CS 351: Solutions to Problem Set 6

Let $(G, +)$ and $(H, \star)$ be two finite groups. A function $g : G \longrightarrow H$ is said to be a homomorphism if for all $x, y \in G$, $g(x + y) = g(x) \star g(y)$. For $f, g : G \longrightarrow H$, let $\mathbf{Dist}(f, g) = \Pr_{x \in G}[f(x) \neq g(x)]$, when x is uniformly disributed in G. Let $\mathbf{Lindist}(f)$ denote the distance of f to the nearest homomorphism, that is $\mathbf{Lindist}(f) = \min\{\mathbf{Dist}(f, g) \mid g \text{ is a homomorphism from } G \text{ into } H\}$.

For $f : G \longrightarrow H$, let $\mathbf{Fail}(f) = \Pr[f(x) \star f(y) \neq f(x + y)]$.

Notation. We will freely apply the group inverse operator "a^{-1}" to elements a of G as well as of H; the context will clearly indicate which group inverse is being used.

(1) If $d = \mathbf{Lindist}(f)$, then $\mathbf{Fail}(f) \geq (2/9)d$ for all $d \leq 1/2$.

The proof is essentially what we did in class, reworked in the more general context of finite groups. The only improvement in parameters occurs in the proof of Claim 3. Let $r = \mathbf{Fail}(f)$. If $r \geq 2/9$, then since $d \leq 1$, we trivially have $r \geq 2/9 \geq (2/9)d$. For the rest of the proof we will assume that $r < 2/9$.

For $x \in G$, let $V(x)$ denote the multiset $\{f(x+y) \star f(y)^{-1} \mid y \in G\}$. Define the function $g : G \longrightarrow H$ by letting $g(x)$ denote the most popular value in $V(x)$, with ties broken arbitrarily.

Claim 1. $\mathbf{Dist}(f, g) \leq 2r$.

Let $B = \{x \in G \mid f(x) \neq g(x)\}$, that is, those x for which the most popular element from the multiset $V(x)$ disagrees with $f(x)$. In particular, for every $x \in B$, at least half the values in $V(x)$ disagree with $f(x)$. (Reason: if the most popular element in $V(x)$ occurs at least $|G|/2$ times, then by definitions of B and g, at least half the elements of $V(x)$ disagree with $f(x)$; on the other hand, if the most popular element occurs fewer than $|G|/2$ times, then all the other elements of $V(x)$ disagree with $f(x)$.) Note that $\mathbf{Dist}(f, g) = |B|/|G|$.

Therefore, we have

$$\begin{aligned}
r &= \Pr_{x,y}[f(x) \star f(y) \neq f(x + y)] \\
&= \mathbf{Dist}(f, g) \cdot \Pr_{y}[f(x) \star f(y) \neq f(x + y) \mid x \in B] \\
&\quad + (1 - \mathbf{Dist}(f, g)) \cdot \Pr_{y}[f(x) \star f(y) \neq f(x + y) \mid x \notin B] \\
&\geq \mathbf{Dist}(f, g) \cdot (1/2) + 0 \\
&= \mathbf{Dist}(f, g)/2.
\end{aligned}$$

Claim 2. For every $x \in G$, $\Pr_{y,z\in G}[f(x+y) \star f(y)^{-1} = f(x+z) \star f(z)^{-1}] \geq 1 - 2r$, where y and z are independent and uniformly chosen elements of G.

Let y, w be chosen independently and uniformly at random from G. Let $E_1 \equiv [f(x+y) = f(x+y+w) \star f(w)^{-1}]$, and let $E_2 \equiv [f(y) = f(y+w) \star f(w)^{-1}]$. If E_1 and E_2 both hold, then we have

$$\begin{aligned}
&f(x+y) \star f(y)^{-1} \\
&= (f(x+y+w) \star f(w)^{-1}) \star (f(y+w) \star f(w)^{-1})^{-1} \\
&= f(x+y+w) \star f(w)^{-1} \star f(w) \star f(y+w)^{-1} \qquad \text{since } (a \star b)^{-1} = b^{-1} \star a^{-1} \\
&= f(x+y+w) \star f(y+w)^{-1}.
\end{aligned}$$

