Quelques résultats supplémentaires de théorie des nombres

Christophe Antonini[1], Olivier Teytaud[2], Pierre Borgnat[3], Annie Chateau[4], and Edouard Lebeau[5]

[1]Enseignant en CPGE, Institut Stanislas, Cannes
[2]Chargé de rechercher INRIA, Université d'Orsay, Orsay
[3]Chargé de recherche CNRS, ENS Lyon, Lyon
[4]Maitre de conférence, Université Montpellier-2, Montpellier
[5]Enseignant en CPGE, Lycée Henri Poincaré, Nancy

7 avril 2023

Compléments sur les groupes, quelques applications dont une à la cryptographie.

1 Quelques résultats supplémentaires de théorie des nombres

On présentera ici quelques résultats supplémentaires, de bon aloi pour illustrer une leçon : la très classique étude des sous-groupes de $\mathbb{R}$ pour l'addition (1.1), représentation p-adique des réels (1.2), fractions continues (1.3), cryptographie à clé révélée (1.4).

1.1 Sous-groupes additifs de $\mathbb{R}$

La proposition qui suit n'est pas difficile, mais tellement plus joliment rédigée quand on a vu une fois qu'il fallait introduire $\inf G \cap \mathbb{R}^{+*}$.

Proposition 0.1

Tout sous-groupe G du groupe $(\mathbb{R}, +)$ vérifie l'une et une seule des deux conditions suivantes :
- $\exists x \;\; G = x\mathbb{Z}$
- G est dense dans $\mathbb{R}$.

Démonstration *On considère $\alpha = \inf G \cap R^{+*}$. On distingue les deux cas $\alpha > 0$ et $\alpha = 0$.*

1.2 Représentation p-adique des réels

DÉFINITION 0.1 **fraction continue**

On se donne un entier $p > 1$. On appelle représentation p-adique du réel x la suite d'entiers $(c_n)_{n\in\mathbb{N}}$ définie par

$$c_n = \begin{cases} E(x) \text{ si } n = 0 \\ \frac{1}{p^n}[E(p^n x) - pE(p^{n-1}x)] \text{ sinon} \end{cases}$$

($E(y)$ désignant la partie entière de y).

Intuition Le développement p-adique des réels est simplement la façon quotidienne de parler des nombres réels : le développement p-adique de π est simplement $c_0 = 3$, $c_1 = 1$, $c_2 = 4$, $c_3 = 1$,... Les ordinateurs utilisent la même chose en binaire ($p = 2$).

Une propriété importante est

$$x = \sum_{n=0}^{+\infty} c_n p^{-n}$$

THÉORÈME 0.2

Le développement p-adique de $x \in \mathbb{R}$ est périodique à partir d'un certain rang si et seulement si x est rationnel.

Démonstration

•*Supposons tout d'abord le développement périodique.*

Alors x est somme des $c_n p^{-n}$ pour $n \in \mathbb{N}$. Vue la périodicité, cette somme se réécrit comme somme d'un rationnel et de $\sum_{n\geq N} \frac{a}{(p^{-k})^n}$, avec a dans $\mathbb{N}$ et $k > 0$, et donc x est somme d'un rationnel et de $\frac{ap^{-kN}}{1-p^{-k}}$, et donc x est rationnel.

•*Réciproquement supposons que x soit rationnel.*

- On peut écrire $x = a/b$ avec a et b dans $\mathbb{N}$ (on se limite au cas $x > 0$, les autres cas étant similaires)

- On définit $x_0 = a$, et par récurrence $x_{n+1} = (x_n - bc_n)p$, avec c_n le quotient dans la division euclidienne de x_n par b.

- On montre facilement par récurrence que $0 \leq x_i < bp$ pour tout i et que les c_i sont le développement p-adique de x.

- les x_i étant bornés, on passe nécessairement deux fois par la même valeur ; à partir de ce moment, le développement est clairement périodique.

1.3 Fractions continues

DÉFINITION 0.2 **Fractions continues**

Une **fraction continue** est un objet de la forme suivante :

$$[a_0, a_1, \ldots, a_n, \ldots] = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \ldots}}}$$

Elle est caractérisée par une suite d'entiers qui est finie ou infinie.

On appelle **convergents** d'une fraction continue la suite de numérateurs p_n et de dénominateurs q_n définis par :

— $p_0 = a_0$, $q_0 = 1$

— $p_1 = a_0 a_1 + 1$, $q_1 = 1$

— $\vdots$

— $p_n = a_n p_{n-1} + p_{n-2}$,$q_n = a_n q_{n-1} + q_{n-2}$.

Quelques propriétés :

— A tout nombre réel on peut associer un et un seul développement en fraction continue.

— Tout nombre rationnel peut être représenté par une fraction continue finie (ex $\frac{1}{3} = 0 + \frac{1}{2+1}$).

