楕円曲線暗号解読で10時間を1秒に短縮する学習λ法

2019年6月10日

後　保範（インダストリスパコン推進センター）

はじめに

- 現在の暗号システム
  公開鍵暗号と共通鍵暗号のハイブリット
  公開鍵暗号：鍵は2種類、鍵交換可能、処理低速
  共通鍵暗号：鍵は1種類、鍵交換不可、処理高速
- 現在の公開鍵暗号
  ほとんど下記のどちらか
  256ビット楕円曲線暗号（ECC、鍵交換:ECDH)
  2048ビットRSA暗号

楕円曲線暗号

- 楕円曲線暗号(ECC)
  $y^2 = x^3 + ax + b \pmod p$で定義される暗号
  x,y,a,bは整数でpとr(位数)は素数
- ρ法(並列はλ法)による解読(ECDLP)
  現在最速の解読法
  計算量 = $O(r^{1/2})$: rがnビットなら$O(2^{n/2})$
  間引き(特徴点)可能で記憶量が少なくできる

ρ法の計算手順 (シリアル)

- $Q=n \times P$となる整数nを求める
- 初期値$S_0=(x_0,y_0)$をランダムに与える
  $S_0=\alpha_0 \times P+\beta_0 \times Q$で整数$\alpha_0,\beta_0$を与える
- $S_h=S_j$ となるまで下記を計算$(k=1,2,..,j,..,h)$
  $S_k=f(S_{k-1})$,　$\alpha_k=f(\alpha_{k-1})$, $\beta_k=f(\beta_{k-1})$
- nを$\alpha_j,\alpha_h,\beta_j,\beta_h$より計算
  $n = (\alpha_h - \alpha_j) / (\beta_j - \beta_h) \pmod r$

λ法の計算手順 (並列)

- $Q=n\times P$となる整数nを求める
- 並列tに対しランダム初期値${}_tS_0=({}_tx_0,{}_ty_0)$
  ${}_tS_0={}_t\alpha_0\times P+{}_t\beta_0\times Q$で整数${}_t\alpha_0,{}_t\beta_0$を与える
- ${}_tS_h={}_uS_j$ となるまで下記を計算$(k=1,2,..,j,..,h)$
  ${}_tS_k=f({}_tS_{k-1})$,　${}_t\alpha_k=f({}_t\alpha_{k-1})$, ${}_t\beta_k=f({}_t\beta_{k-1})$
- nを${}_u\alpha_j,{}_t\alpha_h,{}_u\beta_j,{}_t\beta_h$より計算(u,tは並列番号)
  $n = ({}_t\alpha_h - {}_u\alpha_j) / ({}_u\beta_j - {}_t\beta_h) \pmod r$

ρ法(λ法)の$S_k=f(S_{k-1})$の例1

- 準備計算

  c_jを決め、$P_j=c_j \times P$, $Q_j=c_j \times Q$, $j=0,\ldots,7$

- $S_k=f(S_{k-1})$, $S_{k-1}=(x,y)$, $i = x \pmod 8$

  $x =< p/2$の場合

  $S_k = S_{k-1} + P_i$, $\alpha_k=\alpha_{k-1}+c_i \pmod r$, $\beta_k=\beta_{k-1}$

  $x > p/2$の場合

  $S_k = S_{k-1} + Q_i$, $\beta_k=\beta_{k-1}+c_i \pmod r$, $\alpha_k=\alpha_{k-1}$

