Stratégies de cofactorisation pour l'algorithme Function Field Sieve

Benoît Gaudel Encadrant : Emmanuel Thomé

Equipe CARAMEL
LORIA

Plan de l'exposé

1. Sujet du stage
2. Déroulement du stage
   - Objectifs
   - Outils utilisés
   - Modélisations
   - Stratégies et observations
3. Conclusion

Logarithme discret

Dans un corps fini F :

- Il exite α tel que $(F^*, \times) =< \alpha >$
- Logarithme en base α de $\beta \in F^*$: m tel que $\beta = \alpha^m$

Calcul d'index

Calcul du logarithme dans $\mathbb{F}_{p^n}$

- Calcul du logarithme de "petits éléments" b_i
  - Détermination de relations : $m_i = \sum_j e_{ij} log(b_j)$
  - Résolution du système linéaire obtenu
- Calcul du logarithme voulu
  - Si m est tel que $\beta\alpha^m = \prod_i b_i^{\alpha_i}$
  - $\beta\alpha^m$ est dit "friable"
  - $log_\alpha(\beta) = (\sum_i \alpha_i log(b_i)) - m$

Calcul d'index

Calcul du logarithme dans $\mathbb{F}_{p^n}$

- Calcul du logarithme de "petits éléments" b_i
  - Détermination de relations : $m_i = \sum_j e_{ij} log(b_j)$
  - Résolution du système linéaire obtenu
- Calcul du logarithme voulu
  - Si m est tel que $\beta\alpha^m = \prod_i b_i^{\alpha_i}$
  - $\beta\alpha^m$ est dit "friable"
  - $log_\alpha(\beta) = (\sum_i \alpha_i log(b_i)) - m$

Calcul d'index

Calcul du logarithme dans $\mathbb{F}_{p^n}$

- Calcul du logarithme de "petits éléments" b_i
  - Détermination de relations : $m_i = \sum_j e_{ij} log(b_j)$
  - Résolution du système linéaire obtenu
- Calcul du logarithme voulu
  - Si m est tel que $\beta\alpha^m = \prod_i b_i^{\alpha_i}$
  - $\beta\alpha^m$ est dit "friable"
  - $log_\alpha(\beta) = (\sum_i \alpha_i log(b_i)) - m$

Calcul d'index

Calcul du logarithme dans $\mathbb{F}_{p^n}$

- Calcul du logarithme de "petits éléments" b_i
  - Détermination de relations : $m_i = \sum_j e_{ij} log(b_j)$
  - Résolution du système linéaire obtenu
- Calcul du logarithme voulu
  - Si m est tel que $\beta\alpha^m = \prod_i b_i^{\alpha_i}$
  - $\beta\alpha^m$ est dit "friable"
  - $log_\alpha(\beta) = (\sum_i \alpha_i log(b_i)) - m$

Calcul d'index

Calcul du logarithme dans $\mathbb{F}_{p^n}$

- Calcul du logarithme de "petits éléments" b_i
  - Détermination de relations : $m_i = \sum_j e_{ij} log(b_j)$
  - Résolution du système linéaire obtenu
- Calcul du logarithme voulu
  - Si m est tel que $\beta\alpha^m = \prod_i b_i^{\alpha_i}$
  - $\beta\alpha^m$ est dit "friable"
  - $log_\alpha(\beta) = (\sum_i \alpha_i log(b_i)) - m$

Calcul d'index

Calcul du logarithme dans $\mathbb{F}_{p^n}$

- Calcul du logarithme de "petits éléments" b_i
  - Détermination de relations : $m_i = \sum_j e_{ij} log(b_j)$
  - Résolution du système linéaire obtenu
- Calcul du logarithme voulu
  - Si m est tel que $\beta\alpha^m = \prod_i b_i^{\alpha_i}$
  - $\beta\alpha^m$ est dit "friable"
  - $log_\alpha(\beta) = (\sum_i \alpha_i log(b_i)) - m$

Début du crible du corps de fonctions

En caractéristique 2

1. Choix de f tel que $\mathbb{F}_{2^n} = \mathbb{F}_2[X]/\langle f \rangle$
2. Création de couples :

$$\phi = A(X)Y + B(X) \in \mathbb{F}_2[X][Y]$$

$$\phi_{ra} \in \mathbb{F}_2[X] \qquad \phi_{alg} \in \mathbb{F}_2[X]$$

$$\phi_{ra} \equiv \phi_{alg} \ mod \ f$$

3. Crible sur les couples
4. Cofactorisation des couples ayant passé le crible

cofactorisation

But de la cofactorisation :

- Sélectionner les couples doublement friables (aux "large primes" près)
- Factoriser ces couples

Plan de l'exposé

1 Sujet du stage

2 Déroulement du stage
- Objectifs
- Outils utilisés
- Modélisations
- Stratégies et observations

3 Conclusion

Objectifs

- Modéliser et étudier l'étape de cofactorisation de FFS
- Elaborer des stratégies pour cette étape

Plan de l'exposé

1 Sujet du stage

2 Déroulement du stage
- Objectifs
- Outils utilisés
- Modélisations
- Stratégies et observations

3 Conclusion

Outils mathématiques

Test de friabilité

- Proposition :

  Dans $\mathbb{F}_q[X]$, $X^{q^i} - X$ est le produit de tous les polynômes irréductibles de degré divisant i.
- Test : Si $\prod_{i=1}^{b}(X^{q^i} - X) = 0 \ mod \ P$, alors P est b-friable.

Outils mathématiques

Factorisation : Cantor-Zassenhaus

- Étape 1 : squarefree factorization
- Étape 2 : distinct degree factorization
- Étape 3 : equal degree factorization

Outils mathématiques

Factorisation : Berlekamp (en caractéristique 2)

- Étape 1 : squarefree factorization
- Étape 2 : Détermination des facteurs
  - σ : élévation au carré modulo P (Frobenius)
  - $\mathbb{F}_2[X]/ < P >= \prod_i \mathbb{F}_2[X]/ < g_i >$
  - Si $\alpha \in \mathbb{F}_2[X]/ < P >$ vérifie $\alpha^2 = \alpha$ alors $\alpha = 0$ ou 1 $mod\ g_i$
  - $1 \in Ker(\sigma - Id)$
  - Pour $h \in Ker(\sigma - Id)$ et $h \neq 0$ ou 1 on a $h(h+1) = 0\ mod\ P$ donc $pgcd(P, h)$ et $pgcd(P, h+1)$ sont deux facteurs différents de P

Outils mathématiques

Factorisation : Niederreiter (cas de $P \in \mathbb{F}_2[X]$)

- Principe :

  Utiliser les solutions de l'équation en h : $(Ph)' = h^2$
  Si $P = \prod g_i^{\alpha_i}$ alors $h_i = \frac{P}{g_i} g_i'$ est solution $\rightarrow g_i = \frac{P}{pgcd(P,h_i)}$
- Détermination des solutions :
  - $P = A^2 + XB^2$
  - N :
  - Si $i = 2$
  - Si $i = 2$
  - $N(h) = h^2$ se ramène à un système linéaire
- Remarque : P n'a pas besoin d'être sans facteurs carrés

Outils mathématiques

Calcul du quotient et du reste à l'aide d'un précalcul

- $F = PQ + R$, Q?, R?
- Précalcul : $X^k = PQ_k + R_k$ avec $k \geq deg(F)$
- $FQ_k = PQ_kQ + RQ_k = (X^k + R_k)Q + RQ_k$

  $\rightarrow Q = FQ_k \gg k$

  $\rightarrow R = F + PQ$

Outils informatiques

Ressources et outils de travail

- cado-nfs
- bibliothèque NTL de C++
- magma

Plan de l'exposé

1 Sujet du stage

2 Déroulement du stage
- Objectifs
- Outils utilisés
- Modélisations
- Stratégies et observations

3 Conclusion

Modélisation de l'entrée de la cofactorisation

À l'aide de cado-nfs...