— Seuls les nombres rationnels peuvent être représentés par une fraction continue finie.

— Un nombre est quadratique (i.e. solution d'une équation du second degré à coefficients dans $\mathbb{Z}$) si et seulement si son développement en fraction continue est périodique.

— Une fraction continue est liée à ses convergents par les relations $[a_0, \ldots, a_n] = p_n/q_n$ et $[a_0, \ldots, a_n, \ldots] = \lim_n \frac{p_n}{q_n}$. En outre avec $[a_0, \ldots, a_n] = \frac{p_n}{q_n}$, $|[a_0, \ldots, a_n] - [a_0, \ldots, a_n, \ldots]| < 1/q_n^2$.

THÉORÈME 0.3 **Formule d'Euler**

Supposons les a_i tous non nuls.
Alors $1/a_1 - 1/a_2 + 1/a_3 - 1/a_4 + ... + (-1)^n/a_n =$

$$\cfrac{1}{a_1 + \cfrac{a_1^2}{a_2 - a_1 + \cfrac{a_1^2}{a_3 - a_2 + \cfrac{a_3^2}{\ddots + \cfrac{\ddots}{a_{n-1} - a_{n-2} + \cfrac{a_{n-1}^2}{a_n - a_{n-1}}}}}}}.$$

Démonstration

•*Pour $n = 1$, le résultat est clair.*

•*Au rang 2, un calcul rapide montre que le résultat est encore valable.*

•*On procède ensuite par récurrence, en supposant l'égalité vraie pour $n - 1$ et les rangs inférieurs.*

•*Dans l'égalité pour $n - 1$, on remplace a_n par $\frac{a_n . a_{n+1}}{a_{n+1} - a_n}$.*

•*Le résultat en découle tout seul.*

1.4 Cryptographie à clé révélée : RSA

Nous nous limiterons ici à une brève introduction. On pourra consulter [1] pour plus d'informations.

Précisons que l'on parle aussi parfois de clé 'publique' ; il s'agit de la même notion que la clé révélée.

L'objectif de la **cryptographie** est de permettre de communiquer par des messages codés, qui ne pourront être lus que par leur destinataire.

Pour cela, un « superviseur » donne à chaque receveur potentiel un « décodeur » et une « clé ». La clé, comme son nom ne l'indique peut-être pas, est quelque chose qui peut être diffusé à tout le monde.

Pour envoyer un message M crypté à un individu I, il suffit de passer le message M par la moulinette de la clé correspondante à I. Cela n'est pas difficile, puisque I diffuse abondamment sa clé, à tous ses correspondants éventuels. Lorsque I reçoit un message, il peut alors utiliser son décodeur, qu'il est seul à posséder, pour transformer le message crypté en le message original.

La difficulté est que, formellement, il est toujours possible de reconstruire le message initial à partir du message crypté, pourvu que vous ayez la clé. Pour cela, il suffit de tester tous les messages possibles, l'un après l'autre (ils sont bien en bijection avec $\mathbb{N}$, comme on peut s'en convaincre facilement en considérant l'ordre lexicographique sur les messages possibles), et de les passer par la moulinette de la clé jusqu'à ce que l'on retrouve le message crypté. Mais il reste un espoir de fabriquer une cryptographie efficace, car bien sûr, cette méthode prendrait un temps énorme. La cryptographie est ainsi basée sur l'hypothèse de base que certaines tâches, faciles à faire dans un sens (le sens du cryptage par une clé), sont difficiles à faire dans l'autre (décryptage à l'aide d'une clé).

On note bien que la difficulté réside dans le fait que la fonction « clé » est publique. Si on cache la clé, il est très facile de réaliser des cryptographies parfaites. Par exemple, on peut utiliser le protocole suivant pour que A envoie un message à B :

- A signale à B qu'il veut lui envoyer un message, que l'on supposera constitué uniquement de 0 et de 1 (par un codage quelconque on peut facilement se ramener à cela), et de longueur 1000.
- B fournit à A une liste L de 1000 chiffres 0 ou 1, 0 et 1 étant équiprobables.
- le protocole recommence à l'étape précédente jusqu'à ce que la liste de chiffres soit passée sans être interceptée ; on tire au sort une nouvelle liste de 1000 chiffres à chaque nouvel essai.
- A transforme le message M en un message M', par $M' = M + L$ dans $\mathbb{Z}/2\mathbb{Z}$.
- A envoit M' à B ; si M' est intercepté, il ne sera pas décodable, puisque L n'est pas connu.
- B décode M' par $M = M' + L$ dans $\mathbb{Z}/2\mathbb{Z}$.

