6.841: Advanced Complexity Theory Fall 2012

Lecture 20 — November 20, 2012

Prof. Dana Moshkovitz *Scribe: Kate Rudolph*

1 Overview

In this lecture, we move towards completing our proof of the PCP theorem by analyzing the soundness case of the Long Code Test. To do this, we introduce and use Fourier analysis.

Over the past few lectures, we've been going through the proof of the PCP theorem. We used the powering operation to improve the gap in the PCP theorem. However, powering increases the alphabet by a lot– from Σ to Σ^{d^t}. This would be okay if we only needed to do it a few times, but we need to do this operation over and over, so this is too much for the alphabet to grow. We used the composition step to decrease the alphabet size at each step.

2 The Long Code Test

After powering, the assignment to u is some $\sigma(u) \in \Sigma^{d^t}$. The alphabet is too big, so we solve this by replacing u with many vertices, and use those vertices to represent an assignment to u. We do the same for all vertices v, and all edges e connecting vertices u and v. We need to construct edges between these new vertices so we can tell if the original constraint between u and v, ϕ_e, is satisfied. Essentially, we want to check if all of the constraints among all the vertices are satisfied, but we want to do this by only querying a constant number of the new vertices for their assignment. But this is exactly in the spirit of probabilistically checkable proofs – PCPs give a way to write a proof that you only need to query a constant number of locations to verify correctness; here, we're not verifying a proof, but probabilistically checking whether an assignment satisfies a given constraint. Essentially we're using a "mini-PCP" (called an *assignment tester* in the literature) to accomplish this. The "mini-PCP" we use here is known as the *Long Code Test* or the *Dictator Test*.

In the last lecture we proved completeness: if ϕ_e is satisfied by $\sigma(v)$ and $\sigma(u)$, then the long code test must always accept. In this lecture we will prove soundness. We can decompose the assignment table into three tables, one for the vertices that replace u, one for the vertices that replace v, and one for the vertices that replace e. To prove soundness of the overall PCP theorem, we need to use the Hastad test that presented in the last lecture, but today, we're going to talk about a simplified version of the long code test where we only look at the table of vertices replacing e. Let that assignment table be A. The simplified test is as follows:

Definition 1 (The Long Code Test / Dictator Test)**.** *Given $A : f \to \{\pm 1\}$, such that $\mathbb{E}_f[A(f)] = 0$, the long code test consists of the following:*

- *Pick f,g uniformly at random*
- *Pick μ. For every x, if $f(x) = 1$, set $\mu(x) = 1$. If $f(x) = -1$, set $\mu(x) = 1$ with probability $1 - \tau$, and $\mu(x) = -1$ with probability τ.*

- *Accept unless $A(f) = A(g) = A(g\mu) = 1$.*

3 Analysis of Soundness

The main lemma of the lecture follows:

Lemma 2 (Main Lemma)**.** *If the long code test accepts with probability $\geq 1 - \delta$, then there exists some $\sigma \in \Sigma^{d^t} = \Sigma'$ such that $A(f) = f(\sigma)$ for a $1 - \delta'$ fraction of all functions $f : \Sigma' \to \{\pm 1\}$, where δ' is a function depending on τ and δ.*

That is, if the long code test accepts with high probability, then the table is actually very close to a long code. The proof of this lemma can be extended to an analysis of the actual 3-table test, even though this lemma only applies to one table.

It seems very hard to analyze soundness. How will we determine the relationship between A on g and A on a slightly perturbed input, $g\mu$? The answer is to use Fourier Analysis.

4 Fourier Analysis

If there's one fundamental idea you should get out of this lecture, it is this: use Fourier analysis when you want to analyze the correlation between functions and shifted functions. This is exactly what we want to do, to analyze the relationship between $A(g)$ and $A(g\mu)$.

Now, we introduce the basic ideas of Fourier Analysis.

Definition 3 (Convolution)**.** *The* Convolution *of two functions A and B is denoted $(A * B)$, and*

$$(A * B)(f) = \mathbb{E}_g[A(g)B(fg^{-1})]$$

Suppose we have a function A. We can always represent A by its truth table, a vector describing, for each input, what the output of A is. Functions with this representation induce a vector space. So, we could chose a different basis to represent the elements of the vector space, i.e. the function A.

One such alternative basis is the Fourier basis. Let N denote the size of A's input. For all $\alpha \subseteq \{\pm 1\}^N$, define the function

$$\chi_\alpha(A) = \prod_{x \in \alpha} A(x).$$

Then we can represent the function A by

$$A = \sum_\alpha \widehat{A}(\alpha)\chi_\alpha.$$

The $\widehat{A}(\alpha)$ are called the Fourier coefficients, and the Fourier Transform is the transformation from A to $\widehat{A}(\alpha)$. It was proved in the homework that $\widehat{A * B} = \widehat{A}\widehat{B}$.