Étape similaire dans NFS (Number Field Sieve), déjà implémentée (cado-nfs) :

- trial division et nouveau crible
- test de friabilité pour les large primes
- factorisation du produit de large primes
- stratégie de cofactorisation de NFS $\rightarrow$ modifier l'ordre de ces étapes dans cado-nfs pour obtenir tous les couples

Modélisation de l'entrée de la cofactorisation

À l'aide de cado-nfs...

Étape similaire dans NFS (Number Field Sieve), déjà implémentée (cado-nfs) :

- trial division et nouveau crible
- test de friabilité pour les large primes
- factorisation du produit de large primes
- stratégie de cofactorisation de NFS $\rightarrow$ modifier l'ordre de ces étapes dans cado-nfs pour obtenir tous les couples

Modélisation de l'entrée de la cofactorisation

...et de magma

Dans l'ordre chronologique :

- Méthode 1 :
  - récupérer l'entrée de "factor_leftover_norm"
  - factoriser à l'aide de magma
  - transformer les facteurs premiers en polynômes irréductibles
  - remultiplier les facteurs irréductibles avec multiplicités
- Méthode 2 :
  - récupérer les couples avant "trial_div"
  - les transformer en polynômes (comme en 1)
  - stocker les polynômes sous forme de nombres
- Méthode 3 :
  - récupérer les couples après "trial_div"

Modélisation de l'entrée de la cofactorisation

...et de magma

Dans l'ordre chronologique :

- Méthode 1 :
  - récupérer l'entrée de "factor_leftover_norm"
  - factoriser à l'aide de magma
  - transformer les facteurs premiers en polynômes irréductibles
  - remultiplier les facteurs irréductibles avec multiplicités
- Méthode 2 :
  - récupérer les couples avant "trial_div"
  - les transformer en polynômes (comme en 1)
  - stocker les polynômes sous forme de nombres
- Méthode 3 :
  - récupérer les couples après "trial_div"

Modélisation de l'entrée de la cofactorisation

...et de magma

Dans l'ordre chronologique :

- Méthode 1 :
  - récupérer l'entrée de "factor_leftover_norm"
  - factoriser à l'aide de magma
  - transformer les facteurs premiers en polynômes irréductibles
  - remultiplier les facteurs irréductibles avec multiplicités
- Méthode 2 :
  - récupérer les couples avant "trial_div"
  - les transformer en polynômes (comme en 1)
  - stocker les polynômes sous forme de nombres
- Méthode 3 :
  - récupérer les couples après "trial_div"

Modélisation de l'entrée de la cofactorisation

...et de magma

Dans l'ordre chronologique :

- Méthode 1 :
  - récupérer l'entrée de "factor_leftover_norm"
  - factoriser à l'aide de magma
  - transformer les facteurs premiers en polynômes irréductibles
  - remultiplier les facteurs irréductibles avec multiplicités
- Méthode 2 :
  - récupérer les couples avant "trial_div"
  - les transformer en polynômes (comme en 1)
  - stocker les polynômes sous forme de nombres
- Méthode 3 :
  - récupérer les couples après "trial_div"

Plan de l'exposé

1 Sujet du stage

2 Déroulement du stage
- Objectifs
- Outils utilisés
- Modélisations
- Stratégies et observations

3 Conclusion

Premières observations

Avec la méthode 2 de modélisation de l'entrée

Observations après une cofactorisation(basique) du type :

1. Test de friabilité des deux polynômes
2. Si tests positifs, factorisation des deux polynômes

Peu de relations, donc de factorisations
ex : pour un nombre de 130 chiffres en entrée de cado-nfs,
$\sim$ 40000 relations parmi les 1200000 premiers couples soit $1/30$
Temps des tests : de l'ordre de dix fois le temps de factorisation

Premières observations

Avec la méthode 2 de modélisation de l'entrée

Observations après une cofactorisation(basique) du type :

1. Test de friabilité des deux polynômes
2. Si tests positifs, factorisation des deux polynômes

Peu de relations, donc de factorisations
ex : pour un nombre de 130 chiffres en entrée de cado-nfs,
$\sim$ 40000 relations parmi les 1200000 premiers couples soit $1/30$
Temps des tests : de l'ordre de dix fois le temps de factorisation

Premières observations

Avec la méthode 2 de modélisation de l'entrée

Observations après une cofactorisation(basique) du type :

1. Test de friabilité des deux polynômes
2. Si tests positifs, factorisation des deux polynômes

Peu de relations, donc de factorisations
ex : pour un nombre de 130 chiffres en entrée de cado-nfs,
$\sim$ 40000 relations parmi les 1200000 premiers couples soit $1/30$
Temps des tests : de l'ordre de dix fois le temps de factorisation

Stratégies envisagées

Avec la méthode 2 de modélisation de l'entrée

Sur les tests de friabilité :

- Tester d'abord du côté où il y a le plus de tests négatifs
- Écarter sans tester les couples où un polynôme a un degré plus grand qu'une borne raisonnable
- Diminuer le degré des polynômes à tester par trial division

Sur les factorisations :

- Trial division (car méthode 2 de modélisation)
- Séparation des polynômes en deux facteurs grâce au test :
  - par dichotomie
  - par un choix unique en fonction des observations
- Choix de l'algorithme de factorisation en fonction des degrés et des observations
- Combiner les stratégies précédentes

Trial division

Chronologiquement :

- Collecte des facteurs au fur et à mesure $\rightarrow$ pas de grosse trial division rentable
- Préstockage des facteurs de degré inférieur à 5
  - Quotients et restes obtenus par la méthode du précalcul
  - Usage de polynômes creux pour le reste et du précalcul pour le quotient

Trial division

Chronologiquement :

- Collecte des facteurs au fur et à mesure → pas de grosse trial division rentable
- Préstockage des facteurs de degré inférieur à 5
  - Quotients et restes obtenus par la méthode du précalcul
  - Usage de polynômes creux pour le reste et du précalcul pour le quotient

Trial division

Chronologiquement :

