ARITHMÉTIQUE, GROUPES ET PROBABILITÉS (INFO S4, MATHS)

ADRIEN BROCHIER

TABLE DES MATIÈRES

1. ARITHMÉTIQUE

L'arithmétique est l'étude des nombres entiers. On note $\mathbb{Z}$ l'ensemble des entiers relatifs, et $\mathbb{N}$ l'ensemble des entiers naturels, càd positifs ou nuls. Si $a \in \mathbb{Z}$, on note $|a|$ sa valeur absolue qui est donc dans $\mathbb{N}$.

1.1. **Divisibilité.**

Définition 1.1. *Soient $a, b \in \mathbb{Z}$. On dit que a* divise *b, ou que a* est un diviseur de *b, ou que b* est un multiple de *a, et on note $a|b$ si il existe $k \in \mathbb{Z}$ tel que*

$$b = k \times a.$$

Proposition 1.2. *Soient $a, b, c \in \mathbb{Z}$.*

(1) $1|a$.

Date: 5 mai 2021.

(2) $a|0$.
(3) $a|a$.
(4) *Si* $a|b$ *et* $b|c$, *alors* $a|c$.
(5) *Si* $a|b$ *et* $b|a$ *alors* $|a| = |b|$, *c'est-à-dire* $a = \pm b$.
(6) *Si* $a|b$ *et* $a|c$ *alors* $a|(kb + lc)$ *pour tous* $k, l \in \mathbb{Z}$.

Démonstration. Exercice. □

Définition 1.3. *Un nombre premier est un entier naturel qui a exactement deux diviseurs positifs, 1 et lui-même. Un nombre qui n'est pas premier est appelé nombre composé.*

Attention : 1 n'est donc pas un nombre premier. 2, 3, 5, 7, 13 et 34565989798090989829 sont des nombres premiers. 6 n'est pas premier car $2|6$.

1.2. **Division euclidienne.**

Proposition 1.4. *Soient* $a, b \in \mathbb{N}$, $b \neq 0$. *Alors il existe un unique entier naturel* q, *le quotient de* a *par* b, *et un unique entier naturel* r, *le reste de la division de* a *par* b, *tels que*

(1) $a = bq + r$
(2) $0 \leq r < b$.

Démonstration. Si $b > a$, alors on pose $q = 0$ et $r = a$. Sinon, on pose

$$q = \max\{k \in \mathbb{N}, kb \leq a\},$$

et $r = a - bq$. Puisque par construction $bq \leq a < b(q+1)$, en soustrayant bq partout on a bien que $0 \leq r < b$.

Prouvons l'unicité. Supposons que (q', r') satisfait aussi ces conditions. On a donc

$$bq + r - (bq' + r') = a - a = 0$$

donc

$$b(q' - q) = r' - r.$$

Sans perte de généralité on peut supposer $r \leq r'$. On a donc que $0 \leq r' - r < b$, et par le calcul précédent b divise $r' - r$. Donc $r' - r = 0$, impliquant $r = r'$ et finalement $q = q'$.

□

Remarque 1.5. On peut facilement étendre ce théorème au cas ou a et b sont des entiers relatifs. Il faut simplement faire attention à ce que le reste soit positif, par exemple la division de -11 par 3 donne $-11 = -4 \times 3 + 1$.

Remarque 1.6. On a que $b|a$ si et seulement si le reste de la division de a par b est 0.

1.3. **Application : écriture en base N.** L'écriture habituelle des nombres est ce qu'on appelle écriture en base 10, c'est-à-dire qu'on utilise 10 symboles pour les représenter (les chiffres de 0 à 9), et par exemple l'écriture 213 représente le nombre

$$3 \times 10^0 + 1 \times 10^1 + 2 \times 10^2.$$

Théorème 1.7. *Soit $N \geq 2$ un entier, tout entier naturel n s'écrit de façon unique*

$$n = i_0 + i_1 N + i_2 N^2 \cdots + i_k N^k \tag{1.1}$$

avec $0 \leq i_k, \ldots, i_0 < N$ et $i_k \neq 0$.

Démonstration. On fait la division euclidienne de n par N

$$n = Nq_0 + r_0$$

et on pose $i_0 = r_0$. Si $q_0 = 0$ c'est terminé. Sinon on fait ensuite la division euclidienne de q_0 par N

$$q_0 = Nq_1 + r_1$$

et on pose $i_1 = r_1$. Par construction, on a

$$n = N(Nq_1 + r_1) + r_0 = N^2 q_1 + Nr_1 + r_0,$$

donc si $q_1 = 0$ on a fini. Sinon on fait la division euclidienne de q_1 par N et ainsi de suite. Comme les q_i sont positifs et de plus en plus petit, ce processus s'arrête au bout d'un moment. Prouvons maintenant l'unicité : supposons qu'il existe une autre écriture de n sous cette forme

$$n = i_0 + i_1 N \cdots + i_k N^k = i'_0 + i'_1 N \cdots + i'_{k'} N^{k'}.$$

Sans perte de généralité on peut supposer que $i_0 \leq i'_0$. On a alors que N divise $i'_0 - i_0$, et par ailleurs $0 \leq i'_0 - i_0 < N$ donc forcément $i'_0 = i_0$. On peut donc simplifier par i_0 des deux côtés, diviser par N et appliquer le même argument. □

Remarque 1.8. Si $N = 10$ alors $(i_k, \ldots, i_0)$ sont par définition juste les chiffres de l'écriture habituelle de n. En général, on appelle la liste $(i_k, \ldots, i_0)$ *l'écriture en base N* de n. On l'écrit souvent $(i_k i_{k-1} \ldots i_0)_N$. Si $N > 9$, pour éviter toute ambiguïté on choisit des symboles pour représenter les nombres de 10 à N (voir exemple ci-dessous).

Inversement, si on se donne l'écriture d'un nombre en base N, il suffit pour l'écrire en base 10 de calculer (1.1).

Deux bases sont fondamentales en informatique : la base 2, dite binaire, et la base 16, dite hexadécimale. Dans cette dernière on utilise comme "chiffre" les chiffres de 0 à 9, puis les lettres de A à F.

Exemple 1.9. On a $213 = 106 \times 2 + 1$, $106 = 2 \times 53 + 0$, $53 = 2 \times 26 + 1$, $26 = 2 \times 13 + 0$, $13 = 2 \times 6 + 1$, $6 = 2 \times 3 + 0$, $3 = 2 \times 1 + 1$ et finalement $1 = 0 \times 2 + 1$. Donc

$$213 = (11010101)_2.$$

Exemple 1.10. On a $213 = 13 \times 16 + 5$, et $13 = 0 \times 16 + 13$. Donc
$$213 = (D5)_{16}.$$

Remarque 1.11. Avec deux "chiffres" hexadécimaux, on peut représenter les nombres de 0 à $16^2 - 1 = 2^8 - 1 = 255$, c'est-à-dire les valeurs possibles pour ce qu'on peut stocker dans un octet, un bloc de 8 bits.

1.4. **PGCD, PPCM.**

Définition 1.12. *Soient $a, b \in \mathbb{Z}$ non nuls. On dit que $d \in \mathbb{Z}$ est un diviseur commun de a et b si $d|a$ et $d|b$. On dit que $m \in \mathbb{Z}$ est un multiple commun de a et b si $a|m$ et $b|m$.*

Exemple 1.13. Les diviseurs communs de 60 et 90 sont 1, 2, 3, 5, 6, 10, 15 et 30.

-1 et 1 sont toujours des diviseurs communs de a et b, et le produit ab est toujours un multiple commun de a et b. L'ensemble des diviseurs et multiples communs de deux entiers donnés n'est donc jamais vide. Si c est un diviseur commun de a, b, alors $|c| \leq |a|$ et $|c| \leq |b|$, l'ensemble des diviseurs communs est borné et donc fini, donc il en existe un plus grand. De même, l'ensemble des multiples communs *positifs* est minoré, donc il en existe un plus petit.

Définition 1.14. *On appelle plus grand diviseur commun de a et b, et on note $\mathrm{PGCD}(a, b)$ ou $a \wedge b$, le plus grand des diviseurs commun de a et b. On appelle plus petit multiple commun de a et b, et on note $\mathrm{PPCM}(a, b)$ ou $a \vee b$ le plus petit des multiples communs* positifs *de a et b.*

Remarque 1.15. Si disons b est nul, alors $\mathrm{PGCD}(a, 0) = a$. Si a et b sont tous les deux nuls, alors le PGCD n'est pas défini.

Définition 1.16. *Soient a, b non nuls. On dit que a et b sont* premiers entre eux *si $\mathrm{PGCD}(a, b) = 1$.*

Lemme 1.17. *Soient $a, b \in \mathbb{Z}$ et $d = \mathrm{PGCD}(a, b)$. Alors $a' = a/d$ et $b' = b/d$ sont premiers entre eux.*

Démonstration. Soit d' un diviseur commun positif de a' et b'. Puisque $d'|a'$, alors $dd'|a$, et de même $dd'|b$. Mais puisque d' est positif, on a $dd' \geq d$, or d est le plus

grand diviseur commun donc on a forcément $dd' = d$, donc $d' = 1$ puisque il est positif. □

1.5. **Calcul du PGCD, théorème de Bezout.** Une technique efficace pour calculer le PGCD de deux entiers repose sur le lemme suivant :

Lemme 1.18. *Soient $a, b \in \mathbb{N}$ non nuls et r le reste de la division de a par b. Alors*

$$\mathrm{PGCD}(a, b) = \mathrm{PGCD}(b, r).$$

Démonstration. Soit (q, r) le résultat de la division euclidienne de a par b, et d un diviseur commun de a et b. Alors d divise $r = a - bq$, donc c'est un diviseur commun de b et r. Réciproquement, si d est un diviseur commun de b et r, alors d divise bq donc divise $a = bq + r$. On en conclut que l'ensemble des diviseurs communs de a et b, est le même que l'ensemble des diviseurs communs de b et r. En particulier $\mathrm{PGCD}(a, b) = \mathrm{PGCD}(b, r)$. □

Ce Lemme donne une méthode pratique pour calculer le PGCD de deux entiers, qu'on appelle l'algorithme d'Euclide. Pour calculer $\mathrm{PGCD}(a, b)$, on calcule le reste r_1 de la division de a par b. Si $r_1 = 0$, alors c'est que $b|a$ donc $\mathrm{PGCD}(a, b) = b$. Sinon, on calcule le reste r_2 de la division de b par r_1. Si r_2 est nul c'est que r_1 divise b, donc $\mathrm{PGCD}(a, b) = \mathrm{PGCD}(b, r_1) = r_1$. Sinon on recommence en divisant r_1 par r_1 etc... Le PGCD de a et b est le dernier reste non nul dans cette succession de division.

Exemple 1.19. Calculons le PGCD de 21 et 15. On a :

$$21 = 15 \times 1 + 6$$
$$15 = 6 \times 2 + 3$$
$$6 = 3 \times 2 + 0.$$

Le PGCD de 21 et 15 est le dernier reste non nul, c'est-à-dire dans ce cas 3.

L'algorithme d'Euclide nous donne en fait un résultat beaucoup plus puissant.

Théorème 1.20 (Bezout)**.** *Soient a, b des entiers relatifs qui ne sont pas tous les deux nuls. Alors il existe des entiers relatifs u, v tels que*

$$ua + vb = \mathrm{PGCD}(a, b).$$

Nous allons donner une preuve constructive de ce théorème, c'est-à-dire que nous allons montrer que l'algorithme d'Euclide permet de calculer efficacement une paire (u, v) qui fait marcher le théorème.

Démonstration. Informellement, l'idée consiste à appliquer l'algorithme d'Euclide, puis à "remonter" à partir de la fin.

Remarquons que quitte à remplacer u par $-u$ et/ou v par $-v$ on peut supposer que a, b sont positifs. On peut aussi supposer que $a \geq b$. Comme dans l'algorithme d'Euclide on écrit $a = bq_1 + r_1$. Si $r_1 = 0$, c'est que $b|a$ et donc on peut écrire

$$\mathrm{PGCD}(a, b) = b = 0 \times a + 1 \times b.$$

sinon, on écrit $b = r_1 q_2 + r_2$. Si $r_2 = 0$, alors $\mathrm{PGCD}(a, b) = r_1$ et

$$r_1 = 1 \times a + -q_1 \times b.$$

Sinon, on écrit $r_1 = r_2q_3 + r_3$. Si $r_3 = 0$, alors $\mathrm{PGCD}(a,b) = r_2$ et

$$\begin{aligned} r_2 &= b - r_1q_2 \\ &= b - (a - bq_1)q_2 \\ &= -q_2 \times a + (1 + q_1) \times b \end{aligned}$$

Sinon on refait une division euclidienne etc... □

Exemple 1.21. Reprenons l'exemple de 21 et 15.

$$\begin{aligned} 3 &= 15 - 6 \times 2 \\ &= 15 - (21 - 15 \times 1) \times 2 \\ &= -2 \times 21 + 3 \times 15 \end{aligned}$$

Nous allons voir quelques conséquences importantes du Théorème de Bezout, qui seraient pénible à démontrer directement.

Proposition 1.22. *Soient $a, b \in \mathbb{Z}$. Alors a et b sont premiers entre eux si et seulement si il existe $u, v \in \mathbb{Z}$ tels que*

$$au + bv = 1.$$

Démonstration. Si a, b sont premiers entre eux alors u, v existent d'après Bezout. Réciproquement, si une telle paire $(u, v) \in \mathbb{Z}$ existe, alors puisque $\mathrm{PGCD}(a, b)$ divise a et b, il divise $au + bv$ et donc il divise 1. Donc $\mathrm{PGCD}(a, b) = 1$. □

Le lemme suivant dit que le PGCD de deux nombres n'est pas seulement le plus grand "en taille", mais aussi le plus grand pour la relation "divise" ce qui est beaucoup plus fort (et un peu étonnant quand on y pense, en général si on prend deux diviseurs communs de a et b, il n'y a pas de raisons que l'un divise l'autre).

