Chapitre 2
Relations

Les relations sont l'une des notions les plus fondamentales associées aux ensembles. Tout le monde a une compréhension intuitive d'une relation : par exemple, si A est l'ensemble de tous les êtres vivants, alors la relation C (« être le fils de ») peut être définie sur les éléments de cet ensemble. On dit que deux êtres humains a et b dans A sont liés par cette relation si b est le fils de a, et l'on note $a \sim_C b$. Dans ce cas, l'ordre est important, ainsi $a \sim_C b$ et $b \sim_C a$ sont incompatibles. Un autre exemple est donné par l'ensemble des entiers et la relation D (pour « divisibilité »), où $a \sim_D b$ si b divise a.

Notre compréhension intuitive d'une relation est rendue formelle dans ce chapitre. De plus, nous allons discuter dans ce chapitre d'un type particulier de relations, les *relations d'ordre*. De telles relations apparaissent dans des domaines variés des mathématiques. Elles ont également un grand nombre de propriétés combinatoires très intéressantes que nous allons voir.

2.1. Relations

Définition 2.1. Une relation R entre les ensembles A et B est un sous-ensemble de $A \times B$:

$$R \subseteq A \times B.$$

On dit que $a \in A$ et $b \in B$ sont liés par R, ce qui se note $a \sim_R b$, si $(a, b) \in R$. Le *domaine* $\mathrm{Dom}(R)$ est l'ensemble de tous les éléments de A qui sont en relation avec certains éléments de B :

$$\mathrm{Dom}(R) = \{a \in A \mid \exists b \in B \colon a \sim_R b\}.$$

L'*image* (ici *range* en anglais) $\mathrm{Ran}(R)$ est l'ensemble de tous les éléments de B qui sont en relation avec certains éléments de A :

$$\mathrm{Ran}(R) = \{b \in B \mid \exists a \in A \colon a \sim_R b\}.$$

Exemple 2.2.

1. Soient A l'ensemble des professeurs de l'EPFL et B l'ensemble des cours donnés à l'EPFL ce semestre. La relation $R = \{(a, b) \mid a \text{ enseigne le cours } b\}$ décrit le lien entre les cours et les professeurs. $\mathrm{Dom}(R)$ est l'ensemble de tous les professeurs qui donnent un cours ce semestre. L'image $\mathrm{Ran}(R)$ est égale à B.
2. Soient $A = \mathbb{Z}$ et $B = \{0, 1\}$. Soit l'ensemble $R \subseteq A \times B$ défini comme l'ensemble de tous les couples (a, b) tels que $b = 1$ si a est un nombre premier et 0 si a est composé d'exactement deux facteurs premiers distincts. Ainsi, par exemple, $(5, 1) \in R$, mais $(7, 0)$, $(6, 1)$ et $(30, 0)$ ne sont pas dans R. Alors $\mathrm{Dom}(R)$ est l'ensemble des entiers qui ont au plus 2 facteurs premiers et $\mathrm{Ran}(R)$ est $\{0, 1\}$.
3. Soient A l'ensemble de tous les entiers naturels et B l'ensemble de tous les sous-ensembles finis de A. La relation $R = \{(a, b) \in A \times B \mid a = \sum_{x \in b} x\}$ représente les partitions de a comme une somme d'entiers naturels distincts. Alors $\mathrm{Dom}(R) = A$, et $\mathrm{Ran}(R) = B$.

◇

Très souvent dans ce cours, nous considérerons des relations définies sur un seul ensemble.

Définition 2.3. Soit $R \subseteq A \times A$ une relation.

(1) R est dite *symétrique* si $(a, b) \in R$ implique $(b, a) \in R$.

(2) R est dite *réflexive* si $(a, a) \in R$ pour tout $a \in A$.

(3) R est dite *transitive* si

$$\forall a, b, c \in A \colon (a, b) \in R \wedge (b, c) \in R \implies (a, c) \in R.$$

(4) R est une *relation d'équivalence* si elle est symétrique, réflexive et transitive.

Exemple 2.4.

1. Supposons que A est l'ensemble de tous les humains. L'ensemble $R = \{(a, b) \in A \times A \mid a \text{ est marié à b}\}$ décrit la relation "être marié". Il s'agit d'une relation symétrique mais pas réflexive. Que signifie la transitivité pour cette relation ?
2. L'ensemble $\{(a, b) \in \mathbb{R} \times \mathbb{R} \mid a \leq b\}$ décrit la relation d'ordre usuelle sur $\mathbb{R}$. Elle n'est pas symétrique, mais elle est réflexive et transitive.
3. Formalisons la relation sur les entiers donnée par $\ll m$ divise $n \gg$. Pour cela, considérons, $A = B = \mathbb{Z}$, et $R = \{(m, n) \in \mathbb{Z} \times \mathbb{Z} \mid m \text{ divise } n\}$. Cette relation n'est toujours pas symétrique, mais elle est réflexive et transitive.
4. Prenons encore $A = B = \mathbb{Z}$ l'ensemble des entiers et $m \in \mathbb{Z}$. La relation de $\ll$ congruence $\gg \equiv$ est définie par $R_m = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid m \text{ divise } a - b\}$. On écrit $a \equiv b \bmod m$ si $(a, b) \in R_m$. Cette fois, il s'agit d'une relation d'équivalence. En effet, $(a, a) \in R_m$, car m divise $a - a = 0$. Ensuite, si m divise $a - b$, il divise aussi $b - a$, et donc R_m est réflexive. Finalement, supposons que $(a, b) \in R_m$ et $(b, c) \in R_m$. Alors m divise $a - b$ et divise aussi $b - c$, il doit donc diviser la somme de ces nombres, c'est à dire $a - c$. On en déduit $(a, c) \in R_m$ et la transitivité de R_m.
5. Soit V un espace vectoriel sur un corps K, et soit U un sous-espace vectoriel de V. On peut définir une relation de congruence sur $V \times V$ par $R_U := \{(a, b) \in V \times V \mid a - b \in U\}$. On dit que $a \equiv b \bmod U$ si $(a, b) \in R_U$. Cette relation est aussi une relation d'équivalence. En effet, U est un sous-espace, il contient donc l'élément 0, et pour tout élément a, il contient aussi son opposé $-a$. Ainsi, $(a, a) \in R_U$, et si $(a, b) \in R_U$, alors on a aussi $(b, a) \in R_U$, ce qui montre que R_U est à la fois réflexive et symétrique. Si (a, b) et (b, c) sont dans R_U, alors $a - b$ et $b - c$ sont dans U. Comme U est un sous-espace, il contient $a - b$ et $b - c$ et leur somme, c'est-à-dire $a - c \in U$ et $(a, c) \in R_U$. Finalement U est bien transitive.
6. Soient $A = \mathbb{R}$, et $R = \{(a, b) \mid b^2 = a\}$. Alors R n'est ni symétrique, ni réflexive, ni transitive.
7. Soit A l'ensemble de tous les polynômes à une variable et à coefficients entiers. On définit la relation R sur A par $(a, b) \in R$ ssi $\deg(a) = \deg(b)$. C'est une relation d'équivalence (preuve laissée en exercice).

◇

Définition 2.5. Soit R une relation sur $A \times B$. Alors, pour tout $a \in A$, l'ensemble

$$[a] := \{b \in B \mid (a, b) \in R\}$$

est appelé la *classe* de a.

Exemple 2.6. Soit A l'ensemble {Analyse, Mathématiques numériques, Théorie des probabilités, Mathématiques discrètes} et B l'ensemble {Lundi, Mardi, Mercredi, Jeudi, Vendredi}. Considérons la table suivante :

	Lundi	Mardi	Mercredi	Jeudi	Vendredi
Analyse	×		×		
Mathématiques numériques		×		×	
Théorie des probabilités		×			×
Mathématiques discrètes				×	

Cette table définit deux relations, l'une, appelons-la R, sur $A \times B$, et l'autre, appelons-la R', sur $B \times A$. Ces relations sont définies de manière évidente : $(a, b) \in R$ ssi il y a une croix dans la case correspondante à l'intersection de la ligne correspondant à a et de la colonne correspondant à b. Ainsi, [Théorie de probabilité] $= \{$Mardi, Vendredi$\}$. De manière similaire, $(b, a) \in R'$ ssi $(a, b) \in R$. Ainsi, par exemple, [Mercredi] $= \{$Analyse$\}$. $\diamond$

Les classes d'une relation d'équivalence ont des propriétés intéressantes : Deux telles classes sont soit disjointes (c'est-à-dire d'intersection vide), soit égales. Ce résultat, qui est montré plus bas, n'est en général pas vrai pour les relations. Ainsi, dans l'exemple précédent, [Théorie des probabilité] et [Mathématiques numériques] ont une intersection non vide $\{$Mardi$\}$ et ne sont pas égales.