- Collecte des facteurs au fur et à mesure $\rightarrow$ pas de grosse trial division rentable
- Préstockage des facteurs de degré inférieur à 5
  - Quotients et restes obtenus par la méthode du précalcul
  - Usage de polynômes creux pour le reste et du précalcul pour le quotient

Trial division

Chronologiquement :

- Collecte des facteurs au fur et à mesure $\rightarrow$ pas de grosse trial division rentable
- Préstockage des facteurs de degré inférieur à 5
  - Quotients et restes obtenus par la méthode du précalcul
  - Usage de polynômes creux pour le reste et du précalcul pour le quotient

Trial division

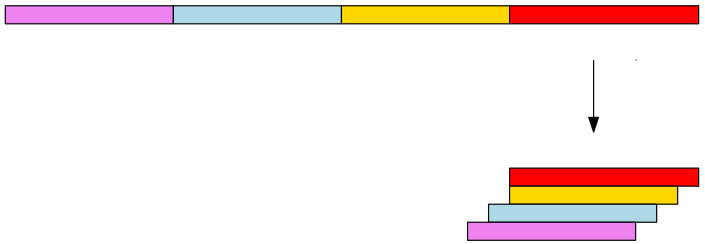
Chronologiquement :

- Collecte des facteurs au fur et à mesure $\rightarrow$ pas de grosse trial division rentable
- Préstockage des facteurs de degré inférieur à 5
  - Quotients et restes obtenus par la méthode du précalcul
  - Usage de polynômes creux pour le reste et du précalcul pour le quotient

Trial division

Chronologiquement :

- Collecte des facteurs au fur et à mesure $\rightarrow$ pas de grosse trial division rentable
- Préstockage des facteurs de degré inférieur à 5
  - Quotients et restes obtenus par la méthode du précalcul
  - Usage de polynômes creux pour le reste et du précalcul pour le quotient



Remarque sur le test de friabilité

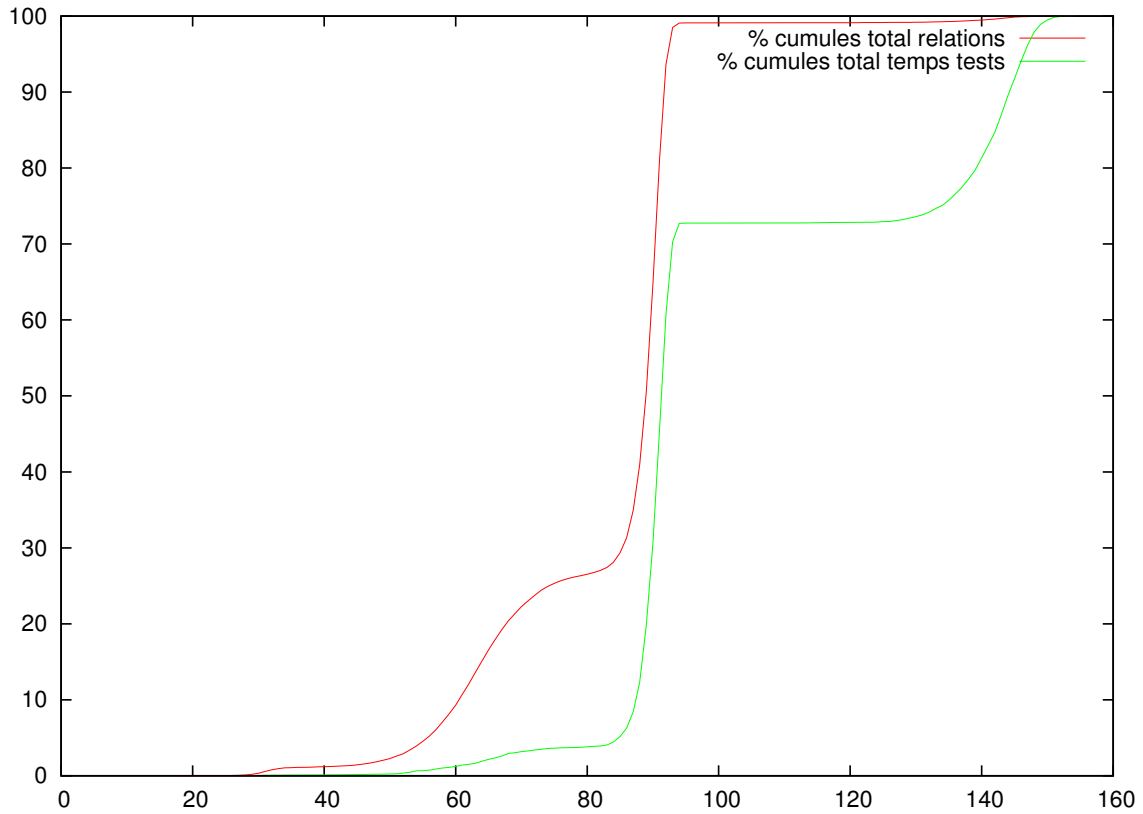
Test de friabilité et factorisation

- le degré de $pgcd(\prod_{i=1}^{k}(X^{2^i}+X), P)$ croît avec k
- on mémorise les $\prod_{i=1}^{k}(X^{2^i}+X)\ mod\ P$
- on sépare P en deux facteurs de degrés proche de $\frac{deg(P)}{2}$

Observations

Nombre à 130 chiffres

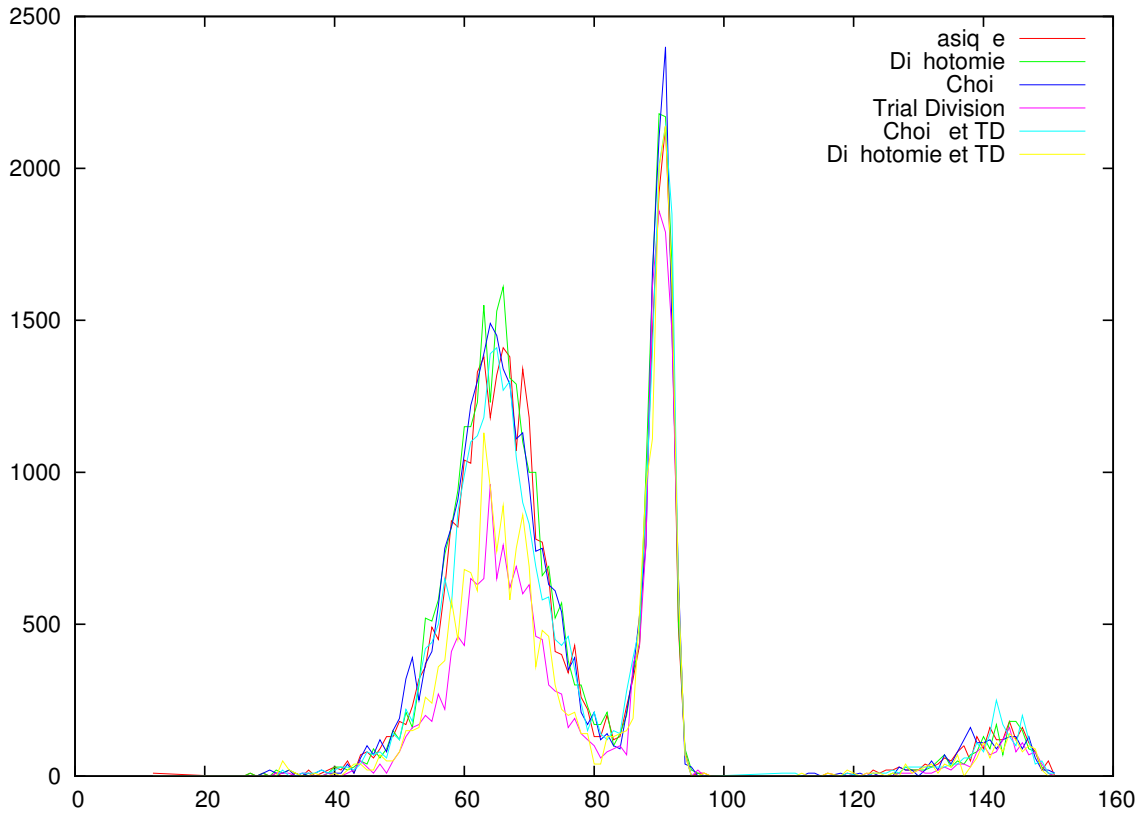
Répartition des relations et du temps de test