Lemme 1.23. *Soient $a, b \in \mathbb{Z}$, et $c \in \mathbb{Z}$ un diviseur commun de a et b. Alors c divise $\mathrm{PGCD}(a, b)$.*

Démonstration. D'après Bezout, il existe $u, v \in \mathbb{Z}$ tels que

$$\mathrm{PGCD}(a, b) = ua + vb.$$

Puisque c divise a et b, il divise $ua + vb$, donc il divise $\mathrm{PGCD}(a, b)$. □

Cette propriété caractérise le PGCD : $d = \mathrm{PGCD}(a, b)$ si et seulement si d est un diviseur commun de a et b et tous les diviseurs commun de a et b divisent d.

Lemme 1.24. *Soient $a, b, c \in \mathbb{Z}$. Alors $\mathrm{PGCD}(ca, cb) = |c| \mathrm{PGCD}(a, b)$.*

Démonstration. Exercice. □

Lemme 1.25 (Gauss)**.** *Soient $a, b \in \mathbb{Z}$ et c un diviseur du produit ab. Si c est premier avec a, alors c divise b.*

Démonstration. D'après Bezout, puisque par hypothèse $\mathrm{PGCD}(a, c) = 1$, il existe u, v tels que

$$1 = ua + vc.$$

En multipliant tout par b, on obtient

$$b = uab + vcb.$$

Par hypothèse, c divise le produit ab, donc divise aussi uab. Par ailleurs c divise évidemment vcb, donc c divise b. □

Corollaire 1.26 (Lemme d'Euclide)**.** *Soient $a, b \in \mathbb{Z}$ et p un nombre premier. Si $p|ab$, alors $p|a$ ou $p|b$.*

Démonstration. De deux choses l'une :
- Soit p divise a et on a gagné.
- Soit p ne divise pas a, dans ce cas a et p sont premiers entre eux (en effet les seuls diviseurs positifs de p sont 1 et p par définition, donc si p ne divise pas a, leur seul diviseur positif commun est 1). Le Lemme de Gauss implique donc que p divise b.

□

Corollaire 1.27. *Soient $a, b, c \in \mathbb{Z}$ non nuls. alors si a divise c, b divise c et a, b sont premiers entre eux, alors ab divise c.*

Démonstration. Puisque a divise c, il existe k tel que $c = ak$. Puisque a et b sont premiers entre eux, et puisque b divise c, d'après le Lemme de Gauss b divise k. Donc ab divise c. □

La proposition suivante est importante : elle donne un lien entre le PGCD et le PPCM et donc un moyen de calculer ce dernier.

Lemme 1.28. *Soient $a, b \in \mathbb{Z}$. Alors*

$$\mathrm{PGCD}(a, b)\,\mathrm{PPCM}(a, b) = ab.$$

En particulier, si a, b sont premiers entre eux, alors $\mathrm{PPCM}(a, b) = ab$.

Démonstration. Soit $d = \mathrm{PGCD}(a, b)$ et m un multiple commun de a, b. On pose $a' = a/d$ et $b' = b/d$, a', b' sont donc premiers entre eux. Il existe $k, l \in \mathbb{Z}$ tels que $ka = m = lb$, et donc en divisant par d $ka' = lb'$. Par le Lemme de Gauss, a' divise l, donc il existe q tel que $l = qa'$. On en déduit qu'un multiple commun de a et b est forcément de la forme $qa'b'd$ pour un certain $q \in \mathbb{Z}$. Le plus petit d'entre eux est donc $a'b'd = ab/d$.

□

Remarque 1.29. On a prouvé au passage que si m est un multiple commun de a, b, alors $\mathrm{PPCM}(a, b)|m$.

1.6. **Théorème fondamental de l'arithmétique.** Dans cette section, on montre que tout entier naturel peut s'écrire de façon unique comme un produit de nombres premiers. Notons que la preuve que nous en donnons est anachronique, puisque nous avons utilisé le théorème de Bezout pour démontrer le Lemme d'Euclide. Ce lemme (et le théorème fondamental ci-dessous) était bien sûr connu d'Euclide qui en avait une preuve plus élémentaire, Bezout n'est donc pas nécessaire pour démontrer ce théorème, il est juste pratique.

Théorème 1.30. *Pour tout entier naturel $N \geq 2$, il existe une unique*[a] *liste de nombres premiers distincts $(p_1, \dots, p_k)$, et une unique liste d'entiers strictement positifs $(m_1, \dots, m_k)$, telles que*

$$n = p_1^{m_1} \dots p_k^{m_k}.$$

a. plus exactement unique à l'ordre près

Démonstration. Nous allons montrer ce théorème par récurrence. L'énoncé est vrai pour $N = 2$ puisque 2 est lui-même premier. Soit $N > 2$ et supposons que nous avons montré cette propriété pour $2, \dots, N-1$. De deux choses l'une : soit N est premier et dans ce cas il est évident qu'il admet une unique décomposition en produit de facteurs premiers. Soit N n'est pas premier. Dans ce cas il existe $1 < a, b < N$ tels que

$$N = ab.$$

Par hypothèse de récurrence a et b admettent une décomposition en facteurs premiers, donc N aussi. Supposons maintenant qu'on a deux décompositions

$$N = p_1^{m_1} \dots p_k^{m_k} = q_1^{n_1} \dots q_l^{n_l}.$$

On voit que p_1 divise $q_1^{n_1} \dots q_l^{n_l}$, donc par le Lemme d'Euclide il divise q_i pour un certain i, et quitte à réordonner le produit on peut supposer $i = 1$.. Puisque q_1 est premier par hypothèse, et comme $p_1 \neq 1$, on a forcément $p_1 = q_1$. Posons

$$M := \frac{N}{p_1} = p_1^{m_1 - 1} \dots p_k^{m_k}.$$

On en déduit que

$$M := p_1^{m_1 - 1} \dots p_k^{m_k} = q_1^{n_1 - 1} \dots q_l^{n_l}.$$

Mais par construction $M < N$, donc par hypothèse de récurrence il admet une unique décomposition en facteurs premiers. Donc $k = l$, et quitte à réordonner les facteurs on a $p_i = q_i$. □

En particulier, on obtient une autre façon d'exprimer le PGCD et le PPCM de deux nombres a, b. En effet on peut écrire

$$a = p_1^{k_1} \dots p_n^{k_n}$$

et

$$b = p_1^{l_1} \dots p_n^{l_n}$$

avec $k_i, l_i \in \mathbb{N}$. Attention ! Ici on utilise la même liste de nombres premiers pour les deux nombres, c'est-à-dire qu'on a combiné les diviseurs premiers de a et b (et donc on autorise $k_i = 0$ ou $l_i = 0$).

Dans ce cas on a

$$\mathrm{PGCD}(a, b) = p_1^{\min(k_1, l_1)} \dots p_n^{\min(k_n, l_n)}$$

et

$$\mathrm{PPCM}(a,b) = p_1^{\max(k_1,l_1)} \dots p_n^{\max(k_n,l_n)}.$$

Les nombres premiers sont donc les "briques élémentaires" qui permettent de construire tous les autres nombres. Un théorème important est le suivant :

Théorème 1.31. *Il existe une infinité de nombres premiers.*

Démonstration. Supposons qu'il en existe un nombre fini $\{p_1, \dots, p_n\}$ pour un certain $n \in \mathbb{N}$. Soit $N = P_1 \times P_2 \times \cdots \times p_n + 1$, et soit q u nombre premier qui divise N (il en existe forcément au moins un). Si q était égal disons à p_1, alors q diviserait N (par hypothèse)et $N - 1$ (puisque c'est le cas de p_1), et donc q diviserait 1 ce qui est absurde. Donc q n'appartient pas à la liste de nombres premiers dont nous somme partis, cette liste ne contient donc pas tous les nombres premiers. □

1.7. **Application du théorème de Bezout à la résolution d'équation.** Soient $a, b, c \in \mathbb{Z}$. On cherche dans cette section les paires $(x, y) \in \mathbb{Z}$ qui sont solutions de l'équation

$$ax + by = c.$$

Si a et/ou b vaut 0 c'est trivial, donc on suppose qu'ils sont tous les deux non nuls. Soit $d = \mathrm{PGCD}(a, b)$. Puisque par définition d divise a et b, si (x, y) est une solution alors d divise $ax + by$, donc doit diviser c. Par conséquent :

Première règle : si d ne divise pas c, alors l'équation n'a pas de solutions.

Supposons donc que $d|c$, c'est-à-dire qu'il existe c' tel que $c = c'd$.. Le théorème de Bezout nous dit qu'il existe $u, v \in \mathbb{Z}$ tels que

$$au + bv = d.$$

En multipliant par c' on obtient donc une solution particulière de l'équation

$$a(uc') + b(vc') = c'd = c.$$

On a donc pour l'instant trouvé *une* solution, mais on veut trouver *toutes* les solutions. Pour cela, posons $a' = a/d$, $b' = b/d$, $c' = c/d$. En divisant par deux deux côtés, on voit que l'équation qu'on veut ressoude a les mêmes solutions que

$$a'x + b'y = c'.$$

Mais cette fois a' et b' sont premiers entre eux. Soit donc (x_0, y_0) une solution particulière de cette équation, et (x_1, y_1) une autre solution. En soustrayant membre à membre on constate que

$$a'(x_1 - x_0) + b'(y_1 - y_0) = 0$$

c'est-à-dire

$$a'(x_1 - x_0) = -b'(y_1 - y_0).$$

Mais, a' et b' sont premiers entre eux (c'est pour ça qu'on a divisé par d), donc par le lemme de Gauss, a' divise $(y_1 - y_0)$, donc il existe $k \in \mathbb{Z}$ tel que

$$-(y_1 - y_0) = a'k.$$

En remplaçant on obtient

$$a'(x_1 - x_0) = b'a'k$$

et finalement en divisant des deux côtés par a'

$$(x_1 - x_0) = b'k.$$

En résumé :

Deuxième règle : si (x_0, y_0) est une solution particulière de l'équation

$$ax + by = c$$

et si $d = \mathrm{PGCD}(a, b)$, alors les solutions générales sont données par

$$(x_0 + \frac{bk}{d}, y_0 - \frac{ak}{d})\ k \in \mathbb{Z}.$$

Exemple 1.32. Essayons de résoudre les équations

$$102x + 38y = 47 \qquad\qquad 102x + 38y = 100.$$

On commence par appliquer l'algorithme d'Euclide pour calculer le PGCD

$$\begin{aligned} 102 &= 2 \times 38 + 26 \\ 38 &= 1 \times 26 + 12 \\ 26 &= 2 \times 12 + 2 \\ 12 &= 2 \times 6 + 0 \end{aligned}$$

donc $\mathrm{PGCD}(102, 38) = 2$. Puisque 2 ne divise pas 47, la première équation n'a aucune solution. EN revanche 2 divise 100, donc la seconde a des solutions. Cherchons les, en calculant d'abord une identité de Bezout.

$$\begin{aligned} 2 &= 26 - 2 \times 12 \\ 2 &= 26 - 2 \times (38 - 1 \times 26) \\ 2 &= 3 \times 26 - 2 \times 38 \\ 2 &= 3 \times (102 - 2 \times 38) - 2 \times 38 \\ 2 &= 3 \times 102 - 8 \times 38. \end{aligned}$$

Puisque $100 = 2 \times 50$, on multiplie par 50

$$100 = 150 \times 102 + (-400) \times 38$$

et on obtient ainsi une solution particulière. On a ensuite $102 = 2 \times 51$ et $38 = 2 \times 19$, donc les solutions sont

$$(150 + 19k, -400 - 51k)\ k \in \mathbb{Z}.$$

1.8. **Congruences.**

Définition 1.33. *Soient $a, b \in \mathbb{Z}$, $n \geq 1$ un entier. On dit que a est* congru à b modulo n*, et on note $a \equiv b \mod n$, si n divise $b - a$.*

Exemple 1.34. On a par exemple $2 \equiv -3 \mod 5$, $7 \equiv 0 \mod 7$, mais $2 \not\equiv 3 \mod 9$.

Remarque 1.35. Il est facile de voir que $a \equiv b \mod n$ si et seulement si a et b ont le même reste dans la division par n. En particulier, $a \equiv 0 \mod n$ ssi a est un multiple de n.

La notion de congruence est ce qu'on appelle une relation d'équivalence : deux nombres qui sont congruents ne sont pas égaux en général, mais ils sont "égaux", ou plutôt interchangeables, du point de vue du reste dans la division par n. La relation de congruence est donc similaire à la relation d'égalité :

Proposition 1.36. *Soient $a, b, c, n \in \mathbb{Z}$, $n \geq 1$, $k \in \mathbb{N}$. On a :*

(1) $a \equiv a \mod n$

(2) $a \equiv b \mod n \Leftrightarrow b \equiv a \mod n$

(3) $(a \equiv b \mod n$ *et* $b \equiv c \mod n) \Rightarrow a \equiv c \mod n$.

(4) si $k|n$, $a \equiv b \mod n \Rightarrow a \equiv b \mod k$.

Par ailleurs, la notion de congruence est compatible avec les opérations algébriques sur les entiers au sens suivant :

Proposition 1.37. *Soient $a, b, c, d, n \in \mathbb{Z}$, $n \neq 0$, et supposons*

$$a \equiv b \mod n \qquad\qquad c \equiv d \mod n.$$

Alors :

(1) $a + c \equiv b + d \mod n$

(2) $a - c \equiv b - d \mod n$

(3) $ac \equiv bd \mod n$

(4) $a \equiv b \mod n \Rightarrow a^k \equiv b^k \mod n$

Démonstration. $(a - b)c + b(c - d) = ac - bc + bc - bd$ □

Remarque 1.38. Une bizarrerie par rapport au calcul habituel est qu'un produit de deux nombres peut être nul (plus exactement, congru à 0) sans que l'un des facteurs ne soit nuls. Par exemple $2 \times 3 \equiv 0 \mod 6$.

