COURS D'ARITHMÉTIQUE

LM220

Alain Kraus

Université de Paris VI

2006/07

Table des matières

Introduction

L'objectif de ce cours est de présenter les concepts de base de l'arithmétique, de la théorie des corps finis, d'en déduire quelques applications à la cryptographie, et de donner une introduction à la théorie des codes correcteurs d'erreurs.

On ne se préoccupera pas de la construction de l'ensemble $\mathbb{N}$ des entiers naturels, ni de celle de l'ensemble $\mathbb{Z}$ des entiers relatifs. Nous admettrons donc que ces ensembles existent, et qu'ils sont munis de la relation d'ordre et des lois de composition usuelles que le lecteur connaît depuis longtemps. Le début du cours est consacré à l'étude de la divisibilité dans $\mathbb{Z}$, qui est le point de départ de l'arithmétique, et à une introduction à la théorie des groupes finis. Les premiers résultats de cette théorie sont indispensables dans la plupart des applications arithmétiques ultérieures. Un accent particulier est mis sur l'étude des groupes abéliens, pour lesquels on définit la notion « non évidente » de groupe quotient. On l'étend ensuite au cas des anneaux commutatifs, en vue d'étudier les quotients de $\mathbb{Z}$, les quotients des anneaux de polynômes, et notamment les corps finis. On aborde à la fin la notion de code correcteur, en particulier celle de code linéaire sur un corps fini.

J'ai développé à certains endroits, en complément, quelques résultats pour les étudiants qui pourraient être intéressés. Les parties concernées seront précisées pendant le semestre. À titre indicatif, je signale les ouvrages suivants comme bibliographie complémentaire du cours :

1) X. Buff, J. Garnier, E. Halberstadt, T. Lachand-Robert, F. Moulin, J. Sauloy, Mathématiques, Tout-en-un pour la Licence, Niveau L1, Sous la direction de J.-P. Ramis et A. Warusfel, Dunod, 2006.
2) M. Demazure, Cours d'algèbre, primalité divisibilité codes, Nouvelle bibliothèque mathématique, Cassini, 1997.
3) J. Dixmier, Cours de mathématiques du premier cycle, 1^{er} année, gauthier-villars, deuxième édition, 1976.
4) R. Godement, Cours d'algèbre, enseignement des sciences, Hermann, troisième édition, 1980.

Chapitre I — Arithmétique sur $\mathbb{Z}$

Soient $\mathbb{N}$ l'ensemble des entiers naturels et $\mathbb{Z}$ l'ensemble des entiers relatifs. On dispose sur $\mathbb{Z}$ des trois lois de composition[1], qui à tout couple (x, y) de $\mathbb{Z} \times \mathbb{Z}$ associent la somme $x+y$, la différence $x-y$ et le produit xy. Nous noterons comme il est d'usage $\leq$ la relation d'ordre[2] usuelle sur $\mathbb{Z}$. Pour tous $x, y \in \mathbb{Z}$, si l'on a $x \leq y$, on dit que x est plus petit que y, ou que x est inférieur à y. La relation $x \leq y$ s'écrit aussi $y \geq x$. Si l'on a $x \leq y$ avec $x \neq y$, on écrira parfois que l'on a $x < y$ ou bien $y > x$. Cette relation d'ordre, induite sur $\mathbb{N}$, munit $\mathbb{N}$ d'une structure d'ensemble ordonné, pour laquelle la propriété fondamentale suivante est vérifiée :

Propriété fondamentale. *Toute partie non vide de $\mathbb{N}$ a un plus petit élément*[3].

On l'utilisera à de nombreuses reprises.

1. Division euclidienne

Pour tout $x \in \mathbb{Z}$, on note $|x|$ le plus grand des entiers $-x$ et x.

Théorème 1.1. (Division euclidienne) *Soient a et b deux éléments de $\mathbb{Z}$ avec $b \neq 0$. Il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{Z}$ tel que l'on ait*

$$a = bq + r \quad et \quad 0 \leq r < |b|.$$

On dit que q est le quotient et que r est le reste de la division euclidienne de a par b.

Démonstration : 1) Démontrons l'assertion d'unicité. Supposons pour cela qu'il existe deux couples (q, r) et (q', r') d'entiers relatifs tels que l'on ait

$$a = bq + r = bq' + r' \quad \text{avec} \quad 0 \leq r < |b| \quad \text{et} \quad 0 \leq r' < |b|.$$

---

[1] Une loi de composition sur un ensemble E est une application du produit cartésien $E \times E$ à valeurs dans E. Combien existe-t-il de lois de composition sur un ensemble fini de cardinal n ?

[2] Une relation d'ordre sur un ensemble E est une relation binaire $\mathcal{R}$ sur E telle que pour tous x, y et z dans E, les conditions suivantes soient remplies :

1) on a $x\mathcal{R}x$ (réflexivité).
2) Si $x\mathcal{R}y$ et $y\mathcal{R}x$, alors $x = y$ (antisymétrie).
3) Si $x\mathcal{R}y$ et $y\mathcal{R}z$, alors $x\mathcal{R}z$ (transitivité).

Une relation d'ordre est souvent notée $\leq$ par commodité.

[3] Soient $\leq$ une relation d'ordre sur un ensemble E et F une partie de E. Un élément $a \in E$ est appelé plus petit élément de F, si a appartient à F et si pour tout $x \in F$, on a $a \leq x$. D'après la propriété d'antisymétrie, s'il existe un plus petit élément de F, il est unique. On parle alors du plus petit élément de F. Par exemple, 0 est le plus petit élément de $\mathbb{N}$. Bien entendu, un tel élément n'existe pas toujours. À titre indicatif, $\mathbb{Z}$ n'a pas de plus petit élément. Il en est de même de l'intervalle $]0, 1]$ dans l'ensemble des nombres réels muni de sa relation d'ordre usuelle.

On a l'égalité

$$|q - q'||b| = |r' - r|. \tag{1}$$

Par ailleurs, on a les inégalités

$$-|b| < -r \leq 0.$$

Puisque r et r' sont positifs, on a donc $-|b| < r' - r < |b|$, autrement dit, on a

$$|r - r'| < |b|.$$

Il résulte alors de l'égalité (1) que l'on a $|q - q'| < 1$, d'où $q = q'$ puis $r = r'$, ce qui établit l'unicité.

2) Démontrons l'assertion d'existence. Considérons pour cela l'ensemble

$$A = \Big\{a - bk \mid k \in \mathbb{Z}\Big\} \cap \mathbb{N}.$$

Vérifions que A n'est pas vide. Tel est le cas si $a \geq 0$, car dans ce cas a est dans A (on prend $k = 0$). Supposons $a < 0$. Si $b \geq 1$, on constate que $a(1-b) \in A$ (prendre $k = a$) et si $b \leq -1$, alors $a(1 + b) \in A$ (prendre $k = -a$). D'après la propriété fondamentale satisfaite par $\mathbb{N}$, l'ensemble A possède donc un plus petit élément r. Puisque r appartient à A, on a $r \geq 0$ et il existe $q \in \mathbb{Z}$ tel que l'on ait $a - bq = r$. Il reste à vérifier que l'on a $r < |b|$. Supposons le contraire. On a alors

$$0 \leq r - |b| = a - b(q + \varepsilon) \in A \quad \text{avec} \quad \varepsilon = \pm 1.$$

L'inégalité $r - |b| < r$ contredit alors le caractère minimal de r. D'où le résultat.

Définition 1.1. *Soient a et b deux éléments de $\mathbb{Z}$. On dit que b divise a ou que b est un diviseur de a, ou bien encore que a est un multiple de b (dans $\mathbb{Z}$) s'il existe $k \in \mathbb{Z}$ tel que $a = bk$. Si b est non nul, cette condition signifie que le reste de la division euclidienne de a par b est nul.*

Exercice 1.

1) Quels sont le quotient et le reste de la division euclidienne de 56798 par 23 ?
2) Démontrer que 14443 est divisible par 101.
3) Soient a et n deux entiers naturels non nuls. Montrer que l'on a l'égalité

$$a^n - 1 = (a - 1)(a^{n-1} + a^{n-2} + \cdots + a + 1).$$

En particulier, $a - 1$ divise $a^n - 1$.

4) Déterminer tous les entiers naturels n tels que $n + 1$ divise $n^2 + 1$.

2. Nombres premiers

Définition 1.2. *On appelle nombre premier tout entier $p \geq 2$ dont les seuls diviseurs positifs sont 1 et p.*

Par exemple 2, 3, 5, 7, 11, 13, $\cdots$ sont des nombres premiers.

Lemme 1.1. *Soit p un entier ≥ 2. Alors, p est premier si et seulement si p n'est pas le produit de deux entiers strictement plus grands que 1.*

Démonstration : Si l'on a $p = ab$ avec a et b strictement plus grands que 1, alors a divise p et a est distinct de 1 et p, autrement dit, p n'est pas premier. Inversement, si p n'est pas premier, il possède un diviseur positif a autre que 1 et p. On a alors $p = ab$, où a et b sont ≥ 2.

Théorème 1.2. *Tout entier $n \geq 2$ est un produit de nombres premiers. En particulier, tout entier $n \geq 2$ possède un diviseur premier.*

Démonstration : On procède par récurrence[4] sur n. Notons $P(n)$ la propriété : n est un produit de nombres premiers. D'abord $P(2)$ est vraie, car 2 est premier. Considérons alors un entier $n \geq 3$ tel que $P(k)$ soit vraie pour tout entier k tel que $2 \leq k < n$. Il s'agit de démontrer que $P(n)$ est vraie. Tel est le cas si n est premier. Si n n'est pas premier, il existe deux entiers a et b strictement plus grands que 1 tels que $n = ab$ (lemme 1.1). Puisque l'on a $2 \leq a < n$ et $2 \leq b < n$, les propriétés $P(a)$ et $P(b)$ sont vraies, ce qui entraîne le résultat.

Notation. On notera désormais $\mathbf{P}$ l'ensemble des nombres premiers.

Le résultat suivant est dû à Euclide qui vécut au III^e siècle avant J. C. :

Théorème 1.3. *L'ensemble $\mathbf{P}$ est infini.*

Démonstration : Supposons que $\mathbf{P}$ soit fini de cardinal n. Soient $p_1, \cdots, p_n$ ses éléments. Posons $N = 1 + p_1 \cdots p_n$. On a $N \geq 2$, donc N possède un diviseur premier p. L'entier p divise $p_1 \cdots p_n$, d'où l'on déduit que p divise 1, ce qui conduit à une contradiction.

---

[4] Rappelons le principe du raisonnement par récurrence. Soit n_0 un entier naturel. Il s'agit de démontrer qu'une certaine propriété $P(n)$ de l'entier n est vraie pour tout $n \geq n_0$. Supposons vérifiées les deux assertions suivantes :

1) la propriété $P(n_0)$ est vraie.
2) Pour tout entier $n \geq n_0$, si la propriété $P(n)$ est vraie, alors $P(n+1)$ l'est aussi.

Sous ces hypothèse, la propriété $P(n)$ est vraie pour tout entier $n \geq n_0$.

Une variante consiste à remplacer la deuxième condition par la suivante, qui est parfois plus facile à utiliser, et qui, avec la première condition, conduit à la même conclusion :

2') Pour tout entier $n > n_0$, si la propriéte $P(k)$ est vraie pour tout k tel que $n_0 \leq k < n$, alors $P(n)$ l'est aussi.

Théorème 1.4. (Lemme d'Euclide) *Soient a, b deux entiers relatifs et p un nombre premier tels que p divise ab. Alors, p divise l'un des entiers a et b.*

Démonstration. La démonstration qui suit est due à Gauss[5]. Supposons que p ne divise pas a. Il s'agit de montrer que p divise b. Considérons pour cela l'ensemble

$$A = \big\{n \geq 1 \mid p \text{ divise } an\big\}.$$

Il est non vide, car par exemple p appartient à A. Soit m le plus petit élément de A. D'après l'hypothèse faite sur a, on a l'inégalité

$$m \geq 2. \tag{2}$$

Soit n un élément de A. Vérifions que m divise n. D'après le théorème de la division euclidienne, il existe deux entiers q et r tels que l'on ait $n = mq + r$ avec $0 \leq r < m$. On a l'égalité $an - (am)q = ar$, d'où l'on déduit que p divise ar (car n et m sont dans A). Puisque l'on a $r < m$, r n'est pas dans A, d'où $r = 0$ et notre assertion. Les entiers p et b étant dans A, il en résulte que m divise p et b. L'inégalité (2) et le fait que p soit premier entraînent alors $p = m$. Par suite, p divise b.

Corollaire 1.1. *Si un nombre premier divise un produit d'entiers relatifs, il divise l'un de ces entiers. En particulier, si un nombre premier divise un produit de nombres premiers, il est égal à l'un d'eux.*

Démonstration : C'est une conséquence directe du théorème 1.4, en procédant par récurrence sur le nombre de facteurs du produit (exercice).

Le théorème suivant s'appelle parfois le théorème fondamental de l'arithmétique :

Théorème 1.5. *Tout entier $n \geq 2$ s'écrit de façon unique sous la forme*

$$n = p_1^{n_1} \cdots p_r^{n_r}, \tag{3}$$

où les n_i sont des entiers naturels non nuls, et où les p_i sont des nombres premiers vérifiant $p_{i-1} < p_i$ pour tout $i = 2, \cdots, r$. On dit que l'égalité (3) *est la décomposition de n en produit de nombres premiers.*

[5] Carl Friedrich Gauss, surnommé le prince des mathématiciens, est né à Brunswick en 1777 et décède à Göttingen en 1855. On lui doit une quantité massive de résultats en arithmétique, ainsi que dans d'autres domaines. Son ouvrage, Disquisitiones Arithmeticae, est resté célèbre en théorie des nombres. On pourra trouver les arguments de la démonstration du théorème 1.4 à la page 6 de ce livre. À dix ans, le maître d'école lui demanda de calculer la somme des cent premiers entiers naturels. Il donna de façon surprenante la réponse très rapidement, à savoir $50 \times 101 = 5050$. Quelle formule avait-il utilisée ?

Démonstration : L'assertion d'existence provient du théorème 1.2 en regroupant les facteurs égaux par ordre croissant. Prouvons l'assertion d'unicité. Supposons que l'on ait

$$n = p_1^{n_1} \cdots p_r^{n_r} = q_1^{m_1} \cdots q_s^{m_s},$$

où les p_i et q_i sont premiers tels que $p_1 < \cdots < p_r$, $q_1 < \cdots, q_s$ et où les n_i et m_i sont des entiers naturels non nuls. On déduit du corollaire 1.1 que l'on a

$$\{p_1, \cdots, p_r\} = \{q_1, \cdots, q_s\}.$$

Par suite, on a $r = s$. De plus, p_1 est le plus petit élément de $\{p_1, \cdots, p_r\}$ et q_1 est le plus petit élément de $\{q_1, \cdots, q_r\}$, d'où $p_1 = q_1$, puis $p_i = q_i$ pour tout i. Par ailleurs, s'il existe un indice i tel que $n_i \neq m_i$, par exemple $n_i < m_i$, alors p_i divise 1 ou bien un produit de nombres premiers tous distincts de lui même, ce qui contredit le corollaire 1.1 et établit le résultat.

Exercice 2. (Petit théorème de Fermat[6]) Soient a un entier ≥ 1 et p un nombre premier.

1) Soit k un entier tel que $1 \leq k \leq p-1$. Montrer que p divise le coefficient binomial C_p^k.
2) En déduire, par récurrence sur a, que p divise $a^p - a$.

On démontrera ce résultat de façon plus conceptuelle au chapitre IV.

Une problème naturel qui se pose est le suivant :

Problème. *Soit n un entier ≥ 2. Comment décider si n est un nombre premier ou non ?*

Il existe de nombreux tests permettant parfois de reconnaître si un entier n est premier. Nous n'aborderons pas cette étude dans ce cours. C'est la théorie des tests de primalité. Signalons seulement à ce sujet le résultat ci-dessous :

[6] Pierre de Fermat est né près de Toulouse en 1601 et mourut à Castres en 1665. Bien qu'il consacra une partie de sa carrière à sa fonction de conseiller à la Cour de Toulouse, il restera comme l'un des grands mathématiciens de son temps, notamment pour ses travaux en théorie de nombres et en probabilité. Il existe aussi un « grand théorème de Fermat », qui en réalité n'est devenu un théorème qu'en 1994. Il s'agit de l'énoncé suivant : pour tout entier $n \geq 3$, il n'existe pas d'entiers relatifs x, y et z tels que $x^n + y^n = z^n$ avec $xyz \neq 0$. L'entier $n = 2$ doit évidemment être exclu vu que pour tous a et b dans $\mathbb{Z}$, on a l'égalité $(a^2 - b^2)^2 + (2ab)^2 = (a^2 + b^2)^2$, ce qui géométriquement signifie qu'il existe une infinité de triangles rectangles dont les longueurs des côtés sont des entiers. La recherche d'une démonstration, ne serait-ce que pour des valeurs particulières de l'exposant n, a par exemple donné naissance à la notion d'idéal d'un anneau, puis à toute la théorie algébrique des nombres.

Lemme 1.2. *Soit n un entier ≥ 2. Si n n'est pas premier, alors n possède un diviseur premier p vérifiant l'inégalité $p^2 \leq n$.*

Démonstration : Si n n'est pas premier, il existe deux entiers a et b strictement plus grands que 1 tels que $n = ab$ (lemme 1.1). Supposons par exemple $a \leq b$. Puisque $a \geq 2$, a possède un diviseur premier p (th. 1.2). En particulier, p divise n et l'on a $p^2 \leq ap \leq ab = n$.

En utilisant ce résultat, on constate par exemple que 641 est premier. En effet, s'il ne l'était pas, il devrait exister un nombre premier $p < 25$ divisant 641. Les nombres premiers plus petits que 25 sont $2, 3, 5, 7, 11, 13, 17, 19$ et 23. On verifie alors qu'aucun de ces nombres ne divise 641 en utilisant le théorème de la division euclidienne.

Étant donné un entier N, il existe un procédé de criblage, appelé crible d'Ératosthène (il vécut au III^e siècle), qui permet de déterminer tous les nombres premiers inférieurs à N, en utilisant seulement l'opération de multiplication, et pas celle de division, ce qui est plus facile. Son principe est le suivant. On écrit d'abord dans un tableau tous les entiers jusqu'à N. On raye ensuite tous les multiples de 2, autres que 2, puis tous les multiples de 3, autres que 3, etc, autrement dit, à chaque étape on raye tous les multiples du plus petit entier qui n'a pas encore été rayé. Pour démontrer que N est premier, si tel est le cas, il suffit d'après le lemme 1.2 de cribler tous les entiers plus petits que $\sqrt{N}$. Par exemple, en rayant tous les multiples de 2, 3, 5 et 7, on constate que 101 est premier.

Remarquons par ailleurs que le petit théorème de Fermat permet parfois de démontrer qu'un entier $n \geq 2$ n'est pas premier, si tel est le cas. Compte tenu de ce théorème, il suffit en effet d'expliciter un entier naturel a tel que n ne divise pas $a^n - a$. Cela étant, il existe des entiers n non premiers pour lesquels quel que soit $a \in \mathbb{Z}$, l'entier $a^n - a$ est divisible par n. Ces entiers s'appellent les nombres de Carmichael (1879-1967). Tel est par exemple le cas de $n = 561$ (c'est le plus petit) et de $n = 1105$ (c'est le suivant). On sait par ailleurs démontrer, depuis 1992, qu'il existe une infinité de nombres de Carmichael. La démonstration de ce résultat, qui dépasse de loin le niveau de ce cours, utilise la théorie analytique des nombres. Signalons qu'il est néanmoins assez facile de prouver que pour tout entier $n \geq 1$, si les trois nombres $p = 6n+1$, $q = 12n+1$ et $r = 18n+1$ sont premiers, alors pqr est un nombre de Carmichael. Il en est ainsi avec $n = 1$, auquel cas $pqr = 1729$. La question de savoir s'il existe une infinité de tels entiers n est ouverte.

Exercice 3. Montrer que 3571 est un nombre premier (c'est le cinq centième nombre premier).

Exercice 4. Soit n un entier ≥ 1.

1) Montrer que si $2^n - 1$ est un nombre premier, il en est de même de n. Un nombre de la forme $2^n - 1$, avec n premier, s'appelle un nombre de Mersenne (c'était un moine français qui vécut de 1588 à 1648). On connaît « beaucoup » de nombres premiers de

Mersenne (en fait une quarantaine),

$$2^2-1=3,\ 2^3-1=7,\ 2^5-1=31,\cdots,2^{19}-1,\cdots,2^{61}-1,\cdots,$$

mais on ne sait pas prouver qu'il en existe une infinité. On notera que $2^{11}-1=23\times 89$ n'est pas premier. On ne sait pas non plus démontrer l'existence d'une infinité de nombres de Mersenne qui ne sont pas premiers.

2) Montrer que si 2^n+1 est premier, alors n est une puissance de 2. Un nombre de la forme $2^{2^n}+1$ s'appelle un nombre de Fermat. On ne connaît que cinq nombres premiers de Fermat : $3, 5, 17, 257$ et $65537=2^{2^4}+1$. On conjecture qu'il n'en existe qu'un nombre fini. En fait les nombres de Fermat croissent rapidement, et il est difficile de décider s'ils sont premiers y compris pour des petites valeurs de n.
Prouvons ici que $2^{32}+1$ n'est pas premier en montrant qu'il est divisible par 641. La démonstration qui suit est due à Euler. Posons $p=641$. On écrit que l'on a

$$p=5^4+2^4 \quad \text{et} \quad p=5.2^7+1.$$

Il existe donc $k\in\mathbb{Z}$ tel que l'on ait

$$5^4.2^{28}=(p-1)^4=1+kp.$$

On en déduit l'égalité $(p-2^4)2^{28}=1+kp$ i.e. $p\big(2^{28}-k\big)=2^{32}+1$, d'où l'assertion. Cela étant, on ne sait toujours pas démontré l'existence d'une infinité de nombres de Fermat qui ne sont pas premiers.

Exercice 5. Déterminer tous les nombres premiers p tels que p divise 2^p+1 (utiliser le petit théorème de Fermat).

Exercice 6. Soit n un entier naturel. Posons $p=2n+1$. Démontrer que p est un nombre premier si et seulement si n ne figure pas dans le tableau infini suivant :

$$\begin{pmatrix} 4 & 7 & 10 & 13 & 16 & \dots \\ 7 & 12 & 17 & 22 & 27 & \dots \\ 10 & 17 & 24 & 31 & 38 & \dots \\ 13 & 22 & 31 & 40 & 49 & \dots \\ 16 & 27 & 38 & 49 & 60 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

dans lequel la première colonne est une suite arithmétique de premier terme 4 et de raison 3 et la k-ième ligne est une suite arithmétique de raison $2k+1$. (On vérifiera que le coefficient de la k-ième ligne et de la j-ième colonne de ce tableau est $2kj+k+j$).

Exercice 7. Démontrer que l'entier $1+2+2^2+2^3+\cdots+2^{26}$ n'est pas premier.

3. Valuation p-adique d'un entier relatif

On considère dans ce paragraphe l'ensemble $\mathbb{N} \cup \{+\infty\}$ obtenu en adjoignant à $\mathbb{N}$ un élément noté $+\infty$, que l'on munit de la structure d'ensemble ordonné qui induit l'ordre usuel sur $\mathbb{N}$ et telle que $+\infty \geq n$ pour tout entier naturel n. On prolonge par ailleurs la loi additive de $\mathbb{N}$ à cet ensemble en posant $(+\infty) + n = n + (+\infty) = +\infty$ et $(+\infty) + (+\infty) = +\infty$. Pour tout nombre premier p, on va définir ici une application, appelée valuation p-adique,

$$v_p : \mathbb{Z} \to \mathbb{N} \cup \{+\infty\}.$$

Définition 1.2. *Soient n un entier relatif et p un nombre premier.*

1) Si l'on a $n \geq 2$, alors $v_p(n)$ est l'exposant de p dans la décomposition de n en produit de nombres premiers. Autrement dit :

1.1) si p ne divise pas n, on a $v_p(n) = 0$.

1.2) Si $n = p_1^{n_1} \cdots p_r^{n_r}$ est la décomposition de n en produit de nombres premiers ($n_i \geq 1$), on a

$$v_{p_i}(n) = n_i \quad \textit{pour} \quad i = 1, \cdots, r.$$

2) On pose $v_p(0) = +\infty$ et $v_p(1) = 0$.

3) Pour tout $n \geq 1$, on pose $v_p(-n) = v_p(n)$.

On dit que $v_p(n)$ est la valuation p-adique de n.

Pour tout nombre premier p et tout entier $n \geq 1$, zéro est divisible par p^n, ce qui justifie l'égalité $v_p(0) = +\infty$.

Exemple 1.1. Posons $n = 539000$. On a $n = 2^3.5^3.7^2.11$, de sorte que l'on a $v_2(n) = 3$, $v_5(n) = 3$, $v_7(n) = 2$, $v_{11}(n) = 1$ et pour tout nombre premier p distinct de $2, 5, 7$ et 11, on a $v_p(n) = 0$.

Avec cette définition, le théorème 1.5 s'écrit comme suit :

Théorème 1.6. *Tout entier relatif n non nul s'écrit de manière unique, à l'ordre près des facteurs, sous la forme*

$$n = \varepsilon \prod_{p \in \mathbf{P}} p^{v_p(n)} \quad \textit{avec} \quad \varepsilon = \pm 1. \tag{4}$$

On a $\varepsilon = 1$ si $n \geq 1$ et $\varepsilon = -1$ si $n \leq -1$. Contrairement à ce que laisse supposer cette formule, il s'agit d'un produit fini car on a $v_p(n) = 0$ pour presque tout $p \in \mathbf{P}$ (au sens tous sauf un nombre fini). De plus, pour tout $n \in \mathbb{Z}$, on a l'équivalence

$$v_p(n) \geq 1 \Longleftrightarrow p \text{ divise } n. \tag{5}$$

Pour tous x et y dans $\mathbb{Z}$, on note dans la suite $\mathrm{Min}(x, y)$ le plus petit d'entre eux.

Proposition 1.1. *Soient a et b deux entiers relatifs et p un nombre premier.*

1) On a $v_p(ab) = v_p(a) + v_p(b)$.

2) On a $v_p(a+b) \geq \mathrm{Min}\big(v_p(a), v_p(b)\big)$. De plus, si $v_p(a) \neq v_p(b)$, alors on a l'égalité $v_p(a+b) = \mathrm{Min}\big(v_p(a), v_p(b)\big)$.

3) Pour que a divise b, il faut et il suffit que l'on ait $v_p(a) \leq v_p(b)$ pour tout $p \in \mathbf{P}$.

Démonstration : On vérifie directement que ces assertions sont vraies si $ab = 0$. Supposons donc $ab \neq 0$.

1) Il résulte du théorème 1.6 que l'on a les égalités

$$ab = \varepsilon \prod_{p \in \mathbf{P}} p^{v_p(ab)} = \varepsilon \prod_{p \in \mathbf{P}} p^{v_p(a)+v_p(b)} \quad \text{avec} \quad \varepsilon = \pm 1.$$

L'unicité de la décomposition d'un entier sous la forme (4) entraîne alors l'égalité annoncée.

2) ll existe des entiers r et s, qui ne sont pas divisibles par p, tels que l'on ait

$$a = p^{v_p(a)} r \quad \text{et} \quad b = p^{v_p(b)} s.$$

Supposons par exemple $v_p(a) \geq v_p(b)$. On a

$$a + b = p^{v_p(b)} \big(p^{v_p(a)-v_p(b)} r + s\big), \tag{6}$$

d'où l'on déduit, d'après la première assertion, que l'on a

$$v_p(a+b) = v_p(b) + v_p\big(p^{v_p(a)-v_p(b)} r + s\big) \geq v_p(b) = \mathrm{Min}\big(v_p(a), v_p(b)\big).$$

Supposons de plus $v_p(a) > v_p(b)$. Dans ce cas, p ne divise pas $p^{v_p(a)-v_p(b)} r + s$. D'après (6), cela conduit à l'égalité $v_p(a+b) = v_p(b)$.

3) Supposons que a divise b. Il existe $k \in \mathbb{Z}$ tel que $b = ak$. Pour tout nombre premier p, on a $v_p(b) = v_p(a) + v_p(k)$, d'où $v_p(b) \geq v_p(a)$. Inversement, d'après l'hypothèse faite, pour tout $p \in \mathbf{P}$ il existe $t_p \geq 0$ tel que l'on ait $v_p(b) = v_p(a) + t_p$. Pour presque tout p, on a $v_p(a) = v_p(b) = t_p = 0$. Posons

$$t = \prod_{p \in \mathbf{P}} p^{t_p}.$$

Pour tout $p \in \mathbf{P}$, on a donc $v_p(b) = v_p(at)$, d'où $b = \pm at$ et a divise b.

Exercice 8. Calculer $v_2(1056)$. Pour tout $p \in \mathbf{P}$ calculer $v_p(196000)$.

Exercice 9. Démontrer que $\sqrt{2}$ n'appartient pas à l'ensemble $\mathbb{Q}$ des nombres rationnels. Soit a un nombre rationnel strictement positif. Donner une condition nécessaire et suffisante pour que $\sqrt{a}$ appartienne à $\mathbb{Q}$.

Définition 1.3. *Soient a et b deux entiers relatifs. On dit que a et b sont premiers entre eux s'il n'existe pas de nombres premiers divisant à la fois a et b, autrement dit, si pour tout nombre premier p, on a $\mathrm{Min}\big(v_p(a), v_p(b)\big) = 0$. Dans ce cas, on dit aussi que a est premier avec b.*

Théorème 1.7. (Lemme de Gauss) *Soient a, b, c trois entiers relatifs tels que a divise bc et que a soit premier avec b. Alors, a divise c.*

Démonstration : Soit p un nombre premier. Compte tenu de l'assertion 3 de la proposition 1.1, il s'agit de démontrer que l'on a l'inégalité $v_p(a) \leq v_p(c)$. Elle est évidente si $v_p(a) = 0$. Supposons $v_p(a) \geq 1$ i.e. que p divise a. L'entier a étant premier avec b, on alors $v_p(b) = 0$. Puisque a divise bc, on a $v_p(a) \leq v_p(bc)$, d'où $v_p(a) \leq v_p(c)$ et le résultat.

Corollaire 1.2. *Soient a un entier relatif et r, s deux entiers premiers entre eux. Si a est divisible par r et s, alors a est divisible par rs.*

Démonstration : Il existe u et v dans $\mathbb{Z}$ tels que l'on ait les égalités $a = ur = vs$. D'après le lemme de Gauss, r divise donc v, d'où l'assertion.

Exercice 9. Soient p et q deux nombres premiers distincts. Montrer que pq divise $p^{q-1} + q^{p-1} - 1$.

4. Plus grand commun diviseur

On considère dans ce paragraphe deux entiers relatifs a et b non tous les deux nuls.

Théorème 1.8. *Il existe un unique entier $d > 0$ vérifiant les deux conditions suivantes :*

1) l'entier d est un diviseur commun à a et b.

2) Tout diviseur commun à a et b divise d.

On a l'égalité

$$d = \prod_{p \in \mathbf{P}} p^{\mathrm{Min}\left(v_p(a), v_p(b)\right)}. \tag{7}$$

Définition 1.4. *L'entier d défini par l'égalité* (7) *est appelé le plus grand commun diviseur de a et b, ou en abrégé le pgcd de a et b. On le note* $\mathrm{pgcd}(a, b)$ *ou* $a \wedge b$.

Démonstration : Considérons l'entier d défini par l'égalité (7) (d est bien défini car a et b ne sont pas tous les deux nuls). Pour tout nombre premier p, $v_p(a)$ et $v_p(b)$ sont plus grands que $\mathrm{Min}\big(v_p(a), v_p(b)\big)$, donc d est un diviseur commun à a et b (assertion 3 de la prop. 1.1). Par ailleurs, si c est un diviseur commun à a et b, alors pour tout nombre premier p on a $v_p(c) \leq v_p(a)$ et $v_p(c) \leq v_p(b)$, d'où $v_p(c) \leq \mathrm{Min}\big(v_p(a), v_p(b)\big)$, donc c divise d (*loc. cit.*). Ainsi d vérifie les conditions 1 et 2. Par ailleurs, si d' est un entier naturel non nul vérifiant ces conditions, alors d divise d' et d' divise d, d'où $d = d'$.

Lemme 1.3. *Les entiers $\frac{a}{d}$ et $\frac{b}{d}$ sont premiers entre eux.*

Démonstration : Pour tout $p \in \mathbf{P}$, $v_p(d)$ est égal à $v_p(a)$ ou $v_p(b)$. Par ailleurs, on a $v_p(a/d) = v_p(a) - v_p(d)$ et $v_p(b/d) = v_p(b) - v_p(d)$, donc le minimum de $v_p(a/d)$ et $v_p(b/d)$ est nul, d'où le lemme.

Lemme 1.4. *Les entiers a et b sont premiers entre eux si et seulement si leur pgcd est 1.*

Démonstration : Les entiers a et b sont premiers entre eux si et seulement si pour tout $p \in \mathbf{P}$, on a $\mathrm{Min}\big(v_p(a), v_p(b)\big) = 0$. D'après (7), cela est équivalent à l'égalité $\mathrm{pgcd}(a, b) = 1$.

Exercice 10. Déterminer le pgcd de 2800 et 120.

Exercice 11. Soient a et m deux entiers tels que $a \geq 2$ et $m \geq 1$. Montrer que l'on a

$$\mathrm{pgcd}\left(\frac{a^m - 1}{a - 1}, a - 1\right) = \mathrm{pgcd}(a - 1, m).$$

Exercice 12. Déterminer l'ensemble des couples $(x, y) \in \mathbb{Z}^2$ pour lesquels on a l'égalité $x + y - 1 = \mathrm{pgcd}(x, y)$.

Exercice 13. Soient a, b et c trois entiers relatifs non nuls. Montrer que l'on a

$$\big(a \wedge b\big) \wedge c = a \wedge (b \wedge c).$$

Théorème 1.9. (Théorème de Bézout[7]) *Il existe deux entiers relatifs u et v tels que l'on ait*

$$\mathrm{pgcd}(a, b) = au + bv.$$

Démonstration : On considère l'ensemble

$$A = \Big\{au + bv \mid u, v \in \mathbb{Z}\Big\} \cap \Big(\mathbb{N} - \big\{0\big\}\Big).$$

C'est une partie non vide de $\mathbb{N}$. Soit c son plus petit élément. On a $c \geq 1$. Vérifions que l'on a

$$A = \Big\{ck \mid k \geq 1\Big\}. \tag{8}$$

D'abord, c étant dans A, les éléments de la forme ck, avec $k \geq 1$, sont aussi dans A. Inversement, soit n un élément de A. D'après le théorème de la division euclidienne, il existe $q, r \in \mathbb{Z}$ tels que l'on ait $n = cq + r$ avec $0 \leq r < c$. Supposons $r \neq 0$. On a alors

[7] Étienne Bézout fut un mathématicien français qui vécut de 1730 à 1783. Il fut chargé de l'enseignement des élèves du corps de l'artillerie. Il publia une théorie générale des équations algébriques à Paris en 1779. Outre le théorème 1.9, un autre théorème célèbre porte son nom concernant l'intersection de deux « courbes algébriques ».

$r = n - cq \geq 1$. Les entiers n et c étant dans A, $n - cq$ est aussi de la forme $a\alpha + b\beta$ avec $\alpha, \beta \in \mathbb{Z}$, en particulier r appartient à A. Le caractère minimal de c conduit alors à une contradiction. On a donc $r = 0$, puis $n = cq$ avec $q \geq 1$. Cela établit l'égalité (8). Démontrons alors que l'on a

$$\text{pgcd}(a, b) = c, \tag{9}$$

ce qui entraînera le résultat. Si $ab \neq 0$ les entiers $|a|$ et $|b|$ sont dans A, et d'après (8), c divise donc a et b. On a la même conclusion si $ab = 0$. Ainsi, c est un diviseur commun à a et b. Par ailleurs, il existe u et v dans $\mathbb{Z}$ tels que $c = au + bv$, de sorte que tout diviseur commun à a et b divise c. L'égalité (9) en résulte, car c vérifie les deux conditions du théorème 1.8.

On en déduit l'énoncé suivant :

Corollaire 1.3. *Les entiers a et b sont premiers entre eux si et seulement si il existe u et v dans $\mathbb{Z}$ tels que l'on ait $1 = au + bv$.*

Exercice 14. Pour tout $n \in \mathbb{Z}$, montrer que les entiers $5n+2$ et $12n+5$ sont premiers entre eux.

Remarque 1.1. On peut généraliser la notion de pgcd au cas d'une famille finie d'entiers relatifs non tous nuls, et les résultats de ce paragraphe s'étendent à cette situation. Si $(x_i)_{1 \leq i \leq n}$ est une famille d'entiers relatifs non tous nuls, le pgcd des x_i est l'unique entier $d > 0$ tel que d divise tous les x_i et que tout diviseur commun aux x_i divise d. Pour tout nombre premier p, on vérifie que l'on a (avec la notation évidente)

$$v_p(d) = \text{Min}\big(v_p(x_1), \cdots, v_p(x_n)\big).$$

On démontre que d peut s'écrire sous la forme $d = u_1x_1 + \cdots + u_nx_n$ pour des entiers u_i convenablement choisis. On dit que les x_i sont premiers entre eux dans leur ensemble s'ils n'ont pas de diviseurs premiers communs, ce qui signifie que leur pgcd vaut 1. Il convient de noter que cela ne signifie pas qu'ils soient premiers entre eux deux à deux. En pratique, le calcul du pgcd d'une famille d'entiers se ramène à des calculs de pgcd de deux entiers. Par exemple, le pgcd de trois entiers non nuls a, b, c n'est autre que $(a \wedge b) \wedge c$.

5. L'algorithme d'Euclide

Considérons dans ce paragraphe deux entiers naturels a et b non nuls tels que $a \geq b$. On va détailler ici un algorithme, qui utilise seulement le théorème de la division euclidienne, permettant d'une part de déterminer le pgcd de a et b, et d'autre part d'expliciter une relation de Bézout entre a et b, autrement dit, de déterminer deux entiers relatifs u et v tels que l'on ait $\text{pgcd}(a, b) = au + bv$.

On construit pour cela une suite finie d'entiers naturels $(r_i)_{i\geq 0}$, que l'on appelle la suite des restes (associée à a et b), par le procédé suivant : on pose d'abord

$$r_0 = a \quad \text{et} \quad r_1 = b.$$

Supposons construits $r_0, r_1, \cdots, r_i$ où $i \geq 1$. Si $r_i \neq 0$, on définit alors r_{i+1} comme étant le reste de la division euclidienne de r_{i-1} par r_i. Si $r_i = 0$, le procédé s'arrête et la suite des restes est alors formée des entiers $r_0, r_1, \cdots, r_{i-1}, r_i = 0$. Il existe un unique indice $n \geq 1$ tel que la condition suivante soit satisfaite :

$$0 < r_n < r_{n-1} < \cdots < r_1 \leq r_0 \quad \text{et} \quad r_{n+1} = 0.$$

Proposition 1.2. *On a* $r_n = \text{pgcd}(a, b)$.

Démonstration : Soit i un entier tel que $1 \leq i \leq n$. Il existe $q_i \in \mathbb{Z}$ tel que l'on ait

$$r_{i-1} = q_i r_i + r_{i+1} \quad \text{avec} \quad 0 \leq r_{i+1} < r_i. \tag{10}$$

Il résulte directement du théorème 1.8 que l'on a

$$\text{pgcd}(r_{i-1}, r_i) = \text{pgcd}(r_i, r_{i+1}).$$

Par suite, on a $\text{pgcd}(a, b) = \text{pgcd}(r_0, r_1) = \text{pgcd}(r_1, r_2) = \cdots = \text{pgcd}(r_{n-1}, r_n) = r_n$.

On a ainsi démontré que le pgcd de a et b est le dernier reste non nul r_n dans la suite des restes que l'on a construite. Il existe donc u et v dans $\mathbb{Z}$ tels que l'on ait

$$r_n = au + bv.$$

Le problème qui nous intéresse maintenant est d'expliciter un tel couple (u, v). On construit pour cela deux suites d'entiers $(u_i)_{0\leq i\leq n}$ et $(v_i)_{0\leq i\leq n}$ en posant

$$u_0 = 1, \quad u_1 = 0 \quad \text{et} \quad v_0 = 0, \quad v_1 = 1,$$

$$u_{i+1} = u_{i-1} - u_i q_i \quad \text{et} \quad v_{i+1} = v_{i-1} - v_i q_i \quad \text{pour tout } i = 1, \cdots, n-1,$$

où q_i est défini par l'égalité (10), autrement dit, où q_i est le quotient de la division euclidienne de r_{i-1} par r_i.

Proposition 1.3. *On a* $r_n = au_n + bv_n$.

Démonstration : Il suffit de vérifier que pour tout i tel que $0 \leq i \leq n$, on a l'égalité $r_i = au_i + bv_i$. Elle est vraie si $i = 0$ et $i = 1$. Considérons un entier k vérifiant les inégalités $1 \leq k < n$ tel que l'on ait $r_i = au_i + bv_i$ pour tout $i \leq k$. On a alors

$$r_{k+1} = r_{k-1} - q_k r_k = (u_{k-1}a + v_{k-1}b) - q_k(u_k a + v_k b) = au_{k+1} + bv_{k+1},$$

d'où l'égalité annoncée.

Il peut être commode de présenter les étapes de calculs sous la forme du tableau suivant :

	q_1	q_2	$\cdots$	q_{n-1}	q_n
$r_0 = a$	$r_1 = b$	r_2	$\cdots$	r_{n-1}	r_n
1	0	u_2	$\cdots$	u_{n-1}	u_n
0	1	v_2	$\cdots$	v_{n-1}	v_n

Exemple 1.2.

Appliquons ce qui précède au calcul du pgcd des entiers $a = 17640$ et $b = 525$. On obtient le tableau :

	33	1	1	2
17640	525	315	210	105
1	0	1	-1	2
0	1	-33	34	-67

Ainsi $105 = \text{pgcd}(a, b)$ et l'on obtient la relation de Bézout

$$2 \times 17640 - 67 \times 525 = 105.$$

Bien entendu, on peut aussi expliciter les décompositions de a et b en produit de nombres premiers. On trouve $a = 2^3.3^2.5.7^2$ et $b = 3.5^2.7$, d'où $\text{pgcd}(a, b) = 3.5.7 = 105$ comme attendu (th. 1.8).

Exercice 15. Soit n un entier ≥ 1. Déterminer le pgcd de $9n + 4$ et $2n - 1$.

Exercice 16. Soient a et b deux entiers ≥ 1. Déterminer le pgcd de $2^a - 1$ et $2^b - 1$.

Exercice 17. Déterminer les entiers n de quatre chiffres tels que les restes des divisions euclidiennes de 21685 et 33509 par n soient respectivement 37 et 53.

6. L'équation $ax + by = c$

Considérons trois entiers relatifs non nuls a, b et c. On se propose ici de décrire l'ensemble S formé des couples $(x, y) \in \mathbb{Z}^2$ tels que l'on ait

$$ax + by = c. \tag{11}$$

Proposition 1.4. *Soit d le pgcd de a et b. Posons $a' = \frac{a}{d}$ et $b' = \frac{b}{d}$.*

1) L'ensemble S est non vide si et seulement si d divise c.

2) Supposons que d divise c. Soit (x_0, y_0) un élément de $\mathbb{Z}^2$ tel que $ax_0 + by_0 = c$. On a

$$S = \Big\{ \big(x_0 + kb', y_0 - ka'\big) \mid k \in \mathbb{Z} \Big\}.$$

Démonstration : 1) S'il existe des entiers x et y vérifiant (11), d doit diviser c vu que d est un diviseur de a et b. Inversement, supposons que d divise c. Il existe c' dans $\mathbb{Z}$ tel que $c = dc'$. Puisque a' et b'sont premiers entre eux (lemme 1.3), il existe u et v dans $\mathbb{Z}$ tels que $a'u + b'v = 1$ (cor. 1.3). On obtient alors l'égalité $c = a(c'u) + b(c'v)$, ce qui prouve que S n'est pas vide.

2) Soit (x, y) un élément de S. On a l'égalité

$$a'(x - x_0) = b'(y_0 - y).$$

Puisque a' est premier avec b', on déduit du lemme de Gauss que a' divise $y - y_0$. Il existe donc $k \in \mathbb{Z}$ tel que l'on ait $y = y_0 - ka'$, puis $x = x_0 + kb'$. Inversement, pour tout $k \in \mathbb{Z}$, on a $a(x_0 + kb') + b(y_0 - ka') = c$, d'où le résultat.

Exercice 18. Déterminer les couples $(x, y) \in \mathbb{Z}^2$ tels que $47x + 111y = 1$.

7. Plus petit commun multiple

Soient a et b deux entiers relatifs non nuls. Pour tous $x, y \in \mathbb{Z}$, on note $\mathrm{Max}(x, y)$ le plus grand d'entre eux.

Théorème 1.10. *Il existe un unique entier $m > 0$ vérifiant les deux conditions suivantes :*

1) l'entier m est un multiple commun à a et b.

2) Tout multiple commun à a et b est un multiple de m.

On a l'égalité

$$m = \prod_{p \in \mathbf{P}} p^{\mathrm{Max}\left(v_p(a), v_p(b)\right)}. \tag{12}$$

Définition 1.5. *L'entier m défini par l'égalité* (12) *est appelé le plus petit commun multiple de a et b, ou en abrégé le ppcm de a et b. On le note* $\mathrm{ppcm}(a, b)$ *ou $a \vee b$.*

Démonstration : Considérons l'entier m défini par l'égalité (12). Pour tout $p \in \mathbf{P}$, on a $\mathrm{Max}\big(v_p(a), v_p(b)\big) \geq v_p(a), v_p(b)$. Ainsi, m est un multiple commun à a et b (prop. 1.1). Par ailleurs, si c est un multiple commun à a et b, on a pour tout $p \in \mathbf{P}$ les inégalités $v_p(c) \geq v_p(a)$ et $v_p(c) \geq v_p(b)$, d'où $v_p(c) \geq \mathrm{Max}\big(v_p(a), v_p(b)\big)$, donc c est un multiple de

m (*loc. cit.*). L'entier m vérifie donc les conditions 1 et 2. Si m' est un entier ≥ 1 vérifiant ces conditions, alors m' est un multiple de m et m est un multiple de m', d'où $m = m'$.

Exercice 19. Calculer le ppcm de 1080 et de 3600.

Proposition 1.5. *On a l'égalité* $\mathrm{pgcd}(a,b)\,\mathrm{ppcm}(a,b) = |ab|$.

Démonstration : Pour tout $p \in \mathbf{P}$, on a

$$v_p(ab) = v_p(a) + v_p(b) = \mathrm{Max}\big(v_p(a), v_p(b)\big) + \mathrm{Min}\big(v_p(a), v_p(b)\big).$$

Les théorèmes 1.8 et 1.10 entraînent alors le résultat.

Exercice 20. Trouver tous les couples d'entiers naturels (a,b) pour lesquels on a $\mathrm{pgcd}(a,b) = 5$ et $\mathrm{ppcm}(a,b) = 8160$.

Remarque 1.2. On peut, comme pour le pgcd, généraliser la notion de ppcm au cas d'une famille finie d'entiers. Étant donnés des entiers non nuls $x_1, \cdots, x_n$ leur ppcm est l'unique entier $m > 0$ multiple des x_i, tel que tout multiple des x_i soit multiple de m. Pour tout $p \in \mathbf{P}$, on a comme attendu

$$v_p(m) = \mathrm{Max}\big(v_p(x_1), \cdots, v_p(x_n)\big).$$

Cela étant, on notera que la proposition 1.5 est fausse dans ce cadre général, ce qui s'explique par le fait que l'entier $\mathrm{Min}\big(v_p(x_1), \cdots, v_p(x_n)\big) + \mathrm{Max}\big(v_p(x_1), \cdots, v_p(x_n)\big)$ n'est pas en général la somme des $v_p(x_i)$: prendre par exemple $(x_1, x_2, x_3) = (2, 3, 4)$ et $p = 2$. Le pgcd des x_i est 1, leur ppcm est 12 et leur produit vaut 24.

8. Numération en base b

On considère dans ce paragraphe un entier $b \geq 2$.

Théorème 1.11. *Soit x un entier naturel non nul. On peut écrire x de manière unique sous la forme*

$$x = a_n b^n + a_{n-1} b^{n-1} + \cdots + a_1 b + a_0, \tag{13}$$

où n est un entier naturel, où $a_0, \cdots, a_n$ sont des entiers tels que $0 \leq a_i \leq b-1$ et où a_n est non nul. On dit que $x = a_n a_{n-1} \cdots a_1 a_0$ est l'écriture de x en base b et l'on écrit parfois $x = (a_n \cdots a_0)_b$.

Démonstration : Démontrons l'assertion d'existence. Notons pour cela $P(x)$ la propriété : x possède une écriture de la forme (13) comme indiquée dans l'énoncé. La propriété $P(1)$ est vraie, avec $n = 0$ et $a_0 = 1$. Considérons alors un entier $x \geq 2$ et supposons que

la propriété $P(k)$ soit vraie pour tout entier k tel que $1 \leq k < x$. Il s'agit de démontrer que $P(x)$ est vraie. Tel est le cas si l'on a $x < b$, en prenant $n = 0$ et $a_0 = x$ dans (13). Supposons donc $x \geq b$. Il existe des entiers q et a_0 tels que l'on ait $x = bq + a_0$ avec $0 \leq a_0 < b$. L'inégalité $x \geq b$ entraîne $q \geq 1$. Par suite, on a $q < bq \leq x$. La propriété $P(q)$ étant vraie, il existe un entier $n \geq 1$ tel que l'on ait $q = a_n b^{n-1} + \cdots + a_2 b + a_1$, où les a_i sont entiers vérifiant les inégalités $0 \leq a_i \leq b - 1$ et où $a_n \neq 0$. L'égalité $x = bq + a_0$ entraîne alors que $P(x)$ est vraie, d'où l'assertion d'existence.

Prouvons l'assertion d'unicité. On remarque pour cela que l'entier n intervenant dans (13) vérifie les inégalités

$$b^n \leq x < b^{n+1}. \tag{14}$$

En effet, la première inégalité est immédiate et le fait que les a_i soient compris entre 0 et $b - 1$ entraîne que l'on a $x \leq (b-1)(b^n + b^{n-1} + \cdots + b + 1) = b^{n+1} - 1 < b^{n+1}$. Par ailleurs, les inégalités (14) caractérisent l'entier n (n est le plus grand entier inférieur ou égal à $\log x / \log b$). Tout revient donc à démontrer que si l'on a

$$x = a_n b^n + a_{n-1} b^{n-1} + \cdots + a_1 b + a_0 = c_n b^n + c_{n-1} b^{n-1} + \cdots + c_1 b + c_0,$$

avec $a_n c_n \neq 0$ et $0 \leq a_i, c_i \leq b - 1$, alors $a_i = c_i$ pour tout i. Vu le caractère d'unicité du reste de la division euclidienne de x par b, on a $a_0 = c_0$. On en déduit ensuite l'assertion en procédant par récurrence finie sur les indices des coefficients.