Exemple 2.7. Considérons la relation de congruence R_m définie dans l'exemple 2.4(4). Alors, pour tout $a \in \mathbb{Z}$, on a $[a] = \{a + mz \mid z \in \mathbb{Z}\}$. Si $a \equiv b \bmod m$, alors $a - b$ est divisible par m, disons $a - b = zm$, et donc $a = b + zm$, de telle manière que $a \in [b]$. Puisque $b = a - zm$, $b \in [a]$ et donc $[a] = [b]$. D'un autre côté, si $a \not\equiv b \bmod m$, alors $[a]$ et $[b]$ sont disjointes : sinon, si $x = a + zm = b + z'm$, alors $a - b = (z' - z)m$, et donc $a \equiv b \bmod m$, une contradiction. Il se trouve que les classes distinctes sont dans ce cas $[0], [1], \ldots, [m-1]$. $\diamond$

Pour prouver le résultat que nous venons de mentionner sur les relations d'équivalences nous avons besoin d'une définition :

Définition 2.8. Soit S un ensemble. Une *partition* de S est un sous-ensemble $\Pi \subseteq P(S)$ de l'ensemble des parties de S tel que

(1) Pour tout $A \in \Pi$: $A \neq \emptyset$,

(2) Pour tout $A, B \in \Pi$, si $A \neq B$, alors $A \cap B = \emptyset$.

(3) $S = \cup_{A \in \Pi} A$.

Proposition 2.9. *Soit R une relation d'équivalence sur A.*

(1) *Si $a \sim_R b$, alors $[a] = [b]$.*

(2) *Si $a \not\sim_R b$, alors $[a] \cap [b] = \emptyset$.*

(3) *Les ensembles distincts parmi les classes $[a]$ pour $a \in A$ forment une partition de A.*

Démonstration. (1) Si $a \sim_R b$, alors $b \in [a]$, et par symétrie on a aussi $b \sim_R a$ et $a \in [b]$. Si $x \in [a]$, alors $a \sim_R x$ et comme $b \sim_R a$, par transitivité, on a $b \sim_R x$. Finalement, $[a] \subseteq [b]$. Par symétrie, l'autre inclusion se montre de manière similaire et $[a] = [b]$.

(2) Supposons que $a \not\sim_R b$ et qu'il existe $x \in [a] \cap [b]$. Alors $a \sim_R x$ et $b \sim_R x$. Par symétrie, $x \sim_R b$, et par transitivité, $a \sim_R x$ et $x \sim_R b$ impliquent $a \sim_R b$, et donc $[a] = [b]$, ce qui est une contradiction.

(3) Soit $I \subseteq A$ tel que les classes $[a], a \in I$ soient toutes distinctes et représentent toutes les classes de R. (I est appelé un ensemble de représentants de classes.) On prouve d'abord que $A = \cup_{a \in I} [a]$. Soit $x \in A$. Alors, comme $(x, x) \in R$ par symétrie, $x \in [x]$, ce qui montre l'assertion. Comme par (2) les classes $[a]$ et $[b]$ pour $a, b \in I$, $a \neq b$, sont disjointes, on voit que $\{[a] \mid a \in I\}$ forment une partition de A. $\square$

Définition 2.10. Soit R une relation d'équivalence sur un ensemble A. Alors tout sous-ensemble $I \subseteq A$ tel que les classes $[a]$, $a \in I$ forment une partition de A est appelé *un ensemble de représentants de classes* pour R.

Les relations d'équivalence sont les mêmes objets que les partitions comme le montre le théorème suivant.

Théorème 2.11. *Soit A un ensemble. Il y a une bijection entre les relations d'équivalence sur A et les partitions de A.*

Démonstration. Soient L l'ensemble de toutes les relations sur A, et P l'ensemble de toutes les partitions de A. Définissons l'application $f \colon L \to P$ par $f(R) = \{[a] \mid a \in I\}$, où I est un ensemble de représentants des classes de R. Si on peut trouver une application $g \colon P \to L$ telle que $g(f(R)) = R$ pour tout $R \in L$ et $f(g(\Pi)) = \Pi$ pour tout $\Pi \in P$, alors d'après l'exercice 2.1 on vient de montrer la bijectivité de f. En effet, il est facile de vérifier que $g \circ f$ est injective, ce qui implique que f est injective, et que $f \circ g$ est surjective, ce qui implique que f est surjective.

Soit $\Pi = \{A_1, \ldots, A_t\}$ une partition de A. On définit la relation $g(\Pi) = R_\Pi$ par

$$a \sim_{R_\Pi} b \iff \exists i \colon a \in A_i \wedge b \in A_i.$$

En d'autres termes, a est lié à b si tous deux sont dans le même ensemble de la partition. On affirme que R_Π est une relation d'équivalence. La symétrie et la réflexivité étant faciles à voir, on va se concentrer sur la transitivité. Supposons que $(a,b) \in R_\Pi$ et $(b,c) \in R_\Pi$. Alors il existe un i tel que $a, b \in A_i$ et il existe un j tel que $b, c \in A_j$. Puisque $A_i \cap A_j = \emptyset$ pour $i \neq j$ (cela découle de la définition d'une partition), on en déduit $i = j$, et donc $(a,c) \in R_\Pi$.

Remarquez que les classes de R_Π sont par définition précisément les ensembles A_i, et donc $f(g(\Pi)) = \Pi$. D'autre part, si R est une relation d'équivalence et que $\Pi = f(R)$ est l'ensemble des classe distinctes de R, alors $g(\Pi) = R$, ce dont on peut se convaincre après une courte réflexion. □

Si $R \subseteq A \times A$ est une relation sur l'ensemble A, alors l'ensemble des classes distinctes de R a un nom particulier.

Définition 2.12. Soient A un ensemble et $R \subseteq A \times A$ une relation d'équivalence sur A. Alors l'ensemble des classes de R est noté A/R et est appelé l'ensemble quotient de A par rapport à R.

Exemple 2.13. Soient $A = \mathbb{Z}$, n un entier non nul et R_m la relation de congruence définie dans l'example 2.4 (4). Alors $\mathbb{Z}/R_m$ est formé des m éléments $\{i + m\ell \mid m \in \mathbb{Z}\}$ pour $i = 0, 1, \ldots, m-1$. L'ensemble $\mathbb{Z}/R_m$ et usuellement noté $\mathbb{Z}/m\mathbb{Z}$ dans la littérature.

2.2. Ensemble partiellement ordonné (poset)

Définition 2.14. Soient A un ensemble et R une relation sur cet ensemble.

1. R est appelée *une relation d'ordre* ou un *ordre partiel* si les conditions suivantes sont vérifiées :
   (a) Pour tout $a \in A$ on a $(a,a) \in R$ (réflexivité).
   (b) Pour tout $a, b \in A$ si $(a,b) \in R$ et $(b,a) \in R$ alors $a = b$ (antisymétrie).
   (c) pour tout $a, b, c \in A$ si $(a,b), (b,c) \in R$, alors $(a,c) \in R$ (transitivité).
2. R est appelé un *ordre total* si c'est un ordre partiel et si pour tous $a, b \in A$ tels que $a \neq b$, soit $(a,b) \in R$, soit $(b,a) \in R$.
3. Si R est un ordre partiel sur A, alors on appelle la paire $\mathcal{P} = (A, R)$ un *ensemble partiellement ordonné* ou un *poset* et on dit que A est un ensemble partiellement ordonné par rapport à R. Si R est un ordre total, alors on appelle $\mathcal{P}$ un *ensemble totalement ordonné* et on dit que A est totalement ordonné par rapport à R.
4. Si $\mathcal{P} = (A, R)$ est un poset, on écrit $a \leq_\mathcal{P} b$ pour $(a,b) \in R$, et on écrit $a <_\mathcal{P} b$ si $a \leq_\mathcal{P} b$ et $a \neq b$. De plus, on écrit $a \in \mathcal{P}$ si $a \in A$.
5. Si a et b sont des éléments d'un poset $\mathcal{P}$ et $a \not\leq_\mathcal{P} b$ et $b \not\leq_\mathcal{P} a$, alors a et b sont dits *incomparables* ; sinon, il sont dits *comparables*.