Informellement quand on calcule modulo n, deux nombres qui sont congrus doivent être pensés comme étant "les mêmes". Ceci motive la définition suivante :

Définition 1.39. *Soit $a \in \mathbb{Z}$, $n > 1$ un entier. On appelle* classe d'équivalence modulo n de a, *et on note souvent $\bar{a}$, l'ensemble*

$$\bar{a} = \{k \in \mathbb{Z}, k \equiv a \mod n\}.$$

Un élément de cet ensemble s'appelle un représentant *de la classe d'équivalence.*

Remarque 1.40. Soit $a \in \mathbb{Z}$ et r le reste de la division de a par n. Alors $a \equiv r \mod n$. En particulier, l'ensemble $\{0, \ldots, n-1\}$ est ce qu'on appelle un *système de représentants* : tout entier est congru modulo n à exactement un élément de cet ensemble.

Remarque 1.41. Il est donc souvent pratique quand on fait des calculs de toujours prendre le reste dans la division par n à chaque étape. Ça n'est pas toujours le plus malin ! Si par exemple on veut calculer le reste dans la division par 8 de $7^{1254568765467653233}$. Dans ce cas il est utile de remarquer que $7 \equiv -1 \mod 8$, donc puisque la puissance précédente est impaire ce nombre est encore congru à -1, donc à 7, modulo 8, donc le reste recherché est 7.

1.9. **Systèmes d'équations et théorème des restes chinois.** On a vu plus haut que le produit de deux nombres qui ne sont pas congrus à $9 \mod n$, peut être congru à $0 \mod n$. Plus généralement, si $ab \equiv ac \mod n$, ça ne veut *pas* dire en général que $b \equiv c \mod n$! En revanche, il est souvent possible de "diviser modulo n" :

Interlude : que veut dire "diviser". Soit l'équation

$$7x = 2.$$

Si on travaille dans $\mathbb{Q}$, on sait que pour la résoudre on doit "diviser par 7 des deux côtés". Ce que ça veut dire, en fait, c'est qu'il existe un nombre rationnel, qu'on note $\frac{1}{7}$ et qu'on appelle l'inverse de 7, défini par le fait que

$$7 \times \frac{1}{7} = 1.$$

On multiplie donc des deux côtés de l'équation par ce nombre

$$\frac{1}{7} \times 7x = \frac{1}{7}2$$

donc $x = \frac{2}{7}$.

Si on travaille, par exemple, modulo 13, et donc qu'on veut résoudre

$$7x \equiv 2 \mod 13$$

on applique le même raisonnement : on peut remarquer que 2 est l'inverse de 7 au sens que

$$2 \times 7 \equiv 1 \mod 13.$$

Attention ! Un inverse n'existe pas toujours, c'est justement l'objet du théorème qui suit. Mais dans ce cas on en a un, et on peut donc résoudre en multipliant des deux côtés par 2 :

$$2 \times 7x \equiv 2 \times 2 \mod 13$$

et puisque $2 \times 7 \equiv 1 \mod 13$, on trouve $x \equiv 4 \mod 13$.

Théorème 1.42. *Soient $a, b, n \in \mathbb{Z}$, $n \neq 0$ et supposons que a et n sont premiers entre eux. Alors il existe une solution à l'équation*

$$ax \equiv b \mod n$$

et cette solution est unique modulo n.

Démonstration. Puisque a et n sont premiers entre eux, par Bezout il existe $u, v \in \mathbb{Z}$ tels que

$$au + nv = 1.$$

On a clairement $nv \equiv 0 \mod n$, donc

$$au \equiv 1 \mod n.$$

On a donc que $x = ub$ est une solution, en effet

$$a(ub) \equiv 1 \times b \mod n.$$

Par ailleurs, si y est aussi une solution de l'équation, alors on a

$$a(x-y) \equiv b - b \equiv 0 \mod n$$

donc n divise $a(x-y)$. Mais a et n sont premiers entre eux, donc par le lemme de Gauss n divise $x-y$, autrement dit

$$x \equiv y \mod n.$$

□

Remarque 1.43. On appelle u dans ce cas l'inverse de a modulo n, au sens ou multiplier par u, c'est pareil que "diviser/simplifier par a".

Si on ne suppose pas que a, n sont premiers entre eux, on a pas toujours des solutions, mais s'il y en a on peut se ramener au cas précédent.

Proposition 1.44. *Soient $a, b \in \mathbb{Z}$, $n > 1$ et $d = \mathrm{PGCD}(a, n)$. L'équation*

$$ax \equiv b \mod n$$

a des solutions si et seulement si $d|b$, et dans ce cas ses solutions sont les mêmes que celles de l'équation

$$a'x \equiv b' \mod n'$$

avec $a' = a/d$, $b' = b/d$, $n' = n/d$.

Démonstration. Si x est une solution de l'équation donnée, alors n divise $ax - b$. Or d divise n par définition, donc divise aussi $ax-b$. Mais comme d divise a encore par définition, on doit forcément avoir d divise b, donc c'est une condition nécessaire. Dans ce cas on a qu'il existe $k \in \mathbb{Z}$ tel que

$$\begin{aligned} & ax - b = kn \\ \Leftrightarrow & a'x - b' = kn' \\ \Leftrightarrow & a'x \equiv b' \mod n'. \end{aligned}$$

□

Remarque 1.45. Attention ! Dans la version précédente on a donc une solution unique modulo n', pas modulo n. Si un exercice demande les solutions modulo n (en général cela sous entend : entre 0 et $n-1$) alors il faut travailler un peu plus.

Exemple 1.46. Résolvons

$$6x \equiv 9 \mod 15.$$

On a $\mathrm{PGCD}(6, 15) = 3$ et $3|9$ donc il y a des solutions, et elles sont exactement les solutions de

$$2x \equiv 3 \mod 5.$$

En appliquant l'algorithme de Bezout étendu, on trouve $1 = -2 \times 2 + 1 \times 5$ donc l'inverse de 2 modulo 5 est -2, ou si on préfère 3. Donc les solutions

sont

$$x \equiv 3 \times 3 \equiv 4 \mod 5.$$

Si on demandait les solutions modulo 15, on aurait $4, 9, 14$ puisque ce sont les entiers entres 0 et 14, qui sont congrus à 4 modulo 5.

Le théorème suivant permet en gros de réduire le calcul de congruence modulo un grand nombre, au calcul modulo ses facteurs.

Théorème 1.47 (Théorème des restes chinois)**.** *Soient m, n des entiers naturels non nuls* premiers entre eux*, et a, b des entiers quelconques. Soient u, v des entiers tels que*

$$um + vn = 1$$

(qui existent d'après Bezout). Alors $x = bum + avn$ est une solution du système

$$\begin{cases} x & \equiv & a & \mod m \\ x & \equiv & b & \mod n. \end{cases}$$

De plus, cette solution est unique modulo mn.

Démonstration. La difficulté consiste seulement à inventer la formule donnée dans l'énoncé. Vérifier que c'est une solution est trivial : par définition $bum \equiv 0 \mod m$, donc le x donné est congru à avn modulo m. Mais la relation de Bezout implique (comme dans le théorème précédent) que $vn \equiv 1 \mod m$, donc $x \equiv a \mod m$. On prouve de la même façon que $x \equiv b \mod n$.

Par ailleurs, si y est une autre solution du système, alors $x - y \equiv 0 \mod m$, et $x - y \equiv 0 \mod n$. Puisque m, n sont premiers entre eux, on en déduit que le produit mn divise $x - y$. Réciproquement, si x est solution et si $y \equiv x \mod mn$, alors en particulier $x \equiv y \equiv a \mod m$, et pareil pour b. □

Remarque 1.48. En utilisant le théorème d'avant, on peut résoudre plus généralement les systèmes

$$\begin{cases} cx & \equiv & a & \mod m \\ dx & \equiv & b & \mod n. \end{cases}$$

En général : si m et n ne sont pas forcément premiers entre eux, posons $d = \mathrm{PGCD}(m, n)$ et $M = \mathrm{PPCM}(m, n)$. Si x est solution du système

$$\begin{cases} x & \equiv & a & \mod m \\ x & \equiv & b & \mod n. \end{cases}$$

alors d divise $b - a$, donc c'est une condition nécessaire pour l'existence de solutions. Montrons que c'est aussi suffisant, soit

$$mu + nv = d$$

une relation de Bezout. Posons $m' = m/d$ et $n' = n/d$. Alors

$$x = bum' + avn'$$

est solution : par hypothèse on a $a \equiv b \mod d$ donc $b = a + kd$ pour un certain $k \in \mathbb{Z}$ (éventuellement nul). Donc

$$\begin{aligned} bum' &= aum' + kudm' \\ &= aum' + kum \text{ puisque } m = m'd \\ &\equiv aum' \mod m. \end{aligned}$$

donc

$$x \equiv aum' + avn' \equiv a \mod m$$

comme voulu. On prouve de la même façon que $x \equiv b \mod n$. Finalement o,n vérifie que si x' est une autre solution, $x - x'$ est divisible par M, donc la solution qu'on a trouvé est unique modulo M (et pas mn, attention !).

Exemple 1.49. Résolvons

$$\begin{cases} 3x \equiv 9 \mod 14 \\ x \equiv 7 \mod 10. \end{cases}$$

On remarque d'abord que $\mathrm{PGCD}(14, 3) = 1$. Une relation de Bezout est

$$(-1) \times 14 + 5 \times 3 = 1$$

donc 5 est l'inverse de 3 modulo 14. On multiplie donc la première ligne par 5 :

$$\begin{cases} x \equiv 5 \times 9 \equiv 3 \mod 14 \\ x \equiv 7 \mod 10. \end{cases}$$

On calcule $\mathrm{PGCD}(14, 10) = 2$, et 2 divise bien $7 - 3 = 4$ donc on a des solutions. Une relation de Bezout est :

$$(-2) \times 14 + 3 \times 10 = 2.$$

On a donc une solution

$$x = 7 \times (-2) \times 7 + 3 \times 3 \times 5 = -53 \equiv 17 \mod 70 (= \mathrm{PPCM}(14, 10)).$$

Remarque 1.50. En travaillant de proche en proche, on peut résoudre des systèmes de plus de deux équations.

1.10. $\mathbb{Z}/n\mathbb{Z}$**.** Pour se simplifier la vie, plutôt que de travailler avec des entiers modulo, il est souvent pratique de travailler avec les classes modulo n. On rappelle que si $a \in \mathbb{Z}$, alors la classe $\bar{a}$ de a modulo n est l'ensemble des entiers congrus à a modulo n. Par exemple, modulo 7 on a $\bar{2} = \{\cdots - 5, 2, 9, 16, \ldots\}$ et $\bar{0} = \{\cdots - 14, -7, 0, 7, 14, \ldots\}$. L'idée c'est de remarquer que

$$a \equiv b \mod n \Leftrightarrow \bar{a} = \bar{b}.$$

En quelque sorte, le but est de remplacer une congruence qui "ressemble" à une égalité, par une vraie égalité.

Définition 1.51. *On note $\mathbb{Z}/n\mathbb{Z}$ et on appelle "anneau des entiers modulo n" l'ensemble des classes d'équivalences modulo n.*

Remarque 1.52. On expliquera pas le mot "anneau", mais en gros cela veut dire qu'on peut additionner et multiplier ses éléments. Le symbole "/" signifie qu'on prend $\mathbb{Z}$ et qu'on le "divise" en n paquets.

Concrètement on peut identifier $\mathbb{Z}/n\mathbb{Z}$ avec $\{\bar{0}, \bar{1}, \dots, \overline{n-1}\}$. C'est donc un ensemble fini contenant n éléments. Calculer modulo n c'est la même chose que de calculer dans $\mathbb{Z}/n\mathbb{Z}$. Par exemple si $n = 13$ on peut écrire $\bar{2} \times \bar{7} = \bar{1}$ ou $\bar{5} + \bar{7} = \bar{0}$, et dans ce cas c'est bien un signe égal et non une congruence.

Un rôle particulier est joué par l'ensemble des éléments inversibles :

Définition 1.53. *On note $(\mathbb{Z}/n\mathbb{Z})^\times$ l'ensemble des éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$, ou de façon équivalente l'ensemble des classes d'éléments premiers avec n.*

Remarque 1.54. Vous avez peut-être déjà vu la notation $\mathbb{R}^\times$ qui désigne les réels non nuls. En réalité cette notation désigne les nombres réels qui sont inversibles pour la multiplication, il se trouve simplement que dans ce cas là ce sont exactement les réels non nuls.

2. Éléments de théorie des groupes

C'est peut-être étonnant, mais en mathématique il est souvent plus facile de démontrer quelque chose de général que quelque chose de particulier. L'idée c'est que plutôt que de démontrer un ensemble de choses au cas par cas, on peut essayer d'identifier une structure commune à plusieurs objets qui nous intéressent, puis à travailler abstraitement avec cette structure, ce qui permet d'isoler ce qui est important de ce qui est trop spécifique. C'est ce que nous allons faire en introduisant la notion de *groupe*.

2.1. Généralités.

Définition 2.1. *Soit G un ensemble non vide et $\cdot$ une application de $G \times G$ dans G. On dit que $(G, \cdot)$ est un* groupe *si*

(1) $\cdot$ est associative *: $\forall x, y, z \in G, (x \cdot y) \cdot z = x \cdot (y \cdot z)$*

(2) Il existe un élément neutre : $\exists e \in G, \forall x \in G, x \cdot e = e \cdot x = x$

(3) tous les éléments de G sont inversibles : $\forall x \in G, \exists y, x \cdot y = y \cdot x = e$.

On dit que G est commutatif (ou abélien, du nom de Niels Abel) si $\forall x, y \in G, x \cdot y = y \cdot x$.

Remarque 2.2. On appelle $\cdot$ la loi de composition de G. En pratique on désigne souvent un groupe par son ensemble G sous-jacent, la loi de composition est sous-entendue.

Exemple 2.3. Les principaux exemples sont :

— $(\mathbb{Z}, +)$, $(\mathbb{R}^\times, \times)$
— $(\mathbb{Z}/n\mathbb{Z}, +)$. L'élément neutre est $\bar{0}$, et l'inverse de $\bar{a}$ est $-\bar{a} = \overline{-a}$.
— $((\mathbb{Z}/n\mathbb{Z})^\times, \times)$. L'élément neutre est $\bar{1}$, l'inverse de $\bar{a}$ est son inverse modulo n

Remarque 2.4. Une source possible de confusion : quand on parle d'un groupe abstrait, on note toujours la loi multiplicativement, càd qu'on note x^k pour $x \cdot x \dots x$ etc.