Exemple 1.3. On vérifie que l'on a $101 = 1 + 2^2 + 2^5 + 2^6$, de sorte que l'écriture de 101 en base 2 est 1100101 i.e. on a $101 = (1100101)_2$.

Exercice 21. Déterminer l'écriture en base 3 de 7456.

Exercice 22. Soit n un entier naturel.

1) Trouver une condition nécessaire et suffisante simple pour que n soit divisible par 3, respectivement par 9.
2) Soit $n = (a_k a_{k-1} \cdots a_1 a_0)_{10}$ l'écriture de n en base 10. Montrer l'équivalence

$$n \text{ est divisible par } 11 \iff \sum_{i=0}^{k} (-1)^i a_i = 0.$$

Exercice 23. Déterminer les nombres de deux chiffres qui s'écrivent uv en base 10 et vu en base 7.

Terminons ce chapitre en donnant deux applications de ce théorème.

1. Calcul «rapide» de la puissance d'un entier

L'existence de l'écriture en base 2 des entiers permet d'accélérer le calcul de la puissance d'un entier. Plus précisément, considérons deux entiers $x \geq 1$ et $n \geq 1$. Afin de calculer x^n, il faut a priori effectuer $n-1$ multiplications. En fait, la détermination de l'écriture de n en base 2 permet de calculer x^n en effectuant au plus $2\log n/\log 2$ multiplications. On procède comme suit. Soit

$$n = 2^{i_k} + 2^{i_{k-1}} + \cdots + 2^{i_1} + 2^{i_0},$$

le développement de n en base 2 avec $i_0 < i_1 \cdots < i_k$. On a l'égalité

$$x^n = x^{2^{i_k}} \times x^{2^{i_{k-1}}} \times \cdots \times x^{2^{i_1}} \times x^{2^{i_0}}.$$

On peut effectuer le calcul de $x^{2^{i_k}}$ avec i_k multiplications, ce qui fournit aussi le calcul des autres termes $x^{2^{i_j}}$ pour $0 \leq j \leq k$. Il en résulte que l'on peut calculer x^n avec $i_k + k$ multiplications. Par ailleurs, on a

$$k \leq i_k \quad \text{et} \quad 2^{i_k} \leq n \quad \text{i.e.} \quad i_k \leq \frac{\log n}{\log 2}.$$

On en déduit que

$$i_k + k \leq \frac{2\log n}{\log 2},$$

ce qui établit notre assertion.

Exemple 1.4. On a vu plus haut que l'on a $101 = 2^6 + 2^5 + 2^2 + 1$. Par suite, le calcul de x^{101} peut se faire avec neuf multiplications (au lieu de cent a priori).

2. Formule de Legendre[8] donnant $v_p(n!)$

Soient n un entier naturel non nul et p un nombre premier. On se propose de déterminer ici la valuation p-adique de $n!$. On va démontrer le résultat suivant, qui a été obtenu par Legendre en 1808 :

Théorème 1.11. *Soit $n = (a_k \cdots a_0)_p$ l'écriture de n en base p. On a*

$$v_p(n!) = \frac{n - S}{p-1} \quad \text{où} \quad S = \sum_{i=0}^{k} a_i. \tag{15}$$

Démonstration : Pour tout $x \in \mathbb{R}$, notons $[x]$ la partie entière de x, i.e. le plus grand entier relatif inférieur ou égal à x. On utilise le lemme suivant :

[8] Adrien-Marie Legendre est né à Paris en 1752 et décède en 1833. Il fut professeur à l'École Militaire de Paris de 1775 à 1780. Il apporta sa contribution mathématique dans de nombreux domaines, notamment au calcul intégral, à la théorie des fonctions elliptiques ainsi qu'à la théorie des nombres. Il démontra le grand théorème de Fermat pour $n = 5$.

Lemme 1.5. *Soient a et b deux entiers ≥ 1. On a*

$$\left[\frac{a+1}{b}\right] - \left[\frac{a}{b}\right] = \begin{cases} 1 & \textit{si } b \textit{ divise } a+1 \\ 0 & \textit{sinon.} \end{cases}$$

Démonstration : Il existe deux entiers q et r tels que l'on ait

$$a = bq + r \quad \text{avec} \quad 0 \leq r < b.$$

On a en particulier

$$\left[\frac{a}{b}\right] = q \quad \text{et} \quad a+1 = bq + (r+1) \quad \text{avec} \quad r+1 \leq b.$$

On en déduit que b divise $a+1$ si et seulement si $b = r+1$. Si b divise $a+1$, on a donc $a+1 = b(q+1)$, d'où

$$\left[\frac{a+1}{b}\right] = q+1,$$

et l'on obtient dans ce cas l'égalité annoncée. Si b ne divise pas $a+1$, on a $r+1 < b$, d'où

$$\left[\frac{a+1}{b}\right] = q.$$

et le résultat.

Pour tout entier $N \geq 1$, posons

$$S_N = \sum_{i \geq 1} \left[\frac{N}{p^i}\right].$$

Il s'agit d'une somme finie car $\left[\frac{N}{p^i}\right]$ est nul dès que i est assez grand.

Corollaire 1.4. *Pour tout entier $N \geq 1$, on a $v_p(N!) = S_N$.*

Démonstration : On procède par récurrence sur N. Le résultat est vrai si $N = 1$. Supposons que tel est aussi le cas pour un entier $N \geq 1$. D'après le lemme 1.5, on a

$$\left[\frac{N+1}{p^i}\right] - \left[\frac{N}{p^i}\right] = \begin{cases} 1 & \text{si } \ i \leq v_p(N+1) \\ 0 & \text{sinon.} \end{cases}$$

Il en résulte que l'on a les égalités

$$S_{N+1} - S_N = \sum_{1 \leq i \leq v_p(N+1)} 1 = v_p(N+1). \tag{16}$$

Par ailleurs, on a

$$v_p\big((N+1)!\big) = \sum_{j=1}^{N+1} v_p(j) = v_p(N!) + v_p(N+1).$$

En utilisant l'hypothèse de récurrence, on obtient ainsi

$$v_p\big((N+1)!\big) = S_N + v_p(N+1),$$

qui d'après (16), n'est autre que S_{N+1}.

Fin de la démonstration du théorème : On a l'égalité

$$n = a_k p^k + \cdots + a_1 p + a_0 \quad \text{avec} \quad 0 \leq a_i < p.$$

Considérons un entier j tel que $1 \leq j \leq k$. On a

$$\frac{n}{p^j} = a_k p^{k-j} + \cdots + a_j + \frac{a_{j-1}p^{j-1} + \cdots a_1 p + a_0}{p^j}.$$

Par ailleurs, on a

$$a_{j-1}p^{j-1} + \cdots a_1 p + a_0 \leq (p-1)(p^{j-1} + \cdots + 1) = p^j - 1 < p^j.$$

Il en résulte que l'on a

$$\left[\frac{n}{p^j}\right] = a_k p^{k-j} + \cdots + a_j. \tag{17}$$

L'inégalité $n < p^{k+1}$ entraîne

$$\left[\frac{n}{p^i}\right] = 0 \quad \text{pour tout} \quad i \geq k+1.$$

D'après le corollaire 1.4, on a donc

$$v_p(n!) = \sum_{j=1}^{k} \left[\frac{n}{p^j}\right].$$

On déduit alors de (17) l'égalité

$$v_p(n!) = a_1 + a_2(p+1) + a_3(p^2+p+1) + \cdots + a_k(p^{k-1} + p^{k-2} + \cdots + p + 1),$$

autrement dit,

$$v_p(n!) = \frac{1}{p-1}\Big(a_1(p-1) + a_2(p^2-1) + a_3(p^3-1) + \cdots a_k(p^k-1)\Big) = \frac{n-S}{p-1},$$

ce qui établit la formule (15).

Exemple 1.5. On a $100 = 2^2 + 2^5 + 2^6 = 4.5^2$, autrement dit, $100 = (1100100)_2$ et $100 = (400)_5$. On a donc $v_2(100!) = 97$ et $v_5(100!) = 24$. Il en résulte que 100! se termine par vingt-quatre zéros dans son écriture en base 10.

Chapitre II — La notion de groupe

1. Définition d'un groupe

Afin de définir un groupe, il convient de se donner un ensemble G ainsi qu'une loi de composition sur G vérifiant certaines conditions[9] :

Définition 2.1. *On appelle groupe un couple $(G, *)$ formé d'un ensemble G et d'une loi de composition $(x, y) \mapsto x * y$ sur G, tels que les trois conditions suivantes soient vérifiées :*

1) *on a $x * (y * z) = (x * y) * z$ quels que soient $x, y, z \in G$ (associativité).*
2) *Il existe un élément $e \in G$ tel que $e * x = x * e = x$ pour tout $x \in G$ (existence d'un élément neutre).*
3) *Pour tout $x \in G$, il existe un élément $y \in G$ tel que $x * y = y * x = e$ (existence d'un symétrique pour tout élément de G).*

Si de plus, quels que soient $x, y \in G$, on a $x * y = y * x$ (commutativité), on dit que G est un groupe commutatif ou abélien. C'est la situation que l'on rencontrera principalement dans ce cours.

Un groupe peut être fini ou infini. S'il est fini, on appelle ordre du groupe le nombre de ses éléments i.e. son cardinal.

Notation. Dans la définition 2.1, on a utilisé la notation abstraite $*$ pour définir la loi de composition sur G. En théorie des groupes, on note en fait la plupart du temps la loi de composition sous-jacente multiplicativement $(x, y) \mapsto xy$, ou bien additivement $(x, y) \mapsto x + y$. En notation multiplicative, on emploie le mot inverse au lieu du mot symétrique et l'inverse d'un élément x se note x^{-1}. Pour tous $x, y \in G$, on a alors la formule $(xy)^{-1} = y^{-1}x^{-1}$. En notation additive, on dit opposé au lieu de symétrique, et l'on note généralement 0 l'élément neutre et $-x$ l'opposé de x. Dans la pratique, la notation additive

[9] Soient E un ensemble et $* : E \times E \to E$ une loi de composition sur E. On dit que $*$ est associative si pour tous $x, y, z \in E$, on a $x * (y * z) = (x * y) * z$. Elle est dite commutative si pour tous $x, y \in E$, on a $x * y = y * x$. On appelle élément neutre pour $*$, tout élément $e \in E$ tel que pour tout $x \in E$ on ait $x * e = e * x = x$. Si une loi de composition admet un élément neutre, elle en admet un seul (exercice). Dans le cas où E possède un élément neutre e, un élément $x \in E$ est dit symétrisable s'il existe $y \in E$ tel que $x * y = y * x = e$. Un tel élément y est appelé symétrique de x. Si $*$ est associative et si E a un élément neutre pour $*$, alors pour tout $x \in E$, si x admet un symétrique, il en admet un seul (exercice)

Signalons qu'il existe des lois de composition non associatives, ayant un élément neutre, telles que les éléments de l'ensemble sous-jacent, autres que l'élément neutre, aient une infinité de symétriques. Tel est par exemple le cas de l'ensemble $\mathbb{N}$ des entiers naturels muni de la loi de composition définie comme suit : pour tous $a, b \in \mathbb{N}$, on pose $a * 0 = 0 * a = a$ et $a * b = 0$ si $ab \neq 0$.

est utilisée uniquement pour les groupes abéliens. Cela étant, la notation multiplicative est aussi très souvent employée pour les groupes abéliens. Dans toute la suite, on appellera groupe multiplicatif, un groupe dont la loi de composition est notée multiplicativement, et groupe additif, un groupe dont la loi de composition est notée addtitivement.

Exercice 1. Soient G un groupe multiplicatif et e son élément neutre.

1) Montrer que si pour tout $x \in G$, on a $x^2 = e$, alors G est abélien.

2) Supposons que G soit un groupe fini d'ordre pair. Montrer qu'il existe un élément $x \in G$, autre que e, tel que $x^2 = e$.

Exemples 2.1.

1) L'ensemble réduit à un seul élément e, avec pour loi de composition $e * e = e$, est un groupe, appelé le groupe trivial.

2) L'ensemble $\mathbb{Z}$ des entiers relatifs muni de la loi de composition $(x, y) \mapsto x + y$ est un groupe commutatif, d'élément neutre 0. On l'appelle le groupe additif des entiers relatifs. En remplaçant $\mathbb{Z}$ par $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$, on obtient le groupe additif des nombres rationnels, le groupe additif des nombres réels et le groupe additif des nombres complexes.

3) L'ensemble $\mathbb{Q}^*$ des nombres rationnels non nuls muni de la loi de composition $(x, y) \mapsto xy$ est un groupe commutatif, d'élément neutre 1. C'est le groupe multiplicatif des nombres rationnels non nuls. On définit de même le groupe multiplicatif $\mathbb{R}^*$ des nombres réels non nuls et le groupe multiplicatif $\mathbb{C}^*$ des nombres complexes non nuls.

4) Si X est un ensemble, soit $\mathbb{S}(X)$ l'ensemble des bijections[10] de X sur X. La loi de composition $(f, g) \mapsto f \circ g$ définit une structure de groupe sur $\mathbb{S}(X)$ que l'on appelle le groupe symétrique de X.

Exercice 2. Supposons que X soit fini de cardinal n. Quel est l'ordre de $\mathbb{S}(X)$? Montrer que X n'est pas commutatif si l'on a $n \geq 3$.

5) **Produit direct de groupes.** Soient $G_1, \cdots, G_n$ des groupes multiplicatifs. Il existe sur le produit cartésien

$$G = G_1 \times \cdots \times G_n$$

une structure de groupe dont la loi de composition est donnée par la formule

$$(x_1, \cdots, x_n).(y_1, \cdots, y_n) = (x_1 y_1, \cdots, x_n y_n).$$

[10] Soient E et F deux ensembles et $f : E \to F$ une application de E à valeurs dans F. On dit que f est une injection, si quels que soient x et y dans E, l'égalité $f(x) = f(y)$ entraîne $x = y$. On dit que f est une surjection, ou une surjection de E sur F, si pour tout $y \in F$, il existe $x \in E$ tel que $f(x) = y$, autrement dit, si tout élément de F à un antécédent par f. On dit que f est une bijection, ou une bijection de E sur F, si f est à la fois une injection et une surjection. Cela signifie que pour tout $y \in F$, il existe un unique élément $x \in E$ tel que $f(x) = y$.

L'élément neutre est $(e_1, \cdots, e_n)$ où e_i est l'élément neutre de G_i. L'inverse d'un élément $x = (x_1, \cdots, x_n)$ est donné par la formule

$$x^{-1} = (x_1^{-1}, \cdots, x_n^{-1}).$$

Le groupe $(G, .)$ est appelé le produit direct des groupes $G_1, \cdots, G_n$, ou bien aussi le groupe produit de $G_1, \cdots, G_n$.

On obtient ainsi (en notation additive) les groupes additifs $\mathbb{Z}^n$, $\mathbb{R}^n$ et $\mathbb{C}^n$, les lois de composition correspondantes étant données par la formule

$$(x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n).$$

2. Sous-groupes d'un groupe

Soit G un groupe multiplicatif d'élément neutre e.

Définition 2.2. *Soit H une partie de G. On dit que H est un sous-groupe de G si les conditions suivantes sont réalisées :*

1) l'élément e appartient à H.

2) pour tous $x, y \in H$, l'élément xy est dans H.

3) Pour tout $x \in H$, l'inverse x^{-1} de x est dans H.

Un sous-groupe de G muni de la loi de composition induite par celle de G est un groupe.

Exercice 3. Soit H une partie de G. Montrer que H est un sous-groupe de G si et seulement si H n'est pas vide, et si pour tous $x, y \in H$ l'élément xy^{-1} est aussi dans H.

Exemples 2.2.

1) Les parties G et $\{e\}$ sont des sous-groupes de G. Le sous-groupe $\{e\}$ s'appelle le sous-groupe trivial de G.

2) Le sous-ensemble de $\mathbb{R}^*$ formé des nombres réels strictement positifs, ainsi que $\{\pm 1\}$, sont des sous-groupes de $\mathbb{R}^*$.

3) L'ensemble des nombres complexes de module 1 est un sous-groupe de $\mathbb{C}^*$.

4) **Sous-groupes de** $(\mathbb{Z}, +)$. Si n est un entier relatif, la partie $n\mathbb{Z} = \{nk \mid k \in \mathbb{Z}\}$ est un sous-groupe de $\mathbb{Z}$. On obtient en fait ainsi tous les sous-groupes de $\mathbb{Z}$:

Proposition 2.1. *Soit H un sous-groupe de $\mathbb{Z}$. Il existe un unique entier naturel n tel que l'on ait $H = n\mathbb{Z}$.*

Démonstration : C'est immédiat si $H = \{0\}$, en prenant $n = 0$. Supposons $H \neq \{0\}$. L'ensemble $A = H \cap \left(\mathbb{N} - \{0\}\right)$ n'est pas vide, car si n est dans H, alors $-n$ l'est aussi.

Soit n le plus petit élément de A. Vérifions que l'on a $H = n\mathbb{Z}$. Tout d'abord, H étant un sous-groupe de $\mathbb{Z}$, et n étant dans H, le sous-groupe $n\mathbb{Z}$ est contenu dans H. Inversement, soit x un élément de H. Il existe q et r dans $\mathbb{Z}$ tels que l'on ait $x = nq + r$ avec $0 \leq r < n$. Puisque x et nq appartiennent à H, il en est de même de $x - nq = r$. Le caractère minimal de n conduit alors à $r = 0$, de sorte que $x = nq$ appartient à $n\mathbb{Z}$. Par ailleurs, si l'on a $n\mathbb{Z} = m\mathbb{Z}$ avec m et n dans $\mathbb{N}$, alors m divise n et n divise m, d'où $m = n$.

5) Soit x un élément de G. Pour tout entier relatif k, on définit x^k comme suit[11] :

$$x^k = \begin{cases} x \cdots x \text{ (k facteurs)} & \text{si } k \geq 1 \\ e & \text{si } k = 0 \\ (x^{-1})^{-k} & \text{si } k < 0. \end{cases} \tag{1}$$

Quels que soient les entiers relatifs k et k', on vérifie que l'on a les égalités

$$x^k x^{k'} = x^{k+k'}, \quad (x^k)^{-1} = x^{-k}, \quad (x^k)^{k'} = x^{kk'}. \tag{2}$$

Il en résulte que l'ensemble $\{x^k \mid k \in \mathbb{Z}\}$ est un sous-groupe de G. Compte tenu de la première égalité de (2), c'est un sous-groupe abélien de G.

Remarque 2.1. On dispose de formules analogues si la loi de composition de G est notée additivement : pour tout entier k, on définit dans ce cas kx par les égalités :

$$kx = \begin{cases} x + \cdots + x \text{ (k facteurs)} & \text{si } k \geq 1 \\ e & \text{si } k = 0 \\ (-k)(-x) & \text{si } k < 0, \end{cases} \tag{3}$$

avec les formules pour tous k et $k' \in \mathbb{Z}$,

$$kx + k'x = (k + k')x, \quad -(kx) = (-k)x, \quad k'(kx) = (kk')x. \tag{4}$$

Cette construction généralise celle des sous-groupes de $\mathbb{Z}$.

---

[11] Soient E un ensemble et $*$ une loi de composition sur E. On définit le composé d'éléments $x_1, \cdots, x_n$ de E par la formule de récurrence :

$$x_1 * x_2 * \cdots * x_n = (x_1 * x_2 * \cdots * x_{n-1}) * x_n.$$

Pour tout $x \in E$ et tout entier $n \geq 1$, on définit la puissance n-ième de x par la formule $x^n = x * \cdots * x$ (n facteurs). Supposons $*$ associative. Vérifions alors que pour tout entier p tel que $1 \leq p \leq n$, on a l'égalité

$$x_1 * \cdots * x_n = (x_1 * \cdots * x_p) * (x_{p+1} * \cdots * x_n).$$

Elle est vraie si $n = 1$. Supposons cette assertion démontrée pour un produit de $n - 1$ éléments où $n \geq 2$. Posons $x = x_1 * \cdots * x_n$. On a $x = (x_1 * \cdots * x_{n-1}) * x_n$. Soit p un entier compris entre 1 et n. L'égalité à prouver étant satisfaite si $p = n$, on peut supposer $p \leq n-1$. D'après l'hypothèse de récurrence, on a donc $x = \big((x_1 * \cdots * x_p) * (x_{p+1} \cdots * x_{n-1})\big) * x_n$. Puisque $*$ est associative, on obtient $x = (x_1 * \cdots * x_p) * \big((x_{p+1} \cdots * x_{n-1}) * x_n\big)$, d'où l'égalité annoncée.

Lemme 2.1. *Soit $(H_i)_{i\in I}$ une famille de sous-groupes de G. L'intersection des H_i est un sous-groupe de G.*

Démonstration : L'élément neutre appartient à chacun des H_i. Par ailleurs, si x et y sont dans l'intersection des H_i, les éléments xy et x^{-1} le sont aussi, car les H_i sont des sous-groupes de G.

Exercice 4. Montrer que la réunion de deux sous-groupes de G est un sous-groupe de G si et seulement si l'un est contenu dans l'autre.

Exercice 5. Pour tous sous-groupes A et B de G, on désigne par AB le sous-ensemble de G formé des éléments de G de la forme ab, où a est dans A et où b est dans B. Soient H et K deux sous-groupes de G. Démontrer que $HK{=}KH$ si et seulement si HK est un sous-groupe de G.

Exercice 6. Soient a et b deux entiers relatifs non nuls. Monter que l'on a l'égalité

$$a\mathbb{Z} \cap b\mathbb{Z} = \mathrm{ppcm}(a,b)\mathbb{Z}.$$

3. Classes modulo un sous-groupe - Théorème de Lagrange

Soient G un groupe multiplicatif d'élément neutre e et H un sous-groupe de G. On associe à H la relation binaire $\mathcal{R}$ sur G définie pour tous $x, y \in G$ par

$$x\mathcal{R}y \Longleftrightarrow x^{-1}y \in H. \tag{5}$$

Proposition 2.2. *La relation $\mathcal{R}$ est une relation d'équivalence*[12] *sur G.*

Démonstration : La propriété de réflexivité résulte du fait que $e \in H$. Considérons des éléments x, y, z de G. L'égalité $(x^{-1}y)^{-1} = y^{-1}x$ entraîne la propriété de symétrie. Par ailleurs, si l'on a $x\mathcal{R}y$ et $y\mathcal{R}z$, les éléments $x^{-1}y$ et $y^{-1}z$ sont dans H, donc le composé $(x^{-1}y)(y^{-1}z) = x^{-1}z$ est aussi dans H, autrement dit, on a $x\mathcal{R}z$. Cela établit la propriété de transitivité et le résultat.

---

[12] Soient E un ensemble et $\mathcal{R}$ une relation binaire sur E. On dit que $\mathcal{R}$ est une relation d'équivalence sur E si, quels que soient x, y, z dans E, les conditions suivantes sont satisfaites :

1) la relation $x\mathcal{R}x$ est vraie (réflexivité).
2) la relation $x\mathcal{R}y$ implique $y\mathcal{R}x$ (symétrie).
3) les relations $x\mathcal{R}y$ et $y\mathcal{R}z$ impliquent $x\mathcal{R}z$ (transitivité).

Soit f une application de E à valeurs dans un ensemble F quelconque. Prenons pour $\mathcal{R}$ la relation suivante : pour tout $(x, y) \in E^2$, on a $x\mathcal{R}y$ si et seulement si $f(x) = f(y)$. C'est une relation d'équivalence sur E, appelée la relation d'équivalence associée à f. En fait, cet exemple conduit à toutes les relations d'équivalence sur E comme on peut le constater ci-après.

Proposition 2.3. *Soit x un élément de G. La classe d'équivalence de x modulo $\mathcal{R}$ est l'ensemble*

$$xH = \Big\{ xh \mid h \in H \Big\}.$$

Démonstration : Soit x un élément de G. Les éléments de la forme xh où $h \in H$ sont en relation avec x, car $x^{-1}xh = h \in H$. Inversement, soit y un élément de G en relation avec x. On a $x\mathcal{R}y$, de sorte que $x^{-1}y \in H$ et y est donc dans xH, d'où l'assertion.

Définition 2.3. *L'ensemble xH s'appelle la classe à gauche de x modulo H. L'ensemble des classes à gauche des éléments de G modulo H se note G/H, et s'appelle l'ensemble quotient à gauche de G modulo H. On a ainsi*

$$G/H = \Big\{ xH \mid x \in G \Big\}.$$

On déduit de ce qui précède le théorème de Lagrange (1736-1813), qui est à la base de toute la théorie des groupes finis. Supposons que G soit un groupe fini. Dans ce cas, il en est de même des ensembles H et G/H. Notons $|G|$, $|H|$ et $|G/H|$ leurs cardinaux respectifs.

Théorème 2.1 (Lagrange). *Supposons G fini. On a l'égalité $|G| = |H| \times |G/H|$. En particulier, l'ordre de H divise celui de G.*

Démonstration : Pour tout $x \in G$, les ensembles H et xH sont en bijection via l'application qui à z associe xz. Le résultat s'en déduit aussitôt compte tenu du fait que $|G/H|$ est le nombre de classes d'équivalence distinctes de G modulo H et que G est la réunion disjointe de ces classes.

Corollaire 2.1. *Supposons que G soit fini d'ordre un nombre premier. Alors, les seuls sous-groupes de G sont G et $\{e\}$.*

Supposons que $\mathcal{R}$ soit une relation d'équivalence sur E. Si x est un élément de E, la classe d'équivalence de x modulo $\mathcal{R}$, ou plus simplement, la classe de x modulo $\mathcal{R}$, est l'ensemble des $y \in E$ en relation avec x, autrement dit, est l'ensemble des $y \in E$ tels que l'on ait $x\mathcal{R}y$. L'ensemble des classes d'équivalence de E modulo $\mathcal{R}$ est appelé l'ensemble quotient de E modulo $\mathcal{R}$, on le note parfois $E/\mathcal{R}$. On dispose de l'application $s : E \to E/\mathcal{R}$ qui à tout $x \in E$ associe sa classe modulo $\mathcal{R}$. On l'appelle la surjection canonique. Par définition, $\mathcal{R}$ est la relation d'équivalence associée à s.
Soit $C(x)$ la classe d'équivalence d'un élément $x \in E$. Alors, x appartient à $C(x)$. Par ailleurs, quels que soient $x, y \in E$, les conditions suivantes sont équivalentes :

1) on a $x\mathcal{R}y$.
2) On a $C(x) = C(y)$.
3) L'ensemble $C(x) \cap C(y)$ n'est pas vide.

L'ensemble des classes d'équivalence distinctes de E modulo $\mathcal{R}$ forme une partition de E. (Une partition de E est un ensemble de parties non vides de E, deux à deux disjointes, dont la réunion est E).

Démonstration : C'est une conséquence directe du théorème de Lagrange et de la définition d'un nombre premier.

On définit de même la classe à droite de Hx de x modulo H comme étant l'ensemble $\{hx \mid h \in H\}$. On démontre comme ci-dessus que Hx est la classe d'équivalence de x associée à la relation d'équivalence sur G pour laquelle x et y sont en relation si et seulement si $xy^{-1} \in H$. Lorsque G est abélien, ce qui sera le cas dans les applications que l'on rencontrera dans ce cours, on a évidemment $xH = Hx$ pour tout $x \in G$. Il n'y a donc pas lieu de faire de distinction dans ce cas, ce qui conduit à la définition suivante :

Définition 2.4. *Supposons G abélien. La relation $\mathcal{R}$ s'appelle la relation d'équivalence modulo H. Pour tout $x \in G$, l'ensemble xH s'appelle la classe de x modulo H et G/H s'appelle l'ensemble quotient de G modulo H.*

Remarque 2.2. Supposons que G soit un groupe abélien additif, i.e. on note additivement la loi de composition sur G. La relation d'équivalence modulo H définie par la condition (5) s'écrit alors sous la forme

$$x\mathcal{R}y \Longleftrightarrow x - y \in H. \tag{6}$$

Pour tout $x \in G$, la classe de x modulo H se note $x + H$. On a

$$x + H = \Big\{x + h \mid h \in H\Big\} \quad \text{et} \quad G/H = \Big\{x + H \mid x \in G\Big\}. \tag{7}$$

Remarquons que si 0 est l'élément neutre de G, on a $0 + H = H$ i.e. la classe modulo H de l'élément neutre de G est H.

Exemple 2.3. L'ensemble quotient $\mathbb{Z}/n\mathbb{Z}$

Prenons $G = \mathbb{Z}$ et $H = n\mathbb{Z}$ où $n \in \mathbb{N}$. D'après (6), quels que soient $x, y \in \mathbb{Z}$ on a l'équivalence

$$x \text{ et } y \text{ sont en relation modulo } n\mathbb{Z} \Longleftrightarrow x - y \in n\mathbb{Z}. \tag{8}$$

Deux entiers relatifs x et y sont donc en relation modulo $n\mathbb{Z}$ si et seulement si n divise $x - y$. Dans ce cas, on dit que x et y sont congrus modulo n et l'on écrit que l'on a la congruence $x \equiv y$ mod. n. Pour tout $x \in \mathbb{Z}$, la classe de x modulo $n\mathbb{Z}$ est

$$x + n\mathbb{Z} = \Big\{x + nk \mid k \in \mathbb{Z}\Big\}. \tag{9}$$

On la note souvent $\overline{x}$ lorsque que l'entier n est sous-entendu. On dit aussi que c'est la classe de x modulo n. L'ensemble $\mathbb{Z}/n\mathbb{Z}$ est formé des classes d'équivalence modulo n.

Proposition 2.4. *Supposons $n \geq 1$. Alors, $\mathbb{Z}/n\mathbb{Z}$ est fini de cardinal n et l'on a*

$$\mathbb{Z}/n\mathbb{Z} = \Big\{\overline{0}, \overline{1}, \cdots, \overline{n-1}\Big\}.$$

Démonstration : C'est une conséquence du théorème de la division euclidienne. Considérons en effet un élément $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ où $a \in \mathbb{Z}$. Il existe des entiers q et r tels que l'on ait $a = nq + r$ avec $0 \leq r < n$. Puisque $a - r \in n\mathbb{Z}$, on a donc $\overline{a} = \overline{r}$. Par ailleurs, quels que soient a et b distincts compris entre 0 et $n-1$, l'entier n ne divise pas $a - b$, autrement dit, on a $\overline{a} \neq \overline{b}$, d'où le résultat.

On dispose de la surjection canonique $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ qui à un entier a associe sa classe modulo n. Dans le cas où $n = 0$, et dans ce cas seulement, c'est une bijection.

4. Groupe quotient d'un groupe abélien

Soit G un groupe abélien additif, d'élément neutre 0. Soit H un sous-groupe de G. On définit sur l'ensemble quotient

$$G/H = \Big\{x + H \mid x \in G\Big\},$$

une loi de composition $\oplus$ comme suit. Soient u, v deux éléments de G/H. Il existe $x, y \in G$ tels que $u = x + H$ et $v = y + H$. On pose alors

$$u \oplus v = x + y + H, \tag{10}$$

autrement dit, $u \oplus v$ est la classe de $x + y$ modulo H. Il faut bien entendu vérifier que cette définition a un sens, i.e. que $u \oplus v$ ne dépend pas des représentants choisis x et y de u et v. Considérons pour cela des représentants x' et y' respectivement de u et v. Par définition, $x - x'$ et $y - y'$ sont dans H. Puisque G est abélien, il en résulte que

$$(x - x') + (y - y') = x + y - (x' + y') \in H,$$

ce qui signifie que $x + y$ et $x' + y'$ sont en relation modulo H, i.e. que les classes de $x + y$ et $x' + y'$ modulo H sont égales, d'où notre assertion.

Proposition 2.5. *L'ensemble G/H muni de la loi $\oplus$ est un groupe abélien. On l'appelle le groupe quotient de G par H.*

Démonstration : Le fait que la loi $+$ sur G soit associative et commutative entraîne qu'il en est de même de $\oplus$. L'élément neutre de $\oplus$ est $0 + H = H$ et pour tout $x \in G$, l'opposé de $x + H$ est $-x + H$, où $-x$ est l'opposé de x dans G. D'où le résultat.

Comme conséquence du théorème de Lagrange, si G est fini, on a l'énoncé suivant :

Proposition 2.6. *Supposons G fini. Alors, G/H est fini d'ordre $|G|/|H|$.*

Exemple 2.4. Le groupe additif $\big(\mathbb{Z}/n\mathbb{Z}, +\big)$

Soit n un entier naturel. En prenant pour $(G, +)$ le groupe des entiers relatifs $(\mathbb{Z}, +)$ et pour H le sous-groupe $n\mathbb{Z}$ de $\mathbb{Z}$, on obtient le groupe quotient $(\mathbb{Z}/n\mathbb{Z}, \oplus)$. On notera toujours, par abus, la loi $\oplus$ sur $\mathbb{Z}/n\mathbb{Z}$ de nouveau $+$. Le groupe quotient $(\mathbb{Z}/n\mathbb{Z}, +)$ ainsi défini est appelé le groupe additif des entiers relatifs modulo n. Quels que soient $a, b \in \mathbb{Z}$, on a d'après (10),

$$\overline{a} + \overline{b} = \overline{a + b}. \tag{11}$$

L'élément neutre de $\mathbb{Z}/n\mathbb{Z}$ est la classe de 0 modulo $n\mathbb{Z}$ i.e. $0 + n\mathbb{Z} = n\mathbb{Z}$. L'opposé de la classe de a est la classe de $-a$, autrement dit, on a

$$-\overline{a} = \overline{-a}. \tag{12}$$

On a par ailleurs

$$k\overline{a} = \overline{ka} \quad \text{pour tout } k \in \mathbb{Z}. \tag{13}$$

Compte tenu de la proposition 2.4 :

Proposition 2.7. *Supposons $n \geq 1$. Le groupe additif $(\mathbb{Z}/n\mathbb{Z}, +)$ est d'ordre n et ses éléments sont les classes des entiers compris entre 0 et $n - 1$.*

Exercice 7. Montrer que $\big\{\overline{0}, \overline{3}, \overline{6}, \overline{9}\big\}$ est un sous-groupe de $\big(\mathbb{Z}/12\mathbb{Z}, +\big)$. Expliciter un sous-groupe d'ordre 6 de $\big(\mathbb{Z}/12\mathbb{Z}, +\big)$.

Exercice 8. Déterminer tous les sous-groupes de $\big(\mathbb{Z}/6\mathbb{Z}, +\big)$.

5. Sous-groupe engendré par un élément - Ordre d'un élément

Soit G un groupe multiplicatif, d'élément neutre e. Rappelons qu'une intersection de sous-groupes de G est encore un sous-groupe de G (lemme 2.1). Pour tout $x \in G$, il existe donc un plus petit sous-groupe de G qui contient x (au sens de l'inclusion), à savoir l'intersection de tous les sous-groupes de G qui contiennent le singleton $\big\{x\big\}$.

Définition 2.5. *Soit x un élément de G. On appelle sous-groupe engendré par x, l'intersection de tous les sous-groupes de G qui contiennent $\big\{x\big\}$. On le note $< x >$.*

Lemme 2.2. *Soit x un élément de G. On a l'égalité*

$$< x >= \big\{x^k \mid k \in \mathbb{Z}\big\}.$$

Démonstration : Comme on l'a constaté dans les exemples 2.2 du paragraphe 2, $\left\{x^k \mid k \in \mathbb{Z}\right\}$ est un sous-groupe de G qui contient x, donc il contient $< x >$. Inversement, soit H un sous-groupe de G tel que x soit dans H. D'après les propriétés de stabilité d'un sous-groupe (déf. 2.2) l'ensemble $\left\{x^k \mid k \in \mathbb{Z}\right\}$ est contenu dans H. Par suite, il est contenu dans $< x >$, d'où le lemme.

Exemples 2.5.

1) Pour tout entier naturel n, le sous-groupe de $(\mathbb{Z}, +)$ engendré par n est $n\mathbb{Z}$.

2) Pour tout $n \geq 1$, le sous-groupe de $(\mathbb{Z}/n\mathbb{Z}, +)$ engendré par la classe de 1 est $\mathbb{Z}/n\mathbb{Z}$ tout entier. En effet, posons $H =< \overline{1} >$. Pour tout entier k compris entre 0 et $n-1$, on a $k\overline{1} = \overline{k}$ (formule (13)), donc $\overline{k}$ est dans H (lemme 2.2), d'où $H = \mathbb{Z}/n\mathbb{Z}$ (prop. 2.7). On dit que $\overline{1}$ est un générateur de $(\mathbb{Z}/n\mathbb{Z}, +)$. Cette notion sera généralisée plus loin.

On suppose désormais que G est un groupe fini.

Définition 2.6. *Supposons G fini. Soit x un élément de G. On appelle ordre de x, l'ordre du sous-groupe de G engendré par x.*

Exercice 9. Démontrer que pour tout $x \in G$, les éléments x et x^{-1} ont le même ordre.

Théorème 2.2. *Supposons G fini. Soit x un élément de G d'ordre m.*

1) On a $m \geq 1$ et m divise l'ordre de G.

2) On a $x^m = e$ et m est le plus petit entier $k \geq 1$ tel que $x^k = e$.

3) Les éléments $e, x, \cdots, x^{m-1}$ sont distincts deux à deux. En particulier, on a

$$< x >= \left\{e, x, x^2, \cdots, x^{m-1}\right\}.$$

Démonstration : L'assertion 1 résulte du théorème de Lagrange.

Prouvons l'assertion 2. Considérons pour cela l'ensemble A défini par l'égalité

$$A = \left\{k \in \mathbb{N} \mid 1 \leq k \leq m \text{ et } x^k = e\right\}.$$

Il est non vide. En effet, si A était vide, les éléments

$$x, x^2, \cdots, x^m, x^{m+1},$$

seraient distincts deux à deux et l'ordre de $< x >$ serait strictement plus grand que m (lemme 2.2). Soit u le plus petit élément de A. Il s'agit de démontrer que l'on a

$$u = m. \tag{14}$$

Posons

$$B = \left\{e, x, \cdots, x^{u-1}\right\}.$$

D'après le caractère minimal de u, le cardinal de B est u. Vérifions que $< x >$ est contenu dans B. Considérons pour cela un entier relatif k. Il existe q et r dans $\mathbb{Z}$ tels que l'on ait

$$k = uq + r \quad \text{avec} \quad 0 \leq r < u.$$

Vu que l'on a $x^u = e$, on obtient ainsi

$$x^k = (x^u)^q x^r = x^r \in B,$$

d'où l'assertion (lemme 2.2). Par suite, on a $m \leq u$. Puisque u appartient à A, on a aussi $u \leq m$, d'où l'égalité (14).

En ce qui concerne l'assertion 3, on déduit de ce qui précède que le cardinal de l'ensemble $\left\{e, x, \cdots, x^{m-1}\right\}$ est m. Puisqu'il est contenu dans $< x >$, qui est aussi d'ordre m, cela entraîne l'égalité annoncée.

Exercice 10. Soient x et y deux éléments de G. On suppose qu'il existe $g \in G$ tel que l'on ait $y = gxg^{-1}$ (deux tels éléments x et y sont dits conjugués). Montrer que x et y ont le même ordre.
Il en résulte que pour tous a et b dans G, les ordres de ab et de ba sont égaux : on a en effet l'égalité $a^{-1}(ab)a = ba$. Signalons que l'ordre de ab n'est pas en général le produit des ordres de a et de b [13].

Exercice 11. Quel est l'ordre de $\overline{2}$ dans le groupe additif $(\mathbb{Z}/10\mathbb{Z}, +)$?

Exercice 12. Soient G_1 et G_2 deux groupes finis. Soit (x, y) un élément du groupe produit $G_1 \times G_2$. Montrer que l'ordre de (x, y) est le ppcm de l'ordre de x et de celui de y.

Théorème 2.3. *Supposons G fini d'ordre n. Alors, pour tout $x \in G$, on a $x^n = e$.*

Démonstration : Soient x un élément de G et m son ordre. Il existe un entier k tel que $n = km$. Par suite, on a $x^n = (x^m)^k = e$ (th. 2.2).

[13] Soient a et b deux éléments de G. On suppose que a et b commutent et que l'intersection des sous-groupes de G engendrés respectivement par a et b est réduite à l'élément neutre. Sous ces hypothèses, l'ordre de ab est le ppcm des ordres de a et b. En effet, notons α et β les ordres de a et b respectivement, et δ l'ordre de ab. On a $(ab)^\delta = e$. Puisque $ab = ba$, on obtient l'égalité $a^\delta = c^\delta$, où c est l'inverse de b. Ainsi, $a^\delta = c^\delta$ appartient à l'intersection du sous-groupe de G engendré par a et de celui engendré par b. Par suite, on a $a^\delta = c^\delta = e$. Cela implique que δ est un multiple de α et β, donc aussi du ppcm de α et β. Par ailleurs, on a $(ab)^{\mathrm{ppcm}(\alpha,\beta)} = e$ (car $ab = ba$). On a donc $\delta = \mathrm{ppcm}(\alpha, \beta)$. En particulier, si G est abélien et si les ordres de a et de b sont premiers entre eux, alors l'ordre de ab est le produit des ordres de a et de b (vérifier cette dernière assertion).

Remarques 2.3.

1) La démonstration du théorème 2.3 se simplifie notablement si G est un groupe abélien. En effet, supposons G abélien d'ordre n et prouvons que pour tout $x \in G$, on a $x^n = e$. Soit x un élément de G. Puisque l'application de G dans G qui à g associe gx est une bijection, on a l'égalité

$$\prod_{g \in G} gx = \prod_{g \in G} g.$$

Il convient de noter ici que les produits ne dépendent pas de l'ordre choisi des éléments car G est abélien. En utilisant de nouveau cette hypothèse, on obtient

$$x^n \prod_{g \in G} g = \prod_{g \in G} g.$$

En multipliant les deux membres de cette égalité par l'inverse du produit des éléments de G, on en déduit l'égalité annoncée.

2) On peut déduire de l'alinéa précédent (i.e. du théorème 2.3 démontré seulement dans le cas abélien) une autre démonstration des assertions 2 et 3 du théorème 2.2. En effet, soit x un élément de G d'ordre m. Le sous-groupe $< x >$ étant abélien d'ordre m, on a donc $x^m = e$ (alinéa 1). Soit alors k le plus petit entier ≥ 1 tel que $x^k = e$. Vérifions que l'on a $k = m$. Il résulte du lemme 2.2 que l'on a $< x >= \left\{e, x, \cdots, x^{k-1}\right\}$: soient N un entier et x^N un élément de $< x >$. Il existe q et r dans $\mathbb{Z}$ tels que $N = kq + r$ avec $0 \leq r < k$, d'où $x^N = x^r$ et l'assertion. D'après le caractère minimal de k, les éléments x^j pour $0 \leq j \leq k-1$ sont distincts deux à deux, par suite l'ordre de $< x >$ est k, d'où $k = m$. On a en particulier $< x >= \left\{e, x, \cdots, x^{m-1}\right\}$.

3) Les alinéas 1 et 2 fournissent une autre démonstration du théorème 2.3 : si x est un élément de G d'ordre m, on a $x^m = e$ (alinéa 2) et le théorème de Lagrange entraîne alors l'égalité $x^n = e$.

Exercice 13. Supposons G abélien. Soit P le produit de ses éléments. Montrer que l'on a $P^2 = e$. Calculer P si $G = \mathbb{Z}/n\mathbb{Z}$ avec $n \geq 1$ (auquel cas il s'agit ici d'une somme).

Les énoncés suivants sont souvent utilisés en pratique pour déterminer l'ordre d'un élément dans un groupe (abélien ou non).

Proposition 2.8. *Soit x un élément de G d'ordre m.*

1) Soit k un entier tel que $x^k = e$. Alors, m divise k.

2) Pour tout entier k, l'ordre de x^k est $\frac{m}{\mathrm{pgcd}(m,k)}$.

Démonstration : 1) Il existe des entiers q et r tels que l'on ait $k = mq + r$ avec $0 \leq r < m$. On a donc les égalités $x^k = (x^m)^q x^r = x^r = e$. On en déduit $r = 0$ (th. 2.2), d'où l'assertion 1.

2) Posons $d = \text{pgcd}(m,k)$. On a d'abord $(x^k)^{m/d} = (x^m)^{k/d} = e$. Considérons un entier $u \geq 1$ tel que $(x^k)^u = e$. D'après l'assertion 1, m divise uk et donc m/d divise uk/d. Les entiers m/d et k/d étant premiers entre eux (lemme 1.3), on déduit du lemme de Gauss que m/d divise u (th. 1.7). En particulier, on a $m/d \leq u$, d'où le résultat.

Proposition 2.9. *Soient x un élément de G et m un entier ≥ 1. Les conditions suivantes sont équivalentes :*

1) l'ordre de x est m.

2) On a $x^m = e$, et pour tout diviseur premier p de m on a $x^{\frac{m}{p}} \neq e$.

Démonstration : Le fait que la première condition entraîne la seconde est une conséquence directe du théorème 2.2. Inversement, supposons la condition 2 réalisée. Notons δ l'ordre de x. Il existe un entier $k \geq 1$ tel que l'on ait $m = \delta k$: on a $m = \delta k + r$ avec $k \in \mathbb{Z}$ et $0 \leq r < \delta$, d'où $x^r = e$ puis $r = 0$. Supposons $k \geq 2$. Soit p un diviseur premier de k. On a alors les égalités

$$x^{\frac{m}{p}} = \left(x^{\delta}\right)^{\frac{k}{p}} = e,$$

ce qui contredit l'hypothèse faite. Par suite, on a $k = 1$, puis $m = \delta$ et le résultat.

Corollaire 2.2. *Soient n un entier naturel non nul et a un entier tels que $0 \leq a \leq n-1$. Dans le groupe additif $\left(\mathbb{Z}/n\mathbb{Z}, +\right)$, l'ordre de $\overline{a}$ est $\frac{n}{\text{pgcd}(n,a)}$.*

Démonstration : Puisque $\mathbb{Z}/n\mathbb{Z} =< \overline{1} >$, l'ordre de $\overline{1}$ est n. (Pour vérifier que $\overline{1}$ est d'ordre n, on peut aussi remarquer que si $k \geq 1$ est tel que $k\overline{1} = \overline{0}$, on a $\overline{k} = \overline{0}$, donc n divise k, et par ailleurs, on a $n\overline{1} = \overline{0}$). L'égalité $\overline{a} = a\overline{1}$ et l'assertion 2 de la proposition 2.8 (en notation additive) entraînent alors le résultat.

6. Groupes cycliques - Fonction indicatrice d'Euler

Soit G un groupe multiplicatif fini d'ordre n, d'élément neutre e.

Définition 2.7. *On dit que G est cyclique s'il existe un élément $x \in G$ tel que $G =< x >$. Un tel élément x s'appelle un générateur de G.*

Un groupe cyclique est en particulier abélien.

Lemme 2.3. *Le groupe G est cyclique si et seulement si il existe $x \in G$ d'ordre n.*

Démonstration : Supposons G cyclique. Il existe $x \in G$ tel que l'on ait $G =< x >$, et en particulier x est d'ordre n. Inversement, s'il existe $y \in G$ d'ordre n, l'ordre du sous-groupe $< y >$ est n, d'où $G =< y >$.

Remarque 2.4. Pour tout $x \in G$ d'ordre n, on a $G = \left\{e, x, \cdots, x^{n-1}\right\}$ (th. 2.2).

Exemples 2.6. Soit m un entier naturel non nul.

1) Le groupe additif $\left(\mathbb{Z}/m\mathbb{Z}, +\right)$ est cyclique d'orde m.

2) L'ensemble $\left\{ \exp\left(\frac{2k\pi i}{m}\right) \mid 0 \leq k \leq m-1 \right\}$ est un sous-groupe cyclique de $\mathbb{C}^*$ d'ordre m et $\exp\left(\frac{2\pi i}{m}\right)$ en est un générateur. C'est le sous-groupe des racines m-ièmes de l'unité de $\mathbb{C}^*$. Nous verrons plus loin que cet exemple et le précédent sont « essentiellement » les mêmes, et que plus généralement, pour tout entier $m \geq 1$, il y a en un sens précis à définir un unique groupe cyclique d'ordre m.

3) Le groupe $\left(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}, +\right)$ est cyclique d'ordre 6, dont un générateur est $(\overline{1}, \overline{1})$[14].

4) Tout groupe fini d'ordre un nombre premier p est cyclique. En effet, si x est un élément distinct de l'élément neutre, l'ordre du sous-groupe engendré par x divise p (th. de Lagrange), il est donc égal à p. En particulier, tous les éléments autres que l'élément neutre sont générateurs.

Exercice 14. Montrer que tout sous-groupe fini de $(\mathbb{C}^*, \times)$ est cyclique.

Théorème 2.4. *Supposons G cyclique d'ordre n.*

1) Tout sous-groupe de G est cyclique.

2) Pour tout diviseur $d \geq 1$ de n, l'ensemble

$$H_d = \left\{ x \in G \mid x^d = e \right\}$$

est un sous-groupe de G d'ordre d.

3) L'application qui à d associe H_d est une bijection entre l'ensemble des diviseurs positifs de n et l'ensemble des sous-groupes de G. En particulier, pour tout diviseur d de n, H_d est l'unique sous-groupe d'ordre d de G.

Démonstration : Soit x un générateur de G.

1) Soit H un sous-groupe de G. Soit δ le plus petit entier ≥ 1 tel que x^δ appartienne à H. Le sous-groupe de G engendré par x^δ est contenu dans H. Montrons qu'il est égal à H. Soit y un élément de H. Il existe un entier m tel que l'on ait $y = x^m$. Par ailleurs, il existe des entiers q et r tels que l'on ait $m = \delta q + r$, avec $0 \leq r < \delta$. On en déduit que x^r est dans H, et donc que r est nul. D'où $m = \delta q$, et $y = (x^\delta)^q$ appartient à $< x^\delta >$, d'où l'assertion.

[14] Soient G_1 et G_2 deux groupes cycliques d'ordre respectivement n_1 et n_2. Alors, le groupe produit $G_1 \times G_2$ est cyclique si et seulement si on a $\text{pgcd}(n_1, n_2) = 1$. En effet, supposons $G_1 =< x_1 >$ et $G_2 =< x_2 >$. L'élément (x_1, x_2) est d'ordre le ppcm de n_1 et n_2 (exercice 12). Si n_1 et n_2 sont premiers entre eux, cet élément est donc d'ordre $n_1 n_2$, i.e. l'ordre du groupe $G_1 \times G_2$, qui est donc cyclique. Inversement, supposons $G_1 \times G_2$ cyclique. Si (x_1, x_2) un générateur de $G_1 \times G_2$, on constate que l'on a $G_1 =< x_1 >$ et $G_2 =< x_2 >$. Par suite, x_1 est d'ordre n_1 et x_2 est d'ordre n_2 et l'ordre de (x_1, x_2) est le ppcm de n_1 et n_2. On a donc $n_1 n_2 = \text{ppcm}(n_1, n_2)$, ce qui entraîne l'égalité $\text{pgcd}(n_1, n_2) = 1$ (prop. 1.5).

