Provable Collisions in the Pollard ρ Algorithm for Discrete Logarithms

Fields Institute Workshop on Cryptography: Underlying Mathematics, Provability and Foundations

Stephen Miller
Rutgers University

Ramarathnam Venkatesan
Microsoft Research

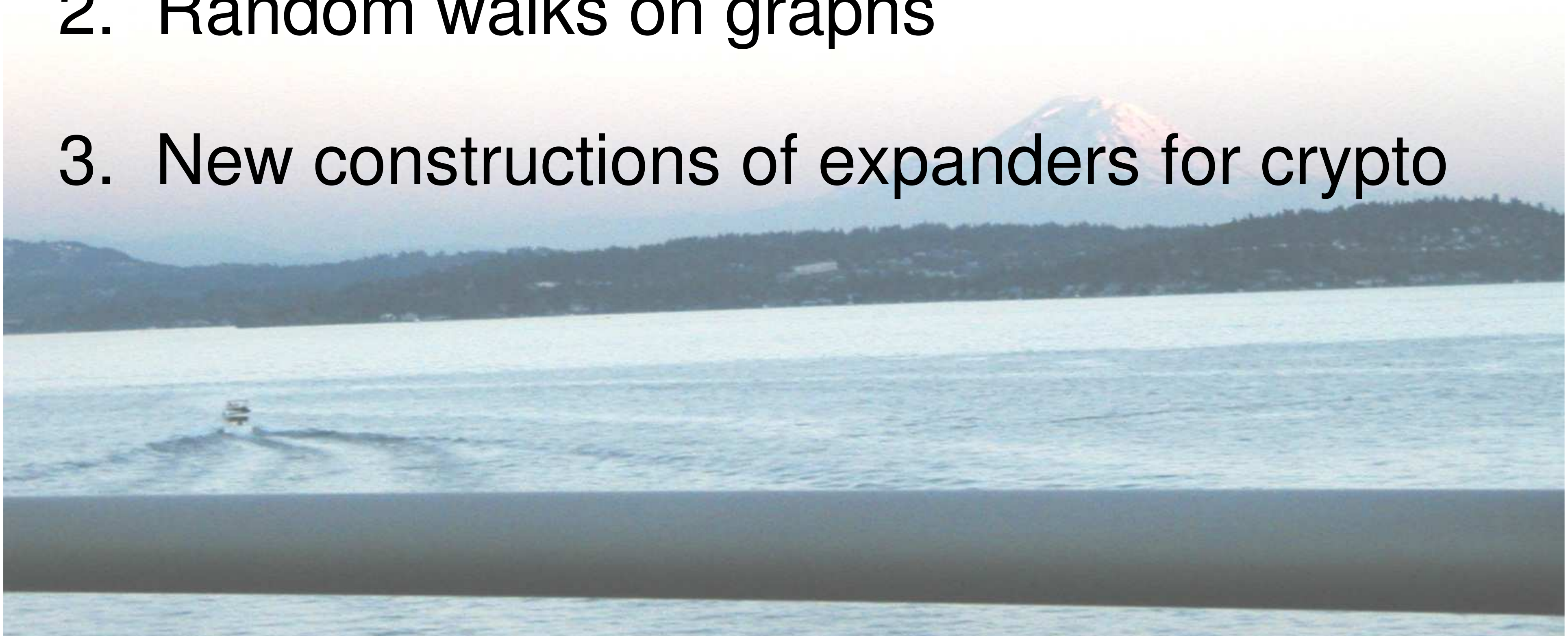
Reference: Procedings of the 7th Algorithmic Number Theory Symposium (Berlin), Springer-Verlag, 2006, pp. 573-581.

Agenda:

1. Results on the Pollard ρ algorithm

2. Random walks on graphs

3. New constructions of expanders for crypto



Agenda:

1. Results on the Pollard ρ algorithm

2. Random walks on graphs

3. New constructions of expanders for crypto

Discrete Logarithm Problem

- Given a cyclic group G with generator g, and power h of it, solve the equation

$$h = g^x \quad \text{for } x.$$

- DLOG is a hard problem which several cryptosystems are based on, for example
  - Diffie-Hellman
  - El-Gamal
  - Elliptic Curve Cryptography

- The difficulty of DLOG is determined by the realization of a group: e.g. trivial for $\mathbb{Z}/n\mathbb{Z}$ but harder for $(\mathbb{Z}/p\mathbb{Z})^*$, and apparently harder yet for elliptic curves.