Il y a de nombreux exemples d'ensembles partiellement ordonnés que vous connaissez déjà. En voici une petite liste :

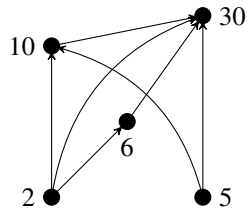
Exemple 2.15.
1. L'ensemble $\mathbb{Z}$ muni de l'ordre naturel est totalement ordonné.
2. L'ensemble $\mathbb{N}$ est partiellement ordonné par la relation de divisibilité. On note ce poset par $(\mathbb{N}, |)$.
3. L'ensemble $\mathbb{Z}$ n'est pas partiellement ordonné par rapport à la divisibilité. Soit a un élément non nul. Alors a divise $-a$, et $-a$ divise a. Mais a n'est pas égal à $-a$, et donc la propriété d'antisymétrie n'est pas valide.
4. Si T est un ensemble, l'ensemble de ses parties $P(T)$ est partiellement ordonné par l'inclusion. On note ce poset $(P(T), \subseteq)$.
5. Supposons que A soit un ensemble totalement ordonné par la relation $\leq$, et supposons que n est un entier positif. Alors l'ensemble A^n est totalement ordonné par la relation $\leq_{\text{lex}}$ définie ci-dessous :
$$(a_1, \ldots, a_n) <_{\text{lex}} (b_1, \ldots, b_n) \iff \exists\, i \in [1,n] : \quad a_1 = b_1, b_2 = a_2, \ldots, a_{i-1} = b_{i-1}, a_i < b_i.$$
Cet ordre est appelée l'ordre *lexicographique*.
6. Soit A l'ensemble totalement ordonné de l'exemple précédent. Alors A^n est partiellement ordonné par l'ordre suivant :
$$(a_1, \ldots, a_n) \leq (b_1, \ldots, b_n) \iff \forall i \geq 1 : \quad a_i \leq b_i.$$

◇

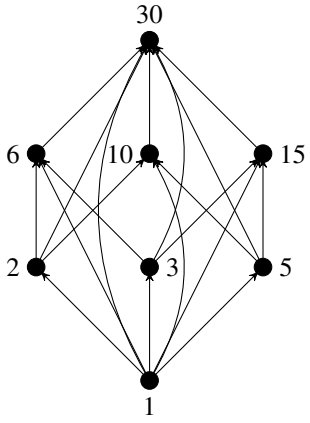
On peut représenter visuellement les posets comme suit : on fait correspondre chaque élément du poset à un cercle (un "sommet") et on dessine une flèche (une "arête") du sommet a au sommet b si et seulement si les éléments correspondants du poset vérifient $a \leq b$. On remarque que le réflexivité de la relation revient à dessiner des "boucles" sur tous les sommets, c'est-à-dire des flèches qui partent d'un sommet et arrivent au même sommet. Une telle structure est appelée "graphe orienté" ou DAG (de l'anglais "directed acyclic graph"). Nous reparlerons des graphes dans le chapitre 3.

Quand on dessine le DAG d'un poset, on omet usuellement la réflexivité et donc on ne dessine pas de boucles sur les sommets du graphe. L'exemple suivant illustre cela.

Exemple 2.16. (1) Soient $A = \{2, 5, 6, 10, 30\}$ et R la relation de divisibilité sur A. son DAG est donné ci-dessous :



(2) Soient $A = \{1, 2, 3, 5, 6, 10, 15, 30\}$ et R la relation de divisibilité sur A. La représentation par DAG de ce poset est donné ci-dessous :



(3) En général, soit n un entier positif. On note $(\text{Div}(n), |)$ le poset dont les éléments sont tous les diviseurs de n, et la relation est celle de divisibilité. La partie (2) de cet exemple est le DAG de $(\text{Div}(30), |)$.

◇

Le lecteur s'est peut-être déjà aperçu que la représentation par un DAG d'un poset contient un grand nombre de redondance. Par exemple, dans le DAG de la première partie de l'exemple précédent, l'arête de 2 à 30 est superflue puisqu'il y a déjà une arête de 6 à 30 et une entre 2 et 6. L'arête entre 2 et 30 est implicite. La définition suivante rend cela plus précis :

Définition 2.17. Soit $\mathcal{P}$ un poset.

1. Pour $a, b \in \mathcal{P}$ tel que $a \leq_{\mathcal{P}} b$ on appelle b un *successeur* de a et a un *prédécesseur* de b.
2. Un élément de $a \in \mathcal{P}$ est dit *minimal* s'il n'a aucun prédécesseur. Un élément de $b \in \mathcal{P}$ est dit *maximal* s'il n'a aucun successeur.
3. Un *prédécesseur immédiat* de $a \in \mathcal{P}$ est un élément $b \in \mathcal{P}$ tel que $b \leq_{\mathcal{P}} a$ et tel qu'il n'existe pas de $c \neq b$ distinct de a tel que $b \leq_{\mathcal{P}} c \leq_{\mathcal{P}} a$.

Exemple 2.18. (1) Dans l'exemple 2.16 (1), les éléments 2 et 5 n'ont pas de prédécesseur, ils sont donc minimaux. L'élément 30 n'a pas de successeur, il s'agit donc d'un élément maximal. Les éléments 6 et 10 sont des prédécesseurs immédiats de l'élément 30.

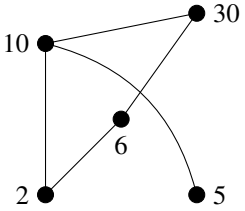
(2) Dans l'exemple 2.16 (2), le poset a exactement un élément minimal 1 et exactement un élément maximal 30.

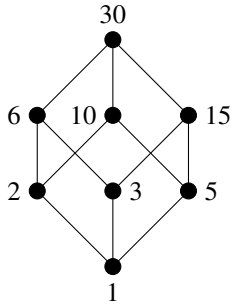
◇

Les prédécesseurs immédiats n'existent pas dans tous les posets. Par exemple, aucun des éléments du poset $\mathcal{P} = (\mathbb{R}, \leq)$ formé des nombres réels muni de l'ordre naturel n'admet de prédécesseur immédiat.

Les prédécesseurs immédiats existent de manière évidente pour les posets finis. Pour de tel posets, il y a une représentation très pratique qui utilise les *diagrammes de Hasse*. Comme pour la représentation par DAG, on dessine les éléments du poset comme les sommets d'un graphe, mais on ne relie l'élément a à l'élément b que si a est un prédécesseur immédiat de b. On oublie également l'orientation des arêtes ; à la place, si $a \leq_{\mathcal{P}} b$, alors on dessine le sommet correspondant à a sous le sommet correspondant à b. Le diagramme final est alors équivalent à tout le poset.

Exemple 2.19. Les diagrammes de Hasse des posets des exemples 2.16 (1) et (2) sont donnés ci-dessous :





◇

Précisons à présent le sens de certaines notions que vous avez déjà dû utiliser par ailleurs.

Définition 2.20. Soient $\mathcal{P} = (A, \leq)$ un poset et $C \subseteq A$ une partie de A.

1. On appelle *minorant* de C tout élément m de A tel que pour tout $c \in C$, $m \leq_{\mathcal{P}} c$.
2. On appelle *majorant* de C tout élément M de A tel que pour tout $c \in C$, $c \leq_{\mathcal{P}} M$.
3. On dit d'un élément x de C qu'il est *minimal* dans C s'il n'a pas de prédécesseur dans C, i.e. si pour tout $c \in C$, $c \leq x$ implique $c = x$.
4. On dit d'un élément y de C qu'il est *maximal* dans C s'il n'a pas de successeur dans C, i.e. si pour tout $c \in C$, $c \geq x$ implique $c = x$.
5. On appelle *plus petit élément* ou *minimum* de C tout élément m de C tel que pour tout $c \in C$, $m \leq_{\mathcal{P}} c$.
6. On appelle *plus grand élément* ou *maximum* de C tout élément M de C tel que pour tout $c \in C$, $c \leq_{\mathcal{P}} M$.
7. On appelle *borne inférieure* ou *infimum* de C le plus grand des minorants.
8. On appelle *borne supérieure* ou *supremum* de C le plus petit des majorants.

Remarque 2.21. *Lorsqu'ils existent, le minimum, le maximum, la borne inférieure et la borne supérieure sont uniques. Par contre, un ensemble peut contenir plusieurs éléments minimaux et plusieurs éléments maximaux distincts. Lorsqu'il existe, le minimum de C est aussi un minorant de C et la borne inférieure de C ; de même, lorsqu'il existe, le maximum de C est aussi un majorant de C et la borne supérieure de C.*

2.3. Chaînes et lemme de Zorn

Définition 2.22. Soit $\mathcal{P}$ un poset. Une *chaîne* de $\mathcal{P}$ est un sous-ensemble qui est totalement ordonné par $\leq_{\mathcal{P}}$. Un élément $x \in \mathcal{P}$ est appelé un *majorant* pour une chaîne C si pour tout $a \in C$ on a $a \leq_{\mathcal{P}} x$.