2) Soit d un diviseur ≥ 1 de n. L'ensemble H_d est un sous-groupe de G. En effet, e appartient à H_d, pour tous $a, b \in H_d$, on a $(ab)^d = a^d b^d = e$ (car G est abélien) et compte tenu de (2), on a $(a^{-1})^d = a^{-d} = (a^d)^{-1} = e$, de sorte que ab et a^{-1} sont dans H_d. On a

$$\left(x^{\frac{n}{d}}\right)^d = x^n = e,$$

par suite, $x^{n/d}$ appartient à H_d. L'élément $x^{n/d}$ est d'ordre $\frac{n}{\mathrm{pgcd}(n/d,n)} = d$ (prop. 2.8), ainsi d divise l'ordre de H_d (th. de Lagrange). Par ailleurs, H_d est cyclique (assertion 1). Si y est un générateur de H_d on a $y^d = e$, donc l'ordre de y, qui est celui de H_d, divise d (prop. 2.8), d'où le fait que H_d soit d'ordre d.

3) Soit H un sous-groupe de G. Vérifions que l'on a $H = H_d$ où d est l'ordre de H, ce qui prouvera que l'application considérée est une surjection. Pour tout $z \in H$ on a $z^d = e$, donc H est contenu dans H_d. Puisque H_d est d'ordre d, on a donc $H = H_d$. Il reste à montrer que cette application est une injection : soient d et d' deux diviseurs positifs de n tels que $H_d = H_{d'}$. Les groupes H_d et $H_{d'}$ ont le même ordre, d'où $d = d'$ et le résultat.

Une question importante concerne la description des générateurs d'un groupe cyclique : en particulier, combien il y a-t-il de générateurs dans un groupe cyclique d'ordre n ? Introduisons pour cela la fonction indicatrice d'Euler[15] φ, qui est définie sur l'ensemble des entiers naturels ≥ 1 à valeurs dans $\mathbb{N}$.

Définition 2.8. (Fonction indicatrice d'Euler) *Pour tout entier $m \geq 1$, l'entier $\varphi(m)$ est le nombre des entiers naturels plus petits que m et premiers avec m. Autrement dit, $\varphi(m)$ est le nombre des entiers k pour lesquels on a :*

$$1 \leq k \leq m \quad et \quad \mathrm{pgcd}(k, m) = 1.$$

Par exemple, on a $\varphi(1) = 1$, $\varphi(2) = 1$, $\varphi(3) = 2$, et pour tout nombre premier p, on a $\varphi(p) = p - 1$. Plus généralement :

Lemme 2.4. *Pour tout nombre premier p et tout entier naturel $r \geq 1$, on a*

$$\varphi(p^r) = p^r - p^{r-1} = p^r\left(1 - \frac{1}{p}\right).$$

Démonstration : Les entiers naturels plus petits que p^r et premiers avec p^r sont ceux plus petits que p^r qui ne sont pas multiples de p. Cela entraîne l'assertion, vu qu'il existe exactement p^{r-1} multiples de p compris entre 1 et p^r.

[15] Leonhard Euler était un mathématicien suisse. Il est né à Bâle en 1707 et décède à Saint-Petersbourg en 1783. Il apporta d'importantes contributions en théorie des nombres et en analyse. Il établit sa renommée en calculant la somme des inverses des carrés des entiers, en démontrant l'égalité $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$, où n parcourt les entiers ≥ 1.

Nous déterminerons dans le chapitre IV l'entier $\varphi(m)$ pour tout m. Signalons la formule suivante :

Lemme 2.5. *Pour tout entier $m \geq 1$, on a l'égalité*

$$m = \sum_{d|m} \varphi(d).$$

Démonstration : Considérons l'ensemble $F = \left\{\frac{1}{n}, \frac{2}{n}, \cdots, \frac{n-1}{n}, \frac{n}{n} = 1\right\}$. Pour tout diviseur d de n, posons $F_d = \left\{\frac{a}{d} \mid 1 \leq a \leq d \text{ et } \operatorname{pgcd}(a,d) = 1\right\}$. L'ensemble F est la réunion disjointe des F_d, où d parcourt l'ensemble des diviseurs (positifs) de n. En effet, tout élément de $\frac{a}{d} \in F_d$ s'écrit sous la forme $\frac{ka}{n}$ avec $kd = n$, qui appartient à F (car $a \leq d$), et inversement, chaque élément de F a un représentant irréductible dans l'un des F_d. Par ailleurs, si $\frac{a}{d} = \frac{a'}{d'}$ est dans $F_d \cap F_{d'}$, on a $d'a = a'd$, ce qui conduit à $a = a'$ et $d = d'$ car on a $\operatorname{pgcd}(a,d) = \operatorname{pgcd}(a',d') = 1$ (lemme de Gauss), d'où notre assertion. Cela entraîne le résultat vu que le cardinal de F est n et que celui de F_d est $\varphi(d)$.

Théorème 2.5. *Supposons G cyclique d'ordre n. Soit x un générateur de G. Alors, l'ensemble des générateurs de G est*

$$\left\{x^k \mid 1 \leq k \leq n \text{ et } \operatorname{pgcd}(k,n) = 1\right\}.$$

En particulier, G possède exactement $\varphi(n)$ générateurs.

Démonstration : On a $G = \left\{x, \cdots, x^{n-1}, x^n\right\}$. Pour tout k compris entre 1 et n, l'ordre de x^k est $\frac{n}{\operatorname{pgcd}(n,k)}$ (prop. 2.8). Par suite, x^k est d'ordre n si et seulement si on a $\operatorname{pgcd}(n,k) = 1$, ce qui établit le résultat.

Corollaire 2.3. *Supposons G cyclique d'ordre n. Pour tout diviseur positif d de n, il y a exactement $\varphi(d)$ éléments d'ordre d dans G*[16].

Démonstration : Il existe un unique sous-groupe H_d d'ordre d de G, qui n'est autre que l'ensemble des $x \in G$ tels que $x^d = e$ (th. 2.4). Par suite, l'ensemble des éléments d'ordre d de G est contenu dans H_d, et cet ensemble est formé des générateurs de H_d. Puisque H_d est cyclique (th. 2.4), il possède exactement $\varphi(d)$ générateurs (th. 2.5), d'où le résultat.

[16] Ce résultat permet de retrouver l'égalité du lemme 2.5. En effet, pour tout diviseur d de n, soit Φ_d l'ensemble des éléments de G d'ordre d. Le groupe G est la réunion disjointe des Φ_d, où d parcourt l'ensemble des diviseurs de n. Le cardinal de Φ_d étant $\varphi(d)$ (cor. 2.3), l'égalité s'en déduit aussitôt.

Corollaire 2.4. *Soit n un entier naturel non nul. L'ensemble des générateurs du groupe $(\mathbb{Z}/n\mathbb{Z}, +)$ est $\left\{\overline{a} \mid 1 \leq a \leq n \text{ et } \operatorname{pgcd}(a,n) = 1\right\}$.*

Démonstration : On sait que $\overline{1}$ est un générateur de $\mathbb{Z}/n\mathbb{Z}$. Compte tenu du théorème 2.5, cela entraîne l'assertion.

Exercice 15. Pour tout $n \geq 1$, montrer que $\varphi(n)$ divise $n!$.

Exercice 16. Soit n un entier≥ 1. Notons $d(n)$ le nombre de diviseurs positifs de n.

1) Déterminer $d(n)$ en fonction de la décomposition en facteurs premiers de n.
2) Trouver tous les entiers $n \leq 30$ tels que $d(n) = \varphi(n)$. (On peut démontrer que pour $n > 30$, on a $\varphi(n) > d(n)$).

La fonction indicatrice d'Euler a suscité de nombreux travaux et il reste encore beaucoup de problèmes non résolus à son sujet[17].

Indiquons une démonstration du théorème suivant, dû à Euler, qui apparaîtra de nouveau dans le chapitre IV, et qui généralise le petit théorème de Fermat (exercice 2 du chap. I) :

Théorème 2.6. (Euler) *Soient a et n deux entiers naturels non nuls premiers entre eux. On a la congruence*

$$a^{\varphi(n)} \equiv 1 \text{ mod. } n.$$

Démonstration : On peut supposer $n \geq 2$. Posons $t = \varphi(n)$. Notons $b_1, \cdots, b_t$ les entiers compris entre 1 et n et premiers avec n. Pour tout $i = 1, \cdots, t$, il existe des entiers q_i et r_i tels que l'on ait $ab_i = nq_i + r_i$ avec $0 \leq r_i < n$. On a la congruence

$$\prod_{i=1}^{t} ab_i \equiv \prod_{i=1}^{t} r_i \text{ mod. } n.$$

Par ailleurs, les entiers ab_i étant premiers avec n, on a $\operatorname{pgcd}(r_i, n) = 1$ et $1 \leq r_i \leq n$ (on a $r_i \neq 0$ car $n \geq 2$). Vérifions alors que l'on a

$$\left\{b_1, \cdots, b_t\right\} = \left\{r_1, \cdots, r_t\right\}.$$

[17] Citons-en deux : le problème de Lehmer. On a vu que si p premier, on a $\varphi(p) = p-1$. En 1932, Lehmer, a posé la question suivante : existe-t-il des entiers n, qui ne sont pas premiers, tels que $\varphi(n)$ divise $n-1$? On conjecture que non. Soit $\omega(n)$ le nombre de diviseurs premiers de n. Lehmer a démontré que si un tel entier n existe, alors n est impair sans facteurs carrés (pour tout p premier, on a $v_p(n) \leq 1$) et $\omega(n) \geq 7$. Ce résultat a été amélioré par la suite. Sans être exhaustif, signalons qu'il a été démontré en 1980, que pour un tel entier n on a $n > 10^{20}$ et $\omega(n) \geq 14$. En fait, si $\varphi(n)$ divise $n-1$, alors n est un nombre de Carmichael (c'est une conséquence du théorème d'Euler ci-dessus).
Voici une autre conjecture sur la fonction φ, qui a été énoncée par Carmichael, et qui n'est toujours pas élucidée : pour tout entier pair $n \geq 1$, il existe un entier $m \neq n$ tel que $\varphi(n) = \varphi(m)$ (si n est impair, on a en fait $\varphi(n) = \varphi(2n)$ comme on le verra plus loin).

Il suffit pour cela de montrer que si i et j sont distincts, on a $r_i \neq r_j$. Dans le cas contraire, on aurait $ab_i \equiv ab_j$ mod. n i.e. n diviserait $a(b_i - b_j)$. Puisque a et n sont premiers entre eux, d'après le lemme de Gauss, $b_i - b_j$ serait divisible par n, et l'on aurait $b_i = b_j$, d'où une contradiction. On en déduit les congruences

$$\prod_{i=1}^{t} ab_i \equiv \prod_{i=1}^{t} b_i \text{ mod. } n \quad \text{i.e.} \quad (a^t - 1)\prod_{i=1}^{t} b_i \equiv 0 \text{ mod. } n.$$

Puisque les b_i sont premiers avec n, il en est de même du produit des b_i. Le lemme de Gauss entraîne alors le résultat.

7. Homomorphismes de groupes

Sauf précision contraire, les groupes considérés dans ce paragraphe sont implicitement supposés multiplicatifs.

Définition 2.9. *Soient G et G' deux groupes. On appelle homomorphisme (de groupes), ou morphisme (de groupes), de G dans G', toute application $f : G \to G'$ telle que l'on ait*

$$f(xy) = f(x)f(y). \tag{15}$$

Remarque 2.5. Cette formule convient évidemment d'être modifiée si les lois de composition de G et G' ne sont pas notées multiplicativement. Par exemple, si les lois de G et G' sont notées additivement, la formule (15) s'écrit alors $f(x+y) = f(x)+f(y)$, et si la loi de G est multiplicative et celle de G' additive, cette formule devient $f(xy) = f(x)+f(y)$.

Exemples 2.7.

1) Soient $\mathbb{R}_+^*$ le sous-groupe de $\mathbb{R}^*$ formé des nombres réels strictement positifs et $\log : \mathbb{R}_+^* \to \mathbb{R}$ la fonction logarithme Népérien. La formule bien connue en Analyse

$$\log(xy) = \log(x) + \log(y) \quad \text{quels que soient } x, y > 0,$$

définit un homomorphisme de $(\mathbb{R}_+^*, \times)$ à valeurs dans $(\mathbb{R}, +)$. De même la fonction exponentielle définit un homomorphisme de $(\mathbb{R}, +)$ à valeurs dans $(\mathbb{R}_+^*, \times)$.

2) Soit G un groupe multiplicatif. Pour tout $a \in G$, l'application $f_a : \mathbb{Z} \to G$ définie par $f(n) = a^n$ est homomorphisme du groupe $(\mathbb{Z}, +)$ à valeurs dans G. Cela résulte de la première égalité de (2). En fait, pour tout homomorphisme f de $\mathbb{Z}$ dans G, il existe $a \in G$ tel que $f = f_a$, comme on le constate en posant $f(1) = a$.

3) Pour tout $n \geq 1$, la surjection canonique $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ est un homomorphisme de $(\mathbb{Z}, +)$ à valeurs dans le groupe additif $(\mathbb{Z}/n\mathbb{Z}, +)$ (formule (11)). Plus généralement, pour tout groupe abélien additif G et tout sous-groupe H de G, l'application $s : G \to G/H$

définie par $s(x) = x + H$ est un homomorphisme de groupes. Cela résulte de la définition de la loi de groupe sur G/H.

Lemme 2.6. *Soit $f : G \to G'$ un homomorphisme de groupes. Soient e et e' les éléments neutres de G et G' respectivement.*

1) On a $f(e) = e'$.

2) Pour tout $x \in G$, on a $f(x^{-1}) = f(x)^{-1}$.

Démonstration : Pour tout $x \in G$, on a $f(x) = f(xe) = f(x)f(e)$. Par suite, on a $f(x)^{-1}f(x) = f(x)^{-1}f(x)f(e)$, ce qui conduit à $e' = f(e)$. Par ailleurs, compte tenu de la première assertion, on a les égalités

$$e' = f(xx^{-1}) = f(x)f(x^{-1}) \quad \text{et} \quad e' = f(x^{-1}x) = f(x^{-1})f(x),$$

d'où le lemme.

Lemme 2.7. *Soient $f : M \to N$ et $g : N \to P$ des homomorphismes de groupes. Alors, l'application composée $g \circ f : M \to P$ est encore un homomorphisme de groupes. Si un homomorphisme de groupe $f : M \to N$ est une bijection de M sur N, alors l'application réciproque est aussi un homomorphisme de groupes.*

Démonstration : Soient x et y des éléments de M. On a les égalités

$$(g \circ f)(xy) = g\big(f(xy)\big) = g\big(f(x)f(y)\big) = g\big(f(x)\big)g\big(f(y)\big),$$

d'où la première assertion. En ce qui concerne la seconde, considérons deux éléments u et v de N. Il s'agit de montrer que l'on a

$$f^{-1}(uv) = f^{-1}(u)f^{-1}(v). \tag{16}$$

Puisque f est une bijection de M sur N, c'est en particulier une injection, et suffit donc de montrer que les images par f des deux membres de (16) sont égales, autrement dit que l'on a $uv = f\big(f^{-1}(u)f^{-1}(v)\big)$, ce qui résulte du fait que f soit un homomorphisme.

Définition 2.10. *Soient G et G' deux groupes. On appelle isomorphisme de G sur G', tout homomorphisme bijectif de G sur G'. On dit que G et G' sont isomorphes s'il existe un isomorphisme de G sur G'. On appelle automorphisme de G tout isomorphisme de G sur lui-même.*

Compte tenu du lemme 2.7, s'il existe un isomorphisme de G sur G', il en existe aussi un de G' sur G, à savoir l'isomorphisme réciproque.

Remarque 2.7. Il est important de noter que deux groupes isomorphes possèdent exactement les mêmes propriétés. La connaissance d'un isomorphisme f entre deux groupes

G et G' permet de traduire explicitement toute assertion relative à G en la même relative à G'. Par exemple, si G est cyclique engendré par g, alors G' est cyclique engendré par $f(g)$. Si $x \in G$ est d'ordre k, alors $f(x)$ est aussi d'ordre k, etc...

Exemples 2.8.

1) La fonction logarithme réalise un isomorphisme du groupe multiplicatif $\mathbb{R}_+^*$ sur le groupe additif $\mathbb{R}$.

2) Soient G un groupe et a un élément de G. L'application $f_a : G \to G$ définie par $f_a(x) = axa^{-1}$ est un automorphisme de G sur G (exercice). Les f_a où $a \in G$ s'appellent les automorphismes intérieurs de G.

3) Décrivons, à isomorphisme près, les groupes cycliques d'ordre n.

Lemme 2.8. *Tout groupe cyclique d'ordre n est isomorphe au groupe additif $\big(\mathbb{Z}/n\mathbb{Z}, +\big)$.*

Démonstration : Soit G un groupe cyclique d'ordre n. Choisissons un générateur g de G. On a $G = \big\{e, g, \cdots, g^{n-1}\big\}$. Considérons l'application $f : G \to \mathbb{Z}/n\mathbb{Z}$ définie par

$$f\big(g^k\big) = \overline{k} \quad \text{pour tout } k \in \mathbb{Z}.$$

Elle est bien définie car si $g^k = g^{k'}$, alors n divise $k - k'$ i.e. $\overline{k} = \overline{k'}$. Par ailleurs, on a

$$f\big(g^{k+k'}\big) = \overline{k + k'} = \overline{k} + \overline{k'} = f\big(g^k\big) + f\big(g^{k'}\big),$$

donc f est un homomorphisme. Il est surjectif (par définition), il est donc aussi injectif car G et $\mathbb{Z}/n\mathbb{Z}$ ont le même ordre (on peut aussi le vérifier directement), d'où le lemme.

En particulier, deux groupes cycliques sont isomorphes si et seulement si ils ont le même ordre. Puisque tout groupe fini d'ordre un nombre premier est cyclique, on en déduit l'énoncé suivant :

Corollaire 2.5. *Tout groupe fini d'ordre un nombre premier p est isomorphe au groupe additif $\big(\mathbb{Z}/p\mathbb{Z}, +\big)$.*

4) Démontrons que tout groupe d'ordre 4 est isomorphe à $\big(\mathbb{Z}/4\mathbb{Z}, +\big)$ ou bien au groupe produit $\big(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, +\big)$. Soit G un groupe d'ordre 4. Supposons qu'il ne soit pas cyclique. Soit x un élément de G autre que l'élément neutre e. D'après le théorème de Lagrange, l'ordre δ de x divise 4. On a donc $\delta = 2$ (car si $\delta = 1$, on a $x = e$, et si $\delta = 4$, G est cyclique engendré par x). Il existe dans G un élément y distinct de e et x (car G est d'ordre 4), et nécessairement y est d'ordre 2. On a ainsi $G = \{e, x, y, xy\}$. L'application $f : G \to \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ définie par $f(e) = (0, 0)$, $f(x) = (1, 0)$, $f(y) = (0, 1)$ et $f(xy) = (1, 1)$, réalise alors un isomorphisme de groupes de G sur le groupe produit $(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, +)$, d'où le résultat. Les groupes $\big(\mathbb{Z}/4\mathbb{Z}, +\big)$ et $\big(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, +\big)$ n'étant pas isomorphes ($\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ n'est pas cyclique car ses éléments autres que l'élément neutre sont d'ordre 2), il y a donc, à isomorphisme près, exactement deux groupes d'ordre 4.

5) Si X est un ensemble de cardinal n, le groupe symétrique $\mathbb{S}(X)$ de X est isomorphe au groupe S_n des bijections de $\{1, 2, \cdots, n\}$. Plus précisément, si $f : X \to Y$ est une bijection de X sur un ensemble Y, alors l'application $\psi : \mathbb{S}(X) \to \mathbb{S}(Y)$ définie par l'égalité $\psi(g) = fgf^{-1}$ est un isomorphisme de groupes.

Noyau et image d'un homomorphisme

Lemme 2.9. *Soit f un homomorphisme d'un groupe G dans un groupe G'.*

1) *Pour tout sous-groupe H de G, l'image $f(H)$[18] de H par f est un sous-groupe de G'.*
2) *Pour tout sous-groupe H' de G', l'image réciproque $f^{-1}(H')$[19] de H' par f est un sous-groupe de G.*

Démonstration : La première assertion est laissée au lecteur en exercice. Démontrons la seconde. Notons e et e' les éléments neutres de G et G' respectivement. Soit H'un sous-groupe de G'. On a $f(e) = e' \in H'$ donc e appartient à $f^{-1}(H')$. Par ailleurs, si x et y sont dans $f^{-1}(H')$, alors $f(x)$ et $f(y)$ sont dans H' et $f(xy) = f(x)f(y)$ appartient aussi à H', d'où $xy \in f^{-1}(H')$. De même, on a $f(x^{-1}) = f(x)^{-1} \in H'$, donc $x^{-1} \in f^{-1}(H')$, d'où l'assertion.

Définition 2.11. *Soit $f : G \to G'$ un homomorphisme de groupes. On appelle image de f le sous-groupe $f(G)$ de G'. On appelle noyau de f le sous-groupe $f^{-1}\big(\{e'\}\big)$ de G (où e' est l'élément neutre de G'), on le note souvent* $\mathrm{Ker}(f)$. *On a donc*

$$\mathrm{Ker}(f) = \Big\{x \in G \mid f(x) = e'\Big\}.$$

Lemme 2.10. *Soit f un homomorphisme d'un groupe G dans un groupe G'. Pour que f soit injectif il faut et il suffit que* $\mathrm{Ker}(f)$ *soit réduit à l'élément neutre de G.*

Démonstration : Supposons f injectif. Soit x un élément de $\mathrm{Ker}(f)$. On a les égalités $f(x) = e' = f(e)$, d'où $x = e$. Inversement, supposons $\mathrm{Ker}(f) = \{e\}$. Soient x et y deux éléments de G tels que $f(x) = f(y)$. On a$f(x)f(y)^{-1} = e'$ i.e. $f(xy^{-1}) = e'$, d'où $xy^{-1} = e$ puis $x = y$.

---

[18] Soit $f : E \to F$ une application entre deux ensembles. Rappelons que pour toute partie A de E, l'image (directe) de A par f, que l'on note $f(A)$, est l'ensemble des $f(x)$ tels que x soit dans A. Pour toute famille de parties A_i de E, on a $f\big(\bigcup A_i\big) = \bigcup f(A_i)$. L'égalité analogue obtenue en remplaçant « réunion » par « intersection » est fausse en général.

[19] Rappelons que pour toute partie B de F, l'image réciproque de B par f, que l'on note $f^{-1}(B)$, est l'ensemble des $x \in E$ tels que $f(x)$ appartienne à B. Pour toute famille de parties B_i de F, on a les égalités $f^{-1}\big(\bigcup B_i\big) = \bigcup f^{-1}(B_i)$ et $f^{-1}\big(\bigcap B_i\big) = \bigcap f^{-1}(B_i)$.

Théorème 2.7. *Soient G un groupe abélien additif et f un homomorphisme de G dans un groupe G'. Alors, le groupe quotient $G/\operatorname{Ker}(f)$ est isomorphe à $f(G)$, via l'application qui à $x + \operatorname{Ker}(f)$ associe $f(x)$.*

Démonstration : Soient x et y deux éléments de G tels que $x-y$ appartienne à $\operatorname{Ker}(f)$. On a $f(x-y) = e'$, autrement dit, on a $f(x) = f(y)$. On obtient ainsi une application

$$h : G/\operatorname{Ker}(f) \to f(G)$$

définie pour tout $x \in G$ par l'égalité

$$h\big(x + \operatorname{Ker}(f)\big) = f(x).$$

C'est un homomorphisme. En effet, quels que soient $x, y \in G$, on a les égalités

$$h\big(x + y + \operatorname{Ker}(f)\big) = f(x+y) = f(x)f(y) = h\big(x + \operatorname{Ker}(f)\big)h\big(y + \operatorname{Ker}(f)\big).$$

Par ailleurs, si $f(x) = e'$, x appartient à $\operatorname{Ker}(f)$, de sorte que $x + \operatorname{Ker}(f) = \operatorname{Ker}(f)$. Cela prouve que le noyau de h est réduit à l'élément neutre de $G/\operatorname{Ker}(f)$ et donc que h est injectif (lemme 2.10). Par définition, f est aussi une surjection de $G/\operatorname{Ker}(f)$ sur $f(G)$, d'où le résultat.

Exemple 2.9. Soient G un groupe cyclique d'ordre n et x un générateur de G. Soit $f : \mathbb{Z} \to G$ l'application définie par $f(k) = x^k$. C'est un homomorphisme surjectif. Son noyau est formé des entiers $k \in \mathbb{Z}$ tels que $x^k = e$. Puisque x est un générateur de G, il en résulte que $\operatorname{Ker}(f)$ est formé des entiers multiples de n, autrement dit, on a $\operatorname{Ker}(f) = n\mathbb{Z}$. D'après le théorème 2.7, on en déduit que les groupes $\mathbb{Z}/n\mathbb{Z}$ et G sont isomorphes. Cela fournit une autre démonstration du résultat énoncé dans le lemme 2.8.

Terminons ce chapitre en donnant deux applications du théorème 2.7.

1) **Théorème de Cauchy[20] pour les groupes abéliens finis**

Il s'agit du résultat suivant :

Théorème 2.8. *Soient G un groupe abélien fini d'ordre n et p un diviseur premier de n. Il existe dans G un élément d'ordre p.*

Démonstration : Posons

$$G = \Big\{a_1, \cdots, a_n\Big\}.$$

[20] Augustin Louis Cauchy est né à Paris en 1789 et décède à Sceaux en 1857. Il fut professeur à l'école Polytechnique de 1838 jusqu'à sa mort. On lui doit notamment la mise en place de la théorie des fonctions de variables complexes et son application au calcul intégral.

Pour tout $a_i \in G$, notons α_i son ordre et considérons l'application

$$f : \mathbb{Z}/\alpha_1\mathbb{Z} \times \cdots \times \mathbb{Z}/\alpha_n\mathbb{Z} \to G,$$

définie par

$$f\big((\overline{k_1}, \cdots, \overline{k_n})\big) = a_1^{k_1} \cdots a_n^{k_n}.$$

On vérifie d'abord que f est bien définie, en remarquant que pour tout $i = 1, \cdots, n$, si l'on a $\overline{k_i} = \overline{k_i'}$, alors $a_i^{k_i} = a_i^{k_i'}$. C'est un homomorphisme car G est supposé abélien. Par ailleurs, il est surjectif, car pour tout i, on a $f\big((\overline{0}, \cdots, \overline{1}, \cdots, \overline{0})\big) = a_i$, où $\overline{1}$ est la i-ème composante de l'élément considéré. En notant Γ le groupe produit des $\mathbb{Z}/\alpha_i\mathbb{Z}$, on déduit du théorème 2.7 que G est isomorphe à $\Gamma/\operatorname{Ker}(f)$. On a en particulier l'égalité

$$|\Gamma| = |\operatorname{Ker}(f)|.|G|.$$

Puisque p divise $|G|$, cela entraîne que p divise $|\Gamma| = \alpha_1 \cdots \alpha_n$. Par suite, il existe i tel que p divise α_i. L'élément a_i étant d'ordre α_i, on a

$$a_i^{\frac{\alpha_i}{p}} \neq e \quad \text{et} \quad \left(a_i^{\frac{\alpha_i}{p}}\right)^p = e,$$

ce qui prouve que $a_i^{\frac{\alpha_i}{p}}$ est d'ordre p dans G. D'où le résultat[21].

Le théorème 2.9 vaut en fait pour tous les groupes finis, mais la démonstration du cas général nous entraînerait trop loin. Cela étant, on en déduit l'énoncé suivant :

Corollaire 2.6. *Soit G un groupe abélien d'ordre n sans facteurs carrés, i.e. pour tout nombre premier p on a $v_p(n) \leq 1$. Alors, G est cyclique d'ordre n.*

Démonstration : On a $n = p_1 \cdots p_r$ où les p_i sont des nombres premiers deux à deux distincts. D'après ce qui précède, pour tout $i = 1, \cdots, r$, il existe $a_i \in G$ d'ordre p_i. Par ailleurs, pour tout $s \geq 1$ et $s \leq r$, l'élément $a_1 \cdots a_s$ est d'ordre $p_1 \cdots p_s$ (on vérifie cette assertion par récurrence et on utilise le résultat[13] du bas de la page 35). En particulier, $a_1 \cdots a_r$ est d'ordre n, ce qui établit l'assertion.

Signalons que l'on peut démontrer le résultat suivant : soit n un entier ≥ 1. Pour que tout groupe fini d'ordre n soit cyclique il faut et il suffit que n et $\varphi(n)$ soient premiers entre eux.

---

[21] Indiquons une autre démonstration du théorème 2.8, en procédant par récurrence sur l'ordre n de G. L'énoncé est vrai si $n = 1$ (puisque p ne divise pas 1). Considérons un entier $n \geq 2$ et supposons que le résultat est vrai pour tous les groupes d'ordre $< n$. Il existe dans G un élément x d'ordre $m > 1$. Si p divise m, on a $m = pr$ et l'élément $y = x^r \in G$ est alors d'ordre p. Supposons que p ne divise pas m. Soit H le sous-groupe de G engendré par x. L'ordre du groupe G/H, qui est $n/m < n$, est divisible par p (car p ne divise pas m). D'après l'hypothèse de récurrence, il existe un élément $zH \in G/H$ qui est d'ordre p. L'élément $w = z^m$ est alors un élément de G d'ordre p. En effet, z^p appartient à H et H est d'ordre m, donc on a $w^p = e$. Par ailleurs, on a $w \neq e$: supposons $w = e$. Puisque p ne divise pas m, il existe $a, b \in \mathbb{Z}$ tels que $ap + bm = 1$. Par suite, on a $z = z^{ap} z^{bm} = z^{ap} \in H$ car z^p est dans H. Cela conduit à une contradiction car zH étant d'ordre p dans G/H, l'élément z n'est pas dans H, d'où le résultat.

2) **Puissances dans un groupe cyclique**

Considérons un groupe cyclique G d'ordre n et un élément a de G.

Proposition 2.10. *Soit k un entier naturel. Pour qu'il existe $x \in G$ vérifiant l'égalité $x^k = a$ il faut et il suffit que l'on ait*

$$a^{\frac{n}{d}} = e \quad \text{où} \quad d = \text{pgcd}(k, n).$$

Supposons cette condition réalisée. Soit x_0 un élément de G tel que $x_0^k = a$. Alors, l'ensemble des éléments $x \in G$ tels que $x^k = a$ est

$$\Big\{x_0 z \mid z \in G \text{ et } z^d = e\Big\}.$$

C'est un ensemble de cardinal d.

Démonstration : On considère l'homomorphisme de groupes $\psi : G \to G$ défini par $\psi(x) = x^k$. Vérifions que son noyau est d'ordre d. Soit x un élément de $\text{Ker}(\psi)$. On a $x^k = e$ et $x^n = e$ (th. 2.3), d'où en utilisant le théorème de Bézout, $x^d = e$. On en déduit que les éléments de $\text{Ker}(\psi)$ sont exactement les éléments $x \in G$ pour lesquels on a $x^d = e$. D'après l'assertion 2 du théorème 2.4, on a donc $|\text{Ker}(\psi)| = d$. D'après le théorème 2.7, l'ordre de l'image de ψ est donc n/d. Par suite, si a est dans l'image de ψ, on a $a^{n/d} = e$ (th. 2.3). Inversement, si l'on a l'égalité $a^{n/d} = e$, alors a appartient à l'unique sous-groupe de G d'ordre n/d (th. 2.5), qui est précisément l'image de ψ, d'où la condition annoncée.

On suppose alors qu'il existe $x_0 \in G$ tel que $x_0^k = a$. Si $x \in G$ vérifie l'égalité $x^k = a$, on a $(xx_0^{-1})^k = e$, d'où $x = x_0 z$ avec $z^k = e$, et comme on l'a constaté ci-dessus, on a alors $z^d = e$. Inversement, pour tout $z \in G$ tel que $z^d = e$, on a $(x_0 z)^k = a$ car d divise k, d'où l'ensemble des solutions annoncé. Par ailleurs, il y a exactement d éléments $z \in G$ tels que $z^d = e$ (th. 2.4). Cela établit le résultat.

Exemple 2.10. Prenons pour G le groupe additif $\mathbb{Z}/25\mathbb{Z}$. Déterminons l'ensemble des solutions de l'équation $5x = \overline{15}$. On remarque pour cela que $x_0 = \overline{3}$ est une solution particulière. Par ailleurs, les éléments $x \in G$ qui vérifient $5x = \overline{0}$ sont les classes de 0, 5, 10, 15 et 20. L'ensemble des solutions cherché est donc $\big\{\overline{3}, \overline{8}, \overline{13}, \overline{18}, \overline{23}\big\}$.

Exercice 17. Dans le groupe additif $\mathbb{Z}/1000\mathbb{Z}$, résoudre l'équation $5x = \overline{50}$.

Chapitre III — Anneaux et corps

1. Définition d'un anneau

Définition 3.1. *On appelle anneau un triplet formé d'un ensemble A et de deux lois de composition sur A, une addition $(x, y) \mapsto x + y$ et une multiplication $(x, y) \mapsto xy$, tels que les conditions suivantes soient vérifiées :*

1) le couple $(A, +)$ est un groupe commutatif.

2) La multiplication est associative et possède un élément neutre.

3) La multiplication est distributive par rapport à l'addition, ce qui signifie que l'on a

$$x(y + z) = xy + xz \quad et \quad (x + y)z = xz + yz \quad quels\ que\ soient\ x, y, z \in A.$$

Si de plus la multiplication est commutative, autrement dit, si l'on a $xy = yx$ quels que soient $x, y \in A$, on dit que A est un anneau commutatif.

On notera 0 l'élément neutre de $(A, +)$ et 1, ou 1_A, l'élément neutre de A pour la multiplication. Rappelons que pour tout $x \in A$, il existe un élément de A, noté $-x$, tel que l'on ait $x + (-x) = 0$ ($-x$ est l'opposé de x).

Lemme 3.1. *Quels que soient $x, y, z \in A$, on a*

$$x(y - z) = xy - xz \quad et \quad (y - z)x = yx - zx.$$

Démonstration : D'après la condition 3, on a $x(y - z) + xz = x(y - z + z) = xy$ et $(y - z)x + zx = (y - z + z)x = yx$, d'où le lemme.

Corollaire 3.1. *Quels que soient $x, y \in A$, on a*

$$x0 = 0x = 0, \quad x(-y) = -xy \quad et \quad (-y)x = -yx.$$

En particulier, on a $(-1)x = -x$.

Par convention, on a

$$x^0 = 1 \quad \text{pour tout } x \in A.$$

Un anneau réduit à un élément, i.e. pour lequel on a $1 = 0$, est dit nul.

Exemples 3.1.

1) **L'anneau $\mathbb{Z}$.** En munissant $\mathbb{Z}$ des deux lois de composition usuelles (addition et multiplication) on obtient l'anneau des entiers relatifs, qui est évidemment commutatif. Les ensembles $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ munis de l'addition de la multiplication usuelles sont aussi des anneaux commutatifs.

2) **L'anneau $F(X, A)$.** Soient X un ensemble et A un anneau. Notons $F(X, A)$ l'ensemble des applications de X à valeurs dans A. On définit la somme et le produit de deux éléments $f, g \in F(X, A)$ par les égalités

$$(f + g)(x) = f(x) + g(x) \quad \text{et} \quad (fg)(x) = f(x)g(x) \quad \text{pour tout } x \in A.$$

L'ensemble $F(X, A)$ muni de ces deux lois est un anneau, qui est commutatif si A l'est, et s'appelle l'anneau des applications de X dans A.

3) **L'anneau $\mathbb{M}_n(A)$.** Soient A un anneau, n un entier ≥ 1 et $\mathbb{M}_n(A)$ l'ensemble des matrices carrées de taille (n, n) (on dit aussi d'ordre n) à coefficients dans A. Rappelons que l'on définit sur $\mathbb{M}_n(A)$ une addition et une multiplication comme suit. Soient $M = (a_{ij})$ et $N = (b_{ij})$ deux matrices de $\mathbb{M}_n(A)$. Par définition, le coefficient de la i-ème ligne et de la j-ième colonne de $M + N$ est $a_{ij} + b_{ij}$, et celui de MN est donné par la somme

$$\sum_{k=1}^{n} a_{ik} b_{kj}.$$

Muni de ces deux lois de composition, $\mathbb{M}_n(A)$ est un anneau, qui n'est jamais commutatif si $n \geq 2$ et si A n'est pas nul. Pour $n = 2$, on le constate par exemple en remarquant que l'on a pour tout $x \in A$ les égalités

$$\begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x & x \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x & 1 \\ 0 & 1 \end{pmatrix},$$

et les résultats obtenus sont distincts si $x \neq 1$.

4) **L'anneau $A[X]$.** Soit A un anneau. Rappelons que les polynômes à une indéterminée à coefficients dans A sont les suites $(a_n)_{n \geq 0}$ d'éléments de A qui sont nulles à partir d'un certain rang. Les a_n sont appelés les coefficients du polynôme. Sur cet ensemble de polynômes, on définit les deux lois de compositions suivantes : si $P = (p_0, p_1, \cdots)$ et $Q = (q_0, q_1, \cdots)$, alors l'addition et la multiplication de P et Q sont définies respectivement par les égalités

$$P + Q = (p_0 + q_0, p_1 + q_1, \cdots) \quad \text{et} \quad PQ = (s_0, s_1, \cdots) \quad \text{avec} \quad s_n = \sum_{i+j=n} p_i q_j.$$

On vérifie que l'ensemble des polynômes à coefficients dans A est ainsi muni d'une structure d'anneau, qui est commutatif si A l'est. Pour tout $a \in A$, on note a le polynôme $(a, 0, \cdots, 0, \cdots)$. Posons $X = (0, 1, 0, \cdots, 0, \cdots)$. Pour tout entier $n \geq 1$, et tout $a \in A$, on vérifie que $aX^n = (0, \cdots, 0, a, 0, \cdots)$, où le $n + 1$-ième terme de la suite est a et où tous les autres sont nuls. Avec ces notations, tout polynôme $P = (p_0, p_1, \cdots, p_n, 0, \cdots)$, dont les coefficients sont nuls pour les entiers $> n$, s'écrit alors

$$P = p_0 + p_1 X + \cdots p_n X^n,$$

qui est la notation polynômiale de P et que l'on utilise exclusivement. On note $A[X]$ l'anneau ainsi obtenu. Il s'appelle l'anneau des polynômes en une indéterminée à coefficients dans A. Bien entendu, on peut désigner le polynôme $(0, 1, 0, \cdots)$ par d'autres lettres que X, par exemple Y, Z ou T, à condition que la lettre choisie n'ait pas été utilisée par ailleurs.

5) **Produit direct d'anneaux.** Soient $A_1, \cdots, A_n$ des anneaux. Il existe sur le produit cartésien

$$A = A_1 \times \cdots \times A_n$$

une structure d'anneau, l'addition et la multiplication étant données par les formules

$$(x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n), \tag{1}$$

$$(x_1, \cdots, x_n)(y_1, \cdots, y_n) = (x_1 y_1, \cdots, x_n y_n). \tag{2}$$

Si tous les anneaux A_i sont commutatifs, il en est de même de A. On dit que A est le produit direct des A_i, ou encore l'anneau produit des A_i. Notons que l'élément neutre multiplicatif de A est $(1_{A_1}, \cdots, 1_{A_n})$, où 1_{A_i} l'élément neutre multiplicatif de A_i.

Exercice 1. Soit A un anneau tel que pour tous $a, b \in A$, on ait $(ab)^2 = a^2 b^2$. Montrer que A est commutatif.

2. Sous-anneaux - Idéaux

Soient A un anneau et B une partie de A.

Définition 3.2. *On dit que B est un sous-anneau de A si les conditions suivantes sont vérifiées :*

1) B est un sous-groupe additif de A.

2) Quels que soient x et y dans B, le produit xy est dans B.

3) L'élément neutre multiplicatif 1 *appartient à B.*

On vérifie que si B est un sous-anneau de A, alors B muni des deux lois de composition induites par celles de A est un anneau.

Exemples 3.2.

1) $\mathbb{Z}$ est un sous-anneau de $\mathbb{R}$, lui-même étant un sous-anneau de $\mathbb{C}$.

2) L'ensemble des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ est un sous-anneau de $F(\mathbb{R}, \mathbb{R})$.

3) Soit i une racine carrée de -1 dans $\mathbb{C}$. L'ensemble $\mathbb{Z}[i]$ des éléments de la forme $a + ib$ avec $a, b \in \mathbb{Z}$ est sous-anneau de $\mathbb{C}$. On l'appelle l'anneau des entiers de Gauss.

Définissons maintenant la notion d'idéal de A dans le cas où A est commutatif.

Définition 3.3. *Supposons A commutatif. On dit que B est un idéal de A si les deux conditions suivantes sont vérifiées :*

1) B est un sous-groupe additif de A.

2) Quels que soient $x \in B$ et $y \in A$, le produit xy est dans B.

Exemples 3.3.

1) **Idéaux de $\mathbb{Z}$.** Les idéaux de $\mathbb{Z}$ sont les $n\mathbb{Z}$, où n parcourt $\mathbb{Z}$ (ou $\mathbb{N}$). En effet, ce sont exactement les sous-groupes de $\mathbb{Z}$, et ils vérifient la condition 2 de la définition.

2) Soient X un ensemble et Y une partie de X. Le sous-ensemble de $F(X, \mathbb{R})$ formé des applications qui s'annulent sur Y est un idéal de $F(X, \mathbb{R})$.

3) **Idéaux principaux.** Supposons A commutatif. Soit a un élément de A. L'ensemble des éléments de la forme ax, où x parcourt A, est un idéal de A. On l'appelle l'idéal principal engendré par a. On le note aA ou (a). En particulier, les parties $\{0\}$ et A sont des idéaux de A. Tous les idéaux de $\mathbb{Z}$ sont principaux.

4) L'ensemble $\mathbb{Z}$ n'est pas un idéal de $\mathbb{Q}$ (ni de $\mathbb{R}$, ni de $\mathbb{C}$).

5) **Idéaux d'un anneau produit.** Soient K et L deux anneaux commutatifs.

Lemme 3.2. *Les idéaux de l'anneau produit $K \times L$ (qui est commutatif) sont exactement les $I \times J$, où I est un idéal de K et J un idéal de L.*

Démonstration : Soient I et J deux idéaux de K et L respectivement. Il est immédiat de vérifier que $I \times J$ est un sous-groupe additif de $K \times L$. Par ailleurs, si (a, b) est un élément de $K \times L$, alors pour tout $(i, j) \in I \times J$, on a $(a, b)(i, j) = (ai, bj)$ qui appartient à $I \times J$, donc $I \times J$ est un idéal de $K \times L$. Considérons maintenant un idéal $\mathfrak{I}$ de $K \times L$. Soit I l'image de $\mathfrak{I}$ par la première projection $K \times L \to K$ qui à (u, v) associe u. De même, soit J l'image de $\mathfrak{I}$ par l'application $K \times L \to L$ qui à (u, v) associe v. On vérifie que I (resp. J) est un idéal de K (resp. de L). Montrons que l'on a $\mathfrak{I} = I \times J$. Par définition, $\mathfrak{I}$ est contenu dans $I \times J$. Inversement, soit (x, y) un élément de $I \times J$. Il existe $r \in K$ et $s \in L$ tels que (x, s) et (r, y) soient dans $\mathfrak{I}$. Si 1_K (resp. 1_L) désigne l'élément neutre multiplicatif de K (resp. de L), l'égalité

$$(x, y) = (1_K, 0)(x, s) + (0, 1_L)(r, y)$$

entraîne alors que (x, y) est dans $\mathfrak{I}$, d'où l'assertion[22].

[22] L'assertion analogue obtenue en remplaçant «anneau produit» par «groupe produit» est fausse. Autrement dit, les sous-groupes d'un groupe produit $G_1 \times G_2$ ne sont pas exclusivement les produits $H_1 \times H_2$, où H_1 et H_2 sont des sous-groupes de G_1 et G_2 respectivement. Il y en a d'autres en général. Par exemple, soit G un groupe d'ordre 2. Posons $G = \{e, a\}$, où e est l'élément neutre de G. Alors, $H = \{(e, e), (a, a)\}$ est un sous-groupe de $G \times G$ et n'est pas de la forme $H_1 \times H_2$, où H_1 et H_2 sont des sous-groupes de G, vu que les sous-groupes de G sont $\{e\}$ et G.

Exercice 2. Un idéal $\mathfrak{p}$ d'un anneau commutatif A est dit premier s'il vérifie les deux conditions suivantes :

1) on a $\mathfrak{p} \neq A$.

2) Quels que soient $x, y \in A$, si xy est dans $\mathfrak{p}$, alors x ou bien y est dans $\mathfrak{p}$.

Quels sont les idéaux premiers de $\mathbb{Z}$?

3. Anneau quotient d'un anneau commutatif

Considérons un anneau commutatif A et I un idéal de A. Compte tenu du chapitre II, puisque I est un sous-groupe de A et que $(A, +)$ est un groupe abélien, on peut associer à I la relation d'équivalence $\mathcal{R}$ définie pour tous $x, y \in A$ par la condition

$$x \mathcal{R} y \Longleftrightarrow x - y \in I.$$

L'ensemble quotient A/I, muni de la loi de composition définie pour tous $x, y \in A$ par l'égalité

$$(x + I) + (y + I) = (x + y) + I, \tag{3}$$

est alors un groupe abélien, d'élément neutre I i.e. la classe de 0. On va définir une seconde loi de composition sur A/I, appelée multiplication, de sorte que A/I soit, avec l'addition précedente, muni d'une structure d'anneau commutatif. Soient $x+I$ et $y+I$ deux éléments de A/I. On définit la multiplication par la formule

$$(x + I)(y + I) = xy + I. \tag{4}$$

Pour que cette définition ait sens, il convient de vérifier qu'elle ne dépend pas des représentants x et y de $x + I$ et de $y + I$. Soient x' et y' dans A tels que l'on ait $x + I = x' + I$ et $y + I = y' + I$. Il existe r et t dans I tels que $x = x' + r$ et $y = y' + t$. On a

$$xy = x'y' + (x't + ry' + rt).$$

Puisque r et t sont dans I, il en est de même de $x't+ry'+rt$, par suite, $xy-x'y'$ appartient à I, ce qui établit notre assertion.

Théorème 3.1. *L'ensemble A/I muni de l'addition et la multiplication définies par les formules* (3) *et* (4) *est un anneau commutatif. On l'appelle l'anneau quotient de A par I.*

Démonstration : On sait déjà que $(A/I, +)$ est un groupe abélien. La multiplication dans A étant associative et commutative, il en est de même dans A/I comme on le constate directement. Par ailleurs, $1+I$ est l'élément neutre multiplicatif de A/I (car 1 est l'élément

neutre multiplicatif de A). Il reste à vérifier que la multiplication est distributive par rapport à l'addition. Soient x, y, z des éléments de A. On a les égalités

$$(x+I)\big((y+I)+(z+I)\big) = (x+I)\big((y+z)+I\big) = x(y+z)+I = xy+xz+I = (xy+I)+(xz+I),$$

par suite, on a

$$(x+I)\big((y+I)+(z+I)\big) = (x+I)(y+I)+(x+I)(z+I).$$

La deuxième égalité de définition de la distributivité se vérifie de la même façon. D'où le résultat.

Exemples 3.4.

1) **L'anneau quotient $\mathbb{Z}/n\mathbb{Z}$.** Soit n un entier naturel non nul. On a vu que $n\mathbb{Z}$ est un idéal de $\mathbb{Z}$. D'après le théorème 3.1, l'ensemble $\mathbb{Z}/n\mathbb{Z}$ est donc muni d'une structure d'anneau commutatif, pour laquelle l'addition et la multiplication sont données par les égalités

$$\overline{a}+\overline{b} = \overline{a+b} \quad \text{et} \quad \overline{a}\overline{b} = \overline{ab} \quad \text{quels que soient } a, b \in \mathbb{Z}. \tag{5}$$

Rappelons que l'élément neutre additif est $\overline{0} = n\mathbb{Z}$. L'élément neutre multiplicatif est $\overline{1} = 1 + n\mathbb{Z}$, i.e. l'ensemble des entiers a tels que n divise $a-1$. L'anneau $\mathbb{Z}/n\mathbb{Z}$ s'appelle l'anneau des entiers modulo n.

2) Soit $F \in A[X]$ un polynôme à coefficients dans un anneau commutatif A. On peut considérer l'anneau quotient $A[X]/(F)$, où (F) est l'idéal principal de $A[X]$ engendré par F. Nous étudierons plus loin ces anneaux, par exemple si $A = \mathbb{Z}/p\mathbb{Z}$, où p est premier.

4. Groupe des éléments inversibles - Corps - Anneaux intègres

Soit A un anneau.

Définition 3.4. *Soit x un élément de A. On dit que x est un élément inversible de A s'il possède un inverse pour la multiplication, autrement dit, s'il existe un élément $b \in A$ tel que l'on ait $ab = ba = 1$. On notera A^* l'ensemble des éléments inversibles de A.*

Rappelons que si $a \in A$ est inversible, il existe un unique élément $b \in A$ tel que $ab = ba = 1$ et on le note a^{-1}. Par ailleurs, si x et y sont deux éléments de A^* alors le produit xy est aussi dans A^* et son inverse est $y^{-1}x^{-1}$. En particulier, la multiplication induit sur A^* une loi de composition.

Proposition 3.1. *L'ensemble A^*, muni de la multiplication induite par celle de A, est un groupe. On l'appelle le groupe des éléments inversibles de A, ou le groupe des unités de A.*

Démonstration : C'est une conséquence directe des définitions 2.1 et 3.4, l'élément neutre de A^* étant l'élément neutre multiplicatif 1 de A.

Par exemple, le groupe des élements inversibles de l'anneau $\mathbb{Z}$ est $\{\pm 1\}$. On étudiera en détails dans le chapitre suivant l'anneau $\mathbb{Z}/n\mathbb{Z}$ et l'on décrira en particulier le groupe $(\mathbb{Z}/n\mathbb{Z})^*$ de ses éléments inversibles.

Exercice 3. Démontrer que le groupe des éléments inversibles de l'anneau $\mathbb{Z}[i]$ est $\{-1, 1, -i, i\}$. Montrer que ce groupe est isomorphe à $(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, +)$.

Exercice 4. Soient x et y deux éléments de A tels que $1 - xy$ soit inversible. Montrer que $1 - yx$ est aussi inversible.

