ARITHMÉTIQUE

JEAN-PAUL CALVI

Le thème de cette séquence correspond à un sous-ensemble du Chapitre 3 par. a) du programme officiel de l'agrégation interne de mathématiques : anneau $\mathbb{Z}$ des entiers relatifs; division euclidienne; sous-groupes additifs de $\mathbb{Z}$; nombres premiers; décomposition en facteurs premiers; plus grand commun diviseur (PGCD, *ici* gcd) et plus petit commun multiple (PPCM); théorème de Bézout; algorithme d'Euclide; congruences; applications arithmétiques des anneaux quotients $\mathbb{Z}/n\mathbb{Z}$; théorème chinois; groupe des éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$ (*ajout : petit théorème de Fermat*); applications à des problèmes de calendriers; exemples de méthodes de codage et de cryptage.

1. Arithmétique élémentaire

1.1. **Algorithme(s) d'Euclide.** Il se trouve dans le second livre d'Euclide (IIIe siècle av. J.-C.).

Algorithme 1 (Euclide)**.** Soient a et b deux entiers positifs ou nuls. Le résultat ($\gcd(a,b)$) se trouve dans la variable a à la fin de l'algorithme. Celui-ci utilise les variables a,b et la variable auxiliaire r.

Step 1.1 (Arrêt ?). Si $b = 0$ alors afficher la réponse a et FIN.

Step 1.2. FAIRE $r \leftarrow a \mod b$, $a \leftarrow b$, $b \leftarrow r$ et RETOURNER à 1.1.

Ici on utile la notation $a \mod b$ pour désigner le reste dans la division de a par b. La notation reprendra le sens arithmétique usuel plus bas. On convient que $a \mod 0 = a$.

On peut aussi énoncer l'algorithme de la manière suivante.

Théorème 1 (Euclide)**.** *Soient a et b deux entiers positifs ou nuls avec $a \geq b$. On définit une suite d'entires u_n de la manière suivante $u_0 = a$, $u_1 = b$ et pour tout $n \geq 0$, $u_{n+2} = u_n \mod u_{n-1}$. La suite u_n est nulle à partir d'un certain rang et $\gcd(a,b)$ est le dernier terme non nul de cette suite.*

1 Montrer le théorème.

1.2. **L'algorithme d'Euclide modifié.** Le théorème de Bézout dit qu'il existe toujours un couple[1] (u,v) tel que $ua + vb = \gcd(a,b)$. Un tel couple est obtenu de manière très efficace en "remontant" les étapes de l'algorithme 1. Mais cette méthode a l'inconvénient, quand on la programme, de devoir garder en mémoire un grand nombre de données. Il existe une méthode (un peu plus coûteuse en temps) mais qui donne directement le gcd et un couple (u,v).

Théorème 2. *Soient $a > b$ deux entiers strictement positifs. On définit la suite de triplets $W_n = (t_n,u_n,v_n)$ par $W_0 = (a,1,0)$, $W_1 = (b,0,1)$ et pour $n \geq 2$,*

$$W_n = W_{n-2} - [t_{n-2}/t_{n-1}]W_{n-1} \tag{1.1}$$

Date: Octobre 2005.

1. En réalité une infinité de couples voir 1.4.

où la notation qW_{n-1} désigne le triplet $(qt_{n-1},qu_{n-1},qv_{n-1})$ et $[p/s]$ est la partie entière de p/s, autrement dit le quotient dans la division euclidienne de p par s.

Avec ces notations, la suite (t_n) est nulle à partir d'un certain rang N. On a alors

$$\gcd(a,b) = t_{N-1} \quad et \quad \gcd(a,b) = u_{N-1}a + v_{N-1}b. \tag{1.2}$$

2 (Démonstration)

(1) Montrer que $t_{n+2} = t_{n+1} \mod t_n$ et en déduire l'existence de N.

(2) Montrer que pour tout $n \geq 0$, on a

$$t_n = u_n a + v_n b$$

et en déduire que $\gcd(a,b)$ divise t_{N-1}.

