CAPITOLO G

Teoria dell'Inversione di Möbius-Rota

L'obiettivo di questo capitolo è di presentare la teoria dell'inversione basata su *insiemi parzialmente ordinati*, con particolare riferimento alle applicazioni algebriche e combinatorie.

G1. Insiemi parzialmente ordinati ed inversione di Möbius

La parte teorica dell'inversione di Möbius costituisce l'argomento di questa sezione. Partendo dai concetti fondamentali di *insiemi parzialmente ordinati*, *reticoli* e di *algebra di incidenza*, si stabiliscono la funzione di Möbius e le proprietà (ortogonalità e relazione ricorrente), le quali sono indispensabili per costruire la teoria centrale dell'inversione, cioè il teorema dell'inversione. Tale teorema ci permette di determinare una funzione, definita su un insieme finito parzialmente ordinato, quando si conoscono le sue somme parziali e la funzione di Möbius corrispondente. Per facilitare il calcolo delle funzioni di Möbius, sono state incluse due formule concernenti il prodotto diretto e l'isomorfismo di insiemi parzialmente ordinati.

G1.1. Insiemi parzialmente ordinati. Il concetto più generale che considereremo in questa sezione è quello di un insieme parzialmente ordinato.

Ricordiamo che una relazione binaria su un insieme S è un sottoinsieme R dell'insieme prodotto $S \times S$. Diciamo che $a \in S$ è in relazione R con $b \in S$ e scriviamo $a\,R\,b$ se e solo se $(a, b) \in R$.

Definizione G1.1 (Poset). *Sia S un insieme su cui è definita una relazione binaria denotata con $\leq$* (oppure con $\leq_S$ quando c'è possibilità di confusione) *soddisfacente i seguenti tre assiomi:*

- *riflessività:* $\forall a \in S: \quad a \leq a.$
- *antisimmetria:* $\forall a, b \in S: \quad a \leq b,\ b \leq a \ \Rightarrow\ a = b.$
- *transitività:* $\forall a, b, c \in S: \quad a \leq b,\ b \leq c \ \Rightarrow\ a \leq c.$

La coppia $(S, \leq)$ *si chiama* insieme parzialmente ordinato *o semplicemente* "poset" (partially ordered set).

- Due elementi a e b di S, distinti ($a \neq b$), si dicono incomparabili se $a \not\leq b$ e $b \not\leq a$. Se $a \leq b$ e $a \neq b$, allora scriviamo $a < b$.
- Inoltre scriviamo $a \geq b$ in alternativa di $b \leq a$ e $a > b$ per $b < a$.
- In generale, se $(S, \leq_S)$ è un insieme parzialmente ordinato, allora ogni sottoinsieme T di S è parzialmente ordinato attraverso la stessa relazione $\leq_S$ di S ristretta a T ($\leq_T$), così definita:

$$\forall a, b \in T: \quad a \leq_T b \quad \Longleftrightarrow \quad a \leq_S b.$$

Così, se $(S, \leq_S)$ è un insieme finito parzialmente ordinato, allora esso ha esattamente $2^{|S|}$ sottoinsiemi parzialmente ordinati attraverso la relazione sopra definita.

Definizione G1.2 (Poset localmente finito). *Sia* $(S, \leq)$ *un insieme parzialmente ordinato:*

- *Un intervallo* (chiuso) *è un sottoinsieme parzialmente ordinato così definito*

$$[a, b] := \Big\{ s \in S \mid a \leq s \leq b \Big\} \quad con \quad a \leq b.$$

- *Analogamente definiamo l'intervallo* (aperto)

$$(a, b) := \{ s \in S \mid a < s < b \} \quad con \quad a \leq b.$$

- *In particolare, si ha* $[a, a] = \{a\}$ *e* $(a, a) = \emptyset$.

Allora, $(S, \leq)$ *si dice* localmente finito *se* $\forall a, b \in S : [a, b]$ *è finito.*

Definizione G1.3 (Cover). *Siano* $(S, \leq)$ *un insieme parzialmente ordinato con* a *e* b *due elementi di* S. *Diciamo che* b è un cover di a *se* $b > a$ *e non esiste* $s \in S$ *tale che* $b > s > a$. *Così,* b *è un "cover" di* a *se* $b > a$ *e* $[a, b] = \{a, b\}$.

Un poset localmente finito $(S, \leq)$ è completamente determinato attraverso le sue relazioni di "covers".

La nozione di "cover" suggerisce un modo di rappresentare un insieme finito parzialmente ordinato $(S, \leq)$ attraverso un *diagramma di Hasse.*

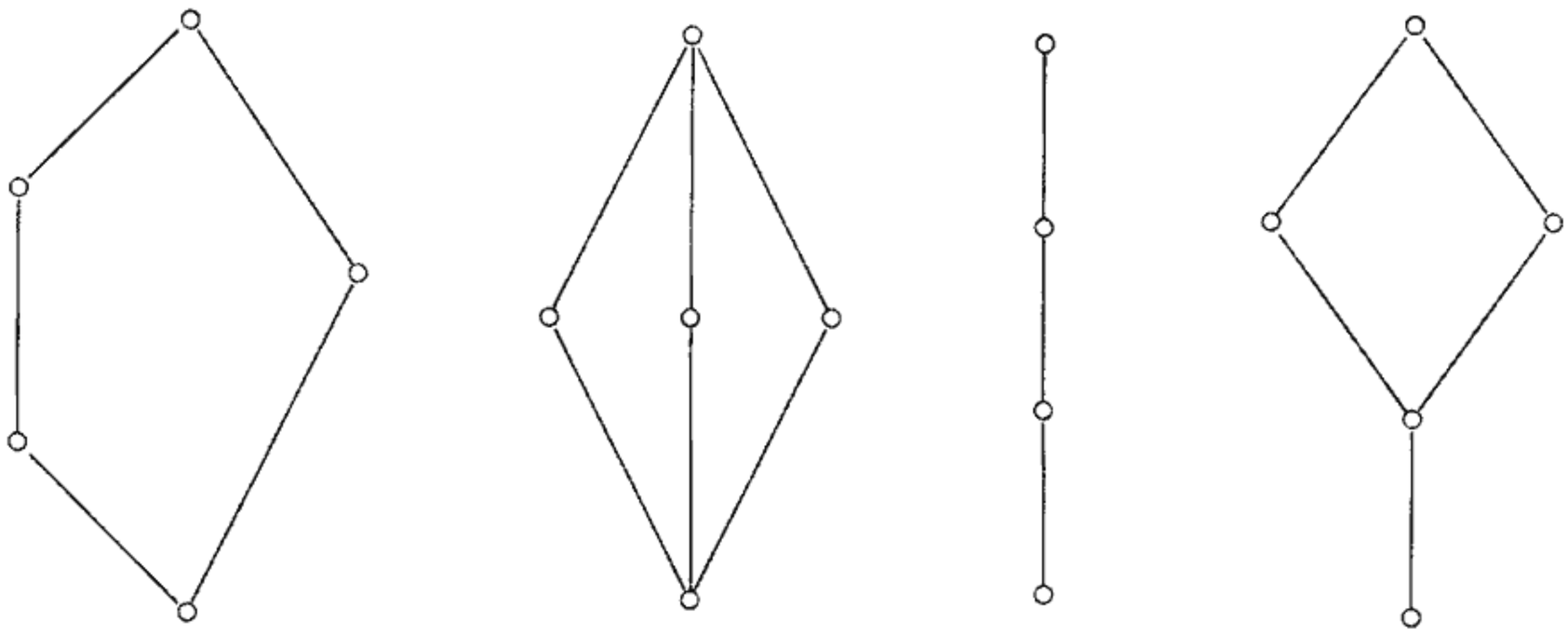
È chiaro che $b > a$ in un insieme finito S se e solo se esiste una sequenza

$$a = s_0, s_1, \cdots, s_n = b$$

tale che ogni s_i è un "cover" di s_{i-1} per $i = 1, 2, \cdots, n$.

Rappresentiamo gli elementi di S nel diagramma di Hasse attraverso punti. Se s_i è un "cover" di $s_{i-1}\,(i = 1, 2, \cdots, n)$, allora collochiamo nel diagramma di Hasse s_i al di sopra di s_{i-1} e congiungiamo i due punti con una linea retta. Allora $b > a$ se e solo se esiste una spezzata discendente che congiunge b ad a. Se nessuna spezzata congiunge b ad $a \neq b$, allora a e b sono incomparabili.

Riportiamo, di seguito, alcuni esempi di diagrammi di Hasse di insiemi finiti parzialmente ordinati.



Definizione G1.4 (Poset totalmente ordinato). *Un insieme parzialmente ordinato* $(S, \leq)$ *si dice* totalmente ordinato *o* catena *se tutti i suoi elementi sono comparabili, cioè se* $\forall a, b \in S : a \neq b \Longleftrightarrow a < b$ *oppure* $b < a$.

- Il terzo diagramma di Hasse sopra riportato rappresenta un insieme finito totalmente ordinato.
- Se $(S, \leq)$ è una catena (insieme finito totalmente ordinato), possiamo immaginarla nella forma $s_0 < s_1 < \cdots < s_n$. In tal caso n è la lunghezza della catena.
- Un'anticatena è un insieme parzialmente ordinato $(S, \leq)$ con tutti gli elementi incomparabili.
- Se un insieme parzialmente ordinato $(S, \leq)$ possiede un minimo (cioè esiste $u \in S$ tale che $\forall s \in S : u \leq s$) e un massimo (cioè esiste $v \in S$ tale che $\forall s \in S : s \leq v$), o soltanto uno di essi, questi vengono indicati usualmente con i simboli 0 e 1 rispettivamente. Gli elementi di S che sono "covers" di 0 si chiamano *atomi*, mentre quelli per i quali 1 è "cover" di quest'ultimi si dicono *coatomi*.

G1.2. Reticoli e proprietà. Un elemento u di un insieme parzialmente ordinato $(S, \leq)$ è un "upper bound" di un sottoinsieme T di S se $\forall t \in T : u \geq t$. L'elemento u è un "least upper bound" oppure $\sup T$ se u è un "upper bound" di T e per ogni "upper bound" v di T: $u \leq v$. È

chiaro che se esiste sup T, allora è unico, grazie all'antisimmetria dell'insieme parzialmente ordinato.

In maniera simile si definiscono i "lower bounds" e i "greatest lower bounds" oppure "infs" di un insieme T. Inoltre, se esiste inf T, allora è unico.

Adesso introduciamo la seguente definizione:

Definizione G1.5 (Lattice). *Un* reticolo *o* "lattice" *è un insieme parzialmente ordinato $(L, \leq)$ nel quale due elementi a e b qualunque hanno il "least upper bound" $a \vee b$ e il "greatest lower bound" $a \wedge b$. Per specificare le due operazioni $\vee$ e $\wedge$, il reticolo L viene denotato con $(L, \vee, \wedge)$.*

- Se a, b e c sono elementi di un reticolo L, allora $(a \vee b) \vee c \geq a, b, c$ e se $v \geq a, b, c$ allora $v \geq (a \vee b)$ e c; così $v \geq (a \vee b) \vee c$. Quindi $(a \vee b) \vee c$ è il sup di a, b e c. Per induzione, si dimostra che qualunque insieme finito di elementi di un reticolo L ha sup.
- Analogamente, qualunque sottoinsieme finito di un reticolo L ha inf. Denotiamo il sup e l'inf di $s_0, s_1, \cdots, s_n$, elementi di un reticolo, L con $s_0 \vee s_1 \vee \cdots \vee s_n$ e $s_0 \wedge s_1 \wedge \cdots \wedge s_n$ rispettivamente.
- Ogni insieme totalmente ordinato è un reticolo; infatti se a e b sono due qualunque elementi di un insieme di questo tipo, abbiamo $a \leq b$ oppure $b \leq a$.

$$\text{Nel } 1^\circ \text{ caso } (a \leq b): \quad a \vee b = b \quad \text{e} \quad a \wedge b = a;$$
$$\text{Nel } 2^\circ \text{ caso } (b \leq a): \quad a \vee b = a \quad \text{e} \quad a \wedge b = b.$$

Per un reticolo $(L, \vee, \wedge)$, le operazioni $\vee$ e $\wedge$ soddisfano le seguenti proprietà:

Associativa	$\forall a, b, c \in L:$	$(a \vee b) \vee c = a \vee (b \vee c)$,
		$(a \wedge b) \wedge c = a \wedge (b \wedge c)$;
Commutativa	$\forall a, b \in L:$	$a \vee b = b \vee a$ e $a \wedge b = b \wedge a$;
Idempotente	$\forall a \in L:$	$a \vee a = a \quad \text{e} \quad a \wedge a = a.$

Inoltre, entrambe le leggi di assorbimento sono valide, ossia:

$$\forall a, b \in L: \quad a \wedge (a \vee b) = a \quad \text{e} \quad a \vee (a \wedge b) = a.$$