Exercice 5. Soient X un ensemble et f un élément de $F(X, A)$. Montrer que f est inversible si et seulement si $f(X)$ est contenu dans A^*.

Le résultat suivant décrit le groupe des éléments inversibles d'un anneau produit, on l'utilisera plus loin.

Lemme 3.3. *Soient A et B deux anneaux. Le groupe des éléments inversibles de l'anneau produit $A \times B$ est $A^* \times B^*$. Autrement dit, on a $(A \times B)^* = A^* \times B^*$. En particulier, si A et B sont finis, on a $|(A \times B)^*| = |A^*||B^*|$.*

Démonstration : Soit (a, b) un élément inversible de $A \times B$. Il existe $(c, d) \in A \times B$ tel que $(a, b)(c, d) = (c, d)(a, b) = (1_A, 1_B)$ où 1_A (resp. 1_B) est l'élément neutre multiplicatif de A (resp. de B). D'après la formule (2), on obtient ainsi les égalités $ac = ca = 1_A$ et $bd = db = 1_B$, ce qui prouve que $a \in A^*$ et que $b \in B^*$. Inversement, si (a, b) est un élément de $A^* \times B^*$, il existe $c \in A$ et $d \in b$ tels que $ac = ca = 1_A$ et $bd = db = 1_B$. Par suite, on a $(a, b)(c, d) = (c, d)(a, b) = (1_A, 1_B)$ donc $(a, b) \in (A \times B)^*$, d'où le résultat.

Lemme 3.4. *Supposons A commutatif. Soit I un idéal de A. Alors, $I = A$ si et seulement si il existe un élément inversible dans I.*

Démonstration : Supposons qu'il existe $x \in I \cap A^*$. Dans ce cas, $xx^{-1} = 1$ est dans I, par suite, pour tout $y \in A$, l'élément $y.1 = y$ est aussi dans I, d'où $I = A$.

Définition 3.5. *On dit que A est un corps si l'on a $1 \neq 0$, et si tout élément non nul de A est inversible i.e. si l'on a $A^* = A - \{0\}$.*

Par définion, un corps possède donc au moins deux éléments, à savoir 0 et 1. Si A est un anneau commutatif et est un corps, on dit que A est un corps commutatif. Les anneaux $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont des corps commutatifs.

Définition 3.6. *Soit K un corps. On appelle sous-corps de K tout sous-anneau L de K qui est un corps. On dit alors que K est un surcorps de L.*

Compte tenu de la définition 3.5, une partie L de K est un sous-corps de K si et seulement si L est un sous-anneau de K dont tous les éléments non nuls sont inversibles.

Exercice 5. Soit A un anneau commutatif. Montrer que A est un corps si et seulement si ses seuls idéaux sont $\{0\}$ et A.

Le produit de deux éléments non nuls dans un corps est non nul. Les corps, tout au moins ceux qui sont commutatifs, font partie de certains anneaux plus généraux, les anneaux intègres :

Définition 3.7. *Un anneau A est dit intègre s'il est commutatif, non réduit à 0 i.e. on a $1 \neq 0$, et si le produit de deux éléments non nuls de A est non nul.*

Par exemple, $\mathbb{Z}$ est un anneau intègre, et plus généralement, tout sous-anneau d'un corps commutatif est un anneau intègre. On utilisera le résultat suivant :

Proposition 3.2. *Soit A un anneau intègre fini. Alors, A est un corps.*

Démonstration : Soit a un élément non nul de A. Il s'agit de montrer que a est inversible. On considère pour cela l'application de A à valeurs dans A qui à x associe ax. Elle est injective, car pour tout $x, y \in A$, si l'on a $ax = ay$, alors, $a(x-y) = 0$ et puisque A est intègre, cela entraîne $x = y$. L'anneau A étant fini, cette application est donc aussi une surjection, en particulier, 1 possède un antécédent, autrement dit, il existe $b \in A$ tel que $ab = 1$ (et $ba = 1$ car A est commutatif), d'où le résultat.

Exemples 3.5. (anneaux non intègres)

1) Soit $n \geq 2$ un entier non premier. Alors, l'anneau $\mathbb{Z}/n\mathbb{Z}$ n'est pas intègre. En effet, on a $n = ab$ avec a et b strictement plus grands que 1, de sorte que l'on a $\overline{a}\overline{b} = \overline{n} = \overline{0}$, bien que $\overline{a}$ et $\overline{b}$ ne soient pas nuls.

2) Si A et B sont deux anneaux intègres, l'anneau produit $A \times B$ n'est jamais intègre, comme le montre l'égalité $(1,0)(0,1) = (0,0)$.

3) L'anneau $F(\mathbb{R}, \mathbb{R})$ des applications de $\mathbb{R}$ dans $\mathbb{R}$ n'est pas intègre. Considérons en effet les deux éléments f et g définis comme suit.

$$f(x) = \begin{cases} x & \text{si } x \geq 0 \\ 0 & \text{si } x \leq 0 \end{cases} \quad \text{et} \quad g(x) = \begin{cases} 0 & \text{si } x \geq 0 \\ x & \text{si } x \leq 0 \end{cases}.$$

On a alors $fg = 0$ i.e. fg est l'application qui en tout $x \in \mathbb{R}$ prend la valeur 0, bien que f et g ne soient pas nuls.

Exercice 6. Soient A un anneau commutatif et I un idéal de A. Montrer que A/I est intègre si et seulement si I est un idéal premier (voir l'exercice 2 pour la définition d'un idéal premier).

Exercice 7. Démontrer qu'un anneau intègre qui ne possède qu'un nombre fini d'idéaux est un corps (étant donné un élément x non nul, afin de montrer qu'il est inversible, on pourra considérer la suite des idéaux principaux (x^n) pour $n \geq 1$).

5. Homomorphisme d'anneaux

Définition 3.8. *Soient A et B deux anneaux. On appelle homomorphisme (d'anneaux), ou morphisme (d'anneaux), de A dans B toute application f de A dans B vérifiant les conditions suivantes :*

1) on a les égalités

$$f(x+y) = f(x) + f(y) \quad et \quad f(xy) = f(x)f(y) \quad quels\ que\ soient\ x, y \in A.$$

2) On a $f(1_A) = 1_B$ (en notant 1_A et 1_B les éléments neutres respectifs de A et B).

Exemples 3.6.

1) Pour tout $n \geq 1$, la surjection canonique $s : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ définie par $s(x) = x + n\mathbb{Z}$ est un homomorphisme.

2) Soient X un ensemble et A un anneau. Pour tout $x \in X$, l'application $F(X, A) \to A$ qui à $f \in F(X, A)$ associe $f(x)$ est un homomorphisme.

Lemme 3.5. *Soient $f : A \to B$ un homomorphisme d'anneaux et A', B' des sous-anneaux de A et B respectivement.*

1) L'image $f(A')$ est un sous-anneau de B.

2) L'image réciproque $f^{-1}(B')$ est un sous-anneau de A.

Démonstration : 1) On sait déjà que $f(A')$ est un sous-groupe additif de B. Par ailleurs, on a $f(1_A) = 1_B$ et $1_A \in A'$ d'où $1_B \in f(A')$. Si x et y sont dans $f(A')$, il existe u et v dans A' tels que $x = f(u)$ et $y = f(v)$, de sorte que $xy = f(u)f(v) = f(uv)$ appartient à $f(A')$.

2) On a vu que $f^{-1}(B')$ est un sous-groupe de A. L'égalité $f(1_A) = 1_B \in B'$, entraîne que $1_A \in f^{-1}(B')$. Si x et y sont dans $f^{-1}(B')$, alors $f(x)$ et $f(y)$ sont dans B', et $f(xy) = f(x)f(y) \in B'$ d'où $xy \in f^{-1}(B')$.

De façon analogue aux homomorphismes de groupes, on démontre que l'application composée de deux homomorphismes d'anneaux est encore un homomorphisme d'anneaux, et que si un homomorphisme d'anneaux est une bijection, son application réciproque est aussi un homomorphisme d'anneaux.

Définition 3.9. *Soient A et B deux anneaux. On appelle isomorphisme de A sur B tout homomorphisme d'anneaux bijectif de A sur B. S'il existe un isomorphisme entre A et B, on dit que A et B sont isomorphes.*

Lemme 3.6. *Soient A et B deux anneaux commutatifs, $f : A \to B$ un homomorphisme et I un idéal de B. Alors, $f^{-1}(I)$ est un idéal de A.*

Démonstration : Considérons deux éléments $x \in A$ et $y \in f^{-1}(I)$. L'élément $f(x)f(y)$ est dans I i.e. $f(xy) \in I$, donc $xy \in f^{-1}(I)$. L'assertion en résulte puisque $f^{-1}(I)$ est un sous-groupe additif de A.

Remarque 3.1. L'image par un homomorphisme d'un idéal n'est pas en général un idéal, comme le montre l'injection $\mathbb{Z} \to \mathbb{Q}$ (car $\mathbb{Z}$ n'est pas un idéal de $\mathbb{Q}$). Cela étant :

Exercice 8. Soient A et B deux anneaux commutatifs, $f : A \to B$ un homomorphisme surjectif de A sur B, et I un idéal de A. Alors, $f(I)$ est un idéal de B.

Définition 3.10. *Soient A et B deux anneaux et $f : A \to B$ un homomorphisme. On appelle noyau de f, et on note* $\mathrm{Ker}(f)$*, l'ensemble des éléments $x \in A$ tels que $f(x) = 0$. Le sous-anneau $f(A)$ de B s'appelle l'image de f.*

On a l'énoncé suivant, analogue au théorème 2.7 :

Théorème 3.2. *Soient A un anneau commutatif, B un anneau et $f : A \to B$ un homomorphisme. Alors,* $\mathrm{Ker}(f)$ *est un idéal de A, et l'anneau quotient $A/\mathrm{Ker}(f)$ est isomorphe à $f(A)$ via l'application qui à $x + \mathrm{Ker}(f)$ associe $f(x)$.*

Démonstration : Le fait que $\mathrm{Ker}(f)$ soit un idéal de A résulte directement des définitions. Notons $h : A/\mathrm{Ker}(f) \to f(A)$ l'application définie par

$$h\big(x + \mathrm{Ker}(f)\big) = f(x).$$

Compte tenu du théorème 2.7, on sait que h est bien définie et que c'est un isomorphisme de groupes. Par ailleurs, si $x + \mathrm{Ker}(f)$ et $y + \mathrm{Ker}(f)$ sont dans $A/\mathrm{Ker}(f)$, on a

$$h\big((x + \mathrm{Ker}(f))(y + \mathrm{Ker}(f))\big) = h\big((xy + \mathrm{Ker}(f))\big) = f(xy) = f(x)f(y),$$

qui n'est autre que $h\big((x + \mathrm{Ker}(f))\big)h\big((y + \mathrm{Ker}(f))\big)$. Puisque l'on a

$$h\big(1_A + \mathrm{Ker}(f)\big) = f(1_A) = 1_B,$$

h est donc un homomorphisme d'anneaux, d'où le résultat.

En illustration de ce qui précède, démontrons l'énoncé suivant qui caractérise, à isomorphisme près, les anneaux quotients de $\mathbb{Z}$.

Proposition 3.3. *Soit A un anneau. Les deux conditions suivantes sont équivalentes :*

1) *l'anneau A ne possède pas de sous-anneaux autres que lui-même.*
2) *Il existe un entier $n \geq 0$ tel que A soit isomorphe à $\mathbb{Z}/n\mathbb{Z}$.*

Démonstration : Pour tout entier $n \geq 0$, l'anneau $\mathbb{Z}/n\mathbb{Z}$ n'a pas de sous-anneaux autres que lui-même. En effet, si B est un sous-anneau de $\mathbb{Z}/n\mathbb{Z}$, alors $\overline{1}$ est dans B,

donc le sous-groupe engendré par $\overline{1}$, i.e. $\mathbb{Z}/n\mathbb{Z}$, est contenu dans B, d'où $B = \mathbb{Z}/n\mathbb{Z}$. En particulier, tout anneau isomorphe à $\mathbb{Z}/n\mathbb{Z}$, pour un certain $n \geq 0$, possède cette propriété. Inversement, supposons la condition 1 réalisée. Considérons l'application $f : \mathbb{Z} \to A$ définie par $f(n) = n1_A$. C'est un homomorphisme d'anneaux. Son image est un sous-anneau de A (lemme 3.5). D'après l'hypothèse faite, on a donc $f(\mathbb{Z}) = A$. Par ailleurs, il existe un entier $n \geq 0$ tel que l'on ait $\mathrm{Ker}(f) = n\mathbb{Z}$. D'après le théorème 3.2, A est donc isomorphe à $\mathbb{Z}/n\mathbb{Z}$.

Exercice 9. Soient K un corps, A un anneau non nul et $f : K \to A$ un homomorphisme. Montrer que f est injectif.

6. La formule du binôme de Newton

On va démontrer ici l'énoncé suivant, connu sous le nom de formule du binôme, qui est très utile dans la théorie des anneaux.

Théorème 3.3. *Soient A un anneau et a, b deux éléments de A tels que $ab = ba$. Alors, pour tout entier $n \geq 0$, on a l'égalité*

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}.$$

Démonstration : Le calcul de $(a+b)^n$ s'obtient en choisissant dans chacun des n facteurs $a+b$, ou bien a ou bien b, en effectuant ensuite le produit des termes ainsi choisis et en additionnant tous les résultats obtenus (il y en a 2^n). Pour tout $k = 0, \cdots, n$, si l'on choisit k fois a et donc $n-k$ fois b, on obtient ainsi, puisque a et b commutent, le terme $a^k b^{n-k}$. Tous les termes du développement de $(a+b)^n$ sont donc de cette forme pour un certain k entre 0 et n. Il reste alors à remarquer que, pour k fixé, le nombre de tels termes est le nombre de façons de choisir k fois a parmi les n possibles, qui n'est autre que le nombre C_n^k de parties à k éléments dans un ensemble à n éléments. D'où le résultat.

Exercice 10. Soit A un anneau commutatif. Un élément $x \in A$ est dit nilpotent s'il existe un entier $n \geq 1$ tel que $x^n = 0$. Montrer que l'ensemble des éléments nilpotents de A, qui est appelé le nilradical de A, est un idéal de A.

Chapitre IV — Arithmétique sur $\mathbb{Z}/n\mathbb{Z}$

Soit n un entier naturel non nul. On a vu dans le chapitre précédent que $\mathbb{Z}/n\mathbb{Z}$ est muni d'une structure d'anneau commutatif, l'addition et la multiplication étant définies pour tous a et b dans $\mathbb{Z}$, par les égalités

$$\overline{a} + \overline{b} = \overline{a+b} \quad \text{et} \quad \overline{a}\overline{b} = \overline{ab}. \tag{1}$$

L'élément neutre additif est $\overline{0} = n\mathbb{Z}$ et l'élément neutre multiplicatif est $\overline{1} = 1 + n\mathbb{Z}$.

1. Le groupe multiplicatif $(\mathbb{Z}/n\mathbb{Z})^*$

Soit n un entier ≥ 1. Rappelons que $\big(\mathbb{Z}/n\mathbb{Z}\big)^*$ désigne le groupe des éléments inversibles pour la multiplication de l'anneau $\mathbb{Z}/n\mathbb{Z}$.

Théorème 4.1. *Soit a un entier. Alors, $\overline{a}$ est inversible si et seulement si a et n sont premiers entre eux. Autrement dit, on a*

$$\big(\mathbb{Z}/n\mathbb{Z}\big)^* = \Big\{\overline{a} \mid 1 \leq a \leq n \text{ et } \operatorname{pgcd}(a,n) = 1\Big\}. \tag{2}$$

Démonstration : Supposons $\overline{a}$ inversible. Il existe alors $b \in \mathbb{Z}$ tel que $\overline{a}\overline{b} = \overline{1}$. Par suite, on a la congruence $ab \equiv 1 \bmod. n$, autrement dit, il existe $c \in \mathbb{Z}$ tel que $ab + nc = 1$, ce qui prouve que a et n sont premiers entre eux. Inversement, d'après le théorème de Bézout, il existe des entiers u et v tels que l'on ait $au + nv = 1$. D'après les égalités (1), on obtient ainsi $\overline{au} = \overline{1}$, ce qui signifie que $\overline{a}$ est inversible. L'égalité (2) en résulte, compte tenu du fait que l'on a $\mathbb{Z}/n\mathbb{Z} = \big\{\overline{1}, \cdots, \overline{n}\big\}$.

En particulier :

Corollaire 4.1. *L'ordre de $\big(\mathbb{Z}/n\mathbb{Z}\big)^*$ est $\varphi(n)$, où $\varphi(n)$ est l'indicateur d'Euler de n.*

Corollaire 4.2. *L'anneau $\mathbb{Z}/n\mathbb{Z}$ est un corps si et seulement si n est un nombre premier.*

Démonstration : Supposons que $\mathbb{Z}/n\mathbb{Z}$ soit un corps. Soit q un diviseur positif de n. On a $n = qk$ où $k \in \mathbb{Z}$, d'où $\overline{n} = \overline{0} = \overline{q}\overline{k}$. Puisque $\mathbb{Z}/n\mathbb{Z}$ est intègre, cela entraîne $\overline{q} = \overline{0}$ ou $\overline{k} = \overline{0}$. Ainsi, n divise q, auquel cas $q = n$, ou bien n divise k, auquel cas $k = n$, puis $q = 1$. Cela prouve que n est premier. Inversement, supposons n premier. D'après le corollaire 4.1, l'ordre de $\big(\mathbb{Z}/n\mathbb{Z}\big)^*$ est alors $n - 1$. Tous les éléments non nuls de $\mathbb{Z}/n\mathbb{Z}$ sont donc inversibles, et $\mathbb{Z}/n\mathbb{Z}$ est un corps.

Remarque 4.1. L'ensemble $\big(\mathbb{Z}/n\mathbb{Z}\big)^*$ est formé des éléments qui sont les générateurs du groupe additif $\big(\mathbb{Z}/n\mathbb{Z}, +\big)$ (cor. 2.4).

Exercice 1. Décrire le groupe $(\mathbb{Z}/20\mathbb{Z})^*$. Déterminer l'inverse de $\overline{17}$.

2. Théorème d'Euler et petit théorème de Fermat

Comme conséquence de ce qui précède, on obtient les résultats suivants :

Théorème 4.2. (Euler) *Soit n un entier ≥ 1. Pour tout entier $a \in \mathbb{Z}$ premier avec n, on a la congruence*

$$a^{\varphi(n)} \equiv 1 \text{ mod. } n. \tag{3}$$

Démonstration : Soit a un entier premier avec n. D'après le théorème 4.1, $\overline{a}$ appartient au groupe $(\mathbb{Z}/n\mathbb{Z})^*$, qui est d'ordre $\varphi(n)$. D'après le théorème 2.3, on a donc dans $\mathbb{Z}/n\mathbb{Z}$ l'égalité $\overline{a}^{\varphi(n)} = \overline{1}$, ce qui signifie que l'on a la congruence (3).

Corollaire 4.3. (Petit théorème de Fermat) *Soit p un nombre premier. Pour tout entier $a \in \mathbb{Z}$, non divisible par p, on a la congruence*

$$a^{p-1} \equiv 1 \text{ mod. } p. \tag{4}$$

En particulier, pour tout $a \in \mathbb{Z}$, on a $a^p \equiv a$ mod. p.

Démonstration : Le théorème 4.2 entraîne l'assertion vu que l'on a $\varphi(p) = p - 1$.

Ces résultats, ainsi que les premières notions de théorie des groupes introduites dans le chapitre II, ont des applications arithmétiques considérables. On présentera au paragraphe 5 une application à la cryptographie, sur l'algorithme RSA. À titre indicatif, illustrons ici ces résultats à travers quelques exemples.

Exemples 4.1.

1) Posons $a = (1035125)^{5642}$. Déterminons le reste de la division euclidienne de a par 17. On a la congruence $1035125 \equiv 12$ mod. 17. D'après le petit théorème de Fermat, on a $12^{16} \equiv 1$ mod. 17. Par ailleurs, on a $5642 \equiv 10$ mod. 16. On en déduit que l'on a $a \equiv 12^{5642} \equiv 12^{10}$ mod. 17. On a $12 \equiv -5$ mod. 17, $12^2 \equiv 8$ mod. 17, $12^4 \equiv -4$ mod. 17, $12^8 \equiv -1$ mod. 17, d'où $a \equiv 9$ mod. 17. Le reste cherché est donc 9.

2) Démontrons que le seul entier impair n qui divise $3^n + 1$ est $n = 1$. Supposons pour cela qu'il existe un entier $n \geq 3$ impair divisant $3^n + 1$. Soit p le plus petit diviseur premier de n. On a $p \geq 5$ (car 3 ne divise pas $3^n + 1$).On a les congruences

$$3^{p-1} \equiv 1 \text{ mod. } p \quad \text{(petit th. de Fermat)} \quad \text{et} \quad 3^{2n} \equiv 1 \text{ mod. } p. \tag{5}$$

Soit δ l'ordre de la classe de 3 dans le groupe multiplicatif $(\mathbb{Z}/p\mathbb{Z})^*$. On déduit de (5) que δ divise $p - 1$ et $2n$. On est alors amené à distinguer deux cas. Supposons δ impair.

D'après le lemme de Gauss, δ divise alors n. L'inégalité $\delta < p$ et le caractère minimal de p entraînent alors $\delta = 1$. Par suite, on a $3^\delta = 3 \equiv 1$ mod. p, d'où $p = 2$ et une contradiction. Supposons δ pair et posons $\delta = 2k$ où $k \in \mathbb{Z}$. L'entier k divise n. Les inégalités $k < \delta < p$ impliquent $k = 1$, d'où $\delta = 2$ puis $3^2 \equiv 1$ mod. p. Cela conduit à $p = 2$ et de nouveau à une contradiction. D'où notre assertion.

Les deux exemples suivants sont des critères permettant parfois de démontrer qu'un entier est premier, si tel est le cas.

3) Soit n un entier naturel ≥ 2. Supposons qu'il existe un entier $a \geq 1$ tel que l'on ait

$$a^{n-1} \equiv 1 \text{ mod. } n \quad \text{et} \quad a^{\frac{n-1}{q}} \not\equiv 1 \text{ mod. } n \quad \text{pour tout diviseur premier } q \text{ de } n-1.$$

Vérifions que n est un nombre premier. La congruence $a^{n-1} \equiv 1$ mod. n entraîne en particulier que a est premier avec n. Il résulte alors de la proposition 2.9 que l'ordre de $a + n\mathbb{Z} \in \left(\mathbb{Z}/n\mathbb{Z}\right)^*$ est $n-1$. Par suite, $n-1$ divise $\varphi(n)$. On a donc les inégalités $n-1 \leq \varphi(n) < n$, d'où $\varphi(n) = n-1$, et le fait que n soit premier (si n n'était pas premier, il posséderait un diviseur autre que 1 et lui-même, et l'on aurait $\varphi(n) < n-1$).

Ce critère, utilisé avec $a = 5$, permet de démontrer que $3 \times 2^{3189} + 1$ est premier. Cet entier possède 961 chiffres dans son écriture décimale (prouver cette assertion).

4) Soit n un entier naturel ≥ 2. Supposons que pour tout diviseur premier q de $n-1$, il existe un entier naturel a (qui dépend de q) tel que l'on ait :

1) $a^{n-1} \equiv 1$ mod. n.

2) $a^{\frac{n-1}{q}} \not\equiv 1$ mod. n.

Démontrons que n est premier. Il suffit de démontrer que l'on a $\varphi(n) = n-1$. Puisque $\varphi(n) \leq n-1$, il suffit donc de prouver que $n-1$ divise $\varphi(n)$. Considérons pour cela un nombre premier q et un entier $r \geq 1$ tels que q^r divise $n-1$. Vérifions que q^r divise $\varphi(n)$. Par hypothèse, il existe un entier a tel que les conditions 1 et 2 soient satisfaites. L'entier a est premier à n. Soit δ l'ordre de la classe de a dans le groupe multiplicatif $(\mathbb{Z}/n\mathbb{Z})^*$. D'après la condition 1, δ divise $n-1$ et d'après la condition 2, δ ne divise pas $(n-1)/q$. Il existe $t \in \mathbb{Z}$ tel que $n-1 = \delta t$. L'entier q ne divise pas t (sinon δ diviserait $(n-1)/q$). Par suite, q^r divise δ. D'après le théorème d'Euler, on a $a^{\varphi(n)} \equiv 1$ mod. n, de sorte que δ divise $\varphi(n)$. Il en résulte que q^r divise $\varphi(n)$, d'où le fait que $n-1$ divise $\varphi(n)$, et le résultat.

On peut démontrer que $3 \times 2^{2816} + 1$ est premier en utilisant ce critère avec les couples $(q, a) = (2, 7)$ et $(3, 2)$, ou bien le critère de l'alinéa 3 avec $a = 13$.

5) Soient p un nombre premier impair et a, b deux entiers non divisibles par p, tels que p divise $a^2 + b^2$. Démontrons que l'on a $p \equiv 1$ mod. 4. Dans le corps $\mathbb{Z}/p\mathbb{Z}$, on a $\bar{a}^2 + \bar{b}^2 = 0$. Puisque p ne divise pas b, on a $\bar{b} \neq 0$. Posons $x = \bar{a}/\bar{b}$. On a l'égalité $x^2 = -1$. Par suite, 4 est l'ordre de x dans le groupe $(\mathbb{Z}/p\mathbb{Z})^*$. D'après le petit théorème de Fermat, on a $x^{p-1} = 1$, donc 4 divise $p-1$.

6) Démontrons qu'il existe une infinité de nombres premiers tels que

$$(i)\ p \equiv 1 \text{ mod. } 4\ ; \quad (ii)\ p \equiv 3 \text{ mod. } 4\ ; \quad (iii)\ p \equiv 5 \text{ mod. } 6.$$

Dans chacun des trois cas envisagés, on suppose que l'ensemble des nombres premiers considérés est fini. On note q le plus grand élément de chacun de ces ensembles.

(i) Posons $N = (q!)^2 + 1$. Soit p un diviseur premier de N. Il ne divise pas $q!$ et en particulier est distinct de 2. D'après l'exemple 5 ci-dessus, on a donc $p \equiv 1$ mod. 4. Par ailleurs, p ne divisant pas $q!$, on a $p > q$, d'où une contradiction et le résultat.

(ii) On pose $N = q! - 1$. Soit p un diviseur premier de N. Comme ci-dessus, p ne divise pas $q!$, donc on a $p > q$. D'après le caractère maximal de q, on a donc $p \equiv 1$ mod. 4. Par suite, tous les diviseurs premiers de N sont congrus à 1 modulo 4, d'où $N \equiv 1$ mod. 4. Cela conduit à une contradiction car $N \equiv -1$ mod. 4.

(iii) L'argument est le même que le précédent. On pose $N = q! - 1$. Pour tout diviseur premier p de N, on a $p > q$, d'où $p \equiv 1$ mod. 6 puis $N \equiv 1$ mod. 6, ce qui n'est pas.

7) Démontrons dans cet exemple qu'il n'existe pas de couples (x, y) dans $\mathbb{Z}^2$ tels que

$$y^2 = x^3 + 7. \tag{6}$$

Supposons qu'il existe $(x, y) \in \mathbb{Z}^2$ vérifiant (6). Supposons de plus x pair. Dans ce cas, on a $y^2 \equiv 7$ mod. 8 et y est impair. On obtient une contradiction car le carré de tout nombre impair est congru à 1 modulo 8 (pour tout $k \in \mathbb{Z}$, on a $(2k+1)^2 = 4k(k+1)+1 \equiv 1$ mod. 8 car $k(k+1)$ est un entier pair). Ainsi, x est impair. On a $y^2 + 1 = (x+2)\big((x-1)^2 + 3\big)$. L'entier $(x-1)^2+3$ est congru à 3 modulo 4. Il possède donc un diviseur premier p congru à 3 modulo 4. Ainsi, p divise $y^2 + 1$, p ne divise pas y et l'assertion de l'exemple 5 conduit de nouveau à une contradiction.

Exercice 2. Démontrer que pour tout entier naturel n impair, on a $2^{n!} \equiv 1$ mod. n (on pourra utiliser le théorème d'Euler et l'exercice 15 du chapitre II).

Exercice 3. Montrer que 13 divise $2^{70} + 3^{70}$.

3. Le théorème chinois

Il s'agit du théorème suivant :

Théorème 4.3. (Théorème chinois) *Soient m et n deux entiers naturels non nuls premiers entre eux. L'application*

$$\psi : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z},$$

définie pour tout $a \in \mathbb{Z}$ par l'égalité

$$\psi(a) = \big(a + m\mathbb{Z}, a + n\mathbb{Z}\big), \tag{7}$$

est un homomorphisme d'anneaux surjectif, de noyau $mn\mathbb{Z}$. En particulier, les anneaux $\mathbb{Z}/mn\mathbb{Z}$ et $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ sont isomorphes via l'application qui à tout élément $a + mn\mathbb{Z}$ de $\mathbb{Z}/mn\mathbb{Z}$ associe le couple $\big(a + m\mathbb{Z}, a + n\mathbb{Z}\big)$ de $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$.

Remarque 4.2. Le contenu essentiel de cet énoncé réside dans le fait que ψ soit une application **surjective** de $\mathbb{Z}$ sur $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. Autrement dit, étant donnés des entiers relatifs a et b, il existe $c \in \mathbb{Z}$ tel que l'on ait

$$c \equiv a \text{ mod. } m \quad \text{et} \quad c \equiv b \text{ mod. } n. \tag{8}$$

Démonstration : Tout d'abord, il résulte des formules (1) et de la définition de la structure d'anneau produit sur $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ que ψ est un homomorphisme d'anneaux. Vérifions que l'on a

$$\mathrm{Ker}(\psi) = mn\mathbb{Z}. \tag{9}$$

Si a est un élément de $\mathrm{Ker}(\psi)$, on a $\big(a + m\mathbb{Z}, a + n\mathbb{Z}\big) = (m\mathbb{Z}, n\mathbb{Z})$, autrement dit, on a $a \equiv 0$ mod. m et $a \equiv 0$ mod. n. Puisque m et n sont premiers entre eux, on en déduit que mn divise a, i.e. que $a \in mn\mathbb{Z}$. Inversement, si a est dans $mn\mathbb{Z}$, alors a est évidemment divisible par m et n, donc a est dans $\mathrm{Ker}(\psi)$, d'où l'égalité (9).

Prouvons que ψ est surjectif. Considérons pour cela un élément $\big(a + m\mathbb{Z}, b + n\mathbb{Z}\big)$ de $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. Puisque m et n sont premiers entre eux, il existe deux entiers u et v tels que l'on ait

$$mu + nv = 1. \tag{10}$$

Posons alors

$$c = b(mu) + a(nv). \tag{11}$$

On vérifie que l'on a les congruences $c \equiv a$ mod. m et $c \equiv b$ mod. n, autrement dit que l'on a $\psi(c) = \big(a + m\mathbb{Z}, b + n\mathbb{Z}\big)$, d'où l'assertion. Compte tenu du théorème 3.2, cela établit le théorème.

Remarque 4.3. La démonstration précédente est effective, au sens où si a et b sont deux entiers relatifs donnés, elle permet d'expliciter un entier c vérifiant les congruences (8). En effet, il suffit pour cela de déterminer deux entiers u et v vérifiant l'égalité (10), ce que l'on peut faire en utilisant par exemple l'algorithme d'Euclide. On peut alors prendre comme entier c celui défini par l'égalité (11). Il est par ailleurs unique modulo $mn\mathbb{Z}$.

Exemple 4.2. Déterminons l'ensemble des entiers $k \in \mathbb{Z}$ tels que

$$k \equiv 2 \text{ mod. } 37 \quad \text{et} \quad k \equiv 9 \text{ mod. } 19. \tag{12}$$

En utilisant l'égalité, $2 \times 19 - 37 = 1$, on constate que l'entier

$$k = 2 \times (19 \times 2) + 9 \times (-37) = -257$$

vérifie les congruences (12). L'ensemble cherché est donc formé des entiers congrus à -257 modulo 703. Le plus petit entier naturel satisfaisant aux congruences (12) est donc 446.

Remarque 4.4. Le théorème chinois peut se reformuler de façon plus générale sans supposer au départ les entiers m et n premiers entre eux. Plus précisément :

Théorème 4.4. *Soient m et n des entiers naturels non nul et $\psi : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ l'application définie pour tout $a \in \mathbb{Z}$ par l'égalité*

$$\psi(a) = (a + m\mathbb{Z}, a + n\mathbb{Z}).$$

Alors, ψ est un homomorphisme d'anneaux de noyau l'idéal $\mathrm{ppcm}(m,n)\mathbb{Z}$*, et son image est formée des couples $\big(a + m\mathbb{Z}, b + n\mathbb{Z}\big)$ tels que le pgcd de m et n divise $b - a$* [23].

On en déduit par exemple qu'il n'existe pas d'entiers relatifs qui sont congrus à la fois à 3 modulo 10 et à 2 modulo 6.

Exercice 4. Déterminer le plus petit entier naturel multiple de 7 et congru à 1 modulo 2, 3, 4, 5 et 6.

---

[23] Indiquons une démonstration de ce résultat. Soit a un élément du noyau de ψ. Il est divisible par m et n donc aussi par le ppcm de m et n, autrement dit, a appartient à l'idéal $\mathrm{ppcm}(m,n)\mathbb{Z}$. Inversement, si a est multiple du ppcm de m et n, il est multiple de m et n d'où $\psi(a) = 0$.
En ce qui concerne l'image de ψ, on remarque d'abord que pour tous a et b dans $\mathbb{Z}$, le fait que le pgcd de m et n divise ou non $b - a$ ne dépend que de la classe de a modulo m et de celle de b modulo n. Soit $\big(a + m\mathbb{Z}, b + n\mathbb{Z}\big)$ un élément de l'image de ψ. Il existe $u \in \mathbb{Z}$ tel que $u \equiv a$ mod. m et $u \equiv b$ mod. n. Il existe donc des entiers λ et μ tels que $u = a + \lambda m$ et $u = b + \mu n$, d'où $b - a = \lambda m - \mu n$ ce qui entraîne que le pgcd de m et n divise $b - a$. Inversement, soit $\big(a + m\mathbb{Z}, b + n\mathbb{Z}\big)$ un élément de $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ vérifiant cette condition. Posons $d = \mathrm{pgcd}(m,n)$. Il existe des entiers u et v tels que $d = mu + nv$ (théorème de Bézout). Considérons l'entier

$$c = u\left(\frac{m}{d}\right)b + v\left(\frac{n}{d}\right)a.$$

En écrivant que l'on a $1 = (m/d)u + (n/d)v$, compte tenu du fait que d divise $b - a$, on obtient

$$c \equiv \left(\frac{b-a}{d}\right)mu + a \equiv a \text{ mod. } m \quad \text{et} \quad c \equiv \left(\frac{a-b}{d}\right)nv + b \equiv b \text{ mod. } n.$$

On a ainsi $\psi(c) = \big(a + m\mathbb{Z}, b + n\mathbb{Z}\big)$ qui est donc dans l'image de ψ. D'où le théorème.

Voici une illustration du théorème chinois.

Lemme 4.1. *Soient m et n deux entiers naturels non nuls tels que m divise n. L'application $f : (\mathbb{Z}/n\mathbb{Z})^* \to (\mathbb{Z}/m\mathbb{Z})^*$ définie par $f(a+n\mathbb{Z}) = a+m\mathbb{Z}$, est une surjection de $(\mathbb{Z}/n\mathbb{Z})^*$ sur $(\mathbb{Z}/m\mathbb{Z})^*$.*

Démonstration : On remarque d'abord que f est bien définie. On écrit ensuite n sous la forme $n = m'r$, où m et m' ont les mêmes facteurs premiers et où r est premier à m'. L'entier m divise m' et r est premier à m. Soit $d + m\mathbb{Z}$ un élément de $(\mathbb{Z}/m\mathbb{Z})^*$. D'après le théorème chinois, il existe un entier a tel que

$$a \equiv d \text{ mod. } m \quad \text{et} \quad a \equiv 1 \text{ mod. } r.$$

Vérifions que a est premier à n. Supposons qu'il existe un nombre premier p qui divise a et n. Alors, p ne divise pas r, donc p divise m'. Par suite, p divise m et d, ce qui contredit le fait que d et m soient premiers entre eux. On a ainsi $f(a+n\mathbb{Z}) = d+m\mathbb{Z}$, d'où l'assertion.

Exercice 5. Avec les notations du lemme 4.1, supposons $m = 10$, $n = 60$. Trouver un antécédent par f de $9 + 10\mathbb{Z}$.

4. Détermination de la fonction indicatrice d'Euler

Pour tout entier $n \geq 2$, on déduit de ce qui précède la valeur de $\varphi(n)$ en termes des diviseurs premiers de n.

Théorème 4.5. *Soit n un entier ≥ 2. On a l'égalité*

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right), \tag{13}$$

où p parcourt l'ensemble des diviseurs premiers de n.

On utilise le fait que la fonction φ est multiplicative, autrement dit :

Proposition 4.1. *Soient m et n deux entiers naturels non nuls premiers entre eux. On a*

$$\varphi(mn) = \varphi(m)\varphi(n).$$

Démonstration : Les entiers m et n étant premiers entre eux, les anneaux $\mathbb{Z}/mn\mathbb{Z}$ et $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ sont isomorphes. Les groupes des éléments inversibles de ces anneaux sont donc aussi isomorphes[24], et en particulier ils ont le même ordre. Le lemme 3.3 et le corollaire 4.1 entraînent alors le résultat.

[24] Soient A et B deux anneaux et $f : A \to B$ un isomorphisme de A sur B. Soit A^* (resp. B^*) le groupe des éléments inversibles de A (resp. de B). Alors, f induit un isomorphisme de groupes de A^* sur B^*. En effet, pour tout $x \in A^*$, on a les égalités $f(x)f(x^{-1}) = f(x^{-1})f(x) = f(1_A) = 1_B$, de sorte que $f(x)$ appartient à B^* (et son inverse est $f(x^{-1})$). Ainsi, f induit un homomorphisme de groupes de A^* dans B^*. De même, l'application réciproque g de f induit un homomorphisme de groupes de B^* dans A^*. L'application $g \circ f$ est l'identité de A^* et $f \circ g$ est l'identité de B^*, d'où notre assertion.

Démonstration du théorème 4.5 : Notons $p_1, \cdots, p_r$ les diviseurs premiers de n. Soit

$$n = \prod_{i=1}^{r} p_i^{n_i},$$

la décomposition en facteurs premiers de n. On déduit de la proposition 4.1 que l'on a

$$\varphi(n) = \prod_{i=1}^{r} \varphi\big(p_i^{n_i}\big).$$

Par ailleurs, d'après le lemme 2.4, on a

$$\varphi\big(p_i^{n_i}\big) = p_i^{n_i}\left(1 - \frac{1}{p_i}\right),$$

d'où l'égalité (13).

Indiquons en corollaires deux propriétés de la fonction φ :

Corollaire 4.4. *Pour tout $n \geq 3$, l'entier $\varphi(n)$ est pair.*

Démonstration : D'après l'égalité (13), on a

$$\varphi(n) = \prod_{p|n} p^{v_p(n)-1}(p-1),$$

où p parcourt l'ensemble des diviseurs premiers de n. Si n possède un diviseur premier impair p, alors $p-1$ est pair, et il en est donc de même de $\varphi(n)$. Si n est une puissance de 2, disons $n = 2^r$ avec $r \geq 2$, alors $\varphi(n) = 2^{r-1}$, d'où l'assertion.

Corollaire 4.5. *Soient m et n deux entiers naturels non nuls tels que m divise n. Alors, $\varphi(m)$ divise $\varphi(n)$.*

Démonstration : Soient P_m (resp. P_n) l'ensemble des diviseurs premiers de m (resp. de n). D'après le théorème 4.5, on a les égalités

$$\frac{\varphi(n)}{\varphi(m)} = \frac{n}{m} \prod_{p \in P_n - P_m} \left(1 - \frac{1}{p}\right). \tag{14}$$

Par ailleurs, pour tout nombre premier $p \in P_n - P_m$, p divise n sans diviser m, donc p divise n/m. Il en résulte que le deuxième membre de l'égalité (14) est un entier, d'où le résultat.

Remarquons que l'implication réciproque du corollaire 4.5 est fausse, comme le montre les égalités $\varphi(3) = \varphi(4) = 2$. Par ailleurs, l'énoncé précédent peut aussi se déduire théorème 2.7 et du lemme 4.1.

5. Application à la cryptographie - Algorithme RSA

La cryptographie est l'étude de la science des communications par des messages codés qui ne pourront être lus que par leur destinataire. Un cryptosystème est un tel mode de communication. La cryptographie suscite en fait de l'intérêt depuis l'antiquité, compte tenu de la nécessité de pouvoir faire parvenir des messages qui ne puissent pas être déchiffrés par un «intrus». L'algorithme RSA, qui a été découvert par Rivest, Shamir et Adleman en 1977, a constitué de ce point de vue une très importante avancée. C'est un algorithme à clé publique. Son efficacité repose sur le fait qu'il est difficile de factoriser un entier ayant disons plus de deux cents chiffres dans son écriture décimale i.e. dans son écriture en base 10. Cet algorithme utilise le résultat suivant, qui est une conséquence du petit théorème de Fermat :

Proposition 4.2. *Soient p et q deux nombres premiers distincts. Posons $n = pq$. Soit t un entier naturel congru à 1 modulo $\varphi(n)$. Alors, on a*

$$a^t \equiv a \text{ mod. } n \quad \textit{quel que soit } a \in \mathbb{Z}.$$

Démonstration : Il existe un entier k tel que l'on ait $t = 1 + k\varphi(n)$. Soit a un entier relatif. Compte tenu de l'égalité $\varphi(n) = (p-1)(q-1)$, on obtient

$$a^t = a\Big(a^{(p-1)(q-1)}\Big)^k = a\Big(a^{p-1}\Big)^{(q-1)k}.$$

Si p ne divise pas a, on a $a^{p-1} \equiv 1$ mod. p, d'où l'on déduit que $a^t \equiv a$ mod. p. Si p divise a, cette congruence est aussi vérifiée. De même, on a $a^t \equiv a$ mod. q. Puisque p et q sont distincts, il en résulte que pq divise $a^t - a$ i.e. que l'on a $a^t \equiv a$ mod. n.

Principe du cryptosystème RSA. Chaque utilisateur de ce cryptosystème procède de la façon suivante :

1) il choisit deux grands nombres premiers p et q et calcule $n = pq$.
2) Il choisit un entier e premier avec $\varphi(n)$ tel que $1 < e < \varphi(n)$. La classe de e modulo $\varphi(n)$ est donc inversible dans $\mathbb{Z}/\varphi(n)\mathbb{Z}$.
3) Il détermine l'entier d tel que $1 < d < \varphi(n)$ et $ed \equiv 1$ mod. $\varphi(n)$. La classe de d modulo $\varphi(n)$ est donc l'inverse de la classe de e dans $\big(\mathbb{Z}/\varphi(n)\mathbb{Z}\big)^*$. Ce calcul peut être effectué en utilisant l'algorithme d'Euclide.
4) Il publie ensuite le couple (e, n), qui est sa clé publique, et il conserve secret le couple $\big(d, \varphi(n)\big)$, qui est sa clé secrète.

Soit A un utilisateur dont la clé publique est (e, n) et la clé secrète est $\big(d, \varphi(n)\big)$. On dit que l'algorithme de chiffrement de A est l'application $E : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ définie pour tout $x \in \mathbb{Z}/n\mathbb{Z}$ par

$$E(x) = x^e \text{ mod. } n,$$

et que son algorithme de déchiffrement est l'application $D : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ définie par

$$D(x) = x^d \text{ mod. } n.$$

Compte tenu de la proposition 4.2. on a

$$D \circ E = E \circ D = 1_{\mathbb{Z}/n\mathbb{Z}} \quad \text{(l'identité de } \mathbb{Z}/n\mathbb{Z}\text{)}.$$

Supposons qu'un utilisateur B souhaite envoyer un message secret à A sous la forme d'un élément $x_0 \in \mathbb{Z}/n\mathbb{Z}$. Il utilise pour cela l'algorithme de chiffrement de A en lui envoyant l'élément $E(x_0)$. Afin de déchiffrer ce message, A détermine l'image $D\big(E(x_0)\big)$ de $E(x_0)$ par D, qui n'est autre que x_0.

Remarque 4.5. Un intrus C peut casser ce cryptosystème i.e. trouver la factorisation de n (et donc la clé secrète de A) si par exemple le message $x_0 = \widetilde{x_0} + n\mathbb{Z}$ envoyé par B est tel que $\widetilde{x_0}$ ne soit pas premier à n. En effet, C connaît n et x_0^e. Il peut donc déterminer le pgcd de $\widetilde{x_0}$ et de n. Si $\widetilde{x_0}$ et n ne sont pas premiers entre eux, C connaît alors au moins l'un des diviseurs premiers de n, et donc la factorisation de n. Ceci constitue une contrainte sur les messages à envoyer. Cela étant, la «probabilité» pour qu'un élément $a + n\mathbb{Z}$ non nul de $\mathbb{Z}/n\mathbb{Z}$ tiré au hasard soit tel que $\text{pgcd}(a, n) \neq 1$ est

$$1 - \frac{\varphi(n)}{n} = \frac{1}{p} + \frac{1}{q} - \frac{1}{pq},$$

qui est donc très petite si p et q sont grands. D'un point de vue heuristique, il y a donc peu de chance de se trouver dans la situation précédente.

Exercice 6. Soit n un entier naturel produit de deux nombres premiers. Démontrer que la connaissance de n et $\varphi(n)$ est équivalente à celle de la factorisation de n.

Exemple 4.3. Appelons comme il est d'usage Alice l'utilisateur A et Bob l'utilisateur B. Supposons qu'Alice choisisse comme nombres premiers

$$p = 263 \quad \text{et} \quad q = 349.$$

On a $n = 91787$ et $\varphi(n) = 91176$. Elle choisit par ailleurs,

$$e = 641.$$

On vérifie que l'on a

$$d = 21905.$$

La clé publique d'Alice est donc $(641, 91787)$ et sa clé secrète est $(21905, 91176)$.

Supposons que Bob souhaite envoyer à Alice un mot secret de trois lettres, notons-le $A_1A_2A_3$. Pour cela, Alice et Bob numérotent implicitement les lettres de l'alphabet de 0 à 25 (A est numéroté 0, $\cdots$, Z est numéroté 25). Chaque lettre A_i correspond à un unique entier a_i tel que $0 \leq a_i \leq 25$. Bob considère alors l'élément $26^2a_1 + 26a_2 + a_3 + n\mathbb{Z} \in \mathbb{Z}/n\mathbb{Z}$, et suivant l'algorithme de déchiffrement d'Alice, il calcule

$$\alpha = \big(26^2a_1 + 26a_2 + a_3 + n\mathbb{Z}\big)^e.$$

Puisque l'on a $n < 26^4 = 456976$, il existe quatre entiers $b_1, \cdots, b_4$ tels que l'on ait

$$\alpha = 26^3b_1 + 26^2b_2 + 26b_3 + b_4 + n\mathbb{Z} \quad \text{avec} \quad 0 \leq b_1, \cdots, b_4 \leq 25.$$

Chaque b_i correspond à une lettre B_i de l'alphabet. Bob envoie alors à Alice le mot $B_1B_2B_3B_4$. Alice commence par identifier α, puis calcule α^d. Compte tenu du fait que l'on a $n > 26^3 = 17576$, Alice peut ainsi retrouver le triplet (a_1, a_2, a_3), puis le mot $A_1A_2A_3$. Il convient de noter ici que l'inégalité $n > 26^3$ doit être impérativement réalisée pour que ce procédé fonctionne. En effet, si n est plus petit que 26^3, il peut exister deux triplets d'entiers distincts (a_1, a_2, a_3) et (a'_1, a'_2, a'_3) tels que les a_i et a'_i soient compris entre 0 et 25 et que l'on ait $26^2a_1 + 26a_2 + a_3 \equiv 26^2a'_1 + 26a'_2 + a'_3$ mod. n, auquel cas Alice ne peut pas retrouver le message envoyé par Bob.

À titre indicatif, supposons que Bob souhaite envoyer le mot OUI à Alice. Le triplet d'entiers qui correspond à ce mot est $(14, 20, 8)$, l'élément de $\mathbb{Z}/n\mathbb{Z}$ correspondant étant $9992 + n\mathbb{Z}$. Il calcule ensuite

$$E\big(9992 + n\mathbb{Z}\big) = \big(9992 + n\mathbb{Z}\big)^e = 13794 + n\mathbb{Z}.$$

Notons que 13794 étant premier à n, la contrainte des messages envoyés est bien respectée. On a $13794 = 26^2.20 + 26.10 + 14$. Avec les notations précédentes, on a donc $b_1 = 0$, $b_2 = 20$, $b_3 = 10$ et $b_4 = 14$. Bob envoie ainsi à Alice le mot AUKO. Pour déchiffrer ce message, Alice le décode d'abord en l'entier 13794, puis en utilisant sa clé secrète, effectue le calcul

$$D\big(13794 + n\mathbb{Z}\big) = \big(13794 + n\mathbb{Z}\big)^d = 9992 + n\mathbb{Z}.$$

Elle constate que $9992 = 26^2.14 + 26.20 + 8$, ce qui lui permet de retrouver le mot OUI.

Signature. Signalons que l'algorithme RSA fournit un procédé de signature, ou d'authentification, à tout utilisateur qui envoie des messages. Considérons en effet un utilisateur A ayant pour clé publique (e, n) et pour clé secrète $(d, \varphi(n))$. Supposons que A souhaite envoyer un message $x \in \mathbb{Z}/n\mathbb{Z}$, sans confidentialité, à un utilisateur B, de sorte que B soit certain que l'expéditeur du message soit bien A. Pour cela, A envoie à B le couple

$$\big(x, x^d\big) \in \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$

C'est le message signé. L'utilisateur B peut alors authentifié A en calculant $(x^d)^e$. S'il trouve x, il est alors « sûr » que c'est bien A l'expéditeur du message x, vu que A est le seul à connaître d.

Chapitre V — Arithmétique sur $K[X]$ et ses quotients

Dans tout ce chapitre la lettre K désigne un corps commutatif. Rappelons que l'on a défini dans le chapitre III l'anneau $K[X]$ des polynômes à une indéterminée à coefficients dans K. Les propriétés arithmétiques de l'anneau $K[X]$ sont essentiellement les mêmes que celles de l'anneau $\mathbb{Z}$, ce qui s'explique par le fait que ce sont des anneaux intègres, dont tous les idéaux sont principaux. On dit que de tels anneaux sont principaux. On démontrera que $K[X]$ est un anneau principal. Nous n'étudierons pas dans ce cours la théorie générale des anneaux principaux.