(3) Montrer que t_n divise t_{n-2} puis $t_{n-3},\ldots,t_1,t_0$. Conclure la démonstration du théorème.

1.3. **Exercices.**

3 Pour chaque couple d'entiers (a,b) ci-dessous, déterminer, au moyen de l'algorithme d'Euclide, le gcd de a et b ainsi que deux entiers u et v tels que $ua + bv = \gcd(a,b)$.

(1) $(a,b) = (7200,3132)$

(2) $(a,b) = (1812,1572)$

1.4. **Équations diophantiennes $ax + by = c$.** On recherche une méthode pour résoudre des équations de la forme

$$ax + by = c \tag{1.3}$$

où a,b,c sont des entiers donnés et $(x,y) \in \mathbb{Z} \times \mathbb{Z}$ sont les inconnues.

4 Montrer qu'une condition nécessaire et suffisante pour que l'équation ci-dessus ait au moins une solution est que c soit un multiple de $\gcd(a,b)$.

5 Montrer que lorsque cette condition est satisfaite on peut trouver une solution *particulière* de l'équation en utilisant (par exemple) l'algorithme d'Euclide.

6 Résoudre (1.3) dans le cas où $c = 0$.

7 On suppose que l'on connaît une solution particulière (x_0,y_0) de l'équation (1.3). Montrer que si (x_1,y_1) est une autre solution alors $(x_1 - x_0, y_1 - y_0)$ est solution de $ax + by = 0$ et en déduire une méthode générale de résolution de (1.3).

8 Résoudre les équations suivantes :

(1) $3x + 5y = 2$,

(2) $12x + 9y = 6$.

1.5. **Le théorème de Wilson.** On démontre un célèbre théorème sur les nombres premiers de John Wilson (1741-93), avocat et mathématicien anglais, publié en 1770 dans ses *Meditationes Algebraicae*.

Théorème 3 (Wilson)**.** *Pour tout nombre premier p on a*

$$(p-1)! \equiv -1 \mod p.$$

9 Soit m un entier positif non nul.

Montrer que si $a \in \mathbb{Z}$ est premier avec m alors l'application $\varphi : \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ définie (correctement ?) par $\varphi(\overline{x}) = \overline{ax}$ est une bijection.

Rappel : la notation $\overline{x}$ désigne la classe de x modulo m.

10 Montrer que la propriété n'est plus vraie lorsque a n'est pas premier avec m.

11 Déduire de la question précédente que si p est un nombre *premier* alors pour tout $a \in \{1,\ldots,p-1\}$ il existe un unique $a' \in \{1,\ldots,p-1\}$ tel que aa' soit congru a 1 modulo p. Le nombre a' s'appelle le *réciproque* de a modulo p.

12 Déterminer les réciproques de 1,2,3 et 4 modulo 5 et les réciproques de 1,2,3,4,5 et 6 modulo 7.

13 Dans le cas général, à quelle(s) condition(s) a est-il son propre réciproque modulo p?

14 En regroupant, dans le produit $1\times 2\times 3\times\cdots\times(p-2)\times(p-1)$ chaque nombre avec son réciproque, montrer le théorème de Wilson.

1.6. **Propriétés arithmétiques des coefficients binomiaux.**

15 Montrer que le produit de n entiers positifs *consécutifs* est toujours divisible par $n!$.

16 Démontrer que si p est un nombre premier alors les nombres

$$C_p^1, C_p^2, C_p^3, \ldots C_p^{p-1}$$

sont divisibles par p.

17 Démontrer que si p est premier et si a_1,$a_2,\ldots,a_k$ sont k nombres entiers alors on a

$$(a_1+a_2+\cdots+a_k)^p \equiv a_1^p+a_2^p+\cdots+a_k^p \mod p.$$

18 Soit α un entier positif et p un nombre premier. Montrer que

$$m=1 \mod p^\alpha \quad\Longrightarrow\quad m^p \equiv 1 \mod p^{\alpha+1}.$$

2. FONCTION D'EULER

2.1. **Définition de la fonction d'Euler.** Soit $m \in \mathbb{N}/\{0,1\}$.

