Lattice reduction techniques and predicting pseudorandom number generators

Jaime Gutierrez

ALGORITHMIC **M**ATHEMATICS **A**ND **C**RYPTOGRAPHY

Universidad de Cantabria, Santander

ORGANIZATION

- Lattices:
  - Shortest Vector Problem, **SVP**.
  - Closest Vector Problem, **CVP**.
  - LLL reduced bases.
  - Lattice reduction in computer algebra:
    - Polynomial factorization.
    - Ideal decomposition and intermediate fields.

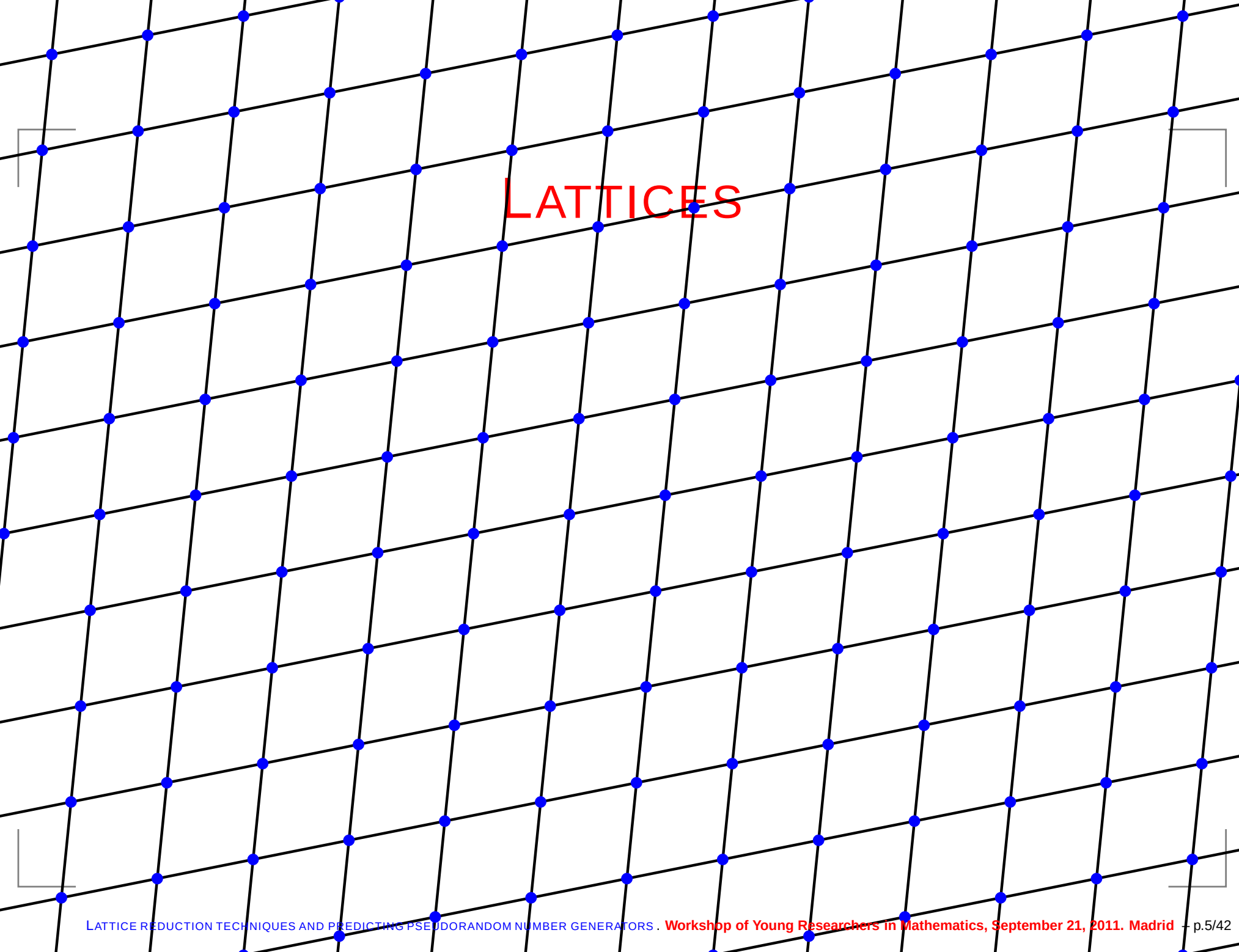
ORGANIZATION

- Lattice reduction in Cryptography:
  - Pseudorandom number generators
    - Linear and non-linear congruential generator.
    - Cryptographically secure generator and predicting pseudorandom number generators.
    - Linear generator over elliptic curves.
  - Integer factorization with extra information.