ρ法(λ法)の$S_k=f(S_{k-1})$の例2

- $S_k=f(S_{k-1})$, $S_{k-1}=(x,y)$

  x =< p/3の場合

  $S_k = S_{k-1} + P$, $\alpha_k=\alpha_{k-1}+1 \pmod r$, $\beta_k=\beta_{k-1}$

  p/3 < x <= 2p/3の場合

  $S_k = S_{k-1} + Q$, $\beta_k=\beta_{k-1}+1 \pmod r$, $\alpha_k=\alpha_{k-1}$

  x > 2p/3の場合

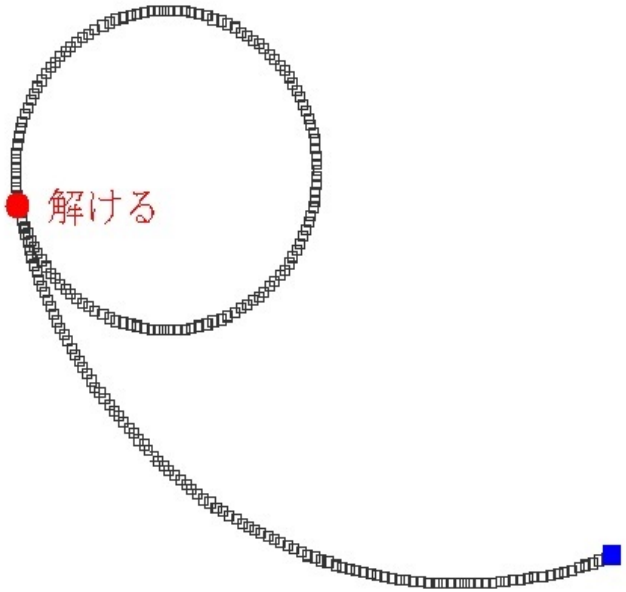
  $S_k = 2 \times S_{k-1}$, $\alpha_k=2\alpha_{k-1}$, $\beta_k=2\beta_{k-1} \pmod r$

ρ法による解読(軌跡)例

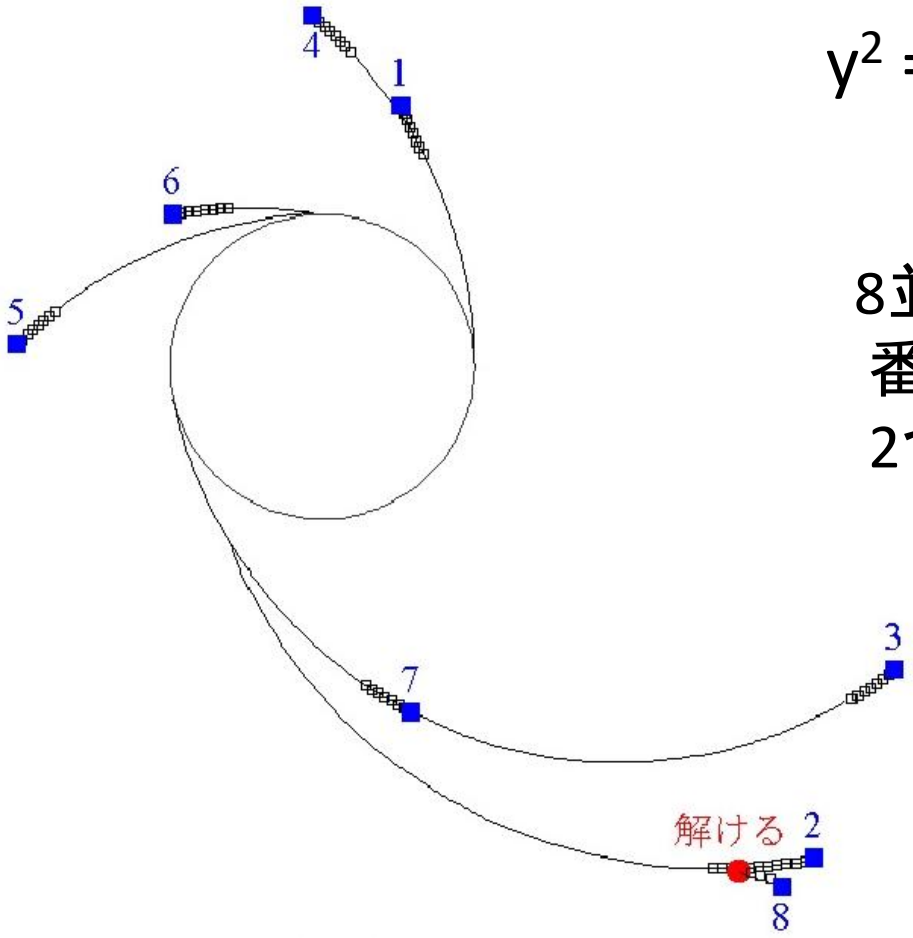
$$y^2 = x^3 + 9973x + 8108 \pmod{10007}$$

