How close can we come to a parity function when there isn’t one?

Cristopher Moore  
Computer Science Department  
University of New Mexico  
and the Santa Fe Institute  
moore@cs.unm.edu

Alexander Russell  
Dept. of Computer Science and Engineering  
University of Connecticut  
acr@cse.uconn.edu

May 27, 2010

Abstract

Consider a group \( G \) such that there is no homomorphism \( f : G \to \{\pm 1\} \). In that case, how close can we come to such a homomorphism? We show that if \( f \) has zero expectation, then the probability that \( f(xy) = f(x)f(y) \), where \( x, y \) are chosen uniformly and independently from \( G \), is at most \( 1/2 \left( 1 + 1/\sqrt{d} \right) \), where \( d \) is the dimension of \( G \)'s smallest nontrivial irreducible representation. For the alternating group \( A_n \), for instance, \( d = n - 1 \). On the other hand, \( A_n \) contains a subgroup isomorphic to \( S_{n-2} \), whose parity function we can extend to obtain an \( f \) for which this probability is \( 1/2 \left( 1 + 1/(n^2) \right) \). Thus the extent to which \( f \) can be “more homomorphic” than a random function from \( A_n \) to \( \{\pm 1\} \) lies between \( O(n^{-1/2}) \) and \( \Omega(n^{-2}) \).

The symmetric group \( S_n \) has a parity function, i.e., a homomorphism \( f : S_n \to \{\pm 1\} \), sending even and odd permutations to +1 and −1 respectively. The alternating group \( A_n \), which consists of the even permutations, has no such homomorphism. How close can we come to one? What is the maximum, over all functions \( f : A_n \to \{\pm 1\} \) with zero expectation, of the probability

\[
\Pr_{x,y} [f(x)f(y) = f(xy)],
\]

where \( x \) and \( y \) are chosen independently and uniformly from \( A_n \)?

We give simple upper and lower bounds on this quantity, for groups in general and for \( A_n \) in particular. Our results are easily extended to functions \( f : G \to \mathbb{C} \), but we do not do this here. Our main result is the following:

**Theorem 1.** Let \( G \) be a group, and let \( f : G \to \{\pm 1\} \) such that \( \mathbb{E} f = 0 \). Then

\[
\Pr_{x,y} [f(x)f(y) = f(xy)] \leq \frac{1}{2} \left( 1 + \frac{1}{\sqrt{d}} \right),
\]

where \( d = \min_{\rho \neq 1} d_{\rho} \) is the dimension of the smallest nontrivial irreducible representation of \( G \).

Thus if \( G \) is *quasirandom* in Gowers' sense [1]—that is, if \( \min_{\rho \neq 1} d_{\rho} \) is large—it is impossible for \( f \) to be much more homomorphic than a uniformly random function. For \( A_n \) in particular, the dimension of the smallest nontrivial representation is \( d = n - 1 \), so \( \Pr_{x,y} [f(x)f(y) = f(xy)] - 1/2 = O(1/\sqrt{n}) \).

If \( f \) is a *class function*, i.e., if \( f \) is invariant under conjugation so that \( f(x^{-1}yx) = f(y) \) for all \( x, y \in G \), then we can tighten this bound from \( 1/\sqrt{d} \) to \( 1/d \):
**Theorem 2.** Let $G$ be a group, and let $f : G \to \{\pm 1\}$ be a class function such that $Ef = 0$. Then

$$\Pr_{x,y}[f(x)f(y) = f(xy)] \leq \frac{1}{2} \left( 1 + \frac{1}{d} \right),$$

where $d = \min_{\rho \neq 1} d_\rho$ is the dimension of the smallest nontrivial irreducible representation of $G$.

As a partial converse to these upper bounds, we have

**Theorem 3.** Suppose $G$ has a subgroup $H$ with a nontrivial homomorphism $\phi : H \to \{\pm 1\}$. Then there is a function $f : G \to \{\pm 1\}$ such that $Ef = 0$ and

$$\Pr_{x,y}[f(x)f(y) = f(xy)] \geq \frac{1}{2} \left( 1 + \frac{1}{2} \frac{|H|}{|G|} \left( 1 - \frac{\Norm H}{|G|} \right) + \frac{|H|^2}{|G|^2} \right),$$

where $\Norm H = \{ c : cHc^{-1} \}$ denotes the normalizer of $H$.