Observations

Nombre à 130 chiffres

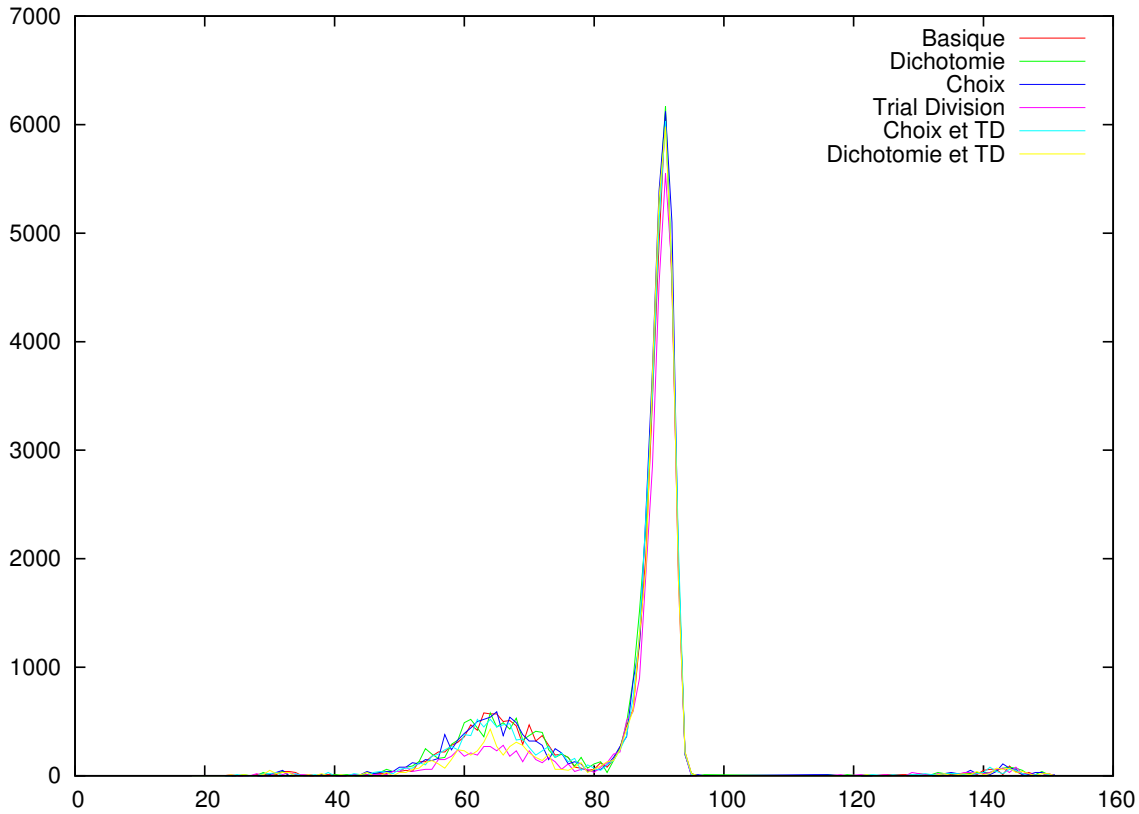
Factorisation Cantor-Zassenhaus (côté algébrique)



Observations

Nombre à 130 chiffres

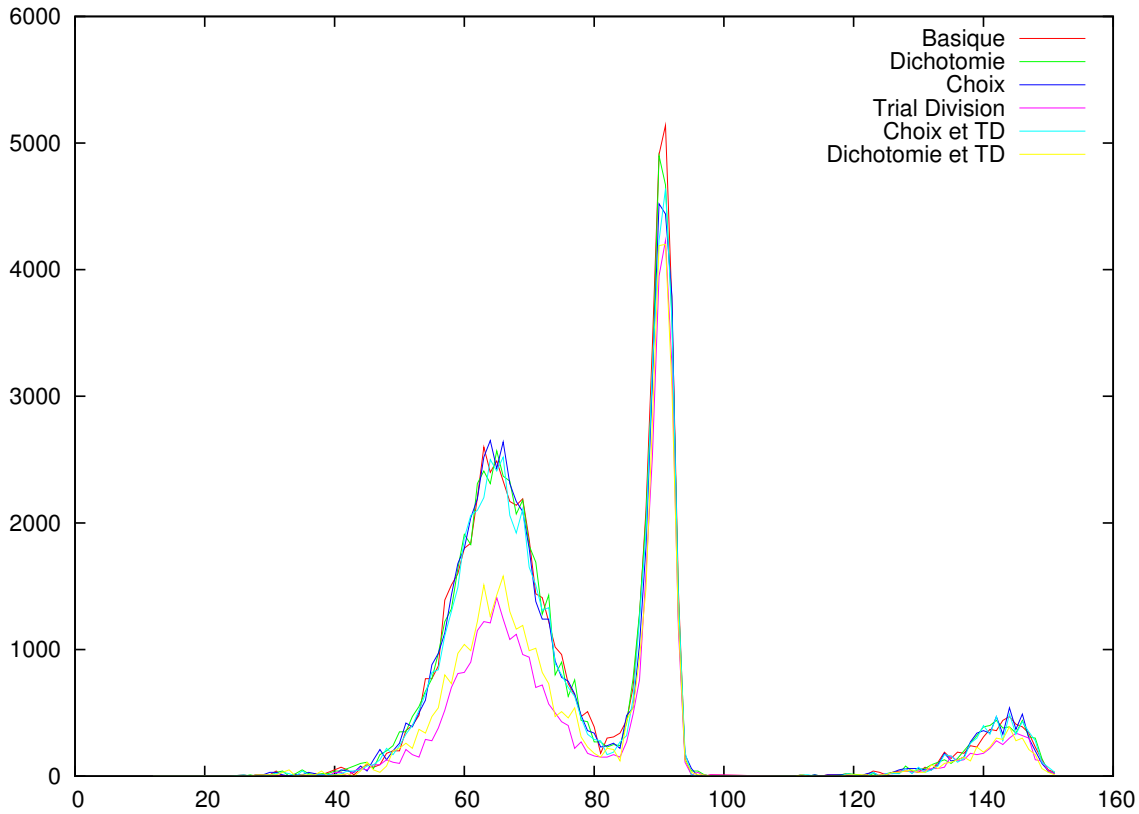
Factorisation Cantor-Zassenhaus (côté rationnel)