- ■ 初期値から出発し
- □ で進み円を一周する

と解ける



λ法による解読(軌跡)例



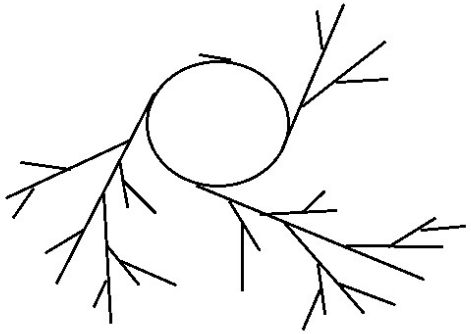
$$y^2 = x^3 + 9973x + 8108 \pmod{10007}$$

8並列λ法
番号は並列の初期値
2つが出会うと解ける

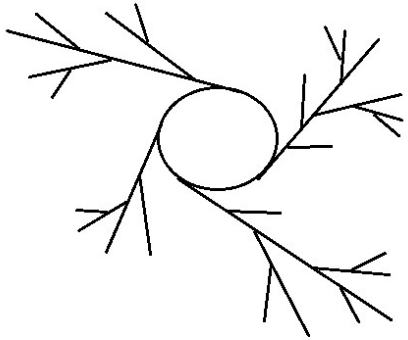
学習ρ法の発見

- ρ法の軌跡群は変動する(P,QはECC上の任意点)
  Q=n×Pからρ法で整数nを求める
  P及びQに依存し軌跡群は変動する
- 点Bで固定軌跡群となるρ法を発見(2018/7)
- R=s×Bをρ法で解き軌跡群を学習
- Q=n×Pのnは下記で計算
  P=k×B, Q=l×Bで整数k,lを学習ρ法で計算
  → n=l/k (mod r)