If $H$ is normal so that $\Norm H = G$, Theorem 3 gives a bias which is quadratically small as a function of the index $|G|/|H|$. However, in some cases we can do better—for instance, if we can find a set of coset representatives which are involutions:

**Theorem 4.** Suppose $G$ has a subgroup $H$ with a nontrivial homomorphism $\phi : H \to \{\pm 1\}$. Suppose further that it has a set of coset representatives $T$ such that $c^2 = 1$ for all $c \in T$. Then there is a function $f : G \to \{\pm 1\}$ such that $Ef = 0$ and

$$\Pr_{x,y}[f(x)f(y) = f(xy)] \geq \frac{1}{2} \left( 1 + \frac{|H|}{|G|} \right).$$

For instance, $A_n$ has a subgroup $H$ isomorphic to $S_{n-2}$, consisting of permutations of the last $n-2$ elements, with the first two elements switched if necessary to keep the parity even. The index of this subgroup is $|H|/|G| = \binom{n}{2}$. Moreover, there is a set of coset representatives $c$ such that $c^2 = 1$; namely, the permutations which switch the first two elements, setwise, with some other pair. Thus Theorem 4 applies, and the extent to which $f : A_n \to \{\pm 1\}$ can be more homomorphic than a uniformly random function is between $O(n^{-1/2})$ and $\Omega(n^{-2})$. It would be nice to close this gap.

**Proof of Theorem 4.** We rely on nonabelian Fourier analysis, for which we refer the reader to [4]. In order to establish our notation and choice of normalizations, let $f : G \to \mathbb{C}$ and let $\rho : G \to U(d)$ be an irreducible unitary representation of $G$. We adopt the Fourier transform $\hat{f}(\rho) = \sum_x f(x) \rho(x)$ in which case we have the Fourier inversion formula

$$f(x) = \frac{1}{|G|} \sum_\rho d_\rho \text{tr}(\hat{f}(\rho) \rho(x)^\dagger)$$

and the Plancherel formula

$$\langle f, g \rangle = \sum_x f(x)^* g(x) = \frac{1}{|G|} \sum_\rho d_\rho \text{tr}(\hat{f}^\dagger \hat{g}). \quad (1)$$

For two functions $f, g : G \to \mathbb{C}$ we define their convolution $(f * g)(x) = \sum_y f(y) g(y^{-1} x)$. With the above normalization,

$$\hat{f} * \hat{g}(\rho) = \hat{f}(\rho) \cdot \hat{g}(\rho).$$
Now consider a function $f : G \to \{\pm 1\}$ such that $\mathbb{E}f = 0$. We can write the probability that $f$ acts homomorphically on a random pair of elements as an expectation,

$$\Pr_{x,y}[f(x)f(y) = f(xy)] = \frac{1}{2} \left( 1 + \mathbb{E}_{x,y} [f(x)f(y)f(xy)] \right). \quad (2)$$

We have

$$\mathbb{E}_{x,y} [f(x)f(y)f(xy)] = \mathbb{E}_{x,y} [f(x)f(y)g(y^{-1}x^{-1})] = \frac{1}{|G|^2} (f * f * g)(1),$$

where $g(z) = f(z^{-1})$. Observe that $\hat{g}(\rho) = \sum_x f(x^{-1})\rho(x) = \sum_x f(x)\rho(x)^\dagger = \hat{f}(\rho)^\dagger$ and hence, by Fourier inversion,

$$\mathbb{E}_{x,y} [f(x)f(y)f(xy)] = \frac{1}{|G|^3} \sum_{\rho \neq 1} d_{\rho} \text{tr}(\hat{f}(\rho)^\dagger \hat{f}(\rho)^\dagger), \quad (3)$$

where we used the fact that $\hat{f}(1) = |G|\mathbb{E}f = 0$. Everything up to here is essentially identical to the Fourier-analytic treatment of the Blum-Luby-Rubinfeld linearity test [2, 3].