Tous ces exemples sont commutatifs. UN exemple de groupe non commutatif est l'ensemble des matrices $n \times n$ inversibles pour $n \geq 2$, avec comme loi la multiplication. Un autre exemple est l'ensemble des bijections d'un ensemble fini vers lui-même, avec comme loi la composition.

Lemme 2.5. *L'élément neutre d'un groupe est unique. Pour tout x de G, l'inverse de s est aussi unique (et sera noté x^{-1}).*

Démonstration. Soient e, e' deux éléments neutres. Alors on a $e \cdot e' = e$ d'une part, et $e \cdot e' = e'$ d'autre part, donc $e = e'$. □

Définition 2.6. *Soit H un sous ensemble* non-vide *de G. On dit que H est un* sous-groupe *de G si*

(1) $x, y \in H \Rightarrow x \cdot y \in H$
(2) $x \in H \Rightarrow x^{-1} \in H$

Exemple 2.7. Les sous-groupes de $\mathbb{Z}$ sont exactement les $n\mathbb{Z}$. $\{\bar{0}, \bar{2}\}$ est un sous-groupe de $\mathbb{Z}/4\mathbb{Z}$.

Remarque 2.8. Un sous-groupe H d'un groupe G, est lui-même un groupe.

Définition 2.9. *On appelle ordre d'un groupe son cardinal. On appelle groupe fini un groupe dont le cardinal est fini.*

Définition 2.10. *Soit $S = \{x1, \dots, x_k\}$ un ensemble d'éléments de G. On appelle sous-groupe engendré par S et on note $\langle S \rangle$ ou $\langle x_1, \dots, x_k \rangle$ l'ensemble de tous les produits possibles des x_i* et de leurs inverses.

Exemple 2.11. le sous-groupe de $\mathbb{Z}/4\mathbb{Z}$ engendré par $\{\bar{2}\}$ est $\{\bar{0}, \bar{2}\}$.

Définition 2.12. *ON dit que G est engendré par l'exemble S si $< S >= G$, c'es-à-dire si tous les éléments de G peuvent s'écrire comme un produit d'éléments de S et de leurs inverses. On dit que G est* cyclique *s'il existe un sous-ensemble S de G qui contient un seul élément et tel que G est engendré par S.*

Exemple 2.13. Le groupe $\mathbb{Z}/n\mathbb{Z}$ est toujours cyclique : il est engendré par $\bar{1}$. Attention, ça n'est en général pas le seul choix pssible, par exemple $\mathbb{Z}/5\mathbb{Z}$ est aussi engendré par $\bar{2}$.

2.2. **Théorème de Lagrange.** Le résultat suivant est très important.

Théorème 2.14. *Soit G un groupe fini, et H un sous-groupe de G. Alors l'ordre de H divise l'ordre de G.*

Démonstration. Pour tout $x \in G$ on introduit l'ensemble

$$xH := \{x \cdot h, h \in H\}$$

Si $x, y \in G$ alors de deux choses l'une :

- soit il existe $h \in H$ tel que $x = yh$ et dans ce cas $xH = yH$. En effet un élément z appartient à xH ssi il existe h' tel que $z = xh' = yhh'$ donc $z \in yH$. Réciproquement, puisque h est inversible par définition, on a aussi $y = xh^{-1}$ donc $yH \subset xH$.
- soit un tel h n'existe pas, et dans ce cas $xH \cap yH = \emptyset$.

  On peut donc choisir une famille finie $x_1, \dots, x_k$ d'élements telles que tout élément x de G appartiennent a un et un seul des $x_i H$. On obtient ce qu'on appelle une *partition* de G :

$$G = \sqcup_{i=1,\dots,k} x_i H.$$

  Par ailleurs, tous les ensemble xH ont le même nombre d'élements, puisque l'application de H dans xH donnée par $z \mapsto xz$ est une bijection, d'inverse donnée par $z \mapsto x^{-1}z$. On a donc divisé/partitionné G en k paquets, qui ont tous le même nombre d'éléments que H. DOnc le nombre d'élements de G est k fois le nombre d'éléments de H.

□

Ce théorème montre qu'être une structure de groupe est quelque chose de contraignant, par exemple on peut en déduire le résultat suivant pas du tout évident à priori :

Proposition 2.15. *Soit G un groupe fini d'ordre p un nombre premier. Alors G est cyclique.*

Démonstration. Soit x un élément de G différent de l'élément neutre, et soit $H = \langle x \rangle$. Par le thm de Lagrange, l'ordre de H divise p. Mais H contient au moins deux éléments (l'élément neutre et x), donc puisque p est premier on en déduit que l'ordre de H est p, et donc que $H = G$. □

Remarquons que dans $\mathbb{Z}/8\mathbb{Z}$ par exemple, on a $\bar{2}+\bar{2}+\bar{2}+\bar{2}=\bar{0}$, et dans $(\mathbb{Z}/8\mathbb{Z})^\times$ on a $\bar{3}\times\bar{3}=1$, autrement dit en partant d'un élément et en itérant la loi de composition on retombe sur l'élément neutre. C'est une conséquence d'un résultat plus général :

Définition 2.16. *Soit x un élément d'un groupe G d'élément neutre e. Si k est un entier, on note $x^k := x\cdot x\cdot\cdots\cdot x$ k fois, et par convention $x^0 = e$. Le plus petit entier strictement positif k, s'il existe, tel que $x^k = e$ est appelé* ordre de x. *Si un tel entier n'existe pas, on dit que k est d'ordre infini.*

Proposition 2.17. *Soit G un groupe et $x \in G$.*

(1) S'il existe un entier n tel que $x^n = e$, alors x est d'ordre fini et n est un multiple de l'ordre de x.

(2) L'ordre de x est égal à l'ordre du sous-groupe engendré par x.

(3) Dans un groupe fini, tout élément est d'ordre fini.

Démonstration. (1) Soit $n = kq + r$ la division euclidienne de n par k. On a $x^n = (x^k)^q\cdot x^r = e$ par hypothèse. Or $x^k = e$, donc $x^n = x^r$. Mais $0 \leq r < k$, et k est par définition le plus petit entier positif tel que $x^k = e$, donc $r = 0$.

(2) Tout produit de x et de x^{-1} peut se simplifier et s'écrire x^m pour $m \in \mathbb{Z}$. De deux choses l'une :
- soit tous les x^m sont différents les uns des autres, dans ce cas l'ordre de $\langle x\rangle$ est infini, et l'ordre de x aussi puisque $x^m \neq e$ si $m \neq 0$.
- soit il existe m, m' tels que $m \neq m'$ et $x^m = x^{m'}$. Dans ce cas on a $x^{m-m'} = e$ donc x est d'ordre fini. Dans ce cas, soit k l'ordre de x, alors on a
$$\langle w\rangle = \{e, x, x^2, \ldots, x^{k-1}\}$$
en remarquant par exemple que $x^{-1} = x^{k-1}$.
- découle trivialement du point précédent.

□

Un corollaire important de cette proposition et du théorème de Lagrange :

Proposition 2.18. *Soit G un groupe fini et $x \in G$. Alors l'ordre de x divise l'ordre de G.*

2.3. **Application : arithmétique modulaire et cryptographie.** Il découle de la section précédente que si x est élément de $(\mathbb{Z}/n\mathbb{Z})^\times$, alors x est d'ordre fini, donc il existe un plus petit entier positif x tel que $x^k = \bar{1}$. Si a est un entier premier avec n, on appelle l'ordre de $\bar{a}$ dans $\mathbb{Z}/n\mathbb{Z}$ *l'ordre multiplicatif de a modulo n*.

Définition 2.19. *Soit n un entier. On appelle indicatrice d'Euler de n, et on note $\phi(n)$, le nombre d'entiers entre 1 et n qui sont premiers avec n. Autrement dit, on a*
$$\phi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|.$$

On en déduit le théorème d'Euler :

Théorème 2.20. *Soit a et n premiers entre eux. Alors*

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

Si p est un nombre premier, alors par définition $\phi(p) = p - 1$. On en déduit le petit théorème de Fermat :

Théorème 2.21. *Soit p un premier, et a un entier. Alors*

$$a^p = a \mod p.$$

Si de plus p ne divise pas a, alors

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

Démonstration. Le 2e point est une conséquence directe du théorème d'Euler : puique p est premier, a et p sont premiers entre eux si et seulement si p ne divise pas a. On obtient dans ce cas le premier point en multipliant par a des deux côtés. Si p divise a, alors $a \equiv 0 \mod p$ et dans ce cas on a toujours trivialement $a^p \equiv 0 \equiv a \mod p$, donc le premier point est bien vrai quelq eu soit a. □

Proposition 2.22. *Si p est premier, alors $\phi(p) = p - 1$. Si m, n sont premiers entre eux, alors $\phi(mn) = \phi(m)\phi(n)$.*

Démonstration. Le second découle du théorème chinois : un entier est inversible modulo mn ssi il est inversible modulo m et modulo n, on a donc une bijection

$$(\mathbb{Z}/mn\mathbb{Z})^\times \to (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times.$$

En fait cette bijection est compatible avec la multiplication, c'est ce qu'on appelle un isomorphisme de groupe. □

Ce formalisme a de nombreuses applications. Nous allons voir l'exemple de l'algorithme RSA, du nom de ses inventeurs Ronald Rivest, Adi Shamir et Leonard Adleman. SUpposons qu'Alice veuille envoyer un message secret à Bob. Il semble à priori qu'Alice et Bob doivent convenir d'un mot de passe, mais pour l'envoyer de façon sécurisée il faudrait déjà en avoir un... C'est le point de départ de la cryptographie dit assymétrique, om le mot de passe (on dit plutôt la clé) pour chiffrer le message est différent de celui nécessaire pour déchiffrer.

L'algorithme RSA repose sur le fait qu'il est facile (en terme de temps de calcul) de multiplier deux nombres, mais difficile de factoriser un entier. Grossièrememnt, la complexité de la multiplication est de l'ordre le nombre de chiffres des entiers considérés, tandis que la factorisation est de l'ordre le nombre lui-même.

Le protocole est donc le suivant : Alice choisit deux grands (plus de 300 chiffres) nombres premiers p, q. Elle

- calcule $n = pq$
- calcule $\phi(n) = (p-1)(q-1)$. C'est l'ingrédient clé ! C'est facile de calculer $\phi(n)$ si on connait p et q, mais très, très difficile sinon.
- choisit un entier e premier avec $\phi(n)$ et plus petit que $\phi(n)$.
- elle calcule l'inverse d de e modulo $\phi(n)$, ce qui est facile avec l'algo d'Euclide.

Elle donne à tout le monde, et donc à Bob, la paire (e, n). Bob encode son message sous forme de bloc d'entiers plus petits que n. Soit M un tel bloc. On suppose

que ni p ni q ne divise M, ce qui implique que M et N sont premiers entre eux. Si p et q

Alors il calculs

$$S = M^e \mod n$$

c'est le message secret. Pour le déchiffrer, Alice calcule

$$S^d \mod n.$$

En effet, par hypothèse il existe k tel que $ed = 1 + k\phi(n)$. Donc

$$M^{ed} \equiv M \times (M^{\phi(n)})^k \equiv M \mod n.$$

2.4. **Homomorphismes, isomorphismes.**

Définition 2.23. *Un homomorphisme (ou juste morphisme) d'un groupe $(G, \cdot)$ vers un groupe $(H, \star)$ est une application*

$$f : G \longrightarrow H$$

telle que pour tout $x, y \in G$, on a

$$f(x \cdot y) = f(x) \star f(y).$$

Un homomorphisme bijectif est appelé un isomorphisme.

Remarque 2.24. Dire qu'il existe un isomorphisme entre deux groupes, c'est dire que toute question sur l'un peut être traduite en une question sur l'autre, autrement dit les deux groupes sont essentiellement "les mêmes".

Proposition 2.25. *Si f est un homomorphisme d'un groupe $(G, \cdot)$ vers un groupe $(H, \star)$, alors automatiquement*

(1) $f(e_G) = e_H$

(2) $f(x^{-1}) = f(x)$

Démonstration. On a $f(e_G) = f(e_g \cdot e_g) = f(e_G) \star f(e_G)$. Or par définition $e_H = f(e_g) \star f(e_G)^{-1} = f(e_G)$.

Ensuite on a $\forall x \in G$, $e_H = f(e_G) = f(x \cdot x^-1) = f(x) \star f(x^{-1})$. Donc $f(x^{-1})$ est l'inverse de $f(x)$. □

Exemple 2.26.

(1) L'application f de $\mathbb{Z}/4\mathbb{Z}$ dans $\mathbb{Z}/6\mathbb{Z}$ définie par

$$f(\bar{0}) = \bar{0} \qquad f(\bar{1}) = \bar{3}$$
$$f(\bar{2}) = \bar{0} \qquad f(\bar{3}) = \bar{3}$$

est un morphisme de groupe.

(2) Soit $n \geq 1$. L'application

$$\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$$

qui envoie un entier a sur $\bar{a}$ est un morphisme.

(3) si m divise n, alors on a un morphisme
$$\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}.$$
(4) l'inclusion $K \subset G$ d'un sous-groupe K d'un groupe G est un morphisme.
(5) Si G est un groupe cyclique d'ordre n, et si x est un générateur de g, alors l'application de $\mathbb{Z}/n\mathbb{Z}$ vers G qui envoie $\bar{a}$ sur x^a est un isomorphisme de groupe. Par exemple $(\mathbb{Z}/5\mathbb{Z})^\times$ est cyclique d'ordre 4, engendré par $\bar{2}$.