Definizione G1.6 (Complete lattice). *Un insieme parzialmente ordinato è detto un* reticolo completo *o* "complete lattice" *se ogni sottoinsieme T di S ha* sup *e* inf. *Denotiamo questi rispettivamente con*

$$\bigvee_{t \in T} t \qquad e \qquad \bigwedge_{t \in T} t.$$

G1.3. Esempi di "lattices". Diamo alcuni esempi importanti di reticoli:

Esempio G1.7 $(\mathbb{Z}, \leq)$. *L'insieme* $\mathbb{Z}$ *dei numeri interi relativi con la relazione d'ordine per grandezza è un reticolo. Infatti* $(\mathbb{Z}, \leq)$ *è un poset totalmente ordinato poiché*

$$\forall a, b \in \mathbb{Z} \quad con \quad a \neq b: \quad a < b \quad oppure \quad b < a.$$

Notando che l'insieme dei numeri naturali $\mathbb{N}$ *è incluso in* $\mathbb{Z}$*, si ha che* $\mathbb{N}$ *è un reticolo sotto la stessa relazione d'ordine "*$\leq$*".*

Esempio G1.8 $(\mathbb{N}, \mid)$. *L'insieme* $\mathbb{N}$ *dei numeri naturali con la relazione di divisibilità, che indicheremo con "*$\mid$*", così definita:*

$$\forall a, b \in \mathbb{N}: \quad a \mid b \iff a \text{ divide } b$$

è un reticolo dove

$$\forall a, b \in \mathbb{N}: \quad a \wedge b = \mathrm{mcd}(a, b) \quad e \quad a \vee b = \mathrm{mcm}(a, b).$$

Lo 0 *del reticolo coincide col numero naturale* 1 *e gli atomi sono i primi.*

Esempio G1.9 $(\mathcal{P}(S), \subseteq)$. *L'insieme* $\mathcal{P}(S)$ *delle parti di un insieme finito* S *con la relazione di inclusione, così definita:*

$$\forall A, B \in \mathcal{P}(S): \quad A \subseteq B \iff A \text{ è una parte di } B$$

è un reticolo dove le operazioni $\wedge$ *e* $\vee$ *corrispondono all'intersezione e all'unione rispettivamente. Gli insiemi* $\emptyset$ (vuoto) *e* S *sono rispettivamente* 0 *e* 1 *del reticolo.*

G1.4. Funzioni su insiemi parzialmente ordinati.

Definizione G1.10 (L'algebra di incidenza). *Siano* $(S, \leq)$ *un poset localmente finito e* K *un campo* (si pensi in pratica ai reali). *Indicheremo con* $\mathfrak{F}(S)$ *l'insieme delle funzioni* $f : S \times S \to K$ *tali che*

$$\forall a, b \in S: \quad a \not\leq b \implies f(a, b) = 0.$$

In $\mathfrak{F}(S)$ definiamo l'operazione di *convoluzione* "$\star$" come segue:

$$\forall f,\ g \in \mathfrak{F}(S): \quad f \star g(a, b) := \sum_{a \leq t \leq b} f(a, t)\, g(t, b).$$

L'insieme $\mathfrak{F}(S)$ con la convoluzione, con le ordinarie operazioni di addizione e moltiplicazione scalare tra funzioni, diventa una K-algebra associativa, ma in generale non commutativa. Essa è chiamata *algebra d'incidenza* di $(S, \leq)$ in K.

Si verifica facilmente che la funzione di Kronecker

$$\begin{aligned} \delta : S \times S &\longrightarrow K; \\ (a,b) &\longmapsto \delta(a,b) := \begin{cases} 1, & \text{se} \quad a = b; \\ 0, & \text{se} \quad a \neq b; \end{cases} \end{aligned}$$

è l'elemento neutro per l'operazione "$\star$", ciò significa che

$$\forall f \in \mathfrak{F}(S): \quad f \star \delta = \delta \star f = f.$$

Infatti, siano $a,\ b \in S$; allora si vede facilmente

$$\begin{aligned} f \star \delta\,(a,b) &= \sum_{a \le t \le b} f(a,t)\,\delta(t,b) \;=\; f(a,b); \\ \delta \star\ f(a,b) &= \sum_{a \le t \le b} \delta(a,t)\,f(t,b) \;=\; f(a,b). \end{aligned}$$

Per quanto riguarda l'inversa di una funzione rispetto all'operazione di convoluzione "$\star$", bisogna tener presente il seguente risultato.

Lemma G1.11. *Sia $f \in \mathfrak{F}(S)$, con $f(a,a) \neq 0$ per ogni $a \in S$. Allora esiste l'unica inversa g di f tale che $f \star g = g \star f = \delta$. Per due elementi $a,\ b \in S$ l'inversa g è definita per induzione come segue:*

$$g(a,b) = \begin{cases} \dfrac{1}{f(a,a)}, & \textit{se} \quad a = b; \\ -\displaystyle\sum_{a \le t < b} \frac{f(t,b)}{f(b,b)} g(a,t), & \textit{se} \quad a < b. \end{cases}$$

Inoltre, vale anche la seguente formula duale:

$$g(a,b) = \begin{cases} \dfrac{1}{f(a,a)}, & \textit{se} \quad a = b; \\ -\displaystyle\sum_{a < t \le b} \frac{f(a,t)}{f(a,a)} g(t,b), & \textit{se} \quad a < b. \end{cases}$$

Dimostrazione. Se $a = b$, risulta

$$g \star f(a,a) = g(a,a) f(a,a) = 1.$$

Invece se $a < b$, abbiamo che

$$\begin{aligned} 0 &= \delta(a,b) \;=\; g \star f(a,b) \\ &= g(a,b)\, f(b,b) + \sum_{a \le t < b} g(a,t) f(t,b). \end{aligned}$$

Così abbiamo dimostrato che $g \star f = \delta$, cioè che g è un'inversa sinistra di f. Analogamente, possiamo provare l'esistenza di un'inversa destra h di f.

Ora si verifica facilmente che $g = h$. Infatti, multiplicando $g \star f = \delta$ alla destra per h ed richiamando la proprietà associativa, si ha che

$$h = \delta \star h = (g \star f) \star h = g \star (f \star h) = g \star \delta = g.$$

Inoltre, supponendo che g e g' siano le inverse sinistre di f, multiplicando $\delta = g \star f = g' \star f$ alla destra per h si conferma $g = g'$ come segue:

$$\begin{aligned} h = \delta \star h &= (g \star f) \star h = g \star (f \star h) = g \star \delta = g \\ &= (g' \star f) \star h = g' \star (f \star h) = g' \star \delta = g'. \end{aligned}$$

Invece, supponendo che sia h che h' sono le inverse destre di f, si dimostra ugualmente che $g = h = h'$. Quindi f ammette l'unica inversa (sinistra e destra). Dall'equazione $f \star g = \delta$, segue la formula duale esplicitamente. □

G1.5. Funzione di Möbius μ. Prima di definire la funzione di Möbius, abbiamo bisogno di introdurre delle funzioni ξ, η, λ.

Una delle funzioni più importanti in $\mathfrak{F}(S)$ è la *zeta* (o *funzione d'incidenza*) così definita:

$$\xi : \ S \times S \longrightarrow K;$$

$$(a,b) \longmapsto \xi(a,b) := \begin{cases} 1, & \text{se} \quad a \leq b; \\ 0, & \text{altrimenti.} \end{cases}$$

Essa infatti caratterizza la relazione d'ordine parziale.

Altre due funzioni interessanti dell'algebra d'incidenza sono:

$$\eta : \ S \times S \longrightarrow K;$$

$$(a,b) \longmapsto \eta(a,b) := \begin{cases} 1, & \text{se} \quad a < b; \\ 0, & \text{altrimenti.} \end{cases}$$

ovvero $\eta = \xi - \delta$; la funzione di ricoprimento

$$\lambda : \ S \times S \longrightarrow K$$

$$(a,b) \longmapsto \lambda(a,b) := \begin{cases} 1, & \text{se } b \text{ è un "cover" di } a; \\ 0, & \text{altrimenti.} \end{cases}$$

Le funzioni ξ e λ hanno un significato combinatorio derivante dalla loro definizione. Esse danno alcune informazioni sugli intervalli dell'insieme parzialmente ordinato $(S, \leq)$ espresse nel seguente:

Teorema G1.12. *Sia $(S, \leq)$ un poset localmente finito. Per ogni intervallo chiuso $[a, b]$ con $a \leq b$, risulta:*

$$\begin{aligned}
&\text{(a)} \quad \xi^2(a,b) = \textit{cardinalità di } [a,b].\\
&\text{(b)} \quad \lambda \star \xi(a,b) = \textit{numero di atomi in } [a,b].\\
&\text{(c)} \quad \xi \star \lambda(a,b) = \textit{numero di coatomi in } [a,b].
\end{aligned}$$

DIMOSTRAZIONE. Procediamo con la dimostrazione caso per caso.

[a] La prima formula si dimostra osservando i seguenti passaggi:

$$\xi^2(a,b) = \xi \star \xi(a,b) = \sum_{a \leq t \leq b} \xi(a,t)\xi(t,b) = \sum_{t \in [a,b]} 1 = \Big|[a,b]\Big|.$$

[b] La seconda formula segue dalla definizione dell'operazione "$\star$":

$$\begin{aligned}
\lambda \star \xi(a,b) &= \sum_{a \leq t \leq b} \lambda(a,t)\xi(t,b) = \sum_{t \in (a,b]:\ |[a,t]|=2} 1\\
&= \text{numero di atomi in } [a,b].
\end{aligned}$$

[c] Per la terza ed ultima formula si ha:

$$\begin{aligned}
\xi \star \lambda(a,b) &= \sum_{a \leq t \leq b} \xi(a,t)\lambda(t,b) = \sum_{t \in [a,b):\ |[t,b]|=2} 1\\
&= \text{numero di coatomi in } [a,b].
\end{aligned}$$

Così tutte le affermazioni del teorema sono verificate. □

La *funzione di Möbius* viene definita come l'inversa della funzione zeta (o funzione d'incidenza) $\mu = \xi^{-1}$. In virtù del Lemma **G1.11** si può esprimere per ricorrenza nel modo seguente:

$$\mu : S \times S \longrightarrow K;$$

$$(a,b) \longmapsto \mu(a,b) := \xi^{-1}(a,b) = \begin{cases} 1, & \text{se} \quad a = b;\\ -\sum_{a \leq t < b} \mu(a,t), & \text{se} \quad a < b. \end{cases}$$

Evidentemente abbiamo le seguenti ortogonalità:

$$\begin{aligned}
\xi \star \mu = \delta \quad &\Longleftrightarrow \quad \sum_{a \leq t \leq b} \mu(t,b) = \delta(a,b);\\
\mu \star \xi = \delta \quad &\Longleftrightarrow \quad \sum_{a \leq t \leq b} \mu(a,t) = \delta(a,b).
\end{aligned}$$

Per poter enunciare altre proprietà delle funzioni ξ, η e λ bisogna dare ulteriori definizioni.

G1.6. Catene. Gli elementi $s_0, s_1, \cdots, s_n$ non necessariamente distinti di un insieme parzialmente ordinato $(S, \leq)$ formano una multicatena di lunghezza n quando $s_0 \leq s_1 \leq \cdots \leq s_n$ (così una multicatena è ovviamente una catena con elementi ripetuti).

Una catena (multicatena) con primo elemento a ed ultimo elemento b si denota con $\langle a, b\rangle$-catena (multicatena). Una $\langle a, b\rangle$-catena è massimale se non esiste alcun elemento dell'insieme parzialmente ordinato $(S, \leq)$ da poter aggiungere per ottenere una catena più lunga.

Ecco allora come le precedenti funzioni permettono di rispondere a problemi combinatori riguardanti catene e multicatene.

Teorema G1.13. *Sia $(S, \leq)$ un poset localmente finito. Per ogni a, $b \in S$, valgono le seguenti identità:*

(a) $\eta^k(a, b)$ – *numero di $\langle a, b\rangle$-catene lunghe k.*
(b) $\lambda^k(a, b)$ – *numero di $\langle a, b\rangle$-catene massimali lunghe k.*
(c) $\xi^k(a, b)$ – *numero di $\langle a, b\rangle$-multicatene lunghe k.*

Le formule elencate nel teorema sono ovvie. □

La prima identità ci permette di dare un significato combinatorio alla funzione di Möbius. Infatti $\xi = \delta + \eta$, mentre si può dimostrare che nell'algebra d'incidenza, vale la relazione

$$\mu = \xi^{-1} = (\delta + \eta)^{-1} = \sum_{k \geq 0} (-1)^k \eta^k.$$

Quindi, per ogni intervallo $[a, b]$ di S con $a \leq b$, risulta

$$\mu(a, b) = 1 - \eta(a, b) + \eta^2(a, b) - \cdots$$

dove $\eta^k(a, b)$ è il numero di $\langle a, b\rangle$-catene lunghe k. Questa formula, in alcuni casi, permette di ricavare $\mu(a, b)$ una volta noti i numeri $\eta^k(a, b)$.