Observations

Nombre à 130 chiffres

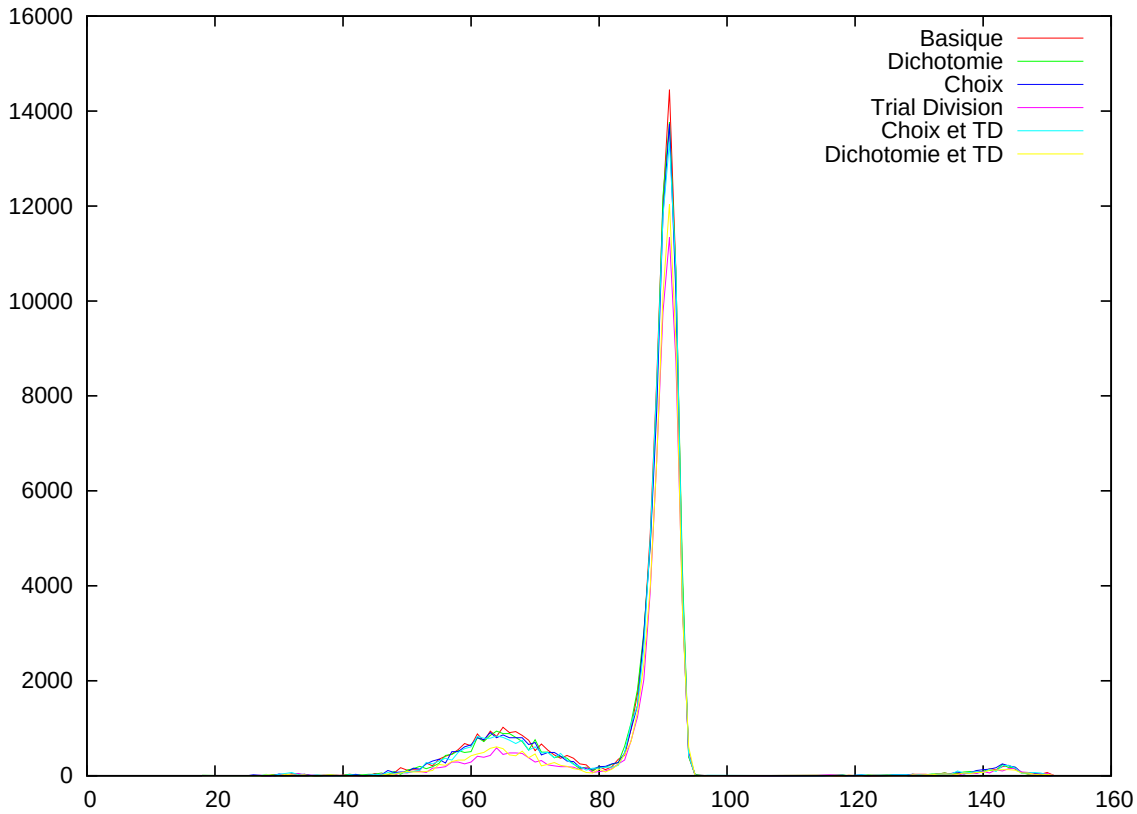
Factorisation Berlekamp (côté algébrique)



Observations

Nombre à 130 chiffres

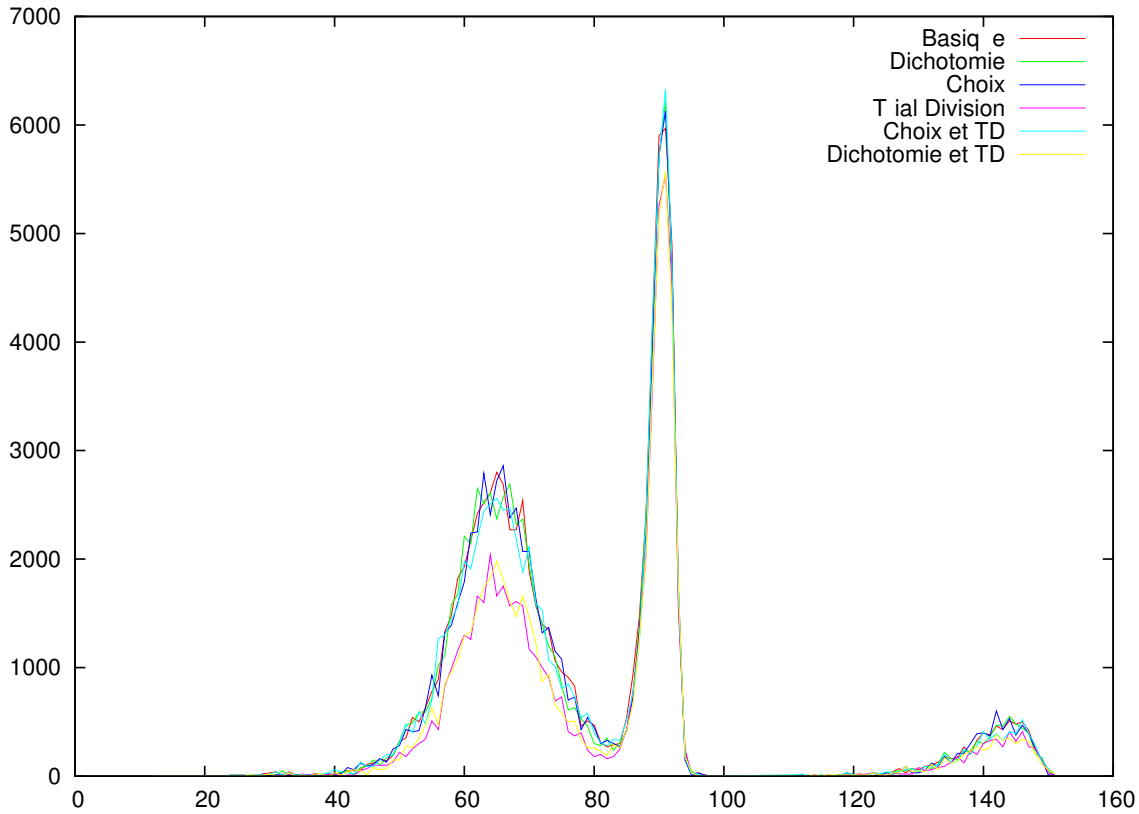
Factorisation Berlekamp (côté rationnel)