Définition 2.27. *L'image d'un morphisme $f : G \to H$ est*
$$\operatorname{Im} f = \{y \in H, \exists x \in G, f(x) = y\}$$
et le noyau est
$$\ker f = \{x \in G, f(x) = e_H\}.$$

Exemple 2.28. Reprenons l'exemple de l'application f de $\mathbb{Z}/4\mathbb{Z}$ dans $\mathbb{Z}/6\mathbb{Z}$ définie par
$$f(\bar{0}) = \bar{0} \qquad f(\bar{1}) = \bar{3}$$
$$f(\bar{2}) = \bar{0} \qquad f(\bar{3}) = \bar{3}.$$
Dans ce cas, $\operatorname{Im} f = \{\bar{0}, \bar{3}\} \subset \mathbb{Z}/6\mathbb{Z}$, et $\ker f = \{\bar{0}, \bar{2}\} \subset \mathbb{Z}/4\mathbb{Z}$.

Proposition 2.29. *Soit $f : (G, \cdot) \longrightarrow (H, \star)$ un morphisme de groupes.*
(1) $\operatorname{Im} f$ et $\ker f$ sont des sous-groupes de H et de G respectivement.
(2) un morphisme de groupe est injectif ssi son noyau est le sous-ensemble réduit à $\{e_G\}$

Démonstration. Pour le premier point : soient y, y' des éléments de $\operatorname{Im} f \subset H$. Par définition, ça signifie qu'il existe x, x' tels que
$$f(x) = y \qquad f(x') = y'.$$
Puisque f est un moprphisme de groupes, on a
$$f(x \cdot x') = y \star y'$$
ce qui siginifie que le produit $y \star y'$ est bien l'image par f d'un élément de G, et donc que $y \star y' \in \operatorname{Im} f$. Par ailleurs on a, encore une fois parce que f est un morphisme de groupe,
$$f(x^{-1}) = f(x)^{-1} = y^{-1}$$
ce qui prouve que $y^{-1} \in \operatorname{Im} f$, et finalement que $\operatorname{Im} f$ est un sous-groupe de H.

Soient maintenant $x, x' \in \ker f$. Par définition, cela signifie que
$$f(x) = e_H \qquad f(x') = e_H.$$
On a donc
$$f(x \cdot x') = f(x) \star f(x') = e_H \star e_H = e_H$$

et

$$f(x^{-1}) = f(x)^{-1} = e_H^{-1} = e_H$$

donc $x \cdot x' \in \ker f$ et $x^{-1} \in \ker f$, donc $\ker f$ est un sous-groupe de G.

Pour le second point, rappelons que par définition l'application f est injective si et seulement si

$$\forall x, x' \in G, f(x) = f(x') \Rightarrow x = x'$$

(autrement dit une application est injective si des éléments distincts ont forcément des images distinctes).

Mais puisque dans ce cas H est un groupe, $f(x) = f(x')$ est équivalent à $f(x) \star f(x')^{-1} = e_H$. Et puisque f est un morphisme ça revient à dire que

$$f(x \cdot x'^{-1}) = e_H$$

autrement dit que $x \cdot x'^{-1} \in \ker f$. f est injective si et seulement si cela implique que $x = x'$, c'est-à-dire $x \cdot x'^{-1} = e_G$. Donc f est injective si et seulement si $\ker f = \{e_G\}$. □

Remarque 2.30. Le deuxième point est très utile en pratique pour tester si un morphisme est injectif.

2.5. **Groupe symétrique et permutations.**

Définition 2.31. *Soit $n \geq 1$ un entier. On appelle n-ième groupe symétrique, et on note S_n, le groupe des bijections de l'ensemble $[n] = \{1, \ldots, n\}$ dans lui-même, muni de la composition des applications. Son élément neutre est* Id, *l'application identité définie par*

$$\mathrm{Id}(x) = x.$$

On rappelle que si f est une bijection d'un ensemble E vers E, alors par définition pour tout $b \in E$ il existe un *unique* $a \in E$ tel que

$$f(a) = b.$$

Autrement dit, tout élément de E a un unique antécédent par l'application f. Il existe donc une application inverse notée f^{-1} qui envoit b sur a (qui envoit tout élément de E sur son unique antécédent par f). Informellement, f^{-1} défait ce que f fait. Plus formellement, on a

$$f \circ f^{-1} = f^{-1} \circ f = \mathrm{Id},$$

donc S_n est bien un groupe.

Une bijection de l'ensemble $[n]$, c'est une façon de "mélanger" ses éléments. Un élément de S_n est donc une *permutation* de la liste ordonnée $(1, \ldots, n)$. On note une permutation σ de la façon suivante :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ \sigma(1) & \sigma(2) & & \cdots & \sigma(n-1) & \sigma(n). \end{pmatrix}$$

Exemple 2.32. Dans S_3 la permutation identité est

$$\mathrm{Id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}.$$

La permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

est l'application σ définie par $\sigma(1) = 3, \sigma(2) = 1, \sigma(2) = 3$.

Ce groupe a $3! = 6$ éléments. En plus des deux qu'on vient d'écrire il y'a :

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Remarque 2.33. On note souvent les applications avec les lettres $f, g, \ldots$. En revanche, on a l'habitude de noter les permutations $\sigma, \tau, \ldots$. On omet aussi souvent le symbole $\circ$, càd qu'on note $\sigma\tau$ le produit au lieu de $\sigma \circ \tau$.

On rappelle qu'on note $n!$ (la factorielle de n) le produit $n \times n - 1 \times \cdots \times 1$.

Proposition 2.34. *Le groupe S_n contient $n!$ éléments. En particulier c'est donc un groupe fini.*

Démonstration. On a n choix pour l'image de 1, puis $n - 1$ choix pour l'image de 2, etc.... □

Ces groupes sont importants en combinatoire, mais aussi pour l'étude des groupes finis en général pour la raison suivante :

Théorème 2.35 (Cayley, 1854)**.** *Soit $(G, \cdot)$ un groupe fini d'ordre n. Alors il existe un homomorphisme de groupe injectif de G vers S_n. Autrement dit, G est isomorphe à un sous-groupe de S_n.*

Démonstration. Tout élément a de G induit une application de l'ensemble G vers lui même définie par

$$x \longmapsto a \cdot x.$$

Puisque a est inversible, c'est une bijection.

Par ailleurs, on peut numérote les éléments de x par les entiers de 1 à n : $g_1, g_2, \ldots, g_n$. On peut donc définir pour tout a dans G une bijection l_a de $[n]$ vers $[n]$ en posant :

$$l_a(i) = \text{le numéro de } a \cdot g_i.$$

Soient $a, b \in G$ et soit j le numéro de l'élement $b \cdot g_i$, c'est-à-dire qu'on a

$$l_b(i) = j.$$

ALors on a d'une part

$$(l_a \circ l_b)(i) = l_a(j) = \text{le numéro de } a \cdot g_j$$

et d'autre part

$$l_{a\cdot b}(i) = \text{le numéro de } (a\cdot b)\cdot g_i = \text{le numéro de } a\cdot(b\cdot g_i) = \text{le numéro de } a\cdot g_j.$$

Ceci implique que l'application de G vers S_n définie par $a \mapsto l_a$ est bien un morphisme de groupe.

Finalement, un élément a de G est dans le noyau de cette appliaction si et seulement si pour tout $i \in [n]$ on a

$$a \cdot g_i = g_i.$$

On en déduit que a est forcément égal à l'élément neutre de G, et donc par la Proposition 2.29 que ce morphisme est injectif. □

Remarque 2.36. Une bijection de $[n]$ induit une bijection de $[m]$ pour $m \geq n$ en déclarant que $\sigma(i) = i$ pour $i > n$. C'est clairement compatible avec la composition, autrement dit S_n est naturellement (identifié à) un sous-groupe de S_m.

Définition 2.37. *On appelle k-cycle une permutation σ de l'ensemble $[n]$ telle qu'il existe $i_1, \ldots, i_k$ tels que $\sigma(i_1) = i_2, \sigma(i_2) = i_3, \ldots, \sigma(i_{k-1}) = i_k, \sigma(i_k) = i_1$, et les autres éléments sont fixés. On note un tel cycle $(i_1 \ldots i_k)$. Un 2-cycle est souvent appelé une transposition. Le* support *d'un cycle est l'ensemble $\{i_1, \ldots, i_k\}$.*

Exemple 2.38. La permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 3 & 1 & 5 \end{pmatrix}$$

est un 3-cycle, puisque elle envoit 1 sur 2, 2 sur 4 et 4 sur 1, et laisse 3 et 5 inchangés. On peut donc aussi la noter $(1\ 2\ 4)$.

Dans l'exemple précédent, on voit que $\sigma^2(1) = \sigma(2) = 4$, donc prendre le carré revient à parcourir le cycle de deux en deux. De même $\sigma^3(1) = \sigma(4) = 1$ donc on est revenus au point de départ. Plus généralement, on a :

Proposition 2.39. *Un k-cycle est d'ordre k. En particulier, si τ est une transposition on a $\tau^2 = \mathrm{Id}$, et donc $\tau = \tau^{-1}$.*

Démonstration. Soit $\sigma = (i_1, i_2, \ldots, i_k)$ un k-cycle. Par construction, $\sigma^l(i_j) = i_{j+l \mod k}$, qui est différent de i_j sauf si $l = k$. Par ailleurs, pour tous les entiers $j \notin \{i_1? \ldots, i_k\}$, on a $\sigma(j) = j$ donc $\sigma^k(j) = j$. □

Proposition 2.40. *Deux cycles dont les supports sont disjoints commutent.*

Démonstration. C'est évident puisque si leurs supports sont disjoints, cela signifie qu'ils ne "touchent" pas aux mêmes entiers, on obtient donc le même résultat en appliquant l'une puis l'autre, ou l'autre puis l'une. □

Exemple 2.41. Soient $\sigma = (1\ 2\ 4)$ et $\tau = (3\ 5\ 7)$ des cycles à supports disjoints de S_7. On a

$$(1,2,3,4,5,6,7) \xrightarrow{\sigma} (2,4,3,1,5,6,7) \xrightarrow{\tau} (2,4,5,1,7,6,3)$$

et

$$(1,2,3,4,5,6,7) \xrightarrow{\tau} (1,2,5,4,7,6,3) \xrightarrow{\sigma} (2,4,5,1,7,6,3).$$

On a donc bien $\tau\sigma = \sigma\tau$.

Le théorème suivant est intuitif parce qu'on a que deux mains ! Si j'ai n objets devant moi, je peux toujours les mettre dans un ordre quelconque en utilisant seulement mes deux mains, càd en n'échangeant que deux objets à la fois.

Théorème 2.42. *L'ensemble des transpositions est un ensemble de générateurs pour S_n. Autrement dit, toute permutation peut s'écrire comme un produit de transpositions.*

Démonstration. On montre le résultat par récurrence. C'est clair pour S_2. Supposons que c'est vrai pour S_n. Soit $\sigma \in S_{n+1}$. De deux choses l'une :

— Ou bien $\sigma(n+1) = n+1$, c'est-à-dire que σ ne permute que les n premiers entiers et ne touche pas à $n+1$, donc on peut la voir comme un élément de S_n (auquel on pense comme à un sous-groupe de S_{n+1}). Par hypothèse de récurrence, on peut donc l'écrire comme un produit de transpositions dans S_n, donc dans S_{n+1}.

— Sinon, soit $k = \sigma(n+1)$, et on pose

$$\sigma' = (k\ n+1) \circ \sigma.$$

Par construction $\sigma'(n+1) = n+1$, donc par le point précédent σ' peut s'écrire comme un produit de transpositions. Par construction on a

$$\sigma = (k\ n+1) \circ \sigma'$$

puisque

$$(k\ n+1) \circ (k\ n+1) = \mathrm{Id}$$

(c'est vrai pour toutes les transpositions). Donc on peut écrire σ comme un produit de transpositions.

□

Remarque 2.43. Ce théorème fournit un algorithme pour obtenir une décomposition particulière. En revanche, celle-ci n'est pas unique !

Exemple 2.44. Soit

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 6 & 7 & 1 & 2 & 4 \end{pmatrix} \in S_7.$$

On voit que σ envoit 7 sur 4. On considère la transposition $(4\ 7)$. Rappelons que les transpositions sont d'ordre 2 (échanger deux entiers, puis les échanger de nouveau c'est comme ne rien faire du tout). Formellement on a $(4\ 7) \circ$

$(4\ 7) = \mathrm{Id}$. On a donc

$$\sigma = (4\ 7)(4\ 7) \circ \sigma = (4\ 7) \circ \sigma'$$

avec

$$\sigma' = (4\ 7) \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 6 & 4 & 1 & 2 & 7 \end{pmatrix}.$$

Prosaïquement on a donc, par rapport à σ, remis 7 à sa place et changé le reste en conséquence.

En particulier $\sigma'(7) = 7$ donc on peut l'identifier à la permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 6 & 4 & 1 & 2 \end{pmatrix} \in S_6.$$

On recommence : cette permutation envoit 6 sur 2, donc on écrit

$$\sigma = (4\ 7) \circ \sigma' = (4\ 7) \circ (2\ 6) \circ \sigma''$$

avec

$$\sigma'' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 2 & 4 & 1 & 6 & 7 \end{pmatrix}.$$

qu'on peut identifier à une permutation de S_5. En continuant on obtient

$$\sigma = (4\ 7)(2\ 6)(1\ 5)(2\ 3)(1\ 2).$$

La définition suivante est un peu ad hoc mais nous verrons qu'elle est utile.

Définition 2.45. *Soit σ une permutation. Une* inversion *pour σ est une paire d'entiers i, j tels que $i < j$ et $\sigma(i) > \sigma(j)$. La* signature *de σ est $\epsilon(\sigma) = (-1)^N$ où N est le nombre d'inversions pour σ. Une permutation est dite paire si $\epsilon(\sigma) = 1$, impaire sinon.*

Remarque 2.46. Si k est un nombre, on note

$$\mathrm{sgn}(k) = \begin{cases} -1 & \text{si } k < 0 \\ 0 & \text{si } k = 0 \\ 1 & \text{si } k > 0 \end{cases}$$

Dans ce cas la signature d'une permutation est donnée par la formule

$$\epsilon(\sigma) = \prod_{1 \leq i < j \leq n} \mathrm{sgn}(\sigma(i) - \sigma(j)).$$

Exemple 2.47.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 4 & 2 \end{pmatrix}$$

On compte le nb de paires en désordre sur la seconde ligne : on a $3 > 2$, $5 > 4$, $5 > 2$, $4 > 2$. On a donc 4 inversions, donc σ est paire.