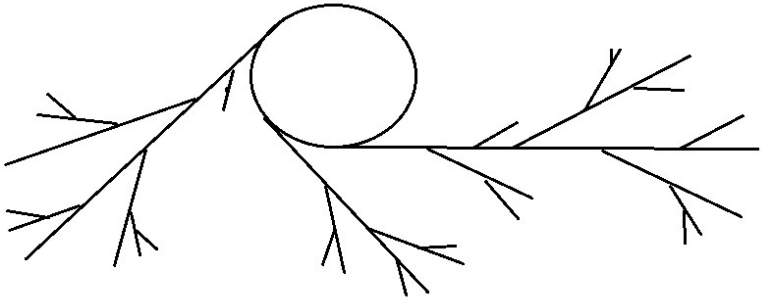
変動軌跡群の例



軌跡群-A

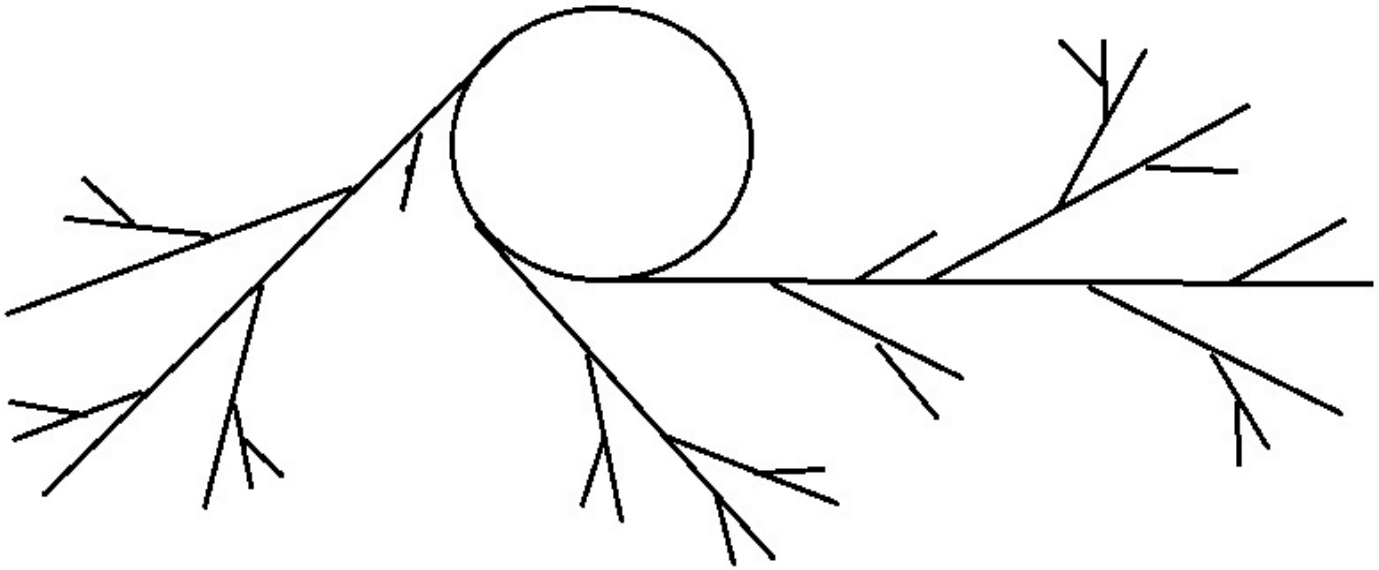


軌跡群-B



軌跡群-C

固定軌跡群の例



ρ法の軌跡は初期値で変るが、必ず上記軌跡群上をたどり、円周を一周する

固定軌跡群の学習ρ法

- ρ法はQ=n × Bに適用
  可変の点でなく、固定の点Bに変更
- ρ法の$S_k=f(S_{k-1})$, $S_{k-1}=(x,y)$を下記に変更
  x <= p/2の場合

$$S_k = S_{k-1} + B,\ \ \alpha_k = \alpha_{k-1}+1 \pmod r,\ \ \beta_k = \beta_{k-1}$$

  x > p/2の場合

$$S_k = 2 \times S_{k-1},\ \ \alpha_k = 2\alpha_{k-1},\ \beta_k = 2\beta_{k-1} \pmod r$$

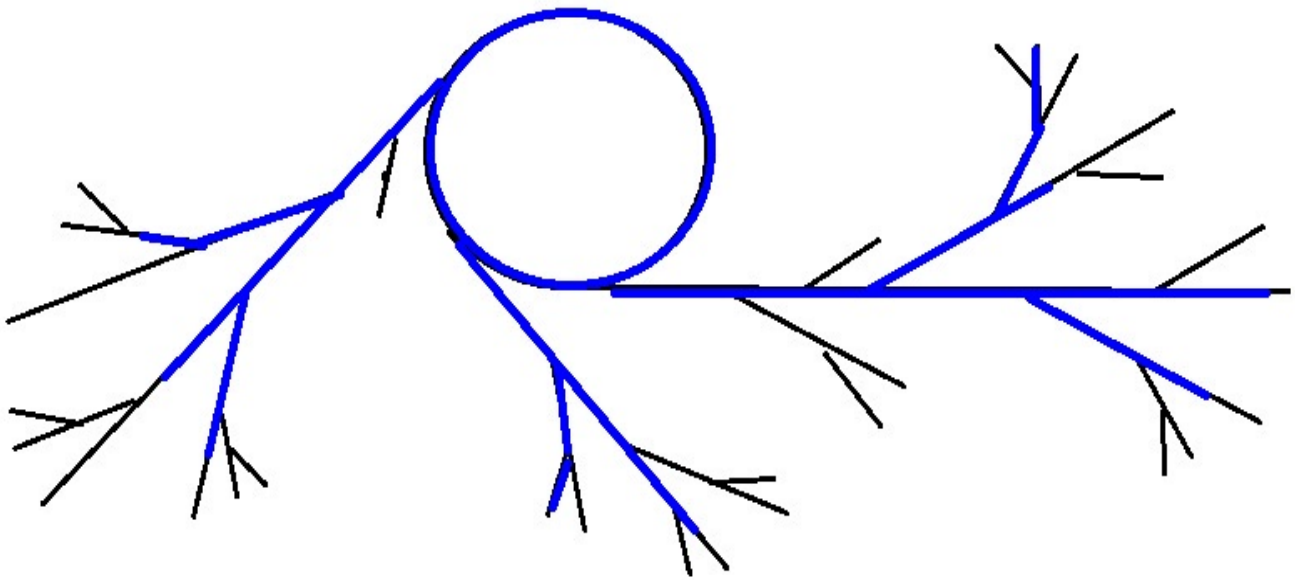
固定軌跡群となる理由(学習ρ法)