19 Montrer que $\overline{r} \in (\mathbb{Z}/m\mathbb{Z})^*$ si et seulement si m et r sont premiers entre eux. On note $\phi(m)$ le cardinal de $(\mathbb{Z}/m\mathbb{Z})^*$, le sous-groupe des éléments inversibles de $\mathbb{Z}/m\mathbb{Z}$.

Definition 1. L'application ϕ s'appelle L'INDICATRICE D'EULER. On convient que $\phi(0)=0$ et $\phi(1)=1$.

20 Montrer le théorème suivant qui est une généralisation du petit théorème de Fermat.

Théorème 4. *Let $m \geq 1$. Pour tout $a \in \mathbb{Z}$ premier avec m, on a $a^{\phi(m)} = 1$ mod m.*

21 (Calcul de ϕ sur les puissances de nombres premiers.)

(1) Que vaut $\phi(p)$ lorsque p est un nombre premier?

(2) Que vaut $\phi(m)$ lorsque $m=p^s$ avec p premier?

Definition 2. On dit qu'une fonction f définie sur $\mathbb{N}$ à valeurs dans $\mathbb{Z}$ est *arithmétique* si $f(m\cdot n) = f(m)\cdot f(n)$ chaque fois que m et n sont premiers entre eux.

On se propose de démontrer le théorème suivant.

Théorème 5. *La fonction d'Euler est arithmétique.*

22 (Question préliminaire sur le produit de deux anneaux.) Soient $(A, \oplus_A, \odot_A)$ et $(B, \oplus_B, \odot_B)$ deux anneaux commutatifs unitaires. On définit sur $A \times B$ les lois $+$ et $\cdot$ de la manière suivante

$$(a,b) + (a',b') = (a \oplus_A a', b \oplus_B b') \quad \text{et} \quad (a,b) \cdot (a',b') = (a \odot_A a', b \odot_B b').$$

(1) Montrer que $(A \times B, +, \cdot)$ est un anneau commutatif unitaire. Est-il intègre ?
(2) Déterminer les éléments inversibles de $A \times B$ en fonction des éléments inversibles de A et de B.

23 On suppose que $A = \frac{\mathbb{Z}}{m\mathbb{Z}}$ et $B = \frac{\mathbb{Z}}{n\mathbb{Z}}$ où m et n sont deux entiers premiers entre eux.

(1) Montrer que l'application f ci-dessous est bien définie :

$$f : \begin{array}{ccc} \mathbb{Z}/mn\mathbb{Z} & \longrightarrow & \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \\ \mathbf{cl}_{mn}(a) & \longmapsto & (\mathbf{cl}_m(a), \mathbf{cl}_n(a)) \end{array}$$

où on utilise la notation $\mathbf{cl}_s(a)$ pour représenter la classe de a dans $\mathbb{Z}/s\mathbb{Z}$.

(2) Montrer que f est un isomorphisme d'anneau : c'est le théorème chinois.
(3) En déduire que lorsque m et n sont premiers entre eux on a $\phi(mn) = \phi(m)\phi(n)$ (Théorème 5).

24 Donner une formule générale pour le calcul de $\phi(m)$, m entier positif quelconque.

2.2. **Un problème de calendrier.** Trois professeurs commencent leur cours respectivement le lundi, le mardi et le jeudi. Le premier fait un cours tous les deux jours, le second tous les trois jours, le dernier tous les cinq jours. Le règlement de l'école exige que les cours qui tombent un dimanche soient annulés. Quand, pour la première fois, les trois professeurs auront à supprimer leur cours le même jour ?