Aucune interception ne permet de décoder le message ; mais les étapes 2 et 3 peuvent prendre du temps ou n'être pas réalisables. Il est indispensable de changer de liste L à chaque nouveau message, ou du moins régulièrement - sinon, en considérant les fréquences des 0 et des 1, un observateur des différents M' pourrait finir par reconstituer L.

L'algorithme **RSA**, du nom de ses inventeurs, Rivest, Shamir et Adleman, est basé sur la difficulté de la factorisation d'un nombre entier en nombres premiers.

Supposons que A souhaite envoyer des messages cryptés RSA à B. Alors B se donne deux grands nombres premiers p et q. Maple permet aisément de construire de tels nombres, par exemple ; il suffit par exemple de tirer des nombres au sort et de recommencer jusqu'à ce qu'ils soient premiers, grâce à un algorithme permettant de dire si oui ou non un nombre est premier. En fait les algorithmes utilisés pour cela sont généralement probabilistes, c'est-à-dire qu'ils ont une probabilité non nulle de se tromper ; mais les erreurs sont extrêmement rares. La fonction Maple `isprime` permet de tester la primalité d'un nombre ; par exemple, `isprime(12345678901234567890123456 7)` renvoit `false`, donc `12345678901234567890123456 7` n'est pas premier. `nextprime(12345678901234567890123456 7)` renvoit `12345678901234567890123465 1`, qui est donc le nombre premier le plus petit plus grand

que celui-ci. On constate donc que Maple permet très rapidement de trouver de grands nombres premiers ; les exemples ici fournis ne sont pas du tout à la limite du faisable, on peut aller largement au delà.

Ces deux nombres premiers seront notés p et q. La première partie de la clé publique, notée c, sera le produit de p et q. A utilise alors un nombre d (d est la seconde partie de la clé publique, qui peut donc être fournie par B éventuellement, si A n'a pas eu l'occasion d'accéder de manière sûre à p et q), premier avec $\phi(c) = (p-1)(q-1)$, où ϕ est la fonction d'Euler, c'est-à-dire le nombre de nombres premiers plus petits que c, donc $(p-1)(q-1)$. Il est facile de choisir un nombre qui soit premier avec un autre : il suffit d'en piocher un au hasard, et de recommencer jusqu'à ce que l'algorithme de Bezout confirme que ces nombres sont premiers entre eux. On peut aussi déterminer facilement d^{-1}, inverse de d dans $(\mathbb{Z}/c\mathbb{Z})^* \simeq \mathbb{Z}/\phi(c)\mathbb{Z}$ (il y a isomorphisme car $(\mathbb{Z}/c\mathbb{Z})^* \simeq (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/q\mathbb{Z})^*$ par le Corollaire **??** et isomorphisme entre $(\mathbb{Z}/p\mathbb{Z})^*$ et $\mathbb{Z}/(p-1)\mathbb{Z}$ (resp. $(\mathbb{Z}/q\mathbb{Z})^*$ et $\mathbb{Z}/(q-1)\mathbb{Z}$) par le Lemme **??**) : il suffit d'utiliser l'algorithme de Bezout.

Cryptage :

- On suppose le message suffisamment court pour être codable par un élément inversible de $\mathbb{Z}/c\mathbb{Z}$, ce qui est possible en remplaçant le message par des tranches successives suffisamment petites (si on a un alphabet de taille α, il suffit de prendre des tranches de longueur l avec $\alpha^l < \phi(c)$). Cette méthode de codage n'a pas à être compliquée ni à être cachée. Il suffit donc d'avoir une injection de $[1, \alpha^l]$ dans $(\mathbb{Z}/c\mathbb{Z})^*$.

- chaque message de A est donc remplacé (par A) par un élément n inversible dans $\mathbb{Z}/c\mathbb{Z}$.

- A envoie alors à B le nombre $e = n^d$ dans $\mathbb{Z}/c\mathbb{Z}$.

Décryptage :

- B, qui dispose de d et de d^{-1}, effectue simplement le calcul de $e^{d^{-1}}$, qui lui donne n, et donc le message initial.

D'autres systèmes de cryptographie à clé publique font intervenir des structures plus complexes que $\mathbb{Z}/n\mathbb{Z}$, comme par exemple les courbes elliptiques.

Au total :

— Information accessible à B : p, q, $c = pq$, d, d^{-1} (pour $\mathbb{Z}/c\mathbb{Z}$), e (version codée du message, fournie par A).

— Information accessible à A : n, c, d, $e = n^d$ dans $\mathbb{Z}/c\mathbb{Z}$.

— Information accessible à tout l'univers : c, d, e. c et d permettent théoriquement de calculer d^{-1} dans $\mathbb{Z}/d\mathbb{Z}$, mais ce calcul est énorme ; alors qu'en disposant de q et p ce calcul est facile, grâce à notre théorème de Bezout.

Références

[1] F. Combes *Algèbre et géométrie*, Bréal, 1998.