G1.7. Formula di inversione. Presentiamo il risultato più importante della teoria delle algebre di incidenza che è fondamentale nella combinatoria enumerativa e nell'algebra quantitativa.

Teorema G1.14. *Siano $(S, \leq)$ un poset finito, f e g funzioni definite da S a valori in un campo K con $f, g : S \longrightarrow K$. Allora per ogni $a \in S$ vale*

l'equivalenza:

$$f(a) = \sum_{t \le a} g(t) \iff g(a) = \sum_{t \le a} \mu(t,a)\, f(t)$$

dove μ è la funzione di Möbius dell'algebra $\mathfrak{F}(S)$.

DIMOSTRAZIONE. Osserviamo che

$$f(a) = \sum_{t \le a} g(t) \quad \text{per tutti} \quad a \in S$$

è un sistema di equazioni; allora il teorema afferma che questo sistema è equivalente al sistema delle equazioni

$$g(a) = \sum_{t \le a} \mu(t,a)\, f(t) \quad \text{per tutti} \quad a \in S.$$

Supponiamo che il primo sistema sia valido; dimostriamo, allora, che il secondo è vero. Sostituendo $f(t)$ nel secondo sistema, possiamo riscrivere il membro destro dell'equivalenza come segue:

$$\begin{aligned} \sum_{t \le a} \mu(t,a)\, f(t) &= \sum_{t \le a} \mu(t,a) \sum_{c \le t} g(c) \\ &= \sum_{c \le a} g(c) \sum_{c \le t \le a} \mu(t,a). \end{aligned}$$

Questa somma si riduce alla funzione $g(a)$ poiché

$$\sum_{c \le t \le a} \mu(t,a) = \delta(c,a)$$

in virtù delle ortogonalità della funzione di Möbius. □

In modo analogo si può dimostrare la seguente forma duale:

Teorema G1.15. *Siano $(S, \le)$ un poset finito, f e g funzioni definite da S a valori in un campo K con $f, g : S \to K$. Allora per ogni $a \in S$ vale l'equivalenza:*

$$f(a) = \sum_{t \ge a} g(t) \iff g(a) = \sum_{t \ge a} \mu(a,t)\, f(t)$$

dove μ è la funzione di Möbius dell'algebra $\mathfrak{F}(S)$. □

L'utilità del teorema di inversione consiste principalmente nel fatto seguente: permette di determinare la funzione g quando si conoscono le sue somme parziali, cioè la funzione f e la funzione μ di Möbius corrispondente.

Introduciamo ora i concetti di prodotto diretto e di isomorfismo di insiemi parzialmente ordinati per dare vari esempi di applicazione del teorema.

G1.8. Prodotto diretto di posets. Siano $(S_1, \leq_{S_1})$ e $(S_2, \leq_{S_2})$ due insiemi parzialmente ordinati. Il loro *prodotto diretto* $(S_1 \otimes S_2, \leq)$ è costituito dalle coppie ordinate di elementi di S_1 e S_2 con la relazione d'ordine seguente:

$$(a,b) \leq (c,d) \iff \begin{cases} a \leq_{S_1} c, & \forall a, c \in S_1; \\ b \leq_{S_2} d, & \forall b, d \in S_2. \end{cases}$$

Il prodotto diretto di un numero finito, o numerabilmente infinito, di insiemi parzialmente ordinati è introdotto in maniera analoga.

Teorema G1.16. *Siano μ_{S_1} e μ_{S_2} le funzioni di Möbius per gli insiemi parzialmente ordinati $(S_1, \leq_{S_1})$ e $(S_2, \leq_{S_2})$ rispettivamente, allora per ogni $a, c \in S_1$ e $b, d \in S_2$, la funzione di Möbius $\mu_{S_1 \otimes S_2}$ del loro prodotto diretto $(S_1 \otimes S_2, \leq)$ è data dalla formula:*

$$\mu_{S_1 \otimes S_2}\big((a,b),(c,d)\big) = \mu_{S_1}(a,c)\,\mu_{S_2}(b,d).$$

DIMOSTRAZIONE. Per $f = \delta$ e $f = \xi$ vale ovviamente l'identità:

$$f_{S_1 \otimes S_2}\big((a,b),(c,d)\big) = f_{S_1}(a,c) f_{S_2}(b,d)$$

dove $a, c \in S_1$ e $b, d \in S_2$. Pertanto si può scrivere:

$$\begin{aligned} & \sum_{\substack{a \leq c \leq e \\ b \leq d \leq f}} \mu_{S_1 \otimes S_2}\big((a,b),(c,d)\big) \xi_{S_1 \otimes S_2}\big((c,d),(e,f)\big) \\ = \quad & \delta_{S_1 \otimes S_2}\big((a,b),(e,f)\big) = \delta_{S_1}(a,e)\delta_{S_2}(b,f) \\ = \quad & \sum_{a \leq c \leq e} \mu_{S_1}(a,c)\,\xi_{S_1}(c,e) \sum_{b \leq d \leq f} \mu_{S_2}(b,d)\,\xi_{S_2}(d,f) \\ = \quad & \sum_{\substack{a \leq c \leq e \\ b \leq d \leq f}} \mu_{S_1}(a,c)\,\mu_{S_2}(b,d)\,\xi_{S_1 \otimes S_2}\big((c,d)(e,f)\big) \end{aligned}$$

che conferma la tesi del teorema. □

G1.9. Isomorfismo di poset. Siano $(S_1, \leq_{S_1})$ e $(S_2, \leq_{S_2})$ due insiemi parzialmente ordinati. Essi sono *isomorfi* se esiste una biiezione $\psi : S_1 \to S_2$ tale che per ogni a e b di S_1 risulti:

$$a \leq_{S_1} b \implies \psi(a) \leq_{S_2} \psi(b).$$

Teorema G1.17. *Se ψ è un isomorfismo tra $(S_1, \leq_{S_1})$ e $(S_2, \leq_{S_2})$ allora per ogni $a, b \in S_1$, vale*

$$\mu_{S_1}(a,b) = \mu_{S_2}\big(\psi(a), \psi(b)\big).$$

DIMOSTRAZIONE. Denotiamo per ogni $a, b \in S_1$ con

$$\begin{aligned} f(a,b) &:= \mu_{S_1}(a,b), \\ g(a,b) &:= \mu_{S_2}\big(\psi(a),\psi(b)\big). \end{aligned}$$

Allora sia $f(a,b)$ che $g(a,b)$ sono funzioni su $S_1 \otimes S_1$. Entrambe le funzioni hanno i seguenti valori particolari:

$$f(a,b) = g(a,b) = 1, \quad \text{se} \quad a = b;$$
$$f(a,b) = g(a,b) = 0, \quad \text{se} \quad a \not\leq b.$$

Inoltre, le due funzioni soddisfano le stesse relazioni ricorrenti

$$\begin{aligned} \delta(a,b) &= \sum_{a\leq t\leq b} f(a,t) = \sum_{a\leq t\leq b} g(a,t), \\ \delta(a,b) &= \sum_{a\leq t\leq b} f(t,b) = \sum_{a\leq t\leq b} g(t,b); \end{aligned}$$

grazie all'isomorfismo tra $(S_1, \leq_{S_1})$ e $(S_2, \leq_{S_2})$ e alle ortogonalità della funzione di Möbius mostrate nella sezione **G1.5**. Per induzione matematica, f e g coincidono; cioè la tesi. □

Esempio G1.18 (Funzione di Möbius su $(\mathbb{N}_0, \leq)$). *Denotiamo con* $(\mathbb{N}_0, \leq)$ *l'insieme dei numeri interi relativi ordinati per grandezza. La funzione di Möbius* μ *per* $(\mathbb{N}_0, \leq)$ *dell'Esempio* **G1.7** *è chiaramente data da:*

$$\forall a, b \in \mathbb{N}_0: \quad \mu(a,b) = \begin{cases} 1, & \textit{se} \quad b = a; \\ -1, & \textit{se} \quad b = a+1; \\ 0, & \textit{altrimenti.} \end{cases}$$

Allora la inversione di Möbius si esprime come segue:

$$\left.\begin{aligned} f(n) = \sum_{k=0}^{n} g(k) &= g(0) + g(1) + \cdots + g(n) \\ g(n) = \nabla f(n) &= f(n) - f(n-1) \end{aligned}\right\} \quad (n \in \mathbb{N}_0).$$

Più in generale, l'insieme composto da una catena $(n_0 < n_1 < n_2 < \cdots)$ con $n_k \in \mathbb{N}_0$ ha la seguente funzione di Möbius:

$$\mu(n_i, n_j) = \begin{cases} +1, & j = i; \\ -1, & j = i+1; \\ 0, & \text{altrimenti.} \end{cases}$$

Ricordando il Teorema **G1.16** del prodotto diretto, possiamo ottenere la funzione di Möbius per $(\mathbb{N}_0^\ell, \leq)$ come segue:

$$\mu\big((n_1, n_2, \cdots, n_\ell), (m_1, m_2, \cdots, m_\ell)\big) = \begin{cases} (-1)^{\sum_{k=1}^{\ell}(m_k - n_k)}, & \substack{m_k - n_k = 0,\,1 \\ \text{per } 1\leq k\leq \ell;} \\ 0, & \text{altrimenti;} \end{cases}$$

dove $(m_1, m_2, \cdots, m_\ell)$ e $(n_1, n_2, \cdots, n_\ell)$ sono due ℓ-uple di $\mathbb{N}_0^\ell$.

G2. Funzioni aritmetiche ed applicazione

Per l'insieme $(\mathbb{N}, |)$ dei numeri naturali ordinati per divisibilità, la funzione di Möbius coincide con quella classica. In questa sezione, studieremo due funzioni aritmetiche, la classica funzione di Möbius e la funzione di Eulero, utili per le applicazioni relative all'inversione di Möbius nella teoria dei numeri. Esporremo la proprietà moltiplicativa di entrambe e le relazioni che le legano. Successivamente, verrà trattato il problema enumerativo delle permutazioni circolari.

G2.1. Funzione classica μ di Möbius e proprietà. La funzione di Möbius μ per $(\mathbb{N}, |)$, l'insieme dei numeri naturali ordinati per divisibilità, dell'Esempio **G1.8** è data da:

$$\forall a, b \in \mathbb{N}: \quad \mu(a,b) = \begin{cases} (-1)^k, & \text{se } b/a \text{ è un prodotto di } k \text{ primi distinti;} \\ 0, & \text{se } b/a \text{ ha un primo fattore quadrato.} \end{cases}$$

OSSERVAZIONE: Per $n \in \mathbb{N}$ il teorema fondamentale dell'aritmetica afferma che n ammette una scomposizione unica in fattori primi. Quindi $n = p_1^{n_1} p_2^{n_2} \cdots p_\ell^{n_\ell}$ dove i p_i sono primi distinti e gli $n_i > 0$ con $i = 1, 2, \cdots, \ell$.

Indichiamo con D_n il reticolo dei divisori interi positivi di n ordinato con la relazione di divisibilità. D_n è isomorfo al prodotto diretto $C_1 \otimes C_2 \otimes \cdots \otimes C_\ell$ dove C_i è la catena $C_i = \{0, 1, \cdots, n_i\}$ con $i = 1, 2, \cdots, \ell$. Per $b|n$ con $b = p_1^{b_1} p_2^{b_2} \cdots p_\ell^{b_\ell}$, quest'isomorfismo è così definito:

$$\begin{array}{ccc} D_n & \longrightarrow & C_1 \otimes C_2 \otimes \cdots \otimes C_\ell; \\ b & \longmapsto & (b_1, b_2, \cdots, b_\ell). \end{array}$$

Se $a|b$ con $a = p_1^{a_1} p_2^{a_2} \cdots p_\ell^{a_\ell}$, si ha che $a_i \leq b_i$ per ogni $i = 1, 2, \cdots, \ell$; allora combinando i Teoremi **G1.16** e **G1.17**, ne segue che:

$$\begin{aligned} \mu(a,b) &= \prod_{i=1}^{\ell} \mu(a_i, b_i) \\ &= \begin{cases} (-1)^k, & \text{se } b/a \text{ è un prodotto dei } k \text{ primi distinti;} \\ 0, & \text{se } b/a \text{ ha un primo fattore quadrato.} \end{cases} \end{aligned}$$

Ponendo $\mu(b/a) = \mu(a,b)$, otteniamo la classica funzione di Möbius, che è una delle notevoli funzioni dell'aritmetica introdotta da Möbius nel 1832 per lo studio della distribuzione dei numeri primi.