Observations

Nombre à 130 chiffres

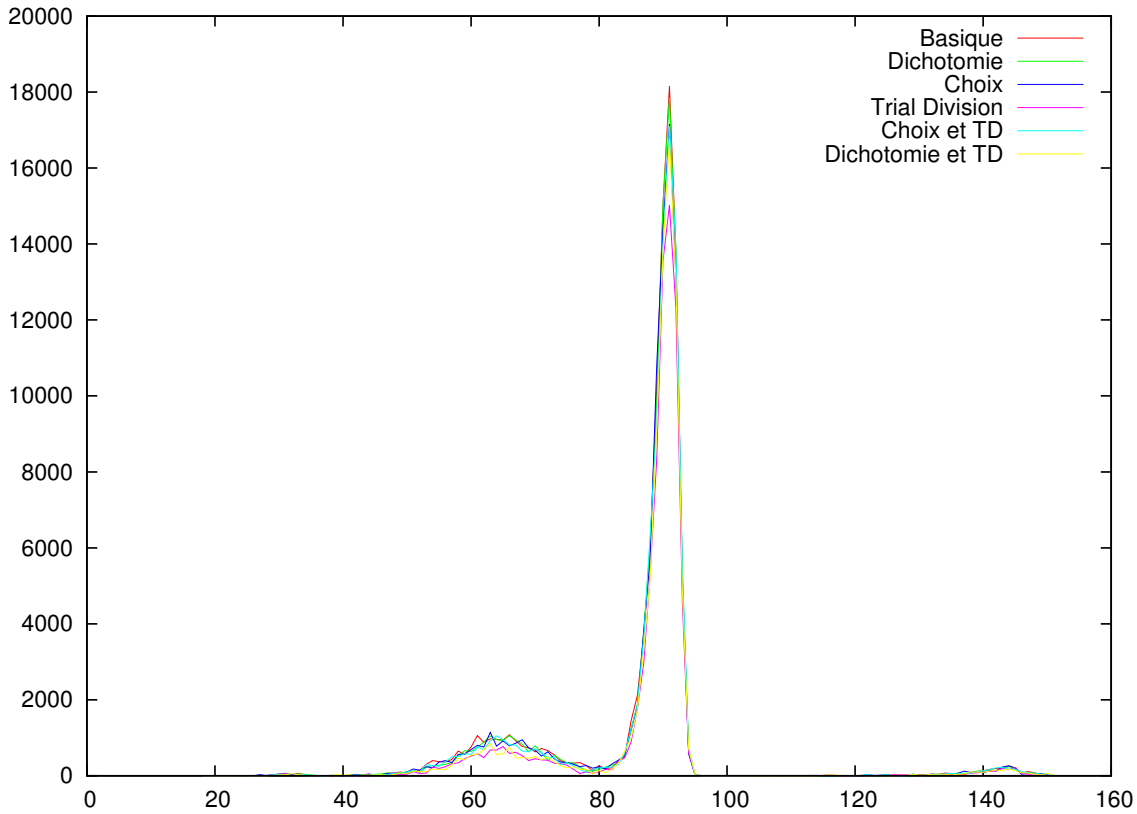
Factorisation Niederreiter (côté algébrique)



Observations

Nombre à 130 chiffres

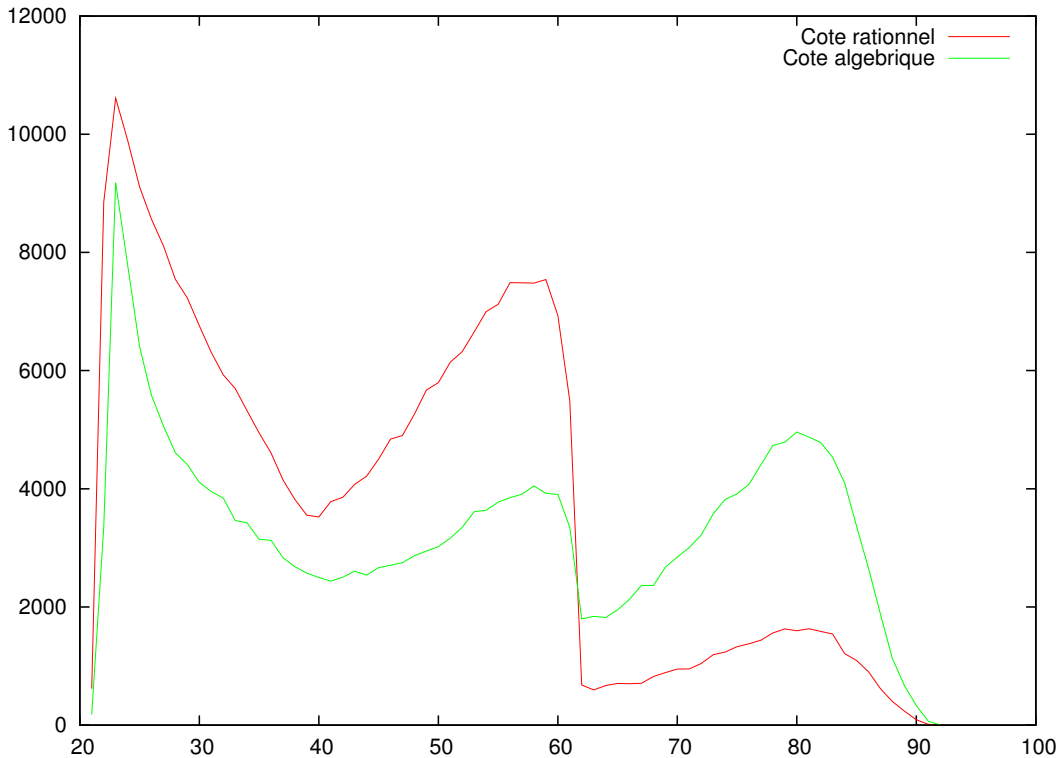
Factorisation Niederreiter (côté rationnel)



