Factorisation d'entiers (première partie)

Francesco Pappalardi

Théorie des nombres et algorithmique

22 novembre, Bamako (Mali)

Quelle est la taille des "grands nombres"

- ☞ NOMBRE DE COMBINAISONS À LA LOTERIE : 622.614.630
- ☞ NOMBRE DE CELLULES DANS UN CORPS HUMAIN : 10^{15}
- ☞ NOMBRE D'ATOMES DANS L'UNIVERS : 10^{80}
- ☞ NOMBRE DE PARTICULES SUBATOMIQUES : 10^{120}
- ☞ NOMBRE D'ATOMES DANS LE CERVEAU HUMAIN : 10^{27}
- ☞ NOMBRE D'ATOMES DANS UN CHAT : 10^{26}

$$RSA_{2048} = 25195908475657893494027183240048398571429282126204$$
$$032027777137836043662020707595556264018525880784406918290641249$$
$$515082189298559149176184502808489120072844992687392807287776735$$
$$971418347270261896375014971824691165077613379859095700097330459$$
$$748808428401797429100642458691817195118746121515172654632282216$$
$$869987549182422433637259085141865462043576798423387184774447920$$
$$739934236584823824281198163815010674810451660377306056201619676$$
$$256133844143603833904414952634432190114657544454178424020924616$$
$$515723350778707749817125772467962926386356373289912154831438167$$
$$899885040445364023527381951378636564391212010397122822120720357$$

RSA_{2048} est un nombre avec 617 chiffres (décimaux)

`http://www.rsa.com/rsalabs/challenges/factoring/challengenumbers.txt`

$$RSA_{2048} = p \cdot q, \quad p, q \approx 10^{308}$$

$\boxed{\textbf{PROBLÈME: } \textit{Calculer } p \textit{ et } q}$

PRIX: 200.000 US\$ (~ 94.580.000 XOF)!!

Théorème. Si $a \in \mathbb{N}$, il y a $p_1 < p_2 < \cdots < p_k$ des nombres *premiers* uniques tels que $a = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$

Malheureusement: RSAlabs estime que pour factoriser ces nombres en un an nous avons besoin:

nombre	ordinateurs	mémoire
RSA_{1620}	1.6×10^{15}	120 Tb
RSA_{1024}	$342,000,000$	170 Gb
RSA_{760}	215,000	4Gb.

`http://www.rsa.com/rsalabs/challenges/factoring/challengenumbers.txt`

Nombre	Prix ($US)
RSA_{576}	$10,000
RSA_{640}	$20,000
RSA_{704}	$30,000
RSA_{768}	$50,000
RSA_{896}	$75,000
RSA_{1024}	$100,000
RSA_{1536}	$150,000
RSA_{2048}	$200,000

`http://www.rsa.com/rsalabs/challenges/factoring/challengenumbers.txt`

Nombre	Prix ($US)	Etat
RSA_{576}	\$10,000	Factorisé Décembre 2003
RSA_{640}	\$20,000	Factorisé Novembre 2005
RSA_{704}	\$30,000	pas factorisé
RSA_{768}	\$50,000	pas factorisé
RSA_{896}	\$75,000	pas factorisé
RSA_{1024}	\$100,000	pas factorisé
RSA_{1536}	\$150,000	pas factorisé
RSA_{2048}	\$200,000	pas factorisé

Célèbre citation!!!

Un phénomène dont la probabilité est 10^{-50} *ne se produira jamais, et de plus ne sera jamais observé.*

- Émil Borel (La probabilité et sa vie)

L'École d'Athènes (Raffaello Sanzio)