- Difficulty is determined by the largest prime divisor of n = #G. From now on we assume n is a large prime.

- We will consider the DLOG problem on a "black-box" group, i.e. one which uses no specific features of its embedding. In this case, it is a theorem of Nechaev, Shoup that no DLOG algorithm can run in time $o(n^{1/2})$.

Pollard ρ algorithm's attributes

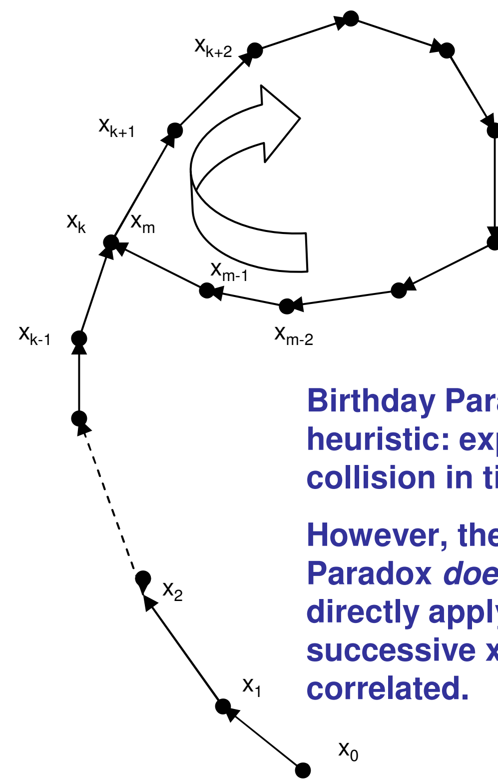
- Uses birthday paradox to run in time $\sim n^{1/2}$ (heuristically). Can this be proved?
- Up to constant factors, the fastest known on general groups, and the only with low storage requirements.
- In particular, it is the fastest known algorithm for general elliptic curve groups (i.e. aside from special curves which have subexponential algorithms).
- Therefore the $n^{1/2}$ running time is used to gauge the relative bit-by-bit strengths of ECC vs. RSA cryptosystems (e.g. 160-bit ECC $\approx$ 1024-bit RSA).

The actual algorithm

1. Let g equal the generator, $h = g^y$ the element whose discrete logarithm (y) we wish to recover.
2. Partition the group G into 3 random sets S_1, S_2, and S_3 (each element of G has, independently, a 1/3 probability of being in each S_j).
3. Set $x_0 = h$ (or more generally a random power $g^{r_1}h^{r_2}$).
4. Iterate $x_{k+1} = f(x_k)$, where

$$f(x) = \begin{cases} gx\,, & x \in S_1\,; \\ hx\,, & x \in S_2\,; \\ x^2\,, & x \in S_3\,. \end{cases}$$

5. Find a collision $x_k = x_m$ (or more properly $x_i = x_{2i}$ to save on storage).
6. Use collision information to find y (next slide).



Birthday Paradox heuristic: expect collision in time $O(n^{1/2})$.

However, the Birthday Paradox *does not* directly apply since the successive x_k are highly correlated.

What to do with a collision

(If you walk into your own back, you may learn something about yourself)

- At each step, x_k may be written as $g^{\alpha_k y+\beta_k}$

- The iteration f(x) sends the coefficients (α,β) to one of:
  - $(\alpha+1,\beta)$ [The move $x \mapsto hx$]
  - $(\alpha,\beta+1)$ [$x \mapsto gx$]
  - $(2\alpha,2\beta)$ [$x \mapsto x^2$]

- Given a collision $x_k = x_m$, we must have that $\alpha_k y + \beta_k = \alpha_m y + \beta_m$
  Since the exponents are taken mod n, we can solve this:

$$y = \frac{\beta_m - \beta_k}{\alpha_k - \alpha_m}$$

  provided that $\alpha_k \neq \alpha_m$ (mod n). [Non-degeneracy condition]. Expect this with high probability $\approx$ 1-1/n. This is even more likely than a collision (heuristically).

