Primalité et factorisation des entiers

Soit $N > 1$ un entier. Nous considérons 3 problèmes « élémentaires » :

1. (Primalité) En supposant que N est premier, le prouver.
2. (Non primalité) En supposant que N est composé, le prouver.
3. (Factorisation) En supposant que N est composé, trouver un facteur non trivial $d \mid N$, $d \neq 1, N$.

Ils permettent de résoudre le problème global de la factorisation complète :

4. Factoriser N complètement : déterminer les diviseurs premiers p de N, et pour chacun d'entre eux afficher $v_p(N)$ et une preuve de primalité.

En pratique, (2) est très facile, (1) admet une solution raisonnable, et (3) est toujours largement ouvert.

Théorème (Rabin, 1980)

(2) est résolu en temps probabiliste $\widetilde{O}(\log N)^2$.

Théorème (Agrawal-Kayal-Saxena, 2002)

(1) et (2) sont résolus en temps déterministe $O(\log N)^{10.5}$.

Lenstra et Pomerance ont amélioré l'estimation (et l'algorithme) : $O_\varepsilon(\log N)^{6+\varepsilon}$ mais la constante du O n'est plus effective si $\varepsilon > 0$ est trop petit !

Théorème (Miller, 1976)

En supposant une Hypothèse de Riemann convenable, (1) et (2) sont résolus en temps $\widetilde{O}(\log N)^4$.

C'est un enoncé prouvé mais conditionnel !

Conjecture (Goldwasser-Killian 1986, Atkin-Morain 1993, Shallit ≈ 1990)

L'algorithme ECPP (Elliptic Curves Primality Proving) résout (1) en temps probabiliste $\widetilde{O}(\log N)^4$.

C'est bien une conjecture ! Par contre, si l'algorithme s'arrête, il produit une *preuve* (ou certificat) de primalité : un algorithme auxiliaire permettant de vérifier rapidement la primalité de N.

Définition

Soit $0 \leqslant \alpha \leqslant 1$, on définit

$$L_\alpha(N) = \exp\left((\log N)^\alpha (\log\log N)^{1-\alpha}\right).$$

N.B. On a $L_0(N) = \log N$ et $L_1(N) = N$.

Théorème (Dixon, 1981)

L'entier N est factorisé en temps probabiliste $L_{1/2}(N)^{O(1)}$.

Conjecture (Pollard, 1988)

Le crible algébrique (NFS) factorise N en temps probabiliste $L_{1/3}(N)^{O(1)}$.

Record actuel (2020) pour un entier type RSA : 250 chiffres

Théorème

Si $N > 2$ est premier et $0 < a < N$, alors on a dans $\mathbb{Z}/N\mathbb{Z}$:

1. $a^{N-1} = 1$;
2. *Écrivons $N - 1 = 2^e q$ où q impair et $e \geqslant 1$ et posons $b := a^q$. Alors $b = 1$ ou il existe un unique indice $0 \leqslant i < e$ tel que $b^{2^i} = -1$.*

Preuve.

(1) est le test de Fermat (1640). (2) (Rabin-Miller) suit de l'identité polynomiale

$$X^{2^e} - 1 = \prod_{d|2^e} \Phi_d(X) = (X-1)(X+1)(X^2+1)\dots(X^{2^{e-1}}+1)$$

évaluée en $X = b$. □

Test probabiliste : on tire $a \in]0, N[$ uniformément au hasard (et indépendamment des tests précédents) et on vérifie (1) ou (2). Si l'assertion n'est pas vérifiée pour le a choisi, alors N est composé. Un tel a est appelé témoin. N.B. Si a est témoin pour le test de Fermat alors il l'est aussi pour Rabin-Miller.

Complexité

$O(\log N)^3$ opérations binaires.

Théorème (Korselt, 1899)

Si $N > 1$ est composé, alors

$$\left\{0 < a < N : a^{N-1} \equiv 1 \pmod{N}\right\}$$

est un sous-groupe de $(\mathbb{Z}/N\mathbb{Z})^*$. Il est égal à $(\mathbb{Z}/N\mathbb{Z})^*$ tout entier si et seulement si N est sans facteurs carrés et $p \mid N \Rightarrow p - 1 \mid N$ pour tout diviseur premier p de N.

Par exemple 561, 1105, 1729, 2465... Il y en a une infinité. Si le sous-groupe est d'indice $\geqslant 2$, au moins 50% des a sont témoins pour le test de Fermat.