Euclide d'Alexandrie

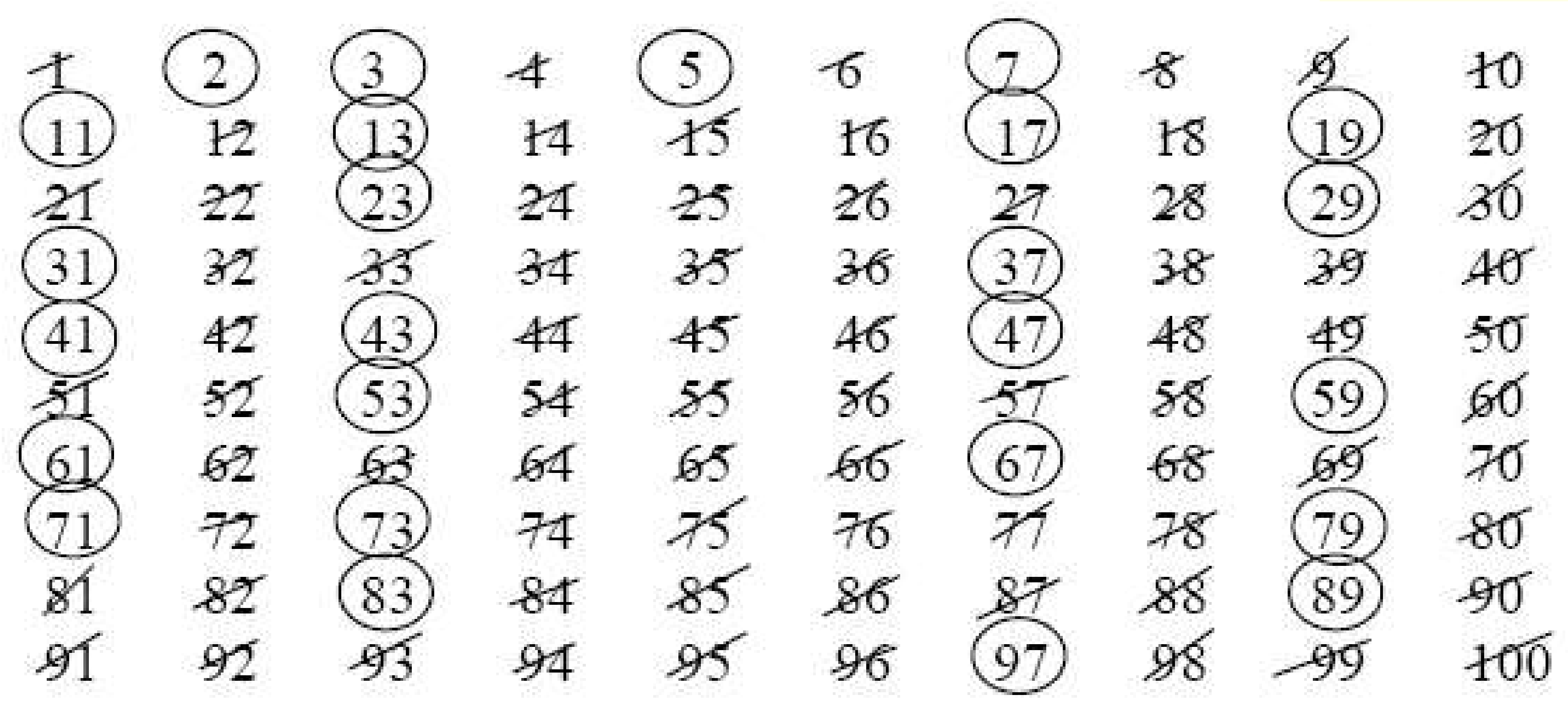
Naissance: 325 avant JC (Approx.)
Décès: 265 avant JC (Approx.)

Il y a une infinité de nombres premier!!

Histoire de "l'art de la factorisation"

220AC (Ératosthène de Cyrène)

Histoire de "l'art de la factorisation"



Le Crible d'Ératosthène

Histoire de "l'art de la factorisation"

1730 Euler $2^{2^5} + 1 = 641 \cdot 6700417$

Comment Euler a-t-il factorisé $2^{2^5}+1$?

PROPOSITION *Supposons que $p \mid b^n+1$. Il en résulte que soit*

1. $p \mid b^d+1$ *pour un certain diviseur propre d de n tel que n/d est impair, ou bien*
2. $p \equiv 1 \bmod 2n$.

PREUVE. Soit $d = \gcd(n, \frac{p-1}{2})$. On écrit l'identité de Bezout, $d = \alpha n + \beta \frac{p-1}{2}$.

De $b^n \equiv -1 \bmod p$ et de $b^{(p-1)/2} \equiv \pm 1 \bmod p$ (par le Théorème de Euler), il en résulte que $b^d = b^{\alpha n + \beta \frac{p-1}{2}} = (b^n)^\alpha \cdot (b^{\frac{p-1}{2}})^\beta \equiv \pm 1 \bmod p$.

