that this occurs *every* time is at most 2^{-100}. So, if the algorithm does not prove that $f \neq 0$, we can be pretty certain that actually $f = 0$. Not 100% certain, but if we lose the bet, we would know that an experiment that had only two possible outcomes ended with the one that had probability 2^{-100}. This should compensate for our trouble: we found a needle in a haystack!

As our next example, consider the following situation. We have two friends, Alice and Bob. Alice maintains a large database of information. Bob maintains a second copy of the database. Periodically, they must compare their databases for consistency. Because the transmission between Alice and Bob is expensive, they would like to discover the presence of inconsistency without transmitting the entire database between them. Denote Alice's data by the sequence $a = a_0 \cdots a_{n-1}$ and Bob's data by the sequence $b = b_0 \cdots b_{n-1}$ where $a_i, b_i \in \{0, 1\}$. It is clear that any deterministic consistency check that transmits fewer than n bits will fail (just because an adversary can modify the unsent bits). Using randomness it is possible to design a strategy that detects an inconsistency with high probability (at least $1 - n^{-1}$) while transmitting many fewer than n bits, namely only $O(\log n)$ bits.

Think of the strings a and b as (strings of coefficients of) univariate polynomials over the field $\mathbb{F}_p$ where p is a prime such that $n^2 < p < 2n^2$ (theorems regarding the density of primes guarantee the existence of such p). That is, consider polynomials

$$\begin{aligned} A(x) &= a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} \pmod{p}, \\ B(x) &= b_0 + b_1 x + \ldots + b_{n-1} x^{n-1} \pmod{p}. \end{aligned}$$

In order to detect whether $a = b$, Alice and Bob use the following strategy:

> Alice picks uniformly at random a number $\boldsymbol{r}$ in $\mathbb{F}$ and sends to Bob the numbers $\boldsymbol{r}$ and $A(\boldsymbol{r})$. Bob responds with 1 if $A(\boldsymbol{r}) = B(\boldsymbol{r})$ and with 0 otherwise. The number of bits transmitted is $1 + 2 \log p = O(\log n)$.

If $a = b$ then $A(\boldsymbol{r}) = B(\boldsymbol{r})$ for all $\boldsymbol{r}$, so the output is always 1. If $a \neq b$ we have two distinct polynomials $A(x)$ and $B(x)$ of degree at most $n - 1$. By Lemma 16.4, the probability of error is

$$\Pr\left[A(\boldsymbol{r}) = B(\boldsymbol{r})\right] \leq \frac{n-1}{|\mathbb{F}|} = \frac{n-1}{p} \leq \frac{1}{n}\,.$$

16.2 Solution of Kakeya's problem in finite fields

A famous unsolved problem in mathematics is the Kakeya conjecture in geometric measure theory. This conjecture is descended from the following question asked in 1917 by Japanese mathematician Soichi Kakeya: What is the smallest set in the plane in which one can rotate a needle around completely? He likened this to a samurai turning his lance around in a small toilet. For

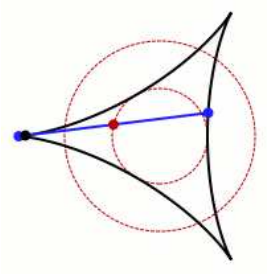

Fig. 16.1 As the middle (flexible) point moves around the smaller circle, the needle rotates through 360°.

instance, one can rotate a unit needle inside a unit disk, which has area $\pi/4$. By using a deltoid one requires only $\pi/8$ area (see Fig. 16.1).

The Kakeya conjecture in more dimensions states that any subset of $\mathbb{R}^n$ that contains a unit line segment in every direction has Hausdorff dimension equal to n. This conjecture remains open in dimensions three and higher, and gets more difficult as the dimension increases.

To approach this question, Wolff (1999) proposed a simpler *finite* field analogue of the Kakeya conjecture. If $\mathbb{F}^n$ is a vector space over a finite field $\mathbb{F}$, define a Kakeya set to be a subset $K \subseteq \mathbb{F}^n$ which contains a line in every direction, namely for any $v \in \mathbb{F}^n$ there exists a vector $w \in \mathbb{F}^n$ such that the line $\{w + tv : t \in \mathbb{F}\}$ is contained in K; here, vector w is the origin and vector v the direction of the line. The finite field Kakeya conjecture stated that there exists a constant $c > 0$ depending only on the dimension n such that every Kakeya set $K \subseteq \mathbb{F}^n$ has cardinality $|K| \geq c|\mathbb{F}|^n$.