Infatti, per ogni $n \in \mathbb{N}$ avente la scomposizione in primi $n = p_1^{e_1} p_2^{e_2} \cdots p_\ell^{e_\ell}$, possiamo tradurre $\mu(n) = \mu(1,n)$ come segue:

$$\mu(n) := \begin{cases} 1, & \text{se } n = 1; \\ (-1)^k, & \text{se } e_k = 1 \text{ per } 1 \le k \le \ell; \\ 0, & \text{se esiste } k \text{ con } 1 \le k \le \ell \text{ tale che } e_k > 1. \end{cases}$$

Questa espressione è esattamente quella classica per definire la funzione di Möbius μ.

Teorema G2.1. *La funzione μ è moltiplicativa; cioè per ogni n ed $m \in \mathbb{N}$ tale che* $\mathrm{mcd}(n,m) = 1$ *si ha*

$$\mu(n \cdot m) \quad = \quad \mu(n) \cdot \mu(m).$$

Dimostrazione. Siano $n, m \in \mathbb{N}$ tali che $\mathrm{mcd}(n,m) = 1$. L'equazione viene dimostrata distinguendo i tre casi:

- $n = 1$ o $m = 1$: Senza perdere di generalità supponiamo che $n = 1$. In tale caso risulta
$$\mu(n \cdot m) = \mu(1 \cdot m) = \mu(m) = \mu(n) \cdot \mu(m).$$
- n o m contiene un fattore primo quadrato: Assumendo che la decomposizione di n ha un fattore primo quadrato, allora anche $n \cdot m$ contiene tale fattore primo quadrato; perciò
$$\mu(n \cdot m) = 0 = \mu(n) = \mu(n) \cdot \mu(m).$$
- $n > 1$, $m > 1$ e nessuno dei due contiene un fattore primo quadrato: Sia n che m sono prodotti di primi distinti
$$n = p_1 p_2 \cdots p_k \quad \text{e} \quad m = q_1 q_2 \cdots q_\ell.$$
Allora
$$n \cdot m = p_1 p_2 \cdots p_k \, q_1 q_2 \cdots q_\ell$$
costituisce la decomposizione di $n \cdot m$ in primi distinti poiché per ipotesi $\mathrm{mcd}(n,m) = 1$. Secondo la definizione di funzione di Möbius, si conclude che
$$\mu(n \cdot m) = (-1)^{k+\ell} = (-1)^k \cdot (-1)^\ell = \mu(n) \cdot \mu(m).$$

Così abbiamo provato il teorema. □

Teorema G2.2. *Sia $n \in \mathbb{N}$. Abbiamo:*

$$\sum_{d|n} \mu(d) = \begin{cases} 1, & se \quad n = 1; \\ 0, & se \quad n > 1. \end{cases}$$

DIMOSTRAZIONE. La formula è chiaramente vera se $n = 1$.

Assumiamo, allora, che $n > 1$ e scriviamo $n = p_1^{n_1} p_2^{n_2} \cdots p_\ell^{n_\ell}$ con p_i primi distinti e $n_i > 0$ per ogni $i = 1, 2, \cdots, \ell$. Tutti i divisori d di n saranno del tipo: $p_1^{e_1} p_2^{e_2} \cdots p_\ell^{e_\ell}$ dove $0 \leq e_i \leq n_i$ con $i = 1, 2, \cdots, \ell$. Quando d contiene una potenza di un primo, si ha $\mu(d) = 0$. Si noti che in questo caso tali divisori danno contributo nullo alla somma del problema.

Invece se d è un prodotto di primi distinti, $d = p_1^{\epsilon_1} p_2^{\epsilon_2} \cdots p_\ell^{\epsilon_\ell}$, dove $\epsilon_i \in \{0, 1\}$ con $i = 1, 2, \cdots, \ell$, indichiamo con r il numero di potenze diverse da zero, cioè

$$k := \Big|\{i : \ 1 \leq i \leq \ell \mid \epsilon_i = 1\}\Big|.$$

Allora, ricordando che il numero di modi per scegliere gli k fattori tra ℓ fattori è $\binom{\ell}{k}$, si ha:

$$\sum_{d|n} \mu(d) = \sum_{k=0}^{\ell} (-1)^k \binom{\ell}{k} = (1-1)^\ell = 0$$

dove l'ultimo passaggio è stato giustificato dal teorema binomiale. □

Sulla base del teorema appena dimostrato, possiamo stabilire facilmente la versione classica del Teorema **G1.14**.

Teorema G2.3. *Siano f e g due sequenze complesse. Allora il sistema delle equazioni*

$$f(n) = \sum_{d|n} g(d) \quad per \quad n = 1, 2, \cdots$$

è equivalente al sistema

$$g(n) = \sum_{d|n} \mu(\frac{n}{d})\, f(d) \quad per \quad n = 1, 2, \cdots$$

dove μ è la funzione classica di Möbius. □

G2.2. Funzione φ di Eulero e proprietà. Ricordiamo che la funzione di Eulero $\varphi(n)$ indica il numero di interi positivi $k \leq n$ coprimi con n:

$$\varphi(n) = \sum_{\substack{1 \leq k \leq n \\ \text{mcd}(k,n)=1}} 1.$$

Come nel caso della funzione di Möbius $\mu(n)$, esiste una formula semplice per la somma di questa funzione su tutti i divisori d di n.

Teorema G2.4. *Per $n \in \mathbb{N}$, vale la seguente identità:*

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

Questa formula è stata già provata nella sezione A5.3 e verrà ridimostrata tramite partizione insiemistica. .

DIMOSTRAZIONE. Ripartiamo l'insieme $\{1, 2, \cdots, n\}$ considerando per ogni divisore d di n il seguente insieme:

$$\Phi_d := \Big\{ k \in \{1, 2, \cdots, n\} \mid \text{mcd}(k, n) = d \Big\}.$$

Allora

$$\{1, 2, \cdots, n\} = \biguplus_{d|n} \Phi_d$$

è un'unione disgiunta. D'altra parte

$$\Phi_d = \Big\{ kd : \ k \in \big\{1, 2, \cdots, n/d\big\} \mid \text{mcd}(k, n/d) = 1 \Big\}.$$

Quindi, per la definizione di funzione di Eulero, risulta $|\Phi_d| = \varphi(\frac{n}{d})$, pertanto

$$\begin{aligned} n &= \Big|\{1, 2, \cdots, n\}\Big| = \Big|\biguplus_{d|n} \Phi_d\Big| \\ &= \sum_{d|n} |\ \Phi_d\ | = \sum_{d|n} \varphi(n/d) \end{aligned}$$

la quale è equivalente all'espressione $\sum_{d|n} \varphi(d) = n$, invertendo l'ordine della somma. □

La funzione di Eulero è in relazione con la funzione di Möbius attraverso la seguente formula:

Teorema G2.5. *Per $n \in \mathbb{N}$, vale la seguente formula:*

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

DIMOSTRAZIONE. La definizione di $\varphi(n)$ può essere riformulata nella forma

$$\varphi(n) = \sum_{k=1}^{n} \Big\lfloor \frac{1}{\text{mcd}(k, n)} \Big\rfloor$$

dove con $\lfloor x \rfloor$ denotiamo la parte intera di un numero reale x. Adesso usiamo il Teorema **G2.2** sostituendo ad n il $\mathrm{mcd}(k,n)$ per ottenere

$$\varphi(n) = \sum_{k=1}^{n} \sum_{d \mid \mathrm{mcd}(k,n)} \mu(d) = \sum_{d \mid n} \mu(d) \sum_{\substack{k=1 \\ d \mid k}}^{n} 1 = \sum_{d \mid n} \mu(d) \frac{n}{d}.$$

Ciò completa la dimostrazione. □

La somma per $\varphi(n)$ nel teorema precedente può essere espressa anche come un prodotto esteso ai divisori primi distinti di n.

Teorema G2.6. *Per $n \in \mathbb{N}$, vale la seguente formula:*

$$\varphi(n) = n \prod_{p \mid n} \left(1 - \frac{1}{p}\right)$$

dove p corre su tutti i fattori primi di n.

DIMOSTRAZIONE. Per $n = 1$ il prodotto è nullo poiché non esistono primi che dividano 1. In questo caso assegnamo per convenzione $\varphi(n) = 1$.

Supponiamo allora che $n > 1$ e siano $p_1, p_2, \cdots, p_k$ divisori primi distinti di n. Il prodotto può essere riscritto come:

$$\begin{aligned} n \prod_{p \mid n} \left(1 - \frac{1}{p}\right) &= n \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right) \\ &= n + \sum_{\iota=1}^{k} \sum_{1 \le i_1 < i_2 < \cdots < i_\iota \le k} \frac{(-1)^\iota \, n}{p_{i_1} p_{i_2} \cdots p_{i_\iota}}. \end{aligned}$$

Notiamo che ogni termine dell'ultima espressione è della forma $\pm n/d$ dove d è un divisore di n che è un primo o un prodotto di primi distinti. Il numeratore $\pm n$ è esattamente $n \cdot \mu(d)$. Poiché $\mu(d) = 0$ se d è divisibile dal quadrato di qualche p_i, si conclude che l'ultima espressione è esattamente $\sum_{d \mid n} \mu(d) \, \frac{n}{d}$; cioè la tesi grazie al teorema precedente. □

Molte proprietà di $\varphi(n)$ possono essere dedotte da quest'ultimo teorema.

Teorema G2.7. *La funzione di Eulero soddisfa le seguenti proprietà:*

(a) *Per ogni primo p e per ogni $e \ge 1$, risulta*

$$\varphi(p^e) = p^e - p^{e-1}.$$

(b) *Per ogni $m, n \in \mathbb{N}$ tale che $d = \mathrm{mcd}(m,n)$ si ha*

$$\varphi(m \cdot n) = \varphi(m) \cdot \varphi(n) \frac{d}{\varphi(d)}.$$

(c) *La funzione φ è moltiplicativa; cioè per ogni $m, n \in \mathbb{N}$ si ha*

$$\mathrm{mcd}(m,n) = 1 \quad \Longleftrightarrow \quad \varphi(m \cdot n) \;=\; \varphi(m) \cdot \varphi(n).$$

(d) *Per ogni $m, n \in \mathbb{N}$ si ha*

$$m \mid n \qquad \Longleftrightarrow \qquad \varphi(m) \mid \varphi(n).$$

DIMOSTRAZIONE. La prima parte segue ponendo $n = p^e$ nel teorema precedente. Per provare la seconda parte scriviamo:

$$\frac{\varphi(n)}{n} = \prod_{p|n} \Big(1 - \frac{1}{p}\Big).$$

Notiamo che ogni divisore primo di $m \cdot n$ è un divisore primo di m o di n o di ambedue, e questi primi che dividono sia m che n dividono anche il $\mathrm{mcd}(m,n)$. Donde:

$$\frac{\varphi(m \cdot n)}{m \cdot n} = \prod_{p|m\cdot n} \Big(1 - \frac{1}{p}\Big)$$

$$= \quad \frac{\prod_{p|m} \left(1 - \frac{1}{p}\right) \prod_{p|n} \left(1 - \frac{1}{p}\right)}{\prod_{p|\mathrm{mcd}(m,n)} \left(1 - \frac{1}{p}\right)}$$

$$= \quad \left\{ \frac{\varphi(m)}{m} \times \frac{\varphi(n)}{n} \right\} \Big/ \left\{ \frac{\varphi(d)}{d} \right\}$$

e quindi vale la seconda formula.

La terza parte è un caso speciale della seconda. L'ultima parte segue dalla formula esplicita nel Teorema **G2.6**. □

G2.3. Permutazioni circolari e conteggio. Adesso rivediamo il problema di collane trattato nell'Esempio **F1.7**. Denotiamo la permutazione ciclica con $\pi := (1\,2 \cdots m)$. Nell'insieme delle parole lineari di lunghezza m su un alfabeto di n lettere, definiamo la seguente relazione di equivalenza: le parole $a_1\,a_2 \cdots a_m$ e $b_1\,b_2 \cdots b_m$ sono equivalenti se esiste una permutazione σ, potenza del ciclo π tale che

$$a_{\sigma(1)}\; a_{\sigma(2)} \cdots a_{\sigma(m)} = b_1\; b_2 \cdots b_m.$$

Una classe di equivalenza sarà chiamata *parola circolare.*

Sia d un divisore di m; se la parola circolare $\mathcal{C}$ è una sequenza di altre m/d parole circolari di lunghezza d, uguali tra loro, $\mathcal{C}$ è detta *periodica.* Si definisce *periodo* di $\mathcal{C}$ la più piccola delle lunghezze delle parole circolari che rendono $\mathcal{C}$ periodica.

Teorema G2.8. *Il numero di parole circolari aperiodiche di lunghezza m su n lettere è:*

$$w(m,n) \;=\; \frac{1}{m}\sum_{d|m}\mu(d)\,n^{m/d}$$

mentre il numero di parole circolari di lunghezza m su n lettere è:

$$W(m,n) \;=\; \frac{1}{m}\sum\varphi(d)\,n^{m/d}.$$
$$d|m$$

DIMOSTRAZIONE. Sia $w(d,n)$ il numero di parole circolari aperiodiche di lunghezza d, a ciascuna delle quali corrispondono d parole lineari distinte. Tutte le parole lineari di lunghezza m che si possono formare con n lettere sono in numero n^m, come sappiamo. Classificandole secondo il periodo "d", otteniamo l'identità seguente:

$$n^m \;=\; \sum_{d|m} d\,w(d,n) \qquad (\#)$$

dove d varia su tutti i divisori di m.