En outre, de $b^n = (b^d)^{n/d} \equiv -1 \bmod p$, nous avons que $b^d \equiv -1 \bmod p$ et que n/d est impair.

Si $d < n$ nous sommes dans le cas (1) et si $n = d$, alors $n \mid (p-1)/2$ et nous sommes dans le cas (2). $\square$

Comment avez Euler factorisé $2^{2^5}+1$?

Application. Soit $b=2$ et $n=2^5=64$. Alors $2^{2^5}+1$ est soit an nombre premier ou bien est divisible par un nombre premier $p \equiv 1 \bmod 64$.

Notez que

$1+1\times 64 = 5\times 13$	$1+2\times 64 = 3\times 43,$	$1+3\times 64 = 193$ est premier
$1+4\times 64 = 257$ est premier	$1+5\times 64 = 3\times 107$	$1+6\times 64 = 5\times 7\times 11$
$1+7\times 64 = 449$ est premier	$1+8\times 64 = 3^3\times 19$	$1+9\times 64 = 577$ est premier
$1+10\times 64 = 641$ est premier		

Enfin

$$\frac{2^{2^5}+1}{641} = \frac{4294967297}{641} = 6700417$$

.

Application. Landry & Le Lasseur (1880): $2^{2^6}+1 = 274177\times 67280421310721$.

En fait, $274177 = 1+2^7\times 2142$.

Histoire de "l'art de la factorisation"

1730 Euler $2^{2^5} + 1 = 641 \cdot 6700417$

Histoire de "l'art de la factorisation"

1750–1800 Fermat, Gauss (Cribles - Tableaux)

L'idée d'un crible

La méthode de base

- Soit $n = pq$, avec $p > q$ nombres impairs
- On écrit $a = \frac{p+q}{2}, b = \frac{p-q}{2}$
- D'où $p = a + b, q = a - b$
- $n = (a+b)(a-b) = a^2 - b^2$
- Chaque nombre impair composé est une différence de carrés
- Cela est efficace si les 2 facteurs p et q sont proches

Exemple: $n = 200819$. $[\sqrt{n}] + 1 = 449$. $449^2 - n = 782$, ce n'est pas un carré. $450^2 - n = 1681 = 41^2$. D'où $n = 450^2 - 41^2 = 491 \times 409$.

L'idée d'un crible

➽ 1919 Pierre et Eugène Carissan (Machine pour Factoriser)

Exemple: $n = 611 = x^2 - y^2 = (x - y)(x + y)$

- Chaque carré parfait est 0, 1 ou 4 modulo 8
- $611^2 \equiv 3 \bmod 8$; par conséquent $x^2 = n + y^2 \equiv 4(\mathrm{mod} 8)$.
- Il suit que $x \equiv 2(\mathrm{mod} 4)$ et $y^2 \equiv 1(\mathrm{mod} 8)$
- De même, puisque chaque carré parfait est 0 ou 1 $(\mathrm{mod} 3)$, nous constatons que $x \equiv 0(\mathrm{mod} 3)$.
- La même idée pour les congruences supplémentaires modulo 5 et 7 conduit à: $x^2 \equiv 611 + 0, 1, 4(\mathrm{mod} 5)$ et $x^2 \equiv 611 + 0, 1, 2, 4(\mathrm{mod} 7)$

L'idée d'un crible

Donc, le problème de trouver un diviseur de 611 se réduit à trouver les valeurs de x tel que $0 \leq x < 611$ et

$$\begin{cases} x \equiv 2 & \pmod 4, \\ x \equiv 0 & \pmod 3, \\ x \equiv 0, 1, \text{ou} 4 & \pmod 5, \\ x \equiv 2, 3, 4, \text{ou} 5 & \pmod 7. \end{cases}$$

Enfin le plus petit x qui satisfait en même temps tout ce qui précède est 30 et, en fait, $611 = 30^2 - 17^2 = 13 \cdot 47$.

➼ En 1919, Pierre et Eugène Carissan ont construit une machine mécanique pour trouver un tel x automatiquement.