Remarquons que cette façon de calculer la signature est un peu pénible ! Voici un ingrédient essentiel d'une approche plus efficace.

Théorème 2.48. *La signature du produit de deux permutations, est le produit de leurs signatures. Autrement dit : soit $\mathcal{E}$ le groupe $\{\pm 1\}$ muni de la multiplication. Alors la signature est un morphisme de groupe*

$$\epsilon : S_n \longrightarrow \mathcal{E}.$$

Démonstration. Calcul un peu pénible. □

Définition 2.49. *L'ensemble des permutations paires est donc un sous-groupe de S_n (l'identité est paire, le produit de deux permutations paires est paire). Plus formellement, c'est le noyau de ce morphisme. On appelle ce groupe le* groupe alterné *A_n.*

Remarque 2.50. Si σ est une permutation impaire quelconque, alors l'ensemble des permutations impaires n'est autre que l'ensemble σA_n. En raisonnant comme dans la preuve du théorème de Lagrange, on voit que A_n et σA_n ont le même nombre d'éléments. On en déduit que la moitié des permutations sont paires, et l'autre moitié impaires. Par conséquent le cardinal de A_n est $n!/2$.

L'intérêt de ce théorème est de pouvoir calculer la signature une fois qu'on a décomposé une permutation en produit d'éléments plus simples. En particulier, nous avons vu que toute permutation se décompose en un produit de transpositions.

Proposition 2.51. *La signature d'une transposition est -1.*

Démonstration. Soit $\tau = (i\ j)$ une transposition avec $i < j$. Soit $\{k, l\}$ une paire d'entiers avec $k < l$. Pour que ça soit une inversion pour τ, on vérifie facilement qu'il est nécessaire que $k = i$ ou bien que $l = j$. On voit par exemple que la paire $\{i, j\}$ est toujours une inversion pour τ. Comptons les autres.

La paire $\{i, l\}$, $l \neq j$, est une inversion pour τ si et seulement si $i < ll < j$. Il y'a donc $j - i - 1$ possibilités. De même, il y'a $j - i - 1$ inversions de la forme $\{k, j\}$. Il y'en a donc $2(j - i - 1) + 1$ au total, et ce nombre est bien impair. □

Corollaire 2.52. *La signature d'une permutation σ est égale à $(-1)^N$ où N est le nombre de transpositions apparaissant dans* une *décomposition de σ en produit de transpositions. Autrement dit une permutation est paire (resp. impaire) si et seulement si ce nombre de transpositions est paire (resp. impair)*[a].

a. C'est d'ailleurs l'origine de cette terminologie

Remarque 2.53. Nous avons évoqué le fait que la décomposition d'une permutation en produit de transposition n'est pas unique. En fait même le nombre

de transpositions apparaissant n'est pas unique. Par contre, le corollaire précédent implique que la parité de ce nombre est toujours la même.

Exemple 2.54. Soit
$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$
une permutation (remarquons que c'est le cycle $(1\ 3\ 1)$ ce qui permet de calculer sa signature encore plus facilement comme nous verrons plus tard). UN calcul rapide montre que
$$\sigma = (1\ 2)(2\ 3)$$
mais aussi que
$$\sigma = (2\ 3)(1\ 2)(2\ 3)(1\ 2).$$
Ces deux décompositions ont des longueurs différentes, mais elles sont toutes les deux de longueurs paires, donc σ est une permutation paire, càd de signature 1.

Avant de poursuivre, voyons un exemple d'application de la signature :

Exemple 2.55 (Jeu de taquin)**.** Le jeu de taquin est un jeu créé aux États-unis à la fin du 19e siècle que vous connaissez sûrement :

Le but du jeu est de faire coulisser les cases afin de les remettre dans le bon ordre. Si on décide que la case vide porte le numéro 16, on peut voir chaque configuration du jeu comme une permutation de S_{16}. Ce groupe contient 16!=20.922.789.888.000 éléments! Pourtant il est facile de montrer le résultat suivant : à partir d'une configuration où la case vide est en bas à droite (comme sur l'image), alors le jeu est résoluble si et seulement si la permutation associée à la configuration est paire. Or, dans une des versions originales du jeu seule la case 15 et 14 sont inversées, autrement dit la configuration correspond à la transposition $(14\ 15)$ qui est impaire, ce qui avouons le est un peu vicieux. Ce qui est moins évident c'est que la réciproque est vraie : si deux configurations ont la même parité, alors il est possible de passer de l'une à l'autre. Par conséquent à partir de n'importe quelle configuration, il y'a 16!/2=10.461.394.944.000 configurations qu'on peut atteindre! Je vous conseille de jeter un œil à `http://images.math.cnrs.fr/Le-jeu-de-taquin-du-cote-de-chez` pour en savoir plus.

Définition 2.56. *Soit $\sigma \in S_n$ une permutation, et $x \in [n]$ un entier. On appelle orbite de x sous l'action de σ, et on note $O_\sigma(x)$ l'ensemble des entiers qu'on obtient*

en appliquant répétitivement σ à x. Formellement

$$O_\sigma(x) = \{\sigma^k(x), k \in \mathbb{Z}\}.$$

Remarque 2.57. Soit l l'ordre de σ. Alors par définition $\sigma^l(x) = x$, par conséquent

$$O_\sigma(x) = \{x, \sigma(x), \sigma^2(x), \dots, \sigma^{l-1}(x)\}.$$

Attention ! Ça ne veut pas dire que l'orbite contient toujours l éléments, seulement qu'elle a au plus l éléments.

Remarque 2.58. Si y est un autre élément de $[n]$ qui appartient à l'orbite de x, alors par définition il existe m tel que $\sigma^m(x) = y$. En particulier, on a égalité entre les ensemble

$$O_\sigma(x) = O_\sigma(y).$$

On peut donc partitionner l'ensemble $[n]$ en orbites disjointes..

Exemple 2.59. Soit

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 1 & 2 & 5 & 7 & 6 & 8 & 4 \end{pmatrix}$$

une permutation de S_7. Alors :

- l'orbite de 1 est $\{1, 2, 3\}$. C'est du coup aussi l'orbite de 2 et 3.
- l'orbite de 4 est $\{4, 5, 7, 8\}$. C'est donc aussi l'orbite de 5, 7 et 8
- l'orbite de 6 est $\{6\}$.

Proposition 2.60. *Soit $\sigma \in S_n$ et $k \geq 2$. Alors σ est un k-cycle si et seulement si elle admet une orbite qui contient k éléments et que toutes les autres contiennent exactement un élément.*

Démonstration. Si σ est le cycle $(i_1 \ \dots i_k)$, alors par définition l'orbite de i_1 est l'ensemble $\{i_1, \dots, i_k\}$, et toujours par définition pour tout entier x de $[n]$ qui n'est pas égal à l'un des i_k on a $\sigma(x) = x$, donc l'orbite de x est $\{x\}$, elle est donc réduite à un point. Réciproquement, si σ n'a qu'une seule orbite O non réduite à un point et que celle-ci contient k éléments, alors soit x n'importe quel élément de O, par définition de l'orbite on a

$$O = \{x, \sigma(x), \dots, \sigma^{k-1}(x)\}.$$

Par ailleurs, tout entier qui n'est pas dans cette orbite est fixé par σ puisque les autres orbites ne contiennent qu'un seule élément. En posant $i_l = \sigma^l(x)$ on voit bien que σ est le cycle $(i_1 \ i_2 \ \dots i_k)$. □

On a tous les ingrédients pour prouver le théorème suivant :

Théorème 2.61. *Soit $\sigma \in S_n$ une permutation. Alors il existe une unique, à l'ordre près, décomposition de σ en produit de cycles à supports deux à deux disjoints*[a].

a. puisque deux cycles à support disjoints commutent on peut les mettre dans n'importe quelle ordre

Démonstration. Si $\sigma = \mathrm{Id}$ alors c'est trivial, donc supposons que ça n'est pas le cas. σ a donc au moins une orbite qui n'est pas réduite a un point. Soit $O_1, \dots, O_m$ la liste des orbites qui contiennent au moins deux éléments. D'après la proposition précédente, il existe un unique cycle γ_i dont l'unique orbite non réduite à un point est O_i. On pose

$$\sigma' = \gamma_1 \circ \gamma_1 \circ \cdots \circ \gamma_m.$$

On remarque que O_i est par construction le support de γ_i, et que les O_i n'ont pas d'éléments en commun (c'est toujours le cas des orbites d'une permutation). Il en découle que les γ_i ont bien des supports disjoints.

Montrons que $\sigma' = \sigma$. Soit $x \in [n]$. De deux choses l'une :

— soit $\sigma(x) = x$. Dans ce cas l'orbite de x ne contient qu'un élément, donc x n'appartient à aucune des O_i, donc $\sigma'(x) = x$ également.

— soit $\sigma(x) \neq x$. Dans ce cas x appartient à l'une des orbites de la liste ci-dessus, disons $x \in O_j$. Dans ce cas, puisque les γ_i ont des supports disjoints par construction, on a

$$\gamma_i(x) = \begin{cases} x & \text{si } i \neq j \\ \sigma(x) \text{ si } i = j. \end{cases}$$

Donc $\sigma = \sigma'$. cette décomposition est clairement unique : si on a une autre décomposition en produit de cycles à supports disjoints, et si δ est un cycle de cette décomposition, alors soit $\delta(x) = x$, soit δ est l'unique cycle dont l'orbite non réduite à un point coïncide avec ml'orbite de x sous l'action de σ, donc δ est égale à l'un des γ_i.

□

Exemple 2.62. Dans l'exemple précédent, on a donc $\sigma = (1\ 2\ 3)(4\ 5\ 7\ 8)$.

La décomposition en produit de cycles disjoints est à la fois plus facile à obtenir, et souvent plus pratique que la décomposition en produit de transpositions. En plus elle a le bon goüt d'être unique.

Théorème 2.63. *Soit γ un cycle de longueur k. Alors γ peut s'écrire comme un produit de $k-1$ transpositions. En particulier, d'après le Corollaire 2.52, sa signature est $(-1)^{k-1}$.*

Démonstration. Montrons ce résultat par récurrence.

Si $k = 1$, alors $\gamma = \mathrm{Id}$ qui est bien un produit de zéro transposition(s).

Si $k = 2$, alors γ est une transposition, donc peut s'écrire comme un produit de $2 - 1 = 1$ transposition(s).

Supposons donc qu'on a montré le résultat pour tous les entiers jusqu'à $k - 1$ pour un certain entier $k > 2$, et supposons que γ est un cycle de longueur k. Alors il existe des entiers $i_1, \dots, i_k$ tels que γ est le cycle $(i_1\ i_2\ \dots\ i_k)$. Autrement dit $\gamma(i_1) = i_2$, $\gamma(i_2) = i_3$, etc...et $\gamma(i_k) = i_1$.

On pose

$$\gamma' = (i_1\ i_k) \circ \gamma.$$

Encore une fois, on a donc aussi

$$\gamma = (i_1\ i_k) \circ \gamma'.$$

(le "prime" a changé de place!). Montrons que γ' est un cycle de longueur $k-1$. Moralement, on a simplement "remis i_k à sa place". En effet :

— pour tout $1 \leq j \leq k-2$ on a

$$\gamma'(i_j) = \gamma(i_j) = i_{j+1}$$

— on a $\gamma'(i_{k-1}) = i_1$ et $\gamma'(i_k) = i_k$ par construction. Donc γ' n'est autre que le cycle

$$\gamma' = (i_1\ i_2\ \dots\ i_{k-1})$$

qui est bien de longeur $k-1$. Par hypothèse de récurrence, on peut écrire γ' comme un produit de $k-2$ transpositions, donc on peut écrire γ comme un produit de $k-1$ transpositions.

□

Proposition 2.64. *Soit $\sigma = \gamma_1 \dots \gamma_m$ une permutation décomposé en produit de cycles à support disjoint, et soit $\ell(\gamma_i)$ la longueur du cycle γ_i. Alors l'ordre de σ est* $\mathrm{PPCM}(\ell(\gamma_1), \dots, \ell(\gamma_m))$.

Démonstration. Rappelons d'abord que l'ordre d'un cycle n'est autre que sa longueur. Puisque les γ_i commutent, on a pour tout entier k

$$\sigma^k = \gamma_1^k \dots \gamma_m^k.$$

En particulier, puisque les γ_i ont des supports disjoints, $\sigma^k = 1$ si et seulement si $\gamma_i^k = 1$ pour tout $1 \leq i \leq m$. D'après le corollaire du théorème de Lagrange 2.18, on a donc $\sigma^k = \mathrm{Id}$ si et seulement si k est un multiple commun des ordres des γ_i. L'ordre de σ est donc le plus petit multiple positif de ces ordres. □

3. Probabilités

La théorie des probabilité est l'étude mathématique des phénomènes liés au hasard. Elle joue un rôle important dans la plupart des sciences, y compris en informatique. C'est peut-être surprenant, mais beaucoup d'algorithmes fonctionnent en tirant des solutions au hasard et en les combinant d'une façon intelligente. La théorie des probabilités est donc essentielles pour concevoir et analyser ces algorithmes.

3.1. Notions de base.

Intuitivement la question de base qu'on se pose en théorie des probabilités c'est : quel est l'ensemble des évènements possibles (c'est-à-dire des choses qui peuvent se passer), et "combien de chances" y a-t-il que chacun d'eux se produise. Nous allons formaliser cette idée.

Définition 3.1. *On appelle* univers *l'ensemble des résultats possibles d'une expérience aléatoire, qu'on note souvent Ω.*

Remarque 3.2. Dans la première partie de ce cours on considérera principalement des exemples où Ω est un ensemble fini. Il y a beaucoup de choses qui se simplifient dans ce cas. Néanmoins, dans les définitions qui suivent on ne supposera *pas* qu'il est fini.