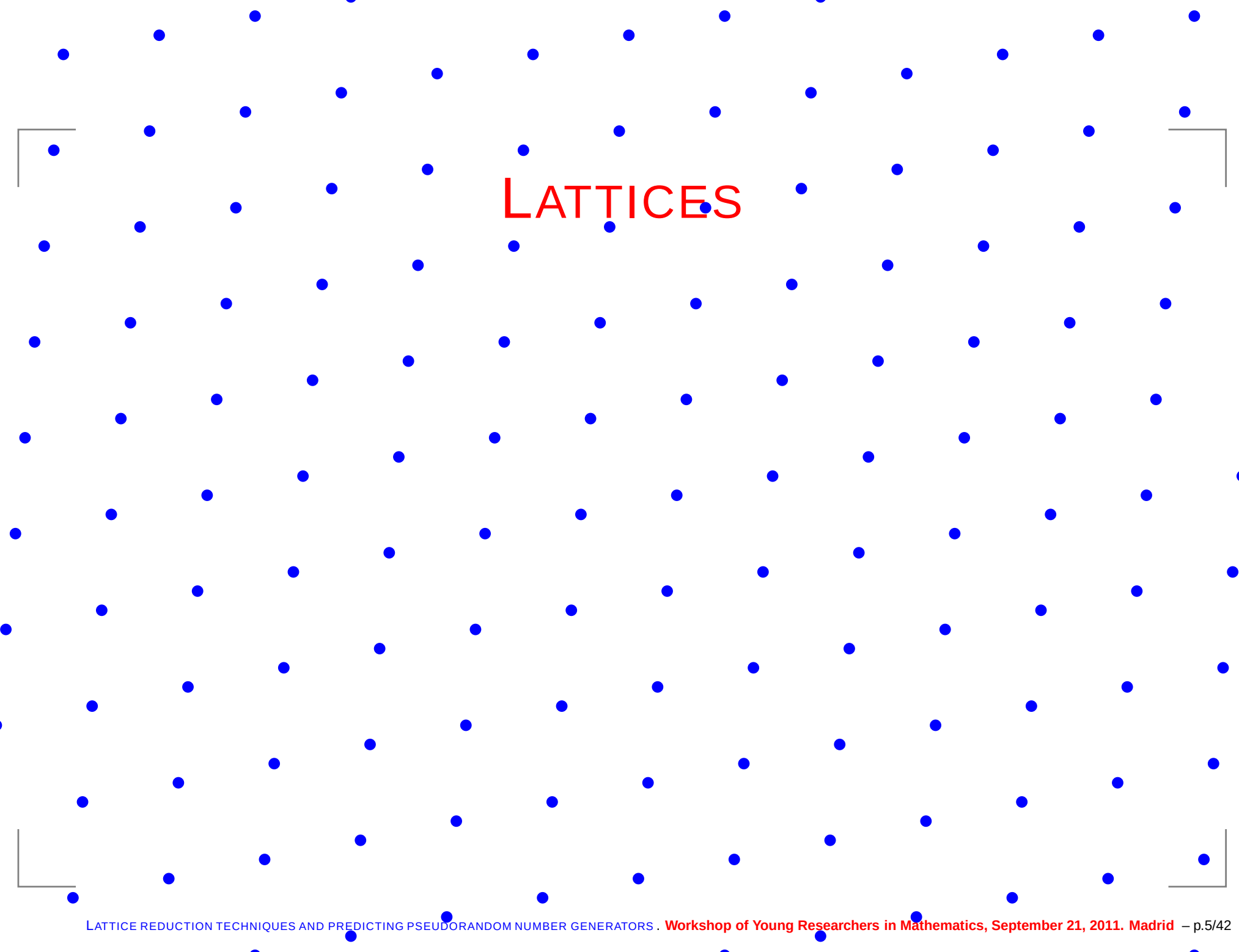
LATTICES

LATTICES

LATTICES



LATTICES



$$\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^m, \text{ l.i.}$$

$$\mathcal{L} = \mathcal{L}([\mathbf{b}_1 | \cdots | \mathbf{b}_n]) =$$

$$= \left\{ \sum_{i=1}^{n} \lambda_i \mathbf{b}_i \;\middle|\; \lambda_i \in \mathbb{Z} \right\}.$$

$B = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ is a **basis** of $\mathcal{L}$.

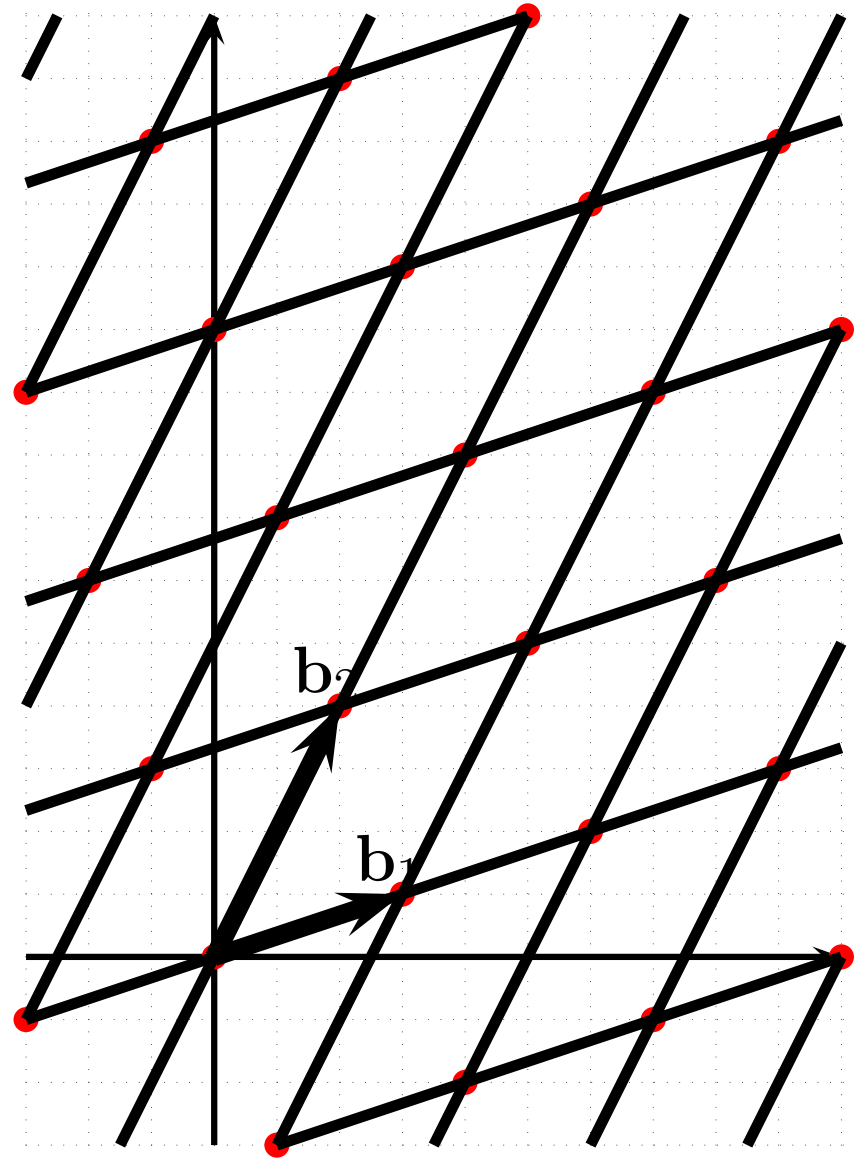
$$\mathcal{L}([(1,0),(0,1)] =$$

$$\mathcal{L}([(2010,1),(2011,1)]) = \mathbb{Z}^2$$

$$\mathcal{L}([(3,1),(2,4)]) =$$

$$\{(x,y) \in \mathbb{Z}^2 : 3x + y \equiv 0 \bmod 10\}$$

[LAGRANGE, GAUSS, MINSKOWSKI]



DISCRETE SUBGROUPS

In general, $\mathcal{L}(B)$ is not a lattice:

DISCRETE SUBGROUPS

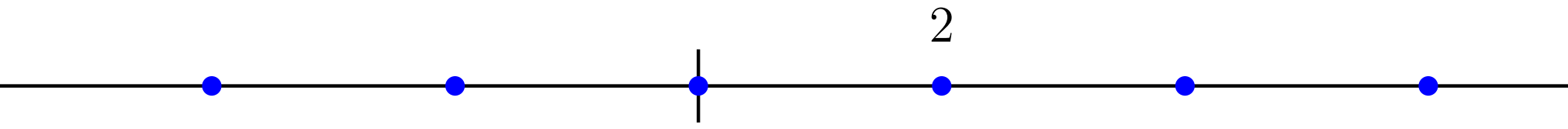
In general, $\mathcal{L}(B)$ is not a lattice:

$$B := (2, \sqrt{2})$$

DISCRETE SUBGROUPS

In general, $\mathcal{L}(B)$ is not a lattice:

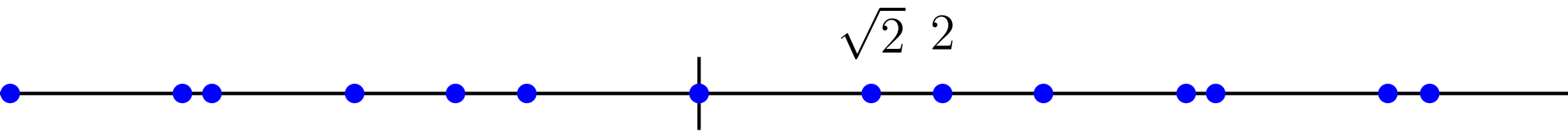
$$B := (2, \sqrt{2})$$



DISCRETE SUBGROUPS

In general, $\mathcal{L}(B)$ is not a lattice:

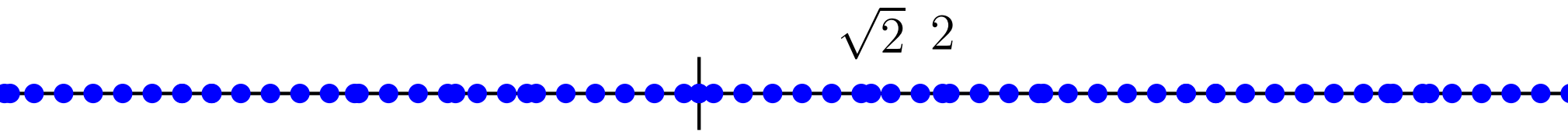
$$B := (2, \sqrt{2})$$



DISCRETE SUBGROUPS

In general, $\mathcal{L}(B)$ is not a lattice:

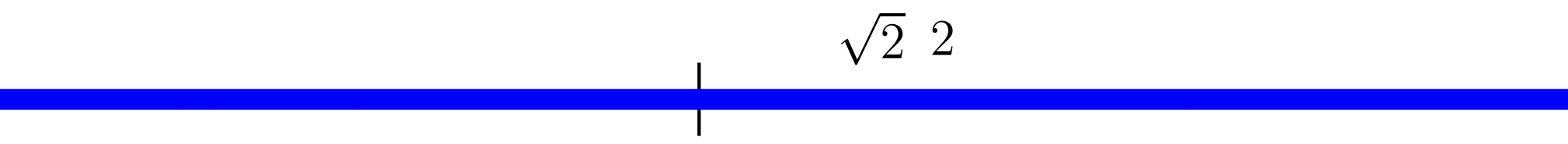
$$B := (2, \sqrt{2})$$



DISCRETE SUBGROUPS

In general, $\mathcal{L}(B)$ is not a lattice:

$$B := (2, \sqrt{2})$$



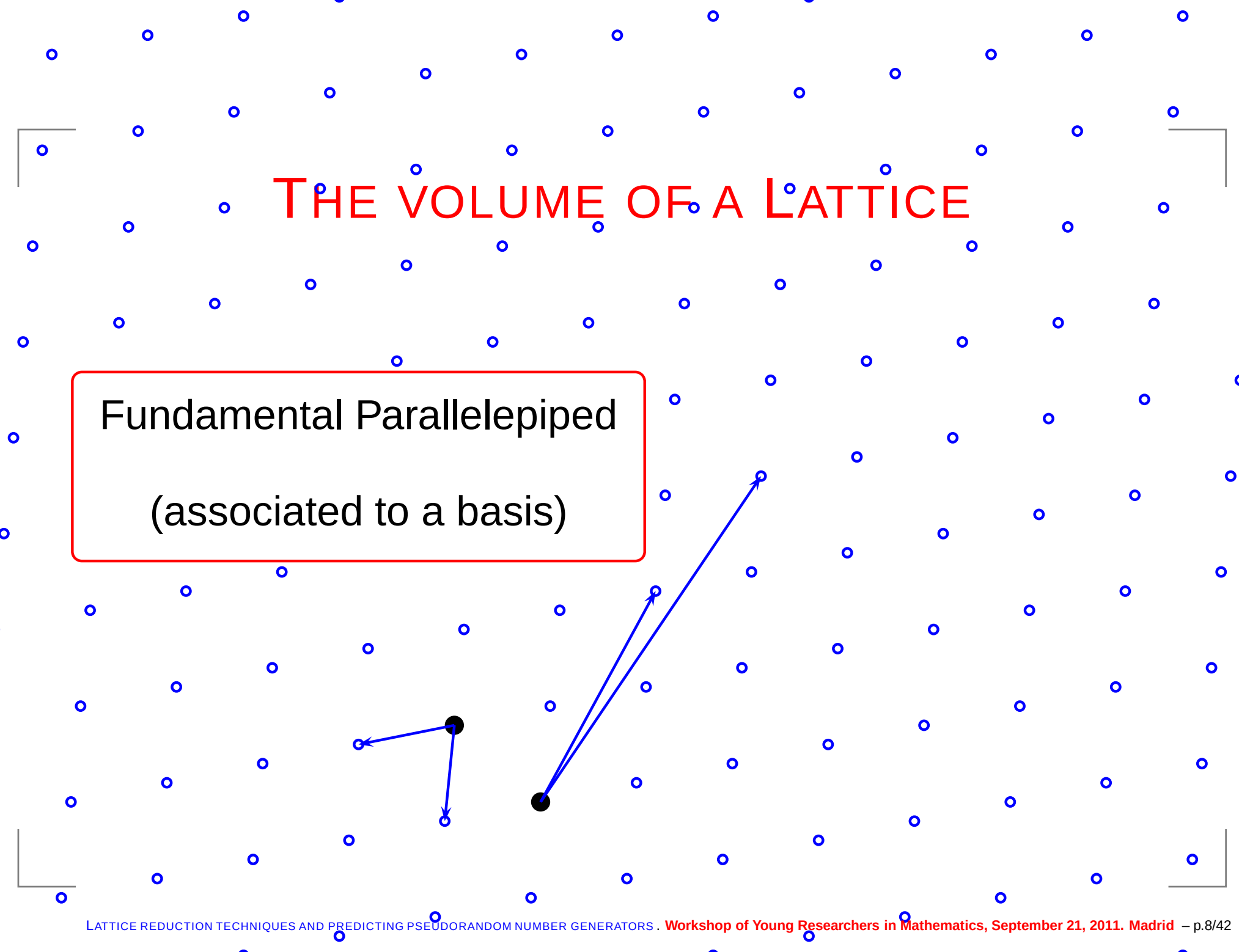
Lattices are **discrete** subgroups.

$$\lambda_1 := \min\{\|\mathbf{v}\| \ / \ \mathbf{v} \in \mathcal{L}\backslash\{0\}\}$$

The volume of a Lattice

Fundamental Parallelepiped

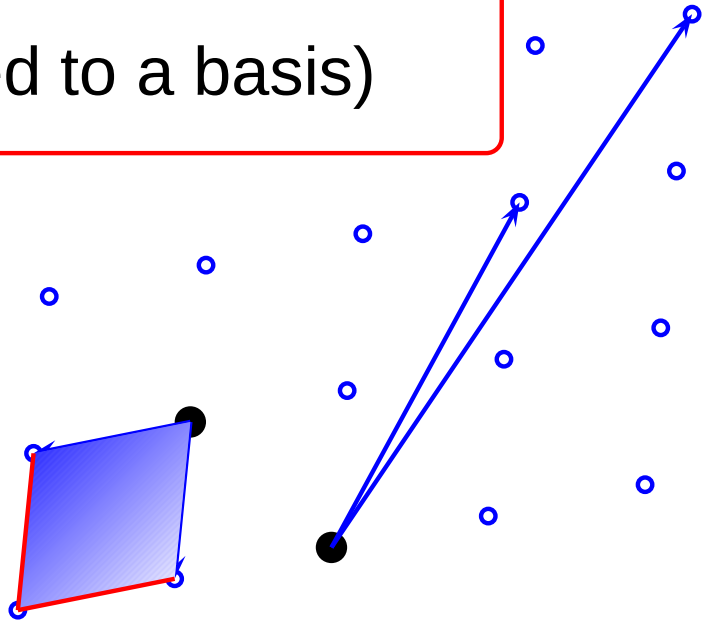
(associated to a basis)