Da questa relazione si può ricavare $w(d,n)$ utilizzando la formula dell'inversione di Möbius nel Teorema **G2.3**. Infatti, se poniamo

$$\begin{aligned} f(m) &:= n^m, \\ g(d) &:= d\,w(d,n); \end{aligned}$$

si ottiene, invertendo la relazione (#):

$$w(m,n) \;=\; \frac{1}{m}\sum_{d|m}\mu(d)\,n^{m/d}$$

cioè la prima formula desiderata del teorema.

Classificando tutte le parole circolari di lunghezza n secondo il periodo, otteniamo immediatamente

$$W(m,n) = \sum_{d|m} w(d,n)$$

da cui procediamo come segue:

$$W(m,n) = \sum_{d|m}\frac{1}{d}\sum_{c|d} n^c\,\mu(d/c) = \sum_{c|m}\frac{n^c}{m}\sum_{\frac{d}{c}\,|\,\frac{m}{c}}\frac{m/c}{d/c}\,\mu(d/c).$$

L'ultima somma interna si riduce alla funzione di Eulero $\varphi(m/c)$ in virtù del Teorema **G2.5**. Così abbiamo completato la dimostrazione del teorema. □

G3. Principio di inclusione ed esclusione

In questa sezione vogliamo studiare una generalizzazione della formula sulla cardinalità degli insiemi finiti A e B:

$$\mid A \cup B \mid = \mid A \mid + \mid B \mid - \mid A \cap B \mid .$$

Per il reticolo $(P(S), \subseteq)$ dell'insieme delle parti (ordinate per inclusione) di un insieme finito S, vengono determinata la funzione di Möbius e presentata il principio di inclusione-esclusione (come conseguenza del teorema di inversione). Come applicazioni, vengono affrontati tre problemi:

- le permutazioni senza punti fissi, viene cioè determinato il numero delle permutazioni di un insieme finito senza punti fissi.
- il problema dei Ménages (Lucas, 1891) nel quale si chiede il numero di modi di assegnare i posti intorno ad un tavolo rotondo a n coppie in modo alternato, tale che nessun marito abbia al proprio fianco la moglie.
- la presentazione di una soluzione al problema 10770 di *American Mathematical Monthly*, riguardante la divisibilità di una somma binomiale.

G3.1. Funzione di Möbius su $(\mathcal{P}(S), \subseteq)$. Per S insieme finito, la funzione di Möbius μ per $(\mathcal{P}(S), \subseteq)$, l'insieme delle parti di S dell'Esempio **G1.9**, è data da:

$$\forall A, B \in \mathcal{P}(S): \quad \mu(A,B) = \begin{cases} (-1)^{|B\backslash A|}, & \text{se } A \subseteq B; \\ 0, & \text{se } A \not\subseteq B. \end{cases}$$

dove $B\backslash A = B \cap A^c$ (A^c è il complementare di A in S). Osserviamo che: $\mathcal{P}(S)$ è isomorfo, quando $|S| = n$, al prodotto diretto $C^n := C \otimes C \otimes \cdots \otimes C$ per n volte, dove C è la catena $C = \{0, 1\}$. Infatti, se A è un sottoinsieme di S, allora associamo ad A la sua funzione caratteristica χ_A così definita:

$$\chi_A(s) = \begin{cases} 1, & \text{se } s \in A; \\ 0, & \text{se } s \notin A. \end{cases}$$

La funzione

$$\begin{array}{rcl} \mathcal{P}(S) & \longrightarrow & C^n = C \otimes C \otimes \cdots \otimes C, \\ A & \longmapsto & \big(\chi_A(s_1), \chi_A(s_2), \cdots, \chi_A(s_n)\big); \end{array}$$

è un isomorfismo. Allora, combinando i Teoremi **G1.16** e **G1.17**, ne segue che:

$$\mu(A,B) = \prod_{i=1}^{n} \mu(\chi_A(s_i), \chi_B(s_i)) = \begin{cases} (-1)^{|B\backslash A|}, & \text{se } A \subseteq B; \\ 0, & \text{se } A \not\subseteq B. \end{cases}$$

G3.2. Principio di inclusione-esclusione. Per un numero naturale n ed $[n] = \{1, 2, \cdots, n\}$, siano Ω un insieme finito e $\Phi = \{A_1, A_2, \cdots, A_n\}$ una classe di sottoinsiemi di Ω. Definiamo i coefficienti

$$S_m := \sum_{\substack{\sigma \subset [n] \\ |\sigma| = m}} \Big| \bigcap_{j \in \sigma} A_j \Big| \quad \text{con} \quad S_0 := |\ \Omega\ | \ .$$

Inoltre, denotiamo con ω_m la cardinalità del sottoinsieme di Ω definito da

$$\Omega_m = \big\{ x \in \Omega \big| x \text{ appartiene esattamente a } m \text{ sottoinsiemi di } \{A_k\}_{k=1}^n \big\}.$$

Allora abbiamo le seguenti formule:

Teorema G3.1 (Principio di inclusione-esclusione).

(a) **Formula di Sylvester**:

$$\Big| \bigcap_{i=1}^n A_i^c \Big| \ = \ \omega_0 \ = \ \sum_{k=0}^n (-1)^k S_k.$$

(b) **Formula di Da Silva**:

$$\Big| \bigcup_{i=1}^n A_i \Big| \ = \ |\Omega| - \omega_0 \ = \ \sum_{k=1}^n (-1)^{k-1} S_k.$$

(c) **Formula di Jordan**:

$$\omega_m \ = \ \sum_{k=m}^n (-1)^{k+m} \binom{k}{m} S_k.$$

DIMOSTRAZIONE Il principio di inclusione ed esclusione si può ricavare come caso particolare del teorema dell'inversione **G1.15**. Sull'insieme delle parti di $[n]$ definiamo la funzione g nel modo seguente:

$$\forall \sigma \subseteq [n]: \quad g(\sigma) := \Big| \Big(\bigcap_{i \in \sigma} A_i \Big) \bigcap \Big(\bigcap_{i \notin \sigma} A_i^c \Big) \Big|.$$

Per come è definita, $g(\sigma)$ è il numero di elementi di Ω che appartengono agli insiemi A_i aventi indice in σ e in nessun altro. Per ogni $\sigma \subseteq [n]$, determiniamo

$$f(\sigma) \ = \ \Big| \bigcap_{i \in \sigma} A_i \Big| \ = \ \sum_{\sigma \subseteq \tau \subseteq [n]} g(\tau).$$

Il teorema d'inversione **G1.15**, costruito sull'insieme parzialmente ordinato $(\mathcal{P}([n]), \subseteq)$ ci dà:

$$\begin{aligned} g(\sigma) &= \sum_{\sigma \subseteq \tau \subseteq [n]} \mu(\sigma, \tau) f(\tau) \\ &= \sum_{\sigma \subseteq \tau \subseteq [n]} (-1)^{|\tau \setminus \sigma|} \ | \ (\cap_{i \in \tau} A_i) \ | \ . \end{aligned}$$

Per $\sigma = \emptyset$ si ha:

$$\begin{aligned} g(\emptyset) &= \left|A_1^c \cap A_2^c \cap \cdots \cap A_n^c\right| = \sum_{\emptyset\subseteq\tau\subseteq[n]} (-1)^{|\tau|}\left|(\cap_{i\in\tau} A_i)\right| \\ &= \sum_{k=0}^{n}(-1)^k \sum_{|\tau|=k} |\ (\cap_{i\in\tau} A_i)\ | = \sum_{k=0}^{n}(-1)^k S_k \end{aligned}$$

che è la formula di Sylvester.

Notando la relazione insiemistica

$$\left|A_1 \cup A_2 \cup \cdots \cup A_n\right| = |\Omega| - \left|A_1^c \cap A_2^c \cap \cdots \cap A_n^c\right|$$

otteniamo subito, dalla formula di Sylvester, la formula di Da Silva.

Se vogliamo, invece, il numero ω_m di elementi di Ω che si trovano in esattamente m degli insiemi A_k con $1 \le k \le n$, bisogna manipolare la seguente somma:

$$\begin{aligned} \omega_m &= \sum_{|\sigma|=m} g(\sigma) = \sum_{|\sigma|=m}\sum_{\sigma\subseteq\tau\subseteq[n]} (-1)^{|\tau\setminus\sigma|}\left|(\cap_{i\in\tau} A_i)\right| \\ &= \sum_{\substack{\tau\subseteq[n]\\|\tau|\ge m}}\sum_{\substack{\sigma\subseteq\tau\\|\sigma|=m}} (-1)^{|\tau|-m}\left|(\cap_{i\in\tau} A_i)\right| \\ &= \sum_{k=m}^{n}(-1)^{k-m}\binom{k}{m}\sum_{|\tau|=k}\left|(\cap_{i\in\tau} A_i)\right| \\ &= \sum_{k=m}^{n}(-1)^{k-m}\binom{k}{m} S_k \end{aligned}$$

che è la formula di Charles Jordan.

OSSERVAZIONE: La formula di Jordan è in effetti duale alla seguente:

$$S_m = \sum_{\sigma\subset[n]} \left|\bigcap_{j\in\sigma} A_j\right| = \sum_{k=m}^{n}\binom{k}{m}\omega_k$$
$$|\sigma|=m$$

che può essere stabilita tramite ragionamento combinatorio. □

Teorema G3.2 (Principio di inclusione-esclusione con funzione peso). *Introducendo una* funzione peso *su* Ω*:*

$$\begin{aligned} w : \Omega &\longrightarrow A, \\ x &\longmapsto w(c); \end{aligned}$$

dove A un anello commutativo. Per ogni sottoinsieme $X \subseteq \Omega$, definiamo il suo enumeratore *come segue:*

$$\mathcal{W}(X) = \sum_{x \in X} w(x).$$

Allora il principio d'inclusione ed esclusione ha una forma pesata con la funzione w:

$$\mathcal{W}(\omega_m) \;=\; \sum_{k=m}^{n} (-1)^{k+m} \binom{k}{m} \sum_{\substack{\sigma \subset [n] \\ |\sigma|=k}} \mathcal{W}\Big(\bigcap_{j \in \sigma} A_j \Big). \qquad \square$$

G3.3. Permutazioni senza punti fissi. Un'applicazione della formula di Sylvester si ha risolvendo il seguente problema: vogliamo calcolare il numero di permutazioni su n elementi, tra le $n!$ possibili, che non abbiano punti fissi.

Sia Ω l'insieme di tutte le permutazioni di $[n]$ con $|\Omega| = n!$. Una permutazione π di Ω ha un punto fisso $j \in [n]$ se $\pi(j) = j$. Indichiamo con $\mathcal{D}_n$ il numero delle permutazioni di $[n]$ senza elementi fissi, cioè

$$\mathcal{D}_n := \Big| \{ \pi \in \Omega \mid \forall j \in [n] : \pi(j) \neq j \} \Big|.$$

Per calcolare $\mathcal{D}_n$, consideriamo gli insiemi A_i definiti per ogni $i \in [n]$ come segue:

$$A_i := \{ \pi \in \Omega \mid \pi(i) = i \}.$$

Così dovrà essere

$$\mathcal{D}_n = \Big| A_1^c \cap A_2^c \cap \cdots \cap A_n^c \Big|$$

e possiamo applicare la formula di Sylvester.

Per definizione $S_0 := |\Omega| = n!$, mentre per ogni parte σ di $[n]$ si ha

$$\Big| \bigcap_{i \in \sigma} A_i \Big| = (n - |\sigma|)!$$

che conta il numero delle permutazioni di Ω che lasciano fissi gli elementi appartenenti a σ. Allora

$$S_k := \sum_{\substack{\sigma \subseteq [n] \\ |\sigma|=k}} \Big| \bigcap_{i \in \sigma} A_i \Big| \;=\; \binom{n}{k} (n-k)! \;=\; \frac{n!}{k!}$$

e quindi per la formula di Sylvester, si ha:

$$\mathcal{D}_n \;=\; \sum_{k=0}^{n} (-1)^k \, \frac{n!}{k!} \;=\; n! \sum_{k=0}^{n} (-1)^k \, \frac{1}{k!} \;\approx\; \frac{n!}{e}.$$

I numeri $\mathcal{D}_n$ al variare di $n \in \mathbb{N}$ sono detti *subfattoriali* in quanto verificano la ricorrenza

$$\mathcal{D}_n \;=\; n\,\mathcal{D}_{n-1} + (-1)^n$$

che è una conseguenza immediata della formula esplicita appena stabilita.

