Lecture 23

In a previous lecture we showed that $\mathbf{NP} = \mathbf{PCP}[O(\log n), poly(\log n)]$. Today we will begin a proof that $\mathbf{NP}$ is contained in a variant of $\mathbf{PCP}[O(\log n), poly(\log n)]$ where we make only a *constant* number of queries into the proof, although each query will read $poly(\log n)$ bits of the proof. At first this may not appear to be a useful improvement, but this restricted variant is the foundation of tighter results on the relationship between $\mathbf{NP}$ and $\mathbf{PCP}$. For today we will content ourselves with stating the main results and laying the foundation for the proof. The analysis and justification will follow in the next lecture.

1 Main Lemma

Lemma 1
Given a 3CNF formula ϕ, with m clauses on n variables, fix a field $\mathbb{F}$ of size $|\mathbb{F}| = O(\log^3 n)$. Define $m \equiv \log n / \log\log n$, $m' \equiv 3m+3$, $h \equiv \log n$, and $d \equiv hm' + m' + 2$. Then, there is a sequence of $poly(n)$ tests of the form

$$(\underline{x_1}, \ldots, \underline{x_c}, T)$$

where $\underline{x_i} \in \mathbb{F}^{m'+1}$, $T : \mathbb{F}^c \to \{0,1\}$, and $c = O(poly(\log n))$, such that:

- *If ϕ is satisfiable, then there exists a polynomial $p : \mathbb{F}^{m'+1} \to \mathbb{F}$ of degree d such that $T(p(\underline{x_1}), \ldots, p(\underline{x_c})) = 1$ for every test.*
- *If ϕ is not satisfiable, then for all polynomials $p : \mathbb{F}^{m'+1} \to \mathbb{F}$ of degree d, for at least $1/2$ of the tests, $T(p(\underline{x_1}), \ldots, p(\underline{x_c})) = 0$.*

2 Another Low-Degree Test

Our goal is to test whether a function for which we have oracle access, $f : \mathbb{F}^{m'+1} \to \mathbb{F}$, is a low-degree polynomial. In addition to the oracle for f, we assume we are given a parameter d (the degree for which we are testing) as well as oracle access to a function $L : (\mathbb{F}^{m'+1} \times \mathbb{F}^{m'+1}) \to \mathbb{F}^{d+1}$. The claim to be tested is that:

1. f is a degree d polynomial
2. For all $\underline{x}, \underline{y} \in \mathbb{F}^{m'+1}$, $L(\underline{x}, \underline{y})$ is the description of the degree d univariate polynomial q such that $q(t) = f(\underline{x} + t\underline{y})$ for every $t \in \mathbb{F}$.

Our test is as follows:

pick at random $\underline{x}, \underline{y} \in \mathbb{F}^{m'+1}$ and $t \in \mathbb{F}$
$q \leftarrow L(\underline{x}, \underline{y})$
reject if $f(\underline{x} + t\underline{y}) \neq q(t)$
accept

We claim that there exists a constant δ_0 such that if the preceding test accepts (f, L) with probability $1-\delta$ where $\delta < \delta_0$, then f has agreement at least $1-2\delta$ with some degree d polynomial.

3 Curves in $\mathbb{F}^m$

Definition 1 *A curve in $\mathbb{F}^m$ is a function $C : \mathbb{F} \to \mathbb{F}^m$. We say C is a curve of degree d if there are m polynomials of degree d, $\{c_1, c_2, \ldots, c_m\}$, such that $C(t) = (c_1(t), c_2(t), \ldots, c_m(t))$.*

For example, $C(t) = \underline{x} + t\underline{y}$ is a curve of degree 1. We note the following easily verified facts about curves.

1. For any $d+1$ points $\underline{x_0}, \ldots, \underline{x_d} \in \mathbb{F}^m$, and for any $t_0, \ldots, t_d \in \mathbb{F}$, there is a curve C of degree d such that $C(t_i) = \underline{x_i}$ for all i. This follows from component-wise polynomial interpolation.