- 固定軌跡群となる条件はS_{k-1}が定まると、S_kは一意に定まること
- 学習ρ法はこの条件を満たす

  Bが固定なので、$S_{k-1}+B$も$2\times S_{k-1}$も一意
- 通常のρ法はこの条件を満たさない

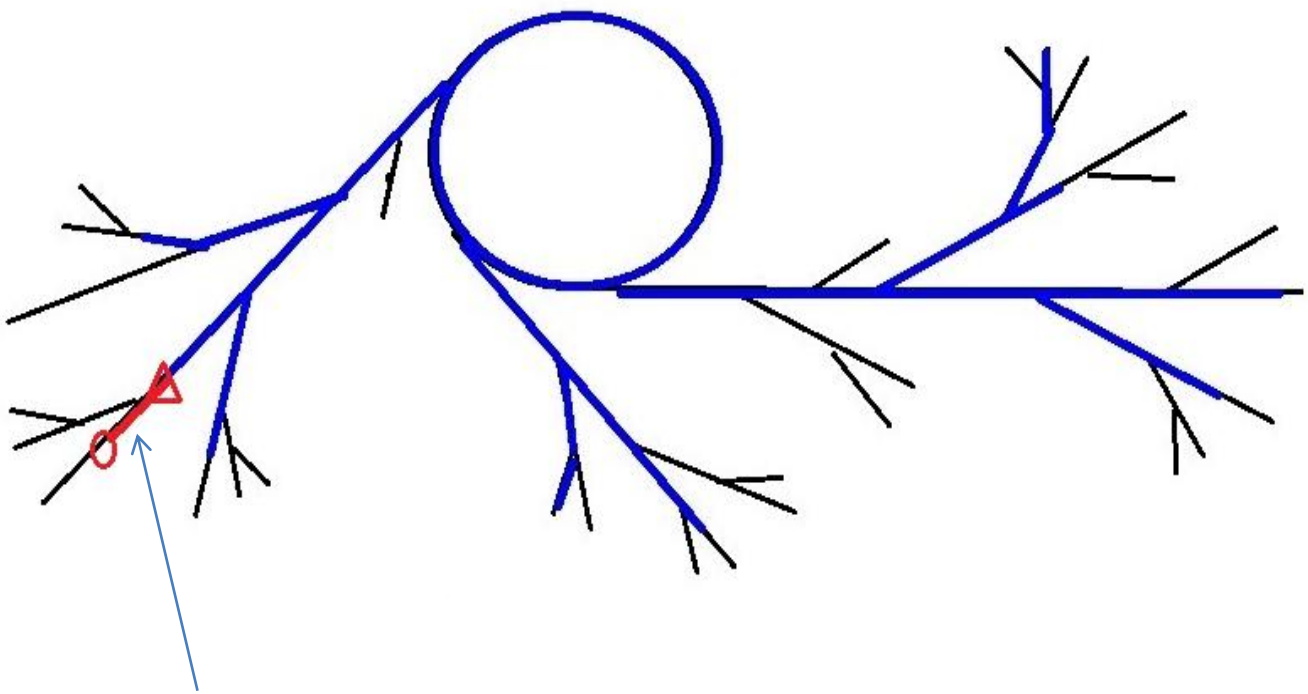
  P,Qは変わるので、$S_{k-1}+P$や$S_{k-1}+Q$は一意でなく、変動する

学習ρ法による学習



青太線は学習した軌跡
学習が進むと青太線が増加
黒細線は学習未完の軌跡群

学習ρ法による解読



青線の軌跡群は既知なので、赤丸から出発し赤線上を移動し、赤△に到達すれば解読

固定軌跡群の学習λ法

- λ法は$Q=n \times B$に適用
  Bは可変の点でなく、固定の点Bに変更
- λ法の$S_k=f(S_{k-1})$, $S_{k-1}=(x,y)$を下記に変更

(1) 準備計算

c_jを決め、$B_j=c_j \times B$,　$j=0,\ldots,15$

(2) 各ステップでの計算 (βは不変)