G3.4. Problema dei Ménages. Questo problema ci chiede il numero di modi di assegnare i posti intorno ad un tavolo rotondo di n signori numerati da 1 a n e delle loro rispettive consorti numerate da $1'$ a n' in modo alternato tale che nessun marito abbia al proprio fianco la moglie.

Supponiamo gli uomini già seduti; con questa ipotesi stiamo considerando il *problema dei Ménages ridotto*. Una disposizione al tavolo si può descrivere con una biiezione

$$f : [n] \;\longrightarrow\; \{1', 2', \cdots, n'\}.$$

L'uomo numero 1 si siede e alla sua destra sta la donna $f(1)$; a destra della donna $f(1)$ si siede l'uomo numero 2 alla cui destra si siede la donna $f(2)$ e così via. Poniamo

$$\begin{aligned}
1 \le i \le n: \quad & A_{2i-1} := \big\{f : [n] \to \{1', 2', \cdots, n'\} \mid f(i) = i'\big\};\\
1 \le i < n: \quad & A_{2i} \quad := \big\{f : [n] \to \{1', 2', \cdots, n'\} \mid f(i) = (i+1)'\big\};\\
i = n: \quad & A_{2n} \quad := \big\{f : [n] \to \{1', 2', \cdots, n'\} \mid f(n) = 1'\big\}.
\end{aligned}$$

Per un fissato $\sigma \subseteq [n]$, denotiamo

$$\theta(\sigma) \quad = \quad \Big|\bigcap_{i\in\sigma} A_i\Big|.$$

Se $\sigma \subseteq [n]$ non contiene due interi consecutivi della successione circolare $(1, 2, \cdots, 2n)$, abbiamo che

$$\theta(\sigma) \quad = \quad (n - |\sigma|)!$$

altrimenti $\theta(\sigma)$ si riduce allo zero.

È noto che le k-parti non contenenti due interi consecutivi della successione circolare $(1, 2, \cdots, m)$ sono in numero $\frac{m}{m-k}\binom{m-k}{k}$. Applicando la formula di Sylvester, si ha la formula di *Touchard* come segue:

$$\begin{aligned}
\big|A_1^c \cap A_2^c \cap \cdots \cap A_{2n}^c\big| &= \sum_{\sigma\subseteq[2n]}^{*} (-1)^{|\sigma|}\,\theta(\sigma) \;=\; \sum_{\sigma\subseteq[2n]}^{*} (-1)^{|\sigma|}\,\Big|\big(\bigcap_{i\in\sigma} A_i\big)\Big|\\
&= \sum_{k=0}^{n} (-1)^k \frac{2n}{2n-k}\binom{2n-k}{k}(n-k)!
\end{aligned}$$

dove l'indice della somma condizionata da "$*$" varia nei sottoinsiemi compatibili in $[n]$; cioè, σ non contiene due interi consecutivi della successione circolare $(1, 2, \cdots, 2n)$.

Quindi la soluzione finale dei Ménages risulta

$$2 \cdot n! \sum_{k=0}^{n} (-1)^k \frac{2n}{2n-k} \binom{2n-k}{k} (n-k)!.$$

G3.5. Problema 10770 in American Mathematical Monthly. Siano $m, n \in \mathbb{N}$ tali che $1 < m < n + \varphi(m)$, dove φ è la funzione di Eulero. Il problema 10770 di *American Mathematical Monthly* chiede di dimostrare che m divide la somma binomiale definita da

$$\sum_{k=1}^{n} (-1)^k \binom{n}{k} k^m.$$

Forniamo una soluzione utilizzando il principio di inclusione-esclusione.

La somma binomiale nel problema può essere trasformata nella somma multipla definita da

$$T(m,n) := \sum_{\substack{i_1+i_2+\cdots+i_n=m \\ i_k>0:\, k=1,2,\cdots,n}} \binom{m}{i_1, i_2, \cdots, i_n}$$

dove l'argomento della sommatoria è l'usuale coefficiente multinomiale.

Per l'equazione $x_1 + x_2 + \cdots + x_n = m$ con due numeri naturali fissati m ed n, sia Ω l'insieme delle sue soluzioni intere non negative, cioè:

$$\Omega = \Big\{ (i_1, i_2, \cdots, i_n) \in \mathbb{N}_0^n \mid i_1 + i_2 + \cdots + i_n = m \Big\}.$$

Definiamo la funzione peso $\mathcal{W}$ su Ω attraverso

$$\mathcal{W}\left[(i_1, i_2, \cdots, i_n)\right] := \binom{m}{i_1, i_2, \cdots, i_n} \quad \text{per} \quad (i_1, i_2, \cdots, i_n) \in \Omega.$$

Per $k \in [n]$, sia S_k il sottoinsieme delle n-uple di Ω nelle quali la k-esima coordinata è uguale a zero. Per il Teorema **G3.2** del principio di inclusione-esclusione con funzione di peso, abbiamo

$$T(m,n) = \mathcal{W}\Big[\bigcap_{k=1}^{n} S_k^c\Big] = \sum_{\sigma \subset [n]} (-1)^{|\sigma|}\, \mathcal{W}\Big[\bigcap_{i \in \sigma} S_i\Big]$$

dove con $|\sigma|$ denotiamo la cardinalità di $\sigma \subset [n]$.

Con $|\sigma| = n - k$ specificato da $[n] \backslash \sigma = \{\nu_1, \nu_2, \cdots, \nu_k\}$, possiamo valutare $\mathcal{W}\left[\cap_{i \in \sigma} S_i\right]$ grazie al teorema multinomiale, come segue:

$$\mathcal{W}\left[\cap_{i \in \sigma} S_i\right] = \sum_{\substack{j_1+j_2+\cdots+j_k=m \\ j_i>0:\, i=1,2,\cdots,k}} \binom{m}{j_1, j_2, \cdots, j_k} = k^m$$

che dipende solo dalla cardinalità di σ.

Classificando la multisomma rispettando la cardinalità di $\sigma \subset [n]$, abbiamo:

$$T(m,n) \;=\; \mathcal{W}\Big[\bigcap_{k=1}^{n} S_k^c\Big] \;=\; (-1)^n \sum_{k=1}^{n} (-1)^k \binom{n}{k} k^m$$

che è esattamente la trasformazione anticipata all'inizio.

In base al teorema fondamentale dell'aritmetica, possiamo scrivere:

$$m = p_1^{m_1}\, p_2^{m_2} \cdots p_\ell^{m_\ell}$$

dove i p_i sono primi distinti e gli m_i interi positivi per $i = 1, 2, \cdots, \ell$. Se possiamo mostrare che per ogni p_k della decomposizione di m, l'argomento della sommatoria (coefficiente multinomiale) di $T(m,n)$ è un multiplo di $p_k^{m_k}$, allora esso è anche un multiplo di m. Segue immediatamente che $T(m,n)$ è divisibile per m come desiderato.

Per m fissato come prima, richiamando la funzione di Eulero

$$\varphi(m) \;=\; m \prod_{k=1}^{\ell} \Big(1 - \frac{1}{p_k}\Big) \le m\, \frac{p_k - 1}{p_k} \quad \text{per} \quad k = 1, 2, \cdots, \ell$$

possiamo riformulare $m < n + \varphi(m)$ come

$$m < n\, p_k \quad \text{per} \quad k = 1, 2, \cdots, \ell$$

Allora per ogni p_k assegnato, osserviamo che esiste un indice di $\{i_1, i_2, \cdots, i_n\}$ nell'argomento della sommatoria di $T(m,n)$ il quale è più piccolo di p_k. Altrimenti, dovremmo avere

$$m = i_1 + i_2 + \cdots + i_n \ge n p_k$$

che ci induce in contraddizione con $m < n\, p_k$ già stabilito. Senza perdere di generalità, supponiamo che l'indice specificato sia i_1 con $0 < i_1 < p_k$. È facile vedere che $\binom{m}{i_1}$ è divisibile per $p_k^{m_k}$ per $m = \prod_{i=1}^{\ell} p_i^{m_i}$. Come conseguenza, l'argomento della sommatoria

$$\binom{m}{i_1, i_2, \cdots, i_n} = \binom{m}{i_1}\binom{m}{i_2, \cdots, i_n}$$

è un multiplo di $p_k^{m_k}$.

Questo conferma che $T(m,n)$ è divisibile per m con $m < n + \varphi(m)$, il quale fornisce una soluzione al problema.

G4. Spazi vettoriali e coefficiente Gaussiano

Lo scopo di questa sezione è lo studio del reticolo degli spazi vettoriali di dimensione finita su un campo finito. Si dimostra che il numero di sottospazi

è uguale al coefficiente Gaussiano. Si valuta la funzione di Möbius per il reticolo di tutti i sottospazi. Infine viene calcolato il numero delle trasformazioni lineari (iniettive, suriettive e biettive) tra due spazi vettoriali di dimensione finita, che risulta conseguentemente in due identità q-binomiali.

G4.1. Spazi vettoriali finiti su campi finiti. Di seguito riportiamo un teorema che ci permette di determinare il numero di tutti i sottospazi di uno spazio vettoriale finito.

Teorema G4.1. *Denotiamo con $V_n(q)$ lo spazio vettoriale di dimensione finita n sul campo finito $K(q)$ di q elementi* (dove q è una potenza di un primo). *Allora, per ogni k con $0 < k \leq n$, il numero dei sottospazi di $V_n(q)$ di dimensione k è il coefficiente Gaussiano:*

$$\begin{bmatrix} n \\ k \end{bmatrix} := \frac{(1-q^n)(1-q^{n-1})\cdots(1-q^{n-k+1})}{(1-q^k)(1-q^{k-1})(1-q)}.$$

DIMOSTRAZIONE. Dalle ipotesi segue che $V_n(q)$ ha q^n vettori distinti. Per determinare i sottospazi k-dimensionali di $V_n(q)$, prima determiniamo tutti i possibili insiemi $\{v_1, v_2, \cdots, v_k\}$ di k vettori linearmente indipendenti.

Tali insiemi possono essere scelti come segue: v_1 può essere qualunque dei $q^n - 1$ non zero elementi di $V_n(q)$; v_2 può essere qualunque dei $q^n - q$ vettori situato al di fuori del sottospazio generato da v_1; v_3 può essere qualunque dei $q^n - q^2$ vettori situato al di fuori del sottospazio generato da $\{v_1, v_2\}$; e così via. Donde il numero di modi in cui può essere scelto l'insieme $\{v_1, v_2, \cdots, v_k\}$ è $(q^n-1)(q^n-q)\cdots(q^n-q^{k-1})$. Adesso, ognuna di tali k-uple $\{v_1, v_2, \cdots, v_k\}$ genera un sottospazio k-dimensionale di $V_n(q)$; tuttavia, diverse k-uple possono generare lo stesso sottospazio. Infatti, più precisamente

$$(q^k-1)(q^k-q)\cdots(q^k-q^{k-1})$$

è il numero di modi di scegliere un sottoinsieme di k elementi linearmente indipendenti di $V_k(q)$ che genera lo stesso sottospazio k-dimensionale.

Donde il numero di sottospazi k-dimensionali di $V_n(q)$ è

$$\begin{aligned} & \frac{(q^n-1)(q^n-q)\cdots(q^n-q^{k-1})}{(q^k-1)(q^k-q)\cdots(q^k-q^{k-1})} \\ = \quad & \frac{(1-q^n)(1-q^{n-1})\cdots(1-q^{n-k+1})}{(1-q^k)(1-q^{k-1})\cdots(1-q)} \end{aligned}$$

che è quanto volevamo dimostrare. □

- Il numero $\begin{bmatrix} n \\ k \end{bmatrix}$ è detto *coefficiente Gaussiano* che si riduce, quando $q \to 1$, al coefficiente binomiale ordinario.
- Dal Teorema **G4.1** segue che il numero di tutti i sottospazi di $V_n(q)$ è

$$G_n(q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}$$

che è denominato numero di Galois.
- Indichiamo con $(L(V_n), \subseteq)$ il reticolo costituito da tutti i sottospazi dello spazio vettoriale $V_n(q)$ di dimensione finita n sul campo finito $K(q)$ con la relazione di inclusione tra sottospazi.

G4.2. Funzione di Möbius sugli spazi vettoriali. Introduciamo il lemma di Weisner (1935) per determinare la funzione di Möbius degli spazi vettoriali finiti.

Lemma G4.2 (Weisner, 1935). *Sia μ la funzione di Möbius di un "lattice" finito L con $0_L := \inf L$ e $1_L := \sup L$. Per ogni elemento $a \in L$ con $a > 0_L$ si ha*

$$\sum_{x \in L:\, x \vee a = 1_L} \mu(0_L, x) = 0.$$

DIMOSTRAZIONE. Per $a \in L$, consideriamo una doppia somma formale

$$S := \sum_{x,\, y \in L} \mu(0_L, x)\xi(x, y)\xi(a, y)\mu(y, 1_L).$$

Manipoliamo questa doppia somma in due modi diversi per arrivare al risultato desiderato.