L'idée d'un crible

$$611 = x^2 - y^2 \qquad \begin{cases} x \equiv 2 & (\bmod 4), \\ x \equiv 0 & (\bmod 3), \\ x \equiv 0, 1, \text{ou} 4 & (\bmod 5), \\ x \equiv 2, 3, 4, \text{ou} 5 & (\bmod 7). \end{cases}$$

Chacune des 4 conditions est cyclique (c'est à dire qu'elles se répètent sans cesse).

x	0	1	2	3	4	5	6	7	8	9	10	11	12	$\cdots$	30
$x(\bmod 4)$	$\times$	$\times$	$\surd$	$\times$	$\times$	$\times$	$\surd$	$\times$	$\times$	$\times$	$\surd$	$\times$	$\times$	$\cdots$	$\surd$
$x(\bmod 3)$	$\surd$	$\times$	$\times$	$\surd$	$\times$	$\times$	$\surd$	$\times$	$\times$	$\surd$	$\times$	$\times$	$\surd$	$\cdots$	$\surd$
$x(\bmod 5)$	$\surd$	$\surd$	$\times$	$\times$	$\surd$	$\surd$	$\surd$	$\times$	$\times$	$\surd$	$\surd$	$\surd$	$\times$	$\cdots$	$\surd$
$x(\bmod 7)$	$\times$	$\times$	$\surd$	$\surd$	$\surd$	$\surd$	$\times$	$\times$	$\times$	$\surd$	$\surd$	$\surd$	$\surd$	$\cdots$	$\surd$

On peut représenter chaque condition favorable comme des clous dans un disque.

Ancien Machine pour factoriser dei Carissan

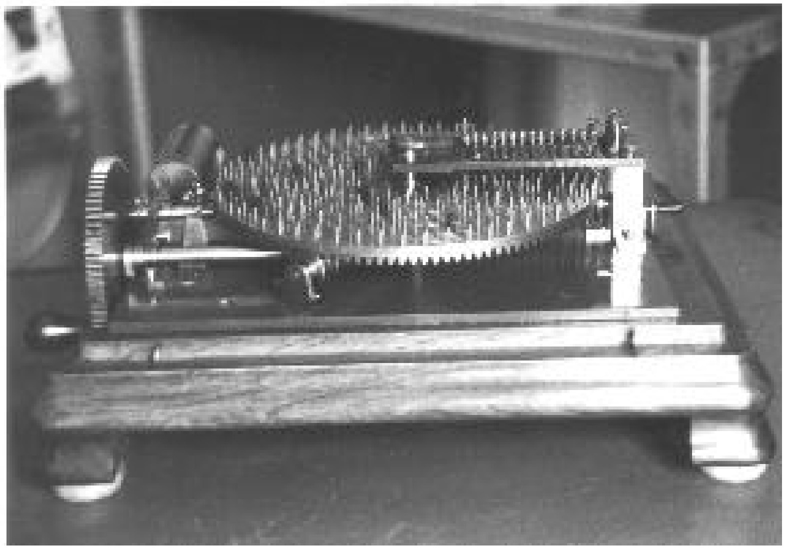


Figure 1: Conservatoire Nationale des Arts et Métiers in Paris

`http://www.cs.uwaterloo.ca/~shallit/Papers/carissan.html`

Ancien Machine pour factoriser dei Carissan

- L'idée de construire une machine mettant en œuvre des cribles est plus âgée.
- Charles Babbage a imaginé une machine similaire, mais ne l'a jamais construite
- Morain, Shallit et Williams ont découvert cette machine à l'Observatoire de Floriac, près de Bordeaux
- F. Morain, J. Shallit and H.C. Williams, *La machine à congruences*, Musée des Arts et Métiers La Revue, **14**, 1996, pp. 14-1