The volume of a Lattice

Fundamental Parallelepiped

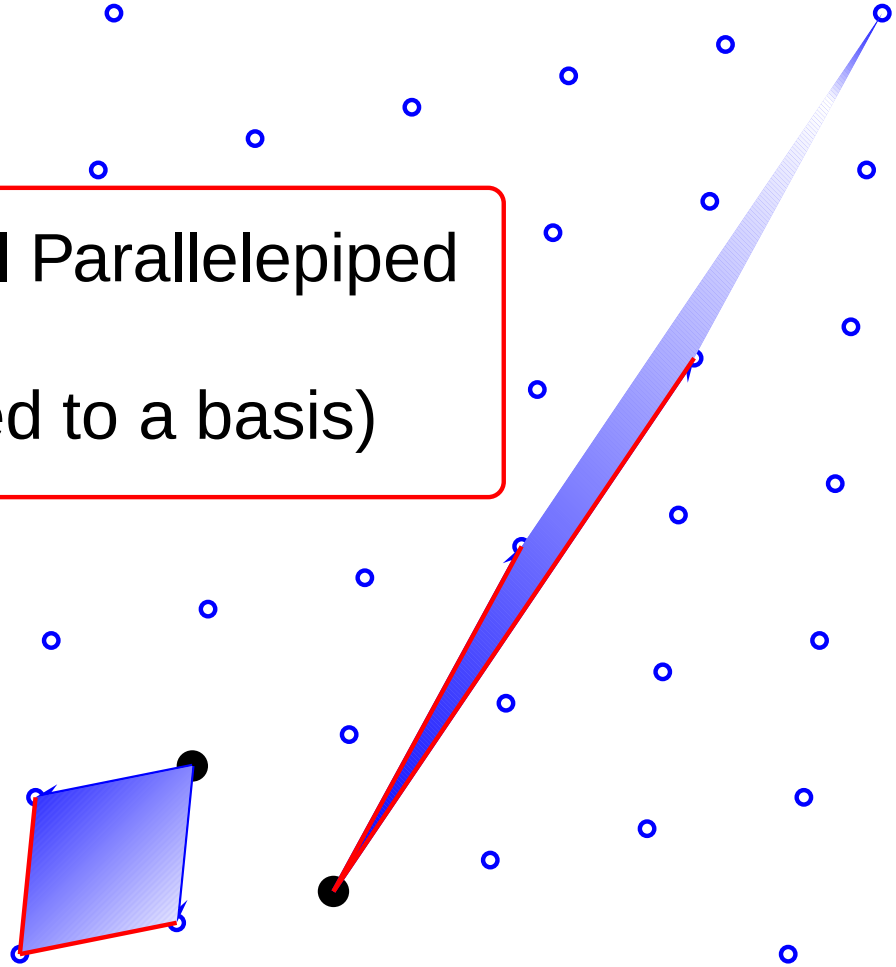
(associated to a basis)



The volume of a Lattice

Fundamental Parallelepiped

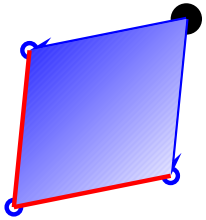
(associated to a basis)

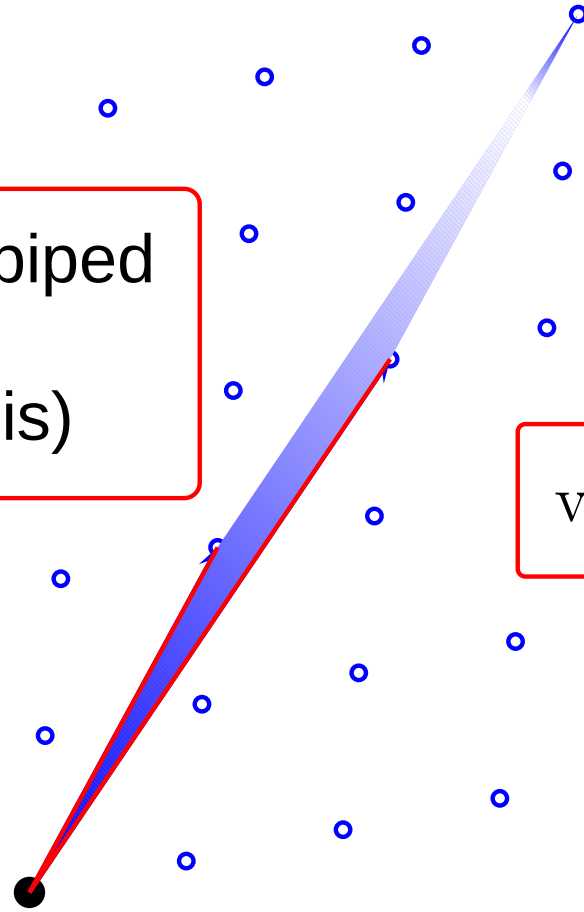


The volume of a Lattice

Fundamental Parallelepiped

(associated to a basis)





$$\operatorname{vol} \mathcal{L} = \sqrt{|B^t B|}$$

MAIN BOUNDS

- Norm ℓ_∞:

$$\lambda_1 \leq (\mathrm{vol}\,\mathcal{L})^{1/n}$$

- Norm ℓ_1:

$$\lambda_1 \leq (n!\,\mathrm{vol}\,\mathcal{L})^{1/n}$$

- Norm ℓ_2:

$$\lambda_1 \leq \sqrt{\gamma_n}(\mathrm{vol}\,\mathcal{L})^{1/n}$$

$$\frac{n}{2e\pi} + o(n) \leq \gamma_n \leq \frac{1'744}{e\pi}n + o(n)$$

Gauss Heuristic

MAIN PROBLEMS

- SVP

$$\min\{\|\mathbf{v}\| \ / \ \mathbf{v} \in \mathcal{L}\backslash\{0\}\}$$

NP hard ??

- CVP

$$\min\{\|\mathbf{v} - \mathbf{t}\| \ / \ \mathbf{v} \in \mathcal{L}\}$$

NP hard

Fixed n is polynomial: [KANNAN (1987)]

Approximation: [LLL = A. LENSTRA, H. LENSTRA, L. LOVÁCS (1982)],

[T. BABAI (1986)], [M. AJTAI (1998)]

LLL-REDUCTION

$B = [\mathbf{b}_1 | \cdots | \mathbf{b}_n]$ is a δ-LLL reduced basis of $\mathcal{L}$ if

- $|\mu_{i,j}| \leq 1/2$
- $\delta \|\mathbf{b}_i^*\| \geq \|\mu_{i+1,i}\mathbf{b}_i^* + \mathbf{b}_{i+1}^*\|,$

where $1/4 < \delta < 1$ y $\mathbf{b}_1^*, \ldots, \mathbf{b}_n^*$ is the Gram-Schmidt orthogonal basis : $\mathbf{b}_i^* := \mathbf{b}_i - \sum_{j<i} \mu_{ij}\mathbf{b}_j, \mu_{i,j} := \frac{<\mathbf{b}_i, \mathbf{b}_j^*>}{<\mathbf{b}_j^*, \mathbf{b}_j^*>}.$

$$\|\mathbf{b}_1\| \leq \left(\frac{1}{\delta - 1/4}\right)^{\frac{n-1}{2}} \lambda_1$$

LLL-algorithm is polynomial in $M = \max\{n, \log(\max_i \|\mathbf{b}_i\|)\}$.