$i = x \pmod{16}$

$S_k = S_{k-1} + B_i$,　$\alpha_k = \alpha_{k-1} + c_i \pmod{r}$

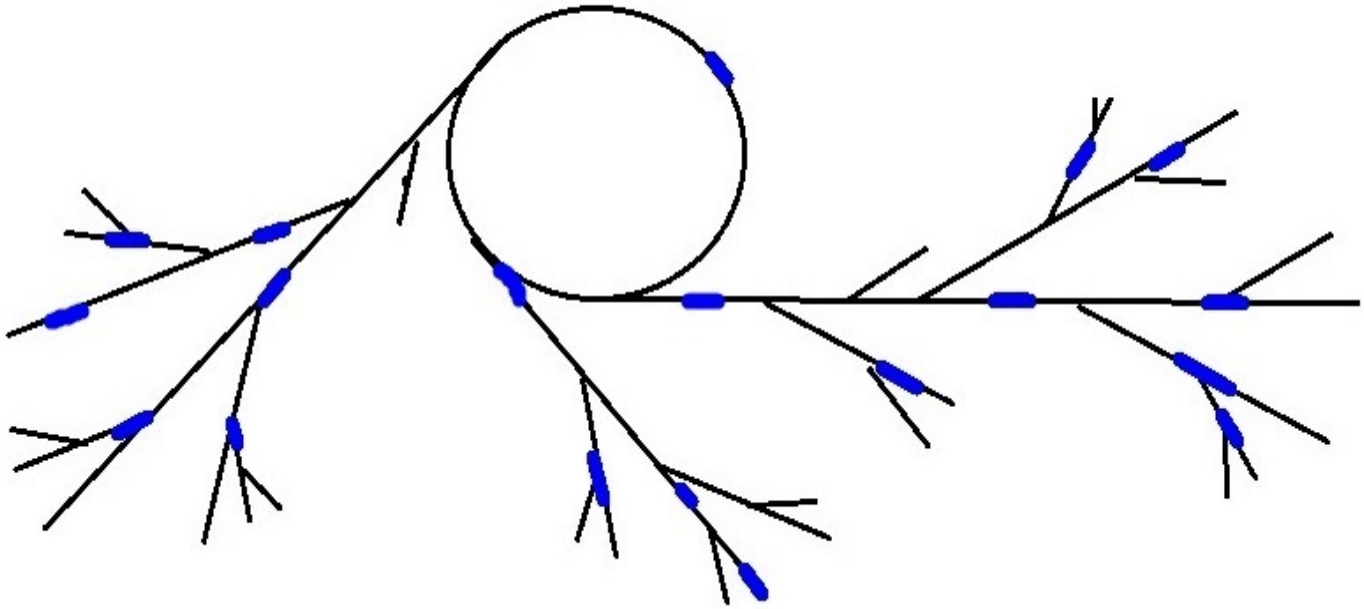
固定軌跡群となる理由(学習λ法)