Since $x+y$ and w are independent of each other and distributed uniformly in G, the probability that E_1 fails is at most r. Similarly, since y and w are independent of each other and distributed uniformly in G, the probability that E_2 fails is at most r. Thus with probability at least $1-2r$, we have $f(x+y) \star f(y)^{-1} = f(x+y+w) \star f(y+w)^{-1}$. Finally, if z is uniformly distributed in G and independent of y, then the pair (y, z) has the same joint distribution as the pair $(y, y+w)$; therefore, we have

$$\begin{aligned}
&\Pr_{y,z\in G}[f(x+y) \star f(y)^{-1} = f(x+z) \star f(z)^{-1}] \\
&= \Pr_{y,w\in G}[f(x+y) \star f(y)^{-1} = f(x+y+w) \star f(y+w)^{-1}] \\
&\geq 1 - 2r.
\end{aligned}$$

Claim 3. (Self-correction Lemma) For every $x \in G$, $\Pr_{y\in G}[g(x) = f(x+y) \star f(y)^{-1}] \geq 1 - c$, where $c < 1/2$ satisfies $c(1-c) \leq r$.

Note: this is the only step where the proof from class has been tweaked.

The value $g(x)$ is defined to be the most popular value in $V(x)$. Let $q_1 \geq q_2 \geq \ldots$ denote the probability values for each element in $V(x)$, sorted in decreasing order. Thus $\Pr_{y\in G}[g(x) = f(x+y) \star f(y)^{-1}] = q_1$. On the other hand, Claim 2 shows that two randomly and independently chosen values in $V(x)$ agree with probability at least $1 - 2r$, that is, $\sum_i q_i^2 \geq 1 - 2r$. Thus we have $1 - 2r \leq q_1^2 + \sum_{i>1} q_i^2 \leq q_1^2 + \left(\sum_{i>1} q_i\right)^2 = q_1^2 + (1-q_1)^2 = 1 - 2q_1 + 2q_1^2$. In other words, $q_1(1-q_1) \leq r$; furthermore, since $q_1 = q_1 \sum_i q_i \geq \sum_i q_i^2 \geq 1 - 2r \geq 5/9$ since $r < 2/9$, we also have $q_1 > 1/2$. Letting $c = 1 - q_1 < 1/2$, Claim 3 is proved.

Claim 4. For every $x \in G$, $\Pr_{y\in G}[g(x) \neq f(x+y) \star f(y)^{-1}] < 1/3$.

An easy consequence of Claim 3, since $c(1-c) \leq r < 2/9$ and $c < 1/2$, we have $c < 1/3$.

Claim 5. The function $g : G \longrightarrow H$ is a homomorphism, that is, for every $x, y \in G$, $g(x + y) = g(x) \star g(y)$.

Fix $x, y \in G$. Suppose there exists an element $z \in G$ such that the following conditions hold.

(1) $g(x + y) = f(x + y + z) \star f(z)^{-1}$;

(2) $g(x) = f(x + y + z) \star f(y + z)^{-1}$;

(3) $g(y) = f(y + z) \star f(z)^{-1}$.

Then we have $g(x) + g(y) = f(x + y + z) \star f(y + z)^{-1} \star f(y + z) \star f(z)^{-1} = g(x + y)$. We will show that such a z exists by an application of the probabilistic method. If z is chosen uniformly at random from G, then by Claim 4 we have:

(1) $\Pr[g(x + y) \neq f(x + y + z) \star f(z)^{-1}] < 1/3$;

(2) $\Pr[g(x) \neq f(x + y + z) \star f(y + z)^{-1}] < 1/3$;

(3) $\Pr[g(y) \neq f(y + z) \star f(z)^{-1}] < 1/3$.

Thus the probability that any one of them fails to happen is strictly less than 1, which implies that with probability strictly greater than zero, all three events occur; in other words, an element z with the three desired properties exists, and the proof of Claim 5 is complete.

To complete the proof of $\mathbf{Fail}(f) \geq (2/9)\mathbf{Lindist}(f)$, note that it is trivially true if $\mathbf{Fail}(f) \geq 2/9$, and if $\mathbf{Fail}(f) < 2/9$, then Claim 5 shows that g, as defined above, is a homomorphism, and Claim 1 shows that $\mathbf{Lindist}(f) \leq \mathbf{Dist}(f, g) \leq 2\mathbf{Fail}(f)$. Thus, in either case, $\mathbf{Fail}(f) \geq (2/9)\mathbf{Lindist}(f)$.

(2) For $f : \mathbf{Z}_2^n \longrightarrow \mathbf{Z}_2$, $\mathbf{Fail}(f) \geq 3d - 6d^2$ for all $d \leq 1/2$, where $d = \mathbf{Lindist}(f)$, where $\mathbf{Fail}(f)$ and $\mathbf{Lindist}(f)$ are as defined before.

One can work through the steps of the previous proof, but here is a more "finished product" that encapsulates the argument succinctly.