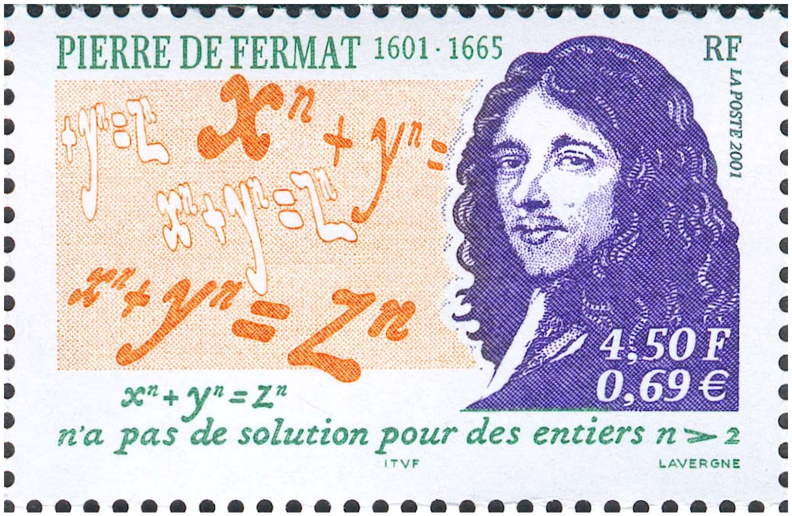
Figure 2: Lieutenant Eugène Carissan

$$225058681 = 229 \times 982789 \qquad \text{2 minutes}$$

$$3450315521 = 1409 \times 2418769 \qquad \text{3 minutes}$$

$$3570537526921 = 841249 \times 4244329 \qquad \text{18 minutes}$$

Histoire de "l'art de la factorisation"



1750–1800 Fermat, Gauss (Cribles - Tableaux)

Premier algorithme de factorisation par crible $N = x^2 - y^2 = (x - y)(x + y)$

Histoire de "l'art de la factorisation"

- 220AC (Ératosthène de Cyrène)
- 1730 Euler $2^{2^5}+1 = 641 \cdot 6700417$
- 1750–1800 Fermat, Gauss (Cribles - Tebleaux)
- 1880 Landry & Le Lasseur:

$$2^{2^6}+1 = 274177 \times 67280421310721$$

- 1919 Pierre et Eugène Carissan (Machine pour Factoriser)
- 1970 Morrison & Brillhart - *Le premier à utiliser des ordinateurs "modernes"*

$$2^{2^7}+1 = 59649589127497217 \times 5704689200685129054721$$

Histoire de "l'art de la factorisation"

1970 - John Brillhart & Michael A. Morrison

$$2^{2^7}+1 = 59649589127497217 \times 5704689200685129054721$$

La méthode ρ de J. Pollard (1975)

- PROBLÈME: Étant donné $n \in \mathbb{N}$, trouver un diviseur propre de n
- Un problème très ancien et très difficile;
- La plupart du temps la méthode de division "brute force" ou la méthode de crible de Fermat nècessite une quantité inacceptable de calculs
- Il y a plusieurs algorithmes différents
- nous passons en revue la méthode élégante de Pollard (méthode ρ).
- Il est basé sur le paradoxe des anniversaires

Dans une groupe de 23 amis il y a une bonne probabilité qu'au moins deux aient le même anniversaire

La méthode ρ de J. Pollard (1975)

- Soit n un nombre composé, et p un facteur premier de n
- Supposons que nous trouvons deux nombres a, b tels que: $a \equiv b \bmod p$ mais $a \not\equiv b \bmod n$
- Alors $\gcd(a-b, n)$ est un diviseur strict de n
- Cela nous permet de factoriser n
- Comment pouvons-nous trouver une telle paire?

Itération des fonctions aléatoires

- Soit E un ensemble fini d'ordre n
- Soit $f : E \to E$ une fonction aléatoire et x_0 un élément aléatoire de E
- Nous définissons une séquence par $x_{n+1} = f(x_n)$
- Cette séquence est une suite périodique, de période au plus n
- Á quelle période peut-on s'attendre, pour une fonction aléatoire ?
- Nous supposons que n n'est pas une puissance parfaite
  et nous considérons: $f : \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}, \quad x \mapsto f(x) = x^2 + 1.$
- Ainsi, la k-ième itération de f est $f^k(x) = f^{k-1}(f(x))$ avec $f^1(x) = f(x)$.
- Si $x_0 \in \mathbb{Z}/n\mathbb{Z}$ est choisi suffisamment "aléatoire",
  la séquence $\{f^k(x_0)\}$ se comporte comme une séquence aléatoire d'éléments de $\mathbb{Z}/n\mathbb{Z}$
- nous allons exploiter ce fait