Le lemme de Zorn, donné ci-dessous, garantit l'existence d'éléments maximaux et minimaux d'un poset qui a certaines propriétés. Malgré son nom de « lemme », il s'agit en fait d'un axiome qui est équivalent à l'axiome de bon ordre et à l'axiome du choix (axiomes dont on ne discutera pas ici).

Théorème 2.23 (Lemme de Zorn)**.** *Un poset pour lequel chaque chaîne admet un majorant contient au moins un élément maximal. De manière équivalente, un poset pour lequel chaque chaîne admet un minorant contient au moins un élément minimal.*

Le lemme de Zorn est un outil puissant pour prouver un bon nombre d'assertions que nous considérons usuellement comme données. Par exemple, on montre que tout espace vectoriel admet une base.

Exemple 2.24. Soit K un corps (par exemple $\mathbb{R}$) et soit V un espace vectoriel sur K. Un ensemble $S \subset V$ est dit *linéairement indépendant* si pour tout nombre $n \in \mathbb{N}$, tous $a_1, \ldots, a_n \in K$ non tous nuls et tous $s_1, \ldots, s_n \in S$ distincts deux à deux, on a $\sum_{i=1}^{n} a_i s_i \neq 0$. Une *base* de V est un sous-ensemble B linéairement indépendant tel que

$$V = \left\{ a \;\middle|\; \exists n \in \mathbb{N}, a_1, \ldots, a_n \in K, b_1, \ldots, b_n \in B : \quad a = \sum_{i=1}^{n} a_i b_i \right\}.$$

Soit L un sous-ensemble indépendant de V (par exemple $\emptyset$). L'ensemble de tous les sous-ensembles indépendants de V contenant L est un poset G par rapport à l'inclusion. (C'est un sous-poset de $(P(V), \subseteq)$.) Supposons que C est une chaîne dans G. Soit alors $T = \cup_{c \in C} c$. Notez que T est un majorant de C (c'est clair) et que $T \in G$ (pourquoi ?). Ainsi, en utilisant le lemme de Zorn, le poset G admet au moins un élément maximal B. Par définition, B est linéairement indépendant. On affirme qu'il s'agit également d'un ensemble de générateurs pour V et donc d'une base. Supposons le contraire et supposons que W est l'espace vectoriel généré par B. Soit alors x un élément de $V \setminus W$. Par maximalité de B, l'ensemble $B \cup \{x\}$ n'est pas linéairement indépendant, donc il existe $n \in \mathbb{N}$, $b_1, \ldots, b_n \in B$ et $a_1, \ldots, a_n, a \in K$ non tous nuls tels que

$$a_1 b_1 + \cdots + a_n b_n + ax = 0.$$

Clairement, $a \neq 0$ car les b_i sont linéairement indépendants. Ainsi, x est une combinaison linéaire des b_i et appartient à W, une contradiction.

2.4. Treillis

Définition 2.25. Soient $\mathcal{P}$ un poset et a, b des éléments de $\mathcal{P}$.

1. Une *borne supérieure* ou *plus petit majorant* ou encore *supremum* de a et b, notée $a \vee b$ est un élément c tel que $a \leq_{\mathcal{P}} c$ et $b \leq_{\mathcal{P}} c$ et tel que si d est un autre élément de $\mathcal{P}$ qui vérifie $a \leq_{\mathcal{P}} d$ et $b \leq_{\mathcal{P}} d$, alors $c \leq_{\mathcal{P}} d$.
2. Une *borne inférieure* ou *plus grand minorant* ou *infimum* de a et b, notée $a \wedge b$ est un élément c tel que $c \leq_{\mathcal{P}} a$ et $c \leq_{\mathcal{P}} b$ et tel que si d est un autre élément de $\mathcal{P}$ qui vérifie $d \leq_{\mathcal{P}} a$ et $d \leq_{\mathcal{P}} b$, alors $d \leq_{\mathcal{P}} c$.
3. Un *treillis* est un poset pour lequel chaque paire d'éléments admet un infimum et un supremum.

La preuve de la remarque suivante est laissé au lecteur.

Remarque 2.26. *Soit $\mathcal{L}$ un treillis. Alors la borne supérieure et la borne inférieure de n'importe quelle paire d'éléments est unique.*

Un treillis n'a pas nécessairement d'élément maximal ou minimal, mais si c'est le cas alors ces éléments sont uniques :

Proposition 2.27. *Soit $\mathcal{L}$ un treillis. Alors $\mathcal{L}$ a au plus un élément minimal et un élément maximal.*

Démonstration. Supposons que $\mathcal{L}$ a deux éléments minimaux a et b. Il s'ensuit que si $c \in \mathcal{L}$ est tel que $c \leq_{\mathcal{L}} a$ alors $c = a$, il en est de même pour b. Soit alors $c = a \wedge b$ la borne inférieure de a et b. Par définition $c \leq_{\mathcal{L}} a$ et $c \leq_{\mathcal{L}} b$, ce qui implique $c = a = b$.

La preuve pour l'unicité d'un élément maximal est analogue. □

Exemple 2.28. (1) Considérons l'ensemble totalement ordonné $\mathbb{R}$ avec l'ordre naturel. Alors $\mathbb{R}$ est un treillis. La borne inférieure de a et b est $\min(a, b)$, et la borne supérieure de a et b est $\max(a, b)$. Néanmoins $\mathbb{R}$ n'admet ni d'élément maximal, ni d'élément minimal.

(2) Considérons le poset $(\mathbb{N}^*, |)$. Il n'est pas totalement ordonné mais c'est un treillis : La borne inférieure de a et b est $\text{pgcd}(a, b)$, et la borne supérieure de a et b est $\text{ppcm}(a, b)$. Ce treillis possède un unique élément minimal 1, mais n'admet pas d'élément maximal.

(3) Soit le poset $(\mathbb{N}, |)$. C'est un treillis et il admet un unique élément maximal, 0. Même si ce treillis a un élément maximal, il contient des chaînes de longueur arbitraire.

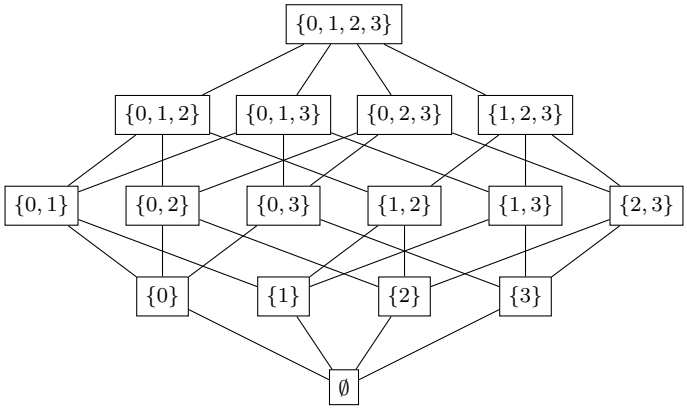


Figure 2.1 – Diagramme de Hasse de $\mathcal{B}_4$.

(4) Considérons maintenant le poset $(\mathrm{Div}(n), |)$ introduit dans l'exemple 2.16 (3). C'est aussi un treillis avec $a \wedge b = \mathrm{pgcd}(a, b)$ et $a \vee b = \mathrm{ppcm}(a, b)$. Il admet un unique élément minimal 1, et un unique élément maximal n. Le diagramme de droite de l'exemple 2.19 est le diagramme de Hasse de $(\mathrm{Div}(30), |)$.

(5) Le treillis $(P(\underline{n}), \subseteq)$ est appelé le *treillis booléen d'ordre* n, et est usuellement noté $\mathcal{B}_n$. Il admet un unique élément maximal, $\underline{n}$, et un unique élément minimal, $\emptyset$. La figure 2.1 montre le diagramme de Hasse de $\mathcal{B}_4$.

◇

2.5. La fonction de Möbius

Soit $\mathcal{L}$ un treillis fini et soit g une fonction $\mathcal{L} \to \mathbb{R}$. On définit la fonction $f : \mathcal{L} \to \mathbb{R}$ par

$$\forall a \in \mathcal{L} : \quad f(a) := \sum_{\mathcal{L} \ni b \leq_{\mathcal{L}} a} g(b).$$

Le but de cette section est de trouver une expression pour g en termes de f.