1. Degré - Division euclidienne

Définissons ce que l'on appelle le degré d'un élément de $K[X]$. On considère pour cela l'ensemble $\mathbb{N} \cup \{-\infty\}$ obtenu en adjoignant à $\mathbb{N}$ un élément noté $-\infty$, que l'on munit de la structure d'ensemble ordonné qui induit l'ordre usuel sur $\mathbb{N}$ et telle que $-\infty \leq n$ pour tout entier naturel n. Tout ensemble non vide de $\mathbb{N} \cup \{-\infty\}$ possède ainsi un plus petit élément. On prolonge par ailleurs la loi additive de $\mathbb{N}$ à cet ensemble en posant $(-\infty)+n = n+(-\infty) = -\infty$ et $(-\infty)+(-\infty) = -\infty$. On définit alors l'application degré

$$\deg : K[X] \to \mathbb{N} \cup \{-\infty\},$$

de la façon suivante :

Définition 5.1. *Soit $F = (a_i)_{i\geq 0}$ un polynôme à coefficients dans K.*

1) Si $F = 0$ i.e. si tous les a_i sont nuls, on pose $\deg(F) = -\infty$.

2) Si F n'est pas nul, $\deg(F)$ est le plus grand entier $n \geq 0$ tel que $a_n \neq 0$.

On dit que $\deg(F)$ est le degré de F.

Si $F \in K[X]$ est non nul, le coefficient de $X^{\deg(F)}$ est appelé le coefficient dominant de F. C'est le coefficient du terme de plus haut degré de F. S'il vaut 1, le polynôme F est dit unitaire.

Lemme 5.1. *Soient P et Q deux éléments de $K[X]$.*

1) On a $\deg(P+Q) \leq \mathrm{Max}\Big(\deg(P), \deg(Q)\Big)$ avec égalité si $\deg(P) \neq \deg(Q)$.

2) On a $\deg(PQ) = \deg(P) + \deg(Q)$.

Démonstration : On vérifie que ces assertions sont vraies si l'un des polynômes P et Q est nul. Supposons P et Q non nuls. Posons

$$P = a_0 + \cdots + a_r X^r \quad \text{et} \quad Q = b_0 + \cdots + b_s X^s \quad \text{avec } a_r b_s \neq 0.$$

On a évidemment $\deg(P+Q) \leq \mathrm{Max}(r,s)$ avec égalité si $r \neq s$. Par ailleurs, on a une égalité de la forme $PQ = a_r b_s X^{r+s} + R$ où $R \in K[X]$ est de degré $< r+s$. Puisque K est un anneau intègre (c'est un corps), on a $a_r b_s \neq 0$ de sorte que le degré de PQ est $r+s$.

Corollaire 5.1. *L'anneau $K[X]$ est intègre et le groupe de ses éléments inversibles est K^* i.e. est l'ensemble des éléments non nuls de K*[25].

Démonstration : L'anneau $K[X]$ n'est pas nul et est commutatif. D'après l'assertion 2 du lemme 5.1, quels que soient P et Q dans $K[X]$, si $PQ = 0$ alors $\deg(P)$ ou bien $\deg(Q)$ vaut $-\infty$, autrement dit, on a $P = 0$ ou bien $Q = 0$, donc $K[X]$ est intègre. Par ailleurs, si $PQ = 1$, on a $\deg(P) + \deg(Q) = 0$, ce qui entraîne $\deg(P) = \deg(Q) = 0$, d'où le résultat.

Le théorème de division euclidienne est le suivant :

Theorème 5.1. (Division euclidienne) *Soient A et B deux polynômes de $K[X]$ tels que $B \neq 0$. Alors, il existe un unique couple (Q,R) de polynômes de $K[X]$ tel que*

$$A = BQ + R \quad \textit{avec} \quad \deg R < \deg B.$$

On dit que Q est le quotient et que R est le reste de la division euclidienne de A par B.

Démonstration : On prouve le lemme suivant :

Lemme 5.2. *Soient U et V des polynômes de $K[X]$ tels que $V \neq 0$ et que l'on ait $\deg(U) \geq \deg(V)$. Alors, il existe $Q \in K[X]$ tel que l'on ait*

$$\deg(U - VQ) < \deg(U).$$

Démonstration : Puisque V est non nul, il en est de même de U. Soit a_k le coefficient dominant de U et b_q celui de V (de sorte que $\deg(U) = k$ et $\deg(V) = q$). On a par hypothèse $k \geq q$. Posons

$$Q = \frac{a_k}{b_q} X^{k-q}.$$

On constate que l'on a $\deg(U - VQ) < \deg(U)$, d'où le lemme.

Le théorème 5.1 se déduit comme suit. Démontrons l'assertion d'existence. On considère l'ensemble

$$S = \Big\{ A - BQ \mid Q \in K[X] \Big\}.$$

Le sous-ensemble de $\mathbb{N} \cup \big\{ -\infty \big\}$ formé des degrés des éléments de S possède un plus petit élément r. Considérons un polynôme $Q \in K[X]$ tel que

$$\deg(A - BQ) = r.$$

[25] Ce résultat est faux si K est remplacé par un anneau non intègre. Par exemple, le polynôme $F = 2X + 1 \in \big(\mathbb{Z}/4\mathbb{Z}\big)[X]$ vérifie l'égalité $F^2 = 1$ (on note ici 2 la classe de 2 et 1 la classe de 1 modulo $4\mathbb{Z}$). On a aussi dans cet anneau $(2X)^2 = 0$.

Il s'agit de montrer que l'on a $r < \deg(B)$. Supposons le contraire i.e. que l'on a $r \geq \deg(B)$. On a $B \neq 0$. Compte tenu du lemme 5.2, il existe $Q' \in K[X]$ tel que le polynôme

$$A - BQ - Q'B = A - B(Q + Q')$$

soit de degré strictement plus petit que r. Puisque ce polynôme est dans S, le caractère minimal de r conduit alors à une contradiction.

Prouvons l'assertion d'unicité. Soient Q et Q_1 éléments de $K[X]$ tels que l'on ait

$$\deg(A - BQ) < \deg(B) \quad \text{et} \quad \deg(A - BQ_1) < \deg(B).$$

On déduit de l'assertion 1 du lemme 5.1 que l'on a

$$\deg\big((A - BQ) - (A - BQ_1)\big) = \deg\big(B(Q_1 - Q)\big) < \deg(B).$$

D'après l'assertion 2 de ce lemme, on a ainsi $\deg(Q_1 - Q) < 0$, ce qui entraîne $Q_1 = Q$ et le résultat[26].

Exercice 1. Déterminer le quotient et le reste de la division euclidienne de X^3+X^2+1 par $X^2 + X + 1$ dans $\big(\mathbb{Z}/2\mathbb{Z}\big)[X]$.

Définition 5.2. *Soient A et B deux polynômes de $K[X]$. On dit que B divise A, ou que A est multiple de B, s'il existe $Q \in K[X]$ tel que $A = BQ$. Si $B \neq 0$ cela signifie que le reste de la division euclidienne de A par B est nul.*

Lemme 5.3. *Soient A et B deux polynômes non nuls de $K[X]$ tels que A divise B et que B divise A. Il existe $\lambda \in K$ non nul tel que $A = \lambda B$. On dit alors que A et B sont associés.*

Démonstration : Il existe Q et Q_1 dans $K[X]$ tels que $A = BQ$ et $B = AQ_1$, d'où $A(1 - QQ_1) = 0$. Par suite, on a $QQ_1 = 1$, autrement dit Q est inversible, d'où l'assertion (cor. 5.1).

2. Idéaux de $K[X]$ - pgcd - ppcm

Commençons par décrire tous les idéaux de $K[X]$. Rappelons que si P est un polynôme de $K[X]$, l'ensemble $(P) = \Big\{PR \mid R \in K[X]\Big\}$ des multiples de P est un idéal de $K[X]$. C'est l'idéal de $K[X]$ engendré par P.

[26] Cette démonstration montre que le théorème de division euclidienne est aussi valable si l'on remplace K par un anneau commutatif quelconque à condition de supposer que le coefficient dominant de B soit un élément inversible.

Théorème 5.2. *Soit I un idéal non nul de $K[X]$. Il existe un unique polynôme unitaire $P \in K[X]$ tel que l'on ait $I = (P)$.*

Démonstration : Il existe P non nul dans I de degré minimum (parmi les éléments non nuls de I). Quitte à multiplier P par un élément convenable de K, on peut supposer que P est unitaire. Vérifions que l'on a $I = (P)$. D'abord tout multiple de P appartient à I. Inversement, soit A un élément de I. D'après le théorème de division euclidienne, il existe Q et R dans $K[X]$ tels que l'on ait $A = PQ + R$ avec $\deg(R) < \deg(P)$. Puisque A est dans I, le polynôme $R = A - PQ$ est aussi dans I. Le caractère minimal de P entraîne alors $R = 0$, d'où $A = PQ \in (P)$. Cela établit l'assertion d'existence. Considérons alors des polynômes unitaires F et G de $K[X]$ tes que $I = (F) = (G)$. En exprimant le fait que F est dans (G) et que G est dans (F), on constate que F et G sont associés (lemme 5.3). Puisqu'ils sont unitaires, on a donc $F = G$.

pgcd de deux polynômes

Considérons deux polynômes A et B de $K[X]$ non tous les deux nuls, ainsi que l'ensemble

$$I = \Big\{ AU + BV \mid U, V \in K[X] \Big\}.$$

On vérifie que I est un idéal non nul de $K[X]$. D'après le théorème 5.2, il existe donc un unique polynôme unitaire $D \in K[X]$ tel que l'on ait

$$I = (D). \tag{1}$$

Définition 5.3. *On dit que D est le plus grand commun diviseur de A et B, ou en abrégé, le pgcd de A et B.*

Avec cette définition, on a de fait la propriété de Bézout dans $K[X]$, à savoir qu'il existe U et V dans $K[X]$ tels que l'on ait

$$D = AU + BV. \tag{2}$$

Cela étant, il convient de vérifier la propriété attendue du pgcd de deux polynômes :

Théorème 5.3. *Soit F un polynôme unitaire de $K[X]$. Alors, F est le pgcd de A et B si et seulement si les deux conditions suivantes sont vérifiées :*

1) le polynôme F divise A et B.

2) Tout diviseur de A et B dans $K[X]$ divise F.

Démonstration : En exprimant le fait que A et B appartiennent à I, on constate que D divise A et B. Par ailleurs, d'après (2), si un polynôme de $K[X]$ divise A et B, alors il divise D. Par suite, D vérifie les conditions 1 et 2. Inversement, soit $F \in K[X]$ réalisant ces

conditions. L'égalité (2) et la condition 1 entraînent que F divise D. D'après la condition 2, D divise F. Puisque D et F sont unitaires, on a donc $F = D$.

Définition 5.4. *On dit que A et B sont premiers entre eux, ou que A est premier avec B, si l'on a $D = 1$.*

On a ainsi l'énoncé suivant (égalité (2) et th. 5.5) :

Corollaire 5.2. *Les polynômes A et B sont premiers entre eux si et seulement si il existe U et V dans $K[X]$ tels que $AU + BV = 1$.*

Théorème 5.4. (Gauss) *Soient F, G et H des polynômes de $K[X]$ tels que F divise GH et que F soit premier avec G. Alors, F divise H.*

Démonstration : Il existe $U, V \in K[X]$ tels que $UF + VG = 1$ (cor. 5.2), d'où l'égalité $(UH)F + V(GH) = H$, donc F divise H.

Remarque 5.1. Compte tenu du théorème de division euclidienne, afin de déterminer le pgcd de deux polynômes, et d'obtenir explicitement une relation de Bézout, on peut utiliser, comme dans le cas de l'anneau $\mathbb{Z}$, l'algorithme d'Euclide.

Exercice 2. Déterminer une relation de Bézout entre les polynômes $(X-1)^3$ et $(X+1)^3$ dans $\mathbb{Q}[X]$.

Exercice 3. Soient m et n deux entiers naturels non nuls. Montrer que le pgcd des polynômes $X^m - 1$ et $X^n - 1$ est $X^d - 1$ où $d = \text{pgcd}(m, n)$.

ppcm de deux polynômes

Considérons deux polynômes non nuls A et B de $K[X]$. L'ensemble $(A) \cap (B)$ est un idéal de $K[X]$. Il existe donc un unique polynôme unitaire $M \in K[X]$ tel que l'on ait

$$(A) \cap (B) = (M). \tag{3}$$

Définition 5.5. *On dit que M est le plus petit commun multiple de A et B, ou en abrégé, le ppcm de A et B.*

Théorème 5.5. *Soit F un polynôme unitaire de $K[X]$. Alors, F est le ppcm de A et B si et seulement si les deux conditions suivantes sont vérifiées :*

1) *le polynôme F est un multiple de A et B.*
2) *Tout multiple de A et B dans $K[X]$ est un multiple de F.*

Démonstration : Supposons $F = M$. Puisque M appartient à (A) et (B), le polynôme M est un multiple de A et B. Par ailleurs, si un polynôme de $K[X]$ est multiple de A et

B, il est dans (M), c'est donc un multiple de M. Inversement, soit $F \in K[X]$ réalisant les conditions 1 et 2. On déduit de la condition 1 que F est dans (M). D'après la condition 2, M est dans (F). Par suite, on a $(M) = (F)$, puis $F = M$ vu que F et M sont unitaires.

Proposition 5.1. *Soit D le pgcd de A et B. On a $(AB) = (DM)$.*

Démonstration : Il s'agit de démontrer l'égalité d'idéaux

$$\Big(\frac{AB}{D^2}\Big) = \Big(\frac{M}{D}\Big).$$

L'entier M/D est le ppcm de A/D et B/D (cf. th. 5.5). Par ailleurs, les polynômes A/D et B/D sont premiers entre eux. On se ramène ainsi à prouver l'assertion dans le cas où $D = 1$. Supposons donc A et B premiers entre eux et vérifions que l'on a $(AB) = (M)$. Le polynôme AB est un multiple de A et B. Par ailleurs, soit C un multiple de A et B. Compte tenu du théorème 5.5, tout revient à vérifier que C est un multiple de AB. Il existe R et S dans $K[X]$ tels que $C = RA$ et $C = SB$. On a $RA = SB$, donc A divise SB. Puisque A est par hypothèse premier avec B, on déduit du théorème de Gauss que A divise S, ce qui entraîne le résultat.

3. Polynômes irréductibles

Définition 5.6. *Un polynôme de $K[X]$ est dit irréductible (dans $K[X]$) si son degré est supérieur ou égal à 1 et si l'ensemble de ses diviseurs est formé des éléments non nuls de K et des polynômes qui lui sont associés*[27].

Autrement dit, un polynôme $P \in K[X]$ de degré ≥ 1 est irréductible s'il ne possède pas de diviseur $Q \in K[X]$ tel que $1 \leq \deg(Q) \leq \deg(P) - 1$. Tel est le cas des polynômes de degré 1. Rappelons que ce sont les seuls si K est le corps $\mathbb{C}$ des nombres complexes. Deux polynômes irréductibles de $K[X]$ sont premiers entre eux ou sont associés. Un polynôme qui n'est pas irréductible est dit réductible.

Exercice 4. Montrer que le seul polynôme irréductible de degré 2 dans $\big(\mathbb{Z}/2\mathbb{Z}\big)[X]$ est $X^2 + X + 1$.

Exercice 5. Soit p un nombre premier. Quel est le nombre de polynômes unitaires de degré 2 dans l'anneau $\big(\mathbb{Z}/p\mathbb{Z}\big)[X]$? Montrer que le nombre de polynômes irréductibles unitaires de degré 2 dans cet anneau est $\frac{p(p-1)}{2}$.

Exercice 6. Montrer que le polynôme $X^2 - 2$ est irréductible dans $\mathbb{Q}[X]$.

[27] Cette définition est un cas particulier de la notion générale d'élément irréductible dans un anneau commutatif. Si A est un tel anneau, un élément $a \in A$ est dit irréductible s'il n'est pas inversible et si ses seuls diviseurs sont les éléments inversibles et les ua où u est inversible. Les nombres premiers et leurs opposés sont les éléments irréductibles de $\mathbb{Z}$. À titre indicatif, $2X$ n'est pas irréductible dans $\mathbb{Z}[X]$.

Soit $\mathbb{P}$ l'ensemble des polynômes irréductibles unitaires de $K[X]$. Comme dans le cas de l'anneau $\mathbb{Z}$, on a le résultat suivant, qui est le théorème fondamental de l'arithmétique de $K[X]$:

Théorème 5.6. *Soit P un polynôme non nul de $K[X]$. Alors, P s'écrit de manière unique sous la forme*

$$P = \lambda \prod_{F \in \mathbb{P}} F^{n_F}, \tag{4}$$

où $\lambda \in K$, et où les n_F sont des entiers naturels nuls sauf un nombre fini d'entre eux.

Démonstration : Cet énoncé est vrai si le degré de P est nul, auquel cas on prend $\lambda = P$ et tous les n_F nuls. Considérons alors un entier $n \geq 1$. Supposons que le résultat soit vrai pour tous les polynômes de degré $\leq n-1$ et que l'on ait $\deg(P) = n$. Soit E l'ensemble de tous les diviseurs de P de degré ≥ 1. Cet ensemble n'est pas vide car P est dans E. Il existe donc un élément $Q \in E$ de degré minimum. Ce polynôme est irréductible. Il existe $R \in K[X]$ tel que $P = QR$. On a $\deg(R) \leq n-1$. D'après l'hypothèse de récurrence, R possède une décomposition de la forme (4), et il en est donc de même de P. Cela établit l'assertion d'existence. Vérifions l'assertion d'unicité. Prouvons pour cela le résultat suivant :

Lemme 5.3. *Soit A un polynôme irréductible divisant un produit de polynômes $A_1 \cdots A_r$ dans $K[X]$. Alors, A divise l'un des A_i.*

Démonstration : Supposons le contraire. Puisque A est irréductible, cela signifie que pour tout $i = 1, \cdots, r$, les polynômes A et A_i sont premiers entre eux. Il existe donc des polynômes U_i et V_i dans $K[X]$ tels que l'on ait

$$U_i A + V_i A_i = 1 \quad \text{pour } i = 1, \cdots, r.$$

Par ailleurs, il existe $R \in K[X]$ tel que l'on ait

$$1 = \prod_{i=1}^{r} (U_i A + V_i A_i) = RA + \prod_{i=1}^{r} V_i A_i,$$

ce qui entraîne que A divise 1, et conduit à une contradiction.

Le théorème se déduit comme suit. Supposons que l'on ait deux décompositions de la forme (4) :

$$P = \lambda \prod_{F \in \mathbb{P}} F^{n_F} = \mu \prod_{F \in \mathbb{P}} F^{m_F}. \tag{5}$$

Soit F un élément de $\mathbb{P}$ tel que $n_F = 0$. Il résulte du lemme 5.3 que l'on a $m_F = 0$. Par suite, pour tout $F \in \mathbb{P}$, n_F est nul si et seulement si tel est le cas de m_F. Considérons

alors $F \in \mathbb{P}$ tel que $n_F > 0$. On a donc $m_F > 0$. En divisant les membres (5) par F, on obtient des égalités analogues avec un polynôme de degré $\leq n-1$. D'après l'hypothèse de récurrence, on a donc $\lambda = \mu$, $n_G = m_G$ pour tout $G \neq F$, et $n_F - 1 = m_F - 1$ i.e. $n_F = m_F$. D'où le théorème.

Comme conséquence des théorèmes 5.3, 5.5 et 5.6, on obtient l'énoncé suivant :

Corollaire 5.3. *Soient P et Q deux polynômes non nuls de $K[X]$. Soient*

$$P = \lambda \prod_{F \in \mathbb{P}} F^{n_F} \quad et \quad Q = \mu \prod_{F \in \mathbb{P}} F^{m_F},$$

les décompositions de P et Q en produit d'éléments de $\mathbb{P}$. Soient D et M respectivement le pgcd et le ppcm de P et Q. On a alors

$$D = \prod_{F \in \mathbb{P}} F^{\mathrm{Min}(n_F, m_F)} \quad et \quad M = \prod_{F \in \mathbb{P}} F^{\mathrm{Max}(n_F, m_F)}.$$

Pour tout $F \in \mathbb{P}$, l'égalité

$$\mathrm{Min}(n_F, m_F) + \mathrm{Max}(n_F, m_F) = n_F + m_F,$$

entraîne alors l'égalité d'idéaux

$$(DM) = (PQ),$$

déjà démontrée dans la proposition 5.1.

Exercice 7. Déterminer la décomposition en produit de facteurs irréductibles du polynôme $X^4 + 1$ dans $\mathbb{R}[X]$, $\big(\mathbb{Z}/2\mathbb{Z}\big)[X]$, $\big(\mathbb{Z}/3\mathbb{Z}\big)[X]$ et $\big(\mathbb{Z}/5\mathbb{Z}\big)[X]$. En utilisant la théorie des corps finis, qui fera l'objet du chapitre suivant, on peut en fait démontrer que le polynôme $X^4 + 1$ est réductible dans $\big(\mathbb{Z}/p\mathbb{Z}\big)[X]$ pour tout nombre premier p, tout en étant irréductible dans l'anneau $\mathbb{Z}[X]$.

4. Racines d'un polynôme

Définition 5.7. *Soit $P = a_0 + \cdots + a_n X^n$ un polynôme de $K[X]$. On appelle fonction polynôme associée à P l'application $\widetilde{P} : K \to K$ définie par*

$$\widetilde{P}(x) = \sum_{i=0}^{n} a_i x^i \quad quel\ que\ soit\ x \in K^{28}.$$

[28] On obtient ainsi une application $\Phi : K[X] \to F(K, K)$, de $K[X]$ à valeurs dans l'ensemble des applications de K dans K, définie par $\Phi(P) = \widetilde{P}$. C'est un homomorphisme d'anneaux. Si K est un corps fini, cette application n'est pas injective : par exemple si $K = \mathbb{Z}/p\mathbb{Z}$ où p est premier, on a $x^p - x = 0$ pour tout $x \in K$, pour autant le polynôme $X^p - X$ n'est pas nul. Si K est infini, l'application Φ est injective comme on le constatera plus loin.

Pour tout $a \in K$ fixé, l'application $K[X] \to K$ qui à $P \in K$ associe $\widetilde{P}(a)$ est un homomorphisme d'anneaux. Étant donnés $P \in K[X]$ et $a \in K$, on notera par abus $P(a)$ l'élément $\widetilde{P}(a)$.

Définition 5.8. *Soient P un élément de $K[X]$ et a un élément de K. On dit que a est une racine de P si l'on a $P(a) = 0$.*

Lemme 5.4. *Soient P un élément de $K[X]$ et a un élément de K. On a $P(a) = 0$ si et seulement si $X - a$ divise P*[29]*.*

Démonstration : Supposons $P(a) = 0$. D'après le théorème de division euclidienne, il existe Q et R dans $K[X]$ tels que $P = (X-a)Q+R$ avec $\deg(R) < 1$. On a $P(a) = 0$, d'où $R(a) = 0$. Puisque R est un élément de K, on a donc $R = 0$. La réciproque est immédiate.

Remarques 5.2.

1) Un polynôme de $K[X]$ de degré ≥ 2 qui possède une racine dans K est réductible (lemme 5.4). Par ailleurs, les polynômes de $K[X]$ de degré 1 sont irréductibles et cependant ils ont une racine dans K.

2) Soit P un polynôme de $K[X]$ de degré 2 ou 3. Alors, P est irréductible si et seulement si P n'a pas de racines dans K. C'est une conséquence de la première remarque et du fait que si $P = AB$ où $A, B \in K[X]$ sont non inversibles, alors le degré de P étant 2 ou 3, on a $\deg(A) = 1$ ou bien $\deg(B) = 1$, de sorte que P a une racine dans K.

3) Il est faux en général que la condition « P n'a pas de racines dans K » entraîne que P soit irréductible, comme le montre le polynôme $(X^2+1)^2 \in \mathbb{R}[X]$: il est réductible dans $\mathbb{R}[X]$ et sans racines dans $\mathbb{R}$.

Exercice 8. Démontrer que les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 ayant un discriminant négatif (on utilisera le fait que tout polynôme à coefficients dans $\mathbb{R}$ possède une racine dans $\mathbb{C}$).

Si $a \in K$ est une racine de $P \in K[X]$, il importe souvent de connaître la plus grande puissance de $X - a$ qui divise P. Cela conduit à la notion d'ordre de multiplicité.

---

[29] On peut évidemment définir, comme ci-dessus, la notion de racine d'un polynôme à coefficients dans un anneau commutatif K quelconque. Vérifions que cet énoncé est encore valable dans ce cas. Soient P un élément de $K[X]$ et $a \in K$ tels que $P(a) = 0$, et Y une autre indéterminée. En substituant à X le polynôme $a + Y \in K[Y]$, on obtient dans cet anneau le polynôme

$$P(a + Y) = a_0 + a_1 Y + \cdots + a_n Y^n$$

avec des $a_i \in K$. On a donc $P(a) = a_0$, d'où l'on déduit que $P(a + Y) = P(a) + YQ(Y)$ où $Q \in K[Y]$. En substituant Y par $X - a$, on a ainsi les égalités

$$P(X) = P(a) + (X - a)H(X) = (X - a)H(X),$$

où $H \in K[X]$, par suite $X - a$ divise P.

Définition 5.9. (Ordre de multiplicité d'une racine) *Soient P un polynôme non nul de $K[X]$ et $a \in K$ une racine de P. On appelle ordre de multiplicité de a (dans P) le plus grand entier naturel r tel que P soit divisible par $(X-a)^r$. Si $r=1$, on dit que a est racine simple de P, et si $r \geq 2$, on dit que a est une racine multiple de P.*

Afin de calculer r, il convient de définir la notion de polynôme dérivé :

Définition 5.10. (Polynôme dérivé) *Soit $P = a_0 + a_1X + \cdots + a_nX^n$ un polynôme de $K[X]$. On appelle polynôme dérivé de P, et on le note P', le polynôme*

$$P' = \sum_{i=1}^{n} i a_i X^{i-1}.$$

En particulier, si $\deg(P) = 0$, on a $P' = 0$. On vérifie que toutes les règles de dérivation usuelles sur les fonctions d'une variable réelle restent valables dans ce contexte. En effet, pour tous P et Q dans $K[X]$, $\lambda \in K$ et $n \geq 1$, on a les relations

$$(P+Q)' = P' + Q', \quad (\lambda P)' = \lambda P', \quad (PQ)' = P'Q + PQ', \quad P^n = nP^{n-1}P'.$$

Proposition 5.2. *Soit P un polynôme de $K[X]$. Pour qu'un élément $a \in K$ soit racine simple de P, il faut et il suffit que l'on ait*

$$P(a) = 0 \quad et \quad P'(a) \neq 0.$$

Démonstration : Soit a une racine de P. On a $P = (X-a)Q$ où $Q \in K[X]$. Par ailleurs, on a $P' = Q + (X-a)Q'$, d'où $Q(a) = P'(a)$. Si a est une racine simple, on a $Q(a) \neq 0$, d'où $P'(a) \neq 0$. Inversement, si $P'(a) \neq 0$, il en est de même de $Q(a)$. Par suite, $X-a$ ne divise pas Q, ce qui entraîne que $(X-a)^2$ ne divise pas P i.e. que a est racine simple de P.

Théorème 5.7. *Soient P un polynôme de $K[X]$ et $a_1, \cdots, a_k$ des éléments de K, distincts deux à deux, qui sont racines de P d'ordre de multiplicité $n_1, \cdots, n_k$ respectivement. Alors, il existe un polynôme $Q \in K[X]$, tel que l'on ait*

$$P = Q \prod_{i=1}^{k} (X-a_i)^{n_i},$$

et que $Q(a_i)$ soit non nul pour tout $i = 1, \cdots, k$.

Démonstration : On procède par récurrence sur k. L'énoncé est vrai si $k=1$. Considérons un entier $k \geq 2$ tel que cet énoncé soit vrai pour l'entier $k-1$. Il existe donc $R \in K[X]$ tel que l'on ait

$$P = R \prod_{i=1}^{k-1} (X-a_i)^{n_i}.$$

Par ailleurs, $(X-a_k)^{n_k}$ divise P et est premier avec le produit des $(X-a_i)^{n_i}$ pour i compris entre 1 et $k-1$. En effet dans le cas contraire, d'après le lemme 5.3, $X-a_k$ devrait diviser l'un des facteurs $(X-a_i)^{n_i}$ ce qui conduit à une contradiction vu que $a_k \neq a_i$ pour $i = 1, \cdots, k-1$. D'après le théorème de Gauss (th. 5.4), $(X-a_k)^{n_k}$ divise donc R, ce qui entraîne le résultat.

Corollaire 5.4. *Soit P un polynôme non nul de degré n dans $K[X]$. Alors, P possède au plus n racines distinctes dans K*[30].

Démonstration : Si P possédait (au moins) $n+1$ racines dans K il serait divisible par un polynôme de $K[X]$ de degré $n+1$, ce qui contredit le fait que P soit de degré n.

Ce résultat entraîne le fait que l'application $\Phi : K[X] \to F(K,K)$ définie par l'égalité $\Phi(P) = \widetilde{P}$ est injective si K est infini. En effet, si la fonction polynôme $\widetilde{P}$ associée à P est nulle sur K, cela signifie que P a une infinité de racines (car K est infini), et d'après le résultat précédent, P doit être le polynôme nul. Par suite, sur un corps infini, on peut identifier $K[X]$ et l'anneau des fonctions polynômes sur K i.e. l'image de Φ.

Comme application de ce qui précède, démontrons le théorème de Wilson (1741-1793), qui est une caractérisation des nombres premiers :

---

[30] En fait, ce résultat est encore vrai si l'anneau de base est un anneau intègre quelconque : soit F un polynôme de degré $n \geq 0$ à coefficients dans un anneau intègre A. Alors, F possède au plus n racines dans A. Pour le démontrer, on procède par récurrence sur n. Si $n = 0$, alors F est un élément non nul de A, donc ne possède aucune racine et le résultat est démontré dans ce cas. Supposons alors F de degré $n \geq 1$ et le résultat démontré pour tous les polynômes de degré $\leq n-1$. Soit $a \in A$ une racine de F. Il existe $Q \in A[X]$ tel que $F = (X-a)Q$ (lemme 5.4 et le bas de page numéro 29). Puisque A est intègre, le degré de Q est $n-1$ (cela se justifie comme dans le lemme 5.1). Par ailleurs, si $b \in A$ est une racine de F distincte de a, on a $(b-a)Q(b) = 0$, d'où $Q(b)$ car A est intègre. Ainsi les racines de F autres que a sont celles de Q. D'après l'hypothèse de récurrence, Q possède au plus $n-1$ racines dans A, donc F en possède au plus n, d'où le résultat.
En revanche, ce résultat est faux en général si A n'est pas un anneau intègre. On le constate par exemple en considérant le polynôme $(X-2)(X-3) \in \big(\mathbb{Z}/6\mathbb{Z}\big)[X]$, qui est de degré 2, et qui possède quatre racines dans $\mathbb{Z}/6\mathbb{Z}$, à savoir les classes de 0, 2, 3 et 5. Tel est aussi le cas du polynôme $2X \in \big(\mathbb{Z}/4\mathbb{Z}\big)[X]$ qui possède comme racines les classes de 0 et de 2 modulo $4\mathbb{Z}$. Il existe aussi des anneaux non intègres sur lesquels il existe des polynômes de degré 2 ayant une infinité de racines. En effet, soient E un ensemble infini et A l'anneau formé de l'ensemble des parties de E muni de la différence symétrique $\triangle$ et de l'intersection $\cap$ comme addition et multiplication. Rappelons que si U et V sont deux parties de E, on a par définition $U \triangle V = U \cup V - U \cap V$. L'élément neutre additif est l'ensemble vide et l'élément neutre multiplicatif est l'ensemble E. Toute partie U de E est alors racine du polynôme $X^2 + X \in A[X]$, qui a donc une infinité de racines si E est infini : on a ici $U^2 + U = (U \cap U) \triangle U = U \triangle U = \emptyset$.
Signalons le resultat suivant (exercice) : soit A un anneau commutatif non nul qui n'est pas isomorphe à $\mathbb{Z}/4\mathbb{Z}$ ni à $\big(\mathbb{Z}/2\mathbb{Z}\big)[X]/(X^2)$. Alors, A est intègre si et seulement si tout polynôme unitaire de $A[X]$ de degré 2 a au plus deux racines dans A.

Théorème 5.8. (Wilson) *Soit p un entier ≥ 2. Alors, p est premier si et seulement si $(p-1)!+1$ est divisible par p.*

Démonstration : Supposons que p soit un nombre premier. Il résulte du petit théorème de Fermat que pout tout a compris entre 1 et $p-1$, l'élément $a+p\mathbb{Z}$ est racine du polynôme $X^{p-1}-1 \in \big(\mathbb{Z}/p\mathbb{Z}\big)[X]$. D'après le théorème 5.7, on en déduit que l'on a dans cet anneau

$$X^{p-1}-1 = \prod_{a=1}^{p-1}(X-\overline{a}).$$

En exprimant le fait que les termes constants sont les mêmes, on obtient l'égalité dans le corps $\mathbb{Z}/p\mathbb{Z}$

$$-\overline{1} = (-1)^{p-1}\prod_{a=1}^{p-1}\overline{a}.$$

Autrement dit, on a la congruence

$$-1 \equiv (-1)^{p-1}(p-1)! \text{ mod. } p,$$

par suite p divise $(p-1)!+1$. Inversement, supposons $(p-1)!+1$ divisible par p. Si ℓ est un diviseur positif de p autre que p, alors ℓ divise $(p-1)!$, d'où $\ell = 1$ et le fait que p soit un nombre premier.

Remarques 5.3.

1) Pour démontrer le théorème 5.8 on peut aussi utiliser l'argument suivant. Supposons p premier. Dans le corps $\mathbb{Z}/p\mathbb{Z}$, les seuls éléments égaux à leur inverse sont ± 1. Il en résulte que l'on a (en regroupant chaque terme du produit avec son inverse modulo p)

$$\prod_{k=2}^{p-2} k \equiv 1 \text{ mod. } p.$$

Par suite, on a $(p-1)! \equiv p-1$ mod. p i.e. $(p-1)!+1 \equiv 0$ mod. p.

2) Le théorème de Wilson est un test de primalité, mais il n'est pas efficace car le calcul de $(p-1)!$ nécessite beaucoup d'opérations. Signalons que si p est un nombre premier, on définit le quotient de Wilson

$$W(p) = \frac{(p-1)!+1}{p},$$

qui est donc un entier. On dit que p est un nombre premier de Wilson si p divise $W(p)$, autrement dit, si l'on a $(p-1)!+1 \equiv 0$ mod. p^2. Par exemple, 5 et 13 sont des nombres premiers de Wilson. La question de savoir s'il existe une infinité de tels nombres premiers est ouverte. En dehors de 5 et 13, on ne connaît qu'un seul autre nombre premier de Wilson, à savoir 563 (découvert en 1953). Il n'y en a pas d'autres plus petits que 5.10^8.

5. Les algèbres quotients de $K[X]$ modulo un idéal

Étant donné un idéal I de $K[X]$, on va définir dans ce paragraphe une structure naturelle de K-algèbre sur l'ensemble quotient $K[X]/I$. Afin de définir une structure de K-algèbre, il faut définir une structure d'anneau et une structure de K-espace vectoriel. Rappelons la définition de cette notion en commençant par celle d'espace vectoriel sur K.

Définition 5.11. *On appelle K-espace vectoriel tout ensemble E muni d'une structure définie par la donnée :*

1) *d'une loi de groupe abélien sur E ;*
2) *d'une application $K \times E \to E$, notée $(\lambda, x) \mapsto \lambda x$, souvent appelée loi de composition externe, pour laquelle on a les relations*

$$(\lambda + \mu)x = \lambda x + \mu x, \quad \lambda(x + y) = \lambda x + \lambda y,$$

$$\lambda(\mu x) = (\lambda\mu)x, \quad 1x = x,$$

quels que soient $x, y \in E$ et $\lambda, \mu \in K$.

Définition 5.12. *On appelle K-algèbre tout ensemble E muni d'une structure définie par la donnée :*

1) *de deux lois de composition sur E , une addition $+$ et une multiplication $\times$, telles que $(E, +, \times)$ soit un anneau,*
2) *d'une loi de composition externe $K \times E \to E$, notée $(\lambda, x) \mapsto \lambda x$, qui avec la loi additive $+$, munie E d'une structure de K-espace vectoriel,*

de telle sorte que la condition suivante soit satisfaite :

3) *quels que soient $\lambda \in K$ et $x, y \in E$, on a*

$$\lambda(x \times y) = (\lambda x) \times y = x \times (\lambda y). \tag{6}$$

On définit la notion d'homomorphisme de K-algèbres entre deux K-algèbres. Ce sont les applications K-linéaires qui sont en même temps des homomorphismes d'anneaux. Deux K-algèbres sont dites isomorphes si elles sont liées par un homomorphisme bijectif. On omet très souvent, par abus, la notation $\times$ dans les calculs dans les K-algèbres. Cela ne prête pas à confusion pourvu que l'on précise la nature des éléments que l'on compose. La condition (6) s'écrit alors

$$\lambda(xy) = (\lambda x)y = x(\lambda y) \quad \text{quels que soient } \lambda \in K \quad \text{et} \quad x, y \in E.$$

Exemples 5.1.

1) Pour tout $n \geq 1$, l'ensemble $K^n = K \times \cdots \times K$ (n facteurs) est muni d'une structure de K-algèbre, pour laquelle la structure d'anneau est celle définie p. 51, et la loi externe est donnée par l'égalité

$$\lambda(x_1, \cdots, x_n) = (\lambda x_1, \cdots, \lambda x_n) \quad \text{quels que soient } \lambda \in K \quad \text{et} \quad (x_1, \cdots, x_n) \in K^n.$$

Notons que l'on a $\lambda(x_1, \cdots, x_n) = (\lambda, \cdots, \lambda)(x_1, \cdots, x_n)$, et que l'application $K \to K^n$ qui à λ associe $(\lambda, \cdots, \lambda)$ est un homomorphisme de K-algèbres injectif, que l'on appelle le plongement diagonal de K dans K^n.

2) Si L est un surcorps (commutatif) de K, alors L est muni de la structure de K-algèbre, pour laquelle la structure d'anneau sur L est celle donnée dans sa définition de corps et la loi externe $K \times L \to L$ est celle qui au couple $(\lambda, x) \in K \times L$ associe le produit λx dans L.

3) L'anneau $K[X]$ est muni de la structure de K-algèbre pour laquelle la structure d'anneau est celle définie p. 50, et la loi externe $K \times K[X] \to K[X]$ est celle qui au couple (λ, P) associe le polynôme λP obtenu en multipliant dans K les coefficients de P par λ.

4) L'anneau des matrices carrées $\mathbb{M}_n(K)$ est aussi naturellement muni d'une structure de K-algèbre (exercice : expliciter cette structure).

Structure de K-algèbre sur le quotient de $K[X]$ modulo un idéal

Considérons un idéal I de $K[X]$. Nous allons munir ici l'ensemble quotient $K[X]/I$ d'une structure de K-algèbre. Pour tout $P \in K[X]$, posons $\overline{P} = P + I$ la classe de P modulo I. Rappelons que $\overline{P}$ est le sous-ensemble de $K[X]$ formé des polynômes Q tels que $P - Q$ appartienne à I. Conformément au paragraphe 3 du chapitre III, l'ensemble $K[X]/I$ est muni de la structure d'anneau définie par les égalités

$$\overline{P} + \overline{Q} = \overline{P + Q}, \tag{7}$$

$$\overline{P}\ \overline{Q} = \overline{PQ}, \tag{8}$$

quels que soient P et Q dans $K[X]$. L'élément neutre additif est $\overline{0} = I$ et l'élément neutre multiplicatif est $\overline{1}$.

Cet anneau est aussi muni d'une structure de K-espace vectoriel définie comme suit. Tout d'abord, muni de son addition définie par l'égalité (7), $K[X]/I$ est un groupe abélien. Par ailleurs, l'application

$$K \times K[X]/I \to K[X]/I$$

qui au couple $(\lambda, \overline{P}) \in K \times K[X]/I$ associe $\overline{\lambda P}$ réalise la condition 2 de la définition 5.11. Avec les notations de cette définition, on a donc l'égalité

$$\lambda \overline{P} = \overline{\lambda P}. \tag{9}$$

Il convient de vérifier que cette définition a bien un sens, autrement dit, qu'elle ne dépend que de la classe de P et non pas d'un de ses représentants. Considérons pour cela P et Q dans $K[X]$ tels que $\overline{P} = \overline{Q}$. Le polynôme $P - Q$ appartient à I, par suite $\lambda(P - Q)$ est aussi dans I. De l'égalité $\lambda(P - Q) = \lambda P - \lambda Q$, on déduit alors que $\overline{\lambda P} = \overline{\lambda Q}$, d'où notre assertion. Les relations de la condition 2 sont alors des conséquences directes des égalités (7) et (9). Les égalités (7), (8) et (9) munissent ainsi $K[X]/I$ d'un structure de K-algèbre.

Remarques 5.4.

1) Il résulte des égalités (8) et (9) que l'on a pour tous $\lambda \in K$ et $P \in K[X]$,

$$(\lambda + I)(P + I) = \lambda P + I = \lambda(P + I). \tag{10}$$

2) Supposons I distinct de $K[X]$. Le corps K est isomorphe à un sous-anneau de $K[X]/I$. En effet, soit $i : K \to K[X]/I$ l'application définie par

$$i(\lambda) = \lambda + I.$$

Puisque l'on a $I \neq K[X]$, i est un homomorphisme d'anneaux injectif. En effet, il résulte des définitions que i est un homomorphisme d'anneaux. Par ailleurs, soit λ dans K tel que $\lambda + I = 0$ i.e. tel que λ soit dans I. Puisque l'on a $I \neq K[X]$, et que les éléments non nuls de K sont inversibles dans $K[X]$, on a donc $\lambda = 0$ et i est injectif. (cf. on peut aussi utiliser directement l'exercice 9 du chapitre III). Ainsi $i(K)$ est un sous-anneau de $K[X]/I$ isomorphe à K. Au cours des calculs dans l'algèbre $K[X]/I$, on identifie toujours K et $i(K)$. Avec cette identification, la multiplication dans $K[X]/I$ par un élément de K se note de la même façon que la loi externe de K sur $K[X]/I$ (cf. l'égalité (10) dans laquelle $\lambda + I$ est identifié à λ).

Si I n'est pas nul, il existe alors un polynôme P tel que l'on ait $I = (P)$ (th. 5.2) (on peut si on le souhaite demander que P soit unitaire, auquel cas il y a unicité de P, mais peu importe ici). Le chapitre suivant est consacré, entre autres, à l'étude des algèbres $K[X]/(P)$ dans le cas où K est un corps fini. Décrivons maintenant quelques propriétés des algèbres $K[X]/(P)$ en termes du polynôme P, que l'on utilisera dans la suite. Commençons par une propriété fondamentale concernant la structure de K-espace vectoriel.

Théorème 5.9. *Soit P un polynôme de $K[X]$ degré $n \geq 0$. Le K-espace vectoriel $K[X]/(P)$ est de dimension finie n. Plus précisément, en posant $\alpha = X + (P)$, le système $\left(\alpha^i\right)_{0 \leq i \leq n-1}$ est une K-base de $K[X]/(P)$.*

Démonstration : Vérifions que $\left(\alpha^i\right)_{0 \leq i \leq n-1}$ est un système libre. Soient $\lambda_0, \cdots, \lambda_{n-1}$ des éléments de K tels que

$$\sum_{i=0}^{n-1} \lambda_i \alpha^i = 0.$$

Cette égalité signifie que l'on a

$$\sum_{i=0}^{n-1} \lambda_i X^i \in (P).$$

Puisque P est de degré n, cela entraîne que tous les λ_i sont nuls. Démontrons que le système considéré est générateur. Soit ξ un élément de $K[X]/(P)$. Il existe un polynôme $F \in K[X]$ tel que $\xi = F + (P)$. On a $P \neq 0$. D'après le théorème de division euclidienne, il existe Q et R dans $K[X]$ tels que l'on ait $F = PQ + R$ avec $\deg(R) < \deg(P)$. On a donc $\xi = \overline{R}$, ce qui entraîne notre assertion et le résultat.

Le lemme suivant est d'un usage constant pour effectuer des calculs dans les K-algèbres quotients de $K[X]$:

Lemme 5.5. *Soit $P = a_0 + a_1 X + \cdots + a_n X^n \in K[X]$ un polynôme degré $n \geq 0$. Posons $\alpha = X + (P) \in K[X]/(P)$. On a l'égalité*

$$\sum_{i=0}^{n} a_i \alpha^i = 0. \tag{11}$$

Démonstration : On a $\overline{P} = 0$. Cette égalité signifie que l'on a

$$\sum_{i=0}^{n} \overline{a_i X^i} = \sum_{i=0}^{n} a_i \overline{X}^i = 0,$$

d'où l'assertion.

Remarques 5.5.

1) Dans l'énoncé du théorème 5.9, si $n = 0$, alors la dimension de $K[X]/(P)$ est nulle et la base vide est une base de cet espace vectoriel. Cela était prévisible vu que P est dans ce cas un élément non nul de K, donc est inversible dans $K[X]$, par suite $(P) = K[X]$ et le quotient $K[X]/(P)$ est l'anneau nul.

2) Le théorème 5.9 se traduit par l'égalité

$$K[X]/(P) = \left\{ \sum_{i=0}^{n-1} a_i \alpha^i \mid a_i \in K \right\}.$$

3) En identifiant K et son image dans $K[X]/(P)$, l'égalité (11) s'obtient alors en substituant X par α dans P (cf. (10)). Elle peut donc aussi s'écrire $P(\alpha) = 0$.

4) Si I est l'idéal nul, $K[X]/I$ est isomorphe comme K-algèbre à $K[X]$. En particulier, $K[X]/I$ est dans ce cas un K-espace vectoriel de dimension infinie (une K-base de $K[X]$ est formée des $(X^n)_{n\geq 0}$).

5) Soient x et y deux élément de $K[X]/(P)$ ($P \in K[X]$ de degré $n \geq 1$). Ils s'écrivent de manière unique sous la forme

$$x = \sum_{i=0}^{n-1} a_i \alpha^i \quad \text{et} \quad y = \sum_{i=0}^{n-1} b_i \alpha^i,$$

où les a_i et b_i sont dans K. Les coordonnées de $x + y$ dans la base $\left(\alpha^i\right)$ sont les $a_i + b_i$. Il est moins simple d'obtenir les coordonnées du produit xy. Afin de les déterminer, on utilise l'égalité fondamentale (11). On peut alors procéder de deux façons. La première consiste par exemple à déterminer le reste R de la division euclidienne du polynôme produit $\left(\sum a_i X^i\right)\left(\sum b_i X^i\right)$ par P. On a alors

$$xy = R(\alpha),$$

et R étant de degré $\leq n - 1$, on obtient les coordonnées cherchées. On peut aussi utiliser directement l'égalité (11), afin de déduire les coordonnées des α^k pour $k \geq n$, puis celles de xy.

Le résultat qui suit concerne la structure d'anneau de $K[X]/(P)$, qui n'est autre que l'analogue du théorème 4.1 sur la description des éléments inversibles des quotients de $\mathbb{Z}$.

Théorème 5.10. *Soit P un polynôme de $K[X]$ de degré $n \geq 0$. Le groupe des éléments inversibles de l'anneau $K[X]/(P)$ est formé des classes de polynômes $F \in K[X]$ telles que F soit premier avec P.*

Démonstration : Soit F un polynôme de $K[X]$ premier avec P. Il existe U et V dans $K[X]$ tels que $UP + VF = 1$. On a donc $\overline{V}\ \overline{F} = 1$, ce qui prouve que $\overline{F}$ est inversible. Inversement, soit $\overline{F}$ un élément inversible de $K[X]/(P)$. Il existe alors $Q \in K[X]$ tel que $\overline{F}\ \overline{Q} = 1$. Cette égalité signifie que $FQ - 1$ appartient à (P), autrement dit qu'il existe $U \in K[X]$ tel que $FQ + UP = 1$, donc F et P sont premiers entre eux, d'où le résultat.

Corollaire 5.5. *Soit P un polynôme de $K[X]$ de degré $n \geq 0$. Les conditions suivantes sont équivalentes :*

1) l'anneau $K[X]/(P)$ est intègre.

2) Le polynôme P est irréductible dans $K[X]$.

3) L'anneau $K[X]/(P)$ est un corps.

Démonstration : Supposons que $K[X]/(P)$ soit intègre. Tout d'abord, P n'est pas inversible, sinon on a $(P) = K[X]$ et $K[X]/(P)$ est l'anneau nul, ce qui est exclu par définition. Soit F un diviseur de P. Il s'agit de montrer que F est inversible ou bien que F et P sont associés. Il existe $Q \in K[X]$ tel que $P = FQ$, d'où $\overline{F}\ \overline{Q} = 0$. Par hypothèse, cela entraîne que $\overline{F} = 0$ ou bien que $\overline{Q} = 0$. Si $\overline{F} = 0$, alors F est dans (P) i.e. P

divise F, donc P et F sont associés. Si $\overline{Q} = 0$, alors Q et P sont associés, par suite on a $\deg(F) = 0$ i.e. F est inversible. Cela prouve que P est irréductible dans $K[X]$. Supposons alors P irréductible dans $K[X]$ et prouvons que tout élément non nul $\overline{F}$ de $K[X]/(P)$ est inversible. Puisque P est irréductible et que P ne divise pas F, les polynômes F et P sont premiers entre eux. D'après le théorème 5.10, $\overline{F}$ est donc inversible, donc $K[X]/(P)$ est un corps. La dernière implication est immédiate.

On déduit de ce qui précède le résultat fondamental suivant :

Théorème 5.11. *Soit P un polynôme irréductible de $K[X]$. Il existe un corps commutatif L contenant K comme sous-corps et possédant les deux propriétés suivantes :*

1) le polynôme P a une racine dans L.

2) Le K-espace vectoriel L est de dimension finie sur K, égale au degré de P.

Démonstration : Puisque P est irréductible, l'anneau $A = K[X]/(P)$ est un corps (cor. 5.5). Soit $i : K \to A$ l'application définie pour tout $\lambda \in K$ par $i(\lambda) = \lambda + (P)$. Soient Z le complémentaire de $i(K)$ dans A, et L la réunion des ensembles K et Z. L'application $\psi : A \to L$ définie pour tout $\lambda \in K$ et tout $x \in Z$ par

$$\psi\big(i(\lambda)\big) = \lambda \quad \text{et} \quad \psi(x) = x,$$

est une bijection de A sur L. Par transport de structure via ψ, on peut donc munir L d'une structure de K-algèbre : pour tous ξ_1 et ξ_2 dans L, on définit l'addition et la multiplication par les formules

$$\xi_1 + \xi_2 = \psi\big(\psi^{-1}(\xi_1) + \psi^{-1}(\xi_2)\big) \quad \text{et} \quad \xi_1 \times \xi_2 = \psi\big(\psi^{-1}(\xi_1) \times \psi^{-1}(\xi_2)\big),$$

la loi externe de K sur L étant définie pour tous $\lambda \in K$ et $\xi \in L$ par l'égalité

$$\lambda.\xi = \psi\big(\lambda.\psi^{-1}(\xi)\big).$$

Par définition des lois de composition sur L, les K-algèbres A et L sont isomorphes via ψ et l'on a $\psi\big(i(K)\big) = K$. En particulier, L est de dimension finie sur K, égale au degré de P (th. 5.9), et L est un surcorps de K. Il reste à vérifier que P a une racine dans L. Soit α la classe de X modulo (P). On a $\overline{P} = 0$, autrement dit, si $P = a_0 + \cdots + a_n X^n$, on a $i(a_0) + \cdots + i(a_n)\alpha^n = 0$ et en prenant l'image par ψ des deux membres de cette égalité, on obtient $P\big(\psi(\alpha)\big) = 0$ i.e. $\psi(\alpha)$ est une racine de P dans L, d'où le résultat.