Fissando $x \in L$ possiamo riscriverla nel modo seguente:

$$S = \sum_{x \in L} \mu(0_L, x) \sum_{a,\, x \le y \le 1_L} \mu(y, 1_L)$$

ma $y \geq a$ e $y \geq x$ se e solo se $y \geq x \vee a$ e la somma interna si riduce:

$$\sum_{a \vee x \le y \le 1_L} \mu(y, 1_L) = \begin{cases} 1, & \text{se } x \vee a = 1; \\ 0, & \text{se } x \vee a < 1. \end{cases}$$

Così S diventa la somma nell'espressione del teorema.

Invece, fissando $y \in L$ riformuliamo la doppia somma come segue:

$$S = \sum_{a \le y \le 1_L} \mu(y, 1_L) \sum_{0_L \le x \le y} \mu(0_L, x)$$

e la somma interna viene annullata poiché $y > 0_L$. Perciò $S = 0$ grazie alla condizione $0_L < a \leq y$. Confrontando le due espressioni ottenute, abbiamo il risultato di Weisner. □

Ora siamo in grado di dimostrare il seguente teorema che ci dà una formula esplicita per la funzione di Möbius sul reticolo degli spazi vettoriali finiti.

Teorema G4.3 (Funzioni di Möbius). *Sia $(L(V_n), \subseteq)$ il reticolo di tutti i sottospazi dello spazio vettoriale $V_n(q)$ di dimensione finita n sul campo finito $K(q)$ con la relazione di inclusione tra sottospazi. Per $U, W \in L(V_n)$, si ha la funzione di Möbius:*

$$\mu(U,W) = \begin{cases} (-1)^k\, q^{\binom{k}{2}}, & \text{se } U \subseteq W \text{ e } \dim(W) - \dim(U) = k; \\ 0, & \text{se } U \not\subseteq W. \end{cases}$$

DIMOSTRAZIONE. Per l'isomorfismo tra lo spazio vettoriale V_k di dimensione k e lo spazio quoziente W/U, sarà sufficiente mostrare che, per uno spazio vettoriale V di dimensione n, vale

$$\mu(0,V) = (-1)^n\, q^{\binom{n}{2}}.$$

Procediamo per induzione sulla dimensione n dello spazio vettoriale V.

Sia P un sottospazio unidimensionale di V. Per il lemma di Weisner risulta

$$\mu(0,V) = -\sum_{U \subset V:\, U \vee P = V} \mu(0,U).$$

Dall'ipotesi induttiva, il termine generale nella somma uguaglia

$$\mu(0,U) = (-1)^{\dim U}\, q^{\binom{\dim U}{2}} = (-1)^{n-1}\, q^{\binom{n-1}{2}}.$$

I soli sottospazi U oltre a V tali che $U \vee P = V$ sono quelli U di dimensione $n-1$ che non contengono P; il loro numero è uguale a

$$\begin{bmatrix} n \\ n-1 \end{bmatrix} - \begin{bmatrix} n-1 \\ n-2 \end{bmatrix} = q^{n-1}.$$

Quest'espressione ci porta alla formula

$$\begin{aligned} \mu(0,V) &= -\sum_{U \subset V:\, U \vee P = V} \mu(0,U) \\ &= -q^{n-1} \times (-1)^{n-1}\, q^{\binom{n-1}{2}}. \end{aligned}$$

Ciò completa la dimostrazione. □

G4.3. Teorema q-binomiale. Sia $V_n(q)$ uno spazio vettoriale di dimensione finita n sul campo finito $K(q)$ di q elementi (q potenza di un primo); così $V_n(q)$ ha complessivamente q^n vettori distinti.

Sia $X(q)$ uno spazio vettoriale sullo stesso campo avente x vettori. Conteremo in due modi l'insieme di tutte le trasformazioni lineari da $V_n(q)$ a $X(q)$, ottenendo in questo modo un'identità.

Sia $T : V_n(q) \to X(q)$ una tale trasformazione lineare e sia $\{v_1, v_2, \cdots, v_n\}$ una base per lo spazio vettoriale $V_n(q)$. La trasformazione lineare T è univocamente determinata una volta che le immagini dei v_i con $i = 1, 2, \cdots, n$ sono assegnate. L' immagine di ogni v_i $(i = 1, 2, \cdots, n)$ può essere uno degli x vettori di $X(q)$, donde ci sono x^n scelte per T.

Adesso contiamo l'insieme di tutte le trasformazioni lineari T, in accordo con la dimensione dei loro nuclei. Avendo scelto un sottospazio N di dimensione k, $k \leq n$ di $V_n(q)$, l'insieme delle trasformazioni lineari T in $X(q)$ il cui nucleo è N viene contato come segue.

Siano $\{v_1, v_2, \cdots, v_n\}$ una base per $V_n(q)$ e $\{v_1, v_2, \cdots, v_k\}$ una base per il sottospazio N. Una trasformazione lineare T ha N come suo nucleo se e solo se le immagini dei vettori $\{v_1, v_2, \cdots, v_k\}$ sono nulle e nessun'altra combinazione lineare non banale dei v_i $(i = 1 + k, \cdots, n)$ viene annullata da T. Questo dà, per l'immagine di v_{k+1}, la scelta di $x - 1$ vettori, tutti appartenenti a $X(q)$ tranne il vettore nullo; per l'immagine di v_{k+2}, la scelta di $x - q$ vettori, tutti appartenenti a $X(q)$ tranne quelli della combinazione lineare ottenuta dall'immagine di v_{k+1}; per l'immagine di v_{k+3}, la scelta di $x - q^2$ vettori, tutti appartenenti a $X(q)$ tranne quelli della combinazione lineare ottenuta dalle immagini di v_{k+1} e di v_{k+2}; e così via.

Allora le trasformazioni lineari T con nucleo N sono in numero

$$(x-1)(x-q)\cdots(x-q^{n-k-1}).$$

Per il Teorema **G4.1** sappiamo che il numero di sottospazi k-dimensionali di $V_n(q)$ è $\begin{bmatrix} n \\ k \end{bmatrix}$; quindi, combinando i due conti, otteniamo l'identità cercata.

Teorema G4.4 (Teorema q-binomiale).

$$x^n = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} (x-1)(x-q)\cdots(x-q^{n-k-1}). \qquad \square$$

Applicando la formula esplicita per la funzione di Möbius sul reticolo degli spazi vettoriali finiti, possiamo ora determinare il numero delle trasformazioni fra due spazi vettoriali di dimensioni finite.

Teorema G4.5. *Il numero delle trasformazioni lineari suriettive da uno spazio n-dimensionale V_n ad uno spazio m-dimensionale V_m sullo stesso campo finito $K(q)$ è:*

$$\sum_{k=0}^{m}(-1)^{m-k}\begin{bmatrix} m \\ k \end{bmatrix} q^{nk+\binom{m-k}{2}}.$$

DIMOSTRAZIONE. Per un sottospazio $U \subseteq V_m$ denotiamo con $g(U)$ il numero delle trasformazioni lineari da V_n a V_m, la cui immagine è U, e con $f(U)$ il numero delle trasformazioni lineari da V_n a V_m, la cui immagine è contenuta in U. Chiaramente

$$f(U) = q^{n \dim U} \qquad \text{e} \qquad f(U) = \sum_{W \subseteq U} g(W).$$

Per il Teorema di inversione di Möbius **G1.14** risulta:

$$g(U) = \sum_{W \subseteq U} \mu(W, U)\, q^{n\, dim(W)}.$$

Prendiamo $U = V_m$ e usiamo i Teoremi **G4.1** e **G4.3** per avere la tesi. □

Ricordando che i sottospazi di dimensione ℓ in V_m sono in numero $\begin{bmatrix} m \\ \ell \end{bmatrix}$, si ha subito il seguente risultato.

Corollario G4.6. *Il numero delle matrici $n \times m$ sul campo finito $K(q)$ di rango ℓ ugualia*

$$\begin{bmatrix} m \\ \ell \end{bmatrix} \sum_{k=0}^{\ell}(-1)^{\ell-k}\begin{bmatrix} \ell \\ k \end{bmatrix} q^{nk+\binom{\ell-k}{2}}.$$

Notiamo che il numero delle trasformazioni lineari iniettive ha una forma relativamente semplice. Se fissiamo una base per V_n e consideriamo le iniezioni in V_m, l'immagine dell'i-esimo vettore base ($i = 1, 2, \cdots, n$) deve essere scelto come uno fra i $(q^m - q^{i-1})$ vettori che non appartengono allo spazio delle immagini dei precedenti vettori base. In conclusione ci sono

$$(q^m - 1)(q^m - q) \cdots (q^m - q^{n-1})$$

trasformazioni lineari iniettive. Poiché il corollario precedente con $\ell = n$ dà anche un'espressione per questo numero, abbiamo provato la seguente identità:

$$(q^m - 1)(q^m - q) \cdots (q^m - q^{n-1}) = q^{n^2}\begin{bmatrix} m \\ n \end{bmatrix} \sum_{k=0}^{n}(-1)^k\begin{bmatrix} n \\ k \end{bmatrix} q^{\binom{k}{2}-kn}.$$

G5. Funzione di Möbius del reticolo delle partizioni

Sulla base del reticolo $(\mathbb{P}(X), \leq)$ delle partizioni dell'insieme finito X, studiamo, in questa sezione, la funzione di Möbius e un'applicazione al calcolo del permanente per una matrice rettangolare.

G5.1. Funzione di Möbius su $(\mathbb{P}(X), \leq)$. Sia $(\mathbb{P}(X), \leq)$ il reticolo delle partizioni dell'insieme finito X, ordinato per rifinitezza. Se due partizioni $\mathcal{B}, \mathcal{F} \in \mathbb{P}(X)$ sono ordinate come $\mathcal{B} \leq \mathcal{F}$, allora per ogni parte $F \in \mathcal{F}$, si ha per rifinitezza che $\{B \in \mathcal{B} | B \subseteq F\} \in \mathbb{P}(F)$. Questa partizione viene chiamata la restrizione di $\mathcal{B}$ a F e denotata con $\mathcal{B}(F)$.

Proposizione G5.1 (Funzione di Möbius delle partizioni). *La funzione di Möbius per il reticolo $(\mathbb{P}(X), \leq)$ è data da*

$$\mu(\mathcal{B}, \mathcal{F}) = \prod_{F \in \mathcal{F}} (-1)^{\ell(\mathcal{B}(F))-1} (\ell(\mathcal{B}(F)) - 1)!$$

dove con $\ell(\mathcal{B}(F))$ si denota il numero delle parti della partizione $\mathcal{B}(F)$.

DIMOSTRAZIONE. Data $\mathcal{F} = \{F_1, F_2, \cdots, F_\ell\} \in \mathbb{P}(X)$, supponiamo che $\mathcal{B} \leq \mathcal{F}$. Per ogni $k = 1, 2, \cdots, \ell$, si definisce $\mathcal{B}_k = \mathcal{B}(F_k)$. Si verifica facilmente che l'intervallo $[\mathcal{B}, \mathcal{F}]$ nel $(\mathbb{P}(X), \leq)$ è isomorfo al prodotto diretto:

$$[\mathcal{B}_1, F_1] \otimes [\mathcal{B}_2, F_2] \otimes \cdots \otimes [\mathcal{B}_\ell, F_\ell]$$

che è a sua volta un intervallo del reticolo

$$(\mathbb{P}(F_1), \leq) \otimes (\mathbb{P}(F_2), \leq) \otimes \cdots \otimes (\mathbb{P}(F_\ell), \leq).$$

Notiamo che ogni intervallo $[\mathcal{B}_k, F_k]$ nel $(\mathbb{P}(F_k), \leq)$ è isomorfo all'intervallo $[\mathfrak{B}_k, \mathcal{B}_k]$ del reticolo $(\mathbb{P}(\mathcal{B}_k), \leq)$, dove $\mathfrak{B}_k$ è la partizione minima (ogni parte è composta da un solo membro) dell'insieme $\mathcal{B}_k$ (delle parti di $\mathcal{B}_k$ come membri). Allora l'intervallo $[\mathcal{B}, \mathcal{F}]$ nel $(\mathbb{P}(X), \leq)$ è isomorfo al prodotto diretto:

$$[\mathfrak{B}_1, \mathcal{B}_1] \otimes [\mathfrak{B}_2, \mathcal{B}_2] \otimes \cdots \otimes [\mathfrak{B}_\ell, \mathcal{B}_\ell].$$

Quindi si ha che

$$\mu(\mathcal{B}, \mathcal{F}) = \prod_{k=1}^{\ell} \mu_k(\mathfrak{B}_k, \mathcal{B}_k)$$

dove μ_k è la funzione di Möbius del reticolo $(\mathbb{P}(\mathcal{B}_k), \leq)$ per $k = 1, 2, \cdots, \ell$. Dunque la dimostrazione si riduce a confermare che, per ogni insieme finito B, la funzione di Möbius μ_o del reticolo $(\mathbb{P}(B), \leq)$ soddisfa la relazione:

$$\mu_o(\mathfrak{B}, B) = (-1)^{|B|-1}(|B| - 1)!. \qquad \text{(A)}$$

Possiamo procedere tramite il principio di induzione su $|B|$. Quando $|B| = 1$, sia sinistra che destra di (A) si riduce ovviamente a uno. Ora sia B con

$|B| > 1$. Supponiamo, come l'ipotesi dell'induzione, che vale (A) per tutti i sottoinsiemi $F \subset B$ con $|F| < |B|$.