Thus, analyzing a shift in a function, while hard in the original basis, is just reduced to pointwise multiplication after the Fourier transform, so it is much easier. To prove the main lemma, we will represent everything in the Fourier basis and calculate what happens.

5 Proof of Main Lemma

The proof of the main lemma requires the following two lemmas

Lemma 4. *If the long code test accepts with probability $\geq 1-\delta$, then*

$$\sum_{|\alpha|>1} |\widehat{A}(\alpha)|^2 \leq O(\delta).$$

Lemma 5 (FKN)**.** *If $\sum_{|\alpha|>1} |\widehat{A}(\alpha)|^2 \leq \rho$ and $\widehat{A}(\emptyset) = 0$, then $\exists \sigma$ such that $|\widehat{A}(\{\sigma\})| \geq 1 - O(\rho)$.*

No proof of Lemma 5 is given here, but one can be found at [2]. Lemma 4 deals with sets α of size greater than 1, and Lemma 5 deals with α of size ≤ 1. Combining them, we will have our result for the sum over all α of $\widehat{A}(\alpha)$, which is what we want. We will show that for all sets α of size larger than 1, their Fourier coefficient is so small as to be negligible, and thus Lemma 5 proves the theorem.

To understand the lemmas, first we must determine the values of $\widehat{A}(\emptyset) = 0$ and $\widehat{A}(\{\sigma\})$ for a singleton set $\{\sigma\}$. We have

$$\begin{aligned} \widehat{A}(\emptyset) &= \langle A, \chi_\emptyset \rangle \\ &= \mathbb{E}_f[A(f)] \\ &= 0, \end{aligned}$$

and

$$\begin{aligned} \widehat{A}(\{\sigma\}) &= \langle A, \chi_{\{\sigma\}} \rangle \\ &= \langle A, f(\sigma) \rangle \\ &= \mathbb{E}_f[A(f)f(\sigma)] \\ &= (1) \cdot \Pr_f[A(f) = f(\sigma)] + (-1) \cdot \Pr_f[A(f) \neq f(\sigma)] \\ &= 2\Pr_f[A(f) = f(\sigma)] - 1. \end{aligned}$$

This says that A is close to a long code! So if we can apply the HKN lemma we will be done. We only need to show that the contributions from subsets besides $\{\sigma\}$ are so small as to be negligible, and we will do this by proving lemma 4.

Proof of Lemma 4. The probability that the test accepts is

$$\Pr(\text{test accepts}) = 1 - \mathbb{E}_{f,g,\mu}\left[\left(\frac{1+A(f)}{2}\right)\left(\frac{1+A(g)}{2}\right)\left(\frac{1+A(g\mu)}{2}\right)\right]$$

because if the test passes, at least one of $A(f)$, $A(g)$ and $A(g\mu)$ will be -1 so the corresponding term will be zero, and the expectation for that case will be zero. Otherwise the expectation will be 1. We've simply transformed the range of $A(f)$ from $\{-1,1\}$ to $\{1,0\}$ and multiplied.

Expanding this expression, we have that the probability is

$$\frac{7}{8} - \frac{1}{8}\left(\mathbb{E}[A(f)A(g)] + \mathbb{E}[A(f)] + \mathbb{E}[A(g)] + \mathbb{E}[A(g\mu)] + \mathbb{E}[A(f)A(g\mu)] + \mathbb{E}[A(g)A(g\mu)] + \mathbb{E}[A(f)A(g)A(g\mu)]\right).$$

Of the seven expectation terms, the first five terms are zero because the functions are not correlated with each other and by assumption, the expected value of $A(f)$ is zero. Thus, the probability is

$$\frac{7}{8} - \frac{1}{8}\mathbb{E}[A(g)A(g\mu)] - \frac{1}{8}\mathbb{E}[A(f)A(g)A(g\mu)].$$

We will analyze the two expectations in the expression above.

5.1 Analysis of $\mathbb{E}[A(g)A(g\mu)]$

First we analyze the expression $\mathbb{E}_{g,\mu}[A(g)A(g\mu)]$. We move to the Fourier basis, so the expression is

$$\begin{aligned}
\mathbb{E}_{g,\mu}[A(g)A(g\mu)] &= \mathbb{E}_{g,\mu}\left[\left(\sum_{\alpha}\widehat{A}(\alpha)\chi_\alpha(g)\right)\left(\sum_{\beta}\widehat{A}(\beta)\chi_\beta(g\mu)\right)\right] \\
&= \sum_{\alpha,\beta}\widehat{A}(\alpha)\widehat{A}(\beta)\mathbb{E}_{g,\mu}[\chi_\alpha(g)\chi_\beta(g\mu)]
\end{aligned}$$

We know that $\chi_\beta(g\mu) = \chi_\beta(g)\chi_\beta(\mu)$, and that $\chi_\alpha(g)\chi_\beta(g) = \chi_{\alpha\oplus\beta}(g)$. Thus we have

$$\mathbb{E}[A(g)A(g\mu)] = \sum_{\alpha,\beta}\widehat{A}(\alpha)\widehat{A}(\beta)\mathbb{E}_g[\chi_{\alpha\oplus\beta}(g)]\mathbb{E}_\mu[\chi_\beta(\mu)]$$