- Note that if $x_k = x_m$ is the first collision and if $\alpha_k = \alpha_m$ (degenerate) there, the α's will be equal at any subsequent collision in the loop (because they evolve the same way under the iterating function f).

- Likewise (since each step is invertible) if the α's are distinct at the first collision, they remain distinct at all future collisions.

An estimate on the collision time

- Theorem 1: Fix any $p < 1$. Then the Pollard ρ algorithm finds a collision in time $O_p(n^{1/2}(\log n)^3)$ with probability $\geq p$, where the probability is taken over all partitions of G into the three subsets S_1, S_2, and S_3.

  $\left(\tilde{O}_p(n^{1/2})\right)$

- This is the first nontrivial rigorous result on the runtime of the algorithm.
- $O(n^{1/2})$ is the expected optimal collision time.
- Montenegro observation improves this to $O_p(n^{1/2}(\log n)^{3/2})$.

A complete runtime estimate
(for almost all n)

- *Multiplicative order of 2 modulo n:* the least $k > 0$ for which $2^k = 1 \pmod{n}$.
- Theorem 2. Assume that the multiplicative order of 2 modulo n is $\Omega(\log(n)^3)$. Then the Pollard ρ collisions guaranteed by the previous theorem are nondegenerate with probability $1 - O(\log(n)^6 / n)$.
- Almost all primes have this property:
  e.g. if 2 is a generator of $(\mathbb{Z}/n\mathbb{Z})^*$, then the multiplicative order is n-1.
- More precisely, at most $O(\log(X)^5)$ such primes n exist in the interval between X and 2X.
- So the theorem, in practice, proves the Pollard ρ runtime for random group orders.
- In general, one can quickly test if n has this property. If it doesn't, the theorem works if the squaring step x^2 is replaced by another small power x^a for which **a** has large multiplicative order.

Agenda:

1. Results on the Pollard ρ algorithm
2. Random walks on graphs
3. New constructions of expanders for crypto

Random Walks

- The Pollard ρ iteration $x_{k+1} = f(x_k)$ is definitely *not* a random walk, since it goes into a loop after its first collision. This loop is a key feature of the algorithm.
- However, since membership of x_k in S_1, S_2, or S_3 is random, the walk behaves *exactly* as a random walk until the first collision occurs – thereafter it is decidedly non-random.
- Upshot: to prove collisions occur it suffices to consider random walks on G whose steps have the form

  $\mathbf{x \mapsto gx}$, $\mathbf{x \mapsto hx}$, **and** $\mathbf{x \mapsto x^2}$ (each with probability 1/3).

- In analyzing a random walk on a cyclic group, we can use additive notation (i.e. simply consider $\mathbb{Z}/n\mathbb{Z}$).
- We consider the equivalent random walk on $\mathbb{Z}/n\mathbb{Z}$ given by moves of the form

  $\mathbf{x \mapsto x+1}$, $\mathbf{x \mapsto x+y}$, **or** $\mathbf{x \mapsto 2x}$ (each with probability 1/3).

- Our result: this random walk has mixing time $\mathbf{(\log n)^3}$ (more precise on next slide).
  - Mixing time is a measure of how many steps it takes for a random walk to become equidistributed, and thereby "forget" where it started.
  - Comparison: if one could make totally random walks (like in the Birthday paradox), the mixing time would be 1.
  - Without the squaring step the mixing time would be $\mathbf{n^{power}}$ ([Teske]).

Graph reformulation

- Define the *Pollard ρ graph* Γ to have vertices $\mathbb{Z}/n\mathbb{Z}$ and directed edges connecting **$x \to x+1$, $x \to x+y$, and $x \to 2x$** for each **$x \in \mathbb{Z}/n\mathbb{Z}$**.
- The previous random walk is now a random walk on this graph, where we move from vertex to vertex by picking one of the three edges that starts there with equal probability = 1/3.
- We prove this walk has mixing time **$(\log n)^3$**, and collisions in time **$O(n^{1/2}(\log n)^{3/2})$**.