- 固定軌跡群となる条件はS_{k-1}が定まると、S_kは一意に定まること
- 学習λ法ではこの条件を満たす

  $B(B_i)$が固定なので、$S_{k-1}+B_i$も一意
- 通常のλ法はこの条件を満たさない

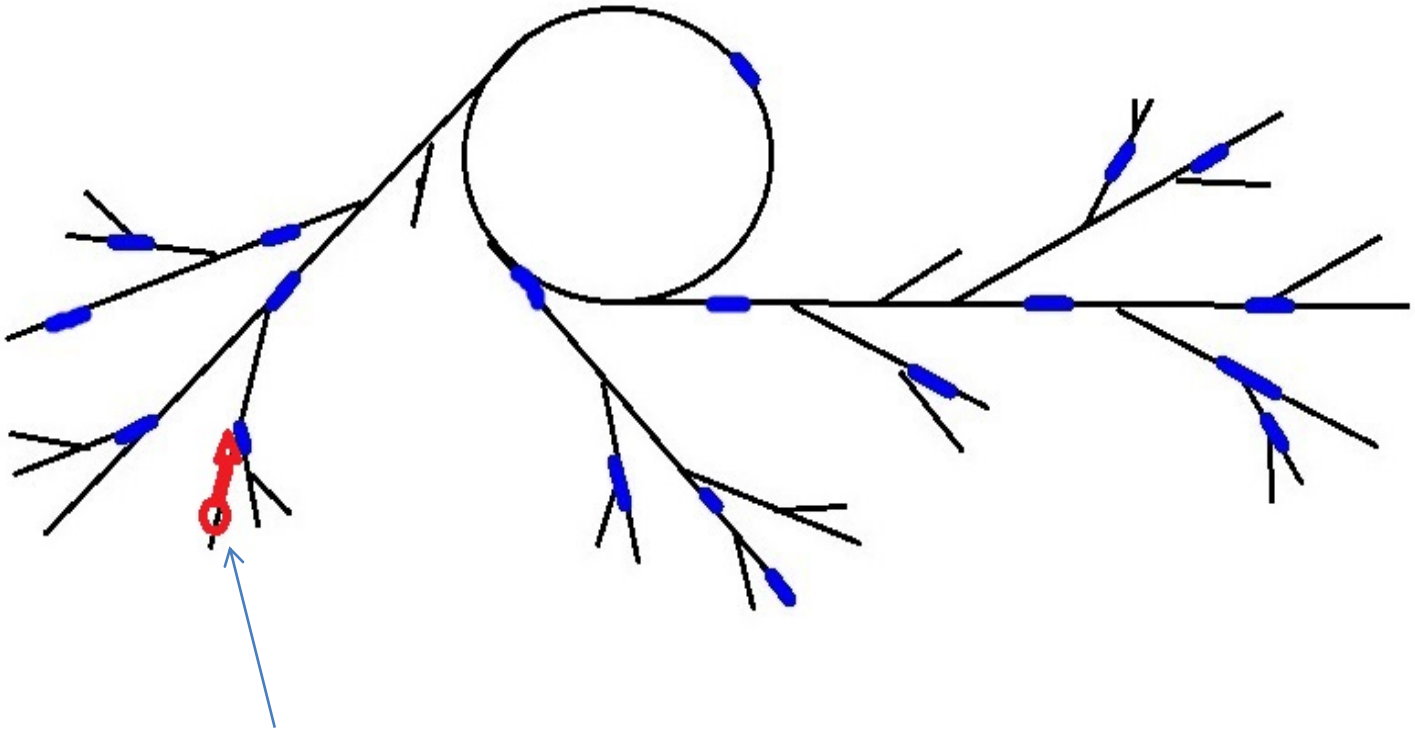
  P,Qは変わるので、$S_{k-1}+P_i$や$S_{k-1}+Q_i$は一意でなく、変動する

学習λ法による学習



青太線は学習した軌跡
学習が進むと青太線が増加
黒細線は学習未完の軌跡群

学習λ法による解読(1コア計算)



青線の軌跡群は既知なので、赤丸から出発し赤線上を移動し、赤△に到達すれば解読

学習λ法の軌跡群記憶

- 求めた軌跡群を間引きして記憶(特徴点)
  bビットの解読の平均軌跡長 := $2^{b/2}$
  点S_k=(x,y)でx<p/2^hなら記憶(1/2^h)を間引き
- 間引きによる高速化倍率限界(n倍)
  $n = 2^{b/2-h}$， b=60でn=8千倍ならh=17
- 高速化限界(間引き)による記憶量の増加
  間引き軌跡点(n)×学習量(n): $O(n^2)$

学習λ法の必要記憶量

高速化 (対ρ,λ法)	暗号ビット数 (以下)		
	64ビット	128ビット	256ビット
1万倍	2.5 GB	4.5 GB	8.5GB
4万倍	36 GB	70 GB	140 GB
10万倍	220 GB	430 GB	850 GB
100万倍	20 TB	40 TB	80 TB

軌跡群の記憶量は高速化の2乗に比例
検索情報の記憶も必要

SSD使用学習λ法

- 外部記憶装置のランダムアクセス性能

  HDD: 80～160回/s,　SSD: 50000回/s
- 学習λ法のアクセス数と時間(60ビット44件)

  約300回:　SSD: 300/50000 (s) = 0.006 (s)
- 学習λ法の解読時間(60ビット44件,4Ghz PC)

  4万倍高速(対ρ法) = 0.9 (s)

  → SSDなら解読の経過時間への影響無し

  → HDDなら影響(0.9秒が数秒に増加)

数値実験環境(解読)

パソコン: 解読は1コア使用

HP Desktop 870 (Intel Core i7 6700k)

4Ghz, 8GB, SSD:512G(SD8SBU-512G)

システム: Windows10 (64ビット)

コンパイラー: (Cで作成)