3. Introduction à la théorie du Cryptage (chiffrage)

3.1. **Une courte description du problème.** Un individu A_1 envoie des messages $m \in M$ à son allié A_2 éloigné de lui. Le contenu de ce message ne doit pas être connu de leur ennemi commun X lequel cherche au contraire à intercepter le message. A_1 crypte son message m, c'est-à-dire le transforme en $m' = E(m)$ et envoie m' à A_2 qui connaît la méthode de cryptage utilisée (la méthode E) et est capable de décrypter c'est-à-dire de retrouver $m = D(m')$ à partir de m'. En pratique, aujourd'hui le plus souvent, le cryptage et décryptage sont effectués par un calculateur électronique. Il faut donc choisir une méthode de cryptage que puisse gérer un ordinateur.

Le problème est formalisé de la manière suivante.

(1) M est un ensemble (fini mais très grand) de messages m formés à partir de lettres l appartenant un alphabet Δ;
(2) L'application de cryptage E, que j'appellerai le *chiffreur*, est une application *injective* de M dans M, l'application de décryptage D, le *déchiffreur*, est définie sur $E(M)$ par $(D \circ E)(m) = m$ pour tout $m \in M$, c'est simplement la réciproque de la bijection $E : M \to E(M)$.

Pour être acceptable une méthode de cryptage doit posséder la propriété suivante : *il n'est pas possible de déterminer E (et D) connaissant un petit nombre de $E(m)$*. Plus précisément, le temps qu'il faut à X pour découvrir la méthode de cryptage doit être supérieur au temps pendant lequel le message doit demeurer secret.

La méthode courante (même si elle limite la complexité du cryptage) consiste à crypter les lettres de l'alphabet Δ et d'utiliser sur M le cryptage induit.

Voici un exemple classique.

Exemple 1.

(1) Δ est l'alphabet latin ordinaire composé de 26 lettres et M est l'ensemble des textes (disons de moins de N caractères) sans ponctuation et sans espace.
(2) A chaque lettre on associe un élément de $\mathbb{Z}/26\mathbb{Z}$ en suivant l'ordre alphabétique de sorte $a = \bar{0}$, $b = \bar{1}$, $c = \bar{2}$, ... , $z = \bar{25}$.
(3) Le cryptage $E = E_k$ est défini par

$$E_k(\text{lettre}) = \text{lettre} + \bar{k}.$$

25

(1) Crypter par E_2 le message PLEINELUNE.
(2) Décrypter le message YHQLYLGLYLFL sachant qu'il a été crypté par E_3.
(3) Montrer que si X sait (par exemple grâce à un espion) que A_1 utilise un seul chiffreur de type E_k alors la connaissance d'un seul $E(l)$, $l \in \Delta$, permet de déterminer E.
(4) Connaissez-vous la méthode classique pour découvrir E lorsque A_1 n'utilise qu'une méthode de cryptage ?

3.2. **Exemple.** Lorsqu'on travaille avec une famille de chiffreurs E_λ, comme dans l'exemple précédent, on dit que λ est une *clé* du cryptage E. (Il est parfois coûteux en temps de déterminer le déchiffreur correspondant $D_{\lambda'}$, dans ce cas une clé est un couple (λ,λ').)

Exemple 2. Soit $\Delta = \mathbb{Z}/n\mathbb{Z}$. Pour tout $(a,b) \in (\mathbb{Z}/n\mathbb{Z})^\star \times \mathbb{Z}/n\mathbb{Z}$ on définit le chiffreur

$$E_{a,b} : x \in \mathbb{Z}/n\mathbb{Z} \longrightarrow ax + b \in \mathbb{Z}/n\mathbb{Z}.$$

26 On prend $n = 641$. Déterminer le déchiffreur associé à $E_{\bar{41},\bar{26}}$.

27 Montrer que, en général, si Y réussit à obtenir deux couples (m,m') chiffrés à l'aide d'un $E_{a,b}$ alors le cryptage est mis à jour.