Premières observations

Avec la méthode 3 de modélisation de l'entrée

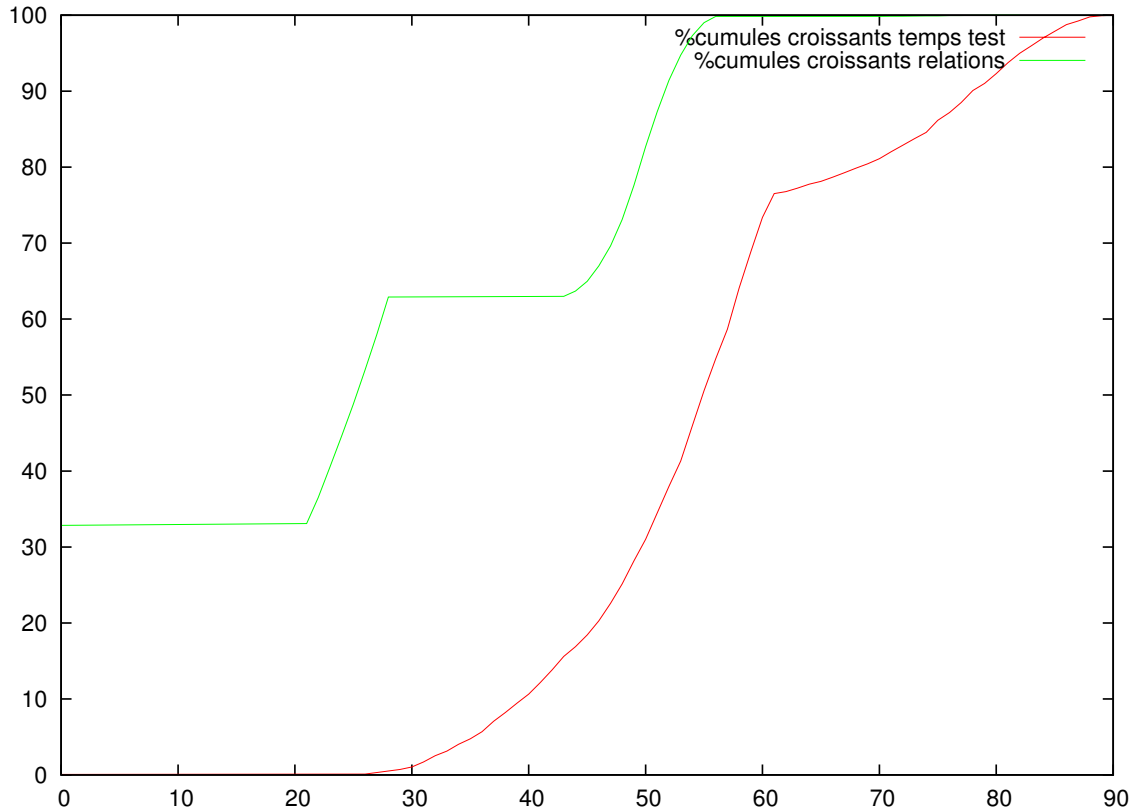
Degrés des facteurs des polynômes (friables ou non) obtenus pour un nombre de 130 chiffres



Premières observations

Avec la méthode 3 de modélisation de l'entrée

Répartition des relations et du temps de test



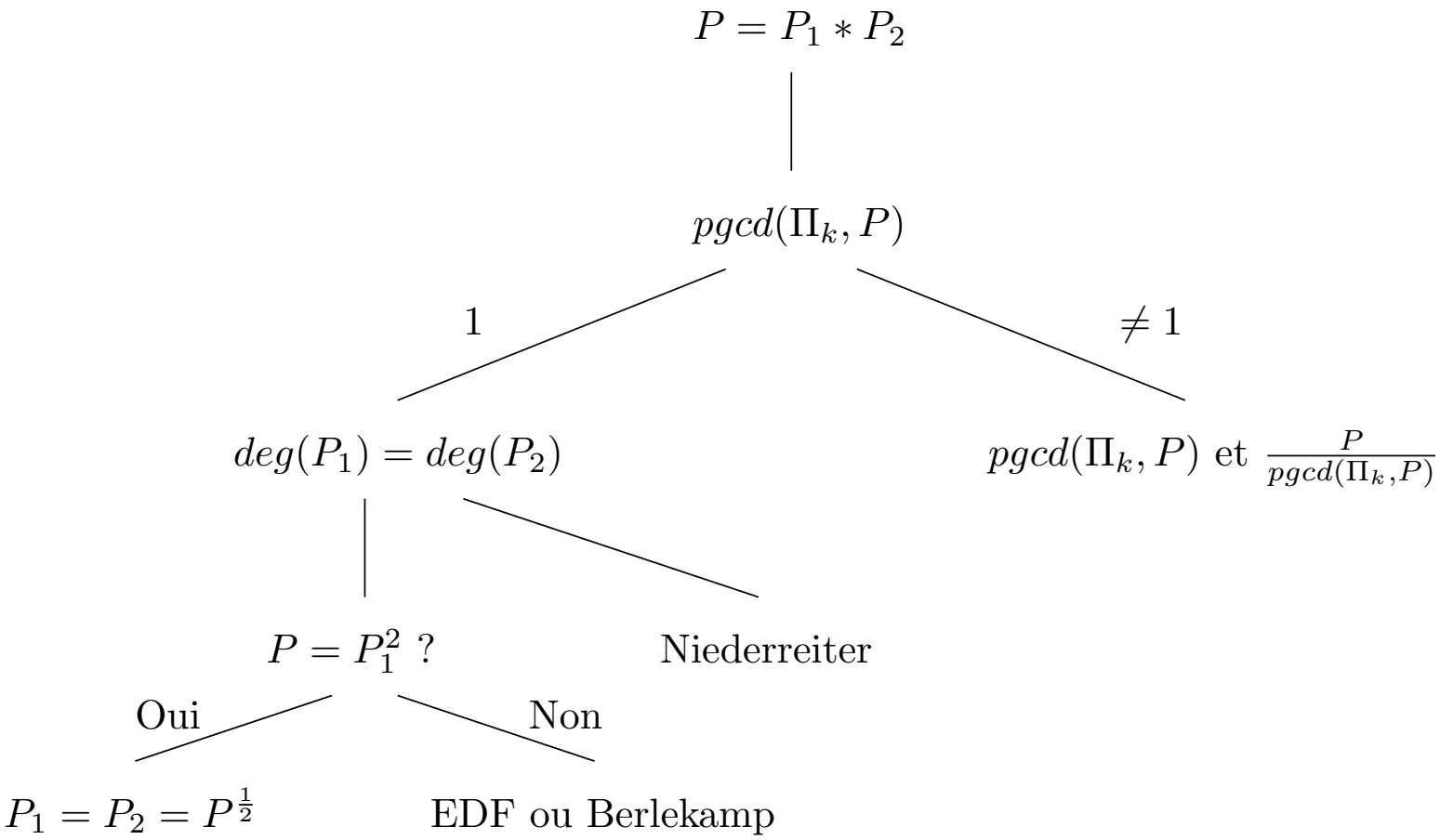
Stratégies envisagées

Avec la méthode 3 de modélisation de l'entrée

- Se limiter à deux "large primes"
- Stratégie sur les tests de friabilité :
  tester uniquement entre $2I$ et $2L$
- Stratégie sur la factorisation :
  utiliser le dernier $\Pi_k = \prod_{i=1}^{k}(X^{2^i} + X)\ mod\ P$ non nul

Stratégies envisagées

Avec la méthode 3 de modélisation de l'entrée



Résultats

Avec la méthode 3 de modélisation de l'entrée

Temps totaux des tests et des factorisations en millisecondes :

	Cantor-Zassenhaus	Berlekamp	Niederreiter
Temps tests	12510	11940	12010
Rat. Basique	721	889	1269
Rat. Méthode	78	139	196
Alg. Basique	271	351	503
Alg. Méthode	32	49	67