This finite field version of the conjecture has had a significant influence on the subject, in particular inspiring work on the sum-product phenomenon in finite fields, which has since proved to have many applications in number theory and computer science. Modulo minor technicalities, the progress on the finite field Kakeya conjecture was, however, essentially the same as that of the original "Euclidean" Kakeya conjecture.

Recently Dvir (2009) used a surprisingly simple application of the polynomial method to prove the finite field Kakeya conjecture.

Lemma 16.5. *Let $f \in \mathbb{F}[x_1, \ldots, x_n]$ be a polynomial of degree at most $q - 1$ over a finite field with $q = |\mathbb{F}|$ elements. If f vanishes on a Kakeya set K, then f is the zero polynomial.*

Proof. The argument is similar to that in the proof of Lemma 16.2. Suppose for a contradiction that f is nonzero. We can write $f = \sum_{i=0}^{d} f_i$, where $0 \leq d \leq q - 1$ is the degree of f and f_i is the i-th homogeneous component; thus f_d is nonzero. Since f vanishes on K, d cannot be zero. Hence, f_d is a nonzero polynomial.

Let $v \in \mathbb{F}^n \setminus \{\mathbf{0}\}$ be an arbitrary direction. As K is a Kakeya set, K contains a line $\{w + tv : t \in \mathbb{F}\}$ for some $w \in \mathbb{F}^n$, thus $f(w + tv) = 0$ for all $t \in \mathbb{F}$. The left-hand side is a polynomial $g_{w,v}(t)$ in t of degree at most $q-1$, and must be the zero polynomial by the factor theorem, that is, all its coefficients are zero. In particular, the coefficient of t^d, which is $f_d(v)$, must be zero. Since v was arbitrary, it follows that the polynomial $f_d(x)$ vanishes on all points in $\mathbb{F}^n$. But since $dq^{n-1} \le (q-1)q^{n-1} < q^n$, Lemma 16.2 implies that f_d must be a zero polynomial. □

Theorem 16.6 (Dvir 2009)**.** *Let $K \subset \mathbb{F}^n$ be a Kakeya set. Then*

$$|K| \ge \binom{|F| + n - 1}{n} \ge \frac{|\mathbb{F}|^n}{n!}\,.$$

Proof. Let $q = |\mathbb{F}|$ and suppose that $|K| < \binom{n+q-1}{n}$. Then, by Lemma 16.1, there exists a *nonzero* polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ of degree at most $q-1$ that vanishes on K, which contradicts Lemma 16.5. □

16.3 Combinatorial Nullstellensatz

The following special case of Hilbert's Nullstellensatz has found numerous applications in combinatorics.

Theorem 16.7 (Nullstellensatz)**.** *Let $f \in \mathbb{F}[x_1, \ldots, x_n]$, and let $S_1, \ldots, S_n$ be nonempty subsets of $\mathbb{F}$. If $f(x) = 0$ for all $x \in S_1 \times \cdots \times S_n$, then there are polynomials $h_1, \ldots, h_n \in \mathbb{F}[x_1, \ldots, x_n]$ such that* $\deg(h_i) \le \deg(f) - |S_i|$ *and*

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} h_i(x_1, \ldots, x_n) \prod_{s \in S_i} (x_i - s)\,.$$

Proof (due to Alon 1999). Define $d_i = |S_i| - 1$ for all i, and consider polynomials

$$g_i(x_i) = \prod_{s \in S_i} (x_i - s) = x_i^{d_i+1} - \sum_{j=0}^{d_i} a_{ij} x_i^j.$$

Observe that if $x_i \in S_i$ then $g_i(x_i) = 0$, that is,

$$x_i^{d_i+1} = \sum_{j=0}^{d_i} a_{ij} x_i^j\,. \tag{16.1}$$

Let $\overline{f}$ be the polynomial obtained by writing f as a linear combination of monomials and replacing, repeatedly, each occurrence of $x_i^{t_i}$ $(1 \le i \le n)$, where $t_i > d_i$, by a linear combination of smaller powers of x_i, using the relations (16.1). The resulting polynomial $\overline{f}$ is clearly of degree at most d_i in