INTEGER LATTICES

$\mathcal{L}$ consist of integer solutions $\mathbf{x} = (x_0, \ldots, x_{s-1}) \in \mathbb{Z}^s$ of a system of congruences

$$\sum_{i=0}^{s-1} a_{ij} x_i \equiv 0 \bmod q_j, \qquad j = 1, \ldots, m,$$

modulo the intergers $q_1, \ldots, q_m$.

- Typically, $\mathrm{vol}(\mathcal{L}) = Q = q_1 \ldots q_m$.
- SVP and CVP are polynomials in $\log Q$.

THE PROBLEM

Given:

$f(x)$ *polynomial in* $\mathbb{Q}[x]$

Find:

$g(x), h(x), \bar{f}(x) \in \mathbb{Q}[x]$ *verifying*

$$f(x)\bar{f}(x) = g(h(x)),$$

$$1 < \deg g, \deg h < \deg f.$$

POLYNOMIAL DECOMPOSITION AND RATIONAL SUBFIELDS

If $\deg \bar{f} = 0$, then

$$f(x) = g(h(x)).$$

Lüroth's Theorem. [A. SCHINZEL (1982)]

Let $\mathbb{F}$ be a field such that $\mathbb{K} \subset \mathbb{F} \subset \mathbb{K}(x_1, \ldots, x_n)$ and $\mathrm{tr.deg.}(\mathbb{F}/\mathbb{K}) = 1$. Then there exists $h \in \mathbb{K}(x_1, \ldots, x_n)$ such that $\mathbb{F} = \mathbb{K}(h)$. Also, if the field contains a polynomial, then a polynomial generator exists.

$$\begin{array}{ccc} \{[(g,h)] : f = g(h)\} & \longleftrightarrow & \{\mathbb{F} : \mathbb{K}(f) \subset \mathbb{F} \subset \mathbb{K}(x)\} \\ [(g,h)] & \longleftrightarrow & \mathbb{F} = \mathbb{K}(h). \end{array}$$

[GUTIERREZ- LANDAU-KOZEN-VON ZUR GATHEN (1989)],

[J. G. AND R. RUBIO AND D. SEVILLA (2005)]

MOTIVATION

Simplifying of equations. Evaluating functions

$$f(x) = x^6 - 6\,x^5 + 5\,x^4 + 12\,x^3 + 4\,x^2 + 1$$

$$f(x) = g(x) \circ h(x) = g(h(x))$$

$$g(x) = x^2 - 2\,x + 2, \quad h(x) = x^3 - 3\,x^2 - 2\,x + 1$$

$$g(x) = 0 \Rightarrow x = \beta_1, \quad x = \beta_2$$

$$h(x) = \beta_i, \quad (i = 1, 2) \quad \Rightarrow x = \alpha_i, \quad (i = 1, 2, 3)$$

MOTIVATION

Reparametrizing algebraic parametric curves

Parametric equations	Implicit equations
$x = t^2, \quad y = t^4$	$y - x^2$

$$x = t, \quad y = t^2$$

[SEDEBERG (1989)], [J. G. AND RECIO (1996)]

MOTIVATION: ALGEBRAIC SUBFIELDS

If $f(x)$ is irreducible and $f(\alpha) = 0$, then the following statements are equivalent:

- f is an ideal decomposable.
- $Gal_{\mathbb{Q}}(f)$ acts imprimitively on the roots of f.
- A proper subfield, $\mathbb{Q}(\beta)$, exists with

$$\mathbb{Q} \subset \mathbb{Q}(\beta) \subset \mathbb{Q}(\alpha).$$

$$\begin{array}{ccc} \{[(\bar{f}, g, h)] : f\bar{f} = g(h)\} & \longleftrightarrow & \{\mathbb{F} : \mathbb{Q} \subset \mathbb{F} \subset \mathbb{Q}(\alpha)\} \\ \left[(\bar{f}, g, h)\right] & \longleftrightarrow & \mathbb{F} = \mathbb{Q}(h(\alpha)). \end{array}$$

[S. LANDAU AND G. MILLER (1985)], [J. KLÜNERS AND M. POHST (1997)]
[J. G. AND D. SEVILLA (2006)]

THE ALGORITHM

We divide the problem into two parts:

1. Compute candidates $h(x)$.
2. Given $f(x)$ and $h(x)$, compute (if it exists) $\bar{f}(x), g(x)$:

$$f(x)\bar{f}(x) = g(h(x))$$

- Compute $\bar{f}(x), g(x)$ from $f(x)$ and $h(x)$ is solving a linear system of equations.
- The hard part is compute $h(x)$.

THE BASIC IDEA

From

$$f(x)\bar{f}(x) = g(h(x)),$$

There are two distintict roots of $f(x)$, say α_i and α_j for which

$$h(\alpha_i) = h(\alpha_j)$$

$h(x) = \sum_{k=1}^{s} h_k x^k \in \mathbb{Z}[x]$ for some s with: $1 < s < \deg f$

Find the coefficients $h_1, h_2, \ldots, h_s$ in $\mathbb{Z}$ satisfying

$$h_1(\alpha_i - \alpha_j) + h_2(\alpha_i^2 - \alpha_j^2) + \cdots + h_s(\alpha_i^s - \alpha_j^s) = 0.$$

INTEGER RELATIONS

A nonzero vector $\mathbf{h} = (h_1, \ldots, h_s) \in \mathbb{Z}^s$ is called an INTEGER RELATION for the real numbers $\gamma_1, \ldots, \gamma_s$ if

$$h_1\gamma_1 + \cdots + h_s\gamma_s = 0.$$

Given $\bar{\gamma}_1, \ldots, \bar{\gamma}_s$ complex numbers approximating to the algebraic numbers $\gamma_1, \ldots, \gamma_s$, and a parameter ϵ

- either finds an integer relation for $\gamma_1, \ldots, \gamma_s$ or
- proves that no relation of Euclidean length shorter than $1/\epsilon$ exists.

INTEGER RELATIONS AMONG ALGEBRAIC NUMBERS

$\mathcal{L}([\mathbf{b}_1|\cdots|\mathbf{b}_s]) \subset \mathbb{Q}^{s+2}$ the lattice spanned

$$\begin{cases} \mathbf{b}_1 = (1, 0, \ldots, 0, C \cdot Re(\bar{\gamma}_1), C \cdot Im(\bar{\gamma}_1)) \\ \quad \vdots \\ \mathbf{b}_s = (0, \ldots, 0, 1, C \cdot Re(\bar{\gamma}_s), C \cdot Im(\bar{\gamma}_s) \end{cases}$$

C is a large integer depend on ϵ, the height of $(\gamma_1, \ldots, \gamma_s)$ and $[\mathbb{Q}(\gamma_1, \ldots, \gamma_s) : \mathbb{Q}]$.

- Let $\mathbf{b}$ be the first vector of the LLL-basis:
  $$\mathbf{b} = (m_1, \ldots, m_s, C \cdot \sum_{i=1}^{s} m_i Re(\bar{\gamma}_i), C \cdot \sum_{i=1}^{s} m_i Im(\bar{\gamma}_i)).$$
- If $\|\mathbf{b}\| < 2^n/\epsilon^2$, then $(m_1, \ldots, m_s)$ is a solution. Otherwise, no solution shorter than $1/\epsilon$ exists.

[J. HÄSTAD AND B. JUST AND J. LAGARIAS AND C. SCHNORR (1989)]

BOUNDS ON THE COEFFICIENTS OF $h(x)$

Given a non-trivial ideal decomposition of polynomial $f \in \mathbb{Z}[x]$, i.e

$$f(x)\bar{f}(x) = g(h(x))$$

find an upper bound on the height $Ht(h(x))$ of $h(x)$.