Let $g : \mathbf{Z}_2^n \longrightarrow \mathbf{Z}_2$ denote a linear function closest to f, that is, $\mathbf{Dist}(f, g) = d$. (Note: if $d < 1/4$, then g is unique; Exercise: prove this. If it is not unique, fix g arbitrarily to be one of them.) Let $D = \{x \mid f(x) \neq g(x)\}$, and let $\overline{D} = \mathbf{Z}_2^n \setminus D$; note that $d = |D|/2^n$. For each $x \in \mathbf{Z}_2^n$, define the label of x, $\ell(x)$, to be "B" (for "bad") if $x \in D$, and "G" (for "good") otherwise. Each element (x, y) of $\mathbf{Z}_2^n \times \mathbf{Z}_2^n$ can now be given a label from the set $\{\text{G}, \text{B}\}^3$ as follows: the label of (x, y) is $(\ell(x), \ell(y), \ell(x+y))$. Partition $\mathbf{Z}_2^n \times \mathbf{Z}_2^n$ based on the labels, as described below. For each set in the partition, we use the corresponding lower case letter to denote the fraction of pairs in $\mathbf{Z}_2^n \times \mathbf{Z}_2^n$ that belongs to that set. Thus, for example, $p = |P|/2^{2n}$, and so on.

$P : (\text{B}, \text{B}, \text{B})$; $Q : (\text{B}, \text{G}, \text{G})$; $R : (\text{G}, \text{B}, \text{G})$; $S : (\text{G}, \text{G}, \text{B})$;

$U : (\text{B}, \text{B}, \text{G})$; $V : (\text{B}, \text{G}, \text{B})$; $W : (\text{G}, \text{B}, \text{B})$; $X : (\text{G}, \text{G}, \text{G})$.

Note that the pair (x, y) fails the test "$f(x) + f(y) = f(x + y)$" iff it belongs to one of P, Q, R, S. We will lower bound the number of such pairs as follows. On the one hand, $q + r + v + w = 2d(1 - d)$ (since if (x, y) belongs to one of these four sets, then exactly one of x, y belongs to D). On the other hand, since $p + s + v + w = d$ (these are the pairs (x, y) s.t. $x + y \in D$), we have $v + w \leq d$; therefore, $q + r \geq 2d(1 - d) - d = d - 2d^2$. We will argue that for each of these pairs, we can find two more distinct pairs that will fail the test. Specifically, if $(x, y) \in Q$, then the pairs $(y, x+y)$ and $(x+y, y)$ both have the label $(\text{G}, \text{G}, \text{B})$, and therefore belong to S; similarly, if $(x, y) \in R$, then both $(x, x+y)$ and $(x+y, x)$ belong to S. Thus, a total of at least $3(d - 2d^2) = 3d - 6d^2$ fraction of all pairs fail the linearity test.

Linearity Testing Analysis via the Fourier Transform

Consider the space of functions from $\mathbf{Z}_2^n$ into $\mathbf{R}$, that is, the set $\mathcal{F}_n = \{f : \mathbf{Z}_2^n \longrightarrow \mathbf{R}\}$. Each $f \in \mathcal{F}_n$ can be thought of as a vector of 2^n real values. Thus it is easy to see that $\mathcal{F}_n$ is a vector space of dimension 2^n, whose standard basis consists of the so-called "delta" functions: for each $x \in \mathbf{Z}_2^n$, $e_x : \mathbf{Z}_2^n \longrightarrow \mathbf{Z}_2$ is defined by

$$e_x(y) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$$

We will endow this space with the inner product defined by $(f, g) = 2^{-n} \sum_{x \in \mathbf{Z}_2^n} f(x)g(x)$. For convenience, we will view a Boolean function on $\mathbf{Z}_2^n$ as taking on the values $\{-1, +1\}$.

The *Fourier basis* for $\mathcal{F}_n$ is given by the functions χ_α, for $\alpha \in \mathbf{Z}_2^n$, defined as follows. Let $\langle \cdot, \cdot \rangle$ denote the inner product in $\mathbf{Z}_2^n$, that is, for $x, y \in \mathbf{Z}_2^n$, $\langle x, y \rangle \in \mathbf{Z}_2$, and is given by $\langle x, y \rangle = \sum_i x_i y_i$ (where the sum and product are modulo 2 operations). For $\alpha \in \mathbf{Z}_2^n$, the function $\chi_\alpha : \mathbf{Z}_2^n \longrightarrow \mathbf{R}$ is defined by $\chi_\alpha(x) = (-1)^{\langle x, \alpha \rangle}$.