3.3. **Le système RSA.** Pour utiliser un cryptage du type précédent les deux alliés A_1 et A_2 doivent d'abord convenir d'une clé. On dit alors qu'on a affaire un cryptage à *clés privées*. Ce système à plusieurs inconvénients spécialement quand il est appliqué à l'échange de données sur internet. D'abord il faut que la clé passe de A_1 à A_2 et si elle est interceptée tous les messages qui l'utiliseront seront "ouverts". Le second inconvénient, plus ennuyeux, est que la gestion des clés devient difficile si l'échange de messages concerne N personnes avec N grand car alors il faut convenir d'une clé particulière pour chaque paire de communicants (A_i,A_j), soit $(N(N-1))/2$ clés. Le système RSA breveté en 1977 par Ron Rivest, Adi Shamir et Leon Adleman utilise le concept, à première vue paradoxal de *clé publique*. La clé du codage, comme nous allons le voir est publiée, automatiquement envoyée à qui la demande, par A_2. En réalité, il faudrait parler de clé semi-publique car A_1 reçoit non pas la clé entière mais un "morceau" de la clé suffisant pour effectuer le codage mais non le décodage (du moins en un laps de temps limité). On dit que la partie transmise est la clé publique et la partie non transmise est la clé privée.

Algorithme 2 (Description des cryptage et décryptage RSA)**.**

Step 2.1 (Demande). A_1 signale à A_2 qu'il à une communication à effectuer.

Step 2.2 (Génération de clé privée et fourniture de clé publique par A_2). A_2 sélectionne 2 nombres premiers p et q et un entier e premier avec $p-1$ et avec $q-1$.

Ensuite A_2 calcule l'inverse d de e modulo $(p-1)(q-1)$. La clé privée que seul A_2 connaît est la paire $(d,\{p,q\})$. La clé publique que A_2 communique à A_1 et que chacun est libre d'intercepter est la paire (e,n) où $n = pq$.

Clés	
publique	(e,n)
privée	$(d,\{p,q\})$

Relations
$n = pq$, p, q, premiers
e premier avec p et q
$ed = 1 \mod (p-1)(q-1)$

Step 2.3 (Cryptage par A_1 à l'aide de la clé publique). A_1 transforme son message (numérisé) $m = a_0a_1 \dots a_k$ avec $a_i \in 0, \dots, n-1$ en

$$m' = E^{RSA}_{e,n}(m) := (a_0^e \mod n)(a_1^e \mod n) \dots (a_k^e \mod n).$$

Step 2.4 (Décryptage par A_2 à l'aide la clé privée). Recevant $m' = b_0b_1 \dots b_k$, A_2 récupère $m = D(m')$ par

$$m = D^{RSA}_{e,\{p,q\}} = (b_0^d \mod n)(b_1^d \mod n) \dots (b_k^d \mod n).$$

Pour que le cryptage RSA soit sûr, il faut que X ne soit pas capable d'obtenir p et q en un temps raisonnable à partir de n qui est public. Pour que cette propriété soit vérifiée on prend en général des nombres premiers très grands, disons de l'ordre de 10^{150}. On peut calculer très rapidement n à partir de p et q (travail de A_2) par contre le travail de X l'intercepteur qui consiste à factoriser n de l'ordre 10^{300} est très long. Remarquer que la connaissance de p et q est nécessaire pour avoir d qui est essentiel dans la phase déchiffrage. Dans les exemples qui suivent on prendre des valeurs de p et q petites pour comprendre le fonctionnement du cryptage RSA mais il faut bien avoir avoir à l'esprit que ces valeurs ne sont pas utilisables en pratique : le cryptage associé est trop facile à percer.

28 Crypter le message 367 avec la clé publique (7,55).

29 Décrypter le message 784 avec la clé privée (23,{5,11}) correspondante à la clé publique (7,55).

30 Montrer que le cryptage RSA est en réalité bien défini sur un alphabet Δ strictement plus petit que $\mathbb{Z}/n\mathbb{Z}$. Montrer cependant, par un raisonnement probabiliste, que ce problème peut être négligé.