Itération des fonctions aléatoires

- La durée moyenne n'est pas grande, elle est de l'ordre de $\sqrt{n}$

- Plus précisément:
  Soit $\lambda > 0$**, et** $l = 1 + [\sqrt{2\lambda n}]$**. La probabilité de** $(f; x_0)$ **telle que** $x_0, x_1, \ldots, x_l$ **sont toutes distinctes est inférieure à** $e^{-\lambda}$

- Si la fonction est de definie sur $\mathbb{Z}/n\mathbb{Z}$, et se dirige vers le quotient $\mathrm{mod} p$, on peut s'attendre à ce que la période est beaucoup plus courte $\mathrm{mod} p$.
- Nous supposons que le la fonction $X^2 + 1$ est une fonction aléatoire

La méthode de factorisation ρ de Pollard

```
Input:   n ∈ ℕ odd and not a perfect power (to be factored)
Output:  a non trivial factor of n
1.  Choose at random x ∈ ℤ/nℤ = {0, 1, ..., n − 1}
2.  For i = 1, 2 ....
          g := gcd(f^i(x) − f^{2i}(x), n)
          If g = 1, goto next i
          If 1 < g < n then output g and halt
          If g = n then go to Step 1 and choose another x.
```

Qu'est-ce qui se passe ici?

Est est évidemment un algorithme probabiliste, mais il n'est même pas certain qu'il sera jamais fin.

Mais en fait, il termine avec environ $\sqrt[4]{n}$ itérations qui est atteint dans le pire des cas (iwhen n est un module RSA).

Le paradoxe des anniversaires

Question de théorie élémentaire des probabilités : *quelle est la probabilitè que dans une séquence des k éléments (où les rèpètitions sont autorisés) è partir d'un ensemble de n éléments, il y a une répétition?*

Réponse : Le hasard est $1 - \dfrac{n!}{n^k(n-k)!} \approx 1 - e^{-k(k-1)/2n}$

> Dans une groupe de 23 amis il y a une probabilité 50.04% qu'au moins deux aient le même anniversaire!!

Pertinence de la méthode ρ :

> Si d est un diviseur de n, alors dans l'ordre de $O(\sqrt{d}) = O(\sqrt[4]{n})$ steps il y a une forte hasard que dans la séquence $\{f^k(x_0) \bmod d\}$ il y a une répétition modulo d.

REMARQUE (POURQUOI ρ?). Si

$y_1, \ldots, y_m, y_{m+1}, \ldots, y_{m+k} = y_m, y_{m+k+1} = y_{m+1}, \ldots$ et i est le plus petit multiple de k avec $i \geq m$, alors $y_i = y_{2i}$ (le truc du cycle de l'Floyd).

Example :

- $n = 4171$, $f(X) = X^2 + 1$, $x_0 = 2$;
- $X_1 = 5$, $X_2 = 26$, $\gcd(X_2 - X_1, n) = 1$;
- $X_3 = 677$, $X_4 = 3691$, $\gcd(X_4 - X_2, n) = 1$;
- $X_5 = 996$, $X_6 = 3490$, $\gcd(X_6 - X_3, n) = 1$;
- $X_7 = 781$, $X_8 = 996$, $\gcd(X_8 - X_4, n) = 1$;
- $X_9 = 3490$, $X_10 = 781$, $\gcd(X_{10} - X_5, n) = 43$;

Références pour ce cours

[1] J. Buhler & S. Wagon *Basic algorithms in number theory* Algorithmic Number Theory, MSRI Publications Volume **44**, 2008
http://www.msri.org/communications/books/Book44/files/02buhler.pdf

[2] C. Pomerance *Smooth numbers and the quadratic sieve* Algorithmic Number Theory, MSRI Publications Volume **44**, 2008
http://www.msri.org/communications/books/Book44/files/03carl.pdf

[3] R. Crandall and C. Pomerance, *Prime numbers*, 2nd ed., Springer-Verlag, New York, 2005.

[4] E. Bach and J. Shallit, *Algorithmic number theory, I: Efficient algorithms*, MIT Press, Cambridge, MA, 1996.

[5] J. von zur Gathen and J. Gerhard, *Modern computer algebra*, 2nd ed., Cambridge University Press, Cambridge, 2003.