Remarques 5.6.

1) Dans la démonstration précédente, on a utilisé le fait qu'étant donnés deux ensembles X et Y, il en existe un autre contenant à la fois X et Y, à savoir leur réunion. Il convient de noter que cela est un axiome de la théorie des ensembles.

2) Si $P \in K[X]$ n'est pas irréductible, de degré ≥ 1, il existe aussi un surcorps de K dans lequel P a une racine, car P est produit de polynômes irréductibles (th. 5.6).

3) Considérons la question suivante : soit L un corps commutatif contenant K dans lequel le polynôme P a une racine. Alors, P a-t-il toutes ses racines dans L, autrement dit, P est-il produit de polynômes de degré 1 dans $L[X]$? La réponse est négative en général. En effet, prenons par exemple $K = \mathbb{Q}$ et $P = X^3 - 2 \in \mathbb{Q}[X]$. Soient α la racine réelle de P et L le sous-ensemble de $\mathbb{R}$ formé des éléments $a + b\alpha + c\alpha^2$ où $a, b, c \in \mathbb{Q}$. Alors, L est un sous-corps de $\mathbb{R}$ et α est la seule racine de P dans L vu que ses deux autres racines, $j\alpha$ et $j^2\alpha$ avec $j^3 = 1$ et $j \neq 1$, ne sont pas réelles. Cela étant, on démontrera au chapitre suivant que la réponse est positive si K et L sont des corps finis.

Exemples 5.2.

1) Prenons $K = \mathbb{Z}/2\mathbb{Z}$ et $P = X^3 + X + 1 \in K[X]$. Puisque P est de de degré 3 et qu'il n'a pas de racines dans $\mathbb{Z}/2\mathbb{Z}$, il est irréductible dans $\big(\mathbb{Z}/2\mathbb{Z}\big)[X]$. Le quotient $K[X]/(P)$ est donc un corps et un espace vectoriel de dimension 3 sur $\mathbb{Z}/2\mathbb{Z}$. En particulier, c'est un corps à huit éléments (cf. la décomposition des éléments dans une base). Vérifions que P a toutes ses racines dans K. Si α est la classe de X modulo (P), on a $P(\alpha) = 0$, ce qui entraîne

$$P = (X - \alpha)(X^2 + \alpha X + 1 + \alpha^2).$$

On vérifie ensuite que α^2 et $\alpha^4 = \alpha + \alpha^2$ sont racines de $X^2 + \alpha X + 1 + \alpha^2$. (Le corps considéré ayant huit éléments, on peut par exemple tester tous les éléments pour trouver les racines. Notons que les formules classiques de résolution d'une équation du second degré ne fonctionnent pas ici car $2 = 0$. On peut aussi remarquer que si β est racine d'un polynôme de $\big(\mathbb{Z}/2\mathbb{Z}\big)[X]$, il en est de même de β^2 : cf. la formule du binôme de Newton. Cette remarque sera généralisée au chapitre suivant). On a donc (puisque $-1 = 1$ dans K)

$$P = (X + \alpha)(X + \alpha^2)(X + \alpha + \alpha^2).$$

2) Prenons $K = \mathbb{Q}$ et $P = X^3 + X + 1 \in \mathbb{Q}[X]$. La $\mathbb{Q}$-algèbre $\mathbb{Q}[X]/(P)$ est de dimension 3 (pour tout corps K, la dimension d'une K-algèbre est sa dimension en tant que K-espace vectoriel). Si $\alpha = X + (P)$, le système $(1, \alpha, \alpha^2)$ est une $\mathbb{Q}$-base de $\mathbb{Q}[X]/(P)$. On a $P(\alpha) = 0$. Explicitons les coordonnées de l'élément $\alpha^5 + 1$ dans la base $(1, \alpha, \alpha^2)$. On peut effectuer pour cela la division euclidienne du polynôme $X^5 + 1$ par P. On trouve

$$X^5 + 1 = (X^2 - 1)P + (-X^2 + X + 2),$$

de sorte que l'on obtient $\alpha^5 + 1 = -\alpha^2 + \alpha + 2$ et les coordonnées cherchées sont donc $(2, 1, -1)$. Vérifions que $\mathbb{Q}[X]/(P)$ est un corps, autrement dit que P est irréductible dans $\mathbb{Q}[X]$. Puisque P est de degré 3, il s'agit de montrer que P n'a pas de racines dans $\mathbb{Q}$. On utilise pour cela le lemme suivant très utile en pratique :

Lemme 5.6. *Soit $F = a_0 + \cdots a_{n-1}X^{n-1} + X^n \in \mathbb{Z}[X]$ un polynôme unitaire de degré n. Alors, si F a une racine dans $\mathbb{Q}$, elle est dans $\mathbb{Z}$ et elle divise a_0.*

Démonstration : Soit β une racine de F dans $\mathbb{Q}$. Posons $\beta = u/v$, où u et v sont deux entiers premiers entre eux. On a l'égalité

$$v^n a_0 + a_1 v^{n-1} u + \cdots + a_{n-1} v u^{n-1} + u^n = 0.$$

Il en résulte que v divise u^n, puis que $v = \pm 1$ car u et v sont premiers entre eux. Ainsi β est dans $\mathbb{Z}$. Par ailleurs, u divise $v^n a_0 = \pm a_0$, d'où le lemme.

Notre assertion s'en déduit aussitôt : si F a une racine dans $\mathbb{Q}$, cette racine est ± 1, qui n'est pas racine de F.

À titre indicatif, déterminons les coordonnées de l'inverse de l'élément $1 + \alpha$ dans la base $(1, \alpha, \alpha^2)$. Voici une méthode possible. On considère le $\mathbb{Q}$-endomorphisme ψ de $\mathbb{Q}[X]/(P)$ défini par $a \mapsto a(1 + \alpha)$ (c'est l'endomorphisme de multiplictaion par $1 + \alpha$). C'est un endomorphisme bijectif. La matrice de ψ dans la base considérée est

$$M = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

On vérifie que l'inverse de M est

$$M^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

L'image de 1 par ψ^{-1} est alors l'élément cherché. On trouve ainsi

$$(1 + \alpha)^{-1} = 2 - \alpha + \alpha^2.$$

Une autre méthode consiste à utiliser l'algorithme d'Euclide. On trouve que l'on a la relation de Bezout

$$(1 + X)(X^2 - X + 2) - (X^3 + X + 1) = 1,$$

d'où de nouveau l'égalité ci-dessus.

3) Vérifions que le corps $\mathbb{C}$ des nombres complexes est isomorphe à $\mathbb{R}[X]/(X^2+1)$. On remarque d'abord que $X^2 + 1$, n'ayant pas de racines dans $\mathbb{R}$, est irréductible dans $\mathbb{R}[X]$, donc $\mathbb{R}[X]/(X^2+1)$ est un corps. C'est par ailleurs un espace vectoriel de dimension 2 sur $\mathbb{R}$. Considérons alors l'application $\psi : \mathbb{R}[X] \to \mathbb{C}$ définie par $\psi(F) = F(i)$ où $i^2 = -1$ (cette égalité signifie que l'on a choisi implicitement une racine carrée i de -1 dans $\mathbb{C}$). C'est un homomorphisme de corps. Il est surjectif puisque $\psi(a + bX) = a + ib$ pour tous $a, b \in \mathbb{R}$. Vérifions que son noyau est (X^2+1), ce qui, compte tenu du théorème 3.2, prouvera notre

assertion. Tout d'abord, il est immédiat de constater que l'idéal (X^2+1) est contenu dans $\mathrm{Ker}(\psi)$. Inversement, soit F un élément de $\mathrm{Ker}(\psi)$. Il existe deux polynômes Q et R de $\mathbb{R}[X]$ tels que $F = (X^2+1)Q+R$ avec $\deg(R) \leq 1$. On a donc $\psi(F) = \psi(R) = 0$. Par suite, si $R = aX + b$, on a l'égalité $ai + b = 0$, d'où $a = b = 0$ puis $R = 0$, donc F appartient à (X^2+1). On a ainsi $\ker(\psi) = (X^2+1)$. Vu que ψ est $\mathbb{R}$-linéaire, on a en fait montré que les $\mathbb{R}$-algèbres $\mathbb{C}$ et $\mathbb{R}[X]/(X^2+1)$ sont isomorphes. On pourrait ainsi poser par définition $\mathbb{C} = \mathbb{R}[X]/(X^2+1)$[31].

Nous verrons dans le chapitre suivant que tous les corps finis s'obtiennent par une construction analogue à la précédente qui fait passer de $\mathbb{R}$ à $\mathbb{C}$.

[31] Rappelons que l'on peut définir le corps $\mathbb{C}$ comme le produit cartésien $\mathbb{R} \times \mathbb{R}$ muni des deux lois de compositions suivantes : pour tous (x, y) et $(z, t) \in \mathbb{R} \times \mathbb{R}$, on pose

$$(x, y) + (z, t) = (x + z, y + t) \quad \text{(l'addition)},$$

$$(x, y) \times (z, t) = (xz - yt, xt + yz) \quad \text{(la multiplication)}.$$

On vérifie alors que le triplet $(\mathbb{C}, +, \times)$ est un corps commutatif, appelé corps des nombres complexes. L'application de $\mathbb{R}$ dans $\mathbb{C}$ qui à x associe $(x, 0)$ est un homomorphisme de corps, donc est injectif. On note i l'élément $(0, 1)$. En identifiant $\mathbb{R}$ et son image dans $\mathbb{C}$, on a alors l'égalité attendue $i^2 = -1$ et tout nombre complexe s'écrit de façon unique sous la forme $a + ib$ avec $a, b \in \mathbb{R}$.

Chapitre VI — Corps finis - Construction

L'objectif de ce chapitre est de construire les corps finis et de donner quelques applications à la cryptographie. On admettra dans toute la suite le résultat suivant, dont la démonstration dépasserait le niveau de ce cours :

Théorème 6.1. (Wedderburn 1882 - 1948) *Tout corps fini est commutatif.*

Les premiers exemples de corps finis sont les quotients de l'anneau $\mathbb{Z}$

$$\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z},$$

où p est un nombre premier. Compte tenu du chapitre précédent, d'autres exemples sont fournis par les quotients

$$\mathbb{F}_p[X]/(F),$$

où F est un polynôme irréductible de $\mathbb{F}_p[X]$. Ce sont en effet des corps (cor. 5.5), et ils sont finis, puisqu'ils sont de dimension finie sur $\mathbb{F}_p$. Nous reviendrons sur ce point, et démontrerons que l'on obtient de la sorte tous les corps finis. On prouvera dans les sept premiers paragraphes les énoncés suivants :

1) tout corps fini contient un sous-corps isomorphe à un corps $\mathbb{F}_p$.
2) Le cardinal d'un corps fini est une puissance d'un nombre premier.
3) Le groupe multiplicatif d'un corps fini est cyclique.
4) Tout corps fini K de cardinal p^n est isomorphe à $\mathbb{F}_p[X]/(F)$, où F est un polynôme irréductible de degré n dans $\mathbb{F}_p[X]$.
5) Pour tout nombre premier p et tout entier $n \geq 1$, il existe un corps à p^n éléments et il est unique à isomorphisme près.

Il est assez simple de démontrer les quatre premiers résultats, notamment les 1, 2 et 4. Le cinquième l'est beaucoup moins.

1. Caractéristique d'un anneau

Soit A un anneau. Notons 1_A l'élément neutre multiplicatif de A. Soit $f : \mathbb{Z} \to A$ l'application de $\mathbb{Z}$ dans A définie par

$$f(m) = m1_A \quad \text{pour tout } m \in \mathbb{Z}. \tag{1}$$

C'est un homomorphisme d'anneaux de $\mathbb{Z}$ dans A (et d'ailleurs le seul). Son noyau est un idéal de $\mathbb{Z}$. Il existe donc un unique entier naturel n tel que l'on ait

$$\mathrm{Ker}(f) = n\mathbb{Z}.$$

Définition 6.1. *L'entier n est la caractéristique de A.*

Lemme 6.1. *Si A est intègre, sa caractéristique est nulle ou est un nombre premier. Tel est en particulier le cas si A est un corps commutatif.*

Démonstration : L'anneau $\mathbb{Z}/n\mathbb{Z}$ est isomorphe à un sous-anneau de A, à savoir l'image de f (th. 3.2). Puisque A est intègre, il en est donc de même de $\mathbb{Z}/n\mathbb{Z}$. Si n n'est pas nul, $\mathbb{Z}/n\mathbb{Z}$ est alors un corps (prop. 3.2), et n est un nombre premier (cor. 4.2).

Théorème 6.2. *Soit K un corps commutatif d'élément neutre multiplicatif 1_K. Soit m un entier relatif.*

1) *Supposons K de caractéristique zéro. On a $m1_K = 0$ si et seulement si $m = 0$. Dans ce cas, K contient un sous-corps isomorphe à $\mathbb{Q}$.*
2) *Supposons K de caractéristique un nombre premier p. On a $m1_K = 0$ si et seulement si p divise m. Dans ce cas, K contient un sous-corps isomorphe à $\mathbb{F}_p$.*

Démonstration : Supposons K de caractéristique 0. L'homomorphisme f défini par (1) est alors injectif, d'où l'équivalence annoncée. L'application de $\mathbb{Q}$ dans K qui à $a/b \in \mathbb{Q}$ associe $a1_K(b1_K)^{-1}$, prolonge f de $\mathbb{Z}$ à $\mathbb{Q}$, et est un homomorphisme de corps. (Notons que b étant non nul, on a $b1_K \neq 0$ et l'on vérifie que f est bien définie). Son image est donc un sous-corps de K isomorphe à $\mathbb{Q}$. Si K est de caractéristique p, le noyau de f est l'idéal $p\mathbb{Z}$. Par suite, on a $m1_K = 0$ i.e. m appartient au noyau de f si et seulement si p divise m. L'image de f est alors un sous-corps de K isomorphe à $\mathbb{F}_p$.

Par exemple $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ sont de caractéristique 0. Pour tout p premier, $\mathbb{F}_p$ est de caractéristique p. On obtient aussitôt les résultats 1 et 2 énoncés précédemment :

Corollaire 6.1. *Soit K un corps fini. La caractéristique de K est un nombre premier p et K contient un sous-corps isomorphe à $\mathbb{F}_p$. De plus, il existe un entier $n \geq 1$ tel que le cardinal de K soit p^n.*

Démonstration : Puisque K est fini, K ne contient pas de sous-corps isomorphe à $\mathbb{Q}$. La caractéristique de K est donc un nombre premier p et K contient un sous-corps isomorphe à $\mathbb{F}_p$ (th. 6.2). Par suite, K est naturellement muni d'une structure d'espace vectoriel sur $\mathbb{F}_p$ (cf. exemples 5.1, 3). Le corps K étant fini, la dimension de K sur $\mathbb{F}_p$ est aussi finie. Si n est cette dimension, K est donc isomorphe, comme espace vectoriel, à $\mathbb{F}_p^n$, et K est de cardinal p^n [32].

[32] Rappelons que deux espaces vectoriels sur un corps sont isomorphes si et seulement si ils ont la même dimension. En particulier, tout espace vectoriel de dimension n sur K est isomorphe à K^n (le fait que K soit fini n'intervient pas ici). Pour justifier que $|K| = p^n$, on peut aussi choisir une base de K sur $\mathbb{F}_p$. Les coordonnées des éléments de K étant dans $\mathbb{F}_p$, il y a p choix possibles pour chaque coordonnée, d'où les p^n éléments attendus, puisque tout élément de K s'écrit de façon unique comme une combinaison linéaire des vecteurs d'une base.

Corollaire 6.2. *Soit K un corps fini de cardinal p. Alors, K est isomorphe à $\mathbb{F}_p$.*

Démonstration : C'est immédiat vu que K contient un sous-corps isomorphe à $\mathbb{F}_p$.

Corollaire 6.3. *Soient K un corps fini, de caractéristique p, et F un polynôme irréductible de degré n dans $K[X]$. Alors, $K[X]/(F)$ est un corps fini, de caractéristique p, et son cardinal est $|K|^n$.*

Démonstration : Le corps $K[X]/(F)$ contenant un sous-corps isomorphe à K (remarque 5.6), sa caractéristique est la même que celle de K i.e. est p. Par ailleurs, le K-espace vectoriel $K[X]/(F)$ est de dimension n (th. 5.9), donc est isomorphe à K^n, d'où le résultat.

2. Groupe multiplicatif d'un corps fini

Démontrons dans ce paragraphe le résultat 3 annoncé.

Théorème 6.3. *Soient K un corps commutatif et H un sous-groupe fini de K^*. Alors, H est un groupe cyclique.*

En particulier :

Corollaire 6.4. *Si K est un corps fini, le groupe multiplicatif K^* est cyclique.*

Démonstration du théorème 6.3 : On utilise les deux lemmes ci-dessous.

Lemme 6.2. *Soient G un groupe abélien multiplicatif, et x, y deux éléments de G d'ordre m et n premiers entre eux. Alors, xy est d'ordre mn.*

Démonstration : On a déjà démontré ce résultat dans un cadre plus général au bas de page numéro 13. Démontrons ici ce lemme directement. Puisque G est abélien, on a $(xy)^{mn} = e$, où e est l'élément neutre de G. L'ordre de xy divise donc mn. Par ailleurs, il existe u et v dans $\mathbb{Z}$ tels que l'on ait $mu + nv = 1$ (Bézout). On a

$$(xy)^{um} = y^{um} = y^{1-vn} = y \quad \text{et} \quad (xy)^{vn} = x^{vn} = x^{1-um} = x.$$

Considérons alors un entier $r \geq 1$ tel que $(xy)^r = e$. On a $(xy)^{rum} = y^r = e$, et de même $x^r = e$. Il en résulte que r est un multiple de m et n, donc aussi de mn vu que l'on a $\text{pgcd}(m, n) = 1$, d'où le résultat.

Lemme 6.3. *Soient G un groupe abélien fini et x, y deux éléments de G. Il existe dans G un élément dont l'ordre est le ppcm des ordres de x et y.*

Démonstration : Notons multiplicativement la loi de composition de G. Soient α l'ordre de x et β celui de y. Soit R l'ensemble des diviseurs premiers de α pour lesquels on a $v_p(\alpha) > v_p(\beta)$, et S l'ensemble des diviseurs premiers de β pour lesquels on a $v_p(\beta) \geq v_p(\alpha)$. Posons

$$a = \prod_{p \in R} p^{v_p(\alpha)} \quad \text{et} \quad b = \prod_{p \in S} p^{v_p(\beta)}.$$

Il existe deux entiers r et s tels que l'on ait

$$\alpha = ar \quad \text{et} \quad \beta = bs.$$

Posons alors

$$z = x^r y^s \in G.$$

L'élément x^r est d'ordre a et y^s est d'ordre b (prop. 2.8). Par ailleurs, a et b sont premiers entre eux. Puisque G est abélien, z est donc d'ordre ab (lemme 6.2), qui n'est autre que le ppcm de α et β.

Le théorème 6.3 se déduit comme suit : soit m l'ordre de H. Puisque H est fini, il existe un élément de H d'ordre maximum n. Le corps K étant commutatif, H est en particulier abélien, donc pour tout élément de H d'ordre d, il existe un élément de H d'ordre le ppcm de d et n (lemme 6.3). Par suite, on a $\text{ppcm}(d, n) = n$, donc d divise n. Ainsi, les ordres de tous les éléments de H divisent n. Le polynôme $X^n - 1 \in K[X]$ ayant au plus n racines dans K, on en déduit que l'on a $m \leq n$. Puisque n divise m, on a donc $m = n$. Il existe ainsi un élément d'ordre m dans H, ce qui établit le théorème[33].

---

[33] On peut aussi utiliser le résultat suivant : soit G un groupe multiplicatif fini d'ordre n, d'élément neutre e. On suppose que pour tout diviseur d de n, l'ensemble des éléments $y \in G$ tels que $y^d = e$, est de cardinal au plus d. Alors, G est cyclique d'ordre n.

En effet, Soit d un diviseur de n. Vérifions que l'ensemble des éléments de G d'ordre d est vide ou bien que son cardinal est $\varphi(d)$, où φ est la fonction indicatrice d'Euler. Supposons qu'il existe $x \in G$ d'ordre d. Le sous-groupe $< x >$ de G engendré par x est cyclique d'ordre d. Soit T l'ensemble des éléments $y \in G$ tels que $y^d = e$. Le groupe $< x >$ est contenu dans T, et d'après l'hypothèse faite sur G, on a donc $T =< x >$. Il en résulte que l'ensemble des éléments d'ordre d de G est formé des générateurs de $< x >$, et il y en a $\varphi(d)$ (th. 2.5). D'où l'assertion. Pour tout diviseur d de n, notons alors Φ_d l'ensemble des éléments d'ordre d de G. Le groupe G étant la réunion disjointe de Φ_d, on a donc

$$n = \sum_{d|n} \left|\Phi_d\right| \leq \sum_{d|n} \varphi(d).$$

S'il existait un diviseur d de n tel que $\left|\Phi_d\right| = 0$, on aurait ainsi

$$n < \sum_{d|n} \varphi(d),$$

et une contradiction (lemme 2.5). En particulier, Φ_n n'est pas vide, autrement dit, il existe dans G un élément d'ordre n i.e. G est cyclique d'ordre n.

Le théorème 6.3. se déduit alors du fait que pour tout diviseur d de l'ordre de H, le polynôme $X^d - 1 \in K[X]$ a au plus d racines dans K, donc en particulier dans H.

Corollaire 6.5. *Soit K un corps fini de cardinal q. Le groupe K^* possède exactement $\varphi(q-1)$ générateurs, où φ est la fonction indicatrice d'Euler. De plus, si α est un générateur de K^*, alors l'ensemble des générateurs de K^* est*

$$\Big\{\alpha^k \mid 1 \leq k \leq q-1 \text{ et } \operatorname{pgcd}(k, q-1) = 1\Big\}.$$

Démonstration : C'est une conséquence directe du théorème 2.5 et du corollaire 6.4.

Exercice 1. On considère le polynôme $P = X^4 + X + 1 \in \mathbb{F}_2[X]$.

1) Montrer que $K = \mathbb{F}_2[X]/(P)$ est un corps.
2) Quelle est la caractéristique de K, le cardinal de K ?
3) Soit α la classe de X modulo (P). Montrer que α est un générateur de K^*. Combien il y a-t-il de générateurs dans K^* ? Déterminer leurs coordonnées dans la base $(1, \alpha, \alpha^2, \alpha^3)$ de K sur $\mathbb{F}_2$.

3. Corps finis comme quotients de $\mathbb{F}_p[X]$

Voici le quatrième résultat annoncé :

Théorème 6.4. *Soit K un corps fini de cardinal p^n. Il existe un polynôme $F \in \mathbb{F}_p[X]$ irréductible de degré n tel que les corps K et $\mathbb{F}_p[X]/(F)$ soient isomorphes.*

Démonstration : Soit α un générateur de K^*. On considère l'application

$$\psi : \mathbb{F}_p[X] \to K$$

définie pour tout $P = \sum a_i X^i \in \mathbb{F}_p[X]$ par l'égalité

$$\psi(P) = \sum a_i \alpha^i,$$

où l'on identifie ici $a_i \in \mathbb{F}_p$ avec n'importe quel entier relatif dont la classe modulo p est a_i. Cela est licite car K est de caractéristique p. C'est un homomorphisme d'anneaux. Il est surjectif vu que α est un générateur de K^*. Le noyau de ψ est un idéal I de $\mathbb{F}_p[X]$ et $\mathbb{F}_p[X]/I$ est donc un anneau isomorphe à K. L'idéal I n'est pas nul, sinon K serait isomorphe à $\mathbb{F}_p[X]$, or $\mathbb{F}_p[X]$ n'est pas un corps. Il existe donc un polynôme $F \in \mathbb{F}_p[X]$ tel que $I = (F)$ (th. 5.2). Puisque $\mathbb{F}_p[X]/(F)$ est un corps, F est donc irréductible (cor. 5.4). Par ailleurs, si m est le degré de F, le cardinal de $\mathbb{F}_p[X]/(F)$ est p^m (cor. 6.3), d'où $m = n$ et le résultat.

Il en résulte que les corps finis de cardinal p^n s'obtiennent exclusivement à partir de polynômes irréductibles de degré n dans $\mathbb{F}_p[X]$. Par conséquent, compte tenu du corollaire 6.3 et du théorème 6.4, on obtient l'énoncé suivant :

Proposition 6.1. *Soient p un nombre premier et n un entier ≥ 1. Les deux assertions suivantes sont équivalentes :*

1) il existe un corps à p^n éléments.

2) Il existe un polynôme irréductible de degré n dans $\mathbb{F}_p[X]$.

Il s'agit donc maintenant de démontrer l'existence de polynômes irréductibles de tout degré $n \geq 1$ dans $\mathbb{F}_p[X]$. Il s'agira aussi de démontrer que si U et V sont deux polynômes irréductibles de degré n dans $\mathbb{F}_p[X]$, alors les corps $\mathbb{F}_p[X]/(U)$ et $\mathbb{F}_p[X]/(V)$ sont isomorphes (unicité à isomorphisme près des corps à p^n éléments).

Exercice 2. Démontrer l'existence de corps à 27, puis à 125 éléments.

4. Construction et unicité des corps à p^2 éléments

Soit p un nombre pemier. On va démontrer directement l'énoncé suivant, qui est un cas particulier de celui que l'on a en vue.

Théorème 6.5. *Il existe, à isomorphisme près, un unique corps de cardinal p^2.*

Démonstration : Supposons $p = 2$. Le polynôme $X^2+X+1 \in \mathbb{F}_2[X]$ étant irréductible sur $\mathbb{F}_2$, l'anneau

$$\mathbb{F}_4 = \mathbb{F}_2[X]/(X^2 + X + 1),$$

est donc un corps à quatre éléments. Puisqu'il n'existe qu'un seul polynôme irréductible de degré 2 de $\mathbb{F}_2[X]$, le corps $\mathbb{F}_4$ est donc le seul corps à isomorphisme près de cardinal 4.

Supposons désormais p impair. On a vu en exercice qu'il existe $p(p-1)/2$ polynômes irréductibles unitaires de degré 2 dans $\mathbb{F}_p[X]$[34]. Il existe donc des corps à p^2 éléments.

Vérifions l'unicité annoncée. Considérons pour cela deux corps de cardinal p^2. Il s'agit de montrer qu'ils sont isomorphes, autrement dit, que si U et V sont deux polynômes irréductibles unitaires de degré 2 dans $\mathbb{F}_p[X]$, les corps

$$K = \mathbb{F}_p[X]/(U) \quad \text{et} \quad K' = \mathbb{F}_p[X]/(V),$$

sont isomorphes (th. 6.4). Posons $U = X^2 + bX + c \in \mathbb{F}_p[X]$. Puisque p est impair, on a l'égalité

$$U = \Big(X + \frac{b}{2}\Big)^2 - \frac{b^2 - 4c}{4},$$

ce qui permet de se ramener au cas où U et V sont de la forme

$$U = X^2 - d \quad \text{et} \quad V = X^2 - d',$$

[34] On procède comme suit. Tout d'abord, il y a p^2 polynômes unitaires de degré 2 dans $\mathbb{F}_p[X]$. Ceux qui sont réductibles sur $\mathbb{F}_p$ sont de la $(X-a)(X-b)$ avec a et b dans $\mathbb{F}_p$. Il y en a p pour lesquels $a = b$ et $p(p-1)/2$ pour lesquels $a \neq b$. On obtient ainsi $p^2 - p - p(p-1)/2 = p(p-1)/2$ polynômes irréductibles unitaires de degré 2 dans $\mathbb{F}_p[X]$.

avec d et d' deux éléments de $\mathbb{F}_p$ qui ne sont pas des carrés dans $\mathbb{F}_p$. Posons

$$\alpha = X + (U) \in K \quad \text{et} \quad \alpha' = X + (V) \in K'.$$

On va démontrer que

$$d' \in K^2,$$

i.e. qu'il existe $\zeta \in K$ tel que $\zeta^2 = d'$. Cette assertion entraîne le résultat. En effet, une fois cette condition démontrée, on vérifie alors que l'application $\psi : K' \to K$ définie par

$$\psi(a + b\alpha') = a + b\zeta,$$

est un isomorphisme de corps, de K' sur K. (C'est un homomorphisme de corps, il est donc injectif, puis surjectif car K et K' ont le même cardinal). On cherche ainsi x et y dans $\mathbb{F}_p$ tels que l'on ait

$$d' = (x + y\alpha)^2.$$

Compte tenu du fait que $(1, \alpha)$ est une base de K sur $\mathbb{F}_p$, on obtient les relations

$$d' = x^2 + dy^2 \quad \text{et} \quad 2xy = 0.$$

On a $2 \neq 0$ car p est impair, et nécessairement $y \neq 0$, car d' n'est pas un carré dans $\mathbb{F}_p$. Par suite, x doit être nul, et tout revient donc à montrer qu'il existe $y \in \mathbb{F}_p$ tel que

$$y^2 = \frac{d'}{d},$$

autrement dit, à montrer que le produit dd' est un carré dans $\mathbb{F}_p$. Cela résulte du lemme suivant :

Lemme 6.4. *Le produit de deux éléments de $\mathbb{F}_p^*$, qui ne sont pas des carrés dans $\mathbb{F}_p^*$, est un carré dans $\mathbb{F}_p^*$.*

Démonstration : Posons $G = \mathbb{F}_p^*$ et considérons l'application $f : G \to G$ qui à x associe x^2. C'est un homomorphisme de groupes dont l'image est le sous-groupe G^+ des carrés de $\mathbb{F}_p^*$. Son noyau est $\{\pm 1\}$ (le polynôme $X^2 - 1$ a deux racines dans $\mathbb{F}_p$ qui sont ± 1). On en déduit que (p est impair)

$$|G^+| = \frac{p-1}{2} \quad \text{et} \quad \left|G/G^+\right| = 2. \tag{1}$$

Le groupe G/G^+ est donc d'ordre 2, d'élément neutre G^+, et si G^- désigne l'ensemble des éléments de $\mathbb{F}_p^*$ qui ne sont pas des carrés, on a

$$G/G^+ = \left\{G^+, G^-\right\}.$$

En particulier, on a $G^-.G^- = G^+$, ce qui établit le lemme.

Cela termine la démonstration du théorème 6.5.

Remarque 6.1. Dans le cas où p est impair congru à 3 modulo 4, il est facile d'expliciter concrètement un corps à p^2 éléments. En effet, dans ce cas -1 n'est pas un carré dans $\mathbb{F}_p$ et $\mathbb{F}_p[X]/(X^2+1)$ est donc un corps à p^2 éléments. Pour le vérifier, il suffit de démontrer l'énoncé qui suit :

Lemme 6.5. *Soit p un nombre premier impair. On a l'équivalence*

$$-1 \in \mathbb{F}_p^2 \Longleftrightarrow p \equiv 1 \text{ mod. } 4.$$

Démonstration : On peut utiliser directement la première égalité de (1). Si -1 est un carré dans $\mathbb{F}_p$, on a ainsi dans $\mathbb{F}_p$ l'égalité (th. 2.3)

$$(-1)^{\frac{p-1}{2}} = 1,$$

ce qui entraîne $p \equiv 1$ mod. 4. Inversement, supposons $p \equiv 1$ mod. 4. Puisque 4 divise $p-1$ et que $\mathbb{F}_p^*$ est cyclique d'ordre $p-1$, le groupe $\mathbb{F}_p^*$ possède donc un sous-groupe H d'ordre 4 cyclique (th. 2.4). Si x est un générateur de H, on a $x^4 = 1$, d'où $x^2 = -1$ et le résultat.

Exercice 3. Soit K un corps fini de cardinal q et de caractéristique p. On note K^2 l'ensemble des éléments de K qui sont des carrés dans K i.e. l'ensemble des x^2 où x est dans K.

1) Si $p = 2$ montrer que $K^2 = K$.
2) Si p est distinct de 2, montrer que $|K^2| = (q+1)/2$.
3) En déduire que tout élément de K est la somme de deux carrés dans K.
4) Supposons $p \geq 3$. Montrer qu'un élément non nul $x \in K$ est un carré dans K si et seulement si on a $x^{\frac{q-1}{2}} = 1$.

5. Polynômes irréductibles sur un corps fini

On va démontrer dans ce paragraphe le résultat suivant :

Théorème 6.6. *Soient K un corps de cardinal q et n un entier naturel non nul. L'ensemble des diviseurs irréductibles du polynôme $X^{q^n} - X \in K[X]$ est formé des polynômes irréductibles de $K[X]$ de degré divisant n. Plus précisément, on a l'égalité*

$$X^{q^n} - X = \prod F, \tag{2}$$

où F parcourt l'ensemble des polynômes irréductibles unitaires de $K[X]$ de degré divisant n.

La démonstration repose sur plusieurs lemmes intermédiaires, qui sont par eux mêmes intéressants d'un point de vue pratique. On suppose dans ce qui suit que K est un corps fini de caractéristique p et de cardinal q (qui est donc une puissance de p).

Lemme 6.6. *Soit k un entier naturel. On a l'égalité*

$$(x+y)^{p^k} = x^{p^k} + y^{p^k} \quad \textit{quels que soient } x, y \in K.$$

Démonstration : Procédons par récurrence sur k. L'énoncé est vrai si $k = 0$. Soit alors k un entier ≥ 0 tel que l'égalité annoncée soit vérifiée. Pour tous $x, y \in K$, on a

$$(x+y)^{p^{k+1}} = \left((x+y)^{p^k}\right)^p = \left(x^{p^k} + y^{p^k}\right)^p,$$

la dernière égalité provenant de l'hypothèse de récurrence. Par ailleurs, pour tout entier $j = 1, \cdots, p-1$, le coefficient binômial C_p^j est divisible par p [35]. La formule du binôme de Newton entraîne alors l'égalité $(x+y)^{p^{k+1}} = x^{p^{k+1}} + y^{p^{k+1}}$, et le résultat[36] .

Lemme 6.7. *Soit L un corps fini contenant K.*

1) Pour tout $x \in L$, x appartient à K si et seulement si on a $x^q = x$.

2) Pour tout $F \in L[X]$, F appartient à $K[X]$ si et seulement si on a $F(X^q) = F(X)^q$.

Démonstration : 1) Soit x un élément de L. Si x est dans K^*, vu que K^* est un groupe d'ordre $q-1$, on a $x^{q-1} = 1$ (th. 2.3), d'où $x^q = x$. Par ailleurs, le polynôme $X^q - X \in L[X]$ possède au plus q racines, donc K est l'ensemble de ses racines, d'où l'assertion 1.

2) Soit $F = \sum a_k X^k$ un polynôme de $L[X]$. On a

$$F(X)^q = \left(\sum a_k X^k\right)^q = \sum a_k^q X^{kq} \text{ [37]}.$$

D'après la première assertion, F appartient à $K[X]$ si et seulement si $a_k^q = a_k$ pour tout k, ce qui entraîne le résultat.

---

[35] Rappelons l'argument. Pour tout $j = 1, \cdots, p-1$, on a en effet, $j!(p-j)!C_p^j = p!$, et puisque p ne divise pas $j!(p-j)!$, il en résulte que p divise C_p^j (lemme de Gauss).

[36] Bien qu'inutile pour la démonstration du théorème 6.6, signalons une conséquence du lemme 6.6. Soit $f : K \to K$ l'application définie pour tout $x \in K$ par $f(x) = x^p$. Alors, f est automorphisme de K. On l'appelle l'automorphisme de Frobenius de K. En effet, f est un homomorphisme de groupes (lemme 6.6). Par ailleurs, on a l'égalité $(xy)^p = x^p y^p$ pour tous $x, y \in K$ (K est commutatif) et $f(1) = 1$, donc f est un homomorphisme de corps. Son noyau est un idéal de K, donc est nul puisque ce dernier est distinct de K. Ainsi, f est une injection de K dans K, donc aussi une surjection car K est fini. C'est donc un automorphisme de K. Si $|K| = p^N$, on a $f^N = \mathrm{id}$ (on note f^N l'application itérée N fois de f et id l'identité de K). De plus, N est le plus petit entier $i \geq 1$ tel que $f^i = \mathrm{id}$, car si $i < N$, le polynôme $X^{p^i} - X$ ne peut avoir $p^N > p^i$ racines dans K. Il en résulte que les automorphismes $\mathrm{id}, f, \cdots, f^{N-1}$ sont deux à deux distincts. Ils forment un groupe cyclique d'ordre N, que l'on appelle le groupe de Galois de K (sur $\mathbb{F}_p$).

[37] Cette dernière égalité se démontre en utilisant le lemme 6.6, et en procédant par récurrence sur le nombre de monômes de F.

Lemme 6.8. *Soient F un polynôme unitaire irréductible de $K[X]$ et L un corps fini contenant K dans lequel F a une racine α. Il existe un plus petit entier $r \geq 1$ tel que l'on ait $\alpha^{q^r} = \alpha$. On a $r = \deg(F)$ et l'égalité*

$$F = \prod_{i=0}^{r-1} \Big(X - \alpha^{q^i} \Big).$$

Démonstration : Il existe un entier $m \geq 1$ tel que le cardinal de L soit q^m. On a $\alpha^{q^m} = \alpha$, donc il existe un plus petit entier $r \geq 1$ tel que l'on ait $\alpha^{q^r} = \alpha$. Par ailleurs, α étant racine de F, on déduit du lemme 6.7 que les éléments α^{q^i} pour $i = 1, \cdots, r-1$ sont aussi des racines de F. Posons

$$G = \prod_{i=0}^{r-1} \Big(X - \alpha^{q^i} \Big).$$

Puisque les α^{q^i} pour $i = 0, \cdots, r-1$ sont distincts deux à deux[38], il en résulte G divise F dans $L[X]$ (th. 5.7). De plus, on a les égalités

$$G(X)^q = \prod_{i=0}^{r-1} \Big(X - \alpha^{q^i} \Big)^q = \prod_{i=0}^{r-1} \Big(X^q - \alpha^{q^{i+1}} \Big) = G(X^q).$$

On en déduit que G appartient à $K[X]$ (lemme 6.7). Le quotient et le reste de la division euclidienne de F par G étant indépendants du corps de base, vu leur caractère d'unicité, on en déduit que G divise F dans $K[X]$. Le polynôme F étant irréductible, G étant de degré au moins 1, et F et G étant unitaires, on a donc $F = G$, d'où le résultat.

Remarque 6.2. Le lemme 6.8 montre qu'un polynôme irréductible F de $K[X]$ qui a une racine dans un surcorps fini L de K, a toutes ses racines dans L, comme annoncé dans les remarques 5.6. De plus, si r est le degré de F, et si $\alpha \in L$ est une racine de F, alors les racines de F sont les α^{q^i} pour $i = 0, \cdots, r-1$.

---

[38] On peut justifier cette assertion comme suit. Supposons qu'il existe deux entiers i et j compris entre 0 et $r-1$ tels que $i < j$ et $\alpha^{q^i} = \alpha^{q^j}$. On a alors l'égalité

$$\left(\frac{\alpha^{q^{j-i}}}{\alpha} \right)^{q^i} = 1.$$

Il s'agit d'en déduire que $\alpha^{q^{j-i}} = \alpha$, ce qui conduira à une contradiction vu le caractère minimal de r. Tout revient ainsi à démontrer que pour tout $y \in L$, l'égalité $y^q = 1$ entraîne $y = 1$. Les entiers q et $q^m - 1$ étant premiers entre eux, il existe u et v dans $\mathbb{Z}$ tels que l'on ait $uq + v(q^m - 1) = 1$. Pour tout $y \in L$, on a $y^{q^m - 1} = 1$. Par suite, si $y^q = 1$, on obtient $y = y^{uq + v(q^m - 1)} = 1$, et notre assertion.

Fin de la démonstration du théorème 6.6.

Soit $F \in K[X]$ un polynôme irréductible unitaire de degré r. Il s'agit de démontrer l'équivalence suivante :

$$F \text{ divise } X^{q^n} - X \iff r \text{ divise } n. \tag{3}$$

Considérons un corps fini L contenant K dans lequel F a une racine α : un tel corps L existe (th. 5.11). D'après le lemme 6.8, r est le plus petit entier ≥ 1 tel que $\alpha^{q^r} = \alpha$ et l'on a l'égalité

$$F = \prod_{i=0}^{r-1} \Big(X - \alpha^{q^i} \Big). \tag{4}$$

Par ailleurs, on a

$$\alpha^{q^{ir}} = \alpha \quad \text{pour tout } i \geq 0. \tag{5}$$

En effet, cette égalité est vraie si $i = 0$, et si elle est vérifiée pour un entier $i \geq 0$, on a

$$\alpha^{q^{(i+1)r}} = \big(\alpha^{q^{ir}}\big)^{q^r} = \alpha^{q^r} = \alpha,$$

donc elle l'est aussi pour $i + 1$.

Supposons alors que F divise $X^{q^n} - X$. Puisque $F(\alpha) = 0$, on a $\alpha^{q^n} = \alpha$. Il existe deux entiers naturels t et s tels que l'on ait $n = rt + s$ avec $0 \leq s < r$. On a donc

$$\alpha^{q^n} = \big(\alpha^{q^{tr}}\big)^{q^s}.$$

D'après (5), on a $\alpha^{q^{tr}} = \alpha$. Par suite, on a $\alpha = \alpha^{q^s}$, ce qui d'après le caractère minimal de r, entraîne $s = 0$, ainsi r divise n.

Inversement, supposons que r divise n. On déduit de (5) que l'on a $\alpha^{q^n} = \alpha$, autrement dit, que α est racine du polynôme $X^{q^n} - X$. Les éléments

$$\alpha, \alpha^q, \cdots, \alpha^{q^{r-1}},$$

sont donc des racines deux à deux distinctes de $X^{q^n} - X$. Il résulte de (4) que F divise $X^{q^n} - X$ dans $L[X]$, donc aussi dans $K[X]$ car F est à coefficients dans K. Cela prouve l'équivalence (3).

On déduit de ce qui précède, et du théorème 5.6, une égalité de la forme

$$X^{q^n} - X = \prod_F F^{n_F},$$

où F parcourt l'ensemble des polynômes irréductibles unitaires de $K[X]$ de degré divisant n, et où les n_F sont des entiers naturels non nuls. Tout revient alors à démontrer que les

n_F sont égaux à 1. Étant donné un tel polynôme F, on a $X^{q^n} - X = F^{n_F} Q$ où $Q \in K[X]$, d'où l'on déduit, en considérant les polynômes dérivés des deux membres de cette égalité (on a $q1_K = 0$ car K est de caractéristique p),

$$-1 = n_F F^{n_F - 1} F' Q + F^{n_F} Q'.$$

Par suite, $F^{n_F - 1}$ divise -1 dans $K[X]$, ce qui entraîne $n_F = 1$ et le résultat.

Exercice 4. Factoriser $X^8 - X \in \mathbb{F}_2[X]$ en produit de polynômes irréductibles de $\mathbb{F}_2[X]$.

6. Théorème d'existence et dénombrement

On considère dans ce paragraphe un corps fini K de cardinal q.

Notation. Pour tout entier $m \geq 1$, on note $I_m(q)$ le nombre de polynômes irréductibles unitaires de degré m de $K[X]$.

Pour tout $n \geq 1$, la formule (2) permet de calculer $I_n(q)$. En effet, dans le produit intervenant dans (2) il y a $I_d(q)$ facteurs de degré d pour chaque diviseur d de n. En considérant les degrés des polynômes de chaque membre, on obtient ainsi

$$q^n = \sum_{d|n} I_d(q) d. \tag{6}$$

Le théorème d'existence annoncé au début sur les corps finis est une conséquence de l'énoncé suivant :

Théorème 6.7. *Pour tout $n \geq 1$, on a $I_n(q) > 0$.*

Démonstration : On procède par récurrence sur n. Le résultat est vrai si $n = 1$ (pour tout $a \in K$, $X - a$ est irréductible dans $K[X]$). Considérons alors un entier $n \geq 2$, et supposons le résultat démontré pour tout entier $d < n$. En utilisant la formule (6), on obtient l'égalité

$$q^d = d I_d(q) + \sum_{d' | d,\ d' < d} I_{d'}(q) d' \quad \text{pour tout} \quad d < n.$$

D'après l'hypothèse de récurrence, on a donc

$$q^d > d I_d(q) \quad \text{pour tout} \quad d < n.$$

La formule

$$q^n = n I_n(q) + \sum_{d|n,\ d<n} d I_d(q)$$

entraîne alors les inégalités

$$q^n < nI_n(q) + \sum_{d|n,\ d<n} q^d \leq nI_n(q) + \sum_{k=0}^{n-1} q^k = nI_n(q) + \frac{q^n - 1}{q-1} < nI_n(q) + q^n,$$

d'où $I_n(q) > 0$ et le résultat.

Corollaire 6.6. *Pour tout entier $n \geq 1$ et tout nombre premier p, il existe un corps de cardinal p^n.*

Démonstration : C'est une conséquence directe du théorème 6.7, appliqué avec $q = p$, et de la proposition 6.1.

On va maintenant démontrer une formule permettant de calculer directement $I_n(q)$. Il nous faut pour cela démontrer une formule d'inversion, que l'on appelle la formule d'inversion de Möbius. Définissons d'abord ce que l'on appelle la fonction de Möbius.

Définition 6.2. *La fonction de Möbius $\mu : \mathbb{N} \to \{0, \pm 1\}$ est définie pour tout $n \in \mathbb{N}$ par les égalités*

$$\mu(n) = \begin{cases} (-1)^r & \text{si } n \text{ est le produit de } r \text{ nombres premiers distincts} \\ 0 & \text{sinon.} \end{cases}$$

Notons que l'on a par définition $\mu(1) = 1$. Par ailleurs, pour tout $n \geq 1$, on a la formule

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{sinon.} \end{cases} \tag{7}$$

En effet, supposons $n \geq 2$ et la décomposition en facteurs premiers de n de la forme

$$n = p_1^{n_1} \cdots p_r^{n_r},$$

avec des entiers $n_i \geq 1$, les p_i étant des nombres premiers distincts deux à deux. Parmi les diviseurs de n, seuls ceux qui sont sans facteurs carrés ont une contribution non nulle dans la somme des $\mu(d)$. Par suite, on a

$$\sum_{d|n} \mu(d) = C_r^0 - C_r^1 + \cdots + (-1)^r C_r^r = (1-1)^r = 0,$$

d'où la formule (7).

On va démontrer l'énoncé suivant :

Théorème 6.8. *Pour tout entier $n \geq 1$, on a l'égalité*

$$nI_n(q) = \sum_{d|n} \mu(d)\ q^{\frac{n}{d}}. \tag{8}$$

C'est une conséquence directe de la formule (6) et du résultat qui suit, connu sous le nom de formule d'inversion de Möbius. On pose $\mathbb{N}^* = \mathbb{N} - \{0\}$.

Théorème 6.9. *Soient $(G, +)$ un groupe abélien et f une fonction de $\mathbb{N}^*$ à valeurs dans G. Soit $g : \mathbb{N}^* \to G$ la fonction définie pour tout $n \in \mathbb{N}^*$ par l'égalité*

$$g(n) = \sum_{d|n} f(d).$$

Alors, pour tout $n \in \mathbb{N}^*$, on a

$$f(n) = \sum_{d|n} \mu(d) g\Big(\frac{n}{d}\Big). \tag{9}$$

Remarque 6.3. L'application $d \mapsto n/d$ permute entre eux les diviseurs de n. Par suite, la formule (9) s'écrit aussi

$$f(n) = \sum_{d|n} \mu\Big(\frac{n}{d}\Big) g(d).$$

Démonstration : Soit n un entier naturel non nul. Pour tout $r \geq 1$, posons

$$\delta(r) = \sum_{k|r} \mu(k).$$

On déduit de la formule (7) l'égalité

$$f(n) = \sum_{d|n} \delta(d) f\Big(\frac{n}{d}\Big),$$

autrement dit,

$$f(n) = \sum_{d|n} \sum_{k|d} \mu(k) f\Big(\frac{n}{d}\Big).$$

On en déduit l'égalité

$$f(n) = \sum_{k|n} \sum_{k|d|n} \mu(k) f\Big(\frac{n}{d}\Big).$$

Par ailleurs, pour k fixé divisant n, on a

$$\sum_{k|d|n} f\Big(\frac{n}{d}\Big) = \sum_{j|\frac{n}{k}} f\Big(\frac{n}{kj}\Big).$$

Il en résulte l'égalité

$$f(n) = \sum_{k|n} \mu(k) \sum_{j|\frac{n}{k}} f\Big(\frac{n}{kj}\Big).$$

La formule (9) en résulte, vu que pour tout k divisant n, on a

$$g\Big(\frac{n}{k}\Big) = \sum_{j|\frac{n}{k}} f\Big(\frac{n}{kj}\Big).$$

Remarques 6.4.

1) Si $n = 1$, on retrouve l'égalité $I_1(q) = q$ comme attendu. Pour $n = 2$, on obtient

$$I_2(q) = \frac{q(q-1)}{2},$$

formule que l'on avait déjà établie si q est un nombre premier. De même, avec $n = 3$, on constate que l'on a

$$I_3(q) = \frac{q(q^2-1)}{3}.$$

Avec $q = 2$, on obtient ainsi $I_3(2) = 2$, les deux polynômes irréductibles (unitaires) de degré 3 dans $\mathbb{F}_2[X]$ étant $X^3 + X^2 + 1$ et $X^3 + X + 1$.

2) En utilisant la formule (8), on vérifie que

$$I_{50}(2) = 22.517.997.465.744.$$

Cela fournit de nombreuses façons de construire un corps de cardinal 2^{50}, qui est par ailleurs unique à isomorphisme près, comme on le constatera dans le paragraphe suivant .