Brief Review of Graph Theory

- Definitions: A graph Γ is a collection of vertices V, and (directed) edges E connecting the vertices.
- A k-regular graph has exactly k edges meeting at each vertex (k in, k out).
- Adjacency operator A on $L^2(V)$ averages the function over its neighbors

$$A: f(x) \mapsto \sum_{x \to y} f(y)$$

- The constant functions on V are eigenfunctions with the *trivial eigenvalue* $\lambda_{triv} = k$.

Operator Norm Mixing Lemma

- Often people look at the spectral gap (for undirected graphs).
- Our graph is directed, so instead we will look at restriction of A to the orthogonal complement of $\mathbb{1}$, and its operator norm [= the most it distorts lengths].
- **Lemma**: Suppose that the operator norm of A's restriction to {the orthogonal complement of the constant function} is bounded by $\mu < k$. Let x be an arbitrary vertex and S be an arbitrary subset of vertices of Γ. Then the probability of a random walk of length $r \geq \log(2n)/\log(k/\mu)$ starting from x landing in S is between $\frac{1}{2}|S|/n$ and $3/2\ |S|/n$ (i.e. $(1\pm\frac{1}{2})\cdot|S|/n$).
- This assumption is met for the Pollard ρ graphs, with k = 3 and $\mu = 3 - c/(\log n)^2$ for some c>0 (see next slide). So the mixing happens for $r \gg (\log n)^3$.
- Method of proof: study action of A^r on the characteristic function of {x}, which tells you where walks end. Take inner product with χ_S, apply Cauchy-Schwartz inequality, and finally the operator norm bound.

Operator Norm Bound for A

- We need to show that $\|A_{\text{restricted to } \mathbb{1}^{\perp}}\| \leq 3 - \frac{c}{(\log n)^2}$
  i.e. that
  $$\|Af\| \leq \left(3 - \frac{c}{(\log n)^2}\right) \|f\|$$
  for all $f \in L^2(V)$ which are orthogonal to $\mathbb{1}$.

- Recall $(Af)(x) = f(x+1) + f(x+y) + f(2x)$.

- Look at basis of the orthogonal complement consisting of nontrivial characters of $(\mathbb{Z}/n\mathbb{Z})$ given by
  $\chi_k(x) = e^{2 \pi i k x/n}$, $k \neq 0$.
  $$\langle \chi_k, \chi_\ell \rangle = \begin{cases} n, & k = \ell \\ 0, & \text{otherwise}. \end{cases}$$

- We write $f = \sum_{k\neq 0} c_k \chi_k$, so that $\|f\|^2 = n \sum |c_k|^2$.

- We have $A\chi_k = d_k\chi_k + \chi_{2k}$, where $d_k = \chi_k(1) + \chi_k(y)$,
  $|d_k| = 2|\cos(\pi k(y-1)/n)|$.

Operator Norm Bound for A (ctd.)

- We compute:

$$\|Af\|^2 \;=\; \langle Af, Af\rangle \;=\; \langle \sum c_k A\chi_k, \sum c_k A\chi_k \rangle \;=$$

$$\sum_{k,\ell\neq 0} c_k\,\overline{c_\ell}\,[\langle d_k\chi_k, d_\ell\chi_\ell\rangle + \langle \chi_{2k}, \chi_{2\ell}\rangle + \langle d_k\chi_k, \chi_{2\ell}\rangle + \langle \chi_{2k}, d_\ell\chi_\ell\rangle]$$

$(|d_k| \leq 2)$

$$\leq \; n\left(5\sum |c_k|^2 + 2\sum |c_k||c_{2k}||d_{2k}|\right).$$

- This needs to be $\leq$ n (9 – c/(log n)^2) $\sum$ |c
- The savings is gained from the second sum and the following quadratic form bound:
  - if $Q(x_1,\ldots,x_{n-1}) \;:=\; \sum_{k=1}^{n-1} |x_k|\,|x_{2k}|\,\lambda_k$, where $\lambda_k = |\cos(\pi k/n)|$
  - then $|Q(x_1,\ldots,x_{n-1})| \;\leq\; \left(1 - \frac{c}{(\log n)^2}\right) \sum_{k=1}^{n-1} x_k^2\,.$
- This can be viewed as a “reciprocal” Hilbert inequality (continued…).