Sia $\mathcal{F}$ una partizione di B con $\mathfrak{B} \leq \mathcal{F} < B$. Seguendo lo stesso ragionamento di prima, possiamo verificare che l'intervallo $[\mathfrak{B}, \mathcal{F}]$ nel $(\mathbb{P}(B), \leq)$ è isomorfo al prodotto cartesiano degli intervalli $[\mathfrak{F}, F]$ per $F \in \mathcal{F}$. Allora per l'ipotesi dell'induzione, si ha che

$$\mu_o(\mathfrak{F}, \mathcal{F}) = \prod_{F \in \mathcal{F}} (-1)^{|F|-1}(|F| - 1)!. \qquad \text{(B)}$$

Ricordiamo che la funzione di Möbius soddisfa la seguente proprietà:

$$\sum_{\mathcal{F} \in \mathbb{P}(B)} \mu(\mathfrak{B}, \mathcal{F}) = 0 \quad per \quad |B| > 1. \qquad \text{(C)}$$

Fissando $x \in B$, ogni partizione $\mathcal{F}$ di B può essere considerata come un'unione della parte D contenente x con una partizione del complemento di D in B. Allora (C) equivale alla seguente ($|B| > 1$):

$$\mu_o(\mathfrak{B}, B) = - \sum_{x \in D \subset B} \mu_o(\mathfrak{D}, D) \sum_{\mathcal{F} \in \mathbb{P}(B \setminus D)} \mu_o(\mathfrak{B} \setminus \mathfrak{D}, \mathcal{F}). \qquad \text{(D)}$$

Se $|B \setminus D| > 1$, l'ipotesi dell'induzione implica che la seconda somma in (D) si riduce a zero. Qulla somma è uguale a uno se $|B \setminus D| = 1$. A questo punto, possiamo riformulare (D) come segue:

$$\begin{aligned} \mu_o(\mathfrak{B}, B) &= - \sum_{\substack{x \in D \subset B \\ |B \setminus D| = 1}} (-1)^{|D|-1}(|D| - 1)! \\ &= - \sum_{\substack{x \in D \subset B \\ |B \setminus D| = 1}} (-1)^{|B|-2}(|B| - 2)! \\ &= (-1)^{|B|-1}(|B| - 2)!\,(|B| - 1) \end{aligned}$$

che è equivalente a (A). Secondo il principio dell'induzione, abbiamo completato la dimostrazione della Proposizione. □

G5.2. Permanente. Armati con la funzione di Möbius delle partizioni, procediamo a calcolare il permanente $\mathrm{per}(A)$ per una matrice $A = [a_{ij}]$ di ordine $n \times m$ su campo complesso $\mathbb{C}$.

Denotiamo con $(\mathbb{P}[n], \leq)$ il reticolo delle partizioni di $[n] := \{1, 2, \cdots, n\}$ con la minima partizione $\mathfrak{N} = \uplus_{k=1}^{n} \{k\}$. Per ogni partizione $\mathcal{D} \in \mathbb{P}[n])$, indichiamo con $\Omega(\mathcal{D})$ tutte le applicazioni da $[n]$ a $[m]$ che sono costanti in ogni parte $D \in \mathcal{D}$ e hanno valori distinti in tutte le parti di $\mathcal{D}$. Consideriamo

la somma definita da

$$\alpha(\mathcal{D}) = \sum_{\sigma\in\Omega(\mathcal{D})}\prod_{i=1}^{n} a_{i\sigma(i)}.$$

Osservando che $\Omega(\mathfrak{N})$ consiste di tutte le n-permutazioni di $[m]$, allora vale $\mathrm{per}(A) = \alpha(\mathfrak{N})$.

Introduciamo un'altra funzione di $\mathbb{P}[n]$ tramite la matrice A:

$$\beta(\mathcal{D}) = \prod_{D\in\mathcal{D}}\sum_{j=1}^{m}\prod_{i\in D} a_{ij}.$$

Se $|D| = 2$ con $D = \{i, i\}$, allora la somma $\sum_{j=1}^{m}\prod_{i\in D} a_{ij}$ si riduce al prodotto scalare delle due righe di A indicizzate con $\imath$ e i rispettivamente.

Per ogni partizione $\mathcal{B} \in \mathbb{P}[n]$, è possibile verificare le seguenti relazioni:

$$\beta(\mathcal{B}) = \sum_{\mathcal{B}\leq\mathcal{D}} \alpha(\mathcal{D}), \tag{$\triangle$}$$

$$\alpha(\mathcal{B}) = \sum_{\mathcal{B}\leq\mathcal{D}} \beta(\mathcal{D})\mu(\mathcal{B},\mathcal{D}). \tag{∇}$$

Secondo il teorema di inversione di Möbius, (∇) è la conseguenza immediata di $(\triangle)$. Quindi dobbiamo solo stabilire $(\triangle)$.

Data una partizione $\mathcal{B} = \{B_1, B_2, \cdots, B_\ell\} \in \mathbb{P}[n]$, possiamo riscrivere la funzione $\beta(\mathcal{B})$ come segue:

$$\begin{aligned}
\beta(\mathcal{B}) &= \prod_{B\in\mathcal{B}}\sum_{j=1}^{m}\prod_{i\in B} a_{ij}\\
&= \prod_{k=1}^{\ell}\sum_{j_k=1}^{m}\prod_{i\in B_k} a_{ij_k}\\
&= \sum_{\sigma\in\Lambda(\mathcal{B})}\prod_{i=1}^{n} a_{i\sigma(i)}
\end{aligned}$$

dove $\Lambda(\mathcal{B})$ è l'insieme delle applicazioni da $[n]$ ad $[m]$ che sono costanti in ogni parte B_k della partizione $\mathcal{B} \in \mathbb{P}[n]$.

Per ogni funzione $f \in \Lambda(\mathcal{B})$, la sua immagine induce una partizione $\mathcal{D}$ del dominio, cioè un raggruppamento delle parti di $\mathcal{B}$. Si ha ovviamente che $\mathcal{B} \leq \mathcal{D}$ in $\mathbb{P}[n]$ e $f \in \Omega(\mathcal{D})$. Classificando $\Lambda(\mathcal{B})$ secondo le partizioni $\mathcal{D}$ con $\mathcal{B} \leq \mathcal{D}$:

$$\Lambda(\mathcal{B}) = \biguplus_{\mathcal{B}\leq\mathcal{D}} \Omega(\mathcal{D})$$

otteniamo la seguente relazione:

$$\beta(\mathcal{B}) = \sum_{\sigma\in\Lambda(\mathcal{B})}\prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\mathcal{B}\le\mathcal{D}}\sum_{\sigma\in\Omega(\mathcal{D})}\prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\mathcal{B}\le\mathcal{D}}\alpha(\mathcal{D}).$$

Questa è esattamente $(\triangle)$ che volevamo dimostrare. □

Sostituendo $\alpha(A)$, $\beta(A)$ e $\mu(\mathcal{B},\mathcal{D})$ con le loro espressioni esplicite, ricaviamo da (∇) la seguente formule per il calcolo del permanente.

Teorema G5.2 (Permanente). *Per una matrice $A = [a_{ij}]$ di ordine $n \times m$ su campo complesso $\mathbb{C}$, il permanente* $\mathrm{per}(A)$ *è uguale alla seguente:*

$$\mathrm{per}(A) = \sum_{\mathcal{D}\in\mathbb{P}[n]}\prod_{D\in\mathcal{D}}(-1)^{|D|-1}(|D|-1)!\sum_{j=1}^{m}\prod_{i\in D} a_{ij}.$$

G5.3. Formula di Ryser. Possiamo trattare il permanente anche per mezzo del principio d'inclusione ed esclusione.

Sia $A = (a_{ij})$ una matrice $n \times m$ su un anello commutativo. Il *permanente* di A, scritto $\mathrm{per}(A)$, è definito dalla formula

$$\mathrm{per}(A) = \sum_{\pi}\prod_{i=1}^{n} a_{i\pi(i)}$$

dove la somma è estesa a tutte le applicazioni iniettive da $[n]$ ad $[m]$ (o tutte le n-permutazioni di $[m]$). Quando $m = n$, $\mathrm{per}(A)$ è la somma dei termini (a parte il fattore alternato) che compaiono nello sviluppo del determinante di A. Inoltre, definiamo una funzione sulla matrice

$$\rho(A) = \prod_{i=1}^{n}\sum_{j=1}^{m} a_{ij}$$

che è data dal prodotto delle somme di ogni riga della matrice A.

Per $\Omega = [m]^{[n]}$, introduciamo la funzione peso

$$w(\pi) = \prod_{i=1}^{n} a_{i\pi(i)} \quad \text{per ogni} \quad \pi\in\Omega.$$

Fissando $\kappa \in [m]$, consideriamo il sottoinsieme

$$B_\kappa = \{\pi\in\Omega \mid \pi^{-1}(\kappa) = \emptyset\}.$$

Per $\sigma \subseteq [m]$, indichiamo con σ^c il complemento di σ in $[m]$. Allora non è difficile verificare che

$$\mathcal{W}\Big(\bigcap_{\kappa\in\sigma} B_\kappa\Big) = \rho(A_{\sigma^c})$$

dove A_σ è la sottomatrice di A avente gli indici delle colonne in σ; perciò, A_{σ^c} è la sottomatrice di A senza le colonne indicizzate con σ.

Per l'insieme Ω e la classe dei sottoinsiemi $\{B_1, B_2, \cdots, B_m\}$, notiamo che tutte le applicazioni iniettive da $[n]$ ad $[m]$ costituiscono Ω_{m-n}. Secondo la formula pesata del principio d'inclusione ed esclusione (vedi il Teorema **G3.2**), si ha che

$$\begin{aligned}
\mathrm{per}(A) = \mathcal{W}(\Omega_{m-n}) &= \sum_{k=m-n}^{m} (-1)^{m-n+k} \binom{k}{m-n} \sum_{\substack{\sigma\subseteq[m]\\|\sigma|=k}} \mathcal{W}\Big(\bigcap_{\kappa\in\sigma} B_\kappa\Big) \\
&= \sum_{k=m-n}^{m} (-1)^{m-n+k} \binom{k}{m-n} \sum_{\substack{\sigma\subseteq[m]\\|\sigma|=k}} \rho(A_{\sigma^c}) \\
&= \sum_{k=1}^{n} (-1)^{n-k} \binom{m-k}{m-n} \sum_{\substack{\sigma\subseteq[m]\\|\sigma|=k}} \rho(A_\sigma)
\end{aligned}$$

dove l'ultimo passaggio viene giustificato dal fatto che $\rho(A_\emptyset) = 0$ e dalla sostituzione $k \to m-k$ sull'indice della somma. Così abbiamo stabilito il seguente importante risultato.

Teorema G5.3 (Formula di Ryser). *Sia $A = (a_{ij})$ una matrice $n \times m$ su un anello commutativo. Vale la seguente formula:*

$$\mathrm{per}(A) = \sum_{\sigma\subseteq[m]} (-1)^{n+|\sigma|} \binom{m-|\sigma|}{m-n} \rho(A_\sigma).$$

Esempio G5.4. *Sia $J[n\times m]$ la matrice $n\times m$ con tutti gli elementi uguali ad uno e $I[n\times m]$ la matrice con gli elementi diagonali uguali ad uno ed altri a zero. Allora*

$$\begin{aligned}
&\mathrm{per}\,(I[n\times m]) = 1; \\
&\mathrm{per}\,(J[n\times m]) = \langle m\rangle_n = m(m-1)\cdots(m-n+1); \\
&\mathrm{per}\,(J[n\times n] - I[n\times n]) = D_n.
\end{aligned}$$

Secondo il Teorema **G5.3** e l'espressione

$$\rho(J_\sigma[n\times m]) = |\sigma|^n$$

abbiamo la seguente formula

$$\begin{aligned}\mathrm{per}(J[n\times m]) &= \sum_{k=1}^{n}(-1)^{n-k}\binom{m-k}{m-n}\sum_{\substack{\sigma\subseteq[m]\\|\sigma|=k}}\rho(A_\sigma)\\ &= \sum_{k=1}^{n}(-1)^{n-k}\binom{m}{k}\binom{m-k}{m-n}k^n\\ &= \binom{m}{n}\sum_{k=1}^{n}(-1)^{n-k}\binom{n}{k}k^n\end{aligned}$$

che ci porta conseguentemente all'identità combinatoria:

$$\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}k^n = n!.$$

□