- POLYNOMIAL DECOMPOSITION, i.e., $\bar{f}(x) = 1$, then
$$Ht(h(x)) < CHt(f(x)).$$
- If $f(x)$ is irreducible: [J. DIXON (1990)], [J. MCKAY (1996)].
- In general, ???

$$Ht(h(x)) < CHt(f(x))^{\deg f},$$

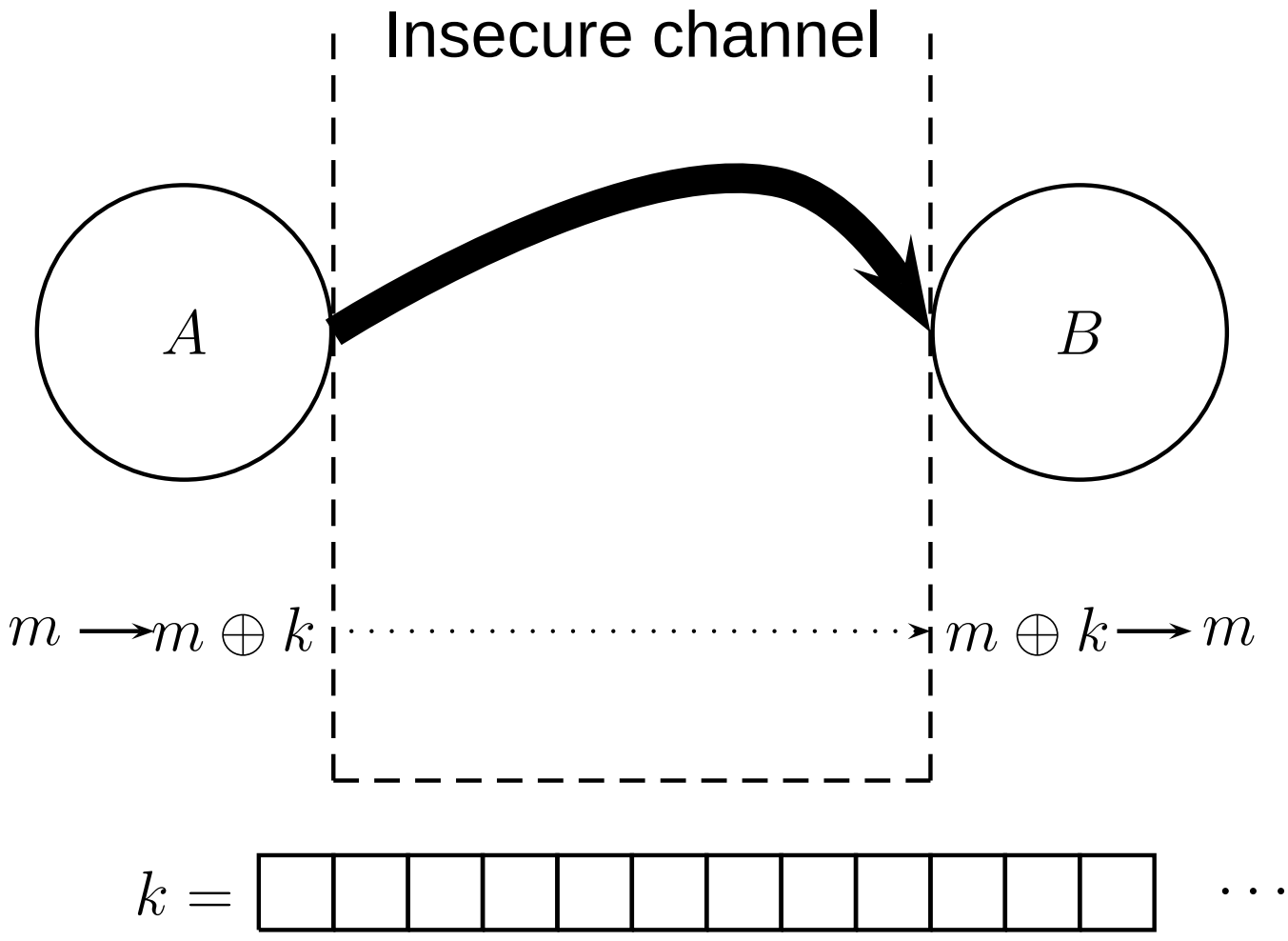
where C is a constant.

LATTICE REDUCTION

in

CRYPTOGRAPHY

STREAM CIPHERS



PSEUDORANDOM NUMBER GENERATORS

$$\varphi : \quad \mathbb{Z}_N \quad \to \quad \mathbb{Z}_N$$

$\varphi(u_i) \equiv u_{i+1} \bmod N, \qquad u_0$ is the seed.

$$\begin{array}{rccc} \varphi_L : & \mathbb{Z}_N & \to & \mathbb{Z}_N \\ & u & \rightsquigarrow & a\,u + b \end{array}$$

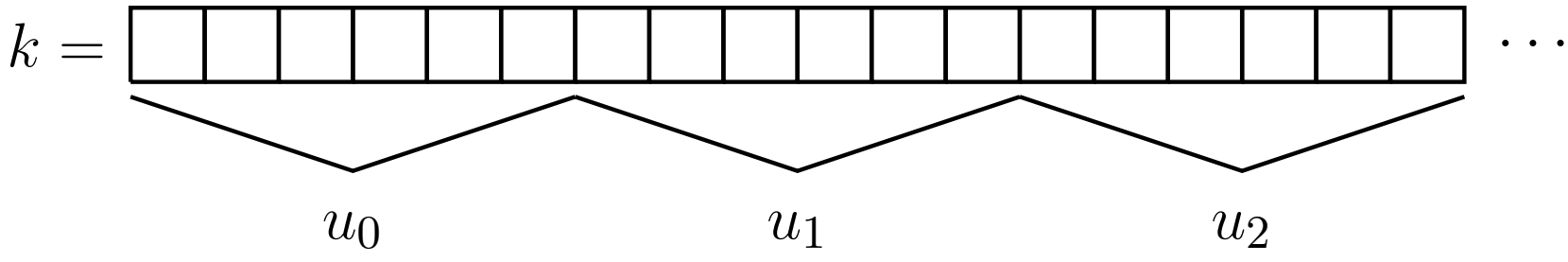
$$\begin{array}{rccc} \varphi_I : & \mathbb{F}_p & \to & \mathbb{F}_p \\ & 0 \neq u & \rightsquigarrow & a\,u^{-1} + b \\ & 0 & \rightsquigarrow & b \end{array}$$