Théorème (Rabin, 1980)

Si $N > 1$ est composé, un $a \in]0, N[$ uniforme est témoin pour le test de Rabin-Miller avec probabilité $\geqslant 3/4$.

Théorème (Burthe, 1997)

La moyenne des plus petit témoins quand N parcourt les entiers composés impairs $\leqslant X$ est $2 + o(1)$.

Preuve de primalité

L'idée est la suivante : $N > 1$ est premier si et seulement si $(\mathbb{Z}/N\mathbb{Z})^*$ est cyclique d'ordre $N-1$. Il suffit donc d'exhiber un élément $a \in (\mathbb{Z}/N\mathbb{Z})^*$ d'ordre $N-1$. Il faut que $a^{N-1} = 1$ et $\mathrm{pgcd}(a^{(N-1)/p} - 1, N) \neq 1$ pour tout diviseur premier p de $N-1$.
En fait, le même test élémentaire permet de faire mieux :

Théorème (Pocklington, 1914)

Soit $N > 1$ un entier et $p \mid N-1$ un premier tel que $v_p(N-1) = e$. On suppose que l'entier a satisfait

- $a^{N-1} \equiv 1 \pmod{N}$,
- $\mathrm{pgcd}(a^{(N-1)/p} - 1, N) = 1$.

Alors tout diviseur $d \mid N$ vérifie $d \equiv 1 \pmod{p^e}$.

Preuve.

On peut supposer que d est premier, puisqu'un produit d'entiers $\equiv 1 \pmod{M}$ est congru à $1 \pmod{M}$. Puisque $d \mid N$, on a $a^{N-1} \equiv 1 \pmod{d}$, qui implique que $\mathrm{pgcd}(a, d) = 1$ et donc $a^{d-1} \equiv 1 \pmod{d}$ (Fermat).
Puisque $a^{(N-1)/p} \not\equiv 1 \pmod{d}$, l'ordre r de a dans $(\mathbb{Z}/d\mathbb{Z})^*$ vérifie $r \mid N-1$, mais $r \nmid (N-1)/p$, et donc $p^e \mid r$. On conclut puisque $r \mid d-1$. □

Corollaire

Soit $N > 1$ un entier. Supposons $N - 1 = FU$ avec $F \geqslant \sqrt{N}$, où les diviseurs premiers de F sont connus. On suppose que pour chaque $p \mid F$, un entier a_p comme ci-dessus est donné. Alors N est premier.

Preuve.

Si $d \mid N$, alors $d \equiv 1 \pmod{F}$ par le Lemme Chinois. Donc $d = 1$ ou $d \geqslant F + 1 > \sqrt{N}$. Cette dernière inégalité implique que $d = N$, puisque $N/d < \sqrt{N}$ doit être 1. □

Produire un grand nombre premier

Théorème (Théorème des Nombres Premiers (TNP))

Quand $x \to +\infty$, *on a*

$$\pi(x) := \#\{p \leqslant x : p \text{ premier}\} \sim \frac{x}{\log x}.$$

On a des bornes effectives :

$$\frac{x}{\log x}\left(1 + \frac{1}{2\log x}\right) \leqslant \pi(x) \leqslant \frac{x}{\log x}\left(1 + \frac{3}{2\log x}\right),$$

où l'inégalité de gauche est valide pour $x \geqslant 59$ *et celle de droite pour tout* $x > 1$.

Le nombre de premiers dans $]N, 2N]$ est

$$\pi(2N) - \pi(N) \sim \frac{N}{\log N} \ .$$

En tirant des entiers uniformément au hasard dans cet intervalle, l'espérance du nombre de tirage avant de produire un premier est donc $\log N$.

Produire beaucoup de petits premiers

Algorithme 1. Crible d'Ératosthène

Entrées: Un entier B.

Sorties: L'ensemble des premiers $p \leqslant B$.

1: Créer un tableau $A[2] = \cdots = A[B] = \texttt{vrai}$.
2: **pour** $n = 2, \ldots, \sqrt{B}$ **faire**
3: **si** $A[n] = \texttt{vrai}$ **alors** $\{n$ *est premier*$\}$
4: **pour** $k = 2, \ldots, B/n$ **faire** $\{$*barrer les multiples de* $n\}$
5: $A[kn] \leftarrow \texttt{faux}$, $\{kn$ *n'est pas premier*$\}$
6: Afficher les n tels que $A[n] = \texttt{vrai}$.