Now, if $\alpha \neq \emptyset$, then $\mathbb{E}_f[\chi_\alpha(f)] = 0$. Thus, the only terms that are nonzero occur when $\alpha \oplus \beta = \emptyset$, or when $\alpha = \beta$. So this sum is

$$\sum_{\alpha}\widehat{A}(\alpha)^2\mathbb{E}_\mu[\chi_\alpha(\mu)].$$

Now, we need to calculate $\mathbb{E}_\mu[\chi_\alpha(\mu)]$:

$$\begin{aligned}
\mathbb{E}_\mu[\chi_\alpha(\mu)] &= \mathbb{E}_\mu\left[\prod_{x\in\alpha}\mu(x)\right] \\
&= \prod_{x\in\alpha}\mathbb{E}_\mu[\mu(x)] \\
&= \prod_{x\in\alpha}(-\tau) \\
&= (-\tau)^{|\alpha|}
\end{aligned}$$

So, we find that

$$\mathbb{E}[A(g)A(g\mu)] = \sum_{\alpha}\widehat{A}(\alpha)^2(-\tau)^{|\alpha|}.$$

5.2 Analysis of $\mathbb{E}[A(f)A(g)A(g\mu)]$

We are interested now in the quantity $\mathbb{E}[A(f)A(g)A(g\mu)]$. In the Fourier basis, this is

$$\sum_{\alpha,\beta,\gamma} \widehat{A}(\alpha)\widehat{A}(\beta)\widehat{A}(\gamma)\mathbb{E}_{f,g,\mu}[\chi_\alpha(f)\chi_\beta(g)\chi_\gamma(g\mu)].$$

As before, we write $\chi_\gamma(g\mu) = \chi_\gamma(g)\chi_\gamma(\mu)$. We can combine the g's to rewrite the expectation term as

$$\mathbb{E}_g[\chi_\beta(g)\chi_\gamma(g)] = \mathbb{E}_g[\chi_{\beta\oplus\gamma}(g)]$$

which is 0 unless $\beta = \gamma$, in which case it is 1. So we may eliminate this term and substitute $\beta = \gamma$ in the original expression, which becomes

$$\sum_{\alpha,\gamma} \widehat{A}(\alpha)\widehat{A}(\gamma)^2\mathbb{E}_{f,\mu}[\chi_\alpha(f)\chi_\gamma(\mu)].$$

Now, we want to understand the expectation term $\mathbb{E}_{f,\mu}[\chi_\alpha(f)\chi_\gamma(\mu)]$.

$$\begin{aligned}\mathbb{E}_{f,\mu}[\chi_\alpha(f)\chi_\gamma(\mu)] &= \mathbb{E}_{f,\mu}\left[\left(\prod_{x\in\alpha} f(x)\right)\left(\prod_{y\in\gamma}\mu(y)\right)\right]\\ &= \mathbb{E}_{f,\mu}\left[\left(\prod_{x\in\alpha\cap\gamma} f(x)\mu(x)\right)\left(\prod_{x\in\alpha-\gamma} f(x)\right)\left(\prod_{x\in\gamma-\alpha}\mu(x)\right)\right]\end{aligned}$$

There are three products in this term. The first product is $(-1+\tau)^{|\alpha|}$. For the second product, if $\alpha \not\subseteq \gamma$, the second product is zero. For the third product, as before we calculate that the expectation is $(-\tau)^{|\gamma-\alpha|}$.

Thus we have

$$\mathbb{E}[A(f)A(g)A(g\mu)] = \sum_{\alpha\subseteq\gamma} \widehat{A}(\alpha)\widehat{A}(\gamma)^2(-1+\tau)^{|\alpha|}(-\tau)^{|\gamma-\alpha|}.$$

Now we must simply bound $\sum_{|\alpha|>1}\widehat{A}(\alpha)^2$. We have run out of time to do this in lecture, but in about 5 more lines of computation, we can invoke Cauchy-Schwartz, to find that

$$\sum_{|\alpha|>1}\widehat{A}(\alpha)^2 \le \frac{8\delta}{(1-\tau)(1-\sqrt{1-\tau})}.$$

For more details, refer to [1]. The relevant proof is in the appendix.

The key point of the lecture was, once we moved to the Fourier basis, everything was just arithmetic. This is the way to analyze the long code to get exactly the tight results.

□

References

[1] Irit Dinur. 2007. *The PCP theorem by gap amplification.* J. ACM 54, 3, Article 12 (June 2007).

[2] Friedgut, Kalai, and Noor. *Boolean Functions whose Fourier Transform is Concentrated on the First Two Levels.* Advances in Applied Mathematics 29, no. 3, (2002) 427-437.