Lemma 3.

$$E\left(\sum_{\substack{x \in G \\ x \neq 0}} \left(m(x) - \frac{\alpha n}{4}\right)^{2k}\right) \leqq K n^{k+1}$$

*(K depends only on k).*

Proof. This is a calculation using Lemmas 1, 2.

Proof of Theorem II. We must make a computation. Let $\beta = 1/n$. Using the argument of Lemma 1 we have

$$E\left(\left(m(0) - \beta \frac{n^2}{4}\right)^2\right) = \beta \left(\frac{n}{2}\right)^2 \left(\frac{n}{2} - 1\right)^2 \frac{1}{n-3} + \beta \left(\frac{n}{2}\right)^2 - \beta^2 \left(\frac{n^2}{4}\right)^2 = O(n).$$

The Theorem now follows from this computation and Lemma 3.

## Reference

1. P. Erdös and A. Rényi, *Probabilistic methods in group theory*, J. Analyse Math. **14** (1965), 127–138.

Mathematics Institute
Realtanoda U. 13–15
Budapest, Hungary

Department of Mathematics
University of Kentucky
Lexington, KY 40506 USA