Quadratic Form Bound

- Let $\mathbb{S}$ be the set of k between –n/4 and n/4 (mod n). Then
  $\lambda_k = |\cos(\pi k/n)| \leq 1$ for $k \in \mathbb{S}$, and $\lambda_k \leq \text{sqrt}(1/2)$ for $k \notin \mathbb{S}$.
- So we need to show $$\sum |x_k|\,|x_{2k}|\,\varepsilon_k \;\leq\; \left(1 - \frac{c}{(\log n)^2}\right) \sum x_k^2$$
  where ε_k = indicator function of $k \in \mathbb{S}$.
- Let $\gamma_k > 0$. Then $$0 \;\leq\; (\gamma_k x_k \pm \gamma_k^{-1} x_{2k})^2 \;=\; \gamma_k^2\, x_k^2 \;+\; \gamma_k^{-2}\, x_{2k}^2 \;\pm\; 2 x_k\, x_{2k}$$
  and so the quadratic form is bounded by the *diagonal* quadratic form
  $$\tfrac{1}{2} \sum x_k^2 \left(\varepsilon_k \gamma_k^2 + \varepsilon_{\bar{2}k} \gamma_{\bar{2}k}^{-2}\right)$$
- At this point, one needs simply to chose the γ_k such that the expression in parentheses is $\leq 2 - \Omega((\log n)^{-2})$.
- In a moment we will choose γ_k between 1 and 1.5. Observe that with such small γ_k our desired inequality automatically holds unless both k and $2^{-1}k$ lie in the residues in $\mathbb{S}$ (mod n).
- So we take γ_k to be 1 for $k \notin \mathbb{S}$, and otherwise equal to $1 - sd/(\log n)^2$, in which 2^s is the exact power of 2 dividing k (viewed as an integer in [-n/4,n/4]). Here d is a constant.
- It is easy to check that $$\gamma_k^2 \;+\; \gamma_{\bar{2}k}^{-2} \;\;\approx\;\; 1 - 2\frac{sd}{(\log n)^2} + 1 + 2\frac{(s-1)d}{(\log n)^2} \;\;\leq\;\; 2 - \Omega((\log n)^{-2})$$
  because if you double the integer representing $2^{-1}k$ between –n/4 and n/4, you get exactly the integer representing k in that range. So $s(k) = s(2^{-1}k)+1$.

Putting Together: Graph Mixing Theorem

- **Theorem:** Let **x** be any vertex and **S** be any subset of vertices of Γ. Then there exists an explicit constant **c** such that the probability that

  {a random walk of length $\geq$ **c(log n)^3** starting at **x** ends in **S**}

  is between **½|S|/n and 3/2 |S|/n**.

- This implies the collision time estimate of **O(n$^{1/2}$(log n)^3)** as follows:
  - Let S = the set of the first t = $\lfloor n^{1/2} \rfloor$ iterates $\mathbf{x_1, \cdots, x_t}$ of the random walk.
  - We may assume that |S| = t, for otherwise a collision has already occurred.
  - Let r = c(log n)^3 above. Then the probability of $\mathbf{x_{t+r}, x_{t+2r}, x_{t+3r}, \cdots, x_{t+kr}}$ lying in S are each independently at least 1/(3t).
  - Choose k = 3bt, b fixed. The probability that none of those points lies in S is bounded by $(1\text{-}1/(3t))^{3bt} \approx e^{-b}$, which can be arbitrarily small if b is large.
  - Thus, with high probability $\geq 1\text{-}e^{-b}$, there is a collision in time **O(n$^{1/2}$(log n)^3)**.

- Montenegro's observation: if t = $\lfloor n^{1/2}(\log n)^{\mathbf{3/2}} \rfloor$, then the collision exponent is reduced to $O(n^{1/2}(\log n)^{\mathbf{3/2}})$.

Agenda:

1. Results on the Pollard ρ algorithm

2. Random walks on graphs

3. New constructions of expanders for crypto

Brief History of Expander Graphs

- Definitions vary: usually undirected graphs with $\lambda \leq c\ \lambda_{triv}$ for some positive constant $c < 1$. Random walks mix rapidly.
- Originally shown to exist by counting methods by Pinsker: There are far more graphs than there are non-expander graphs.
- Margulis (70s, 80s), Lubotzky-Phillips-Sarnak (1986) give first constructions.
- LPS "Ramanujan graphs" use the (known) Ramanujan conjectures in their proof. The Ramanujan conjectures in number theory are a statement about optimal cancellation in random sums.
- Other constructions: Reingold-Vadhan-Wigderson "Zig-Zag", algebraic geometry. Have algebraic flavor.
- Unfortunately existing constructions are not suitable for implementation (either too slow, or too large a probability of returning quickly to a previously visited node because undirected).