Définition 2.29. Soit $\mathcal{L}$ un treillis fini. La *fonction de Möbius bivariée* sur $\mathcal{L}$ est la fonction $\mu_{\mathcal{L}} : \mathcal{L} \times \mathcal{L} \to \mathbb{Z}$ caractérisée par les propriétés suivantes :

1. Si $a \not\leq_{\mathcal{L}} b$ alors $\mu_{\mathcal{L}}(a, b) = 0$.
2. Pour tout $a, b \in \mathcal{L}$ avec $a \leq_{\mathcal{L}} b$ on a

$$\sum_{a \leq_{\mathcal{L}} x \leq_{\mathcal{L}} b} \mu_{\mathcal{L}}(x, b) = \begin{cases} 1 & \text{si } a = b \\ 0 & \text{sinon.} \end{cases}$$

On note que cette propriété implique $\mu_{\mathcal{L}}(a, a) = 1$ pour $a \in \mathcal{L}$.

Si $\mathcal{L}$ admet un élément minimal t (et c'est toujours le cas pour un treillis fini), alors la *fonction de Möbius univariée*, notée encore $\mu_{\mathcal{L}}$ est définie par $\mu_{\mathcal{L}}(a) := \mu_{\mathcal{L}}(t, a)$, pour tout $a \in \mathcal{L}$.

Commençons par voir pourquoi les propriétés définies plus haut définissent de façon unique la fonction de Möbius.

Proposition 2.30. *La fonction de Möbius d'un treillis fini $\mathcal{L}$ est définie de manière unique par les propriétés de la définition précédente.*

Démonstration. Supposons qu'il existe deux telles fonctions de Möbius μ_1 et μ_2, et que $a, b \in \mathcal{L}$ sont tels que $\mu_1(a, b) \neq \mu_2(a, b)$. Soit S l'ensemble de tous les $x \leq_{\mathcal{L}} b$ tels que $\mu_1(x, b) \neq \mu_2(x, b)$. S est un poset, il est fini, non vide ($a \in S$) et admet donc un élément maximal. Appelons-le c. Notez que $c <_{\mathcal{L}} b$ puisque $\mu_1(b, b) = \mu_2(b, b) = 1$. On sait par définition de μ que pour $i = 1, 2$

$$\mu_i(c, b) = -\sum_{c <_{\mathcal{L}} x \leq_{\mathcal{L}} b} \mu_i(x, b).$$

Le terme de droite de cette équation est le même pour les deux valeurs de i, car c est un élément maximal de S. Ainsi, le terme de gauche doit être le même aussi, une contradiction. □

On peut maintenant énoncer le théorème principal de cette partie :

Théorème 2.31 (Inversion de Möbius sur les treillis)**.** *Soient $\mathcal{L}$ un treillis fini et $g, f : \mathcal{L} \to \mathbb{R}$ des fonctions telles que pour tout $a \in \mathcal{L}$ on a $f(a) = \sum_{b \leq_{\mathcal{L}} a} g(b)$. Alors pour tout $a \in \mathcal{L}$:*

$$g(a) = \sum_{b \leq_{\mathcal{L}} a} \mu_{\mathcal{L}}(b, a) f(b).$$

De manière similaire, si $f(a) = \sum_{a \leq_{\mathcal{L}} b} g(b)$ pour chaque $a \in \mathcal{L}$, alors pour tout $a \in \mathcal{L}$:

$$g(a) = \sum_{a \leq_{\mathcal{L}} b} \mu_{\mathcal{L}}(a, b) f(b).$$

Démonstration. On a

$$\begin{aligned} \sum_{b \leq_{\mathcal{L}} a} \mu_{\mathcal{L}}(b, a) f(b) &= \sum_{b \leq_{\mathcal{L}} a} \mu_{\mathcal{L}}(b, a) \sum_{c \leq_{\mathcal{L}} b} g(c) \\ &= \sum_{c \leq_{\mathcal{L}} a} \left(\sum_{\substack{b \\ c \leq_{\mathcal{L}} b \leq_{\mathcal{L}} a}} \mu_{\mathcal{L}}(b, a) \right) g(c) \\ &= g(a), \end{aligned}$$

ce qui prouve le théorème. La démonstration de la deuxième forme est similaire. □

Il est parfois utile de considérer une description « duale » de la fonction de Möbius. La définition ci-dessus utilise le fait que la somme de tous les $\mu_{\mathcal{L}}(x, b)$ entre a et b de $\mathcal{L}$ vaut 0. Dans la formulation duale, on montre que la somme de tous les $\mu_{\mathcal{L}}(a, x)$ pour x entre a et b vaut également 0.

Proposition 2.32. *Supposons que $\mathcal{L}$ est un treillis fini avec comme fonction de Möbius $\mu_{\mathcal{L}}$. Soient a et b des éléments de $\mathcal{L}$ avec $a \leq_{\mathcal{L}} b$. On a alors*

$$\sum_{a \leq_{\mathcal{L}} x \leq_{\mathcal{L}} b} \mu_{\mathcal{L}}(a, x) = \begin{cases} 1 & \text{si } a = b \\ 0 & \text{sinon.} \end{cases}$$

De plus, si une fonction $\mu_{\mathcal{L}}$ satisfait cette condition et vérifie de plus $\mu_{\mathcal{L}}(a, b) = 0$ pour tous $a, b \in \mathcal{L}$ tels que $a \not\leq_{\mathcal{L}} b$, alors $\mu_{\mathcal{L}}$ est la fonction de Möbius de $\mathcal{L}$.

Démonstration. (Idée.) Considérons une matrice M dont les lignes et les colonnes sont indexées par les éléments de $\mathcal{L}$, et telle que $M_{a,b} = \mu_{\mathcal{L}}(a, b)$. De plus, considérons la matrice Z dont les lignes et les colonnes sont aussi indexées par $\mathcal{L}$, et telle que $Z_{a,b}$ est à 1 si $a \leq_{\mathcal{L}} b$, et est à 0 sinon. (Ainsi, Z est la matrice des relations du treillis sous-jacent.) On laisse en exercice la preuve que la propriété 2 de la définition 2.29 de la fonction de Möbius est équivalente à dire que ZM est la matrice identité. C'est-à-dire que M est l'inverse de Z, mais alors MZ est aussi la matrice identité. Le coefficient (a, b) de ce produit est nul si $a \neq b$ et vaut 1 si $a = b$. Ce coefficient égale

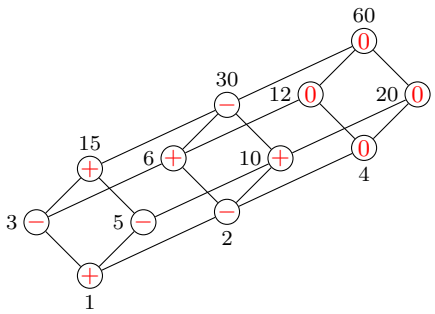
$$\sum_{x \in \mathcal{L}} M_{a,x} Z_{x,b} = \sum_{x \leq_{\mathcal{L}} b} \mu_{\mathcal{L}}(a, x) = \sum_{a \leq_{\mathcal{L}} x \leq_{\mathcal{L}} b} \mu_{\mathcal{L}}(a, x).$$

Ce qui termine (l'idée de) la preuve. □

L'existence de la fonction de Möbius sur tout treillis fini peut être facilement obtenue par récurrence en utilisant cette forme duale. Pour calculer $\mu_{\mathcal{L}}(a, b)$, on commence par $\mu_{\mathcal{L}}(a, a) = 1$. Ensuite, si on connaît tout les $\mu_{\mathcal{L}}(a, c)$ pour $a \leq_{\mathcal{L}} c <_{\mathcal{L}} b$, on définit $\mu_{\mathcal{L}}(a, b)$ par $-\sum_{a \leq_{\mathcal{L}} c <_{\mathcal{L}} b} \mu_{\mathcal{L}}(a, c)$.

Exemple 2.33. Soit $\mathcal{L}$ le treillis $(\mathrm{Div}(60), |)$. Dans cet exemple on va calculer la valeur de $\mu_{\mathcal{L}}(1, x)$ pour plusieurs éléments x dans ce treillis et pour cela on va utiliser la proposition précédente. Premièrement il est clair que $\mu_{\mathcal{L}}(1, 1) = 1$. Ensuite, supposons que x est un nombre premier (dans ce cas, soit 2, 3, ou 5). Alors en utilisant la proposition 2.32 on a $0 = \sum_{y|x} \mu_{\mathcal{L}}(1, y) = \mu_{\mathcal{L}}(1, 1) + \mu_{\mathcal{L}}(1, x)$, et donc $\mu_{\mathcal{L}}(1, x) = -1$. Ensuite, on calcule $\mu_{\mathcal{L}}(1, 4)$: on a $0 = \mu_{\mathcal{L}}(1, 1) + \mu_{\mathcal{L}}(1, 2) + \mu_{\mathcal{L}}(1, 4)$, et donc $\mu_{\mathcal{L}}(1, 4) = 0$.