Cygwin Ver. 2.7 gcc (64ビット), 最適化(-O3)

多倍長関数: gmp のmpz_t 及びCで自作

数値実験環境(学習)

並列計算機(GPU): 学習にだけ使用

NVIDIA GeForce GTX 1080

2560コア, 8GB

システム: Windows10 (64ビット)

コンパイラー: (Cudaで作成)

NVIDIA Cuda Ver.8.0, 最適化(-O3)

多倍長関数: Cudaで自作

楕円曲線暗号解読

- 60ビット楕円曲線暗号
  $y^2 = x^3 + ax + b \pmod p$,　位数: r,　p,rは素数
  p=b98cc7dffc8ebbf,　r=b98cc7e0020c6af
  a=3dd997f5542f826,　b=7bb32feaa873e61
- 評価データ(Q=n×Pより整数nを求める)
  ランダムな22個のPと44個のQで44件
  44件は70ビット暗号相当。上記を5組使用。

ρ法と学習λ法の解読方法

- ρ法解読

  Q=n×Pから直接ρ法でnを求める
- 学習λ法解読（シリアル計算、点Bは固定）

  Q=n×Pのnは下記で計算

  P=k×B, Q=l×Bで整数k,lを学習λ法で計算

  → n=l/k (mod r)

  学習で蓄積した軌跡群データを利用

ρ法による解読結果

- データ5組の計算時間(CPUの1コア)

データA: 38295秒　　データB: 35813秒

データC: 36728秒　　データD: 34661秒

データE: 38183秒

平均: 36736秒 >= 10時間

学習λ法による解読結果

- データ5組の計算時間(CPUの1コア)

  データA: 3.3秒　　データB: 3.3秒

  データC: 3.0秒　　データD: 3.1秒

  データE: 3.3秒　　平均: 3.2秒
- 学習λ法の効果(60ビット44件)

  ρ法/学習λ法 = 10時間/3.2秒 = 11,000倍

  学習量: 2GB,　学習時間: 1.0日(GTX 1080)

SSD使用の学習λ法による解読結果

- データ5組の計算時間(CPUの1コア)

  データA: 0.94秒　　データB: 0.81秒

  データC: 0.92秒　　データD: 0.92秒

  データE: 0.86秒　　平均: 0.9秒
- 学習λ法の効果(60ビット44件)

  ρ法/学習λ法 = 10時間/0.9秒 = 40,000倍

  学習量: 36GB,　学習時間: 4日(GTX 1080)

予稿のあと(AI学習λ法)

- 学習λ法に人口知能(AI)などの高速化を追加
- AI学習λ法は、学習も解読も共に高速化する
- 学習λ法は、ρ法やλ法よりAI化し易い
  任意点Bで固定し、軌跡群が固定されるため
- AI学習λ法
  (1) 初期値選定やステップ選択の最適化
  (2) 逆元計算の改良等でのλ法自体の高速化

AI学習λ法の目標

- AI学習λ法高速化の条件

(1) 学習量は100GB以内

(2) ECC-79ビット相当(1050件のECC-60)を4Ghz PCの1コアで解読(ρ法で10日:864,000秒)

- AI学習λ法の高速化目標（個人研究の場合）

半年後(2019年末) ： 20万倍 (4.2秒)

1.5年後(2020年末) ： 50万倍 (1.68秒)

2.5年後(2021年末) ： 100万倍 (0.84秒)

AI学習λ法の現状

- PC(4Ghz)で1050件のECC-60の解読時間
  ECC-79(1997年に世界初解読)相当
  ρ法(gmp多倍長)で10日(86万4千秒)
- 改良経過(2019年末目標: 4.2秒:20万倍)

	解読時間	高速化
(1) 2019/5/20:	7.59秒	11万倍
(2) 2019/5/25:	5.76秒	15万倍

おわりに

- 単純学習で楕円曲線暗号の解読高速化に成功
- 60ビット44件の解読で10時間を0.9秒に短縮
- 学習λ法でρ法の4万倍まで高速化した
- 軌跡群の固定で、AIの活用が期待できる
- AI学習λ法でρ法の15万倍まで高速化した
- AIで大幅性能向上は個人では難しい

米、中は人、金に糸目つけずAI化を進めるはず

→ AI学習λ法の研究は国の主導が望ましい