A simpler version: “GRH Graphs”

New, conditional construction of expander graphs.

- Let Q be a large integer.

- Let S = { primes $p < (\log Q)^B$, $p \nmid Q$ } , for B > 2.

- Define the graph Γ to have
  - vertices $V=(\mathbb{Z}/Q\mathbb{Z})^*$.
  - edges connecting **v** to **pv**, for each **v** $\in$ **V** and **p** $\in$ **S**.
  - (Γ is the *Cayley graph* of the group $(\mathbb{Z}/Q\mathbb{Z})^*$ with respect to the generating set S).

- Theorem – Assuming GRH, Γ is an expander: its nontrivial eigenvalues satisfy the bound

$$|\lambda| = O(k^{1/2+1/B}). \quad [k = \text{degree} = \lambda_{\text{triv}}]$$

Sketch of Proof

- The graph is a Cayley graph of $G = (\mathbb{Z}/Q\mathbb{Z})^*$, and so characters G of are eigenfunctions of A:

$$(A\chi)(g) = \sum_{s\in S} \chi(sg) = (\sum_{s\in S} \chi(s))\cdot\chi(g)$$

- In our case, the eigenvalue is

$$\lambda_\chi \;=\; \sum_{p\in S} [\, \chi(p) + \chi(p^{-1})\,] \;=\; 2 \text{ Re} \sum_{p\in S} \chi(p).$$

- Since G is abelian, there are as many characters as eigenfunctions, so the entire spectrum is obtained this way.
- Trivial character: trivial eigenvalue = degree = k.
- Nontrivial character: bound $|\lambda| = O(k^{1/2+1/B})$.
  This follows from GRH (it is a problem about primes in progressions).
  All one needs for *some* expansion is

  $$|\lambda| < k\cdot(1\text{-}c/(\log Q)^{\text{power}})$$

  – E.g. when χ is the quadratic character, this is basically the question of finding the least prime nonresidue mod Q – a difficult analytic problem.
- Could also use Lindelöf Hypothesis to get weaker bounds.

Generalization to other groups

- Argument applies also to other groups from number theory.
- Example (just mentioned in Couveignes' talk):
  - Let G = $\mathcal{I}$ be the ideal class group of an order in an imaginary quadratic number field $\mathbb{Q}$(-D).
  - Let S = ideal classes represented by prime ideals of norm $O((\log D)^B)$, B>2.
  - Then (assuming GRH) the Cayley graph generated by G and S is an expander. In particular S generates G!
    - uses Hecke's theory of Grossencharacter L-functions, cancellation of Fourier coefficients of θ-functions.
  - This is connected to elliptic curves: G represents ordinary elliptic curves, and S represents computable isogenies between them.
  - Connection proves the [JMV] isogeny result from David Jao's talk (continued…).

Random Reducibility of EC DLOG

- In elliptic curve crypto, curves are randomly selected based on the field and point count. *Is this justified?*
- Using previous graph, we show it is – modulo technical assumptions which do not arise in practice:

<u>Jao, M-, Venkatesan (2004):</u> Assuming GRH, the DLOG problem on ECs over the same field with the same point count is "random reducible" in the following sense:

> Given any algorithm A that solves DLOG on a fraction of curves in a "level", one can probabilistically solve DLOG on any curve in the same level with polylog(q) queries to A with random inputs.

"Level" means same End(E) – doesn't matter in practice. Hence this shows that the difficulty of DLOG is solely determined by the ground field and point count.