POLLARD GENERATOR

$$c \in \mathbb{F}_p$$

$$u_0 \in \mathbb{F}_p \text{ (seed)}$$

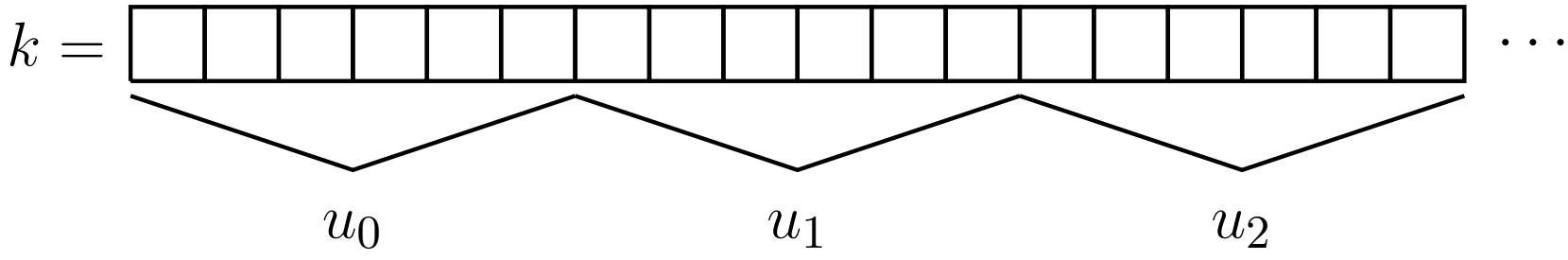
$$u_{n+1} \equiv_p u_n^2 + c$$



CRYPTOGRAPHICALLY SECURE GENERATOR

A **PRBG** *is cryptographically secure if there is no polynomial time algorithm which on input of the first l bits of an output sequence s can predict the $(l+1)^{st}$ bit of s with probability significant greater than $1/2$.*

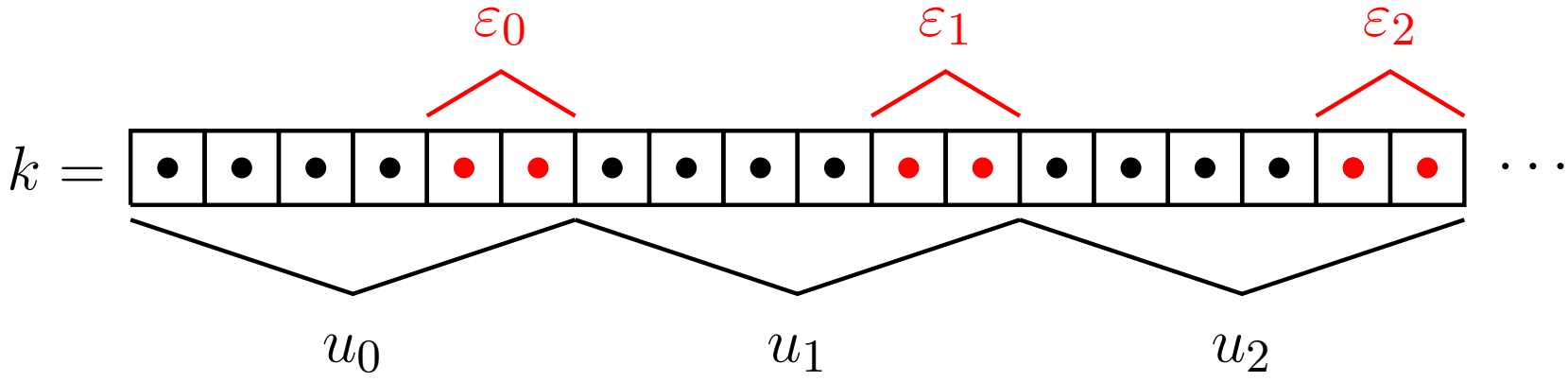
[YAO (1982)], [BLUM AND BLUM AND SHUB (1998)]



CRYPTOGRAPHICALLY SECURE GENERATOR

A **PRBG** *is cryptographically secure if there is no polynomial time algorithm which on input of the first l bits of an output sequence s can predict the $(l+1)^{st}$ bit of s with probability significant greater than $1/2$.*

[YAO (1982)], [BLUM AND BLUM AND SHUB (1998)]

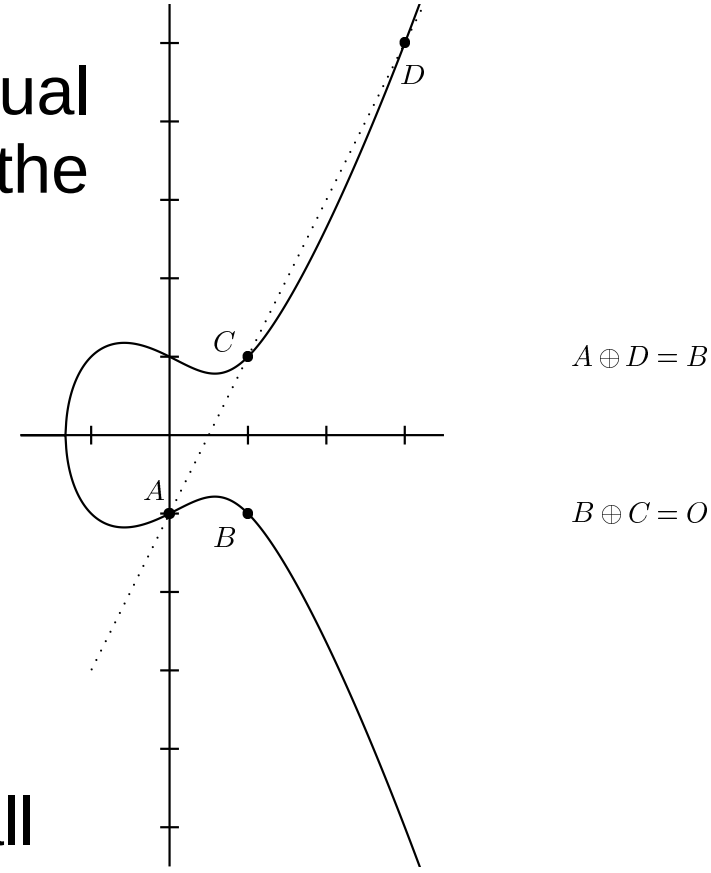


[S. BLACKBURN AND D. GÓMEZ AND J. G. AND I. SHPARLINSKI (2005)]

[D. GÓMEZ AND J. G. AND A. IBEAS (2006)], [KNUTH (1985)]

LINEAR GENERATOR ON ELLIPTIC CURVES

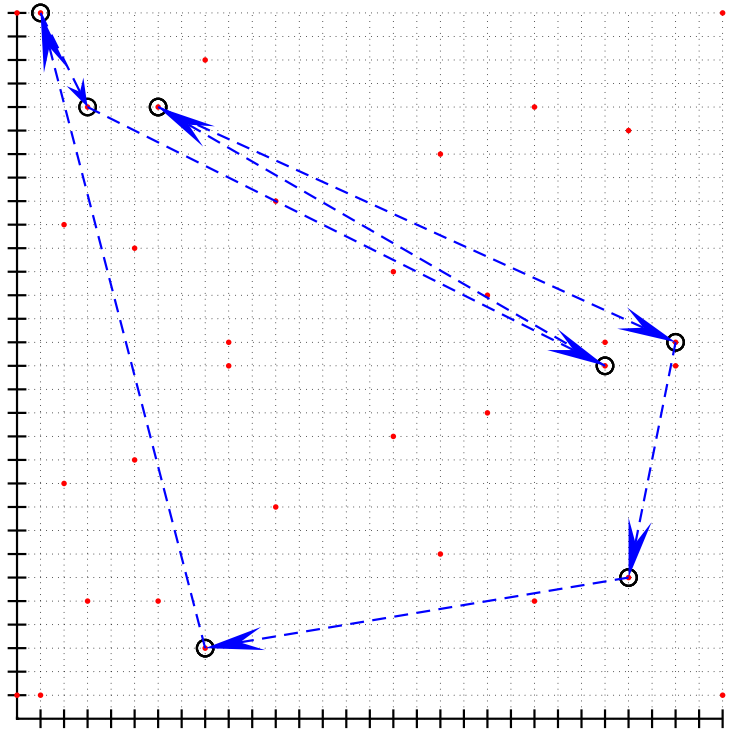
This generator employs the usual abelian group operation $(\oplus)$ on the set of points of an elliptic curve:



- A prime p.
- An elliptic curve $\mathbb{E} : Y^2 = X^3 + AX^2 + B$ over $\mathbb{F}_p$.
- The seed $U_0 \in \mathbb{E}$.
- The parameter $G \in \mathbb{E}$, which we call **composer**.

$$U_{n+1} = U_n \oplus G, \forall n \geq 0.$$

LINEAR GENERATOR ON ELLIPTIC CURVES



Toy example with a 7-periodic generator in an elliptic curve with 35 points over $\mathbb{F}_{31}$.

$$\mathbb{E} : Y^2 = X^3 - X^2 + 1$$

$$U_{n+1} = n(5, 11) \oplus (8, 3)$$

METHOD SKETCH

We assume access to:

- Approximations W_0, W_1 to the first two values U_0, U_1:
  - $U_i = (x_i, y_i), W_i = (\alpha_i, \beta_i)$,
  - $x_i = \alpha_i + e_i,\ y_i = \beta_i + f_i,\ |e_i|, |f_i| \leq \Delta$.
- The composer $G = (x_G, y_G)$.

From $U_0 \oplus G = U_1$ when $U_0 \notin \{G, -G\}$, we obtain (over $\mathbb{F}_p$):

$$x_G^3 + x_1 x_G^2 - x_0 x_G^2 - 2x_1 x_G x_0 - x_G x_0^2 + x_0^3 + 2y_G y_0 + x_1 = 0,$$

$$y_1 x_G - y_1 x_0 - y_G x_0 + y_G x_1 - y_0 x_1 + y_0 x_G = 0.$$

METHOD SKETCH

- Translate previous equations into a linear system in the approximation errors e_0, e_1, f_0, f_1.
- Find the smaller integer solution to the system (CVP)
- Check if the obtained solution is valid.

THEOREM. [J. G. AND A. IBEAS (2007)]

If the algorithm outputs a wrong solution, the first coordinate x_0 of the first value must satisfy a certain equation. This leads to the bound $O(\Delta^6)$ for the possibilities for x_0 in a failure example. So, when $\Delta < p^{1/6}$ we can expect with high probability the success of the guessing algorithm.

UNKNOWN COMPOSER

We take three approximations $W_i = (\alpha_i, \beta_i)$ to any three values (consecutive or not): $U_i = (x_i, y_i)$

$$y_i^2 = x_i^3 + Ax_i + B, i = 0, 1, 2.$$

Eliminating the curve parameters A, B and assuming that $U_0 \notin \{U_1, -U_1\}$ (that is, $x_0 \neq x_1$), we obtain the following equation:

$$-y_2^2x_1+y_2^2x_0+x_2^3x_1-x_2^3x_0-x_2y_0^2+x_2x_0^3+x_2y_1^2-x_2x_1^3-y_1^2x_0+x_1^3x_0+x_1$$

Substituting $x_i = \alpha_i + e_i$, $y_i = \beta_i + f_i$ for $i = 0, 1, 2$, we obtain a threshold of $p^{1/46}$ for the tolerance below wich we can expect successful guessing.

THE PROBLEM

INPUT : $N = PQ$ and the high-order h bits of P.

THE PROBLEM

INPUT : $N = PQ$ and the high-order h bits of P.

OUTPUT: The factorization of N, i.e, P and Q.

WHY STUDY THIS PROBLEM ?

WHY STUDY THIS PROBLEM ?

- because I like it,

WHY STUDY THIS PROBLEM ?

- because I like it,
- RSA
  - loss of the equipment that generated P and Q,
  - explicit release of partial extra information as part of a protocol, for instance exchange of secret,
  - timing measurements,
  - routine usage of P and Q to decrypt mail, sign messages, etc.,
  - poor physical security to guard P and Q,
  - any other heuristic attack . . .

[RIVEST AND SHAMIR (1986)], [COPPERSMITH (1995-1998)], [BONEH AND HOWGRAVE-GRAHAM (1999)], [MAY AND CORON (2005)]

ATTACKING RSA

DEFINITION. *We say that an integer w is a Δ-approximation to the integer u when $|w-u| \leq \Delta$.*

We can build a Δ-approximation P_0 to P, by taking the h high-order bits of P and $\lfloor \log P \rfloor + 1 - h$ zeroes. In this case, $\Delta = 2^{\lfloor \log P \rfloor + 1 - h} - 1$, that is,

$$P - P_0 \leq \Delta \cong \frac{P}{2^h}.$$

By dividing N into P_0, we obtain a Δ_1-approximation Q_0 to Q:

$$|Q - Q_0| \leq \Delta_1 \cong \frac{Q\Delta}{P}.$$

ATTACKING RSA

Let $\varepsilon_0 = P - P_0$ and $\varepsilon_1 = Q - Q_0$. From $N = PQ$ we obtain:

$$f(\varepsilon_0, \varepsilon_1) = 0,$$

where

$$f(\varepsilon_0, \varepsilon_1) = (P_0 + \varepsilon_0)(Q_0 + \varepsilon_1) - N.$$

And with

$$|\varepsilon_0| \leq \Delta, \quad |\varepsilon_1| \leq \Delta_1.$$

The main objective is to find small roots of this innocent polynomial $f(\varepsilon_0, \varepsilon_1)$.

SMALL ROOTS OF INTEGER BIVARIATE POLYNOMIALS

THEOREM. [D. COPPERSMITH (1997)]
Let $p(\varepsilon_0, \varepsilon_1)$ be an irreducible polynomial in two variables over $\mathbb{Z}$, of maximum degree δ in each variable separately. Let Δ, Δ_1 be bounds on the desired solutions x_0, y_0. Define $p^*(\varepsilon_0, \varepsilon_1) = p(\varepsilon_0\Delta, \varepsilon_1\Delta_1)$ and let W be the absolute value of the largest coefficient of $p^*(\varepsilon_0, \varepsilon_1)$. If

$$\Delta\Delta_1 \leq W^{2/(3\delta)-\epsilon}2^{-14\delta/3},$$

then in polynomial time in $(\log W, \delta, 1/\epsilon)$ we can find all integer pairs (x_0, y_0) with $p(x_0, y_0) = 0$ bounded by $|x_0| \leq \Delta, |y_0| \leq \Delta_1$.

ATTACKING RSA

We suppose that we know $N = PQ$ and the high-order $h = \frac{1}{4}\log_2 N$ bits of P. We apply the previous result to polynomial $f(\varepsilon_0, \varepsilon_1)$ and take:

$$|\varepsilon_0| < P_0 N^{-1/4} = \Delta,$$
$$|\varepsilon_1| < Q_0 N^{-1/4} = \Delta_1,$$
$$\delta = 1, \quad W = N^{3/4}.$$

Corollary. [D. COPPERSMITH (1997)]
In polynomial time we can find the factorization of $N = PQ$ if we know the high-order $(\frac{1}{4}\log_2 N)$ bits of P