Le nombre d'opérations est

$$2B + \sum_{p \leqslant \sqrt{B}} \lfloor B/p \rfloor = 2B + O(\sqrt{B}) + B \sum_{p \leqslant \sqrt{B}} \frac{1}{p} \sim B \log\log B,$$

en utilisant $\sum_{p \leqslant x} p^{-1} \sim \log\log x$, qui suit par exemple du TNP et d'une sommation d'Abel.

Entiers friables

Définition

Soit $B > 0$. Un entier $N > 0$ est

- B-friable au sens restreint si (p premier divise N) implique $p \leqslant B$;
- B-friable si (p^k divise N) implique $p^k \leqslant B$ (pour tout p premier et tout entier k).

Théorème (de Bruijn)

Soit $\psi(x, y) = \#\{n \leqslant x : n$ *est* y*-friable*$\}$. *On choisit* $y = x^{1/u}$ *et on suppose que* $(\log x)^{\varepsilon} \leqslant u \leqslant (\log x)^{1-\varepsilon}$ *pour un* $0 < \varepsilon < 1$, *alors*

$$\frac{\psi(x, x^{1/u})}{x} = u^{-u+o(u)}$$

quand $x \to \infty$.

On se permettra le raisonnement non rigoureux consistant à supposer que la proportion d'entiers friables dans certains ensembles est proche des u^{-u} du théorème. Voici un exemple simple :

Algorithme 2. La méthode $p-1$ de Pollard

Entrées: N un entier, B une borne de friabilité.
Sorties: Un facteur non trivial de N ou `Échec`.

1: À l'aide du crible d'Ératosthène calculer les premiers $\ell \leqslant B$.
2: Tirer $a \in \mathbb{Z}/N\mathbb{Z}$ uniformément au hasard. Soit $b \leftarrow a$.
3: **pour** $\ell \leqslant B$ **faire** $\quad \{$*calculer* $b = a^{\mathrm{ppcm}(2,\ldots,B)}\}$
4: Soit k l'exposant maximal tel que $\ell^k \leqslant B$.
5: $b \leftarrow b^{\ell^k}$.
6: **si** $d := \gcd(b-1, N)$ est un diviseur non trivial de N **alors**
7: Afficher d.
8: **sinon**
9: `Échec`.

L'algorithme trouve un facteur quand l'ordre de a modulo un diviseur de N est B-friable, ssi l'ordre de a modulo un diviseur *premier* de N est B-friable. Ceci se produit en particulier si $p-1$ est B-friable et a peu de chance de se produire sinon :

Lemme

Soit ℓ est un diviseur premier de $p-1$, un $a \in \mathbb{F}_p^*$ a un ordre divisible par ℓ dans $\mathbb{F}_p^*$ avec probabilité $\geqslant 1-1/\ell$.

Preuve.

Dans un groupe cyclique $\langle g \rangle$ d'ordre n divisible (exactement) par ℓ^e, l'ordre de g^k est $n/\operatorname{pgcd}(n,k)$. Il est premier à ℓ ssi $v_\ell(k) \geqslant e$, soit n/ℓ^e valeurs de k. La probabilité est donc exactement $1-1/\ell^e$. □

En particulier, si $\ell > B$ est un grand diviseur premier de $p-1$, cette probabilité $\geqslant 1-1/\ell \approx 1$.

Donc à la fin de la boucle **for**, $b = a^{\mathrm{ppcm}(2,\ldots,B)}$ est congru à 1 modulo p si $p-1$ est B-friable. Et donc $p \mid b-1$. Avec un peu de chance, $N \nmid b-1$ et d est un facteur non trivial. L'algorithme peut échouer pour 2 raisons :

- $d = 1$: il n'y a pas de diviseur premier $p \mid N$ tel que that $p-1$ est B-friable. Il faut augmenter B et recommencer.
- $d = N$: l'ordre de $a \in (\mathbb{Z}/N\mathbb{Z})^*$ est B-friable. Ce qui veut probablement dire que $\varphi(N)$ est B-friable. Il faut diminuer B et recommencer.

Complexité

Le coût de la méthode est dominé par le calcul de b :

$$\sum_{\ell^k \leqslant B} (\log \ell^k) O(\log N)^2 = O(B(\log N)^2).$$