Fact (1) For any $\alpha, \gamma, x, y \in \mathbf{Z}_2^n$, we have:

$\chi_\alpha(x)\chi_\gamma(x) = (-1)^{\langle x,\alpha\rangle}(-1)^{\langle x,\gamma\rangle} = (-1)^{\langle x,\alpha\rangle + \langle x,\gamma\rangle} = (-1)^{\langle x,\alpha+\gamma\rangle} = \chi_{\alpha+\gamma}(x)$.

Similarly, $\chi_\alpha(x)\chi_\alpha(y) = (-1)^{\langle x,\alpha\rangle}(-1)^{\langle y,\alpha\rangle} = (-1)^{\langle x,\alpha\rangle + \langle y,\alpha\rangle} = (-1)^{\langle x+y,\alpha\rangle} = \chi_\alpha(x + y)$.

Problem (3c), aka **Fact (2)** For $\alpha \neq 0^n$, $\sum_x \chi_\alpha(x) = 0$.

This is really trivial; for completeness, here is a messy argument:

$2^{-n} \sum_x \chi_\alpha(x) = \Pr_x[\langle x, \alpha \rangle = 0] - \Pr_x[\langle x, \alpha \rangle = 1]$. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $x = (x_1, \ldots, x_n)$. Let i denote the largest integer such that $\alpha_i \neq 0$; since $\alpha \neq 0^n$, such an i exists. Note that $\alpha_i = 1$. Now $\langle x, \alpha \rangle = \sum_j \alpha_j x_j = 0$ iff $\alpha_i x_i = \sum_{j<i} \alpha_j x_j$, where the sums are in $\mathbf{Z}_2$. For any fixed value $b \in \{0, 1\}$ of $\sum_{j<i} \alpha_j x_j$, $\alpha_i x_i = b$ iff $x_i = b$, which happens with probability exactly 1/2. Thus $\Pr_x[\langle x, \alpha \rangle = 0] = \Pr_x[\langle x, \alpha \rangle = 1]$, which implies that $2^{-n} \sum_x \chi_\alpha(x) = 0$.

(3a) For every $\alpha \in \mathbf{Z}_2^n$, since $\chi_\alpha(x) \in \{-1, +1\}$ for any x, we have:

$$(\chi_\alpha, \chi_\alpha) = 2^{-n} \sum_x \chi_\alpha(x)\chi_\alpha(x) = 2^{-n} \sum_x 1 = 1.$$

(3b) For $\alpha \neq \beta \in \mathbf{Z}_2^n$, Fact (1) implies that

$$(\chi_\alpha, \chi_\beta) = 2^{-n} \sum_x \chi_\alpha(x)\chi_\beta(x) = 2^{-n} \sum_x \chi_{\alpha+\beta}(x),$$

which equals 0 by Fact (2) since $\alpha \neq \beta$ implies that $\alpha + \beta \neq 0^n$.

Thus there are 2^n ($= \dim(\mathcal{F}_n)$) distinct functions χ_α, which form an orthonormal basis of $\mathcal{F}_n$. This is called the *Fourier basis* of $\mathcal{F}_n$. Now any function $f : \mathbf{Z}_2^n \longrightarrow \mathbf{R}$ may be expressed in this basis as $f = \sum_{\alpha \in \mathbf{Z}_2^n} \widehat{f}_\alpha \chi_\alpha$. The orthonormality of the Fourier basis implies that for each α, $\widehat{f}_\alpha = (f, \chi_\alpha) = 2^{-n} \sum_x f(x)\chi_\alpha(x)$. The coefficients $\{\widehat{f}_\alpha\}_{\alpha \in \mathbf{Z}_2^n}$ are called the *Fourier coefficients* of f.

(3d) Plancherel theorem: For $f, g \in \mathcal{F}_n$, we have:

$$\begin{aligned}
\sum_\alpha \widehat{f}_\alpha \widehat{g}_\alpha &= \sum_\alpha \left(2^{-n}\sum_x f(x)\chi_\alpha(x)\right)\left(2^{-n}\sum_y g(y)\chi_\alpha(y)\right) \\
&= \sum_\alpha 2^{-2n}\left(\sum_x f(x)g(x)\chi_\alpha(x)\chi_\alpha(x) + \sum_{x\neq y} f(x)g(y)\chi_\alpha(x)\chi_\alpha(y)\right) \\
&= \sum_\alpha 2^{-2n}\sum_x f(x)g(x) + \sum_\alpha 2^{-2n}\sum_{x\neq y}\chi_\alpha(x+y) \\
&= \left(2^{-n}\sum_\alpha (f,g)\right) + 2^{-2n}\sum_{x\neq y}\sum_\alpha \chi_\alpha(x+y) \\
&= (f,g) + 2^{-2n}\sum_{x\neq y} 0 \qquad \text{similar to Fact (2) since } x+y\neq 0^n \\
&= (f,g).
\end{aligned}$$

A consequence of the Plancherel theorem is *Parseval's identity,* which, for the special case of Boolean functions $f : \mathbf{Z}_2^n \longrightarrow \{-1,+1\}$, implies that $\sum_\alpha \widehat{f}_\alpha^2 = 1$.