2. For any d points $\underline{x_1}, \ldots, \underline{x_d} \in \mathbb{F}^m$, fix $t_0, \ldots, t_d \in \mathbb{F}$. Pick at random $\underline{x_0} \in \mathbb{F}^m$ and construct C of degree d such that $C(t_i) = \underline{x_i}$ for $i = 0, 1, \ldots, d$. Then $C(t)$ is uniformly distributed in $\mathbb{F}^m$ for every $t \notin \{t_1, \ldots, t_d\}$.

   Why is this true? That coordinate i of $C(t)$ is uniformly distributed in $\mathbb{F}$ follows from the facts that there is a bijection between coordinate i of $C(t_0) = \underline{x_0}$ and coordinate i of $C(t)$ for any $t \notin \{t_1, \ldots, t_d\}$. (The latter point in turn follows from the fact that a polynomial of degree d is uniquely specified by $d+1$ points.) Finally, since the coordinates of $\underline{x_0}$ are chosen independently, it follows that $C(t)$ is uniformly distributed over $\mathbb{F}^m$.

3. If $p : \mathbb{F}^m \to \mathbb{F}$ is a degree d polynomial and $C : \mathbb{F} \to \mathbb{F}^m$ is a degree c curve, then $p(C(t))$ is a degree cd univariate polynomial. This follows from basic facts about polynomial composition.

4 Components of the Proof

The prover must provide the following 3 items.

1. A description of the function $f : \mathbb{F}^{m'+1} \to \mathbb{F}$. The claim of the prover is that $f = p$, the degree d polynomial whose existence is asserted by lemma 1 for satisfiable instances ϕ.

2. A description of the function $L : (\mathbb{F}^{m'+1} \times \mathbb{F}^{m'+1}) \to \mathbb{F}^{d+1}$. The claim of the prover is that for all $\underline{x}, \underline{y} \in \mathbb{F}^{m'+1}$, $L(x, y)$ describes the degree d univariate polynomial q satisfying $q(t) = f(\underline{x} + t\underline{y})$.

3. A description of the function $\mathcal{C} : (\mathbb{F}^{m'+1})^{c+1} \to \mathbb{F}^{cd+1}$. For all $\underline{x_0}, \ldots, \underline{x_c} \in \mathbb{F}^{m'+1}$, let C be the unique degree c curve $C : \mathbb{F} \to \mathbb{F}^{m'+1}$ such that $C(t_i) = \underline{x_i}$ for all $i \in \{0, \ldots, c\}$. Then the claim of the prover is that $\mathcal{C}(\underline{x_0}, \ldots, \underline{x_c})$ is the degree cd polynomial equal to $f(C(t))$.

5 Verification

The verifier performs the following test whose correctness will be demonstrated in the next lecture.

input: a proof as described above
output: if proof is valid, **accept** with probability 1; otherwise **accept** with probability $\leq \frac{1}{2}$.

pick at random $\underline{x}, \underline{y} \in \mathbb{F}^{m'+1}$ and $t \in \mathbb{F}$
$q \leftarrow L(\underline{x}, \underline{y})$
reject if $f(\underline{x} + t\underline{y}) \neq q(t)$
pick at random a test $(\underline{x_1}, \ldots, \underline{x_c}, T)$
pick at random $\underline{x_0} \in \mathbb{F}^{m'+1}$
$q \leftarrow \mathcal{C}(\underline{x_0}, \ldots, \underline{x_c})$
compute C such that $C(t_i) = \underline{x_i}$ for all i
pick random $t \in \mathbb{F}$
reject if $f(C(t)) \neq q(t)$
reject if $T(q(t_1), \ldots, q(t_c)) \neq 1$
accept

Finally we note that to choose a random element of $\mathbb{F}$ requires $O(\log\log n)$ random bits and hence choosing an element of $\mathbb{F}^{m'+1}$ requires $O(m' \log\log n) = O(\log n)$ random bits. Likewise there are only polynomially many tests to choose from so we need only $O(\log n)$ random bits to make such a choice. Thus the total number of random bits required is $O(\log n)$. Furthermore it is evident that the verification algorithm makes only a constant number of queries to each of $f, L, \mathcal{C}$ and hence only a constant number of queries total are required, although we again note that the total number of bit queries is still $poly(\log n)$.