As $NN^\dagger$ is positive semidefinite,

$$\left| \text{tr}(NN^\dagger) \right| \leq \|N\|_{\text{op}} \text{tr}(N^\dagger N) \leq \|N\|_{\text{op}} \cdot \|N\|_{\text{frob}}, \quad (4)$$

where $\|N\|_{\text{op}}$ denotes the operator norm

$$\|N\|_{\text{op}} = \max_v \frac{\langle v, Nv \rangle}{\langle v, v \rangle},$$

and $\|N\|_{\text{frob}}$ denotes the Frobenius norm,

$$\|N\|_{\text{frob}} = \sqrt{\text{tr}(N^\dagger N)} = \sqrt{\sum_{ij} |N_{ij}|^2}.$$

Considering also that, from equation (1),

$$\|f\|^2 = |G|\langle f, f \rangle = \frac{1}{|G|} \sum_{\rho} d_{\rho} \|\hat{f}(\rho)\|_{\text{frob}}^2, \quad (5)$$

we conclude from (3) and (4) that

$$\mathbb{E}_{x,y} [f(x)f(y)f(xy)] \leq \frac{1}{|G|^3} \sum_{\rho \neq 1} d_{\rho} \left\|\hat{f}(\rho)\right\|_{\text{op}} \cdot \left\|\hat{f}(\rho)\right\|_{\text{frob}}^2 \leq \max_{\rho \neq 1} \frac{\|\hat{f}(\rho)\|_{\text{op}}}{|G|^3} \sum_{\rho \neq 1} d_{\rho} \left\|\hat{f}(\rho)\right\|_{\text{frob}}^2 = \max_{\rho \neq 1} \frac{\|\hat{f}(\rho)\|_{\text{op}}}{|G|}. \quad (6)$$
Equation (5) also implies that, for any $\rho$,

$$\|f(\rho)\|_{\text{frob}} \leq \frac{|G|}{\sqrt{d_\rho}}.$$  \hfill (7)  

Since $\|N\|_{\text{op}}$ is $N$’s largest singular value and $\|N\|_{\text{frob}}^2$ is the sum of their squares,

$$\|f(\rho)\|_{\text{op}} \leq \|f(\rho)\|_{\text{frob}}.$$  \hfill (8)  

Equation (7) then becomes

$$\|f(\rho)\|_{\text{op}} \leq \frac{|G|}{\sqrt{d_\rho}}.$$  \hfill (9)  

Along with (6), this implies that

$$\mathbb{E}_{x,y} [f(x)f(y)f(xy)] \leq \max_{\rho \neq 1} \frac{1}{\sqrt{d_\rho}},$$

and combining this with (2) completes the proof. \hfill \Box

Proof of Theorem 2. The proof is the same as that for Theorem 1 except that if $f$ is a class function, then $\hat{f}(\rho)$ is a scalar. That is, for each $\rho$ there is a $c$ such that $\hat{f}(\rho) = c \mathbb{1}$. Equation (8) then becomes

$$\|\hat{f}(\rho)\|_{\text{op}} = |c| = \frac{1}{\sqrt{d_\rho}} \|\hat{f}(\rho)\|_{\text{frob}},$$

and (9) becomes

$$\|\hat{f}(\rho)\|_{\text{op}} \leq \frac{|G|}{d_\rho}.$$  

Along with (6), this implies that

$$\mathbb{E}_{x,y} [f(x)f(y)f(xy)] \leq \max_{\rho \neq 1} \frac{1}{d_\rho},$$

and combining this with (2) completes the proof as before. \hfill \Box

Proof of Theorem 3. Let $\phi : H \rightarrow \{\pm 1\}$ be a homomorphism. We extend $\phi$ to a function $f : G \rightarrow \{\pm 1\}$ in the following way. We choose a set $T$ of coset representatives such that $G$ is a disjoint union of left cosets, $G = \bigcup_{c \in T} cH$, including the trivial coset $H$ where $c = 1$. Note that $T = |G|/|H|$. For the trivial coset, we define $f(h) = \phi(h)$ for all $h \in H$. For each $c \neq 1$, we choose $f(c)$ uniformly from $\{\pm 1\}$, and define $f(ch) = f(c)\phi(h)$ for all $h \in H$. Since $\phi$ is nontrivial, we have $\mathbb{E}_H[\phi] = 0$ and therefore $\mathbb{E}_G[f] = 0$.