Le dessin suivant montre le diagramme de Hasse de ce treillis et les valeurs de $\mu_{\mathcal{L}}(1, x)$ Pour tous les éléments du treillis :



Ici, « + » veut dire 1, et « − » veut dire -1. ◇

De manière similaire, on peut calculer la valeur de la fonction de Möbius sur nos deux treillis favoris :

Théorème 2.34. (1) *Soit $\mathcal{L}$ le treillis $(Div(n), |)$. Alors on a pour des nombres naturels a, b avec $a \mid b \mid n$:*

$$\mu(a, b) = \begin{cases} (-1)^t & \text{si } b/a \text{ est le produit de } t \text{ nombres premiers distincts,} \\ 0 & \text{sinon.} \end{cases}$$

(2) *Soient S un ensemble fini et $\mathcal{L}$ le treillis $(P(S), \subseteq)$. Alors on a pour des sous-ensembles $X \subseteq Y \subseteq S$:*

$$\mu_{\mathcal{L}}(X, Y) = (-1)^{|Y \setminus X|}.$$

Démonstration. (1) Pour prouver le théorème, on utilise l'unicité de la fonction de Möbius (cf proposition 2.30). Il nous suffit donc de montrer que la fonction donnée dans l'énoncé satisfait les deux propriétés de la proposition 2.32. La seconde est trivialement vérifiée donc on se concentre sur la première. On doit donc montrer que

$$\sum_{\substack{d \\ a \mid d \mid b}} \mu_{\mathcal{L}}(a, d) = \begin{cases} 1 & \text{si } a = b, \\ 0 & \text{sinon.} \end{cases}$$

Les nombres d tels que $a \mid d \mid b$ sont en bijection avec les diviseurs δ de b/a en écrivant $a\delta = d$. En d'autres termes, le treillis de division formé par les éléments entre a et b est isomorphe à $(\mathrm{Div}(b/a), |)$. La fonction de Möbius prend donc les mêmes valeurs sur les deux treillis. Il faut donc montrer que pour $m \in \mathbb{N}$ (qui correspond à b/a) on a

$$\sum_{\delta \mid m} \mu_{\mathcal{L}}(1, \delta) = \begin{cases} 1 & \text{si } b/a = 1, \\ 0 & \text{sinon.} \end{cases}$$

Pour cela, supposons que $m = \prod_{i=1}^{t} p_i^{e_i}$ où les p_i sont des nombres premiers et $e_i \geq 1$ pour tout i. Ainsi tout $\delta \mid m$ est de la forme $\prod_{i=1}^{t} p_i^{\varepsilon_i}$, avec $0 \leq \varepsilon_i \leq e_i$. En utilisant la formule du théorème, $\mu_{\mathcal{L}}(1, \delta) = 0$ s'il existe un $\varepsilon_i > 1$. On doit donc montrer que

$$\sum_{0 \leq \varepsilon_1 \leq 1, \ldots, 0 \leq \varepsilon_t \leq 1} \mu_{\mathcal{L}}\left(1, \prod_{i=1}^{t} p_i^{\varepsilon_i}\right) = 0 \,.$$

En utilisant encore une fois la définition de μ du théorème, on voit que dans cette somme, $\mu_{\mathcal{L}}(1, \prod_{i=1}^{t} p_i^{\varepsilon_i}) = (-1)^s$ où s est le nombre de ε_i non nuls. L'ensemble des vecteurs $(\varepsilon_1, \ldots, \varepsilon_t)$ qui ont s composantes non nulles est de cardinal $\binom{t}{s}$. On en déduit

$$\sum_{0 \leq \varepsilon_1 \leq 1, \ldots, 0 \leq \varepsilon_t \leq 1} \mu_{\mathcal{L}}\left(1, \prod_{i=1}^{t} p_i^{\varepsilon_i}\right) = \sum_{s=0}^{t} \binom{t}{s} (-1)^s = (1-1)^t = \begin{cases} 1 & \text{si } t = 0, \\ 0 & \text{sinon,} \end{cases}$$

et on a terminé.

(2) On raisonne par récurrence sur $n = |Y \setminus X|$. On commence avec $n = 0$, ce qui est trivial. Supposons maintenant que $|Y \setminus X| = n+1$, et que l'assertion est vraie pour tous les couples Y, Z avec $Z \subseteq Y$ et $|Y \setminus Z| \leq n$. On utilise maintenant le fait que

$$\sum_{X \subseteq Z \subseteq Y} \mu_{\mathcal{L}}(Z, Y) = 0.$$

Mais

$$\begin{aligned} \sum_{X \subseteq Z \subseteq Y} \mu_{\mathcal{L}}(Z, Y) &= \mu_{\mathcal{L}}(X, Y) + \sum_{X \subset Z \subseteq Y} \mu_{\mathcal{L}}(Z, Y) \\ &= \mu_{\mathcal{L}}(X, Y) + \sum_{i=1}^{n+1} \binom{n+1}{i} (-1)^{n+1-i}. \end{aligned}$$

Pour voir cela, supposons que $Y \setminus X = \{b_1, \ldots, b_{n+1}\}$. Alors tous les ensembles possibles Z sont obtenus comme $Z = X \cup T$ où T est un sous-ensemble non vide de $Y - X$. Le nombre de tels sous-ensembles à i éléments est $\binom{n+1}{i}$, et si Z est un tel sous-ensemble, on a $\mu_{\mathcal{L}}(Z, Y) = (-1)^{n+1-i}$ par hypothèse de récurrence.

En continuant avec la dernière expression, on note que

$$\sum_{i=0}^{n+1} \binom{n+1}{i} (-1)^{n+1-i} = (1-1)^{n+1} = 0,$$

ainsi $\mu_{\mathcal{L}}(X, Y) = (-1)^{n+1}$, et le résultat en découle. □

Définition 2.35. Comme la fonction de Möbius sur le treillis $(\mathrm{Div}(n), |)$ ne dépend que du quotient de ses arguments et en particulier ne dépend pas de n, on définit la fonction de Möbius à une variable (c'est la fonction originale) μ de $\mathbb{N}$ dans $\{-1, 0, 1\}$ par

$$\mu(x) := \begin{cases} (-1)^t & \text{si } x \text{ est le produit de } t \text{ nombres premiers distincts,} \\ 0 & \text{sinon.} \end{cases}$$

Ainsi, $\mu(x) = \mu_{\mathcal{L}}(1, x)$ où $\mu_{\mathcal{L}}$ est la fonction de Möbius de $(\mathrm{Div}(n), |)$, pour un n multiple de x.

La preuve de la remarque suivante est laissé en exercice.

Remarque 2.36. *La fonction de Möbius μ est* faiblement multiplicative *: si n et m sont des entiers tels que $pgcd(n, m) = 1$, alors $\mu(nm) = \mu(n)\mu(m)$.*

Corollaire 2.37 (Inversion de Möbius originelle)**.** *Soient $f, g \colon \mathbb{N} \to \mathbb{R}$ des fonctions telles que pour tout $n \in \mathbb{N}$ on a $g(n) = \sum_{d|n} f(d)$. Alors on a pour tout $n \in \mathbb{N}$:*

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d).$$

Démonstration. Soit $\mu_{\mathcal{L}}$ la fonction de Möbius du treillis $(\mathrm{Div}(n), |)$. On a $f(n) = \sum_{d|n} \mu_{\mathcal{L}}(d, n) g(d)$. Notez que $\mu_{\mathcal{L}}(d, n) = \mu_{\mathcal{L}}(1, n/d) = \mu(n/d)$. □

2.6. Exemple : la fonction φ d'Euler

On donne dans cette section une application de la formule d'inversion de Möbius. Pour un entier $n \in \mathbb{N}$ on note $\varphi(n)$ le nombre d'entiers d, $1 \leq d \leq n$, tels que $\mathrm{pgcd}(d, n) = 1$. Cette fonction est appelé la *fonction d'Euler*. Par exemple, $\varphi(10) = 4$, car les seuls entiers plus petits que 10 qui sont premiers avec 10 sont $1, 3, 7, 9$.

On commence par montrer que

$$n = \sum_{d|n} \varphi(d). \tag{2.1}$$

Pour cela, on définit pour chaque diviseur d de n l'ensemble $F_d := \{x \mid 0 \leq x < n, \mathrm{pgcd}(x, n) = d\}$. Il est clair que les ensembles F_d sont disjoints. De plus l'ensemble $\{0, 1, \ldots, n-1\}$ est égal à l'union des F_d : si x est un

élément de l'ensemble précédent, alors $x \in F_d$ avec $d = \text{pgcd}(x, n)$. On en déduit que $n = \sum_{d|n} |F_d|$. On montre maintenant que $|F_d| = \varphi(n/d)$. Pour cela, remarquez que $x \in F_d$ ssi $\text{pgcd}(x/d, n/d) = 1$. Ce qui prouve (2.1).