Exemple 3.3.
— Si on lance une pièce de monnaie, alors $\Omega = \{\text{pile}, \text{face}\}$.
— Si on lance un dé normal, alors $\Omega = \{1, 2, 3, 4, 5, 6\}$.
— SI on lance un dé à 10 faces numérotées de 0 à 9 une infinité de fois, alors Ω est l'ensemble des suites infinies de chiffre.
— Si on joue à chifoumi, alors $\Omega = \{\text{pierre}, \text{feuille}, \text{ciseau}\}$.

Définition 3.4. *Un évènement associé à un univers Ω est un sous-ensemble de l'ensemble Ω. Un singleton, c'est-à-dire un sous-ensemble qui contient un seul élément, est appelé évènement élémentaire.*

Exemple 3.5. Si on lance un dé
— l'évènement "faire un 6" correspond au singleton $\{6\}$.
— l'évènement "le résultat est pair" correspond au sous-ensemble $\{2, 4, 6\} \subset \{1, 2, 3, 4, 5, 6\}$

Si on lance 2 dés, alors $\Omega = \{1, \dots, 6\}^2$. Dans ce cas :
— l'évènement "la somme des dés vaut 4" correspond au sous-ensemble
$$\{(1, 3), (3, 1), (2, 2)\}$$
— l'évènement "on obtient une paire" correspond au sous-ensemble
$$\{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$$
.

Une fois qu'on a effectué l'expérience "pour de vrai", c'est-à-dire qu'on obtient un élément ω de Ω, on dit qu'un événement A est réalisé si $\omega \in A$. Par exemple si je lance mes deux dés pour de vrai et que j'obtiens $(1, 1)$, alors des deux derniers exemples le premier n'est pas réalisé, et le second est réalisé.

On a une correspondance entre langage usuel/probabiliste, et langage ensembliste :

Définition 3.6. *Soit Ω un univers et $A, B \subset \Omega$ deux évènements.*
— *On dit que A implique B, ou que B est réalisé dès que A l'est, si $A \subset B$.*
— *L'évènement "A ou B" correspond au sous-ensemble $A \cup B$.*
— *L'évènement "A et B " correspond au sous-ensemble $A \cap B$.*
— *L'évènement contraire de A est le complémentaire $\bar{A}$ de A dans Ω.*
— *L'évènement Ω est dit certain, tandis que l'évènement $\emptyset$ est dit impossible.*
— *On dit que A et B sont incompatibles si $A \cap B = \emptyset$.*

Exemple 3.7. Si on lance un dé :
- L'évènement "le résultat est pair OU un multiple de 3" correspond à $\{2,4,6\} \cup \{3,6\} = \{2,3,4,6\}$. L'évènement "le résultat est pair ET un multiple de 3" correspond à $\{2,4,6\} \cap \{3,6\} = \{6\}$.
- L'évènement "le résultat n'est pas pair" correspond au complémentaire de $\{2,4,6\}$, c'est-à-dire $\{1,3,5\}$.
- Les évènements "le résultat est pair" et "le résultat est impair" sont incompatibles.

Définition 3.8. *Soit Ω un univers et $\mathcal{T}$ un sous ensemble de $\mathcal{P}(\Omega)$, c'est-à-dire un ensemble de sous-ensemble de Ω, qu'on appelle ensemble des évènements. On dit que $\mathcal{T}$ est une* tribu *si*

(1) $\Omega \in \mathcal{T}$: l'ensemble des résultats possible est un évènement, l'évènement certain.

(2) Si $A \in \mathcal{T}$, alors son complémentaire $\bar{A}$ appartient à $\mathcal{T}$ aussi. Autrement dit, si un certain ensemble de résultats de l'expérience est un évènement, alors le fait qu'aucun de ces résultats n'arrive est un évènement aussi.

(3) $\mathcal{T}$ est stable par réunion finie ou dénombrable :
- *si $A, B \in \mathcal{T}$, alors $A \cup B \in \mathcal{T}$: si A, B sont des événements, alors "A ou B" est aussi un évènement.*
- *si $(A_i)_{i\in\mathbb{N}}$ est une famille infinie dénombrable d'évènements, alors*

$$\cup_{i\in\mathbb{N}} A_i \in \mathcal{T}.$$

Remarque 3.9. Très souvent, on prendra $\mathcal{T} = \mathcal{P}(\Omega)$, mais pas toujours. On peut aussi, par exemple, toujours choisir $\mathcal{T} = \{\emptyset, \Omega\}$ (c'est facile de vérifier que c'est bien une tribu).

On peut enfin définir la notion de probabilité, c'est-à-dire "combien de chances y a-t-il qu'un évènement se produise". Par définition c'est un nombre entre 0 et 1, ou alors un pourcentage (dire qu'il y a 30% de chances qu'un évènement survienne, c'est dire que sa probabilité est 0.3).

Définition 3.10. *Soit Ω un univers, et $\mathcal{T}$ une tribu sur Ω. Une loi de probabilité est une application P de $\mathcal{T}$ dans l'intervalle $[0,1]$ qui satisfait :*

(1) $P(\Omega) = 1$ (Ω tout entier a par définition 100% de chances d'être réalisé)

(2) Pour toute famille éventuellement infinie $(A_i)_{i\in\mathbb{N}}$ d'évènements incompatibles, c'est-à-dire que $\forall i \neq j, A_i \cap A_j = \emptyset$, on a

$$P(\cup_{i\in\mathbb{N}} A_i) = \sum_{i=1}^{\infty} P(A_i).$$

On appelle le triplet $(\Omega, \mathcal{T}, P)$ un espace probabilisé.

Remarque 3.11. Le deuxième point est en particulier valable pour les familles finies, par exemple pour deux évènements : si $A \cap B = \emptyset$, alors

$$P(A \cup B) = P(A) + P(B).$$

D'une façon générale, on peut déduire des axiomes que :

(1) $P(\bar{A}) = 1 - P(A)$, puisque $\Omega = A \cup \bar{A}$ et $A \cap \bar{A} = \emptyset$ par définition.
(2) En particulier, $P(\emptyset) = 0$.
(3) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
(4) Si $A \subset B$, alors $P(A) \leq P(B)$.

Exemple 3.12. Reprenons en détail l'exemple du lancer d'un dé. Rappelons qu'on a $\Omega = \{1, 2, 3, 4, 5, 6\}$. Le choix naturel est $\mathcal{T} = \mathcal{P}(\Omega)$. On vérifie qu'il existe une unique loi de probabilité telle que

$$\forall 1 \leq i \leq 6, P(\{i\}) = \frac{1}{6}.$$

Les règles de calcul nous disent que :

— La probabilité que le résultat soit pair, est la somme des probabilités qu'on obtienne 2,4 et 6 respectivement, c'est à dire $\frac{1}{2}$.
— la probabilité de l'évènement "obtenir 5 ou un nombre pair" est, puisque ces évènements sont incompatibles, $\frac{1}{2} + \frac{1}{6} = \frac{2}{3}$.
— la probabilité de l'évènement "obtenir un nombre pair ou un multiple de 3" est la somme de la probabilité d'obtenir un nombre pair et de la probabilité d'obtenir un multiple de trois, à laquelle on retranche la probabilité d'obtenir un nombre qui est à la fois pair et un multiple de 3 (c'est-à-dire 6). Donc c'est

$$\frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{2}{3}.$$

Remarquons que c'est aussi, sans surprise, la probabilité de l'évènement "obtenir un nombre pair ou un 3" puisque cet évènement correspond au même sous-ensemble de $\mathcal{P}(\Omega)$.

L'exemple précédent illustre deux phénomènes généraux dans le cas où Ω est un ensemble fini. D'abord, la loi de probabilité est caractérisée par sa valeur sur les singletons (les ensembles à un seul élément). Ensuite, on observe que la probabilité d'un évènement A est simplement la somme des probabilités (des singletons formés de) chacun de ses éléments. En particulier, si tous les singletons sont *équiprobables*, c'est-à-dire s'ils ont la même probabilité, alors

$$P(A) = \frac{|A|}{|\Omega|}$$

où $|A|$ désigne le cardinal (le nombre d'éléments) de A.

Une astuce bien utile : quand on cherche la probabilité d'un évènement A, il est parfois plus facile de calculer la probabilité de son complémentaire. Illustrons cette idée sur un exemple classique :

Exemple 3.13 (Paradoxe des anniversaires)**.** Si on réunit 30 personnes dans une pièce, quelle est la probabilité qu'au moins deux d'entre elles aient leur anniversaire le même jour ? Pour simplifier on ne tiendra pas compte des années bissextiles. On a donc 365 jours possibles pour chaque personnes, et donc

$$\Omega = \{1, \dots, 365\}^{30}.$$

On pose $\mathcal{T} = \mathcal{P}(\Omega)$. On fait l'hypothèse (pas tout à fait réaliste) que toutes les dates possibles sont équiprobables. La problème c'est que calculer la probabilité qu'au moins deux personnes ont la même date d'anniversaire revient à calculer la probabilité de beaucoup d'évènements différents (par exemple il faudrait tenir compte de la possibilité que trois d'entre elles ont leur anniversaire le même jour, mais que cinq d'entre elles ont aussi la même date d'anniversaire, un autre jour).

Il est du coup plus simple de calculer la probabilité que toutes les dates d'anniversaires sont différentes. Formellement, l'évènement A qu'on considère est donc l'élément de $\mathcal{T}$, dont les éléments sont tous les ensembles de 30 dates (donc 30 entiers entre 1 et 365) qui sont toutes distinctes. On a donc 365 choix possibles pour le premier élément d'un tel ensemble. Puisque le second élément doit être différent du premier, on a plus que 364 possibilités, puis 363 pour le troisième et ainsi de suite. On a donc

$$\begin{aligned}|A| &= 365 \times (365-1) \times \dots (365-29) \\ &= 217103018350855706605753347724808139946552034366767459652335681771929600000000\end{aligned}$$

La probabilité que tous les anniversaires soient distincts est donc

$$\begin{aligned}P(A) &= \frac{217103018350855706605753347724808139946552034366767459652335681771929600000000}{365^{30}} \\ &\approx 0.2936.\end{aligned}$$

Finalement, on en déduit que la probabilité que deux personnes au moins aient leur anniversaire le même jour est

$$P(\bar{A}) = 1 - P(A) \approx 0.7063.$$

Autrement dit il y'a plus de 70% de chances que cela se produise, ce qui est assez surprenant !

3.2. **Combinatoire.** Pour pouvoir calculer des probabilités avec des ensembles finis, on a en général besoin de compter le nombre de manière d'effectuer un certain tirage au sort. Une situation classique est celle où on tire k éléments parmi un ensemble fini E (par exemple 3 boules numérotées dans une urne, ou 5 cartes dans un jeu, etc..). On distingue 4 types de tirages suivant :

— si l'ordre dans lequel les éléments sont tirés compte ou pas
— si le tirage se fait avec ou sans remise, c'est-à-dire si les éléments de E peuvent être tirés plusieurs fois ou non.

Le cas le plus simple est celui d'un tirage avec remise, où l'ordre compte. Dans ce cas si E a n éléments, on a n^k tirages possibles.

Définition 3.14. *Soit k un entier et E un ensemble fini de cardinal n. Un k-arrangement de E est une façon de choisir k éléments dans E telle que :*

— *on fait un tirage sans remise, c'est-à-dire que chaque élément de E peut être choisi au plus une fois*

— *l'ordre compte*
On note A_n^k le nombre de k-arrangement de E.

Proposition 3.15. *Si $k > n$, alors $A_n^k = 0$. Sinon,*

$$A_n^k = \frac{n!}{(n-k)!}.$$

Définition 3.16. *Soit E un ensemble fini de cardinal n. Une k-combinaison de E est simplement un sous-ensemble de E de cardinal k. AUtrement dit c'est une façon de choisir k éléments de E, sans remise, mais cette fois sans tenir compte de l'ordre. On note C_n^k le nombre de k-combinaisons.*

Remarque 3.17. On dit souvent k parmi n pour parler de C_n^k. On appelle aussi ces nombres "coefficients bnomiaux".

Proposition 3.18. *Si $k > n$, $C_n^k = 0$. Sinon*

$$C_n^k = \frac{n!}{(n-k)!k!}.$$

Exemple 3.19. Au loto, on choisit 5 numéros parmi 49, et un numéro "chance" parmi 10. C'est donc un tirage sans remise, et l'ordre ne compte pas, on a donc

$$10 \times C_{49}^5 = 19\,068\,840$$

combinaisons possibles. Dans la version précédente, on choisissait 6 numéros parmi 49, soit un total de 13 983 816 combinaisons. On remarque qu'on a moins de chances de gagner avec la nouvelle formule.

Exemple 3.20. À la belote, chaque joueur ou joueuse a une main de 8 cartes dans un jeu de 32. L'ordre n'a ici pas d'importance, donc on a

$$C_{32}^4 = 10518300$$

mains possibles.

Finalement, on peut procéder à un tirage avec remise, et sans tenir compte de l'ordre. Cette situation est plus compliquée puisque on devrait diviser par le nombre de façon d'ordonner les objets qu'on a tirés, mais ce nombre dépend du tirage ! Il faut donc être un peu astucieux.

Proposition 3.21. *Soit E un ensemble fini de cardinal n. Le nombre de façons de tirer k éléments de E, avec remise, est*

$$D_n^k = C_{n+k-1}^k.$$

Démonstration. L'idée c'est de représenter un tirage de la façon suivante : on associe à chaque élément de E le nombre de fois qu'il a été tiré. Par exemple si $E = \{1, 2, 3, 4, 5\}$ et qu'on a tiré $1, 2, 1, 5$, on représente ça par la suite

$$2, 1, 0, 0, 1.$$

On peut aussi représenter ce tirage de la façon suivante :

$$**|*|||*$$

où les barres séparent des "cases" associé à chaque élément de E, et le nombre d'étoile dans chaque case nous dit combien de fois cet élément a été tiré. On a donc, en général, $n-1$ barres et k étoiles. Le problème se ramène donc à choisir k éléments sans ordre et sans remise (la position des étoiles) parmi $n-1+k$ (la longueur de la suite de symboles). C'est donc bien C_{n-1+k}^k. □

Exemple 3.22. Supposons que je lance 4 dés de couleurs différentes. Dans ce cas le nombre de résultats possibles est $6^4 = 1295$. Si je suppose maintenant que les dés sont tous exactement identiques, et que je lance les 4 en même temps, il y a seulement $D_6^4 = 126$ possibilités.