We will show that, in expectation over $x, y$ and over our choices of $f(c)$, we have

$$\mathbb{E}[f(x)f(y)f(xy)] \geq \frac{1}{2} \left( \frac{|H|}{|G|} \left( 1 - \frac{|\text{Norm } H|}{|G|} \right) + \frac{|H|^2}{|G|^2} \right).$$  \hfill (10)  

The theorem then follows from (2).
Lemma 1. Let $H < G$. Write $z = xy$, and consider whether $f(z) = f(x)f(y)$. There are two cases. If $y \in H$, then writing $x = ch$ we have
\[
f(z) = f(chy) = f(c)\phi(hy) = f(c)\phi(h)\phi(y) = f(x)f(y).
\]
The probability of this event is $|H|/|G|$, contributing $|H|/|G|$ to the expectation $\mathbb{E}[f(x)f(y)f(xy)]$.

In the other case, $y \in cH$ for some $c \neq 1$. Then $x$ and $z$ cannot be in the same left coset $c'H$ as each other, since writing $x = c'h$, $y = ck$, and $z = c'\ell$ we would have
\[
c'hck = c'\ell
\]
for some $h, k, \ell \in H$. This would imply that $hck \in H$ and therefore $c \in H$, a contradiction.

Now, if $x$ and $z$ are in distinct nontrivial cosets, or if one of $x, z$ is in $H$ but the other is in a nontrivial coset other than $cH$, then $f(x)f(y)f(xy)$ is uniformly random in $\{\pm 1\}$. Thus these events contribute zero to $\mathbb{E}[f(x)f(y)f(xy)]$. This leaves us with two cases: $x \in H$ and $y, z \in cH$, or $x, y \in cH$ and $z \in H$.

We deal with the case $x \in H$ and $y, z \in cH$ first. Writing $x = h$, $y = ck$, and $z = c\ell$ gives
\[
hck = c\ell,
\]
or, rearranging,
\[
c^{-1}hc = \ell k^{-1}.
\]
Then we have
\[
f(z) = f(c)\phi(\ell) = f(c)\phi(\ell k^{-1})\phi(k) = f(c)\phi(c^{-1}hc)\phi(k),
\]
while
\[
f(x)f(y) = f(c)\phi(h)\phi(k).
\]
Thus the question is whether or not
\[
\phi(c^{-1}hc) = \phi(h). \tag{11}
\]
The following lemma shows that this is true with probability at least 1/2 if $h$ is chosen uniformly from $H$ conditioned on $c^{-1}hc \in H$, i.e., uniformly from $H \cap cHc^{-1}$. Therefore, this event contributes at least zero to $\mathbb{E}[f(x)f(y)f(xy)]$.

**Lemma 1.** Let $\phi : H \to \{\pm 1\}$ be a homomorphism and let $c \in G$. Then (11) holds for at least half the elements of $H \cap cHc^{-1}$.

**Proof.** We can define a homomorphism $\psi : H \cap cHc^{-1} \to \{\pm 1\}$ as
\[
\psi(h) = \phi(h)\phi(c^{-1}hc).
\]
Clearly (11) holds if and only if $h \in \ker \psi$, i.e., if $\phi(h) = 1$. But $\ker \psi$ comprises at least half the elements of $H \cap cHc^{-1}$. \hfill \Box

The case $x, y \in cH$ and $z \in H$ is more troublesome. Writing $x = ch$, $y = ck$, and $z = \ell$, we have
\[
chck = \ell.
\]
This event occurs if and only if \( chc \in H \). We then have

\[
f(x)f(y) = f(c)^2 \phi(h)\phi(k) = \phi(h)\phi(k),
\]

while

\[
f(z) = \phi(\ell) = \phi(chc)\phi(k).
\]

Then, analogous to (11), the question is whether

\[
\phi(chc) = \phi(h) . \tag{12}
\]

Unfortunately, it can be the case that \( \phi(chc) = -\phi(c) \) for all \( h \in H \) and all \( 1 \neq c \in T \). For example, let \( G = \{1, c, c^2, c^3\} \cong \mathbb{Z}_4 \) and \( H = \{1, c^2\} \cong \mathbb{Z}_2 \), and let \( \phi \) be the isomorphism from \( H \) to \( \{\pm1\} \). Then \( \phi(chc) = -\phi(h) \) for all \( h \in H \).

This event, that \( chc \in H \) and \( \phi(chc) = -\phi(c) \), contributes a negative term to \( \mathbb{E}[f(x)f(y)f(xy)] \).

We will bound this term by bounding the probability that \( chc \in H \) but \( c \neq 1 \). First consider the following lemma.