En appliquant la formule d'inversion de Möbius du corollaire 2.37 à (2.1), on voit que

$$\varphi(n) = \sum_{d|n} \mu(n/d)d.$$

De cette formule, on peut facilement en déduire les faits suivants :

1. si n est un nombre premier, alors $\varphi(n) = n - 1$: Dans ce cas les seuls diviseur sont n et 1, et comme $\mu(n) = -1$, le résultat en découle.
2. Si $n = p^t$ pour un nombre premier p, alors $\varphi(n) = (p-1)p^{t-1}$: dans ce cas les diviseurs de n sont $1, p, \ldots, p^t$, et $\mu(n/d) \neq 0$ seulement si $d = p^{t-1}$ ou $d = p^t$; dans le premier cas $\mu(n/d) = -1$, et dans le second cas $\mu(n/d) = 1$.
3. Si $n = PQ$ où $\text{pgcd}(P, Q) = 1$, alors $\varphi(n) = \varphi(P)\varphi(Q)$: dans ce cas, tout diviseur d de n peut s'écrire de manière unique comme $d = d_1 d_2$ avec d_1 un diviseur de P et d_2 un diviseur de Q (preuve ?). Ainsi

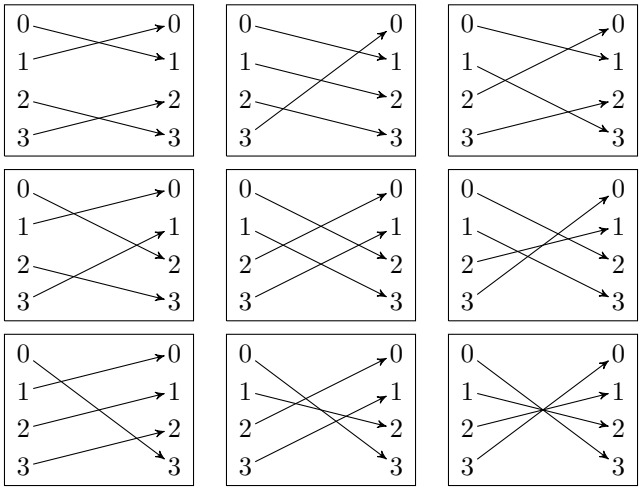
$$\varphi(n) = \sum_{d_1|P, d_2|Q} \mu\left(\frac{PQ}{d_1 d_2}\right) d_1 d_2.$$

Comme P/d_1 et Q/d_2 sont premiers entre eux et que μ est une fonction faiblement multiplicative d'après la remarque 2.36, la dernière expression est égale à

$$\left(\sum_{d_1|P} \mu\left(\frac{P}{d_1}\right) d_1\right)\left(\sum_{d_2|Q} \mu\left(\frac{Q}{d_2}\right) d_2\right) = \varphi(P)\varphi(Q).$$

2.7. Exemple : nombre de dérangements

Un *dérangement* sur $\underline{n}$ est une application bijective de $\underline{n}$ sur lui-même qui ne fixe aucun élément de $\underline{n}$. Par exemple les applications suivantes sont toutes des dérangements de $\underline{4}$:



Dans cette section, on va calculer le nombre de dérangements de $\underline{n}$ en utilisant la formule d'inversion de Möbius sur $\mathcal{B}_n$. Soit S un sous-ensemble de $\underline{n}$, et soit $p_n(S)$ le nombre de bijections qui fixent tous les éléments de S et aucun des éléments hors de S. De plus, soit $q_n(S)$ le nombre d'applications qui fixent tous les éléments de S sans contrainte sur les éléments hors de S. Clairement, $q_n(S) = (n - |S|)!$, puisqu'une application fixant les éléments de S est une permutation quelconque hors de S. De plus, on a

$$q_n(S) = \sum_{S \subseteq T} p_n(T),$$

car les applications fixant S peuvent être partitionnées en celles fixant T et ne fixant aucun élément hors de T, pour tout T contenant S. Par application de la deuxième forme de la formule d'inversion de Möbius, on obtient

$$\begin{aligned} p_n(S) &= \sum_{S\subseteq T} \mu_{\mathcal{B}_n}(S,T) q_n(T) \\ &= \sum_{S\subseteq T} (-1)^{|T\setminus S|}(n-|T|)! \qquad \text{(par le théorème 2.34 (2))} \\ &= \sum_{\ell=|S|}^{n} (-1)^{\ell-|S|}\binom{n-|S|}{\ell-|S|}(n-\ell)!. \end{aligned}$$

La dernière étape est obtenue en choisissant $\ell := |T|$ et en remarquant que le nombre d'ensembles de taille ℓ contenant S est $\binom{n-|S|}{\ell-|S|}$. On est intéressé par la valeur de cette expression pour $S = \emptyset$, puisque dans ce cas $p_n(\emptyset)$ nous donne le nombre de dérangements. On obtient

$$p_n(\emptyset) = n!\left(\frac{1}{2!} - \frac{1}{3!} \pm \cdots + (-1)^n \frac{1}{n!}\right).$$

Avec $\mathrm{e} \simeq 2.7182818\cdots$ qui est le nombre d'Euler, on a encore, selon la parité de n,

$$p_n(\emptyset) = \left\lceil \frac{n!}{\mathrm{e}} \right\rceil \text{ si } n \text{ est pair} \qquad p_n(\emptyset) = \left\lfloor \frac{n!}{\mathrm{e}} \right\rfloor \text{ si } n \text{ est impair,}$$

parce que $\sum_{k=0}^{\infty} \frac{(-1)^k}{k!}$ est une série alternée qui converge vers e^{-1}. Comme la série est alternée, son terme de reste $\sum_{k=n+1}^{\infty} \frac{(-1)^k}{k!}$ est compris entre 0 et $\frac{1}{(n+1)!}$ quand n est impair et $-\frac{1}{(n+1)!}$ et 0 quand n est pair. Comme l'intervalle $[0, \frac{1}{n+1}]$ est de longueur inférieure à 1 et $p_n(\emptyset)$ est entier, on peut se contenter d'arrondir $\frac{n!}{\mathrm{e}}$ vers le bas (respectivement vers le haut) pour retrouver sa valeur quand n est impair (respectivement quand n est pair). En particulier, quand $n = 4$, on obtient

$$p_4(\emptyset) = \frac{4!}{2!} - \frac{4!}{3!} + \frac{4!}{4!} = 12 - 4 + 1 = 9.$$

2.8. Décomposition en chaînes, antichaînes et largeur

Définition 2.38. Soit $\mathcal{P}$ un poset. Une *antichaîne* de $\mathcal{P}$ est une suite $c_1, c_2, \ldots, c_m$ d'éléments de $\mathcal{P}$ telle que pour tout i, j, $i \neq j$, c_i et c_j ne sont pas comparables dans $\mathcal{P}$. La *largeur* d'un poset fini $\mathcal{P}$ est la longueur maximale d'une antichaîne de $\mathcal{P}$.

Exemple 2.39.

1. Considérons le poset $(\underline{10}, |)$. (Notez que ce poset n'est pas egal à $(\mathrm{Div}(10), |)$. Les suites $(1,2,4,8)$, $(3,6)$, $(5,0)$, (7), et (9) forment des chaînes disjointes de ce poset. L'ensemble $\{5,6,7,8,9\}$ est une antichaîne de ce poset.
2. Considérons le treillis $\mathcal{B}_4$. Voici une liste des chaînes disjointes de ce poset :

$$\begin{gathered} (\emptyset, \{0\}, \{0,1\}, \{0,1,2\}, \{0,1,2,3\}) \\ (\{1\}, \{1,2\}, \{1,2,3\}) \\ (\{2\}, \{2,3\}, \{0,2,3\}) \\ (\{3\}, \{1,3\}, \{0,1,3\}) \\ (\{0,2\}) \\ (\{0,3\}). \end{gathered}$$

L'ensemble suivant est une antichaîne :

$$\{\{0,1\}, \{0,2\}, \{0,3\}, \{1,2\}, \{1,3\}, \{2,3\}\}.$$

◇

Soit $\mathcal{P}$ un poset. Un ensemble $\{C_1, \ldots, C_m\}$ de chaînes disjointes de $\mathcal{P}$ tel que $\mathcal{P} = \cup_{i=1}^{m} C_i$ est appelé *une décomposition en chaînes* de $\mathcal{P}$. Une *décomposition minimale en chaînes* de $\mathcal{P}$ est une décomposition pour laquelle le nombre de chaînes est minimal.