31 Expliquer comment obtenir d à partir de e en utilisant l'algorithme d'Euclide.

4. Introduction à la théorie des codes

4.1. **Introduction et définition des notions fondamentales.** Supposons que A_1 doive envoyer (à l'aide d'une quelconque technologie) un message — ici on parle plutôt de *mot* — contenant 5 signaux, disons,

$$m \in \{\text{ROUGE, VERTE, NOIRE}\}.$$

A cause de problèmes de brouillage, de parasite, d'insuffisance du matériel, chaque signal du mot est susceptible d'être modifié. Si on sait qu'il y aura au maximum deux erreurs et si A_2 reçoit $m' = \text{MORTE}$ alors il est sûr que le mot envoyé était VERTE. La raison est que MORTE a seulement deux signaux en commun avec ROUGE et NOIRE donc si ROUGE (ou NOIRE) avait été transformé en MORTE, il y aurait eu 3 erreurs dans la transmission ce qui est contraire à l'hypothèse. Par contre si l'ensemble de mots de départ est {ROUGE, FOULE, POULE} et que A_2 reçoit SOUPE alors A_2 ne peut pas déterminer le mot de départ car, en acceptant deux erreurs, aussi bien ROUGE, FOULE que POULE peuvent se transformer en SOUPE. L'objet principal de la théorie des codes est de construire des ensembles de mots qui ont

les avantages (contradictoires) d'être petits (en cardinal) et de d'autoriser le plus grand nombre d'erreurs possibles. Cette propriété sera précisée plus bas. Les mots seront composés de nombres plutôt que de lettres. En pratique, parce que c'est ce qui est utile en électronique, l'alphabet est $F_2 := \mathbb{Z}/2\mathbb{Z}$ et l'ensemble C des mots est un sous-ensemble du F_2-espace vectoriel F_2^n qui est l'ensemble de tous les F_2-mots de longueur n. On pourrait considérer plus généralement l'alphabet $F_p = \mathbb{Z}/p\mathbb{Z}$, p premier et le F_p-espace vectoriel F_p^n.

Definition 3. Tout ensemble de mots $C \subset F_2^n$ est appelé un CODE (d'ordre n sur F_2). Lorsque C est un sous-espace vectoriel de F_2^n on parle de CODE LINÉAIRE.

32 Soit $C = \{00110, 10011, 01101, 11000\} \subset F_2^5$. Présenter dans un tableau tous les mots obtenus en autorisant au plus une erreur et en déduire que C permet de corriger (au plus) une erreur. S'agit-il d'un code linéaire ?

Definition 4 (Distance de Hamming)**.** Soient a et b deux F_2-mots de longueur n. La distance de Hamming $d(a,b)$ est définie par

$$d(a,b) := \sharp\{i \in \{1,\dots,n\} \,:\, a_i \neq b_i\}.$$

où $\sharp$ est utilisé pour désigner le cardinal. Autrement dit, $d(a,b)$ est le nombre de position en lesquelles les lettres de a et b diffèrent.

En pratique, il ne suffit pas de détecter l'erreur mais il faut pouvoir retrouver le message de départ à partir du message erroné reçu. On parle alors de CODE CORRECTEUR.

Théorème 6. *Un code C permet de* détecter *la présence d'un nombre d'erreurs k si pour tous $(a,b) \in C^2$ on a $d(a,b) \geq k+1$.*

Démonstration. En effet si le mot m est transformé en m' avec $1 \leq l \leq k$ erreurs, le mot $m' \notin C$ et on peut donc en déduire la présence d'au moins une erreur. □

Théorème 7. *Un code C permet de* corriger *l erreurs, $1 \leq l \leq k$, si pour tous $(a,b) \in C^2$ on a $d(a,b) \geq 2k+1$.*

33 Justifier le théorème.

4.2. **Le F_2-code de longueur 7 de Hamming.** Soit M_H la matrice à coefficients dans F_2 (on écrit 0 pour $\bar{0}$ et 1 pour $\bar{1}$).

$$M_H := \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Théorème 8. *Soit $C_H = \ker M_H \subset F_2^7$. Alors C_H est un code linéaire capable de détecter jusqu'à deux erreurs et capable de corriger une erreur.*

34 Démontrer le théorème.