Another style of expanders

- Notion of additive reversalization:
  - Suppose that $A = A^t$ is the adjacency matrix of an undirected graph with good separation: $\|Af\| \leq c\|f\|$ for any $f \perp \mathbb{1}$, where $c < k =$ degree.
  - If P is any permutation, then

$$\|(AP + P^t A)f\| \;\leq\; \|APf\| + \|P^t Af\| \;\leq\; c\|Pf\| + \|Af\| \;\leq\; c\|f\| + c\|f\|$$

  so its operator norm on $f \perp \mathbb{1}$ is bounded by 2c (vs. $\lambda_{triv} = 2k$).
  - This still has significant eigenvalue separation.
- Idea – even if A has bad eigenvalue separation, $AP+(AP)^t$ might have excellent separation.
- Application: the circle graph on $\mathbb{Z}/n\mathbb{Z}$, with edges connecting $x-1 \leftrightarrow x \leftrightarrow x+1$, has the poorest possible spectral gap $\approx 1/n$.
  - Apply permutation P: $x \mapsto r \cdot x$, with $(r,n)=1$.
  - Get good separation (next slide).
- This shows a fundamental randomness property of integers: adding and multiplying mixes very quickly. Since these are basic operations, it has some applications.
- One of them is the Pollard ρ expansion used earlier in this talk.

Making stream ciphers: Goal is speed

- Theorem: Let N and r be relatively prime integers > 1. Form a 4-regular graph on $\mathbb{Z}/N\mathbb{Z}$ by connecting x to $r(x+1)$ and $r(x-1)$.
  Then the eigenvalues of the adjacency matrix either satisfy:
  - $\lambda = 4\cos(2\pi k/N)$ for those k with $r\cdot k = k \pmod N$ or
  - $|\lambda| \leq 4 - c(\log N)^{-2}$ for some $c > 0$.

  (Good expanders if N=2^k; fast nonlinearity on machine hardware)

- For group theoretic reasons, a bounded set of affine transformations on $\mathbb{Z}/N\mathbb{Z}$ cannot have fixed eigenvalue separation, so the logs are necessary.

- Related graph: take $(x+1)^r$, $(x-1)^r$ instead. This seems to have bounded separation (above constraints do not apply).

- This graph is used in a new stream cipher ("MV3") which is twice as fast as RC4, and whose statistical properties can be proven from the expansion.
  [Keller, M-,Mironov,Venkatesan – CT-RSA 2007]

Conclusions:

- Mathematics of expanders can be used to prove common beliefs about important crypto algorithms.
- Pollard ρ algorithm finds collisions in $O(n^{1/2}(\log n)^{3/2})$ time with arbitrarily high probability.
- For typical primes n, this collision is nondegenerate, i.e. the algorithm solves DLOG with high probability in this many steps.
- EC crypto selection practice of relying on point count is justified, assuming GRH.
- Principles from the proofs can be used to design other expanders with cryptographic applications.

In general $f = \sum c_k \chi_k$

We want to compute

$< Af, Af >$

$$\begin{bmatrix} f(x+1)+f(x+y)+f(2x) \end{bmatrix} = \begin{bmatrix} & A & \end{bmatrix} \begin{bmatrix} f \end{bmatrix}$$

If $f = \chi_k$ then $\chi_k(x) = e^{\frac{\pi i k x}{n}}$

$$< A\sum c_k \chi_k , A\sum c_k \chi_k >$$

$$=$$

$$\sum_{k,\ell\neq 0} c_k \overline{c_\ell} \left[\langle d_k\chi_k, d_\ell\chi_\ell\rangle + \langle \chi_{2k}, \chi_{2\ell}\rangle + \langle d_k\chi_k, \chi_{2\ell}\rangle + \langle \chi_{2k}, d_\ell\chi_\ell\rangle\right]$$

$$\begin{bmatrix} \chi_k(x+1)+\chi_k(x+y)+\chi_k(2x) \end{bmatrix} = \begin{bmatrix} & A & \end{bmatrix} \begin{bmatrix} \chi_k \end{bmatrix}$$

Using $< \chi_i, \chi_j > = 0$, if $i \neq j$ and n else

$$\leq \quad n\left(5\sum |c_k|^2 + 2\sum |c_k||c_{2k}||d_{2k}|\right).$$

$d_k\chi_k + \chi_{2k}$

where $d_k = \chi_k(1) + \chi_k(y)$,

$|d_k| = 2|\cos(\pi\, k(y-1)/n)|$.