41241 relations pour 1200000 couples

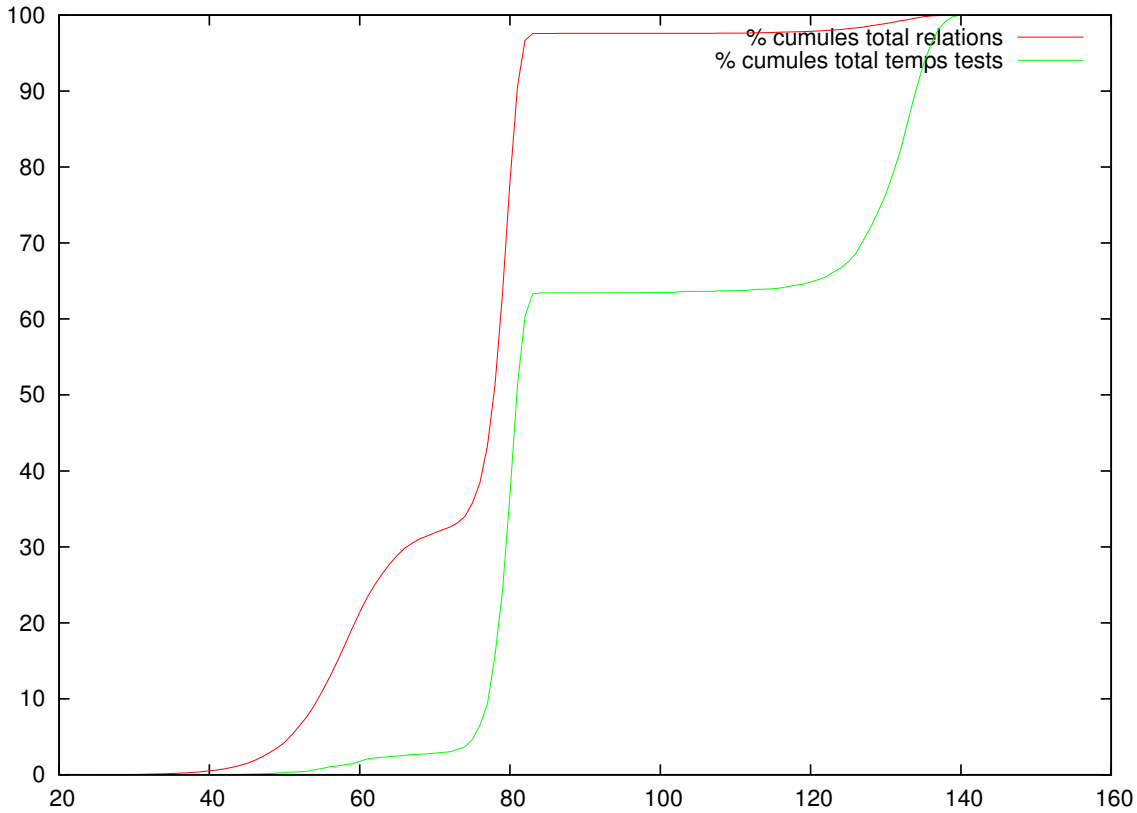
Conclusion

- Se concentrer sur la rapidité du test de friabilité
- Si "resieving" se limiter à deux "large primes"

Observations

Nombre à 111 chiffres

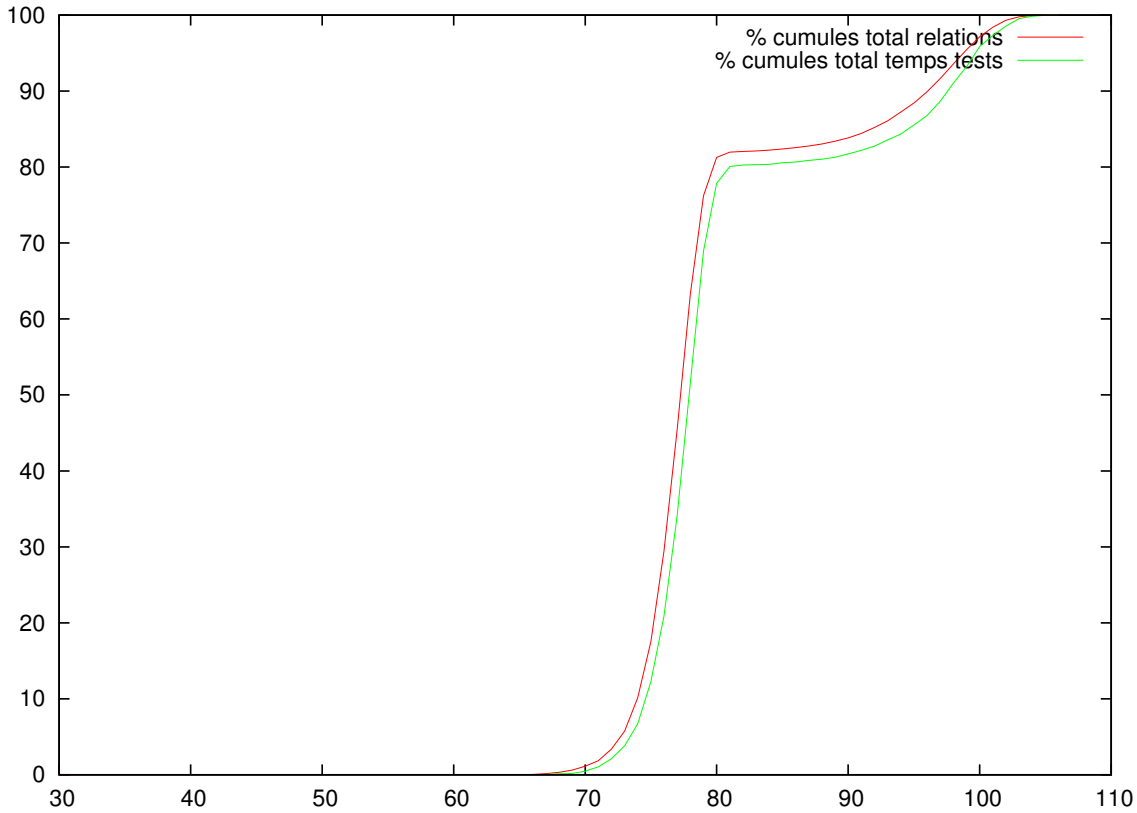
Test de friabilité :

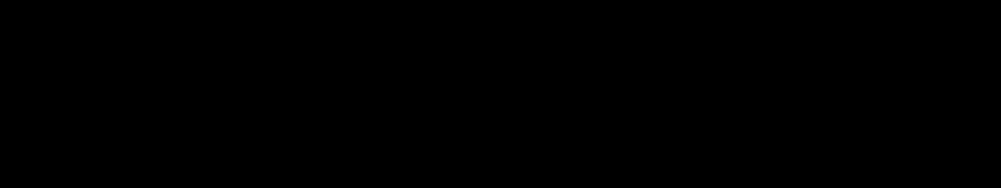


Observations

Nombre à 91 chiffres

Test de friabilité :





10

9

8

7

6

5

4

3

2

1

Repartition des polynômes selon le degré :