Lemme 2.40. *Soit $\mathcal{P}$ un poset.*

(1) *Alors le cardinal de la décomposition minimale en chaînes de $\mathcal{P}$ est au moins égal à la largeur de $\mathcal{P}$.*

(2) *Si w est la largeur de $\mathcal{P}$, $C_1, \ldots, C_w$ est une décomposition en chaînes de $\mathcal{P}$, et A est une antichaîne qui comporte w éléments, alors $|A \cap C_i| = 1$ pour tout $i = 1, \ldots, w$.*

Démonstration. (1) Soient A une antichaîne de $\mathcal{P}$, et $C_1, \ldots, C_t$ des chaînes disjointes couvrant $\mathcal{P}$. Alors pour tout i on a $|A \cap C_i| \leq 1$. Sinon, s'il existe $x, y \in A \cap C_i$ avec $x \neq y$, x et y sont comparables et donc ne peuvent appartenir tous deux à A, une contradiction. Comme les C_i sont disjoints, on a $|A| = \sum_{i=1}^{t} |A \cap C_i|$ et ce dernier nombre est au plus t.

(2) Si $t = w$, alors $w = |A| = \sum_{i=1}^{w} |A \cap C_i| \leq w$, et nous avons une égalité des deux côtés. De plus, $|A \cap C_i| = 1$ pour tout i. □

En fait, le théorème de Dilworth montre que le cardinal d'une décomposition minimale en chaînes et la largeur d'un poset sont les mêmes. La preuve, donnée ci-dessous, construit une décomposition en chaînes du poset et une antichaîne qui sont toutes deux de même taille. En utilisant le lemme précédent, il est alors possible de conclure à l'égalité de la largeur du poset avec le cardinal d'une décomposition minimale en chaînes.

Théorème 2.41 (Théorème de Dilworth)**.** *Soit $\mathcal{P}$ un poset fini. La largeur de $\mathcal{P}$ est égale au cardinal d'une décomposition minimale en chaînes de $\mathcal{P}$.*

Démonstration. La preuve utilise un raisonnement par récurrence sur les éléments du poset $\mathcal{P}$. Si le nombre d'éléments est 1, l'assertion est triviale.

Supposons maintenant que l'assertion soit vraie pour tout poset de n éléments. On aimerait montrer qu'elle est également satisfaite par un poset à $n+1$ éléments. Pour cela, soit $\mathcal{P}$ un poset à $n+1$ éléments. Soit y un élément maximal de $\mathcal{P}$, et considérons le poset $\mathcal{P}' := \mathcal{P} \setminus \{y\}$ avec n éléments. Par hypothèse de récurrence, le cardinal d'une décomposition minimale en chaînes et la largeur de $\mathcal{P}'$ sont les mêmes, disons w.

Soient $C_1, \ldots, C_w$ des chaînes disjointes couvrant $\mathcal{P}'$. On va construire une antichaîne A de $\mathcal{P}'$ avec w éléments de la manière suivante : pour tout i, on choisit dans C_i un élément maximal x_i qui appartient à une antichaîne de longueur w.

Commençons par nous convaincre que les x_i sont deux à deux incomparables. Supposons que i et j sont deux à deux différents et que $x_i \geq x_j$. Soit B une antichaîne de longueur w qui contient x_i. D'après le lemme 2.40 (2) on a $B \cap C_j = \{y_j\}$ pour un certain y_j. On a $y_j \leq x_j$, car y_j appartient à une antichaîne de longueur w, et que x_j est un élément maximal par rapport à cette condition. Comme y_j est dans l'antichaîne B, x_i et y_j ne sont pas comparables, on ne peut donc avoir $x_i \geq x_j$. Comme i et j ont été choisis arbitrairement, on vient de montrer que pour tous i et j les éléments x_i et x_j ne sont pas comparables, donc que A est une antichaîne.

On va maintenant utiliser A pour construire une antichaîne et une décomposition en chaînes de $\mathcal{P}$ de la même taille, ce qui terminera le raisonnement par récurrence. On va distinguer 2 cas.

(1) Supposons d'abord que $y \geq x_i$ pour un certain i. Dans ce cas, on va trouver une décomposition en chaînes de $\mathcal{P}$ avec w éléments et une antichaîne de $\mathcal{P}$ avec w éléments. Le lemme 2.40 (1) nous montre alors que le cardinal d'une décomposition minimale de $\mathcal{P}$ est le même que sa largeur. Soit $C := \{y\} \cup \{z \in C_i \mid z \leq x_i\}$. C est une chaîne. Considérons le poset $\mathcal{P}'' := \mathcal{P} - C$.

On va montrer que $\mathcal{P}''$ n'a pas d'antichaîne de longueur w. Supposons d'abord que x_i est l'élément maximal de C_i. Alors, enlever C de $\mathcal{P}$ est la même chose qu'enlever C_i de $\mathcal{P}'$, de telle manière que $\mathcal{P}''$ admet une décomposition en $w-1$ chaînes, ce qui implique que la largeur de $\mathcal{P}''$ ne peut dépasser $w-1$. Ensuite, supposons que x_i n'est pas l'élément maximal de C_i. Alors, les autres éléments de C_i qui sont plus grand que x_i ne peuvent être dans une antichaîne de longueur w par définition de x_i. Il s'ensuit que la largeur de $\mathcal{P}''$ est d'au plus $w-1$.

On en déduit que $\mathcal{P}''$ peut être couvert par au plus $w-1$ chaînes disjointes et que donc $\mathcal{P}$ peut être couvert par au plus w chaînes disjointes (celles de $\mathcal{P}''$ et C), ainsi la largeur de $\mathcal{P}$ est d'au plus w. Comme A est une antichaîne de $\mathcal{P}$ de taille w, on vient de montrer qu'il y a une décomposition en chaînes de $\mathcal{P}$ de la même taille qu'une antichaîne, ce qui prouve le théorème.

(2) Supposons maintenant que $y \not\geq x_i$ pour tout i. Comme y est un élément maximal de $\mathcal{P}$, y est incomparable avec tous les x_i, et donc $A' := \{y, x_1, \ldots, x_w\}$ est une antichaîne. D'un autre côté, $\{y\}, C_1, \ldots, C_w$ est une décomposition en chaînes de $\mathcal{P}$ de taille $w + 1 = |A'|$, on en déduit donc que dans ce cas également le cardinal d'une décomposition minimale en chaînes de $\mathcal{P}$ est le même que sa largeur.

En regroupant ces deux cas, la preuve est terminée. □

Théorème 2.42 (Théorème de Sperner)**.** *La largeur de $\mathcal{B}_n$ est $\binom{n}{\lfloor n/2 \rfloor}$.*

Démonstration. Premièrement, remarquez que l'ensemble de tous les ensembles de taille $\lfloor n/2 \rfloor$ forme une antichaîne de $\underline{n}$. Il reste donc à prouver qu'il n'existe pas d'antichaîne qui a plus d'éléments. Soit $A = \{S_1, \ldots, S_w\}$ une antichaîne maximale de $\mathcal{B}_n$. Une chaîne maximale de $\mathcal{B}_n$ est une chaîne qu'on ne peut augmenter. Remarquez que $\mathcal{B}_n$ a $n!$ chaînes maximales, une pour chaque permutation de $\underline{n}$. Chacune de ces chaînes intersecte l'ensemble A en au plus un élément. (Sinon A contiendrait au moins deux éléments comparables ce qui n'est pas possible.) Si la taille de S_i est n_i, alors le nombre de chaînes maximales contenant S_i est $n_i!(n - n_i)!$ ce qui correspond au nombre de toutes les permutations de $\underline{n}$ qui laissent l'ensemble S_i invariant. On a alors

$$\sum_{i=1}^{w} n_i!(n - n_i)! \leq n!,$$

car on a à gauche le nombre de chaînes maximales qui contiennent un des éléments de A, et on a à droite le nombre de toutes les chaînes maximales. En divisant les deux côtés par $n!$, on obtient l'inégalité LYM

$$\sum_{i=1}^{w} \frac{1}{\binom{n}{n_i}} \leq 1.$$

Remarquez que $\binom{n}{n_i} \leq \binom{n}{\lfloor n/2 \rfloor}$, et donc

$$w \frac{1}{\binom{n}{\lfloor n/2 \rfloor}} \leq \sum_{i=1}^{w} \frac{1}{\binom{n}{n_i}} \leq 1.$$

On en déduit

$$w \leq \binom{n}{\lfloor n/2 \rfloor},$$

ce qui prouve le théorème. □