Statistiques sur le PGCD de nombres au hasard

1 Présentation du problème, Notations

On tire au sort deux nombres entiers naturels a et b ayant au plus 1024 bits. On suppose que chaque entier n tel que :

$$0 \leq n \leq 2^{1024} - 1$$

a la même chance que les autres d'être tiré. On se fixe un entier h.

Le problème posé est le suivant : quelle est la probabilité $P(h)$ pour que $\gcd(a, b) \geq h$?

Soit N un entier naturel. On pose :

$$I = \{1, 2, \cdots, N\}.$$

On définit alors les ensembles suivants :

$$E = \{(a, b) \in I^2 \mid a \wedge b = 1\},$$
$$A = \{(a, b) \in I^2 \mid b < a \text{ et } a \wedge b = 1\},$$
$$B = \{(a, b) \in I^2 \mid a < b \text{ et } a \wedge b = 1\}.$$

On remarque que le couple $(1, 1)$ est le seul couple de I^2, de nombres premiers entre eux qui ne soit ni dans A ni dans B. Avec les notations précédentes, on dispose des relations suivantes :

$$A \cap B = \emptyset,$$
$$E = A \cup B \cup \{(1, 1\},$$
$$\#A = \#B,$$

et donc :

$$\#E = 2\#A + 1.$$

Il convient de calculer le nombre d'éléments de A. Pour cela on introduit pour tout a fixé :

$$A_a = \{b \in I \mid b < a \text{ et } a \wedge b = 1\}.$$

Alors :

$$A = \bigcup_{a \in I} A_a \times \{a\}$$

De plus les ensembles $A_a \times \{a\}$ sont disjoints. Pour tout $a \geq 2$, on a :

$$\#A_a \times \{a\} = \#A_a = \phi(a),$$

ou ϕ est la fonction indicatrice d'Euler, et clairement $A_1 = \emptyset$. Donc :

$$\#A = \sum_{1 \leq a \leq N} \phi(a) - 1.$$

On conclut que :

$$\#E = 2 \sum_{1 \leq a \leq N} \phi(a) - 1.$$

2 Nombres premiers entre eux

Mettons sur l'intervalle I la probabilité équirépartie. On effectue l'expérience suivante : on tire au sort un nombre $a \in I$ et on tire au sort un nombre $b \in I$, ces deux tirages étant faits indépendamment. On cherche la probabilité p pour que a et b soient premiers entre eux. On a bien entendu :

$$p = \frac{\#E}{N^2},$$

ce qui donne :

$$p = \frac{2\sum_{1\leq a\leq N}\phi(a) - 1}{N^2}.$$

On sait que

$$\frac{\sum_{1\leq a\leq N}\phi(a)}{N^2}$$

est à peu près égal à $3/\pi^2$.

On pourra voir sur des essais que par exemple pour $N \geq 10$ la probabilité p peut être prise raisonnablement pour :

$$p \simeq \frac{6}{\pi^2}.$$

3 PGCD de deux nombres au hasard

Nous en venons ici au problème posé, c'est-à-dire que nous étudions le $\gcd$ de deux nombres de l'intervalle I pris au hasard.

Posons :

$$G_k = \{(a,b) \in I^2 \mid gcd(a,b) = k\}.$$

Il est clair que :

$$G_k = \{(a,b) \in I^2 \mid a = k_1 k, b = k_2 k, gcd(k_1,k_2) = 1\}.$$

Notons aussi :

$$H_k = \{(k_1,k_2) \in I^2 \mid 1 \leq k_1,k_2 \leq \frac{N}{k},\ k_1 \wedge k_2 = 1\}.$$

Alors d'après les définitions on voit que :

$$\#G_k = \#H_k.$$

Mais on a approximativement :

$$\#H_k \simeq \frac{6}{\pi^2}\frac{N^2}{k^2}.$$

Si on introduit maintenant :

$$G = \{(a,b) \in I^2 \mid gcd(a,b) \geq h\},$$

alors :

$$\#G = \sum_{h\leq k\leq N} \#G_k \simeq \frac{6}{\pi^2}\,N^2 \sum_{h\leq k\leq N}\frac{1}{k^2}.$$

Et la probabilité $P(h)$ vaut approximativement :

$$P(h) \simeq \frac{6}{\pi^2}\sum_{h\leq k\leq N}\frac{1}{k^2},$$

ce qui, compte tenu de l'évaluation asymptotique du reste de la série de terme général $1/k^2$, est encore approximativement égal à :

$$P(h) \simeq \frac{6}{\pi^2}\frac{1}{h}$$