3) La formule (8) montre que son membre de droite est divisible par n, ce qui n'est pas évident a priori. Il est instructif de le démontrer directement. Vérifions en fait, que l'on a

$$\sum_{d|n} \mu(d) x^{\frac{n}{d}} \equiv 0 \text{ mod. } n \quad \text{pour tout } x \in \mathbb{Z}. \tag{10}$$

Établissons d'abord (10) dans le cas où n est une puissance d'un nombre premier. Posons $n = p^r$ avec $r \geq 1$. On a (avec $d = 1$ et $d = p$)

$$\sum_{d|n} \mu(d) x^{\frac{n}{d}} = x^{p^r} - x^{p^{r-1}}.$$

Supposons que p ne divise pas x. Dans ce cas, x est inversible modulo p^r, et le groupe $\big(\mathbb{Z}/p^r\mathbb{Z}\big)^*$ étant d'ordre $\varphi(p^r) = p^r - p^{r-1}$, on a

$$x^{p^r-p^{r-1}} \equiv 1 \text{ mod. } p^r \quad \text{i.e.} \quad x^{p^r} \equiv x^{p^{r-1}} \text{ mod. } p^r,$$

d'où la congruence (10) dans ce cas. Si p divise x, alors $x^{p^r} - x^{p^{r-1}}$ est divisible par $p^{p^{r-1}}$. Par ailleurs, on vérifie (par récurrence sur r) que l'on a $p^{r-1} \geq r$, d'où de nouveau la condition (10). Passons au cas général. Il suffit de démontrer que pour tout nombre premier p divisant n, on a

$$\sum_{d|n} \mu(d) x^{\frac{n}{d}} \equiv 0 \text{ mod. } p^{v_p(n)},$$

où $v_p(n)$ est la valuation p-adique de n. Considérons donc un nombre premier p qui divise n. Posons

$$r = v_p(n) \quad \text{et} \quad n = p^r m.$$

Dans la somme $\sum \mu(d) x^{\frac{n}{d}}$, les termes donnant une contribution éventuellement non nulle sont ceux pour lesquels d est premier à p ou bien ceux pour lesquels d est de la forme pj avec j premier avec p. Par suite, on a

$$\sum_{d|n} \mu(d) x^{\frac{n}{d}} = \sum_{j|n,(p,j)=1} \mu(j) \Big(x^{\frac{n}{j}} - x^{\frac{n}{pj}} \Big),$$

vu que $\mu(pj) = -\mu(j)$. Chaque entier j divisant n et premier avec p est un diviseur de m. Pour un tel entier j, on a donc

$$x^{\frac{n}{j}} - x^{\frac{n}{pj}} = y^{p^r} - y^{p^{r-1}} \quad \text{avec} \quad y = x^{\frac{m}{j}}.$$

Le cas particulier déjà traité entraîne alors le résultat.

4) À titre d'exemple, calculons $I_{70}(q)$. On a $70 = 2.5.7$ et le tableau suivant :

d	1	2	5	7	10	14	35	70
$\mu(d)$	1	-1	-1	-1	1	1	1	-1
$\frac{n}{d}$	70	35	14	10	7	5	2	1

On en déduit que l'on a

$$I_{70}(q) = \frac{1}{70} \Big(q^{70} - q^{35} - q^{14} - q^{10} + q^7 + q^5 + q^2 - q \Big).$$

On obtient au passage l'identité (congruence (10))

$$x^{70} - x^{35} - x^{14} - x^{10} + x^7 + x^5 + x^2 - x = 0 \quad \text{pour tout } x \in \mathbb{Z}/70\mathbb{Z}.$$

5) La fonction φ étant la fonction indicatrice d'Euler, on a vu que tout entier $n \geq 1$ est la somme des $\varphi(d)$, où d parcourt l'ensemble des diviseurs de n. La formule d'inversion de Möbius entraîne alors l'égalité

$$\frac{\varphi(n)}{n} = \sum_{d|n} \frac{\mu(d)}{d} \quad \text{pour tout } n \geq 1.$$

7. Théorème d'unicité

Il s'agit de démontrer l'énoncé suivant :

Théorème 6.10. *Deux corps finis ayant le même nombre d'éléments sont isomorphes.*

Démonstration : Soient K et L deux corps à q éléments et p leur caractéristique. On a $q = p^n$ pour un entier $n \geq 1$. Il existe un polynôme irréductible $F \in \mathbb{F}_p[X]$, de degré n, tel que K soit isomorphe à $\mathbb{F}_p[X]/(F)$ (th. 6.4). Le polynôme F divise $X^q - X \in \mathbb{F}_p[X]$ (th. 6.6). Par ailleurs, pour tout $x \in L$, on a $x^q = x$, autrement dit, on a

$$X^q - X = \prod_{a \in L} (X - a).$$

On en déduit que F possède une racine $a \in L$. Considérons l'application $\psi : \mathbb{F}_p[X] \to L$ définie par $\psi(P) = P(a)$ pour tout $P \in \mathbb{F}_p[X]$. C'est un homomorphisme d'anneaux. Vu que F est irréductible dans $\mathbb{F}_p[X]$ et que $F(a) = 0$, le noyau de ψ est l'idéal (F). Il en résulte que $\mathbb{F}_p[X]/(F)$ est isomorphe à l'image de ψ, qui n'est autre que L, car L et $\mathbb{F}_p[X]/(F)$ ont le même cardinal. Cela entraîne que les corps K et L sont isomorphes.

Ce résultat justifie l'abus courant consistant à parler « du » corps à q éléments. On le note alors souvent $\mathbb{F}_q$, y compris si q n'est pas premier (mais une puissance d'un nombre premier). Il faut prendre garde ici qu'il ne s'agit pas de l'anneau $\mathbb{Z}/q\mathbb{Z}$, qui n'est pas un corps si q n'est pas premier ! On a par exemple

$$\mathbb{F}_8 = \mathbb{F}_2[X]/(X^3 + X + 1), \quad \mathbb{F}_{81} = \mathbb{F}_3[X]/(X^4 + X^3 + 2), \cdots$$

Exercice 5. Déterminer $\mathbb{F}_{25}$ et $\mathbb{F}_{49}$.

8. Le problème du logarithme discret - Algorithme de Silver, Pohlig et Hellman

Soit K un corps à q éléments. On sait que le groupe multiplicatif K^* est cyclique (cor. 6.4). Soit g un de ses générateurs (il y en a $\varphi(q-1)$ (cor. 6.5)). Tous les éléments de K^* sont des puissances de g. Plus précisément, on a

$$K^* = \Big\{ g^i \mid 0 \leq i \leq q-2 \Big\}.$$

Le problème du logarithme discret de base g dans K^* est le suivant :

Problème. Étant donné un élément $x \in K^*$, trouver l'entier i tel que l'on ait

$$x = g^i \quad \text{et} \quad 0 \leq i \leq q-2.$$

On note parfois cet entier i, $\log_g(x)$ ou bien $\mathrm{ind}_g(x)$.

Certains algorithmes de cryptographie sont basés sur le fait que, pour certains corps finis de grand cardinal q, ce problème soit difficile à résoudre. Tel est par exemple le cas si l'on ne sait pas factoriser $q-1$, notamment si $q-1$ possède un grand diviseur premier. L'algorithme de Silver, Pohlig et Hellman, que l'on décrit plus bas, permet de résoudre ce problème si l'on connaît la décomposition de $q-1$ en facteurs premiers.

Abordons ce problème dans un cas simple. Prenons le corps, de cardinal 27,

$$K = \mathbb{F}_3[X]/(X^3+2X+1).$$

(Le polynôme X^3+2X+1 est irréductible dans $\mathbb{F}_3[X]$ vu qu'il est de degré 3 et qu'il n'a pas de racines dans $\mathbb{F}_3$). Le groupe K^* est d'ordre 26. Les ordres de ses éléments autres que l'élément neutre, sont donc 2, 13 ou 26. En fait, -1 est le seul élément d'ordre 2 de K^*, car par exemple ± 1 sont les seules racines du polynôme $X^2-1 \in K[X]$. Notons α la classe de X modulo X^3+2X+1. Vérifions que α est un générateur de K^*. On a en effet, $\alpha^3 = \alpha - 1$, d'où $\alpha^9 = \alpha^3 - 1$ (car K est de caractéristique 3) i.e. $\alpha^9 = \alpha + 1$, d'où $\alpha^{12} = \alpha^2 - 1$ puis $\alpha^{13} = -1$ et notre assertion.

Résolvons alors le problème du logarithme discret de base α dans K^*. Tout élément de K^* s'écrit de manière unique sous la forme $a + b\alpha + c\alpha^2$ avec $a, b, c \in \mathbb{F}_3$. Il s'agit donc pour chacun de ces éléments de déterminer l'entier i tel que l'on ait

$$a + b\alpha + c\alpha^2 = \alpha^i \quad \text{avec} \quad 0 \leq i \leq 25.$$

On vérifie les calculs suivants :

x	1	2	α	$\alpha+1$	$\alpha+2$	2α	$2\alpha+1$	$2\alpha+2$	α^2	α^2+1	α^2+2	$2\alpha^2$	$2\alpha^2+1$
$\log_\alpha(x)$	0	13	1	9	3	14	16	22	2	21	12	15	25

x	$2\alpha^2+2$	$\alpha^2+\alpha+1$	$\alpha^2+2\alpha+1$	$\alpha^2+2\alpha+2$	$\alpha^2+\alpha+2$	$2\alpha^2+2\alpha+1$
$\log_\alpha(x)$	8	6	18	7	11	24

x	$2\alpha^2+2\alpha+2$	$2\alpha^2+\alpha+2$	$2\alpha^2+\alpha+1$	$\alpha^2+\alpha$	$\alpha^2+2\alpha$	$2\alpha^2+2\alpha$	$2\alpha^2+\alpha$
$\log_\alpha(x)$	19	5	20	10	4	23	17

Compte tenu du fait que les générateurs de K^* sont les α^k avec $1 \leq k \leq 26$ et k premier avec 26, on connaît ainsi explicitement leurs coordonnées dans la base $(1, \alpha, \alpha^2)$ de K sur $\mathbb{F}_3$. Par exemple, $\alpha^{21} = 1 + \alpha^2$ est un générateur de K^* :

Exercice 6. Résoudre le problème du logarithme discret de base $1 + \alpha^2$ dans K^*.

Algorithme de Silver, Pohlig et Hellman

Soit K un corps fini à q éléments. Soit g un générateur de K^*. On va décrire ici un algorithme permettant de résoudre le problème du logarithme discret de base g dans K^*, dans le cas où l'on connaît la factorisation de $q-1$ en facteurs premiers. Cet algorithme est d'autant plus efficace que les diviseurs premiers de $q-1$ sont petits.

Partons d'un élément $x \in K^*$. Il s'agit de déterminer l'entier n tel que l'on ait

$$x = g^n \quad \text{et} \quad 0 \leq n \leq q-2.$$

On suppose connue la décomposition de $q-1$ en facteurs premiers. Soit

$$q - 1 = \prod_{p | q-1} p^{r_p}$$

cette décomposition. Afin de calculer n, l'idée de base est qu'il suffit de connaître n modulo p^{r_p} pour chaque nombre premier p qui divise $q-1$, le théorème chinois permettant ensuite de retrouver n.

L'algorithme est le suivant. Soit p un diviseur premier de $q-1$. Puisque K^* est cyclique d'ordre $q-1$, il existe un unique sous-groupe μ_p de K^* d'ordre p, qui n'est autre que l'ensemble des racines p-ièmes de l'unité de K^* (th. 2.4). Un générateur ζ de μ_p est

$$\zeta = g^{\frac{q-1}{p}}, \tag{11}$$

vu que cet élément est d'ordre p dans K^* (prop. 2.8). On a donc

$$\mu_p = \left\{1, \zeta, \zeta^2, \cdots, \zeta^{p-1}\right\}.$$

Par ailleurs, il existe des entiers n_i tels que l'on ait

$$n \equiv n_0 + n_1 p + \cdots + n_{r_p - 1} p^{r_p - 1} \text{ mod. } p^{r_p} \quad \text{avec} \quad 0 \leq n_i < p. \tag{12}$$

Conformément à la stratégie annoncée, on est confontré au problème de la détermination des n_i. On calcule n_0 à partir de l'élément $x^{\frac{q-1}{p}}$. En effet, on a

$$\left(x^{\frac{q-1}{p}}\right)^p = 1,$$

de sorte $x^{\frac{q-1}{p}}$ appartient à μ_p. En écrivant que l'on a $x = g^n$, et en utilisant le fait que $g^{q-1} = 1$, on obtient alors les égalités

$$x^{\frac{q-1}{p}} = g^{n\frac{q-1}{p}} = g^{n_0\frac{q-1}{p}} = \zeta^{n_0}, \tag{13}$$

ce qui permet d'obtenir n_0. On procède de même pour les autres coefficients n_i. Posons

$$x_i = \frac{x}{g^{n_0+\cdots+n_{i-1}p^{i-1}}} \quad \text{pour tout} \quad i \geq 1 \quad \text{et} \quad i \leq r_p - 1. \tag{14}$$

En écrivant de nouveau que $x = g^n$, on a alors

$$x_i^{\frac{q-1}{p^{i+1}}} = g^{n_i\frac{q-1}{p}} = \zeta^{n_i}, \tag{15}$$

et l'on détermine ainsi n_i, d'où la connaissance de n modulo p^{r_p}. En effectuant ces calculs pour chaque diviseur premier de $q-1$, et en utilisant le fait que les anneaux $\mathbb{Z}/(q-1)\mathbb{Z}$ et le produit des $\mathbb{Z}/p^{r_p}\mathbb{Z}$ sont isomorphes pour p divisant $q-1$ (théorème chinois), on peut ainsi obtenir n modulo $q-1$, d'où l'entier n cherché.

Exemples 6.1.

1) Prenons le corps $K = \mathbb{F}_{53}$, de cardinal $q = 53$. On a $q - 1 = 4 \times 13$. Un générateur de K^* est $g = 2$ (plus exactement la classe de 2 modulo 53). On vérifie cette assertion en remarquant que l'on a

$$2^{13} = 8192 \equiv 30 \text{ mod. } 53 \quad \text{d'où} \quad 2^{26} \equiv 900 \equiv -1 \text{ mod. } 53,$$

de sorte que l'ordre de 2 dans K^* est 52. On prend $x = 23$. Déterminons le logarithme discret $\log_2(23)$ de base 2 de 23 dans K^*. On cherche donc l'entier n tel que

$$23 \equiv 2^n \text{ mod. } 53 \quad \text{et} \quad 0 \leq n \leq 51.$$

Reprenons les notations utilisées ci-dessus. On détermine d'abord n modulo 4. On a l'égalité $\mu_2 =< -1 >$. Par ailleurs, il existe des entiers n_0 et n_1 égaux à 0 ou 1, tels que l'on ait (formule (12))

$$n \equiv n_0 + 2n_1 \text{ mod. } 4.$$

On a dans K^* (formule (13))

$$23^{26} = (-1)^{n_0}.$$

On vérifie que l'on a $23^2 = 529 \equiv -1$ mod. 53 d'où $23^{26} \equiv -1$ mod. 53, ce qui entraîne $n_0 = 1$. L'élément $x_1 \in K^*$, défini par la formule (14), est ici

$$x_1 = \frac{23}{2} = 38.$$

On a alors (formule (15))

$$38^{13} \equiv (-1)^{n_1} \text{ mod. } 53.$$

On a $38^2 \equiv 13$ mod. 53, d'où $38^{12} \equiv 46$ mod. 53, puis $38^{13} \equiv -1$ mod. 53, d'où $n_1 = 1$. Finalement, on obtient

$$n \equiv 3 \text{ mod. } 4. \tag{16}$$

Déterminons maintenant la congruence de n modulo 13. On a $\mu_{13} =< 2^4 >$ (formule (11)), d'où l'on déduit que

$$\mu_{13} = \Big\{1, 16, 44, 15, 28, 24, 13, 49, 42, 36, 46, 47, 10\Big\},$$

où les éléments sont rangés par ordre croissant des puissances du générateur g. On cherche l'entier n_0 compris entre 0 et 12 tel que l'on ait $n \equiv n_0$ mod. 13. D'après la formule (13), on a

$$23^4 \equiv 16^{n_0} \text{ mod. } 53.$$

Puisque l'on a $23^4 \equiv 1$ mod. 53, on en déduit que $n_0 = 0$, i.e. que l'on a

$$n \equiv 0 \text{ mod. } 13. \tag{17}$$

L'entier n cherché est donc l'unique entier compris entre 0 et 51 tel que les congruences (16) et (17) soient satisfaites. Conformément à la démonstration du théorème chinois, on doit écrire une relation de Bézout entre 4 et 13, ce qui est immédiat vu l'égalité $(-3)\times 4 + 13 = 1$. On en déduit que $n = 39$. Autrement dit, on a dans K^* l'égalité $23 = 2^{39}$, d'où

$$\log_2(23) = 39.$$

2) Prenons le corps $K = \mathbb{F}_3[X]/(X^3 + 2X + 1)$ étudié précédemment et $x = \alpha^2 + 1$, où α est la classe de X modulo $(X^3 + 2X + 1)$. L'élément $g = \alpha$ est un générateur de K^*. (Re)déterminons l'entier

$$n = \log_\alpha(\alpha^2 + 1).$$

On vérifie que l'on a $x^{13} = 2 = -1$, d'où il résulte que n est impair. Par ailleurs, on a $\mu_{13} =< \alpha^2 >$ et l'on vérifie les égalités $x^2 = 2\alpha + 1 = (\alpha^2)^8$. On a donc $n \equiv 8$ mod. 13, ce qui conduit à $n = 21$ comme attendu.

Remarque 6.5. Soit K un corps de cardinal q tel que q ne soit pas une puissance de 2. Dans ce cas, 2 divise $q - 1$. Étant donné un élément $x \in K^*$, dans la formule (13) pour $p = 2$, l'entier n_0 vaut 1 si et seulement si x est un carré dans K (exercice 3 de ce chapitre).

9. Application à la cryptographie - Protocole de Diffie-Hellman - Algorithme de El Gamal

Donnons ici deux applications de la théorie des corps finis à la cryptographie, le protocole d'échange de clés de Diffie-Helman et l'algorithme de chiffrement à clé publique de El Gamal.

1. Protocole de Diffie-Hellman

Deux personnes, Alice et Bob, souhaitent se construire une clé secrète commune, qu'ils seront en principe les seuls à connaître, afin de communiquer sur un canal non sûr en utilisant cette clé pour chiffrer leur correspondance. Leur procédé de fabrication est basé sur le fait que le problème du logarithme discret soit difficile à résoudre dans certains corps finis bien choisis. Ce principe, qui date de 1976, est le suivant. Soient K un corps fini de cardinal q, dans lequel le problème du logarithme discret soit a priori difficile à résoudre, et g un générateur de K^*. Le couple (K, g) est public.

1) Alice choisit secrètement et aléatoirement un entier a tel que $1 < a < q-1$, et elle transmet à Bob publiquement l'élément g^a.
2) Bob choisit aussi secrètement et aléatoirement un entier b tel que $1 < b < q-1$, et il transmet à Alice publiquement l'élément g^b.
3) Alice élève g^b à la puissance a, et elle obtient ainsi l'élément g^{ab}.
4) Bob élève g^a à la puissance b, obtenant de même g^{ab}.

Leur clé secrète commune est alors g^{ab}. Ils sont les seuls à la connaître, car quiconque disposant du couple (K, g), ainsi que des éléments g^a et g^b, ne peut pas en déduire g^{ab}, sauf à déterminer a ou b i.e. $\log_g(g^a)$ ou $\log_g(g^b)$.

Exercice 7. Alice et Bob souhaitent se construire une clé secrète commune, suivant le protocole de Diffie-Hellman, au moyen du corps $K = \mathbb{F}_3[X]/(X^3+2X+1)$ et de l'élément $\alpha = X + (X^3+2X+1)$ qui, on l'a vu, est un générateur de K^*. Pour cela, Alice choisit l'entier $a = 9$ et transmet à Bob α^9. Ce dernier choisit un entier b compris entre 1 et 25 et lui renvoie l'élément $\alpha^b = 2 + \alpha + 2\alpha^2$. Quelle est la clé secrète d'Alice et Bob ?

2. Algorithme de chiffrement à clé publique de El Gamal

Cet algorithme concerne le problème de la confidentialité des messages envoyés, et son efficacité est aussi basée sur la difficulté du problème du logarithme discret. Une personne, Alice, souhaite permettre à quiconque de lui envoyer des messages confidentiels. Pour cela, elle choisit au départ un couple public (K, g), formé d'un corps fini K et d'un générateur g de K^*. Soit q le cardinal de K. Le procédé est alors le suivant :

1) Alice choisit aléatoirement un entier a tel que $1 < a < q-1$, qui sera sa clé secrète. Elle calcule g^a qu'elle publie, et qui sera sa clé publique.

La clé publique de l'algorithme est donc au départ le triplet (K, g, g^a).

2) Afin d'envoyer un message $m \in K$ à Alice, Bob choisit aléatoirement un entier x tel que $1 < x < q-1$, et transmet à Alice le couple

$$(g^x, mg^{ax}).$$

C'est la phase d'encryptage du message m.

3) Afin de décrypter le message reçu, il s'agit donc de la phase de décryptage, Alice, connaissant a et g^x, détermine alors l'inverse dans K de l'élément $(g^x)^a$ i.e. g^{-ax} [39]. Elle effectue ensuite la multiplication de g^{-ax} par mg^{ax}, ce qui, vu l'égalité

$$g^{-ax}\big(mg^{ax}\big) = m,$$

lui permet de retrouver m.

Exercice 8. Alice souhaite se faire envoyer des messages confidentiellement en utilisant l'algorithme de El Gamal. Pour cela elle utilise le corps $K = \mathbb{F}_2[X]/(X^4+X+1)$ étudié dans l'exercice 1 et l'élément $\alpha = X + (X^4 + X + 1)$, qui rappelons-le est un générateur de K^*. Elle choisit par ailleurs un entier a compris entre 1 et 14 pour lequel $\alpha^a = 1 + \alpha^2$.

1) Bob veut envoyer le message $1 + \alpha$ à Alice. Que lui transmet-il ?

2) Vous interceptez le message $(\alpha^3, \alpha+\alpha^2+\alpha^3)$. Quel était dans ce cas le message envoyé par Bob ?

3. Algorithme de signature à clé publique de El Gamal

Cet algorithme concerne le problème de l'authentification de l'expéditeur d'un message, et non pas le problème de la confidentialité du message. Il s'agit pour l'expéditeur de «signer» son message afin que son destinataire puisse l'authentifier. On se place ici dans la situation où le corps de base est de cardinal premier. On dispose au départ d'un nombre premier p pour lequel le problème du logarithme discret soit difficile à résoudre dans $\mathbb{F}_p$ et d'un générateur g de $\mathbb{F}_p^*$, le couple $(\mathbb{F}_p, g)$ étant public. Supposons qu'Alice souhaite envoyer à Bob un message signé $m \in \mathbb{F}_p^*$. Le procédé est le suivant :

1) elle choisit un entier a tel que $1 < a < p-1$ et publie l'élément $g^a \in \mathbb{F}_p^*$. L'entier a est sa clé secrète et g^a est sa clé publique.

2) Elle choisit ensuite un entier e tel que $1 < e < p-1$, premier avec $p-1$, et calcule son inverse d modulo $p-1$: on a $ed \equiv 1$ mod. $p-1$.

Notons $\widetilde{m}$ l'entier compris entre 1 et $p-1$ tel que $m = \widetilde{m} + p\mathbb{Z}$.

[39] Notons que pour tout $y \in K^*$, on a $y^{q-1} = 1$, donc l'inverse de y est $y^{-1} = y^{q-2}$. On a ainsi $g^{-ax} = (g^x)^{a(q-2)}$, de sorte qu'Alice peut trouver l'inverse de g^{ax} en élevant directement g^x à la puissance $a(q-2)$.

3) Alice calcule les entiers r et s définis par les conditions :

$$g^e = r \text{ mod. } p \quad \text{avec} \quad 1 \leq r < p \quad \text{et} \quad s \equiv d(\widetilde{m} - ar) \text{ mod. } p-1 \quad \text{avec} \quad 1 \leq s \leq p-1.$$

On a ainsi la congruence

$$\widetilde{m} \equiv es + ar \text{ mod. } p-1. \tag{18}$$

4) Alice envoie alors à Bob le message signé $\big(m, (r,s)\big)$.

Bob peut alors « a priori » authentifier Alice, en calculant l'élément $\big(g^a\big)^r r^s \in \mathbb{F}_p^*$, puis en vérifiant que l'on a l'égalité attendue (cf. (18)) :

$$g^{\widetilde{m}} = r^s \big(g^a\big)^r \in \mathbb{F}_p^*. \tag{19}$$

On pourra trouver ci-dessous une justification heuristique de cet algorithme dans un cas particulier.

Exercice 9. On illustre l'algorithme précédent avec le nombre premier $p = 31$.

1) Montrer que (la classe de) 3 est un générateur de $\mathbb{F}_{31}^*$.
2) Alice souhaite envoyer le message $m = 23 \in \mathbb{F}_{31}^*$ en utilisant le couple $(\mathbb{F}_{31}, 3)$ de façon à ce que Bob puisse l'identifier. Pour cela, Alice publie l'élément $20 \in \mathbb{F}_{31}^*$. Bob reçoit le message $\big(23, (17,1)\big)$. A priori, Alice est-elle l'expéditrice de ce message[40] ?
3) Déterminer la clé secrète a d'Alice. En déduire les entiers e et d.

[40] En fait, la réponse à cette question est oui. Cela étant, on ne peut pas être certain que ce soit vraiment Alice l'expéditrice du message. En effet, on vérifie qu'il y a exactement dix-huit couples (r,s) vérifiant (19) : $(17,1), (3,29), (6,5), (6,11), (11,5), \cdots$. Par ailleurs, il y a neuf cents couples (r,s) possibles dans l'envoi du message $\big(m, (r,s)\big)$. Il en résulte que la « probabilité » pour ce soit Alice l'auteur du message est $1 - \frac{18}{900} = \frac{49}{50}$, ce qui est néanmoins rassurant.

Chapitre VII — Codes correcteurs d'erreurs

1. Problématique des codes correcteurs

La science du codage peut se situer dans plusieurs contextes. Par exemple, les codes secrets concernent plus particulièrement le domaine de la cryptographie. On a vu à ce sujet les algorithmes RSA et de El Gamal ainsi que le protocole de Diffie-Hellman. On peut aussi coder des informations de façon à optimiser leur stockage par exemple par compression. Le point de vue que l'on considère ici est celui des codes correcteurs d'erreurs, que l'on appelle aussi codes correcteurs. Leur problématique de base est la suivante :

Problème. *Une information est transmise via un canal bruité, et elle parvient au récepteur avec éventuellement un certain nombre d'erreurs. Comment celui-ci peut-il les détecter, et si possible les corriger, de façon automatique, pourvu qu'elles ne soient pas trop nombreuses ?*

La théorie des codes correcteurs, qui est née vers 1950, consiste ainsi en la mise en oeuvre de codages, avant la transmission d'un message, dans le but que les erreurs éventuellement introduites dans le message lors de sa transmission soient détectées, et si possible corrigées. Il convient de supposer au départ que les erreurs éventuelles ne sont pas trop nombreuses, autrement dit que les canaux de transmission ont une fiabilité minimum, sans quoi aucune communication n'est possible. Le principe de ces codages est de rajouter à un message à transmettre une information supplémentaire, que l'on appelle information redondante ou de contrôle, et qui permet parfois la détection ou la correction des erreurs survenues. Cette opération est la phase d'encodage du message et son résultat est un mot de code. Par définition, le code est l'ensemble des mots de code ainsi obtenus.

Exemple 7.1. Illustrons sur un exemple ce que cela signifie dans son principe. Supposons qu'un individu A souhaite transmettre des messages à un individu B par les mots (de longueur 2) 01, 10, 11, et 00. Si A envoie à B le mot 01 et que le canal de transmission soit affecté de sorte que B reçoive le message 11, alors B n'a a priori aucune possibilité de savoir qu'une erreur soit survenue. Pour cette raison, il est recommandé que A encode ces messages par divers procédés qu'il choisit en fonction des contraintes imposées. Indiquons trois exemples possibles d'encodages.

1) Une première possibilité consiste à dédoubler chaque message, c'est-à-dire d'encoder les messages 01, 10, 11 et 00 respectivement par les mots de code (de longueur 4)

$$0101, \quad 1010, \quad 1111 \quad \text{et} \quad 0000.$$

On suppose au départ que le canal de transmission est suffisamment fiable de sorte que chaque mot soit affecté d'au plus une erreur. Autrement dit, au plus une composante de chaque mot peut être éventuellement modifiée lors de sa transmission. Dans ce cas,

B recevant un de ces mots de code, est en mesure de détecter une erreur éventuelle. En effet, si une erreur est survenue lors de la transmission d'un de ces mots, alors ce mot est transformé en un autre qui n'est pas un mot de code. Cela étant, il ne pourra pas corriger cette erreur, autrement dit, identifier le mot de code envoyer par A. Par exemple, si le message reçu est 1101, alors il se peut que A ait envoyé le mot 0101 ou bien 1111. De même, si B reçoit 0100, alors A peut avoir envoyé 0101 ou bien 0000. La seule possibilité pour B est alors de demander à A de lui réexpédier son message. Bien entendu, l'hypothèse faite sur la fiabilité du canal est fondamentale, car si deux erreurs étaient susceptibles de survenir lors d'une transmission, B pourrait recevoir le mot 0101 alors que A lui a envoyé 0000 et dans ce cas B n'a de nouveau aucun moyen de détecter la moindre erreur.

2) Une autre possibilité pour A est d'encoder ses messages par des mots plus courts. Par exemple, A peut encoder chacun des messages 01, 10, 11 et 00 en lui rajoutant une troisième composante de façon que la somme des composantes de chaque mot obtenu soit paire. On obtient dans ce cas le code formé des quatre mots (de longueur 3)

$$011, \quad 101, \quad 110 \quad \text{et} \quad 000.$$

Toujours sous l'hypothèse précédente de fiabilité du canal, B est de nouveau en mesure de détecter une erreur, vu qu'une erreur survenue modifie la parité du mot envoyé. Comme ci-dessus, ce procédé d'encodage ne permet pas à B de corriger une erreur éventuelle.

3) Un autre encodage possible, plus coûteux, consiste à tripler chaque message 01, 10, 11 et 00. On obtient alors les mots de code (de longueur 6)

$$010101, \quad 101010, \quad 111111 \quad \text{et} \quad 000000.$$

Dans ce cas, sous l'hypothèse d'au plus une erreur survenue, B peut la détecter, et il peut de plus la corriger. En effet, si x est le message reçu il existe un unique mot de code qui diffère de x d'une seule composante.

On généralise l'exemple précédent comme suit :

1) On choisit au départ un ensemble fini non vide F de cardinal q, on parle alors de l'alphabet F, de sorte que tous les messages à transmette soient des éléments de F^k. On dit que ce sont des messages de longueur k. Chaque message $(x_1, \cdots, x_k) \in F^k$ est la plupart du temps noté $x_1x_2 \cdots x_k$. L'ensemble F^k s'appelle l'espace des messages. Il est de cardinal q^k. Bien entendu, on peut aussi supposer que l'espace des messages est un sous-ensemble de F^k.

2) Supposons que l'on souhaite encoder les éléments d'un sous-ensemble I de F^k. Cette phase d'encodage consiste alors à choisir un entier $n > k$, puis à associer à chaque élément de I un mot de code de longueur n i.e. un élément de F^n, et cela de façon injective. On

obtient ainsi, ce que l'on appelle un code de F^n. Un encodage des éléments de I est donc une application injective

$$E : I \to F^n,$$

dont l'image est le code C associé à I au moyen de E. En particulier, E est une bijection de I sur C. Les mots de C sont des éléments de F^n. L'entier n est la longueur de C. On dit que C est un code q-aire de longueur n et de cardinal celui de I.

3) Les messages transmis au récepteur sont les éléments de C. Autrement dit, si l'on souhaite faire parvenir au récepteur l'élément $c \in I$, alors on lui transmet l'élément $E(c) \in F^n$ (et non pas c). On dit que l'on a introduit une redondance égale à $n-k$. Comme conséquence de la phase d'encodage, le récepteur obtient ainsi des mots plus longs que ceux de I, par suite les mots de C sont plus distinguables, i.e. plus distants les uns des autres en sens que l'on va préciser ci-après, ce qui peut lui permettre de détecter, voir de corriger, les erreurs éventuelles de transmission. Le rapport k/n s'appelle le taux d'information de C. Lors de la transmission d'un mot $m = (m_1, \cdots, m_n) \in C$, on dit que m est affecté de r erreurs, ou d'une erreur de poids r, si exactement r de ses composantes ont été modifiées.

4) Il reste la phase de décodage qui doit être effectuée si possible à la réception de chaque mot $m \in F^n$. Comme on l'a déjà signalé, ce décodage suppose certaines propriétés de fiabilité du canal de transmission, de façon à ce que le nombre d'erreurs éventuelles susceptibles d'affecter m soit limité. Il s'agit d'abord de déterminer si m appartient à C i.e. si m est un mot de code. Si tel est le cas, on considère qu'aucune erreur n'a affecté m lors de la transmission et le décodage de m est terminé (si $m \in C$, alors m se décode par lui-même). Dans le cas où m n'est pas un mot de code, il s'agit alors d'identifier autant que possible le mot de code émis en utilisant la redondance. Un décodage associé à C est donc une application

$$D : F^n \to F^n \quad \text{telle que} \quad D\big(F^n\big) = C,$$

ses points fixes étant exactement les éléments de C. Autrement dit, pour tout $x \in F^n$, on a $D(x) = x$ si et seulement si x appartient à C. Si le décodage est correct, le récepteur peut alors retrouver, au moyen de $E^{-1} : C \to I$, le message de I émis.

2. Distance de Hamming - Définition et paramètres d'un code

Considérons un ensemble fini non vide F et un entier $n \geq 1$. R. Hamming (1915-1998) a défini la distance entre deux éléments de F^n comme suit :

Définition 7.1. *Soient $x = (x_1, \cdots, x_n)$ et $y = (y_1, \cdots, y_n)$ deux éléments de F^n. On appelle distance de Hamming entre x et y, et on note $d_H(x, y)$, le nombre d'indices i tels que $1 \leq i \leq n$ et $x_i \neq y_i$.*

Proposition 7.1. *L'application $d_H : F^n \times F^n \to \mathbb{N}$ qui à tout couple $(x, y) \in F^n \times F^n$ associe $d_H(x, y)$ est une distance*[41] *sur F^n. On l'appelle la distance de Hamming sur F^n.*

Démonstration : L'application d_H, à valeurs en particulier dans $\mathbb{R}_+$, vérifie évidemment les deux premières conditions ci-dessous de la définition d'une distance. Vérifions l'inégalité triangulaire. Soient x, y, z des éléments de F^n. Soit i un indice entre 1 et n tel que $x_i \neq y_i$. On a nécessairement $x_i \neq z_i$ ou bien (non exclusif) $z_i \neq y_i$, ce qui entraîne le résultat.

Suivant la terminologie employée dans la problématique des codes correcteurs, voici ce que l'on entend par code sur un alphabet (i.e. un ensemble) fini F :

Définition 7.2. *Un ensemble C est un code sur F s'il existe un entier $n \geq 1$ tel que C soit une partie de F^n. On dit que n est la longueur de F. On appelle code de F^n tout code de longueur n sur F. Ce sont donc exactement les parties de F^n. Les éléments d'un code C s'appellent les mots de code de C. Si F et de cardinal q, on dit que C est un code q-aire. Il est dit binaire si $q = 2$ et ternaire si $q = 3$.*

La notion de distance minimum d'un code est essentielle :

Définition 7.3. *Soit C un code sur F. On appelle distance minimum de C l'entier d défini par l'égalité*

$$d = \mathrm{Min}\Big\{d_H(x, y) \mid x \in C, y \in C \text{ et } x \neq y\Big\}.$$

Autrement dit, d est la plus petite des distances entre deux éléments distincts de C. Si C est un singleton, on convient que $d = 0$.

Les paramètres fondamentaux associés à un code sont sa longueur, son cardinal et sa distance minimum.

Définition 7.4. *Un code sur un ensemble de cardinal q, de longueur n, de cardinal M et de distance minimum d, est appelé un code q-aire de paramètres (n, M, d).*

---

[41] Soit E un ensemble. On appelle distance sur E toute application $d : E \times E \to \mathbb{R}_+$ ($\mathbb{R}_+$ étant l'ensemble des nombres réels positifs ou nuls), telle que pour tous $x, y, z \in E$ les conditions suivantes soient vérifiées :

1) on a $d(x, y) = 0$ si et seulement si $x = y$.
2) On a $d(x, y) = d(y, x)$.
3) On a $d(x, y) \leq d(x, z) + d(z, y)$ (inégalité triangulaire).

Un tel couple (E, d) s'appelle un espace métrique. À titre d'exemple, l'application $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ définie pour tous $x, y \in \mathbb{R}$ par l'égalité $d(x, y) = |x - y|$ est une distance sur $\mathbb{R}$. Plus généralement, si $f : \mathbb{R} \to \mathbb{R}$ est une application injective de $\mathbb{R}$ dans $\mathbb{R}$, l'application $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ définie par $d(x, y) = |f(x) - f(y)|$ est une distance sur $\mathbb{R}$.

Exercice 1. Déterminer les paramètres associés au code binaire

$$C = \Big\{00001100, 00001111, 01010101, 11011101\Big\}.$$

3. Décodage par maximum de vraisemblance - Capacité de correction

Soit C un code dont on transmet les éléments à travers un canal bruité. Afin de décoder des messages reçus, on adopte généralement le principe de décodage par maximum de vraisemblance. Sa signification est la suivante. Si x est un mot reçu, on décode x par le mot, ou l'un des mots, le plus proches de x au sens de la distance de Hamming. C'est ce principe qui rend l'observation de x la plus plausible. Autrement dit, Si r est la plus petite distance de x à un mot de code, i.e. si r est le plus petit entier ≥ 0 pour lequel il existe $m \in C$ tel que $d_H(m, x) = r$, on part alors du principe qu'il y eu exactement r erreurs introduites dans le message émis. Il se présentent alors deux possibilités :

1) Il existe un unique élément $m \in C$ tel que l'on ait $d(m, x) = r$. Dans ce cas, on décode x par m i.e. on corrige x par m.
2) Il existe plusieurs mots de code qui sont à une distance r de x. On décode alors x par l'un de ces mots.

Le deuxième cas n'est évidemment pas favorable. Ce principe de décodage peut aussi conduire à des corrections erronées dans le premier cas. En effet, supposons que lors de la transmission d'un message $m \in C$, le mot x reçu soit affecté de s erreurs et qu'il existe un unique $m' \in C$ tel que l'on ait

$$d_H(m', x) < s. \tag{1}$$

Dans ce cas, m' est le mot de code le plus proche de x, et l'on effectuera un décodage incorrect en décodant x par m'. Pour effectuer un décodage correct il convient donc les erreurs introduites lors de la transmission d'un message ne soient pas trop nombreuses. On va préciser dans ce qui suit ce que l'on entend par là.

Exemple 7.2. Soit d est la distance minimum de C. Une condition suffisante pour éviter la présence d'un élément $m' \in C$ vérifiant (1) est que le message reçu soit affecté de s erreurs avec

$$s \leq \left[\frac{d}{2}\right].$$

En effet, par hypothèse il existe $m \in C$ tels que l'on ait $d_H(m, x) = s$. S'il existe $m' \in C$ tel que l'on ait $d_H(m', x) < s$, il résulte alors de l'inégalité triangulaire que l'on a

$$d_H(m, m') < 2s \leq d,$$

ce qui contredit la définition de d, vu que m et m' sont distincts.

Cela étant, si d est pair et si l'on a $s = \left[\frac{d}{2}\right]$, il n'est pas exclu que x se trouve à une distance $\frac{d}{2}$ de deux mots distincts de C. On peut alors seulement détecter qu'il y a eu $\frac{d}{2}$ erreurs de transmission, et l'on se trouve dans le deuxième cas envisagé ci-dessus. On va constater maintenant que l'on peut effectuer un décodage correct pourvu que l'on ait

$$s \leq \left[\frac{d-1}{2}\right].$$

Introduisons pour cela la notion de capacité de correction d'un code.

Capacité de correction

Soient F un alphabet fini, n un entier ≥ 1 et C un code de F^n.

Notation. Pour tout $x \in F^n$ et tout entier $r \geq 0$, on notera $B(x, r)$ la boule fermée de centre x et de rayon r pour la distance d_H. Autrement dit, on a

$$B(x, r) = \Big\{ y \in F^n \mid d_H(x, y) \leq r \Big\}.$$

Définition 7.5. *Soit d la distance minimum de C. La capacité de correction de C, notée souvent t, est l'entier*

$$t = \left[\frac{d-1}{2}\right].$$

On dit alors que C est un code t-correcteur.

Lemme 7.1. *Soit t la capacité de correction de C.*

1) *Pour tous m et m' dans C tels que $m \neq m'$, les boules $B(m, t)$ et $B(m', t)$ sont disjointes.*
2) *Soit x un élément de F^n. Il existe au plus un élément de C qui appartienne à la boule $B(x, t)$.*

Démonstration : Supposons qu'il existe un élément $x \in F^n$ qui appartienne à $B(m, t)$ et $B(m', t)$ avec $m \neq m'$. Dans ce cas, on a

$$d_H(m, m') \leq d_H(m, x) + d_H(x, m') \leq 2t \leq d - 1.$$

Les mots m et m' étant distincts et d étant la distance minimum de C, on obtient ainsi une contradiction. L'argument est le même pour la deuxième assertion : supposons qu'il existe deux éléments m et m' de C dans la boule $B(x, t)$. On a alors $d_H(m, m') \leq d - 1$, d'où $m = m'$.

On en déduit le résultat annoncé :

Corollaire 7.1. *Soit x un message reçu. Supposons que x soit affecté de r erreurs avec $r \leq t$. Alors, il existe un unique élément $m \in C$ tel que $d_H(m,x) = r$, et l'on effectue un décodage correct en décodant x par m.*

Démonstration : Soit m un mot de C tel que $d_H(m,x) = r$. Puisque l'on a l'inégalité $r \leq t$, l'élément m appartient en particulier à $B(x,t)$. D'après l'assertion 2 du lemme 7.1, c'est l'unique mot de code dans cette boule et m est donc l'unique mot de code à une distance r de x. Cela entraîne le résultat en vertu du principe de décodage par maximum de vraisemblance.

Ce résultat justifie la terminologie utilisée dans la définition 7.5.

Exemple 7.3. On considère le code sur $\mathbb{F}_2$ rencontré dans l'exemple 7.1 :

$$C = \Big\{010101, 101010, 111111, 000000\Big\}.$$

La distance minimum de C est $d = 3$ et sa capacité de correction est $t = 1$. On retrouve ici le fait que l'on puisse corriger une erreur lors de la transmission d'un de ces quatre mots. Ainsi, C est un code 1-correcteur. On dit aussi que C corrige une erreur.

4. Codes parfaits

Soit C un code q-aire sur un alphabet fini F de paramètres (n, M, d). Rappelons que cela signifie que F est de cardinal q et que C est un code dans F^n, de cardinal M et de distance minimum d. Soit t la capacité de correction de C.

Définition 7.6. *On dit que C est un code parfait si les boules fermées centrées en les mots de C et de rayon t recouvrent F^n. Autrement dit, C est parfait si et seulement si on a la réunion disjointe*

$$F^n = \bigcup_{m \in C} B(m,t).$$

C'est une situation idéale de décodage, vu que pour tout message x reçu, il existe dans ce cas un unique élément $m \in C$ tel que x appartienne à la boule $B(m,t)$, et l'on décode alors x par m. Les codes parfaits sont malheureusement rares. On en verra des exemples au paragraphe 11.

Proposition 7.2. *Soit r un entier naturel. Pour tout $x \in F^n$, on a l'égalité*

$$\big|B(x,r)\big| = \sum_{i=0}^{r} C_n^i (q-1)^i.$$

En particulier, le cardinal de $B(x,r)$ ne dépend pas de x.

Démonstration : Soit i un entier vérifiant les inégalités $0 \leq i \leq r$. Il existe exactement $C_n^i(q-1)^i$ éléments $y \in F^n$ tels que l'on ait $d_H(x,y) = i$, d'où l'assertion.

Corollaire 7.2. (Borne d'empilement des sphères) *On a l'inégalité, appelée borne d'empilement des sphères,*

$$M \sum_{i=0}^{t} C_n^i (q-1)^i \leq q^n.$$

Démonstration : Vu que F^n est de cardinal q^n, cela résulte de la proposition 7.2 et de l'assertion 1 du lemme 7.1.

Corollaire 7.3. *Le code C est parfait si et seulement si on a l'égalité*

$$M \sum_{i=0}^{t} C_n^i (q-1)^i = q^n.$$

Démonstration : Supposons que C soit parfait. On alors (lemme 7.1)

$$q^n = \sum_{m \in C} \big|B(m,t)\big|,$$

d'où l'égalité annoncée (prop. 7.2). Inversement, supposons cette égalité réalisée. On a alors

$$\Big| \bigcup_{m \in C} B(m,t) \Big| = \sum_{m \in C} \big|B(m,t)\big| = q^n,$$

ce qui entraîne que F^n est la réunion des boules $B(m,t)$ où m parcourt C, d'où le résultat.

5. Les entiers $A_q(n,d)$

On peut considérer qu'un code est « bon » si son cardinal M et sa distance minimum d sont grands. En effet, plus M est grand, plus on peut transmettre d'informations, et plus d est grand, plus elles se différencient les unes des autres et l'on peut ainsi corriger un grand nombre d'erreurs. En fait, ces exigences sont contradictoires. En effet, si l'on fixe les paramètres n (la longueur) et q (le cardinal de l'alphabet utilisé) alors, quand M croit, d diminue, vu que plus il y a de mots de code, plus ces derniers sont proches. Cela justifie la définition suivante :

Définition 7.7. *Soient n, d et q trois entiers naturels vérifiant les inégalités $q \geq 2$ et $1 \leq d \leq n$. L'entier $A_q(n,d)$ est le plus grand entier naturel M pour lequel il existe un code q-aire de paramètres (n, M, d).*

Remarquons que les hypothèses faites sur n, d et q entraînent l'existence d'un code q-aire de distance minimum d. En effet, étant donné un ensemble F de cardinal q, il suffit de prendre comme code une partie de F^n ayant exactement deux éléments x et y tels que $d_H(x,y) = d$. Par ailleurs, les cardinaux des codes q-aires de longueur n sont majorés par q^n. L'entier $A_q(n,d)$ est donc bien défini.

Par exemple, on a $A_q(n,1) = q^n$. L'entier $A_q(n,d)$ n'est pas connu en général, y compris pour des petites valeurs de q, n et d. À titre d'illustration, on va déterminer ici les entiers $A_2(n,2)$:

Proposition 7.3. *Pour tout entier $n \geq 2$, on a $A_2(n,2) = 2^{n-1}$.*

Démonstration : Vérifions d'abord que l'on a

$$A_2(n,2) \leq 2^{n-1}. \tag{2}$$

Considérons pour cela un code C dans $\mathbb{F}_2^n$ de distance minimum 2 et de cardinal $A_2(n,2)$. Il existe x et y dans C tels que $d_H(x,y) = 2$. Posons $x = (x_1, \cdots, x_n)$ et $y = (y_1, \cdots, y_n)$. Soit i le plus petit indice (par exemple) tel que $x_i \neq y_i$. Soit C' le code de longueur $n-1$ déduit de C en supprimant la i-ème composante des éléments de C. La distance minimum de C' est 1. Par ailleurs, puisque l'on a $d = 2$, le cardinal de C' est celui de C i.e. est $A_2(n,2)$. Le code C' étant un code binaire de longueur $n-1$ et de distance minimum 1, son cardinal $A_2(n,2)$ est donc inférieur ou égal à $A_2(n-1,1)$. L'égalité $A_2(n-1,1) = 2^{n-1}$ entraîne alors (2).

Il suffit alors de construire un code dans $\mathbb{F}_2^n$ de distance minimum 2 et de cardinal 2^{n-1} pour obtenir le résultat. Vérifions que le code C de $\mathbb{F}_2^n$ formé des mots ayant un nombre pair de composantes égales à 1 possède ces propriétés. Il existe deux mots x et y dans C tels que $d_H(x,y) = 2$. En effet, si n est pair il en est ainsi avec $x = (1, \cdots, 1)$ et $y = (1, \cdots, 1, 0, 0)$ et si n est impair, on a la même conclusion en prenant $x = (1, \cdots, 1, 0)$ et $y = (1, \cdots, 1, 0, 0, 0)$. Par ailleurs, il n'existe pas deux éléments distincts de C à une distance 1, sinon l'un d'entre eux aurait un nombre impair de composantes égales à 1. On en déduit que la distance minimum de C vaut 2. Tout revient alors à démontrer que le cardinal de C est 2^{n-1}. On remarque pour cela que l'on a

$$|C| = \sum_{0 \leq k \leq n,\ k\ pair} C_n^k.$$

Par ailleurs, on a

$$2^n = (1+1)^n = \sum_{0 \leq k \leq n,\ k\ pair} C_n^k + \sum_{0 \leq k \leq n,\ k\ impair} C_n^k,$$

et des égalités

$$0 = (1-1)^n = \sum_{k=0}^{n} (-1)^k C_n^k,$$

on déduit que l'on a

$$\sum_{0 \leq k \leq n,\ k\ pair} C_n^k = \sum_{0 \leq k \leq n,\ k\ impair} C_n^k.$$

Cela établit notre assertion et le résultat.

Terminons ce paragraphe en explicitant un encadrement de $\frac{q^n}{A_q(n,d)}$:

Proposition 7.4. *On a les inégalités*

$$\sum_{i=0}^{\left[\frac{d-1}{2}\right]} C_n^i (q-1)^i \leq \frac{q^n}{A_q(n,d)} \leq \sum_{i=0}^{d-1} C_n^i (q-1)^i.$$

Prouvons d'abord le résultat suivant :

Lemme 7.2. *Soient F un ensemble et C un code dans F^n de distance minimum $d \geq 1$. Supposons qu'il n'existe pas de code dans F^n de distance minimum d qui contienne strictement C. Alors, on a l'égalité*

$$F^n = \bigcup_{m \in C} B(m, d-1).$$

Démonstration : Supposons qu'il existe $y \in F^n$ tel que pour tout $m \in C$, y ne soit pas dans la boule $B(m, d-1)$. Pour tout $m \in C$ on a $d_H(y,m) \geq d$, et y n'est pas dans C. Soit alors C' le code

$$C' = C \cup \{y\}.$$

C'est un code contenu dans F^n qui contient strictement C et de distance minimum d. On obtient ainsi une contradiction et le résultat.

Corollaire 7.4. *On a l'inégalité*

$$q^n \leq A_q(n,d) \sum_{i=0}^{d-1} C_n^i (q-1)^i.$$

Démonstration : Soit C un code q-aire de paramètres $(n, A_q(n,d), d)$. Vu le caractère maximal de $A_q(n,d)$, le code C vérifie l'hypothèse faite dans le lemme 7.2. On a donc

$$q^n \leq \sum_{m \in C} \Big| B(m, d-1) \Big| = A_q(n,d) \sum_{i=0}^{d-1} C_n^i (q-1)^i.$$

La proposition 7.4 est alors une conséquence des corollaires 7.2 et 7.4.