**Lemma 2.** Let \( H \) be a subgroup of \( G \), let \( c \in G \), and suppose that \( c \notin \text{Norm}(H) \). Then

\[
|H \cap cHc| \leq \frac{|H|}{2}.
\]

**Proof.** Suppose that \( H \cap cHc \neq \emptyset \). Then there is a pair \( h, k \in H \) such that \( k = chc \), and

\[
cHc = cHc^{-1} \cdot chc = cHc^{-1} \cdot k.
\]

Since \( H = Hk \), we have

\[
H \cap cHc = (H \cap cHc^{-1})k.
\]

In particular,

\[
|H \cap cHc| = |H \cap cHc^{-1}|.
\]

However, if \( c \notin \text{Norm}(H) \) then \( H \cap cHc^{-1} \) is a proper subgroup of \( H \), in which case its cardinality is at most half that of \( H \). \( \square \)

Now note that if \( c' \in H \), then \( c'Hc'^{-1} = cHc^{-1} \). Therefore, each coset \( cH \) is either contained in \( \text{Norm}(H) \) or is disjoint from it. It follows that the probability that a uniformly random \( c \in T \) is in \( \text{Norm}(H) \) is the same as the probability for the entire group, \( |\text{Norm}(H)|/|G| \). Since \( |T| = |G|/|H| \), the number of such \( c \) is

\[
|T \cap \text{Norm}(H)| = \frac{|\text{Norm}(H)|}{|H|}.
\]

Thus we have \( |\text{Norm}(H)|/|H| - 1 \) coset representatives \( c \in \text{Norm}(H) \) other than \( c = 1 \). If we condition on the event that \( x, y \in cH \), each of these \( c \) could conceivably contribute \(-1\) to \( \mathbb{E}[f(x)f(y)f(xy)] \), while Lemma 2 implies that the other \( |G|/|H| - |\text{Norm}(H)|/|H| \) coset representatives contribute at least \(-1/2 \). The total contribution of the case \( x, y \in cH, z \in H \) to \( \mathbb{E}[f(x)f(y)f(xy)] \) is then at least

\[
-\frac{|H|^2}{|G|^2} \left( \frac{|\text{Norm}(H)|}{|H|} - 1 + \frac{1}{2} \left( \frac{|G|}{|H|} - \frac{|\text{Norm}(H)|}{|H|} \right) \right)
\]

\[
= \frac{1}{2} \frac{|H|}{|G|} \left( -1 - \frac{|\text{Norm}(H)|}{|G|} \right) + \frac{|H|^2}{|G|^2}.
\]

Adding the contribution \( |H|/|G| \) from the case \( y \in H \) gives (10) and completes the proof. \( \square \)
Proof of Theorem 4. If $c^2 = 1$, then $cHc = cHc^{-1}$. This changes the troublesome case to the easy one, where (11) holds with probability at least $\frac{1}{2}$ for all $h \in H \cap cHc^{-1}$, and so the event $x, y \in cH$, $z \in H$ contributes at least zero to $\mathbb{E}[f(x)f(y)f(xy)]$. We then have $\mathbb{E}[f(x)f(y)f(xy)] \geq |H|/|G|$ from the case $y \in H$, and the theorem follows from (2).

Note that the premise of Theorem 4 can be weakened considerably: namely, that for all $c$ such that $H \cap cHc \neq \emptyset$, we have $c^2 = k$ for some $k \in H$ with $\phi(k) = 1$.

Acknowledgments We benefited from discussions with Avi Wigderson, who obtained Theorem 1 and generalizations independently, and from his lectures and those of Ben Green at the Bellairs Research Institute of McGill University. This work was supported by the NSF under grants CCF-0829931, 0835735, and 0829917 and by the DTO under contract W911NF-04-R-0009.

References

[1] W. T. Gowers. Quasirandom groups. Combinatorics, Probability and Computing 17(3):363–387, 2008.

[2] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. Journal of Computer and System Sciences, 47(3):549–595, 1993.

[3] Mihir Bellare, Don Coppersmith, Johan Håstad, Marcos A. Kiwi, and Madhu Sudan. Linearity testing in characteristic two. IEEE Transactions on Information Theory, 42(6):1781–1795, 1996.

[4] Jean-Pierre Serre. Linear Representations of Finite Groups. Number 42 in Graduate Texts in Mathematics. Springer-Verlag, 1977.