One has that $|d_k| = 2|\cos(\frac{\pi k(y-1)}{n})| = 2\lambda_{k(y-1)}$.

Let n be an odd integer and $\lambda_k = |\cos(\pi k/n)|$ for $k \in \mathbb{Z}/n\mathbb{Z}$. Consider the quadratic form $Q : \mathbb{R}^{n-1} \to \mathbb{R}$ given by

$$Q(x_1, \ldots, x_{n-1}) \ := \ \sum_{k=1}^{n-1} x_k \, x_{2k} \, \lambda_k \, , \tag{3.1}$$

in which the subscripts are interpreted modulo n.

Proposition 3.1. *There exists an absolute constant $c > 0$ such that*

$$|Q(x_1, \ldots, x_{n-1})| \ \leq \ \left(1 - \frac{c}{(\log n)^2}\right) \sum_{k=1}^{n-1} x_k^2 \, . \tag{3.2}$$

$$\langle A \sum c_k \chi_k , A \sum c_k \chi_k \rangle$$

$=$

$$\leq \ n \left(5 \sum{}' |c_k|^2 + 2 \sum{}' |c_k| |c_{2k}| |d_{2k}| \right) .$$

We now need a lemma on quadratic from minimization

$$\|Af\| \ \leq \ \left(3 - \frac{c}{(\log n)^2}\right) \|f\| \tag{3.8}$$

Note that $|d_k| = 2\lambda_{k(y-1)}$, and that $y-1$ and 2 are invertible in $\mathbb{Z}/n\mathbb{Z}$, by assumption in (3.7). The result now follows from (3.2) with the choice of $x_{2(y-1)k} = |c_k|$.

Operator Norm Bound for A

- Key bound: we show that

$$\|Af\| \;\le\; \left(3 - \frac{c}{(\log n)^2}\right) \|f\|$$

  for all $f \in L^2(V)$ which are orthogonal to $\mathbb{1}$.

- This is equivalent to $\|A_{\text{restricted to } \mathbb{1}^\perp}\| \;\le\; 3 - \frac{c}{(\log n)^2}$
- Method of Proof:
  - Recall $(Af)(x) = f(x+1) + f(x+y) + f(2x)$.
  - Look at basis of the orthogonal complement consisting of nontrivial characters of $(\mathbb{Z}/n\mathbb{Z})$ given by $\chi_k(x) = e^{2\pi i k x/n}$, $k \neq 0$.

$$\langle \chi_k, \chi_\ell \rangle \;=\; \begin{cases} n\,, & k = \ell \\ 0\,, & \text{otherwise}\,. \end{cases}$$

  - We write $f = \sum_{k\neq 0} c_k\chi_k$, so that $||f||^2 = n \sum |c_k|^2$.
  - We have $A\chi_k = d_k\chi_k+\chi_{2k}$, where $d_k = \chi_k(1) + \chi_k(y)$, $|d_k| = 2|\cos(\pi k(y-1)/n)|$.

Operator Norm Bound for A (ctd.)

- We compute:

$$\|Af\|^2 = \langle Af, Af\rangle = \langle \sum c_k A\chi_k, \sum c_k A\chi_k \rangle =$$

$$\sum_{k,\ell\neq 0} c_k\,\overline{c_\ell}\,[\langle d_k\chi_k, d_\ell\chi_\ell\rangle + \langle \chi_{2k}, \chi_{2\ell}\rangle + \langle d_k\chi_k, \chi_{2\ell}\rangle + \langle \chi_{2k}, d_\ell\chi_\ell\rangle]$$

$(|d_k| \leq 2)$

$$\leq \quad n\left(5\sum |c_k|^2 + 2\sum |c_k||c_{2k}||d_{2k}|\right).$$

- The savings is gained from the second sum and the following quadratic form bound:
  - if $Q(x_1,\ldots,x_{n-1}) \; := \; \sum_{k=1}^{n-1} x_k\, x_{2k}\, \lambda_k$, where $\lambda_k = |\cos(\pi k/n)|$
  - then $|Q(x_1,\ldots,x_{n-1})| \; \leq \; \left(1 - \frac{c}{(\log n)^2}\right) \sum_{k=1}^{n-1} x_k^2\,.$
- This can be viewed as a "reciprocal" Hilbert inequality.