6. Codes linéaires

Soient p un nombre premier et q une puissance de p. On notera dans la suite $\mathbb{F}_q$ «le» corps à q éléments. On sait qu'il est unique à isomorphisme près. Pour tout $n \geq 1$, on note d_H la distance de Hamming sur $\mathbb{F}_q^n$.

Définition 7.8. *Un ensemble C est un code linéaire sur $\mathbb{F}_q$ si et seulement si il existe un entier $n \geq 1$ tel que C soit un sous-espace vectoriel non nul de $\mathbb{F}_q^n$* [42].

Terminologie. Soit C un code linéaire sur $\mathbb{F}_q$. On dit que C est un code q-aire de longueur n et de dimension k, ou bien que C un code linéaire de longueur n et de dimension k sur $\mathbb{F}_q$, pour signifier que C est un sous-espace vectoriel de dimension k de $\mathbb{F}_q^n$. Si d est la distance minimum de C, on dit que C est un code linéaire de type $(n, k, d)_q$.

Si $q = 2$, un code q-aire s'appelle comme il est d'usage un code binaire.

Remarque 7.1. Un code linéaire de dimension k sur $\mathbb{F}_q$ est de cardinal q^k.

Exercice 2. Soit C le code binaire défini par

$$C = \Big\{0000, 1100, 1010, 1001, 0110, 0101, 0011, 1111\Big\}.$$

Montrer que C est un code linéaire de longueur 4 et de dimension 3 sur $\mathbb{F}_2$. Quelle est sa distance minimum ?
On constatera en fait que le cardinal de C est $A_2(4, 2) = 8$ (cf. prop. 7.3). De même, on pourra vérifier que $\Big\{000, 110, 011, 101\Big\}$ est un code linéaire binaire de cardinal $A_2(3, 2)$.

Définition 7.9. *Soit x un élément de $\mathbb{F}_q^n$. On appelle poids de x, et on note $w(x)$, le nombre de composantes non nulles de x. Autrement dit, on a $w(x) = d_H(x, 0)$.*

On a l'énoncé immédiat suivant :

Lemme 7.3. *Pour tous x et y dans $\mathbb{F}_q^n$, on a $d_H(x, y) = w(x - y)$.*

Considérons désormais un code linéaire C de type $(n, k, d)_q$.

Lemme 7.4. *La distance minimum d de C est donnée par l'égalité*

$$d = \mathrm{Min}\Big\{w(x) \;\Big|\; x \in C,\ x \neq 0\Big\}.$$

Démonstration : Posons $d' = \mathrm{Min}\Big\{w(x) \;\Big|\; x \in C,\ x \neq 0\Big\}$. Par définition, d' est la plus petite des distances des éléments non nuls de C à 0. Puisque C est un espace vectoriel

[42] Soient K un corps commutatif et E un espace vectoriel sur K. Une partie F de E est un sous-espace vectoriel de E si et seulement si les deux conditions suivantes sont satisfaites :

1) F est un sous-groupe du groupe additif E.
2) Pour tous $x \in F$ et $\lambda \in K$, l'élément λx appartient à F.

Si E est de dimension finie n sur K, alors F l'est aussi et sa dimension est inférieure ou égale à n. Elle vaut n si et seulement si $F = E$.

sur $\mathbb{F}_q$, le mot nul i.e. 0 appartient à C. Il en résulte que l'on a $d \leq d'$. Par ailleurs, il existe deux éléments y et z de C tels que l'on ait $d_H(y,z) = d$. Vu que C est non nul, on a $d \geq 1$, en particulier, y et z sont distincts. L'égalité $w(y-z) = d$ (lemme 7.3) entraîne alors $d' \leq d$, d'où $d = d'$.

On dispose de la majoration suivante de d en fonction de n et k, que l'on appelle la borne de Singleton :

Proposition 7.5. (Borne de Singleton) *On a l'inégalité*

$$d \leq n - k + 1.$$

Démonstration : Considérons le sous-ensemble F de $\mathbb{F}_q^n$ formé des éléments dont les $k-1$ dernières composantes sont nulles. C'est un sous-espace vectoriel de $\mathbb{F}_q^n$. Soit $(e_i)_{1\leq i\leq n}$ la base canonique de $\mathbb{F}_q^n$ sur $\mathbb{F}_q$: toutes les composantes de e_i sont nulles sauf la i-ème qui vaut 1. Le système $\big(e_1, \cdots, e_{n-k+1}\big)$ est une base de F, par suite la dimension de F est $n-k+1$. Il en résulte que $F \cap C$ n'est pas le sous-espace nul[43]. Il existe donc un élément x non nul de $F \cap C$. On a $w(x) \leq n-k+1$ et le lemme 7.4 entraîne alors le résultat.

On définit la distance relative de C comme étant le rapport $\frac{d}{n}$. La borne de Singleton entraîne que la distance relative et le taux d'information $\frac{k}{n}$ de C ne peuvent pas être généralement simultanément proches de 1 vu que leur somme est plus petite que $1 + \frac{1}{n}$.

---

[43] Soit E un espace vectoriel de dimension finie sur corps K. Soient F et G deux sous-espaces vectoriels de E. L'ensemble $F+G$ formé des $a+b$, où $a \in F$ et $b \in G$ est un sous-espace vectoriel de E (c'est le plus petit sous-espace vectoriel de E contenant F et G). Vérifions que si l'on a $F \cap G = \big\{0\big\}$, alors on a l'égalité des dimensions :

$$\dim(F+G) = \dim F + \dim G.$$

Cela entraînera notre assertion dans la démonstration de la proposition 7.5 : avec ses notations, si $F \cap C$ était nul, on aurait $\dim(F+C) = \dim F + \dim C = n-k+1+k = n+1$, ce qui conduit à une contradiction vu que $\dim \mathbb{F}_q^n = n$.
Soient (u_i) une base de F et (v_j) une base de G. Le système formé des u_i et des v_j est évidemment un système générateur de $F+G$. C'est par ailleurs un système libre, car une égalité de la forme $\sum \alpha_i u_i + \sum \beta_j v_j = 0$ entraîne $u_i = v_j = 0$ pour tous i et j, comme on le constate en utilisant le fait que $F \cap G = \big\{0\big\}$, d'où l'égalité annoncée.
Plus généralement, signalons que pour tous sous-espaces vectoriels F et G de E, on a l'égalité

$$\dim F + \dim G = \dim(F+G) + \dim(F \cap G).$$

On peut démontrer cette égalité en considérant l'application linéaire $f : F \times G \to E$ définie par $f\big((x,y)\big) = x+y$. Son image est le sous-espace $F+G$. Son noyau $\mathrm{Ker}(f)$ est formé des couples $(x,-x)$, où x est dans $F \cap G$. L'application $x \mapsto (x,-x)$ étant un isomorphisme de $F \cap G$ sur $\mathrm{Ker}(f)$, on a en particulier $\dim(F \cap G) = \dim\big(\mathrm{Ker}(f)\big)$. Il en résulte que l'on a $\dim(F \times G) = \dim(F \cap G) + \dim(F+G)$. L'égalité $\dim(F \times G) = \dim F + \dim G$ entraîne alors le résultat (rappelons que si U et V sont des espaces vectoriels sur K de dimensions finies, et si $g : U \to V$ est une application linéaire, on a $\dim U = \dim\big(\mathrm{Ker}(g)\big) + \dim\big(g(U)\big)$.

Définition 7.10. *On dit que C est un code MDS, en anglais Maximum Distance Separable, si l'on a l'égalité $d = n - k + 1$.*

Exemple 7.4. Le code linéaire binaire $\Big\{000, 110, 011, 101\Big\}$ est MDS.

Les codes MDS sont des bons codes puisque, à n et k fixés, leur distance minimale est la plus grande possible. Certains codes, utilisés pour la lecture des disques compacts ou les informations satellitaires, sont MDS.

7. Encodages associés aux codes linéaires - Matrices génératrices

Considérons un code linéaire C sur $\mathbb{F}_q$ de longueur n et de dimension k. On suppose que l'espace des messages est $\mathbb{F}_q^k$, autrement dit, que l'ensemble des messages à transmettre est $\mathbb{F}_q^k$. La dimension de C sur $\mathbb{F}_q$ étant k, le choix d'une base de C sur $\mathbb{F}_q$, détermine un isomorphisme E de $\mathbb{F}_q^k$ sur C, ou si l'on préfère une application $\mathbb{F}_q$-linéaire injective de $\mathbb{F}_q^k$ à valeurs dans $\mathbb{F}_q^n$ dont l'image est C. Conformément à la problématique des codes correcteurs présentée au début du chapitre, l'application E est alors un encodage de $\mathbb{F}_q^k$. C'est par définition un encodage associé à C. Chaque base de C sur $\mathbb{F}_q$ permet ainsi de construire un encodage associé à C.

Du point de vue des codeurs, les éléments d'un produit cartésien $\mathbb{F}_q^N$ (où $N \geq 1$) sont des mots de longueur N. Ils identifient ainsi ces éléments, i.e. les N-uplets d'éléments de $\mathbb{F}_q$, avec les matrices à une ligne et N colonnes. Autrement dit, l'élément $(a_1, \cdots, a_N) \in \mathbb{F}_q^N$ est identifié avec le vecteur ligne $(a_1 \cdots a_N)$. Nous ferons implicitement cette identification dans la suite, sans autre précision. Cela est contraire à l'usage en algèbre linéaire classique, où l'on identifie les éléments de $\mathbb{F}_q^N$ avec les matrices à une colonne et N lignes.

Définition 7.11. *Une matrice à coefficients dans $\mathbb{F}_q$, ayant k lignes et n colonnes, est appelée matrice génératrice de C si ses vecteurs lignes forment une base de C.*

Lemme 7.5. *Le rang d'une matrice génératrice de C est k* [44].

Démonstration : Le rang d'une matrice de taille (k, n) à coefficients dans $\mathbb{F}_q$ est la dimension du sous-espace vectoriel de $\mathbb{F}_q^n$ engendré par ses vecteurs lignes, d'où l'assertion.

---

[44] Rappelons que le rang d'une matrice M de taille (r, s), à coefficients dans un corps commutatif K, est la dimension du sous-espace vectoriel de K^r engendré par les vecteurs colonnes de M, qui est aussi la dimension du sous-espace vectoriel de K^s engendré par les vecteurs lignes de M. On démontre que le rang de M est le plus grand entier t pour lequel il existe une matrice carrée de taille (t, t) extraite de M inversible.

Exercice 3. Quel est le rang de la matrice $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$ à coefficients dans K ?

Lemme 7.6. *Soit M une matrice à coefficients dans $\mathbb{F}_q$, ayant k lignes et n colonnes, et de rang k. Alors, les k vecteurs lignes de M forment une base d'un code linéaire sur $\mathbb{F}_q$ de longueur n.*

Démonstration : C'est immédiat compte tenu du rappel du bas de page précédent et du fait qu'en dimension k, un système de k vecteurs forme une base.

Proposition 7.6. *Soit G une matrice génératrice de C. Dans $\mathbb{F}_q^n$, on a l'égalité*

$$C = \Big\{(u_1, \cdots, u_k)G \mid (u_1, \cdots, u_k) \in \mathbb{F}_q^k\Big\}.$$

Démonstration : Notons $(e_i)_{1\leq i\leq k}$ la base canonique de $\mathbb{F}_q^k$ et $(f_j)_{1\leq j\leq n}$ la base canonique de $\mathbb{F}_q^n$. Soit ℓ_i le i-ème vecteur ligne de G et $E : \mathbb{F}_q^k \to \mathbb{F}_q^n$ l'application $\mathbb{F}_q$-linéaire définie par $E(e_i) = \ell_i$ pour $1 \leq i \leq k$. La matrice qui représente E dans les bases précédentes est la transposée de G. Pour tout élément $x \in \mathbb{F}_q^k$, ses coordonnées (u_i) dans la base (e_i) sont donc liées aux coordonnées (v_j) de $E(x)$ dans la base (f_j) par l'égalité $(v_1, \cdots, v_n) = (u_1, \cdots, u_k)G$. Cela entraîne le résultat compte tenu du fait que C est l'image de E.

Exemple 7.5. Considérons la matrice de taille $(3, 5)$ à coefficients dans $\mathbb{F}_2$:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

son rang est 3. Soit $(e_i)_{1\leq i\leq 5}$ la base canonique de $\mathbb{F}_2^5$. Les trois vecteurs lignes $e_1 + e_5$, $e_2 + e_4$ et $e_3 + e_5$ de G engendrent un code linéaire C sur $\mathbb{F}_2$ de longueur 5 et de dimension 3 (lemme 7.6). L'espace des messages est ici $\mathbb{F}_2^3$ et tout élément $(u_1, u_2, u_3) \in \mathbb{F}_2^3$ est encodé par l'élément

$$(u_1, u_2, u_3)G = (u_1, u_2, u_3, u_2, u_1 + u_3) \in C.$$

Avec cet encodage, on constate que les trois premières coordonnées correspondent au message à envoyer et que les deux dernières $(u_2, u_1 + u_3)$ correspondent à la redondance. Cela tient au fait que la matrice extraite de M au moyen de ses trois lignes et de ses trois premières colonnes est l'identité. On vérifie par ailleurs que la distance minimum de C est $d = 2$. En particulier, C ne corrige aucune erreur.

8. Codes systématiques

Pour tout entier $m \geq 1$, notons I_m la matrice identité de taille (m, m). On considère dans tout ce paragraphe un code linéaire C sur $\mathbb{F}_q$ de longueur n et de dimension k.

Définition 7.12. *On dit que C est un code systématique s'il existe une matrice B à coefficients dans $\mathbb{F}_q$, ayant k lignes et $n - k$ colonnes, telle que $\big(I_k|B\big)$ soit une matrice génératrice de C. Une matrice de cette forme est dite sous forme standard, ou normalisée.*

En fait, dans le cas où C est systématique, il existe une unique matrice normalisée qui soit une matrice génératrice de C. Plus précisément :

Lemme 7.7. *Il existe au plus une matrice B à coefficients dans $\mathbb{F}_q$, ayant k lignes et $n-k$ colonnes, telle que $\big(I_k|B\big)$ soit une matrice génératrice de C.*

Démonstration : Soient B et B' deux matrices à coefficients dans $\mathbb{F}_q$, ayant k lignes et $n-k$ colonnes, telles que $\big(I_k|B\big)$ et $\big(I_k|B'\big)$ soient des matrices génératrices de C. Soient $(\ell_i)_{1\leq i\leq k}$ les lignes de $\big(I_k|B\big)$ et $(\ell'_i)_{1\leq i\leq k}$ celles de $\big(I_k|B'\big)$. Soit i un indice entre 1 et k. D'après l'hypothèse faite, il existe des éléments $\lambda_{ij} \in \mathbb{F}_q$ tels que l'on ait

$$\ell_i = \sum_{j=1}^{k} \lambda_{ij}\ell'_i.$$

Les coordonnées des vecteurs lignes des deux membres étant les mêmes, on en déduit que $\lambda_{ii} = 1$ et que $\lambda_{ij} = 0$ pour tout $j \neq i$, d'où $\ell_i = \ell'_i$, puis $B = B'$.

Dans l'exemple ci-dessus le code est défini au départ par une matrice génératrice normalisée, en particulier, il est systématique. Comme on l'a constaté, on peut lire directement le message de $\mathbb{F}_2^3$ à transmettre sur l'encodage de l'élément de $\mathbb{F}_2^5$ correspondant. C'est l'intérêt principal des codes systématiques. On retrouve directement tout message de $\mathbb{F}_q^k$ par les k premières coordonnées de son encodage dans $\mathbb{F}_q^n$.

Il existe des codes qui ne sont pas systématiques. Connaissant une matrice génératrice de C, l'énoncé suivant permet de décider facilement si C est systématique ou non.

Théorème 7.1. *Soient G une matrice génératrice de C et a_{ij} le coefficient de la i-ème ligne et de la j-ième colonne de G (on a $1 \leq i \leq k$ et $1 \leq j \leq n$). Alors, C est systématique si et seulement si la matrice extraite (a_{ij}) de G avec $1 \leq i, j \leq k$, est inversible.*

Avant de démontrer ce résultat, il convient de faire un rappel sur la notion d'opération élémentaire sur les lignes d'une matrice.

Opérations élémentaires sur les lignes d'une matrice

Elles consistent à transformer une matrice en une autre de même taille au moyen de certaines manipulations sur ses lignes. Plus précisément, soient M une matrice à coefficients dans un corps K, et ℓ_i le i-ème vecteur ligne de M. Les transformations suivantes sont appelées « opérations élémentaires sur les lignes » de M. Étant donnés deux indices i et j distincts, toute ligne ℓ_k de M, avec k distinct de i et j, est fixe et l'on effectue sur ℓ_i et ℓ_j l'une des transformations ci-dessous :

1) permutation ou échange de ℓ_i et ℓ_j.

2) Multiplication de ℓ_i par λ, où $\lambda \in K$ est non nul. Cette opération, qui remplace ℓ_i par $\lambda\ell_i$, s'appelle une dilatation de ℓ_i.

3) Addition à ℓ_i de $\lambda\ell_j$, où $\lambda \in K$. Cette opération, qui remplace ℓ_i par $\ell_i + \lambda\ell_j$, s'appelle une transvection de ℓ_i.

On utilise le résultat suivant dans la démonstration du théorème 7.1 :

Proposition 7.7. *Toute opération élémentaire sur les lignes de G transforme G en une autre matrice génératrice de C.*

Démonstration : Il est immédiat de constater qu'une opération de permutation ou de dilatation transforme la base (ℓ_i) de C en une autre base de C. Considérons un élément $\lambda \in \mathbb{F}_q$ et deux lignes ℓ_i et ℓ_j de C, avec $i \neq j$. Il s'agit de vérifier que le système $(\ell_1, \cdots, \ell_{i-1}, \ell_i + \lambda\ell_j, \ell_{i+1}, \cdots, \ell_k)$ est une base de C, ce qui résulte par exemple du fait que ce système de k vecteurs de C est libre.

Toute matrice déduite de M par des opérations élémentaires sur les lignes s'obtient en fait en multipliant à gauche M par une matrice inversible convenable. En particulier, une telle matrice a le même rang que M. On n'utilisera pas dans ce qui suit la traduction matricielle de ces opérations, on illustrera seulement cette remarque à travers un exemple. Le résultat que l'on utilise ici dans la démonstration du théorème 7.1 est le suivant :

Proposition 7.8. *Soit n un entier ≥ 1. Toute matrice carrée de taille (n, n) inversible à coefficients dans K peut être transformée par des opérations élémentaires sur ses lignes en la matrice identité I_n.*

La démonstration algorithmique de ce résultat consiste à utiliser la méthode dite du pivot de Gauss. À l'aide de transvections convenables sur les lignes, on fait apparaître des zéros en dehors de la diagonale de la matrice, et l'on obtient ensuite la matrice I_n par les dilatations associées aux éléments diagonaux, qui sont non nuls, vu que la matrice du départ est supposée inversible. Plutôt que de rappeler sa démonstration formelle, voyons sur un exemple la mise en oeuvre de cette méthode. La situation générale est analogue.

Exemple 7.6. Considérons la matrice à coeficients dans $\mathbb{F}_5$ définie par

$$M = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}.$$

Son déterminant est $-1 \in \mathbb{F}_5$, il est non nul, donc M est inversible. Notons ℓ_1, ℓ_2, ℓ_3 les lignes de M et par abus celles de toute autre matrice déduite de M par des opérations élémentaires. On commence par transformer M de façon à obtenir une matrice avec des zéros sous la diagonale principale.

1) On effectue la transvection sur ℓ_2, qui remplace ℓ_2 par $\ell_2 - \frac{3}{2}\ell_1 = \ell_2 + \ell_1$. On obtient comme nouvelle matrice
$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}.$$

2) On effectue la transvection sur ℓ_3, qui remplace ℓ_3 par $\ell_3 - \frac{1}{2}\ell_1 = \ell_3 + 2\ell_1$. On obtient après cette transformation la matrice
$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 4 & 0 \end{pmatrix}.$$

3) La transvection qui remplace ℓ_3 par $\ell_3 - 2\ell_2$ conduit alors à la matrice
$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

4) La transformation $\ell_1 \to \ell_1 - \frac{1}{2}\ell_2 = \ell_1 + 2\ell_2$ conduit à
$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

5) La transformation $\ell_2 \to \ell_2 - 2\ell_3$ conduit à
$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

6) Avec les multiplications par $\frac{1}{2} = 3$ des deux premières lignes (dilatations de rapport 3), on obtient alors la matrice identité I_3.

Explicitons les multiplications à gauche de M par des matrices convenables, permettant d'obtenir les transformations souhaitées. De façon générale, dans le cas d'une matrice carrée (n, n), afin d'effectuer la transformation $\ell_i \to \ell_i + \lambda\ell_j$, en laissant fixe les autres lignes, on multiplie à gauche M par la matrice T_{ij} dont l'élément de la r-ième ligne et de la s-ième colonne est définie comme suit : on a $t_{rr} = 1$, $t_{ij} = \lambda$ et $t_{rs} = 0$ si $r \neq s$ et $(r, s) \neq (i, j)$. Pour la première étape, on a ainsi
$$T_{21} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad T_{21}M = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}.$$

En posant,

$$T_{31} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}, \quad T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}, \quad T_{12} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad T_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix},$$

on obtient alors

$$T_{23}T_{12}T_{32}T_{31}T_{21}M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad T_{23}T_{12}T_{32}T_{31}T_{21} = \begin{pmatrix} 3 & 2 & 0 \\ 1 & 0 & 3 \\ 0 & 3 & 1 \end{pmatrix}.$$

Par ailleurs, la dilatation de rapport λ sur ℓ_i s'obtient en multipliant M à gauche par la matrice D_i dont l'élément de la r-ième ligne et de la s-ième colonne est 1 si $r = s$ et $r \neq i$, λ si $r = s = i$, et 0 sinon. En posant

$$D_1 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad D_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{on a} \quad D_2D_1 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

ce qui conduit à l'égalité

$$AM = I_3 \quad \text{avec} \quad A = \begin{pmatrix} 4 & 1 & 0 \\ 3 & 0 & 4 \\ 0 & 3 & 1 \end{pmatrix}.$$

Démonstration du théorème 7.1 : Soit A la matrice carrée extraite de G, de taille (k, k), dont l'élément de la i-ème ligne et la j-ième colonne est a_{ij} avec $1 \leq i, j \leq k$.

Supposons le code C systématique. Par définition, il existe une matrice génératrice M de C sous forme normalisée $\big(I_k|*\big)$. Notons ℓ_i le i-ème vecteur ligne de G et ℓ'_i celui de M. Le système $(\ell_i)_{1\leq i\leq k}$ est une base de C et pour tout i le vecteur ligne ℓ'_i appartient à C. Pour tout i entre 1 et k, il existe donc des éléments $\alpha_{ij} \in \mathbb{F}_q$ tels que l'on ait

$$\ell'_i = \sum_{j=1}^{k} \alpha_{ij}\ell_i.$$

On a alors l'égalité matricielle

$$\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1k} \\ \alpha_{21} & \dots & \alpha_{2k} \\ \vdots & \dots & \vdots \\ \alpha_{k1} & \dots & \alpha_{kk} \end{pmatrix} A = I_k,$$

ce qui prouve que A est inversible.

Inversement, supposons la matrice A inversible. Compte tenu de la proposition 7.8, on peut transformer A en la matrice identité I_k par des opérations élémentaires sur les lignes de A. Ces mêmes opérations élémentaires effectuées sur les lignes de G transforment donc G en une matrice de la forme $\big(I_k|*\big)$. D'après la proposition 7.7, c'est une matrice génératrice de G, donc G est systématique. Cela établit le théorème.

Exercice 4. On considère la matrice M à coefficients dans $\mathbb{F}_2$ définie par

$$M = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

1) Quel est le rang de M ?
2) En déduire que M est la matrice génératrice d'un code binaire C de longueur 6 et de dimension 4.
3) Le code C est-il systématique ?
4) Quelle est la distance minimale de C ?

9. Codes équivalents

Bien qu'il existe des codes linéaires non systématiques (le théorème 7.1 permet d'en expliciter facilement), on va voir maintenant que tout code linéaire est « équivalent » en un sens précis à un code systématique.

Soit n un entier ≥ 1. Définissons la notion d'équivalence entre deux codes linéaires de même longueur n. Soit $\mathbb{S}_n$ le groupe symétrique de l'ensemble $\Big\{1, \cdots, n\Big\}$ i.e. le groupe des bijections de cet ensemble. On dispose d'une application

$$\psi : \mathbb{S}_n \times \mathbb{F}_q^n \to \mathbb{F}_q^n,$$

définie par l'égalité

$$\psi\big((\sigma, (x_1, \cdots, x_n)\big) = \big(x_{\sigma(1)}, \cdots, x_{\sigma(n)}\big).$$

Soit C un code linéaire de longueur n sur $\mathbb{F}_q$. Pour tout $\sigma \in \mathbb{S}_n$, posons

$$\sigma(C) = \Big\{\psi\big((\sigma, x)\big) \;\Big|\; x \in C\Big\}.$$

C'est un sous-ensemble de $\mathbb{F}_q^n$.

Lemme 7.8. *L'ensemble $\sigma(C)$ est un code linaire sur $\mathbb{F}_q$ de longueur n et de même dimension que C.*

Démonstration : Tout d'abord 0 est dans $\sigma(C)$. Par ailleurs, quels que soient x et y dans C et $\lambda \in \mathbb{F}_q$, on a les égalités

$$\psi\big((\sigma, x)\big) + \psi\big((\sigma, y)\big) = \psi\big((\sigma, x+y)\big) \quad \text{et} \quad \psi\big((\sigma, \lambda x)\big) = \lambda\psi\big((\sigma, x)\big),$$

ce qui entraîne que $\sigma(C)$ est sous-espace vectoriel de $\mathbb{F}_q^n$. Le lemme en résulte, vu que l'application $C \to \sigma(C)$ qui à x associe $\psi\big((\sigma, x)\big)$ est un isomorphisme de $\mathbb{F}_q$-espaces vectoriels.

Lemme 7.9. *Soient G une matrice génératrice de C et σ un élément de $\mathbb{S}_n$. Pour tout $j = 1, \cdots, n$, soit c_j le j-ième vecteur colonne de G. Alors, la matrice de taille (k, n) dont le j-ième vecteur colonne est $c_{\sigma(j)}$, est une matrice génératrice de $\sigma(C)$.*

Démonstration : Soit k la dimension de C. Notons a_{ij} l'élément de i-ème ligne et de la j-ième colonne de G. Pour tout $i = 1, \cdots, k$ l'élément $\big(a_{i\sigma(1)}, a_{i\sigma(2)}, \cdots, a_{i\sigma(n)}\big)$ appartient à $\sigma(C)$ et le j-ième vecteur colonne de la matrice

$$\begin{pmatrix} a_{1\sigma(1)} & a_{1\sigma(2)} & \cdots & a_{1\sigma(n)} \\ \vdots & \vdots & \cdots & \vdots \\ a_{k\sigma(1)} & a_{k\sigma(2)} & \cdots & a_{k\sigma(n)} \end{pmatrix},$$

est $c_{\sigma(j)}$. On peut en extraire une matrice de taille (k, k) inversible, car tel est le cas de G, c'est donc une matrice de rang k. Par suite, c'est une matrice génératrice de $\sigma(C)$.

Définition 7.13. *Soient C_1 et C_2 deux codes linéaires sur $\mathbb{F}_q$ de longueur n. On dit qu'ils sont équivalents s'il existe $\sigma \in \mathbb{S}_n$ tel que l'on ait $\sigma(C_1) = C_2$.*

On définit ainsi une relation d'équivalence sur l'ensemble des codes linéaires sur $\mathbb{F}_q$ de longueur fixée (vérifier cette assertion en exercice).

Proposition 7.9. *Tout code linéaire est équivalent à un code systématique.*

Démonstration : Soit C un code linéaire sur $\mathbb{F}_q$ de longueur n et de dimension k. Soient G une matrice génératrice de C et c_j les vecteurs colonnes de G. Il existe une matrice de taille (k, k) extraite de G inversible. Notons $c_{j_1}, \cdots, c_{j_k}$ les vecteurs colonnes de cette matrice avec $j_1 < j_2 \cdots < j_k$. Soit σ l'élément de $\mathbb{S}_n$ défini par les égalités $\sigma(r) = j_r$ pour $1 \leq r \leq k$ et $\sigma(a) = a$ pour tout a compris entre $k+1$ et n. Compte tenu du lemme 7.9, la matrice dont les vecteurs colonnes sont les $c_{\sigma(j)}$ pour $j = 1, \cdots, n$, est une matrice génératrice de $\sigma(C)$, qui est donc un code systématique. D'où le résultat.

Exercice 5. Soit M la matrice à coefficients dans $\mathbb{F}_2$ définie par

$$M = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}.$$

1) Montrer que M définit un code linéaire binaire C de dimension 3 et de longueur 5.
2) Quelle est la distance minimum de C ?

3) Montrer que C n'est pas systématique.

4) Trouver un code systématique équivalent à C.

10. Matrices de contrôle

Considérons dans ce paragraphe un code linéaire C sur $\mathbb{F}_q$ de longueur n et de dimension k. On se préoccupe ici du problème d'expliciter une matrice H ayant $n-k$ lignes et n colonnes dont le noyau soit précisément les éléments de C[45]. Étant donné un élément $x \in \mathbb{F}_q^n$, une telle matrice permet ainsi de contrôler si x est dans C ou non. Il suffit en effet, de calculer $H(x)$ et de constater si cet élément est nul ou non.

Tout d'abord, il existe de telles matrices. En effet, C étant un sous-espace vectoriel de $\mathbb{F}_q^n$ de dimension k, C est l'intersection de $n-k$ hyperplans de $\mathbb{F}_q^n$. Autrement dit, il existe $n-k$ formes linéaires $f_i : \mathbb{F}_q^n \to \mathbb{F}_q$ telles que C soit l'ensemble des $x \in \mathbb{F}_q^n$ vérifiant les égalités $f_1(x) = \cdots = f_{n-k}(x) = 0$. L'application $\psi : \mathbb{F}_q^n \to \mathbb{F}_q^{n-k}$ définie par

$$\psi(x) = \big(f_1(x), \cdots, f_{n-k}(x)\big)$$

est alors une application $\mathbb{F}_q$-linéaire de noyau C[46]. Par ailleurs, l'égalité

$$n = \dim C + \dim \psi(\mathbb{F}_q^n)$$

entraîne que l'image de ψ est de dimension $n-k$, par suite, ψ est une application surjective sur $\mathbb{F}_q^{n-k}$. Tout cela justifie la définition suivante :

Définition 7.14. *On appelle matrice de contrôle de C toute matrice ayant $n-k$ lignes et n colonnes à coefficients dans $\mathbb{F}_q$ de noyau C.*

Remarque 7.2. Comme on le signalait ci-dessus, le rang d'une matrice de contrôle de C est $n-k$ i.e. l'application $\mathbb{F}_q$-linéaire $\mathbb{F}_q^n \to \mathbb{F}_q^{n-k}$ qui lui correspond est surjective.

---

[45] On identifie ici la matrice H avec l'application linéaire de $\mathbb{F}_q^n$ à valeurs dans $\mathbb{F}_q^{n-k}$ qui est représentée par H dans les bases canoniques de $\mathbb{F}_q^n$ et $\mathbb{F}_q^{n-k}$. Cette identification sera faite implicitement dans la suite.

[46] On peut aussi procéder autrement en utilisant la notion d'espace vectoriel quotient. Considérons en effet, un espace vectoriel E sur un corps K et F un sous-espace vectoriel de E. On a défini dans le chapitre II le groupe quotient E/F. On peut de plus ici munir E/F d'une structure de K-espace vectoriel (comme d'ailleurs pour les algèbres de polynômes), en considérant comme loi externe l'application $K \times E/F \to E/F$ qui au couple $(\lambda, x+F)$ associe $\lambda x + F$. On vérifie qu'elle est bien définie et l'on peut ainsi parler du K-espace vectoriel E/F. Supposons de plus E de dimension finie sur K. Alors, E/F est aussi de dimension finie sur K et sa dimension est $\dim E - \dim F$. La surjection canonique $E \to E/F$ qui à un élément $x \in E$ associe sa classe modulo F est K-linéaire surjective de noyau F. Dans notre situation, il suffit alors de considérer l'application $\mathbb{F}_q$-linéaire $\mathbb{F}_q^n \to \mathbb{F}_q^n/C$ et de la composer avec un isomorphisme de $\mathbb{F}_q^n/C$ sur $\mathbb{F}_q^{n-k}$ pour obtenir une application $\mathbb{F}_q$-linéaire $\mathbb{F}_q^n \to \mathbb{F}_q^{n-k}$ de noyau C.

Lemme 7.10. *Soit H une matrice à coefficients dans $\mathbb{F}_q$, ayant $n-k$ lignes et n colonnes, et de rang maximum $n-k$. Alors, le noyau de H est un code sur $\mathbb{F}_q$ de longueur n et de dimension k, et H en est une matrice de contrôle.*

Démonstration : Le noyau de H est un sous-espace vectoriel de $\mathbb{F}_q^n$. Puisque le rang de H est $n-k$, la dimension de son noyau est k, d'où l'assertion.

Étant donnée une matrice H à coefficients dans $\mathbb{F}_q$ de taille $(n-k, n)$, et de rang $n-k$, l'énoncé suivant permet de tester si H est une matrice de contrôle de C.

Proposition 7.10. *Soit G une matrice génératrice de C. Soit H une matrice à coefficients dans $\mathbb{F}_q$ ayant $n-k$ lignes et n colonnes, et de rang $n-k$. Alors, H est une matrice de contrôle de C si et seulement on a l'égalité*

$$H\,{}^tG = 0,$$

où ${}^tG$ est la matrice transposée de G.

Démonstration : Supposons que l'on ait $H\,{}^tG = 0$. Pour tout vecteur colonne c de ${}^tG$, on a $H(c) = 0$. Puisque les vecteurs colonnes de ${}^tG$ forment une base de C, il en résulte que C est contenu dans le noyau de H. Le rang de H étant $n-k$, son noyau est donc de dimension k, égale à celle de C. Par suite, C est le noyau de H. Inversement, supposons que C soit le noyau de H. On a alors $H(c) = 0$ pour tout vecteur colonne c de ${}^tG$, d'où $H\,{}^tG = 0$ et le résultat.

Connaissant une matrice de contrôle de C, l'énoncé qui suit précise comment déterminer une matrice de contrôle d'un code équivalent à C.

Lemme 7.11. *Soient H une matrice de contrôle de C et c_j le j-ième vecteur colonne de H $(1 \leq j \leq n)$. Soient σ un élément de $\mathbb{S}_n$ et H' la matrice de taille $(n-k, n)$ dont le j-ième vecteur colonne est $c_{\sigma(j)}$. Alors, H' est une matrice de contrôle de $\sigma(C)$.*

Démonstration : Posons $H = \big(h_{ij}\big)$ avec $1 \leq i \leq n-k$ et $1 \leq j \leq n$. Soit $G = \big(a_{ij}\big)$, avec $1 \leq i \leq k$ et $1 \leq j \leq n$ une matrice génératrice de C. On a l'égalité $H\,{}^tG = 0$ (prop. 7. 10), autrement dit, on a

$$\sum_{r=1}^{n} h_{ir} a_{jr} = 0 \quad \text{pour} \quad i = 1, \cdots, n-k,\ j = 1, \cdots, k.$$

Puisque σ est une bijection de $\big\{1, \cdots, n\big\}$, cette égalité s'écrit aussi

$$\sum_{r=1}^{n} h_{i\sigma(r)} a_{j\sigma(r)} = 0. \tag{3}$$

Soit alors G' la matrice de taille (k, n) dont l'élément de la i-ème ligne et de la j-ième colonne est $a_{i\sigma(j)}$. D'après le lemme 7.9, G' est une matrice génératrice de $\sigma(C)$. Par ailleurs, H étant de rang $n-k$, il en est de même de H', vu que H' se déduit de H par une permutation des colonnes de H. L'élément de la i-ème ligne et de la j-ième colonne de H' est est $h_{i\sigma(j)}$. D'après l'égalité (3) on a donc $H'^t G' = 0$, d'où le résultat (prop. 7.10).

On va s'intéresser maintenant à la détermination d'une matrice de contrôle de C à partir d'une matrice génératrice. On est pour cela amené à considérer deux cas, selon que C soit systématique ou non.

1) Cas où C est systématique :

Proposition 7.11. *Supposons C systématique. Soit $G = (I_k|B)$ la matrice génératrice normalisée de C. Alors, la matrice*

$$H = \left(-{}^tB|I_{n-k}\right)$$

est une matrice de contrôle de C.

Démonstration : Le rang de H est $n-k$. Posons $B = \left(b_{rs}\right)$ avec $1 \leq r \leq k$ et $1 \leq s \leq n-k$. On vérifie alors que pour tous i et j, l'élément de la i-ème ligne et de la j-ième colonne de la matrice $H^t G$ est $-b_{ji} + b_{ji} = 0$, d'où $H^t G = 0$ et l'assertion (prop. 7.10).

Connaissant une matrice génératrice M de C, ce résultat permet alors d'obtenir une de ses matrices contrôle. Il suffit en effet de déterminer la matrice normalisée de C, ce que l'on peut faire en effectuant des opérations élémentaires sur les lignes de M, comme on l'a expliqué au paragraphe 8.

2) Supposons C non systématique. On procède comme suit. D'après la proposition 7.9, il existe $\sigma \in \mathbb{S}_n$ tel que $\sigma(C)$ soit un code systématique. Posons $C' = \sigma(C)$. En déterminant la matrice normalisée de C', on obtient, compte tenu de la proposition 7.11, une matrice de contrôle H' de C'. Soit c'_j le j-ième vecteur colonne de H' $(1 \leq j \leq n)$. L'égalité $C = \sigma^{-1}(C')$ et le lemme 7.11 (appliqué avec C') entraînent alors que la matrice de taille $(n-k, k)$, dont le j-ième vecteur colonne est $c'_{\sigma^{-1}(j)}$, est une matrice de contrôle de C.

Exercice 6. Soit M la matrice à coefficients dans $\mathbb{F}_3$ définie par l'égalité

$$M = \begin{pmatrix} 1 & 2 & 0 & 1 & 1 & 2 \\ 2 & 1 & 0 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

1) Déterminer le rang de M et en déduire que M est la matrice génératrice d'un code linéaire C sur $\mathbb{F}_3$ de longueur 6 et de dimension 3.

2) Déterminer une matrice de contrôle de C.

Le résultat suivant est une réciproque de la proposition 7.11 :

Proposition 7.12. *Supposons qu'il existe une matrice A à coefficients dans $\mathbb{F}_q$, ayant $n-k$ lignes et k colonnes, telle que $\big(A|I_{n-k}\big)$ soit une matrice de contrôle de C. Alors, C est systématique et la matrice génératrice normalisée de C est $\big(I_k|-{}^tA\big)$.*

Démonstration : Posons $G=\big(I_k|-{}^tA\big)$ et $H=\big(A|I_{n-k}\big)$. Notons a_{rs} l'élément de r-ième ligne et de la s-ième colonne de A (on a $1\leq r\leq n-k$ et $1\leq s\leq k$). On vérifie que le coefficient de la i-ème ligne et de la j-ième colonne de la matrice $H{}^tG$ est $a_{ij}-a_{ij}=0$. Ainsi, les vecteurs colonnes de ${}^tG$ sont dans le noyau de H. Ils appartiennent donc à C. Par ailleurs, le rang de ${}^tG$, qui est celui de G, est k. Les k vecteurs colonnes de ${}^tG$ sont donc indépendants. Puisque C est de dimension k, on en déduit qu'ils forment une base de C, i.e. que G est une matrice génératrice de C.

11. Applications des matrices de contrôle

Nous allons voir que la connaissance d'une matrice de contrôle d'un code linéaire permet de calculer directement sa distance minimum, et fournit un algorithme de décodage.

Considérons un code linéaire C sur $\mathbb{F}_q$ de type (n,k,d).

1. Détermination de la distance minimum

La distance minimum d de C se déduit d'une de ses matrices de contrôle comme suit :

Proposition 7.13. *Soit H une matrice de contrôle de C. L'entier d est égal au nombre minimum de vecteurs colonnes de H qui, en tant que vecteurs de $\mathbb{F}_q^{n-k}$, sont linéairement dépendantes.*

C'est une conséquence de l'énoncé suivant :

Lemme 7.12. *Soit r un entier ≥ 1.*

1) *S'il existe dans C un mot de poids r, alors il existe r colonnes de H linéairement dépendantes.*
2) *S'il existe r colonnes de H linéairement dépendantes, alors il existe dans C un mot de poids r' avec $1\leq r'\leq r$.*

Démonstration : 1) Soit c_j la j-ème colonne de H ($1\leq j\leq n$). Par hypothèse, il existe un mot $x=(x_1,\cdots,x_n)\in C$ de poids r. Notons $x_{i_1},\cdots,x_{i_r}$ les r composantes non nulles de x. On a les égalités

$$H(x)=\sum_{j=1}^{n}x_jc_j=\sum_{j=1}^{r}x_{i_j}c_{i_j}=0.$$

Par suite, les r colonnes $c_{i_1}, \cdots, c_{i_r}$ sont $\mathbb{F}_q$-dépendantes, d'où l'assertion.

2) Considérons r colonnes de H, $c_{i_1}, \cdots, c_{i_r}$ linéairement dépendantes. Il existe des éléments $\lambda_{i_1}, \cdots \lambda_{i_r}$ de $\mathbb{F}_q$, qui ne sont pas tous nuls, tels que l'on ait

$$\sum_{j=1}^{r} \lambda_{i_j} c_{i_j} = 0.$$

Soit x l'élément de $\mathbb{F}_q^n$ dont la composante d'indice i_j est λ_{i_j} pour $j = 1, \cdots, r$ et dont toutes les autres composantes sont nulles. L'égalité

$$H(x) = \sum_{j=1}^{r} \lambda_{i_j} c_{i_j}$$

entraîne alors $H(x) = 0$, donc x appartient à C. Le nombre de ses composantes non nuls est égal aux nombres d'éléments λ_{i_j} non nuls, donc le poids r' de x est au plus r. Par ailleurs, on a $r' \geq 1$ vu qu'il existe un λ_{i_j} non nul. D'où le lemme.

Démonstration de la proposition 7.13. : Il existe dans C un mot de poids d. On a $d \geq 1$ car C est non nul (par définition). D'après l'assertion 1 du lemme ci-dessus, il existe d colonnes de H linéairement dépendantes. Supposons alors qu'il existe r colonnes de H dépendantes, avec $1 \leq r < d$. D'après l'assertion 2 du lemme, il existe dans C un mot de poids r' avec $1 \leq r' \leq r$. En particulier, on a $1 \leq r' < d$, d'où une contradiction et le résultat.

Remarque 7.3. il résulte de la proposition 7.13 que tout système de $d-1$ vecteurs colonnes de H est libre.

Exercice 7. Déterminer la distance minimum du code linéaire sur $\mathbb{F}_3$ défini par la matrice M dans l'exercice 6.

Exemple 7.7. Code de Hamming binaire de longueur $2^r - 1$

Soient r un entier ≥ 2 et α un générateur du groupe multiplicatif $\mathbb{F}_{2^r}^*$ du corps $\mathbb{F}_{2^r}$ à 2^r éléments. Posons $n = 2^r - 1$ et considérons l'application $f : \mathbb{F}_2^n \to \mathbb{F}_{2^r}$ définie par l'égalité

$$f\big((x_1, \cdots, x_n)\big) = \sum_{j=1}^{n} x_j \alpha^j.$$

C'est une application $\mathbb{F}_2$-linéaire. Soit C le noyau de f. C'est un code linéaire sur $\mathbb{F}_2$, appelé code de Hamming. On va démontrer en application de ce qui précède le résultat suivant :

Lemme 7.13. *Le code C est un code linéaire sur $\mathbb{F}_2$ de type $(n, n-r, 3)$. De plus, C est un code parfait.*

Démonstration : La longueur de C est n vu que C est un sous-espace vectoriel de $\mathbb{F}_2^n$. Par ailleurs, on a

$$\mathbb{F}_{2^r}^* = \left\{\alpha, \cdots, \alpha^n\right\},$$

d'où l'on déduit que f est une application surjective. Puisque $\mathbb{F}_{2^r}$ est un espace vectoriel sur $\mathbb{F}_2$ de dimension r, il en résulte que l'on a $n = \dim C + r$, d'où $\dim C = n - r$. (Notons que r étant au moins 2, on a $n - r \neq 0$). Vérifions que la distance minimum d de C est égale à 3. Supposons que C contienne un mot $x = (0, \cdots, 0, 1, 0, \cdots 0)$ de poids 1. Supposons que la composante égale à 1 soit en m-ième position. Dans ce cas, on a $f(x) = \alpha^m = 0$, ce qui conduit à une contradiction car $\alpha \neq 0$. Supposons qu'il existe dans C un mot de poids 2. Il existe alors deux entiers i et j tels que $1 \leq i < j \leq n$ et que $\alpha^i + \alpha^j = 0$. On a donc l'égalité $\alpha^{j-i} = 1$ et cela contredit le fait que α soit un élément d'ordre n de $\mathbb{F}_{2^r}^*$. On a donc $d \geq 3$. Soit H une matrice de contrôle de C. Elle possède n colonnes et r lignes. Puisque l'on a $d \geq 3$, il résulte de la proposition 7.13 que toutes les colonnes de H sont non nulles et qu'elles sont distinctes deux à deux. Ainsi, les colonnes de H sont formées de tous les vecteurs colonnes non nuls à r lignes à coefficients dans $\mathbb{F}_2$. Si c_i et c_j sont deux colonnes distinctes de H, alors $c_i + c_j$ est aussi une colonne de H. En particulier, les colonnes c_i, c_j et $c_i + c_j$ sont linéairement dépendantes. On en déduit que $d = 3$ comme annoncé (prop. 7.13).

Il reste à démontrer que C est un code parfait. La capacité de correction de C est $t = 1$. Par ailleurs, pour tout $x \in C$, le cardinal de la boule de Hamming $B(x, t)$ est $n + 1$ (cf. prop. 7.2). Le cardinal de C étant 2^{n-r}, l'égalité

$$2^{n-r}(n+1) = 2^n$$

entraîne alors l'assertion (cor. 7.3) et le résultat.

2. Algorithme de décodage

Soit H une matrice de contrôle de C. Supposons que l'on transmette un mot x de C et que le message reçu y soit affecté de r erreurs. On a alors $y = x + e_y$, où $d_H(y, x) = r$, autrement dit, e_y est un mot de $\mathbb{F}_q^n$ de poids r. Soit t la capacité de correction de C. Dans le cas où l'on a $r \leq t$, on va voir ici un algorithme permettant de décoder y en retrouvant x, qui est l'unique mot de code dans la boule $B(y, t)$ (lemme 7.1 et cor. 7.1). Le principe de décodage est le suivant :

1) on commence par dresser la liste de tous les éléments $e \in \mathbb{F}_q^n$ tels que $w(e) \leq t$ i.e. de poids $\leq t$.
2) On calcule ensuite les images, on dit aussi les syndromes, $H(e)$ des éléments $e \in \mathbb{F}_q^n$ tels que $w(e) \leq t$ [47].

3) On détermine $H(y)$. Il se présente deux cas :

3.1) Il existe $e \in \mathbb{F}_q^n$ tel que $w(e) \leq t$ et $H(y) = H(e)$, ce que les calculs précédents permettent de vérifier (un tel élément e est alors unique). On a $H(y-e) = 0$, l'élément $y - e$ appartient donc à C et l'on décode alors y par $y - e$. En effet, $y - e$ est l'unique mot de C dans la boule $B(y, t)$ (on a $d_H(y - e, y) = d_H(e, 0) \leq t$), d'où $x = y - e$.

3.2) Pour tout $e \in \mathbb{F}_q^n$ tel que $w(e) \leq t$, on a $H(y) \neq H(e)$. Dans ce cas, y est affecté de plus de t erreurs. En effet, supposons qu'il existe un mot de code z tel que $d_H(z, y) \leq t$. On a alors $y = z + e$ avec $w(e) \leq t$ et $H(y) = H(e)$. Dans cette situation, on ne peut pas en général effectuer un décodage correct de y.

Remarque 7.4. Supposons que C soit un code sur $\mathbb{F}_2$ et que l'on ait $t = 1$. Dans l'algorithme de décodage de C, le calcul des syndromes $H(e)$ où e est un élément de $\mathbb{F}_2^n$ de poids au plus 1 (il y en a $n + 1$), se lit alors directement sur les colonnes de H. Plus précisément, si $e = (0, \cdots, 0, 1, 0, \cdots, 0)$, où la composante égale à 1 est en i-ème position, alors on a $H(e) = c_i$, où c_i est le i-ème vecteur colonne de H.

Exemple 7.8. Considérons la matrice H à coefficients dans $\mathbb{F}_2$ définie par l'égalité

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Elle est de rang 3, c'est donc la matrice de contrôle d'un code binaire C sur $\mathbb{F}_2$, de dimension 4, à savoir $C = \mathrm{Ker}(H)$ i.e. C est le noyau de l'application linéaire $\mathbb{F}_2^7 \to \mathbb{F}_2^3$ représentée par H dans les bases canoniques (cf. lemme 7.10). Puisque tous les vecteurs colonnes non nuls de $\mathbb{F}_2^3$ interviennent dans H, la distance minimum de C est 3 et sa capacité de correction est $t = 1$ (cf. l'exemple 7.7). Supposons que le mot reçu soit $y = (1, 1, 0, 1, 1, 1, 1) \in \mathbb{F}_2^7$. On vérifie que l'on a $H(y) = (1, 1, 0) \in \mathbb{F}_2^3$, qui est le troisième vecteur colonne de H. On a donc $H(y) = H(e)$ avec $e = (0, 0, 1, 0, 0, 0, 0)$. Par suite, on décode y par $y - e = (1, 1, 1, 1, 1, 1, 1)$.

[47] Notons que, pour un tel élément $e \in \mathbb{F}_q^n$, la connaissance de $H(e)$ détermine e. En effet, soient e et e' deux éléments de $\mathbb{F}_q^n$ tels que $w(e) \leq t$, $w(e') \leq t$ et $H(e) = H(e')$. On a les inégalités $w(e - e') = d_H(e, e') \leq d_H(e, 0) + d_H(e', 0)) \leq 2t \leq d - 1$. Par ailleurs, on a $H(e - e') = 0$, donc $e - e'$ appartient à C, d'où $e = e'$. En particulier, les syndromes $H(e)$ des éléments $e \in \mathbb{F}_q^n$ de poids au plus t sont distincts deux à deux.