Exemple 3.23. Combien y a t il de dominos dans un jeu ? Chaque domino est un tirage avec répétitions et sans ordre de deux éléments dans l'ensemble

$$\{\text{blanc}, 1, 2, 3, 4, 5, 6\}.$$

Il y en a donc $D_n 7^2 = 28$.

Exemple 3.24. Les formules qu'on a donné couvrent les cas de base, mais on peut ensuite les combiner. Par exemple, combien y a t il à la belote de mains contenant exactement deux as ? C'est le nombre de façons de tirer deux cartes parmi les quatre as, et simultanément 6 cartes parmi les 28 cartes restantes. C'est donc

$$C_5^2 \times C_{28}^6 = 6 \times 3108105 = 18648630.$$

Comme pour le paradoxe des anniversaires, il est parfois plus facile de trouver le cardinal de l'évènement complémentaire. Par exemple, pour compter le nombre de mains contenant au moins un trèfle, il 'est plus facile de compter plutôt le nombre de mains qui ne contiennent aucun trèfles : cela revient à choisir 4 cartes parmi les 24 qui ne sont pas des trèfles, ce qui donne $C_{24}^8 = 735471$. Le nombre de mains qui contiennent au moins un trèfle est donc

$$C_{32}^8 - C_{24^8} = 10518300 - 735471 = 9782829.$$

On en est là.

Une fois qu'on sait compter les différentes façon de faire un certain type de tirage, on peut calculer la probabilité que ce type de tirage se produise. Quand

on est dans une situation où certains des objets sont identiques, il est souvent pratique pour faire les calculs de d'abord faire comme si on savait les distinguer, puis de diviser ensuite par le nombre de façons de les identifier. Dit autrement, quand on cherche à calculer des probabilités dans une situation concrète, il y a potentiellement plusieurs façon de *modéliser* le problème, c'est-à-dire en gros de choisir l'ensemble Ω et une loi de probabilités, et en général il y a un modèle qui est plus pratique que les autres, typiquement un où les évènements élémentaires sont équiprobables.

Exemple 3.25. Une urne contient 5 boules rouges et 4 boules noires. On en tire 3 simultanément (donc sans remise et sans ordre). Quelle est la probabilité d'avoir une boule rouge et deux boules noires ? On a 9 boules au total, et on prétend d'abord qu'on peut les distinguer. Il y a donc

$$C_9^3 = 84$$

tirages possibles. Évidemment ces tirages ne sont justement pas *vraiment* tous différents, si on remplace une boule noire par une autre boule noire on obtient en pratique le même tirage, mais ça simplifie les calculs d'imaginer qu'ils le sont, par exemple on peut imaginer qu'on a collé une étiquette avec un numéro unique sur chaque boule. Il est donc important de comprendre que dans notre façon de modéliser ce problème il y a 84 tirages possibles, qui ont tous la même probabilité, ce qui n'est pas la même chose que le nombre de *résultats* possibles ! En effet il n'y a que 4 résultats possibles :

- 3 boules noires
- 1 noire et 2 rouges
- 2 noires et 1 rouge
- 3 rouges.

Mais ces résultats ont des probabilités différentes d'arriver, il est donc plus simple de modéliser le problème comme on l'a fait en comptant tous les tirages comme si les boules étaient toutes distinctes. Parmi ces tirages, on compte ensuite le nombre de tirages donnant une rouge et deux noires :

$$C_5^1 \times C_4^2 = 5 \times 6 = 30.$$

La probabilité recherchée est donc $30/84 \approx 0.357$.

3.3. **Indépdendances et probabilités conditionnelles.** Étant donnés deux évènement A et B, on cherche à comprendre la relation entre le fait que A soit réalisé et le fait que B soit réalisé. Par exemple ; si je tire une carte dans un jeu, les évènements "c'est un trèfle" et "c'est un as" n'ont pas de "lien" entre eux : on dit qu'ils sont indépendants. Autrement dit, savoir que c'est un trèfle ne vous donne pas d'indice pour savoir si c'est un as ou non. En revanche, si je vous demande de deviner quelle carte c'est, savoir que c'est un trèfle augmente évidemment vos chances !

On suppose fixé un univers Ω, une tribu $\mathcal{T}$ et une loi de probabilité P.

Formellement, on définit l'indépendance de la façon suivante :

Définition 3.26. *Soient $A, B \in \mathcal{T}$. On dit que A et B sont indépendants si*

$$P(A \cap B) = P(A)P(B).$$

Si on veut savoir "à quel point" A déend de B, on utilise :

Définition 3.27. *On suppose que* $P(B) > 0$. *On appelle* "probabilité de A sachant B" *et on note* $P(A|B)$ *le nombre*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Intuitivement, cette formule nous dit : si on sait déjà que l'évènement B est réalisé, alors la seule partie de A qui peut aussi être réalisé c'est celle qui est commune avec B, et on veut faire en sorte que $P(B|B) = 1$.

Toujours sous l'hypothèse que $P(B) > 0$, alors A et B sont indépendants si et seulement si $P(A|B) = P(A)$, autrement dit A et B sont indépendants si savoir que B est réalisé ne donne pas d'information sur A (et vice versa).

Proposition 3.28. *L'application de* $\mathcal{T}$ *vers* $[0, 1]$ *donnée par* $A \mapsto P(A|B)$ *est elle même une loi de probabilité.*

Démonstration. On vérifie d'abord que cette application est bien définie. On a clairement $P(A|B) \geq 0$, et puisque $P(A \cap B) \leq P(B)$ on a aussi $P(A|B) \leq 1$. Ensuite on a bien

— $P(\Omega|B) = 1$

— si $(A_i)_i \in \mathbb{N}$ sont des ensembles mutuellement disjoints, alors

$$\begin{aligned} P(\bigcup_{i\in\mathbb{N}} A_i|B) &= \frac{P(B \cap \bigcup_{i\in\mathbb{N}} A_i)}{P(B)} \\ &= \frac{P(\bigcup_{i\in\mathbb{N}} B \cap A_i)}{P(B)} \\ &= \frac{\sum_{i\in\mathbb{N}} P(B \cap A_i)}{P(B)} \text{ car les } B \cap A_i \text{ mutuellement disjoints} \\ &= \sum_{i\in\mathbb{N}} P(A_i|B). \end{aligned}$$

□

Exemple 3.29. On tire deux boules au hasard, avec remise, dans une urne contenant 10 boules numérotées de 0 à 9. Quelle est la probabilité d'avoir un 3 au premier tirage, si on sait que la somme des numéros des boules vaut 4 ? On a 10^2 tirage possibles. On a 10 tirages tels que la première boule est un 3 : $(3, 0), (3, 1), \ldots$, donc si la probabilité de l'évènement A "la première boule est un 3" est $\frac{10}{160} = \frac{1}{10}$.

On a 3 tirages tels que la somme des boules fait 4 : $(1, 3), (2, 2), (3, 1), (0, 4), (4, 0)$. La probabilité de l'évènement B "la somme des numéros fait 4" est donc $\frac{5}{100}$.

On a $A \cap B = \{(3, 1)\}$ donc $P(A \cap B) = \frac{1}{100}$. Finalement, la probabilité recherchée est

$$P(A|B) = \frac{1/100}{5/100} = \frac{1}{5}.$$

La formule suivante est souvent utile, bien qu'elle découle directement de la définition :

Proposition 3.30 (formule de Bayes). *Soient A, B des évènements de probabilité non nuls. Alors*

$$P(A|B) = P(B|A) \times \frac{P(A)}{P(B)}.$$

Exemple 3.31. On se donne une urne contenant 5 boules bleues et 3 boules vertes, et une autre urne contenant 2 boules bleues et 6 boules vertes. Quelqu'un tire une boule au hasard dans l'un des urnes et me la montre : elle est bleue. Quelle est la probabilité qu'elle vienne de la première urne ?

On pose $A =$"la boule vient de la première urne" et $b =$"la boule est bleue". On veut donc calculer $P(A|B)$, et l'idée c'est que c'est plus facile de calculer $P(B|A)$. En effet cette dernière est simplement la probabilité d'obtenir une boule bleue en faisant un tirage dans la première urne, donc

$$P(B|A) = 5/8.$$

Par ailleurs on a $P(A) = 8/16 = 1/2$, et $P(B) = 7/16$. Donc la formule de Bayes nous donne :

$$P(A|B) = 5/8 * \frac{1/2}{7/16} = 5/7.$$

On remarque que c'est aussi le nombre de boules bleues dans la première urne, sur le nombre total de boules bleues ce qui est logique.

On a aussi la

Proposition 3.32 (Formule des probabilités totales). *Soit $(B_i)_{i\in\mathbb{N}}$ une partition de Ω, et A un évènement. Alors*

$$P(A) = \sum P(A \cap B_i) = \sum_{i\in\mathbb{N}} P(A|B_i)P(B_i).$$

Remarque 3.33. Si $P(B_i) = 0$, alors $P(A|B_i)$ n'est pas défini, mais on pose dans ce cas

$$P(A|B_i)P(B_i) = 0.$$

Démonstration. On pose $A_i = A \cap B_i$. Puisque la famille B_i est une partition de Ω, la famille A_i est une partition de A. En particulier, ils sont deux à deux disjoints, donc la première égalité découle de la définition d'une loi de probabilité. La seconde formule découle directement de la définition de $P(A|B_i)$.

□

Remarque 3.34. Une des applications courante de cette formule est la suivante : soit A, B des évènements, alors

$$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B}).$$

Autrement dit, pour calculer la probabilité que A se produise, il suffit de connaitre la probabilité que A se produise sachant B, et la probabilité que

A se produise sachant que B ne se produit pas. Cette formule permet donc de "découper" le calcul de la probabilité de A selon différentes situation possibles, en pondérant par la probabilité que ces situations se produise.

Exemple 3.35 (Attention : internet war !)**.** Le problème suivant, dit de Monty Hall en référence à une émission de télévision américaine, déchaîne des débats sans fin sur internet !

C'est un jeu qui oppose une présentatrice à une candidate. La candidate est placée devant trois boîtes fermées, dont deux sont vides et dont une contient un prix. La candidate choisit l'une des trois boîtes. La présentatrice ouvre alors l'une des boîtes qui n'est ni celle choisie par la candidate, ni celle contenant le pris. On propose alors à la candidate de changer d'avis et de choisir l'autre boîte restante. La question est : est-ce une bonne idée de changer d'avis ?

Le fait que la présentatrice sait quelle boîte contient le prix, et choisit d'ouvrir une boîte qui ne le contient pas, est une information cruciale qui est souvent sous entendue et source de confusion. La réponse, un peu surprenante, est que la candidate a effectivement intérêt à changer son choix !

C'est contre intuitif : il reste deux boîtes, donc il semble qu'on a une chance sur deux de gagner. Mais en ouvrant une boîte vide, la présentatrice donne en quelque sorte une information à la candidate. En effet, supposons que la candidate décide de changer d'avis. Au départ elle a 2 chances sur 3 de choisir une boîte vide, et 1 chance sur 3 de choisir celle qui contient le prix.

— Si elle a choisi une boîte vide, la présentatrice est forcée d'ouvrir l'autre boîte vide, et donc la candidate gagne forcément en changeant son choix.
— Si elle a choisi la boîte contenant le prix, la présentatrice ouvre n'importe laquelle des boîtes restantes, et cette fois la candidate perd si elle change d'avis.

On constate donc que la candidate a 2 chances sur 3 de gagner si elle décide de changer d'avis !

Essayons de formaliser un peu. Soit A l'évènement "la candidate a choisi la bonne boîte au début" et B l'évènement "la candidate a gagné le prix". Clairement $P(A) = 1/3$, la formule des probabilités totales nous donne

$$\begin{aligned} P(B) &= P(B|A)P(A) + P(B|\bar{A})P(\bar{A}) \\ &= \frac{1}{3}P(B|A) + \frac{2}{3}P(B|\bar{A}). \end{aligned}$$

Maintenant :

— si la candidate ne change pas d'avis, alors $P(B|A) = 1$ par définition, et $P(A|\bar{B}) = 0$.
— si la candidate change d'avis, alors $P(B|A) = 0$ par définition, et $P(A|\bar{B}) = 1$.

Exemple 3.36 (Attention COVID)**.** Un test de détection de la COVID est fiable à 99%. Je fais ce test et il est positif, qulle est la probabilité que je sois éffectivement malade ? On a envie de répondre 99% mais c'est faux ! C'est un biais qui est courant dans beaucoup d'aspect de la vie réelle. La seule chose que l'on sait, c'est que *si* je suis malade, alors le test a 99% de chances d'être positif.

Autrement dit, si on note A "je suis malade" et B "le test est positif", on a

$$P(B|A) = 0.99.$$

Or ce qu'on cherche, c'est au contraire $P(A|B)$, la probabilité que je sois malade sachant que le test est positif. En fait on ne peut pas répondre sans plus d'inromation : à l'heure ou j'écris ces notes de cours, environ 10% de la population est atteinte de la maladie. Comme le test est fiable à 99% également dans le cas où on n'est pas malade, on a aussi

$$P(B|\bar{A}) = 0.99.$$

Par la formule des probabbilités totales, on a

$$P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A}) = 0.99 \times 0.1 + 0.01 \times 0.9 = 0.108.$$

On applique ensuite la formule de Bayes

$$P(A|B) = P(B|A) \times \frac{P(A)}{P(B)} = 0.99 \times \frac{0.1}{0.108} \approx 91\%.$$

Si on fait le même calcul, mais en supposant cette fois qu'il y a seulement 1% de la population qui est malade, alors la probabilité que je sois malade sachant que le test est positif est seulement de 50% ! C'est contre intuitif, pourquoi est-ce que la probabilité que le test donne le bon résultat change selon le nombre de gens qui sont malades ? L'idée c'est que sachant que mon test est positif, moins il y a de chances "dans l'absolu" que je sois malade, plus j'ai de raisons de penser que mon résultat positif est une erreur, et vice versa.