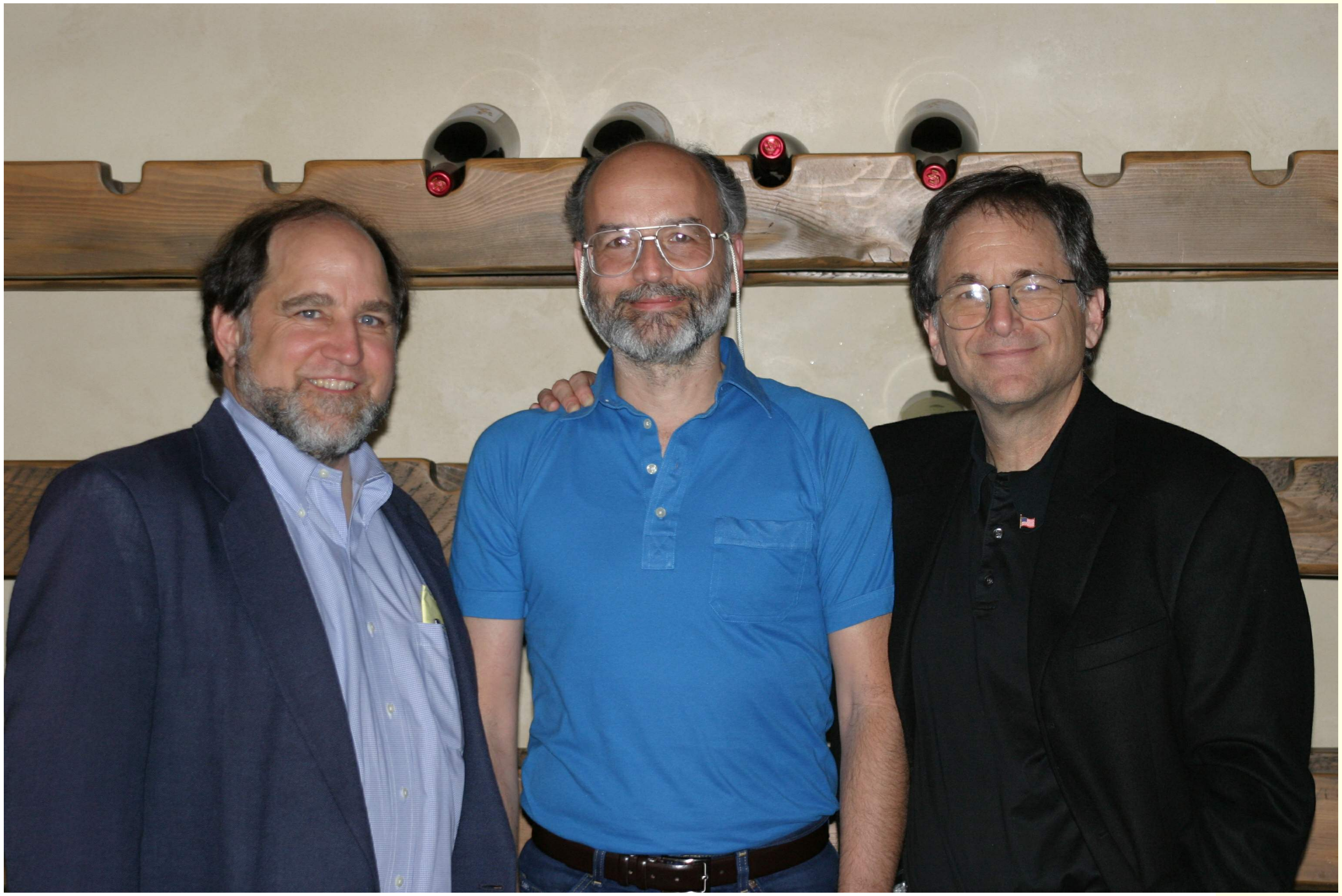
[6] V. Shoup, *A computational introduction to number theory and algebra*, Cambridge University Press, Cambridge, 2005.

[7] These notes http://www.mat.uniroma3.it/users/pappa/bamako2010_A.pdf

RSA

Adi Shamir, Ron L. Rivest, Leonard Adleman (1978)

RSA



Ron L. Rivest, Adi Shamir, Leonard Adleman (2003)