For $f, g : \mathbf{Z}_2^n \longrightarrow \{-1,+1\}$, the following facts are straightforward, using the fact that with our ± 1 convention for Boolean values, the test "$f(x)+f(y) = f(x+y)$" can be rewritten as "$f(x)f(y)f(x+y) = 1$."

- $\mathbf{Fail}(f) = 2^{-2n}\sum_{x,y}\left(\frac{1-f(x)f(y)f(x+y)}{2}\right)$, since the quantity inside the summation is 1 if the test fails at x, y, and zero otherwise;
- $\mathbf{Agreement}(f,g) = 2^{-n}\sum_x (1+f(x)g(x))/2$;
- $\mathbf{Disagreement}(f,g) = 2^{-n}\sum_x (1-f(x)g(x))/2$;
- $\mathbf{Agreement}(f,g) - \mathbf{Disagreement}(f,g) = 2^{-n}\sum_x f(x)g(x) = (f,g)$;
- $\widehat{f}_\alpha = (f,\chi_\alpha) = \mathbf{Agreement}(f,\chi_\alpha) - \mathbf{Disagreement}(f,\chi_\alpha) = 1 - 2\cdot\mathbf{Disagreement}(f,\chi_\alpha)$.
- $\max_\alpha \widehat{f}_\alpha = \max_\alpha (f,\chi_\alpha) = \min_\alpha 1 - 2\cdot\mathbf{Disagreement}(f,\chi_\alpha) = 1 - 2\cdot\mathbf{Lindist}(f)$.

(3e) $\mathbf{Fail}(f) \geq \mathbf{Lindist}(f)$.

$$
\begin{aligned}
1 - 2 \cdot \mathbf{Fail}(f) &= 2^{-2n} \sum_{x,y} f(x) f(y) f(x+y) \\
&= 2^{-2n} \sum_{x,y} \left(\sum_{\alpha} \widehat{f}_\alpha \chi_\alpha(x) \right) \left(\sum_{\beta} \widehat{f}_\beta \chi_\beta(y) \right) \left(\sum_{\gamma} \widehat{f}_\gamma \chi_\gamma(x+y) \right) \\
&= 2^{-2n} \sum_{x,y} \left(\sum_{\alpha} \widehat{f}_\alpha \chi_\alpha(x) \right) \left(\sum_{\beta} \widehat{f}_\beta \chi_\beta(y) \right) \left(\sum_{\gamma} \widehat{f}_\gamma \chi_\gamma(x) \chi_\gamma(y) \right) \\
&= 2^{-2n} \sum_{\alpha,\beta,\gamma} \widehat{f}_\alpha \widehat{f}_\beta \widehat{f}_\gamma \sum_{x,y} \chi_\alpha(x) \chi_\beta(y) \chi_\gamma(x) \chi_\gamma(y) \\
&= \sum_{\alpha,\beta,\gamma} \widehat{f}_\alpha \widehat{f}_\beta \widehat{f}_\gamma \ (\chi_\alpha, \chi_\gamma)(\chi_\beta, \chi_\gamma)
\end{aligned}
$$

By (3a) and (3b), the quantity $(\chi_\alpha, \chi_\gamma)$ equals zero if $\alpha \neq \gamma$, and equals 1 if $\alpha = \gamma$; similarly $(\chi_\beta, \chi_\gamma) = 1$ if $\beta = \gamma$ and is zero otherwise. Therefore,

$$
\begin{aligned}
1 - 2 \cdot \mathbf{Fail}(f) &= \sum_{\alpha} \widehat{f}_\alpha^3 \\
&\leq \sum_{\alpha} |\widehat{f}_\alpha| \widehat{f}_\alpha^2 \\
&\leq (\max_{\alpha} |\widehat{f}_\alpha|) \sum_{\alpha} \widehat{f}_\alpha^2 \\
&= \max_{\alpha} |\widehat{f}_\alpha| \qquad\qquad \text{by Parseval's identity} \\
&= 1 - 2 \cdot \mathbf{Lindist}(f).
\end{aligned}
$$

It follows that $\mathbf{Fail}(f) \geq \mathbf{Lindist}(f)$.